ΠΑΝΕΠΙΣΤΗΜΙΟ ΔΥΤΙΚΗΣ ΑΤΤΙΚΗΣ
ΣΧΟΛΗ ΜΗΧΑΝΙΚΩΝ
ΤΜΗΜΑ ΜΗΧΑΝΙΚΩΝ ΤΟΠΟΓΡΑΦΙΑΣ ΚΑΙ ΓΕΩΠΛΗΡΟΦΟΡΙΚΗΣ
ΠΡΟΓΡΑΜΜΑ ΜΕΤΑΠΤΥΧΙΑΚΩΝ ΣΠΟΥΔΩΝ
«ΓΕΩΧΩΡΙΚΕΣ ΤΕΧΝΟΛΟΓΙΕΣ»

ΔΙΕΡΕΥΝΗΣΗ ΜΟΝΤΕΛΩΝ ΠΑΛΙΝΔΡΟΜΗΣΗΣ ΓΙΑ ΤΗΝ ΕΚΤΙΜΗΣΗ ΑΞΙΑΣ ΑΚΙΝΗΤΩΝ

`

ΜΕΤΑΠΤΥΧΙΑΚΗ ΔΙΠΛΩΜΑΤΙΚΗ ΕΡΓΑΣΙΑ

Σκλήρης Ν. Φώτιος

Μέλη επιτροπής:

Π. Ηλιοπούλου, Καθηγήτρια ΠΑΔΑ (επιβλέπουσα)

Α. Τσάτσαρης, Αν. Καθηγητής ΠΑΔΑ

Ι. Κάτσιος, Αν. Καθηγητής ΠΑΔΑ

Αθήνα, Ιανουάριος 2019

Πίνακας Περιεχομένων

Περιεχόμενα i

Συντομογραφίες vi

Πρόλογος – Ευχαριστίες vii
Περίληψη viii
Abstract x

Κεφάλαιο 1ο- Ακίνητο
1.1 Ορισμός Ιδιοκτησίας Ακινήτου Εδάφους 1
1.2 Αγορά Ακινήτων 2
1.3 Κατηγορίες Χώρων Ακινήτων και Εκτάσεων Γης 3
1.4 Κυριότητα Ακινήτου 5
1.5 Αξία Ακινήτου 5
1.5.1 Οι Τρείς Διαστάσεις του Ακινήτου 6
1.6 Αναγκαιότητα Προσδιορισμού της Αξίας Ακινήτου 9
1.7 Παράγοντες που Επηρεάζουν την Αξία των Ακινήτων 10
1.7.1 Παράγοντες σε Επίπεδο Χώρας 12
1.7.2 Παράγοντες σε Επίπεδο Πόλης 13
1.7.3 Παράγοντες σε Επίπεδο Τμήματος Πόλης 13
1.7.4 Παράγοντες σε Επίπεδο Ακινήτου 14
1.8 Ζήτηση και Προσφορά Ακινήτων 14
1.8.1 Ζήτηση 14
1.8.2 Προσφορά 15
1.8.3 Νόμος Ζήτησης – Προσφοράς 15
1.9 Μορφές αγοράς 15
1.10 Φορείς που Συμμετέχουν στην Αγορά Ακινήτων 16
1.11 Η Ιδιαιτερότητα στη Φυσιογνωμία των Ελληνικών Ακινήτων 16
1.12 Επισκόπηση Παραμέτρων που Επηρεάζουν τις Τιμές των Ακινήτων 16
1.13 Η Αγορά Ακινήτων στην Ελλάδα μέσα από τη Στατιστική Ματιά 18
1.13.1 Πίεση στις Τιμές των Ακινήτων 19
1.13.2 Ελλάδα και Ιδιοκατοίκηση 20
1.13.3 Ελλάδα και Κόστος Κατοίκησης 21
1.13.4 Ελλάδα και Έλλειψη κατοικιών 21
1.13.5 Ελλάδα και Αριθμός Κατοικιών 21
1.14 Εκτίμηση Αξιών Ακινήτων 22
1.15 Είδη και Μέθοδοι Αποτίμησης Αξιών Ακινήτων 23
1.16 Μαζικές Εκτιμήσεις Ακινήτων 29
1.17 Τεχνικές Μαζικών Εκτιμήσεων 33

Κεφάλαιο 2ο- Μεταβλητές

2.1 Προσδιορισμός της Αξίας Ακινήτου και η Επιρροή των Μεταβλητών 35
2.2 Μεταβλητές και Κλίμακες Μέτρησης 36
2.2.1 Ονομαστικές Ιεραρχικές Ισοδιαστημικές Αναλογικές Μεταβλητές 36
2.2.2 Συνεχής Ασυνεχής Διχοτομικές 38
2.2.3 Ανεξάρτητες Εξαρτημένες Μεταβλητές 39
2.2.4 Ψευδομεταβλητές 40
2.3 Εξειδικευμένη Ανάλυση Μεταβλητών Επιρροής στα Ακίνητα 40
2.3.1 Ανεξάρτητες Μεταβλητές Κατασκευαστικών Χαρακτηριστικών 42
2.3.2 Ανεξάρτητες Μεταβλητές Θέσης 49
2.3.3 Ανεξάρτητες Μεταβλητές Γειτονιάς 52
2.3.4 Λοιπές Μεταβλητές 54
2.4 Συνοπτικός Πίνακας Μεταβλητών 55
2.5 Συνηθέστερα Χρησιμοποιούμενες Μεταβλητές στα Ηδονικά Μοντέλα 58

Κεφάλαιο 3ο – Μεθοδολογία
3.1 Βασικές Έννοιες Στατιστικής Θεωρίας 60
3.1.1 Πληθυσμός Στατιστικός Πληθυσμός και Αντικειμενικός Πληθυσμός 60
3.1.2 Δείγμα Δειγματοληψία 60
3.1.2.1 Μέγεθος Δείγματος 61
3.1.3 Κατανομές Συχνοτήτων 62
3.1.4 Μέτρα Κεντρικής Τάσης Περιγραφικής Στατιστικής 63
3.1.5 Μέτρα Δείκτες Διασποράς 63
3.1.6 Δειγματικός Χώρος Κατανομές Πιθανότητας Έλεγχοι Υποθέσεων 65
3.1.6.1 Δειγματικός Χώρος 65
3.1.6.2 Κατανομές 65
3.1.6.3 Διαστήματα Εμπιστοσύνης 66
3.1.6.4 Στατιστικές Υποθέσεις 67
3.1.6.4.1 Παραμετρικά Τεστ 70
3.1.6.4.2 Μη Παραμετρικά Τεστ 72
3.1.6.5 Επαγωγικοί Έλεγχοι 72
3.1.6.6 Διάγραμμα Επιλογής Αποφάσεων 73
3.2 Ανάλυση Παλινδρόμησης 74
3.3 Διάγραμμα Σκεδασμού και Ανάλυση Συσχέτισης 75
3.3.1 Διάγραμμα Σκεδασμού 75
3.3.2 Ανάλυση Συσχέτισης 76
3.4 Μοντέλα Παλινδρόμησης 79
3.4.1 Απλή Γραμμική Παλινδρόμηση (simple linear regression) 79
3.4.2 Πολλαπλή Παλινδρόμηση (multiple regression) 88
3.4.2.1 Μέθοδοι Επιλογής Μεταβλητών 93
3.4.2.2 Κατασκευή ενός Τυπικού Μοντέλου Πολλαπλής Παλινδρόμησης 94
3.4.3 Χωρική Παλινδρόμηση 96
3.4.4 Πολυωνυμική Παλινδρόμηση 102
3.4.5 Παλινδρόμηση κατά Εκατοστημόριο ή Τεταρτημοριακή Παλινδρόμηση (Quantile Regression) 103

3.4.6 Λογιστική Παλινδρόμηση 104
3.4.7 Μετασχηματισμοί δεδομένων στην Παλινδρόμηση ως Εναλλακτικές Μορφές του Γραμμικού Μοντέλου Εγγενείς Γραμμικές Συναρτήσεις Παλινδρόμησης 106
3.4.7.1 Linear- Log Παλινδρόμηση ως Εναλλακτική Μορφή του Γραμμικού Μοντέλου 106
3.4.7.2 Log-Linear Παλινδρόμηση ως Εναλλακτική Μορφή του Γραμμικού Μοντέλου 107
3.4.7.3 Log Log Παλινδρόμηση ως Εναλλακτική Μορφή του Γραμμικού Μοντέλου 108
3.4.7.4 Αντίστροφοι Μετασχηματισμοί και Μετασχηματισμοί Τετραγωνικής Ρίζας και Τετραγώνου 108
3.4.7.5 Λοιπές Εγγενείς Γραμμικές Συναρτήσεις Παλινδρόμησης 109
3.4.7.6 Μοντέλο Μετασχηματισμού Box-Cox 109
3.4.8 Δένδρα Αποφάσεων Δενδρική Παλινδρόμηση και Ταξινόμηση Τεχνική CART 110
3.5 Η Εφαρμογή των Μοντέλων Παλινδρόμησης για την Αξιολόγηση Μεταβλητών Επηρεασμού Τιμών και Προβλεπτικών Μοντέλων Εκτίμησης Αξιών Ακινήτων 112
3.5.1 Ενδεικτική Βιβλιογραφική Ανασκόπηση Μοντέλων Παλινδρόμησης για την Εκτίμηση Αξιών Ακινήτων 113

Κεφάλαιο 4ο – Συλλογή Δεδομένων
4.1 Γενικά Στοιχεία 116
4.2 Στατιστικά Στοιχεία 119
4.3 Επιλογή και Μέγεθος Δείγματος στη Περιοχή Μελέτης 121
4.4 Πηγή Δεδομένων Στοιχείων Δειγματοληψίας στην Περιοχή Μελέτης και Παραδοχές 125
4.5 Συλλεχθείσα Πληροφορία 126

Κεφάλαιο 5ο – Διερεύνηση Μοντέλων Παλινδρόμησης
5.1 Εισαγωγή Στοιχείων στο Λογισμικό SPSS 147
5.2 Ορισμός και Κωδικοποίηση της Εξαρτημένης και των Ανεξάρτητων Μεταβλητών 148
5.3 Περιγραφή Δεδομένων Μέτρα Κεντρικής Τάσης- Διασποράς 149
5.3.1 Περιγραφική στατιστική 149
5.3.2 Κατανομές Μεταβλητών Περιγραφή Δεδομένων 150
5.3.2.1 Υπολογισμός Στατιστικών Μέτρων Ποσοτικών Μεταβλητών 150
5.3.2.2 Υπολογισμός Κατανομών Συχνοτήτων Ποσοτικών Μεταβλητών 155
5.3.2.3 Υπολογισμός Κατανομών Συχνοτήτων Ποιοτικών Μεταβλητών 157
5.4 Δημιουργία και Διερεύνηση Μοντέλων 157
5.4.1 Δημιουργία Πρώτου Μοντέλου 157
5.4.1.1 Προκαταρτικοί Έλεγχοι 158
5.4.1.1.1 Έλεγχος Τυχαιότητας Παρατηρήσεων 158
5.4.1.1.2 Έλεγχος Κανονικότητας 159

5.4.1.1.3 Έλεγχος Ακραίων Παρατηρήσεων 162
5.4.1.2 Έλεγχοι Στατιστικών Υποθέσεων 162
5.4.1.3 Συσχέτιση Μεταβλητών- Έλεγχος x2 163
5.4.1.4 Εξέταση ισομερούς κατανομής ποιοτικών μεταβλητών 165
5.4.1.5 Προσαρμογή Μοντέλου Γραμμικής Παλινδρόμησης 167
5.4.1.5.1 Έλεγχος Ορθότητας Μοντέλου 176
5.4.1.5.2 Εξαγωγή Συμπερασμάτων Αποτελεσμάτων Πρώτου Μοντέλου Μετά από την Επεξεργασία με το Λογισμικό SPSS 181
5.4.1.6 Κλασσική παλινδρόμηση (Ordinary Least Squares Estimation-OLS) σε Περιβάλλον GIS Πρώτο Μοντέλο 182
5.4.1.7 Χωρική Παλινδρόμηση (Geographically Weighted Regression-GWR) σε περιβάλλον GIS Πρώτο Μοντέλο 185
5.4.1.7.1 Εξαγωγή Συμπερασμάτων των Αποτελεσμάτων του Πρώτου Μοντέλου Μετά από την Επεξεργασία με την Χωρική παλινδρόμηση (Geographically Weighted Regression-GWR 189
5.4.2 Δημιουργία Δεύτερου Μοντέλου 190
5.4.2.1 Προκαταρτικοί Έλεγχοι 190
5.4.2.2 Έλεγχος και Αφαίρεση Ακραίων Παρατηρήσεων 191
5.4.2.3 Σύγκριση Βασικών Περιγραφικών Μέτρων 193
5.4.2.4 Προσαρμογή Μοντέλου Γραμμικής Παλινδρόμησης 193
5.4.2.4.1 Έλεγχος ορθότητας μοντέλου 197
5.4.2.4.2 Εξαγωγή Συμπερασμάτων Αποτελεσμάτων Δεύτερου Μοντέλου μετά την επεξεργασία με το λογισμικό SPSS 201
5.4.2.5 Εξέταση του Δεύτερου Μοντέλου Παλινδρόμησης σε Περιβάλλον Γεωγραφικού Συστήματος Πληροφοριών 201
5.4.2.5.1 Χωρική Στατιστική Ανάλυση 202
5.4.2.5.2 Κλασσική παλινδρόμηση (Ordinary Least Squares Estimation-OLS) σε Περιβάλλον GIS 203
5.4.2.5.3 Χωρική Παλινδρόμηση (Geographically Weighted Regression-GWR) σε περιβάλλον GIS 207
5.4.2.5.4 Εξαγωγή Συμπερασμάτων των Αποτελεσμάτων του Δεύτερου Μοντέλου μέσα από την Χωρική παλινδρόμηση (Geographically Weighted Regression-GWR.) 211
5.4.3 Δημιουργία Τρίτου Μοντέλου 212
5.4.3.1 Έλεγχος και Αφαίρεση Ακραίων Παρατηρήσεων 212
5.4.3.2 Μετασχηματισμοί Μεταβλητών 215
5.4.3.3 Ομαδοποίηση Μεταβλητών 216
5.4.3.4 Προσαρμογή Μοντέλου Γραμμικής Παλινδρόμησης ανά Ομάδα Μεταβλητών 217
5.4.3.4.1 Προσαρμογή Τελικού Μοντέλου Γραμμικής Παλινδρόμησης 228
5.4.3.5 Έλεγχος Ορθότητας Μοντέλου 231
5.4.3.6 Εξαγωγή Συμπερασμάτων Αποτελεσμάτων Τρίτου Μοντέλου μετά την επεξεργασία με το λογισμικό SPSS 233

5.4.3.7 Εξέταση του Τρίτου Μοντέλου Παλινδρόμησης σε Περιβάλλον Γεωγραφικού Συστήματος Πληροφοριών 234
5.4.3.7.1 Κλασσική παλινδρόμηση (Ordinary Least Squares Estimation-OLS) σε Περιβάλλον GIS 234
5.4.3.7.2 Χωρική Παλινδρόμηση (Geographically Weighted Regression-GWR) σε περιβάλλον GIS 238
5.4.3.7.3 Εξαγωγή Συμπερασμάτων των Αποτελεσμάτων του Τρίτου Μοντέλου μέσα από την Χωρική παλινδρόμηση (Geographically Weighted Regression-GWR.) 242
5.4.4 Δημιουργία Τέταρτου Μοντέλου 243
5.4.4.1 Προσαρμογή Μοντέλου Γραμμικής Παλινδρόμησης 243
5.4.4.2 Έλεγχος Ορθότητας Μοντέλου 247
5.4.4.3 Εξαγωγή Συμπερασμάτων Αποτελεσμάτων Τέταρτου Μοντέλου μετά την επεξεργασία με το λογισμικό SPSS 250
5.4.4.4 Εξέταση του Τέταρτου Μοντέλου Παλινδρόμησης σε Περιβάλλον Γεωγραφικού Συστήματος Πληροφοριών 250
5.4.4.4.1 Χωρική Στατιστική Ανάλυση 250
5.4.4.4.2 Κλασσική παλινδρόμηση (Ordinary Least Squares Estimation-OLS) 251
5.4.4.4.3 Χωρική Παλινδρόμηση (Geographically Weighted Regression-GWR) σε περιβάλλον GIS 255
5.4.4.4.4 Εξαγωγή Συμπερασμάτων των Αποτελεσμάτων του Τέταρτου Μοντέλου μέσα από την Χωρική παλινδρόμηση (Geographically Weighted Regression-GWR) 259
5.4.5 Σύγκριση Μοντέλων 259

Κεφάλαιο 6ο – Συμπεράσματα 262

Βιβλιογραφία 265

Παράρτημα CD

Συντομογραφίες

ΑΚ	Αστικός Κώδικας
AMV	Automated Mass Valuation
ANOVA	Analysis of Variance
AIC	Akaike Information Criterion
CAMA	Computer Assisted Mass Appraisals
CART	Classification and Regression Trees
CAV	Computer Assisted Valuation
Cook's D	Cook' s Distance
ΔΛΠ	Διεθνές Λογιστικό Πρότυπο
DfFITS	Difference in Fits
ΕΚ	Ευρωπαϊκός Κανονισμός
GIS	Geographic Information Systems
GWR	Geographically weighted regression
ΗΣΑΠ	Ηλεκτρικοί Σιδηρόδρομοι Αθηνών Πειραιώς
IAAO	International Association of Assessing Officers
IVSC	International Valuation Standards Committee
ΚΤΕΛ	Κοινό Ταμείο Είσπραξης Λεωφορείων
MSE	Mean Squared Error
MSR	Regression Mean Squared
ΝΟΚ	Νέος Οικοδομικός Κανονισμός
ΟΑΣΑ	Οργανισμός Αστικών Συγκοινωνιών Αττικής
ΟΚΕ	Οικονομικής Επιτροπής της Ελλάδος
OLS	Ordinary Least Squares
ΟΣΕ	Οργανισμός Σιδηροδρόμων Ελλάδος
SSE	Error Sum of Squares
SSR	Regression Sum of Squares
SST	Total Sum of Squares
VIF	Variance Inflation Factor

Πρόλογος – Ευχαριστίες

Η παρούσα μεταπτυχιακή διπλωματική εργασία με τίτλο ««Διερεύνηση Μοντέλων Παλινδρόμησης για την Εκτίμηση Αξιών Ακινήτων» έγινε στα πλαίσια Μεταπτυχιακού προγράμματος Σπουδών του Πανεπιστημίου Δυτικής Αττικής, στην Σχολή Μηχανικών του τμήματος Μηχανικών Τοπογραφίας και Γεωπληροφορικής, στην γνωστική περιοχή της γεωγραφικής ανάλυσης και χωρικής στατιστικής.

Σίγουρα δεν θα ήταν εφικτό να γίνει πραγματικότητα εάν δεν υπήρχε η πολύτιμη βοήθεια της καθηγήτριας κας Ηλιοπούλου Πολυξένης, την οποία και θα ήθελα να ευχαριστήσω θερμά για την ανάθεση της εργασίας, την εμπιστοσύνη που μου έδειξε και την αδιάκοπη και πολύτιμη καθοδήγησή της.

Σκλήρης Φώτιος 2019

Περίληψη

Η εκτίμηση της αξίας των ακινήτων αποτελεί τμήμα της καθημερινότητας μιας κοινωνίας και αφορά τομείς όπως αγοροπωλησίες, μεταβιβάσεις, λήψεις δανείων και επενδυτικές κινήσεις. Οι διαδικασίες αυτές αφορούν είτε μεμονωμένα άτομα, είτε ιδρύματα και κρατικούς φορείς.

Το ακίνητο αποτελεί ένα κεφαλαιουχικό αγαθό. Ο προσδιορισμός της αξίας του με αντικειμενικό τρόπο αποτελεί πρόβλημα που απασχολεί διαχρονικά τους συμμετέχοντες. Σαφώς και η προσπάθεια αυτοματοποίησης και ελέγχου της εκτίμησης των αγοραίων αξιών ακινήτων μέσω μοντέλων με την βοήθεια της στατιστικής είναι ένας στόχος, η δε επιτυχημένη διαδικασία προσδιορισμού του κατάλληλου μοντέλου είναι ένα επίτευγμα δύσκολο.

Σκοπός της παρούσας εργασίας είναι η διερεύνηση μοντέλων παλινδρόμησης για την εκτίμηση της αξίας των κατοικιών. Αυτό γίνεται με τη βοήθεια ενός στατιστικού δείγματος και την αναζήτηση αυτών των μεταβλητών καθώς και της σχέσης μεταξύ τους που ερμηνεύουν βέλτιστα τον τρόπο με τον οποίο διαμορφώνεται η αγοραία αξία των κατοικιών. Στόχος είναι η διερεύνηση των διαφοροποιήσεων που υφίστανται οι μεταβλητές με βάση την επίδραση του χώρου και της γεωγραφικής θέσης του δείγματος καθώς δεν πρέπει να λησμονούμε ότι πρόκειται για ένα γεωγραφικό φαινόμενο.

Η διερεύνηση πραγματοποιείται στην πόλη των Αθηνών και η συλλογή των ακίνητων του δείγματος και των μεταβλητών γίνεται μέσα από διαδικτυακή πύλη πώλησης ακινήτων αλλά και από στοιχεία της Ελληνικής Στατιστική Αρχής. Επίσης μέσω των Συστημάτων Γεωγραφικών Πληροφοριών δημιουργούνται και νέες μεταβλητές οι οποίες παρέχουν πρόσθετη πληροφορία σχετικά με τη θέση των κατοικιών του δείγματος.

Αφού διερευνηθούν χαρακτηριστικά των μεταβλητών σχηματίζονται τέσσερα μοντέλα παλινδρόμησης με λογισμικό στατιστικής ανάλυσης και με λογισμικά χωρικής στατιστικής ανάλυσης σε περιβάλλον GIS. Δημιουργούνται μοντέλα πολλαπλής γραμμικής παλινδρόμησης (OLS) και γεωγραφικά σταθμισμένης παλινδρόμησης (GWR) χρησιμοποιώντας τις αρχικές μεταβλητές αλλά και μετασχηματισμούς των μεταβλητών (log-log models).

Το πρώτο κεφαλαίο είναι αφιερωμένο στα ακίνητα, το δεύτερο στις μεταβλητές, το τρίτο στη μεθοδολογία, το τέταρτο στα στατιστικά στοιχεία της εργασίας και το πέμπτο στην δημιουργία των μοντέλων παλινδρόμησης. Τα συμπεράσματα που προκύπτουν στο τέλος της εργασίας επικεντρώνονται στη μαθηματική και φυσική ερμηνεία των μοντέλων και στη σύγκριση των μεταβλητών που διαμορφώνουν τις τελικές αξίες ακινήτων. Τέλος αξιολογούνται τα πλεονεκτήματα και μειονεκτήματα της δημιουργίας γραμμικών μοντέλων με την OLS και GWR διαδικασία.

Διαπιστώνουμε ότι τα καλύτερα παραχθέντα μοντέλα είναι κατά σειρά το τέταρτο και δεύτερο μοντέλο. Το τέταρτο μοντέλο είναι της μορφής log- log ενώ το δεύτερο όχι. Παρατηρώντας και τα τέσσερα μοντέλα διαπιστώνουμε ότι οι μεταβλητές «Εμβαδόν», «Κατάσταση κατασκευής», «Όροφος», «Είδος θέρμανσης» , «Τιμή ζώνης» συμμετέχουν σε όλα, αποδεικνύοντας την σημαντικότητά τους.

Λέξεις κλειδιά: ακίνητο, τιμές κατοικιών, πολλαπλή γραμμική παλινδρόμηση, γεωγραφικά σταθμισμένη παλινδρόμηση.

Abstract

The property assessment is part of everyday life in a society for transactions such as buying and selling, loans and investments. These procedures relate to either individuals or institutions or government agencies.

Property is an asset. Determining its value in an objective way is always a problem for each of the involved parties. Clearly the effort for automation of the assessment of real estate market values by modeling with the help of statistics is a goal, however the successful process of determining the appropriate model is a difficult achievement.

The purpose of this thesis is to explore different regression models in order to estimate the value of houses. This is carried out through statistical sampling and selection of those variables that best explain the way in which the market value of the properties is determined. The aim is to investigate the differences of variables, based on the effect of space and location of the sample since property values constitute a geographical phenomenon.

The research has been carried out in the city of Athens. The sample and the variables were collected through an online real estate sales portal. Data were also collected from the Greek Statistical Authority while new variables concerning the location of houses were calculated in a GIS environment.

After describing the variables, four regression models were derived through statistical analysis and GIS software. Multiple linear regression and geographically weighted regression models were calculated employing the original variables but also transformed variables (log-log models).

The first chapter is dedicated to the description of the real estate market, the second concerns the variables, the third presents the methodology, the fourth describes the study region and the fifth includes the calculation of the regression models.

The conclusions drawn at the end of this work focus in the mathematical and physical interpretation of models as well as the comparison of variables that

contribute to final property values. Finally, the advantages and disadvantages of generating linear models addressed through the OLS and GWR process.

It seems that the models with the best fit are the fourth and the second one. For these models both the OLS and the GWR procedures were applied. However the fourth model is a log-log model while the second one is not. Observing all four models we find that the variables «Size» «Condition», «Floor», «Type of heating», «Zone Price» participates in all models, proving their importance.

Keywords: property: housing prices, multiple linear regression, geographically weighted regression, Greece.

Κεφάλαιο 1^{ο} – Ακίνητο

Η εκτίμηση ης αξίας των κατοικιών μέσα από μια αυτοματοποιημένη διαδικασία με την βοήθεια της στατιστικής είναι μια επιστημονική πρόκληση, η δε κατασκευή κατάλληλων προβλεπτικών μοντέλων θα μπορούσε να βοηθήσει στην επίλυση πληθώρας προβλημάτων που σχετίζονται με τον ορθό προσδιορισμό της αξίας των κατοίκων, διαμορφώνοντας ορθά σταθερές βάσεις που θα μπορούν να χρησιμοποιηθούν, όχι μόνο από το ζευγάρι αγοραστή- πωλητή, αλλά και από το κράτος για την ορθή φορολόγηση των κατοικιών.

1.1 Ορισμός Ιδιοκτησίας, Ακινήτου, Εδάφους

Η **Ιδιοκτησία** είναι μία νομική έννοια που περιλαμβάνει όλα τα περιουσιακά δικαιώματα, εμπράγματα και ενοχικά. Διακρίνεται σε ατομική και συλλογική. Αποτελεί θεσμό και απόλυτο δικαίωμα που προστατεύεται από το άρθρο 17 του Συντάγματος, όμως τα δικαιώματα που απορρέουν από αυτή δεν μπορούν να ασκούνται σε βάρος του γενικού συμφέροντος. Ως ανθρώπινος θεσμός υποδηλώνει την κυριαρχία του ανθρώπου πάνω σε ένα αγαθό, εκφράζοντας την ορισμένη σχέση του προς ένα μέρος του κόσμου της ύλης. Το αίσθημα της ιδιοκτησίας θεωρείται σύμφυτο της ανθρώπινης φύσης και η προσπάθεια για την απόκτησή της είναι φυσικό επακόλουθο. ««*Η μορφή της ιδιοκτησίας της γης πέρασε από πολλά στάδια και επηρεάστηκε από τις ιστορικές εξελίξεις. Αρχικά, ο κάτοχος γης ήταν μεν απόλυτος κύριος, χωρίς όμως να έχει και το δικαίωμα της ελεύθερης διάθεσης. Το δικαίωμα αυτό αποκτήθηκε μεταγενέστερα με την εξομοίωση της ιδιοκτησίας ακινήτων προς την ιδιοκτησία των κινητών πραγμάτων*»» (Ζεντέλης, 2015,σελ.39).

Θέλοντας να ορίσουμε εννοιολογικά το **Ακίνητο** θα το προσδιορίζαμε ως πράγμα. Είναι πρακτικά ένας χώρος που τεκμηριώνει ιδιοκτησιακό δικαίωμα, ένα περιουσιακό στοιχείο, που στην υλική του μορφή απεικονίζει το συμβατό δικαίωμα της ιδιοκτησίας (Ζεντέλης, 2015). Σύμφωνα με τον Αστικό Κώδικα, το ακίνητο είναι το έδαφος και τα συστατικά του μέρη (άρ. 948 Α.Κ), δηλαδή εκείνα που δεν μπορεί να αποχωριστεί από το κυρίως πράγμα, χωρίς τη βλάβη αυτού του ιδίου ή του κυρίου του πράγματος (άρ.953 Α.Κ), τα οποία ως αναλώσιμο προϊόν έχουν περιορισμένο χρόνο ζωής. Από οικονομικής απόψεως το ακίνητο περιλαμβάνει όλους τους συντελεστές παραγωγής, τη γη, την εργασία, το κεφάλαιο και την επιχειρηματικότητα, δηλαδή αυτά που αποτελούν τις βελτιώσεις της γης. Το ακίνητο είναι ένας χώρος απτός ή μη,

στον οποίο θεμελιώνεται αυτοτελές ή εξ αδιαιρέτου ιδιοκτησιακό δικαίωμα. Είναι ένα περιουσιακό στοιχείο το οποίο αποτελείται από έδαφος με απεριόριστο χρόνο ζωής και κτίσματα με περιορισμένο.

Η **Ακίνητη περιουσία** σύμφωνα με τους Παγουρτζή, Ασημακόπουλο, Χατζηχρήστο & French (2003) οριοθετείται από τον άνθρωπο έχοντας συγκεκριμένη γεωγραφική θέση και περιγράφει μια δραστηριότητα σε συγκεκριμένη χρονική περίοδο.

Ως **Έδαφος** θεωρείται το ανώτερο στρώμα του φλοιού της γης, περιλαμβάνει την γεωγραφική έκταση καθώς και τις ιδιότητες του που είναι χρήσιμες για τη παραγωγική διαδικασία, ενώ υπό την ευρεία έννοια, πληροί τους όρους της έννοιας του πράγματος και μπορεί να εξατομικευτεί με την ακριβή οριοθέτηση κατά θέση, όρια, πλευρικές διαστάσεις και σχήμα. Το έδαφος περιλαμβάνει όλα εκείνα τα στοιχεία που προέρχονται από τη φύση και μπορεί να χρησιμοποιηθούν για την παραγωγή αγαθών. Ο όρος **Γη**, αναφέρεται στα εντός συναλλαγής τμήματα του εδάφους και ως στοιχείο του περιβάλλοντος που μπορεί να χρησιμοποιηθεί στην παραγωγική διαδικασία, προσδιορίζεται με την έννοια του γεωτεμαχίου, ενώ ως **Εργασία** εννοούμε το Συντελεστή παραγωγής, δηλαδή κάθε ανθρώπινη προσπάθεια που συνειδητά καταβάλλεται και εισρέει στην παραγωγική διαδικασία αποβλέποντας στην επίτευξη οικονομικού αποτελέσματος. Το **Κεφάλαιο** παίρνει την αμοιβή του, με την μορφή τόκου για τη συμμετοχή του στην ενδιάμεση παραγωγική διαδικασία, ενώ η επιχειρηματική εργασία βρίσκει και συντονίζει τους συντελεστές παραγωγής. Ουσιαστικά το ακίνητο περιλαμβάνει τις βελτιώσεις που μπορεί να δεχτεί η κενή γη (Λιανός, Παπαβασιλείου & Χατζηανδρέου, 2015).

1.2 Αγορά Ακινήτων

Αγορά ακινήτων καλείται το σύνολο των θεσμών, σχέσεων και κανόνων που δρουν κατά τη συναλλαγή των ακινήτων μεταξύ αγοραστών και πωλητών μέσα από ένα μηχανισμό διαμόρφωσης τιμών. Εξυπηρετεί οικιστική, εμπορική, βιομηχανική και ειδικής κατάστασης χρήση. Η αγορά ακινήτων είναι χωρικά κατανεμημένη σε όλη τη χώρα και χωρίς συγκεκριμένη γεωγραφική θέση. Είναι περιοδική, καθώς δεν υπάρχει συνέχεια των συναλλαγών, όπως συμβαίνει με πολλές άλλες αγορές αγαθών συνεχούς συναλλακτικής ροής (Ζεντέλης, 2015). Τα ακίνητα καλύπτουν βασικές ανθρώπινες ανάγκες, όπως την παροχή στέγης, είναι όμως και κατά φύση επενδυτικά

στοιχεία με έντονο οικονομικό ενδιαφέρον. Αυτό είναι που θέτει τις βάσεις για την ύπαρξη της έννοιας της αγοράς ακινήτων που διαμορφώνεται και εξετάζεται κατά μεγάλο βαθμό αυτοτελώς, τόσο λόγω της σημαντικότητας της, όσο και λόγω των διαφόρων ιδιαιτεροτήτων που παρουσιάζει. Οι τιμές και γενικότερα οι συνθήκες που επικρατούν στην αγορά ακινήτων και οι επενδύσεις που πραγματοποιούνται σε κάθε είδους ακίνητα, ορίζονται και κινούνται εντός των πλαισίων που ορίζει η κτηματαγορά, δηλαδή του συνόλου των τιμών και όρων που έχουν σχέση με την αγοροπωλησία ακινήτων.

1.3 Κατηγορίες Χώρων, Ακινήτων και Εκτάσεων Γης

Ως **Χώρος** γενικά ονομάζεται το περιβάλλον που είναι άρρηκτα συνδεδεμένο με την απόσταση σε τρεις διαστάσεις, Εντός του χώρου συμβαίνουν και εξελίσσονται τα φυσικά φαινόμενα. Είναι το πλαίσιο μέσα στο οποίο υπάρχουν όλα τα υλικά σώματα και μέσα στο οποίο καθορίζεται η θέση τους.

Ο χώρος διακρίνεται σε τρεις βασικές κατηγορίες χρήσεων:

- Πρώτη κατηγορία είναι ο **Αστικός χώρος**. Το αστικό γεωτεμάχιο, προορίζεται για την κάλυψη των αστικών και ανθρώπινων αναγκών. Σύμφωνα με την Μαντούβαλου (1996) το δημόσιο, ως κύριος φορέας της εγκατάστασης των υποδομών και λειτουργιών βάσης, αναπτύσσει τον αστικό χώρο.
- Δεύτερη κατηγορία είναι ο **Αγροτικός χώρος** που είναι προορισμένος για αγροτική εκμετάλλευση. Αποτελείται από το μη αμιγές τμήμα του, που περιλαμβάνει τις εν δυνάμει αστικοποιούμενες περιοχές και από το αμιγές τμήμα του, το υπολειπόμενο του μη αστικοποιούμενου χώρου, που μπορεί να χρησιμοποιηθεί για αγροτική εκμετάλλευση. Αυτό το τμήμα είναι μεταβαλλόμενο διαχρονικά λόγω της αύξησης των αστικοποιημένων περιοχών.
- Τέλος έχουμε την τρίτη κατηγορία χώρου, τον **Δασικό χώρο,** που αποτελείται από το δάσος και γενικά τις δασικές εκτάσεις και για τον οποίο προβλέπεται συνταγματική προστασία.

Το **Ακίνητο** είναι, όπως προαναφέρθηκε, χώρος που τεκμηριώνει ιδιοκτησιακό δικαίωμα. Σε ότι αφορά την συναλλακτική δραστηριότητα, τα ακίνητα διαχωρίζονται σε:

- **εντός συναλλαγής**, δηλαδή νομικά αυτοτελή και μεταβιβάσιμα αντικείμενα συναλλαγής και τα

- **εκτός συναλλαγής**, δηλαδή τα κοινής χρήσης ή αυτά που είναι εντός κοινοχρήστου χώρου.

Τα ακίνητα διακρίνονται σε τρείς βασικές κατηγορίες χρήσεων :

- Πρώτη κατηγορία είναι τα **Αστικά ακίνητα**, δηλαδή τα προοριζόμενα για διαμονή ανθρώπων, που βρίσκονται εντός εγκεκριμένου ρυμοτομικού σχεδίου πόλης ή οικισμού ή εντός οικιστικής ζώνης.
- Δεύτερη κατηγορία είναι τα **Αγροτικά ακίνητα**, που προορίζονται για καλλιέργεια και γεωργική εκμετάλλευση, βρισκόμενα σε εκτός σχεδίου περιοχές.
- Τρίτη κατηγορία είναι τα **Δασικά ακίνητα**, που βρίσκονται σε δασική έκταση ή μέσα σε δάσος.

Τα ακίνητα ανάλογα με τη χρήση των κτιρίων που βρίσκονται επί της γης διακρίνονται σε ακίνητα για :

- **Οικιστική χρήση**, δηλαδή αυτά που προορίζονται για διαμονή ανθρώπων.
- **Επαγγελματική χρήση**, δηλαδή γραφεία, καταστήματα, αποθήκες.
- **Μικτή χρήση**, δηλαδή επαγγελματική και οικιστική χρήση.
- Ακίνητα **Ειδικής δραστηριότητας,** όπως για παράδειγμα **ν**οσοκομεία, σχολεία, ξενοδοχεία, κτλ.

Θα μπορούσαμε να διαχωρίσουμε τα ακίνητα και με γνώμονα την οικονομική τους δυνατότητα σε:

- **Επενδυτικά**, δηλαδή ακίνητα που είναι ιδιόκτητα ή κατέχονται με χρηματοδοτική μίσθωση και προορίζονται είτε για εκμίσθωση, είτε για αποκόμιση οφέλους από αύξηση της αξίας τους, είτε και για τα δύο, αλλά όχι για ιδιοχρησιμοποίηση ή πώληση.
- **Ιδιοχρησιμοποιούμενα**, δηλαδή ακίνητα κατεχόμενα από τον ιδιοκτήτη ή τον μισθωτή βάσει χρηματοδοτικής μίσθωσης, τα οποία χρησιμοποιούνται για κατοίκηση ή στην παραγωγή αγαθών ή την παροχή υπηρεσιών ή για διοικητικούς σκοπούς .
- **Αναπτυξιακά**, δηλαδή ακίνητα που έχουν σαν σκοπό την ανάπτυξη.

Τέλος οι εκτάσεις γης, υπό την ευρεία χωρική τους έννοια, διακρίνονται στις εκτάσεις:

- **Εντός σχεδίου**, που ονομάζονται οικόπεδα.
- **Εκτός σχεδίου**, που ονομάζονται αγροτεμάχια ή δασοτεμάχια.

1.4 Κυριότητα Ακινήτου

Σύμφωνα με τον Αστικό Κώδικα (άρ. 999 έως 1001) η **Κυριότητα** ενός ακινήτου είναι το πλήρες δικαίωμα που παρέχεται από το νόμο και αναγνωρίζει την σύννομη, άμεση και απόλυτη εξουσία πάνω σε αυτό, η οποία δεν απειλείται από αυτόν ή από δικαιώματα τρίτων. Η κτήση της πραγματοποιείται είτε με την θέληση του προηγούμενου κύριου, είτε με χρησικτησία, είτε με άλλους ειδικούς τρόπους (προσκύρωση, απαλλοτρίωση, επιδίκαση δικαστηρίου κλπ).

Τα βασικότερα είδη κυριότητας είναι:

- η **πλήρης**, δηλαδή η κυριότητα όταν δεν έχει αφαιρεθεί από αυτήν η επικαρπία,
- η **σύνθετη** οριζόντιας ή και κάθετης ιδιοκτησίας,
- η **συγκυριότητα**, η ψιλή κυριότητα, δηλαδή η κυριότητα όταν έχει αφαιρεθεί η επικαρπία και τέλος
- η **μετακλητή**, δηλαδή η κυριότητα που τελεί υπό αίρεση ή προθεσμία.
- Ως ειδικά είδη κυριότητας αναφέρονται το **timesharing και το leasing.**

Σημαντικό είναι να αναφερθεί, ότι εντός ενός νομικά αυτοτελούς και αδιαίρετου γεωτεμαχίου δύναται να δημιουργηθούν πολλά επιμέρους νομικά αυτοτελή ακίνητα, τα οποία συμμετέχουν εξ αδιαιρέτου στο ενιαίο γεωτεμάχιο.

1.5 Αξία Ακινήτου

Ένα θέμα που απασχολεί διαχρονικά την ανθρώπινη σκέψη είναι η αξία της γης και οι τρόποι της σωστής εκτίμησής της. Εκπρόσωποι κοινωνικών και φιλοσοφικών ρευμάτων ασχολήθηκαν με την έννοια αυτή. Πιο συγκεκριμένα ο Αριστοτέλης στα «Πολιτικά» διαχωρίζει την αξία γης, σε αξία χρήσης και ανταλλακτική αξία. Ο Adam Smith (1723-1790), ασχολήθηκε βάζοντας μέτρο στις αξίες, ενώ ο D.Ricardo (1772-1823) διατύπωσε τη θεωρία της γαιοπροσόδου και της φθίνουσας γαιοπροσόδου. Ο Karl Marx (1818-1883) ακολούθησε με την θεωρία της υπεραξίας με βασική θέση ότι το έδαφος δεν είναι προϊόν και συνεπώς δεν έχει αξία. Ο P.J. Proudhon (1809-1865) πρόβαλε την θέση του ότι η ιδιοκτησία είναι κλοπή, ενώ ο η οικονομική σκέψη του J.S. Mill (1806-1873) ανέπτυξε τη θεωρία του κέρδους και της φυσικής αξίας (Ζεντέλης, 2015).

Με τον όρο **Αξία Ακινήτου** προσδιορίζουμε την αξία της γης καθώς και την αξία των ακίνητων κατασκευών που βρίσκονται πάνω ή κάτω από το έδαφος του

ακινήτου. Η αξία του ακινήτου προκύπτει αθροιστικά από τους συντελεστές που τη δημιουργούν, δηλαδή αξία γης συν αξία εργασίας συν κεφαλαίου και επιχειρηματικότητας. Δεν πρέπει να λησμονείται ότι η γη είναι ένα πεπερασμένο αγαθό αλλά όχι πεπερασμένης συνολικής αξίας, κυρίως λόγω των βελτιώσεων που μπορούν να πραγματοποιήσουν σε αυτή οι ανθρώπινες παρεμβάσεις αλλά και λόγω της αυξανόμενης διαχρονικά στενότητάς της. Η δημιουργούμενη κάθε φορά υπεραξία των ακινήτων αποτελεί προωθητική δύναμη και η προσπάθεια δημιουργίας της αποτελεί βασικό κίνητρο στη διαδικασία της ανάπτυξης (Ζεντέλης, 2015).

Ο προσδιορισμός της αξίας ενός ακινήτου εξαρτάται από πολλούς και πολλές φορές αστάθμητους παράγοντες, όπως τα χαρακτηριστικά του δομημένου περιβάλλοντος και της γειτονιάς στην οποία βρίσκονται οι κατοικίες. Τα δίκτυα υποδομής, τα πολεοδομικά δεδομένα της περιοχής, τα δομικά χαρακτηριστικά των ακινήτων, η θέα κτλ. διαδραματίζουν σημαντικό ρόλο στη διαμόρφωση των αξιών τους (Παπαευθυμίου, 2013). Ο Ζεντέλης (2015) υποστηρίζει ότι ο προσδιορισμός της αξίας ενός ακινήτου είναι μια διαδικασία που βρίσκεται μέσα στα όρια που θέτει η ανοιχτή αγορά και προκύπτει τελικά χωρίς υποκειμενικές εκτιμήσεις στα πλαίσια της ανταγωνιστικής αγοράς. Σύμφωνα με την Παγουρτζή (2003) η αξία των ακινήτων εξαρτάται από τέσσερις ανεξάρτητους οικονομικούς παράγοντες: την χρησιμότητα ή ικανότητα να ικανοποιεί τις διαμορφούμενες ανάγκες των ιδιοκτητών, την σπανιότητα, την επιθυμία για κάλυψη αναγκών και την αγοραστική δύναμη.

1.5.1 Οι Τρείς Διαστάσεις του Ακινήτου

Το ακίνητο έχει ουσιαστικά τρείς διαστάσεις: την **τεχνική,** τη **νομική** και την **οικονομική** διάσταση. Και οι τρείς αυτές διαστάσεις λαμβάνονται υπόψη ώστε να καταλήξουμε σε ασφαλή συμπεράσματα όσο αφορά την αξία του.

Η τεχνική φύση του ακινήτου χαρακτηρίζεται από έννοιες όπως η μοναδικότητα, η θέση του, η οικοδομική του άδεια, η παλαιότητα - ηλικία, οι τυχόν ειδικές κατασκευές και οι πιθανές αυθαιρεσίες πάνω σε αυτό. Ακίνητο με αναστρέψιμες αυθαίρετες κατασκευές είναι ακίνητο με μειωμένη ουσιαστικά αξία λόγω του απαιτούμενου κόστους αποκατάστασης, ενώ ακίνητο με μη αναστρέψιμες αυθαίρετες κατασκευές είναι ακίνητο ουσιαστικά μη αξιολογήσιμο.

Το ακίνητο ακολουθεί τις νομικές δεσμεύσεις του, οι οποίες επηρεάζουν την αξία του. Τέτοιες νομικές δεσμεύσεις είναι η συνιδιοκτησία, η δουλεία, η οριζόντια και κάθετη ιδιοκτησία και γενικά τα νομικά βάρη επ’ αυτού. Η νομική διάσταση του

ακινήτου το επηρεάζει ουσιαστικά. Για παράδειγμα ακίνητο με άκυρο τίτλο, είναι πρακτικά ακίνητο χωρίς νομική υπόσταση.

Η αξία των ακινήτων εξαρτάται και από οικονομικούς παράγοντες, δηλαδή αυτά που πρέπει να διαθέτει το ακίνητο ώστε να προσφέρει στον χρήστη του ένα εισόδημα ή μια εξυπηρέτηση (διαμονή, χώρο εργασίας). Αυτοί οι οικονομικοί παράγοντες επηρεάζουν την αξία των ακινήτων καθώς έχουν σχέση με τη χρησιμότητα ή τη δυνατότητα του ακινήτου να ικανοποιεί τις ανάγκες των ιδιοκτητών ή χρηστών που διαμορφώνονται ακόμα και με τη μορφή άμεσης εξυπηρέτησης ή με τη μορφή ροής ρευστού.

Το πώς και πόσο επιδρά η χρησιμότητα ενός ακινήτου στη διαμόρφωση της αξίας του εξαρτάται από τα διάφορα χαρακτηριστικά του, όπως είναι το μέγεθος, η θέση του, η καλύτερη δυνατή χρήση που μπορεί να έχει, η ποιότητα κατασκευής, η πολεοδομική ή και νομική κατάσταση του. Σημειώνεται η σχέση αξίας με την σπανιότητα του καθορίζεται από τη σχέση ζήτησης – προσφοράς, καθώς σταθερή ζήτηση συνεπάγεται σπανιότητα προσφοράς και αύξηση της αξίας. Τέλος η αξία των ακινήτων έχει σχέση όχι μόνο με την επιθυμία για κάλυψη αντικειμενικών ή υποκειμενικών αναγκών του χρήστη αλλά και με την αγοραστική του δύναμη.
Η έννοια της αξίας των ακινήτων προσδιορίζεται διαχρονικά με διάφορα επίθετα και ουσιαστικά, που της προσδίδουν διαφορετικό περιεχόμενο. Έτσι έχουμε την :

- **Αγοραία – Εμπορική Αξία** (Market Value) που είναι « *το αποτιμηθέν ποσό στο οποίο πρέπει να ανταλλάσσεται ένα πάγιο στοιχείο κατά την ημέρα της εκτίμησης μεταξύ ενός πρόθυμου πωλητή και ενός πρόθυμου αγοραστή μετά από ένα λογικό χρόνο διαδικασίας αγοραπωλησίας σε μια συναλλαγή σε συνήθη συναλλακτική κίνηση, όπου κάθε πλευρά ενεργεί με γνώση, σύνεση και χωρίς πίεση*» (Ευρωπαϊκό Εκτιμητικό Πρότυπο 1 EVS, 2012) και η οποία σύμφωνα με τον R. Ratclift είναι η πλέον πιθανή τιμή πώλησης (Αποστολίδης & Μανομενίδης, 2014). Σύμφωνα με τον Καρανικόλα (2010), η αγοραία αξία των ακινήτων επηρεάζεται άμεσα ή έμμεσα από μια πληθώρα φυσικών, κοινωνικών και πολιτικών παραγόντων. Οι παράγοντες αυτοί είναι οι ίδιοι που διαμορφώνουν τον τρόπο λειτουργίας της αγοράς, ρυθμίζουν τον ανταγωνισμό, δημιουργούν τη στενότητα των ακινήτων και με αυτόν τον τρόπο επιδρούν στη σχέση ζήτησης-προσφοράς με αποτέλεσμα τον καθορισμό μιας συγκεκριμένης τιμής για το κάθε ακίνητο.

- **Τρέχουσα Αξία** (Current Value), ως τη τιμή που προκύπτει από τις αποκλίσεις μεταξύ της τιμής πώλησης και της αγοραίας αξίας του ακινήτου.
- **Εύλογη Αξία** (Fair Value), που είναι «*η τιμή που μπορεί να εισπραχθεί για την πώληση ενός παγίου στοιχείου ή για την μεταβίβαση μιας υποχρέωσης στα πλαίσια μιας ομαλής συναλλαγής μεταξύ πρόθυμων μερών κατά τον κρίσιμο χρόνο*» (ΔΛΠ16, Κανονισμός ΕΚ, αρ.1126/2008). Ουσιαστικά είναι το ποσό για το οποίο ανταλλάσσεται ένα περιουσιακό στοιχείο μεταξύ ενδιαφερόμενων μελών σε συνθήκες αγοράς, χωρίς να υπάρχει κάποιο ειδικό συμφέρον ή σχέση μεταξύ των εμπλεκόμενων προσώπων. Η ουσιαστική διαφορά μεταξύ της αγοραίας αξίας και τη εύλογης είναι ότι η αγοραία αναφέρεται πάντα στη βέλτιστη τιμή ενώ η εύλογη μπορεί να είναι και διαφορετική.
- **Πραγματική αξία**, ως τη μέση αξία που καθορίζουν οι διακυμάνσεις της προσφοράς και της ζήτησης σε μια μακρόχρονη χρονική περίοδο.
- **Αξία γης**, ως τη τιμή της μονάδας επιφάνειας, η οποία καθορίζεται από την πλήρη ανάπτυξη ή τη βέλτιστη χρήση, ανεξάρτητα αν η τελευταία έχει υλοποιηθεί και η οποία αυξάνεται μέχρι το όριο εκείνο για το οποίο η δόμηση γίνεται ασύμφορη.
- **Τιμή ζώνης** ή τιμή εκκίνησης ως την τιμή που ορίζεται από την Γενική Γραμματεία Δημόσιας Περιουσίας και είναι η ενιαία τιμή αφετηρίας την οποία έχει κάθε ακίνητο που βρίσκεται σε αυτήν τη ζώνη. Η τιμή ζώνης αναφέρεται στη συνολική αξία κτίσματος και οικοπέδου που αναλογεί σε ένα τετραγωνικό μέτρο επιφάνειας, καινούργιας κατοικίας ή διαμερίσματος στον Α’ όροφο της οικοδομής με πρόσοψη μόνο σε έναν δρόμο. Δηλαδή, η τιμή ζώνης είναι μικτή τιμή αξίας κτίσματος μαζί με την αξία του ποσοστού οικοπέδου που του αναλογεί.
- **Αντικειμενική αξία** ακινήτου, που λειτουργεί ως ένα αμάχητο φορολογικό τεκμήριο της αξίας ενός ακινήτου που εκφράζεται σε αξία ανά τετραγωνικό μέτρο. Οι αντικειμενικές αξίες εκδίδονται από το Υπουργείο Οικονομικών και διαφοροποιούνται ανάλογα με την τοποθεσία του ακινήτου. Σύμφωνα με το άρθρο 41 του Ν.1239/1982, εφαρμόστηκε η αντικειμενική μεθοδολογία προσδιορισμού της αξίας των ακινήτων, βασιζόμενη στη χρήση αντικειμενικών προκαθορισμένων κριτηρίων που προβλέπονται από την διοίκηση. Οι αντικειμενικές αξίες δεν ανταποκρίνονται παρ' όλα αυτά πάντα στην πραγματική αξία της συναλλαγής, αφού στην πράξη είναι χαμηλότερες ή και σε μερικές περιπτώσεις

υψηλότερες από την αγοραία αξία του ακινήτου, ενώ αποτελούν απλώς τεκμήριο για φορολογικούς λόγους, μη εμποδίζοντας τα μέρη να συμφωνήσουν σε υψηλότερο ή χαμηλότερο τίμημα καταβάλλοντας τον αναλογούν φόρο. Σημειώνεται ότι είναι δυνατόν εντός του ίδιου δήμου να υπάρχουν διαφορετικές τιμές ζώνης .

- **Τιμή ζήτησης** (asking price), ως τη τιμή την οποία όταν λάβει ο πωλητής θα ήταν ικανοποιημένος να αποχωριστεί από το ακίνητό του. Στο real estate η τιμή αυτή είναι γνωστή και ως τιμή καταλόγου (listing price).
- **Τιμή προσφοράς** (offer price), ως τη τιμή όπου ο αγοραστής θα ήταν ευχαριστημένος να αποκτήσει το ακίνητο και συνήθως είναι χαμηλότερη από την τιμή ζήτησης. Και οι δύο μαζί αποτελούν την βάση της διαπραγμάτευσης μεταξύ αγοραστού και πωλητού για την αγορά του ακινήτου.
- **Υπολειμματική αξία** ως το καθαρό ποσό που προσδοκά να πάρει μια επιχείρηση για ένα πάγιο στοιχείο στο τέλος της ωφέλιμης ζωής του, φυσικά μετά την αφαίρεση των αναμενόμενων εξόδων διάθεσης.
- **Μισθωτική αξία** το εκτιμώμενο ποσό για το οποίο μια ιδιοκτησία μπορεί να μισθωθεί μεταξύ πρόθυμου μισθωτή και εκμισθωτή σε μια συναλλαγή εντός του πλαισίου της αγοράς.
- **Αξία επένδυσης** ως αξία ενός περιουσιακού στοιχείου με βάση μια μεμονωμένη επένδυση ή συγκεκριμένους επιχειρηματικούς στόχους.

1.6 Αναγκαιότητα Προσδιορισμού της Αξίας των Ακινήτων

Οι συναλλαγές που πραγματοποιούνται στην καθημερινή ζωή απαιτούν την γνώση της πραγματικής αξίας των ακινήτων σε μια δεδομένη χρονική στιγμή. Αυτή η γνώση εξυπηρετεί το πλήθος αναγκών που διαμορφώνονται από τη λειτουργία της αγοράς.

Για τον **δημόσιο τομέα** η ορθή αποτίμηση της αξίας των κατεχόμενων από αυτόν ακινήτων εξυπηρετεί την αποτελεσματική διαχείριση και αξιοποίηση τους, ενώ επιτυγχάνει την δικαιότερη φορολόγηση των ακινήτων που ανήκουν σε ιδιώτες, η οποία και πρέπει να γίνεται βάσει πραγματικών και όχι αντικειμενικών αξιών. Η ορθή αποτίμηση των ακινήτων του δημοσίου τομέα βοηθά στην αξιοποίηση των πόρων και στην βελτίωση των συνθηκών διαβίωσης των πολιτών, καθώς η δημόσια περιουσία αποτελεί μέσο εκπλήρωσης των σκοπών του κράτους. Επιπρόσθετα η ανάγκη άμεσου και ορθού προσδιορισμού της τιμής ενός ακινήτου προς αγορά από τον δημόσιο τομέα, είναι απαραίτητη για την κάλυψη των αναγκών του.

Τα **φυσικά πρόσωπα** που αποτελούν τον ιδιωτικό τομέα ενδιαφέρονται για την πραγματική αξία ενός ακινήτου, όταν συμμετέχουν σε αγοραπωλησίες, κατά την δανειοδότηση και την εγγραφή υποθήκης και κατά τον υπολογισμό της απόδοσης από την τοποθέτηση κεφαλαίων σε ακίνητα εισοδήματος. Ο ιδιωτικός τομέας ενδιαφέρεται για την αποτελεσματικότητα των επενδύσεων και την διευκόλυνση των συναλλαγών, καθώς πλήθος δραστηριοτήτων του συνδέεται άμεσα με την αξία συναλλαγής των ακινήτων. Άλλωστε οι επιχειρήσεις, είναι υποχρεωμένες να πληροφορούν τους ενδιαφερόμενους- μετόχους σχετικά με την αξία των περιουσιακών τους στοιχείων.

Το **τραπεζικό σύστημα** είναι ένας μοχλός ανάπτυξης μέσω των παρεχόμενων χρηματοδοτήσεων και δέχεται εμπράγματες εξασφαλίσεις για την δανειοδότηση επιχειρήσεων και ιδιωτών. Η γνώση λοιπόν της πραγματικής αξίας των εξασφαλίσεων αυτών και η αντίστοιχη αποτίμηση τους αποτελεί υποχρέωση των τραπεζικών ιδρυμάτων, στα πλαίσια των διεθνών προτύπων χρηματοοικονομικής πληροφόρησης. Ας μη ξεχνάμε ότι στις περισσότερες περιπτώσεις η διασφάλιση των πόρων γίνεται με δέσμευση ακινήτων, η ορθή γνώση της αξίας των οποίων αποτελεί βασικό στοιχείο λειτουργίας του τραπεζικού συστήματος.

1.7 Παράγοντες που Επηρεάζουν την Αξία των Ακινήτων

Γενικά είναι παραδεκτό ότι λόγω της αβεβαιότητας που αναπτύσσεται σε περιόδους οικονομικής αστάθειας, ύφεσης, νομισματικών κρίσεων ή πληθωρισμού, είναι προτιμότερη η κατοχή πραγματικών αξιών όπως τα ακίνητα αντί ρευστού. Άλλωστε τα ακίνητα θεωρούνται κατά κανόνα αγαθά σταθερής αξίας, που δεν επηρεάζονται εύκολα από τις μεταβολές άλλων οικονομικών μεγεθών. Αυτό υποδεικνύουν και οι αυξανόμενες ανάγκες, σε μια αυξανόμενη ζήτηση, ενός αυξανόμενου πληθυσμού. Τα φαινόμενα κερδοσκοπίας, με αυξήσεις ζήτησης και τιμών, πολλές φορές οφείλονται στη μεγάλη και αυξανόμενη ταχύτητα συναλλαγών που παρατηρείται στα ακίνητα και στην έλλειψη των απαιτούμενων ή στη μεταβλητότητα των υπαρχόντων θεσμών (Ζεντέλης, 2015).

Η αξία ενός ακινήτου θεωρητικά επηρεάζεται από άπειρους παράγοντες, οι οποίοι διαμορφώνουν τον τρόπο λειτουργίας της αγοράς, καθορίζοντας το βαθμό ανταγωνισμού, επηρεάζοντας την προσφορά και ζήτηση και προσδιορίζοντας μια συγκεκριμένη αγοραία αξία. Επηρεάζεται από ανεξάρτητους παράγοντες όπως είναι η χρησιμότητα, η σπανιότητα, η επιθυμία για κάλυψη αναγκών και η αγοραστική

δύναμη, ενώ η συμβολή τους στην τελική διαμόρφωση της αξίας των ακινήτων είναι δύσκολο να προσδιοριστεί, κυρίως λόγω του μεγάλου πλήθους των, της διαφορετικής συμπεριφοράς τους στο χρόνο ή και στο χώρο καθώς και της δυναμικής τους εξέλιξης.

Σε περίοδο οικονομικής κυρίως κρίσης, οι αξίες των ακινήτων επηρεάζονται κυρίως από τη στεγαστική πίστη, έχοντας ως αποτέλεσμα την εφαρμογή ολοένα και αυστηρότερων κριτηρίων από πλευράς των τραπεζών ως προς τη χορήγηση στεγαστικών δανείων. Όπως αναφέρει ο Χαρδούβελης (2009) επηρεάζονται από την οικονομική κατάσταση των νοικοκυριών, καθώς το διαθέσιμο εισόδημα τους, σύμφωνα με την Ευρωπαϊκή Κεντρική Τράπεζα, αποτελεί το σημαντικότερο παράγοντα που επηρεάζει τις τιμές των ακινήτων. Επηρεάζονται ακόμα από το φορολογικό νομοθετικό πλαίσιο της χώρας, καθώς οι αυξημένες φορολογικές υποχρεώσεις δυσκολεύουν την κατάσταση, επιδρώντας αρνητικά στην αγορά. Οι τιμές των μισθωμάτων προσδιορίζουν την πράξη αγοράς ενός ακινήτου ως επένδυση, αφού σε αρκετές από τις αγοραπωλησίες ακινήτων βασικός στόχος του αγοραστή δεν αποτελεί αναγκαία η κάλυψη κάποιας ανάγκης, αλλά η επένδυση επί του ακινήτου. Η επίδραση του πληθωρισμού είναι σημαντική καθώς κατέχει σπουδαίο ρόλο στη διαμόρφωση τιμών που συμβαδίζουν με αυτόν, ανεξάρτητα με τα χαρακτηριστικά του τραπεζικού συστήματος. Ένας ακόμα ουσιαστικός παράγοντας, είναι τα υπάρχοντα κάθε φορά αποθέματα κατοικιών, καθώς η ύπαρξη μεγάλων αποθεμάτων αδιάθετων κατοικιών λειτουργεί ως ανασταλτικός παράγοντας στην αύξηση των τιμών σε μια στάσιμη αγορά. Η οικοδομική δραστηριότητα και το κόστος κατασκευής είναι ένα ακόμα σημαντικό στοιχείο που συμμετέχει στη διαμόρφωση τιμών ενός ακινήτου, καθώς (όπως αναφέρει ο Τασιόπουλος, 2012) σύμφωνα με την Τράπεζα της Ελλάδος το αυξημένο κόστος κατασκευής «*εκτιμάται ότι λειτουργεί ανασταλτικά στην ανάληψη νέων επενδυτικών πρωτοβουλιών από την πλευρά των κατασκευαστών, καθώς αυξάνει το κόστος αντικατάστασης του διαθέσιμου αποθέματος των προς πώληση ακινήτων*». Η ψυχολογία των επενδυτών είναι ένας σημαντικός παράγοντας, καθώς η αγορά είναι ευαίσθητη στα συναισθήματα τους. Άλλωστε οι άνθρωποι είναι από τη φύση τους φτιαγμένοι για να δρουν ομαδικά επηρεαζόμενοι ο ένας από τον άλλο. Τέλος ο δημογραφικός παράγοντας, η σημαντικότητα του οποίου ασκεί επίδραση στην αξία των ακινήτων, διαδραματίζει μακροπρόθεσμα καθοριστικό ρόλο στη ζήτηση τους.

Επιπρόσθετα παράγοντες που επηρεάζουν την αξία ενός ακινήτου είναι και το

γενικότερο παγκόσμιο οικονομικό κλίμα, η εθνική οικονομία, η νομοθεσία και το φορολογικό πλαίσιο, η γεωγραφική τοποθεσία του ακινήτου, οι προοπτικές ανάπτυξης της περιοχής που βρίσκεται το ακίνητο, η εγγύς κτηματαγορά, τα ιδιαίτερα χαρακτηριστικά του ακινήτου, η κατάσταση του και η ηλικία του, η νομική τάξη του, το τρέχον μίσθωμα που πιθανά υπάρχει και τέλος η μόδα της εποχής (Καρανικόλας, 2010).

Οι βασικοί παράγοντες που επηρεάζουν την αξία των ακινήτων δύναται να ταξινομηθούν σε επίπεδο χώρας, πόλης, τμήματος πόλης και ακινήτου. Με την παραπάνω ταξινόμηση ερμηνεύεται η διαφοροποίηση της αξίας τους μεταξύ χωρικών επιπέδων της ίδιας τάξης μεγέθους, ενώ προσδιορίζεται στο εξεταζόμενο επίπεδο, θεωρώντας την δυνατότητα μεταβολής της ομάδας των παραγόντων. Η κατάταξη σε επίπεδα έχει ιεραρχικό χαρακτήρα, καθώς όσο μεγαλώνει το επίπεδο λεπτομέρειας, τόσο οι παράγοντες των προηγούμενων επιπέδων έχουν ήδη επιδράσει (Ζεντέλης, 2015 · Λαμπρόπουλος, 2017) . Οι παράγοντες που επιδρούν στη διαμόρφωση μιας πόλης, όπως η κατανομή του πληθυσμού, τα δίκτυα υποδομής, οι χρήσεις γης της περιοχής, το πολεοδομικά στοιχεία, τα χαρακτηριστικά θέσης, επιδρούν και στη διαμόρφωση της τιμής ενός ακινήτου. Σύμφωνα με τον Ζεντέλη (2015) η εκτίμηση των δομικών χαρακτηριστικών του ακινήτου, της θέσης του στο οικοδομικό τετράγωνο, της κατάστασης του περιβάλλοντα χώρου του, των πολεοδομικών δεδομένων της περιοχής, του κόστους δόμησης, των αστάθμητων παραγόντων, της προσβασιμότητας, του θορύβου, της γεωμετρίας και μορφής δόμησης, επιδρά στον καθορισμό της αξίας ενός ακινήτου.

1.7.1 Παράγοντες σε Επίπεδο Χώρας

Μεγάλος αριθμός παραγόντων επιδρά στα ακίνητα και στις αξίες τους, διαμορφώνοντας τις τιμές στην αγορά και διαφοροποιώντας αυτές από χώρα σε χώρα, όπως αναπτύσσεται και λειτουργεί ο πολεοδομικός ιστός της κάθε μιας. Σε κάθε κράτος διαμορφώνεται ένα διαφορετικό εύρος αξιών ακινήτων κάθε κατηγορίας, ενώ η αλλαγή στους παράγοντες αυτούς συνεπάγεται και ανάλογη μεταβολή τιμών. Σύμφωνα με τους Chin & Chau (2002) η οικονομική ευμάρεια ενός έθνους έχει αντίκτυπο στην αγορά ακινήτων.

Συνοπτικά λοιπόν οι βασικοί παράγοντες σε επίπεδο χώρας που επιδρούν στα ακίνητα και διαμορφώνουν τις αξίες αυτών ταξινομούνται σε: **Πολιτικούς** (πολιτικής γης, πολιτικής επενδύσεων σε ακίνητα, πολιτικής παρέμβασης στο μηχανισμό της

κτηματαγοράς, πολιτικής για τους οικονομικούς μετανάστες), **Κοινωνικούς** (κοινωνικής διαστρωμάτωσης, σύνθεσης πληθυσμού κατά ηλικία, πυκνότητα πληθυσμού, προτύπων αναγκών), **Οικονομικούς** (οικονομικής κατάστασης της χώρας, βιοτικού επιπέδου, κόστους κατασκευής οικοδομής – υλικών), **Φορολογίας ακινήτων**, **Χωροταξικούς**, **Πολιτιστικούς**, **Ιστορικούς** (ιστορική κληρονομιά και ιδιαιτερότητες), **Γεωγραφικούς** (θέση της χώρας ως γεωγραφική θέση).
Σημειώνεται ότι ο όρος ««Πολιτική Γης» σύμφωνα με τον Αρβανίτη (2009) αναφέρεται στην διατύπωση τεκμηρίωση και εφαρμογή ενός συνόλου κανόνων με κοινωνικοοικονομικές, νομικές και τεχνολογικές προεκτάσεις που υπαγορεύουν πως θα διαχειριστεί η γη και θα καρπωθούν τα οφέλη από αυτή.

1.7.2 Παράγοντες σε Επίπεδο Πόλης

Ένα πλήθος παραγόντων επιδρά, καθώς ο χαρακτήρας μιας πόλης μεταβάλλεται με τη διαδικασία ανάπτυξης και λειτουργίας της. Οι παράγοντες αυτοί διαμορφώνουν τις τιμές, επιδρούν και διαφοροποιούν τις αξίες μεταξύ διαφορετικών πόλεων της ίδιας χώρας, διαμορφώνοντας ένα διαφορετικό εύρος αξιών στις τιμές των ακινήτων κάθε κατηγορίας. Σημειώνεται ότι η αλλαγή των παραγόντων επηρεάζει τις τιμές ακόμα και στην ίδια την πόλη.

Πιο αναλυτικά, σε επίπεδο πόλης οι κύριοι παράγοντες που επηρεάζουν τις αξίες των ακινήτων ταξινομούνται στις κατηγορίες: **Ανάπτυξης** πόλης, **Πολεοδομικής οργάνωσης** (οικιστικής οργάνωσης, προοπτικών ανάπτυξης, δικτύων υποδομής, πολεοδομικού σχεδιασμού, και περιορισμών), **Δικτύων υποδομής** και παροχής υπηρεσιών, **Χωρικών διαφοροποιήσεων** (πολιτική γης και κίνητρων επενδύσεων), **Ρυθμού ένταξης** στο πολεοδομικό ιστό νέων εκτάσεων, **Κατανομής του πληθυσμού** και των επαγγελματικών δραστηριοτήτων του, **Εσωτερικής μετανάστευσης, Γεωλογικών μεταβολών, Μορφών και μεγέθους ρύπανσης**, **Ασφάλειας**, **Σεισμικής δραστηριότητας** της πόλης.

1.7.3 Παράγοντες σε Επίπεδο Τμήματος Πόλης

Μεγάλος αριθμός παραγόντων διαφοροποιεί τις αξίες των ακινήτων και εντός της ίδιας πόλης από τμήμα σε τμήμα, καθώς επιδρά ο διαχωρισμός της σε ζώνες ομοιόμορφης συμπεριφοράς βάσει χαρακτηριστικών όπως η διοικητική της δομή, η διαβάθμιση της ανάπτυξης, ο τύπος των χρήσεων, οι πολεοδομικές διαφοροποιήσεις. Σε επίπεδο τμήματος πόλης οι σημαντικότεροι παράγοντες που επιδρούν είναι: η **Ροή**

ανάπτυξης, τα **Πολεοδομικά** στοιχεία, **Είδος** και δραστηριότητες της περιοχής, **Δίκτυα** υποδομής**, Θέση** της περιοχής στην γενική διάταξη της πόλης, **Χρήσεις** γης, **Είδος δραστηριότητα**ς της περιοχής.

1.7.4 Παράγοντες σε Επίπεδο Ακινήτου

Καθώς η θέση κάθε ακινήτου είναι μοναδική επιδρά με διάφορους τρόπους στην αξία του. Έτσι λοιπόν οι παράγοντες που επηρεάζουν την αξία του είναι: η **Τοποθεσία** του, ο **Περιβάλλον Χώρος**, τα **Γενικά** χαρακτηριστικά του (όπως η μορφολογία, η κλίση και η σύσταση του εδάφους του, η ύπαρξη ρευμάτων και υπόγειων υδάτων), τα **Γεωμετρικά** χαρακτηριστικά του (μέγεθος οικοπέδου, σχήμα, μήκος πρόσοψης), τα **Ειδικά** χαρακτηριστικά του (όπως το πλάτος δρόμων και πεζοδρομίων που το περιβάλλουν), η **Εμπορικότητα** της περιοχής και η **Θέα**. Τα **Πολεοδομικά** δεδομένα, η **Μορφή**, ο **Τρόπος** και το **Κόστος** δόμησης του ακινήτου είναι και αυτοί παράγοντες που επηρεάζουν την αξία του. Τέλος δεν πρέπει να λησμονηθεί η αναφορά σε ειδικούς παράγοντες, όπως οι **Δεσμεύσεις** που μπορεί να έχει ένα ακίνητο λόγω ρυμοτομίας ή απαλλοτρίωσης, ο **Χαρακτηρισμός** κτιρίου ως διατηρητέου, η επίδραση του χρόνου στο ακίνητο ως **Ηλικία**, οι **Φθορές** και η αδυναμία πλήρους **Λειτουργικότητά**ς του και η επιρροή της **Νομικής** κατάστασης – ιδιοκτησιακού καθεστώτος του ακινήτου καθώς και η αναφορά στις γενικότερες προκύπτουσες **Ειδικές** συνθήκες της αγοράς που αφορούν τον αγοραστή ή τον πωλητή.

1.8 Ζήτηση και Προσφορά Ακινήτων

Σύμφωνα με τον Ζεντέλη (2015) η αγοραία αξία των ακινήτων διαμορφώνεται ως αποτέλεσμα των ωφελειών που προσφέρει το ακίνητο και του κόστους του, καθοριζόμενη με βάση το νόμο ζήτησης – προσφοράς. Η μελέτη του ανταγωνισμού της αγοράς στα ακίνητα δείχνει τι ανταγωνισμός επικρατεί στην διαθεσιμότητα συγκρινόμενων μεταξύ τους ακινήτων.

1.8.1 Ζήτηση

Σύμφωνα με το νόμο της ζήτησης, όταν όλοι οι άλλοι παράγοντες είναι σταθεροί, η τιμή είναι μικρότερη όταν η ποσότητα είναι μεγαλύτερη. Η **ζήτηση D (Demand)** αντιπροσωπεύει αυτό που επιθυμούν οι αγοραστές, οι οποίοι θα αγοράσουν σε όλες τις δυνατές τιμές σε μια δεδομένη χρονική περίοδο. Είναι γεγονός ότι η ζήτηση για

αγορά ακινήτων επηρεάζεται από παράγοντες, που μεταβάλλουν την τιμή μιας δεδομένης ποσότητας ή μεταβάλλουν την ποσότητα μιας ορισμένης τιμής, προκαλώντας έξαρση της προτίμησης για ορισμένες κατηγορίες ακινήτων, με διαφοροποίηση κατά περιοχές. Η ζήτηση λοιπόν προσδιορίζει διαφορετική τιμή κατά είδος ακινήτου και κατά περιοχή.

1.8.2 Προσφορά

Η προσφορά ακινήτων επηρεάζεται από παράγοντες όπως οι τιμές των υλικών, οι τιμές εργατικών, η έλλειψη ή υπερπροσφορά ακινήτων, τα επιτόκια δανεισμού κλπ., οι οποίοι μπορεί να προκαλέσουν την αύξηση ή μείωση της. Η πρ**οσφορά S (Supply)** αντιπροσωπεύει λοιπόν την ποσότητα ακινήτων που οι ιδιοκτήτες επιθυμούν να πωλήσουν σε δεδομένη χρονική περίοδο, σε όλες τις δυνατές τιμές.

1.8.3 Νόμος Ζήτησης – Προσφοράς

Στόχος είναι η ζήτηση και η προσφορά να βρεθούν σε μια ισορροπία, με την τιμή του ακινήτου να διαμορφώνεται στο σημείο τομής της. Η ισορροπία αυτή επιτυγχάνεται μέσα από διαδοχικές προσεγγίσεις σε μια χρονική περίοδο προσαρμογής. Οι τιμές που διαμορφώνονται στην περίοδο αυτή, προκαλούν ένα πλεόνασμα ζήτησης ή προσφοράς. Όταν οι παράγοντες επηρεασμού είναι σταθεροί, τότε και η τιμή της ισορροπίας είναι και αυτή σταθερή.

1.9 Μορφές Αγοράς

Η αγορά γενικά κατηγοριοποιείται σε τέσσερις ευρείας μορφής κατηγορίες. Πρώτη είναι η **ανταγωνιστική αγορά** ή τέλεια αγορά ανταγωνισμού. Δεύτερη είναι η **μονοπωλιακή αγορά**, τρίτη είναι η **μονοπωλιακή ανταγωνιστική** και τέλος η **ολιγοπωλιακή** αγορά (Μπακιρτζόγλου, 2012).

Όταν η ζήτηση ακινήτων ισορροπεί με την προσφορά, τότε έχουμε μια πλήρως ανταγωνιστική αγορά, ενώ όταν οι τιμές επηρεάζονται από διάφορους παράγοντες που άλλοτε εξουσιάζονται από την προσφορά και άλλοτε από την ζήτηση, τότε έχουμε μια μονοπωλιακή αγορά. Η ανταγωνιστική αγορά προϋποθέτει μεγάλο αριθμό αγοραστών και πωλητών ώστε να είναι αδύνατη η επιρροή στην τιμή του ακινήτου. Η ανταγωνιστική αγορά αφορά ομογενή αγαθά, πλήρη γνώση των συνθηκών της αγοράς και των εναλλακτικών λύσεων και γενικά ελεύθερη οικονομία. Η μονοπωλιακή αγορά είναι μια αγορά όπου η προσφορά προέρχεται από έναν μόνο

παράγοντα, ο οποίος την ελέγχει πλήρως και διαμορφώνει την τιμή των αγαθών. Στον μονοπωλιακό ανταγωνισμό έχουμε μια αγορά όπου υπάρχει ικανός αριθμός παρεμφερών προϊόντων (στην περίπτωσή μας ακινήτων), ενώ το ολιγοπώλιο είναι η αγορά όπου υπάρχει μικρός αριθμός αγοραστών και πωλητών, ενώ όσο μειώνεται ο αριθμός των πωλητών ή των αγοραστών, τόσο μεγαλώνει η διαπραγματευτική ισχύς τους αντίστοιχα.

1.10 Φορείς που Συμμετέχουν στην Αγορά Ακινήτων

Οι βασικοί φορείς που συμμετέχουν στην αγορά ακινήτων είναι ο **Επενδυτής** (αγοραστής), ο **Δανειστής** (π.χ. τράπεζες), ο **Ιδιοκτήτης** (πωλητής) και η **Πολιτεία** (συμβολαιογράφοι, δικηγόροι, μηχανικοί, φορείς τοπικής αυτοδιοίκησης, εφορεία, κλπ.). Κάθε παράγοντας έχει διαφορετικές και συχνά συγκρουόμενες θέσεις έναντι των άλλων. Οι συμβολαιογράφοι, δικηγόροι, μηχανικοί, είναι παράγοντες που έχουν προσωρινό και δευτερεύοντα ρόλο, ενώ άλλοι παράγοντες όπως μεσίτες, εκτιμητές, σύμβουλοι ακινήτων, εμπλέκονται δυνητικά στην αγοροπωλησία, έχοντας και αυτοί δευτερεύοντα ρόλο.

1.11 Η Ιδιαιτερότητα στη Φυσιογνωμία των Ελληνικών Ακινήτων

Η Ελλάδα είναι μια ιδιαίτερη χώρα και η ιδιαιτερότητα αυτή επιδρά και στην αγορά των ακινήτων της. Το ελληνικό οικογενειακό status χαρακτηρίζεται από την ακίνητη περιουσία του, ενώ σημαντικό χαρακτηριστικό είναι η προσπάθεια της ελληνικής οικογένειας να διατηρήσει την καλής της θέση στη κοινωνία, μέσα από τη διατήρηση και αύξηση των ακίνητων περιουσιακών της στοιχείων. Ένα ακόμα στοιχείο που χαρακτηρίζει την ελληνική κοινωνία είναι η μικροιδιοκτησία. Ο Έλληνας προτιμά τη διαμονή σε ένα μικρό ιδιόκτητο σπίτι παρά την ενοικίαση ενός μεγάλου, γεγονός που ενδυναμώνει την ιδιαιτερότητα αυτή. Επιπρόσθετα θα πρέπει να γίνει αναφορά στο ανένταχτο πνεύμα του έλληνα, που τον οδηγεί σε παραβατική συμπεριφορά, όσο αφορά υπερβάσεις της δόμησης του ακινήτου του, άρνηση παραχώρησης σε κοινή χρήση του προς ρυμοτόμηση τμήματος της ιδιοκτησίας του, μη συμμόρφωσης με τους πολεοδομικούς κανόνες που έχει θεσπίσει η πολιτεία καθώς και το γεγονός ότι δύσκολα απευθύνεται σε ειδικούς (μεσίτες, εκτιμητές) για την ορθή εκτίμηση και προώθηση του προς πώληση ακινήτου του.

1.12 Επισκόπηση Παραμέτρων που Επηρεάζουν τις Τιμές Ακινήτων

Οι παράμετροι που επηρεάζουν κυρίως τις τιμές των ακινήτων είναι: τα **δομικά χαρακτηριστικά** τους, τα οποία περιλαμβάνουν τα στοιχεία που τα χαρακτηρίζουν άμεσα (αριθμό δωματίων, ηλικία, τζάκι, παρκινγκ, κλπ), η **προσβασιμότητα** σε αυτά (ευκολία στην πρόσβαση υπηρεσιών) και τέλος τα **περιβαλλοντικά χαρακτηριστικά** της γειτονιάς τους. Η ποιότητα της γειτονιάς που περιβάλλει μια κατοικία, έχει σχέση με το μέσο εισόδημα ή το ποσοστό ανεργίας. Στην αγορά ενός ακινήτου συνεκτιμούνται μεταβλητές όπως είναι η θέα, που αποτελεί αισθητικό πλεονέκτημα έχοντας σημαντική θετική επιρροή στις τιμές των κατοικιών πόλης από 1% έως και 147% για θέα σε θάλασσα (Μηλάκα 2010). Τα αστικά πάρκα είναι χώροι με θετική επιρροή στις τιμές γειτνιαζόντων σε αυτά κατοικιών, ενώ η ένταση της επίδρασης του αστικού πρασίνου αυξάνεται σε περιοχές κοντά στο εμπορικό κέντρο και εξασθενεί όσο η απόσταση από αυτό μεγαλώνει. Η αλληλεπίδραση του υδάτινου στοιχείου με το πράσινο εκτινάσσει την τιμή του ακινήτου. Η άμεση προσβασιμότητα σε διάφορες άλλες υπηρεσίες, όπως για παράδειγμα ένα εμπορικό κέντρο ή ένα σχολείο, είναι καθοριστικός παράγοντας στη διαμόρφωση των αξιών κατοικίας. Η γειτνίαση με ναό έχει αμφίρροπη επίδραση στην τιμή του ακινήτου, καθώς ρόλο μπορεί να παίξει η αρχιτεκτονική του οικοδομήματος και ο περιβάλλον χώρος αυτού ή ακόμα και η ηλικία των κατοίκων του ακινήτου.

Οι παράμετροι θα μπορούσαν να κατηγοριοποιηθούν σε δύο κατηγορίες, τις **γενικές** και τις **ειδικές** .

Συνοπτικά οι γενικές παράμετροι επίδρασης είναι :

- Η **θέση**, δηλαδή τα χαρακτηριστικά γνωρίσματα της περιοχής (ιστορικά, πολεοδομικά, γεωμορφολογικά, περιβαλλοντικά).
- Το **κοινωνικοοικονομικό** επίπεδο της περιοχής (κοινωνικές δομές και κοινωνική διαστρωμάτωση, σύνθεση πληθυσμού).
- Ο **πολεοδομικός σχεδιασμός** ο οποίος καθορίζει τους κοινόχρηστους χώρους, τα πλάτη οδών και πεζοδρομίων, το εάν υπάρχουν ή όχι προκήπια, τις χρήσεις, την εμπορικότητα, το μέγεθος και το σχήμα των οικοδομικών τετραγώνων, το μέγιστο επιτρεπόμενο ύψος των κτισμάτων.
- Η **οικιστική πυκνότητα**, καθώς υψηλές πυκνότητες οδηγούν σε έντονη συμφόρηση.
- Τα **δίκτυα μεταφορών**, καθώς εύκολη προσβασιμότητα σε μία περιοχή με τη χρήση της δημόσιας συγκοινωνίας, συνεπάγεται μεγαλύτερη αξία ιδιοκτησίας.

- Ο **συντελεστής εμπορικότητας**, δηλαδή η απόσταση από σημεία που συγκεντρώνονται οι υπηρεσίες για την εξυπηρέτηση των πολιτών.
- Η **περιβαλλοντική επιβάρυνση** στη ποιότητα ζωής.
- Το **αστικό πράσινο**.

Οι ειδικές παράμετροι επίδρασης είναι :

- Το **μέγεθος**, που είναι μια ιδιαίτερη παράμετρος επίδρασης, καθώς τα μικρότερα από το σύνηθες μέγεθος ακίνητα λαμβάνουν ένα θετικό συντελεστή που αυξάνει την τιμή τους ανά τετραγωνικό μέτρο, ενώ στην περίπτωση των μεγαλύτερων ακινήτων ο συντελεστής μεγέθους είναι αρνητικός και μπορεί να έχει ιδιαίτερα υψηλό βάρος επί της τελικής διαμόρφωσης της τιμής (Καρανικόλας, 2010).
- Η **παλαιότητα** του κτίσματος, με αρνητική συσχέτιση.
- Ο **όροφος**, με θετική επίδραση.
- Η **ποιότητα** κατασκευής της οικοδομής.
- Η κ**ατάσταση** του ακινήτου (βαθμός συντήρησης).
- Η **ποιότητα υλικών** κατασκευής, κυρίως στις νεόδμητες οικοδομές.
- Το **είδος θέρμανσης** που διαθέτει (ατομική ή κεντρική θέρμανση, πετρέλαιο, φυσικό αέριο, ή ηλεκτρικά σώματα, τζάκι, κλιματισμός).
- Ο **προσανατολισμός** του, ο οποίος έχει άμεση σχέση με την ενεργειακή απόδοση και την θέα.
- Η ύπαρξη θέσης **στάθμευσης** και **αποθήκης.**
- Οι υφιστάμενες **αυθαιρεσίες** επί του ακίνητου.

Επεξήγηση των παραμέτρων επίδρασης - μεταβλητών αναλυτικά θα γίνει στο επόμενο κεφάλαιο της παρούσας εργασίας.

1.13 Η Αγορά Ακινήτων στην Ελλάδα μέσα από τη Στατιστική Ματιά

Σημαντικό κομμάτι της ελληνικής οικονομίας είναι η αγορά ακινήτων. Η εσωτερική μετανάστευση με την μετακίνηση πληθυσμού από την επαρχία στις μεγάλες πόλεις τις δεκαετίες του 50, 60 και 70 αποτέλεσε το εφαλτήριο για την έκρηξη της αγοράς ακινήτων. Η έντονη ανάπτυξη της οικονομίας έδωσε το χρήμα, ενώ τις δεκαετίες που ακολούθησαν, το 80 και 90, η αγορά των ακινήτων παρουσίασε κάμψη καθώς παρουσιάστηκαν υψηλά ποσοστά πληθωρισμού, υψηλές αποδόσεις κρατικών ομολόγων, υψηλά επιτόκια στεγαστικών δανείων καθώς και μια γενικότερη οικονομική ύφεση. Οι αρχές της δεκαετίας του 2000 η αγορά ακινήτων διαδραμάτισε

κεντρικό ρόλο στην ελληνική οικονομία, αποτελώντας το αποθετήριο των χρημάτων που είχαν παραχθεί από άλλες επενδύσεις (επιχειρήσεις, κέρδη χρηματιστηρίου, κλπ) και έπρεπε να επανεπενδυθούν. Στατιστικές αποδεικνύουν ότι τα κεφάλαια που έχουν επενδυθεί σε ακίνητα είναι πολλαπλάσια αυτών σε κρατικούς τίτλους. Το pick των αγοροπωλησιών στον ελλαδικό χώρο έγινε το 2005 με 215.148 πράξεις αγοροπωλησίας ακίνητων. Από το έτος αυτό και μετέπειτα, η οικονομική κρίση επηρέασε και την αγορά ακινήτων. Τα τελευταία χρόνια και μετά από μια δεκαετία πτωτικής πορείας, η αγορά ακινήτων δείχνει και πάλι αμυδρά σημάδια ανάκαμψης. Συγκεκριμένα, σύμφωνα με στοιχεία της ΕΛΣΤΑΤ, το 2016 έγιναν 60.540 αγοροπωλησίες ακινήτων έναντι 54.631 το 2015 και 43.443 το 2014 στο σύνολο της χώρας (Ελληνική Στατιστική Αρχή, 2016). Στο γράφημα 1.1 αποτυπώνεται ο αριθμός συναλλαγών στην αγορά των ακινήτων όλης της χώρας από το 2000 και μετά.

Γράφημα 1.1 .Αριθμός συναλλαγών στην αγορά ακινήτων στο σύνολο της χώρας.
Πηγή: Ελληνική Στατιστική Αρχή, 2016

1.13.1 Πίεση στις Τιμές Ακινήτων

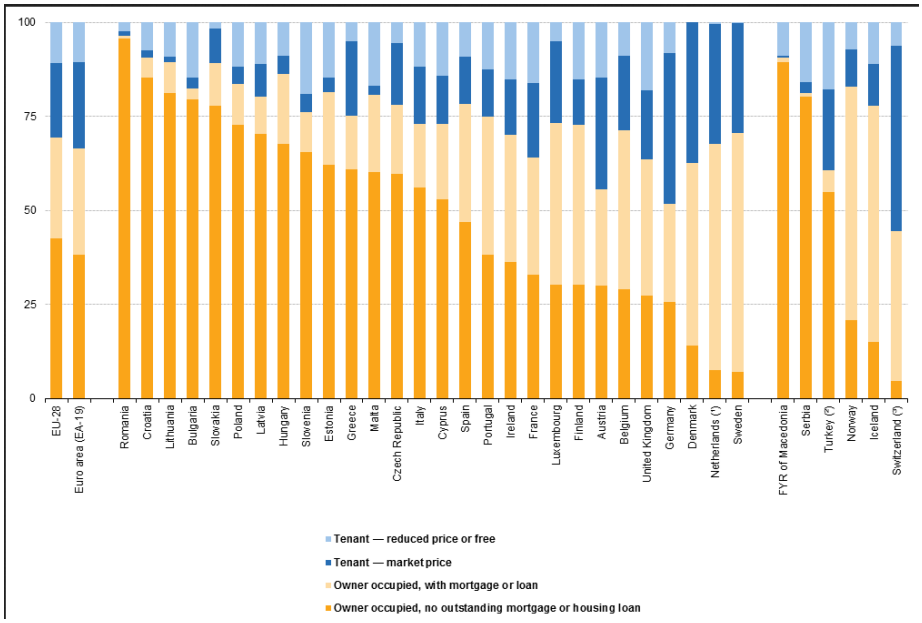
Σύμφωνα με την έκθεση του Διοικητή της Τράπεζας της Ελλάδος κ. Στουρνάρα για το έτος 2016, καταγράφηκε σταθερά επιδείνωση από το 2008 και μετά, τόσο στις εμπορικές αξίες και τα μισθώματα, όσο και στα επίπεδα ζήτησης οικιστικών και επαγγελματικών ακινήτων. Οι πιέσεις στις εμπορικές αξίες και τα μισθώματα των οικιστικών και των επαγγελματικών ακινήτων συνεχίστηκαν και κατά τη διάρκεια του 2016, αν και με μικρότερη ένταση σε σχέση με τα προηγούμενα έτη. Οι πιέσεις αυτές οφείλονται στην δυσμενή οικονομική συγκυρία. Σύμφωνα με την ίδια έκθεση, η αγορά των κατοικιών χαρακτηρίζεται από υπερβάλλουσα προσφορά, εξαιρετικά περιορισμένο αριθμό συναλλαγών και πτωτική τάση των τιμών. Στον τομέα των

επαγγελματικών ακινήτων το επενδυτικό ενδιαφέρον παραμένει, σύμφωνα με την έκθεση, ακόμα σε χαμηλά επίπεδα , ενώ στον τομέα των τουριστικών ακινήτων το έτος 2016 αποτέλεσε το ποιο ανθεκτικό και δυναμικό τμήμα της αγοράς, γεγονός συνυφασμένο με την καλή πορεία του τουρισμού.

Σύμφωνα με πανελλαδική έρευνα της REMAX μεσιτικής για το έτος 2016, το 85% των πωλημένων ακινήτων ήταν κατοικίες, το 6% οικόπεδα και το 9% επαγγελματικά ακίνητα. Εστιάζοντας στα οικιστικά ακίνητα το 40% των κατοικιών είχαν ηλικία άνω των 30 ετών, το 7% είχαν ηλικία από 26 έως 30 έτη, το 6% από 21 έως 25 έτη, το 10% από 16 έως 20 έτη, το 12% από 11 έως 15 έτη, το 17% από 6 έως 10 έτη, το 4% από 1 έως 5 έτη και τέλος το 4% αφορούσε νεόδμητες κατοικίες.
Με βάση το εμβαδόν, το 6% των κατοικιών ήταν πάνω από 171τμ, το 2% είχαν εμβαδόν από 151 έως 170τμ, το 5% εμβαδόν από 131 έως 150τμ, το 8% από 111 έως 130τμ, το 14% από 91 έως 110τμ, το 23% από 71 έως 90τμ, το 19% από 51 έως 70τμ και τέλος το 23% από 1 έως 50τμ. (REMAX Ελλάς, 2016).

1.13.2 Ελλάδα και Ιδιοκατοίκηση

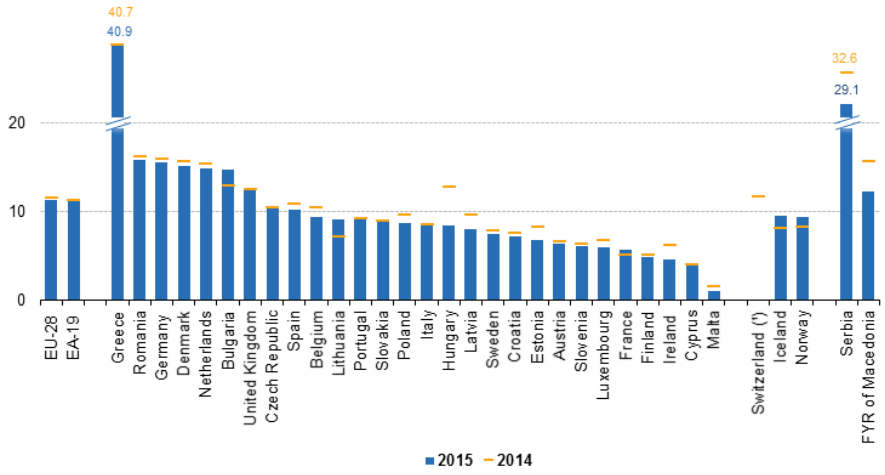
Σύμφωνα με στοιχεία της EUROSTAT ο Έλληνας βρίσκεται στην 14η θέση κρατών μελών με τα υψηλότερα ποσοστά ιδιοκατοίκησης στην Ευρώπη (περίπου 74% για το 2015, σύμφωνα με στοιχεία της European Commission) ενώ στην πρώτη θέση βρίσκεται η Ρουμανία με 96,1% (Eurostat, 2015). Στο γράφημα 1.2 αποτυπώνεται η κατανομή πληθυσμού ανά καθεστώς ιδιοκτησίας για το έτος 2015 στην Ευρώπη.



Γράφημα 1.2. Κατανομή πληθυσμού ανά καθεστώς ιδιοκτησίας για το έτος 2015 στην Ευρώπη. Πηγή: Eurostat ,2015.

1.13.3 Ελλάδα και Κόστος Κατοίκησης

Εξαιρετικά ενδιαφέροντα είναι τα στοιχεία της EUROSTAT για το κόστος κατοίκησης, ως ποσοστό του συνολικού διαθέσιμου εισοδήματος. Αυτό συνδέεται άμεσα με τις μηνιαίες δαπάνες (ΔΕΚΟ και φόρων) που καταβάλλει το κάθε νοικοκυριό. Με ποσοστό 40,7% δαπάνης επί του διαθέσιμου εισοδήματος τους, τα ελληνικά νοικοκυριά βρίσκονται στη πρώτη θέση μεταξύ των κρατών-μελών της Ευρωπαϊκής Ένωσης, ενώ τα γερμανικά νοικοκυριά βρίσκονται στη τρίτη θέση και δαπανούν μόλις το 15,9% (Ευρωπαϊκή Στατιστική Υπηρεσία, 2015). Στο γράφημα 1.3 αποτυπώνεται το ποσοστό δαπάνης επί του διαθέσιμου εισοδήματος για το έτος 2014-2015.



Γράφημα 1.3. Ποσοστό δαπάνης επί του διαθέσιμου εισοδήματος για το έτος 2014-2015. Πηγή: Eurostat, 2015.

1.13.4 Ελλάδα και Έλλειψη Κατοικιών

Στην Ελλάδα καταγράφεται ένα από τα υψηλότερα ποσοστά έλλειψης χώρων κατοικιών σε σχέση με τις ανάγκες των νοικοκυριών (27,4%), με τον μέσο όρο στην Ευρωπαϊκή ένωση να είναι 17,1% (Ελληνική Στατιστική Αρχή, 2015).

1.13.5 Ελλάδα και Αριθμός Κατοικιών

Το σύνολο των κατοικιών στον ελλαδικό χώρο, σύμφωνα με την καταγραφή της ΕΛΣΤΑΤ του 2011, ανέρχεται σε 6.384.353 κατοικίες, από τις οποίες 6.371.901 κανονικές κατοικίες (ποσοστό 99,8%) και 12.452 μη κανονικές κατοικίες (ποσοστό 0,2%), ενώ το μεγαλύτερο ποσοστό κανονικών κατοικιών 22,6% κατασκευάστηκαν την περίοδο 1971 έως 1980 και ποσοστό 96,1% των κανονικών κατοικιών βρίσκεται σε αστικά κέντρα. Από τις κανονικές κατοικίες της χώρας το 64,7% είναι

κατοικούμενες και το 35,3% κενές (Ελληνική Στατιστική Αρχή, 2015).

1.14 Εκτίμηση Αξιών Ακινήτων

Ως ««**Εκτίμηση Αξιών Ακινήτων**»« ορίζεται η διαδικασία προσδιορισμού της αξίας εντός περιουσιακού στοιχείου-ακινήτου, σε δεδομένη χρονική στιγμή, βάση μιας συγκεκριμένης μεθοδολογίας και για ένα συγκεκριμένο σκοπό. Ένας ευρύτερος ορισμός που θα μπορούσε να δοθεί, είναι ότι εκτίμηση είναι η τέχνη ή επιστήμη αποτίμησης της αξίας ενός περιουσιακού στοιχείου, σε μια δεδομένη χρονική στιγμή, λαμβάνοντας υπόψη ιδιαίτερα χαρακτηριστικά του υπό εκτίμηση προϊόντος, τους τρέχοντες οικονομικούς δείκτες της αγοράς και τις εναλλακτικές μορφές επενδύσεων. Αναφορικά με την επιστήμη, αυτή στηρίζεται σε μαθηματικά μοντέλα που χρησιμοποιεί ο εκτιμητής για τους υπολογισμούς του, ενώ όσο αφορά την τέχνη, αυτή βασίζεται στο γεγονός ότι ο εκτιμητής, μέσα από ένα μεγάλο δείγμα στατιστικών στοιχείων, επιλέγει αυτά που θα τον βοηθήσουν να καταλήξει σε ασφαλή συμπεράσματα (Μπακιρτζόγλου, 2012)

Η εκτίμηση λοιπόν είναι η επιστήμη της αποτίμησης της αξίας ενός κινητού ή ακίνητου περιουσιακού στοιχείου (στην εξεταζόμενη μελέτη, ακινήτου περιουσιακού στοιχείου) σε συγκεκριμένο χρόνο, με βάση τα οικονομικά στοιχεία της αγοράς, τις παράμερους επηρεασμού της μεταβολής της αξίας, διαθέτοντας την αναγκαία τεχνική πληροφόρηση που αφορά το προς εκτίμηση προϊόν, λαμβάνοντας υπόψη τα διεθνή και ευρωπαϊκά εκτιμητικά πρότυπα και τις μεθόδους εκτίμησης. Ο εκτιμητής είναι αυτός που αναμένεται να εκτελέσει τις εργασίες αποτίμησης αρμοδίως και με τρόπο ανεξάρτητο, αμερόληπτο και αντικειμενικό. (Uniform Standards of Professional Appraisal Practice, www.brienzaappraisal.com, 2013). Ο εκτιμητής θα συμβουλεύει τον πελάτη για πιθανές επενδυτικές κινήσεις του, για να προσημειωθεί ένα ακίνητο, για δικαστικές αποφάσεις, για να οριστεί τιμή μονάδας αποζημίωσης, για λογιστικές καταστάσεις, για ασφαλιστικούς σκοπούς, για συγχωνεύσεις εταιριών κλπ.

Εκτιμήσεις μπορεί να ζητήσουν ιδιώτες, επενδυτικοί και τραπεζικοί οργανισμοί, ιδιωτικές εταιρίες, δικηγόροι, εταιρίες ορκωτών λογιστών και δημόσιοι οργανισμοί. Για να καταλήξει ένας εκτιμητής σε ασφαλή συμπεράσματα πρέπει να συνυπολογίσει την τεχνική, οικονομική και νομική διάσταση ενός ακίνητου. Η International Valuation Standards Committee- IVSC αποτελεί τον φορέα που εκδίδει τα Διεθνή Εκτιμητικά Πρότυπα και αποτελεί έναν ανεξάρτητο μη κερδοσκοπικό ιδιωτικό οργανισμό που εκπροσωπείται σε 52 χώρες έχοντας ως κύριο σκοπό την

τήρηση των διεθνών εκτιμητικών προτύπων.

1.15 Είδη και Μέθοδοι Αποτίμησης Αξιών Ακινήτων

Τα **συνηθέστερα είδη εκτιμήσεων ακινήτων** που καλείται να διεκπεραιώσει ένας εκτιμητής αφορούν:

- Την αγορά με ενυπόθηκο δανεισμό,
- την πώληση ακινήτου μετά από αίτημα του πωλητή ή διαχειριστή ακινήτου,
- τη μίσθωση ακινήτου,
- τη χορήγηση καταναλωτικού δανείου από χρηματοπιστωτικό ίδρυμα, με εξασφάλιση του ακινήτου,
- την ανέγερση ή και την αποπεράτωση ακινήτου,
- τη βελτίωση ή στις επισκευές επί ακινήτου,
- τις διαδικασίες Leasing ή Sale & Lease Back επί ακινήτου δίνοντας τη δυνατότητα σε επιχειρήσεις να πουληθεί ένα ακίνητο σε μια εταιρία leasing και στην συνέχεια να το μισθώσουν από αυτή,
- τη περιοδική επανεκτίμηση όλων των ακινήτων με ευμετάβλητη εμπορική αξία (επαγγελματικά, ημιτελή κλπ.) από χρηματοπιστωτικό ίδρυμα,
- την αυτοτελή επανεκτίμηση ενυπόθηκου ακινήτου,
- τις διαδικασίες εγγραφής προσυμφωνημένης προσημείωσης σε ακίνητο ιδιοκτήτη μη εξυπηρετούμενου δανείου,
- τις διαδικασίες πλειστηριασμού ακινήτων,
- την απλή εκτίμηση για ενημέρωση εμπορικής αξίας ακινήτου,
- την απλή εκτίμηση για ενημέρωση, χαρτοφυλακίου ακινήτων,
- τη προεκτίμηση αξίας ακινήτου,
- την ασφάλιση ακινήτου για φυσικούς κινδύνους και καταστροφές (Λαμπρόπουλος, 2013).

Για να θεωρηθεί επιτυχημένη μια εκτίμηση θα πρέπει ο εκτιμητής να επιλέξει την σωστή μέθοδο. Καθώς υπάρχουν διάφορες μέθοδοι εκτίμησης αξίας ακινήτου, η επιλογή της καταλληλότερης εξαρτάται από την φύση του ακινήτου και τις προοπτικές του, το λόγο για τον οποίο συντάσσεται η εκτίμηση, τα στοιχεία που έχει στην διάθεσή του ο εκτιμητής και τέλος τις παραδοχές που θα αναγκαστεί να κάνει. Η επιλογή λοιπόν σχετίζεται με τον σκοπό της εκτίμησης και των χαρακτηριστικών του υποκείμενου ακινήτου. Οι παραδοσιακές μέθοδοι αποτίμησης στηρίζονται σε κάποιας

μορφής σύγκριση για τον προσδιορισμό της αγοραίας αξίας και αυτό μπορεί να γίνει στην απλούστερη του μορφή με μία άμεση σύγκριση κεφαλαίου ή με τον συσχετισμό παρατηρήσεων που επιτρέπουν την δημιουργία ενός μοντέλου παλινδρόμησης (regression model).

Οι βασικότερες μέθοδοι εκτίμησης αξίας ενός ακινήτου είναι η συγκριτική, η υπολειμματική μέθοδος, η εισοδηματική, η μέθοδος άμεσης κεφαλαιοποίησης, η μέθοδος προεξόφλησης ταμειακών ροών, η προσαρμοσμένης παρούσας αξίας, η μέθοδος του κόστους, η μέθοδος του υπολειμματικού κόστους αντικατάστασης, η ηδονική μέθοδος, η μέθοδος της πολλαπλής γραμμικής παλινδρόμησης (στατιστική μέθοδος), των νευρωνικών δικτύων και η μέθοδος της ασαφούς λογικής .

Αναλυτικότερα:

- Η **συγκριτική μέθοδος ή μέθοδος συγκριτικών στοιχείων (comparison approach or method) ή μέθοδος της κτηματαγοράς** είναι η παλαιότερη και πιο διαδεδομένη μέθοδος διεθνώς, η οποία αποτιμά κυρίως την εμπορική αξία ή το μίσθωμα ενός ακινήτου, βασίζεται στη θεωρία ότι ένας σώφρον αγοραστής δεν θα πλήρωνε περισσότερο για ένα ακίνητο από όσο θα του κόστιζε να αγοράσει ένα εναλλακτικό παρόμοιο και συγκρίσιμο ακίνητο. Είναι η πιο βασική μέθοδος θεωρούμενη ως η μητέρα όλων των μεθόδων εκτίμησης, ενώ συνήθως εφαρμόζεται ως έλεγχος μετά από άλλες μεθόδους (Καρανικόλας, 2010). Ουσιαστικά στηρίζεται στην ανάλυση πρόσφατων μεταβιβάσεων, πανομοιότυπων με το εκτιμώμενο ακινήτων. Τα στάδια της μεθόδου αυτής περιλαμβάνουν την έρευνα της αγοράς και την συλλογή στοιχείων και πληροφοριών σχετικά με τιμές, αξίες, διαθεσιμότητα ακινήτων, παρόμοιων χαρακτηριστικών σε παρόμοιες συνθήκες. Έπεται η ανάλυση και σύγκριση επί μέρους χαρακτηριστικών και η ολοκλήρωση γίνεται με την διατύπωση και παρουσίαση των αποτελεσμάτων από τις συγκρίσεις. Η συγκριτική μέθοδος εξαρτάται σε μεγάλο βαθμό από την διαθεσιμότητα, την ακρίβεια, την πληρότητα και την επικαιρότητα των δεδομένων της συναλλαγής. Όπως αναφέρει η Παγουρτζή κ.ά. (2003) οι πηγές πληροφόρησης μπορεί να περιλαμβάνουν κρατικά αρχεία, πωλητές δεδομένων και τοπικά δίκτυα εκτιμητών. Σημαντικό σημείο είναι ότι θα πρέπει να δοθεί ιδιαίτερη προσοχή κατά την εκτέλεση της μεθόδου αυτής στην αξιοπιστία των συγκριτικών στοιχείων. Η μέθοδος αυτή εφαρμόζεται συνήθως στην εκτίμηση ιδιοκτησιών κατοικιών, αγροτικής γης, αναπτυσσόμενης γης και αποτιμά κυρίως την αγοραία αξία του ακινήτου (Καρανικόλας, 2010).

- Η **υπολειμματική μέθοδος ή μέθοδος αξιοποίησης (residual method) ή μέθοδος αντιπαροχής,** η οποία χρησιμοποιείται κυρίως για εκτίμηση κενών οικοπέδων – γηπέδων όπου η αξία του ακινήτου ορίζεται ως η διαφορά της αξίας του, που διαμορφώνεται μετά την ολοκλήρωση των παρεμβάσεων για την ανάπτυξη του και του μέσου κόστους εργασιών, λαμβάνοντας υπόψη το κέρδος που θα αποκόμιζε ένας επενδυτής από την αγορά ενός συγκριτικού ακινήτου. Η γενική εξίσωση που απεικονίζει τη μέθοδο είναι:

Αξία ακινήτου =συνολικά έσοδα από την πώλησή του μετά το πέρας της ανάπτυξής του – κόστος κατασκευής του (Καρανικόλας, 2010).

Βασίζεται στο θεωρητικό σενάριο κατασκευής ενός κτίσματος βέλτιστης αξιοποίησης (highest and best use) σύμφωνα με την περιοχή που βρίσκεται η κενή γη. Χρησιμοποιείται κυρίως για τον προσδιορισμό μιας κενής γης και για την εφαρμογής της δημιουργείται ένα σενάριο δόμησης βέλτιστης χρήσης. Πιο ειδικά τα στάδια της μεθόδου αυτής είναι ο προσδιορισμός της βέλτιστης χρήσης, ο προσδιορισμός του κόστους κατασκευής, του εργολαβικού και επιχειρηματικού οφέλους, το κόστος χρηματοδότησης και το κόστος κατασκευής. Έπειτα υπολογίζεται το έσοδο από την πώληση βάση στοιχείων κτηματαγοράς και η αξία προκύπτει από την αφαίρεση της ποσότητας του συνόλου πωλήσεων με το κόστος κατασκευής, αλλά και των λοιπών εξόδων (πιθανός δανεισμός, επιτόκια, παράπλευρα έξοδα, κόστος δέσμευσης κεφαλαίων, εργολαβικό κέρδος κλπ.) ώστε να προσδιοριστεί η αξία της κενής γης .

- Η **εισοδηματική μέθοδος ή μέθοδος κεφαλαιοποίησης του προσόδου (investment method)** στηρίζεται στο γεγονός ότι κάποιες φορές ο σκοπός αγοράς ενός ακινήτου δεν είναι ούτε στεγαστικός ούτε επαγγελματικός, αλλά αυτός της μελλοντικής εκμετάλλευσης του. Βασίζει δηλαδή την αξία του ακινήτου στο εισόδημα που αυτό μπορεί να παράγει. Το βασικό χαρακτηριστικό αυτής της μεθόδου αποτελεί η μελλοντική της διάσταση, με τον εκτιμητή να βασίζεται σε μελλοντικές προσδοκίες, έχοντας υπόψη ότι τα ακίνητα που παράγουν εισόδημα συνήθως αγοράζονται ως επένδυση, ενώ από την οπτική γωνία του επενδυτή η δυνατότητα παραγωγής κερδών από το εκάστοτε ακίνητο αποτελεί τον κρίσιμο παράγοντα που καθορίζει την τιμή του. Η μέθοδος αυτή χρησιμοποιείται κυρίως για την εκτίμηση εμπορικών και επενδυτικών ακινήτων, βασιζόμενη στο ότι η σημερινή αξία ενός ακινήτου εξαρτάται από το μελλοντικό εισόδημα που μπορεί να αποδώσει και το συντελεστή κεφαλαιοποίησης ο οποίος προσδιορίζεται με βάση το πληθωρισμό, τα

τραπεζικά επιτόκια, την ετήσια απόσβεση κεφαλαίου, την επισφάλεια καταβολής μισθωμάτων και του επενδυτικού προφίλ του επενδυτή, ενώ προσαρμόζεται στις συνθήκες της αγοράς (Καρανικόλας, 2010) .

Χρησιμοποιούνται δύο μεθοδολογίες για την εφαρμογής της μεθόδου αυτής: της άμεσης κεφαλαιοποίησης και της μεθόδου προεξόφλησης μελλοντικών χρηματικών ροών.

- Η μέθοδος της **άμεσης κεφαλαιοποίησης (direct capitalization)** χρησιμοποιείται στο πλαίσιο της εισοδηματικής προσέγγισης, ώστε να μετατραπεί μία ετήσια χρηματική ροή σε αξία. Η άμεση κεφαλαιοποίηση αποτελεί μία μέθοδο που χρησιμοποιείται ευρέως όταν το ακίνητο που θα εκτιμηθεί λειτουργεί σε ένα πλαίσιο σταθεροποιημένης βάσης, με σημαντικότερο και πιο δύσκολο σημείο της μεθόδου το σωστό προσδιορισμό του συντελεστή κεφαλαιοποίησης. Η μέθοδος της άμεσης κεφαλαιοποίησης (μισθωμάτων) είναι αυτή που εφαρμόζεται σε ακίνητα τα οποία δύναται να παράγουν κεφάλαιο στον ιδιοκτήτη τους, θεωρώντας τη παραδοχή ότι το μίσθωμα θα λαμβάνεται για πάντα. Μέσω του υπολογισμού του μισθώματος και ενός εύλογου συντελεστή κεφαλαιοποίησης προκύπτει η αξία του ακινήτου μέσω του τύπου: μίσθωμα προς συντελεστής κεφαλαιοποίησης .
- Η μέθοδος της **προεξόφλησης ταμειακών (χρηματικών) ροών (discounted cass flows)** χρησιμοποιείται σε περιπτώσεις όπου δεν υπάρχουν συγκριτικά στοιχεία, ώστε να εκτιμηθεί η τιμή του αρχικού μισθώματος και να χρησιμοποιηθεί η μέθοδος της άμεσης κεφαλαιοποίησης. Με την εφαρμογή του κατάλληλου επιτοκίου μία σειρά από αναμενόμενες μελλοντικές εισροές καθορίζουν την παρούσα αξία του ακινήτου. Κατά την μέθοδο αυτή ο εκτιμητής κάνει μελλοντική πρόβλεψη του εισοδήματος που θα παράγει το ακίνητο, για κάποια δεδομένη χρονική περίοδο, λαμβάνοντας υπόψη ότι η οικονομία της αγοράς θα έχει σταθερό ρυθμό ανάπτυξης. Ουσιαστικά η μέθοδος αυτή υπολογίζει την παρούσα αξία των μελλοντικών εισοδημάτων από την μίσθωση του ακινήτου, αφαιρείται η παρούσα αξία των μελλοντικών εξόδων για συντήρηση, ασφάλιστρα κλπ. για ένα ορισμένο χρονικό διάστημα της επένδυσης και αθροίζονται με την προεξοφλημένη αξία μεταπώλησης στο τέλος του χρονικού διαστήματος που ορίστηκε εξ αρχής .

- Η μέθοδος της **προσαρμοσμένης παρούσας αξίας** όπως αναφέρει η Μηλάκα (2010), αναπτύχθηκε αρχικά από τον Myers και υποθέτει ότι ένα περιουσιακό στοιχείο έχει μία αξία κάτω από συνθήκες τέλειας αγοράς και ενδεχομένως μία πρόσθετη αξία που προκύπτει από τις ατέλειες της .

- Η μέθοδος του **κόστους ή μέθοδος κόστους αποκατάστασης,** βασίζεται στη θεωρία ότι κανένας αγοραστής δεν θα πλήρωνε παραπάνω από το ποσό που θα του κόστιζε να φτιάξει μόνος του ένα παρόμοιο ακίνητο. Στηρίζεται στο ότι η ζητούμενη αξία ενός ακινήτου προκύπτει, αν στο υποθετικό συνολικό κόστος κατασκευής ενός νέου κτιρίου προσθέσουμε την αξία της γης και στην συνέχεια αφαιρέσουμε την συνολική απαξίωση που υπέστη το προς εκτίμηση κτίριο λόγω παλαιότητας, ζημιών, φθορών σε σχέση με το νέο θεωρητικό κτίριο. Σε γενικές γραμμές η μέθοδος του κόστους θεωρείται κατάλληλη μέθοδος για την εκτίμηση νέων ακινήτων. Τα στάδια αυτής της μεθόδου είναι πρώτα ο υπολογισμός της αξίας γης (θεωρώντας την κενή), έπεται ο υπολογισμός του κόστους κατασκευής ενός παρόμοιου κτιρίου και ολοκληρώνεται με τον υπολογισμό της απαξίωσης λόγω παλαιότητας, η οποία και αφαιρείται .
- Η μέθοδος του **υπολειμματικού κόστους αντικατάστασης,** η οποία χρησιμοποιείται σε ειδικής κατηγορίας κτίρια, χωρίς εμπορική χρήση και όταν δεν υπάρχει διαμορφωμένη κτηματαγορά ώστε να αντληθούν πληροφορίες. Βασίζεται στην λογική ότι ένας επενδυτής δεν προτίθεται να καταβάλλει για την αποκατάσταση ενός ακινήτου περισσότερα χρήματα από το κόστος αναπαραγωγής ενός αντίστοιχου ακινήτου, με τα ίδια χαρακτηριστικά, στην ίδια θέση. Η μέθοδος αυτή εφαρμόζεται σε εκτίμηση ακινήτων που βρίσκονται σε τοποθεσίες που πρόκειται να αναπτυχθούν εξ αρχής, ιδιοκτησιών που τα κτίρια επ΄αυτών έχουν εξυπηρετήσει το σκοπό τους και δεν πρόκειται να συνεχίσουν να λειτουργούν με την ίδια χρήση (Καρανικόλας,2010).
- Η **ηδονική μέθοδος (hedonic price method),** βασίζεται στα ηδονικά μοντέλα και είναι μια οικονομετρική μέθοδος για την εκτίμηση των ακινήτων. Η θεωρία των ηδονικών μοντέλων (όπως αναφέρεται από την Μηλάκα, 2010), συγκροτήθηκε από τον Rosen (1974) και βασίστηκε στη θεωρία του Lancaster (1966), σύμφωνα με την οποία η χρηστικότητα ενός αγαθού είναι απόρροια των χαρακτηριστικών και των ιδιοτήτων του. Ο Rosen λοιπόν επέκτεινε την ιδέα στην αγορά κατοικίας και ανέπτυξε τη θεωρία των ηδονικών μοντέλων. Πιο συγκεκριμένα οι ηδονικές τιμολογιακές τεχνικές αποσκοπούν στο να προσδιορίσουν το πώς ορισμένα χαρακτηριστικά των ιδιοκτησιών επηρεάζουν τις τιμές πώλησης και διαμορφώνουν την αγορά γης. Η ηδονική μέθοδος θεωρεί πως κάθε ακίνητο στο σύνολό του είναι τελείως διαφορετικό από κάθε άλλο και αποτελείται από ένα σύνολο χαρακτηριστικών. Κάθε χαρακτηριστικό έχει συγκεκριμένη τιμή αγοράς,

ανάλογα με την χρησιμότητά για τους κατόχους του. Η αξία λοιπόν κάθε ακινήτου προσδιορίζεται από τον συνδυασμό ενός σταθερού αριθμού χαρακτηριστικών και η τελική τιμή είναι το σταθμικό άθροισμα της τιμής των επιμέρους χαρακτηριστικών του. Κάθε τύπος ακινήτου είναι λοιπόν δυνατόν να ερμηνευτεί ως διάνυσμα που αποτελεί συνισταμένη των επιμέρους χαρακτηριστικών. Ένα ηδονικό μοντέλο προσδιορίζει την αξία των επιμέρους χαρακτηριστικών μιας κατοικίας μέσω του προσδιορισμού ενός ολοκληρωμένου μοντέλου γραμμικής παλινδρόμησης. Με την κατάλληλη επιλογή των παραμέτρων που θα περιλαμβάνει και την παλινδρόμηση που θα περιγράψει καλύτερα τη σχέση μεταξύ της εξαρτημένης και των ανεξάρτητων μεταβλητών, εκτιμάται όσο το δυνατόν καλύτερα η αξία του ακινήτου . Η εξαρτημένη μεταβλητή στα ηδονικά μοντέλα είναι η αξία γης ανά ιδιοκτησία, ενώ οι ανεξάρτητες μεταβλητές, είναι οι παράγοντες που επηρεάζουν και καθορίζουν την εξαρτημένη, δηλαδή είναι το σύνολο των ιδιαίτερων χαρακτηριστικών του ακινήτου καθώς και των ιδιαιτεροτήτων ή των ιδιομορφιών της γειτονιάς του. Η αρχική ιδέα των ηδονικών μοντέλων παλινδρόμησης βασίζεται στην υπόθεση ότι οι παράμετροι του μοντέλου διατηρούνται σταθερές στο χώρο ακολουθώντας συναρτήσεις γραμμικής παλινδρόμησης (Μηλάκα, 2010) .

- Η στατιστική μέθοδος ουσιαστικά είναι μια μορφή συγκριτικής μεθόδου, όπου η αξία του ακινήτου προσδιορίζεται με ανάλυση **πολλαπλής γραμμικής παλινδρόμησης.** Η αξία του ακινήτου ως μεταβλητή προσδιορίζεται λαμβάνοντας υπόψη άλλες ανεξάρτητες μεταβλητές. Η διαδικασία αυτή προϋποθέτει ύπαρξη συγκριτικών στοιχείων, ενώ είναι πρακτικά μια αυτοματοποιημένη διαδικασία δημιουργίας μοντέλου που έχει ως στόχο, με την βοήθεια της στατιστικής και των βασικών αρχών εκτίμησης των ακινήτων, την ερμηνεία του τρόπου διαμόρφωσης της αξίας. Χρησιμοποιείται κυρίως σε μαζικές εκτιμήσεις ακινήτων ελέγχοντας την σημαντικότητα των συντελεστών παλινδρόμησης.
- Η μέθοδος των **νευρωνικών δικτύων (Artificial neural networks)** βασίζεται στην εκτίμηση της αξίας ενός ακινήτου, με την εισαγωγή δεδομένων που έχουν να κάνουν με τα ιδιαίτερα χαρακτηριστικά αυτού σε ένα σύστημα. Έπειτα γίνεται η επεξεργασία με την εφαρμογή των κατάλληλων σταθμίσεων και συναρτήσεων μετασχηματισμού και τέλος γίνεται η εξαγωγή του στρώματος εκτιμώμενης αγοραίας αξίας του ακινήτου. Ως μέθοδος πρόβλεψης χρησιμεύει σε μη γραμμικές συναρτήσεις. Τα νευρωνικά δίκτυα είναι τεχνικά συστήματα

προσομοίωσης του ανθρώπινου νευρικού συστήματος, όπου οι νευρώνες έχουν μια συγκεκριμένη διάταξη που οδηγεί σε μια δομή, η οποία ποικίλει στους διάφορους τύπους δικτύων. Όλοι οι τύποι δέχονται σήματα στην είσοδό τους τα οποία πολλαπλασιάζονται επί το αντίστοιχο βάρος, βρίσκουν το άθροισμα όλων των γινομένων και ακολούθως μεταβιβάζουν το άθροισμα αυτό σε μια ειδική συνάρτηση, η οποία παράγει την έξοδο από τον κάθε νευρώνα (Μίμης, Ροβολής & Στάμου, 2012). Ουσιαστικά τα νευρωνικά δίκτυα αποτελούν μια μορφή τεχνητής νοημοσύνης αποτελούμενα από ένα σύνολο απλών διασυνδεδεμένων και προσαρμοστικών μονάδων, που συνιστούν ένα παράλληλο υπολογιστικό μοντέλο, ενώ εκπαιδεύονται ώστε να αναγνωρίζουν και να επιτελούν μια συγκεκριμένη εργασία με την εκπαίδευση τους να γίνεται μέσω αλλαγής των βαρών τους.

- Η μέθοδος της **ασαφούς λογικής (Fuzzy logic)** αποτελεί μια γενίκευση της κλασσικής λογικής και λειτουργεί με μηχανισμούς προσεγγιστικού συλλογισμού εξάγοντας συμπεράσματα (Μαστοροκώστας, 2015). Επικεντρώνεται στην αναπαράσταση της ασάφειας, της ανακρίβειας και της αβεβαιότητας προσπαθώντας να πετύχει λύσεις, μιμούμενη την ικανότητα του ανθρώπινου μυαλού να λειτουργεί και να λαμβάνει αποφάσεις σ’ ένα τέτοιο περιβάλλον. Ουσιαστικά προσομοιώνει τον ανθρώπινο τρόπο σκέψης με σκοπό να εξάγει συμπεράσματα. Τα ασαφή σύνολα περιγράφονται με μαθηματικές αναπαραστάσεις, μέσω ενός υπολογιστικού συστήματος, όπου όλες οι τιμές εισόδου στο σύστημα μεταφράζονται σε γλωσσικές μεταβλητές. Αντίθετα με την δυαδική λογική ένα ασαφές σύνολο δεν έχει αυστηρά καθορισμένα όρια, επιτρέποντας ένα φυσικό τρόπο επεξεργασίας δεδομένων. Τα βασικά στοιχεία ενός ασαφούς συστήματος είναι η ασαφοποίηση-fuzzyfication δηλαδή η μετάφραση των αριθμητικών δεδομένων σε λεκτικά, η ανάπτυξη και η επεξεργασία των κανόνων και τέλος η αποασαφοποίηση -de fuzzyfication, δηλαδή η μετάφραση των συναρτήσεων συμμετοχής σε αριθμούς ή σε λεκτικές εκφράσεις (Παγουρτζή, 2003).

1.16 Μαζικές Εκτιμήσεις Ακινήτων

Οι **μαζικές εκτιμήσεις (Mass Appraisals)** σχετίζονται με τη συστηματική εκτίμηση πληθώρας ακινήτων, αξιοποιώντας πρότυπες διαδικασίες και στατιστική επαλήθευση σε συγκεκριμένο χρονικό διάστημα. Η μαζική αξιολόγηση σε αντίθεση με την εκτίμηση μιας μεμονωμένης περιουσίας, απαιτεί την ανάπτυξη εντός μοντέλου αποτίμησης ικανού να αναπαράγει τις δυνάμεις προσφοράς και ζήτησης σε μια

μεγάλη περιοχή. Οι εκτιμήσεις κρίνονται σε ομάδες και όχι σε μεμονωμένες ιδιότητες. Τις τελευταίες δεκαετίες, η αποτίμηση των ακινήτων εξελίχθηκε από τις απλές εμπειρικές εκτιμήσεις, στα αυτοματοποιημένα μοντέλα αποτίμησης με τις εφαρμογές τους να έχουν επεκταθεί από την αποτίμηση μιας ιδιοκτησίας σε μαζικές εκτιμήσεις (Jahanshiri, Buyong & Shariff, 2011). Οι μη αυτόματες μέθοδοι αξιολόγησης εμπειρογνωμόνων αν και αποτελεσματικές, είναι υποκειμενικές (Benjamin, Guttery & Sirmans, 2004). Για εκτάσεις περιοχών που περιλαμβάνουν δεκάδες ή εκατοντάδες ιδιότητες, η μεμονωμένη αποτίμηση της αξίας κάθε μίας ξεχωριστά, καταναλώνει μεγάλο χρονικό διάστημα επεξεργασίας και επομένως τα αυτοματοποιημένα μοντέλα αποτίμησης προτιμώνται για την επίλυση αυτών των τύπων προβλημάτων (Jahanshiri et al, 2011).

Κατά το International Association of Assessing Officers (IAAO) με τον όρο «μαζικές εκτιμήσεις» (Mass Appraisal ή Mass Valuation), εννοούμε τη συστηματική εκτίμηση ενός πλήθους ακίνητων, σε συγκεκριμένη ημερομηνία, βάσει προτυποποιημένων διαδικασιών-ενεργειών και στατιστικής επαλήθευσης. Κατά τη διαδικασία της μαζικής εκτίμησης, δημιουργείται ένα μοντέλο ικανό να προσομοιώσει το νόμο της αγοράς-ζήτησης, σε μια ευρύτερη περιοχή και αφορά ομάδες και όχι μεμονωμένα ακίνητα. Η επίτευξη του ελέγχου της ακρίβειας, της ποιότητας και της συνεκτικότητας των εκτιμήσεων αυτών πραγματοποιείται με στατιστικές μεθόδους (International Association of Assessing Officers, 2013).

Οι περισσότερες χώρες ανέπτυξαν συστήματα μαζικών εκτιμήσεων βασιζόμενα είτε στο κτηματολόγιο, είτε λειτουργούντα αυτόνομα, αλλά άμεσα συνδεδεμένα με αυτό. Η χρήση των μαζικών εκτιμήσεων γίνεται κυρίως από τον δημόσιο τομέα, για τον προσδιορισμό της φορολογητέας αξίας ακίνητης περιουσίας και για τη αξία των απαλλοτριώσεων.

Όπως αναφέρει ο Λαμπρόπουλος (2013), ο στόχος της δημιουργίας μοντέλων μαζικών εκτιμήσεων είναι να αναπαράγει την αγορά ακινήτων μέσα από ένα αντιπροσωπευτικό μαθηματικό μοντέλο, σαφώς προσδιορισμένο βάσει των αρχών της μικρο-οικονομίας και με τη δυνατότητα να προσαρμόζεται περισσότερο σε ομάδες ακινήτων διαφόρων τύπων παρά σε ανεξάρτητα ακίνητα. Ο έλεγχος του μοντέλου γίνεται βάσει των αποκλίσεων των εκτιμημένων αξιών από τις αγοραίες και η αιτιολόγησή του θεωρείται δυσκολότερη από αυτή του ελέγχου μίας μεμονωμένης εκτίμησης, που γίνεται βάσει συγκεκριμένων συγκριτικών στοιχείων. Μπορεί να είναι αδύνατο να απομονωθούν και να ληφθούν υπ’ όψιν όλοι οι παράγοντες που παίζουν

ρόλο στην επιλογή και αγορά ενός ακινήτου, είναι όμως δυνατή η επιλογή κάποιων από αυτούς και ο προσδιορισμός της συσχέτισής τους με τις αξίες των ακινήτων. Μάλιστα ο προσδιορισμός αυτός, όσο αυθαίρετος κι αν μπορεί να φαίνεται, αρκεί για να μπορεί να προβλέψει την αξία των υπό εκτίμηση ακινήτων με την απαιτούμενη στατιστική ακρίβεια. Η συλλογή των δεδομένων για τη ανάπτυξη μοντέλου αποτίμησης, η μετατροπή τους σε επεξεργάσιμη μορφή και η συντήρηση του αρχείου, είναι οι πιο σημαντικές πτυχές των περισσότερων συστημάτων μαζικής εκτίμησης, ενώ οι αποφάσεις του σχεδιαστή του συστήματος θα καθορίσει και το κόστος λειτουργίας του.

Σε σχετική τους μελέτη, οι John D. Benjamin, Randall S. Guttery και C. F. Sirmans (2004), αναφέρουν ότι ο Lusht το 2001 πρότεινε την χρήση της ανάλυσης πολλαπλής παλινδρόμησης για την αποτίμηση ενός μεγάλου αριθμού ακινήτων, γρήγορα και οικονομικά, πράγμα που εξηγεί την αυξανόμενη δημοτικότητά της μεθόδου αυτής στους φορολογικούς ελεγκτές, ενώ η «The Appraisal of Real Estate», που δημοσιεύθηκε από το The Appraisal Institute το 2001, προσέφερε μια ανάλυση αυτής της μεθόδου σε βάθος, ενώ παράλληλα οι Smith, Root, και Belloit το 1995, οι Downing και Clark το 1997, ο Allison το 1998, ο Baldwin το 1999, οι Bets και Ely το 2001 και ο Ratterman το 2001 ερεύνησαν την αξία της ανάλυσης πολλαπλής παλινδρόμησης και την εφαρμογή της σε μαζικές εκτιμήσεις .

Η πρώτη γνωστή προσπάθεια διενέργειας μαζικών εκτιμήσεων έγινε την άνοιξη του 1896 στην πόλη St. Paul - Minnesota μετά από πρόταση του William A. Somers. Μάλιστα η συγκυρία που προώθησε την ανάγκη μιας τέτοιας διαδικασίας, όπως αυτή περιγράφεται σε σχετικό τότε άρθρο, δεν απέχει πολύ από τη σημερινή οικονομική συγκυρία. Η τότε οικονομική ύφεση είχε επηρεάσει ιδιαίτερα την αξία της γης, όχι όμως ομοιόμορφα, ενώ ταυτόχρονα άλλαζαν οι υφιστάμενες χρήσεις γης και κτισμάτων και πολλά υπερτιμημένα επαγγελματικά ακίνητα έμεναν κενά, κάτι που κατέστησε επιτακτική τη μαζική επανεκτίμηση των περίπου 100.000 ακινήτων της περιοχής (Λαμπρόπουλος, 2013).

Στην Ελλάδα η δημιουργία ενός συστήματος μαζικών εκτιμήσεων βρίσκει εμπόδια, καθώς στην ελληνική αγορά δεν υπάρχουν επίσημα & οργανωμένα δεδομένα τιμών πώλησης (sales prices), η αναζήτηση συγκριτικών στοιχείων γίνεται συνήθως μέσω ζητούμενων τιμών από αγγελίες και μεσίτες, η κύρια προσέγγιση εμπορικών αξιών είναι οι εκτιμήσεις που γίνονται κυρίως από τράπεζες και η βάση αξίας είναι η αγοραία αξία και το κόστος αντικατάστασης για τα αυτοτελή ακίνητα.

Τα πολεοδομικά δεδομένα είναι μοιρασμένα σε κατά τόπους Ο.Τ.Α, χωρίς συνέχεια, ενώ μόλις πριν από λίγο καιρό άρχισε να λειτουργεί πιλοτικά η ηλεκτρονική Πολεοδομία με σοβαρές ελλείψεις και λάθη. Σημειώνεται ότι το Εθνικό Κτηματολόγιο είναι σε φάση εξέλιξης .

Με την ευρεία χρήση και την εξέλιξη των υπολογιστών επιταχύνθηκε η εξέλιξη των αυτοματοποιημένων συστημάτων μαζικών εκτιμήσεων. Έτσι λοιπόν αναπτύχτηκαν το Automated Mass Valuation (AMV), το Computer Assisted Valuation (CAV) και το Computer Assisted Mass Appraisals (CAMA). Το Computer Assisted Mass Appraisal System, είναι ένα αυτοματοποιημένο σύστημα μαζικών εκτιμήσεων το οποίο διαχειρίζεται τις πληροφορίες που αφορούν τα ακίνητα, την ενημέρωση των ιδιοκτητών, την εκτίμηση των ακινήτων και την εξασφάλιση μιας δίκαιης κατανομής φορολογικών βαρών μέσα από τις διαδικασίες εκτίμησης. Είναι απαραίτητο να τροφοδοτείται από τα γνωστά στοιχεία της αγοράς και να χρησιμοποιείται γενικά για τον προσδιορισμό της αξίας των ακινήτων (Ζεντέλης, 2015).

Σύμφωνα με το Ζεντέλη (2015) υπάρχουν τέσσερα βασικά επίπεδα λειτουργίας ενός **συστήματος CAMA**: η διαχείριση της πληροφορίας, η εκτίμηση της αξίας, η ανάλυση και η διοίκηση των εκτιμήσεων .

Ένα καλό σύστημα CAMA εκτελεί αυτές τις λειτουργίες με τρόπο αποτελεσματικό και εύχρηστο. Η διαχείριση της πληροφορίας βασίζεται στην αξιοπιστία της και η πληροφορία αυτή περιλαμβάνει το ιδιοκτησιακό καθεστώς, τα φορολογία δεδομένα, τη χρήση γης, τα στοιχεία εισοδήματος, κλπ. Η εκτίμηση της αξίας έγκειται στην δυνατότητα ενημέρωσης του κόστους, την ευελιξία τον καθορισμό μέτρων σύγκρισης, ενώ η ανάλυση των αξιών πραγματοποιείται ώστε να εξασφαλιστεί η συμφωνία των αποτελεσμάτων με συγκεκριμένα standards . Τέλος η διοίκηση των εκτιμήσεων περιλαμβάνει μια ποικιλία λειτουργιών που σχετίζονται με γενικές διοικητικές δραστηριότητες.

Σε ένα σύστημα CAMA η ακρίβεια πρόβλεψης αποτιμάται από τα μέτρα ποιοτικού ελέγχου και τα στατιστικά εργαλεία. Για την εξακρίβωση της επεξηγηματικότητας λαμβάνονται υπόψη :

- η απλότητα σχεδιασμού των μοντέλων (λειτουργία και λογική επιλογής παραμέτρων του μοντέλου),
- η λογική με την οποία αποδίδεται χρηματική αξία στις μεταβλητές, η

οποία πρέπει να συμφωνεί με τις εκ των προτέρων προσδοκίες,

- η συνοχή μεταξύ των υπομοντέλων,
- η εξασφάλιση ότι οι αξίες μεμονωμένων ακινήτων δε μεταβάλλονται ανεξήγητα από τη μια χρονιά στην επόμενη,
- ο διαχωρισμός της αξίας μεταξύ της αξίας γης και της αξίας των βελτιώσεων αυτής,
- η παρουσίαση με απλούς όρους του τρόπου εκτίμησης της αξίας του ακινήτου και
- η παρουσίαση συγκριτικών αγοραπωλησιών με σκοπό την ενίσχυση της επεξηγηματικότητας (Ζεντέλης 2015).

Η δημιουργία ενός συστήματος CAMA περιλαμβάνει τρία βασικά βήματα έχοντας ως στόχο το μοντέλο να προσεγγίσει την πραγματική αξία σε μεγάλο αριθμό περιπτώσεων με δυνατότητα εφαρμογής σε απούλητες ιδιοκτησίες:

- Συλλογή δεδομένων που αφορούν ακίνητα πουλημένα ή μη και περιλαμβάνουν χαρακτηριστικά των ακινήτων, της τοποθεσίας τους και άλλων παραγόντων που επηρεάζουν την αξία τους.
- Σχεδιασμός/μοντελοποίηση του συστήματος CAMA με διάφορες τεχνικές, έχοντας ως σκοπό την ανάπτυξη ενός εκτιμητικού μοντέλου που αντιγράφει πιστά τις δυνάμεις της αγοράς, τα αποτελέσματα του οποίου πρέπει διαρκώς να ελέγχονται με σκοπό την επίτευξη όσο το δυνατό πιο αξιόπιστων προβλέψεων.
- Εφαρμογή του μοντέλου καθώς με τη δημιουργία του συστήματος CAMA για ένα τύπο ακινήτων, έπεται η εφαρμογή του σε όλα τα ακίνητα του ίδιου τύπου, εξασφαλίζοντας το ότι όλα τα ακίνητα αυτού του ίδιου τύπου θα έχουν την ίδια μεταχείριση σε ότι αφορά τον υπολογισμό της αξίας επί της οποίας θα φορολογηθούν.

1.17 Τεχνικές Μαζικών Εκτιμήσεων

Οι τεχνικές μαζικών εκτιμήσεων έχουν σαν στόχο την επίτευξη αποδεκτών επιπέδων προβλεπτικής ικανότητας και ακρίβειας και γι’ αυτό πρέπει να είναι περιεκτικές και σύνθετες, γεγονός που τις κάνει πολλές φορές μη κατανοητές στο ευρύ κοινό. Οι τεχνικές μαζικών εκτιμήσεων δίνουν έμφαση σε εξισώσεις, πίνακες και χρονοδιαγράμματα. Η κατασκευή τέτοιων μοντέλων μπορεί να θεωρηθεί ως διαδικασία δύο σταδίων:

(1) ο προσδιορισμός της δομής του βασικού μοντέλου και

(2) η βαθμονόμηση μοντέλου.

Σ' όλο τον κόσμο εφαρμόζονται τεχνικές μαζικών εκτιμήσεων, όπως η ανάλυση πολλαπλής παλινδρόμησης, τα τεχνικά νευρωνικά δίκτυα, η ανάλυση συγκριτικών πωλήσεων κλπ. Η τεχνική της **ανάλυσης πολλαπλής παλινδρόμησης** είναι μια διεθνώς αναγνωρισμένη τεχνική ευρείας χρήσης που χρησιμοποιείται σε περιπτώσεις όπου μια μεταβλητή εξαρτάται από αριθμό άλλων μεταβλητών.
Προκειμένου να λειτουργήσει, το σύστημα θα πρέπει να τροφοδοτηθεί με σημαντικό αριθμό ακινήτων γνωστών αξιών. Από τη στιγμή που θα καθοριστεί η καλύτερη δυνατή εξίσωση, μπορεί να εφαρμοστεί και σε ακίνητα για τα οποία δε γνωρίζουμε την αξία τους. Η τεχνική της **Ανάλυσης Συγκριτικών Πωλήσεων** θέτει ερωτήματα στη βάση δεδομένων και επιλέγει εκείνα τα συγκριτικά στοιχεία που είναι πιο άμεσα συγκρίσιμα με το εκτιμώμενο ακίνητο, με σκοπό τα συγκριτικά στοιχεία να αναλυθούν και να υπολογισθεί η ανοιγμένη αξία τους. Η τεχνική των **Νευρωνικών δικτύων** βασίζεται σε αλγορίθμους που μαθαίνουν μέσω εκπαίδευσης. Τα βασικά συστατικά σε ένα νευρωνικό δίκτυο καλούνται κόμβοι ή νευρώνες, και οι συνδέσεις μεταξύ τους καθορίζονται από βάρη. Οι νευρώνες επεξεργάζονται τα δεδομένα εισαγωγής ώστε να παράξουν το αποτέλεσμα. Οι παραγόμενες τιμές διορθώνονται επαναληπτικά, μέχρις ότου το σύστημα επιτύχει την επιθυμητή ακρίβεια. Τα **Ειδικά Συστήματα** πραγματοποιούν αυτοματοποιημένη μοντελοποίηση παλινδρόμησης από την αρχή έως το τέλος, κάνοντας όλους τους απαραίτητους στατιστικούς ελέγχους ανάλογα με τα δεδομένα που έχουν εισαχθεί στο σύστημα από τον χρήστη. Είναι ουσιαστικά συστήματα που έχουν κάποιο βαθμό τεχνητής νοημοσύνης. Η επιθυμητή περιεχόμενη πληροφορία θα μπορούσε να περιέχει τα στοιχεία υπόβαθρου του συστήματος, στοιχεία των ακινήτων και τα τεχνικά χαρακτηριστικά του συστήματος, δηλαδή τη διαχείριση βάσης δεδομένων, το λογισμικό, τη στατιστική ανάλυση και τη δημιουργία αναφορών (Λαμπρόπουλος 2013).

Κεφάλαιο 2^{ο} - Μεταβλητές

Σημαντικό στην κατασκευή ενός μοντέλου παλινδρόμησης είναι η ορθή επιλογή των κατάλληλων μεταβλητών. Παράλληλα η ορθή κατάταξη τους και η σωστή κωδικοποίηση τους θα προσδιορίσουν με τον καλύτερο δυνατό τρόπο την προβλεπτική δυνατότητα του.

2.1 Προσδιορισμός της Αξίας Ακινήτου και η Επιρροή των Μεταβλητών

Ο **προσδιορισμός της αξίας ενός ακινήτου** είναι άμεσα εξαρτώμενος από πολλούς ανεξάρτητους και αστάθμητους παράγοντες. Τα πολεοδομικά δεδομένα της περιοχής, τα δίκτυα υποδομών, τα δομικά-κατασκευαστικά χαρακτηριστικά των ακινήτων, η θέα, η πρόσβαση σε εξυπηρετήσεις και υπηρεσίες δημοσίου χαρακτήρα, τα χαρακτηριστικά του υπόλοιπου δομημένου περιβάλλοντος της περιοχής και ιδιαίτερα της γειτονιάς, ενώ ακόμα και η θέση του κάθε συγκεκριμένου ακινήτου στο χώρο, ως παράγοντες, παίζουν καθοριστικό ρόλο στη διαμόρφωση της αξίας του. Καθίσταται σαφής η αδυναμία συνυπολογισμού όλων των παραμέτρων που δύναται να συμβάλλουν στη διαμόρφωση της αξίας ενός ακινήτου. Άλλωστε η προσπάθεια συνυπολογισμού όλων θα είχε ως αποτέλεσμα μια μελέτη αμφίβολης εγκυρότητας, ενώ το πρόβλημα διαπιστώνεται ως ο καθορισμός αυτών που θα έχουν άμεση επιρροή στην αξία του ακινήτου (Ozgur, Hughes, Rogers & Parvee, 2016). Οι παράγοντες ως σύνολα κατηγοριών ή και περιορισμών εκφράζονται μέσω των μεταβλητών στην επεξεργασία των δεδομένων και για τον λόγο αυτό, ανάλογα με το επίπεδο της έρευνας, επιλέγονται και οι αντίστοιχες μεταβλητές που εξηγούν καλύτερα τον μηχανισμό διαμόρφωσης της αξίας για τη δημιουργία του μοντέλου. Οι Mak, Choy & Ho (2009) σε μελέτη τους αναφέρουν ότι οι μεταβλητές που καθορίζουν την τιμή ενός ακινήτου είναι σχετικές με τα δομικά χαρακτηριστικά του ακινήτου, την προσβασιμότητα σε εξυπηρετήσεις και την ποιότητα του περιβάλλοντος χώρου. Στα χωρικά μοντέλα, δίνεται ιδιαίτερη έμφαση στα χαρακτηριστικά θέσης και στα δομικά χαρακτηριστικά των ακινήτων που εντάσσονται στην έρευνα. Σύμφωνα με τον Tobler (1970) « *Όλα τα πράγματα σχετίζονται το ένα με το άλλο, αλλά αυτά που βρίσκονται κοντά το ένα με το άλλο σχετίζονται περισσότερο από τα άλλα, που βρίσκονται σε μεγαλύτερη απόσταση μεταξύ τους*» (σελ.236). Η τιμή λοιπόν ενός ακινήτου είναι συνήθως συγκρίσιμη με αυτή των

γειτονικών του, έχοντας χωρική συσχέτιση και ως εκ τούτου προκύπτει η ύπαρξη μιας συστηματικής μεταβολής των τιμών των ακινήτων στον χώρο.

2.2 Μεταβλητές και Κλίμακες Μέτρησης

Μεταβλητή είναι κάθε χαρακτηριστικό που μπορεί να μεταβάλλεται και να λαμβάνει διάφορες τιμές, ή να συμβαίνει με διαφορετικές διαβαθμίσεις. Η έννοια της μεταβλητής είναι στενά συνδεδεμένη με αυτή της μέτρησης και μπορεί να μετρηθεί ποσοτικά ή ποιοτικά. Η διαφοροποίηση των τιμών της αποτελεί την κατανομή της (Ηλιοπούλου, 2015).

2.2.1 Ονομαστικές, Ιεραρχικές, Ισοδιαστημικές, Αναλογικές, Μεταβλητές

Σύμφωνα με την πρόταση του Stevens στην επιστημονική επιθεώρηση Science του 1946, διακρίνουμε 4 είδη μεταβλητών και με βάση αυτές 4 είδη μέτρησης (Stevens, 1946):

Πρώτον έχουμε τις **Κατηγορικές ή Ονομαστικές μεταβλητές (nominal or categorical scales),** οι οποίες αφορούν περισσότερο ποιοτικές παρά ποσοτικές μετρήσεις (πχ. διαμέρισμα, μεζονέτα, μονοκατοικία). Για την κατασκευή των κατηγορικών μέτρων υπάρχουν δύο προϋποθέσεις:

α) οι κατηγορίες να είναι αμοιβαία αποκλειόμενες, δηλαδή μια παρατήρηση να μην μπορεί να ανήκει σε περισσότερες από μία κατηγορίες και

β) οι κατηγορίες να είναι εξαντλητικές, δηλαδή να υπάρχουν αρκετές κατηγορίες που να καλύπτουν όλες τις περιπτώσεις.

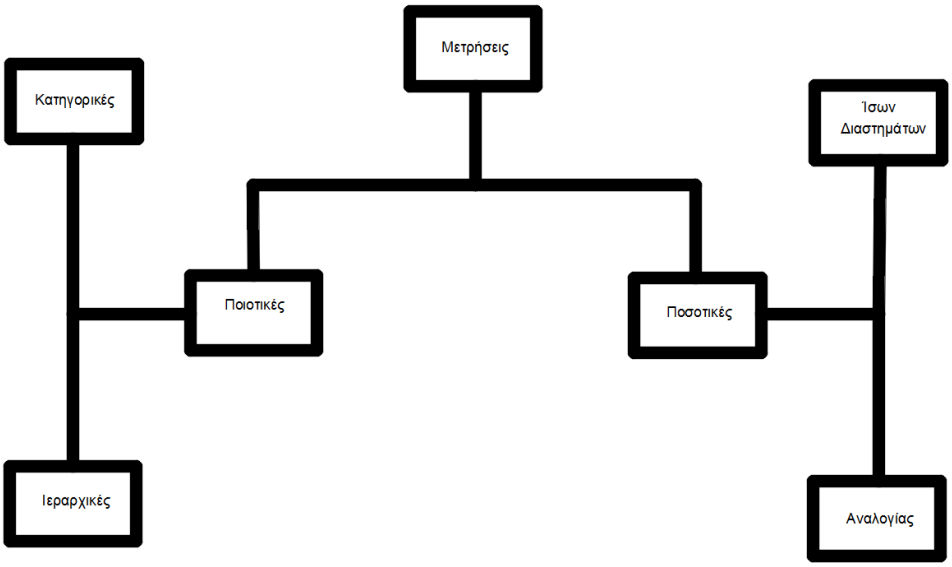
Στις κατηγορικές μεταβλητές δεν υπάρχει η έννοια της ποσότητας ή της μεγαλύτερης ή μικρότερης αξίας, ενώ υπάρχει η έννοια της διαφορετικότητας. Η μεγαλύτερη ή μικρότερη αξία μιας κατηγορίας εξαρτάται από τις υποθέσεις που θα κάνει ο ερευνητής. Γενικά ισχύει ότι οι ονομαστικές μεταβλητές αφορούν ποιοτικές μετρήσεις, παίρνουν αριθμητικές τιμές που αντιστοιχούν σε ονόματα, οι επιτρεπόμενες πράξεις είναι η ισότητα ή η μη ισότητα μεταξύ των τιμών και υπάρχει δυνατότητα διάταξης.

Έπονται οι **Διατάξιμες ή Τακτικές ή Ιεραρχικές μεταβλητές (ordinal scales).** Σε αυτές εφαρμόζονται τα κριτήρια του αμοιβαίου αποκλεισμού, της εξαντλητικότητας και της ιεραρχικής ταξινόμησης. Οι κατηγορίες κατατάσσονται με βάση κάποιων εξωτερικών κριτήριων, ενώ ταξινομώντας σε μία συγκεκριμένη κατηγορία έχουμε περισσότερο ή λιγότερο ένα χαρακτηριστικό, σε σύγκριση με μία

άλλη κατηγορία (πχ. σειρά κατάταξης σε ορόφους: 1ος, 2ος, 3ος, ποιότητα κατασκευής: καλή, μέτρια, κακή, κλπ). Η ιεραρχική ταξινόμηση είναι σύμφωνη με τη φύση της μεταβλητής και οι παρατηρήσεις ταξινομούνται σε κατηγορίες μεταξύ των οποίων υπάρχει σειρά. Οι αποστάσεις μεταξύ δύο σημείων στην ιεραρχική κλίμακα δεν μπορούν να μετρηθούν με κοινή μονάδα μέτρησης. Παράδειγμα είναι η κατανομή των εδαφών σε υψομετρικές κατηγορίες (ορεινό, ημιορεινό κλπ.). Σημειώνεται ότι η ιεράρχηση, η οποία είναι άγνωστη και ασταθής, δεν περιγράφει την ποσότητα της διαφοράς μεταξύ των κατηγοριών. Γενικά ισχύει ότι οι διατάξιμες μεταβλητές αφορούν ποιοτικές μετρήσεις, οι ονομαστικές τιμές έχουν την δυνατότητα διάταξης, ενώ επιτρέπεται η πράξη της ανισότητας.

Ακολουθούν οι **Μεταβλητές Ίσων Διαστημάτων, Ισοδιαστημικές (interval scales),** όπου οι τιμές της μεταβλητής αντικατοπτρίζουν την ύπαρξη μιας ιδιότητας σε μεγαλύτερο ή μικρότερο βαθμό. Στην περίπτωση αυτή ισχύουν όλα τα παραπάνω κριτήρια και επιπλέον ίσα αριθμητικά διαστήματα, που προσδιορίζουν ίσες διαφορές στην ύπαρξη της ιδιότητας αυτής. Γενικά ισχύει ότι οι ισοδιαστημικές μεταβλητές αφορούν κυρίως ποσοτικές μετρήσεις, οι αριθμητικές τιμές έχουν συγκρίσιμες διαφορές, επιτρέπεται η πρόσθεση και η αφαίρεση (παράδειγμα το επίπεδο των ορόφων κτιρίου : -1, 0, +1, +2, και η θερμοκρασία).

Τέλος έχουμε τις **Αναλογικές μεταβλητές (ratio scales)**. Οι αναλογικές κλίμακες έχουν τις ίδιες ιδιότητες με τις κλίμακες ίσων διαστημάτων και επιπλέον τη τιμή 0 (μηδέν) που υποδηλώνει την πραγματική απουσία του χαρακτηριστικού που εκφράζει η μεταβλητή (πχ. απόσταση κατοίκων από την εργασία τους).
Γενικά οι αναλογικές μεταβλητές αφορούν ποσοτικές μετρήσεις, είναι συνεχείς, οι αριθμητικές τιμές έχουν συγκρίσιμες αναλογίες, ενώ επιτρέπονται οι πράξεις του πολλαπλασιασμού και της διαίρεσης (παράδειγμα η παλαιότητα κτιρίου σε έτη 0,1,2,3,κλπ). Στο γράφημα 2.1 παρουσιάζεται η διάταξη των μετρήσεων.



Γράφημα 2.1.Διάταξη Μετρήσεων. Ιδία επεξεργασία.

2.2.2 Συνεχής, Ασυνεχής, Διχοτομικές

Οι ονομαστικές (κατηγορικές) και οι ιεραρχικές μεταβλητές είναι **ασυνεχείς (discrete ή διακριτές)**, παίρνοντας μόνο έναν ορισμένο αριθμό τιμών που αντιστοιχεί στις κατηγορίες των μεταβλητών, ενώ οι ισοδιαστημικές και αναλογικές μεταβλητές δύναται να είναι **συνεχείς ή ασυνεχείς**.

Σε γενικές γραμμές, **συνεχής** (continuous) είναι μια μεταβλητή, όταν παίρνει τιμές από ένα διάστημα, δηλαδή για κάθε δυο τιμές της, υπάρχει μια τιμή ανάμεσά τους (π.χ. ύψος, βάρος, μήκος). Ως συνεχείς λοιπόν ορίζονται οι ποσοτικές μεταβλητές οι οποίες παίρνουν οποιαδήποτε τιμή εντός διαστήματος τιμών, με άπειρο πιθανό αριθμό τιμών, σε ένα δεδομένο διάστημα.

Α**συνεχής (discrete ή διακριτή)** είναι μια μεταβλητή οπου για κάθε δυο τιμές της δεν υπάρχει πάντα μια τιμή της μεταβλητής ανάμεσά τους (π.χ. αριθμός δωματίων μιας κατοικίας). Ως ασυνεχείς ή διακριτές λοιπόν ορίζονται οι μεταβλητές οι οποίες λαμβάνουν διακεκριμένες τιμές και έχουν ελάχιστη μονάδα μέτρησης.

Οι **διχοτομικές μεταβλητές (dichotomous ή binary variables)** είναι μια ειδική περίπτωση μεταβλητών μετρούμενες στην ονομαστική κλίμακα, που παίρνουν μόνο δύο τιμές (πχ. 0-1 , υπάρχει αποθήκη στο ακίνητο ΝΑΙ – ΟΧΙ) και που μπορούν να χρησιμοποιηθούν σε στατιστικές επεξεργασίες για τις οποίες απαιτούνται ισοδιαστημικές ή αναλογικές μεταβλητές. Η ερμηνεία των αποτελεσμάτων πρέπει να λαμβάνει υπόψη τη φύση των απεικονιζόμενων χαρακτηριστικών (Ηλιοπούλου 2015).

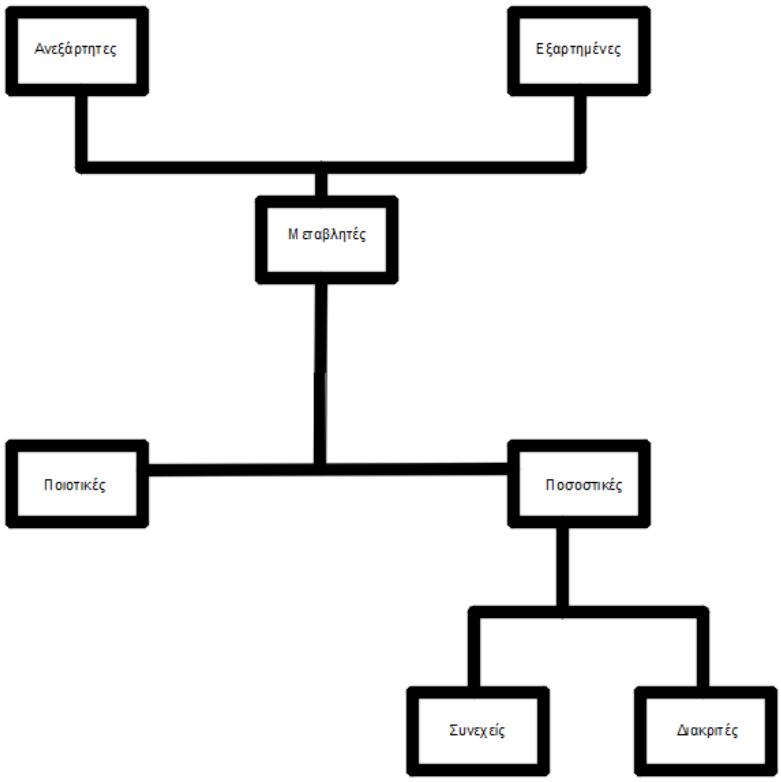
2.2.3 Ανεξάρτητες, Εξαρτημένες Μεταβλητές

Η διερεύνηση αιτιωδών σχέσεων μεταξύ φαινομένων, ώστε να καταστεί δυνατή η διατύπωση πρόβλεψης μελλοντικών φαινομένων σε ερευνητικούς σχεδιασμούς κατατάσσει τις μεταβλητές σε δύο βασικές κατηγορίες: τις ανεξάρτητες και τις εξαρτημένες μεταβλητές (Independent -Dependent) .

Ανεξάρτητη μεταβλητή (independent variable) είναι εκείνη η οποία θεωρούμε ότι προκαλεί την εξαρτημένη μεταβλητή.

Εξαρτημένη μεταβλητή (dependent variable) είναι εκείνη της οποίας η τιμή εξαρτάται από την τιμή της ανεξάρτητης μεταβλητής. Ο ερευνητής μεταβάλλει τις τιμές της ανεξάρτητης ή των ανεξάρτητων μεταβλητών, προκειμένου να προκαλέσει αλλαγές στην εξαρτημένη μεταβλητή.

Κατά την έρευνα συσχέτισης μεταξύ εξαρτημένης και ανεξάρτητης ή ανεξάρτητων μεταβλητών, σε περιπτώσεις παλινδρόμησης με σκοπό την πρόβλεψη, οι ανεξάρτητες μεταβλητές μπορούν να χαρακτηριστούν ως **επεξηγηματικές (explanatory, experimental predictor ή input),** ενώ η εξαρτημένη ως μεταβλητή **απόκρισης (response, output ή outcome).** Στο γράφημα 2.2 παρουσιάζονται οι περιπτώσεις μεταβλητών.



Γράφημα 2.2. Μεταβλητές. Ιδία επεξεργασία.

2.2.4 Ψευδομεταβλητές

Όταν ένα μέγεθος είναι αδύνατο να ποσοτικοποιηθεί, αλλά πρέπει να χρησιμοποιηθεί σ' ένα υπόδειγμα, προσεγγίζεται με την ψευδομεταβλητή. Οι **ψευδομεταβλητές,** συνήθως δίτιμες (dummy variables), είναι μεταβλητές που χρησιμοποιούνται για να ποσοτικοποιηθούν οι ποιοτικές μεταβλητές. Αυτές θα λάβουν τις τιμές (0,1) αναλόγως με το αν η τιμή της παρατήρησης επαληθεύει το χαρακτηριστικό της μεταβλητής ή όχι. Εάν το πλήθος k των ανεξάρτητων κατηγοριών τιμών που μπορεί να λάβει μία μεταβλητή, είναι μεγαλύτερο των δύο, τότε χρησιμοποιούνται k-1 ψευδομεταβλητές, με τη τιμή (1) της καθεμίας να επιβεβαιώνει την ύπαρξη της αντίστοιχης κατηγορίας, ενώ όλες οι άλλες μαζί παίρνουν την τιμή μηδέν (0) (Λαμπρόπουλος, 2013).

2.3 Εξειδικευμένη Ανάλυση Επιρροής Μεταβλητών στα Ακίνητα

Σύμφωνα με τον Dubin, τα χαρακτηριστικά που καθορίζουν τις τιμές των ακινήτων και κατ' επέκταση οι **ανεξάρτητες μεταβλητές** των οικονομετρικών υποδειγμάτων, **ομαδοποιούνται** σε τρεις κατηγορίες:

- στις μεταβλητές θέσης,
- στις μεταβλητές κατασκευαστικών χαρακτηριστικών του ακινήτου και
- στις μεταβλητές της γειτονιάς,

προσδιορίζοντας τη γενική εξίσωση των τιμών των ακινήτων ως εξής:

Τιμή = f (Θέση, Χαρακτηριστικά, Γειτονιά)+Σφάλμα, όπου

–*Τιμή*- είναι το διάνυσμα των τιμών των ακινήτων,

-*Θέση*- είναι οι μεταβλητές θέσης,

-Χ*αρακτηριστικά*- είναι οι μεταβλητές που αφορούν τα κατασκευαστικά χαρακτηριστικά του ακινήτου,

-*Γειτονιά*- είναι οι μεταβλητές της γειτονιάς και

–*Σφάλμα*- είναι το διάνυσμα των τυχαίων όρων σφάλματος. (Dubin, 1988)

Η Can διαχωρίζει τις **επεξηγηματικές μεταβλητές** σε δύο κατηγορίες σύμφωνα με:

- τα δομικά χαρακτηριστικά,
- τη γεωγραφική θέση και γειτονιά (Can, 1992) .

Οι Sirmans & Macherson αναφέρουν, σε βιβλιογραφική ανασκόπηση του 2003, ότι οι περισσότερες ηδονικές μελέτες που έχουν γίνει περιλαμβάνουν

μεταβλητές όπως: μέγεθος οικοπέδου, εμβαδόν κατοικίας, αριθμό δωματίων και μπάνιων, ύπαρξη τζακιού, κεντρικό κλιματισμό, ύπαρξη υπογείου χώρου και γκαράζ, ύπαρξη βεράντας, πισίνας, απόσταση από το κέντρο της πόλης και χρόνο έως τις εξυπηρετήσεις. Θεωρούν ότι οι μεταβλητές αυτές είναι γενικά κατηγοριοποιημένες σε:

- **Κατασκευαστικές,** που περιλαμβάνουν το μέγεθος του προς πώληση χώρου, την ηλικία του, τον αριθμό δωματίων και μπάνιων,
- στη **κατηγορία των εσωτερικών χαρακτηριστικών,** όπως ο προσδιορισμός ύπαρξης ολόκληρων ή μη μπάνιων, WC, η ύπαρξη τζακιού, κλιματισμού, ξύλινων πατωμάτων και υπόγειου,
- **στη κατηγορία των εξωτερικών χαρακτηριστικών,** όπως η ύπαρξη γκαράζ, βεραντών, πισίνας, υπόστεγου,
- στην **κατηγορία των περιβαλλοντικών χαρακτηριστικών**, όπως η θέα ή ακόμα πιο εξειδικευμένα η θέα σε λίμνη, η θέα σε θάλασσα , κλπ.,
- στα **γενικότερα χαρακτηριστικά της γειτονιάς και τοποθεσίας**, όπως η ύπαρξη και τα ποσοστά εγκληματικότητας, η απόσταση από πράσινο, η απόσταση από ελευθέρους χώρους γηπέδων,
- στην κατηγορία των **δημοσίων εξυπηρετήσεων,** όπως η ύπαρξη κοντινού σχολείου, κοντινού μειονοτικού σχολείου, η ύπαρξη κοντινών δημόσιων οργανισμών,
- στους **γενικότερους παράγοντες επηρεασμού της αγοράς,** όπως τα έξτρα αξεσουάρ ποιότητας που μπορεί να έχει ένα ακίνητο, το χρόνο που το ακίνητο είναι στην αγορά προς πώληση, το θετικό timing της αγοράς και τέλος
- στη **χρηματοδοτική δυνατότητα και τα φορολογικά στοιχεία**, δηλαδή την δυνατότητα παροχής δανειοδοτήσεων την περίοδο που το ακίνητο είναι προς πώληση και το γενικότερο οικονομικό κλίμα της εποχής.

Οι Sirmans et al (2003) αναφέρουν στην μελέτη τους και τις μεταβλητές που συνήθως συσχετίζονται αρνητικά με την τιμή του ακινήτου. Τέτοιες είναι: η εγκληματικότητα της περιοχής, η ύπαρξη μειονοτικών σχολείων καθώς και τα πολλά κενά διαμερίσματα στην περιοχή. Επιπρόσθετα αναφέρουν ως συνήθως θετικά συσχετιζόμενες μεταβλητές την κεκλιμένη οροφή του ακινήτου, το ξεχωριστό ντους, τον διπλό φούρνο στην κουζίνα και την αυλή, ενώ με αρνητική συσχέτιση αναφέρουν την μη ύπαρξη ελεύθερου περιμετρικού χώρου, την έντονη σεισμικότητα της περιοχής, την εγγύτητα σε χώρους εκτροφής ζώων, σε χώρους ταφής απορριμμάτων,

σε περιοχές μεγάλων εταιρικών ιδιοκτησιών (βιομηχανικών ζωνών) και σε περιοχές πλησίον γραμμών υψηλής τάσης ρεύματος. Συμπερασματικά επισημαίνουν την σταθερά αρνητική συσχέτιση τιμής με την ηλικία του ακινήτου στη πλειονότητα των μελετών, το γεγονός ότι κάθε επιπλέον δωμάτιο αυξάνει την τιμή πώλησης κατά 10-12% και την αύξηση της τιμής των ακινήτων κατά 6-12% εφόσον υπάρχει τζάκι. Επισημαίνουν ακόμα την σταθερά θετική συσχέτιση με την ύπαρξη κεντρικού κλιματισμού, την αύξηση στην τιμή κατά 12-16% εφόσον υπάρχει υπόγειο ως βοηθητικός χώρος, τη θετική συσχέτιση της πισίνας και του γκαράζ που δίνει αύξηση στην τιμή του ακινήτου από 6-12%. Τέλος γίνεται αναφορά στην θετική συσχέτιση με την ύπαρξη ενός καλού σχολείου στη περιοχή (Sirmans et al.,2003) .

2.3.1 Ανεξάρτητες Μεταβλητές Κατασκευαστικών Χαρακτηριστικών

Το **εμβαδόν** ενός ακινήτου είναι ειδικός παράγοντας επίδρασης, λειτουργεί ως αναλογική, συνεχής, ποσοτική μεταβλητή, η οποία ως μονάδα μέτρησης έχει τα τετραγωνικά μέτρα (τμ) και ανήκει στα γεωμετρικά, κατασκευαστικά χαρακτηριστικά του ακινήτου. Σύμφωνα με τους Rodriguez & Sirmans (1994) το εμβαδόν ενός ακινήτου σχετίζεται θετικά με την τιμή πώλησης και αυτό οφείλεται στην προθυμία των αγοραστών να πληρώσουν περισσότερο για περισσότερο χώρο.
Οι Ozgur et al., (2016) σε εργασία τους διαπιστώνουν και αυτοί την εξαρτημένη σχέση μεταξύ τιμής και επιφάνειας, ενώ οι Mak, et al (2009) σε αντίστοιχη μελέτη επισημαίνουν το ενδιαφέρον των αγοραστών πολυτελών κατοικιών, περισσότερο για τα τετραγωνικά του ακινήτου και τον αριθμό των μπάνιων, παρά για τα υπόλοιπα χαρακτηριστικά του, όπως η προσβασιμότητα σε υπηρεσίες και εξυπηρετήσεις.
Σύμφωνα με τον Καρανικόλα (2010) το μέγεθος έχει πρωταρχικό ρόλο στη διαμόρφωση της αγοραίας αξίας μιας ιδιοκτησίας. Παρόλα αυτά η σχέση μεταξύ αξίας και εμβαδού δεν είναι ανάλογη, με το μέγεθος να επιδρά στα μικρότερα από το συνηθισμένο μέγεθος ακίνητα με ένα θετικό συντελεστή, αυξάνοντας την τιμή τους, ενώ στην περίπτωση των μεγαλύτερων ακινήτων, ο συντελεστής μεγέθους είναι αρνητικός. Σύμφωνα με μελέτη των Raslanas, Tupenaite, Steinbergas (2005) για την περιοχή του Λονδίνου, διαπιστώθηκε και εκεί η αναμενόμενη θετική συσχέτιση εμβαδού τιμής, με παράλληλη όμως ζήτηση μικρών διαμερισμάτων . Οι Chin & Chau, (2002) διαπιστώνουν ότι πιθανοί αγοραστές είναι διατεθειμένοι να πληρώσουν περισσότερα για περισσότερο χώρο, ενώ παράλληλα διαπιστώνουν ότι μεγάλες οικογένειες προτιμούν μεγάλα ακίνητα εφόσον έχουν να τα πληρώσουν.

Η Οικονομικής Επιτροπής της Ελλάδος (ΟΚΕ) σε μελέτη του 2009, σχετικά με την κατασκευαστική δραστηριότητα σε σχέση με την οικονομική κρίση στον ελλαδικό χώρο, επισημαίνει ότι η μείωση του μέσου μεγέθους του νοικοκυριού, έχει επίπτωση στην αγορά κατοικίας, με αποτέλεσμα τη μετατόπιση του αγοραστικού κοινού σε προτίμηση μικρότερων σε εμβαδόν κατοικιών (Οικονομική Επιτροπή της Ελλάδος, 2009).

Βασικό κατασκευαστικό-δομικό χαρακτηριστικό είναι η μεταβλητή που αναφέρεται στον **αριθμό των υπνοδωματίων.** Είναι μία αναλογική, διακριτή, ποσοτική μεταβλητή. Περιλαμβάνεται στην πλειοψηφία των ηδονικών μοντέλων χαρακτηρίζοντας μια κατοικία. Αποτελεί δείκτη πυκνοκατοίκησης, εφόσον αναφέρεται σε άτομα ανά δωμάτιο. Ο αριθμός των υπνοδωματίων είναι μιας μορφής αναλογική μεταβλητή και κατά τους Li & Brown (1980) σχετίζεται με το μέγεθος του ακινήτου. Οι Chrostek και Kopczewska (2013) σε μελέτη τους επισημαίνουν τον συσχετισμό της τιμής με την μεταβλητή χρησιμοποιούμενου χώρου ανά δωμάτιο. Η μεταβλητή αυτή προκύπτει από την διαίρεση της χρήσιμης επιφάνειας με τον αριθμό των υπνοδωματίων της ιδιοκτησίας. Οι Sirmans et al., (2003) αναφέρουν ότι ο αριθμός υπνοδωματίων έχει θετική συσχέτιση με την τιμή. Η μελέτη του 2009 της Οικονομικής Επιτροπής της Ελλάδος (ΟΚΕ), που προαναφέρθηκε, επισημαίνει ότι η μείωση του μέσου μεγέθους του νοικοκυριού έχει επίπτωση στη μετατόπιση του αγοραστικού κοινού σε προτίμηση μικρότερων σε αριθμό δωματίων κατοικιών. Παράδοξο είναι ότι οι Ozgur et al., (2016) σε εργασία τους διαπιστώνουν την εξαρτημένη σχέση μεταξύ τιμής και υπνοδωματίων με αρνητική συσχέτιση. Σημειώνεται ότι η μελέτη έγινε σε δείγμα 480 σπιτιών στην βορειοδυτική Ιντιάνα. Τέλος ότι οι Fan, Ong & Koh (2006) εξετάζοντας το θέμα της παλινδρόμησης με την μη παραμετρική μέθοδο της δενδρικής προσέγγισης, αναφέρουν ότι οι πιθανοί αγοραστές κατοικιών μικρών και μεσαίων διαμερισμάτων ενδιαφέρονται περισσότερο για τα βασικά χαρακτηριστικά στέγασης, όπως η επιφάνεια, ο τύπος του ακινήτου, ο αριθμός δωματίων και η ηλικία, ενώ οι πιθανοί αγοραστές μεγαλύτερων ακινήτων, έχοντας υψηλότερες προσδοκίες και απαιτήσεις ποιότητας κατασκευής, πληρώνουν περισσότερα έχοντας στραμμένη την προσοχή τους σε χαρακτηριστικά ποιότητας και υπηρεσιών, όπως οι εγκαταστάσεις αναψυχής και διαβίωσης.

Η μεταβλητή **παλαιότητα** ως ειδική παράμετρος που επιδρά στην διαμόρφωση της αξίας ενός ακινήτου, είναι μια διακριτή, ποσοτική μεταβλητή διότι αφορά το χρόνο, με μονάδα μέτρησης τα έτη που πέρασαν από την χρονιά αποπεράτωσης της

κατασκευής. Το έτος κατασκευής ως δεδομένο επεξεργάζεται ώστε να προκύψει στην συνέχεια η μεταβλητή με την μορφή αριθμού ετών. Είναι ένα δομικό χαρακτηριστικό με την μορφή αναλογικής (ratio) μεταβλητής διαστήματος. Είναι μια ειδική παράμετρος διαμόρφωσης της αξίας ενός ακινήτου και σύμφωνα με τους Mak, et al (2009) η ηλικία του ακινήτου είναι ένα χαρακτηριστικό που ενδιαφέρει τους αγοραστές σπιτιών χαμηλότερων τιμών. Σημειώνεται ότι η ποιότητα κατασκευής πολλές φορές συνδέεται με την παλαιότητα, καθώς η ποιότητα ενός κτίσματος δεν αφορά μόνο την κατηγοριοποίησή του βάσει πολυτέλειας της κατασκευής αλλά και βάσει της στατικής επάρκειας που έχει προβλεφτεί με την αντισεισμική μελέτη. Μεγαλύτερης παλαιότητας κτίρια εμφανίζονται κυρίως στο κέντρο των πόλεων, ενώ η παλαιότητα του ακινήτου έχει σχέση και με την συντήρηση που έχει γίνει σε αυτό. Μπορεί ακόμα να έχει λιγότερη βαρύτητα στη διαμόρφωση της αγοραίας αξίας ενός ακινήτου η παλαιότητα της οικοδομής όταν ο αγοραστής επιλέγει ένα πλήρως ανακαινισμένο διαμέρισμα στο κέντρο της πόλης, με όλα τα θετικά που αυτό συνεπάγεται. Σύμφωνα με τους Li et al., (1980), διαπιστώθηκε η αρνητική συσχέτιση της τιμής με την ηλικία του ακινήτου η οποία όμως αντιστρέφεται όταν το κτίριο έχει ιστορική σημασία .

Οι Sirmans et al., (2003) αναφέρουν ότι η παλαιότητα έχει αρνητική συσχέτιση με την τιμή, παράλληλα με τους Ozgur et al., (2016) που σε εργασία τους διαπιστώνουν ανάλογα αποτελέσματα, ενώ οι Chrostek et al., (2013) επισημαίνουν ότι το έτος ανέγερσης μπορεί να έχει διφορούμενη επίδραση στην αποτίμηση της αξίας μιας ιδιοκτησίας. Από τη μία μεριά, τα νεώτερα κτίρια μπορούν έχουν υψηλότερη ποιότητα και η απόσβεσή τους μπορεί να είναι χαμηλότερη, ενώ από την άλλη, σε μερικές περιπτώσεις οι τιμές ακινήτων μπορεί να αυξηθούν με την ηλικία, καθώς τα παλαιότερα κτίρια μπορούν να εκτιμηθούν περισσότερο λόγω της αρχιτεκτονικής τους ή της καλύτερης τοποθεσίας τους.

Οι Benjamin,Guttery & Sirmans (2004) σε μελέτη τους που αφορά την περιοχή της Washington διαπιστώνουν αρνητική συσχέτιση μεταξύ της αξίας ενός ακινήτου με την ηλικία του που μπορεί να οφείλεται στην φυσική φθορά των πραγμάτων, στην λειτουργική απαξίωση και στην εξωτερική υποτίμηση.

Η μεταβλητή **όροφος** είναι ένα δομικό χαρακτηριστικό του ακινήτου. Είναι ειδική παράμετρος και ως ποσοτική, ισοδιαστημική, ασυνεχής, μεταβλητή επιδρά στην αξία του ακινήτου έχοντας θετική συσχέτιση. Παραμένει στατιστικά επεξεργάσιμη έπειτα από την ποσοτικοποίηση της. Τα διαμερίσματα που βρίσκονται

στο ισόγειο οικοδομής μπορεί να πάρουν την τιμή μηδέν (0), ο πρώτος όροφος παίρνει την τιμή ένα (1), ο δεύτερος όροφος την τιμή δύο (2) κ.ο.κ. Άμεση συνέπεια του ορόφου ενός διαμερίσματος σε συνδυασμό με τη θέση του σε ένα κτίσμα (πρόσοψη σε οδό ή στον ακάλυπτο χώρο της οικοδομής) είναι η θέα του. Τα διαμερίσματα που βρίσκονται στους υψηλότερους ορόφους και έχουν πρόσοψη σε κύρια οδό, είναι πιθανότερο να έχουν καλύτερη θέα από τα ακριβώς ίδια διαμερίσματα χαμηλότερων ορόφων και κατ' επέκταση μεγαλύτερη αξία. Οι Ozgur et al., (2016) σε εργασία τους διαπιστώνουν την ασθενή θετικά εξαρτημένη σχέση μεταξύ τιμής και ορόφων, ενώ οι Chrostek et al., (2013) σε μελέτη τους επισημαίνουν τον συσχετισμό της τιμής με τον όροφο στον οποίο βρίσκεται η ιδιοκτησία, κατηγοριοποιώντας τα κτίρια σε χαμηλά κτίρια, κτίρια πολυκατοικίας, αρχοντικά, ουρανοξύστες και διαμερίσματα. Παράλληλα διαπιστώνουν ισχυρή συσχέτιση μεταξύ της τιμής και της επίδρασης των χωρικών στοιχείων. Οι Lake, Lovett, Bateman & Day (1999) αναφέρουν σε μελέτη τους ότι τα ψηλά κτίρια έχουν καλύτερη θέα και υπάρχει θετικός συσχετισμός των τριών αυτών μεταβλητών: αξία, όροφος , θέα, ενώ ανάλογη παρατήρηση έκαναν οι So, Tse, & Ganesan (1996) διαπιστώνοντας υψηλότερη αξία διαμερισμάτων σε υψηλότερους ορόφους λόγω της θέας που προσφέρουν. Οι Sirmans et al., (2003) αναφέρουν και αυτοί την θετική συσχέτιση υψηλού ορόφου με την τιμή.

Η μεταβλητή **ανελκυστήρας- ασανσέρ** είναι μια δίτιμη ονομαστική, ποιοτική μεταβλητή, η οποία παίρνει ως τιμές την ύπαρξη και τη μη ύπαρξη. Πρόκειται για ένα κατασκευαστικό χαρακτηριστικό. Για την συμμετοχή της στην διαδικασία της στατιστικής ανάλυσης θα πρέπει να ποσοτικοποιηθεί έτσι ώστε να αντιστοιχίζεται ένας αριθμός σε κάθε μία από τις δύο περιπτώσεις. Χρησιμοποιείται μια ψευδομεταβλητή (dummy variable) και οι τιμές που δίνονται είναι αντίστοιχα 0 και 1, (0 για τη μη ύπαρξη ανελκυστήρα και 1 για την ύπαρξη ανελκυστήρα). Η επιρροή της μεταβλητής αυτής συνήθως είναι μεγαλύτερη στα κέντρα των πόλεων, καθώς λόγω παλαιότητας των οικοδομών η διασπορά των τιμών που δέχεται η μεταβλητή αυτή είναι συνήθως σημαντική. Η μη ύπαρξη ανελκυστήρα σε πολυκατοικία μειώνει την τιμή του διαμερίσματος (Στάμου, 2013).

Η μεταβλητή **αριθμός μπάνιων** είναι μια αναλογική, ασυνεχής, ποσοτική μεταβλητή. Πρόκειται για ένα κατασκευαστικό χαρακτηριστικό. Σύμφωνα με την Can (1992) η μεταβλητή αυτή έχει θετική συσχέτιση με την τιμή του ακινήτου. Μπορεί να δημιουργηθεί και άλλη μια μεταβλητή, ως ανακαινισμένο μπάνιο ή όχι, η

οποία θα μπορούσε να λειτουργήσει ικανοποιητικά σε παλιά κυρίως κτίρια. Οι Ozgur et al., (2016) σε εργασία τους διαπιστώνουν την εξαρτημένη σχέση μεταξύ τιμής και του αριθμού μπάνιων, με θετική συσχέτιση. Σημειώνεται ότι η μελέτη έγινε σε δείγμα 480 σπιτιών στην βορειοδυτική Ιντιάνα των ΗΠΑ ενώ οι Anselin & Gracia (2008) παρατήρησαν θετική επίδραση μεταξύ αριθμού μπάνιων και τιμής ακινήτου (σελ. 21).

Η μεταβλητή **κουζίνα** είναι μια ποιοτική, δίτιμη, ονομαστική, μεταβλητή. Πρόκειται για ένα δομικό χαρακτηριστικό που αφορά κυρίως τις κατοικίες. Η ύπαρξής της είναι απαραίτητη σε ένα σπίτι, γι' αυτό και η μεταβλητή δεν έχει τόσο νόημα να κατηγοριοποιηθεί ως ύπαρξη ή μη ύπαρξη της, ενώ θα μπορούσε να κατηγοριοποιηθεί, σύμφωνα με τη πρόταση του Mach, ως φωτεινή, σκοτεινή, μη απομονωμένη λαμβάνοντας ψευδομεταβλητή με αντιστοίχηση σε τιμές 1,2,3 κ.ο.κ. (Mach, 2017)

Η μεταβλητή της **θέρμανσης** είναι ένα από τα χαρακτηριστικά που κρίνει την γενικότερη ενεργειακή απόδοση του κτιρίου. Μια αντίστοιχη μεταβλητή θα μπορούσε να είναι η μεταβλητή της **ψύξης.** Πρόκειται για ένα κατασκευαστικό χαρακτηριστικό. Είναι μία ποιοτική μεταβλητή που μπορεί να λειτουργήσει ως δίτιμη, ονομαστική με διάκριση την ύπαρξη και τη μη ύπαρξη ή ακόμα και να λαμβάνει παραπάνω από δύο διακριτές τιμές μετρούμενη σε τακτική κλίμακα ώστε τα ακίνητα να ιεραρχούνται ως προς το είδος θέρμανσης και ψύξης που διαθέτουν και στη συνέχεια να γίνεται η ποσοτικοποίηση τους. Δυνατή είναι η κατηγοριοποίησή τους ως: καθόλου θέρμανση, κεντρική θέρμανση, τζάκι, αυτόνομη θέρμανση, αυτόνομη θέρμανση με τζάκι, θέρμανση με κλιματιστικό κλπ, ενώ οι ψευδομεταβλητές που δημιουργούνται αντιστοιχίζονται σε τιμές 0,1,2,3 κλπ., σε σχέση με το ενεργειακό κριτήριο, με αύξουσα σειρά από την δυσμενέστερη στην ευμενέστερη. Αντίστοιχα λειτουργεί και η μεταβλητή της ψύξης, ενώ θα μπορούσε να θεωρηθεί μια μικτή μεταβλητή που να αφορά και τις δύο προαναφερθείσες μεταβλητές με την έννοια του κλιματισμού. Οι Benjamin & al, (2004) σε μελέτη τους διαπιστώνουν θετική συσχέτιση της αξίας ενός ακινήτου με την ύπαρξη τζακιού. Σύμφωνα με την Στάμου (2016) οι μεταβλητές αυτόνομης θέρμανσης και τζακιού, αντίστοιχα συμμετέχουν στην εξίσωση ηδονικών τιμών με θετική συσχέτιση, ενώ οι Anselin et al., (2008) παρατήρησαν θετική επίδραση των μεταβλητών τζάκι, σύστημα θέρμανσης και ψύξης, στη διαμόρφωση της αξίας του ακινήτου.

Η μεταβλητή **αποθήκη** είναι μια δίτιμη ονομαστική, ποιοτική μεταβλητή, η

οποία παίρνει ως τιμές, την ύπαρξη και τη μη ύπαρξη. Πρόκειται για ένα κατασκευαστικό χαρακτηριστικό. Για την συμμετοχή της στην διαδικασία της στατιστικής ανάλυσης απαιτείται η ποσοτικοποίηση της, ώστε να αντιστοιχίζεται ένας αριθμός σε κάθε μία από τις δύο περιπτώσεις. Έτσι η ψευδομεταβλητή (dummy variable) με τις τιμές αντίστοιχα 0 και 1 (0 για τη μη ύπαρξη και 1 για την ύπαρξη) λειτουργεί στην στατιστική ανάλυση. Η μεταβλητή της αποθήκης θα μπορούσε να προσδιοριστεί και ως μεταβλητή **εμβαδού αποθήκης**, δηλαδή ως αναλογική, ποσοτική, συνεχής, συμμετέχοντας στην εξίσωση της παλινδρόμησης.

Οι μεταβλητές **βεράντα, αυλή,** μπορεί να λειτουργήσουν ως μία ενιαία μεταβλητή ή ως ξεχωριστές, διακριτές μεταξύ τους, στο μοντέλο πρόβλεψης και συσχετισμού με την αξία του ακινήτου. Είναι δίτιμες, ονομαστικές ποιοτικές, ενώ πρόκειται για δομικά χαρακτηριστικά που αφορούν κυρίως κατοικίες. Για την συμμετοχή τους στην διαδικασία της στατιστικής ανάλυσης απαιτείται η ποσοτικοποίηση τους, ώστε να αντιστοιχίζεται ένας αριθμός σε κάθε μία από τις δύο περιπτώσεις. Έτσι η ψευδομεταβλητή (dummy variable) με τις τιμές αντίστοιχα 0 και 1 (0 για τη μη ύπαρξη και 1 για την ύπαρξη) λειτουργεί στην διαδικασία της στατιστικής ανάλυσης. Η μεταβλητές θα μπορούσαν να προσδιοριστούν και ως μεταβλητές **εμβαδού βεράντας ή αποθήκης**, δηλαδή ως αναλογικές, ποσοτικές, συνεχείς, συμμετέχοντας στην διαδικασία με διαφορετικό τρόπο. Σύμφωνα με μελέτη των Raslanas et al., (2005) για την περιοχή του Λονδίνου, η μεταβλητή μορφοποιήθηκε σε ευρύτερο πλαίσιο επιπρόσθετων χώρων: μπαλκόνι, μεγάλο μπαλκόνι, ταράτσα, μικρός κήπος, κοινόχρηστος κήπος και ιδιωτικός κήπος .

Η μεταβλητή **παρκινγκ** είναι μια μεταβλητή που αναφέρεται στα δομικά χαρακτηριστικά του ακινήτου. Είναι μια δίτιμη, ονομαστική, ποιοτική μεταβλητή, η οποία παίρνει ως τιμές, την ύπαρξη και τη μη ύπαρξη. Για την συμμετοχή της στην διαδικασία της στατιστικής ανάλυσης ποσοτικοποιείται με τις ψευδομεταβλητές (dummy variables) και τιμές αντίστοιχα 0 και 1 (0 για τη μη ύπαρξη και 1 για την ύπαρξη). Οι Chrostek et al.,(2013) σε μελέτη τους επισημαίνουν τον θετικό συσχετισμό της τιμής με την μεταβλητή ύπαρξης γκαράζ. Οι Benjamin, & al (2004) σε μελέτη τους διαπιστώνουν θετική συσχέτιση της αξίας ενός ακινήτου με την αύξηση των τετραγωνικών του γκαράζ. Ως μεταβλητή **εμβαδού γκαράζ** λειτουργεί με αναλογική, ποσοτική και συνεχή μορφή, συμμετέχοντας στην διαδικασία με διαφορετικό τρόπο.

Η μεταβλητή **σοφίτα** είναι και αυτή μια μεταβλητή που αναφέρεται στα δομικά

χαρακτηριστικά του ακινήτου. Είναι μια δίτιμη, ονομαστική, ποιοτική μεταβλητή, η οποία παίρνει ως τιμές την ύπαρξη και τη μη ύπαρξη ενώ και εδώ ισχύουν τα αναφερόμενα στο παρκινγκ σχετικά με την ποσοτικοποίηση της και τον χαρακτηρισμό της μέσω μιας ψευδομεταβλητής με τιμές αντίστοιχα 0 και 1 (0 για τη μη ύπαρξη και 1 για την ύπαρξη). Ως μεταβλητή **εμβαδού σοφίτας** από την στιγμή που δεν συμμετέχει στα τετραγωνικά του υπολοίπου χώρου, λειτουργεί ως ποσοτική, συνεχής.

Η μεταβλητή **πισίνα** ανήκει στα δομικά χαρακτηριστικά του ακινήτου . Η ύπαρξής της συμπεριφέρεται θετικά ως προς την τιμή και είναι μια δίτιμη ονομαστική, ποιοτική, μεταβλητή, η οποία παίρνει ως τιμές, την ύπαρξη και τη μη ύπαρξη. Η ψευδομεταβλητή (dummy variable) συμμετέχει στη διαδικασία με την μορφή 0 και 1 (0 για τη μη ύπαρξη και 1 για την ύπαρξη). Σύμφωνα με τους Sirmans et al.,συμμετέχει πάντα με θετική συσχέτιση ως προς την τιμή στα ηδονικά μοντέλα και η παρουσία της είναι σημαντική, ιδιαίτερα στις προς βορρά της πόλης ιδιοκτησίες και σε περιοχές με θερμά κλίματα (Sirmans et al., 2003).

Η μεταβλητή **πολεοδομικές αυθαιρεσίες ως** δίτιμη ονομαστική, ποιοτική, μεταβλητή, η οποία παίρνει ως τιμές την ύπαρξη και τη μη ύπαρξη, ανήκει στα δομικά χαρακτηριστικά το ακινήτου. Ακίνητα που έχουν αυθαιρεσίες, βάσει του ΝΟΚ για τον ελλαδικό χώρο αποφεύγονται, με άμεση επίδραση αυτού στη διαμόρφωση της αγοραίας αξίας τους.

Όσο αφορά τη μεταβλητή προσδιορισμού του **είδους ακινήτου, π**ρόκειται για ένα κατασκευαστικό χαρακτηριστικό, μια ποιοτική μεταβλητή παίρνοντας παραπάνω από δύο διακριτές τιμές που μπορεί να μετρηθεί σε ονομαστική ή σε τακτική κλίμακα. Τα ακίνητα ιεραρχούνται ως προς το είδος και στη συνέχεια γίνεται η ποσοτικοποίηση τους. Δυνατή είναι η κατηγοριοποίησή τους ως: διαμέρισμα, μεζονέτα, μονοκατοικία, αρχοντικό ενώ οι τιμές της μεταβλητής αντιστοιχίζονται σε τιμές 1,2,3.

Η μεταβλητή **κατάσταση ακινήτου** αφορά ένα κατασκευαστικό χαρακτηριστικό και είναι ποιοτική, ιεραρχική μεταβλητή που μπορεί να πάρει παραπάνω από δύο διακριτές τιμές. Τα ακίνητα ιεραρχούνται ως προς την κατάστασή τους και κατηγοριοποιούνται με τις ψευδομεταβλητές που δημιουργούνται, ώστε να αντιστοιχίζουν τιμές ανάλογες με την κατάσταση του ακινήτου, κακή, μέτρια καλή, κ.ο.κ. Σύμφωνα με μελέτη για την περιοχή του Λονδίνου διαπιστώθηκε θετική συσχέτιση μεταξύ τιμής και καλής κατάστασης του ακινήτου (Raslanas et al., 2006).

Οι Chin et al., (2002) διαπιστώνουν ότι λίγες μελέτες έχουν γίνει σχετικά με την ποιότητα κατασκευής των ακινήτων και αυτό έχει σχέση με την δυσκολία συλλογής στοιχείων, ενώ παράλληλα διαπιστώνουν ότι η κατάσταση του ακινήτου έχει ισοβαρή ρόλο με τον αριθμό δωματίων και του εμβαδόν όσο αφορά την τιμή του.

2.3.2 Ανεξάρτητες Μεταβλητές Θέσης

Η μεταβλητή **θέση ακινήτου (ως γωνιακό ή όχι)** αποτελεί μια μεταβλητή που εντάσσεται στην γενική κατηγορία των μεταβλητών θέσης. Είναι μία δίτιμη, ονομαστική, ποιοτική μεταβλητή, η οποία ως πληροφορία δίνει τη θέση του ακινήτου σε σχέση με την οικοδομή.

Η μεταβλητή **θέα** αποτελεί μια σημαντική μεταβλητή θέσης. Στη βιβλιογραφία συνήθως αποδίδεται με τη χρήση δίτιμης, ονομαστικής, ποιοτικής μεταβλητής και στις περισσότερες των περιπτώσεων έχει θετική επίδραση στις αξίες γης ενώ ως διατάξιμη (ιεραρχική) και πολυεπίπεδη μπορεί να διαταχθεί και να διαβαθμιστεί. Είναι μια ποιοτική μεταβλητή και ως πολυεπίπεδη μπορεί να πάρει τιμές όπως κακή, τυπική, καλή, εξαιρετική και αυτές να αντιστοιχιστούν σε αριθμητικές τιμές μέσω ψευδομεταβλητών. Θα μπορούσε ακόμα να αποτυπωθεί και ως απόσταση από το σημείο ενδιαφέροντος. Όπως αναφέρει η Στάμου (2016) σε πολλές περιπτώσεις προκύπτει ότι οι αγοραστές προτιμούν τις περιοχές κοντά ή με θέα τη θάλασσα, τις λίμνες, το βουνό, ενώ είναι πρόθυμοι να πληρώσουν παραπάνω για τα ακίνητα που βρίσκονται σε αυτές τις τοποθεσίες. Σημειώνεται ότι τα διαμερίσματα των ψηλότερων ορόφων με πρόσοψη σε κύρια οδό είναι πιθανότερο να έχουν καλύτερη θέα από τα ακριβώς ίδια διαμερίσματα χαμηλότερων ορόφων και αυτό αντικατοπτρίζεται στην αξία τους. Σ' αυτή την περίπτωση υπάρχει συσχέτιση και των τριών μεταβλητών, θέα, όροφος, τιμή. Ισχυρή συσχέτιση μεταξύ της θέας και του ορόφου του ακινήτου βρήκαν σε μελέτη τους και οι So et al., (1996) διαπιστώνοντας ότι διαμερίσματα στους μεγάλους ορόφους συνήθως έχουν υψηλότερη τιμή σε σύγκριση με αυτά των χαμηλότερων επιπέδων λόγω της θέας που προσφέρουν. Οι Rodriguez et al., (1994) διαπιστώνουν και αυτοί τη θετική συσχέτιση της καλής θέας με τη τιμή, ενώ οι Mak, et al (2009) σε μελέτη τους κατηγοριοποιούν τη θέα σε τρείς κατηγορίες: θέα σε άλλο κτίριο, θέα εμποδιζόμενη και ανεμπόδιστη θέα. Οι Lake et al (1999) αναφέρουν ότι η τιμή του ακινήτου αυξάνεται εάν η θέα του είναι προς την παραλία, ενώ αναφέρουν την ύπαρξη θετικής συσχέτισης με τη θέα προς ποταμό . Οι Chin et al., (2002) διαπιστώνουν και αυτοί θετική συσχέτιση μεταξύ τιμής και θέας

ενώ επισημαίνουν μελέτες των Tse & Love που διαπιστώνουν την αρνητική συσχέτιση τιμής με θέα σε νεκροταφείο. Η θέα σίγουρα έχει σχέση και με το φυσικό περιβάλλον της γειτονιάς του ακινήτου ή ακόμα και των γειτονικών του κτιρίων. Στην περίπτωση αυτή, όσο μεγαλώνει η απόσταση του ακινήτου από το σημείο ενδιαφέροντος τόσο η σχέση θέας τιμής είναι αρνητική.

Η μεταβλητή **απόσταση από συγκοινωνίες** αποτελεί μια μεταβλητή θέσης. Είναι μία από τις βασικότερες παραμέτρους επιρροής των αξιών γης σε μια πόλη. Ιδιαίτερα σε μεγάλες πόλεις, τα δημόσια μέσα μεταφοράς είναι ιδιαίτερα δημοφιλή εξυπηρετώντας τις καθημερινές μετακινήσεις των πολιτών και ο τρόπος έκτασης του δικτύου αστικής συγκοινωνίας συμβάλλει στη διαμόρφωση των τιμών των γειτνιαζόντων σε αυτές ακινήτων. Είναι ποσοτική μεταβλητή αλλά μπορεί να αποδοθεί με τη χρήση ποιοτικής, δίτιμης, ονομαστικής μεταβλητής και στις περισσότερες των περιπτώσεων έχει θετική επίδραση στην αξία των ακινήτων και ως διατάξιμη, πολυεπίπεδη, μπορεί να διαταχθεί και να διαβαθμιστεί ή ακόμα μπορεί και να λειτουργήσει ως ποσοτική, αναλογική συνεχής μεταβλητή. Ως πολυεπίπεδη μπορεί να πάρει τιμές όπως πολύ κοντά, κοντά, σε μεσαία απόσταση, μακριά και αυτές να αντιστοιχιστούν σε αριθμητικές τιμές μέσω ψευδομεταβλητών. Μπορεί ακόμα και προσδιοριστεί μέσω αποστάσεων που θα τεθούν από τον μελετητή ανά περίπτωση, προσδιοριζόμενες δε και ως χρόνος πρόσβασης. Σε κάποιες περιπτώσεις εξειδικεύεται το είδος της συγκοινωνίας, όπως για παράδειγμα η απόσταση από τους σταθμούς του Μετρό και του Προαστιακού ή από στάση λεωφορείου. Σε μελέτες διαπιστώνεται ότι η επίδραση των σταθμών μέσων μαζικής μεταφοράς στις τιμές των κατοικιών είναι θετική (So et. al, 1996) ενώ και οι Chin et al.,(2002) διαπιστώνουν θετική συσχέτιση της ύπαρξης καλής συγκοινωνίας πλησίον του ακινήτου. Ο Παπαευθυμίου (2013) διαπιστώνει θετική επίδραση της εγγύτητας των συγκοινωνιών του Μετρό με την τιμή του ακινήτου για τη περιοχή της Αθήνας.

Ανάλογα λειτουργεί και η μεταβλητή **απόσταση από κέντρο πόλης.** Αποδίδεται με τη χρήση δίτιμης ονομαστικής μεταβλητής και στις περισσότερες των περιπτώσεων έχει θετική επίδραση στην αξία των ακινήτων, ενώ ως ιεραρχική διατάξιμη, πολυεπίπεδη, μπορεί να διαταχθεί και να διαβαθμιστεί ή ακόμα μπορεί και να λειτουργήσει και ως ποσοτική, αναλογική, συνεχής μεταβλητή. Είναι μια ποιοτική μεταβλητή και ως πολυεπίπεδη μπορεί να πάρει τιμές όπως πολύ κοντά, κοντά, σε μεσαία απόσταση, μακριά και αυτές να αντιστοιχιστούν σε αριθμητικές τιμές μέσω ψευδομεταβλητών. Μπορεί ακόμα και να προσδιοριστεί μέσω αποστάσεων που θα

τεθούν από τον μελετητή ανά περίπτωση, προσδιοριζόμενες δε και ως χρόνος πρόσβασης. Η απόσταση από το κέντρο της πόλης σίγουρα χρήσει εξειδικευμένης μελέτης και συσχέτισης με άλλες μεταβλητές όπως του μεγέθους διαμερισμάτων, της θέας, αλλά και της εγκληματικότητας και του θορύβου. Σύμφωνα με μελέτη των Raslanas, et al., (2005) για την περιοχή του Λονδίνου, διαπιστώθηκε θετική συσχέτιση της μικρής απόσταση από το κέντρο της πόλης με την τιμή του ακινήτου. Ο Παπαευθυμίου (2013) διαπιστώνει και αυτός θετική επίδραση της εγγύτητας στο κέντρο της πόλης με την τιμή του ακινήτου.

Η μεταβλητή **απόσταση από σχολείο.** Σύμφωνα με την βιβλιογραφία ο συνυπολογισμός της απόστασης από εκπαιδευτικές μονάδες έχει σαφή θετική επίδραση στην αγορά ακινήτου. Οι Cho, Poudyalb & Roland, (2007) αναφέρουν σε μελέτη τους την θετική συσχέτιση της μεταβλητής του ποιοτικού σχολείου με την τιμή του ακινήτου στην περιοχή μελέτης, ενώ οι Lake et al., (1999) διαπίστωσαν θετική συσχέτιση τιμής με ύπαρξη κοντινού σχολείου. Σύμφωνα με τους Clauretie &Neill (2000) η ποιότητα των δημόσιων σχολείων έχει μεγάλη επίδραση στις τιμές των κατοικιών. Η μεταβλητή αυτή ως ποιοτική μεταβλητή θέσης, αποδίδεται στις περισσότερες των περιπτώσεων ως διατάξιμη, πολυεπίπεδη και μπορεί να διαταχθεί και να διαβαθμιστεί, ή ακόμα μπορεί κα να λειτουργήσει και ως ποσοτική, αναλογική, συνεχής μεταβλητή. Χρήζει εξέτασης η επίδραση των πολύ κοντινών με το σχολείο ακινήτων, σε συνδυασμό και με άλλους παράγοντες όπως ο θόρυβος που τυχόν προκαλείται.

Η μεταβλητή **απόσταση από πράσινο.** Το αστικό πράσινο - πάρκο αποτελεί μια δημοφιλή παράμετρο που συχνά συναντάται στον προσδιορισμό αυτών που επηρεάζουν τις τιμές κατοικίας σε μια πόλη. Η μεταβλητή αυτή, ως ποιοτική μεταβλητή θέσης, αποδίδεται στις περισσότερες των περιπτώσεων ως διατάξιμη, πολυεπίπεδη και μπορεί να διαταχθεί και να διαβαθμιστεί ή ακόμα μπορεί κα να λειτουργήσει και ως ποσοτική, αναλογική, συνεχής μεταβλητή. **Σ**ύμφωνα με τους Lake et al., (1999) διαπιστώθηκε θετική συσχέτιση τιμής με πάρκο σε κοντινή απόσταση. Οι Kong, Yin & Nakagoshi (2007) θεωρούν ότι τα αστικά πάρκα είναι χώροι με θετική επιρροή στις τιμές των κατοικιών με τις οποίες γειτνιάζουν σε αυτά, αφού δημιουργούν γειτονιές στις οποίες ο καθένας θα ήθελε να κατοικεί, ενώ ως μεταβλητή εισήχθηκε και η συμμετοχή τους ενισχύθηκε με την εξέλιξη των GIS. Όπως αναφέρει η Μηλάκα, (2010) η ποσοτικοποίηση της επίδρασης στις αξίες των ακινήτων μπορεί να γίνει με διάφορους τρόπους, όπως με βάση το είδος του αστικού

πρασίνου, την απόσταση από το κοντινότερο πάρκο, το μέγεθος του κοντινότερου πάρκου ή ακόμα και με βάση δεικτών που συνδυάζουν το μέγεθος και την απόσταση από το πλησιέστερο πάρκο. Οι Irwin and Bockstael (2001) αναφέρουν σε άρθρο τους ότι προηγούμενες μελέτες δεν κατάφεραν να συσχετίσουν την ύπαρξη κοντινού στην ιδιοκτησία πράσινου με αύξηση της τιμής, αναφερόμενοι στις μελέτες των Cheshire & Sheppard, ενώ κατηγοριοποιούν το πράσινο που περιβάλλει ένα ακίνητο: σε πράσινο ιδιωτικό που μπορεί να εξωραϊστεί, σε πράσινο ιδιωτικό προστατευόμενο και σε δημόσιο πράσινο, διαπιστώνοντας παράλληλα θετικό συσχετισμό της τιμής ακινήτου με το πράσινο και στις τρείς περιπτώσεις. Οι Mayor, Lyons & Duffy (2009) επισημαίνουν την θετική συσχέτιση τιμής και ύπαρξης πάρκου σε κοντινή απόσταση για την περιοχή του Δουβλίνου, ενώ αντίστοιχη διαπίστωση κάνουν και οι Kolbe & Wustemann (2015) για την περιοχή της Κολονίας

Η μεταβλητή **υγειονομική εξυπηρέτηση.** Μία από τις παραμέτρους που μπορεί να ληφθούν υπόψη στο μοντέλο είναι η προσβασιμότητα σε υπηρεσίες υγείας και γενικότερα σε υγειονομικές εξυπηρέτησης, είτε αυτές είναι με την μορφή νοσοκομείων, είτε ιατρικών κέντρων, είτε απλών φαρμακείων, δηλαδή σημεία στα οποία κάποιος μπορεί να έχει άμεση πρόσβαση και να λάβει άμεσα υγειονομική βοήθεια. Η μεταβλητή αυτή ως μια μεταβλητή θέσης είναι ποιοτική (ή ποσοτική εάν μετριέται σε απόσταση) και μπορεί να διαταχθεί και να διαβαθμιστεί όσο αφορά την απόσταση ή ακόμα και τον χρόνο που χρειάζεται κάποιος για να έχει πρόσβαση ή ακόμα μπορεί κα να λειτουργήσει και ως ποσοτική, αναλογική, συνεχής μεταβλητή. Όπως αναφέρει η Στάμου (2013), όσον αφορά τα νοσοκομεία, μελέτες έδειξαν ότι αυτά παρουσιάζουν σημαντική θετική επίδραση στις τιμές των ακινήτων.

2.3.3 Ανεξάρτητες Μεταβλητές Γειτονιάς

Η οικιστική πυκνότητα μέσω του αυξημένου πληθυσμού σε μια περιοχή μπορεί να οδηγήσει σε χαμηλότερη ποιότητα αστικού περιβάλλοντος και σε έντονη συμφόρηση. Βλέποντάς το όμως με άλλη ματιά ο αυξημένος πληθυσμός σε μια περιοχή συνεπάγεται και αύξηση των υπηρεσιών στη θέση αυτή με θετική επίδραση στις αξίες των ακινήτων. Από τα παραπάνω συνάγεται το συμπέρασμα ότι η οικιστική πυκνότητα κάθε περιοχής είναι αρκετά σημαντικός παράγοντας και πρέπει να λαμβάνεται υπόψη στα μοντέλα παλινδρόμησης.

Οι Cho et al (2007) σε μελέτη τους παραθέτουν ως μεταβλητή γειτονιάς την **πληθυσμιακή πυκνότητα** αν τετραγωνικό μέτρο στην περιοχή μελέτης, ενώ οι τιμές

των ακινήτων φαίνεται να εξαρτώνται με σχέση από τις γειτονικές τους ιδιοκτησίες έχοντας άμεση χωρική συσχέτιση (Jahanshiri, Buyong, & Shariff, 2011). Η πληθυσμιακή πυκνότητα λειτουργεί ως αναλογική, συνεχής, ποσοτική μεταβλητή. Οι Chrostek et al., (2013) αναφέρουν σε μελέτη τους τις παραμέτρους γειτονιάς, σχετικά με το ποσοστό των κατοίκων της περιοχής που εργάζονται σε υψηλές θέσεις και αυτών που είναι άνεργοι. Παράλληλα αναφέρουν και άλλες μεταβλητές της γειτονιάς που αφορούν το ποσοστό των κατοίκων με πτυχία, το ποσοστό ιδιοκατοίκησης και το ποσοστό φτώχιας.

Οι μεταβλητές της γειτονιάς, αφορούν την ενσωμάτωση ενός συνόλου χαρακτηριστικών που σχετίζονται με κοινωνικό-οικονομικά κριτήρια και φυσικά χαρακτηριστικά της γειτονιάς του. Τέτοιες μεταβλητές που συνήθως χρησιμοποιούνται για το σκοπό αυτό είναι το **τοπικό ποσοστό εγκληματικότητας (**αναλογική, συνεχής, ποσοτική), το **ποσοστό μεταναστών (**αναλογική, διακριτή, ποσοτική), τα **επίπεδα ρύπανσης (**ιεραρχική, ποιοτική) ή τα **επίπεδα θορύβου (**ιεραρχική, ποιοτική). Οι Chin et al., (2002) διαπιστώνουν ότι οι αγοραστές ανησυχούν για τις περιοχές με μεγάλη εγκληματικότητα και αυτό έχει αντίκτυπο στις τιμές των ακίνητων. Σύμφωνα με τους Li et al.,(1980) ο θόρυβος επηρεάζει αρνητικά τις τιμές των ακίνητων, ενώ σύμφωνα με την Στάμου (2013) τα χαρακτηριστικά της γειτονιάς δεν μπορούν να αποτιμηθούν ρητά στην αγορά, αλλά έμμεσα μέσω της ηδονικής τιμολόγησης συγκρίνοντας σπίτια σε διαφορετικές γειτονιές,

Όσο αφορά την **κοινωνικό-οικονομική μεταβλητή (**ιεραρχική, ποιοτική)**,** διαπιστώνεται ότι η κοινωνική τάξη της γειτονιάς έχει αντίκτυπο στις αξίες των ακινήτων. Η ιστορική διαμόρφωση του αστικού χώρου και η πληθυσμιακή κατανομή κατά περιόδους (π.χ. εγκατάσταση μεταναστών) διαμορφώνει το κοινωνικοοικονομικό επίπεδο της κάθε περιοχής με τις κοινωνικές δομές της, τη κοινωνική διαστρωμάτωση και τη σύνθεση του πληθυσμού με την μορφή ηλικίας, φύλλου και οικογενειακής κατάστασης να έχει έμμεση επίδραση στις αγοραίες αξίες, όπως σημαντική επίδραση έχει και το βιοτικό επίπεδο του μέσου ατόμου της περιοχής ως εισόδημα. Οι Cho et al., (2007) σε μελέτη τους παραθέτουν ως μεταβλητή γειτονιάς το **μέσο εισόδημα (**αναλογική) των κατοίκων της περιοχής μελέτης. Όσο αφορά την **εγκληματικότητα** (αναλογική) οι Sirmans et al., (2003) αναφέρουν ότι ο δείκτης μετριέται σε ποσοστό ανά επιφάνεια και έχει αρνητική επίδραση στη τιμή. Τέλος, σε πολλές περιπτώσεις, οι μεταβλητές θέσης και γειτονιάς ομαδοποιούνται καθώς πολλές από τις μεταβλητές της γειτονιάς θα μπορούσαν να θεωρηθούν και

μεταβλητές που περιγράφουν τη γεωγραφική θέση του ακινήτου (Στάμου, 2013).

2.3.4 Λοιπές Μεταβλητές

Χρόνος παραμονής (αναλογική, διακριτή, ποσοτική) της ιδιοκτησίας **προς πώληση**. Επισημαίνεται η αρνητική συσχέτιση της αξίας ενός ακινήτου με το μεγάλο χρονικό διάστημα που αυτό έχει παραμείνει προς πώληση (Benjamin & al ,2004), ενώ παράλληλα όσο μεγαλύτερο είναι το ακίνητο τόσο μεγαλύτερο χρονικό διάστημα αυτό παραμένει προς διάθεση. Οι Sirmans et al., (2003) αναφέρουν ότι εάν ο χρόνος παραμονής ενός ακινήτου στην αγορά είναι μεγάλος και προσδιορίζεται στατιστικά σημαντικός, τότε η εξίσωση είναι γενικά αρνητική. Επιπρόσθετα αναφέρουν ότι μελέτες των Jud, Seaks και Winkler (1996) εξέτασαν η συσχέτιση μεταξύ μεγάλου χρόνου παραμονής στην αγορά προς πώληση και της τιμής του ακινήτου και διαπιστώθηκε ότι όσο μεγαλύτερο χρόνο παραμένει ένα ακίνητο προς πώληση τόσο μεγαλύτερη ήταν και η τιμή του.

Αντικειμενικές αξίες. Μια άλλη αναλογική, συνεχής, μεταβλητή που θα μπορούσε να ενταχθεί στο μοντέλο πρόβλεψης είναι και οι ισχύουσες **αντικειμενικές αξίες της περιοχής,** διότι κυρίως στο ελλαδικό χώρο η αναντιστοιχία πολλές φορές των αντικειμενικών αξιών με τις τιμές της αγοράς, είτε αυτές είναι προς το ευμενέστερο είτε προς το δυσμενέστερο, κατά περίπτωση, επηρεάζουν την τελική τιμή.

Το **αρχαιολογικό βάρος** σε ένα ακίνητο, η **δυνατότητα άμεσης χρηματοδότησης** για την αγορά του, που έχει σχέση με το οικονομικό momentum της εποχής, αλλά και ο **αριθμός των κενών κατοικιών** στην γειτονιά του προς πώληση ακινήτου, είναι ακόμα μερικές από τις μεταβλητές που θα μπορούσαν να λάβουν μέρος στο μοντέλο πρόβλεψης. Μπορεί να λειτουργήσουν ως δίτιμες, αναλογικές, ποσοτικές μεταβλητές. Όσο αφορά το αρχαιολογικό βάρος, αυτό μπορεί να λειτουργήσει και θετικά, εάν πρόκειται για διατηρητέα κτίρια που προτιμούνται από ιδιαίτερες κατηγορίες αγοραστών, συνήθως μεγάλου εισοδήματος. Η δυνατότητα άμεσης χρηματοδότησης είναι ένας παράγοντας που έχει σχέση κυρίως με την γενικότερη οικονομική κατάσταση της χώρας την περίοδο της αγοράς του ακινήτου, ενώ ο αριθμός των κενών διαμερισμάτων μπορεί να παίξει ρόλο καθώς η υπερπληθώρα ακινήτων κατά κανόνα μειώνει τις τιμές σε μια αγορά με στάσιμο αγοραστικό κοινό.

2.4 Συνοπτικός Πίνακας Μεταβλητών

Παρατίθεται συνοπτικός πίνακας (πίνακας 2.1) με την ονομασία και το είδος μεταβλητών.

Πίνακας 2.1

Ονομασία μεταβλητής	Είδος μεταβλητής
Εξαρτημένη Μεταβλητή	
Αξία κατοικίας	αναλογική, συνεχής
Ανεξάρτητες Μεταβλητές	
Μεταβλητές κατασκευαστικών χαρακτηριστικών	
Βασικές Μεταβλητές	
Εμβαδόν	αναλογική, συνεχής, ποσοτική
Παλαιότητα σε έτη	αναλογική, διακριτή (ασυνεχής), ποσοτική
Όροφος	ισοδιαστημική, διακριτή (ασυνεχής), ποσοτική
Λοιπές Μεταβλητές	
Είδος ακινήτου	ποιοτική, ονομαστική (κατηγορική)
Αριθμός υπνοδωματίων	αναλογική, διακριτή (ασυνεχής), ποσοτική
Κατάσταση ακινήτου	ιεραρχική(τακτική), ποιοτική,
Ανελκυστήρας- ασανσέρ	διχοτομική (δίτιμη), ονομαστική, ποιοτική
Αριθμός μπάνιων	αναλογική, διακριτή (ασυνεχής), ποσοτική
Ξεχωριστό ντους	διχοτομική (δίτιμη), ονομαστική, ποιοτική
Κεκλιμένη οροφή	διχοτομική (δίτιμη), ονομαστική, ποιοτική
Διπλός φούρνος	διχοτομική (δίτιμη), ονομαστική, ποιοτική
Κουζίνα	διχοτομική (δίτιμη,) ονομαστική, ποιοτική
Θέρμανση πετρέλαιο	διχοτομική (δίτιμη), ονομαστική, ποιοτική
Θέρμανση φυσικό αέριο	διχοτομική (δίτιμη), ονομαστική, ποιοτική
Ψύξη	διχοτομική (δίτιμη), ονομαστική, ποιοτική

Τζάκι	διχοτομική (δίτιμη), ονομαστική, ποιοτική
Αποθήκη	διχοτομική (δίτιμη), ονομαστική, ποιοτική
Υπόστεγο	διχοτομική (δίτιμη), ονομαστική, ποιοτική
Εμβαδόν αποθήκης	αναλογική, συνεχής, ποσοτική
Βεράντα	διχοτομική (δίτιμη), ονομαστική, ποιοτική
Αυλή	διχοτομική (δίτιμη), ονομαστική, ποιοτική
Εμβαδόν βεράντας	αναλογική, συνεχής, ποσοτική,
Εμβαδόν αποθήκης	αναλογική, συνεχής, ποσοτική,
Παρκινγκ	διχοτομική (δίτιμη), ονομαστική, ποιοτική
Εμβαδόν γκαράζ	αναλογική, συνεχής, ποσοτική ,
Σοφίτα	διχοτομική (δίτιμη), ονομαστική, ποιοτική
Εμβαδόν σοφίτας	αναλογική, συνεχής, ποσοτική
Πισίνα	διχοτομική (δίτιμη), ονομαστική, ποιοτική
Πολεοδομικές αυθαιρεσίες	διχοτομική (δίτιμη), ονομαστική, ποιοτική
Πρόσφατα ανακαινισμένο	διχοτομική (δίτιμη), ονομαστική, ποιοτική

Μεταβλητές θέσης

Θέση ακινήτου	διχοτομική, ονομαστική, ποιοτική (ιεραρχική, ποιοτική)

Συμπληρωματικές Μεταβλητές

Θέα	διχοτομική (δίτιμη), ονομαστική, ποιοτική
ή (ιεραρχική, ποιοτική)	
Απόσταση από συγκοινωνίες	αναλογική, συνεχής, ποσοτική
Κοντά σε συγκοινωνία	διχοτομική (δίτιμη), ονομαστική , ποιοτική
ή (ιεραρχική, ποιοτική)	
Απόσταση από μετρό	αναλογική, συνεχής, ποσοτική
Κοντά σε μετρό	διχοτομική (δίτιμη), ονομαστική, ποιοτική ή (ιεραρχική, ποιοτική)
Απόσταση από κέντρο πόλης	αναλογική, συνεχής, ποσοτική
Κοντά στο κέντρο της πόλης	διχοτομική (δίτιμη), ονομαστική, ποιοτική

ή (ιεραρχική, ποιοτική)

Απόσταση από σχολείο αναλογική, συνεχής, ποσοτική

Κοντά σε σχολείο διχοτομική (δίτιμη), ονομαστική, ποιοτική ή (ιεραρχική, ποιοτική)

Απόσταση από μειονοτικό σχολείο αναλογική, συνεχής, ποσοτική

Κοντά σε μειονοτικό σχολείο διχοτομική (δίτιμη), ονομαστική, ποιοτική)

ή (ιεραρχική, ποιοτική)

Απόσταση από πράσινο αναλογική, συνεχής, ποσοτική

Κοντά σε πράσινο διχοτομική (δίτιμη), ονομαστική, ποιοτική

ή (ιεραρχική, ποιοτική)

Απόσταση από πάρκο αναλογική, συνεχής, ποσοτική

Κοντά σε πάρκο διχοτομική (δίτιμη), ονομαστική, ποιοτική

ή (ιεραρχική, ποιοτική)

Απόσταση από υγειονομική εξυπηρέτηση αναλογική, συνεχής, ποσοτική

Κοντά σε νοσοκομείο – φαρμακείο διχοτομική (δίτιμη), ονομαστική, ποιοτική)

ή (ιεραρχική ποιοτική)

Απόσταση από παραλία αναλογική, συνεχής, ποσοτική)

Κοντά σε παραλία διχοτομική (δίτιμη), ονομαστική, ποιοτική

ή (ιεραρχική ποιοτική)

Μεταβλητές γειτονιάς

Τοπικό ποσοστό εγκληματικότητας αναλογική, διακριτή (ή ασυνεχής), ποσοτική

Ποσοστό μεταναστών αναλογική, διακριτή (ή ασυνεχής), ποσοτική

Μέσο εισόδημα αναλογική, διακριτή (ή ασυνεχής), ποσοτική

Συμπληρωματικές Μεταβλητές για την Δημιουργία του Μοντέλου

Πληθυσμιακή πυκνότητα αναλογική, διακριτή (ή ασυνεχής), ποσοτική

Επίπεδα ρύπανσης ιεραρχική, ποιοτική

Επίπεδα θορύβου ιεραρχική, ποιοτική

Κοινωνικοοικονομική μεταβλητή	ιεραρχική, ποιοτική
Έντονη σεισμικότητα περιοχής	διχοτομική (δίτιμη), ονομαστική, ποιοτική
Απόσταση από χώρους εκτροφής ζώων	αναλογική, συνεχής, ποσοτική)
Εγγύτητα σε χώρους εκτροφής ζώων	διχοτομική (δίτιμη) ονομαστική , ποιοτική ή (ιεραρχική, ποιοτική)
Απόσταση από χώρους ταφής απορριμμάτων ή νεκροταφεία	αναλογική, συνεχής, ποσοτική
Εγγύτητα σε χώρους ταφής	διχοτομική (δίτιμη), ονομαστική, ποιοτική ή (ιεραρχική ποιοτική)
Απόσταση από περιοχές μεγάλων εταιρικών ιδιοκτησιών	αναλογική, συνεχής, ποσοτική
Εγγύτητα σε περιοχές μεγάλων εταιρικών ιδιοκτησιών	διχοτομική (δίτιμη), ονομαστική, ποιοτική ή (ιεραρχική ποιοτική)
Απόσταση από περιοχές υψηλής τάσης ρεύματος	αναλογική, συνεχής, ποσοτική
Εγγύτητα σε περιοχές υψηλής τάσης ρεύματος	διχοτομική (δίτιμη), ονομαστική, ποιοτική ή (ιεραρχική ποιοτική)
Χρόνος παραμονής προς πώληση	αναλογική, διακριτή (ή ασυνεχής), ποσοτική
Αντικειμενικές αξίες	αναλογική, συνεχής, ποσοτική
Αρχαιολογικό βάρος	διχοτομική (δίτιμη), ονομαστική, ποιοτική
Συμπληρωματικές Μεταβλητές για την Δημιουργία του Μοντέλου	
Δυνατότητα δανειοδότησης	διχοτομική (δίτιμη), ονομαστική, ποιοτική
Κενά διαμερίσματα στην περιοχή	διχοτομική (δίτιμη), ονομαστική, ποσοτική

2.5 Συνηθέστερα Χρησιμοποιούμενες Μεταβλητές στα Ηδονικά Μοντέλα

Παρατίθεται συνοπτικός πίνακας (πίνακας 2.2) με την ονομασία και το είδος επίδρασης στην μεταβλητή «Τιμή ακινήτου» των συνηθέστερα χρησιμοποιουμένων μεταβλητών στα ηδονικά μοντέλα.

Πίνακας 2.2

Ονομασία μεταβλητής	Είδος επίδρασης στην μεταβλητή της τιμής του ακινήτου
Μεταβλητές θέσης	
Απόσταση από κέντρο	Συνήθως θετική (η μικρή απόσταση)
Απόσταση από σχολείο	Συνήθως θετική (η μικρή απόσταση)
Απόσταση από πράσινο	Συνήθως θετική (η μικρή απόσταση)
Απόσταση από συγκοινωνίες	Συνήθως θετική (η μικρή απόσταση)
Απόσταση από υγειονομική εξυπηρέτηση	Συνήθως θετική (η μικρή απόσταση)
Θέα	Συνήθως θετική
Δομικές μεταβλητές	
Αριθμός υπνοδωματίων μεγάλος	Συνήθως θετική
Εμβαδόν μεγάλο	Συνήθως θετική
Ύπαρξη υπογείου, ασανσέρ, κλιματισμού	Συνήθως θετική
Υψηλό επίπεδο ορόφου	Συνήθως θετική
Καλή Ποιότητα κατασκευής	Συνήθως θετική
Ύπαρξη πισίνας και λοιπών εξυπηρετήσεων	Συνήθως θετική
Ηλικία μεγάλη	Συνήθως αρνητική
Μεταβλητές γειτονιάς	
Υψηλό εισόδημα κατοίκων	Συνήθως θετική
Υψηλή εγκληματικότητα	Συνήθως αρνητική
Θόρυβος μεγάλος	Συνήθως αρνητική

Κεφάλαιο 3^{ο} – Μεθοδολογία

Για την διενέργεια μιας στατιστικής μελέτης απαιτείται η προσέγγιση του προβλήματος μέσα από την στατιστική ανάλυση η οποία απαιτεί γνώση των αρχών της στατιστικής θεωρίας.

Στο πρώτο μέρος του παρόντος κεφαλαίου παρουσιάζεται ο βασικός κορμός των εννοιών της στατιστικής θεωρίας. Ακολουθεί το θεωρητικό υπόβαθρο στην ανάλυση παλινδρόμησης, ενώ στο τέλος περιγράφονται τα βασικά μοντέλα παλινδρόμησης και μετασχηματισμών.

Με την χρήση του θεωρητικού υποβάθρου θα πραγματοποιηθεί σε επόμενο κεφάλαιο η διερεύνηση των μοντέλων παλινδρόμησης, για την εκτίμηση αξιών ακινήτων, καθώς αυτό είναι και το θέμα της παρούσας μεταπτυχιακής διπλωματικής εργασίας .

3.1 Βασικές Έννοιες Στατιστικής Θεωρίας

3.1.1 Πληθυσμός, Στατιστικός Πληθυσμός και Αντικειμενικός Πληθυσμός

Πληθυσμός είναι ένα ευρύ σύνολο ατόμων ή αντικειμένων του οποίου μελετάμε τα στοιχεία ως προς τουλάχιστον ένα μετρήσιμο χαρακτηριστικό, ενώ το πλήθος των στοιχείων του συνόλου είναι το μέγεθος του πληθυσμού (Γκαρούτσος, 2008). Σύμφωνα με τον Ζαφειρόπουλο (2015) είναι ένα ευρύ σύνολο υποκειμένων για το οποίο εξάγουμε συμπεράσματα. Σ**τατιστικός πληθυσμός** είναι το τμήμα του ευρύτερου πληθυσμού που μπορεί να συμπεριληφθεί στην έρευνα, δηλαδή το σύνολο των μετρούμενων υπό μελέτη στοιχείων με κοινά χαρακτηριστικά (π.χ. οι κάτοικοι της Ελλάδας), ενώ σύμφωνα με τους Χαλικιά**,** Μανωλέσσου & Λάλου (2015) είναι το σύνολο των ατόμων ή αντικειμένων στα οποία αναφέρονται οι παρατηρήσεις και τα οποία καλούνται **στατιστικές μονάδες**. **Αντικειμενικός πληθυσμός** σύμφωνα με τον Παπαγεωργίου (2015) είναι το σύνολο των μελών για το οποίο μας ενδιαφέρει να εξάγουμε συμπεράσματα και σ**τοιχείο ή μονάδα** δειγματοληψίας είναι η μονάδα για την οποία συλλέγονται τα δεδομένα και μπορεί να επιλεγεί σε κάποιο στάδιο της δειγματοληψίας.

3.1.2 Δείγμα, Δειγματοληψία

Δειγματοληψία είναι το μέρος της στατιστικής που διαπραγματεύεται μεθόδους

επιλογής δείγματος, τα σφάλματα δειγματοληψίας και το κατάλληλο μέγεθος δείγματος, ενώ η στατιστική έρευνα με βάση τα δείγματα ονομάζεται **δειγματοληπτική έρευνα** (Καλαματιανού, 2000). **Δειγματοληπτικό πλαίσιο** είναι ένας κατάλογος ολόκληρου του προς έρευνα πληθυσμού (Ζαφειρόπουλος, 2015)**,** ενώ **δείγμα** είναι κάθε υποσύνολο του πληθυσμού αυτού, που έχει επιλεγεί με τέτοιο τρόπο ώστε να είναι αντιπροσωπευτικό (Ηλιοπούλου, 2015)**. Το μέγεθος του δείγματος** είναι το πλήθος των στοιχείων που το διαμορφώνουν**,** ενώ **τυχαίο δείγμα** είναι ένα αντιπροσωπευτικό και τυχαίο υποσύνολο του υπό μελέτη πληθυσμού, τα στοιχεία του οποίου μελετάμε ως προς τουλάχιστον ένα χαρακτηριστικό. Για τον υπολογισμό του μεγέθους του δείγματος, είναι απαραίτητος ο ορισμός του αποδεκτού περιθωρίου σφάλματος, ώστε να προσδιοριστεί το διάστημα εμπιστοσύνης πέριξ της εκτίμησης. Η **δειγματοληψία** είναι η συγκέντρωση στοιχείων από το μέρος του πληθυσμού που χρησιμοποιεί το δείγμα ώστε να εξάγει συμπεράσματα για το σύνολο του, ενώ η **τυχαία δειγματοληψία** επιτυγχάνεται όταν κάθε πιθανή ομάδα του πληθυσμού έχει την ίδια πιθανότητα επιλογής (Ηλιοπούλου, 2015)

Τα βασικά βήματα σε μια διαδικασία δειγματοληπτικής έρευνας είναι κατά σειρά: ο προσδιορισμός του αντικειμένου της έρευνας, ο ορισμός του υπό δειγματοληψία πληθυσμού και του πληθυσμιακού στόχου, ο ορισμός των δειγματοληπτικών μονάδων, ο προέλεγχος και η οργάνωση της εργασίας πεδίου, η επιλογή της μεθόδου, ο προσδιορισμός του μεγέθους και η επιλογή του δείγματος και τέλος η στατιστική ανάλυση των δεδομένων. Για τον σκοπό της δειγματοληπτικής έρευνας χρησιμοποιούνται κυρίως οι μέθοδοι της απλής τυχαίας δειγματοληψίας, της στρωματοποιημένης δειγματοληψίας, της συστηματικής δειγματοληψίας, και της δειγματοληψίας κατά συστάδες (Καλαματιανού,2000).

3.1.2.1 Μέγεθος Δείγματος

Εξετάζοντας την δειγματική κατανομή μιας εκτιμήτριας διαπιστώνουμε ότι κάποια δείγματα δίνουν σημειακές εκτιμήσεις πλησιέστερα στην πραγματική τιμή της παραμέτρου από ότι άλλα, ενώ τα διάφορα κριτήρια επιλογής των εκτιμητριών μας πληροφορούν για την συμπεριφορά της δειγματικής κατανομής της, αλλά δεν μας δίνουν την πληροφορία της θέσης μιας συγκεκριμένης εκτίμησης ως προς την πραγματική τιμή της παραμέτρου. Με αυτόν τον τρόπο δεν γνωρίσουμε εάν το δείγμα που επιλέχθηκε έχει μικρή ή μεγάλη πιθανότητα να δώσει εκτίμηση κοντά ή μακριά από την πραγματική τιμή της παραμέτρου. Απαραίτητο λοιπόν είναι κάθε σημειακή

εκτίμηση να συνοδεύεται από κάποιο μέτρο του περιεχόμενου πιθανού σφάλματος σε σχέση με την πραγματική τιμή. Αυτό το μέτρο είναι το διάστημα εμπιστοσύνης και τα όρια του ονομάζονται όρια εμπιστοσύνης (Καλαματιανού, 2000). Εκτενέστερη αναφορά και ανάλυση για τα διαστήματα εμπιστοσύνης θα γίνει σε παρακάτω κεφάλαιο.

3.1.3 Κατανομές Συχνοτήτων

Στο χώρο της περιγραφικής στατιστικής οι κατανομές συχνοτήτων αποτελούν μια μέθοδο συστηματοποίησης, σύνοψης και παρουσίασης ποιοτικών ή ποσοτικών δεδομένων. Η κατανομή συχνοτήτων ενός συνόλου δεδομένων απεικονίζεται με την μορφή πίνακα δύο στηλών (διανυσματικός πίνακας), όπου στην πρώτη στήλη παρουσιάζονται οι τιμές της μεταβλητής και στην δεύτερη οι συχνότητες στις οποίες εμφανίζονται οι αντίστοιχες τιμές της. Οι κατανομές συχνοτήτων χωρίζονται σε δύο κατηγορίες: στις κατανομές απόλυτων συχνοτήτων (όπου οι συχνότητες είναι απόλυτοι αριθμοί) και στις κατανομές σχετικών συχνοτήτων (όπου οι συχνότητες είναι σχετικοί αριθμοί, λόγοι, αναλογίες, ποσά) (Καλαματιανού, 2000).

Η **κατανομή συχνότητας** (frequency distribution) είναι ένας τρόπος περιγραφής μιας μεταβλητής, παρουσιάζοντας το πόσο συχνά εμφανίζεται μια τιμή στο σύνολο των δεδομένων. Ενώ για τα ποιοτικά δεδομένα δεν απαιτείται ιδιαίτερη προετοιμασία για την κατασκευή της κατανομής συχνοτήτων, για τις ποσοτικές μεταβλητές απαιτείται η ομαδοποίηση των τιμών της μεταβλητής.

Η δημιουργία πίνακα κατανομής συχνοτήτων ξεκινά με την καταμέτρηση της συχνότητας εμφάνισης κάθε τιμής δηλαδή της **απόλυτης συχνότητας (absolute frequency** ή απλά **frequency)**. Ακολουθεί ο υπολογισμός της **σχετικής συχνότητας (relative frequency)**, δηλαδή της ποσοστιαίας κατανομής των συχνοτήτων και της **αθροιστικής συχνότητας (cumulative frequency)**, που προκύπτει όταν η συχνότητα κάθε κατηγορίας περιλαμβάνει και τις συχνότητες όλων των προηγούμενων κατηγοριών. Θα πρέπει να σημειωθεί ότι το άθροισμα όλων των σχετικών συχνοτήτων είναι ίσο με 1 (Γκαρούτσος,2008). Για την δημιουργία **συχνοτήτων ποσοτικών δεδομένων,** όπου η ποσοτική μεταβλητή μετατρέπεται σε ιεραρχική, απαιτείται πρώτα η ομαδοποίηση των τιμών της σε **αποκλειστικές** (τιμές που ανήκουν σε μία ομάδα και δεν μπορούν να ανήκουν σε άλλη) και **εξαντλητικές** (όπου περιλαμβάνονται όλες τις τιμές της μεταβλητής) τάξεις ή κλάσεις (Ηλιοπούλου, 2017).

3.1.4 Μέτρα Κεντρικής Τάσης Περιγραφικής Στατιστικής

Η περιγραφική Στατιστική (Descriptive Statistics) περιλαμβάνει τις μεθόδους που αφορούν τη σύνοψη, περιγραφή και παρουσίαση των δεδομένων. Αφορά τις διαδικασίες για την οργάνωση των δεδομένων και τη συνοπτική παρουσίασή τους. Τα βασικά αριθμητικά μέτρα που χαρακτηρίζουν διάφορες ιδιότητες του δείγματος, ομαδοποιούνται βασικά σε τρείς κατηγορίες που είναι τα μέτρα θέσης, διασποράς και σχήματος κατανομής. Τα μέτρα θέσης, που είναι η μέση τιμή (mean), η διάμεσος (median), η κορυφή (mode) το πρώτο (first quantile) και τρίτο τεταρτημόριο (third quartile), έχουν σχέση με την θέση των δεδομένων του δείγματος. Τα μέτρα διασποράς, που είναι η διασπορά (variance), η τυπική απόκλιση (standard deviation), η τυπική απόκλιση μέσου (standard error of mean), η μέγιστη (maximum) και ελάχιστη (minimum) τιμή και το ενδοτεταρτομοριακό εύρος (interquardile range), αφορούν το σχήμα της κατανομής των δεδομένων. Τα μέτρα σχήματος, που είναι ο συντελεστής ασυμμετρίας (skewness) και ο συντελεστής κυρτότητας (kurtosis), δείχνουν τον βαθμό ασυμμετρίας που κατανέμονται οι τιμές του δείγματος γύρω από κάποια τιμή (Νικήτα, 2012).

Η τάση συγκέντρωσης των δεδομένων γύρω από μια κεντρική τιμή που εκφράζει την κεντρική ή μέση τιμή τους, είναι ένα χαρακτηριστικό των κατανομών και ονομάζεται κεντρική τάση, ενώ ο βαθμός που ένα σύνολο δεδομένων παρουσιάζει αυτό το χαρακτηριστικό προσδιορίζεται από τα περιγραφικά μέτρα κεντρικής τάσης ή θέσης (measures of central tendency). Τα μέτρα αυτά εκφράζουν το μέσο όρο, λαμβάνοντας υπόψη όλες τις τιμές ποσοτικών δεδομένων, αποδίδοντας μια μόνο τιμή. Τα μέτρα κεντρικής τάσης είναι ο **αριθμητικός μέσος** (arithmetic mean), η **διάμεσος** (median) και η **επικρατούσα τιμή** (mode) (Καλαματιανού, 2000). Ο **χωρικός μέσος** (mean center) είναι αντίστοιχος του αριθμητικού, εξαρτάται από την γεωγραφία του χώρου, ενώ χρησιμοποιείται για την ανάλυση δεδομένων με γεωγραφική αναφορά. Ο **χωρικός διάμεσος** (median center) αποτελεί επέκταση της απλής μονοδιάστατης διαμέσου στο γεωγραφικό χώρο. Οι δύο τελευταίοι είναι δείκτες χωρικής κεντρικότητας, αφορούν κατανομές σημείων στον γεωγραφικό χώρο και είναι γεωστατιστικοί δείκτες (Ηλιοπούλου, 2015).

3.1.5 Μέτρα, Δείκτες Διασποράς

Η **Μεταβλητότητα** (variability) είναι ένα χαρακτηριστικό της κατανομής που αποκαλύπτει την έκταση στην οποία οι τιμές των δεδομένων μεταβάλλονται και

διασκορπίζονται μεταξύ τους σε σχέση με την κεντρική τιμή (Καλαματιανού, 2000). Ουσιαστικά περιγράφει μια κατανομή και είναι η διαφοροποίηση των τιμών μιας μεταβλητής μεταξύ των παρατηρήσεων. Μετριέται με τα **μέτρα διασποράς** (measures of dispersion) που παρέχουν πληροφορίες για την εξάπλωση των τιμών σε σχέση με τα μέτρα κεντρικής τάσης, δίνοντας μια εικόνα για το πόσο συγκεντρωμένες είναι οι παρατηρήσεις σε ένα σύνολο δεδομένων (Αβούρης, Κατσάνος, Τσέλιος & Μουστάκας, 2015) Το **εύρος** (range), είναι η διαφορά μεταξύ της μικρότερης και της μεγαλύτερης τιμής ενός συνόλου δεδομένων, η **μέση απόκλιση** (mean deviation) είναι το άθροισμα των απόλυτων τιμών των αποκλίσεων, διαιρεμένο με τον αριθμό των παρατηρήσεων και προκύπτει από τον τύπο $\frac{\sum_{i=1}^{n}|x_i-\bar{X}|}{n}$, όπου $|x_i - \bar{X}|$ είναι οι απόλυτες τιμές των αποκλίσεων από τον αριθμητικό μέσο**.**

Η **διακύμανση ή διασπορά του δείγματος** στηρίζεται στον υπολογισμό των αποκλίσεων από τον αριθμητικό μέσο, με τις αποκλίσεις να υψώνονται στο τετράγωνο ώστε οι θετικές παραστάσεις που προκύπτουν να υποδηλώνουν την απόσταση της τιμής από αυτόν. Συμβολίζεται με $\mathbf{s^2}$ και προκύπτει από τον τύπο $\frac{\sum_{i=1}^{n}(x_i-\bar{X})^2}{n-1}$. Η **τυπική απόκλιση** (standard deviation) για το δείγμα είναι η τετραγωνική ρίζα της διακύμανσης και εκφράζει την μέση απόκλιση των επιμέρους τιμών από τον αριθμητικό μέσο, συμβολίζεται με **s** και προκύπτει από τον τύπο $\sqrt{\frac{\sum_{i=1}^{n}(x_i-\bar{X})^2}{n-1}}$. Η τυπική απόκλιση για τον πληθυσμό συμβολίζεται με $\boldsymbol{\sigma}$ και είναι ίση με τον τύπο $\sqrt{\frac{\sum_{i=1}^{n}(x_i-\bar{X})^2}{N}}$. Ο **συντελεστής μεταβλητότητας** (coefficient of variation) ισούται με την τυπική απόκλιση, διαιρεμένη με τον αριθμητικό μέσο, εκφραζόμενη με τη μορφή ποσοστού, συμβολίζεται με **C** και προκύπτει από τον τύπο $\frac{s}{\bar{X}}$ ***100**. Η **τυπική απόσταση** είναι δείκτης χωρικής διασποράς και υπολογίζεται για δεδομένα χάρτη, συμβολίζεται με **SD** και προκύπτει από τον τύπο $\sqrt{\frac{\sum d}{n}}$**,** όπου **d** είναι οι αποστάσεις των σημείων στο χωρικό μέσο και n είναι ο αριθμός των σημείων

(Ηλιοπούλου, 2015 · Γκαρούτσος, 2008).

3.1.6 Δειγματικός Χώρος, Κατανομές Πιθανότητας, Έλεγχοι Υποθέσεων

3.1.6.1 Δειγματικός Χώρος

Σε ένα τυχαίο πείραμα το σύνολο των δυνατών αποτελεσμάτων ορίζουν το **δειγματικό χώρο (sample space)** και τα στοιχεία του ονομάζονται **δειγματικά σημεία (sample point),** ενώ οι συνδυασμοί τους ονομάζονται **ενδεχόμενα ή γεγονότα** (event) (Χατζησπύρος,2013).

3.1.6.2 Κατανομές Πιθανότητας

Κατανομή πιθανότητας (probability distribution), της διακριτής τυχαίας μεταβλητής, δηλαδή της μεταβλητής που περιλαμβάνει τις τιμές που προέρχονται από την τυχαία διαδικασία του πειράματος, είναι το σύνολο των τιμών που μπορεί να πάρει η μεταβλητή μαζί με τις πιθανότητες να προκύψει κάθε μία τιμή ή κάποια διαστήματα τιμών.

Η **κανονική κατανομή (normal distribution)** ή **κατανομή σφαλμάτων** ή **κατανομή του Gauss,** αναφέρεται σε συνεχείς μεταβλητές. Βασίζεται στην ιδέα ότι οι περισσότερες τιμές είναι συγκεντρωμένες γύρω από την μέση τιμή, ενώ οι ακραίες εμφανίζονται λιγότερο συχνά. Θωρείται η σπουδαιότερη κατανομή της θεωρίας των πιθανοτήτων και της στατιστικής, καθώς πολλές τυχαίες μεταβλητές περιγράφονται ικανοποιητικά με αυτή, ενώ στηρίζεται στο Κεντρικό Οριακό Θεώρημα, δηλαδή ότι το άθροισμα και επομένως η μέση τιμή μεγάλου αριθμού ανεξάρτητων παρατηρήσεων ακολουθεί κανονική κατανομή, ανεξάρτητα από το ποια κατανομή ακολουθούν οι παρατηρήσεις (Παπαδόπουλος, χ.η).

Σύμφωνα με τους Δαμιανό, Παπαδάτο & Χαραλαμπίδη (2003) **διωνυμική κατανομή** ονομάζεται η κατανομή της τυχαίας μεταβλητής Χ, όπου Χ είναι ο αριθμός των επιτυχιών σε μια ακολουθία **ν** ανεξάρτητων δοκιμών Bernoulli με σταθερή πιθανότητα επιτυχίας **p** σε όλες τις δοκιμές και συμβολίζεται με **B(ν,p).** Η διωνυμική κατανομή, αφορά διακριτές μεταβλητές που λαμβάνουν δύο μόνο τιμές. Περιγράφει ένα τυχαίο επαναλαμβανόμενο πείραμα με δύο πιθανά ενδεχόμενα που ονομάζονται επιτυχία και αποτυχία, χωρίς να έχει σημασία ποιο είναι η επιτυχία, ενώ κάθε δοκιμή του πειράματος έχει μια γνωστή πιθανότητα επιτυχίας **p**, ή πιθανότητα

αποτυχίας είναι **q=1-p**. Τα πιθανά αποτελέσματα του πειράματος αποτελούν τιμές μια νέας μεταβλητής που ονομάζεται διωνυμική και η κατανομή πιθανοτήτων που προσομοιάζει τη συμπεριφορά των τιμών μιας τέτοιας μεταβλητής ονομάζεται διωνυμική κατανομή πιθανοτήτων (binominal probability distribution) (Καλαματιανού, 2000).

Η **κατανομή Poisson** αφορά διακριτές μεταβλητές οι οποίες λαμβάνουν δύο τιμές, με γνωστή πιθανότητα επιτυχίας ή αποτυχίας. Προτείνεται σε περιπτώσεις που η πιθανότητα επιτυχίας είναι πολύ μικρή και το πλήθος των δοκιμών πολύ μεγάλο. Σύμφωνα με την Καλαματιανού (2000) προκύπτει σε ειδικές περιπτώσεις συνδεόμενες με την θεωρεία της δειγματοληψίας, προσδιορίζεται δε ως η κατανομή των σπάνιων φαινομένων ή ως η κατανομή ενδεχομένων σε διαστήματα (χρόνου, εμβαδού μήκους κλπ). Σύμφωνα με τον Παπαδόπουλο (χ.η) η κατανομή Poisson ορίζεται ως η οριακή κατανομή της διωνυμικής, ενώ η διωνυμική κατανομή προσεγγίζεται από την Poisson εάν το πλήθος τείνει θεωρητικά στο άπειρο και η πιθανότητα επιτυχίας τείνει στο μηδέν.

Η **αρνητική διωνυμική κατανομή** ή **κατανομή Pascal,** είναι μια διακριτή συνάρτηση κατανομής τυχαίας μεταβλητής, που περιγράφει ένα τυχαίο πείραμα με δυο πιθανά αποτελέσματα (επιτυχία - αποτυχία) και πιθανότητα επιτυχίας **p** που επαναλαμβάνεται μέχρι να έχουμε **r** επιτυχίες, ενώ το ενδιαφέρον προσδιορίζεται στην υπόθεση ότι τμήματα δεδομένων είναι το αποτέλεσμα διαφορετικών τιμών εμφάνισης, σε αντίθεση με την αρνητική διωνυμική κατανομή όπου η μέση συχνότητα εμφάνισης είναι σταθερή (Ηλιοπούλου, 2015).

3.1.6.3 Διάστημα Εμπιστοσύνης

Το φάσμα πιθανών τιμών μιας υπό εκτίμηση παραμέτρου ονομάζεται **διάστημα εμπιστοσύνης** (confidence interval) και χρησιμοποιείται για την ασφαλέστερη εκτίμηση μιας παραμέτρου ενός πληθυσμού, με βάση ένα τυχαίο δείγμα του. Είναι το ποσοστό συν ή πλην που συνήθως αναφέρεται σε αποτελέσματα δημοσκόπησης και αποτελεί το διάστημα τιμών της προς εκτίμηση μεταβλητής στο οποίο περιλαμβάνεται με συγκεκριμένη πιθανότητα η τιμή της παραμέτρου του πληθυσμού (Ηλιοπούλου, 2015). Για παράδειγμα, σε περίπτωση που τεθεί ένα ερώτημα σε ένα πληθυσμό, εάν χρησιμοποιήσουμε ένα διάστημα εμπιστοσύνης 4% και το 47% του δείγματός μας επιλέξει μια απάντηση, τότε μπορούμε να είμαστε σίγουροι, ότι αν

είχαμε θέσει το ερώτημα σε ολόκληρο το σχετικό πληθυσμό, το ποσοστό για την συγκεκριμένη απάντηση θα ήταν μεταξύ 43% (47-4) και 51% (47 + 4).

Το επίπεδο εμπιστοσύνης (**confidence level)** μας δείχνει πόσο σίγουροι μπορούμε να είμαστε. Εκφράζεται ως ποσοστό και αντιπροσωπεύει την πιθανότητα το πραγματικό ποσοστό του πληθυσμού που θα επιλέξει την απάντηση να βρίσκεται μέσα στο διάστημα εμπιστοσύνης . Το επίπεδο εμπιστοσύνης ουσιαστικά εκφράζει το σφάλμα το οποίο είναι αποδεκτό στην εκτίμηση των αποτελεσμάτων (Ηλιοπούλου, 2017). Για παράδειγμα το επίπεδο εμπιστοσύνης 95% σημαίνει ότι μπορούμε να είμαστε κατά 95% βέβαιοι, ενώ το επίπεδο εμπιστοσύνης 99% σημαίνει ότι μπορούμε να είμαστε κατά 99% αντίστοιχα.

Συνδέοντας το επίπεδο εμπιστοσύνης με το διάστημα εμπιστοσύνης, επιτυγχάνουμε κατά 95% την σιγουριά ότι το πραγματικό ποσοστό του πληθυσμού είναι μεταξύ 43% και 51% (αναφερόμενοι στο παραπάνω παράδειγμα). Όσο ευρύτερο είναι το διάστημα εμπιστοσύνης που είμαστε διατεθειμένοι να αποδεχθούμε, τόσο πιο σίγουρο είναι ότι η εκτίμηση για το σύνολο του πληθυσμού βρίσκεται μέσα σε αυτό το εύρος.

3.1.6.4 Στατιστικές Υποθέσεις

Στην εκτιμητική στατιστική ο βασικός στόχος είναι η εκτίμηση μιας ή περισσοτέρων παραμέτρων. Ως πρόβλημα λοιπόν τίθεται η αποδοχή ή απόρριψη μιας στατιστικής εικασίας που αφορά τον πληθυσμό με δεδομένα περιθώρια σφάλματος (Δαμιανού κ.ά, 2010). Οι **στατιστικές υποθέσεις** ως διαδικασίες που ακολουθούνται για τον έλεγχο θεωρητικών υποθέσεων σχετίζονται με την εξαγωγή συμπερασμάτων σε ότι αφορά τις παραμέτρους του πληθυσμού που ερευνάται. Τα συμπεράσματα αυτά προκύπτουν αφού διαμορφωθούν και ελεγχθούν οι στατιστικές υποθέσεις. Οι **έλεγχοι υποθέσεων** πραγματοποιούνται με στατιστικούς ελέγχους, που εξετάζουν το εάν δεν υπάρχει στατιστικά σημαντική διαφορά μεταξύ ενός, δυο ή περισσοτέρων δειγμάτων, αν τα δείγματα αυτά αντιπροσωπεύουν διαφορετικούς πληθυσμούς. Η γενική ιδέα της διαδικασίας στατιστικού ελέγχου είναι ότι αρχικά θα πρέπει να τεθεί η **μηδενική υπόθεση (null hypothesis) H_0**, η οποία είναι και αυτή για την οποία αμφιβάλλουμε και την οποία ελέγχουμε, ενώ διατυπώνεται σε αυτή το αντίθετο από αυτό που πιστεύουμε ότι συμβαίνει στην πραγματικότητα (Ηλιοπούλου, 2015). Ποιο συγκεκριμένα ας υποθέσουμε ότι έχουμε μια κατανομή $F(x,\theta)$, όπου $\chi \in R$ και $\theta \in \Theta$ του πληθυσμού που εξετάζουμε. Εάν υποθέσουμε ότι τα δύο υποσύνολα του

παραμετρικού χώρου Θ, δηλαδή τα Θ_0 και Θ_1 με $\Theta_0 \cap \Theta_1 \neq 0$, τότε η στατιστική υπόθεση $\mathbf{H_0}$ (μηδενική υπόθεση) είναι ότι το $\theta \in \Theta_0$. Η εικασία ότι $\theta \in \Theta_1$ είναι η εναλλακτική στατιστική υπόθεση $\mathbf{H_1}$. Η διαδικασία που ακολουθούμε για τον έλεγχο της μηδενικής υπόθεσης έναντι της εναλλακτικής είναι να καθορίσουμε τις δύο υποθέσεις, να καθορίσουμε το επίπεδο σημαντικότητας του ελέγχου ως μια μικρή πιθανότητα $\alpha \in (0,1)$ η οποία και προσδιορίζεται ως η μέγιστη αποδεκτή πιθανότητα εσφαλμένης απόρριψης της μηδενικής υπόθεσης, δηλαδή το επίπεδο σημαντικότητας (significance level), το οποίο καθορίζεται συνήθως με την τιμή α=0,05=5%. Έπειτα λαμβάνουμε ένα τυχαίο δείγμα που υποτίθεται ότι δικαιολογεί την μηδενική υπόθεση, να κατασκευάζουμε την κατάλληλη στατιστική συνάρτηση, αφού έχουμε ορίσει τις κρίσιμες τιμές (critical values) και ανάλογα με την τιμή της στατιστικής συνάρτησης, εν σχέση με τα όρια των σταθερών αποκοπής, κρίσιμη τιμή ή όριο απόρριψης και τη μορφή της μηδενικής υπόθεσης, την δεχόμαστε ή να την απορρίπτουμε, αποφασίζοντας υπέρ της εναλλακτικής υπόθεσης (Δαμιανού κ.ά, 2010).

Ο έλεγχος των στατιστικών υποθέσεων γίνεται με βάση το σφάλμα τύπου **I** (type I error), όταν γίνεται απόρριψη μιας σωστής μηδενικής υπόθεσης και τύπου **II** (type II error), όταν γίνεται αποδοχή μιας λανθασμένης μηδενικής υπόθεσης . Το σφάλμα τύπου **I** σημαίνει ότι απορρίπτεται η $\mathbf{H_0}$, όταν αυτή είναι αληθής, με πιθανότητα ίση με **α**, ενώ στο σφάλμα τύπου **II** σημαίνει ότι δεχόμαστε την $\mathbf{H_0}$, όταν η $\mathbf{H_1}$ είναι αληθής, με πιθανότητα ίση με **β**. Η τιμή του **α**, όπως προείπαμε είναι η μέγιστη πιθανότητα να κάνουμε σφάλμα τύπου **I**, δηλαδή αποτελεί το ανεκτό επίπεδο σφάλματος τύπου **I** και ονομάζεται **επίπεδο σημαντικότητας** (significance level). Τα συνήθως επιλεγόμενα επίπεδα α=0,01 και α=0,05, όταν μετατρέπονται σε ποσοστά παραπέμπουν σε αποδοχή 1% και 5% αντίστοιχα της πιθανότητας εσφαλμένων συμπερασμάτων. Τα όρια αυτής της στατιστικής σημαντικότητας αφορούν την πιθανότητα να γίνει τελικά η επιλογή της λανθασμένης υπόθεσης έναντι της σωστής. Εάν για παράδειγμα έχουμε επίπεδο σημαντικότητας 0,01, τότε αυτό σημαίνει ότι υπάρχει λιγότερο από 1% πιθανότητα να επιλεγεί μια υπόθεση λανθασμένα. Η τιμή του β είναι η πιθανότητα σφάλματος τύπου **II** και εξαρτάται από την εναλλακτική υπόθεση που θα επιλέξουμε καθώς εάν θέλουμε να διακρίνουμε μικρές αποκλίσεις από την υποτιθέμενη τιμή της παραμέτρου τότε το **β** θα πρέπει να είναι μεγάλο, ενώ εάν θέλουμε να διαπιστώσουμε τις μεγάλες διαφορές, τότε το **β** θα πρέπει να είναι μικρό (Δαμιανού κ.ά, 2003).

Στον παρακάτω πίνακα (3.1) παρουσιάζονται οι τέσσερις διαφορετικές

περιπτώσεις λήψης απόφασης, ενώ στο σχήμα 3.1 παρουσιάζονται οι περιπτώσεις σφαλμάτων.

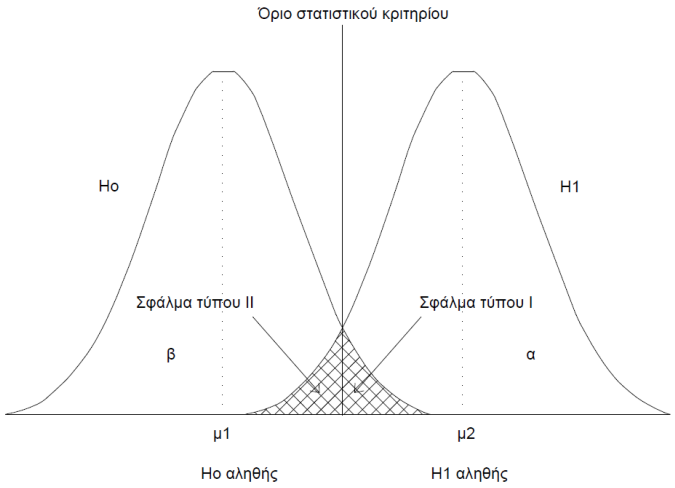
Πίνακας 3.1 Δυνατές περιπτώσεις λήψης απόφασης.

Περιπτώσεις	Ισχύει η μηδενική υπόθεση H_0	Ισχύει η εναλλακτική υπόθεση H_1
Η μηδενική υπόθεση γίνεται δεκτή (απόφαση υπέρ της $H_{0)}$	Ορθή αποδοχή της μηδενικής υπόθεσης H_0	Σφάλμα τύπου II
Η μηδενική υπόθεση απορρίπτεται (απόφαση υπέρ της $H_{1)}$	Σφάλμα τύπου I	Ορθή απόρριψη της μηδενικής υπόθεσης H_0

Πηγή: Δαμιανού κ. ά (2010) σελ. 285 . Ιδία επεξεργασία.

Σφάλμα τύπου I =Εσφαλμένη απόρριψη της $H_0 \leq \alpha$ (Δαμιανού κ.ά, 2003).

Σχήμα 3.1 Περιπτώσεις σφαλμάτων



Ιδία επεξεργασία .

Πηγή Παπαδόπουλος χ.η (στατιστικός έλεγχος υποθέσεων σελ 454)

Σύμφωνα με την Καλαματιανού (2000) το ευρύτερο πλαίσιο ενός στατιστικού ελέγχου ακολουθεί το παρακάτω γενικό περίγραμμα::

- Περιγραφή του πληθυσμού που μελετάται και διατύπωση των περιορισμών και συνθηκών που είναι αναγκαίοι για την μελέτη.
- Προσδιορισμός της παραμέτρου και διατύπωση στατιστικών υποθέσεων.
- Επιλογή του επιπέδου σημαντικότητας και της κατάλληλης στατιστικής.

- Προσδιορισμός στατιστικής ελέγχου και της κατανομής της .
- Προσδιορισμός περιοχών απόρριψης και συνθηκών αποδοχής της μηδενικής υπόθεσης.
- Υπολογισμός της τιμής στατιστικής ελέγχου.
- Στατιστική απόφαση, σύνδεση με την υπόθεση εργασίας.
- Συμπέρασμα.

Σύμφωνα με τους Αβούρη κ.ά.(2015) ο έλεγχος υποθέσεων έχει ως αποτέλεσμα τη εξαγωγή συμπερασμάτων για τους υπό μελέτη πληθυσμούς, στηριζόμενο σε δεδομένα από δείγματα, ενώ οι στατιστικές μέθοδοι που χρησιμοποιούνται πιο συχνά είναι οι παρακάτω:

3.1.6.4.1 Παραμετρικά Τεστ.

Τα παραμετρικά τεστ εφαρμόζονται όταν ικανοποιούνται μόνο ορισμένες προϋποθέσεις και χρησιμοποιούν ποσοτικά δεδομένα. Τα παραμετρικά τεστ είναι :

- Το **z στατιστικό τεστ** (one sample z- test). Όταν η διακύμανση του αρχικού πληθυσμού είναι γνωστή, τότε το z test χρησιμοποιεί δεδομένα που προέρχονται από ένα δείγμα του πληθυσμού αυτού, για να ελέγξει υποθέσεις που σχετίζονται με την μέση τιμή του. Το κριτήριο ελέγχου είναι : $\boldsymbol{Z = \frac{\bar{X} - \mu}{\frac{\sigma}{\sqrt{n}}}}$ **,** όπου **σ** είναι η διακύμανση του πληθυσμού, **n** είναι το μέγεθος του δείγματος, $\boldsymbol{\bar{X}}$ είναι ο αριθμητικός μέσος και **μ** είναι μέση τιμή. Ουσιαστικά συγκρίνουμε ην μέση τιμή μ με μια σταθερή τιμή (πχ. μ_0). Ως μηδενική υπόθεση ορίζεται μια συγκεκριμένη μέση τιμή του πληθυσμού, δηλαδή H_0: $\mu=\mu_o$ με την εναλλακτική να είναι H_1: $\mu\neq\mu_o$, ενώ οι προϋποθέσεις χρησιμοποίησης του στατιστικού τεστ z είναι γνωστή διακύμανση του πληθυσμού, αντιπροσωπευτικό δείγμα με κανονική κατανομή και αμετάβλητη διακύμανση του πληθυσμού με την επίδραση της ανεξάρτητης μεταβλητής (Ηλιοπούλου,2017· Αβούρης κ.ά 2015).
- Το **t στατιστικό τεστ ενός δείγματος** (one sample test) ελέγχει υποθέσεις που σχετίζονται με την μέση τιμή ενός κανονικού πληθυσμού όταν η διακύμανση αυτού είναι άγνωστη, χρησιμοποιώντας δεδομένα από ένα δείγμα του, κάνοντας εκτίμηση της διακύμανσης του πληθυσμού βάση της διακύμανσης του δείγματος. Απαιτείται αντιπροσωπευτικό δείγμα, κανονικότητα του δείγματος και

κανονική κατανομή των τιμών του. Ο τύπος που το καθορίζει είναι: $\boldsymbol{t = \frac{\overline{X} - \mu}{\frac{s}{\sqrt{n}}}}$ όπου **s** είναι η διακύμανση του δείγματος. Ως μηδενική υπόθεση ορίζεται μια συγκεκριμένη μέση τιμή του άγνωστου πληθυσμού , δηλαδή H_0: $\mu=\mu_o$ ενώ η εναλλακτική είναι H_1: $\mu\neq\mu_o$ (Αβούρης κ.ά, 2015).

- Το **t στατιστικό τεστ ανεξάρτητων δειγμάτων** (independentt t- testt) ελέγχει υποθέσεις που σχετίζονται με την διαφορά των μέσων τιμών δύο πληθυσμών, χρησιμοποιώντας δεδομένα από δύο ξεχωριστά δείγματα. Απαιτείται αντιπροσωπευτικότητα και κανονική κατανομή δειγμάτων και ίδια διακύμανση πληθυσμών. Η μηδενική υπόθεση είναι ότι οι μέσοι όροι των δειγμάτων είναι ίσοι, έναντι της εναλλακτικής υπόθεσης που ορίζει στατιστικά σημαντικές διαφορές (Αβούρης κ.ά, 2015).
- Το **t στατιστικό τεστ εξαρτημένων δειγμάτων** (dependent t- test) ελέγχει υποθέσεις που σχετίζονται με την διαφορά των μέσων τιμών δύο συνόλων μετρήσεων από ένα δείγμα. Απαιτείται αντιπροσωπευτικό δείγμα και κανονική κατανομή των διαφορών μετρήσεων. Η μηδενική υπόθεση είναι ότι οι δύο μέσοι όροι είναι ίσοι, έναντι της εναλλακτικής που έχει στατιστικά σημαντικές διαφορές (Αβούρης κ.ά, 2015).
- Όταν θέλουμε να εξετάσουμε εάν μια ανεξάρτητη μεταβλητή με τρείς ή περισσότερες τιμές επιδρά στην εξαρτημένη, τότε χρησιμοποιούμε την ανάλυση διασποράς η οποία μπορεί να εφαρμοστεί είτε σε ανεξάρτητα, είτε σε εξαρτημένα δείγματα. **Η Ανάλυση διασποράς ή ανάλυση διακύμανσης** (Analysis of Variance ή ANOVA) αναλύει την συνολική μεταβλητότητα όλων των τιμών, δηλαδή το πόσο διάσπαρτες είναι οι τιμές σε σχέση με το συνολικό μέσο όρο, ανεξάρτητα από το δείγμα στο οποίο ανήκουν. Αυτή η συνολική μεταβλητότητα παραμένει σταθερή και μπορεί να επιμεριστεί σε δύο μέρη: την μεταβλητότητα μεταξύ δειγμάτων και μέσα στα δείγματα. Η διαδικασία αυτή γίνεται βρίσκοντας πρώτα την διακύμανση κατά παράγοντες, δηλαδή τη διασπορά των μέσων τιμών των δειγμάτων για να εκτιμήσουμε τη διασπορά του πληθυσμού και έπειτα με την εκτίμηση της διακύμανσης σφάλματος, δηλαδή της διασποράς του πληθυσμού με βάση τις διασπορές στο εσωτερικό κάθε δείγματος. Τέλος συγκρίνουμε την πρώτη με τη δεύτερη διασπορά και εάν η πρώτη είναι μικρότερη σε σχέση με τη δεύτερη επαληθεύεται η μηδενική υπόθεση, αλλιώς διαψεύδεται. Με την ανάλυση διασποράς

ελέγχεται η διαφορά των μέσων τιμών για περισσότερα από δύο δείγματα .

Προϋπόθεση που πρέπει να ισχύει για την εφαρμογή των ελέγχων ανάλυσης διακύμανσης, είναι ότι οι πληθυσμοί στους οποίους γίνεται η δειγματοληψία είναι κανονικοί με κοινή διακύμανση (Εμβακωτής, Κάτσης & Σιδέρης, 2006 · Παπαδόπουλος, χ.η ·Αβούρης κ.ά, 2015).

3.1.6.4.2 Μη Παραμετρικά Τεστ.

Τα μη παραμετρικά τεστ εφαρμόζονται όταν δεν ικανοποιούνται οι προϋποθέσεις των παραμετρικών τεστ ή όταν ο ερευνητής έχει στη διάθεσή του ποιοτικά δεδομένα.

Ο έλεγχος X^2 **(Chi-Square Test)** -έλεγχος καλής προσαρμογής **ή** κριτήριο ελέγχου ανεξαρτησίας ως μη παραμετρικό τεστ εφαρμόζεται σε ποιοτικά δεδομένα, σε αντίθεση με τα παραπάνω αναφερόμενα τεστ που αφορούν παραμετρικά δεδομένα και εξετάζει το κατά πόσο περισσότερες από δύο αναλογίες μπορεί να θεωρηθούν ίσες, ενώ ελέγχει εάν δύο μεταβλητές είναι εξαρτημένες ως προς κάποια σχήματα ταξινόμησης, με την μηδενική υπόθεση να είναι ότι οι μεταβλητές είναι ανεξάρτητες μεταξύ τους (Ηλιοπούλου, 2015). Χρησιμοποιείται για να διαπιστωθεί η συσχέτιση μεταξύ δύο κατηγορικών ή διατεταγμένων μεταβλητών, πληροφορώντας τον ερευνητή για την ένταση της συσχέτισης μεταξύ τους, χωρίς όμως να δίνει ένδειξη για την κατεύθυνση της συσχέτισης. Σύμφωνα με τους Εμβαλωτή κ.ά (2006) ο δείκτης x^2 είναι ένας στατιστικός δείκτης που εκτιμά την στατιστική σημαντικότητα, ενώ το σκεπτικό του ελέγχου x^2 είναι ότι αρχικά υποθέτουμε την ανεξαρτησία μεταξύ των μεταβλητών, υπολογίζονται οι αναμενόμενες συχνότητες (δηλαδή οι συχνότητες που θα εμφανίζονταν εάν η μηδενική υπόθεση ήταν αληθής) εάν ίσχυε η αρχική υπόθεση ότι δεν υπάρχει σχέση, συγκρίνουμε με τις πραγματικές συχνότητες και εάν διαφέρουν πολύ, τότε απορρίπτουμε την υπόθεση της ανεξαρτησίας.

3.1.6.5 Επαγωγικοί Έλεγχοι

Για να απαντηθεί το ερώτημα του κατά πόσο είναι η τιμή του συντελεστή Pearson r, για τον οποίο θα μιλήσουμε παρακάτω, είναι σημαντική ή όχι, πρέπει να γίνει ο κατάλληλος έλεγχος του συντελεστή συσχέτισης, ο οποίος έχει ως μηδενική υπόθεση:

H_0*: δεν υπάρχει συσχέτιση μεταξύ των δύο μεταβλητών,*

$r=0$,

έναντι της εναλλακτικής

H_1**:** *υπάρχει γραμμική συσχέτιση μεταξύ των δύο μεταβλητών 0,*

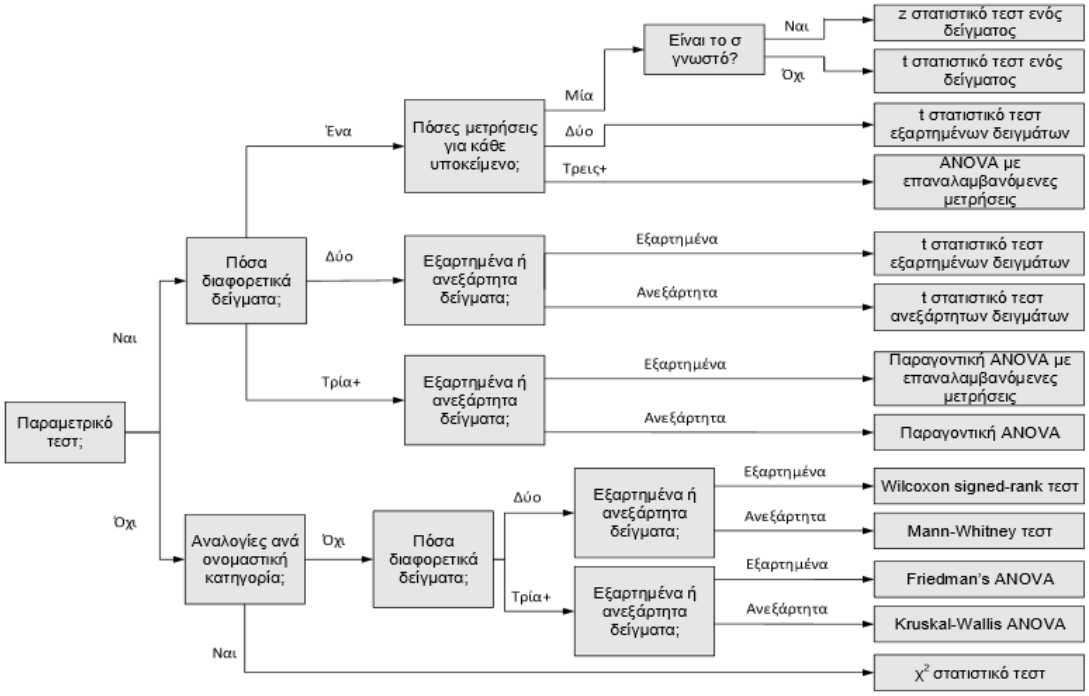
r≠0 (Ηλιοπούλου, 2015).

Για να ελεγχθεί ποια πρόταση απορρίπτεται είναι αναγκαίος ο υπολογισμός της τιμής του κριτηρίου $\boldsymbol{t = r * \sqrt{\frac{n-2}{1-r^2}}}$, όπου **r** είναι ο συντελεστής συσχέτισης Pearson που υπολογίζεται από το δείγμα, το **n** υποδεικνύει το μέγεθος του δείγματος και η σύγκριση γίνεται με την κρίσιμη τιμή της κατανομής του t-student για n-2 βαθμούς ελευθερίας. Η πιθανότητα που υπολογίζεται από το στατιστικό έλεγχο t και ονομάζεται τιμή p (p-value) έχει ένα επιτρεπτό όριο με το οποίο συγκρίνεται. Το επίπεδο αυτό ονομάζεται επίπεδο στατιστικής σημαντικότητας (statistical significance level) και καθορίζεται από τον ερευνητή (Εμβαλωτής κ.ά, 2006 ·Ηλιοπούλου, 2017). Ουσιαστικά η p-value είναι η πιθανότητα η στατιστική συνάρτηση να πάρει τιμή περισσότερο ακραία από αυτή που πήρε στο δείγμα ο ερευνητής για την μηδενική του υπόθεση. Εφόσον η τιμή του t έχει μικρότερη πιθανότητα από το επίπεδο σημαντικότητας (p>0.05 ή p>0.01 για 5% και 1% αντίστοιχα) τότε αποδεικνύεται ότι ο συντελεστής συσχέτισης είναι διάφορος του μηδενός (*r≠0)* και καθώς υπάρχει γραμμική σχέση μεταξύ των μεταβλητών, μπορούμε να προχωρήσουμε στην ανάλυση της παλινδρόμησης (Ηλιοπούλου, 2015).

Τα βήματα που ακολουθούνται για την εξαγωγή συμπερασμάτων από τον έλεγχο t είναι ο έλεγχος κανονικότητας, ο καθορισμός της μηδενικής υπόθεσης και ο ορισμός του επιπέδου σημαντικότητας, η πραγματοποίηση του ελέγχου t, η εύρεση του p και τέλος η σύγκριση του p με το επίπεδο σημαντικότητας. Εάν η τιμή του p είναι μικρότερη ή ίση με το επίπεδο σημαντικότητας τότε απορρίπτουμε την αρχική υπόθεση (Εμβαλωτής κ.ά, 2006).

3.1.6.6 Διάγραμμα Επιλογής Αποφάσεων

Σύμφωνα με τους Αβούρη κ. ά. (2015) το διάγραμμα αποφάσεων επιλογής μεθόδου ελέγχου υποθέσεων έχει οργανωθεί και παρατίθεται στον παρακάτω διάγραμμα (διάγραμμα 3.1).



Διάγραμμα 3.1 . Διάγραμμα αποφάσεων επιλογής μεθόδου ελέγχου υποθέσεων.
Πηγή Αβούρης κ. ά. (2015) σελ. 343

3.2 Ανάλυση Παλινδρόμησης

Ας υποθέσουμε ότι έχουμε δύο μεταβλητές X,Y οι οποίες συνδέονται μεταξύ τους με την σχέση Y=f (X). Με αυτόν τον τρόπο, μέσω της X μπορούμε να προβλέψουμε την τιμή της Y, στην πιο απλή μορφή και εφόσον φυσικά δεν υπάρχουν σφάλματα στη σχέση αυτή. Η παραπάνω σχέση προσδιορίζεται ως συναρτησιακή- προσδιοριστική σχέση και η περιγραφή της εξαρτησιακής σχέσης μεταξύ αυτών των μεταβλητών γίνεται μέσω της ανάλυσης παλινδρόμησης, η οποία είναι από τις σημαντικότερες στατιστικές τεχνικής ανάλυσης δεδομένων.

Ο όρος regression χρησιμοποιήθηκε για πρώτη φορά από τον πολυμαθή Άγγλο ψυχολόγο, ερευνητή, γενετιστή, ανθρωπολόγο Galton το 1886 στο βιβλίο του Regression Towards Mediocrity in Hereditary Stature. Σύμφωνα με τους Δαμιανό κ ά, 2010 ως ανάλυση παλινδρόμησης ορίζεται ο κλάδος αυτός της στατιστικής που εξετάζει τη σχέση δύο ή περισσότερων μεταβλητών σ' ένα στοχαστικό μοντέλο, με σκοπό τη πρόβλεψη μιας από αυτές μέσω των υπολοίπων.

Η ανάλυση παλινδρόμησης (regression analysis), αποτελεί φυσική επέκταση του συντελεστή συσχέτισης, που θα αναπτυχθεί παρακάτω, παρέχοντας εκτός από την δυνατότητα πρόβλεψης και την δυνατότητα συνεκτίμησης περισσότερων των δύο μεταβλητών (Εμβαλωτής κ.ά.,2006). Καθώς ο συντελεστής συσχέτισης εκτιμά

ποσοτικά τη σχέση μεταξύ δύο ή περισσότερων μεταβλητών, η παλινδρόμηση μας βοηθά να προσεγγίσουμε τη μορφή της σχέσης αυτής και να προσδιορίζουμε σχέση αιτίου- αιτιατού ή σχέση αιτίου- αποτελέσματος (Ηλιοπούλου, 2015). Παράλληλα σκοπός της παλινδρόμησης είναι η δημιουργία ρεαλιστικών μοντέλων, ώστε να καταστεί δυνατή η πρόβλεψη και ερμηνεία φαινομένων και η ερμηνεία της συσχέτισης και της δυναμικής των μεταβλητών. Ανάλογα με τον λόγο που θα χρησιμοποιηθεί κάθε φορά η ανάλυση, καθορίζεται και η μεθοδολογία που θα ακολουθηθεί. Ο προσδιορισμός της εξίσωσης με τις λιγότερες δυνατές παραμέτρους και τα λιγότερα λάθη στην πρόβλεψη είναι ο στόχος της παλινδρόμησης εάν ο σκοπός της έρευνας είναι η πρόβλεψη.

Ο προσδιορισμός των ανεξάρτητων ή ελεγχόμενων ή επεξηγηματικών μεταβλητών (independent variables ή predictors ή input variables) που επηρεάζουν την εξαρτημένη ή μεταβλητή απόκρισης (dependent ή response variable (Παπαδόπουλος, χ.η), καθορίζει το σωστό μοντέλο πρόβλεψης, το οποίο όταν δημιουργηθεί, θα χρησιμοποιηθεί για την πρόβλεψη των τιμών της εξαρτημένης από τις τιμές των ανεξάρτητων μεταβλητών για παρατηρήσεις που δεν υπήρχαν στα αρχικά δεδομένα (Ηλιοπούλου, 2015).

Ως **ανεξάρτητη μεταβλητή** προσδιορίζουμε αυτή που επηρεάζει αιτιωδώς την εξαρτημένη μεταβλητή, ενώ **εξαρτημένη** είναι αυτή στη οποία αντανακλάται το αποτέλεσμα των μεταβολών της ή των ανεξάρτητων μεταβλητών (Εμβαλωτής κ.ά 2006). Όταν χρησιμοποιείται μόνο μία ανεξάρτητη μεταβλητή, η παλινδρόμηση ονομάζεται **απλή**, ενώ όταν χρησιμοποιούνται δύο ή περισσότερες ονομάζεται **πολλαπλή**.

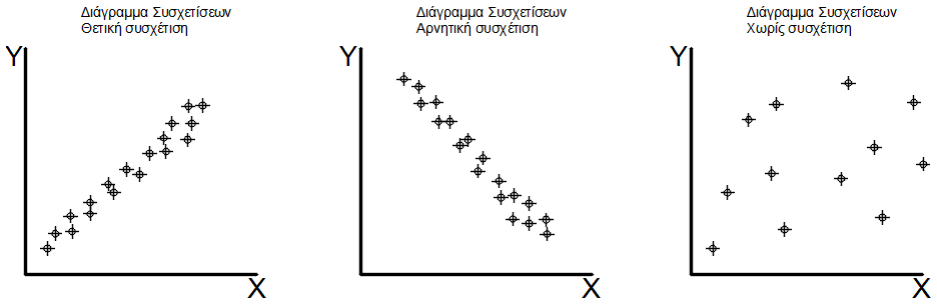
Η μέθοδος που χρησιμοποιείται συνήθως για την περιγραφή της στοχαστικής εξάρτησης μεταξύ δύο μεταβλητών είναι η μέθοδος των ελαχίστων τετραγώνων που θα αναπτυχθεί σε επόμενο κεφάλαιο (Παπαδόπουλος, χη).

3.3 Διάγραμμα Σκεδασμού και Ανάλυση Συσχέτισης

3.3.1 Διάγραμμα Σκεδασμού

Εάν για δύο μεταβλητές Χ και Υ λάβουμε **ν** παρατηρήσεις, με ζεύγη τιμών (x_i, y_i) όπου i =1,2,3,….,ν, τότε το πρώτο βήμα της ανάλυσης συσχέτισης μεταξύ των δύο αυτών μεταβλητών είναι η γραφική απεικόνιση των σημείων αυτών, οπτικοποιώντας τη σχέση μεταξύ τους με το **διάγραμμα σκεδασμού ή διασποράς (scatter diagram)** σε σύστημα καρτεσιανών συντεταγμένων (Διάγραμμα 3.2). Η εξαρτημένη μεταβλητή

τοποθετείται στον άξονα των Y και η ανεξάρτητη στον άξονα των X. Το διάγραμμα σκεδασμού δίνει πληροφορίες για την σχέση εξάρτησης μεταξύ των δύο μεταβλητών αλλά και πέραν από την οπτική ανίχνευση συσχέτισης δίνει πληροφορίες για την ύπαρξη ή μη γραμμικότητας (Εμβαλωτής, κ.ά, 2006 · Δαμιανός, κ.ά, 2010). Δημιουργείται ένα νέφος σημείων που απεικονίζει σε μια εικόνα εάν υπάρχει ή όχι εμφανής συσχέτιση μεταξύ των X και Y. Εάν τα σημεία ακολουθούν μια ευθεία ή μια καμπύλη τότε το διάγραμμα δείχνει συσχέτιση των μεταβλητών, ενώ εάν είναι διάσπαρτα, αποτυπώνεται η μη συσχέτισή τους. Η απεικόνιση αυτή αποτελεί απλά μόνο ένδειξη, χωρίς η οπτική εκτίμηση να αποδεικνύει την ύπαρξη γραμμικής συσχέτισης. Με αυτόν τον τρόπο διαπιστώνουμε εάν έχει νόημα να προχωρήσουμε παρακάτω ή όχι . Όπως είναι προφανές όταν οι δύο μεταβλητές μεταβάλλονται προς την ίδια κατεύθυνση έχουμε θετική συσχέτιση, όταν μεταβάλλονται προς την αντίθετη κατεύθυνση έχουμε αρνητική συσχέτιση και όταν οι τιμές της μίας δεν σχετίζονται με της άλλης, έχουμε μηδενική συσχέτιση.



Διάγραμμα 3.2 . Διαγράμματα σκεδασμού με απεικόνιση θετικής, αρνητικής και ουδέτερης συσχέτισης . Ιδία επεξεργασία

Η ακριβής συσχέτιση αποδεικνύεται με τον υπολογισμό του συντελεστή συσχέτισης και τον έλεγχο της σημαντικότητας του. Η διαδικασία αυτή αναλύεται στο επόμενο βήμα.

3.3.2 Ανάλυση Συσχέτισης

Το δεύτερο βήμα που ακολουθείται στην διαδικασία είναι ο έλεγχος συσχέτισης μεταξύ των μεταβλητών. Είδαμε ότι μια πρώτη ιδέα για την συσχέτιση μεταξύ δύο μεταβλητών μπορούμε να πάρουμε με το διάγραμμα διασποράς. Το πόσο ισχυρή είναι αυτή όμως και το πώς μπορεί να μετρηθεί, διαπιστώνεται από την μέτρηση των συντελεστών συσχέτισης. Ο συντελεστής συσχέτισης μεταξύ δύο μεταβλητών είναι ένα μέτρο βαθμού γραμμικότητας μεταξύ τους. Σύμφωνα με τον Παπαδόπουλο ο

πληθυσμιακός συντελεστής γραμμικής συσχέτισης του Pearson συμβολίζεται με $\boldsymbol{\rho}$. Η **σημειακή εκτίμηση (point estimation)** του συντελεστή συσχέτισης $\boldsymbol{\rho}$ του πληθυσμού από το δείγμα των **ν** ζευγών παρατηρήσεων συμβολίζεται με $\widehat{\boldsymbol{\rho}}$, (Κουγιουτζής, 2011).

Ο **δειγματοληπτικός ή δειγματικός συντελεστής συσχέτισης του Pearson** (sample Pearson correlation coefficient), συμβολίζεται με $\mathbf{r}$, αναφέρεται σε δεδομένα δειγματοληψίας, ενώ είναι ουσιαστικά ο εκτιμητής του $\boldsymbol{\rho}$ όταν τα ζεύγη προέρχονται από τυχαία δειγματοληψία. Ανήκει στην κατηγορία των σταθμισμένων δεικτών και είναι ένας παραμετρικός στατιστικός δείκτης αξιολόγησης της συνάφειας μεταξύ δύο μεταβλητών (Εμβαλωτής κ.ά, 2006), αποδίδοντας ένα μέτρο μεγέθους της γραμμικής συσχέτισης μεταξύ τους. Οι τιμές που λαμβάνει είναι από -1 έως +1. Το -1 φανερώνει την τέλεια αρνητική γραμμική συσχέτιση και το +1 την τέλεια θετική γραμμική συσχέτιση. Εάν ισούται με «0» τότε φανερώνεται η μη ύπαρξη γραμμικής συσχέτισης (Ηλιοπούλου, 2015).

Για την χρήση του συντελεστή Pearson υπάρχουν οι παρακάτω προϋποθέσεις:

- Οι μεταβλητές πρέπει να είναι συνεχείς, σε κλίμακα ίσων διαστημάτων και να υπάρχει κανονική κατανομή καθώς και
- να υπάρχει γραμμική συσχέτιση μεταξύ των μεταβλητών .

Ο συντελεστής Pearson ως παραμετρικός συντελεστής ικανοποιεί την παραπάνω σειρά στατιστικών προϋποθέσεων. Εάν δεν υπάρχουν αυτές οι προϋποθέσεις τότε πρέπει να στραφούμε σε μη παραμετρικούς στατιστικούς δείκτες όπως ο Spearman (Εμβαλωτής κ.ά, 2006).

Ο δειγματικός συντελεστής γραμμικής συσχέτισης του Pearson που συμβολίζεται με το r υπολογίζεται από τον γενικό τύπο: $\boldsymbol{r = \frac{s_{xy}}{s_x \cdot s_y}}$

όπου $\boldsymbol{s_{xy}}$ είναι η δειγματική συνδιασπορά (συμμεταβλητότητα) των δύο μεταβλητών και υπολογίζεται από τον τύπο: $\boldsymbol{s_{xy} = Cov(X,Y) = \frac{1}{\nu-1} \sum_{i=1}^{\nu}(x_i - \bar{x}) \cdot (y_i - \bar{y}) = \frac{1}{\nu-1} \sum_{i=1}^{\nu}(x_i y_i - \nu \cdot \bar{x} \cdot \bar{y}),}$

όπου $\boldsymbol{s_x}$ είναι η τυπική απόκλιση της μεταβλητής Χ και υπολογίζεται απ'τον τύπο:

$$\boldsymbol{s_x = \sqrt{\frac{1}{\nu-1}\sum_{i=1}^{\nu}(x_i - \bar{x})^2}}$$

και όπου $\boldsymbol{s_y}$ είναι η τυπική απόκλιση της μεταβλητής Y και υπολογίζεται από τον τύπο: $\boldsymbol{s_y = \sqrt{\frac{1}{\nu-1}\sum_{i=1}^{\nu}(y_i - \bar{y})^2}}$ (Παπαδόπουλος, χ.η).

Σημειώνεται ότι καλύτερη φυσική ερμηνεία επιτυγχάνεται με τον $\mathbf{R^2}$ δηλαδή τον συντελεστή προσδιορισμού (coefficient of determination) ο οποίος εκφράζεται συνήθως σε ποσοστό **100** $\mathbf{R^2}$**.** Αυτός ο συντελεστής δίνει το ποσοστό μεταβλητότητας των Y που υπολογίζεται από την X (Κουγιουμτζής, 2011) και χρησιμοποιείται για την μέτρηση του βαθμού αξιοπιστίας της γραμμικής παλινδρόμησης. Βέβαια, ο υπολογισμός μόνο του συντελεστή $\mathbf{R}$ δεν σημαίνει ότι η τιμή του είναι στατιστικά σημαντική. Ως παράδειγμα θα μπορούσε να αναφερθεί μεγάλη θετική συσχέτιση σε μικρό δείγμα.

Όπως προαναφέρθηκε στο προηγούμενο κεφάλαιο για να απαντηθεί το ερώτημα του κατά πόσο είναι η τιμή του r σημαντική ή όχι, πρέπει να γίνει ο κατάλληλος έλεγχος του συντελεστή συσχέτισης με το κριτήριο t, που έχει ως μηδενική υπόθεση:

H_0: δεν υπάρχει συσχέτιση μεταξύ των δύο μεταβλητών , *r=0,*

έναντι της εναλλακτικής

H_1: υπάρχει γραμμική συσχέτιση μεταξύ των δύο μεταβλητών *0, r≠0*

Θα πρέπει να σημειωθεί ότι η ύπαρξη ακραίων τιμών στις δύο μεταβλητές στο διάγραμμα διασποράς μπορεί να οδηγήσει σε παραπλανητικά αποτελέσματα του συντελεστή Pearson, ενώ παρόμοια είναι η περίπτωση εμφάνισης ισχυρής συσχέτισης δύο μεταβλητών καθώς μπορεί να έχουν ισχυρή συσχέτιση ως προς μια τρίτη. Αυτές οι περιπτώσεις αναφέρονται ως νόθα συσχέτιση (Ηλιοπούλου, 2015).

Σε περιπτώσεις ιεραρχικών μεταβλητών ή σε περιπτώσεις ύπαρξης ακραίων τιμών ή όταν δεν ικανοποιούνται οι παραμετρικές προϋποθέσεις (κανονικότητα, γραμμικότητα, εύρος παρατηρήσεων, ισοδιαστημική κλίμακα) τότε για την ανίχνευση των σχέσεων μεταξύ των μεταβλητών, χρησιμοποιούνται εναλλακτικοί συντελεστές συσχέτισης, όπως είναι ο **Spearman** (Ηλιοπούλου, 2015). Ανήκει στους μη παραμετρικούς ελέγχους, δεν χρησιμοποιεί εκτίμηση κάποιας παραμέτρου και δεν

προϋποθέτει γνώση της μορφής της κατανομής του πληθυσμού από τον οποίο έγινε η δειγματοληψία (Ηλιοπούλου, 2015). Ο συντελεστής αυτός υπολογίζεται μετατρέποντας τα δεδομένα σε σειρές βάση μεγέθους, με αποτέλεσμα οι αποστάσεις μεταξύ των παρατηρήσεων να χάνουν την σημασία τους, ενώ αξιολογείται η σειρά των συμμετεχόντων στη πρώτη μεταβλητή, σε σχέση με την σειρά που έχουν στη δεύτερη κ.ο.κ . Ο τύπος που δίνει τον συντελεστή **Spearman** είναι:

$\mathbf{r_s = 1 - \frac{6\Sigma D^2}{\nu(\nu^2-1)}}$, όπου **ν** είναι ο το πλήθος των παρατηρήσεων και **D** είναι η διαφορά στην σειρά μεταξύ πρώτης και δεύτερης μέτρησης (Εμβαλωτής κ.ά, 2006). Ο συντελεστής λαμβάνει τιμές από -1 έως +1, δηλώνοντας αντίστοιχα θετική ή αρνητική συσχέτιση. Τιμές κοντά στο 0 δεν υποδεικνύουν συσχέτιση.

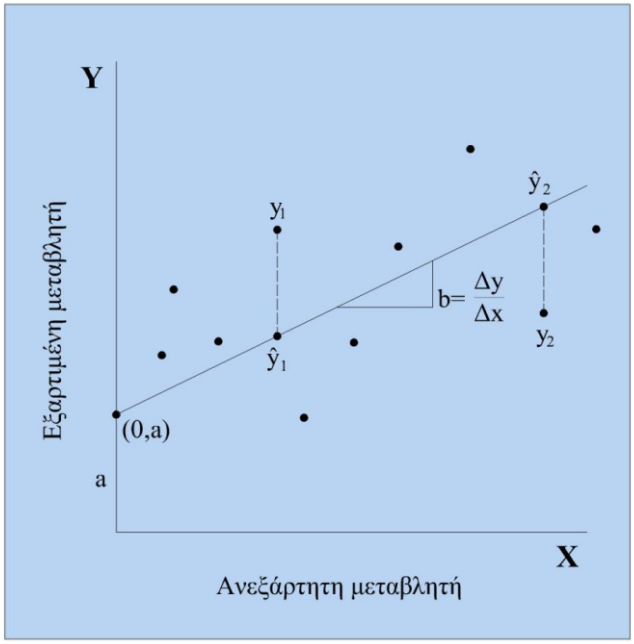
3.4 Μοντέλα Παλινδρόμησης

3.4.1 Απλή Γραμμική Παλινδρόμηση (simple linear regression)

Ας υποθέσουμε ότι σ' έναν άπειρο πληθυσμό λάβουμε δείγμα μεγέθους n και καταγράψουμε για κάθε άτομο του δείγματος δύο μεταβλητές, την **X** και την **Y**. Τα ζεύγη τιμών που θα πάρουμε θα είναι της μορφής (X_1,Y_1), (X_2,Y_2), (X_3,Y_3),.... (X_n,Y_n). Εάν διαπιστώσουμε ότι το διάγραμμα διασποράς των δύο μεταβλητών έχει μια μορφή που παραπέμπει οπτικά σε ευθεία τότε το πιο πιθανό είναι ότι η σχέση των μεταβλητών είναι κατά προσέγγιση γραμμική και είναι της μορφής $\mathbf{Y=\alpha+\beta\cdot X}$. Προφανώς και η γραμμική στοχαστική εξάρτηση των δύο μεταβλητών δεν μπορεί να περιγραφεί έτσι απλά, καθώς στη σχέση αυτή εμπεριέχονται και σφάλματα. Επομένως εάν προσθέσουμε στη παραπάνω σχέση τον όρο **ε**, όπου για μια δεδομένη τιμή της $\mathbf{X}$ περιγράφει την διαφορά της παρατηρούμενης από την θεωρητική τιμή της $\mathbf{Y}$, δηλαδή την $\mathbf{\alpha+\beta\cdot X}$, τότε από την σχέση αυτή το **ε** προσδιορίζεται ως $\mathbf{\varepsilon=Y-(\alpha+\beta\cdot X)}$ και το στοχαστικό μοντέλο που προκύπτει είναι το $\mathbf{Y=\alpha+\beta\cdot X+\varepsilon}$ (Παπαδόπουλος, χη). Το μοντέλο αυτό είναι η εξίσωση της απλής γραμμικής παλινδρόμησης. Η μεταβλητή **X** καλείται ανεξάρτητη και **Y** εξαρτημένη μεταβλητή και ουσιαστικά μεταξύ των δύο μεταβλητών **X** και **Y** υπάρχει σχέση αιτίου – αποτελέσματος. Σύμφωνα με τους Εμβαλωτή κ.ά (2006) το πρόβλημα που αφορά την απλή γραμμική παλινδρόμηση είναι η διερεύνηση της σχέσης μεταξύ των δύο αυτών μεταβλητών και η εφαρμογή του μοντέλου της ευθείας γραμμής με

πρόθεση την πρόβλεψη .

Ας τοποθετήσουμε τώρα σ' ένα καρτεσιανό σύστημα συντεταγμένων τα ζεύγη των σημείων που λάβαμε από τις δύο μεταβλητές εκ των οποίων, η μεν Υ τοποθετείται στον κατακόρυφο άξονα και η δε Χ τοποθετείται στον οριζόντιο (σχήμα 3.2).



Σχήμα 3.2. .Αναπαράσταση εκτιμημένης ευθείας γραμμικής παλινδρόμησης.
Πηγή: Ηλιοπούλου,2015(σελ 159)

Σε κάθε τιμή της εξαρτημένης τιμής της μεταβλητής Υ αντιστοιχεί μια άλλη, η οποία αποτελεί την εκτίμηση από την εξίσωση της παλινδρόμησης και συμβολίζεται με $\widehat{\boldsymbol{Y}}$**.** Η εφαρμογή λοιπόν του μοντέλου της ευθείας γραμμής, με πρόθεση την πρόβλεψη της σχέσης των δύο μεταβλητών, εκφράζεται από την γενική εξίσωση:

$\widehat{\boldsymbol{Y}} = \mathbf{a} + \boldsymbol{b} \cdot \boldsymbol{X}$**,** όπου το $\widehat{\boldsymbol{Y}}$ είναι η προβλεπτική τιμή της μεταβλητής Υ για την τιμή της Χ . Το πρόβλημα που προκύπτει στην απλούστερη του μορφή είναι πλέον ο προσδιορισμός των παραμέτρων της εξίσωσης, καθώς με αυτό τον τρόπο θα εκφραστεί καλύτερα η γραμμική εξάρτηση των δύο μεταβλητών (Κουγιουμτζής, 2011 · Ηλιοπούλου, 2017).

Από την προαναφερθείσα εξίσωση $\mathbf{Y=\alpha+\beta\cdot X+\varepsilon}$ προέρχεται το απλό γραμμικό μοντέλο και καθώς συνηθίζεται ο συμβολισμός των παραμέτρων της παλινδρόμησης α και β αντίστοιχα με β_0 και β_1, απεικονίζουμε το απλό γραμμικό

μοντέλο με την εξίσωση $\mathbf{y=\beta_0+\beta_1.x+\varepsilon}$ όπου το **y** αναπαριστά την εξαρτημένη μεταβλητή, το **x** αναπαριστά την ανεξάρτητη μεταβλητή, το $\boldsymbol{\beta_0}$ είναι το σημείο τομής της ευθείας με το κατακόρυφο άξονα- (intescept term -τεταγμένη), το $\boldsymbol{\beta_1}$ είναι η κλίση-slope, ή αλλιώς συντελεστής παλινδρόμησης (regression coefficient) ή συντελεστής διεύθυνσης και αποτελεί μαζί με το $\boldsymbol{\beta_0}$ τις παραμέτρους της εξίσωσης, ενώ το $\boldsymbol{\varepsilon}$ είναι το σφάλμα της εκτίμησης (υπόλοιπο- residual), ή σφάλμα παλινδρόμησης (regression error) (Ηλιοπούλου, 2015 · Κουγιουμτζής, 2011 · Δαμιανός, κ.ά, 2010 · Montgomery, Peck, & Vining, 2012).

Ο συντελεστής $\boldsymbol{\beta_1}$ εκφράζει την κλίση της ευθείας, δηλώνει την κατεύθυνση της επιρροής της ανεξάρτητης μεταβλητής στην εξαρτημένη και η τιμή του μπορεί να ελεγχθεί με τον έλεγχο **t** και τη μηδενική υπόθεση $\boldsymbol{\beta_1=0}$. Δείχνει πρακτικά πόσο μεταβάλλεται το αποτέλεσμα για κάθε μια μονάδα αλλαγής της ανεξάρτητης μεταβλητής. Εάν ο συντελεστής έχει θετικό πρόσημο η αύξηση της τιμής της ανεξάρτητης προκαλεί αύξηση στην τιμή της εξαρτημένης μεταβλητής (Εμβαλωτής κ. ά, 2006 · Δαμιανός, κ.ά, 2010). Η παράμετρος $\boldsymbol{\beta_0}$ είναι η τιμή της εξαρτημένης μεταβλητής Y για τιμή της ανεξάρτητης X ίσης με μηδέν και δεν μπορεί πάντα να ερμηνευτεί με αξιόπιστο τρόπο. Τα $\boldsymbol{\varepsilon}$ είναι ανεξάρτητες τυχαίες μεταβλητές (τυχαία σφάλματα) που ακολουθούν την κανονική κατανομή, δηλαδή τα σφάλματα της εκτίμησης ή κατάλοιπα ή υπόλοιπα (residual), ή αλλιώς οι διαφορές για κάθε ένα μεταξύ της πραγματικής τιμής της Y και της τιμής πρόβλεψης (παρατηρούμενες τιμές μείον εκτιμώμενες), όπως προκύπτει από το υπόδειγμα. Το $\boldsymbol{\varepsilon}$ παίρνει θετικές και αρνητικές τιμές, όμως το μέσο όρο της τιμής είναι μηδέν, ενώ έχει κανονική κατανομή, με την διακύμανσή του να παραμένει σταθερή όταν μεταβάλλεται η τιμή της X. Τα κατάλοιπα δείχνουν ότι δεν περιλαμβάνονται όλες οι πιθανές ανεξάρτητες μεταβλητές στο μοντέλο μας. Σχεδιαστικά τα κατάλοιπα - υπόλοιπα είναι οι αποστάσεις κάθε σημείου από την ευθεία, ενώ ισχύει $\boldsymbol{\varepsilon_{i=}}$ **yi** $\boldsymbol{-\hat{y}i}$ δηλαδή προσδιορίζεται ως η τιμή παρατήρησης, μείον την εκτιμώμενη τιμή (Ηλιοπούλου, 2015 ·Δαμιανός, κ.ά, 2010·Παπαδόπουλος,χ.η· Κουγιουμτζής, 2011 ·Rawlings, Pantula & Dickey, 1998).

Το συνηθέστερο μαθηματικό κριτήριο για την προσαρμογή της γραμμής

παλινδρόμησης στα εμπειρικά δεδομένα είναι η μέθοδος των ελαχίστων τετραγώνων, σκοπός της οποία είναι η ελαχιστοποίηση του αθροίσματος των τετραγώνων αποκλίσεων παρατηρήσεων από την γραμμή παλινδρόμησης, δηλαδή $\sum_{i=1}^{n}(y_i - \hat{y})^2$=min (Ηλιοπούλου, 2017).

Με την μέθοδο των ελαχίστων τετραγώνων θα ασχοληθούμε αμέσως παρακάτω.

Οι βασικές προϋποθέσεις- παραδοχές της χρήσης της απλής γραμμικής παλινδρόμησης είναι :

- Η μεταβλητή Χ να μετρά χωρίς σφάλματα και η Υ για κάθε επίπεδο της Χ να είναι τυχαία μεταβλητή με πεπερασμένη μέση τιμή και διασπορά.
- Να ισχύει η ύπαρξη γραμμικότητας (linearity), ως γραμμική συσχέτιση εξαρτημένης με ανεξάρτητη μεταβλητή.
- Να ισχύει η ύπαρξη Ομοσκεδαστικότητας- σταθερότητας διασποράς (homoscedasticity – variance stability), ως σταθερή διακύμανση μεταξύ πραγματικών και υπολογισμένων από την παλινδρόμηση τιμών, καθώς η μεταβλητότητα γύρω από την γραμμή παλινδρόμησης πρέπει να είναι σταθερή για όλες τις τιμές του Χ.
- Να ισχύει η ανεξαρτησία (independence), δηλαδή οι τιμές της Υ που αντιστοιχούν στα διάφορα επίπεδα της Χ, πρέπει να είναι ανεξάρτητες μεταξύ τους.
- Να ισχύει η κανονικότητα (normality), δηλαδή κανονική κατανομή της Υ για όλα τα επίπεδα της Χ .
- Οι μεταβλητές να είναι ποσοτικές.
- Τα σφάλματα (κατάλοιπα) να ακολουθούν την κανονική κατανομή, με μέσο μηδέν, σταθερή διακύμανση σ^2 χωρίς πρόβλημα αυτοσυσχέτισης μεταξύ τους.
- Να ισχύει η ανυπαρξία πολυσυγγραμμικότητας (Παπαδόπουλος, χ.η · Δαμιανός, κ.ά, 2010 · Κουγιουμτζής, 2011).

Το μοντέλο της απλής γραμμικής παλινδρόμησης ακολουθεί τρία βασικά στάδια:

- Τον προκαταρτικό έλεγχο,
- το ίδιο το μοντέλο παλινδρόμησης και
- τον έλεγχο των προϋποθέσεων.

Ο προκαταρτικός έλεγχος περιλαμβάνει την διαγραμματική απεικόνιση, τον έλεγχο κανονικότητας και τους δείκτες γραμμικής συσχέτισης. Το μοντέλο παλινδρόμησης περιλαμβάνει τους δείκτες συσχέτισης, τους ελέγχους υποθέσεων και την ερμηνεία

των αποτελεσμάτων. Ο έλεγχος προϋποθέσεων αφορά τη κανονικότητα ανεξαρτησίας, ομοσκεδαστικότητας σφαλμάτων και εξαρτημένης μεταβλητής και τον έλεγχο εκτροπών και ασυνήθιστων τιμών. Όσο περισσότερο τα σημεία παρατηρήσεων πλησιάζουν την ευθεία γραμμή τόσο καλύτερη είναι η γραμμή παλινδρόμησης. Από την στιγμή που θέσαμε στο μοντέλο μας και το θέμα των ανεξάρτητων τυχαίων σφαλμάτων, το πρόβλημα πια της απλής γραμμικής παλινδρόμησης γίνεται η εκτίμηση των παραμέτρων της, δηλαδή της διαφοράς ύψους της ευθείας της παλινδρόμησης $\boldsymbol{\beta_0}$, της κλίσης της ευθείας παλινδρόμησης $\boldsymbol{\beta_1}$ και της διασποράς του σφάλματος παλινδρόμησης $\boldsymbol{\sigma_\varepsilon^2}$, όπου αυτό προσδιορίζει το βαθμό μεταβλητότητας γύρω από την ευθεία παλινδρόμησης εκφράζοντας την αβεβαιότητα της γραμμικής σχέσης.

Αναλυτικότερα:

- ✓ Το πρώτο βήμα για την κατασκευή του προβλεπτικού μοντέλου απλής γραμμικής παλινδρόμησης είναι ο ορισμός της εξαρτημένης **Y** και η ανεξάρτητης **X** μεταβλητής.
- ✓ Με την κατασκευή του διαγράμματος σκεδασμού απεικονίζεται οπτικά η σχέση των δύο μεταβλητών. Το διάγραμμα αυτό όπως προαναφέρθηκε αποτελεί απλά μόνο ένδειξη, χωρίς η οπτική εκτίμηση να αποδεικνύει την ύπαρξη γραμμικής συσχέτισης. Εάν υπάρχει οπτικά γραμμική συσχέτιση έχει νόημα να προχωρήσουμε στο επόμενο βήμα.
- ✓ Ακολουθεί ο υπολογισμός του συντελεστή συσχέτισης r κατά Pearson, σύμφωνα με την σχέση που προαναφέρθηκε: $\boldsymbol{r = \frac{s_{xy}}{s_x \cdot s_y}}$ και όπως περιγράφηκε αναλυτικά στο προηγούμενο κεφάλαιο.
- ✓ Ο έλεγχος της στατιστικής σημαντικότητας του συντελεστή συσχέτισης γίνεται με τον έλεγχο **t** και την επιβεβαίωση από τον έλεγχο υπόθεσης, όπως περιγράφηκε και αυτός στο προηγούμενο κεφάλαιο.
- ✓ Αφού διαπιστωθεί ότι οι δύο μεταβλητές αλληλεξαρτώνται, έπεται ο υπολογισμός της εξίσωσης απλής γραμμικής παλινδρόμησης. Η ευθεία που δίνει την βέλτιστη περιγραφή προσδιορίζεται με την μέθοδο των ελαχίστων τετραγώνων που απαιτεί και τη ελαχιστοποίηση των καταλοίπων, δηλαδή $\boldsymbol{\sum_{i=1}^{n}(y_i - \hat{y})^2}$=min (Ηλιοπούλου, 2017). Ουσιαστικά με την μέθοδο των ελαχίστων τετραγώνων

βρίσκουμε την ευθεία παλινδρόμησης με παραμέτρους $\mathbf{b_0}$ **και** $\mathbf{b_1}$ έτσι ώστε το άθροισμα των τετραγώνων των κατακόρυφων αποστάσεων των σημείων από αυτή να είναι το ελάχιστο (Κουγιουμτζής, 2011).

Έστω λοιπόν ότι έχουμε την ευθεία $\mathbf{y_i=\beta_0+\beta_1.x_i+\varepsilon_i}$

Οι εκτιμήσεις των $\boldsymbol{\beta_0}$ **και** $\boldsymbol{\beta_1}$ δίνονται από την ελαχιστοποίηση του αθροίσματος των τετράγωνων των σφαλμάτων , δηλαδή $\underset{\beta o\, \beta 1}{min} \sum_{i=1}^{n} \varepsilon_i^{\,2}$.

Η εκτίμηση της παραμέτρου $\boldsymbol{\beta_1}$ **προκύπτει ως** $\widehat{\boldsymbol{\beta_1}}$ **και είναι**

$\mathbf{b_1}= \frac{\sum_{i=1}^{n}(x_i-\bar{x})(y_i-\bar{y})}{\sum_{i=1}^{n}(x_i-\bar{x})^2} = \frac{s_{xy}}{s_{xx}}$ και η εκτίμηση της παραμέτρου $\boldsymbol{\beta_0}$ προκύπτει ως

$\widehat{\boldsymbol{\beta_0}}$ **και είναι** $\mathbf{b_0}= \frac{\sum_{i=1}^{n} y_i - b_1 \sum_{i=1}^{n} x_i}{n}$

Οι εκτιμήσεις λοιπόν των $\boldsymbol{\beta_1}$ **και** $\boldsymbol{\beta_0}$ αντίστοιχα είναι :

$\mathbf{b_1}= \frac{\sum_{i=1}^{n}(x_i-\bar{x})(y_i-\bar{y})}{\sum_{i=1}^{n}(x_i-\bar{x})^2}$ **και**

$\mathbf{b_0}= \bar{y} - \mathbf{b_1}\cdot \bar{x}$

και τα $\mathbf{b_0}$ **και** $\mathbf{b_1}$ ορίζουν την

ευθεία των ελαχίστων τετραγώνων $\hat{y}=\mathbf{b_0+b_1.x}$ (Κουγιουμτζής, 2011 · Ηλιοπούλου, 2017)

- ✓ Επόμενο βήμα είναι ο έλεγχος αξιοπιστίας του μοντέλου που δημιουργήθηκε. Αυτό γίνεται με τον **συντελεστή προσδιορισμού** (coefficient of determination) $\mathbf{R^2}$. Ο συντελεστής αυτός εξηγεί το πόσο καλά ερμηνεύει η ανεξάρτητη μεταβλητή την εξαρτημένη, δείχνοντας το πόσο ισχυρή είναι η γραμμική συσχέτιση μεταξύ των μεταβλητών. Ουσιαστικά δίνει τη προσαρμοστικότητα του μοντέλου και εκφράζει το ποσοστό διασποράς της εξαρτημένης Y που ερμηνεύεται από την ανεξάρτητη μεταβλητή X (Ηλιοπούλου, 2015 ·Δαμιανός,κ.ά, 2010). Οι τιμές που λαμβάνει έχουν εύρος από 0 έως 1, ενώ εκφράζεται και ως ποσοστό. Όσο μεγαλύτερη η τιμή του, τόσο καλύτερη και η προσαρμογή του στα δεδομένα. Ουσιαστικά εκφράζει την αναλογία της εξηγήσιμης μεταβλητότητας από το μοντέλο, προς τη συνολική μεταβλητότητα της εξαρτημένης μεταβλητής. Εάν συμβολίσουμε με SSE (error sum of squares) το άθροισμα τετραγώνων σφαλμάτων ή αλλιώς το

υπόλοιπο μεταβλητότητας, που δεν εξηγείται από την παλινδρόμηση και που μετρά την αβεβαιότητα στην πρόβλεψη του y όταν χρησιμοποιείται το x, τότε αυτό προσδιορίζεται από τον τύπο:

$$\mathbf{SSE}=\sum_{i=1}^{n}(y-\widehat{y})^2=\sum_{i=1}^{n}\hat{\varepsilon}^2$$

Εάν συμβολίσουμε με SST (Total sum of squares) την ολική μεταβολή των Y ως το ολικό άθροισμα τετραγώνων που μετρά την μεταβλητότητα των παρατηρήσεων y χωρίς να λαμβάνει υπόψη τις x τότε αυτό είναι ίσο με $\mathbf{SST}=\sum_{i=1}^{n}(y-\overline{y})^2$

Εάν συμβολίσουμε με SSR (regression sum of squares) το άθροισμα τετραγώνων που οφείλεται στην παλινδρόμηση και δίνει ένα μέτρο ελάττωσης της μεταβλητότητας των y όταν γίνεται χρήση της x τότε αυτό δίνεται από τον τύπο

$$\mathbf{SSR}=\sum_{i=1}^{n}(\widehat{y}-\overline{y})^2$$

Ως αποτέλεσμα των παραπάνω έχουμε ότι το $\mathbf{SST=SSE+SSR}$ και ο συντελεστής προσδιορισμού $\mathbf{R^2}=\frac{SSR}{SST}=\mathbf{1}-\frac{SSE}{SST}$ εκφράζει το ποσοστό της ολικής μεταβολής των y που ερμηνεύεται από την παλινδρόμηση. Όταν όλα τα σημεία βρίσκονται πάνω στην ευθεία παλινδρόμησης τότε το SSE =0 και το R^2=1 (τέλεια προσαρμογή).

Το MSE (mean squared error) είναι το μέσο τετραγωνικό σφάλμα, $MSE=\frac{SSE}{n-2}$ και το MSR (regression mean squared) είναι το μέσο τετράγωνο παλινδρόμησης $MSR=\frac{SSR}{1}$.

Για να προσδιοριστεί η σημαντικότητα της τιμής του $\mathbf{R^2}$, δηλαδή το εάν είναι σημαντικά διάφορη του μηδενός, πρέπει να υπολογιστεί ο **δείκτης F** για ένα επίπεδο σημαντικότητας. Η σχέση που προσδιορίζει το F είναι $\mathbf{F}=\frac{MSR}{MSE}$

Με τον υπολογισμό του F από την παραπάνω σχέση διαπιστώνεται εάν είναι μεγαλύτερο από το κρίσιμο F που καθορίστηκε για ένα διάστημα εμπιστοσύνης και κατ' επέκταση εάν υπάρχει συσχέτιση της ανεξάρτητης μεταβλητής X με την εξαρτημένη Y. Για να πραγματοποιηθεί αυτός ο έλεγχος είναι απαραίτητος ο πίνακας που προκύπτει από την ανάλυση διακύμανσης (διασποράς) **AVOVA** και για να αποδειχθεί στατιστικά σημαντική η ύπαρξη του μοντέλου, θα πρέπει να απορριφθεί η

μηδενική υπόθεση του ελέγχου της τιμής του κριτηρίου F. Δηλαδή η τιμή του F που θα υπολογιστεί από τον πίνακα ANOVA πρέπει να είναι μεγαλύτερη από την κρίσιμη τιμή του F για n-2 βαθμούς ελευθερίας και συγκεκριμένο περιθώριο σφάλματος (Δαμιανός, κ.ά,2010 · Χαλικιάς κ.ά., 2015).
Παρακάτω παρουσιάζεται ο πίνακας ανάλυσης διακύμανσης ANOVA (πίνακας 3.2):

3.2 Πίνακας ANOVA

	Βαθμοί Ελευθερίας	**Άθροισμα Τετραγώνων**	**Μέσα Τετράγωνα**	**Στοιχεία Ελέγχου**
Παλινδρόμηση	1	SSR	MSR	F=MSR/MSE
Κατάλοιπα	n-2	SSE	MSE	
Σύνολο	n-1	SST		

Πηγή: Χαλικιάς κ. ά (2015) , σελ 239)

- ✓ Έπεται ο έλεγχος σημαντικότητας του συντελεστή γραμμικής παλινδρόμησης. Η πιο κοινή υπόθεση στην απλή γραμμική παλινδρόμηση είναι ότι η πραγματική τιμή του συντελεστή γραμμικής παλινδρόμησης είναι μηδέν. Αυτό υποδεικνύει ότι η εξαρτημένη μεταβλητή ούτε αυξάνεται ούτε μειώνεται, όσο αλλάξει η ανεξάρτητη. Ο ρόλος της δοκιμής αυτής είναι η προστασία από την παραπλάνηση της τυχαίας διακύμανσης των εκτιμήσεων. Εάν η διαφορά μεταξύ της εκτιμούμενης τιμής β_1 και της υποτιθέμενης είναι μεγαλύτερη από μια εύλογη διακύμανση, τότε η μηδενική υπόθεση απορρίπτεται. Εάν λοιπόν η μηδενική υπόθεση είναι H_0: $\beta_1 = m$, όπου το m είναι οποιοδήποτε τιμή που θα θέσουμε (και φυσικά μπορεί να ισούται με μηδέν), η εναλλακτική υπόθεση είναι ότι είναι διάφορη του μηδενός και εάν η μηδενική υπόθεση είναι ότι η εκτιμώμενη $\beta_1 = m$ είναι αληθής, τότε η εκτιμούμενη β_1-m είναι κανονικά κατανεμημένη με μέσο το μηδέν .

Ο τύπος που ελέγχει την υπόθεση είναι ο t- student $\mathbf{t= (\widehat{\beta_1} -m)/s(\widehat{\beta_1})}$ (Rawlings et. al, 1998)

- ✓ Αφού σε κάθε παρατήρηση του δείγματός υπολογίζεται και ένα κατάλοιπο, και καθώς επιθυμούμε το άθροισμα των τετραγώνων των σφαλμάτων να είναι το ελάχιστο δυνατό, αναζητούμε τους κατάλληλους συντελεστές ώστε να ισχύει

ότι $\sum_{i-1}^{n}\left(Yi-\widehat{Y}i\right)^2 = \boldsymbol{min}$. Ένας τρόπος αξιολόγησης του μοντέλου παλινδρόμησης είναι μέσω αυτών των καταλοίπων με τα οποία υπολογίζεται το τυπικό σφάλμα εκτίμησης του Y (standard error of the estimate). Ο τύπος που ισχύει είναι $\mathbf{S_{e=}}\sqrt{\frac{\sum e_i^2}{n-2}} = \sqrt{\frac{SSE}{n-2}}$ και δείχνει το σφάλμα που αντιστοιχεί κατά μέσο όρο σε κάθε παρατήρηση. Μικρό τυπικό σφάλμα δείχνει ότι οι παρατηρούμενες και οι εκτιμούμενες τιμές δεν διαφέρουν πολύ από την ευθεία παλινδρόμησης (Ηλιοπούλου, 2015 · Παπαδόπουλος, χ.η).

- ✓ Για τον έλεγχο του μοντέλου και για την εξαγωγή ασφαλών συμπερασμάτων όσο αφορά την αξιοπιστία του, βασικό σημείο είναι η εξασφάλιση ότι οι παραδοχές των βασικών υποθέσεων ισχύουν. Δηλαδή ότι ισχύει η κανονικότητα $\varepsilon_i \sim N(0,\sigma^2)$, η ομοσκεδαστικότητα $var(\varepsilon_ι) = \sigma_\varepsilon^2$ και η ανεξαρτησία, δηλαδή ότι πρόκειται για ανεξάρτητες και τυχαίες μεταβλητές $cov(\varepsilon_i,\varepsilon_j)=0$. Η παραβίαση αυτών των παραδοχών δημιουργεί ανασφάλεια και υποδεικνύει αναγκαιότητα αμφισβήτησης των αποτελεσμάτων και ερμηνειών των συντελεστών της παλινδρόμησης καθώς και της ακαταλληλότητας του μοντέλου που δημιουργήθηκε. Οι παραπάνω έλεγχοι βασιζόμενοι στα κατάλοιπα ανακλούν ιδιότητες των σφαλμάτων (Δαμιανός κ.ά, 2010).

Τα κατάλοιπα είναι το ανερμήνευτο κομμάτι της παλινδρόμησης και πρέπει να αποδειχθεί ότι ακολουθούν την κανονική κατανομή. Αυτό μπορεί να εξασφαλιστεί σε πρώτη φάση κατασκευάζοντας το ιστόγραμμα σχετικών συχνοτήτων που αναμένεται να έχει την μορφή καμπύλης ή από έναν οπτικό έλεγχο στο διάγραμμα κανονικότητας καταλοίπων, όπου εάν τα σημεία τείνουν να βρίσκονται σε μια ευθεία αποδεχόμαστε την υπόθεση ότι δεν παραβιάζεται η κανονικότητα (Μηλάκα, 2010 · Δαμιανός, κ.ά,2010). Εξασφαλίζεται επίσης και με ένα έλεγχο υπόθεσης που εφαρμόζεται στις τιμές τους. Στην μηδενική υπόθεση του ελέγχου ορίζεται η ύπαρξη κανονικότητας κατανομής, έναντι της εναλλακτικής υπόθεσης που υποθέτει τη μη-κανονικότητα. Δηλαδή ισχύει ότι:

***H_0**: Τα κατάλοιπα κατανέμονται κανονικά*

&

***H_1**: Τα κατάλοιπα δεν κατανέμονται κανονικά*

(Πραμαγγιούλης, 2008, ·Πετρίδης, 2015)

Ακόμα γίνεται μέσα από τους δείκτες κύρτωσης και στρέβλωσης που υπολογίζονται αυτόματα μέσω στατιστικών πακέτων. Τέλος ο έλεγχος της κανονικότητας συνήθως γίνεται και με το κριτήριο Kolmogorov- Smirnov σε μέτριου μεγέθους δείγματα και Shapiro- Wilk ως πιο αξιόπιστος για μικρά δείγματα, με την μηδενική υπόθεση να απορρίπτεται όταν το p-value είναι μεγαλύτερο του 0,05 (Πραμαγγιούλης, 2008, ·Δαμιανός, κ.ά,2010· Μηλάκα, 2010).

Ένα διάγραμμα σημείων των καταλοίπων (διάγραμμα διασποράς) θα διαπιστώσει την ύπαρξη ομοσκεδαστικότητας εφόσον τα σημεία αυτά κατανέμονται τυχαία χωρίς να παρουσιάζουν κάποια τάση. Αναμενόμενη και επιθυμητή είναι η διάσπαρτη ομοιόμορφα κατανεμημένη απεικόνιση γύρω από το μηδέν. Όταν η διακύμανση των καταλοίπων είναι σταθερή τότε το υπόδειγμα θεωρείται ότι χαρακτηρίζεται από ομοσκεδαστικότητα, δηλαδή ισχύει ότι $var(\varepsilon_t) = \sigma_\varepsilon^2$ σταθερό. Σε αντίθετη περίπτωση στο υπόδειγμα υπάρχει ετεροσκεδαστικότητα, δηλαδή σε διαφορετικές παρατηρήσεις της ανεξάρτητης μεταβλητής Χ η διακύμανση του καταλοίπου δεν είναι σταθερή . Εάν έχουμε ένα μοντέλο στο οποίο ισχύουν όλες οι βασικές προϋποθέσεις εκτός από την ομοσκεδαστικότητα , τότε υπάρχει πρόβλημα στις εκτιμήσεις των διακυμάνσεων των εκτιμητών των συντελεστών και στην αποτελεσματικότητας τους (Πραμαγγιούλης, 2008· Χαλικίας,κ.ά, 2015· Μηλάκα, 2010).

Η ανεξαρτησία των καταλοίπων μπορεί να διαπιστωθεί από ένα διάγραμμα σημείων μεταξύ προβλεπομένων τιμών και καταλοίπων και εφόσον τα σημεία είναι τυχαία, τότε διαπιστώνεται η ανεξαρτησία ενώ η γραμμικότητα διαπιστώνεται από ένα διάγραμμα σημείων (scatter plot) προβλεπόμενων τιμών έναντι των καταλοίπων (Πραμαγγιούλης, 2008).

3.4.2 Πολλαπλή Παλινδρόμηση (multiple regression)

Η διακύμανση ενός μοντέλου σπάνια εξηγείται από μια μόνο μεταβλητή. Η διαφορά της πολλαπλής γραμμικής παλινδρόμησης με την απλή, είναι ότι υπάρχουν περισσότερες της μίας ανεξάρτητες μεταβλητές που λαμβάνουν μέρος στο μοντέλο και λειτουργούν συνδυασμένα. Έτσι η γραμμική σχέση του μοντέλου για k ανεξάρτητες μεταβλητές προσδιορίζεται από την εξίσωση:

$$\boldsymbol{y = \beta_0 + \beta_1 \cdot x_1 + \beta_2 \cdot x_2 + \cdots .. + \beta_k \cdot x_k + \varepsilon}$$

όπου $\boldsymbol{y}$ είναι η τιμή της εξαρτημένης μεταβλητής,

$\mathbf{x_1}$ έως $\mathbf{x_k}$ είναι οι τιμές των ανεξάρτητων μεταβλητών,

$\boldsymbol{\beta_0}$ είναι ο σταθερός όρος, δηλαδή η τιμή της y όταν όλες οι x είναι μηδενικές,

οι $\boldsymbol{\beta_j}$ **, j=1,…,k** είναι οι συντελεστές κλίσης (regression coefficients) ή μερικοί συντελεστές παλινδρόμησης, που αντιστοιχούν σε κάθε μια ανεξάρτητη μεταβλητή, εκφράζοντας το μέγεθος μεταβολής της Y σε κάθε μεταβολή της X κατά μία μονάδα, ενώ τέλος $\boldsymbol{\varepsilon}$ είναι τα κατάλοιπα (Πετρίδης, 2015· Ηλιοπούλου, 2017·Montgomery, et al, 2012). Η εξίσωση της πολλαπλής παλινδρόμησης είναι μια ευθεία γραμμή όταν στο μοντέλο συμμετέχει μία μόνο ανεξάρτητη μεταβλητή, επίπεδο όταν συμμετέχουν δύο και υπερεπίπεδο όταν συμμετέχουν από τρείς και πάνω (Πανάρετος,1994 · Πετρίδης, 2015).

Η μέθοδος των ελαχίστων τετραγώνων χρησιμοποιείται και εδώ όπως και στην απλή γραμμική παλινδρόμηση για τη εκτίμηση των παραμέτρων της εξίσωσης. Σκοπός της είναι η ελαχιστοποίηση του αθροίσματος των τετραγώνων αποκλίσεων παρατηρήσεων από την γραμμή παλινδρόμησης, δηλαδή $\sum_{i=1}^{n}(y_i - \hat{y})^2$=min, ενώ σύμφωνα με τους Montgomery, et al (2012) η απαίτηση των ελαχίστων τετραγώνων είναι η $\mathbf{S(\beta_0,\beta_1,....,\beta_k)}=\sum_{i=1}^{v} \varepsilon_i^{\,2}$ όπου $\varepsilon_i = \sum_{i=1}^{n} y_i - \beta_0 - \sum_{j=1}^{k} \beta_j x_{ij}$ να είναι η μικρότερη δυνατή.

Η λύση επιτυγχάνεται με την επίλυση του συστήματος των κανονικών εξισώσεων (Πανάρετος, 1994).

Σύμφωνα με τον Πετρίδη (2015) τα μοντέλα της πολλαπλής παλινδρόμησης μπορούν να προσδιοριστούν ως επεξηγηματικά, που επιδιώκουν την εδραίωση ενός ισχυρού μοντέλου επιβεβαίωσης των αποτελεσμάτων της πρόβλεψης και προβλεπτικά, τα οποία βασίζονται στη ανάλυση των στοιχείων και επιχειρούν την ανεύρεση της κατάλληλης προβλεπτικής ομάδας μεταβλητών που θα παράξει το καλύτερο προβλεπτικό αποτέλεσμα. Σημειώνεται ότι ο έλεγχος των παραδοχών της κανονικότητας, ομοσκεδαστικότητας, ανεξαρτησίας και γραμμικότητας είναι απαραίτητος και στην πολλαπλή παλινδρόμηση όπως και στην απλή, καθώς εάν κάποια από τις βασικές προϋποθέσεις παραβιάζεται, αυτό υποδεικνύει πρόβλημα στην ασφαλή εξαγωγή συμπερασμάτων και αποτελεσμάτων από το μοντέλο μας. Επιπρόσθετα δεν θα πρέπει να υπάρχει πολυσυγραμμικότητα μεταξύ των ανεξάρτητων μεταβλητών καθώς υπάρχει περίπτωση επικάλυψης της επίδρασή τους

στην διαμόρφωση των τιμών της εξαρτημένης (Ηλιοπούλου, 2015 · Χαλικιάς, κ. ά, 2015).

Πρώτο βήμα της πολλαπλής γραμμικής παλινδρόμησης είναι η μελέτη της εξαρτημένης μεταβλητής ως προς την κατανομή που ακολουθεί και η οποία πρέπει να είναι κανονική με σταθερή διακύμανση. Η κατανομή της εξαρτημένης μεταβλητής αποτυπώνεται και ελέγχεται μέσω ιστογράμματος, ενώ η στατιστική επιβεβαίωση γίνεται μέσω του μη παραμετρικού ελέγχου Kolmogotov Smirnof, ώστε να επιβεβαιωθεί ή να απορριφθεί η αρχική υπόθεση ότι η κατανομή της εξαρτημένης μεταβλητής είναι κανονική. Ένδειξη μη κανονικής κατανομής στην εξαρτημένη μεταβλητή μας οδηγεί σε επιπρόσθετη εργασία κατάλληλων μετασχηματισμών (πχ. λογαρίθμηση) ή ακόμα και αφαίρεση ακραίων τιμών.

Έπεται η εξέταση της σχέσης της κάθε μιας ανεξάρτητης μεταβλητής με την εξαρτημένη που θα πρέπει να είναι ισχυρή και γραμμική. Ο υπολογισμός της συσχέτισης της εξαρτημένης μεταβλητής με τις ανεξάρτητες μπορεί να γίνει μέσω πολλαπλών διαγραμμάτων με ταυτόχρονη απεικόνιση όλων των συνδυασμών.

Ακολουθεί η μελέτη της σχέσης των ανεξάρτητων μεταβλητών μεταξύ τους. Η εξέταση των συσχετίσεων μεταξύ των ανεξάρτητων μεταβλητών γίνεται μέσω πινάκων συσχετίσεων, ώστε να αποφασιστεί το ποιες μεταβλητές τελικά θα συμμετάσχουν στο μοντέλο. Οι μεταβλητές αυτές θα δώσουν ένα ισχυρό μοντέλο εφόσον δεν είναι ισχυρά συνδεδεμένες μεταξύ τους. Εάν είναι, τότε το πρόβλημα που εμφανίζεται είναι η πολυσυγγραμμικότητα, η οποία παρατηρείται όταν οι ανεξάρτητες μεταβλητές έχουν γραμμική συσχέτιση μεταξύ τους. Ως συνέπεια αυτού του γεγονότος είναι οι συντελεστές κλίσης των μεταβλητών που συσχετίζονται γραμμικά να εμφανίζουν μεγάλη μεταβλητότητα, οδηγώντας εσφαλμένα στο συμπέρασμα πως δεν υπάρχει ισχυρή γραμμική συσχέτιση. Αυτό οφείλεται στο γεγονός ότι κάποιες μεταβλητές ερμηνεύουν το ίδιο φαινόμενο. Ένας τρόπος αντιμετώπισης είναι η απόρριψη της μεταβλητής που σχετίζεται εντονότερα με τις άλλες .

Έπεται η κατασκευή του μοντέλου με τον ορισμό της εξαρτημένης και των ανεξάρτητων μεταβλητών που θα επιλέξουμε. Η εξαρτημένη μεταβλητή στο μοντέλο πολλαπλής παλινδρόμησης, είναι αυτή της οποίας θέλουμε να ερμηνεύσουμε τις μεταβολές, αξιολογώντας τις τιμές των υπόλοιπων ανεξάρτητων μεταβλητών. Ο αριθμός των ανεξάρτητων δεν πρέπει να είναι πολύ μεγάλος, ώστε να δημιουργήσουμε ένα μοντέλο εύκολα διαχειρίσιμο. Άλλωστε η τοποθέτηση όλων των

ανεξάρτητων μεταβλητών στο μοντέλο μπορεί να μας οδηγήσει σε πρόβλημα υπερπροσαρμογής . Οι ανεξάρτητες μεταβλητές θα πρέπει πρώτα να εξετασθούν ως προς τη σημαντικότητά τους μία προς μία και μετά οι συνδυασμοί αυτών. Η διαδικασία που ακολουθείται είναι μια από αυτές που θα περιγραφούν αναλυτικά στο επόμενο κεφάλαιο και είναι οι Forward, Backward, Stepwise.
Με τον προσδιορισμό του μοντέλου διαπιστώνουμε και την σχέση μεταξύ των μεταβλητών καθώς και τον τρόπο σύνδεσής τους, ενώ διαπιστώνουμε και την στατιστική σημαντικότητα των ανεξάρτητων μεταβλητών που χρησιμοποιήσαμε. Καθώς από το μοντέλο αυτό προσδιορίζεται τιμή πρόβλεψης για την εξαρτημένη μεταβλητή είναι δυνατή η εξαγωγή συμπερασμάτων, καθώς προκύπτει η διαφορά μεταξύ της τιμής που μπορεί να παρέχει η πρόβλεψη του μοντέλου με την πραγματική τιμή. Η διαφορά αυτή που πιθανά υπάρχει αποτελεί το σφάλμα, η τυχαιότητα του οποίου επιβεβαιώνει την μη ύπαρξη συστηματικού παράγοντα σφάλματος συμμετοχής στο τρόπο που συνδυάζεται η εξαρτημένη μεταβλητή με τις ανεξάρτητες.

Τέλος μετά τον υπολογισμό του μοντέλου της γραμμικής παλινδρόμησης ακολουθεί ο έλεγχος της κανονικότητας των καταλοίπων (Δημητροπουλάκης, χ.η). Οι έλεγχοι υποθέσεων που βοηθούν τον ερευνητή να καθορίσει την σημαντικότητα των μεταβλητών στηρίζονται στον τύπο $\mathbf{t} = \frac{\widehat{\boldsymbol{\beta\iota}}}{\mathbf{S}_{\widehat{\boldsymbol{\beta\iota}}}}$ όπου το $\widehat{\boldsymbol{\beta\iota}}$ είναι η εκτιμήτρια των ελαχίστων τετραγώνων του συντελεστή βi της μεταβλητής xi στο γενικό γραμμικό μοντέλο και το $\mathbf{S}_{\widehat{\boldsymbol{\beta\iota}}}$ είναι η εκτιμούμενη τυπική απόκλιση της εκτιμήτριας $\widehat{\beta\iota}$. (Πανάρετος, 1994). Ο έλεγχος σημαντικότητας της γραμμικότητας της πολλαπλής παλινδρόμησης και κατ' επέκταση η εγκυρότητα του μοντέλου, γίνεται με τον έλεγχο του πίνακα ανάλυσης διακυμάνσεων ANOVA. Είναι ένα τεστ για να προσδιοριστεί η γραμμική σχέση μεταξύ της εξαρτημένης μεταβλητής y και οποιασδήποτε από τις ανεξάρτητες μεταβλητές $x_1,x_2,\ldots x_k$ Η μηδενική υπόθεση είναι ότι

$$\boldsymbol{H_0}: \beta_1=\beta_2=\ldots.=\beta_k=0$$

έναντι της εναλλακτικής

$$\boldsymbol{H1}: \beta j\neq 0 \text{ για τουλάχιστον ένα } j$$

Έτσι αποδεικνύεται ότι τουλάχιστον μια από τις μεταβλητές $x_1,x_2,\ldots.x_k$ συμβάλουν σημαντικά στο μοντέλο. Ο έλεγχος αυτός στηρίζεται στην συνάρτηση F και ελέγχει την ορθότητα της υπόθεσης ότι οι συντελεστές όλων των ανεξάρτητων μεταβλητών, που συμμετέχουν στο μοντέλο είναι ταυτόχρονα μηδέν και αποτυπώνεται με τον τύπο

F=MSR/MSE. Ο έλεγχος της μεμονωμένης συνεισφοράς κάθε ανεξάρτητη μεταβλητής γίνεται με την χρήση των partial F- tests. (Montgomery, et al, 2012· Χαλικίας κ. ά., 2015).

Με σκοπό να αντιληφθούμε πόσο καλά το μοντέλο ταιριάζει στα δεδομένα μας χρησιμοποιούμε τον συντελεστή προσδιορισμού $\mathbf{R^2}$, ο οποίος παίρνει τιμές από 0 έως 1 και μπορεί να εκφραστεί και ως ποσοστό. Ο συντελεστής αυτός δείχνει την αναλογία της εξηγήσιμης μεταβλητότητας από το μοντέλο ως προς την συνολική μεταβλητότητα, ενώ τιμή κοντά στο +1 υποδηλώνει μεγάλο βαθμό ερμηνείας της μεταβλητότητας της εξαρτημένης μεταβλητής από τις ανεξάρτητες μεταβλητές που χρησιμοποιηθήκαν στο μοντέλο (Ηλιοπούλου, 2017).

Ο παράγοντας μεγέθυνσης διασποράς (VIF Variance inflation factor) είναι ένα μέτρο ελέγχου πολυσυγγραμικότητας που μας δείχνει την αύξηση της διασποράς ενός εκτιμημένου συντελεστή παλινδρόμησης. Τιμή του VIF < 5 δείχνει πολυσυγγραμικότητα των ανεξάρτητων μεταβλητών και είναι ένδειξη ότι πρέπει να αφαιρεθούν από το μοντέλο μας οι μη σημαντικοί παράγοντες **(**Χαλικιάς κ. ά, 2015· Montgomery, et al, 2012).

Με την δημιουργία του μοντέλου αποκτούμε μια βέλτιστα εκτιμωμένη ευθεία. Εάν αφαιρέσουμε κάθε τιμή της εξαρτημένης μεταβλητής από την εκτίμηση της, τότε έχουμε υπολογίσει τα κατάλοιπα, η μελέτη της συμπεριφοράς των οποίων μας δίνει πληροφορίες ως προς την εγκυρότητα, του μοντέλου μας. Για να ελέγξουμε την κανονικότητα των καταλοίπων θα πρέπει να ελεγχθεί εάν ακολουθούν κανονική κατανομή. Η ύπαρξη ακραίων τιμών ή μικρό δείγμα είναι μερικοί λόγοι στους οποίους μπορεί να οφείλεται η μη κανονικότητα αυτών. Ως λύση θα πρέπει να εξετασθεί η αφαίρεση των ακραίων τιμών ή το μεγάλωμα του δείγματος. Παράλληλα η μη ύπαρξη αυτοσυσχέτισης των καταλοίπων εξασφαλίζει τη τυχαία διασπορά τους καθώς η ύπαρξη αυτοσυσχέτισης υποδεικνύει ότι κάποια μεταβλητή τα κάνει να αντιδρούν με τέτοιο τρόπο.

Η υπόθεση της κανονικότητας των καταλοίπων γίνεται σε επόμενο στάδιο με οπτική επιβεβαίωση. Όπως και στην απλή, έτσι και στην πολλαπλή παλινδρόμηση η κανονικότητα των καταλοίπων αποτελεί έλεγχο εγκυρότητας του μοντέλου (Χαλικιάς κ. ά, 2015).

Διάφοροι μέθοδοι έχουν αναπτυχθεί ώστε να διευκολύνεται η δημιουργία του μοντέλου γραμμικής παλινδρόμησης, στηριζόμενοι κυρίως σε προγράμματα υπολογιστών. Οι μέθοδοι αυτοί είναι της προοδευτικής προσθήκης μεταβλητών, του αποκλεισμού μεταβλητών και της βηματικής παλινδρόμησης, ενώ αναφέρονται και η μέθοδος της απλής εισαγωγής και αφαίρεσης μεταβλητών από το μοντέλο.

3.4.2.1 Μέθοδοι Επιλογής Μεταβλητών

Επιλογή μεταβλητών είναι αυτή η διαδικασία μέσα από την οποία επιλέγεται μία ομάδα μεταβλητών από το σύνολο των διαθέσιμων, με την οποία θα πραγματοποιηθεί η παλινδρόμηση. Κατά τη διαδικασία αυτή εξαιρούνται οι περιττές, δηλαδή αυτές που δεν προσθέτουν κάποια πληροφορία στο μοντέλο. Ως πλεονέκτημα αυτής της διαδικασίας είναι η μείωση του χρόνου επεξεργασίας και του κόστους, με παράλληλα καλή αποτελεσματικότατα του μοντέλου πρόβλεψης .

Υπάρχουν διάφορες μέθοδοι επιλογής μεταβλητών:

- **Η Forward procedure ή forward selection** (μέθοδος προοδευτικής προσθήκης μεταβλητών), όπου γίνεται εντοπισμός, μεταξύ των ανεξάρτητων μεταβλητών, εκείνης της μεταβλητής που συνδέεται ισχυρότερα με την εξαρτημένη και γίνεται έλεγχος της σημαντικότητας της. Σε επόμενο στάδιο επιλέγεται η μεταβλητή με την αμέσως υψηλότερη συσχέτιση με την εξαρτημένη και την χαμηλότερη συσχέτιση με την ανεξάρτητη μεταβλητή που συμμετέχει ήδη στο μοντέλο, ενώ ακολουθεί ο έλεγχος σημαντικότητας των μεταβλητών και του μοντέλου. Η διαδικασία αυτή επαναλαμβάνεται μέχρι που να μην υπάρχει άλλη σημαντική μεταβλητή ως προς την εξαρτημένη (Κουγιουμτζής, 2011·Montgomery, Peck,1992· Ηλιοπούλου, 2017).

- Η **Backward procedure ή backward elimination ή backward selection** (μέθοδος αποκλεισμού μεταβλητών) ακολουθεί τον αντίθετο δρόμο με την forward. Στη μέθοδο αυτή εισάγονται όλες οι μεταβλητές στο μοντέλο και σε κάθε διαδοχικό βήμα αφαιρείται η λιγότερο σημαντική για την εξαρτημένη μεταβλητή. Η πρώτη που αφαιρείται από τον μοντέλο είναι αυτή με το μικρότερο R2 της απλής, εφόσον δεν συνεπάγεται σημαντική μείωση του R2 της πολλαπλής παλινδρόμησης Η διαδικασία επαναλαμβάνεται τόσες φορές ώστε τελικά μόνο σημαντικές μεταβλητές, ισχυρά συσχετισμένες με την εξαρτημένη και ασυσχέτιστες μεταξύ τους να παραμείνουν στο μοντέλο (Ηλιοπούλου, 2015).

• Η **Stepwise regression (μέθοδος βηματικής παλινδρόμησης),** είναι συνδυασμός της forward και της backward παλινδρόμησης, όπου σε κάθε βήμα προστίθεται ή αφαιρείται μια μεταβλητή από ένα σύνολο, βασιζόμενη σε ένα προκαθορισμένο κριτήριο ενώ ««σε κάθε βήμα ελέγχεται αν οι μεταβλητές οι οποίες έχουν ήδη προστεθεί είναι ακόμα σημαντικές «» (Ηλιοπούλου, 2015· Montgomery et al, 1992).

• Η διαδικασία **Enter** προβλέπει την εισαγωγή όλων ή επιλεγμένων μεταβλητών, στατιστικά σημαντικών και μη, στο μοντέλο. Αποτελεί την προεπιλεγμένη μέθοδο στο λογισμικό SPSS και ουσιαστικά υπολογίζει την επίδραση ταυτόχρονα όλων των ανεξάρτητων μεταβλητών πάνω στην εξηρτημένη (Παυλόπουλος, 2006).

• Η διαδικασία **Remove** όπου οι μεταβλητές απομακρύνονται ως μπλοκ σε ένα βήμα από το μοντέλο μέσα από το στατιστικό λογισμικό (Χαλικιάς, κ.ά 2015).

3.4.2.2 Κατασκευή ενός Τυπικού Μοντέλου Πολλαπλής Παλινδρόμησης Μέσω Στατιστικού Προγράμματος και Προγράμματος Gis

Οι Αρβανίτης, Λαφαζάνη & Σταθάκης (2010) προτείνουν την εφαρμογή της πολλαπλής παλινδρόμησης σε λογισμικό στατιστικής ανάλυσης με τα παρακάτω στάδια επεξεργασίας:

• Εφαρμογή πολλαπλής παλινδρόμησης ώστε να διαπιστωθεί ότι η εξαρτημένη εξαρτάται γραμμικά από τις ανεξάρτητες μεταβλητές.

- Τεκμηρίωση για την λειτουργία του μοντέλου.
- Κατασκευή γραφημάτων συσχέτισης της εξαρτημένης με τις ανεξάρτητες μεταβλητές, ώστε σε πρώτη φάση να απομακρυνθούν από τη διαδικασία οι μη έχουσες καμία γραμμική σχέση.
- Κατασκευή πινάκων συσχέτισης και έλεγχος της γραμμικής σχέσης μέσω του συντελεστή συσχέτισης Pearson και με μηδενική υπόθεση ότι δεν υπάρχει γραμμική συσχέτιση μεταξύ της εξαρτημένης με κάθε ανεξάρτητη (H_0: r=0), έναντι της εναλλακτικής ότι υπάρχει γραμμική συσχέτιση με κάθε ανεξάρτητη (H_1: r≠0) . Εάν το p-value>0,05, δηλαδή η παρατηρούμενη πιθανότητα είναι μικρότερη του α=0,05, τότε εμφανίζεται σημαντική γραμμική συσχέτιση.
- Εισαγωγή μεταβλητών στο μοντέλο με stepwise μέθοδο για μια πρώτη προσέγγιση αυτού.

- Παραδοχές και προϋποθέσεις δημιουργίας μοντέλου.
    - Έλεγχος σχέσης μεγέθους δείγματος και ανεξάρτητων μεταβλητών σύμφωνα με την βιβλιογραφία.
    - Έλεγχος σημαντικότητας ανεξάρτητων μεταβλητών, για την ερμηνευσιμότητα και προβλεπτική ικανότητα του μοντέλου μέσω γραφημάτων και τιμών t . Τα διαγράμματα μερικής παλινδρόμησης ή τα διαγράμματα μερικών υπολοίπων καταδεικνύουν το είδος και την ένταση της σχέσης μεταξύ της εξαρτημένης μεταβλητής και μιας ανεξάρτητης, χωρίς την επιρροή των υπολοίπων μεταβλητών. Στον έλεγχο αυτό αναζητούμε οπτικά γραμμική σχέση. Όσο αφορά τις τιμές t, όσο μεγαλύτερη είναι η τιμή αυτή του t- test για τον αντίστοιχο μερικό συντελεστή παλινδρόμησης β_j, τόσο πιο σημαντική είναι η ανεξάρτητη μεταβλητή που εξετάζεται. Για τιμές $t>\pm 2,00$ ισχύει η σημαντικότητα της εξάρτησης.
    - Έλεγχος της ανεξαρτησία οντοτήτων μέσω του δείκτη Durbin-Watson και ζητούμενο οι τιμές του να βρίσκονται γύρω από το 2, ώστε να αποδειχθεί η ανεξαρτησία μεταξύ διαδοχικών υπολοίπων, με παράλληλο έλεγχο της σημαντικότητας του τεστ μέσω του F- test . Ο έλεγχος της ανεξαρτησία οντοτήτων γίνεται και μέσω του γραφήματος σκέδασης, όπου η μη συστηματική και ανομοιόμορφη συσσώρευση σημείων καθώς και η έλλειψη προτύπων στο διάγραμμα, καταδεικνύει την ανεξαρτησία υπολοίπων.
    - Η μη ύπαρξη συγγραμμικότητας ή πολυσυγγραμμικότητας του μοντέλου ελέγχεται με τον συντελεστή VIF, ενώ τιμές του μεγαλύτερες από 2, στην κλίμακα που υπολογίζεται από το λογισμικό SPSS, αποτελούν ένδειξη προβλήματος συγγραμμικότητας του μοντέλου.
    - Ο έλεγχος της κανονικότητας γίνεται με το ιστόγραμμα ή με το Normal P-P plot of regression standardized residuals ή με το φυλλογράφημα υπολοίπων, ενώ παράλληλα ελέγχεται η ύπαρξη ακραίων οντοτήτων, οι οποίες και θα πρέπει να απομακρυνθούν.
    - Η ομοσκεδαστικότητα ελέγχεται οπτικά με το γράφημα σκέδασης, καθώς εφόσον υπάρχει γραμμική σχέση της εξαρτημένης με τις ανεξάρτητες μεταβλητές, τα κατάλοιπα βρίσκονται στο εσωτερικό μιας οριζόντιας περιοχής σταθερού εύρους. Ελέγχεται ακόμα και με το διάγραμμα διασποράς εκτιμώμενων και παρατηρούμενων τιμών, ενώ υπάρχει η αναγκαιότητα τα σημεία να αναπτύσσονται ομοιόμορφα στο εσωτερικό μιας περιοχής σταθερού εύρους,

εκατέρωθεν της ευθείας των ελαχίστων τετραγώνων, ώστε να διαπιστωθεί η σταθερή διασπορά των υπολοίπων.

- Το μοντέλο ερμηνεύεται με τον δείκτη R^2 . Όσο μεγαλύτερη τιμή έχει τόσο πιο έντονη σχέση έχει η εξαρτημένη μεταβλητή με τις ανεξάρτητες που το απαρτίζουν. Οι συντελεστές Beta δείχνουν τη σχετική βαρύτητα κάθε ανεξάρτητης μεταβλητής στο μοντέλο και προκύπτουν από τα αποτελέσματα της στατιστικής ανάλυσης, ενώ ανάλογα μπορούν να αφαιρεθούν οι λιγότερο σημαντικές, εφόσον φυσικά δεν μεταβάλλεται προς τα κάτω ο συντελεστής R^2.
- Για την περίπτωση χωρικών δεδομένων το μοντέλο που παρήχθη με την χρήση λογισμικού στατιστικής ανάλυσης εισάγεται και σε πρόγραμμα Gis μέσω της διαδικασίας της εφαρμογής του OLS και πρώτο έλεγχο την ύπαρξη ή μη των σωστών πρόσημων των συντελεστών παλινδρόμησης.
- Η στατιστική σημαντικότητα εκφράζεται μέσω της ύπαρξης του αστερίσκου στην στήλη της Πιθανότητας (Probability), ενώ κάθε συντελεστής που δεν έχει αστερίσκο πρέπει να απομακρύνεται.
- Η ύπαρξη του προβλήματος της πολυσυγγραμμικότητας στο μοντέλο διαπιστώνεται με τον δείκτη VIF και αναμενόμενη τιμή κάτω από 7,5, στην κλίμακα που υπολογίζεται από το λογισμικό Arc Gis ώστε να αποδειχθεί ότι το μοντέλο δεν παρουσιάζει τέτοιο πρόβλημα.
- Η εξέταση της συνολικής απόδοσης του μοντέλου γίνεται και εδώ με τον δείκτη R^2 και την τιμή AIC. Όσο μεγαλύτερη τιμή του R^2 και όσο μικρότερη του AIC τόσο καλύτερο το μοντέλο.
- Ο έλεγχος της κανονικής κατανομής των standardized residuals γίνεται με το τεστ Jarque_Beta, όπου αν το p-value <0,05, υποδεικνύεται ότι τα κατάλοιπα ακολουθούν κανονική κατανομή.
- Ο πιο σημαντικός έλεγχος που αφορά την ανεξαρτησία των υπολοίπων είναι αυτός που γίνεται μέσω του δείκτη Moran's I. Το σκεπτικό της όλης διαδικασίας είναι να τεκμηριωθεί ότι το μοντέλο που παράγεται από το στατιστικό πρόγραμμα έχει όμοια προσέγγιση με το μοντέλο που παράγεται από πρόγραμμα Gis.

3.4.3 Χωρική Παλινδρόμηση

Ο Tobler το 1970 στην εργασία του «« A computer movie simulating urban growth in the Detroit region» διατύπωσε τον πρώτο νόμο της Γεωγραφίας ως εξής: *Όλα τα*

πράγματα συνδέονται μεταξύ τους αλλά τα πιο κοντινά συνδέονται περισσότερο απ΄ότι τα μακρινά» (Tobler, 1970,σελ. 236).

Η **χωρική αυτοσυσχέτιση** αναφέρεται στο βαθμό με τον οποίο σχετίζονται, μέσω της χωρικής τους απόστασης στο γεωγραφικό χώρο, ένα ζευγάρι τιμών μιας μεταβλητής. Καθώς τα δεδομένα που αναφέρονται σε τοποθεσίες και βρίσκονται κοντά στον γεωγραφικό χώρο είναι πιθανότερο να παρουσιάζουν ομοιότητες εν σχέση με αυτά που βρίσκονται μακριά, με την ανάλυση των γεωγραφικών φαινόμενων αναζητούμε τα χωρικά πρότυπα και τις διαδικασίες που τα προκαλούν με τη μελέτη της χωρικής αυτοσυσχέτισης. Κατά τη εξέταση των ομοιοτήτων και διαφορών τιμών κάποιου χαρακτηριστικού σε διαφορετικές θέσεις, λαμβάνοντας υπόψη την απόσταση των θέσεων αυτών μεταξύ τους (γεωγραφική εγγύτητα), παρουσιάζεται είτε θετική αυτοσυσχέτιση, που υποδεικνύει παρόμοιες τιμές γειτονικών περιοχών, είτε αρνητική που υποδεικνύει διαφορετικές ιδιότητες, είτε παντελής έλλειψη που υποδεικνύει τυχαίο χωρικό πρότυπο.

Η μέτρηση της χωρικής αυτοσυσχέτισης γίνεται με ένα, συμμετρικό συνήθως, πίνακα χωρικών βαρών (spatial weights matrix) με συμβολισμό W (πίνακας 3.3), που απεικονίζει τη γεωμετρική δομή του γεωγραφικού χώρου, υπολογίζοντας για κάθε θέση των αριθμό των γειτονικών θέσεων ή με την μέτρηση των αποστάσεων μεταξύ πολυγώνων η οποία ορίζεται από τον μελετητή και μέχρι την οποία θεωρεί αυτός ότι τα πολύγωνα γειτνιάζουν. Αν υποθέσουμε ότι η περιοχή μελέτης χωρίζεται σε **n** ζώνες τότε η διάσταση του πίνακα θα είναι **n** x **n**. Οι θέσεις στον χώρο **i** και **j** απεικονίζονται από τους όρους του στοιχείου $\mathbf{w_{ij}}$ που είναι τα χωρικά βάρη (spatial weights). Στον πίνακα των χωρικών βαρών η σειρά των θέσεων είναι μεν αυθαίρετη, είναι δε ίδια στις γραμμές και στις στήλες, με τα στοιχεία της διαγωνίου να είναι μηδενικά, καθώς δεν μας ενδιαφέρει η σχέση μιας θέσης με τον εαυτό της. Ένας από τους πλέον συνηθισμένους τρόπους ορισμού του πίνακα W είναι να ορίσουμε τη τιμή του $\mathbf{w_{ij} = 1}$ εφόσον το πολύγωνο i έχει κοινό σύνορο με το πολύγωνο j (γειτονικά πολύγωνα), ενώ το $\mathbf{w_{ij} = 0}$ σε οποιαδήποτε άλλη περίπτωση. Άλλος τρόπος ορισμού χωρικών βαρών είναι η μέτρηση αποστάσεων μεταξύ γεωγραφικών οντοτήτων σε συγκεκριμένα σημεία. Για παράδειγμα εάν ο γεωμετρικός μέσος ενός πολυγώνου j είναι ένας από τους n πλησιέστερους σε αυτόν του πολυγώνου i τότε η τιμή του $\mathbf{w_{ij} = 1}$, ενώ σε κάθε άλλη περίπτωση είναι **0.** Εναλλακτική λύση είναι ο ορισμός

συγκεκριμένης απόστασης στα όρια της οποίας υπάρχει η θεώρηση ότι τα πολύγωνα γειτνιάζουν, ενώ πέραν αυτής όχι (Ηλιοπούλου, 2017) .

Πίνακας3.3.Πίνακας χωρικών βαρών.

$$W=\begin{bmatrix} W_{11} & W_{12} & \cdots & W_{1n} \\ W_{21} & W_{22} & & \vdots \\ \vdots & & \ddots & \vdots \\ W_{n1} & \cdots & \cdots & W_{nn} \end{bmatrix}$$

Πηγή: Ηλιοπούλου,2017(σελ 169)

Η χωρική αυτοσυσχέτιση μετριέται συνήθως με τον δείκτη **Ι του Moran** ο οποίος λαμβάνει τιμές από -1 έως +1 (ισχυρή αρνητική και ισχυρή θετική αυτοσυσχέτιση αντίστοιχα). Χρησιμεύει για να εξετάσει την ύπαρξη χωρικής αυτοσυσχέτισης σε χωρικά δεδομένα μιας μεταβλητής. Ουσιαστικά, αυτό που δείχνει είναι το κατά πόσο η παρατηρούμενη χωρική κατανομή παρουσιάζεται ως μη τυχαία στο γεωγραφικό χώρο, δηλαδή υπάρχει παρουσία του φαινόμενου συγκέντρωσης cluster.

Ο στατιστικός έλεγχος του δείκτη Ι του Moran γίνεται με διαδικασία πολλών δοκιμών (Monte Carlo simulation), όπου υπολογίζεται η τιμή της εξαρτημένης μεταβλητής. Έπειτα οι ανεξάρτητες μεταβλητές αναδιανέμονται τυχαία στα πολύγωνα και επαναϋπολογίζεται η εξαρτημένη με τον ίδιο δείκτη. (Ηλιοπούλου, 2015· Καλογήρου, 2015).

Τα δείγματα ενός πληθυσμού από χωρικά δεδομένα στη πραγματικότητα δεν είναι τυχαία, όπως προϋποθέτει η κλασική παλινδρόμηση, αλλά συσχετίζονται χωρικά. Η χωρική αυτοσυσχέτιση λοιπόν αναζητά αυτά τα χωρικά πρότυπα και τις διαδικασίες που τα προκαλούν. Η μέθοδος αυτή που αποτελεί ουσιαστικά μιας μορφής τοπική παλίνδρομη, προσαρμοσμένη στα χωρικά δεδομένα, που στηρίζεται στην μέτρηση της χωρικής αυτοσυσχέτισης, αναφέρεται με τον όρο **Χωρική Παλινδρόμηση** (Ηλιοπούλου, 2015).

Σύμφωνα με τη Στάμου (2016), οι παρατηρήσεις οι οποίες λαμβάνονται υπόψη για την δημιουργία ενός υποδείγματος χωρικής παλινδρόμησης σταθμίζονται

με βάρος τη γεωγραφικής τους θέση, ενώ σύμφωνα με τον Παπαευθυμίου (2013), η ύπαρξη χωρικής συσχέτισης οφείλεται να διερευνάται, καθώς τα δεδομένα με χωρική αναφορά έχουν λιγότερους βαθμούς ελευθερίας σε σχέση με αυτά που δεν τη περιέχουν .

Η κλασική παλινδρόμηση δεν λαμβάνει υπόψη την γεωγραφική διάσταση των φαινομένων, σε αντίθεση με την χωρική, που λαμβάνει υπόψη την γειτνίαση των σημείων και τις ομοιότητες που έχουν τα γειτονικά μεταξύ τους, λόγω της φυσικής συνέχειας του γεωγραφικού χώρου. Ουσιαστικά η χωρική παλινδρόμηση στηρίζεται στην χωρική αυτοσυσχέτιση, η δε επίλυση της γίνεται σε περιβάλλον GIS.

Ως βασική παραδοχή στην ανάλυση παλινδρόμησης είναι η ανεξαρτησία και η κανονική κατανομή των καταλοίπων, η οποία όμως δεν ισχύει στην περίπτωση που προσεγγίζουμε το πρόβλημα με την χωρική ανάλυση. Ως λύση θα μπορούσε να εξετασθεί ο υπολογισμός του μοντέλου με την προσθήκη ανεξάρτητων μεταβλητών οι οποίες θα είναι ουσιαστικά οι τιμές της εξαρτημένης μεταβλητής γειτονικών περιοχών. Σε άλλη προσέγγιση θα μπορούσε να αντιμετωπιστεί το πρόβλημα με τον υπολογισμό πολλών τοπικών μοντέλων παλινδρόμησης, με τις παραμέτρους αυτού να διαφοροποιούνται ανάλογα με την θέση στο γεωγραφικό χώρο. Σ’ αυτή την περίπτωση τα σφάλματα επικεντρώνονται στην συγκεκριμένη θέση του κάθε συγκεκριμένου μοντέλου και αφορούν μόνο την θέση αυτή. Αντί λοιπόν να υπολογιστεί ένα γενικό μοντέλο, υπολογίζονται πολλά τοπικά. Αυτή τη προσέγγιση του προβλήματος την πρωτοεξέτασαν και ανέπτυξαν οι Brunsdon, Fotheringham και Charlton και ονομάζεται Γεωγραφικά Σταθμισμένη Παλινδρόμηση (Brunsdon, Fotheringham & Charlton, 1998). Η παλινδρόμηση αυτή μοντελοποιεί χωρικά διαφοροποιημένες σχέσεις, με τη παραδοχή χωρικής ανομοιομορφίας στην περιοχή μελέτης, φτιάχνοντας τοπικά μοντέλα για την εξαρτημένη μεταβλητή, υπολογίζοντας μία εξίσωση παλινδρόμησης για κάθε οντότητα βάση των γειτονικών της οντοτήτων. Κάθε εξίσωση υπολογίζει την ίδια εξαρτημένη μεταβλητή και περιέχει την ίδια ή τις ίδιες ανεξάρτητες μεταβλητές, διαφέροντας όμως ως προς τους συντελεστές των ανεξάρτητων μεταβλητών και το σταθερό όρο. Η διαφοροποίηση των συντελεστών αποδίδεται οπτικά σε κατάλληλα χαρτογραφικά υπόβαθρα, που βοηθούν στην ερμηνεία των αποτελεσμάτων και την εξαγωγή συμπερασμάτων (Ηλιοπούλου, 2015). Εδώ θα πρέπει να σημειωθεί ότι σε αντίθεση με την κλασική παλινδρόμηση, όπου μέσω της μεθόδου των ελαχίστων τετραγώνων θεωρείται ότι τα κατάλοιπα είναι

ανεξάρτητα μεταξύ τους, στην γεωγραφικά σταθμισμένη παλινδρόμηση αυτό δεν ισχύει, λόγω της ύπαρξης αυτοσυσχέτισης (Ηλιοπούλου, 2015).

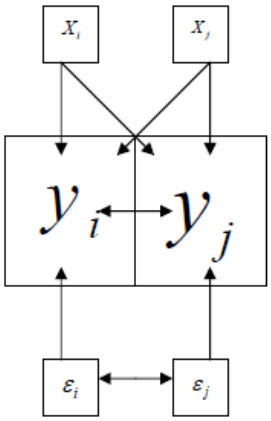
Η αξιολόγηση του μοντέλου της γεωγραφικά σταθμισμένης παλινδρόμησης γίνεται με το $\mathbf{R^2}$ και με πρόσθετα κριτήρια όπως είναι το Akaike Information Criterion (AIC), το οποίο χρησιμοποιεί τα τετράγωνα των καταλοίπων της παλινδρόμησης (Ηλιοπούλου, 2015), ενώ ο έλεγχος της κανονικότητας μπορεί να γίνει με την χρήση ιστογράμματος.

Η γεωγραφικά σταθμισμένη παλινδρόμηση διεξάγεται μέσω λογισμικών σε περιβάλλον GIS, όπως το ArcGIS και το GWR4. Επιπρόσθετα σε εξειδικευμένα λογισμικά πακέτα όπως το GEODA (ακαδημαϊκό λογισμικό) υπολογίζονται διαφορετικά μοντέλα όπως το spatial lag και το spatial error.

Καθώς ως γενικός σκοπός της πολλαπλής γραμμικής παλινδρόμησης είναι η εύρεση της γραμμικής σχέσης μεταξύ μιας εξαρτημένης μεταβλητής και ενός συνόλου ερμηνευτικών- ανεξάρτητων μεταβλητών, η μέθοδος των ελαχίστων τετραγώνων (OLS) καθιστά δυνατό τον υπολογισμό των συντελεστών b της εξίσωσης παλινδρόμησης, ελαχιστοποιώντας το άθροισμα των τετραγώνων των σφαλμάτων της πρόβλεψης. Βασικές παραδοχές για την λειτουργία αυτού του μοντέλου είναι ότι τα τυχαία σφάλματα έχουν μέση τιμή ίση με μηδέν, έχουν ομοσκεδαστικότητα και κανονική κατανομή. Αυτές οι υποθέσεις μπορεί να μην ικανοποιούνται πάντα στην πράξη καθώς πολλές φορές, αφού σύμφωνα με τον νόμο του Tobler όλα τα πράγματα συσχετίζονται χωρικά, με τα πιο κοντινά να συσχετίζονται εντονότερα, μια παρατηρούμενη τιμή σε ένα σημείο εξαρτάται από τις τιμές που παρατηρούνται σε γειτονικές θέσεις. Υπάρχει δηλαδή χωρική εξάρτηση. Αυτό έχει σαν αποτέλεσμα τα χωρικά δεδομένα να παρουσιάζουν εξάρτηση από τις μεταβλητές και τα σφάλματα τους καθώς κατά την συλλογή δεδομένων από τις παρατηρήσεις που σχετίζονται με χωρικές μονάδες αντανακλώνται και τα μετρητικά σφάλματα . Επιπρόσθετα πρέπει να ληφθεί υπόψη και η σημαντικότητα που μπορεί να έχει ένα χαρακτηριστικό στο φαινόμενο που εξετάζουμε .

Το λογισμικό GeoDa περιλαμβάνει μεταξύ άλλων και τους δύο παρακάτω βασικούς τύπους χωρικής εξάρτησης:

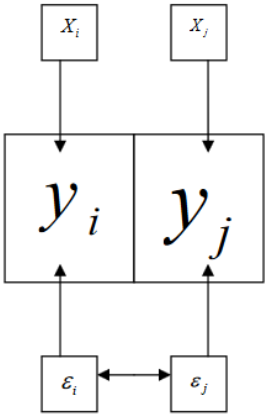
α. Spatial lag Model- Μοντέλο Χωρικής Υστέρησης (σχήμα 3.3). Η γενική του αρχή είναι ότι εξαρτημένη μεταβλητή y στη θέση i επηρεάζεται από τις ανεξάρτητες μεταβλητές, τόσο της θέσης i όσο και της θέσης j.



Σχήμα 3.3 Μοντέλο Χωρικής υστέρησης. Πηγή: www.brown.edu

Η Χωρική υστέρηση είναι ενδεικτικό μιας πιθανής διαδικασίας διάχυσης γεγονότων πρόβλεψης σε ένα μέρος με αυξημένη πιθανότητα παρόμοιων εκδηλώσεων σε γειτονικές θέσεις. Ουσιαστικά το μοντέλο αυτό εξηγεί την γραμμική επίδραση γειτονικών παρατηρήσεων, εισάγοντας μια μεταβλητή που αφορά την χωρική υστέρηση στο μοντέλο. Η μεταβλητή αυτή υπολογίζεται ως μέσος της εξαρτημένης για τα γειτονικά πολύγωνα.

β. Η spatial error (σχήμα 3.4) θεωρεί ότι στο σύστημα υπάρχει μια μεταβλητή με χωρικό χαρακτήρα. Αυτή η μεταβλητή αποτελεί τμήμα του σφάλματος της παλινδρόμησης το οποίο περιγράφεται καθώς οι διαφορετικές χωρικές ενότητες συσχετίζονται.



Σχήμα 3.4 Μοντέλο Spatial error. Πηγή: www.brown.edu

Και στα δύο μοντέλα υπολογίζονται η σταθερά της εξίσωσης, οι συντελεστές παλινδρόμησης και οι εκτιμώμενες τιμές για την εξαρτημένη μεταβλητή και τα

κατάλοιπα. Το GeoDa παρέχει μια σειρά από διαγνωστικά τεστ για την ανίχνευση της χωρικής εξάρτησης μέσω της μεθόδου της μέγιστης πιθανοφάνειας (Ηλιοπούλου, 2015· Spatial regression with Geodata).

3.4.4 Πολυωνυμική Παλινδρόμηση

Η πολυωνυμική παλινδρόμηση είναι μια μορφή της πολλαπλής γραμμικής παλινδρόμησης που προκαλεί γενίκευση του πολυωνύμου δευτέρου βαθμού. Σύμφωνα με τον Κύρκο (2015) μια πολυωνυμική συνάρτηση μπορεί να προσεγγίσει την πραγματική παλινδρόμηση εάν από το διάγραμμα διασποράς προκύπτει ότι είναι καμπυλόγραμμο. Η σχέση που την εκφράζει είναι :

$$\widehat{\boldsymbol{Y}} = \mathbf{a} + \boldsymbol{b_i X} + \boldsymbol{b_2}\ \boldsymbol{X^2} + \ \boldsymbol{b_3}\ \boldsymbol{X^3} + \cdots \ldots + \boldsymbol{b_n}\ \boldsymbol{X^n}$$

όπου n είναι το πλήθος των διαφορετικών δυνάμεων της μεταβλητής X. Ουσιαστικά η σχέση περιγράφει αυξανόμενες δυνάμεις της μεταβλητής X, με διαφορετικό μερικό συντελεστή για κάθε δύναμη της. Όπως είναι φανερό, υπάρχει μόνο μια ανεξάρτητη μεταβλητή η οποία υψώνεται σε διαφορετικές δυνάμεις. Γραφικά κάθε εξίσωση περιγράφεται με καμπύλες διαφορετικών σχημάτων. Το πρότυπο αυτό με τους κατάλληλους μετασχηματισμούς μετατρέπεται σε γραμμικό και είναι δυνατή η εκτίμηση των συντελεστών του μέσα από τις μεθόδους εκτίμησης της γραμμικής παλινδρόμησης. Ο υπολογισμός της πολυωνυμικής εξίσωσης στηρίζεται στην επιλογή της βέλτιστης μεταβλητής X, που η ένταξη αυξανόμενων δυνάμεων σε αυτή θα εμφανίζει στατιστική σημαντικότητα. Ο προσδιορισμός της μέγιστης εκείνης πολυωνυμικής δύναμης, επί της επιθυμητής μεταβλητής, που θα είναι η τελευταία στατιστικά σημαντική είναι το επιθυμητό στην όλη διαδικασία αποτέλεσμα. Περιορισμός είναι ότι ο μέγιστος πολυωνυμικός βαθμός πρέπει να είναι μικρότερος από n-2 παρατηρήσεις παλινδρόμησης, ώστε να υπάρχει στατιστική αξιοπιστία εκτός εάν δικαιολογείται για λόγους που δεν έχουν σχέση με τα δεδομένα. Για την διεκπεραίωση της διαδικασίας, μπορεί να επιλεγεί η μέθοδος της προοδευτικής απόρριψης, ξεκινώντας με υψηλότερες δυνάμεις και μειώνοντας τις, έως ότου βρεθεί μια δύναμη στην φθίνουσα πορεία που να εμφανίζει στατιστικά σημαντικό πολυωνυμικό συντελεστή. Η προοδευτική απόρριψη πραγματοποιείται με τους ελέγχους έλεγχο F ή t για τον πολυωνυμικό συντελεστή. Άλλη μέθοδος είναι η εισαγωγή του μικρότερου δυνατού μοντέλου (απλή ευθεία) και η διαδικασία να εξελιχθεί εισάγοντας τους επόμενους όρους, ελέγχοντας παράλληλα την

σημαντικότητα του μοντέλου. Όταν βρεθεί ο όρος με δύναμη μη στατιστικά σημαντική, τότε αυτή η εξίσωση είναι η αποδεκτή (Πετρίδης, 2015· Montgomery et al, 1992).

3.4.5 Παλινδρόμηση κατά Εκατοστημόριο ή Τεταρτημοριακή Παλινδρόμηση (Quantile Regression)

Η παλινδρόμηση κατά εκατοστημόριο παρουσιάστηκε το 1978 από τους Koenker και Bessett ως μία λύση για την εκτίμηση των συντελεστών παλινδρόμησης του γραμμικού μοντέλου (Koenker, Bessett, 1978). Το μοντέλο αυτό σύμφωνα με τους Zietz, Zietz & Sirmans (2008 σελ. 5) βασίζεται στην ελαχιστοποίηση των σταθμισμένων απόλυτων αποκλίσεων ώστε να εκτιμηθούν υπό όρους οι εκατοστημόριακες λειτουργίες. Οι εκτιμητές της παλινδρόμησης κατά εκατοστημόριο έχουν συγκρίσιμη αποτελεσματικότητα με τον εκτιμώμενο συντελεστή παλινδρόμησης $\hat{\beta}$ όταν η κατανομή είναι κανονική, αλλά είναι καλύτεροι όταν τα κατάλοιπα δεν έχουν κανονική κατανομή. Ο εκτιμητής, σύμφωνα με την μέθοδο των ελαχίστων τετραγώνων, αναφέρεται στο μέσο του δείγματος, ενώ οι εκτιμητές της παλινδρόμησης κατά εκατοστημόριο αντιστοιχούν σε γραμμικούς συνδυασμούς στα εκατοστημόρια του δείγματος.

Σε αντίθεση με την κλασική παλινδρόμηση, η παλινδρόμηση κατά εκατοστημόριο δεν περιορίζεται στο να εξηγήσει τη μέση τιμή της εξαρτημένης μεταβλητής, αλλά στο να εξηγήσει τους καθοριστικούς παράγοντες της εξαρτημένης μεταβλητής σε οποιοδήποτε σημείο της κατανομής (Πομώνης, 2009). Θεωρητικά ο ίδιος στόχος θα μπορούσε να επιτευχθεί με τον τεμαχισμό της εξαρτημένης μεταβλητής, σε υποσύνολα και στη συνέχεια με την εφαρμογή της κλασικής παλινδρόμησης στα υποσύνολα αυτά.

Τα εκατοστημόρια συμβολίζονται με το γράμμα $\boldsymbol{\tau}$ όπου $\mathbf{0>\tau>1}$ ενώ τα εκτιμώμενα εκατοστημόρια συμβολίζονται ως $\widehat{\boldsymbol{\beta}}_{\boldsymbol{\tau}}$ **.** Εάν λοιπόν $\mathbf{(x_i,y_i)}$ είναι ένα δείγμα πληθυσμού όπου i=1,2,…..,n και $\mathbf{y_i}$ είναι η εξαρτημένη μεταβλητή, ενώ $\mathbf{x_i}$ είναι τα διανύσματα των ερμηνευτικών μεταβλητών τότε το μοντέλο της παλινδρόμηση κατά εκατοστημόριο είναι της μορφής $\mathbf{y_i = u_{\tau i} + \beta_\tau x_i}$ όπου $\mathbf{Quant_\tau (y_i \mid x_{i)=} \beta_\tau X_i}$ στο

εξαρτημένο τεταρτημόριο $\mathbf{y_i}$ από το διάνυσμα της μεταβλητής $\mathbf{x_i}$, με το $\boldsymbol{\tau}$ να παίρνει τιμές από 0 έως 1 (Πομώνης, 2009).

Βασικό χαρακτηριστικό της παλινδρόμησης κατά εκατοστημόριο είναι η διερεύνηση της επίδρασης των ερευνητικών μεταβλητών στην εξαρτημένη, σε όλη την έκταση της κατανομής των παρατηρήσεων του δείγματος με τον καθορισμό διαφορετικών τιμών $\boldsymbol{\tau}$. Οι εκτιμητές της είναι πιο αξιόπιστοι από την κλασική παλινδρόμηση εάν δεν υπάρχει κανονική κατανομή διότι βασίζονται στη διάμεσο και όχι στο μέσο του δείγματος. Οι εκτιμητές της παλινδρόμησης κατά εκατοστημόριο έχουν ανεξαρτησία έναντι της κατανομής των μεταβλητών, γιατί τόσο η εκτίμηση του εκατοστημορίου όσο και των παραμέτρων επηρεάζονται μόνο από την γειτονική συμπεριφορά της κατανομής της εξαρτημένης πλησίον του συγκεκριμένου εκατοστημορίου (Πομώνης , 2009, σελ. 98).

3.4.6 Λογιστική Παλινδρόμηση

Η **λογιστική παλινδρόμηση (logistic regression)** καλύπτει την πραγματοποίηση ανάλυσης δεδομένων που αφορούν την μελέτη και πρόβλεψη τιμών μιας κατηγορικής μεταβλητής (τακτικής ή ονομαστικής), χρησιμοποιώντας τόσο ποσοτικές όσο και ποιοτικές ανεξάρτητες μεταβλητές. Αυτή είναι και η διαφορά της λογιστικής με την γραμμική παλινδρόμηση καθώς στην δεύτερη, η μεταβλητή απόκρισης είναι αποκλειστικά ποσοτική (Πετρίδης, 2015· Δημητροπουλάκης, χ.η). Οδηγεί στην ανάπτυξη μιας μη γραμμικής σχέσης που υπολογίζει την πιθανότητα ένα συμβάν να γίνει ή όχι και μπορεί να χρησιμοποιηθεί σε περίπτωση που η μεταβλητή είναι δίτιμη (binary variable) παίρνοντας τιμές 0/1, Ναι / Όχι, Επιτυχία / Αποτυχία κλπ. ακολουθώντας συνήθως διωνυμική κατανομή. Ουσιαστικά είναι μια εναλλακτική μέθοδος παλινδρόμησης κατάλληλη για binary δεδομένα (Chatterjee, Simonoff, 2013).

Σύμφωνα με τον Πετρίδη (2015), αποτελεί ένα μοντέλο ταξινόμησης τιμών της εξαρτημένης μεταβλητής με βάση την θεωρία των πιθανοτήτων, καθώς αυτή εξαρτάται από ένα πλήθος προβλεπτικών μεταβλητών. Κατά την λογιστική παλινδρόμηση, η εκτίμηση των παραμέτρων γίνεται με την μέθοδο του λόγου της πιθανοφάνειας, όπου επιλέγονται οι ποιο πιθανοφανείς τιμές παραμέτρων που θα οδηγήσουν στα παρατηρούμενα αποτελέσματα. Οι τιμές των συντελεστών των ανεξάρτητων μεταβλητών είναι αυτές που κάνουν τις παραχωρηθείσες τιμές της

εξαρτημένης πιο πιθανές. Στο λογιστικό μοντέλο τα σφάλματα δεν υπακούουν σε κανονική κατανομή και η εξαρτημένη μεταβλητή είναι διακριτή. Στα κατάλοιπα των αποκρίσεων αναπτύσσεται πάντα ετεροσκεδαστικότητα, σε κάθε προβλεπόμενη τιμή, λόγω του μεταβαλλόμενου ποσοστού διακύμανσης που αναλογεί. Τέλος σημαντικό να αναφερθεί είναι ότι η αξιοπιστία της επηρεάζεται από το δειγματοληπτικό μέγεθος.

Ουσιαστικά διακρίνονται τρείς τύποι λογιστικής παλινδρόμησης ανάλογα με την φύση της εξαρτημένης μεταβλητής. Όταν αυτή είναι δυαδική (binary) έχουμε δύο κατηγορίες (πχ. Ναι , Όχι). Όταν είναι τακτική (ordinal) έχουμε τρείς ή περισσότερες κατηγορίες για την εξαρτημένη μεταβλητή (πχ. λίγο, μέτρια , πολύ κλπ.). Όταν είναι ονομαστική (nominal) τότε η εξαρτημένη μεταβλητή περιέχει περισσότερες κατηγορίες χωρίς διαβάθμιση (πχ. πράσινο, κόκκινο, πορτοκαλί, κίτρινο κλπ.).

Ως μέθοδος, που κατά κύριο λόγο εφαρμόζεται για να λύσει θέματα κατηγοριοποίησης, συγκεντρώνει πλεονεκτήματα, όπως η σχετική απλότητά της και η ευρεία χρήση της, το γεγονός ότι παρέχει μοντέλα ερμηνεύσιμα και ικανοποιητικές επιδόσεις κατηγοριοποίησης, αλλά έχει και μειονέκτημα, όπως η διατύπωση αυθαίρετων υποθέσεων (πχ. ύπαρξη γραμμικής σχέσης με τον λογάριθμο του κλάσματος των πιθανοτήτων), ενώ θα πρέπει να αποφεύγονται μεταβλητές με μεγάλη συσχέτιση (Κύρκος, 2015).

Η λογιστική παλινδρόμηση εκφράζεται από τη σχέση :

$$\boldsymbol{log(odds) = a + b_1X_1 + b_2X_2 \ + b_3X_3 + \cdots ... + b_kX_k}$$

Δεξιά δημιουργείται ένας γραμμικός συνδυασμός των ανεξάρτητων μεταβλητών που συμμετέχουν στο μοντέλο και αριστερά περιέχονται οι τιμές της εξαρτημένης μεταβλητής με τη μορφή λογαρίθμου των odd όπου είναι ο λόγος των πιθανοτήτων της επιτυχίας p προς την πιθανότητα αποτυχίας 1-p, δηλαδή odds=p/(1-p) ή όπως συνηθίζεται να συμβολίζεται odds=p/1-p. Πχ εάν πούμε ότι τα odds είναι 2 τότε αυτό σημαίνει ότι οι πιθανότητες επιτυχίας είναι διπλάσιες από τις πιθανότητες αποτυχίας (Δημητροπουλάκης, χ.η, κεφάλαιο 12 σελ. 2 · Τζαβέλας, σελ 25· Chatterjee, et al, 2013, σελ 150).

Η μέθοδος των ελαχίστων τετραγώνων για την εκτίμηση των συντελεστών της εξίσωσης δεν είναι η καλύτερη επιλογή, καθώς εδώ δεν έχουμε κανονική κατανομή. Για τον λόγο αυτό χρησιμοποιούμε την μέθοδο εκτίμησης της μέγιστης πιθανοφάνειας, η οποία στηρίζεται στο γεγονός ότι οι παράμετροι υπολογίζονται με

τιμές που δίνουν στα παρατηρούμενα δεδομένα την μέγιστη πιθανότητα να συμβούν. Η συνάρτηση της πιθανοφάνειας έκβασης ενός γεγονότος δείχνει πόσο κατάλληλα περιγράφεται ένα παρατηρούμενο δείγμα από συγκεκριμένες τιμές παραμέτρων και η μεγιστοποίησή της καθορίζει τις πλέον ικανές (Chatterjee, et al, 2013, σελ 156,157 · Πετρίδης, 2015, σελ 92).

3.4.7 Μετασχηματισμοί δεδομένων στην Παλινδρόμηση ως Εναλλακτικές Μορφές του Γραμμικού Μοντέλου, Εγγενείς Γραμμικές Συναρτήσεις Παλινδρόμησης

Εγγενής γραμμική συνάρτηση (intrinsically linear) καλείται η συνάρτηση αυτή που παρόλο που δεν είναι γραμμική με κατάλληλο μετασχηματισμό μπορεί να γίνει. Το βασικό πλεονέκτημα μιας εγγενούς γραμμικής συνάρτησης είναι ότι μπορούμε να εκτιμήσουμε τις παραμέτρους της με την μέθοδο των ελαχίστων τετραγώνων.

Οι μετασχηματισμοί των δεδομένων είναι ένα εργαλείο που μπορεί να χρησιμοποιηθεί με σκοπό την βελτίωση της κανονικότητας της κατανομής, την εξίσωση της διακύμανσης και για την βελτίωση των αποτελεσμάτων στην στατιστική ανάλυση. Οι βασικοί λόγοι για την εφαρμογή μετασχηματισμών στις μεταβλητές μιας εξίσωσης παλινδρόμησης είναι η σταθεροποίηση των διασπορών όταν παραβιάζεται η ομοσκεδαστικότητα, για την κανονικοποίηση της εξαρτημένης μεταβλητής και για την γραμμικοποίηση, όταν τα δεδομένα υποδεικνύουν μη γραμμικό μοντέλο. Καθώς οι στατιστικές προσεγγίσεις κάνουν συνήθως δύο παραδοχές: ότι υπάρχει κανονική κατανομή σφάλματος των μεταβλητών και ότι υπάρχει ομοσκεδαστικότητα, σημαντική παραβίαση αυτών των δύο παραδοχών μπορεί να αυξήσει τις πιθανότητές για σφάλμα τύπου I ή II ανάλογα με τη φύση της ανάλυσης και της παραβίασης της υπόθεσης, η εφαρμογή του κατάλληλου μετασχηματισμού μπορεί να διορθώσει ή απλά να βελτιώσει το αποτέλεσμα της ανάλυσης. Η λογαριθμική μετατροπή των μεταβλητών σ’ ένα μοντέλο παλινδρόμησης είναι πλέον ένας συνηθισμένος τρόπος χειρισμού μιας μη γραμμικής σχέσης ώστε να λειτουργήσει ως γραμμική. Υπάρχουν πληθώρα τύπων μετασχηματισμών δεδομένων. Μερικοί από τους πιο συχνά χρησιμοποιημένους παρουσιάζονται παρακάτω (Osborne, 2010· Benoit, 2010 · Παπαδόπουλος, χη) .

3.4.7.1 Linear- Log Παλινδρόμηση ως Εναλλακτική Μορφή του Γραμμικού Μοντέλου

Εάν χρησιμοποιήσουμε το φυσικό λογάριθμο για την ανεξάρτητη μεταβλητή Χ και κρατήσουμε την εξαρτημένη μεταβλητή Υ στην αρχική κλίμακα, τότε το μοντέλο αυτό ονομάζεται linear-log. Χρησιμοποιείται συνήθως όταν η επίδραση της ανεξάρτητης μεταβλητής στην εξαρτημένη μειώνεται, καθώς η τιμή της ανεξάρτητης αυξάνεται. Το αρχικό μοντέλο δεν είναι γραμμικό αλλά ο λογαριθμικός μετασχηματισμός παράγει τη επιθυμητή γραμμικότητα.

Εάν υποθέσουμε ότι η απλή γραμμική μορφή του μοντέλου είναι η

$\mathbf{Y_i = \alpha+\beta X_i+\varepsilon_i}$ τότε η μορφή του linear- log μοντέλου είναι :

$\mathbf{Y_i= \alpha+ \beta \log X_{i+}\varepsilon_i}$

Στο μοντέλο αυτό η ερμηνεία του εκτιμώμενου συντελεστή $\widehat{\boldsymbol{\beta}}$ είναι ότι η αύξηση κατά μία μονάδα στην $\mathbf{logX}$ παράγει μια αύξηση στην $\mathbf{Y}$ κατά $\widehat{\boldsymbol{\beta}}$ μονάδες.

Σύμφωνα με τον Benoit (2011) εάν θεωρήσουμε ότι

$\mathbf{logX+1= logX+loge=log(eX),}$ δηλαδή το να προσθέσουμε $\mathbf{1}$ στο $\mathbf{logX}$ σημαίνει ότι πολλαπλασιάζουμε το $\mathbf{X}$ με το $\mathbf{e}$ όπου $\mathbf{e =2,72}$, ή αλλιώς $\widehat{\boldsymbol{\beta}}$ είναι η αναμενόμενη αύξηση του $\mathbf{Y}$ όταν το $\mathbf{X}$ αυξάνεται κατά $\mathbf{172\%.}$

Η αναμενόμενη αύξηση του Y σε σχέση με την p% αύξηση του X υπολογίζεται ως

$\widehat{\boldsymbol{\beta}} \cdot \boldsymbol{log}(\frac{[\mathbf{100+p}]}{\mathbf{100}})$ τότε όταν πολλαπλασιάσουμε τα X κατά 1,1 η αναμενόμενη αλλαγή στα Y είναι 0,095$\widehat{\boldsymbol{\beta}}$. Για p=1 τότε για 1% αύξηση του X, το Y αυξάνεται κατά $\hat{\beta}/100$.

3.4.7.2 Log-Linear Παλινδρόμηση ως Εναλλακτική Μορφή του Γραμμικού Μοντέλου

Όπως προαναφέρθηκε μερικές φορές η γραμμική παλινδρόμηση μπορεί να χρησιμοποιηθεί με σχέσεις που δεν είναι εγγενώς γραμμικές, αλλά μπορεί να γίνουν γραμμικές μετά από ένα μετασχηματισμό.

Ειδικότερα, θεωρούμε το ακόλουθο εκθετικό μοντέλο: $\boldsymbol{Y_i = e^{\alpha+\beta X_i+\varepsilon_i}}$

Λαμβάνοντας το φυσικό λογάριθμο έχουμε την εξής ισοδύναμη εξίσωση:

$\mathbf{logY_i=\alpha+\beta X_i+\varepsilon_i}$

Ένα μοντέλο αυτού του τύπου αναφέρεται ως μοντέλο **log- linear** παλινδρόμησης.

Στο μοντέλο αυτό η ερμηνεία του εκτιμώμενου συντελεστή $\widehat{\boldsymbol{\beta}}$ είναι ότι η αύξηση κατά ένα μονάδα στην **X** παράγει μια αύξηση στην **Y** κατά $\boldsymbol{e^{\widehat{\beta}}}$ μονάδες, δηλαδή η αναμενόμενη τιμή του **Y** πολλαπλασιάζεται με $\boldsymbol{e^{\widehat{\beta}}}$ μονάδες. Για μικρές τιμές του $\widehat{\boldsymbol{\beta}}$ περίπου $\boldsymbol{e^{\widehat{\beta}} = 1 + \widehat{\beta}}$ τότε ισχύει ότι: για κάθε αλλαγή του Χ κατά μία μονάδα η αναμενόμενη ποσοστιαία αλλαγή του Υ είναι : $100 \cdot \widehat{\boldsymbol{\beta}}$
(Benoit, 2011· Πανάρετος,1994· Pek, Wong, 2017).

3.4.7.3 Log Log Παλινδρόμηση ως Εναλλακτική Μορφή του Γραμμικού Μοντέλου

Όταν τόσο η εξαρτημένη μεταβλητή όσο και η ανεξάρτητη είναι μετασχηματισμένες λογαριθμικά αντιμετωπίζουμε έναν συνδυασμό linear- log και log- linear μοντέλων. Με άλλα λόγια, η ερμηνεία δίνεται ως μια αναμενόμενη μεταβολή σε ποσοστό επί της εκατό στο Υ, όταν η Χ αυξάνεται κατά ένα ορισμένο ποσοστό. Το μοντέλο αυτό ονομάζεται **log- log** μοντέλο και έχει την μορφή $\boldsymbol{\log Y_i = \alpha + \beta \log X_i + \varepsilon_i}$

Τέτοιες σχέσεις, αναφέρονται στην οικονομετρία ως ελαστικές και ο συντελεστής $\boldsymbol{\log X}$ αναφέρεται ως ελαστικότητα. Πολλαπλασιάζοντας την **X** με **e** θα πολλαπλασιαστεί η αναμενόμενη τιμή του **Y** κατά $\boldsymbol{e^{\widehat{\beta}}}$ (Benoit, 2011).

3.4.7.4 Αντίστροφοι Μετασχηματισμοί και Μετασχηματισμοί Τετραγωνικής Ρίζας και Τετραγώνου

Καθώς διερευνούμε την σχέση των μεταβλητών, πολλές φορές διαπιστώνεται παραβίαση των προϋποθέσεων εφαρμογής του γραμμικού μοντέλου. Η αντιμετώπιση του προβλήματος είναι αναγκαία για την σταθεροποίηση των διασπορών (παραβίαση ομοσκεδαστικότητας), για την κανονικοποίηση όταν οι κατανομές της Υ δεν είναι κανονικές και για την γραμμικοποίηση όταν τα δεδομένα υποδεικνύουν μη γραμμικό μοντέλο. Έτσι ενδείκνυται η εφαρμογή κατάλληλων μετασχηματισμών ώστε το μοντέλο μας να λάβει γραμμική μορφή. Κυριότεροι μετασχηματισμοί είναι οι λογαριθμικοί με τους οποίους ασχοληθήκαμε πιο πάνω, ενώ οι πιο απλές μορφές είναι οι αντίστροφοι και οι μετασχηματισμοί τετραγωνικής ρίζας.

Καθώς η γραμμική σχέση μεταξύ των δύο μεταβλητών εκφράζεται από την γενική εξίσωση: $\boldsymbol{Y_i = \alpha + \beta X_i + \varepsilon_i}$ ο αντίστροφος μετασχηματισμός $\boldsymbol{1/\ Y = Y'}$

χρησιμοποιείται για σταθεροποίηση της διασποράς του $\mathbf{Y}$ όταν έχουμε μεγάλη αύξηση της διασποράς πάνω από κάποια τιμή του $\mathbf{Y}$ και για γραμμικοποίηση του μοντέλου όταν τα δεδομένα δείχνουν το αντίστροφο μοντέλο.

Επιπλέον ο μετασχηματισμός $\sqrt{\boldsymbol{Y}} = \mathbf{Y}'$ χρησιμοποιείται για σταθεροποίηση της διασποράς του $\mathbf{Y},$ όταν η διασπορά είναι ανάλογη της μέσης τιμής του Y και για τη γραμμικοποίηση του μοντέλου.

Ο μετασχηματισμός $\mathbf{Y^2 = Y'}$ χρησιμοποιείται για σταθεροποίηση της διασποράς του Y όταν ελαττώνεται η μέση τιμή της, για την κανονικοποίηση της Y όταν η κατανομή των καταλοίπων παρουσιάζει αρνητική συμμετρία και για την γραμμικοποίηση του μοντέλου όταν τα αρχικά δεδομένα υποδεικνύουν καμπυλόγραμμο μοντέλο (Παπαδόπουλος, χ.η).

3.4.7.5 Λοιπές Εγγενείς Γραμμικές Συναρτήσεις Παλινδρόμησης

Η εκθετική συνάρτηση $\boldsymbol{y = ae^{\beta x}}$ με τον κατάλληλο μετασχηματισμό γίνεται

$$\mathbf{ln(y)=ln(\alpha)+\beta x.}$$

Η συνάρτηση δύναμης $\boldsymbol{y = ax^{\beta}}$ με τον κατάλληλο μετασχηματισμό γίνεται

$$\mathbf{log(y)=log(\alpha)+\beta\ log(x)}$$

Αντεστραμμένο πολυωνυμικό μοντέλο $\mathbf{y=x/\alpha+\beta x+\varepsilon}$ (Κουγιουμτζής, 2011· Rawling et al, 1998).

3.4.7.6 Μοντέλο Μετασχηματισμού Box-Cox

Το μοντέλο μετασχηματισμού Box-Cox πήρε το όνομά του από την εργασία «*An analysis of transformation»* των Box και Cox το 1964. Αν και δεν είναι τόσο ευρέως γνωστή ως τεχνική λόγω της δυσκολίας ερμηνείας των συντελεστών, μια πιο γενική της μορφή εφαρμόζεται στα ηδονικά μοντέλα.

Ο κύριος στόχος κατά την ανάλυση του υποδείγματος μετασχηματισμού Box-Cox είναι να βγάλει κανείς συμπέρασμα σχετικά με την παράμετρο μετασχηματισμού λ μέσω της μεθόδου της μέγιστης πιθανοφάνειας. Σε γενικές γραμμές ο μετασχηματισμός αυτός ανήκει στην οικογένεια των εκθετικών μετασχηματισμών, ενώ καθίσταται δυνατός ο προσδιορισμός του πλέον κατάλληλου μετασχηματισμού

της μεταβλητής απόκρισης, με σκοπό την εξομάλυνση των αποκλίσεων, όταν διαπιστώνεται έλλειψη ομοσκεδαστικότητας, μη κανονικότητας σφαλμάτων, ή μη γραμμικότητας της μέσης απόκρισης. Απώτερος σκοπός είναι η μετατροπή σε μια εξίσωση γραμμικής παλινδρόμησης που να ικανοποιεί τις απαιτήσεις των εξισώσεων ελαχίστων τετραγώνων (Kutner, Nachtsheim, Neter, 2005· Li, 2005· Rawling et al, 1998).

Οι δύο βασικοί μετασχηματισμοί που παρουσίασαν οι Box Cox (1964) είναι: για **y>0**

$\mathbf{y^{(\lambda)}=(y^{\lambda}-1)/\lambda}$, όπου το λ είναι η παράμετρος μετασχηματισμού για $\lambda \neq 0$ &

$\mathbf{y^{(\lambda)}=logy}$, εάν το λ=0.

3.4.8 Δένδρα Αποφάσεων, Δενδρική Παλινδρόμηση και Ταξινόμηση, Τεχνική CART

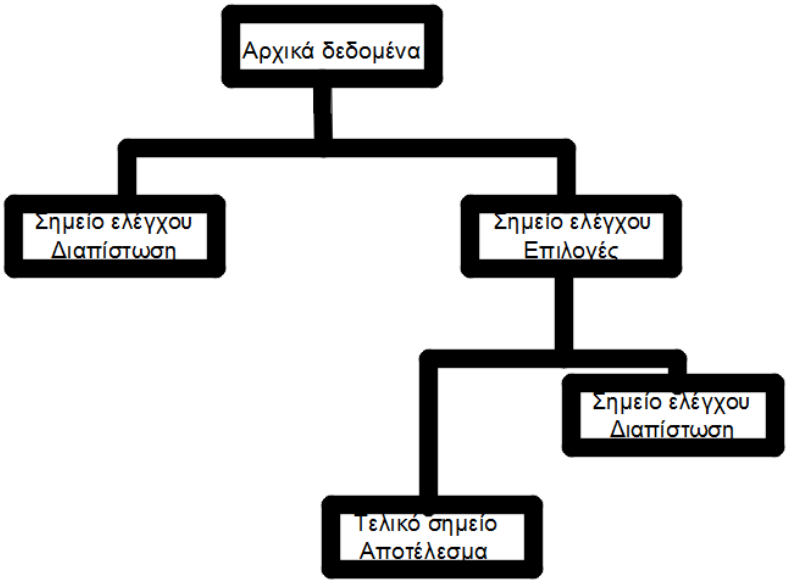
Τα **δένδρα αποφάσεων** αποτελούν ουσιαστικά μέθοδο κατηγοριοποίησης βασιζόμενα στη λογική της διαδοχικής διάσπασης του συνόλου των παρατηρήσεων σε υποσύνολα, με κριτήριο τις τιμές των μεταβλητών. Η διαδικασία αυτή αναπαρίσταται μέσω μιας αντεστραμμένης δενδρικής δομής (σχήμα 3.5). Τα δένδρα αυτά χρησιμοποιούνται εκτός από την κατηγοριοποίηση μεταβλητών και για παλινδρόμηση, προβλέποντας και ονομαστικές τιμές κλάσης και τιμές αριθμητικών πεδίων μέσω λογικών προσεγγίσεων τύπου if- then.

Σε κάθε σημείο διακλάδωσης του δέντρου δημιουργείται ένας κόμβος, ο οποίος προσδιορίζει μια συνθήκη θέλοντας να επιτύχει την βέλτιστη πρόβλεψη. Η τεχνική αυτή επιτυγχάνει απλούστευση των αποτελεσμάτων και ανάπτυξη λογικών διατυπώσεων με μη παραμετρικές καταβολές και μη γραμμική κατεύθυνση, ώστε να μην απαιτούνται έλεγχοι των παραμετρικών υποθέσεων.

Η πιο διαδεδομένη τεχνική δενδρικής προσέγγισης είναι μέσω του αλγορίθμου CART (από το Classification and Regression Trees) η οποία παρουσιάστηκε το 1984 από τους Breiman, Friedman,Olshen & Stone και χρησιμεύει για να προβλέψει ποσοτική εξαρτημένη μεταβλητή ή να ταξινομήσει κατηγορική μεταβλητή. Ξεκινώντας από τον πρωτεύοντα αρχικό κόμβο, η διαδικασία προχωρά δημιουργώντας διαδοχικούς κόμβους ώστε να καταλήξει στον τελικό, όπου δεν θα υπάρχει πλέον δυνατότητα άλλης διακλάδωσης. Οι μη τερματικοί κόμβοι περιέχουν

συνθήκες ελέγχου χαρακτηριστικών, ενώ ο τελικός είναι αυτός, η διάσπαση του οποίου δεν προσθέτει επιπλέον αξία στη πρόβλεψη.

Σε γενικές γραμμές η μέθοδοι της δενδρικής προσέγγισης σε αντίθεση με άλλες μεθόδους δεν κάνουν αυθαίρετες υποθέσεις, είναι μη παραμετρικές, δέχονται ως μεταβλητές εισόδου αριθμητικές τιμές και ονομαστικά γνωρίσματα, αλλά είναι ιδιαίτερα ευπαθείς σε μεταβολές του δείγματος. Πλεονέκτημα είναι η ταχύτητα, η ακρίβεια καθώς και η ικανότητα χειρισμού δεδομένων με χαμηλές τιμές, κατηγοριοποιώντας το σύνολο των δεδομένων, λαμβάνοντας υπόψη το κόστος του σφάλματος. (Κύρκος, 2015· Πετρίδης, 2015).



Σχήμα 3.5. Δενδρική Προσέγγιση.. Ιδία επεξεργασία.

Ας δούμε ένα μικρό και απλό παράδειγμα. Ας υποθέσουμε ότι για τον προσδιορισμό του βέλτιστου μοντέλου πρόβλεψης τιμών αξιών ακινήτων έχουμε στην διάθεσή μας στοιχεία από τρείς ανεξάρτητες μεταβλητές. Το ζητούμενο είναι να καθορίσουμε το ποιες από αυτές τις μεταβλητές θα λάβουν μέρος στο μοντέλο παλινδρόμησης. Η πρώτη είναι ο αριθμός ημερών προς διάθεση του ακινήτου στην αγορά, η δεύτερη είναι το εμβαδόν του ακινήτου και η τρίτη είναι η ηλικία του ακινήτου προς πώληση. Στο πρώτο βήμα όλα τα δεδομένα θα εισαχθούν στο σύστημα. Για να ξεκινήσει η επεξεργασία θα πρέπει να τεθούν οι κανόνες επεξεργασίας στις μεταβλητές. Έτσι ας θέσουμε τον πρώτο κανόνα που είναι ότι στην επεξεργασία του μοντέλου θα λάβουμε υπόψη τις μεταβλητές αυτές που πληρούν την προϋπόθεση ότι τουλάχιστον το 50% του δείγματος έχει τιμές που η απόκλιση τους

από το μέσο όρο τιμών είναι κατώτερη του 10% . Υποθετικά λοιπόν εκκινώντας την διαδικασία επεξεργασίας διαπιστώνουμε ότι η μεταβλητή «αριθμός ημερών προς διάθεση του ακινήτου στην αγορά» δεν πληροί τον κανόνα που έχουμε θέσει, ενώ οι άλλες δύο μεταβλητές τον πληρούν. Τότε στο πρώτο σημείο ελέγχου γίνεται αυτή η διαπίστωση, θέτοντας εκτός διαδικασίας την πρώτη μεταβλητή, ενώ οι άλλες δύο που έχουν μείνει συνεχίζουν κανονικά. Ως δεύτερο κανόνα στο σημείο αυτό, έχουμε ορίσει ότι οι μεταβλητές που θα απομείνουν πρέπει να πληρούν την προϋπόθεση ότι η διαφορά της ελάχιστης με την μέγιστη τιμή του δείγματος δεν πρέπει να αποκλίνει περισσότερο του 5% από την τυπική απόκλιση. Από τον έλεγχο που γίνεται βλέπουμε ότι μόνο η μεταβλητή εμβαδόν πληροί αυτή την προδιαγραφή. Έτσι στο σημείο ελέγχου γίνεται η διαπίστωση αυτή, τίθεται εκτός διαδικασίας η μεταβλητή «ηλικία» και ολοκληρώνεται με την μεταβλητή «εμβαδόν» να συμμετέχει στο μοντέλο παλινδρόμησης που έχουμε ορίσει αρχικά. Σε περίπτωση που ούτε και αυτή η μεταβλητή πληροί τις προϋποθέσεις που έχουμε θέσει, τότε αυτό σημαίνει ότι υπάρχει πρόβλημα στο δείγμα μας και θα πρέπει να εκκινήσουμε από την αρχή την διαδικασία συλλογής του.

3.5 Η Εφαρμογή των Μοντέλων Παλινδρόμησης για την Αξιολόγηση Μεταβλητών Επηρεασμού Τιμών και Προβλεπτικών Μοντέλων Εκτίμησης Αξιών Ακινήτων

Η παλινδρόμηση ως μια στατιστική τεχνική έρευνας και μοντελοποίησης της σχέσης μεταξύ μεταβλητών, βρίσκει εφαρμογές σε κάθε επιστημονικό κλάδο.
Στην διεθνή βιβλιογραφία υπάρχει πληθώρα μελετών που συνδυάζει την παλινδρόμηση με τα ακίνητα, εξετάζοντας την σχέση των μεταβλητών και δημιουργώντας μοντέλα .

Στην περίπτωση της παρούσας εργασίας ερευνάται η εφαρμογή της στην έρευνα ακινήτων και ειδικότερα στην επίδραση διαφόρων χαρακτηριστικών κατοικιών καθώς και στη δημιουργία κατάλληλων προβλεπτικών μοντέλων εκτίμησης της αξίας τους. Με την βοήθεια στατιστικού δείγματος αναζητούνται οι μεταβλητές αυτές που θα ερμηνεύσουν με τον καλύτερο δυνατό τρόπο το πως διαμορφώνεται η αξία των ακινήτων μιας συγκεκριμένης περιοχής, ενώ παράλληλα γίνεται και ο εντοπισμός των διαφοροποιήσεων που υφίστανται οι μεταβλητές ως προς τον αριθμό, το είδος αλλά και την περιοχή μελέτης. Αυτές οι διαφοροποιήσεις αναμένεται να δείξουν και την επίδραση του χώρου στα βάρη των μεταβλητών.

Σημαντική επισήμανση είναι ότι η επιλογή των μεταβλητών δεν πρέπει να είναι τυχαία, αλλά τέτοια ώστε κάθε μεταβλητή που θα επιδράσει στο μοντέλο να έχει τέτοια χαρακτηριστικά που να την διαφοροποιούν από τις υπόλοιπες. Ως εξαρτημένη μεταβλητή θεωρείται η αγοραία αξία των ακινήτων, ενώ γίνεται η αρχική παραδοχή ότι είναι ίδια με την αξία ζήτησης, καθώς δεν είναι δυνατόν να συλλεχθούν στοιχεία για την πραγματική τελική τιμή πώλησης ενός ακινήτου.
Για το σχηματισμό των μοντέλων εκτίμησης αξιοποιείται η μέθοδος της πολλαπλής γραμμικής παλινδρόμησης με τη βοήθεια του στατιστικού εργαλείου SPSS, ενώ σε δεύτερο στάδιο πραγματοποιείται ο σχηματισμός του μοντέλου μέσω του προγράμματος ArcGIS όπου και απεικονίζεται χωρικά το αποτέλεσμα.

3.5.1 Ενδεικτική Βιβλιογραφική Ανασκόπηση Μοντέλων Παλινδρόμησης για την Εκτίμηση Αξιών Ακινήτων

Ενδεικτικά παραθέτουμε ορισμένες μελέτες δημιουργίας μοντέλων παλινδρόμησης για την εκτίμηση αξιών ακινήτων.

- Οι Ηλιοπούλου, Στρατάκης στην μελέτη τους « Spatial analysis of housing prices in the Athens Region, Greece», που αφορούσε τιμές κατοικιών στην ευρύτερη περιοχή της Αθήνας, σε δείγμα 5130 κατοικιών προερχόμενο από προσφορά κατοικιών του πρώτου τετάρτου του 2017, σε 58 δήμους στο Λεκανοπέδιο της Αττικής, προκρίνουν ένα μοντέλο που αποτελείται από το μέγεθος του ακινήτου, την απόσταση από χώρους αναψυχής, την ηλικία, την ύπαρξη θέσης στάθμευσης, την ύπαρξη θέας, τον αριθμό υπνοδωματίων και την απόσταση από το μετρό. Το μοντέλο έδωσε R^2 =0,651 και AIC = 143282 ενώ μέσα από την GWR διαδικασία αντίστοιχα 0,769 και 141700 (Ηλιοπούλου, Στρατάκης, 2018) .

- Η Μηλάκα στην διδακτορική της διατριβή με θέμα «Χωρική ανάλυση της εξέλιξης αστικών περιοχών μέσω πολυμεταβλητής επεξεργασίας εναλλακτικών σεναρίων» αναφερόμενη στο Δήμο Βόλου, προκρίνει ένα semi-log μοντέλο σε μέγεθος 2781 παρατηρήσεων, που περιλαμβάνονται η απόσταση από δημοτικό σχολείο, νηπιαγωγείο, στάση συγκοινωνίας , φροντιστήριο, πάρκο, η ύπαρξη θέας (ως απόσταση από ακτογραμμή), απόσταση από φαρμακείο, τράπεζα, εκκλησία, αριθμό κατοίκων σε κάθε οικοδομικό τετράγωνο και εμβαδόν πλησιέστερης εκκλησίας. Ο συντελεστής προσδιορισμού R^2 έδωσε αποτέλεσμα ίσο με 0,7583 (Μηλάκα, 2010).

- Ο Λαμπρόπουλος στην διδακτορική του διατριβή με θέμα «Αυτοματοποίηση μαζικών εκτιμήσεων αξιών ακίνητων για τον ελληνικό χώρο» αναφερόμενος στην περιοχή του Χαλανδρίου, σε δείγμα 861 ακινήτων, προκρίνει ένα μοντέλο που περιλαμβάνει την εγγύτητα σε νεκροταφείο, την επιφάνεια, τον όροφο, αριθμό παρακολουθημάτων, εμβαδόν Η/Χ και σοφίτας, χρόνο από πρόσφατη ανακαίνιση, με το adjusted R^2 να φτάνει σε επίπεδο 0,5951. Επίσης στην περιοχή της Βούλας σε δείγμα 295 ακινήτων, το μοντέλο περιλαμβάνει το εμβαδόν Η/Χ και σοφίτας, τον εντοπισμό του ακινήτου εντός ζώνης θάλασσας , πλησίον ΚΕΠ, εντός πολεοδομικής ενότητας, την επιφάνεια χώρων του ακινήτου, την εγγύτητα από νοσοκομείο, την ηλικία του ακινήτου, με το adjusted R^2 να φάνει σε επίπεδο 0,3878 (Λαμπρόπουλος, 2013)

- Η Στάμου στην διδακτορικής της διατριβή με θέμα « Διερεύνηση των αξιών ακινήτων στην Μητροπολιτική περιοχή της Αθήνας: εφαρμογή σύγχρονων μεθόδων χωρικής ανάλυσης» αναφερόμενη στην περιοχή της Αττικής, σε δείγμα 11343 κατοικιών, προκρίνει ένα log- linear μοντέλο όπου περιλαμβάνονται η ηλικία , το εμβαδόν, το τζάκι, ο όροφος, η θέα σε θάλασσα, το παρκινγκ το ασανσέρ, η απόσταση από μετρό και ΗΣΑΠ, με το R^2 να φτάνει τα 0,54 (Στάμου, 2016).

- Ο Sopranzetti στην εργασία του «Hedonic Regression Models» εξετάζει μοντέλα παλινδρόμησης με τρεις διαφορετικούς τρόπους αντιμετωπίζοντάς τα με την απλή μορφή γραμμικής παλινδρόμησης, με την ημι-λογαριθμική μορφή και μετά από μετασχηματισμό Box-Cox . Το δείγμα αφορούσε 7088 κατοικίες στην περιοχή του Klein School District, των προαστίων του Houston, (ημερομηνίες καταχωρήσεων μεταξύ του Ιανουαρίου του 1992 και Δεκεμβρίου του 1995). Χρησιμοποιήθηκαν ως μεταβλητές κυρίως η ηλικία, το εμβαδόν, ο αριθμός των υπνοδωματίων, ενώ τα αποτελέσματα στην OLS διαδικασίας έδωσαν R^2 =84,3%, στο ημιλογαριθμικό μοντέλο 87,5% και με τον BOXCOX μετασχηματισμό 87,5% (Sopranzetti, χ.η)

- Οι Dun Li, Xiangqian Zhang σε μελέτη τους «« Analysis of Shanghai housing price based on the hedonic price method » σε 12 περιοχές στην επαρχία Shanghai της Κίνας και σε δείγμα 1240 ακινήτων, διατύπωσαν ένα μοντέλο το οποίο περιελάμβανε την απόσταση από εμπορικό κέντρο, την απόσταση από το κέντρο της πόλης, τον αριθμό των υπνοδωματίων, την επιφάνεια του κυρίως δωματίου, τον αριθμό των μπαλκονιών και την αξία των μηνιαίων εξόδων της κατοικίας. Το R^2 έφτασε στο 0,7728 και το adjusted R^2το 0,766024 .

- Οι Hayrullahoglu, Aliefendioglu, Tanrivermis, στην μελέτη τους ««Estimation of the hedonic valuation model in housing markets: The case of Cukurambar region in Cankaya district of Ankara province» σε ένα δείγμα 327 κατοικιών εκ των οποίων χρησιμοποιήθηκε τελικά το 50% αυτών, πρόκριναν ένα μοντέλο που περιελάβανε το εμβαδόν, το επίπεδο ορόφου, τον προσανατολισμό, το εμβαδόν του συνολικού κτιρίου, την απόσταση από το μετρό και την απόσταση από κεντρικό δρόμο με το adjusted R^2 να φτάνει στο 0,78 .
- Ο Mudit Mankad στη μελέτη του « Analysis of Impact of Accessibility on Residential Property Values in Gotri Area of Vadodara City, India Using OLS and GWR » σε ένα δείγμα 161 κατοικιών προκρίνει ένα μοντέλο που περιλαμβάνει την ηλικία κατασκευής, το εμβαδόν, τον αριθμό δωματίων, την απόσταση από κεντρικό δρόμο, από την αγορά, από το σχολείο, από μέσα μαζικής μεταφοράς, από χώρους αναψυχής, από την κεντρική αγορά, από αυτοκινητόδρομο, από σημείο παραγωγής θορύβου, με το R^2 να φτάνει το 0,67 στην GWR διαδικασία.
- Ο Anatoly Peresetsky στην μελέτη του « The price of Moscow apartments» σε ένα δείγμα 15476 διαμερισμάτων, προκρίνει ένα μοντέλο με σημαντικότερες ανεξάρτητες μεταβλητές την απόσταση από το μετρό, την απόσταση από το κέντρο της πόλης καθώς και το μέγεθος του διαμερίσματος και τον όροφό του με το R^2 να φτάνει το 0,781.
- Ο Παπαευθυμίου στη μελέτη του « Εκτίμηση αξίας ακινήτων. Αξιοποίηση τοπικών μοντέλων παλινδρόμησης» σε δείγμα 763 διαμερισμάτων στην περιοχή της Μεσογαίας προκρίνει ένα μοντέλο με adjusted R^2 ίσο με 0,687 και μεταβλητές την παλαιότητα, την επιφάνεια, αριθμό δωματίων, όροφο, wc, αριθμό μπάνιων, ύπαρξη διαμπερούς προσανατολισμού, παρκινγκ, υπογείου, τζακιού, σοφίτας, πισίνας, θέας, την απόσταση από μετρό, κέντρο Αθήνας. Αεροδρομίου, και των χρόνων προσέγγισης σε μετρό , προαστιακό, και κόμβων της Αττικής οδού.

Κεφάλαιο 4ο –Συλλογή Δεδομένων

Η περιοχή μελέτης που επιλέχθηκε για την παρούσα εργασία είναι αυτή της πόλης των Αθηνών. Το μέγεθος, η πολεοδομικής της διάταξη, τα πληθυσμιακά, δημογραφικά, κοινωνικά και οικονομικά χαρακτηριστικά της καθώς και η πυκνοκατοίκησή της, την καθιστούν ένα άριστο πεδίο έρευνας στην διερεύνηση μοντέλων παλινδρόμησης.

Η στατιστική ανάλυση των δεδομένων υλοποιείται με την βοήθεια των στοιχείων που συλλέγονται. Η συλλογή αυτών είναι τμήμα της προεργασίας της στατιστικής μελέτης που έπεται. Ο προσδιορισμός των μεταβλητών, η κατηγοριοποίησή τους και ο σχηματισμός του δείγματος αποτελούν τα αρχικά βήματα της εργασίας. Το όλο σκεπτικό είναι να προσδιοριστεί το μέγεθος του πληθυσμού, να προσδιοριστούν οι μεταβλητές, στοιχεία για τις οποίες πρέπει να συλλεχθούν και να κωδικοποιηθούν σωστά ώστε να διευκολυνθεί το υπόλοιπο έργο. Ιδιαίτερη σημασία σε εργασίες με γεωχωρική πληροφορία, όπως η παρούσα, έχει η καλύτερη δυνατή απεικόνιση των στοιχείων που συλλέχθηκαν σε χάρτη. Σε γενικές γραμμές μια καλή προεργασία εξοικονομεί χρόνο και ελαττώνει τον κόπο του μελετητή.

4.1 Γενικά Στοιχεία

Η Αθήνα είναι η πρωτεύουσα και η μεγαλύτερη πόλης της Ελλάδας, ενώ βρίσκεται στο κέντρο του ελλαδικού χώρου ανήκοντας στην Αττική. Το μέσο υψόμετρο είναι περίπου 90μ, ενώ καταλαμβάνει έκταση 37,38 τ χιλ. Έχει πληθυσμό 664.046 κατοίκους, σύμφωνα με την απογραφή του 2011 (Ελληνική Στατιστική Αρχή, 2018).

Η Αθήνα έχει μεσογειακό κλίμα, με εναλλαγή παρατεταμένων ζεστών και ξηρών καλοκαιριών και ήπιων, υγρών χειμώνων.

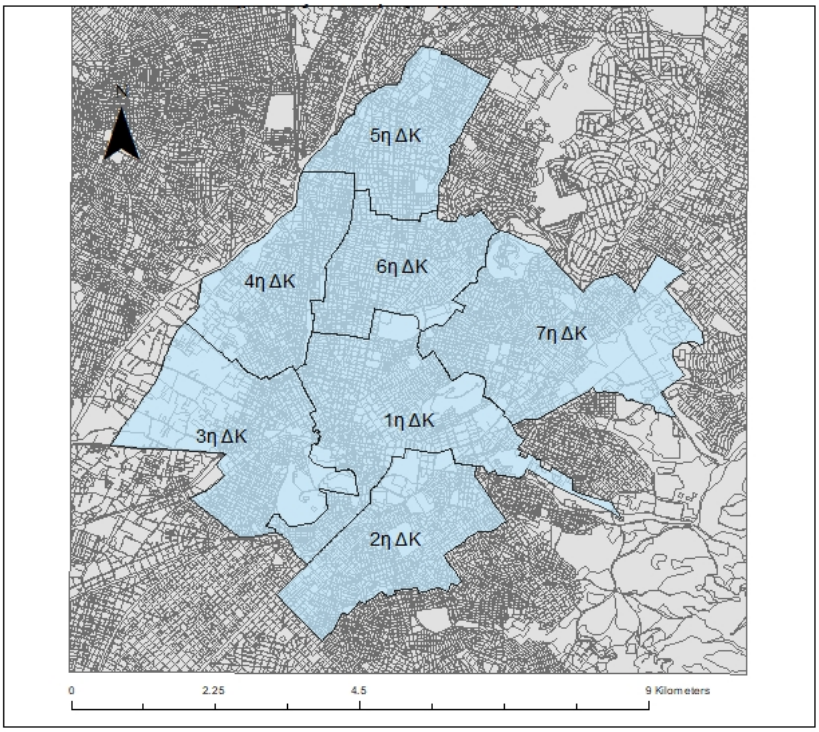
Το όνομα της πόλης προέρχεται από την προστάτιδα θεά Αθηνά. Είναι ιστορικά αποδεκτό ότι προϋπήρχε της Μυκηναϊκής Εποχής. Η Αθήνα παρέμενε μητρόπολη και στον ρωμαϊκό κόσμο, μέχρι και τον 3ο αιώνα μ.Χ, ενώ κατά τη Φραγκοκρατία αναφέρεται ως Δουκάτο των Αθηνών. Το 1458 η πόλη κατακτήθηκε από τους Τούρκους και περιήλθε στην Οθωμανική Αυτοκρατορία. Μετά την επανάσταση του 1821 έγινε και η απελευθέρωση της, ενώ έγινε πρωτεύουσα του Βασιλείου της Ελλάδας το 1833 από τον Βασιλιά Όθωνα.

Κατά τον Β΄ Παγκόσμιο Πόλεμο η πόλη κατακτήθηκε από τους Γερμανούς υπέστη μεγάλες καταστροφές και υπέφερε, ιδιαίτερα κατά τα τελευταία χρόνια του πολέμου από λιμό. Μετά τον πόλεμο η πόλη άρχισε ξανά να μεγαλώνει και κατά τη δεκαετία του '60 παρατηρήθηκε έκρηξη οικοδομικής δραστηριότητας με ανέγερση πολλών πολυκατοικιών, τόσο στο κέντρο όσο και στα προάστια της.

Το ιστορικό κέντρο των Αθηνών εντοπίζεται πέριξ του λόφου της Ακρόπολης, στο Θησείο, την Πλάκα. και το Μοναστηράκι. Το κέντρο της σύγχρονης πόλης είναι η Πλατεία Συντάγματος, όπου είναι εγκατεστημένα τα παλαιά βασιλικά ανάκτορα που σήμερα στεγάζουν το Κοινοβούλιο καθώς και άλλα δημόσια κτίρια του 19ου αιώνα (Ιστορία της πόλης, 2018).

Ο Δήμος Αθηναίων χωρίζεται σήμερα διοικητικά σε επτά δημοτικές ενότητες (πρώην δημοτικά διαμερίσματα) (Χάρτης 4.1):

- Η 1η δημοτική κοινότητα περιλαμβάνει το κέντρο των Αθηνών με το λεγόμενο εμπορικό τρίγωνο (Σύνταγμα, Κολωνάκι, Ιλίσια, Εξάρχεια, Ομόνοια, Μοναστηράκι, Πλάκα, Κουκάκι)
- Η 2η δημοτική κοινότητα περιλαμβάνει τις νοτιοανατολικές συνοικίες (Νέος Κόσμος, Στάδιο, Παγκράτι);
- Η 3η δημοτική κοινότητα περιλαμβάνει τις νοτιοδυτικές συνοικίες (Βοτανικός, Πετράλωνα, Κεραμεικός, Θησείο)
- Η 4η δημοτική κοινότητα περιλαμβάνει τις δυτικές συνοικίες (Κολωνός, Ακαδημία Πλάτωνος, Σεπόλια μέχρι Κάτω Πατήσια)
- Η 5η δημοτική κοινότητα περιλαμβάνει τις βορειοδυτικές συνοικίες (από τα Κάτω Πατήσια μέχρι Προμπονά)
- Η 6η δημοτική κοινότητα περιλαμβάνει τις βόρειες κεντρικές συνοικίες (Πλατεία Αττικής, Πλατεία Αμερικής, Κυψέλη)
- Η 7η δημοτική κοινότητα περιλαμβάνει τις βορειοανατολικές συνοικίες (Αμπελόκηποι, Γκύζη, Πολύγωνο, Ερυθρός Σταυρός, Γουδή, Ελληνορώσων) (Πρόγραμμα Καλλικράτης, 2018)



Χάρτης 4.1. Δημοτικές Κοινότητες Δήμου Αθηναίων. Ιδία επεξεργασία.

Η Αθήνα είναι μια πόλη με ανεπτυγμένες υποδομές από τη δεκαετία του '50 διαθέτοντας πλούσιο οδικό δίκτυο που τη συνδέουν με τους γειτονικούς δήμους, ενώ διαθέτει πλήρες αποχετευτικό σύστημα και ηλεκτρικό φωτισμό από το 1889.

Διαθέτει δίκτυο λεωφορείων, τρόλεϊ και μέσων σταθερής τροχιάς όπως μετρό, ηλεκτρικό σιδηρόδρομο, προαστιακό σιδηρόδρομο και τραμ, που την ενώνουν με τις γειτονιές του λεκανοπεδίου και των περιχώρων των Αθηνών. Παράλληλα τα ΚΤΕΛ που εδρεύουν στο Πεδίον του Άρεως παρέχουν πρόσβαση στις γειτονικές πόλεις και την υπόλοιπη Ελλάδα, ενώ διαθέτει και κεντρικό σιδηροδρομικό σταθμό της Ελλάδας στον Σταθμό Λαρίσης.

Το μετρό που διασχίζει την Αθήνα έχει τρεις γραμμές. Η Γραμμή 1 (πράσινη γραμμή) αφορά το παλαιότερο κομμάτι του δικτύου (ΗΣΑΠ), που συνδέει τον Πειραιά με την Κηφισιά, μέσω του κέντρου της Αθήνας. Το δίκτυο των λεωφορείων αποτελείται από οχήματα ντιζελοκίνητα και φυσικού αερίου καθώς και από ηλεκτροκίνητα τρόλεϊ. Το δίκτυο του τραμ συνδέει το κέντρο της Αθήνας (πλατεία Συντάγματος) με τη Βούλα και το Νέο Φάληρο μέσω δύο διασταυρούμενων γραμμών. Ο προαστιακός σιδηρόδρομος συνδέει το διεθνές αεροδρόμιο με τον Σταθμό Λαρίσης, τον Πειραιά και το Κιάτο. Αεροπορικώς, εξυπηρετείται από το διεθνές αεροδρόμιο Ελευθέριος

Βενιζέλος στην πόλη των Σπάτων, ανατολικά της Αθήνας, ενώ παράλληλα εξυπηρετείται και από μεγάλο αριθμό ταξί.

Η πόλη είναι προσβάσιμη οδικώς από τις δύο μεγάλες εθνικές οδούς που διατρέχουν την Αττική, τον Αυτοκινητόδρομο Ε75 που εισέρχεται από τη βόρεια είσοδο του λεκανοπεδίου και την Εθνική Οδό 8α, Ε65 & Ε94 η οποία εισέρχεται από τα δυτικά, ενώ καθίσταται προσεγγίσιμη επίσης μέσω του λιμανιού του Πειραιά, της Ραφήνας και του Λαυρίου. Από το 2004 ενώνεται με αεροδρόμιο μέσω του κλειστού αυτοκινητοδρόμου της Αττικής Οδού (Αθήνα, https://el.wikipedia.org).

4.2 Στατιστικά Στοιχεία

Ο πληθυσμός της Αθήνας, σύμφωνα με τα στοιχεία της Ελληνικής Στατιστικής Αρχής για την απογραφή του 2011, είναι 664 χιλιάδες κάτοικοι, ενώ σε αυτούς θα πρέπει να προστεθούν και άλλες 209 χιλιάδες μετανάστες. Πιο συγκεκριμένα από τους 664.046 κατοίκους που απογράφηκαν το 2011 ως κάτοικοι Αθηνών, οι 151.660 ήταν αλλοδαποί, με αποτέλεσμα εάν προσθέσουμε σε αυτούς και τους 209.308 μετανάστες, το σύνολο των μη Ελλήνων που ζουν στην Αθήνα να φτάνει τις 360.968, έναντι 512.386 με ελληνική υπηκοότητα και η αναλογία να είναι 1 αλλοδαπός προς 1.41 Έλληνες. Στην απογραφή του 2011 καταγράφηκαν 296.096 νοικοκυριά, 427.825 κανονικές κατοικίες, με μέσο όρο επιφάνειας 71.60τμ , επί 61.764 κτιρίων. Από την διαφορά των νοικοκυριών με το σύνολο των κατοικιών προκύπτει έμμεσα ότι υπάρχουν 131.729 κενές κατοικίες, δηλαδή το 30.79% του συνόλου των κατοικιών προκύπτουν κενές (Ελληνική Στατιστική Αρχή, 2018).

Στον πίνακα 4.1 αποτυπώνονται τα στατιστικά στοιχεία που έχουν σχέση με τον πληθυσμό και την ανεργία ανά δημοτικό διαμέρισμα του Δήμου Αθηναίων, ενώ ανάλογα παρατίθενται στοιχεία για τα κτίρια και τις κατοικίες στον πίνακα 4.2 .

Πίνακας 4.1

Στατιστικά στοιχεία πληθυσμού και ανεργίας ανά δημοτική κοινότητα

Δημοτικό Διαμέρισμα	Πληθυσμός	Νοικοκυριά	Με Ελληνική Υπηκοότητα	Με ξένη Υπηκοότητα	Σύνολο μεταναστών	Ποσοστό Ανεργίας %

1η	75810	36086	55055	20755	32137	20.94
2η	103004	47863	85377	17627	28786	18.50
3η	46508	19821	38905	7603	10222	20.95
4η	85629	34916	68136	17493	21792	21.27
5η	98665	40825	79570	19095	27019	19.10
6η	130582	57615	84549	46033	54326	24.48
7η	123848	58970	100794	23054	35026	17.22

Πηγή: Ελληνική Στατιστική Αρχή (2018).Στοιχεία απογραφής 2011. Ιδία επεξεργασία.

Πίνακας 4.2

Στατιστικά στοιχεία κτιρίων και κατοικιών ανά δημοτική κοινότητα

Δημοτικό Διαμέρισμα	Σύνολο κτιρίων	Κανονικές κατοικίες	Μέσος όρος κατοικιών ανά κτίριο	Μέσος όρος επιφάνειας κατοικιών	Κενές κατοικίες	Ποσοστό κενών κατοικιών επί συνόλου κατοικιών (%)
1η	10404	56208	5	76.01	20122	35.80
2η	9173	67494	7	71.90	19631	29.09
3η	7359	27753	4	68.90	7932	28.58
4η	8306	48127	6	69.23	13211	27.45
5η	6934	56315	8	73.14	15490	27.51

6η	8936	90436	10	70.96	32821	36.29
7η	10652	81492	8	71.07	22522	27.64

Πηγή: Ελληνική Στατιστική Αρχή (2018). Στοιχεία απογραφής 2011. Ιδία επεξεργασία.

4.3 Επιλογή και Μέγεθος Δείγματος στη Περιοχή Μελέτης

Εάν συμβολίσουμε το μέγεθος του πληθυσμού με το γράμμα Ν, τότε αυτό είναι το σύνολο των οριζόντιων ιδιοκτησιών- κατοικιών που βρίσκονται στις αντίστοιχες περιοχές μελέτης και είναι προς πώληση τη χρονική περίοδο της έρευνας. Πρακτικά η συγκέντρωση όλων των στοιχείων είναι αδύνατη ή έστω απαιτείται πάρα πολύ μεγάλο χρονικό διάστημα για την συλλογή τους.

Ο στόχος είναι να σχηματιστεί το κατάλληλο στατιστικό δείγμα που θα πληροί τις προϋποθέσεις μεγέθους που έχουν τεθεί, ώστε να δοθεί ισάξια λύση στα ερωτήματα που καλείται να απαντήσει η μελέτη. Ο σχηματισμός του δείγματος είναι μία σοβαρή διαδικασία, αφού ένα αξιόπιστο αποτέλεσμα εξαρτάται άμεσα από μια σωστή διαδικασία δειγματοληψίας.

Στην τυχαία δειγματοληψία σε κάθε ένα από τα δυνατά δείγματα δεδομένου μεγέθους από έναν πληθυσμό αντιστοιχεί μια γνωστή πιθανότητα επιλογής. Η μη επιλογή του ίδιου δείγματος δύο φορές μας οδηγεί στην τυχαία δειγματοληψία χωρίς επανάθεση, η οποία και εξασφαλίζει ίσες πιθανότητες σε όλα τα δείγματα του πληθυσμού να επιλεγούν. Η δειγματοληψία κατά στρώματα, η οποία αποτελεί μια μέθοδο τυχαίας δειγματοληψίας στηρίζεται στη διαδικασία της καταρχήν διαίρεσης του πληθυσμού σ' ένα συγκεκριμένο αριθμό υποπληθυσμών βασιζόμενοι σε κάποιο κοινό χαρακτηριστικό. Εν συνεχεία επιλέγεται ένα τυχαίο δείγμα από κάθε υποπληθυσμό (στρώμα). Οι υποπληθυσμοί είναι αμοιβαίως αποκλειόμενοι (Καλαματιανού, 2000). Πιο συγκεκριμένα ο πληθυσμός μεγέθους Ν διαιρείται σε k εσωτερικά ομοιογενείς υποπληθυσμούς (στρώματα) μεγέθους N1, N2, ..., Nk. Για τους υποπληθυσμούς αυτούς ισχύει ότι το άθροισμα τους είναι το μέγεθος του Ν δηλαδή N1 + N2 + ... + Nk = N. Στην συνέχεια από κάθ' ένα από τα στρώματα του πληθυσμού επιλέγεται ένα δείγμα μεγέθους ni, i = 1, 2, ..., k ώστε να προκύψει ένα δείγμα μεγέθους n = n1 + n2 + ... + nk το οποίο προκύπτει από την ένωση των k δειγμάτων. Καθώς στη στρωματοποιημένη δειγματοληψία επιλέγεται ένα δείγμα με

ανάλογο μέγεθος του στρώματος στο πληθυσμό, αυτή διαδικασία προκύπτει ότι είναι αναλογική (Ηλιοπούλου, 2015).

Στην συγκεκριμένη περιοχή μελέτης, ο αριθμός των προς πώληση κατοικιών είναι περίπου 17 χιλιάδες οικίες με την μορφή διαμερισμάτων, μονοκατοικιών, και μεζονετών, σύμφωνα πάντα με τις αγγελίες που είναι δημοσιευμένες την περίοδο μελέτης στην συγκεκριμένη ιστοσελίδα (www.spiti24.gr) από την οποία αντλήθηκαν οι πληροφορίες.

Για τη συγκέντρωση του δείγματος στην πόλης της Αθήνας στην παρούσα εργασία έγινε στρωματοποιημένη δειγματοληψία βάση ενός εμπειρικού σχήματος που προσομοιάζει με τη χωρική συστηματική δειγματοληψία.

Σύμφωνα με τον Neuman (2013) όσο μικρότερος είναι ο υπό μελέτη πληθυσμός τόσο μεγαλύτερη είναι η αναλογία δειγματοληψίας. Εάν ο πληθυσμός της μελέτης είναι 1000 η αναλογία του δείγματος θα πρέπει να είναι περίπου 30%. Καθώς ο πληθυσμός της μελέτης μεγαλώνει τόσο η αναλογία δειγματοληψίας μειώνεται. Για παράδειγμα για έναν πληθυσμό 10.000 το μέγεθος του δείγματος πρέπει να είναι 1000 (περίπου το 10%) ενώ για πληθυσμούς πάνω από 150.000, μικρότερες αναλογίες δειγματοληψίας (περίπου 1%) είναι αποδεκτές. Για πληθυσμό 10000 το απαιτούμενο δείγμα είναι 1061 ενώ για πληθυσμό 20000 το απαιτούμενο δείγμα είναι 1121. Σύμφωνα με τον Field (2013, σελ. 392) 10 έως 15 οντότητες είναι αρκετές ανά ανεξάρτητη μεταβλητή για την παραγωγή ενός αξιόπιστου μοντέλου, ενώ σύμφωνα με τον Harrell (2001) ένας πρακτικός τρόπος υπολογισμού του απαιτούμενου δείγματος είναι ο πολλαπλασιασμός του συνόλου των μεταβλητών με το 20 .

Το μέγεθος δείγματος που επιλέχθηκε για την συγκεκριμένη μελέτη είναι πρακτικά υπολογισμένο και ανέρχεται σε 1000 κατοικίες προς πώληση.

Με βάση τα ανωτέρω ο Δήμος της Αθήνας χωρίστηκε σε 23 μεγάλα χωρικά τμήματα τα οποία βρίσκονται μέσα στις 7 δημοτικές κοινότητες που προαναφέρθηκαν. Κάθε ένα από αυτά περιέχει υποπεριοχές, όπως αυτές εμφανίζονται στην ιστοσελίδα από την οποία συλλέχθηκε το δείγμα.

Ποιο συγκεκριμένα τα **23 τμήματα – στρώματα** με τις υποπεριοχές τους, στα οποία χωρίστηκε η Αθήνα για την παρούσα μελέτη είναι :

- **Άγιος Ελευθέριος**, ο οποίος αποτελείται από τις υποπεριοχές του Αγίου Ελευθερίου, του Προμπονά και της Ριζούπολης.

- **Αμπελόκηποι** , που περιλαμβάνουν τις υποπεριοχές των Αμπελοκήπων,των Ελληνορώσων, του Ερυθρού, του Γηροκομείου, της Νέας Φιλοθέης, της Πανόρμου, του Πενταγώνου, της Πλατείας Μαβίλη του Πύργου των Αθηνών και της περιοχής Σούτσου.

- **Άνω Πατήσια**, που περιλαμβάνει τις περιοχές της Κυπριάδου, Κλωναρίδου, Λαμπρινής και Πλατειάς Παπαδιαμάντη.

- **Αττική,** που περιλαμβάνει τις περιοχές της Λεωφόρου Λιοσίων , Πλατείας Αττικής Θυμαράκια.

- **Εξάρχεια** που περιλαμβάνει τις περιοχές Εξάρχεια , Μουσείο, Νεάπολη Εξαρχείων και Πολυτεχνείο.

- **Γκάζι** που περιλαμβάνει τις περιοχές Γκάζι, Κεραμεικός, Μεταξουργείο, Ρούφ, Βοτανικός.

- **Γκύζη**, που περιλαμβάνει τις περιοχές Άρειος Πάγος, Λόφος Φινοπούλου, , Πεδίο Άρεως, Πλατεία Γκύζη.

- **Γουδί,** που περιλαμβάνει τις περιοχές Άγιος Θωμάς, Άλσος, Ιατρική Σχολή, Νοσοκομείο Παίδων.

- **Ιστορικό κέντρο** που περιλαμβάνει τις περιοχές Ακρόπολη, Φιλοπάππου, Μοναστηράκι, Πλάκα, Θησείο.

- **Ιλίσια**, που περιλαμβάνει το Hilton και τα κάτω Ιλίσια.

- Το **Κέντρο** που περιλαμβάνει την Ακαδημία, την Ομόνοια, τη παλιά Βουλή, τις Πλατείες Κλαυθμώνος, Βάθη, Κάνιγγος, Καρύτση, Κουμουνδούρου, την περιοχή του Ψυρρή, το Σύνταγμα και το Ζάππειο.

- **Κυψέλη,** που περιλαμβάνει τη Ευελπίδων, Φωκίωνος Νέγρη, Νέα Κυψέλη και τη πλατεία Κυψέλης.

- **Κολωνό,** που περιλαμβάνει την Κολοκυνθού, Κολωνό, Σταθμό Λαρίσης Ακαδημία Πλάτωνος.

- **Κουκάκι,** που περιλαμβάνει Φιξ, Κουκάκι και Μακρυγιάννη.

- **Μετς,** που περιλαμβάνει Α’ νεκροταφείο, Καλλιμάρμαρο, Μετς.

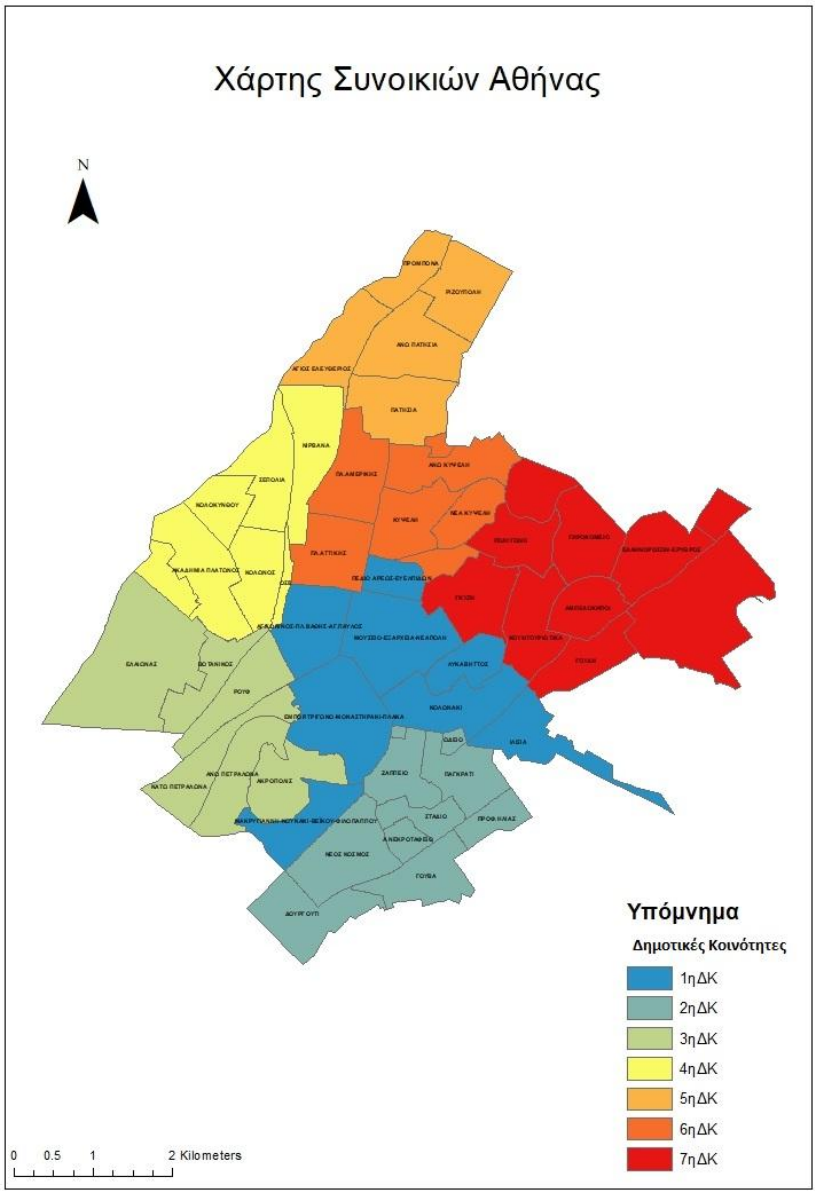
- **Νέο Κόσμο,** που περιλαμβάνει Άγιο Ιωάννη, και Άγιο Σώστη, Νέο Κόσμο και Κυνοσάργες.

- **Παγκράτι,** που περιλαμβάνει Άγιο Αρτέμιο, Άλσος Παγκρατίου, Κέντρο Παγκρατίου Προφήτη Ηλία και Βαρνάβα.

- **Πατήσια,** που περιλαμβάνει Γκράβα Όσο Λουκά και Πλατεία Κολιάτσου.

- **Λεωφόρους Πατησίων και Αχαρνών,** που περιλαμβάνουν Αγίου Μελετίου, Πλατεία Βικτωρία Μάρνη, Άγιο Νικόλαο, Άγιο Παντελεήμονα Κάτω Πατήσια, Πλατεία Αμερικής και Τρείς Γέφυρες.
- **Πετράλωνα,** που περιλαμβάνουν Άνω και Κάτω Πετράλωνα.
- **Πολύγωνο,** που περιλαμβάνει Αττικό Άλσος, Ευελπίδων Πολύγωνο Τουρκοβούνια.
- **Σκουζέ,** που περιλαμβάνει Σκουζέ και Σεπόλια.

Στο χάρτη 4.2 αποτυπώνονται οι συνοικίες της Αθήνας.



Χάρτης 4.2. Οι συνοικίες της Αθήνας. Πηγή: www.spiti24.gr (2018). Ιδία επεξεργασία.

4.4 Πηγή Δεδομένων Στοιχείων Δειγματοληψίας στην Περιοχή Μελέτης και Παραδοχές

Η πηγή συλλογής των πρωτογενών δεδομένων για την παρούσα εργασία αποτέλεσε γνωστή ιστοσελίδα του διαδικτύου (www.spiti24.gr), εμπλουτισμένη με κατά προσέγγιση γεωγραφική θέση των ακινήτων.

Πολύ σημαντικό κομμάτι της διαδικασίας λήψης του δείγματος αποτέλεσε η όσο το δυνατόν ακριβέστερη χωρική τοποθέτηση όλων των ακινήτων (σύμφωνα πάντα με τα στοιχεία που παρείχε η ιστοσελίδα), καθώς η εφαρμογή του δείγματος με γεωγραφικό χαρακτήρα, η απόδοση της θέσης του και η οπτικοποίηση του οδηγεί σε μία ολοκληρωμένη μελέτη, κάνοντας σαφή τα χωρικά της όρια. Επίσης, κρίθηκε απαραίτητη η συλλογή και επεξεργασία μόνο των αγγελιών όπου περιέχουν φωτογραφικό υλικό κατοικιών και ο βασικός λόγος είναι ότι μέσω των φωτογραφιών διαφαίνεται η γενικότερη κατάσταση του ακινήτου που αποτυπώνεται στις μεταβλητές.

Καθώς η ακριβής γνώση της τιμής αγοραπωλησίας του κάθε ακινήτου, είναι αδύνατη στα πλαίσια αυτής της εργασίας, ένα επιπλέον στοιχείο είναι ότι έπρεπε να οριστεί ως παραδοχή ότι οι ζητούμενες τιμές πώλησης ήταν και οι αγοραίες αξίες των ακινήτων, καθώς ήταν αναμενόμενο στην εξαρτημένη μεταβλητή να υπάρχει διαφορά μεταξύ της αγοραίας αξίας του ακινήτου και της τιμής που δινόταν στο διαδίκτυο. Άλλωστε είναι γενικώς παραδεκτό ότι οι ζητούμενες τιμές είναι κατά 10-15% υψηλότερες από τις αγοραίες. Η τιμή αγοροπωλησίας παρέχεται από τις ηλεκτρονικές αγγελίες και ισοδυναμεί με τη ζητούμενη τιμή πώλησης των διαμερισμάτων του δείγματος. Μια δεύτερη παραδοχή ήταν ότι οι πωλητές είχαν προσδιορίσει προσεγγιστικά την θέση των ακινήτων τους μέσω της ιστοσελίδας και είχαν ορθά απεικονίσει πληροφοριακά τις ιδιοκτησίες τους. Τέλος ουσιαστική παραδοχή ήταν ότι τα ακίνητα προς πώληση ήταν μόνο αυτά που προέρχονταν από την συγκεκριμένη ιστοσελίδα.

Από τα στοιχεία που παρείχε η ιστοσελίδα (www.spiti24.gr), υπολογίστηκε το σύνολο του πληθυσμού για κάθε ένα στρώμα ξεχωριστά καθώς και το ποσοστό του δείγματος αναλογικά με το συνολικό πληθυσμό των ακινήτων προς πώληση στο Δήμο Αθηναίων.

Επόμενο βήμα της διαδικασίας ήταν η συστηματική επιλογή δείγματος σε κάθε στρώμα ξεχωριστά, επιλέγοντας τα μέλη του δείγματος από τον κάθε υποπληθυσμό, ένα παρά ένα, χωρίς επανάθεση και χωρίς να μας ενδιαφέρει η σειρά επιλογής.

4.5 Συλλεχθείσα και Υπολογισθείσα Πληροφορία

Για κάθε δείγμα συλλέχθηκαν τα εξής στοιχεία: Αξία, Εμβαδόν, Αριθμός υπνοδωματίων Αριθμός μπάνιων, Αριθμός WC, Όροφος, Έτος κατασκευής, Κατάσταση του ακινήτου, Ύπαρξης ανακαίνισης ή όχι, Ύπαρξη βεράντας ή όχι, Απεικόνιση ως γωνιακό ή διαμπερές, Ύπαρξη ασανσέρ ή όχι, Ύπαρξη θέσης στάθμευσης ή όχι, Ύπαρξη θέρμανσης ή όχι, Είδος θέρμανσης (πετρέλαιο ή φυσικό αέριο, ή τίποτα), Ύπαρξη κλιματισμού ή όχι, Ύπαρξη θέας ή όχι, Ύπαρξη τζακιού ή όχι, Ημερομηνία δημοσίευσης αγγελίας για πρώτη φορά. Επίσης η συλλέχθηκε η πληροφορία που αφορούσε την τοποθεσία του ακινήτου, ενώ έγινε η ανάλογη κωδικοποίηση η οποία ήταν σχετική με την Δημοτική κοινότητα που βρίσκεται το ακίνητο και την επί μέρους περιοχή. Στον πίνακα που ακολουθεί παρατίθεται η κωδικοποίηση των περιοχών (πίνακας 4.3).

Πίνακας 4.3

Κωδικοποίηση περιοχών

Γεωγραφικός κωδικός	Δημοτική Κοινότητα	Περιοχή
101	1η ΔΚ	Μακρυγιάννη Κουκάκι Βείκου Φιλοπάππου
102	1η ΔΚ	Μουσείο Εξάρχεια Νεάπολη
103	1η ΔΚ	Άγιος Κων/νος Πλατεία Βάθη Άγιος Παύλος
104	1η ΔΚ	Κολωνάκι
105	1η ΔΚ	Ιλίσια
106	1η ΔΚ	Εμπορικό τρίγωνο Μοναστηράκι Πλάκα
107	1η ΔΚ	Πεδίο Άρεως Ευελπίδων
201	2η ΔΚ	Νέος Κόσμος
202	2η ΔΚ	Γούβα
203	2η ΔΚ	Στάδιο
204	2η ΔΚ	Παγκράτι
205	2η ΔΚ	Προφήτης Ηλίας
206	2η ΔΚ	Ντουργούντι
301	3η ΔΚ	Άνω Πετράλωνα

302	3η ΔΚ	Κάτω Πετράλωνα
303	3η ΔΚ	Ρούφ
304	3η ΔΚ	Βοτανικός
401	4η ΔΚ	Νιρβάνα
402	4η ΔΚ	Σεπόλια
403	4η ΔΚ	Κολωνός
404	4η ΔΚ	Κολοκυνθού
405	4η ΔΚ	Ακαδημία Πλάτωνος
501	5η ΔΚ	Άνω Πατήσια
502	5η ΔΚ	Άγιος ελευθέριος
503	5η ΔΚ	Πατήσια
504	5η ΔΚ	Προμπονά
505	5η ΔΚ	Ριζούπολη
601	6η ΔΚ	Πλατεία Αμερικής
602	6η ΔΚ	Άνω Κυψέλη
603	6η ΔΚ	Πλατεία Αττικής
604	6η ΔΚ	Κυψέλη
605	6η ΔΚ	Νέα Κυψέλη
701	7η ΔΚ	Γουδί
702	7η ΔΚ	Πολύγωνο
703	7η ΔΚ	Γηροκομείο
704	7η ΔΚ	Γκύζη
705	7η ΔΚ	Κουντουριώτικα
706	7η ΔΚ	Ελληνορώσων Ερυθρός
707	7η ΔΚ	Αμπελόκηποι

Ιδία επεξεργασία

Παράλληλα συλλέχθηκαν δευτερογενή δεδομένα από την Ελληνική Στατιστική Αρχή και επεξεργάστηκαν τα εξής στοιχεία που αφορούν τα δημοτικά διαμερίσματα του δήμου Αθηναίων: Ποσοστό ανεργίας, Συνολικός πληθυσμός, Ποσοστό ξένων

υπηκόων (αλλοδαποί κάτοικοι και μετανάστες), Ποσοστό κενών κατοικιών (*www.statistics.gr)* . Τα στοιχεία αυτά παρουσιάστηκαν προηγουμένως στον πίνακα 4.1 και 4.2 .

Από στοιχεία της Εταιρίας Ανάπτυξης και Τουριστικής Προβολής Αθηνών του Δήμου Αθηναίων που αναφέρονται στο «Κείμενο διαβούλευσης για το έργο: Αθήνα 2020» συλλέχθηκε το Μέσο Φορολογητέο Εισόδημα, Ποσοστό πρασίνου στην επιφάνεια κάθε δημοτικού διαμερίσματος, Τετραγωνικά μέτρα πρασίνου που αντιστοιχούν ανά κάτοικο.

Στον πίνακα 4.4 που παρατίθεται παρουσιάζεται το ποσοστό πρασίνου ανά κάτοικο και τα τετραγωνικά μέτρα επί της συνολικής επιφάνειας κάθε δημοτικού διαμερίσματος.

Πίνακας 4.4

Στατιστικά στοιχεία αναλογούντων τετραγωνικών μέτρων πρασίνου ανά κάτοικο και ανά δημοτική κοινότητα και ποσοστό πρασίνου επί εκτάσεως δημοτικής κοινότητας.

Δημοτική Κοινότητα	Πράσινο ανά κάτοικο σε τμ	Ποσοστό Πρασίνου επί της συνολικής επιφάνειας της δημοτικής κοινότητας (%)
1η	15.87	23
2η	6.48	13
3η	21.12	19
4η	3.20	5

5η	2.33	4
6η	2.50	6
7η	12.93	30

Πηγή: Κείμενο διαβούλευσης για το έργο : Αθήνα 2020 του Δήμου Αθηναίων. (2016). Ιδία επεξεργασία.

Όλες αυτές οι πληροφορίες επεξεργάστηκαν μέσω του λογισμικού Excel, του στατιστικού πακέτου λογισμικού SPSS και του Arc Gis ώστε να γίνουν οι κατάλληλοι υπολογισμοί και κωδικοποιήσεις .

Στον πίνακα 4.5 αποτυπώνεται η συλλεχθείσα και υπολογισθείσα, μετά από επεξεργασία, πληροφορία από την ιστοσελίδα για την παρούσα μελέτη.

Πίνακας 4.5

Συλλεχθείσα και υπολογισθείσα πληροφορία από ιστοσελίδα πώλησης κατοικιών

Μεταβλητές	Παρουσίαση		Μετατροπή με Dummy Value κατά αντιστοιχία	Συλλεχθείσα ή Υπολογισθείσα	Κλίμακα Μέτρησης
Αριθμός	Αύξων αριθμός	Αριθμός	-	Υπολογίστηκε	Ονομαστική
Συντεταγμένες Χ	Αριθμός ποσότητα	Αριθμός	-	Υπολογίστηκε	Αναλογική
Συντεταγμένες Υ	Αριθμός ποσότητα	Αριθμός	-	Υπολογίστηκε	Αναλογική
Γεωγραφικός κωδικός	Αριθμός ποσότητα	Αριθμός	-	Υπολογίστηκε	Ονομαστική
Αξία κατοικίας	Αριθμός σε ευρώ	Αριθμός	-	Συλλέχθηκε	Ονομαστική
Όνομα ΔΚ	Ονομασία ΔΚ	Ονομασία	-	Συλλέχθηκε	Ονομαστική
Περιοχή	Ονομασία περιοχής	Ονομασία	-	Συλλέχθηκε	-
Γειτονιά	Ονομασία γειτονιάς	Ονομασία	-	Συλλέχθηκε	-
Ποσοστό Ελλήνων	Αριθμός ποσότητα	Αριθμός	-	Συλλέχθηκε	Ονομαστική

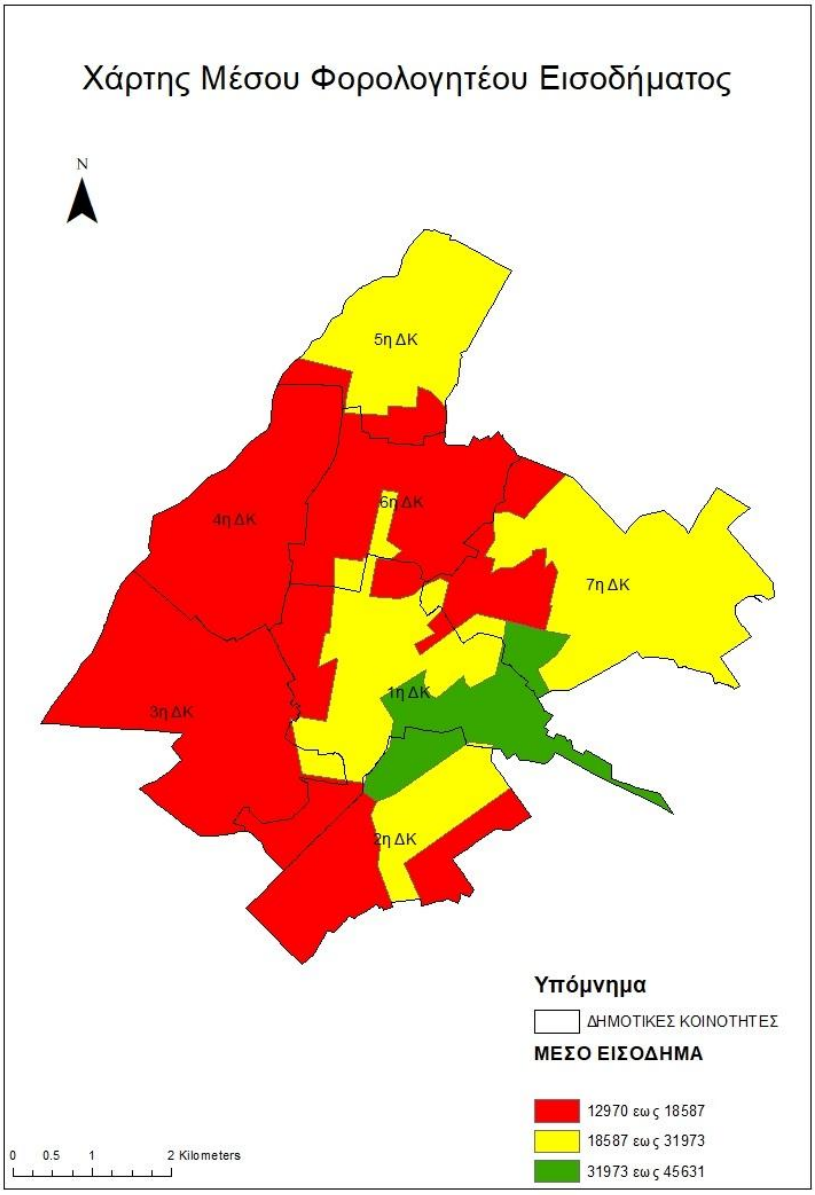
Ποσοστό Αλλοδαπών	Αριθμός ποσότητα	Αριθμός	-	Συλλέχθηκε	Ονομαστική
Ποσοστό Ανεργίας	Αριθμός ποσότητα	Αριθμός	-	Συλλέχθηκε	Ονομαστική
Ποσοστό Κενών κατοικιών	Αριθμός ποσότητα	Αριθμός	-	Συλλέχθηκε	Ονομαστική
Ποσοστό Πρασίνου	Αριθμός ποσότητα	Αριθμός	-	Συλλέχθηκε	Ονομαστική
Ποσοστό Επιμερισμού Κατοίκων ανά ΔΚ	Αριθμός ποσότητα	Αριθμός	-	Συλλέχθηκε	Ονομαστική
Πράσινο ανά κάτοικο τμ	Αριθμός ποσότητα	Αριθμός	-	Συλλέχθηκε	Ονομαστική
Απόσταση από Αστυνομικό τμήμα	Αριθμός ποσότητα	Αριθμός	-	Υπολογίστηκε	Ονομαστική
Απόσταση από Βιομηχανία βιοτεχνία	Αριθμός ποσότητα	Αριθμός	-	Υπολογίστηκε	Ονομαστική
Απόσταση από Βρεφονηπιακό σταθμό	Αριθμός ποσότητα	Αριθμός	-	Υπολογίστηκε	Ονομαστική
Απόσταση από Εμπόριο	Αριθμός ποσότητα	Αριθμός	-	Υπολογίστηκε	Ονομαστική
Απόσταση από Ιστορικό κέντρο	Αριθμός ποσότητα	Αριθμός	-	Υπολογίστηκε	Ονομαστική
Απόσταση από Νεκροταφείο	Αριθμός ποσότητα	Αριθμός	-	Υπολογίστηκε	Ονομαστική
Απόσταση από Νοσοκομείο	Αριθμός ποσότητα	Αριθμός	-	Υπολογίστηκε	Ονομαστική
Πλησιέστερο Είδος Συγκοινωνίας	Λεκτικό	Μετρό- Ηλεκτρικός, Λεωφορείο , Τραμ	1 , 2 , 3	Συλλέχθηκε	Ιεραρχική
Απόσταση από Πλησιέστερη Συγκοινωνία	Αριθμός ποσότητα	Αριθμός	-	Υπολογίστηκε	Ονομαστική
Απόσταση από Ναό	Αριθμός ποσότητα	Αριθμός	-	Υπολογίστηκε	Ονομαστική
Απόσταση από Σχολείο	Αριθμός ποσότητα	Αριθμός	-	Υπολογίστηκε	Ονομαστική
Απόσταση από Αθλητική εγκατάσταση	Αριθμός ποσότητα	Αριθμός	-	Υπολογίστηκε	Ονομαστική
Απόσταση από Πράσινο	Αριθμός ποσότητα	Αριθμός	-	Υπολογίστηκε	Ονομαστική
Απόσταση από περιοχή Πλάκας	Αριθμός ποσότητα	Αριθμός	-	Υπολογίστηκε	Ονομαστική
Απόσταση από Πεζόδρομο	Αριθμός ποσότητα	Αριθμός	-	Υπολογίστηκε	Ονομαστική
Απόσταση από ΟΣΕ	Αριθμός ποσότητα	Αριθμός	-	Υπολογίστηκε	Ονομαστική

Απόσταση από Μουσείο	Αριθμός ποσότητα	Αριθμός	-	Υπολογίστηκε	Ονομαστική
Απόσταση από Πρωτεύον οδικό δίκτυο	Αριθμός ποσότητα	Αριθμός	-	Υπολογίστηκε	Ονομαστική
Απόσταση από Βασικό οδικό δίκτυο	Αριθμός ποσότητα	Αριθμός	-	Υπολογίστηκε	Ονομαστική
Μέσο εισόδημα κατοίκων	Αριθμός ποσότητα	Από 31973 έως 45631, από 18587 έως 31973, από 12970 έως 18587	1 , 2 , 3	Συλλέχθηκε	Ιεραρχική
Αντικειμενική αξία	Αριθμός ποσότητα	Αριθμός	-	Συλλέχθηκε	Ονομαστική
Εμβαδόν	Αριθμός σε τμ	Αριθμός	-	Συλλέχθηκε	Ονομαστική
Τιμή ανά τμ	Αριθμός σε ευρώ	Αριθμός	-	Υπολογίστηκε	Ονομαστική
Αριθμός υπνοδωματίων	Αριθμός ποσότητα	Αριθμός	-	Συλλέχθηκε	Ονομαστική
Αριθμός μπάνιων	Αριθμός ποσότητα	Αριθμός	-	Συλλέχθηκε	Ονομαστική
Ύπαρξη WC	Αριθμός ποσότητα	Ναι, Όχι	1 , 0	Συλλέχθηκε	Ονομαστική
Ύπαρξη αποθήκης	Λεκτικό	Ναι, Όχι	1 , 0	Συλλέχθηκε	Ιεραρχική
Αριθμός ορόφου	Αριθμός ποσότητα	Αριθμός	-	Συλλέχθηκε	Ονομαστική
Έτος κατασκευής	Αριθμός ποσότητα	Αριθμός	-	Συλλέχθηκε	Ονομαστική
Ηλικία κατασκευής	Αριθμός ποσότητα	Αριθμός	-	Υπολογίστηκε	Ονομαστική
Είδος κατοικίας	Λεκτικό	Μονοκατοικία, Μεζονέτα, Διαμέρισμα	1 , 2 , 3	Συλλέχθηκε	Ιεραρχική
Κατάσταση	Λεκτικό	Άριστη, Πολύ καλή, Καλή, Μέτρια, Κακή	1 , 2, 3 , 4 , 5	Συλλέχθηκε	Ιεραρχική
Ύπαρξη ανακαίνισης	Λεκτικό	Ναι, Όχι	1 , 0	Συλλέχθηκε	Ιεραρχική
Ύπαρξη βεράντας	Λεκτικό	Ναι, Όχι	1 , 0	Συλλέχθηκε	Ιεραρχική
Γωνιακό ή Διαμπερές	Λεκτικό	Ναι, Όχι	1 , 0	Συλλέχθηκε	Ιεραρχική
Ύπαρξη ανσασέρ	Λεκτικό	Ναι, Όχι	1 , 0	Συλλέχθηκε	Ιεραρχική
Ύπαρξη θέσης στάθμευσης	Λεκτικό	Ναι, Όχι	1 , 0	Συλλέχθηκε	Ιεραρχική
Είδος θέρμανσης	Λεκτικό	Αυτόνομη, Κεντρική, Χωρίς	1 , 2, 3	Συλλέχθηκε	Ιεραρχική
Καύσιμο θέρμανσης	Λεκτικό	Φυσικό αέριο, Πετρέλαιο, Χωρίς	1 , 2 , 3	Συλλέχθηκε	Ιεραρχική

Ύπαρξη κλιματισμού	Λεκτικό	Ναι, Όχι	1 , 0	Συλλέχθηκε	Ιεραρχική
Ύπαρξη θέας	Λεκτικό	Ναι, Όχι	1 , 0	Συλλέχθηκε	Ιεραρχική
Ύπαρξη τζακιού	Λεκτικό	Ναι, Όχι	1 , 0	Συλλέχθηκε	Ιεραρχική
Ημέρες προς πώληση	Αριθμός ποσότητα	Αριθμός	-	Συλλέχθηκε	Ονομαστική

Ιδία επεξεργασία

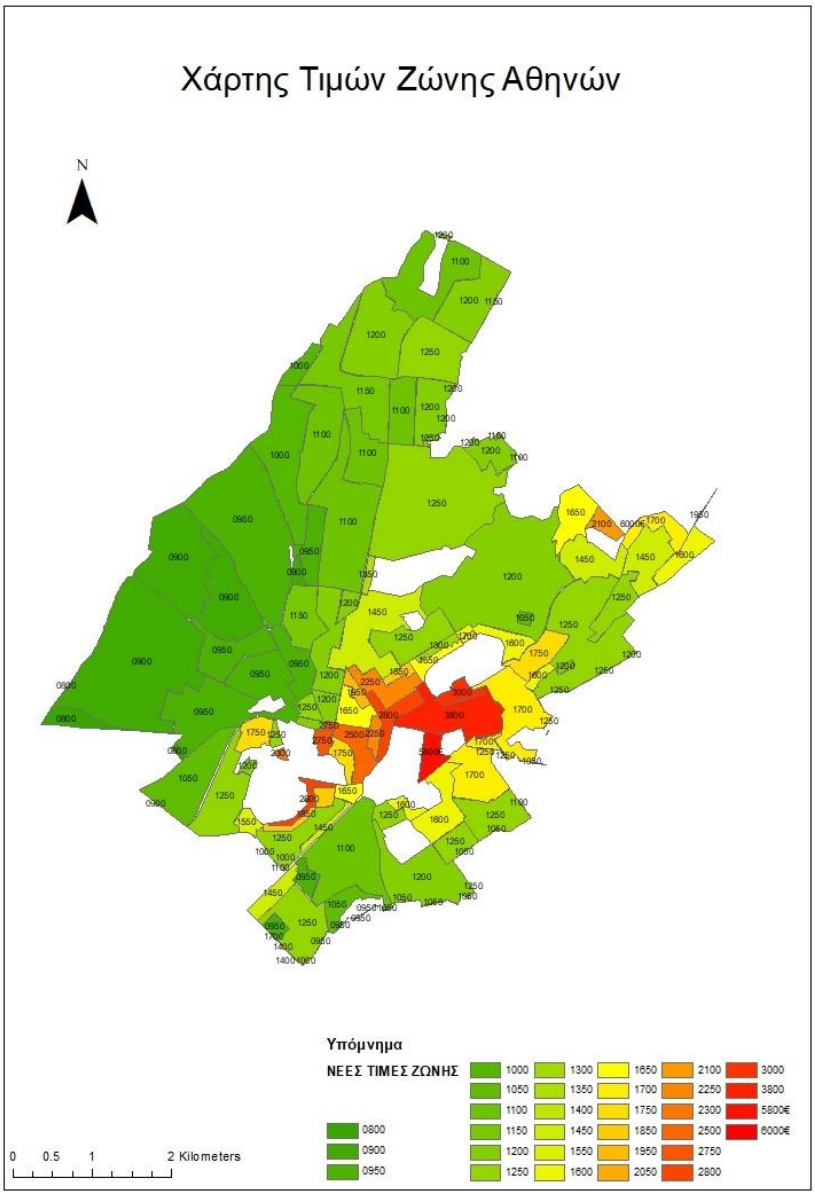
Στο χάρτη που ακολουθεί παρουσιάζεται το μέσο φορολογητέο εισόδημα κατοίκων Αθηνών (Χάρτης 4.3).



Χάρτης 4.3. Μέσο φορολογητέο εισόδημα έτους 2012.

Πηγή: Κείμενο διαβούλευσης για το έργο: Αθήνα 2020 του Δήμου Αθηναίων. (2016). Ιδία επεξεργασία.

Από το Υπουργείο Οικονομικών συλλέχθηκαν στοιχεία που αφορούν τις νέες τιμές ζώνης σε κάθε περιοχή του δήμου (www.minfin.gr) οι οποίες και παρουσιάζονται στο χάρτη 4.4 που ακολουθεί.



Χάρτης 4.4. Τιμές ζώνης Δήμου Αθηναίων. Πηγή Υπουργείο Οικονομικών (2018). Επεξεργασία από Google..

Από την κρατική ιστοσελίδα γεωπληροφοριών data.gov.gr (πρόσβαση 2018) συλλέχθηκαν και επεξεργάστηκαν τα εξής στοιχεία: Γενικό αρχείο χρήσεων γης, όρια

Δήμου Αθηναίων, όρια Δημοτικών κοινοτήτων, όρια Συνοικιών και Γειτονιών, ονομασία Κοινοχρήστων χώρων, ονομασία Όμορων Δήμων, όρια ευθύνης και θέσεις Αστυνομικών τμημάτων, θέσεις Βρεφονηπιακών σταθμών, Μουσείων, Ναών, στάσεις Συγκοινωνιών (Μετρό, ΗΣΑΠ, ΟΑΣΑ), θέση- διαδρομή γραμμών ΟΣΕ, θέσεις Σχολείων (δημοτικά, γυμνάσια, λύκεια, Πανεπιστήμια), θέσεις Δημοσίων κτιρίων, Νοσοκομείων, Νεκροταφείων, Εμπορίου, Βιομηχανικής περιοχής, όρια Ιστορικού κέντρου, όρια περιοχής Πλάκας, Θέσεις πρασίνου (αστικού και περιαστικού), στοιχεία και θέσεις Οδικού δικτύου (βασικό, πρωτεύον, απλό δίκτυο), Σημεία ενδιαφέροντος, θέσεις Αθλητικών εγκαταστάσεων και Πεζόδρομων (www.data.gov.gr). Στον πίνακα 4.6 αποτυπώνεται η γεωγραφική πληροφορία που χρησιμοποιήθηκε στην παρούσα μελέτη καθώς και ο τρόπος της γεωγραφικής της απεικόνιση .

Πίνακας 4.6

Γεωγραφική πληροφορία και γεωγραφική απεικόνιση της.

Γεωγραφική Πληροφορία	Γεωγραφική απεικόνιση
Όρια Δήμου Αθηναίων	Πολύγωνο
Όρια δημοτικών κοινοτήτων	Πολύγωνο
Όρια συνοικιών	Πολύγωνο
Όρια γειτονιών	Πολύγωνο
Όρια ευθύνης αστυνομικών τμημάτων	Πολύγωνο
Θέσεις εκπαίδευσης	Πολύγωνο
Θέσεις νοσοκομείων	Πολύγωνο
Θέσεις νεκροταφείων	Πολύγωνο
Θέσεις εμπορίου	Πολύγωνο
Θέσεις βιομηχανίας	Πολύγωνο
Θέσεις αθλητικών εγκαταστάσεων	Πολύγωνο
Θέση πρασίνου	Πολύγωνο

Θέσεις Ναών	Πολύγωνο
Όρια ιστορικού κέντρου	Πολύγωνο
Όρια περιοχής Πλάκας	Πολύγωνο
Θέσεις πεζόδρομων	Πολύγωνο
Όρια μέσων φορολογητέων εισοδημάτων	Πολύγωνο
Όρια νέων τιμών ζώνης ΥΠ.ΟΙΚ	Πολύγωνο
Θέσεις μουσείων	Σημείο
Θέσεις δημοσίων κτιρίων	Σημείο
Θέσεις σημείων ενδιαφέροντος	Σημείο
Θέσεις στάσεων συγκοινωνίας	Σημείο
Θέσεις αστυνομικών τμημάτων	Σημείο
Θέσεις βρεφονηπιακών σταθμών	Σημείο
Ονομασία όμορων Δήμων	Σημείο
Ονομασία ΚΧ	Σημείο
Θέσεις πρωτεύοντος οδικού δικτύου	Γραμμή
Θέσεις βασικού οδικού δικτύου	Γραμμή
Θέσεις οδικού δικτύου	Γραμμή

Πηγή:data.gov.gr (2018) Ιδία επεξεργασία.

Τα παραπάνω στοιχεία επεξεργάστηκαν και αποτυπώθηκαν με την χρήση του λογισμικού ArcGis και Excel. Στην πρώτη φάση της επεξεργασίας, ενώθηκαν τα στοιχεία της στατιστικής υπηρεσίας με το shp file της γεωγραφικής πληροφορίας των δημοτικών κοινοτήτων το οποίο έλαβε και τις πληροφορίες σχετικά με τα ποσοστά Ελλήνων και αλλοδαπών, το ποσοστό ανεργίας, το ποσοστό κενών κατοικιών, το ποσοστό πρασίνου, το ποσοστό επιμερισμού κατοίκων ανά δημοτική κοινότητα και τα αναλογούντα τετραγωνικά μέτρα ανά κάτοικο σε κάθε δημοτική κοινότητα. Σε δεύτερη φάση υπολογίστηκαν οι αποστάσεις των δεδομένων στοιχείων με τα σημεία που απεικονίζουν τις προς πώληση κατοικίες. Σε τρίτη φάση η γεωχωρική πληροφορία που περιείχε όλα τα στοιχεία που αποτυπώθηκαν και υπολογίστηκαν με

την χρήση του λογισμικού ArcGis ενώθηκε με κάθε σημείο που απεικόνιζε την κάθε μια προς πώληση κατοικία και σε τέταρτη φάση η κάθε μια οντότητα έλαβε τα στοιχεία που συλλέχθηκαν από την ιστοσελίδα. Το τελικό αποτέλεσμα περιείχε την γεωγραφική θέση του κάθε ακινήτου με όλη την συλλεχθείσα και υπολογισθείσα πληροφορία.

Η κωδικοποίηση ήταν απαραίτητη ώστε να καταστεί το τελικό αρχείο εύκολα αναγνώσιμο και διαχειρίσιμο από το λογισμικό SPSS.

Στον πίνακα 4.7 παρουσιάζεται η κωδικοποίηση της πληροφορίας και στους χάρτες 4.7 έως 4.21 παρουσιάζονται τα βασικά στοιχεία γεωχωρικής πληροφορίας που εξετάστηκαν ως μεταβλητές στην παρούσα μελέτη.

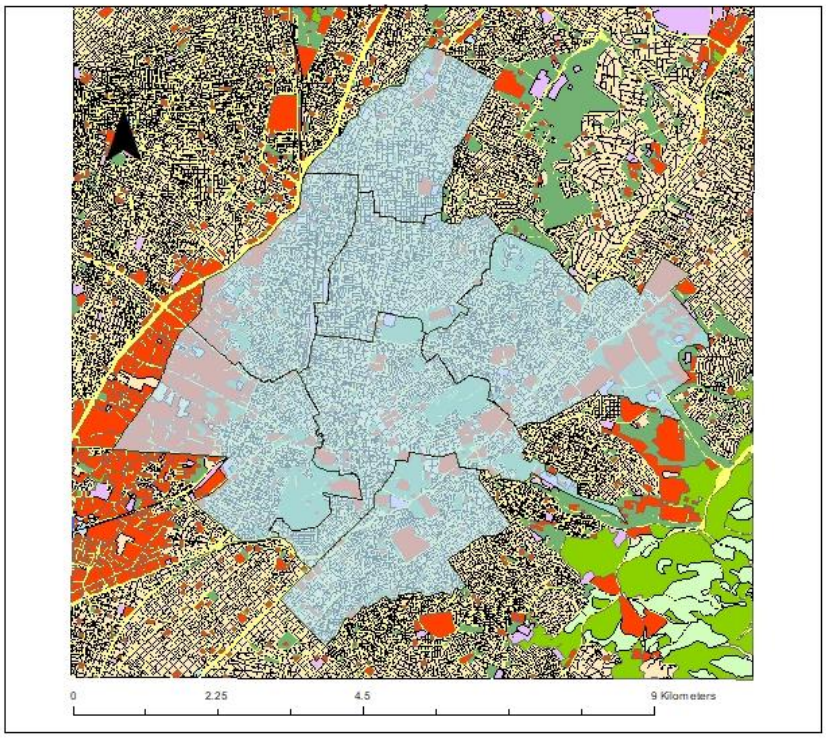
Πίνακας 4.7

Κωδικοποίηση μεταβλητών

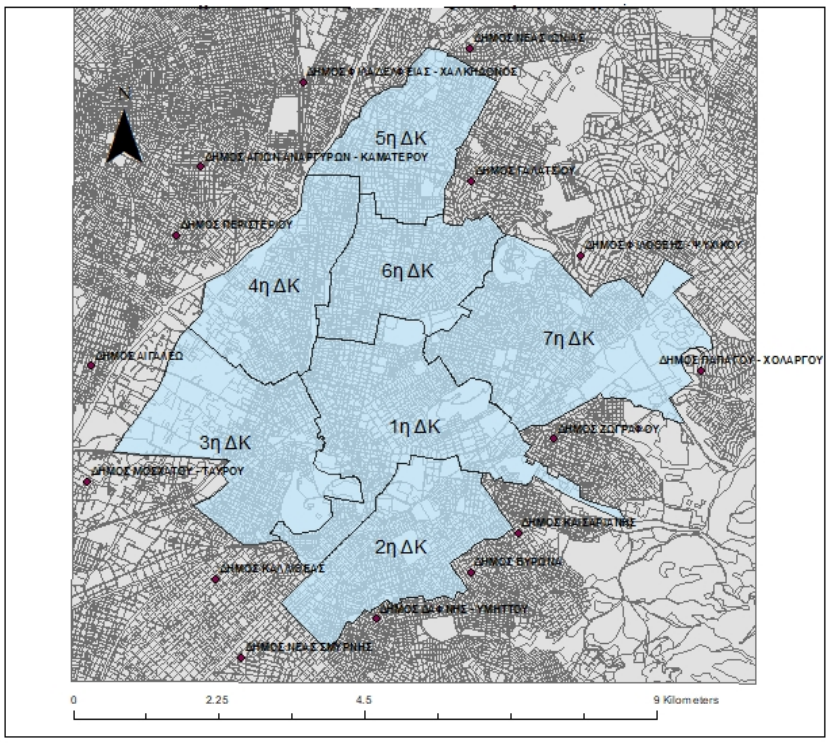
Ονομασία μεταβλητής	Κωδικός	Ονομασία μεταβλητής	Κωδικός	Ονομασία μεταβλητής	Κωδικός
Αξία	Y	Πλησιέστερο Είδος Συγκοινωνίας	X18	Ύπαρξη WC	X36
Όνομα ΔΚ	X1	Απόσταση από Πλησιέστερη Συγκοινωνία	X19	Ύπαρξη Αποθήκης	X37
Περιοχή	X2	Απόσταση από Ναό	X20	Όροφος	X38
Γειτονιά	X3	Απόσταση από Σχολείο	X21	Έτος κατασκευής	X39
Ποσοστό Ελλήνων	X4	Απόσταση από Αθλητική εγκατάσταση	X22	Ηλικία κατοικίας	X40
Ποσοστό Αλλοδαπών	X5	Απόσταση από Πράσινο	X23	Είδος κατοικίας	X41
Ποσοστό Ανεργίας	X6	Απόσταση από περιοχή Πλάκας	X24	Κατάσταση κατασκευής	X42
Ποσοστό Κενών κατοικιών	X7	Απόσταση από Πεζόδρομο	X25	Ύπαρξη ανακαίνισης	X43
Ποσοστό Πρασίνου	X8	Απόσταση από ΟΣΕ	X26	Ύπαρξη Βεράντας	X44
Ποσοστό Επιμερισμού Κατοίκων ανά ΔΚ	X9	Απόσταση από Μουσείο	X27	Γωνιακό ή Διαμπερές	X45
Πράσινο ανά κάτοικο τμ	X10	Απόσταση από Πρωτεύον οδικό δίκτυο	X28	Ύπαρξη ασανσέρ	X46
Απόσταση από Αστυνομικό τμήμα	X11	Απόσταση από Βασικό οδικό δίκτυο	X29	Ύπαρξη θέσης στάθμευσης	X47
Απόσταση από Βιομηχανία βιοτεχνία	X12	Μέσο εισόδημα κατοίκων	X30	Είδος Θέρμανσης	X48
Απόσταση από Βρεφονηπιακό σταθμό	X13	Τιμές Ζώνης	X31	Είδος Καυσίμου θέρμανσης	X49

Απόσταση από Εμπόριο	X14	Εμβαδόν κατοικίας	X32	Ύπαρξη Κλιματισμού	X50
Απόσταση από Ιστορικό κέντρο	X15	Τιμή ανά τμ	X33	Ύπαρξη Θέας	X51
Απόσταση από Νεκροταφείο	X16	Αριθμός υπνοδωματίων	X34	Ύπαρξη Τζακιού	X52
Απόσταση από Νοσοκομείο	X17	Αριθμός Μπάνιων	X35	Αριθμός Ημερών διαθεσιμότητας	X53

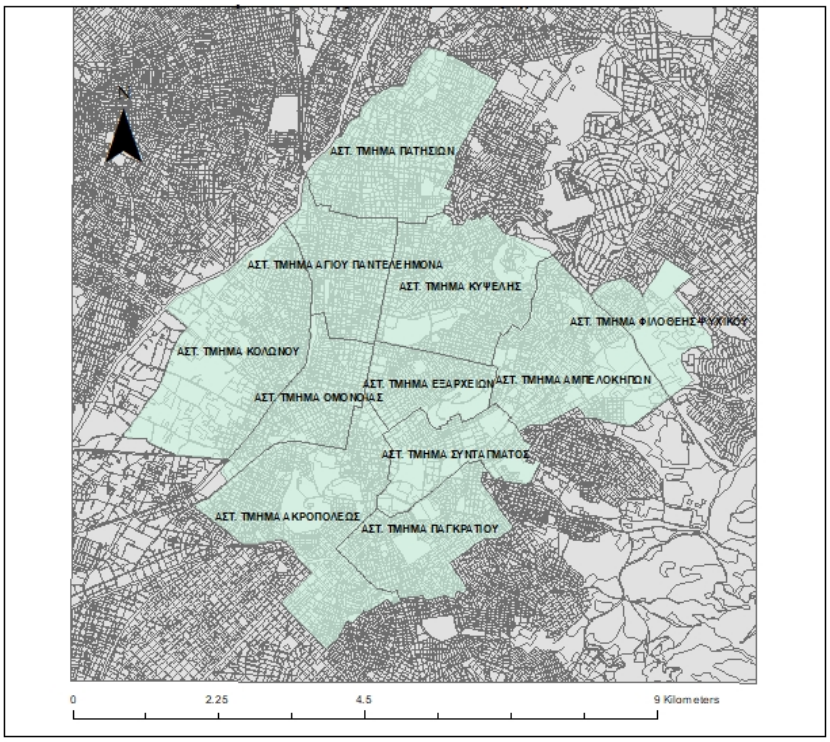
Ιδία επεξεργασία.



Χάρτης 4.5.Γενική άποψη του χώρου μελέτης. Πηγή:data.gov.gr (2018).
Ιδία επεξεργασία

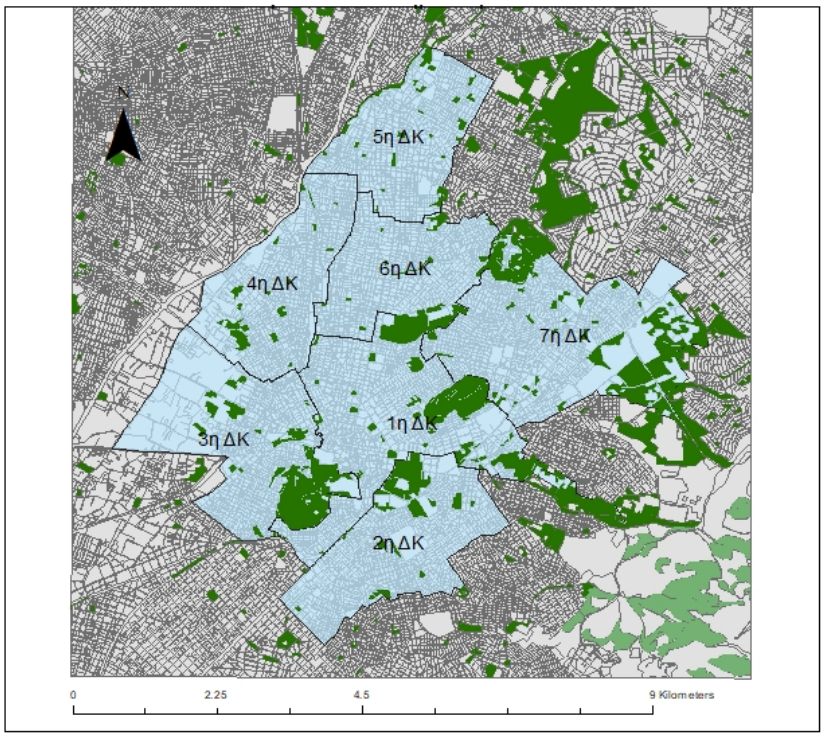


Χάρτης 4.6.Δημοτικές κοινότητες Δήμου Αθηναίων και όμοροι Δήμοι.
Πηγή:data.gov.gr (2018).
Ιδία επεξεργασία.



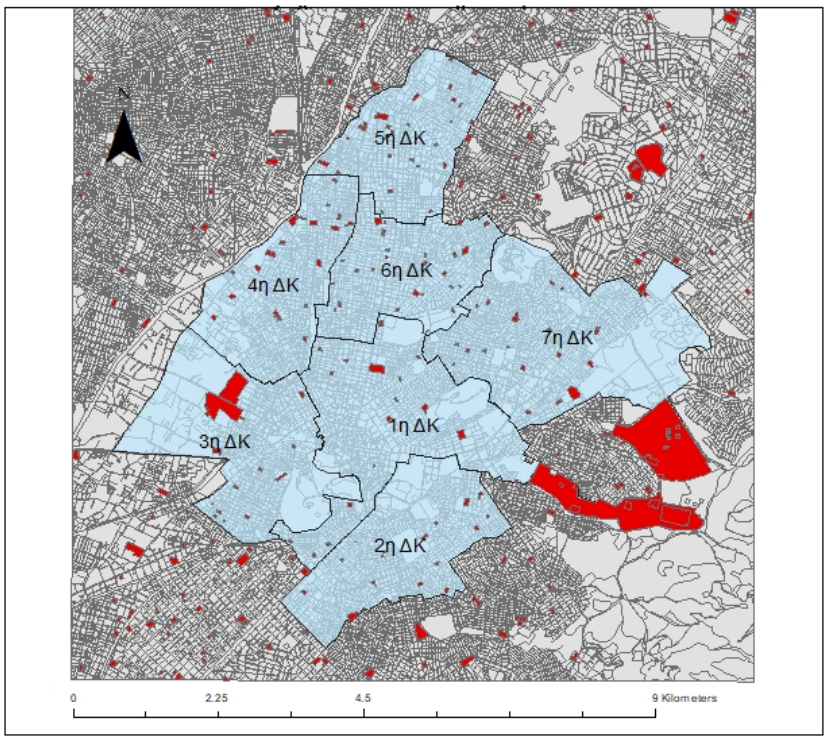
Χάρτης 4.7. Όρια ευθύνης Αστυνομικών τμημάτων Δήμου Αθηναίων.

Πηγή:data.gov.gr (2018). Ιδία επεξεργασία.

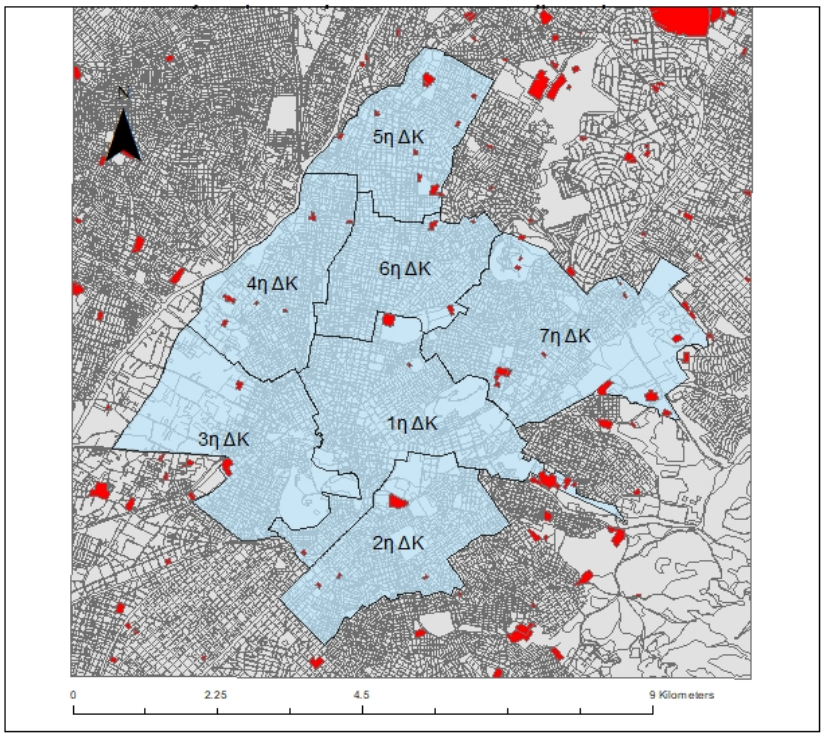


Χάρτης 4.8.Πράσινο στο Δήμο Αθηναίων.

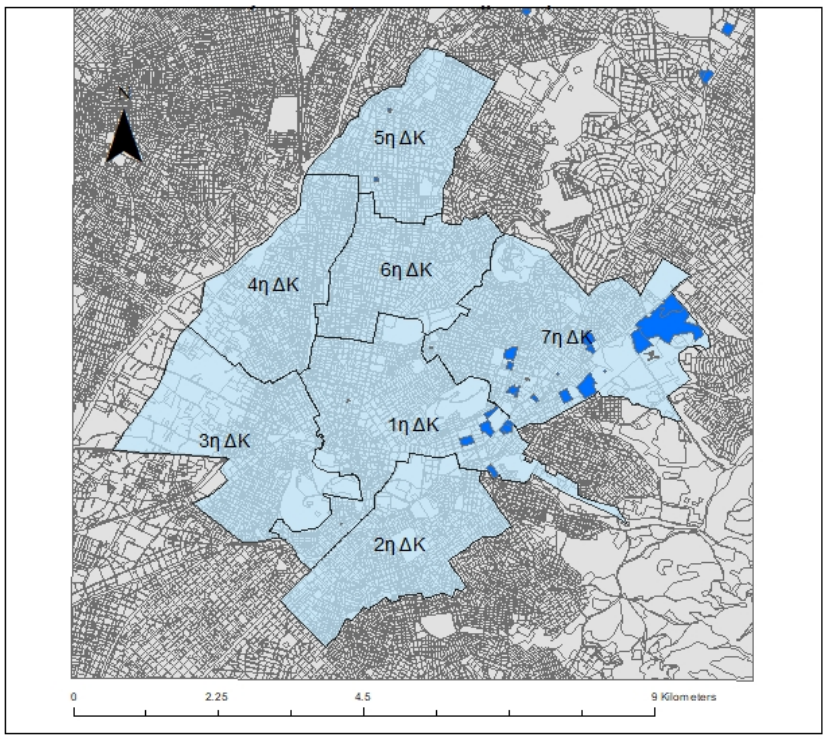
Πηγή:data.gov.gr (2018). Ιδία επεξεργασία.



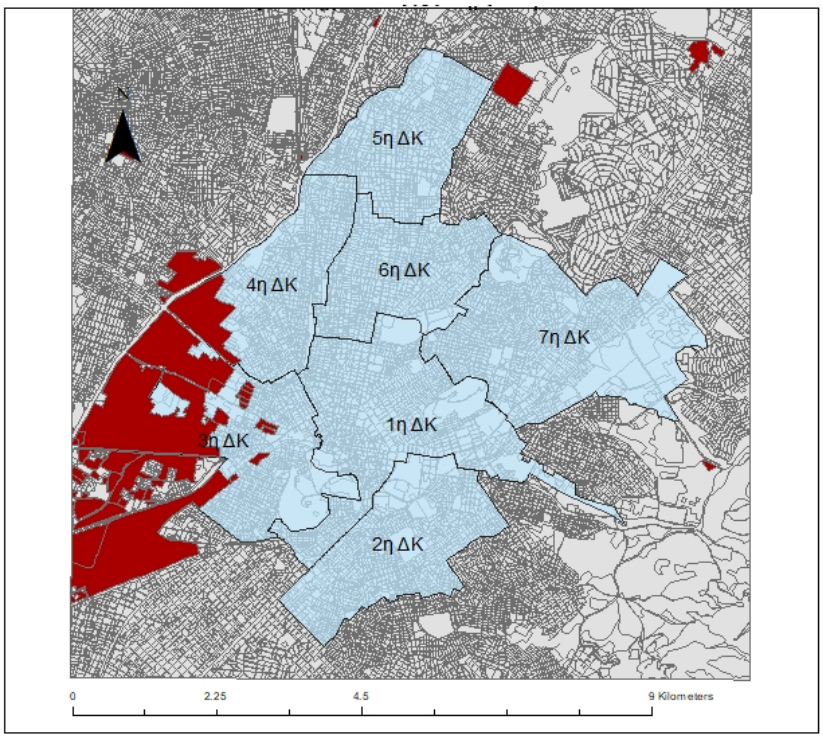
Χάρτης 4.9.Θέσεις σχολείων στο Δήμο Αθηναίων. Πηγή:data.gov.gr (2018). Ιδία επεξεργασία.



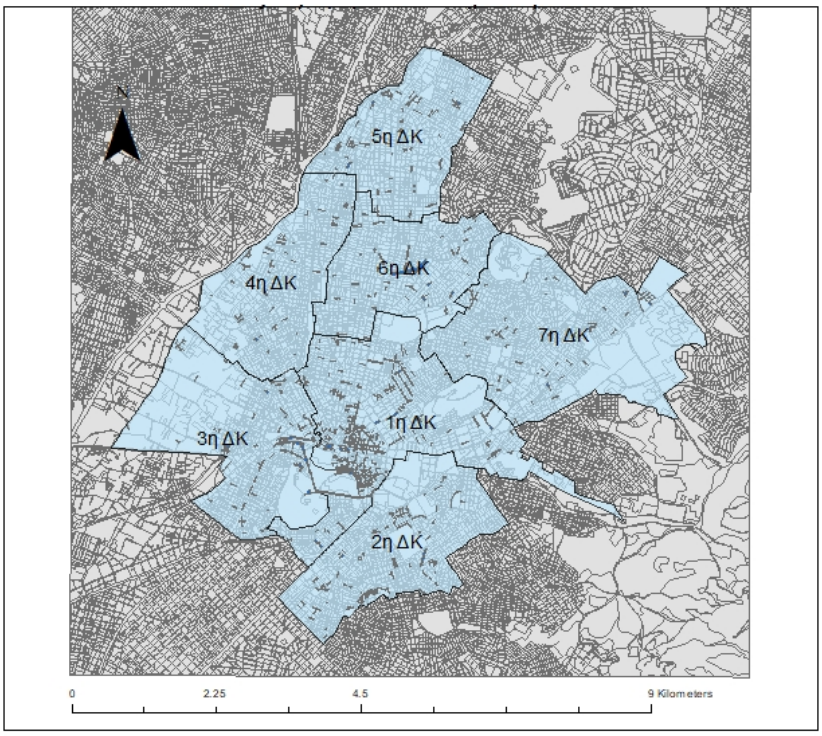
Χάρτης 4.10.Θέσεις αθλητικών εγκαταστάσεων στο Δήμο Αθηναίων. Πηγή:data.gov.gr (2018). Ιδία επεξεργασία.



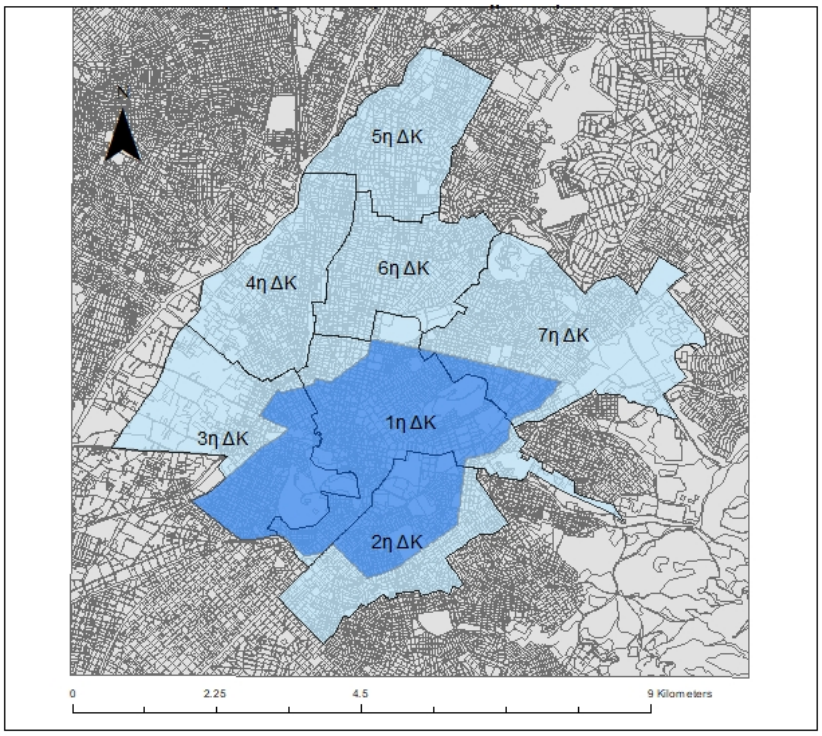
Χάρτης 4.11.Θέσεις νοσηλευτικών ιδρυμάτων στο Δήμο Αθηναίων. Πηγή:data.gov.gr (2018). Ιδία επεξεργασία.



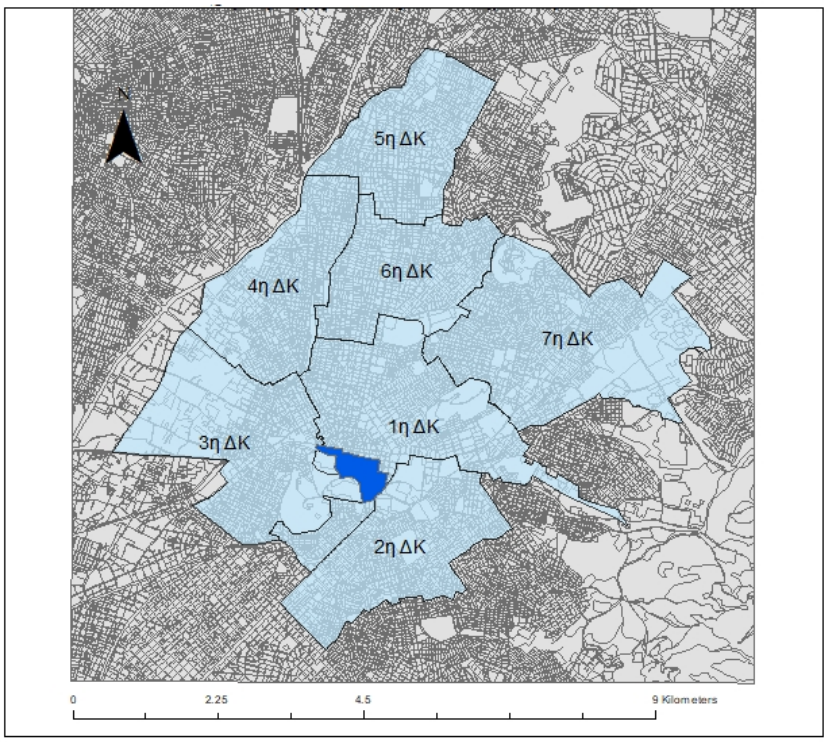
Χάρτης 4.12 Βιομηχανική – βιοτεχνική περιοχή στο Δήμο Αθηναίων. Πηγή:data.gov.gr (2018). Ιδία επεξεργασία.



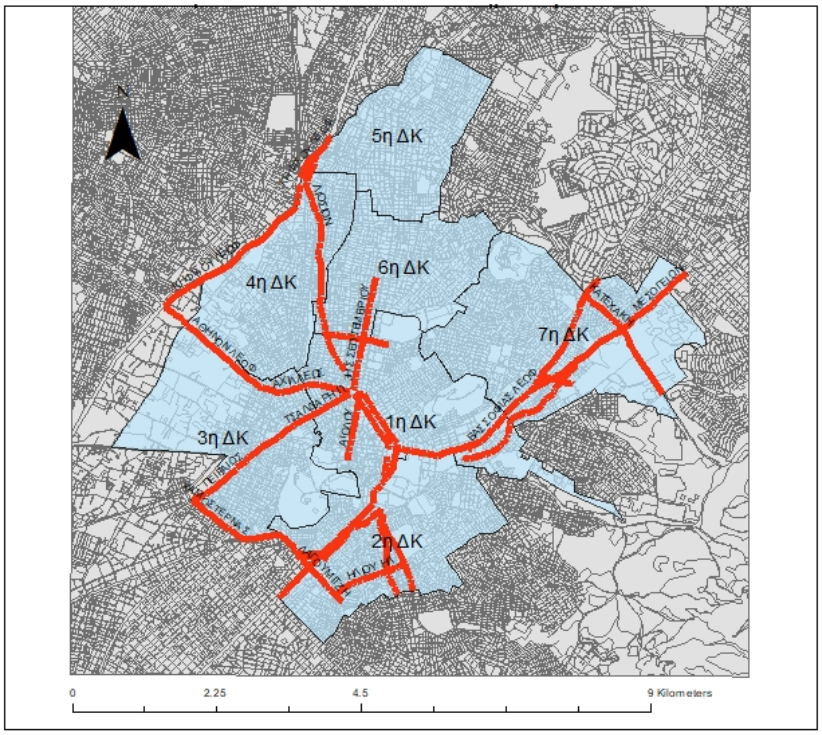
Χάρτης 4.13 Πεζόδρομοι στο Δήμο Αθηναίων. Πηγή:data.gov.gr (2018). Ιδία επεξεργασία.



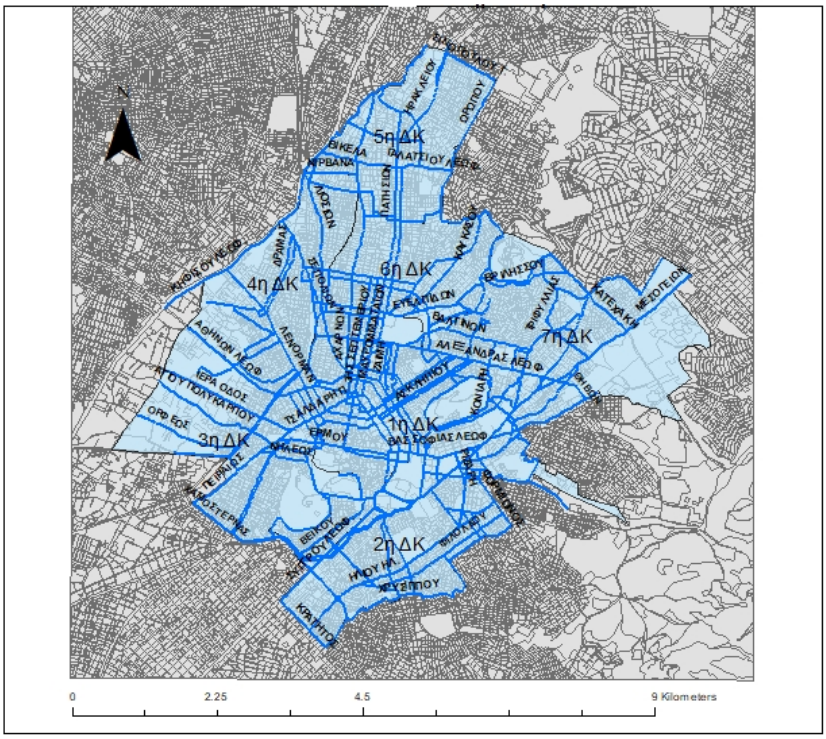
Χάρτης 4.14.Ιστορικό κέντρο στο Δήμο Αθηναίων. Πηγή:data.gov.gr (2018). Ιδία επεξεργασία.



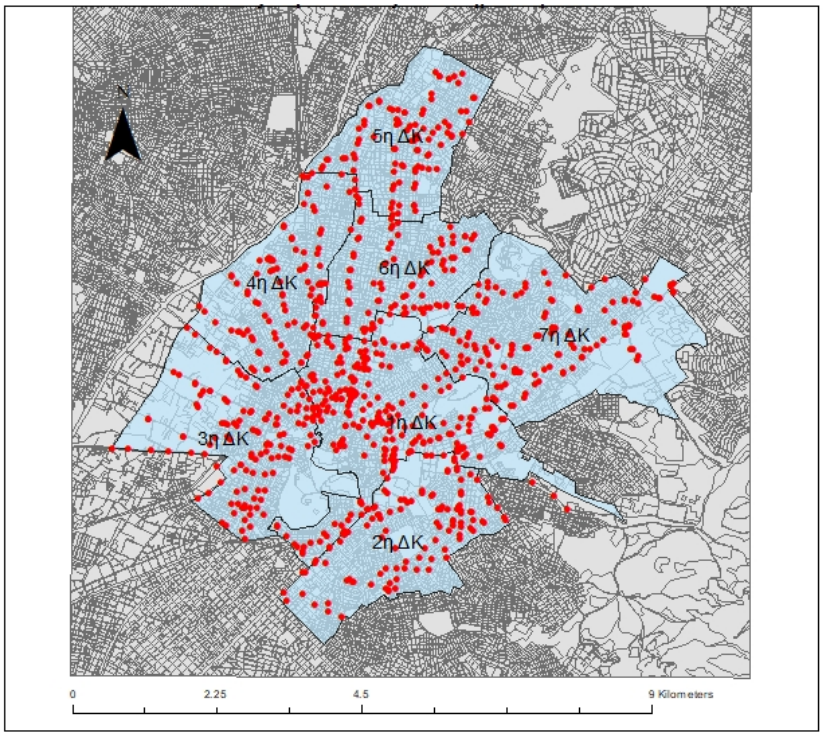
Χάρτης 4.15. Όρια περιοχής Πλάκας στο Δήμο Αθηναίων. Πηγή:data.gov.gr (2018). Ιδία επεξεργασία.



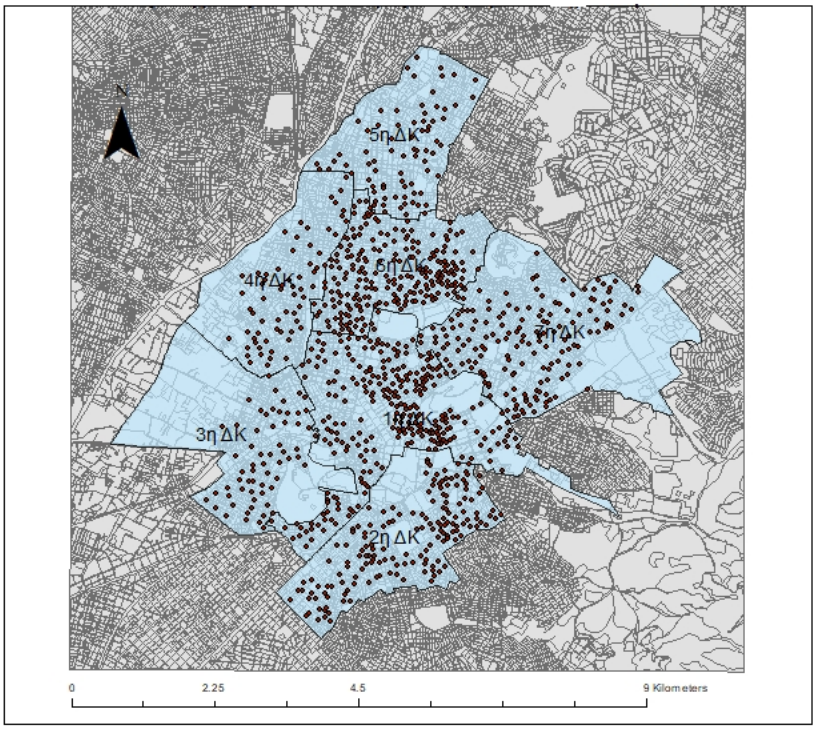
Χάρτης 4.16. Πρωτεύον οδικό δίκτυο στο Δήμο Αθηναίων. Πηγή:data.gov.gr (2018). Ιδία επεξεργασία.



Χάρτης 4.17. Βασικό οδικό δίκτυο στο Δήμο Αθηναίων. Πηγή:data.gov.gr (2018). Ιδία επεξεργασία.



Χάρτης 4.18. Θέσεις στάσεων συγκοινωνίας στο Δήμο Αθηναίων. Πηγή:data.gov.gr (2018). Ιδία επεξεργασία.



Χάρτης 4.19 Θέσεις δείγματος κατοικιών προς πώληση στο Δήμο Αθηναίων. Ιδία επεξεργασία.

Κεφάλαιο 5^{ο} – Διερεύνηση Μοντέλων Παλινδρόμησης

Ένα ορθό μοντέλο παλινδρόμησης είναι αυτό για το οποίο έχουν σαφώς προσδιοριστεί οι ανεξάρτητες μεταβλητές που επηρεάζουν το φαινόμενο που εξετάζεται και αποδίδεται με την μορφή της εξαρτημένης μεταβλητής. Καθώς η πολλαπλή γραμμική παλινδρόμηση είναι η διερεύνηση της συνδιακύμανσης των ανεξάρτητων μεταβλητών μεταξύ τους αλλά και με την εξαρτημένη, τέτοιου είδους ανάλυση αποδίδεται με πολύπλοκα μοντέλα, ενώ η ύπαρξη κατάλληλου θεωρητικού υπόβαθρου είναι απαραίτητη στην επιλογή των μεταβλητών που θα συμμετάσχουν τελικά στο μοντέλο .

5.1 Εισαγωγή Στοιχείων στο Λογισμικό SPSS

Η έναρξη της επεξεργασίας γίνεται με την δημιουργία ενός αρχείου .xls, η δομή του οποίου είναι κατάλληλη για τη μετέπειτα επεξεργασία στο λογισμικό SPSS. Όλα τα δεδομένα του δείγματος τοποθετούνται, ενώ οι τιμές τους ποσοτικοποιούνται όπως αναλυτικά θα αναφερθεί παρακάτω σε 58 στήλες του λογιστικού φύλλου. Η πρώτη έχει χαρακτηρισμό «ΑΡΙΘΜΟΣ» και αφορά τον αύξοντα αριθμό του κάθε ακίνητου. Η δεύτερη και τρίτη έχει την πληροφορία των συντεταγμένων «Χ»,»Υ» (Coord_Point), που προσδιορίζουν την θέση του κάθε ακινήτου σε σύστημα συντεταγμένων ΕΓΣΑ87. Η τέταρτη έχει χαρακτηρισμό «Γεωγραφικός Κωδικός» και είναι η πληροφορία που συνδέει την δημοτική κοινότητα με την περιοχή στην οποία βρίσκεται το ακίνητο μέσω ενός αριθμού. Οι στήλες αυτές που αναφέρθηκαν δεν αποτελούν πιθανές μεταβλητές, αλλά λειτουργούν επικουρικά στην επεξεργασία. Η πέμπτη στήλη έχει χαρακτηρισμό «Αξία Υ» και αφορά την εξαρτημένη μεταβλητή που αποτυπώνει την τιμή πώλησης του ακινήτου. Η έκτη «Χ1» έχει τον αριθμό της δημοτικής κοινότητας, η έβδομη «Χ2» έχει την ονομασία της περιοχής και η όγδοη «Χ3» την ονομασία της γειτονιάς. Και αυτές που αναφέρθηκαν (Χ1,Χ2, Χ3) δεν αποτελούν πιθανές μεταβλητές, αλλά λειτουργούν επικουρικά στην επεξεργασία. Όλες οι υπόλοιπες είναι οι ανεξάρτητες μεταβλητές, που είτε τα στοιχεία τους έχουν συλλεχθεί από την ιστοσελίδα πωλήσεων κατοικιών, είτε αποτυπώνουν τις αποστάσεις από άλλα σημεία ενδιαφέροντος (νοσοκομεία, συγκοινωνία κλπ.) η πληροφορία των οποίων έχει συλλεχθεί διαδικτυακά (Πίνακας 4.7).

5.2 Ορισμός και Κωδικοποίηση της Εξαρτημένης και των Ανεξάρτητων Μεταβλητών

Το πρώτο βήμα για την κατασκευή του προβλεπτικού μοντέλου γραμμικής παλινδρόμησης είναι ο ορισμός της εξαρτημένης μεταβλητής **Y** και των ανεξάρτητων μεταβλητών $\mathbf{X_i}$**.** Από τα στοιχεία που συλλέχθηκαν, η Y (Αξία) ορίζεται ως η εξαρτημένη μεταβλητή, ενώ οι πιθανές ανεξάρτητες μεταβλητές από την μεταβλητή X4 έως και την μεταβλητή X53.

Οι μεταβλητές συλλέχθηκαν επεξεργάστηκαν και επιλέχθηκαν σύμφωνα με την γενικότερη βιβλιογραφία για την κατασκευή ηδονικών μοντέλων όπως αυτά παρουσιάζονται αναλυτικότερα στο κεφάλαιο 2 του παρόντος.

Πιο συγκεκριμένα οι μεταβλητές θέσης που επιλέχθηκαν ως πιθανές ανεξάρτητες μεταβλητές συμμετοχής στο μοντέλο είναι οι: Απόσταση από ΑΤ (X11), Απόσταση από βρεφονηπιακό σταθμό (X13), Απόσταση από εμπόριο (X14), Απόσταση από ιστορικό κέντρο (X15), Απόσταση από νεκροταφείο (X16), Απόσταση από νοσοκομείο (X17), Είδος πλησιέστερης συγκοινωνίας (X18) Απόσταση από πλησιέστερη συγκοινωνία (X19), Απόσταση από ναό (X20), Απόσταση από σχολείο (X21), Απόσταση από αθλητική εγκατάσταση (X22), Απόσταση από πράσινο (X23), Απόσταση από περιοχή Πλάκας (X24), Απόσταση από μουσείο (X27), Ύπαρξη θέας (X51).

Από αυτές οι μικρές αποστάσεις στις μεταβλητές X11, X13, X14, X15, X17, X18, X19, X20, X21, X22, X23, X24, X27, X51 αναμένεται να έχουν θετική επίδραση στην εξαρτημένη μεταβλητή, είτε γιατί αφορούν κοντινή απόσταση από εξυπηρετήσεις όπως σχολική, υγειονομική, τουριστική, είτε γιατί αφορούν την ζητούμενη απ' όλους απόσταση από ασφάλεια, είτε γιατί αφορούν την απόσταση από το κέντρο της πόλης που έχει παρεμπίπτοντος και αρχαιολογικό ενδιαφέρον. Η ύπαρξη θέας (X51) και η κοντινή απόσταση από συγκοινωνία (X19) αναμένεται να έχουν και αυτές θετική επίδραση στο μοντέλο.

Οι δομικές μεταβλητές που επιλέχθηκαν ως πιθανές ανεξάρτητες μεταβλητές συμμετοχής στο μοντέλο είναι : Εμβαδόν κατοικίας (X32), Αριθμός υπνοδωματίων (X34), Αριθμός μπάνιων (X35), Ύπαρξη WC (X36), Αποθήκη (X37), Όροφος (X38), Ηλικία κατασκευής (X40), Είδος κατοικίας (X41), Κατάσταση κατασκευής (X42), Ύπαρξη ανακαίνισης (X43), Ύπαρξη βεράντας (X44), Γωνιακό ή διαμπερές (X45), Ύπαρξη ασανσέρ (X46) , Ύπαρξη θέσης στάθμευσης (X47), Είδος θέρμανσης (X48), Είδος καυσίμου θέρμανσης (X49), Ύπαρξη κλιματισμού (X50), Ύπαρξη τζακιού

(Χ52). Από αυτές αναμένεται να έχουν θετική επίδραση οι μεγάλες τιμές των Χ32, Χ34, Χ35, Χ38, ενώ αναμένεται να έχουν θετική επίδραση στη εξαρτημένη μεταβλητή οι θετικές τιμές των Χ36, Χ37, Χ42, Χ43, Χ44, Χ45, Χ46, Χ47,Χ50, Χ52.

Από τι μεταβλητές γειτονιάς οι υψηλές τιμές των μεταβλητών: Μέσο εισόδημα κατοίκων (Χ30), τμ Πρασίνου ανά κάτοικο (Χ10), Ποσοστό πρασίνου (Χ8), Ποσοστό ελλήνων (Χ4), αναμένεται να έχουν θετική επίδραση στη εξαρτημένη μεταβλητή. Από την άλλη μεριά οι χαμηλές τιμές των μεταβλητών: Ποσοστό επιμερισμού κατοίκων ανά ΔΚ (Χ9), Ποσοστό κενών κατοικιών (Χ7), Ποσοστό ανεργίας (Χ6), Ποσοστό αλλοδαπών (Χ5) αναμένεται να έχουν θετική επίδραση στην εξαρτημένη μεταβλητή. Επιπρόσθετα αρνητική επίδραση αναμένεται να έχει η χαμηλή τιμή της μεταβλητής Απόσταση από βιομηχανία- βιοτεχνία (Χ12), παρόμοια επίδραση η υψηλή τιμή της μεταβλητής Απόσταση από ΟΣΕ (Χ26) και Απόσταση από πρωτεύων οδικό δίκτυο (Χ28). Αντίθετα θετική επίδραση στο μοντέλο αναμένεται να έχει η χαμηλή τιμή της μεταβλητής Απόσταση από βασικό οδικό δίκτυο (Χ29) και Απόσταση από πεζόδρομο (Χ25). Αναλυτικότερη αναφορά σε σχέση με την βιβλιογραφία γίνεται στο κεφάλαιο 2 της παρούσας διατριβής.

5.3 Περιγραφή Δεδομένων, Μέτρα Κεντρικής Τάσης- Διασποράς

5.3.1 Περιγραφική στατιστική

Όπως αναφέρθηκε στο κεφάλαιο 3.1.4 τα κυριότερα μέτρα περιγραφικής στατιστικής είναι τα :

Μέτρα θέσης:

- Μέση τιμή (mean). Γύρω από αυτή βρίσκονται συγκεντρωμένες οι τιμές του δείγματος.
- Διάμεσος(median). Η μεσαία τιμή του δείγματος.
- Κορυφή (mode). Η μέτρηση με την μεγαλύτερη συχνότητα σ' ένα δείγμα.
- Πρώτο (first quantile) και τρίτο τεταρτημόριο (third quartile). Η τιμή του δείγματος για την οποία ισχύει ότι το 25% και 75% αντίστοιχα των τιμών του είναι μικρότερες ή ίσες με αυτή.

Μέτρα διασποράς:

- Διασπορά (variance). Η διασπορά των τιμών του δείγματος γύρω από την μέση τιμή του.
- Τυπική απόκλιση (standard deviation). Η απόκλιση των μετρήσεων από την

μέση τιμή.

- Τυπική απόκλιση μέσου (standard error of mean). Το αποτέλεσμα της διαίρεσης της τυπικής απόκλισης προς την ρίζα του μεγέθους του δείγματος.
- Η ελάχιστη και η μέγιστη τιμή.
- Το ενδοτεταρτημοριακό εύρος (interquartile range). Η διαφορά τιμών του τρίτου με το πρώτο τεταρτημόριο.

Μέτρα σχήματος κατανομής:

- Ο συντελεστής ασυμμετρίας (skewness) και ο συντελεστής κυρτότητας (kurtosis). Δείκτες που χρησιμοποιούνται για το σχήμα της κατανομής του δείγματος.

Η οπτική παρουσίαση των στατιστικών δεδομένων γίνεται μέσα από ραβδογράμματα (bar charts) και κυκλικά διαγράμματα (pie charts) για ποσοτικές (scale) μεταβλητές και μέσα από ιστογράμματα (histograms) και θηκογράμματα (boxplots) για ποιοτικές μεταβλητές (nominal) καθώς και με χάρτες .

5.3.2 Κατανομές Μεταβλητών, Περιγραφή Δεδομένων

5.3.2.1 Υπολογισμός Στατιστικών Μέτρων Ποσοτικών Μεταβλητών

Για μεγάλο αριθμό δείγματος είναι πρακτικά δύσκολο να εξαχθούν ασφαλή συμπεράσματα από τις ποσοτικές μεταβλητές που θα εξυπηρετήσουν την μελέτη χωρίς την χρήση στατιστικών μεθόδων. Απαραίτητο λοιπόν είναι να χρησιμοποιηθούν τρόποι περιληπτικής περιγραφής των δεδομένων του δείγματος και των διαφοροποιήσεων των τιμών της κάθε μεταβλητής. Η διαφοροποίηση αυτή είναι η κατανομή της μεταβλητής η οποία περιγράφεται με τις στατιστικές μεθόδους (Ηλιοπούλου, 2017).

Στο παράρτημα «Α1» παρουσιάζονται αναλυτικά τα αποτελέσματα των στατιστικών μέτρων των ποσοτικών μεταβλητών του δείγματος, ενώ στο κείμενο αναφέροντα περιληπτικά σημαντικά σημεία των αποτελεσμάτων.

Πιο συγκεκριμένα εμφανίζονται οι τιμές προς πώληση των ακινήτων του δείγματος να έχουν μέση τιμή 186833,20 ευρώ, με μέγιστη τιμή τα 4450000 και ελάχιστη τα 4200 ευρώ. Όσο αφορά τα εμβαδά των κατοικιών του δείγματος η μέση τιμή είναι 106,11 τμ και η διάμεσος τα 87τμ με ελάχιστο τα 17τμ και μέγιστο τα 1520 τμ. Η μέση τιμή ανά τετραγωνικό μέτρο στο δείγμα είναι 1572,54 ευρώ, ενώ η μέτρηση με την μεγαλύτερη συχνότητα (mode) είναι τα 1000 ευρώ/τμ .Όσο αφορά τον αριθμό ημερών που τα ακίνητα διατίθενται προς πώληση η μέση τιμή είναι 359,64 ημέρες.

Ο αριθμός των υπνοδωματίων είναι συγκεντρωμένος γύρω από τα 2 (mean=2,03), με ελάχιστα υπνοδωμάτια το 1 και μέγιστο αριθμό τα 7. Η μέτρηση με την μεγαλύτερη συχνότητα στον αριθμό μπάνιων είναι 1 και η μεσαία τιμή του δείγματος για τον αριθμό ορόφου που βρίσκονται οι προς πώληση κατοικίες του δείγματος είναι ο δεύτερος (median=2). Η μέση τιμή που εκφράζει την ηλικία κατασκευής των κατοικιών του δείγματος είναι τα 42,10 έτη, ενώ έχουμε κατοικίες ακόμα και 118 ετών προς πώληση.

Σε ότι αφορά τα ποσοστά των κατοικιών, ανάλογα με τα χαρακτηριστικά της γειτονιάς, ως προς τους αλλοδαπούς και έλληνες κατοίκους η μέση τιμή (mean) είναι 43,06% και 56,94% και αντίστοιχα, ως προς το ποσοστό ανεργίας η διάμεσος είναι 21%, το ποσοστό πρασίνου έχει μέση τιμή 16,05%, ενώ τα τετραγωνικά μέτρα πρασίνου ανά κάτοικο έχουν μέση τιμή 9,51τμ. Η απόσταση των κατοικιών του δείγματος από τα αστυνομικά τμήματα της πόλης των Αθηνών έχουν μέση τιμή 693,63μ, ενώ μεγαλύτερη είναι η μέση τιμή απόστασης από την πλησιέστερη βιομηχανία – βιοτεχνία στα 2067,82μ. Παράλληλα η απόσταση με την μεγαλύτερα εμφανιζόμενη συχνότητα (mode) από το πλησιέστερο νοσοκομείο είναι 663μ. Η μέση απόσταση από εμπορική δραστηριότητα είναι 1177,57μ και από νεκροταφείο είναι 2171,84μ, ενώ αντίστοιχα η μεσαία τιμή (median) είναι 1010,82μ και 2305,44μ.

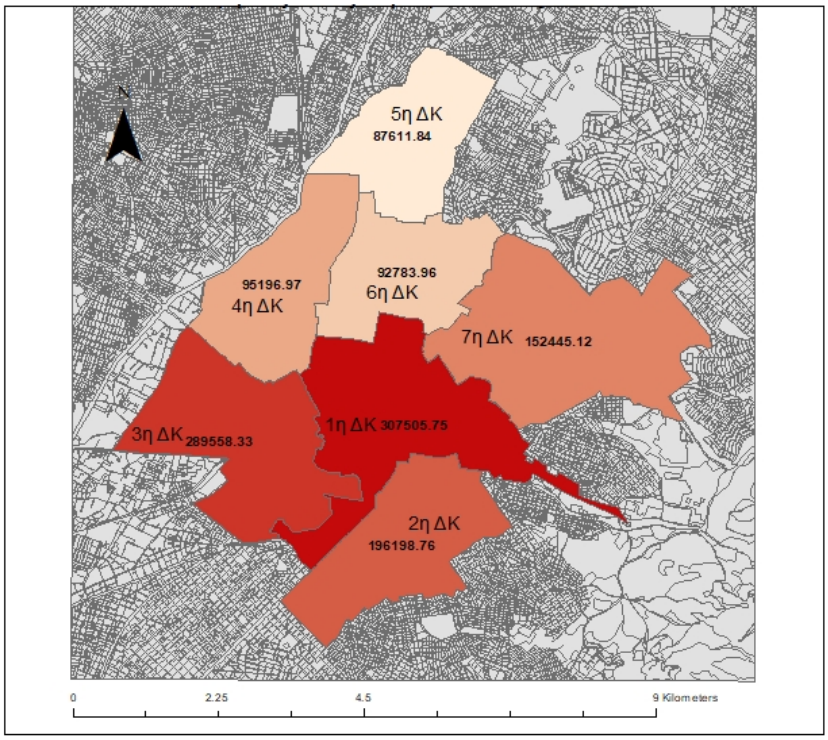
Η μεγαλύτερη απόσταση που παρατηρείται από βρεφονηπιακό σταθμό είναι 922μ, η μέση τιμή απόστασης είναι τα 349,67μ, ενώ η μέση τιμή απόστασης από σχολείο είναι τα 205,13μ. Αντίστοιχα η διάμεσος τιμή απόστασης από αθλητικές δραστηριότητες είναι τα 481,87μ, ενώ η μεγαλύτερη απόσταση από χώρο πρασίνου είναι τα 570μ με τις αποστάσεις να έχουν διάμεσο τα 170,70μ. Η μέση τιμή της απόσταση από το ιστορικό κέντρο της Αθήνας είναι τα 676,48μ, με την μέγιστη να είναι τα 4426μ, η μέγιστη απόσταση από την περιοχή της Πλάκας είναι τα 6144μ, η διάμεσος της απόστασης από θεσμοθετημένο πεζόδρομο είναι τα 105,92μ, ενώ η μεγαλύτερη απόσταση από μουσείο είναι τα 3272μ, με τη μέση τιμή να είναι τα 929,93μ. Η μέγιστη απόσταση από την πλησιέστερη συγκοινωνία είναι τα 439μ, η μέση τιμή απόστασης από τον σταθμό του ΟΣΕ είναι τα 1428,47μ, ενώ οι μέσες τιμές αποστάσεων από το πρωτεύων οδικό δίκτυο και το βασικό οδικό δίκτυο συγκεντρώνονται στα 496,38μ και 80,06μ αντίστοιχα.

Για τον υπολογισμό του αριθμητικού μέσου ανά γεωγραφική περιοχή χρησιμοποιούμε τις μεταβλητές για τις οποίες θέλουμε να ελέγξουμε τη μέση τιμή που παρουσιάζουν σε κάθε δημοτική κοινότηταα του Δήμου Αθηναίων. Στο

παράρτημα «Α2» παρουσιάζονται αναλυτικά τα αποτελέσματα του αριθμητικού μέσου ανά γεωγραφική περιοχή των μεταβλητών του δείγματος ενώ στο κείμενο αναφέροντα περιληπτικά σημαντικά σημεία των αποτελεσμάτων. Παράλληλα παρατίθενται χάρτες (από 5.1 έως 5.4) με τους αριθμητικούς μέσους αξίας κατοικίας, τιμή / τμ, εμβαδόν και ηλικία κατοικίας.

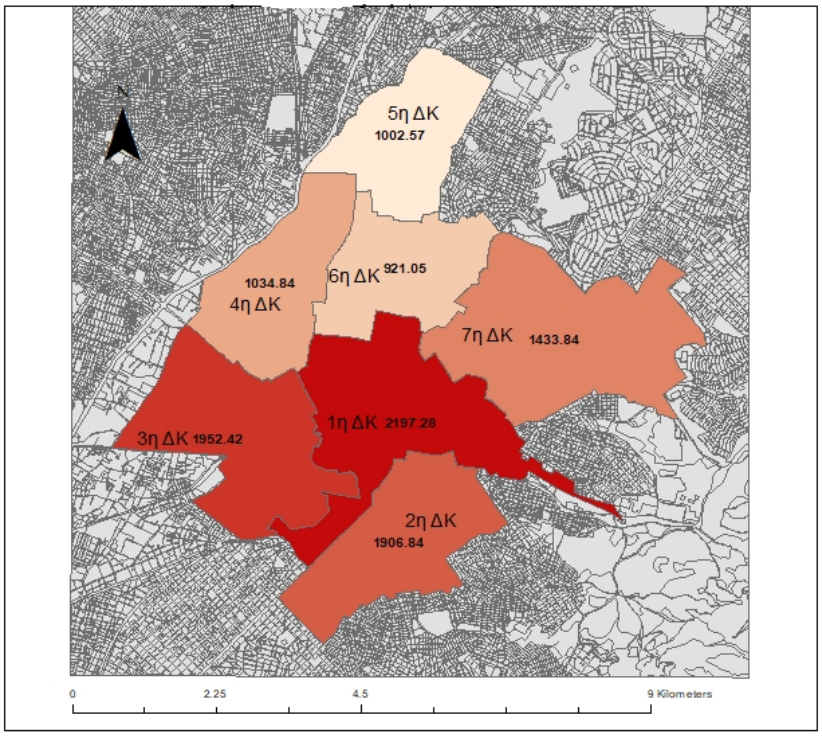
Χάρτης 5.1

Αριθμητικός μέσος στην αξία κατακοικίας ανά Δ.Κ



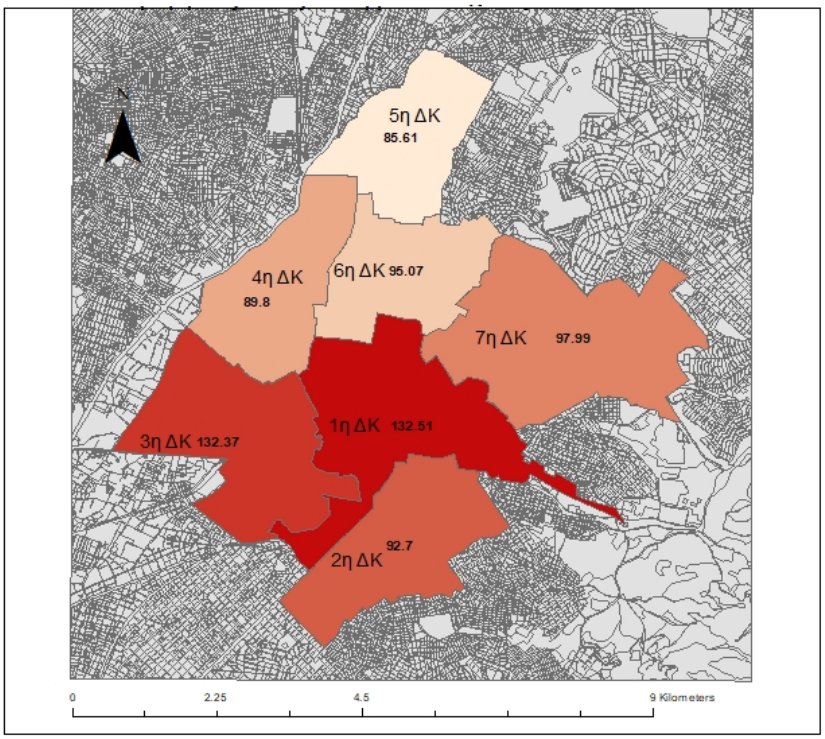
Χάρτης 5.2

Αριθμητικός μέσος τιμής /τμ σε ευρώ ανά Δ.Κ



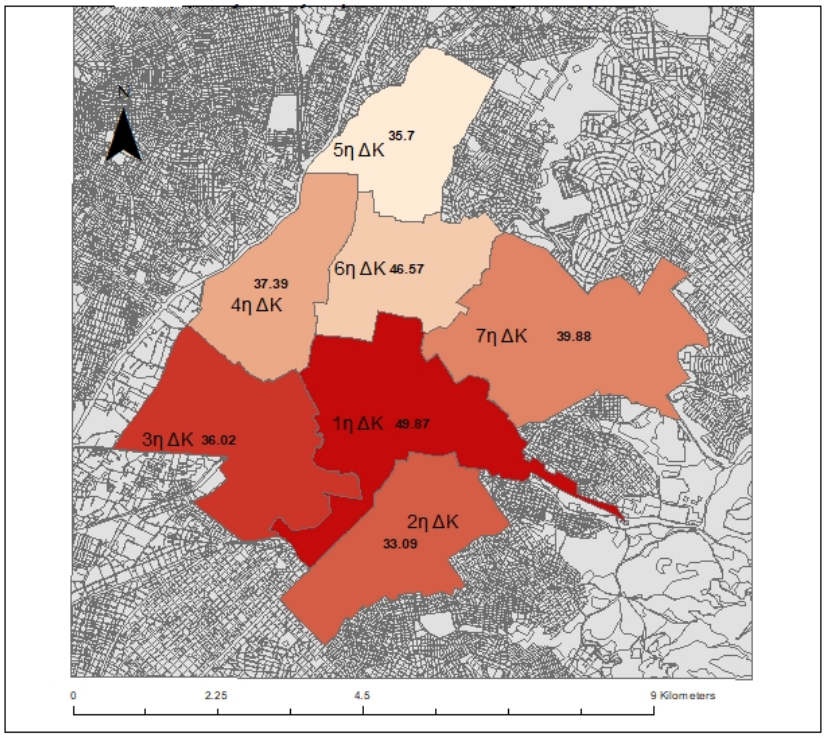
Χάρτης 5.3

Αριθμητικός μέσος στο εμβαδόν κατοικίας ανά Δ.Κ



Χάρτης 5.4

Αριθμητικός μέσος στην ηλικία σε έτη ανά Δ.Κ



Παρατηρούμε ότι οι πιο ακριβές κατοικίες του δείγματος είναι στην πρώτη δημοτική κοινότητα με μέση τιμή τα 307505,75 ευρώ, ενώ έπεται η τρίτη δημοτική κοινότητα με 289558,33 ευρώ. Ανάλογα αποτελέσματα έχουμε και στις τιμές ανά τμ με 2197,28 ευρώ και 1952,42 ευρώ αντίστοιχα καθώς και στα εμβαδά των προς πώληση κατοικιών του δείγματος με 132,51τμ και 132,.37τμ αντίστοιχα. Τέλος οι μεγαλύτερες ηλικίες κατοικιών είναι στην πρώτη και έκτη δημοτική κοινότητα με 49,87 έτη και 46,57 έτη αντίστοιχα.

Δεν παρουσιάζεται κάποια ιδιαίτερη διαφοροποίηση όσο αφορά των αριθμό των ημερών προς διάθεση του ακινήτου. Το μέγιστο ποσοστό των ελλήνων κατοίκων 69% είναι στην τρίτη δημοτική ενότητα, ενώ το μικρότερο ποσοστό ανεργίας παρουσιάζεται στην 7 δημοτική κοινότητα με 17% . Το μέγιστο ποσοστό κενών κατοικιών παρουσιάζεται στην πρώτη και έκτη δημοτική με 36%, ενώ το μέγιστο ποσοστό πρασίνου στην έβδομη δημοτική κοινότητα στο 30%.

Το μέγιστο μέσο ποσοστό κατοίκων ανά δημοτική κοινότητα αφορά την έκτη με 21%, ενώ το μέγιστο μέσο ποσοστό πρασίνου ανά κάτοικο είναι στην τρίτη δημοτική κοινότητα με ποσοστό 21% . Η ελάχιστη μέση απόσταση από αθλητική εγκατάσταση είναι στην πέμπτη δημοτική κοινότητα με 311,48μ.

5.3.2.2 Υπολογισμός Κατανομών Συχνοτήτων Ποσοτικών Μεταβλητών

Η κατανομή συχνοτήτων περιγράφει μια μεταβλητή παρουσιάζοντας την συχνότητα εμφάνισης των τιμών της στα δεδομένα. Καθώς συνήθως ο αριθμός των τιμών των συνεχών ποσοτικών μεταβλητών είναι αρκετά μεγάλος, κρίνεται απαραίτητη η ομαδοποίηση των τιμών σε κλάσεις (τάξεις) οι οποίες πρέπει να είναι αποκλειστικές και εξαντλητικές (Ηλιοπούλου, 2017).

Στο παράρτημα «Α3» παρουσιάζονται αναλυτικά τα αποτελέσματα κατανομών συχνοτήτων των ποσοτικών μεταβλητών του δείγματος.

Ο αριθμός των δύο υπνοδωματίων εμφανίζεται συχνότερα στο δείγμα μας (457 φορές) σε ποσοστό 45,7% ενώ ακολουθούν οι κατοικίες με ένα υπνοδωμάτιο σε ποσοστό 31%. Η πλειονότητα των κατοικιών σε ποσοστό 90% έχουν ένα μπάνιο. Τα περισσότερα διαμερίσματα προς πώληση είναι στο ισόγειο και πρώτο όροφο σε ποσοστό 19,9% και 20,7% αντίστοιχα Η πλειονότητα των κατοικιών του δείγματος σε ποσοστό 49,7% έχει τιμή προς πώληση έως 100.000 ευρώ, ενώ μόνο το 2,3% έχει τιμή μεγαλύτερη από 1000000 ευρώ. Συγκεκριμένα Από 100001 έως 200000 ευρώ είναι το 26,8%, από 200001 έως 300000 ευρώ είναι το 11,4%, από 300001 έως 500000 ευρώ είναι το 5,7%, από 500001 έως 1000000 ευρώ είναι το 4,1% και από 1000001 έως 4500000 ευρώ είναι το 2,3%. Η πλειονότητα κατοικιών του δείγματος σε ποσοστό 50,1% έχει εμβαδόν από 51 έως 100τμ, ενώ το 29,6% έχει εμβαδόν από 101 έως 200τμ. Παράλληλα τα μικρά διαμερίσματα (γκαρσονιέρες) έως 30τμ αποτελούν το 2,4% του δείγματος. Η πλειονότητα των κατοικιών του δείγματος είναι στην αγορά προς πώληση από 0 έως 500 ημέρες σε ποσοστό 77,7% (βλέπε πίνακα Α3). Τα περισσότερα σπίτια είναι ηλικίας από 41 έως 60 έτη σε ποσοστό 52,9%, ενώ συνολικά πάνω από το 62,9% των κατοικιών του δείγματος έχουν ηλικία πάνω από 40 έτη, γεγονός που υποδεικνύει την γήρανση της πόλης. Οι περισσότερες κατοικίες έχουν πολύ κοντά τους σχολείο και μάλιστα σε απόσταση έως 100μ έχει το 21,5% και από 101μ έως 350μ έχει το 67,3% του δείγματος. Οι περισσότερες κατοικίες έχουν σε κοντινή απόσταση κάποιας μορφή αθλητικής εγκατάσταση σε ποσοστό 52,2% για απόσταση έως 500μ.

Σε απόσταση έως 100μ από ναό βρίσκεται το 15,8% του δείγματος, ενώ από 101μ έως 600μ έχουν πλησίον τους ναό το 84% των κατοικιών του δείγματος. Σε απόσταση χιλιομέτρου πάνω από το 65% των κατοικιών έχουν κάποιο νοσοκομείο, γεγονός που υποδεικνύει μια άρτια οργανωμένη πόλη. Πιο συγκεκριμένα έως 500μ έχει το 28,3% του δείγματος, ενώ σε ποσοστό 37,7% του δείγματος των κατοικιών

βρίσκεται σε απόσταση από 501μ έως 1000μ. Σε πολύ κοντινή απόσταση βρίσκεται και η συγκοινωνία για την πλειονότητα των προς πώληση κατοικιών. Πιο συγκεκριμένα το 16,7% έχει συγκοινωνία σε απόσταση έως 50μ, το 30,5% έχει πλησίον του συγκοινωνία σε απόσταση από 51 έως 100μ, ενώ το 49,9% έχει πλησίον του συγκοινωνία σε απόσταση από 101 έως 300μ. Κοντινή απόσταση και για τα αστυνομικά τμήματα αφού πάνω από το 80,7% των προς πώληση κατοικιών έχουν σε απόσταση χιλιόμετρου αστυνομικό τμήμα. Πιο συγκεκριμένα έως 100μ έχει το 2,1% ενώ το 30,9% και το 47,6% έχει αστυνομικό τμήμα σε απόσταση από 101μ έως 500μ και από 501μ έως 1000μ αντίστοιχα. Η πλειονότητα των κατοικιών είναι μακριά από βιομηχανική και βιοτεχνική δραστηριότητα, αφού μόνο το 6,7% βρίσκεται σε απόσταση έως 500μ. Αρκετά μεγάλος αριθμός κατοικιών βρίσκεται κοντά σε εμπορική δραστηριότητα αφού έως 500μ έχουν κοντά τους εμπορική δραστηριότητα το 19,5% των κατοικιών του δείγματος, ενώ από 501 έως 1000μ έχει κοντά του εμπορική δραστηριότητα το 29,5%. Σε απόσταση χιλιομέτρου βρίσκεται η πλειονότητα των κατοικιών από βρεφονηπιακό σταθμό. Πιο συγκεκριμένα το 6,1% έχει σταθμό σε απόσταση έως 100μ, ενώ το 73,2% έχει κοντά του και σε απόσταση από 101μ έως 500μ. Ποσοστό 85,9% των προς πώληση κατοικιών του δείγματος έχουν κοντά τους πράσινο. Το 10,5% είναι σε απόσταση έως 50μ, από 51μ έως 100μ είναι το 15,6% των κατοικιών του δείγματος ενώ το 59% έχει κοντά του πράσινο σε απόσταση από 101μ έως 300μ.

Ποσοστό 38% των προς πώληση κατοικιών βρίσκονται εντός του ιστορικού κέντρου, ενώ το 18% βρίσκεται σε απόσταση έως 500μ από τα όριά του. Ποσοστό 1,5% των προς πώληση κατοικιών του δείγματος βρίσκονται εντός της περιοχής της Πλάκας, ενώ το 4,6% βρίσκεται σε απόσταση από τα όρια της έως τα 500μ. Το 6,4% των προς πώληση κατοικιών του δείγματος έχει πρόσωπο σε πρωτεύων οδικό δίκτυο. Το 23,8% έχει επαφή με θεσμοθετημένο πεζόδρομο, το 41,9% των κατοικιών του δείγματος έχει επαφή με το βασικό οδικό δίκτυο. Σε απόσταση έως 100μ το 2,8% των κατοικιών του δείγματος έχει κοντά του μουσείο. Το 4,5% του δείγματος βρίσκεται κοντά σε σταθμό του ΟΣΕ. Πάνω από το 44% των προς πώληση κατοικιών του δείγματος έχουν τιμή έως 1000 ευρώ ανά τμ . Συγκεκριμένα έως 100 ευρώ/τμ πωλείται το 0,3%, από 100 έως 500 ευρώ το 10,4%, από 501 έως 1000 ευρώ το 33,4%, από 1001 έως 2500 ευρώ το 41,2%, από 2501 έως 4500 ευρώ το 11%, από 4501 έως 10000 ευρώ το 3,1% και από 10001 έως 32000 ευρώ το 0,6% του δείγματος. Τέλος 78% των κατοικιών του δείγματος έχουν τιμή ζώνης έως 1500

ευρώ, ενώ από 901 έως 1500 ευρώ είναι το 75,9% του δείγματος.

5.3.2.3 Υπολογισμός Κατανομών Συχνοτήτων Ποιοτικών Μεταβλητών

Στην περίπτωση των ποιοτικών δεδομένων δεν απαιτείται ιδιαίτερη προετοιμασία αυτών, καθώς περιλαμβάνουν μικρό εύρος τιμών ή είναι λεκτικά αποδιδόμενες. Η κατασκευή του πίνακα κατανομής αποδίδει την καταμέτρηση της συχνότητας εμφάνισης κάθε τιμής της, παρουσιάζοντας την απόλυτη, τη σχετική και την αθροιστική συχνότητα. Στο παράρτημα «Α4» παρουσιάζονται αναλυτικά τα αποτελέσματα κατανομών συχνοτήτων ποιοτικών μεταβλητών του δείγματος.

Το λεωφορείο εμφανίζεται ως η συνηθέστερη μορφή συγκοινωνίας πλησίον των κατοικιών του δείγματος σε ποσοστό 95,4%. Τα εισοδήματα από 12970 έως 18587 ευρώ είναι η συνηθέστερη κατηγορία που εμφανίζονται ως μέσο εισόδημα κατοίκων σε ποσοστό πάνω από 50%. Πιο συγκεκριμένα από 12970 έως 18587 ευρώ είναι το εισόδημα του 54,7% των κατοίκων, ενώ το 35,5% έχει εισόδημα από 18587 έως 31973 ευρώ. Η πλειονότητα των προς πώληση κατοικιών είναι διαμερίσματα και μάλιστα σε ποσοστό 90,3%. Πάνω από το 47% των κατοικιών του δείγματος είναι σε πολύ καλή κατάσταση. Το 52,9% του δείγματος δεν είναι ανακαινισμένες κατοικίες. Τα περισσότερα (ποσοστό 83,3%) δεν έχουν WC. Τα περισσότερα δεν έχουν αποθήκη (ποσοστό 75,1%), γεγονός που προφανώς σχετίζεται με τις γερασμένες κατασκευές. Το ίδιο ισχύει και για τις θέσεις στάθμευσης καθώς οι περισσότερες κατοικίες είναι χωρίς θέση στάθμευσης (ποσοστό 83,9%). Η πλειονότητα των κατοικιών του δείγματος δεν είναι γωνιακές ή διαμπερείς, σε ποσοστό 89%. Σχεδόν το σύνολο των κατοικιών έχουν βεράντα σε ποσοστό 93,5%. Τα περισσότερα δεν έχουν κλιματισμό σε ποσοστό 89,6%. Σχεδόν τα μισά έχουν κεντρική θέρμανση πετρελαίου (ποσοστό 48,6%). Η πλειονότητα δεν έχει θέα (ποσοστό 80,7%) ούτε τζάκι (ποσοστό 93,9%), ενώ η πλειονότητα έχει ασανσέρ (ποσοστό 81,1%).

5.4 Δημιουργία και Διερεύνηση Μοντέλων

5.4.1 Δημιουργία Πρώτου Μοντέλου

Στο πρώτο μοντέλο της παρούσας εργασίας ακολουθείται η διαδικασία της δημιουργίας εξίσωσης παλινδρόμησης χωρίς να αφαιρεθούν οι ακραίες οντότητες του δείγματος. Οι ανεξάρτητες μεταβλητές που λαμβάνουν μέρος σε πρώτη φάση είναι αυτές που δεν εμφανίζουν συσχέτιση μεταξύ τους, απομακρύνοντας αυτές που εμφανίζουν μικρότερη συσχέτιση με την εξαρτημένη. Εισάγονται οι μεταβλητές στο

λογισμικό SPSS με την διαδικασία enter (απλή μαζική εισαγωγή) και ακολουθεί η αφαίρεση αυτών των μεταβλητών που δεν μεταβάλουν σημαντικά τον συντελεστή R^2. Υπολογίζεται το τελικό μοντέλο και η διαδικασία συνεχίζεται με την εξέταση του παραχθέντος μοντέλου μέσω του λογισμικού Arc Gis και της OLS διαδικασίας του και μετέπειτα την εξέταση του μοντέλου μέσω της γεωγραφικά σταθμισμένης παλινδρόμησης.

5.4.1.1 Προκαταρτικοί Έλεγχοι

Βασικές προϋποθέσεις για την χρήση της στατιστικής συμπερασματολογίας είναι η τυχαιότητα του δείγματος, η μη ύπαρξη ακραίων τιμών στις παρατηρήσεις μας, ιδιαίτερα όταν έχουμε μεγάλο αριθμό παρατηρήσεων και η κανονική κατανομή του πληθυσμού από τον οποίο προήλθε το δείγμα μας. Φυσικά και σύμφωνα με το Κεντρικό Οριακό Θεώρημα, δηλαδή ότι το άθροισμα και επομένως η μέση τιμή μεγάλου αριθμού ανεξάρτητων παρατηρήσεων ακολουθεί κανονική κατανομή, ανεξάρτητα από το ποια κατανομή ακολουθούν οι παρατηρήσεις, θα μπορούσε να παραληφθεί αυτό ο έλεγχος (Παπαδόπουλος, χ.η). Επιπρόσθετα η χωρική αυτοσυσχέτιση παραβιάζει ορισμένες παραδοχές των στατιστικών ελέγχων, όπως η υπόθεση τυχαιότητας για την εκτίμηση των παραμέτρων, ενώ συμβάλει στον πλεονασμό στα δεδομένα. Τέλος σε περίπτωση ύπαρξης ακραίων τιμών, ο μη παραμετρικός συντελεστής Spearman χρησιμοποιείται για την εύρεση συσχέτισης μεταβλητών ακόμα και όταν υπάρχουν στο δείγμα μας ακραίες τιμές (Ηλιοπούλου, 2017).

5.4.1.1.1 Έλεγχος Τυχαιότητας Παρατηρήσεων.

Καθώς οι στατιστικοί έλεγχοι προϋποθέτουν ότι η σειρά των παρατηρήσεων στο δείγμα μας είναι τυχαία, πριν την επιλογή της εφαρμογής παραμετρικών ελέγχων, οι οποίοι έχουν ως προϋπόθεση την τυχαιότητα των παρατηρήσεων ή μη παραμετρικών ελέγχων, είναι απαραίτητο να εξετάσουμε την τυχαιότητα των παρατηρήσεων σε αυτό.

Ο έλεγχος της τυχαιότητας έγινε με το στατιστικό λογισμικό SPSS και με μηδενική υπόθεση είναι ότι η σειρά των παρατηρήσεων είναι τυχαία.
Επιλέγοντας την εξαρτημένη μεταβλητή « Αξία Υ» παρατηρούμε τα αποτελέσματα του ελέγχου που εξάγονται από το πρόγραμμα (Πίνακας 5.1).

Πίνακας 5.1

Έλεγχος τυχαιότητας δείγματος με run test.

Runs Test

	Αξία Κατοικίας
Test Value[a]	105000
Cases « Test Value	498
Cases «= Test Value	502
Total Cases	1000
Number of Runs	383
Z	-7,466
Asymp. Sig. (2-tailed)	,000

a. Median

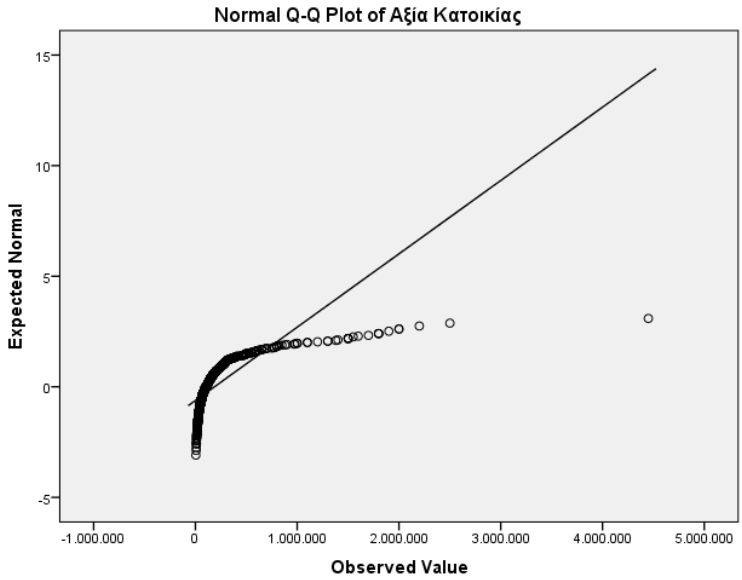
Λαμβάνοντας ως σημείο τομής την διάμεσο, παρατηρούμε ότι τα 105000 ευρώ είναι το σημείο διχοτόμησης στο δείγμα μας. Οι περιπτώσεις πάνω και κάτω από το δείγμα είναι περίπου ίσες, ενώ το z στατιστικό τεστ είναι αρκετά μικρό, το p (Asymp. Sig.) είναι μικρότερο του 0,05 ($p<0.05$) και γι' αυτό το λόγο δεν αποδεχόμαστε την μηδενική υπόθεση ότι υπάρχει τυχαιότητα των παρατηρήσεων του δείγματος. Κατ' επέκταση δεν μπορεί να αποδειχθεί η τυχαιότητα του δείγματος μέσα από αυτόν τον έλεγχο, αλλά ουσιαστικά και από τον τρόπο λήψης του.

5.4.1.1.2 Έλεγχος Κανονικότητας.

Μια ακόμα προϋπόθεση για την εφαρμογή παραμετρικών ελέγχων είναι η ύπαρξη κανονικής κατανομής στο δείγμα μας. Ο έλεγχος αυτός μπορεί να γίνει γραφικά αλλά και με στατιστικά κριτήρια μέσα από τα τεστ των Kolmogorov-Smirnov και Shapiro-Wilk. Ο γραφικός έλεγχος για την εξαρτημένη μεταβλητή γίνεται με το διάγραμμα Normal Q-Q Plot και Detrended Normal Q-Q Plot το οποίο συγκρίνει τα ποσοστιαία σημεία (quantile) του δείγματος έναντι των πληθυσμιακών ποσοστιαίων σημείων της κανονικής κατανομής. Σε μια ιδανική κανονική κατανομή στο Q-Q Plot όλα τα κυκλάκια θα ήταν πάνω στην ευθεία (διχοτόμο των αξόνων), γεγονός το οποίο δεν συμβαίνει και κατ' επέκταση το δείγμα δεν φαίνεται να ακολουθεί την κανονική κατανομή. Εάν οι παρατηρημένες τιμές κατανέμονταν κανονικά τότε τα σημεία στο Detrended Normal Q-Q Plot θα έπρεπε να συγκεντρώνονται σε μια οριζόντια ζώνη κοντά στο μηδέν και χωρίς κάποιο ευδιάκριτο σχέδιο(Εικόνα 5.1, 5.2).

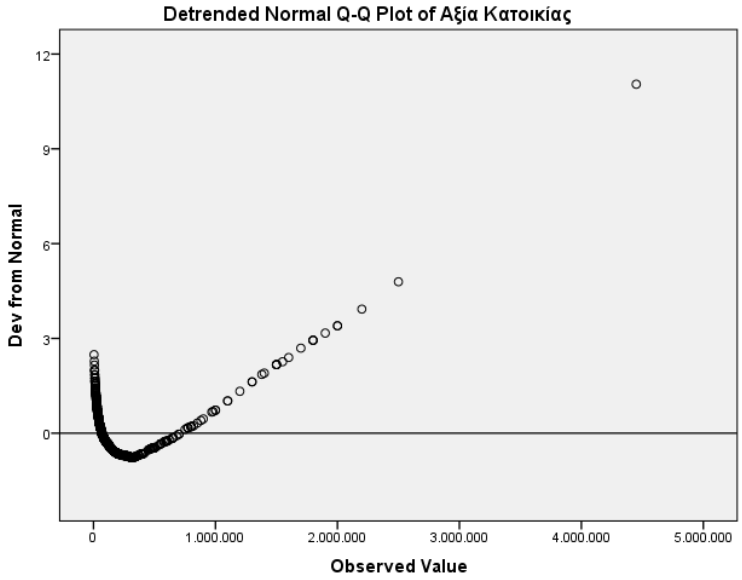
Εικόνα 5.1

Γραφικός έλεγχος κανονικότητας με Normal Q-Q Plot.



Εικόνα 5.2

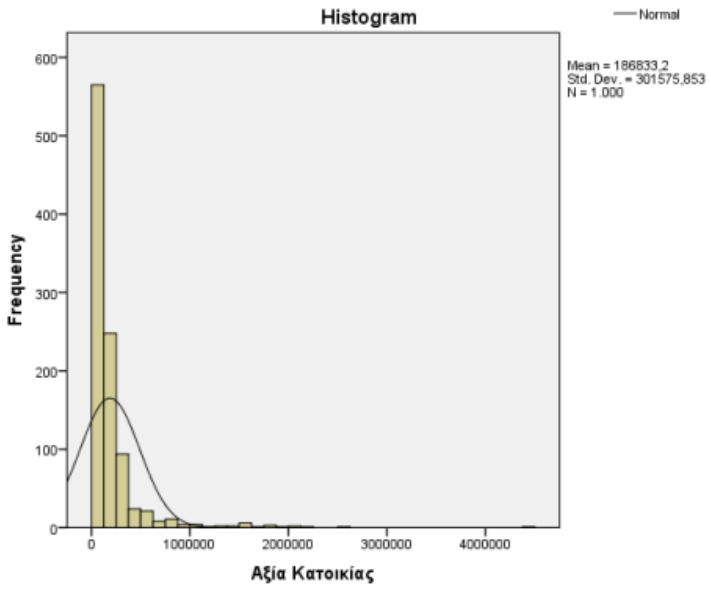
Γραφικός έλεγχος κανονικότητας με Detrended Normal Q-Q Plot.



Επίσης στο ιστόγραμμα παρατηρούμε ότι η καμπύλη δεν είναι συμμετρική γύρω από την μέση τιμή της, παρουσιάζοντας θετική ασυμμετρία, ενώ εμφανίζονται ακραίες τιμές (Εικόνα5.3).

Εικόνα 5.3

Γραφικός έλεγχος κανονικότητας με Histogram συχνοτήτων .



Των διαγραμμάτων έπεται ο έλεγχος ύπαρξης κανονικής κατανομής με στατιστικά κριτήρια. Καθώς το δείγμα μας είναι μεγαλύτερο από 50, μας απασχολεί το κριτήριο Kolmogorov-Smirnov και όχι το Shapiro-Wilk (Δαφέρμος, 2002), ενώ καθώς το p(Sig.)<0,05 απορρίπτουμε τη μηδενική υπόθεση ότι υπάρχει κανονική κατανομή (Πίνακας 5.2).

Πίνακας 5.2

Έλεγχος κανονικότητας δείγματος με το κριτήριο Kolmogorov-Smirnov.

Tests of Normality

	Kolmogorov-Smirnov[a]			Shapiro-Wilk		
	Statistic	df	Sig.	Statistic	df	Sig.
Αξία Κατοικίας	,275	1000	,000	,479	1000	,000

a. Lilliefors Significance Correction

Αν η μηδενική υπόθεση απορρίπτεται, όπως στη συγκεκριμένη περίπτωση, τότε πρέπει να τροποποιήσουμε τα δεδομένα μας έτσι ώστε να ακολουθούν την κανονική κατανομή, χωρίς αυτή η τροποποίηση να είναι πάντα δυνατή. Η τροποποίηση αυτή γίνεται μέσα από μια μαθηματική σχέση όπως ο υπολογισμός του τετραγώνου, της ρίζας ή του λογαρίθμου της εξαρτημένης μεταβλητής. Σε περίπτωση μη τροποποίησης των δεδομένων, καθώς δεν είναι πάντα αυτή δυνατή, τότε η τεχνική εφαρμόζεται και σε μη κανονικά δεδομένα ως μοναδική επιλογή, με τον έλεγχο της

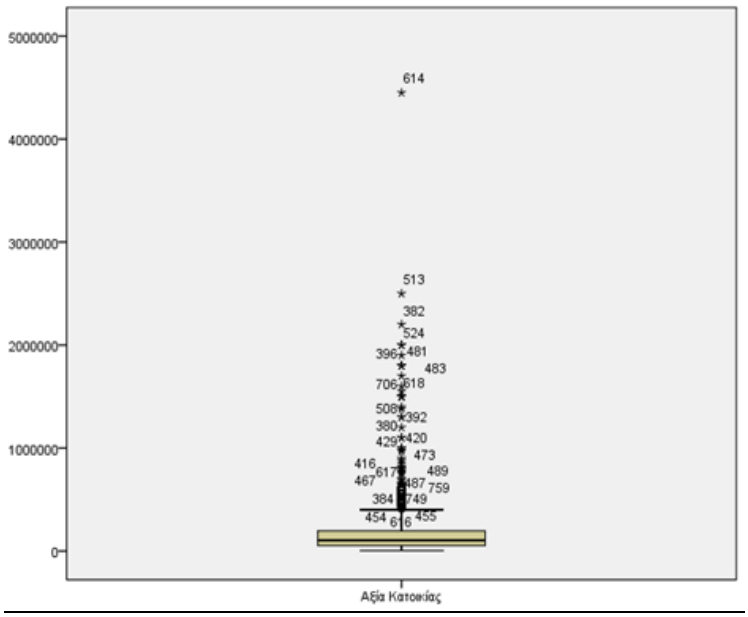
κανονικότητας να αποτελεί απλά ένδειξη εγκυρότητας της μεθόδου (Δημητροπουλάκης, χ.η, σελ. 10). Ανάλογα προχωρούμε και στο συγκεκριμένο μοντέλο.

5.4.1.1.3 Έλεγχος Ακραίων Παρατηρήσεων.

Ο έλεγχος των ακραίων παρατηρήσεων στο δείγμα μας γίνεται με την κατασκευή θηκογράμματος (Εικόνα 5.4) . Παρατηρούμε την ύπαρξη ακραίων τιμών στην εξαρτημένη μεταβλητή του δείγματός μας.

Εικόνα 5.4

Γραφικός έλεγχος ακραίων παρατηρήσεων με θηκόγραμμα.



Από τα παραπάνω δεν επαληθεύονται οι υποθέσεις για τη διενέργεια παραμετρικών ελέγχων στα ποσοτικά δεδομένα.. Καταρχήν θα πραγματοποιηθεί η διενέργεια μη παραμετρικών ελέγχων, τα οποία δεν χρησιμοποιούνται μόνο όταν δεν ικανοποιούνται οι προϋποθέσεις των παραμετρικών αλλά και για τα ποιοτικά δεδομένα.

5.4.1.2 Έλεγχοι Στατιστικών Υποθέσεων

Για την λήψη μιας στατιστικής απόφασης επί της πιθανότητας να συμβεί ένα γεγονός ή όχι, είναι αναγκαίο να κάνουμε υποθέσεις. Έτσι η απόφαση που θα πάρουμε στηρίζεται σε πιθανότητες χωρίς όμως να αποδεικνύεται η απόλυτη ισχύ ή όχι μιας υπόθεσης (Νικήτα, 2012). Καθώς σε προγενέστερο σημείο του παρόντος κεφαλαίου

καταλήξαμε στο συμπέρασμα ότι η εξαρτημένη μεταβλητή «Αξία Υ» δεν ακολουθεί την κανονική κατανομή, ξεκινούμε την ανάλυση με μη παραμετρικές μεθόδους, προκειμένου να εξακριβωθεί η συσχέτιση μεταξύ αυτής και των υπολοίπων ανεξάρτητων μεταβλητών.

5.4.1.3 Συσχέτιση Μεταβλητών- Έλεγχος x^2

Οι διασταυρώσεις μεταβλητών προκύπτουν από τους πίνακες συνάφειας οι οποίοι εμφανίζουν την σχέση μεταξύ δύο ή και περισσότερων μεταβλητών. Είναι σαφές ότι σε περίπτωση εμπλοκής ποσοτικών μεταβλητών είναι απαραίτητη η πρότερη ομαδοποίηση τους μέσα από κλάσεις. Η μηδενική υπόθεση είναι ότι οι μεταβλητές είναι ανεξάρτητες, ενώ αυτή απορρίπτεται όταν το p είναι μικρότερο του 0,05.

Τα εμβαδά ακινήτων ομαδοποιηθήκαν με βάση το μέγεθος που ακολουθείται κυρίως στα εμβαδά των ακινήτων στον ελλαδικό χώρο, δηλαδή πολύ μικρά (μικρότερα των 30τμ), μικρά (31-50τμ), μεσαία κλπ. Αντίστοιχη ομαδοποίηση έγινε και με τις ηλικίες των κατασκευών σε νεόδμητα, σε νέα (6 έως 15 έτη), κλπ. Όσο αφορά την ομαδοποίηση των αποστάσεων, αυτή ακολούθησε την λογική του πολύ κοντά, μέτρια απόσταση, μεγάλη, κ.ο.κ, όπως για παράδειγμα οι ομάδες αποστάσεων από αστυνομικά τμήματα. Αντίστοιχη λογική ομαδοποίησης είχε και η απόσταση από συγκοινωνία , με πρώτη ομάδα τη μηδενική απόσταση έως 50μ, δηλαδή σε εμβέλεια βαδίσματος κάτω του λεπτού . Η λογική της ομαδοποίησης των αποστάσεων από νεκροταφείο με πρώτη ομάδα αυτή όπου υπάρχει επαφή με το νεκροταφείο, είναι ότι αυτά τα ακίνητα που εφάπτονται σε αυτό λογικά έχουν μικρότερη τιμή από αυτά που απέχουν για παράδειγμα 500μ καθώς λογικά παρεμβάλλονται άλλες πολυκατοικίες και δεν έχουν οπτική επαφή μαζί του. Τέλος η λογική ομαδοποίησης αποστάσεων από βιομηχανία βιοτεχνία έγινε σε μεγαλύτερη κλίμακα αποστάσεων, έχοντας υπόψη ότι συνήθως και οι εκτάσεις που καταλαμβάνουν οι βιομηχανίες είναι πολύ μεγάλες . Έτσι ως πρώτη ομάδα τέθηκε αυτή έως 500μ , δεύτερη αυτή από 500 έως 1000μ, κοκ.

Στον πίνακα 5.3 παρουσιάζονται συγκεντρωτικά τα αποτελέσματα συσχέτισης της ομαδοποιημένης εξαρτημένης μεταβλητής, που αφορά την αξία του ακινήτου, με τις υπόλοιπες ποσοτικές μεταβλητές, όπως αυτές έχουν ομαδοποιηθεί (δηλαδή έχουν μετατραπεί σε ιεραρχικές μεταβλητές) ώστε να είναι επεξεργάσιμες για τον έλεγχο chi square test και τις ποιοτικές μεταβλητές. Σκοπός μας είναι να εντοπίσουμε τις μεταβλητές αυτές που δεν εμφανίζουν εξάρτηση με την εξαρτημένη μεταβλητή, ώστε να μη λάβουν μέρος στην κατασκευή του μοντέλου μας. Τα αναλυτικά αποτελέσματα

παρουσιάζονται στο παράρτημα «Α5».

Πίνακας 5.3

Συνοπτική παρουσίαση αποτελεσμάτων ελέγχου Pearson Chi- Square

ΟΜΑΔΕΣ ΑΞΙΑΣ ΑΚΙΝΗΤΩΝ		
	Pearson Chi- Square	p
ΟΜΑΔΕΣ ΕΜΒΑΔΩΝ ΑΚΙΝΗΤΩΝ Χ32α	495.201	0
ΟΜΑΔΕΣ ΗΜΕΡΩΝ ΠΡΟΣ ΔΙΑΘΕΣΗ Χ53α	20988	0.137
ΟΜΑΔΕΣ ΗΛΙΚΙΑΣ ΚΑΤΑΣΚΕΥΗΣ Χ40α	151.879	0
ΟΜΑΔΕΣ ΑΠΟΣΤΑΣΕΩΝ ΑΠΌ ΣΧΟΛΕΙΟ Χ21α	40.269	0
ΟΜΑΔΕΣ ΑΠΟΣΤΑΣΗ ΑΠΌ ΑΘΛΗΤΙΣΜΟ Χ22α	97.909	0
ΟΜΑΔΕΣ ΑΠΟΣΤΑΣΗ ΑΠΌ ΝΑΟ Χ20α	27.389	0
ΟΜΑΔΕΣ ΑΠΟΣΤΑΣΗ ΑΠΌ ΝΟΣΟΚΟΜΕΙΟ Χ17α	13.011	0.601
ΟΜΑΔΕΣ ΑΠΟΣΤΑΣΗ ΑΠΌ ΒΙΟΜ/ΝΙΑ ΒΙΟΤ/ΝΙΑ Χ12α	41.438	0
ΟΜΑΔΕΣ ΑΠΟΣΤΑΣΗ ΑΠΌ ΕΜΠΟΡΙΟ Χ14α	125.853	0
ΟΜΑΔΕΣ ΑΠΟΣΤΑΣΗ ΑΠΌ ΒΡΕΦ/ΚΟ ΣΤΑΘΜΟ Χ13α	79.027	0
ΟΜΑΔΕΣ ΑΠΟΣΤΑΣΗ ΑΠΌ ΠΡΑΣΙΝΟ Χ8α	51.562	0
ΟΜΑΔΕΣ ΑΠΟΣΤΑΣΗ ΑΠΌ ΙΣΤΟΡΙΚΟ ΚΕΝΤΡΟ Χ15α	223.594	0
ΟΜΑΔΕΣ ΑΠΟΣΤΑΣΗ ΑΠΌ ΠΕΖΟΔΡΟΜΟ Χ25α	21.91	0.016
ΟΜΑΔΕΣ ΑΠΟΣΤΑΣΗ ΑΠΌ ΒΑΣΙΚΟ ΟΔ. ΔΙΚΤΥΟ Χ29α	7.349	0.692
ΟΜΑΔΕΣ ΑΠΟΣΤΑΣΗ ΑΠΌ ΜΟΥΣΕΙΟ Χ27α	172.548	0
ΟΜΑΔΕΣ ΑΠΟΣΤΑΣΗ ΑΠΌ ΟΣΕ Χ26α	85.067	0
ΟΜΑΔΕΣ ΤΙΜΩΝ ΑΝΑ τμ Χ33α	1215.023	0
ΟΜΑΔΕΣ ΤΙΜΩΝ ΖΩΝΗΣ Χ31α	341.999	0
Είδος πλησιέστερης συγκοινωνίας Χ18	9.928	0.447
Μέσο εισόδημα κατοίκων Χ30	157.718	0
Ύπαρξη WC Χ36	52.382	0
Ύπαρξη αποθήκης Χ37	108.318	0
Είδος κατοικίας Χ41	142.728	0
Κατάσταση κατασκευής Χ42	225.733	0
Ύπαρξη ανακαίνισης Χ43	131.882	0
Ύπαρξη βεράντας Χ44	16.09	0.007
Γωνιακό ή διαμπερές Χ45	14.314	0.014
Ύπαρξη ασανσέρ Χ46	10.03	0.074
Ύπαρξη θέσης στάθμευσης Χ47	91.942	0
Είδος θέρμανσης Χ48	118.306	0
Είδος καυσίμου Θέρμανσης Χ49	94.82	0
Ύπαρξη κλιματισμού Χ50	9.268	0.099
Ύπαρξη θέας Χ51	124.052	0
Ύπαρξη τζακιού Χ52	63.066	0
Ποσοστό ελλήνων Χ4	209.899	0
Ποσοστό αλλοδαπών Χ5	209.899	0
Ποσοστό ανεργίας Χ6	151.009	0
Ποσοστό κενών κατοικιών Χ7	64.049	0

Ποσοστό πρασίνου Χ8	231.255	0
Ποσοστό επιμερισμού κατοίκων ανά Δ.Κ Χ9	153.999	0
Πράσινο τμ ανά κάτοικο Χ10	220.889	0
ΟΜΑΔΕΣ ΑΠΟΣΤΑΣΗΣ ΑΠΌ ΑΤ Χ11α	28.301	0.02
Αριθμός υπνοδωματίων Χ34	358.948	0
Αριθμός μπάνιων Χ35	262.025	0
Όροφος Χ38	75.535	0.001
ΟΜΑΔΕΣ ΑΠΟΣΤΑΣΗΣ ΑΠΌ ΣΥΓΚΟΙΝΩΝΙΑ Χ19α	13011	0.601
ΟΜΑΔΕΣ ΑΠΟΣΤΑΣΗΣ ΑΠΌ ΠΡ. ΟΔΙΚΟ ΔΙΚΤΥΟ Χ28α	45.756	0
ΟΜΑΔΕΣ ΑΠΟΣΤΑΣΗΣ ΑΠΌ ΠΛΑΚΑ Χ24α	223.594	0
ΟΜΑΔΕΣ ΑΠΟΣΤΑΣΗΣ ΑΠΌ ΝΕΚΡΟΤΑΦΕΙΟ Χ16α	100.331	0

Καθώς η μηδενική υπόθεση είναι ότι οι εξεταζόμενες με τον έλεγχο chi square test μεταβλητές είναι ανεξάρτητες, ενώ η μηδενική υπόθεση απορρίπτεται όταν το p είναι μικρότερο του 0,05, από τον παραπάνω πίνακα προκύπτει ότι θα πρέπει να αφαιρεθούν από την συμμετοχή τους στο μοντέλο μας οι μεταβλητές : Αριθμός ημερών προς διάθεση του ακινήτου Χ53, Απόσταση από Νοσοκομείο Χ17, Απόσταση από το βασικό οδικό δίκτυο Χ29, Είδος πλησιέστερης συγκοινωνίας Χ18, Ύπαρξη ασανσέρ Χ46, Ύπαρξη κλιματισμού Χ50, Απόσταση από πλησιέστερη συγκοινωνία Χ19, διότι το p που εμφανίζουν είναι μεγαλύτερο από 0,05.

Επιπρόσθετα θα πρέπει στο σημείο αυτό να αφαιρεθούν οι μεταβλητές που δεν έχουν να δώσουν κάτι επιπλέον στο μοντέλο καθώς είτε αποτελούν απόρροια κάποιας άλλης μεταβλητής, είτε γιατί λειτουργούν ως συμπληρωματικές κάποιας άλλης. Έτσι θα πρέπει να αφαιρεθεί η μεταβλητή « Έτος κατασκευής Χ39» διότι αποτελεί πρακτικά την μεταβλητή ηλικία ακινήτου καθώς αυτή προκύπτει από το υπόλοιπο της αφαίρεσης του έτους κατασκευής με την σύγχρονη ημερομηνία. Επιπλέον το ποσοστό ελλήνων με το ποσοστό αλλοδαπών λειτουργούν ως συμπληρωματικές μεταβλητές και έτσι θα πρέπει να αφαιρεθεί η μία εξ αυτών. Επιλέγεται να αφαιρεθεί η μεταβλητή « Ποσοστό αλλοδαπών Χ5». Τέλος θα πρέπει να αφαιρεθεί και η μεταβλητή «Τιμή / τμ Χ33» καθώς πρόκειται για την τιμή που απορρέει από την διαίρεση της αξίας ακινήτου με το εμβαδόν.

5.4.1.4 Εξέταση Ισομερούς Κατανομής Ποιοτικών Μεταβλητών

Είναι γνωστό ότι οι μεταβλητές που χρησιμοποιούνται κατά κύριο λόγο στην ανάλυση παλινδρόμησης είναι ποσοτικές μεταβλητές. Οι ποιοτικές μεταβλητές, ενώ γενικά δεν χρησιμοποιούνται, εμφανίζουν μερικές φορές σημαντικότατη συσχέτιση

με την εξαρτημένη μεταβλητή και γι’ αυτό τον λόγο είναι σημαντικό να περιληφθούν στο μοντέλο (Ηλιοπούλου, 2017). Σύμφωνα με την βιβλιογραφία (Αρβανίτης κ.ά , 2010), οι δίτιμες ποιοτικές μεταβλητές, εφόσον λαμβάνουν μέρος στο μοντέλο, πρέπει να εκπροσωπούνται ισομερώς στο δείγμα. Επόμενο βήμα λοιπόν είναι η εξέταση της ισομερούς κατανομής των ποιοτικών μεταβλητών που έχουν απομείνει ώστε να αποφασιστεί η αφαίρεσή τους ή όχι. Στο πίνακα «Α6» παρουσιάζονται αναλυτικά τα αποτελέσματα ενώ προκύπτει ότι : η μεταβλητή «Ύπαρξη WC Χ36» εμφανίζει το Όχι σε ποσοστό 83,3%, η μεταβλητή «Ύπαρξη αποθήκης Χ37» εμφανίζει το Όχι σε ποσοστό 75,1%, η μεταβλητή «Ύπαρξη βεράντας Χ44» εμφανίζει το Ναι σε ποσοστό 93,5%, η μεταβλητή «Γωνιακό ή διαμπερές Χ45» εμφανίζει το Όχι σε ποσοστό 89% , η μεταβλητή «Ύπαρξη θέσης στάθμευσης Χ47» εμφανίζει το Όχι σε ποσοστό 83,9% , η μεταβλητή «Ύπαρξη θέας Χ51» εμφανίζει το Όχι σε ποσοστό 80,7%, η μεταβλητή «Ύπαρξη τζακιού Χ52» εμφανίζει το Όχι σε ποσοστό 93,9%. Από την στιγμή που δεν εκπροσωπούνται ισομερώς στο δείγμα μας , σύμφωνα με την μεθοδολογία που ακολουθούμε στο συγκεκριμένο μοντέλο θα πρέπει να αφαιρεθούν.

Στον πίνακα 5.4 παρουσιάζεται ενημερωμένος ο αρχικός πίνακας με το σύνολο των μεταβλητών, με αντίστοιχη διαγραφή των αφαιρούμενων μεταβλητών της μέχρι τώρα επεξεργασίας.

Πίνακας 5.4

Πρώτη φάση επεξεργασίας. Παρουσίαση πίνακα μεταβλητών.

Ονομασία μεταβλητής	Κωδικός	Ονομασία μεταβλητής	Κωδικός	Ονομασία μεταβλητής	Κωδικός
Αξία	Υ	~~Πλησιέστερο Είδος Συγκοινωνίας~~	~~Χ18~~	~~Ύπαρξη WC~~	~~Χ36~~
Όνομα ΔΚ	Χ1	~~Απόσταση από Πλησιέστερη Συγκοινωνία~~	~~Χ19~~	~~Ύπαρξη Αποθήκης~~	~~Χ37~~
Περιοχή	Χ2	Απόσταση από Ναό	Χ20	Όροφος	Χ38
Γειτονιά	Χ3	Απόσταση από Σχολείο	Χ21	~~Έτος κατασκευής~~	~~Χ39~~
Ποσοστό Ελλήνων	Χ4	Απόσταση από Αθλητική εγκατάσταση	Χ22	Ηλικία κατοικίας	Χ40
~~Ποσοστό Αλλοδαπών~~	~~Χ5~~	Απόσταση από Πράσινο	Χ23	Είδος κατοικίας	Χ41

Ποσοστό Ανεργίας	X6	Απόσταση από περιοχή Πλάκας	X24	Κατάσταση κατασκευής	X42
Ποσοστό Κενών κατοικιών	X7	Απόσταση από Πεζόδρομο	X25	Ύπαρξη ανακαίνισης	X43
Ποσοστό Πρασίνου	X8	Απόσταση από ΟΣΕ	X26	~~Ύπαρξη Βεράντας~~	~~X44~~
Ποσοστό Επιμερισμού Κατοίκων ανά ΔΚ	X9	Απόσταση από Μουσείο	X27	~~Γωνιακό ή Διαμπερές~~	~~X45~~
Πράσινο ανά κάτοικο τμ	X10	Απόσταση από Πρωτεύον οδικό δίκτυο	X28	~~Ύπαρξη ασανσέρ~~	~~X46~~
Απόσταση από Αστυνομικό τμήμα	X11	~~Απόσταση από Βασικό οδικό δίκτυο~~	~~X29~~	~~Ύπαρξη θέσης στάθμευσης~~	~~X47~~
Απόσταση από Βιομηχανία βιοτεχνία	X12	Μέσο εισόδημα κατοίκων	X30	Είδος Θέρμανσης	X48
Απόσταση από Βρεφονηπιακό σταθμό	X13	Αντικειμενική αξία	X31	Είδος Καυσίμου θέρμανσης	X49
Απόσταση από Εμπόριο	X14	Εμβαδόν κατοικίας	X32	~~Ύπαρξη Κλιματισμού~~	~~X50~~
Απόσταση από Ιστορικό κέντρο	X15	~~Τιμή ανά τμ~~	~~X33~~	~~Ύπαρξη Θέας~~	~~X51~~
Απόσταση από Νεκροταφείο	X16	Αριθμός υπνοδωματίων	X34	~~Ύπαρξη Τζακιού~~	~~X52~~
~~Απόσταση από Νοσοκομείο~~	~~X17~~	Αριθμός Μπάνιων	X35	~~Αριθμός Ημερών διαθεσιμότητας~~	~~X53~~

5.4.1.5 Προσαρμογή Μοντέλου Γραμμικής Παλινδρόμησης

Η προσαρμογή του μοντέλου της πολλαπλής γραμμικής παλινδρόμησης γίνεται μέσα από το στατιστικό πρόγραμμα SPSS με τις ανεξάρτητες μεταβλητές X4, X6, X7, X8, X9, X10, X11, X12, X13, X14, X15, X16, X20, X21, X22, X23, X24, X25, X26, X27, X28, X30, X31 ,X32, X34, X35, X38, X40, X41, X42, X43, X48, X49 και την εξαρτημένη μεταβλητή Y. Ουσιαστικά ορίζουμε το μέγιστο δυνατό μοντέλο που θα έχει όλους τους δυνατούς predictors, όλους τους μεγαλύτερου βαθμού όρους και όλες τις μεταξύ των μεταβλητών αλληλεπιδράσεις (Χαλικιάς, κ.ά, 2015, σελ. 246).

Επιλέγουμε την μέθοδο enter και τρέχουμε το πρόγραμμα. Σημειώνεται ότι με τη διαδικασία αυτή εισάγονται στο μοντέλο όλες οι ανεξάρτητες μεταβλητές που έχουν απομείνει και θα προκύψει στη συνέχεια αν είναι στατιστικά σημαντικές σε σχέση με την εξαρτημένη μεταβλητή. Στόχος της κατασκευής του μοντέλου μας είναι η αξιοπιστία και η εγκυρότητα (Χαλικιάς, κ.ά, 2015, σελ. 246).

Τα αποτελέσματα παρουσιάζονται αναλυτικά στο παράρτημα «Α7» . Τα συνοπτικά αποτελέσματα του μοντέλου που παρουσιάζονται στον πίνακα 5.5 έχουν τα εξής στοιχεία: Το R=0,665 είναι η απόλυτη τιμή του συντελεστή γραμμικής συσχέτισης, ενώ το R Square=0,443 είναι το τετράγωνο του και ονομάζεται συντελεστής προσδιορισμού. Αυτό που δείχνει είναι το ποσοστό μεταβλητότητας των δεδομένων που εξηγείται από το γραμμικό μοντέλο που προσαρμόστηκε έχοντας λάβει υπόψη του και το μέγεθος του δείγματος.

Πίνακας 5.5

Γραμμική παλινδρόμηση: Συντελεστής προσδιορισμού

Model Summary[b]

Model	R	R Square	Adjusted R Square	Std. Error of the Estimate
1	,665[a]	,443	,424	228931,324

a. Predictors: (Constant), X49 Είδος Καυσίμου θέρμανσης, X35 Αριθμός Μπάνιων, X11 Απόσταση από Αστυνομικό τμήμα, X26 Απόσταση από ΟΣΕ, X21 Απόσταση από Σχολείο, X20 Απόσταση από Ναό, X23 Απόσταση από Πράσινο, X38 Όροφος, X25 Απόσταση από Πεζόδρομο, X13 Απόσταση από Βρεφονηπιακό σταθμό, X16 Απόσταση από Νεκροταφείο, X41 Είδος Κατοικίας, X43 Ύπαρξη Ανακαίνισης, X28 Απόσταση από Πρωτεύων οδικό δίκτυο, X40 Ηλικία Κατασκευής, X10 Πράσινο τμ ανά κάτοικο, X7 Ποσοστό Κενών κατοικιών, X34 Αριθμός Υπνοδωματίων, X22 Απόσταση από Αθλητισμό, X31 Τιμή ζώνης, X48 Είδος Θέρμανσης, X32 Εμβαδόν Κατοικίας, X30 Μέσο Εισόδημα κατοίκων, X14 Απόσταση από Εμπορική δραστηριότητα, X27 Απόσταση από Μουσείο, X42 Κατάσταση Κατασκευής, X9 Ποσοστό Κατοίκων ανά ΔΚ, X12 Απόσταση από Βιομηχανία, X6 Ποσοστό Ανεργίας, X15 Απόσταση από Ιστορικό κέντρο Αθήνας, X24 Απόσταση από την Πλάκα, X8 Ποσοστό Πρασίνου, X4 Ποσοστό Ελλήνων

b. Dependent Variable: Y Αξία Κατοικίας

Το F=23,261 στο πίνακα 5.6 δείχνει την ποιότητα της παλινδρόμησης, καθώς καλή παλινδρόμηση απορρέει από μεγάλο F και σημαντικότητα (Sig. F Change) μικρότερη από το 0,05. Η συνάρτηση F ελέγχει την υπόθεση ότι οι συντελεστές των ανεξάρτητων μεταβλητών που συμμετέχουν στο μοντέλο είναι ταυτόχρονα μηδέν.

Το σύνολο της διακύμανσης του δείγματος (Total= 90857047017758,000) είναι άθροισμα της διακύμανσης της παλινδρόμησης (regression = 40229420565924,190) και της διακύμανσης του λάθους (Residual = 50627626451833,810). Όσο καλύτερο

είναι το μοντέλο τόσο μεγαλύτερο μέρος της συνολικής διακύμανσης του δείγματος επεξηγεί. Το R^2 και το F είναι οι δείκτες καλής προσαρμογής. Ο έλεγχος F (βασίζεται στην F κατανομή) και ελέγχει αν όλοι οι παράμετροι του μοντέλου είναι μηδέν ή αν έστω και ένας είναι διάφορος του μηδενός. Ο αριθμός 40229420565924,190 δείχνει τη διακύμανση που εξηγείται από το μοντέλο όπως αυτό προσαρμόστηκε και ο 90857047017758,000 δείχνει τη συνολική διακύμανση των δεδομένων. Η διαφορά τους είναι η διακύμανση που δεν εξηγείται από το μοντέλο , ενώ το πηλίκο τους είναι ο συντελεστής προσδιορισμού. Στον πίνακα ANOVA μας ενδιαφέρει το F. Αν το Sig. <0,05 και το F είναι μεγάλο τότε η γραμμική παλινδρόμηση είναι στατιστικά σημαντική.

Πίνακας 5.6

Γραμμική παλινδρόμηση: Ανάλυση διασποράς

ANOVA[a]

Model		Sum of Squares	df	Mean Square	F	Sig.
1	Regression	40229420565924,520	33	1219073350482,561	23,261	,000[b]
	Residual	50627626451833,270	966	52409551192,374		
	Total	90857047017757,800	999			

Στον πίνακα 5.7 παρουσιάζονται τα αποτελέσματα της παλινδρόμησης. Τα Beta (b) είναι οι συντελεστές της κάθε ανεξάρτητης μεταβλητής Xi, αν στο μοντέλο μας δεν χρησιμοποιήσουμε τις αυθεντικές τιμές (Unstandardized Coefficients), αλλά τις τυποποιημένες (Standardized Coefficients). Τα ελέγχουμε σε μοντέλα με πολλές ανεξάρτητες μεταβλητές, χρησιμεύοντας στην ταξινόμηση τους ως προς την ερμηνευτική τους ικανότητα. Δεν λαμβάνουμε υπόψη τα πρόσημα για την ταξινόμηση, αλλά τις απόλυτες τιμές. Τα t και significance συνιστούν τον έλεγχο σημαντικότητας της κάθε ανεξάρτητης μεταβλητής. Ελέγχουν αν το B της ανεξάρτητης μεταβλητής μπορεί ποτέ να είναι μηδέν. Αν το p Sig. <0.05 η μεταβλητή Xi είναι στατιστικά σημαντική. Η τιμή 5225061,959 (Constant) είναι η τιμή στην οποία η ευθεία των ελαχίστων τετραγώνων τέμνει τον άξονα της εξαρτημένης μεταβλητής. Οι υπόλοιπες τιμές δείχνουν την κύρια επίδρασή τους στην εξαρτημένη μεταβλητή. Η στήλη Sig. περιέχει τα παρατηρημένα επίπεδα στατιστικής

σημαντικότητας, ώστε να εξαχθούν συμπεράσματα επί της στατιστικής σημαντικότητας των παραμέτρων του μοντέλου. Καθώς η μηδενική υπόθεση είναι ότι οι συντελεστές είναι μηδέν, εάν τα p-value είναι μικρότερα του 0,05 τότε αυτή απορρίπτεται, υποδεικνύοντας την στατιστική σημαντικότητα των συντελεστών. Ένδειξη προβλήματος πολυσυγγραμικότητας έχουμε όταν οι τιμές του VIF (Variation Inflation Factor) ως μέτρου διάγνωσης συγγραμμικότητας είναι μεγαλύτερες του 5, ενώ η τιμή της Tolerance παίρνει τιμές από 0-1 και για μία τιμή δείχνει το ποσοστό της διακύμανσης της μεταβλητής που εξηγείται από τις υπόλοιπες ανεξάρτητες μεταβλητές του μοντέλου. Τιμές μικρότερες του 0,5 αποτελούν ένδειξη προβλήματος, δηλαδή ότι η μεταβλητή είναι σε γραμμικό σχεδόν συνδυασμό με τις άλλες ανεξάρτητες μεταβλητές (Τσαγρής,2014 σελ.95). Ο VIF μεγαλώνει όσο μικραίνει ο Tolerance. Συνήθως ανεξάρτητες μεταβλητές με VIF>5 αποβάλλονται (Χαλικιάς, κ.ά, 2015).

Παρατηρούμε ότι το R =0,665 και το R^2 =0,443 ενώ το F=23,261 το p<0.05. Το δείγμα έχει μέτρια προσαρμογή αφού εξηγεί το 44,3% της διασποράς της εξαρτημένης μεταβλητής . Επιπρόσθετα παρατηρούμε από τον πίνακα 5.8 ότι οι μεταβλητές X6, X8, X9, X11, X12, X13, X14, X20, X21, X25, X26, X27, X28, X30, X34, X35, X41, X43, X49 φαίνονται στατιστικά μη σημαντικές καθώς το αντίστοιχο p- value»0.05 σε επίπεδο σημαντικότητας 5%.

Παράλληλα παρατηρούμε ότι ο δείκτης VIF είναι μεγαλύτερος του 5 για τις μεταβλητές X4, X6, X7, X8, X9, X10, X12, X14, X15, X16, X24, X26, X27 και θεωρείται ως ένδειξη πολυσυγγραμικότητας.

Πίνακας 5.7

Αποτελέσματα παλινδρόμησης: Συντελεστές παλινδρόμησης

Coefficients[a]

Model		Unstandardized Coefficients		Standardized Coefficients	t	Sig.	Collinearity Statistics	
		B	Std. Error	Beta			Tolerance	VIF
1	(Constant)	5352433,857	2021460,299		2,648	,008		
	X4 Ποσοστό Ελλήνων	-40257,879	15084,304	-1,078	-2,669	,008	,004	282,586
	X6 Ποσοστό Ανεργίας	17655,441	27202,900	,134	,649	,516	,013	74,098

Χ7 Ποσοστό Κενών κατοικιών	-101670,742	29594,499	-1,282	-3,435	,001	,004	241,529
Χ8 Ποσοστό Πρασίνου	-11951,233	10419,200	-,370	-1,147	,252	,006	180,419
Χ9 Ποσοστό Κατοίκων ανά ΔΚ	6560,169	7905,078	,087	,830	,407	,053	18,864
Χ10 Πράσινο τμ ανά κάτοικο	32405,256	14119,643	,671	2,295	,022	,007	148,218
Χ11 Απόσταση από Αστυνομικό τμήμα	-30,092	26,365	-,037	-1,141	,254	,563	1,777
Χ12 Απόσταση από Βιομηχανία	20,013	25,059	,058	,799	,425	,108	9,274
Χ13 Απόσταση από Βρεφονηπιακό σταθμό	-16,466	47,932	-,010	-,344	,731	,623	1,604
Χ14 Απόσταση από Εμπορική δραστηριότητα	-44,174	30,668	-,118	-1,440	,150	,086	11,616
Χ15 Απόσταση από Ιστορικό κέντρο Αθήνας	113,583	51,726	,334	2,196	,028	,025	40,149
Χ16 Απόσταση από Νεκροταφείο	-62,566	19,350	-,213	-3,233	,001	,133	7,519
Χ20 Απόσταση από Ναό	15,193	66,082	,006	,230	,818	,797	1,254
Χ21 Απόσταση από Σχολείο	31,700	66,273	,013	,478	,633	,785	1,274
Χ22 Απόσταση από Αθλητισμό	-75,437	37,280	-,066	-2,024	,043	,549	1,820
Χ23 Απόσταση από Πράσινο	-237,430	74,428	-,087	-3,190	,001	,782	1,279
Χ24 Απόσταση από την Πλάκα	-87,399	42,005	-,361	-2,081	,038	,019	52,096
Χ25 Απόσταση από Πεζόδρομο	54,847	89,754	,017	,611	,541	,750	1,334
Χ26 Απόσταση από ΟΣΕ	35,837	27,797	,124	1,289	,198	,062	16,106
Χ27 Απόσταση από Μουσείο	-40,196	35,018	-,086	-1,148	,251	,102	9,760
Χ28 Απόσταση από Πρωτεύων οδικό δίκτυο	-18,177	30,868	-,026	-,589	,556	,297	3,362

	Χ30 Μέσο Εισόδημα κατοίκων	29531,448	19333,474	,065	1,527	,127	,316	3,162
	Χ31 Τιμή ζώνης	101,725	18,189	,218	5,593	,000	,381	2,624
	Χ32 Εμβαδόν Κατοικίας	1327,043	126,728	,385	10,472	,000	,427	2,343
	Χ34 Αριθμός Υπνοδωματίων	14224,519	10759,181	,047	1,322	,186	,464	2,156
	Χ35 Αριθμός Μπάνιων	-5959,755	19033,584	-,009	-,313	,754	,666	1,502
	Χ38 Όροφος	11112,138	3955,046	,073	2,810	,005	,866	1,154
	Χ40 Ηλικία Κατασκευής	1732,577	419,015	,130	4,135	,000	,586	1,706
	Χ41 Είδος Κατοικίας	-31842,974	21433,695	-,042	-1,486	,138	,739	1,353
	Χ42 Κατάσταση Κατασκευής	-53927,501	12122,939	-,230	-4,448	,000	,217	4,619
	Χ43 Ύπαρξη Ανακαίνισης	-44531,739	29695,077	-,074	-1,500	,134	,239	4,192
	Χ48 Είδος Θέρμανσης	-42381,728	15684,369	-,094	-2,702	,007	,472	2,119
	Χ49 Είδος Καυσίμου θέρμανσης	990,942	14671,975	,002	,068	,946	,540	1,850

a. Dependent Variable: Y Αξία Κατοικίας

Σε αυτό το βήμα αφαιρούνται οι μεταβλητές Χ4, Χ6, Χ7, Χ8, Χ9, Χ10, Χ11, Χ12, Χ13, Χ14, Χ15, Χ16, Χ20, Χ21, Χ24, Χ25, Χ26, Χ27, Χ28, Χ30, Χ34, Χ35, Χ41, Χ43, Χ49 και το μοντέλο θα επαναϋπολογιστεί χωρίς αυτές που είναι μη στατιστικά σημαντικές ή εμφανίζουν ένδειξη πολυσυγγραμικότητας.

Ξανατρέχουμε το μοντέλο και τα αποτελέσματα παρουσιάζονται στο παράρτημα «Α8». Στους πίνακες 5.8 , 5.9, 5.10 παρουσιάζονται τα αποτελέσματα των συσχετίσεων και της παλινδρόμησης. Παρατηρούμε ότι το R=0,630 και το R^2 =0,397, το F=81,531, ενώ όλες οι ανεξάρτητες μεταβλητές παρουσιάζονται σημαντικές πλην της μεταβλητής « Απόσταση από αθλητισμό Χ22» η οποία και θα πρέπει να απομακρυνθεί.

Πίνακας 5.8

Γραμμική παλινδρόμηση: Συντελεστής προσδιορισμού

Model Summary[b]

Model	R	R Square	Adjusted R Square	Std. Error of the Estimate
1	,630[a]	,397	,392	235140,605

a. Predictors: (Constant), X48 Είδος Θέρμανσης, X23 Απόσταση από Πράσινο, X32 Εμβαδόν Κατοικίας, X22 Απόσταση από Αθλητισμό, X38 Όροφος, X42 Κατάσταση Κατασκευής, X31 Τιμή ζώνης, X40 Ηλικία Κατασκευής

b. Dependent Variable: Y Αξία Κατοικίας

Πίνακας 5.9

Γραμμική παλινδρόμηση: Ανάλυση διασποράς

ANOVA[a]

Model		Sum of Squares	df	Mean Square	F	Sig.
1	Regression	36063562809175,480	8	4507945351146,935	81,531	,000[b]
	Residual	54793484208582,320	991	55291104145,895		
	Total	90857047017757,800	999			

Πίνακας 5.10

Αποτελέσματα παλινδρόμησης: Συντελεστές παλινδρόμησης

Coefficients[a]

Model		Unstandardized Coefficients		Standardized Coefficients	t	Sig.	Collinearity Statistics	
		B	Std. Error	Beta			Tolerance	VIF
1	(Constant)	8866,631	38610,980		,230	,818		
	X22 Απόσταση από Αθλητισμό	-1,402	31,160	-,001	-,045	,964	,830	1,205
	X23 Απόσταση από Πράσινο	-204,099	69,833	-,074	-2,923	,004	,937	1,067
	X31 Τιμή ζώνης	108,198	13,283	,231	8,146	,000	,754	1,326
	X32 Εμβαδόν Κατοικίας	1519,949	89,013	,441	17,076	,000	,913	1,095
	X38 Όροφος	8644,899	3928,886	,056	2,200	,028	,926	1,080
	X40 Ηλικία Κατασκευής	1627,238	403,928	,122	4,029	,000	,666	1,502
	X42 Κατάσταση Κατασκευής	-42752,229	6780,808	-,182	-6,305	,000	,730	1,370
	X48 Είδος Θέρμανσης	-39521,148	13189,056	-,088	-2,997	,003	,704	1,420

a. Dependent Variable: Y Αξία Κατοικίας

Επαναλαμβάνουμε την διαδικασία ξανατρέχοντας το μοντέλο έχοντας αφαιρέσει την μεταβλητή X22. Τα αποτελέσματα παρουσιάζονται αναλυτικά στο παράρτημα «Α9», ενώ παρατίθενται περιληπτικά τα αποτελέσματα της

παλινδρόμησης στους πίνακες 5.11, 5.12, 5.13. Παρατηρούμε ότι το R=0,630 και το R^2 =0,397ενώ το F=93,272, ενώ όλες οι ανεξάρτητες μεταβλητές παρουσιάζονται πλέον σημαντικές.

Πίνακας 5.11

Γραμμική παλινδρόμηση: Συντελεστής προσδιορισμού

Model Summary[b]

Model	R	R Square	Adjusted R Square	Std. Error of the Estimate
1	,630[a]	,397	,393	235022,297

a. Predictors: (Constant), X48 Είδος Θέρμανσης, X23 Απόσταση από Πράσινο, X32 Εμβαδόν Κατοικίας, X38 Όροφος, X31 Τιμή ζώνης, X42 Κατάσταση Κατασκευής, X40 Ηλικία Κατασκευής

b. Dependent Variable: Y Αξία Κατοικίας

Πίνακας 5.12

Γραμμική παλινδρόμηση: Ανάλυση διασποράς

ANOVA[a]

Model		Sum of Squares	df	Mean Square	F	Sig.
1	Regression	36063450938239,164	7	5151921562605,595	93,272	,000[b]
	Residual	54793596079518,630	992	55235479918,870		
	Total	90857047017757,800	999			

Πίνακας 5.13

Αποτελέσματα παλινδρόμησης: Συντελεστές παλινδρόμησης

Coefficients[a]

Model		Unstandardized Coefficients		Standardized Coefficients			Collinearity Statistics	
		B	Std. Error	Beta	t	Sig.	Tolerance	VIF
1	(Constant)	8615,690	38186,598		,226	,822		
	X23 Απόσταση από Πράσινο	-204,232	69,736	-,075	-2,929	,003	,939	1,065
	X31 Τιμή ζώνης	107,978	12,342	,231	8,749	,000	,872	1,146
	X32 Εμβαδόν Κατοικίας	1519,812	88,916	,441	17,093	,000	,914	1,094

	Χ38 Όροφος	8640,136	3925,483	,056	2,201	,028	,927	1,079
	Χ40 Ηλικία Κατασκευής	1626,652	403,515	,122	4,031	,000	,666	1,501
	Χ42 Κατάσταση Κατασκευής	-42735,088	6766,685	-,182	-6,316	,000	,732	1,365
	Χ48 Είδος Θέρμανσης	-39587,824	13098,900	-,088	-3,022	,003	,713	1,403

a. Dependent Variable: Υ Αξία Κατοικίας

<u>Η τελική εξίσωση παλινδρόμησης</u> που προκύπτει μετά από την παραπάνω επεξεργασία είναι η :

Y = 8615,690 - 204,232 (X23) + 107,978 (X31) + 1519,812 (X32) + 8640,136 (X38) + 1626,652 (X40) - 42735,088 (X42) - 39587,824 (X48)

Παρατηρούμε το αρνητικό πρόσημο της μεταβλητής « Απόσταση από πράσινο Χ23». Αυτό υποδεικνύει ότι όσο μεγαλώνει δηλαδή η απόσταση από πράσινο τόσο μειώνεται η τιμή της εξαρτημένης μεταβλητής « Αξία Υ» γεγονός που είναι λογικό. Θετικά συμβάλει στο μοντέλο η μεταβλητή « Τιμή ζώνης Χ31» . Δηλαδή μεγαλύτερη τιμή σε αυτή την μεταβλητή αποδίδει μεγαλύτερη τιμή στην εξαρτημένη μεταβλητή. Η μεταβλητή « Εμβαδόν κατοικίας Χ32» συμβάλει και αυτή θετικά στο μοντέλο, όπως και η μεταβλητή « Όροφος Χ38» καθώς όσο μεγαλύτερος είναι ο όροφος τόσο μεγαλύτερη τιμή έχει το ακίνητο. Προβληματισμό προκαλεί το θετικό πρόσημο της μεταβλητής « Ηλικία κατασκευής Χ 40» καθώς υποδεικνύει ότι όσο μεγαλύτερη ηλικία έχει η κατασκευή τόσο περισσότερο συμβάλει θετικά στην τιμή της κατοικίας. Αυτό θα μπορούσε να δικαιολογηθεί καθώς οι περισσότερες κατοικίες στην Αθήνα είναι μεγάλης ηλικίας ή ακόμα και από τις υψηλότατες τιμές στο δείγμα μας που εμφάνιζαν ακίνητα μεγάλης ηλικίας, κυρίως σε καλές περιοχές των Αθηνών. Η μεταβλητή « Κατάσταση κατασκευής Χ42» που έλαβε μέρος στο μοντέλο ως ιεραρχική μεταβλητή έχει αρνητικό πρόσημο. Υπενθυμίζουμε ότι η ιεράρχηση ήταν Άριστη=1, Πολύ καλή = 2, Καλή = 3,Μέτρια =4 και Κακή = 5. Δηλαδή όσο μεγαλύτερος ο αριθμός τόσο πιο αρνητική επιρροή έχει στο μοντέλο μας, γεγονός που θεωρείται λογικό. Τέλος η μεταβλητή « Είδος θέρμανσης Χ48» έχει και αυτή αρνητική επίδραση . Συμμετείχε στο μοντέλο ως ιεραρχική μεταβλητή και είχε τιμές Φυσικό αέριο= 1, Πετρέλαιο= 2, Χωρίς θέρμανση = 3. Δηλαδή όσο μεγαλύτερος

ήταν ο αριθμός τόσο αρνητικότερη επίδραση είχε στην εξαρτημένη μεταβλητή, γεγονός που θεωρείται λογικό.

5.4.1.5.1 Έλεγχος Ορθότητας Μοντέλου.

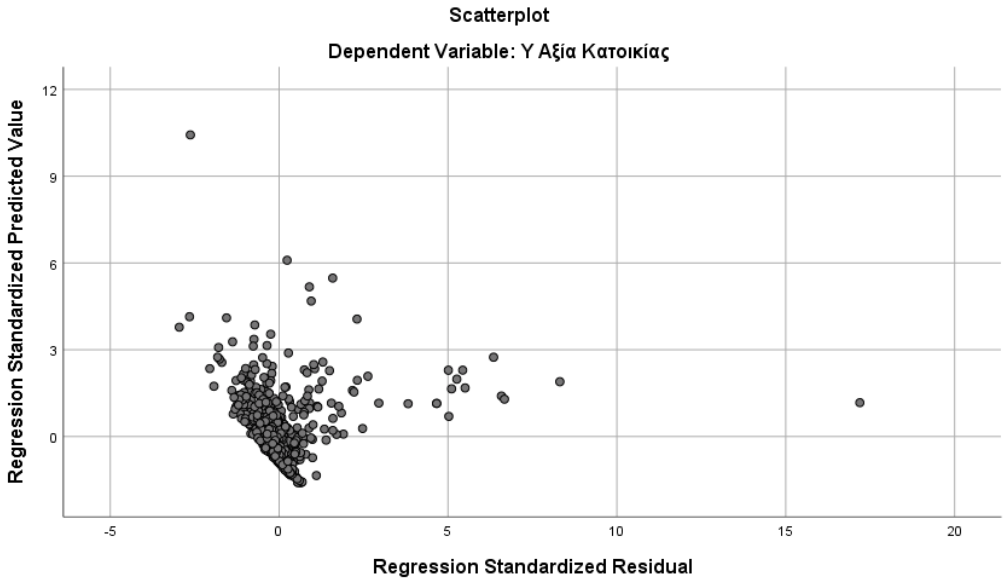
Σ' αυτό το σημείο θα πρέπει να ελέγξουμε την ορθότητα του μοντέλου, δηλαδή το εάν οι παρατηρήσεις μας δίνουν αξιόπιστα αποτελέσματα, προσαρμοζόμενες ικανοποιητικά στο μοντέλο. Εάν ναι τότε ισχυροποιείται η άποψη ότι τα προκύπτοντα συμπεράσματα είναι και αξιόπιστα. Οι υποθέσεις θα πρέπει να γίνουν στα κατάλοιπα, δηλαδή στις αποκλίσεις των τιμών των ανεξάρτητων μεταβλητών από τις αντίστοιχες εκτιμώμενες. Η χρήση των γραφημάτων των σφαλμάτων έναντι των προσαρμοσμένων τιμών που μελετάμε και ο έλεγχος των προτύπων (patterns) που εμφανίζουν γίνεται από τα plots μέσα από το λογισμικό SPSS. Τα γραφήματα σφαλμάτων έναντι των προσαρμοσμένων τιμών και όχι των παρατηρούμενων, είναι γιατί τα σφάλματα συσχετίζονται με τις παρατηρούμενες και όχι με τις προσαρμοσμένες τιμές (Χαλικίας, κ.ά, 2015).

Ομοσκεδαστικότητα υπολοίπων.

Παρατηρούμε στο διάγραμμα σκεδασμού, όπως αυτό έχει προκύψει από την επεξεργασία στην εικόνα 5.5, ότι δεν υπάρχει σταθερή διασπορά, ενώ το σχήμα δεν έχει παντού το ίδιο πάχος με το γράφημα να ανοίγει σαν χωνί (Χαλικιάς κ.ά, 2015) γεγονός που φανερώνει ετεροσκεδαστικότητα. Αυτό υποδεικνύει παραβίαση της βασικής υπόθεσης της παλινδρόμησης ότι η διακύμανση του διαταρακτικού όρου ε_i παραμένει σταθερή, όποιες και να είναι οι τιμές των ερμηνευτικών μεταβλητών. Σημειώνεται ότι η ομοσκεδαστικότητα και η ανεξαρτησία των καταλοίπων συχνά παραβιάζεται στην μελέτη γεωγραφικών δεδομένων (Ηλιοπούλου , 2017)

Εικόνα 5.5

Διάγραμμα σκεδασμού υπολοίπων



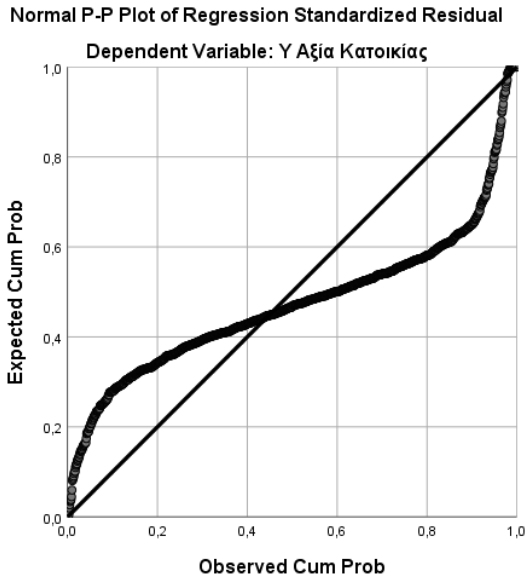
Ακολουθούν οι έλεγχοι της κανονικότητας και ανεξαρτησίας των υπολοίπων.

Κανονικότητα υπολοίπων.

Η κανονικότητα των σφαλμάτων, δηλαδή ότι τα κατάλοιπα (υπολοίπων) πρέπει να ακολουθούν κανονική κατανομή με μέσο 0 και διακύμανση γνωστή, ελέγχεται γραφικά από τα διάγραμμα Q-Q Plot και P-P Plot (Εικόνα 5.6), από το ιστόγραμμα των τυποποιημένων υπολοίπων (εικόνα 5.7) και μέσω του ελέγχου Kolmogorov-Smirnov (Πίνακας 5.14). Η μη ύπαρξη συσχέτισης στα σφάλματα δημιουργεί συνθήκες κανονικής κατανομής. Σημειώνεται ότι τα τυποποιημένα υπόλοιπα (standardized residuals) είναι τα υπόλοιπα διαιρούμενα με τη τυπική τους απόκλιση. Από το ιστόγραμμα δεν προκύπτει ότι υπάρχει κανονική κατανομή καθώς παρατηρούμε εξάρσεις κεντρικά και στήλες εκτός περιοχής στο δεξί μέρος του διαγράμματος. Το ίδιο προκύπτει και από το P-P Plot καθώς η επιθυμία μας τα σημεία να είναι κοντά στην ευθεία, ώστε να αποδειχτεί η κανονικότητα, δεν συμβαίνει στο διάγραμμά μας.

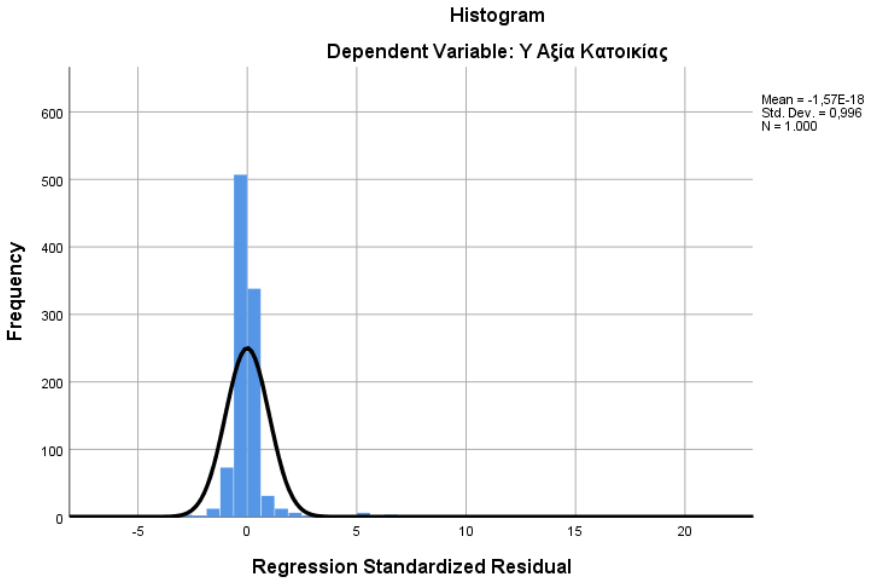
Εικόνα 5.6

Γραφικός έλεγχος κανονικότητας. Διάγραμμα P-P_ PLot



Εικόνα 5.7

Γραφικός έλεγχος κανονικότητας. Ιστόγραμμα υπολοίπων



Τέλος με το Kolmogorov-Smirnov (K-S) Test, η μηδενική υπόθεση είναι ότι τα κατάλοιπα ακολουθούν κανονική κατανομή, ενώ δεν φαίνεται να ικανοποιείται ως απαίτηση καθώς τα p-value=<0.05.

Πίνακας 5.14

Έλεγχος κανονικότητας με Kolmogorov-Smirnov (K-S) Test

Tests of Normality

	Kolmogorov-Smirnov[a]			Shapiro-Wilk		
	Statistic	df	Sig.	Statistic	df	Sig.
Standardized Residual	,251	1000	,000	,501	1000	,000

a. Lilliefors Significance Correction

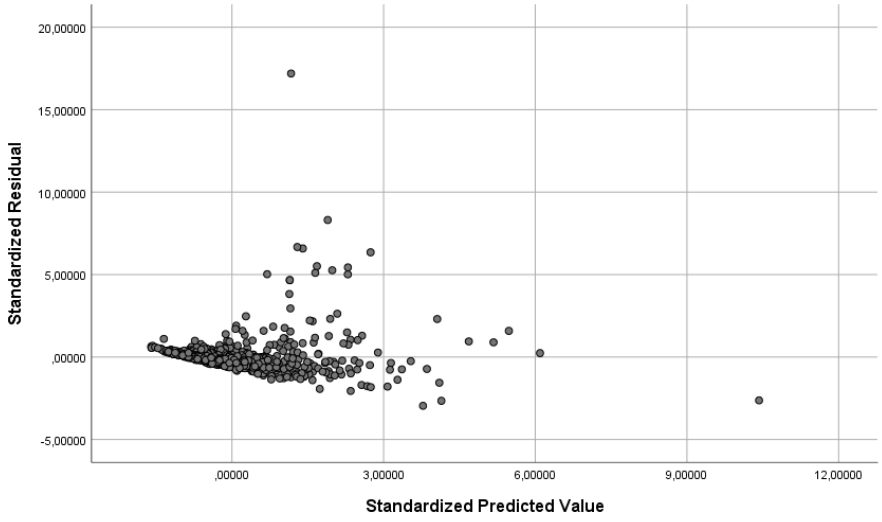
Σημειώνεται ότι σύμφωνα με τον Μπατσίδη (2014) εάν το μέγεθος του δείγματος είναι μεγάλο, λόγω του κεντρικού οριακού θεωρήματος χρησιμοποιούμε την κανονικότητα των σφαλμάτων προσεγγιστικά με την διαφοροποίηση ότι οι κρίσιμες πιθανότητες , δηλαδή οι p τιμές είναι προσεγγιστικές και όχι ακριβείς.

Ανεξαρτησία υπολοίπων.

Εφόσον τα κατάλοιπα εμφανίζουν ανεξαρτησία δεν υπάρχει αυτοσυσχέτιση μεταξύ τους. Εάν δεν εμφανίζουν ανεξαρτησία τότε το πιο πιθανόν είναι να έχει παραληφθεί κάποια μεταβλητή σημαντική για το μοντέλο μας. Όσο αφορά την ανεξαρτησία των υπολοίπων, αυτή βασίζεται στο γεγονός ότι η συμμετοχή μιας οντότητας στο δείγμα δεν πρέπει να επηρεάζει την συμμετοχή μιας άλλης. Γίνεται με γράφημα σκέδασης scatter plot (Εικόνα 5.8). Στο διάγραμμα μας δεν θέλουμε συστηματικές και ανομοιόμορφες συσσωρεύσεις σημείων (clustering), ούτε πρότυπα (patterns). Παρατηρούμε συσσωρεύσεις τιμών μη ομοιόμορφα κατανεμημένων.

Εικόνα 5.8

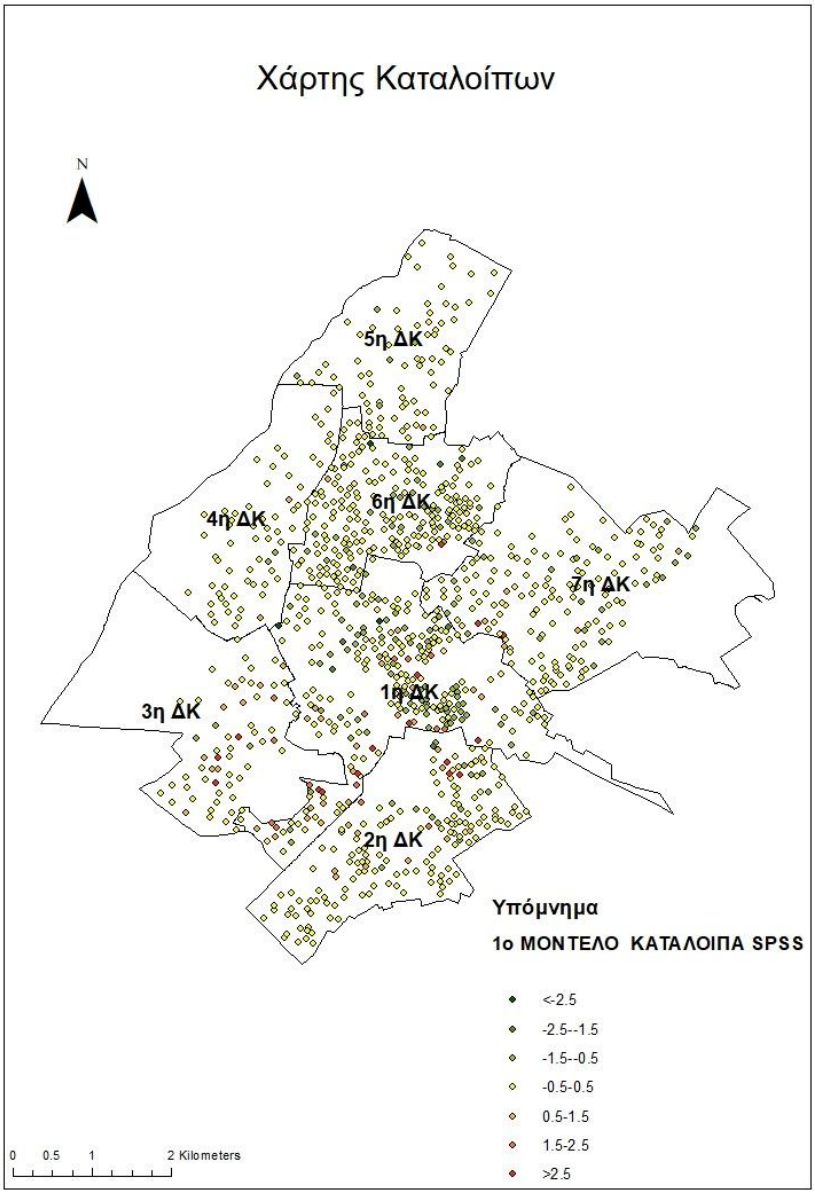
Γραφικός έλεγχος ανεξαρτησίας. Scatter plot residuals / predicted



Στο χάρτη 5.5 φαίνεται η διάταξη των υπολοίπων με την χαρτογράφηση να γίνεται μέσα από το λογισμικό ArcGis.. Η ταξινόμηση των ακινήτων γίνεται με βάση την τιμή των τυποποιημένων υπολοίπων (Std Residuals) σε επτά κλάσεις που αποτελούν ένδειξη για το ποια είναι η κατανομή των υπολοίπων και εάν το μοντέλο ταιριάζει στα δεδομένα. Παρατηρούμε συσσωρεύσεις αρνητικών καταλοίπων κυρίως στην 6η και 1η δημοτική ενότητα, υπονοώντας υπερεκτίμηση των τιμών κατοικιών από το μοντέλο μας .

Χάρτης 5.5

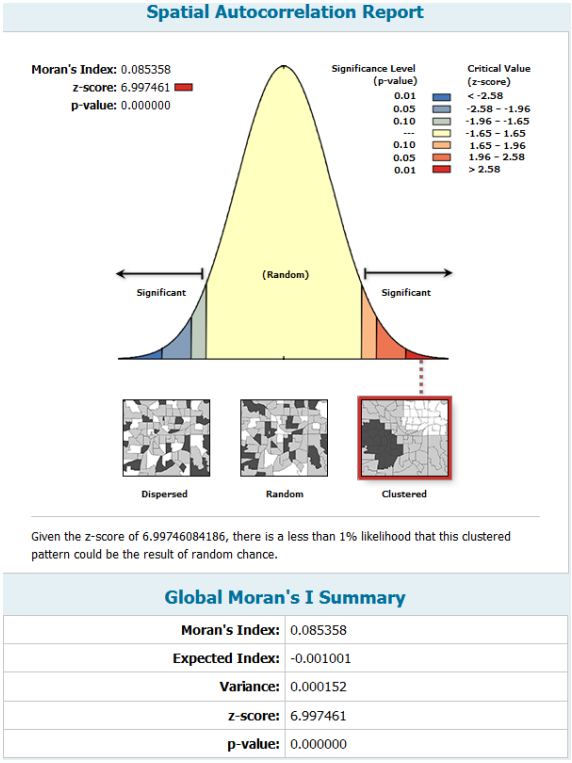
Χάρτης καταλοίπων της παλινδρόμησης μέσω SPSS (πρώτο μοντέλο)



Καθώς δεν αρκεί η παρατήρηση ενός χάρτη για να εξακριβώσουμε την συγκέντρωση ενός φαινομένου στον γεωγραφικό χώρο, πρέπει να είμαστε σε θέση να μετρήσουμε με κάποιο τρόπο αν το παρατηρούμενο χωρικό πρότυπο είναι τυχαίο ή συγκεντρωμένο . Αυτό γίνεται μέσα από το ίδιο λογισμικό υπολογίζοντας τον δείκτη Morans I (Εικόνα 5.9) όπου υποδηλώνεται η ύπαρξη της χωρικής αυτοσυσχέτισης των υπολοίπων.

Εικόνα 5.9

Έλεγχος χωρικής αυτοσυσχέτισης καταλοίπων. Υπολογισμός Morans I .



Global Moran's I Summary

Moran's Index:	0.085358
Expected Index:	-0.001001
Variance:	0.000152
z-score:	6.997461
p-value:	0.000000

5.4.1.5.2 Εξαγωγή Συμπερασμάτων, Αποτελεσμάτων Πρώτου Μοντέλου Μετά από την Επεξεργασία με το Λογισμικό SPSS.

Συμπερασματικά από την δημιουργία του πρώτου μοντέλου μέσω του λογισμικού SPSS διαπιστώνουμε ότι πρόκειται για ένα πολύ μέτριο μοντέλο. Η εξαρτημένη μεταβλητή δεν φαίνεται να ακολουθεί την κανονική κατανομή, ενώ ούτε και τα κατάλοιπα της παλινδρόμησης την ακολουθούν. Τα σφάλματα δεν κατανέμονται με τυχαίο τρόπο, δεν έχουμε ανεξαρτησία, δεν έχουμε ομοσκεδαστικότητα, ενώ δεν μπορεί να αποδειχθεί η τυχαιότητα της μεταβλητής «Αξία_Υ» μέσα από τον κατάλληλο έλεγχο αλλά και μέσα από τον τρόπο λήψης του

δείγματος. Άλλωστε θα πρέπει να λάβουμε υπόψη μας ότι πρόκειται για δείγμα χωρικά κατανεμημένο. Υπάρχουν ακραίες τιμές στο δείγμα μας και στα κατάλοιπα της παλινδρόμησης.

5.4.1.6 Κλασσική παλινδρόμηση (Ordinary Least Squares Estimation-OLS) σε Περιβάλλον GIS για το Πρώτο Μοντέλο.

Τα βασικά βήματα της διαδικασίας δημιουργίας του μοντέλου κλασικής παλινδρόμησης μέσα από το ArcGis δεν διαφέρουν ουσιαστικά από αυτά που πραγματοποιούνται στο στατιστικό πακέτο SPSS. Η λογική είναι η ίδια. Ως πρώτο βήμα στην διαδικασία OLS (Ordinary least square) γίνεται η επιλογή της εξαρτημένης μεταβλητής Y, έπεται η επιλογή των ανεξάρτητων μεταβλητών Xi , ακολουθεί η διερεύνηση των ανεξάρτητων μεταβλητών. Η διαδικασία συνεχίζεται έως την κατασκευή του καταλληλότερου μοντέλου. Καθώς έχουμε ήδη κατασκευάσει το μοντέλο στο SPSS και για να αποφευχθεί όλη αυτή η διαδικασία εισάγονται στο μοντέλο μόνο εκείνες οι μεταβλητές που λειτούργησαν για την δημιουργία του μοντέλου στο SPSS. Με αυτό τον τρόπο θα επιτευχθεί και η σύγκριση των αποτελεσμάτων ώστε να φανούν οι ομοιότητες ή οι διαφορές στις μαθηματικές εξισώσεις των δύο λογισμικών πακέτων. Τα αποτελέσματα της διαδικασίας παρουσιάζονται στον πίνακα 5.15 ενώ στον πίνακα 5.16 παρουσιάζονται τα διαγνωστικά της παλινδρόμησης. Παρατηρούμε ότι το adjusted R^2 =0,397 το F=93,272 και το AIC είναι ίσο με 27582,89. Τα αποτελέσματα (Πίνακας 5.15) ταυτίζονται με το μοντέλο όπως αυτό είχε δημιουργεί με το SPSS και παρουσιάζονται αναλυτικά στο παράρτημα «Α10».

Πίνακας 5.15

Αποτελέσματα OLS διαδικασίας. Συντελεστές παλινδρόμησης

Summary of OLS Results - Model Variables

Variable	Coefficient [a]	StdError	t-Statistic	Probability [b]	Robust_SE	Robust_t	Robust_Pr [b]	VIF [c]
Intercept	8615.690100	38186.598371	0.225621	0.821542	40890.602717	0.210701	0.833161	--------
APOSTASH_A	-204.232190	69.735546	-2.928667	0.003490*	57.265821	-3.566389	0.000393*	1.065205
TIMH_ZWNHS	107.978131	12.341568	8.749142	0.000000*	15.225826	7.091775	0.000000*	1.146146
EMBADON_X3	1519.811760	88.916232	17.092625	0.000000*	180.854154	8.403521	0.000000*	1.094189
OROFOS_X38	8640.136385	3925.482643	2.201038	0.027948*	3982.433022	2.169562	0.030262*	1.078983
HLIKIA_KAT	1626.651781	403.515141	4.031204	0.000067*	702.061157	2.316966	0.020693*	1.500791
KATASTASH_	-42735.08828	6766.685357	-6.315513	0.000000*	11472.535042	-3.724991	0.000218*	1.365357
EIDOS_UERM	-39587.82358	13098.899990	-3.022225	0.002585*	14779.567833	-2.678551	0.007513*	1.402521

Πίνακας 5.16

Διαγνωστικά παλινδρόμησης OLS διαδικασίας

OLS Diagnostics

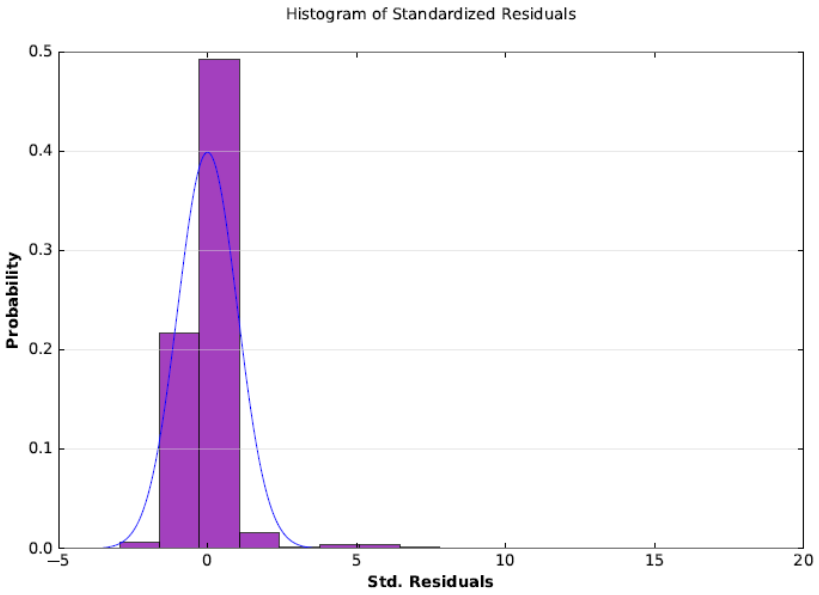
Input Features:	1o__MONTELO_OLS_TELIKA	Dependent Variable:	AXIA_Y
Number of Observations:	1000	Akaike's Information Criterion (AICc) [d]:	27582.898049
Multiple R-Squared [d]:	0.396925	Adjusted R-Squared [d]:	0.392670
Joint F-Statistic [e]:	93.271962	Prob(>F), (7,992) degrees of freedom:	0.000000*
Joint Wald Statistic [e]:	270.363816	Prob(>chi-squared), (7) degrees of freedom:	0.000000*
Koenker (BP) Statistic [f]:	31.169033	Prob(>chi-squared), (7) degrees of freedom:	0.000058*
Jarque-Bera Statistic [g]:	449338.826775	Prob(>chi-squared), (2) degrees of freedom:	0.000000*

Όπως έχει ήδη αναφερθεί μία από τις παραδοχές για την ανάλυση παλινδρόμησης είναι η ανεξαρτησία των καταλοίπων, η οποία όμως συνήθως παραβιάζεται στην ανάλυση των γεωγραφικών φαινομένων λόγω του φαινομένου της χωρικής αυτοσυσχέτισης. Η ανάλυση των καταλοίπων γίνεται αρχικά με τη χαρτογράφησή τους (Χάρτης 5.5). Η ταξινόμηση των ακινήτων γίνεται με βάση την τιμή των τυποποιημένων υπολοίπων (Std Residuals) σε επτά κλάσεις που αποτελούν ένδειξη για το ποια είναι η κατανομή των υπολοίπων και εάν το μοντέλο ταιριάζει στα δεδομένα.

Στην εικόνα 5.10 παρουσιάζεται το ιστόγραμμα αυτών και στην εικόνα 5.11 παρουσιάζεται το διάγραμμα διασποράς τους. Υποδεικνύεται η μη κανονική κατανομή τους.

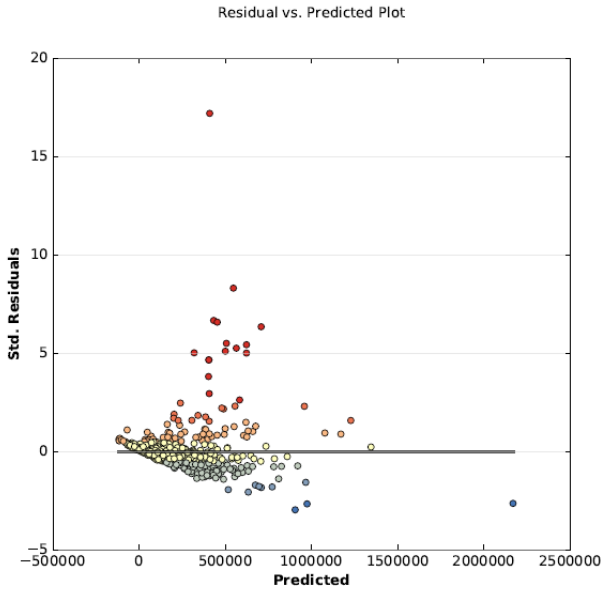
Εικόνα 5.10

Γραφικός έλεγχος κανονικότητας. Ιστόγραμμα καταλοίπων



Εικόνα 5.11

Διάγραμμα διασποράς καταλοίπων



Για τον έλεγχο τη χωρικής αυτοσυσχέτισης των υπολοίπων χρησιμοποιούμε τον δείκτη Morans I (Εικόνα 5.9). Παρατηρούμε ότι τα κατάλοιπα της παλινδρόμησης παρουσιάζουν τιμή στο δείκτη ίση με 0,0853 και z-score ίσο με 6,9974, ενώ το p=0,0000. Παράλληλα παρουσιάζεται φαινόμενο συγκέντρωσης (cluster). Η σημαντικότητα του φαινομένου υποδεικνύεται από την τιμή του z score, όπου

δηλώνεται ότι το φαινόμενο cluster έχει πιθανότητα μικρότερη του 1% να οφείλεται σε τυχαίους παράγοντες. Η τιμή Moran's I είναι θετική αλλά κοντά στο μηδέν και έτσι υποδεικνύεται ότι υπάρχει μικρή συγκέντρωση χωρικά παρόμοιων τιμών. Γενικά τιμές κοντά στο μηδέν υποδεικνύουν έλλειψη χωρικής εξάρτησης μεταξύ των γειτνιαζόντων παρατηρήσεων. Από τα κατάλοιπα της OLS διαδικασίας, με τον υπολογισμό του Moran I, διαπιστώνουμε την ύπαρξη χωρικής αυτοσυσχέτισης, δηλαδή cluster χωρικό πρότυπο. Κρίνεται λοιπόν σκόπιμο να προχωρήσουμε στη γεωγραφικά σταθμισμένη παλινδρόμηση.

5.4.1.7 Χωρική Παλινδρόμηση (Geographically Weighted Regression-GWR) σε Περιβάλλον GIS για το Πρώτο Μοντέλο

Η ιδέα πάνω στην οποία στηρίζεται η γεωγραφικά σταθμισμένη παλινδρόμηση είναι η αντιμετώπιση του προβλήματος της διαφορετικής σημασίας που μπορεί να έχει μια ανεξάρτητη μεταβλητή στις υποπεριοχές της περιοχής μελέτης και έτσι μπορεί να μελετηθεί η γεωγραφική διαφοροποίηση της επίδρασης των ανεξάρτητων μεταβλητών στην εξαρτημένη.

Ανάλογα με την κλασική παλινδρόμηση έτσι και στην γεωγραφικά σταθμισμένη, προκύπτουν οι εκτιμώμενες τιμές για την εξαρτημένη μεταβλητή και τα κατάλοιπα. Οι παράμετροι της έχουν γεωγραφική κατανομή, ενώ δεν είναι σταθεροί για το σύνολο της περιοχής μελέτης. Χαρτογραφώντας τις παραμέτρους μπορεί να προκύψουν χρήσιμα συμπεράσματα για την ερμηνεία της εξαρτημένης μεταβλητής. Η αξιολόγηση του μοντέλου της χωρικής παλινδρόμησης εκτός από το R^2 γίνεται με το κριτήριο Akaike Information Criterion (AIC),το οποίο χρησιμοποιεί τα τετράγωνα των καταλοίπων της παλινδρόμησης , τον αριθμό τν παραμέτρων της εξίσωσης και το μέγεθος του δείγματος και το οποίο συγκρίνει διαφορετικά μοντέλα παλινδρόμησης μεταξύ τους, χρησιμοποιώντας τα τετράγωνα των καταλοίπων της παλινδρόμησης, τον αριθμό των παραμέτρων της εξίσωσης και το μέγεθος του δείγματος. Η μείωση του κριτηρίου αυτού, από την OLS στην GWR, δείχνει βελτίωση του μοντέλου (Ηλιοπούλου, 2017).

Σημαντικό είναι να αναφερθεί ότι στα αποτελέσματα της GWR προκύπτει μία εξίσωση για κάθε σημείο και όχι μία εξίσωση για όλη την περιοχή μελέτης. Η σταθερά της παλινδρόμησης, ο συντελεστής προσδιορισμού R^2 και οι συντελεστές παλινδρόμησης μεταβάλλονται γεωγραφικά από παρατήρηση σε παρατήρηση και

έτσι προκύπτουν εκτιμήσεις των παραμέτρων για όλα τα σημεία της περιοχής μελέτης και όχι μόνο για αυτά στα οποία υπάρχουν δεδομένα.

Η εφαρμογή της GWR στο μοντέλο μας γίνεται μέσω του λογισμικού ArcGis και της κατάλληλης εφαρμογής. Τα αποτελέσματα φαίνονται στον πίνακα 5.17 και αναλυτικά παρουσιάζονται στο παράρτημα «Α11» , ενώ αναλυτικά αποτελέσματα των υπολογισμών μέσω της διαδικασίας για κάθε ακίνητο παρουσιάζονται στο παράτημα «Α12». Παρατηρούμε ότι μέσω της διαδικασίας αυτής το $R^{2,}$ που στην OLS ήταν 0,397, τώρα είναι 0,456, το Adjusted R square που στην OLS ήταν 0,393, τώρα είναι 0,4415, ενώ παράλληλα μειώθηκε και το AICc από 27582,716 που ήταν στην OLS διαδικασία, σε 27503,93 Συνολικά λοιπόν το μοντέλο δείχνει να έχει βελτιωθεί ελαφρώς βελτιωθεί..

Πίνακας 5.17

Αποτελέσματα GWR διαδικασίας Πρώτου Μοντέλου

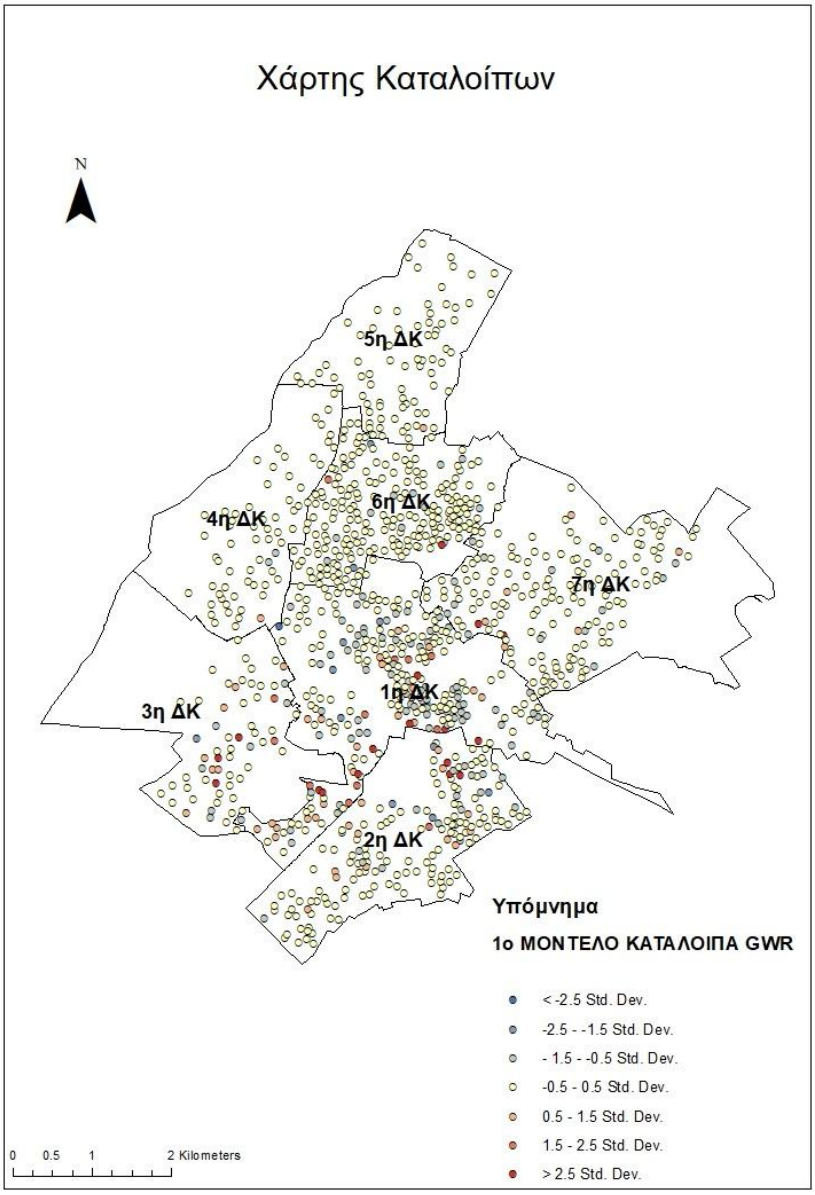
GWR (Geographically weighted regression) result

OID	VARNAME	VARIABLE	DEFINITION
0	Bandwidth	2531.854788	
1	ResidualSquares	49394952231200	
2	EffectiveNumber	27.43202	
3	Sigma	225362.318141	
4	AICc	27503.931746	
5	R2	0.456344	
6	R2Adjusted	0.441569	
7	Dependent Field	0	AXIA_Y
8	Explanatory Field	1	APOSTAS_11
9	Explanatory Field	2	TIMH_ZWNHS
10	Explanatory Field	3	EMBADON_X3
11	Explanatory Field	4	OROFOS_X38
12	Explanatory Field	5	HLIKIA_KAT
13	Explanatory Field	6	KATASTASH_
14	Explanatory Field	7	EIDOS_UERM

Στον χάρτη 5.6 φαίνονται τα τυποποιημένα κατάλοιπα (standardized) της διαδικασίας GWR. Από τον χάρτη αυτό φαίνεται ότι σε σχέση με τον χάρτη 5.6 παρατηρείται μείωση του μεγέθους των καταλοίπων. Το ίδιο προκύπτει και από το ιστόγραμμα (Εικόνα 5.12) παρά το γεγονός ότι εμφανίζονται κεντρικά του διαγράμματος εξάρσεις.

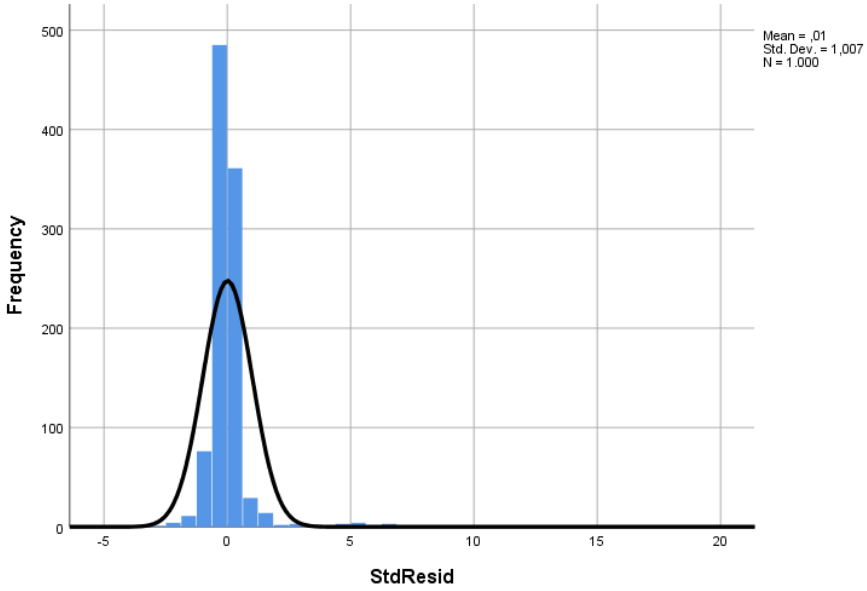
Χάρτης 5.6

Χαρτογράφηση καταλοίπων GWR διαδικασίας πρώτου μοντέλου



Εικόνα 5.12

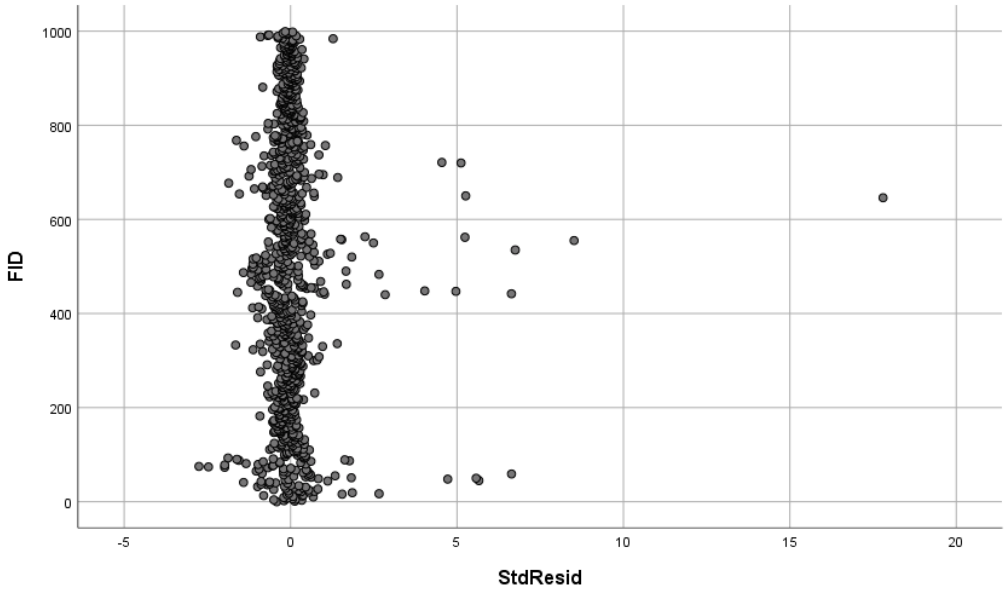
Ιστόγραμμα καταλοίπων GWR διαδικασίας πρώτου μοντέλου



Απεικονίσαμε το scatter plot (Εικόνα 5.13) και υπολογίσαμε τον συντελεστή αυτοσυσχέτισης Moran I (Εικόνα 5.14) για τα κατάλοιπα της GWR διαδικασίας. Στο διάγραμμα διασποράς φαίνεται να υπάρχει ακόμα χωρική συσχέτιση ενώ το ίδιο δείχνει και για τον δείκτη Moran που υποδεικνύει μη τυχαίο φαινόμενο με το z να είναι ίσο με 4,344 και το pattern που δημιουργείται να μην εμφανίζεται τυχαίο.

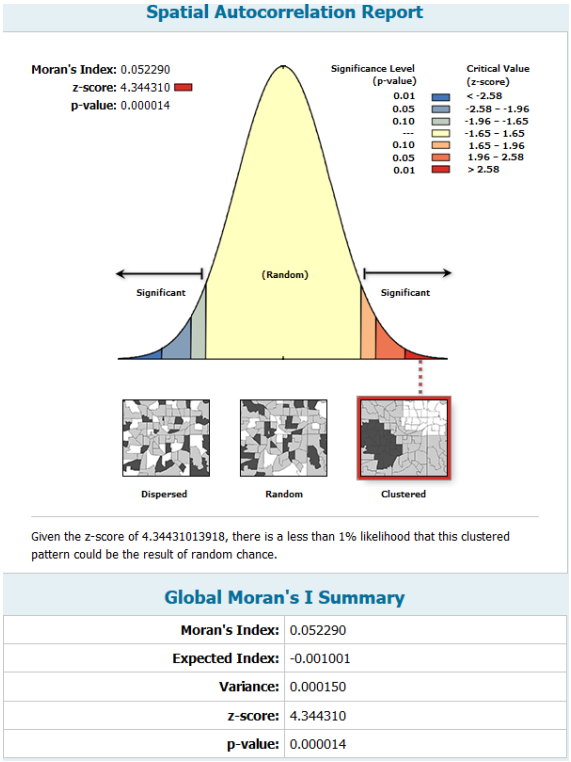
Εικόνα 5.13

Διάγραμμα διασποράς υπολοίπων πρώτου μοντέλου μέσω GWR διαδικασίας



Εικόνα 5.14

Υπολογισμός Morans I για τα κατάλοιπα της GWR διαδικασίας του πρώτου μοντέλου



5.4.1.7.1 Εξαγωγή Συμπερασμάτων Αποτελεσμάτων του Πρώτου Μοντέλου Μετά από την Επεξεργασία με την Χωρική παλινδρόμηση (Geographically Weighted Regression-GWR).

Συμπερασματικά από την επεξεργασία του πρώτου μοντέλου μέσα από την χρήση της χωρικής παλινδρόμησης διαπιστώνουμε ότι αυτό δείχνει να βελτιώθηκε ελαφρά (βλέπε πίνακα 5.18). Ο συντελεστής συσχέτισης βελτιώθηκε λίγο, ενώ μικρές ανάλογες ενδείξεις λάβαμε και από το κριτήριο Akaike Information Criterion (AIC). Παρατηρείται μείωση του μεγέθους των καταλοίπων, ενώ φαίνεται να υπάρχει ακόμα χωρική συσχέτιση καθώς ο δείκτης Morans υποδεικνύει μη τυχαίο φαινόμενο.

Πίνακας 5.18

Σύγκριση OLS & GWR διαδικασίας πρώτου μοντέλου .

Συντελεστές	OLS	GWR
Adjusted R Square	0,393	0,4416

Morans I residuals	0,0853	0,0522
AICc	27582,716	27503.93

Είναι φανερό λοιπόν ότι από το συγκεκριμένο μοντέλο δεν προέκυψε καλύτερη ερμηνεία της εξαρτημένης μεταβλητής και ακριβέστερος προσδιορισμός του μοντέλου της παλινδρόμησης.

Ως λύση στο πρόβλημα θα μπορούσε να προταθεί η επεξεργασία του μοντέλου με ανάλυση κατά συστάδες έτσι ώστε να ταξινομηθούν σύνολα παρατηρήσεων με παρόμοια χαρακτηριστικά ανήκοντα στην ίδια ομάδα. Καθώς η παραγοντική ανάλυση ομαδοποιεί παρόμοιες μεταβλητές, η ανάλυση συστάδων ομαδοποιεί παρόμοιες παρατηρήσεις με βάση κάποιο μέτρο ομοιότητας(Ηλιοπούλου, 2017). Η διαδικασία αυτή είναι αρκετά περίπλοκη και δεν άπτεται της παρούσα μελέτης.

5.4.2 Δημιουργία Δεύτερου Μοντέλου

Στο δεύτερο μοντέλο της παρούσας εργασίας ακολουθείται η διαδικασία της δημιουργίας εξίσωσης παλινδρόμησης αφού πρώτα αφαιρεθούν οι ακραίες οντότητες σε επιλεγμένες μεταβλητές του δείγματος. Απομακρύνονται οι ανεξάρτητες μεταβλητές που ουσιαστικά περιέχουν και επαναλαμβάνουν πληροφορία που υπάρχει σε μια άλλη μεταβλητή. Λαμβάνεται υπόψη το κεντρικό οριακό θεώρημα ώστε να αντιμετωπιστεί τα πρόβλημα της κανονικότητας στην μεταβλητή «Αξία_Υ», ενώ το πρόβλημα της τυχαιότητας αντιμετωπίζεται στα πλαίσια της χωρικής ανάλυσης των δεδομένων έχοντας υπόψη την χωρική αυτοσυσχέτιση της αξίας των ακινήτων του δείγματος. Στο μοντέλο αυτό εισάγονται οι ανεξάρτητες μεταβλητές με την διαδικασία Stepwise, η οποία ως η πιο σύνθετη διαδικασία από όλες συνδυάζει την Forward και την Backward μέθοδο, αντιμετωπίζοντας τα μειονεκτήματά τους. Η μέθοδος αυτή εντοπίζει την καλύτερη μεταβλητή για να εισέλθει στο μοντέλο και αφού αυτό γίνει, διενεργεί ελέγχους μήπως πρέπει να αφαιρεθεί από αυτό. Επαναλαμβάνει την διαδικασία για όλες τις ανεξάρτητες μεταβλητές. Πρακτικά ξεκινάει με την μέθοδο Forward και συνεχίζει με την Backward.

5.4.2.1 Προκαταρτικοί Έλεγχοι

Όπως είναι γνωστό βασικές προϋποθέσεις για την χρήση της στατιστικής είναι η τυχαιότητα του δείγματος, στην περίπτωση αυτή η τυχαιότητα της εξαρτημένης

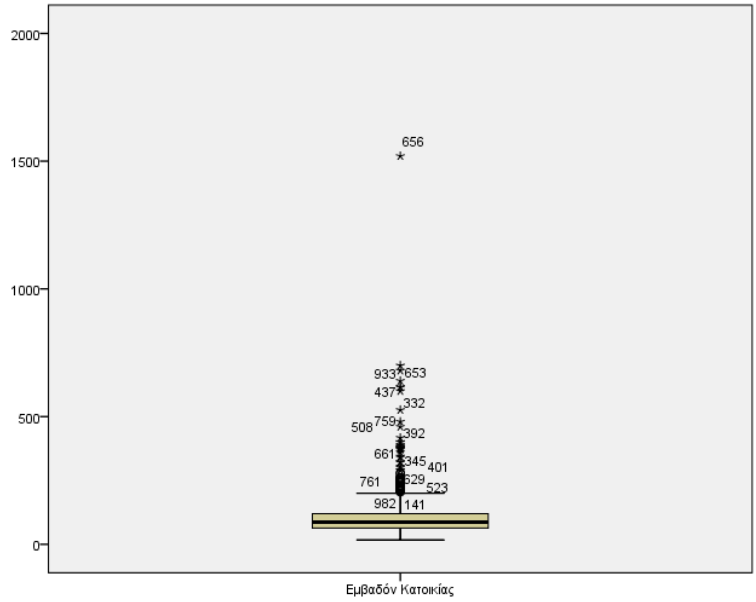
μεταβλητής «Αξία_Υ», η μη ύπαρξη ακραίων τιμών στις παρατηρήσεις μας και η κανονική κατανομή του πληθυσμού από τον οποίο προήλθε το δείγμα μας. Σύμφωνα με το Κεντρικό Οριακό Θεώρημα το άθροισμα και επομένως η μέση τιμή μεγάλου αριθμού ανεξάρτητων παρατηρήσεων ακολουθεί κανονική κατανομή, ανεξάρτητα από το ποια κατανομή ακολουθούν οι παρατηρήσεις. Στην συγκεκριμένη μελέτη το δείγμα είναι μεγάλο και ως εκ τούτου θεωρούμε, για την κατασκευή του δεύτερου μοντέλου, ότι δεν εμποδίζει την ανάλυση η μη κανονικότητα ορισμένων μεταβλητών. Επιπρόσθετα η χωρική αυτοσυσχέτιση παραβιάζει ορισμένες παραδοχές των στατιστικών ελέγχων, όπως η υπόθεση τυχαιότητας για την εκτίμηση των παραμέτρων, ενώ συμβάλει στο πλεονασμό των δεδομένων. Τέλος η προϋπόθεση της ύπαρξης του τυχαίου δείγματος συνήθως παραβιάζεται στην ανάλυση των γεωγραφικών φαινομένων λόγω της συνέχειας του γεωγραφικού χώρου και του φαινομένου της χωρικής αυτοσυσχέτισης (Ηλιοπούλου, 2017).Το πρόβλημα αυτό αντιμετωπίζεται στην δόμηση του 2ου μοντέλου με μεθόδους χωρικής παλινδρόμησης.

5.4.2.2 Έλεγχος και Αφαίρεση Ακραίων Παρατηρήσεων

Η ύπαρξη ακραίων τιμών στο δείγμα παραβιάζει μια από τις βασικές προϋποθέσεις του συντελεστή συσχέτισης Pearson . Εξάλλου όταν υπάρχουν ακραίες τιμές (outliers) στις μεταβλητές, ο συντελεστής συσχέτισης μπορεί να δώσει παραπλανητικά αποτελέσματα και έτσι είναι σκόπιμο να ελεγχθεί χωρίς τις ακραίες ή τις απομονωμένες αυτές τιμές (Ηλιοπούλου, 2017). Σε προηγούμενη ενότητα εξετάστηκε και παρουσιάστηκε σε εικόνα (Εικόνα 5.4) η ύπαρξη ακραίων τιμών στην εξαρτημένη μεταβλητή. Στις παρακάτω εικόνες (Εικόνα 5.15 , 5.16) παρουσιάζονται τα box plot δύο ακόμα σημαντικών ανεξάρτητων μεταβλητών «Εμβαδόν κατοικίας», «Τιμή / τμ», όπου και σε αυτές βλέπουμε την ύπαρξη πολλών ακραίων τιμών. Στην ροή της διαδικασίας γίνεται η αφαίρεση των ακραίων τιμών από τις μεταβλητές «Αξία», « Εμβαδόν κατοικίας» και « Τιμή / τμ» η οποία γίνεται συνδυαστικά και σταδιακά. Σκοπός μας είναι το μοντέλο της παλινδρόμησης να έχει την καλύτερη δυνατή μορφή και αποτέλεσμα.

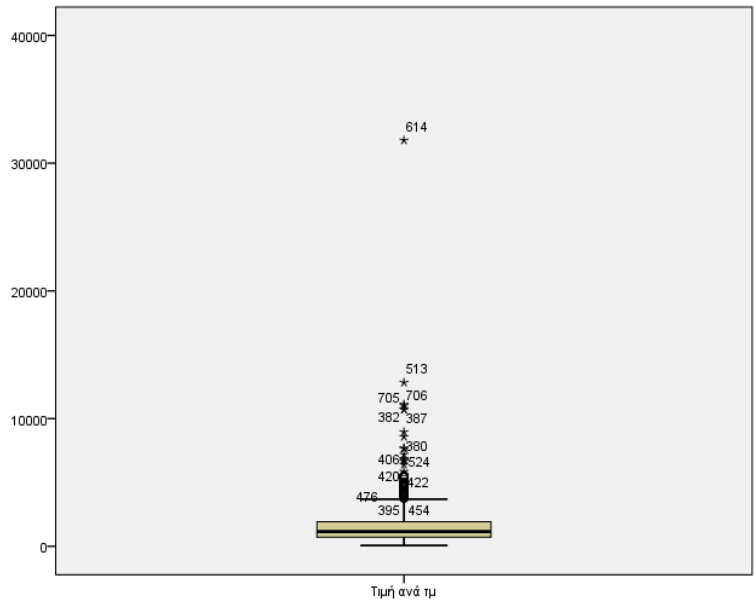
Εικόνα 5.15

Έλεγχος ακραίων παρατηρήσεων στην μεταβλητή Εμβαδόν κατοικίας



Εικόνα 5.16

Έλεγχος ακραίων παρατηρήσεων στην μεταβλητή Τιμή / τμ



Μετά από πληθώρα δοκιμών καταλήγουμε στην αφαίρεση 145 παρατηρήσεων . Πιο συγκεκριμένα αφαιρούμε τιμές της μεταβλητής « Αξία» από 400000 ευρώ και πάνω, τιμές της μεταβλητής « Εμβαδόν κατοικίας « από 205 τμ και πάνω και τιμές της μεταβλητής « τιμή /τμ» από 3333 ευρώ και πάνω.

5.4.2.3 Σύγκριση Βασικών Περιγραφικών Μέτρων

Μετά την αφαίρεση των ακραίων τιμών παρατηρούμε τις διαφορές των στατιστικών μέτρων μεταξύ του αρχικού συνολικού δείγματος και αυτού, όπως έχει διαμορφωθεί πλέον, μετά την αφαίρεση των outliers τιμών.

Παρατηρούμε ότι η μεταβλητή « Αξία» είχε μέση τιμή 186833,20, διάμεσο 10500,00 και κορυφή 45000 στο αρχικό δείγμα, ενώ μετά την αφαίρεση των ακραίων τιμών έχει μέση τιμή 107242,34, διάμεσο 85000 και κορυφή 45000. Το εμβαδόν κατοικίας είχε αρχικά μέση τιμή 106,11τμ, διάμεσο 87,00τμ και κορυφή 50τμ, ενώ πλέον έχει 85,85τμ , 82,00τμ και 75τμ αντίστοιχα. Πλήρη αναφορά στοιχείων βρίσκεται στον πίνακα 5.19.

Πίνακας 5.19

Σύγκριση βασικών περιγραφικών μέτρων

Μεταβλητή	δείγμα 1000	δείγμα 855	δείγμα 1000	δείγμα 855	δείγμα 1000	δείγμα 855
	Mean		Median		Mode	
Αξία κατοικίας	186833,20	107242,34	10500,00	85000	45000	45000
Εμβαδόν κατοικίας	106,11	85,85	87,00	82,00	50	75
Τιμή ανά τμ	1572,54	1220,30	1151,92	1037,50	1000	1000
Αριθμός υπνοδωματίων	2,03	1,86	2,00	2,00	2	2
Αριθμός μπάνιων	1,13	1,06	1,00	1,00	1	1
Όροφος	2,43	2,46	2,00	2,00	1	1
Ηλικία κατασκευής	42,10	40,00	46,00	45,00	48	48

5.4.2.4 Προσαρμογή Μοντέλου Γραμμικής Παλινδρόμησης

Παρατηρούμε ότι υπάρχουν μερικές εμφανέστατες συσχετίσεις μεταξύ ανεξάρτητων μεταβλητών. Έτσι προκύπτει ότι πρέπει να αφαιρεθούν οι μεταβλητές « Έτος κατασκευής Χ39» καθώς από αυτή προκύπτει και έχει απόλυτη συσχέτιση με την μεταβλητή «Ηλικία κατασκευής» και η μεταβλητή « Ποσοστό αλλοδαπών Χ5» καθώς έχει απόλυτη αρνητική συσχέτιση με την μεταβλητή « Ποσοστό Ελλήνων». Τέλος θα πρέπει να αφαιρεθεί η μεταβλητή « Τιμή ανά τ.μ. Χ33» καθώς ουσιαστικά αποτελεί την διαίρεση της εξαρτημένης με την μεταβλητή « Εμβαδόν ακινήτου» .

Εισάγουμε όλες τις υπόλοιπες μεταβλητές στο πρόγραμμα. Οι ανεξάρτητες μεταβλητές που λαμβάνουν μέρος στην επεξεργασία είναι οι: Χ4, Χ6, Χ7, Χ8, Χ9, Χ10, Χ11, Χ12, Χ13, Χ14, Χ15, Χ16, Χ17, Χ18, Χ19, Χ20, Χ21, Χ22, Χ23, Χ24, Χ25, Χ26, Χ27, Χ28, Χ29, Χ30, Χ31, Χ32, Χ34, Χ35, Χ36, Χ37, Χ38, Χ40, Χ41, Χ42, Χ43, Χ44, Χ45, Χ46, Χ47, Χ48, Χ49, Χ50, Χ51, Χ52, Χ53

Η μέθοδος επεξεργασίας για το συγκεκριμένο μοντέλο, μέσω του στατιστικού πακέτου SPSS είναι η Stepwise. Πρόκειται για μια παραλλαγή της διαδικασίας Forward. Μας επιτρέπει να επανεξετάσουμε αν κάποια από τις μεταβλητές, εκτός της νεοεισαχθείσας, μπορεί να αφαιρεθεί χρησιμοποιώντας τη διαδικασία Backward. Στο πρώτο βήμα της διαδικασίας αυτής ορίζεται ένα μέγιστο και ένα ελάχιστο p-value για την αφαίρεση και την είσοδο αντίστοιχα κάποιας μεταβλητής, τέτοιο ώστε να εξασφαλίζεται ότι καμία μεταβλητή που θα αφαιρεθεί δεν θα επανατοποθετηθεί στο μοντέλο στο ίδιο βήμα της διαδικασίας. Σε δεύτερη φάση υπολογίζονται τα partial F-tests για κάθε υποψήφια μεταβλητή με όσες είναι ήδη μέσα στο μοντέλο. Εάν το χαμηλότερο p-value είναι χαμηλότερο από το αντίστοιχο της εισόδου τότε εισέρχεται στο μοντέλο η μεταβλητή. Σε τρίτη φάση υπολογίζονται όλα τα partial F-tests για όλες τις μεταβλητές. Εφόσον το μεγαλύτερο p-value είναι και μεγαλύτερο από το αντίστοιχο p-value της εξόδου τότε η μεταβλητή απομακρύνεται. Σε επόμενο στάδιο η διαδικασία πηγαίνει ξανά στην δεύτερη φάση και επαναλαμβάνεται μέχρι να μην είναι κατάλληλη για προσθήκη ή απαλοιφή καμία μεταβλητή του σετ (Χαλικίας, κ.ά 2015 σελ 249). Τα αναλυτικά βήματα της διαδικασίας περιγράφονται στο παράρτημα «Β2» ενώ παρατίθεται η καλύτερη λύση της διαδικασίας. Η διαδικασία ολοκληρώνεται στο 18ο μοντέλο που δημιουργείται. Τα διαδοχικά βήματα με τα αποτελέσματα του R^2 παρουσιάζονται στον παρακάτω συνοπτικό πίνακα (Πίνακας 5.20) . Όπως έχουμε αναφέρει το R είναι η απόλυτη τιμή του συντελεστή γραμμικής συσχέτισης, ενώ το R Square (R^2) είναι το τετράγωνο του και ονομάζεται συντελεστής προσδιορισμού, δείχνοντας το ποσοστό μεταβλητότητας των δεδομένων που εξηγείται από το γραμμικό μοντέλο που προσαρμόστηκε, έχοντας λάβει υπόψη του και το μέγεθος του δείγματος. Στο 18ο αποτέλεσμα του μοντέλου που δημιουργήθηκε το R= 0,810 ενώ ο R^2 είναι ίσο με 0,656. Παράλληλα το Std. Error of the Estimate είναι 44981,821.

Πίνακας 5.20

Γραμμική Παλινδρόμηση: Συντελεστές προσδιορισμού

Model Summary[s]

Model	R	R Square	Adjusted R Square	Std. Error of the Estimate
18	,810[r]	,656	,650	44981,821

r. Predictors: (Constant), X32 Εμβαδόν Κατοικίας, X42 Κατάσταση Κατασκευής, X31 Τιμή ζώνης, X47 Ύπαρξη Θέσης στάθμευσης, X41 Είδος Κατοικίας, X38 Όροφος, X48 Είδος Θέρμανσης, X51 Ύπαρξη Θέας, X35 Αριθμός Μπάνιων, X16 Απόσταση από Νεκροταφείο, X23 Απόσταση από Πράσινο, X40 Ηλικία Κατασκευής, X8 Ποσοστό Πρασίνου, X17 Απόσταση από Νοσοκομείο

s. Dependent Variable: Y Αξία Κατοικίας

Το σύνολο της διακύμανσης του δείγματος είναι ίσο με 4943984627321.640. Η διακύμανση της παλινδρόμησης είναι ίση με 3244358700497.380 και η διακύμανση του λάθους (Residual) είναι ίση με 1699625926824.270. Από τον πίνακα ANOVA μας ενδιαφέρει το F, το οποίο στην περίπτωσή μας είναι ίσο με 114,532, ενώ το Sig. <0,05, και η γραμμική παλινδρόμηση φαίνεται στατιστικά σημαντική (Πίνακας 5.21) .

Πίνακας 5.21

Γραμμική παλινδρόμηση: Ανάλυση διασποράς

ANOVA[a]

Model		Sum of Squares	df	Mean Square	F	Sig.
18	Regression	3244358700497,371	14	231739907178,384	114,532	,000[s]
	Residual	1699625926824,274	840	2023364198,600		
	Total	4943984627321,645	854			

Οι μεταβλητές που λαμβάνουν μέρος στο μοντέλο είναι οι : Εμβαδόν Κατοικίας X32, Κατάσταση Κατασκευής X42, Τιμής Ζώνης X31, Ύπαρξη Θέσης στάθμευσης X47, Είδος Κατοικίας X41, Όροφος X38, Είδος Θέρμανσης X48, Ύπαρξη Θέας X51, Αριθμός Μπάνιων X35 , Απόσταση από Νεκροταφείο X16, Απόσταση από Πράσινο X23, Ηλικία Κατασκευής X40, Ποσοστό Πρασίνου X8, Απόσταση από Νοσοκομείο X17.

Στον πίνακα 5.22 παρουσιάζονται τα αποτελέσματα της παλινδρόμησης με την διαδικασία stepwise. Όλες οι μεταβλητές που χρησιμοποιηθήκαν στο μοντέλο δείχνουν σημαντικές και το δείγμα φαίνεται να έχει καλή προσαρμογή.

Πίνακας 5.22

Αποτελέσματα παλινδρόμησης: Συντελεστές προσδιορισμού

Coefficients[a]

Model		Unstandardized Coefficients		Standardized Coefficients	t	Sig.
		B	Std. Error	Beta		
18	(Constant)	61155,984	21113,564		2,897	,004
	X32 Εμβαδόν Κατοικίας	987,702	49,591	,447	19,917	,000
	X42 Κατάσταση Κατασκευής	-9311,769	1496,072	-,151	-6,224	,000
	X31 Τιμή ζώνης	40,261	3,733	,252	10,784	,000
	X47 Ύπαρξη Θέσης στάθμευσης	18862,627	5036,792	,091	3,745	,000
	X41 Είδος Κατοικίας	-27373,447	4844,988	-,122	-5,650	,000
	X38 Όροφος	3952,767	877,296	,099	4,506	,000
	X48 Είδος Θέρμανσης	-11447,851	2911,343	-,098	-3,932	,000
	X51 Ύπαρξη Θέας	17508,945	4612,597	,084	3,796	,000
	X35 Αριθμός Μπάνιων	19515,449	6437,140	,065	3,032	,003
	X16 Απόσταση από Νεκροταφείο	-9,086	1,664	-,123	-5,461	,000
	X23 Απόσταση από Πράσινο	-37,590	14,439	-,055	-2,603	,009
	X40 Ηλικία Κατασκευής	-263,667	100,650	-,071	-2,620	,009
	X8 Ποσοστό Πρασίνου	1622,009	220,051	,204	7,371	,000
	X17 Απόσταση από Νοσοκομείο	14,723	4,284	,092	3,437	,001

a. Dependent Variable: Y Αξία Κατοικίας

<u>Η τελική εξίσωση παλινδρόμησης</u> που προκύπτει μετά από την παραπάνω επεξεργασία είναι :

Y= 61155,984 + 987,702 (X32) – 9311,769 (X42) + 40,261 (X31) + 18862,627 (X47) -27373,447 (X41) + 3952,767 (X38) – 11447,851 (X48) +17508,945 (X51) + 19515,449(X35) – 9,086 (X16) – 37,590 (X23) – 263,667 (X40) + 1622,009 (X8) + 14,723 (X17)

Με μια πρώτη ματιά τα αποτελέσματα κρίνονται λογικά. Το αρνητικό πρόσημο στην μεταβλητή «Κατάσταση κατασκευής X42» οφείλεται στον αρχικό προσδιορισμό της ιεράρχησης των τιμών ως: άριστη κατάσταση= 1, πολύ καλή κατάσταση=2, καλή κατάσταση= 3, μέτρια κατάσταση= 4, κακή κατάσταση =5. Το

αρνητικό πρόσημο στην μεταβλητή « Είδος θέρμανσης Χ48» οφείλεται και αυτό στον αρχικό προσδιορισμό της ιεράρχησης των τιμών ως Αυτόνομη=1, Κεντρική =2 και Χωρίς θέρμανση =3.

Το αρνητικό πρόσημο στην μεταβλητή «Απόσταση από πράσινο Χ23» και στην μεταβλητή « Ηλικία κατασκευής Χ40» κρίνονται και αυτά λογικά, καθώς όσο μεγαλώνουν οι τιμές τόσο περισσότερο αρνητικό ρόλο παίζουν στο μοντέλο. Προβληματισμό προκαλεί το αρνητικό πρόσημο της μεταβλητής « Απόσταση από Νεκροταφείο Χ16» καθώς υποδεικνύει ότι όσο μεγαλώνει η απόσταση από Νεκροταφείο τόσο αρνητικό ρόλο παίζει αυτή στο μοντέλο μας. Βέβαια εδώ θα πρέπει να σημειωθεί ότι πιθανόν έχει σημασία το γεγονός ότι η περιοχή στην οποία βρίσκεται το Πρώτο Νεκροταφείο θεωρείται μια από τις καλές περιοχές της Αθήνας.

Παρατηρούμε ότι σημαντικό θετικό ρόλο έχει η μεταβλητή «Ύπαρξη θέσης στάθμευσης Χ47», η μεταβλητή «Ύπαρξη θέας Χ51» και η μεταβλητή « Αριθμός μπάνιων Χ35» καθώς και η μεταβλητή « Είδος κατοικίας Χ41». Εδώ σημειώνεται ότι στην ιεράρχηση της μεταβλητής αυτής, οι τιμές που δόθηκαν ήταν μονοκατοικία =1, μεζονέτα =2, διαμέρισμα= 3, γεγονός που η λογική υποδεικνύει ότι λειτουργεί ορθά στο μοντέλο μας.

5.4.2.4.1 Έλεγχος ορθότητας μοντέλου.

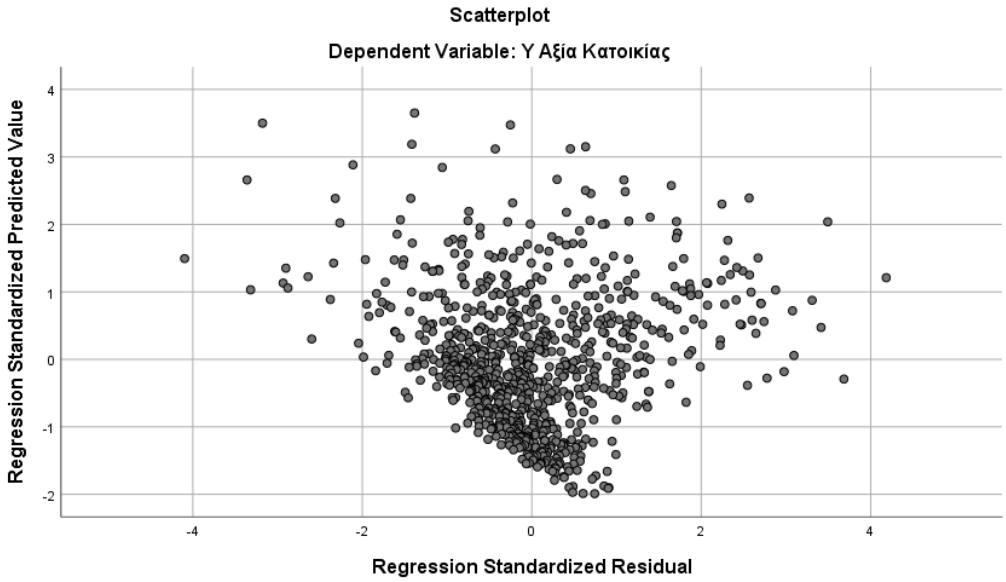
Έπεται ο έλεγχος της ορθότητας του μοντέλου, αναφορικά με την αξιοπιστία των αποτελεσμάτων που δίνει .και την αξιόπιστη προσαρμογή των παρατηρήσεων σε αυτό. Οι υποθέσεις γίνονται στα κατάλοιπα.

Ομοσκεδαστικότητα υπολοίπων.

Παρατηρούμε στο διάγραμμα σκεδασμού, όπως αυτό έχει προκύψει από την επεξεργασία στην εικόνα 5.17, ότι δεν υπάρχει σταθερή διασπορά και τα σφάλματα δεν κατανέμονται με τυχαίο τρόπο, γεγονός που φανερώνει την έλλειψη ομοσκεδαστικότητας. Όπως έχει προαναφερθεί η ομοσκεδαστικότητα και η ανεξαρτησία των καταλοίπων συχνά παραβιάζεται στην μελέτη γεωγραφικών δεδομένων (Ηλιοπούλου, 2017)

Εικόνα 5.17

Διάγραμμα διασποράς υπολοίπων δεύτερου μοντέλου μέσω SPSS.

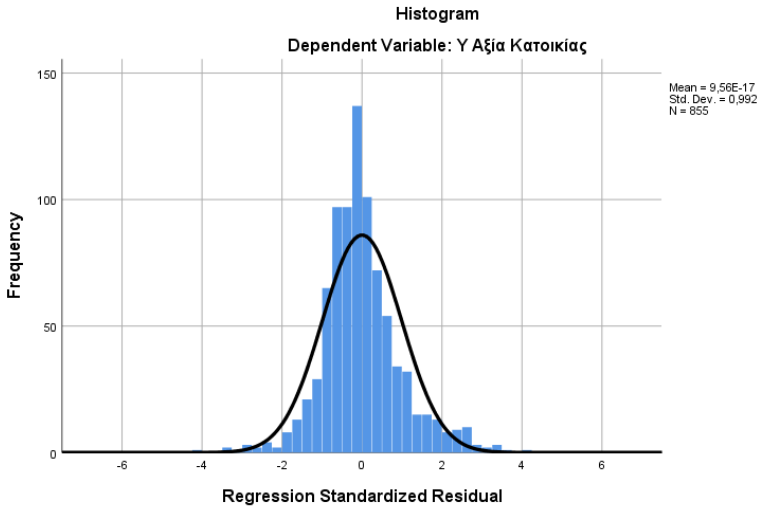


Κανονικότητα υπολοίπων.

Από το ιστόγραμμα (Εικόνα 5.18) δεν προκύπτει η ύπαρξη κανονικής κατανομής καθώς παρατηρούμε εξάρσεις κεντρικά και στήλες εκτός περιοχής στο δεξί μέρος του διαγράμματος. Το ίδιο προκύπτει και από το P-P Plot (Εικόνα 5.19) καθώς η επιθυμία μας τα σημεία να είναι κοντά στην ευθεία, ώστε να αποδειχτεί η κανονικότητα, δεν συμβαίνει στο διάγραμμά μας. Τέλος με το Kolmogorov-Smirnov (K-S) Test (Πίνακας 5.23) όπου η μηδενική υπόθεση είναι ότι τα κατάλοιπα ακολουθούν κανονική κατανομή, δεν φαίνεται αυτή να ικανοποιείται ως απαίτηση καθώς τα p-value=<0.05.

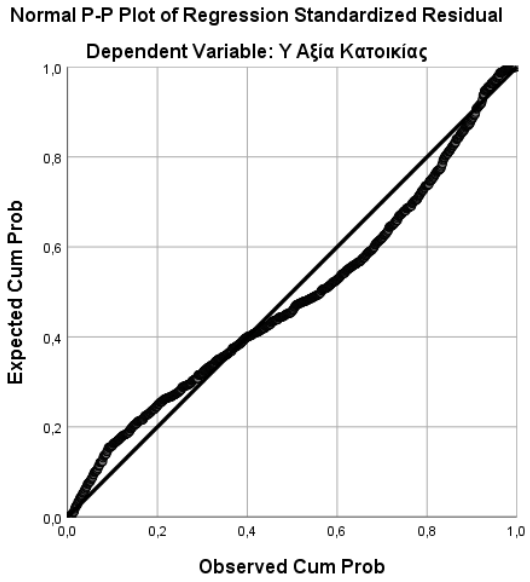
Εικόνα 5.18

Ιστόγραμμα Υπολοίπων μέσω SPSS δεύτερου μοντέλου



Εικόνα 5.19

P-P plot υπολοίπων δεύτερου μοντέλου μέσω SPSS



Πίνακας 5.23

Kolmogorov-Smirnov (K-S) Test για τα κατάλοιπα. δεύτερου μοντέλου μέσω SPSS

Tests of Normality

	Kolmogorov-Smirnov[a]			Shapiro-Wilk		
	Statistic	df	Sig.	Statistic	df	Sig.
Standardized Residual	,085	855	,000	,959	855	,000

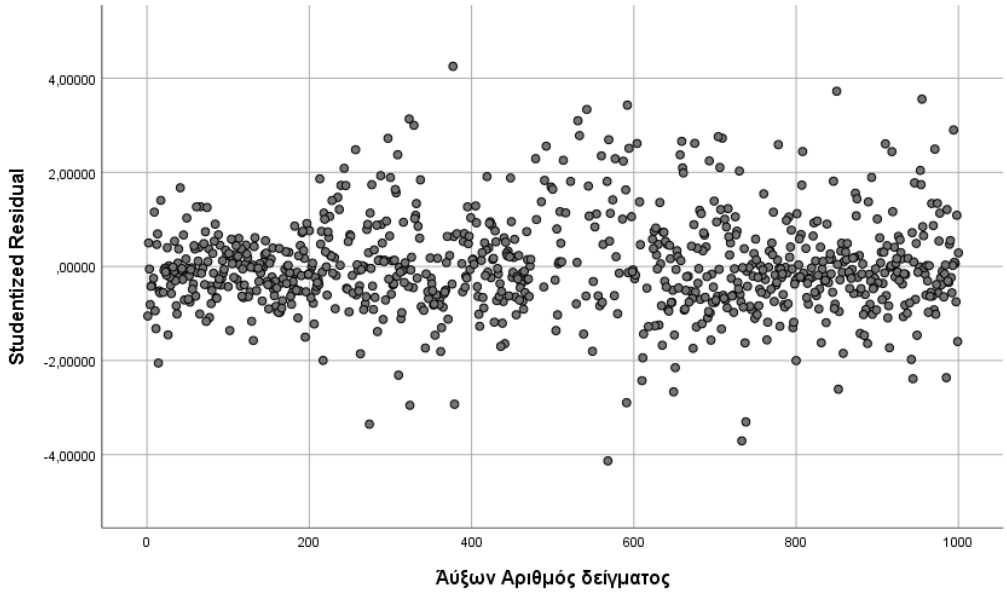
a. Lilliefors Significance Correction

Ανεξαρτησία υπολοίπων.

Παρατηρούμε ότι οι τιμές στο γράφημα σκέδασης scatter plot (Εικόνα 5.20) κυμαίνονται εκατέρωθεν της γραμμής του μηδενός ομοιόμορφα εκτός ελαχίστων εξαιρέσεων. Επιπρόσθετα από τον έλεγχο ροών (Πίνακας 5.24) και επιλέγοντας σαν μεταβλητή τα studentized residuals παρατηρούμε το $p<0.05$ άρα απορρίπτουμε την μηδενική υπόθεση ότι τα κατάλοιπα είναι τυχαία. Σημειώνεται ότι τα κατάλοιπα συμπεριφέρονται καλύτερα ως προς την κανονικότητα και την ομοσκεδαστικότητα σε σχέση με το πρώτο μοντέλο.

Εικόνα 5.20

Διάγραμμα σκεδασμού υπολοίπων δεύτερου μοντέλου μέσω SPSS



Πίνακας 5.24

Run test των Studentized residuals δεύτερου μοντέλου μέσω SPSS

	Studentized Residual
Test Value[a]	-,10589
Cases « Test Value	427
Cases «= Test Value	428
Total Cases	855
Number of Runs	350
Z	-5,372
Asymp. Sig. (2-tailed)	,000

a. Median

5.4.2.4.2 Εξαγωγή Συμπερασμάτων Αποτελεσμάτων Δεύτερου Μοντέλου μετά την Επεξεργασία με το Λογισμικό SPSS.

Συμπερασματικά από την δημιουργία του δεύτερου μοντέλου μέσα από το λογισμικό SPSS διαπιστώνουμε ότι πρόκειται για ένα αρκετά καλό μοντέλο. Παρόλα αυτά η εξαρτημένη μεταβλητή δεν φαίνεται να ακολουθεί την κανονική κατανομή, ενώ ούτε και τα κατάλοιπα της παλινδρόμησης την ακολουθούν. Τα σφάλματα δεν κατανέμονται με τυχαίο τρόπο, δεν έχουμε ομοσκεδαστικότητα, ενώ δεν μπορεί να αποδειχθεί η τυχαιότητα της μεταβλητής «Αξία» μέσα από τον κατάλληλο έλεγχο, αλλά μέσα από τον τρόπο λήψης του. Το δείγμα είναι χωρικά κατανεμημένο και αυτό θα πρέπει να ληφθεί υπ’ όψιν. Στο συγκεκριμένο μοντέλο γίνεται η αποδοχή του κεντρικού οριακού θεωρήματος, ώστε να αντιμετωπιστεί το πρόβλημα της κανονικότητας στην μεταβλητή «Αξία», ενώ το πρόβλημα της τυχαιότητας θα αντιμετωπιστεί ουσιαστικά στα πλαίσια της χωρικής ανάλυσης των δεδομένων, έχοντας υπόψη την χωρική αυτοσυσχέτιση της αξίας των ακινήτων του δείγματος η οποία είναι αναπόφευκτη. Για τον λόγο αυτό παραβλέπουμε τα αποτελέσματα που παρουσιάστηκαν, όπως αυτά προέκυψαν από την επεξεργασία με το λογισμικό SPSS, όσο αφορά την ορθότητα του μοντέλου και προχωράμε στην επεξεργασία μέσω του λογισμικού Arc Gis.

5.4.2.5 Εξέταση του Δεύτερου Μοντέλου Παλινδρόμησης σε Περιβάλλον Γεωγραφικού Συστήματος Πληροφοριών

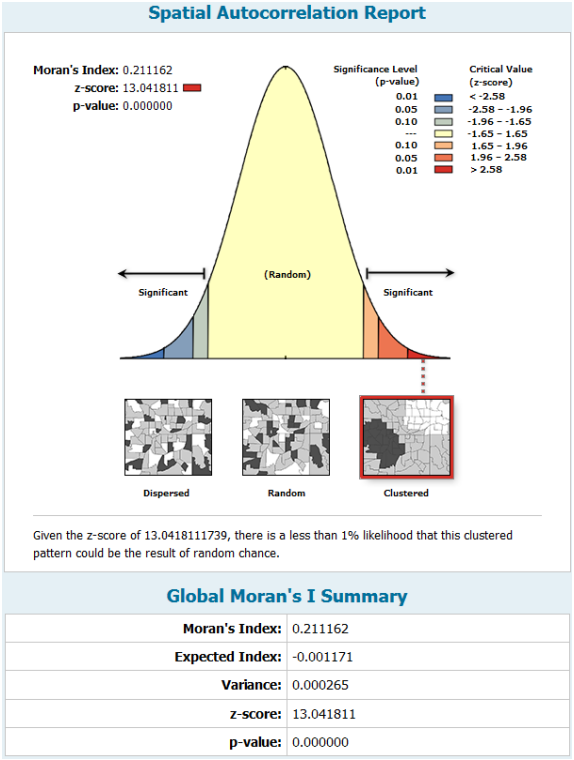
Το φαινόμενο που εξετάζεται εντάσσεται σε μία γεωγραφική περιοχή και γι’ αυτό τον λόγο είναι επιθυμητό να εξεταστεί η εφαρμογή της πολλαπλής γραμμικής παλινδρόμησης σε περιβάλλον Γεωγραφικών Συστημάτων Πληροφοριών GIS. Το μοντέλο που θα εξεταστεί είναι αυτό που παρουσίασε καλύτερα αποτελέσματα και αυτό είναι το δεύτερο. Σ΄αυτό το περιβάλλον γίνεται συνδυασμένη χρήση των πληροφοριών των χαρακτηριστικών του κάθε στοιχείου με την ταυτόχρονη πληροφορία της θέσης του. Το λογισμικό ArcGis αποτελεί ένα πακέτο που δίνει την δυνατότητα αυτή, δηλαδή της προσαρμογής της κλασικής παλινδρόμησης στα χωρικά δεδομένα. Καθώς η χωρική παλινδρόμηση στηρίζεται στη μέτρηση της χωρικής αυτοσυσχέτισης εξετάζει τις ομοιότητες ή τις διαφορές στις τιμές κάποιου χαρακτηριστικού σε διαφορετικές θέσεις, ανάλογα με την απόσταση των θέσεων αυτών μεταξύ τους.

5.4.2.5.1 Χωρική Στατιστική Ανάλυση.

Ο συντελεστής Moran (Moran's I) αποτελεί το πιο διαδεδομένο μέτρο της χωρικής αυτοσυσχέτισης και δείχνει κατά πόσο η παρατηρούμενη χωρική κατανομή παρουσιάζεται ως μη τυχαία στον γεωγραφικό χώρο, δηλαδή παρουσιάζει φαινόμενο cluster. Σε πρώτη φάση εφαρμόζουμε τον συντελεστή Moran's I για την εξαρτημένη μεταβλητή «Αξία Υ». Τα αποτελέσματα παρουσιάζονται στην εικόνα 5.21. Παρατηρούμε ότι η εξαρτημένη μεταβλητή παρουσιάζει τιμή στο δείκτη ίση με 0,211162 και z-score ίσο με 13,041811, ενώ το p=0,000. Παράλληλα παρουσιάζεται φαινόμενο συγκέντρωσης (cluster), η σημαντικότητα δε του φαινομένου αυτού ενδυναμώνεται και από την υψηλή τιμή του z score, όπου δηλώνεται ότι το φαινόμενο cluster έχει πιθανότητα μικρότερη του 1% να οφείλεται σε τυχαίους παράγοντες. Καθώς η τιμή Moran's I είναι θετική αλλά κοντά στο μηδέν, υποδεικνύεται ότι υπάρχει μικρή συγκέντρωση χωρικά παρόμοιων τιμών. Μια αρνητική τιμή θα υποδείκνυε συγκέντρωση χωρικά διαφορετικών τιμών. Γενικά τιμές κοντά στο μηδέν υποδεικνύουν έλλειψη χωρικής εξάρτησης μεταξύ των γειτνιαζόντων παρατηρήσεων , τιμές κοντά στο +1 υποδεικνύουν ότι η εξεταζόμενη θέση έχει πολύ κοντινές τιμές με τις γειτονικές της , ενώ τιμές κοντά στο -1 υποδεικνύουν ότι η θέση παρουσιάζει ακραία τιμή σε σχέση με τις γειτονικές της. Τα όρια των τιμών z-score, είναι από -1,96 έως +1,96. Τιμές χαμηλότερες από την κατώτερη τιμή σημαίνουν ότι το σημείο θεωρείται ότι έχει ακραία τιμή και τιμές μεγαλύτερες της ανώτερης τιμής σημαίνουν ότι το εξεταζόμενο σημείο δε διαφέρει σημαντικά από τα γειτονικά του.

Εικόνα 5.21

Εξέταση συντελεστή Moran's I δεύτερου μοντέλου για την εξαρτημένη μεταβλητή



Global Moran's I Summary

Moran's Index:	0.211162
Expected Index:	-0.001171
Variance:	0.000265
z-score:	13.041811
p-value:	0.000000

5.4.2.5.2 Κλασσική παλινδρόμηση (Ordinary Least Squares Estimation-OLS) σε Περιβάλλον GIS

Εφαρμόζουμε την κλασική παλινδρόμηση σε περιβάλλον GIS και τα αποτελέσματα της διαδικασίας παρουσιάζονται στον πίνακα 5.25, ενώ στον πίνακα 5.19 παρουσιάζονται τα διαγνωστικά της παλινδρόμησης. Παρατηρούμε ότι το R^2 =0,656223 το F=114,531980 και το AIC είναι ίσο με 20764,864352.

Τα αποτελέσματα (Πίνακας 5.26) ταυτίζονται απόλυτα με το μοντέλο όπως αυτό είχε δημιουργεί με το SPSS και παρουσιάζονται αναλυτικά στο παράρτημα «Β3».

Πίνακας 5.25

Αποτελέσματα OLS διαδικασίας δεύτερου μοντέλου

Summary of OLS Results - Model Variables

Variable	Coefficient [a]	StdError	t-Statistic	Probability [b]	Robust_SE	Robust_t	Robust_Pr [b]	VIF [c]
Intercept	61155.983783	21113.564071	2.896526	0.003878*	25073.237874	2.439094	0.014917*	--------
8_POSOST	1622.008887	220.050765	7.371067	0.000000*	204.856167	7.917794	0.000000*	1.872893
16_APOST	-9.086113	1.663863	-5.460854	0.000000*	1.664915	-5.457405	0.000000*	1.247052
17_APOST	14.722978	4.284175	3.436596	0.000633*	3.761044	3.914599	0.000107*	1.746520
23_APOST	-37.590338	14.438574	-2.603466	0.009385*	13.912192	-2.701971	0.007030*	1.085662
31_TIMH_	40.261279	3.733487	10.783829	0.000000*	4.434346	9.079417	0.000000*	1.335576
32_EMBAD	987.701885	49.590901	19.916998	0.000000*	57.351427	17.221923	0.000000*	1.231667
35_ARIUM	19515.449245	6437.140008	3.031696	0.002518*	11842.527704	1.647912	0.099756	1.132605
38_OROFO	3952.767494	877.295980	4.505626	0.000010*	927.976416	4.259556	0.000027*	1.190727
40_HLIKI	-263.667084	100.650092	-2.619641	0.008955*	112.523171	-2.343225	0.019335*	1.811871
41_EIDOS	-27373.44721	4844.987899	-5.649848	0.000000*	6467.376535	-4.232543	0.000030*	1.141092
42_KATAS	-9311.768948	1496.072175	-6.224144	0.000000*	1521.579540	-6.119804	0.000000*	1.433691
47_YPARX	18862.626845	5036.791579	3.744969	0.000204*	5112.152340	3.689762	0.000251*	1.450995
48_EIDOS	-11447.85147	2911.342798	-3.932155	0.000100*	2971.780979	-3.852185	0.000136*	1.524167
51_UEA_X	17508.945443	4612.597009	3.795897	0.000168*	5632.289295	3.108673	0.001956*	1.202593

Πίνακας 5.26

Διαγνωστικά παλινδρόμησης OLS διαδικασίας δεύτερου μοντέλου

OLS Diagnostics

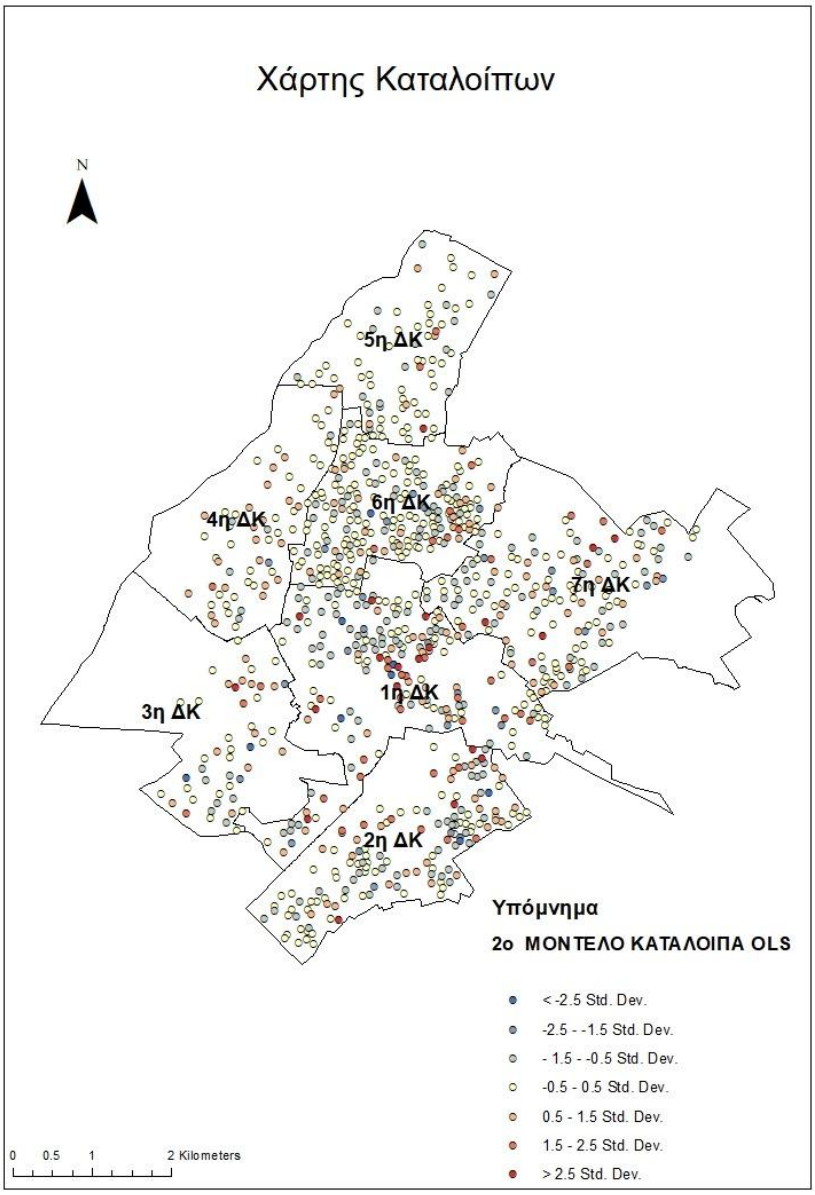
Input Features:	2_OLS_MONTELO_TELIKO	Dependent Variable:	_AXIA_Y
Number of Observations:	855	Akaike's Information Criterion (AICc) [d]:	20764.864352
Multiple R-Squared [d]:	0.656223	Adjusted R-Squared [d]:	0.650494
Joint F-Statistic [e]:	114.531980	Prob(>F), (14,840) degrees of freedom:	0.000000*
Joint Wald Statistic [e]:	1473.278272	Prob(>chi-squared), (14) degrees of freedom:	0.000000*
Koenker (BP) Statistic [f]:	159.969962	Prob(>chi-squared), (14) degrees of freedom:	0.000000*
Jarque-Bera Statistic [g]:	170.773293	Prob(>chi-squared), (2) degrees of freedom:	0.000000*

Όπως έχει ήδη αναφερθεί μία από τις παραδοχές για την ανάλυση παλινδρόμησης είναι η ανεξαρτησία των καταλοίπων, η οποία όμως συνήθως παραβιάζεται στην ανάλυση των γεωγραφικών φαινομένων λόγω του φαινομένου της χωρικής αυτοσυσχέτισης. Η ανάλυση των καταλοίπων γίνεται αρχικά με τη χαρτογράφησή τους. Η ταξινόμηση των ακινήτων γίνεται με βάση την τιμή των τυποποιημένων υπολοίπων (Std Residuals) σε επτά κλάσεις που αποτελούν ένδειξη για το ποια είναι η κατανομή των υπολοίπων και εάν το μοντέλο ταιριάζει στα

δεδομένα. Στη χάρτη 5.7 παρουσιάζεται η χαρτογράφηση των καταλοίπων, ενώ στην εικόνα 5.22 παρουσιάζεται το ιστόγραμμα αυτών και στην εικόνα 5.23 παρουσιάζεται το διάγραμμα διασποράς τους. Υποδεικνύεται η μη κανονική κατανομή τους.

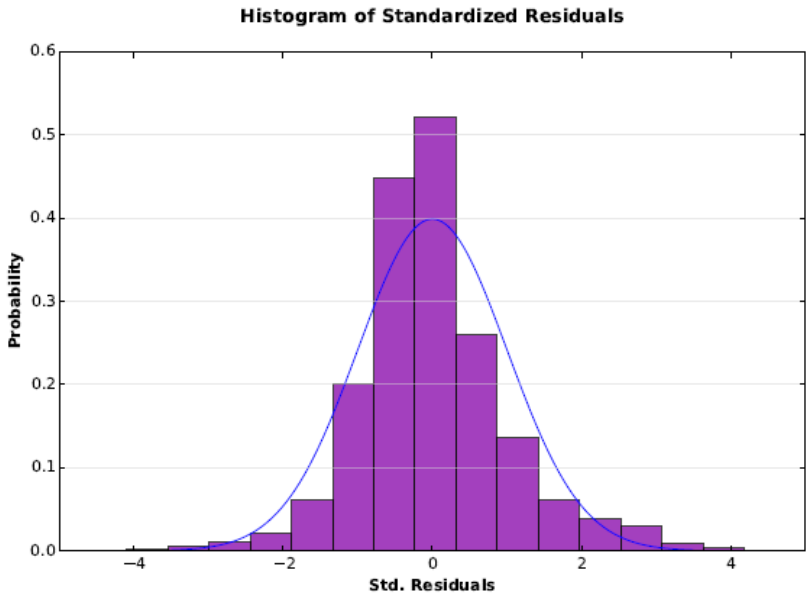
Χάρτης 5.7

Χαρτογράφηση Καταλοίπων OLS διαδικασίας δεύτερου μοντέλου



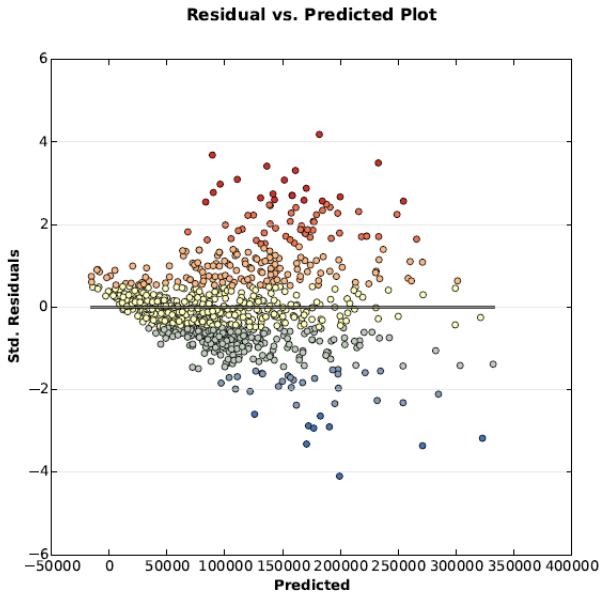
Εικόνα 5.22

Ιστόγραμμα καταλοίπων δεύτερου μοντέλου



Εικόνα 5.23

Διάγραμμα διασποράς καταλοίπων δεύτερου μοντέλου

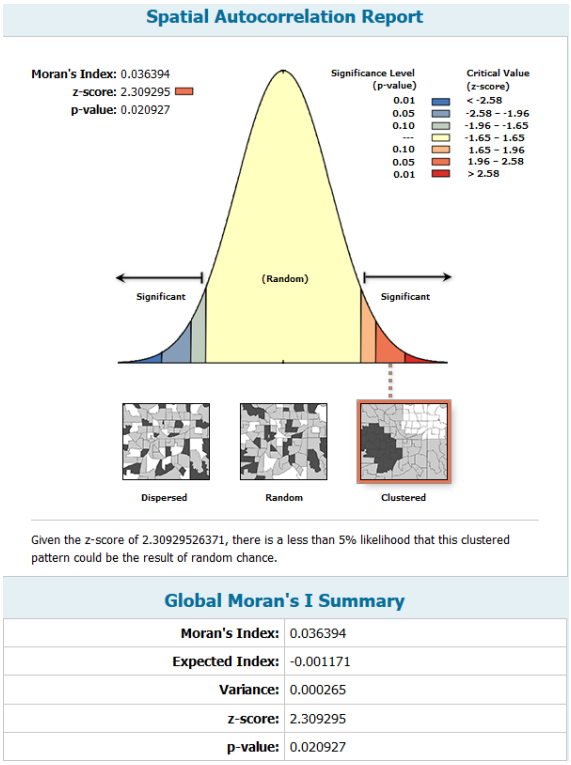


Για τον έλεγχο τη χωρικής αυτοσυσχέτισης των υπολοίπων χρησιμοποιούμε τον δείκτη Morans I (Εικόνα 5.24). Παρατηρούμε ότι τα κατάλοιπα της παλινδρόμησης παρουσιάζουν τιμή στο δείκτη ίση με 0,036394 και z-score ίσο με 2,309295, ενώ το p=0,020927. Παράλληλα παρουσιάζεται φαινόμενο συγκέντρωσης (cluster). Η σημαντικότητα του φαινομένου υποδεικνύεται από την τιμή του z score, όπου

δηλώνεται ότι το φαινόμενο cluster έχει πιθανότητα μικρότερη του 5% να οφείλεται σε τυχαίους παράγοντες. Η τιμή Moran's I είναι θετική αλλά κοντά στο μηδέν και έτσι υποδεικνύεται ότι υπάρχει μικρή συγκέντρωση χωρικά παρόμοιων τιμών.

Εικόνα 5.24

Έλεγχος χωρικής αυτοσυσχέτισης καταλοίπων δεύτερου μοντέλου μέσω OLS διαδικασίας



Global Moran's I Summary

Moran's Index:	0.036394
Expected Index:	-0.001171
Variance:	0.000265
z-score:	2.309295
p-value:	0.020927

Ελέγχοντας τα κατάλοιπα της OLS διαδικασίας, με τον υπολογισμό του Moran I, διαπιστώνουμε την ύπαρξη χωρικής αυτοσυσχέτισης, δηλαδή cluster χωρικό πρότυπο. Κρίνεται λοιπόν σκόπιμο να προχωρήσουμε στη γεωγραφικά σταθμισμένη παλινδρόμηση.

5.4.2.5.3 Χωρική Παλινδρόμηση (Geographically Weighted Regression-GWR) σε περιβάλλον GIS.

Η εφαρμογή της GWR στο μοντέλο μας γίνεται και εδώ μέσω του open source λογισμικού GWR4. Τα αποτελέσματα φαίνονται στον πίνακα 5.27 και αναλυτικά παρουσιάζονται στο παράρτημα «Β4» , ενώ αναλυτικά αποτελέσματα της διαδικασίας για κάθε ακίνητο παρουσιάζονται στο παράτημα «Β5». Παρατηρούμε ότι

μέσω της διαδικασίας αυτής το Adjusted R square που στην OLS ήταν 0,650494, τώρα είναι 0,689138, ενώ παράλληλα μειώθηκε και το AICc από 20764,8643 που ήταν στην OLS διαδικασία, σε 20703,2800. Συνολικά λοιπόν το μοντέλο δείχνει να έχει βελτιωθεί αισθητά.

Πίνακας 5.27

Αποτελέσματα GWR διαδικασίας

GWR (Geographically weighted regression) result

Bandwidth and geographic ranges
Bandwidth size: 419.519440

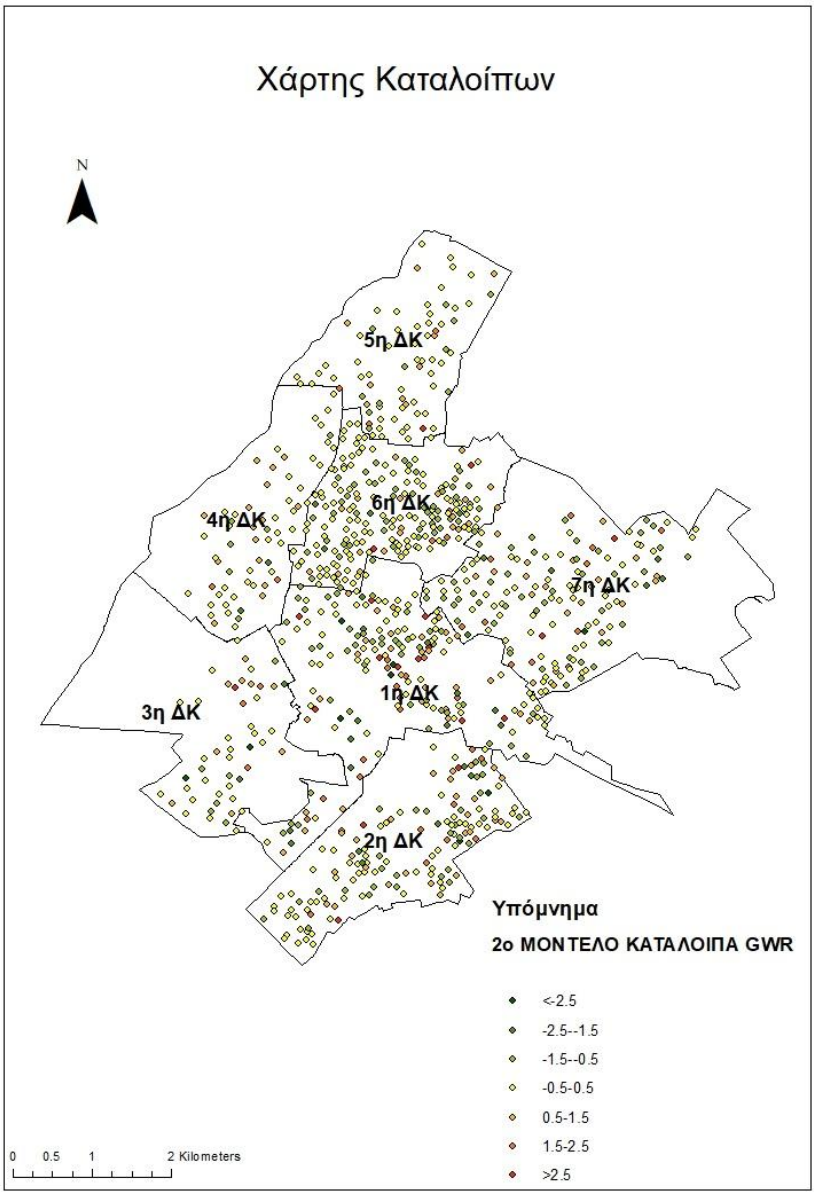
Coordinate	Min	Max	Range
X-coord	473725.000000	480487.000000	6762.000000
Y-coord	4200150.000000	4208998.000000	8848.000000

Diagnostic information
Residual sum of squares: 1375198052363.830000
Effective number of parameters (model: trace(S)): 68.782863
Effective number of parameters (variance: trace(S'S)): 47.716405
Degree of freedom (model: n - trace(S)): 786.217137
Degree of freedom (residual: n - 2trace(S) + trace(S'S)): 765.150679
ML based sigma estimate: 40105.096617
Unbiased sigma estimate: 42394.463160
-2 log-likelihood: 20551.117273
Classic AIC: 20690.682998
AICc: 20703.280098
BIC/MDL: 21022.228458
CV: 2017026767.861450
R square: 0.721844
Adjusted R square: 0.689138

Στο χάρτη 5.8 φαίνονται τα τυποποιημένα κατάλοιπα (standardized) της διαδικασίας GWR. Από τον χάρτη αυτό φαίνεται ότι σε σχέση με τον προηγούμενο χάρτη παρατηρείται μείωση του μεγέθους των καταλοίπων . Το ίδιο προκύπτει και από το ιστόγραμμα (Εικόνα 5.25)

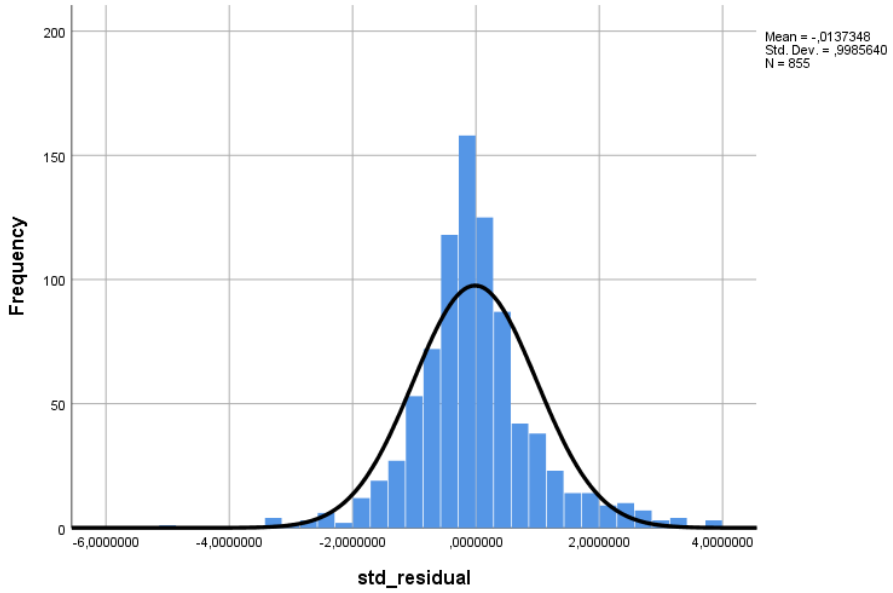
Χάρτης 5.8

Χάρτης καταλοίπων GWR διαδικασίας



Εικόνα 5.25

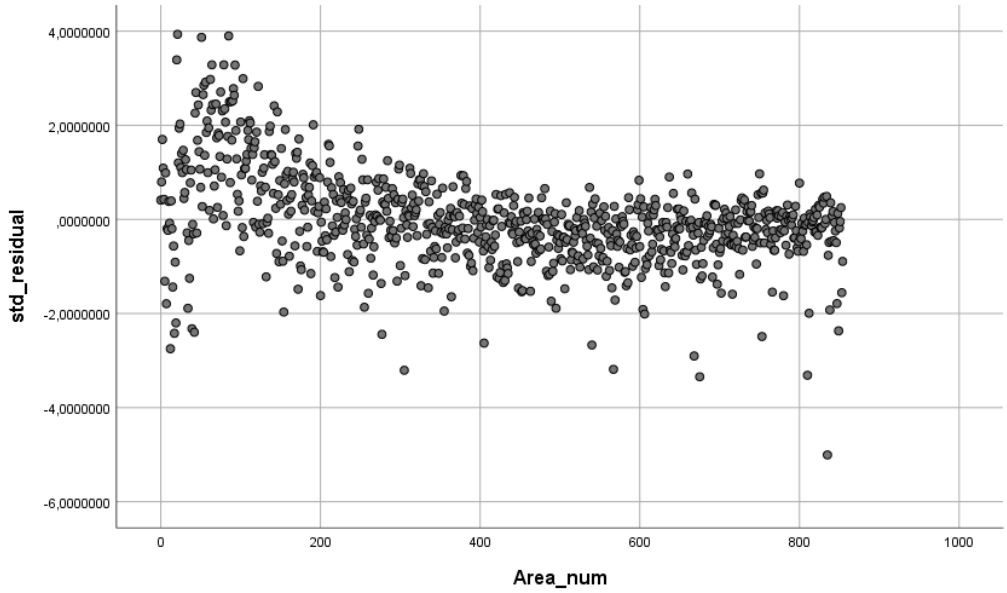
Ιστόγραμμα καταλοίπων



Αν απεικονίσουμε το scatter plot (Εικόνα 5.26) και υπολογίσουμε τον συντελεστή αυτοσυσχέτισης Moran I (Εικόνα 5.27) για τα κατάλοιπα της GWR θα προκύψει τυχαίο πρότυπο, δηλαδή τα κατάλοιπα δεν παρουσιάζουν πλέον χωρική αυτοσυσχέτιση και οφείλονται σε τυχαίους παράγοντες.

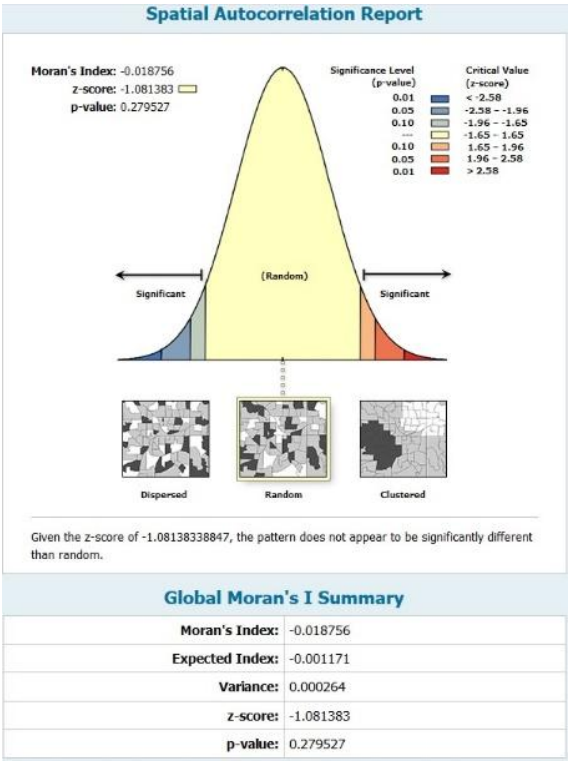
Εικόνα 5.26

Διάγραμμα διασποράς υπολοίπων.



Εικόνα 5.27

Υπολογισμός Morans I για τα κατάλοιπα



5.4.2.5.4 Εξαγωγή Συμπερασμάτων των Αποτελεσμάτων του Δεύτερου Μοντέλου μέσα από την Χωρική Παλινδρόμηση (Geographically Weighted Regression-GWR.)

Συμπερασματικά από την επεξεργασία του δεύτερου μοντέλου μέσα από την χρήση της χωρικής παλινδρόμησης διαπιστώνουμε ότι αυτό βελτιώθηκε αισθητά (βλέπε πίνακα 5.28). Ο συντελεστής συσχέτισης βελτιώθηκε αισθητά ενώ ανάλογες ενδείξεις λάβαμε και από το κριτήριο Akaike Information Criterion (AIC). Παρατηρείται μείωση του μεγέθους των καταλοίπων ενώ αυτά δεν παρουσιάζουν πλέον χωρική αυτοσυσχέτιση και οφείλονται σε τυχαίους παράγοντες.

Πίνακας 5.28

Σύγκριση αποτελεσμάτων δεύτερου μοντέλου μέσω OLS & GWR διαδικασίας.

Συντελεστές	OLS	GWR
Adjusted R Square	0.65	0.6891
Morans I residuals	0.036394	-0.0187
AICc	20764.8643	20703.280

Είναι φανερό πλέον ότι από το μοντέλο της γεωγραφικά σταθμισμένης παλινδρόμησης προκύπτει καλύτερη ερμηνεία της εξαρτημένης μεταβλητής και ακριβέστερος προσδιορισμός του μοντέλου της παλινδρόμησης.

5.4.3 Δημιουργία Τρίτου Μοντέλου

Στο τρίτο μοντέλο της παρούσας εργασίας ακολουθείται η διαδικασία της δημιουργίας εξίσωσης παλινδρόμησης με την χρήση μετασχηματισμών με την βοήθεια λογαρίθμων και ρίζας, αφού πρώτα αφαιρεθούν οι ακραίες οντότητες σε επιλεγμένες μεταβλητές του δείγματος κάνοντας χρήση των δεικτών DfFITS (Difference in Fits) και Cook's D (Cook' s distance). Η επεξεργασία γίνεται αφού πρώτα ομαδοποιηθούν οι ανεξάρτητες μεταβλητές σύμφωνα με τα αναφερόμενα στο δεύτερο κεφάλαιο. Η κάθε ομάδα μεταβλητών δημιουργεί μια εξίσωση παλινδρόμησης. Επιλέγονται οι μεταβλητές που έλαβαν μέρος στην κάθε μια επί μέρους εξίσωση και συμμετέχουν στην δημιουργία της τελικής εξίσωσης παλινδρόμησης με την OLS διαδικασία, η οποία μετά επεξεργάζεται και με την GWR διαδικασία. Η διαδικασία εισαγωγής των ανεξάρτητων μεταβλητών στο λογισμικό SPSS γίνεται με την μέθοδο Stepwise ενώ η GWR διαδικασία γίνεται με το λογισμικό Arc Gis και το λογισμικό GWR4.

5.4.3.1 Έλεγχος και Αφαίρεση Ακραίων Παρατηρήσεων

Στη στατιστική ακραία τιμή είναι μια παρατήρηση η οποία είναι αριθμητικά απόμακρη από το υπόλοιπο σύνολο των δεδομένων. Τα ακραία σημεία μπορούν να υποδείξουν ελαττωματικά στοιχεία ή εσφαλμένες διαδικασίες, όπου μια συγκεκριμένη θεωρία μπορεί τελικά να μην είναι έγκυρη. Βέβαια σε μεγάλα δείγματα, ένας μικρός αριθμός εσφαλμένων αποτελεσμάτων είναι αναμενόμενος, καθώς μπορεί να οφείλεται σε οποιαδήποτε ανώμαλη κατάσταση. Θέλοντας να δώσουμε έναν γενικό ορισμό θα λέγαμε ότι ακραία τιμή είναι μια παρατήρηση ή υποσύνολο των παρατηρήσεων που φαίνεται να είναι ασυνεπής με το υπόλοιπο του συνόλου των δεδομένων .

Την ύπαρξη ακραίων τιμών στην εξαρτημένη μεταβλητή « Αξία_ Υ» είδαμε σε boxplots (Εικόνα 5.4) του προηγούμενου μοντέλου. Η μεθοδολογία που ακολουθούμε για την αφαίρεση των ακραίων τιμών στο τρίτο μοντέλο είναι η χρησιμοποίηση των δεικτών DfFITS (Difference in Fits) και Cook's D (Cook' s

distance) με όρια αποκοπής τις τιμές 2*√(p/n) και 4/(n-p) αντίστοιχα (Khaleelur Rahman, K., Sathik, M., Kannan.,S, 2012, σελ. 24), όπου n είναι το μέγεθος του δείγματος και p είναι ο αριθμός των μεταβλητών. Ο δείκτης DfFITS μετρά πόσο μια παρατήρηση έχει επηρεάσει την προσαρμοσμένη της τιμή από το μοντέλο παλινδρόμησης, ενώ ο δείκτης Cook'sD μετρά το συνολικό αντίκτυπο της κάθε παρατήρησης πάνω στην ομάδα των συντελεστών παλινδρόμησης, καθώς και την ομάδα των προσαρμοσμένων τιμών. Για να δημιουργηθούν οι τιμές των δεικτών μέσω του λογισμικού SPSS τρέχουμε την γραμμική παλινδρόμηση με εξαρτημένη μεταβλητή την μεταβλητή «Αξία Y « και ανεξάρτητες μεταβλητές όλες τις υπόλοιπες εκτός των μεταβλητών : «Έτος κατασκευής X39» καθώς από αυτήν προκύπτει και έχει απόλυτη συσχέτιση με την μεταβλητή «Ηλικία κατασκευής», « Ποσοστό αλλοδαπών X5» καθώς έχει απόλυτη αρνητική συσχέτιση με την μεταβλητή «Ποσοστό Ελλήνων» και της μεταβλητής «Τιμή ανά τ.μ. X33» καθώς ουσιαστικά αποτελεί την διαίρεση της εξαρτημένης με την μεταβλητή «Εμβαδόν ακινήτου». Εισάγουμε όλες τις υπόλοιπες μεταβλητές στο πρόγραμμα με την διαδικασία enter. Οι ανεξάρτητες μεταβλητές που λαμβάνουν μέρος στην επεξεργασία είναι οι: X4, X6, X7, X8, X9, X10, X11, X12, X13, X14, X15, X16, X17, X18, X19, X20, X21, X22, X23, X24, X25, X26, X27, X28, X29, X30, X31, X32, X34, X35, X36, X37, X38, X40, X41, X42, X43, X44, X45, X46, X47, X48, X49, X50, X51, X52, X53. Σκοπός μας είναι σε πρώτη φάση να βρούμε τα outliers και να τα αφαιρέσουμε εφαρμόζοντας τις τιμές αποκοπής για τους δείκτες DfFITS και Cook’s D που υπολογίζονται σε ± 0,433 και ±0,00419 αντίστοιχα, έχοντας υπόψη ότι το δείγμα μας είναι 1000 και οι συμμετέχουσες παράμετροι 48. Παρατηρούμε τα αποτελέσματα της διαδικασίας στους πίνακες 5.29 και 5.30. Το R^2 είναι ίσο με 0,677 και το Adjusted R^2 είναι ίσο με 0,431. Η εκτίμηση σφάλματος είναι 227.482,454 και το F είναι ίσο με 17,101.

Πίνακας 5.29

Γραμμική παλινδρόμηση: Συντελεστές προσδιορισμού

Model Summary

Model	R	R Square	Adjusted R Square	Std. Error of the Estimate
1	,677a	,458	,431	227482,454

Πίνακας 5.30

Γραμμική παλινδρόμηση: Ανάλυση διασποράς

ANOVAa

Model		Sum of Squares	df	Mean Square	F	Sig.
1	Regression	41592696812634,960	47	884950996013,510	17,101	,000b
	Residual	49264350205123,040	952	51748267022,188		
	Total	90857047017758,000	999			

Στο τέλος αυτής της διαδικασίας και αφού αφαιρέθηκαν 36 παρατηρήσεις, απέμειναν 964 προς περαιτέρω επεξεργασία.

Ξανατρέχουμε το μοντέλο μας με την απλή διαδικασία enter και παρατηρούμε αυτομάτως ότι έχουν βελτιωθεί αισθητά τα αποτελέσματα (πίνακες 5,31 & 5.32). Το R είναι ίσο με 0,854 και το R^2 είναι ίσο με 0,751. Η εκτίμηση σφάλματος είναι 86.002,966 και το F είναι ίσο με 52,367.

Πίνακας 5.31

Γραμμική παλινδρόμηση: Συντελεστές προσδιορισμού

Model Summary

Model	R	R Square	Adjusted R Square	Std. Error of the Estimate
1	,854[a]	,729	,715	86002,966

Πίνακας 5.32

Γραμμική παλινδρόμηση: Ανάλυση διασποράς

ANOVA[a]

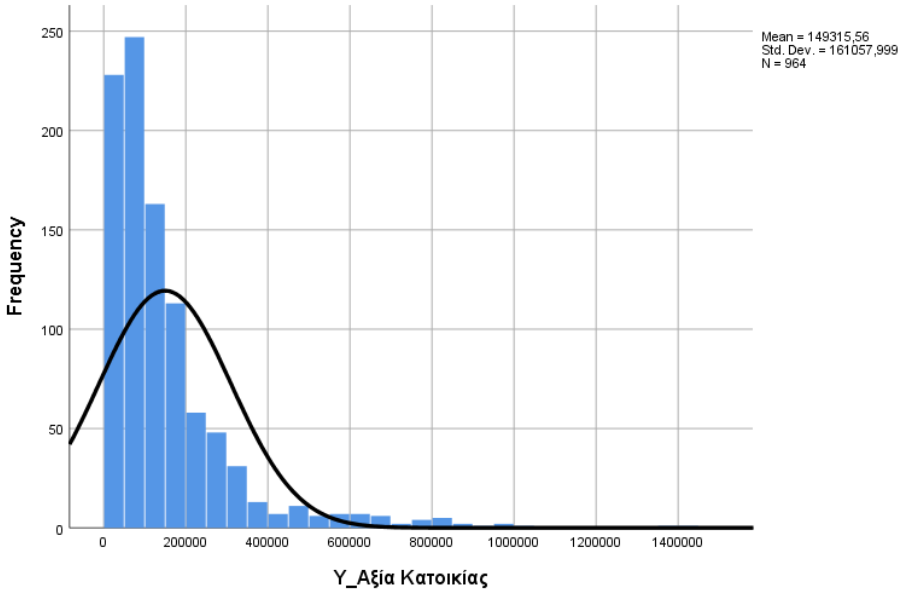
Model		Sum of Squares	df	Mean Square	F	Sig.
1	Regression	18204707715911,125	47	387334206721,513	52,367	,000[b]
	Residual	6775203330686,365	916	7396510186,339		
	Total	24979911046597,492	963			

5.4.3.2 Μετασχηματισμοί Μεταβλητών

Μια βασική παραδοχή στην κατασκευή ενός μοντέλου παλινδρόμησης είναι ότι υπάρχει κανονικότητα . Όπως αναφέραμε στο κεφάλαιο 3.4.7 οι μετασχηματισμοί των δεδομένων είναι ένα εργαλείο που μπορεί να χρησιμοποιηθεί με σκοπό την βελτίωση της κανονικότητας της κατανομής, της εξίσωση της διακύμανσης και για την βελτίωση των αποτελεσμάτων στην στατιστική ανάλυση. Εξετάζοντας το ιστόγραμμα της εξαρτημένης μεταβλητής παρατηρούμε την μη ύπαρξη κανονικής κατανομής καθώς έχουμε μια πολύ μεγαλύτερη ουρά στη δεξιά πλευρά από ό, τι στα αριστερά (Εικόνα 5.28).

Εικόνα 5.28

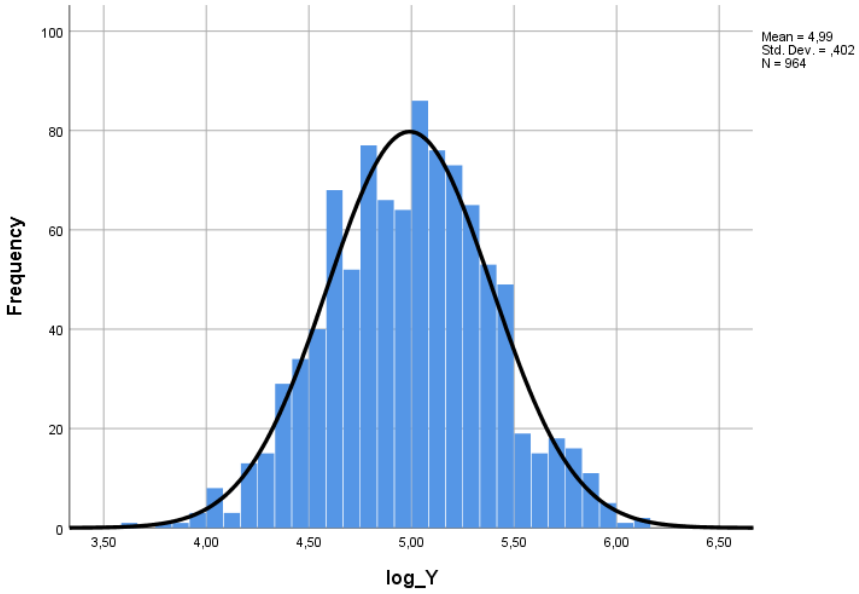
Ιστόγραμμα εξαρτημένης μεταβλητής Y_ Αξία



Για την διόρθωση αυτού του προβλήματος κατασκευάζουμε με τον φυσικό λογάριθμο της μεταβλητής «Αξία_Y» μια νέα μεταβλητή την «log_Y» και παρατηρούμε στην εικόνα 5.29 την αισθητή βελτίωση του αποτελέσματος.

Εικόνα 5.29

Ιστόγραμμα λογαριθμησμένης εξαρτημένης μεταβλητής log_Y Αξία



Η ίδια διαδικασία εφαρμόζεται και στις υπόλοιπες ποσοτικές μεταβλητές που το ιστόγραμμά τους δεν παρουσιάζεται ικανοποιητικό, ενώ δεν έχει νόημα να γίνει η ίδια διαδικασία σε δίτιμες και ιεραρχικές μεταβλητές. Τα ιστογράμματα παρουσιάζονται στο παράρτημα «Γ1», ενώ αναφέρουμε ότι έγιναν οι λογαριθμήσεις των ανεξάρτητων μεταβλητών X15, X16, X19, X20, X22, X23, X26, X27, X28, X29, X32, X53. Είναι σημαντικό να αναφερθεί ότι η προσπάθεια για λογαριθμήσεις πολλών μεταβλητών δεν έδωσε το αναμενόμενο αποτέλεσμα και γι' αυτό εφαρμόστηκε ο μετασχηματισμός της ρίζας στις μεταβλητές X11, X13, X14, X17, X21, X24, X25, τα δε ιστογράμματα αυτών παρουσιάζονται αναλυτικά στο παράρτημα «Γ1».

5.4.3.3 Ομαδοποίηση Μεταβλητών

Στο κεφάλαιο 2.3 παρουσιάστηκε μια ομαδοποίηση των ανεξάρτητων μεταβλητών σύμφωνα με την βιβλιογραφία. Έτσι οι ανεξάρτητες μεταβλητές ομαδοποιήθηκαν ουσιαστικά σε τέσσερις κατηγορίες: Των κατασκευαστικών χαρακτηριστικών, της θέσης, της γειτονίας και των λοιπών μεταβλητών.

Σε αυτές τις ομάδες θα ενταχθούν και οι μεταβλητές του παρόντος τρίτου μοντέλου, όπως αυτές έχουν μετασχηματιστεί, ώστε να επεξεργαστούν ανά κατηγορία, δημιουργώντας επιμέρους μοντέλα παλινδρόμησης. Το κάθε επιμέρους μοντέλο μέσα από την OLS διαδικασία θα δώσει τις πλέον κατάλληλες μεταβλητές που θα λάβουν μέρος στο τελικό συνολικό μοντέλο μας.

Πιο συγκεκριμένα στην πρώτη ομάδα των κατασκευαστικών χαρακτηριστικών θα λάβουν μέρος οι μεταβλητές: log_ εμβαδόνΧ32, Αριθμός υπνοδωματίων_Χ34, Αριθμός μπάνιων_Χ35, Ύπαρξη WC_Χ36, Ύπαρξη αποθήκης_Χ37, Όροφος_Χ38, Ηλικία κατασκευής_Χ40, Είδος κατοικίας_Χ41, Κατάσταση κατασκευής_Χ42, Ύπαρξη ανακαίνισης_Χ43, Ύπαρξη βεράντας_Χ44, Ύπαρξη ανσασέρ_Χ46, Ύπαρξη θέσης στάθμευσης_Χ47, Είδος θέρμανσης_Χ48, Είδος καυσίμου θέρμανσης_Χ49, Ύπαρξη κλιματισμού_Χ50, Ύπαρξη τζακιού_Χ52

Στην δεύτερη ομάδα των μεταβλητών θέσης θα λάβουν μέρος οι μεταβλητές: sqrt_Απόσταση από αστυνομικό τμήμα11, Απόσταση από βιομηχανία_Χ12, sqrt_Απόσταση από βρεφονηπιακό σταθμό13, sqrt_Απόσταση από εμπόριο14, log_Απόσταση από ιστορικό κέντρο15, log_Απόσταση από νεκροταφείο16, sqrt_Απόσταση από νοσοκομείο17, Είδος πλησιέστερης συγκοινωνίας_Χ18, log_Απόσταση από πλησιέστερη συγκοινωνία19, log_Απόσταση από ναό20, sqrt_ Απόσταση από σχολείο21, log_ Απόσταση από αθλητισμό22, log_Απόστασης από πράσινο23, sqrt_Απόσταση από Πλάκα24, sqrt_Απόσταση από πεζόδρομο25,log_Απόσταση από ΟΣΕ26, log_Απόσταση από Μουσείο27, log_Απόσταση από πρωτεύον δίκτυο28, log_Απόσταση από βασικό δίκτυο29,Γωνιακό ή διαμπερές_Χ45, Ύπαρξη θέας_Χ51

Στην τρίτη κατηγορία των μεταβλητών γειτονιάς θα λάβουν μέρος οι μεταβλητές : Ποσοστό ελλήνων_Χ4, Ποσοστό ανεργίας_Χ6, Ποσοστό πρασίνου_Χ8, Ποσοστό κατοίκων ανά ΔΚ_ Χ9, τμ πρασίνου ανά κάτοικο_Χ10, Μέσο εισόδημα κατοίκων_Χ30.

Στη τέταρτη κατηγορία των λοιπών μεταβλητών θα λάβουν μέρος οι μεταβλητές: Ποσοστό κενών κατοικιών_Χ7, Τιμή ζώνης_Χ31, log_Αριθμός ημερών προς διάθεση_Χ53

5.4.3.4 Προσαρμογή Μοντέλου Γραμμικής Παλινδρόμησης ανά Ομάδα Μεταβλητών

Η επεξεργασία γίνεται μέσω του SPSS και με την μέθοδο εισόδου μεταβλητών stepwise, δίνοντας τα παρακάτω αποτελέσματα ενώ αυτά παρουσιάζονται αναλυτικά στα παραρτήματα Γ2, Γ3, Γ4, Γ5.

Η πρώτη ομάδα των κατασκευαστικών χαρακτηριστικών έδωσε τα παρακάτω αποτελέσματα για την εξαρτημένη μεταβλητή log_Y (παράρτημα Γ2). Στον πίνακα 5.33 φαίνεται το R να είναι ίσο με 0,799 και το Adjusted R^2 να είναι ίσο με 0,636. Η

εκτίμηση σφάλματος είναι 0,24242 , ενώ το F είναι ίσο με 211,134 όπως φαίνεται στον πίνακα 5.34.

Πίνακας 5.33

Γραμμική παλινδρόμηση: Συντελεστές προσδιορισμού(κατασκευαστικά χαρακτηριστικά)

Model Summary[i]

Model	R	R Square	Adjusted R Square	Std. Error of the Estimate
8	,799[h]	,639	,636	,24242

h. Predictors: (Constant), log_32, X42_Katasταση Κατασκευής, X38_Όροφος, X48_Είδος Θέρμανσης, X41_Είδος Κατοικίας, X44_Ύπαρξη Βεράντας, X37_Ύπαρξη Αποθήκης, X49_Είδος Καυσίμου θέρμανσης

i. Dependent Variable: log_Y

Πίνακας 5.34

Γραμμική παλινδρόμηση: Ανάλυση διασποράς

ANOVA[a]

Model		Sum of Squares	df	Mean Square	F	Sig.
8	Regression	99,262	8	12,408	211,134	,000[i]
	Residual	56,123	955	,059		
	Total	155,384	963			

Στον πίνακα 5.35 παρουσιάζονται οι συμμετέχουσες μεταβλητές οι οποίες είναι οι : log32, X42, X38, X48, X41, X44, X37, X49. Παράλληλα στις εικόνες 5.30, 5.31, 5.32 παρουσιάζονται το ιστόγραμμα το P- P Plot και το διάγραμμα διασποράς των καταλοίπων.

Πίνακας 5.35

Αποτελέσματα παλινδρόμησης. Συντελεστές παλινδρόμησης.

Coefficients[a]

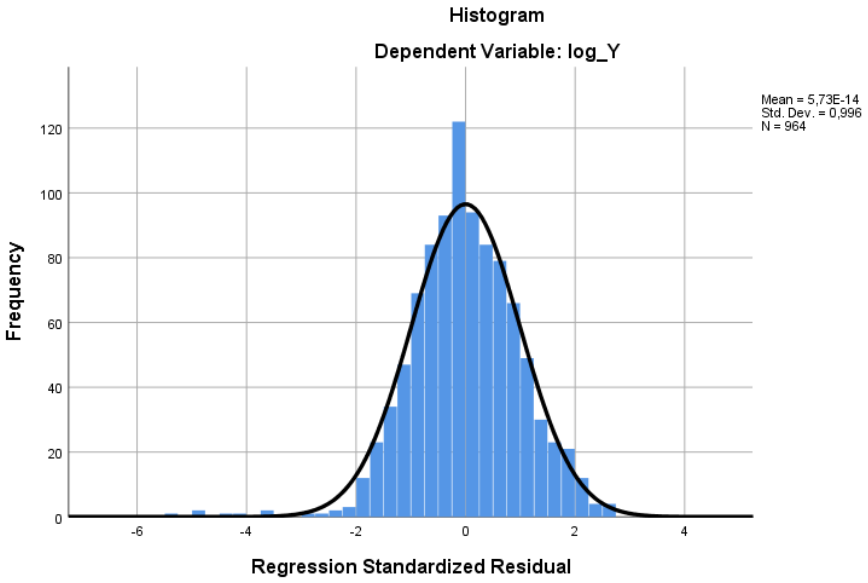
Model		Unstandardized Coefficients		Standardized Coefficients	t	Sig.
		B	Std. Error	Beta		
8	(Constant)	3,280	,124		26,365	,000
	log_32_Εμβαδόν	1,121	,040	,605	27,963	,000
	X42_Katasταση Κατασκευής	-,078	,007	-,246	-11,008	,000
	X38_Όροφος	,021	,004	,103	4,986	,000
	X48_Είδος Θέρμανσης	-,041	,016	-,069	-2,575	,010

X41_Είδος Κατοικίας	-,093	,023	-,087	-4,092	,000
X44_Ύπαρξη Βεράντας	,097	,034	,059	2,902	,004
X37_Ύπαρξη Αποθήκης	,057	,020	,061	2,801	,005
X49_Είδος Καυσίμου θέρμανσης	-,030	,015	-,049	-1,972	,049

a. Dependent Variable: log_Y

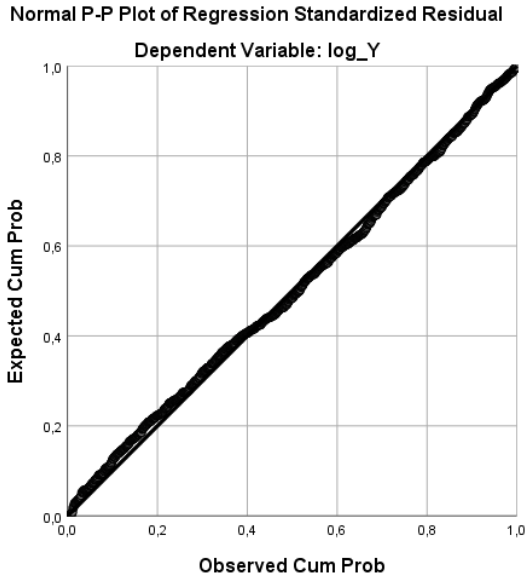
Εικόνα 5.30

Ιστόγραμμα υπολοίπων



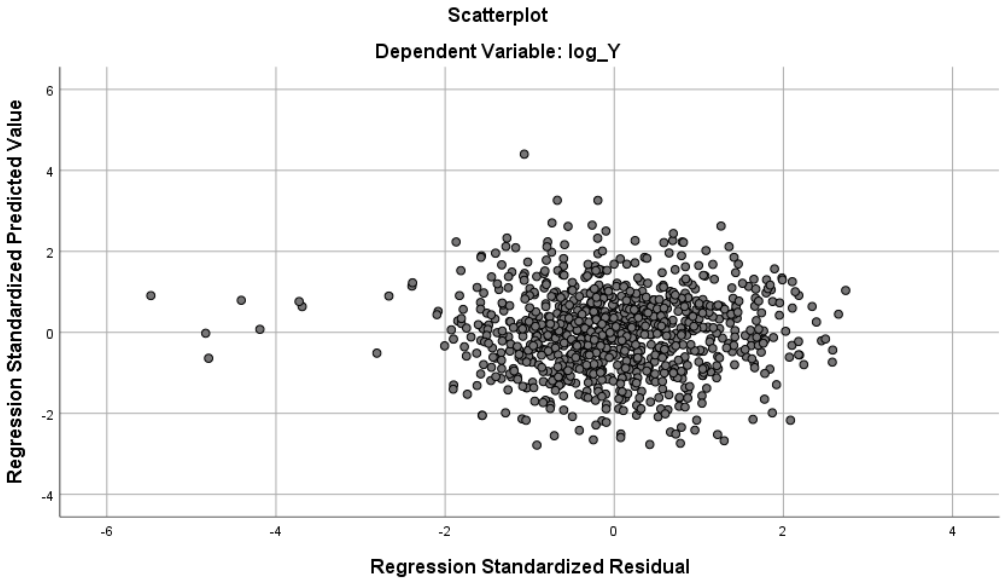
Εικόνα 5.31

Διάγραμμα P-P_ PLot



Εικόνα 5.32

Διάγραμμα διασποράς υπολοίπων.



Η δεύτερη ομάδα των χαρακτηριστικών θέσης έδωσε τα παρακάτω αποτελέσματα για την εξαρτημένη log_Y (παράρτημα Γ3). Στον πίνακα 5.36 φαίνεται το R να είναι ίσο με 0,410 και το Adjusted R^2 να είναι ίσο με 0,163. Η εκτίμηση σφάλματος είναι 0,31328 , ενώ το F είναι ίσο με 30,459 όπως φαίνεται στον πίνακα 5.37.

Πίνακας 5.36

Γραμμική παλινδρόμηση: Συντελεστές προσδιορισμού(χαρακτηριστικά θέσης)

Model Summary[e]

Model	R	R Square	Adjusted R Square	Std. Error of the Estimate
4	,410[d]	,168	,163	,31328

d. Predictors: (Constant), X51_Ύπαρξη Θέας, log_26_Απόσταση από ΟΣΕ, SQRT_11_ Απόσταση από αστυνομικό τμήμα, X45_Γωνιακό ή Διαμπερές

e. Dependent Variable: log_Y

Πίνακας 5.37

Γραμμική παλινδρόμηση: Ανάλυση διασποράς

ANOVA[a]

Model		Sum of Squares	df	Mean Square	F	Sig.
4	Regression	11,957	4	2,989	30,459	,000[e]
	Residual	59,180	603	,098		
	Total	71,137	607			

Στον πίνακα 5.38 παρουσιάζονται οι συμμετέχουσες μεταβλητές οι οποίες είναι οι : X51, log26, sqrt11, X45. Παράλληλα στις εικόνες 5.33, 5.34, 5.35 παρουσιάζονται το ιστόγραμμα το P- P Plot και το διάγραμμα διασποράς των καταλοίπων.

Πίνακας 5.38

Αποτελέσματα παλινδρόμησης: Συντελεστές παλινδρόμησης

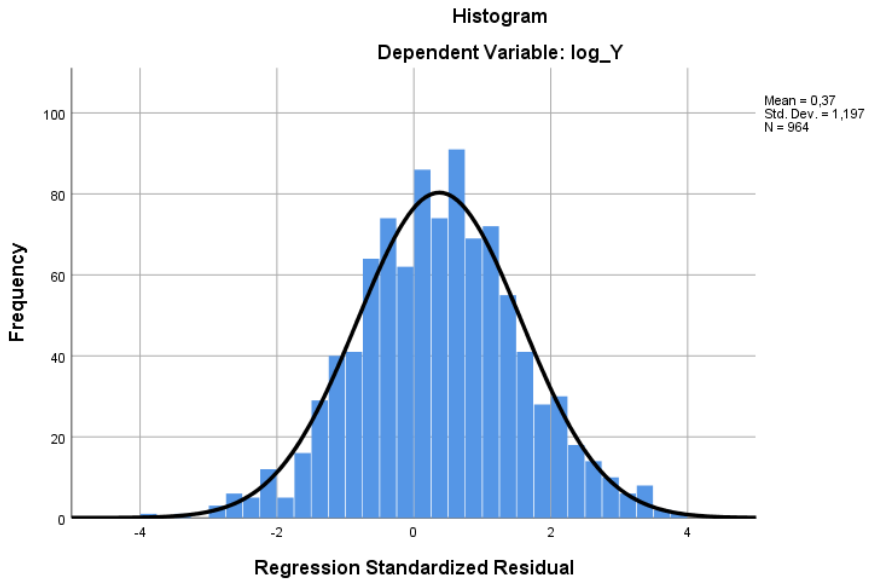
Coefficients[a]

Model		Unstandardized Coefficients		Standardized Coefficients		
		B	Std. Error	Beta	t	Sig.
4	(Constant)	4,266	,084		50,540	,000
	X51_Ύπαρξη Θέας	,262	,035	,280	7,456	,000
	log_26_ Απόσταση από ΟΣΕ	,149	,026	,217	5,705	,000
	SQRT_11_ Απόσταση από ΑΤ	,004	,002	,086	2,288	,022
	X45_Γωνιακό ή Διαμπερές	,090	,041	,083	2,218	,027

a. Dependent Variable: log_Y

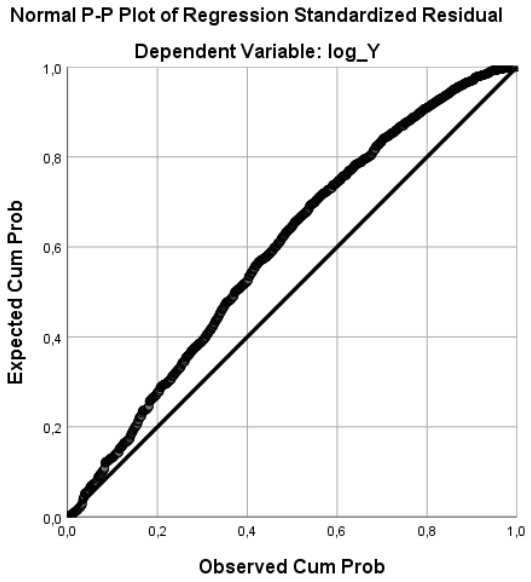
Εικόνα 5.33

Ιστόγραμμα υπολοίπων



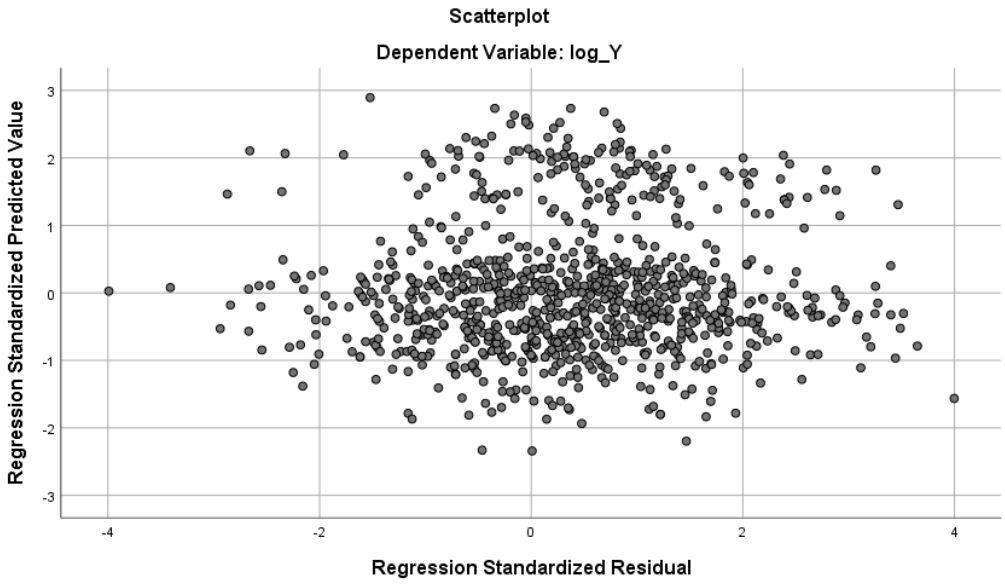
Εικόνα 5.34

Διάγραμμα P-P_ PLot



Εικόνα 5.35

Διάγραμμα διασποράς υπολοίπων.



Η τρίτη ομάδα των χαρακτηριστικών γειτονιάς έδωσε τα παρακάτω αποτελέσματα για την εξαρτημένη log_Y (παράρτημα Γ4). Στον πίνακα 5.39 φαίνεται το R να είναι ίσο με 0,436 και το Adjusted R^2 να είναι ίσο με 0,188. Η εκτίμηση σφάλματος είναι 0,36207, ενώ το F είναι ίσο με 30,459 όπως φαίνεται στον πίνακα 5.40.

Πίνακας 5.39

Γραμμική παλινδρόμηση: Συντελεστές προσδιορισμού

Model Summary[d]

Model	R	R Square	Adjusted R Square	Std. Error of the Estimate
3	,436[c]	,190	,188	,36207

c. Predictors: (Constant), X10_Πρασίνο τμ ανά κάτοικο, X30_Μέσο Εισόδημα κατοίκων, X9_Ποσοστό Κατοίκων ανά ΔΚ

d. Dependent Variable: log_Y

Πίνακας 5.40

Γραμμική παλινδρόμηση: Ανάλυση διασποράς

ANOVAa

Model		Sum of Squares	df	Mean Square	F	Sig.
3	Regression	29,532	3	9,844	75,091	,000[d]
	Residual	125,852	960	,131		
	Total	155,384	963			

Στον πίνακα 5.41 παρουσιάζονται οι συμμετέχουσες μεταβλητές οι οποίες είναι οι : X10, X30, X9. Παράλληλα στις εικόνες 5.36, 5.37, 5.38 παρουσιάζονται το ιστόγραμμα το P- P Plot και το διάγραμμα διασποράς των καταλοίπων.

Πίνακας 5.41

Αποτελέσματα παλινδρόμησης: Συντελεστές παλινδρόμησης

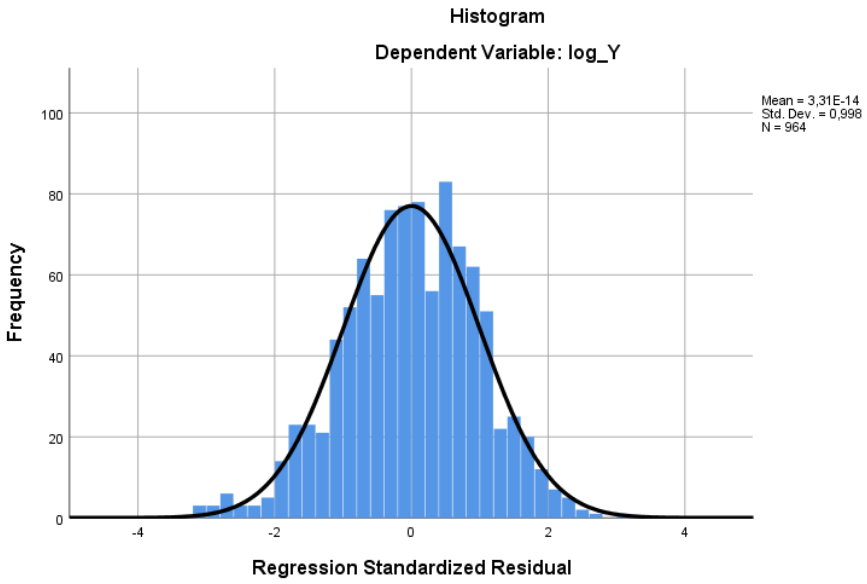
Coefficients[a]

Model		Unstandardized Coefficients		Standardized Coefficients	t	Sig.
		B	Std. Error	Beta		
3	(Constant)	5,494	,096		57,411	,000
	X10_Πρασίνο τμ ανά κάτοικο	,011	,003	,172	4,335	,000
	X30_Μέσο Εισόδημα κατοίκων	-,146	,019	-,241	-7,727	,000
	X9_Ποσοστό Κατοίκων ανά ΔΚ	-,016	,004	-,159	-4,239	,000

a. Dependent Variable: log_Y

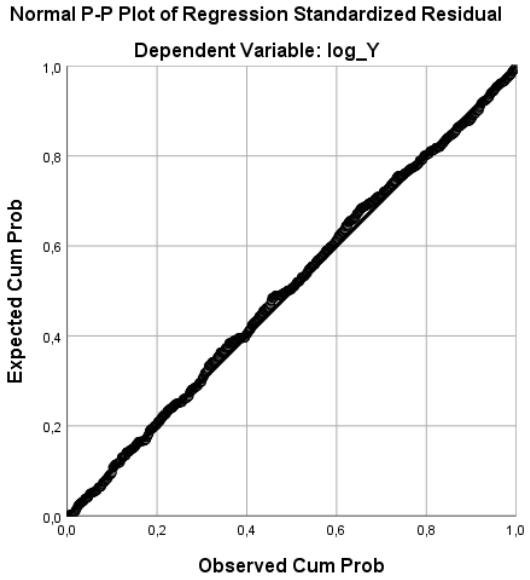
Εικόνα 5.36

Ιστόγραμμα υπολοίπων



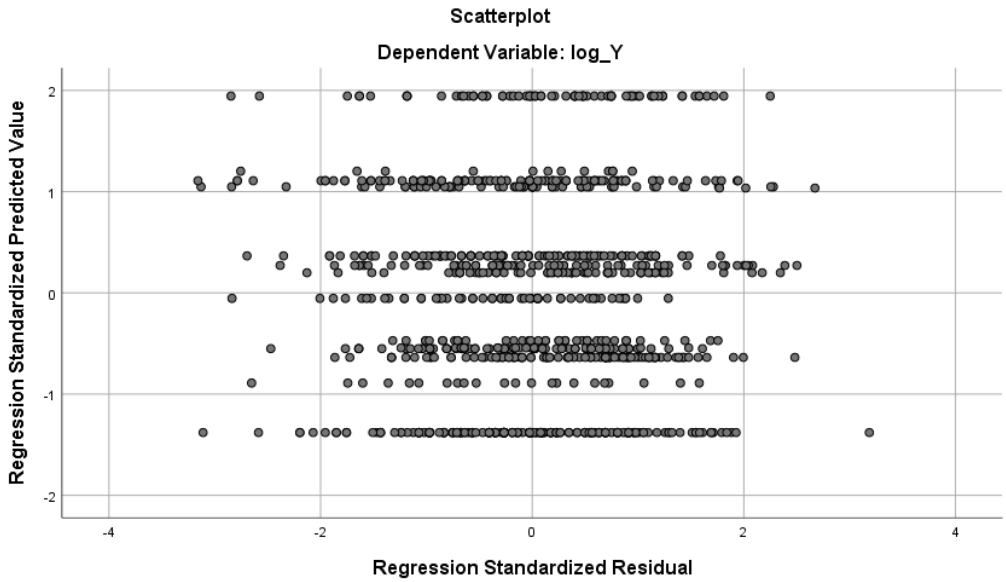
Εικόνα 5.37

Διάγραμμα P-P_ PLot



Εικόνα 5.38

Διάγραμμα διασποράς υπολοίπων.



Η τέταρτη ομάδα των λοιπών μεταβλητών έδωσε τα παρακάτω αποτελέσματα για την εξαρτημένη log_Y (παράρτημα Γ5). Στον πίνακα 5.42 φαίνεται το R να είναι ίσο με 0,440 και το Adjusted R^2 να είναι ίσο με 0,192. Η εκτίμηση σφάλματος είναι 0,36109, ενώ το F είναι ίσο με 115,370 όπως φαίνεται στον πίνακα 5.43.

Πίνακας 5.42

Γραμμική παλινδρόμηση: Συντελεστές προσδιορισμού

Model Summary[c]

Model	R	R Square	Adjusted R Square	Std. Error of the Estimate
2	,440[b]	,194	,192	,36109

b. Predictors: (Constant), X31_Τιμή ζώνης, X7_Ποσοστό Κενών κατοικιών

c. Dependent Variable: log_Y

Πίνακας 5,43

Γραμμική παλινδρόμηση: Ανάλυση διαπσοράς

ANOVA[a]

Model	Sum of Squares	df	Mean Square	F	Sig.

2	Regression	30,085	2	15,042	115,370	,000[c]
	Residual	125,299	961	,130		
	Total	155,384	963			

Στον πίνακα 5.44 παρουσιάζονται οι συμμετέχουσες μεταβλητές οι οποίες είναι οι : X31, X7. Παράλληλα στις εικόνες 5.39, 5.40, 5.41 παρουσιάζονται το ιστόγραμμα το P- P Plot και το διάγραμμα διασποράς των καταλοίπων.

Πίνακας 5.44

Αποτελέσματα παλινδρόμησης: Συντελεστές παλινδρόμησης

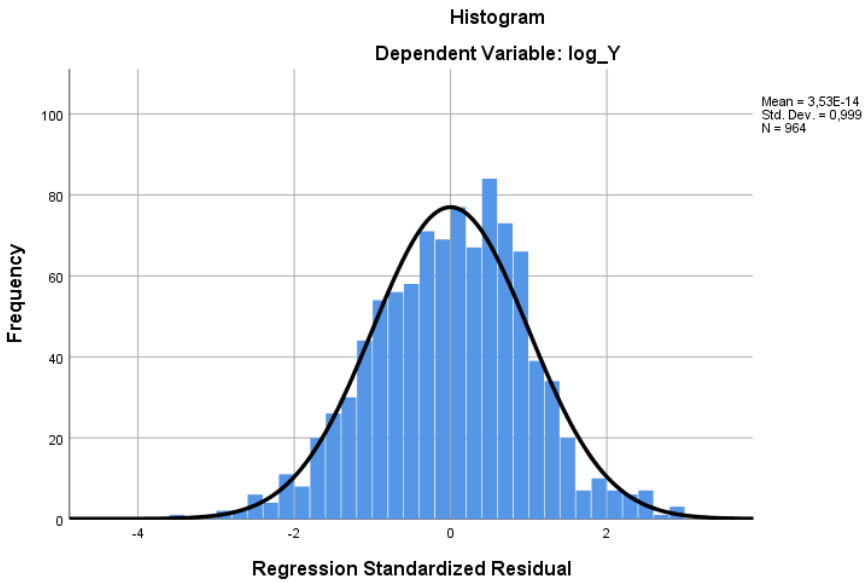
Coefficients[a]

Model		Unstandardized Coefficients		Standardized Coefficients		
		B	Std. Error	Beta	t	Sig.
2	(Constant)	4,947	,098		50,244	,000
	X31_Τιμή ζώνης	,000	,000	,456	15,188	,000
	X7_Ποσοστό Κενών κατοικιών	-,012	,003	-,111	-3,693	,000

a. Dependent Variable: log_Y

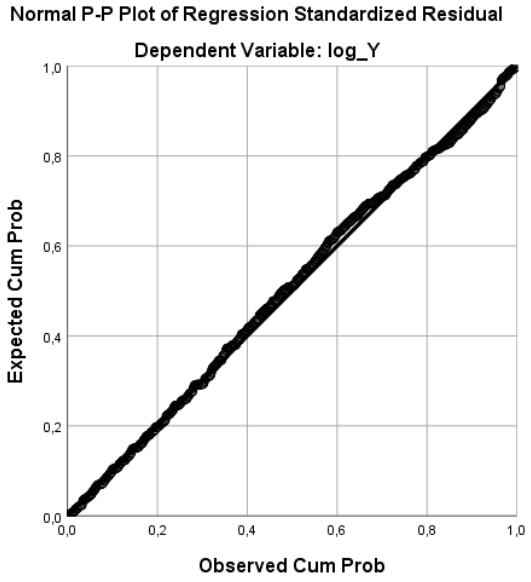
Εικόνα 5.39

Ιστόγραμμα υπολοίπων



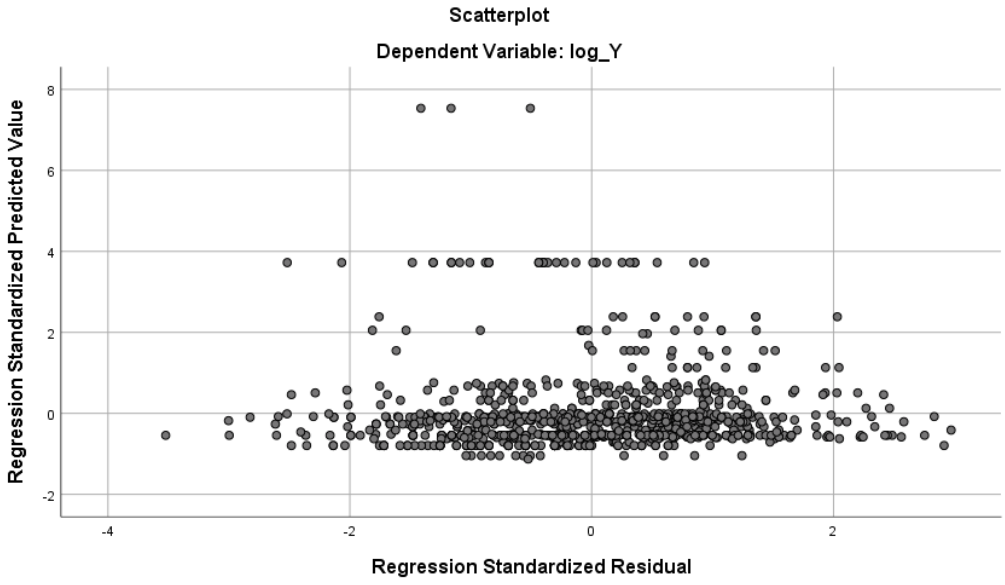
Εικόνα 5.40

Διάγραμμα P-P_ PLot



Εικόνα 5.41

Διάγραμμα διασποράς υπολοίπων.



Από τα παραπάνω προκύπτει ότι από τις τέσσερις ομάδες η ομάδα των κατασκευαστικών χαρακτηριστικών εξηγεί σε μεγαλύτερο βαθμό την εξαρτημένη μεταβλητή της αξίας σε σχέση με την ομάδα μεταβλητών θέσης γειτονιάς και των λοιπών μεταβλητών.

5.4.3.4.1 Προσαρμογή Τελικού Μοντέλου Γραμμικής Παλινδρόμησης.

Στο επόμενο στάδιο χρησιμοποιούμε τις μεταβλητές που έμειναν ως σημαντικές από τις επί μέρους διαδικασίες και δημιουργούμε το τελικό μας μοντέλο για την εξαρτημένη log_Y (Παράρτημα Γ6). Στον πίνακα 5.45 φαίνεται το R να είναι ίσο με 0,860 και το Adjusted R^2 να είναι ίσο με 0,736. Η εκτίμηση σφάλματος είναι 0,20635, ενώ το F είναι ίσο με 235,163 όπως φαίνεται στον πίνακα 5.46.

Πίνακας 5.45

Γραμμική παλινδρόμηση: Συντελεστές προσδιορισμού.

Model Summary[m]

Model	R	R Square	Adjusted R Square	Std. Error of the Estimate
12	,860[l]	,739	,736	,20635

l. Predictors: (Constant), log_32, X42_Katasταση Κατασκευής, X31_Τιμή ζώνης, X9_Ποσοστό Κατοίκων ανά ΔΚ, X38_Όροφος, X48_Είδος Θέρμανσης, X41_Είδος Κατοικίας, log_26, X51_Ύπαρξη Θέας, X37_Ύπαρξη Αποθήκης, X10_Πρασίνο τμ ανά κάτοικο, X7_Ποσοστό Κενών κατοικιών

m. Dependent Variable: log_Y

Πίνακας 5,46

Γραμμική παλινδρόμηση : Ανάλυση διασποράς

ANOVA[a]

Model		Sum of Squares	df	Mean Square	F	Sig.
12	Regression	114,892	12	9,574	224,863	,000[m]
	Residual	40,492	951	,043		
	Total	155,384	963			

Στον πίνακα 5.47 παρουσιάζονται οι συμμετέχουσες μεταβλητές οι οποίες είναι οι : log32, X42, X31, X9, X38, X48, X41, log26, X51, X37, X10, X7.

Πίνακας 5.47

Αποτελέσματα παλινδρόμησης : Συντελεστές παλινδρόμησης

Coefficients[a]

Model		Unstandardized Coefficients		Standardized Coefficients	t	Sig.
		B	Std. Error	Beta		
12	(Constant)	3,423	,125		27,423	,000
	log_32_ Εμβαδόν	1,000	,035	,539	28,594	,000
	X42_Katasταση Κατασκευής	-,054	,006	-,171	-8,844	,000

X31_Τιμή ζώνης	,00014	,000	,215	10,885	,000
X9_Ποσοστό Κατοίκων ανά ΔΚ	-,012	,003	-,122	-4,620	,000
X38_Όροφος	,025	,004	,124	7,002	,000
X48_Είδος Θέρμανσης	-,065	,012	-,109	-5,545	,000
X41_Είδος Κατοικίας	-,101	,019	-,095	-5,191	,000
log_26_ Απόσταση από ΟΣΕ	,063	,018	,073	3,580	,000
X51_Ύπαρξη Θέας	,059	,019	,056	3,075	,002
X37_Ύπαρξη Αποθήκης	,048	,017	,051	2,727	,007
X10_Πρασίνο τμ ανά κάτοικο	,005	,002	,075	3,005	,003
X7_Ποσοστό Κενών κατοικιών	-,004	,002	-,039	-1,996	,046

a. Dependent Variable: log_Y

<u>Η τελική παλινδρόμηση έχει εξίσωση</u> :

log_Y= 3,423 + 1,000* log_32 – 0,054 * X_42 + 0,00014 * X_31 - 0.012 * X_9 + 0.025 * X_38 – 0.065 * X_48 - 0,101 * X_41 + 0,063 * log_26 + 0,059 * X_51 + 0,048 * X_37 + 0.005 * X_10 - 0.004 * X_7

Για να μπορέσουμε να εξάγουμε συμπεράσματα θα πρέπει να θυμηθούμε ότι σε ένα μοντέλο μορφής:

- linear → [Y=α +β·X+ε], κάθε μονάδα αύξησης στην ανεξάρτητη μεταβλητή X μεταβάλει την εξαρτημένη μεταβλητή Y κατά μονάδα ,
- linear- log → [Y=α+β· log(X)+ε], όταν το X μεταβάλλεται κατά 10% το Y μεταβάλλεται κατά log(1,1)·β= 0,095· β
- log- linear → [log(Y)=α+β·X+ε], όταν το X μεταβάλλεται κατά c μονάδες τότε το Y μεταβάλλεται κατά $e^{c\cdot\beta}$, όπου το e≈2,72,
- log-log → [log(Y)=α+β·log·(X)+ε], όταν το X μεταβάλλεται κατά 1% τότε το Y μεταβάλλεται κατά $e^{\beta\cdot\log(1,01)}$,όπου το e≈2,72 (Benoit, 2011).

Από τους συντελεστές της παλινδρόμησης και αναφορικά με τις λογαριθμήσεις παρατηρούμε ότι:

i. η αύξηση κατά 1% στο εμβαδόν αυξάνει την τιμή του ακινήτου κατά 0.433% (όπως προκύπτει από $e^{1\cdot\log(1,01)}$ =1,004333→ 0,433%,

ii. αύξηση κατά 1% στην απόσταση από ΟΣΕ αυξάνει την τιμή του ακινήτου κατά 0,027% (όπως προκύπτει από $e^{0,063\cdot\log(1,01)}$ =1,000272→ 0,0272%

iii. αύξηση κατά μία μονάδα στην κατάσταση κατασκευής μειώνει την τιμή του ακινήτου κατά 5,26% (όπως προκύπτει από $e^{-0,054}$=0,9473→(1-0,9473=0,526)→ -5,26%

iv. αύξηση κατά μία μονάδα στην τιμή ζώνης αυξάνει την τιμή του ακινήτου κατά 0,014% (όπως προκύπτει από $e^{0,00014)}$=1,00014→0,014%

v. αύξηση κατά μία μονάδα στο ποσοστό κατοίκων ανά ΔΚ μειώνει την τιμή του ακινήτου κατά 1,19% (όπως προκύπτει από $e^{-0,012)}$=0,98806→(1-0,98806=0,0119)→-1,19%

vi. αύξηση κατά μία μονάδα στην τιμή του ορόφου (επιπλέον όροφος) αυξάνει την τιμή του ακινήτου κατά 2,533%

vii. αύξηση κατά μία μονάδα στο είδος θέρμανσης μειώνει την τιμή του ακινήτου κατά 6,29%, (υπενθυμίζουμε ότι η αντιστοίχηση ήταν 1=φυσικό αέριο, 2= πετρέλαιο, 3=χωρίς θέρμανση),

viii. αύξηση κατά μία μονάδα στο είδος κατοικίας μειώνει την τιμή του ακινήτου κατά 9,612% (υπενθυμίζουμε ότι η αντιστοίχηση ήταν 1= μονοκατοικία, 2= μεζονέτα, 3= διαμέρισμα),

ix. αύξηση κατά μία μονάδα στην ύπαρξη θέας αυξάνει την τιμή του ακινήτου κατά 6,081%, (υπενθυμίζουμε ότι η αντιστοίχηση ήταν 0= χωρίς, 1=με θέα),

x. αύξηση κατά μία μονάδα στην ύπαρξη αποθήκης αυξάνει την τιμή του ακινήτου κατά 4,92%, (υπενθυμίζουμε ότι η αντιστοίχηση ήταν 0= χωρίς, 1=με αποθήκη),

xi. αύξηση κατά μία μονάδα στην τιμή του τμ πράσινο ανά κάτοικο αυξάνει την τιμή του ακινήτου κατά 0,50%

xii. αύξηση κατά μία μονάδα στο ποσοστό κενών κατοικιών μειώνει την τιμή της κατοικίας κατά 0,399%

Το μοντέλο μας εξηγεί τον λογάριθμο της μεταβλητής « Αξία Υ» και λαμβάνει μέρος ο λογάριθμος της μεταβλητής « Εμβαδόν Χ32» με θετικό πρόσημο καθώς η αύξηση του εμβαδού αυξάνει και την τιμή του ακινήτου. Η ιεραρχική μεταβλητή « Κατάσταση κατασκευής Χ42» συμβάλει με αρνητικό πρόσημο καθώς η ιεράρχηση της μεταβλητής ήταν από άριστη= 1 έως κακή = 5 υποδεικνύοντας ότι όσο αυξάνεται ο αριθμός τόσο αρνητικότερη συνεισφορά έχει στην τιμή της εξαρτημένης μεταβλητής. Η αύξηση των τιμών των μεταβλητών «Τιμή ζώνης Χ31», « Όροφος Χ38», « Ύπαρξη αποθήκης Χ37», « Ύπαρξη θέας Χ51», «Πράσινο ανά κάτοικο

Χ10» συμβάλουν ορθώς θετικά στην εξαρτημένη μεταβλητή. Η μεταβλητή « Ποσοστό κατοίκων ανά ΔΚ Χ9» συμβάλει αρνητικά στην εξαρτημένη μεταβλητή αφού υποδεικνύει μειωμένες τιμές σε πυκνοκατοικημένες περιοχές. Η μεταβλητή « Ποσοστό κενών κατοικιών Χ7» συμβάλει αρνητικά στην εξαρτημένη μεταβλητή υποδεικνύοντας ότι όσο μεγαλύτερο είναι το ποσοστό των κενών κατοικιών τόσο περισσότερο μειώνονται οι τιμές των κατοικιών. Ο λογάριθμος της μεταβλητής « Απόσταση από ΟΣΕ Χ26» συμβάλει θετικά καθώς υποδεικνύει ότι όσο μεγαλύτερη απόσταση έχουν τα ακίνητα από τον σταθμό του ΟΣΕ τόσο μεγαλύτερη τιμή έχουν. Τέλος η ιεραρχική μεταβλητή « Είδος κατοικίας Χ41» συμβάλει αρνητικά . Υπενθυμίζουμε ότι η ιεράρχηση ήταν από μονοκατοικία= 1 μεζονέτα =2 και διαμέρισμα =3 .

5.4.3.5 Έλεγχος Ορθότητας Μοντέλου

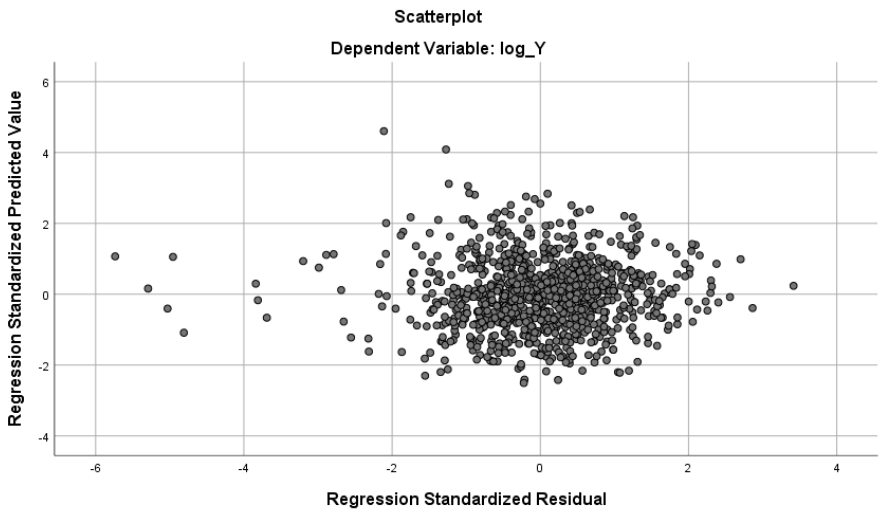
Έπεται ο έλεγχος της ορθότητας του μοντέλου, αναφορικά με την αξιοπιστία των αποτελεσμάτων που δίνει .και την αξιόπιστη προσαρμογή των παρατηρήσεων σε αυτό. Οι υποθέσεις γίνονται στα κατάλοιπα.

Ομοσκεδαστικότητα υπολοίπων.

Παρατηρούμε στο διάγραμμα σκεδασμού, όπως αυτό έχει προκύψει από την επεξεργασία στην εικόνα. 5.42, ότι δεν υπάρχει σταθερή διασπορά και τα σφάλματα δεν φαίνεται να κατανέμονται με τυχαίο τρόπο, γεγονός που φανερώνει ετεροσκεδαστικότητα .

Εικόνα 5.42

Διάγραμμα σκεδασμού υπολοίπων

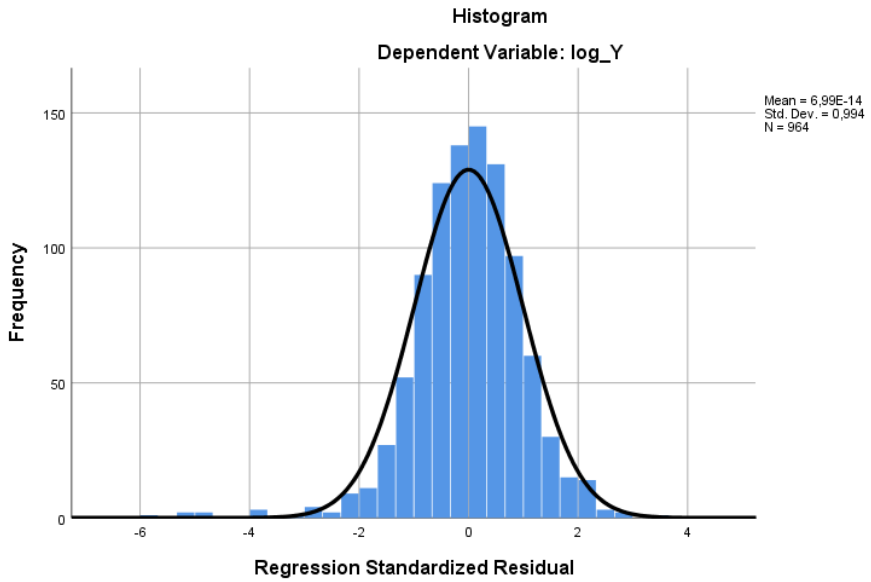


Κανονικότητα υπολοίπων.

Από το ιστόγραμμα (Εικόνα 5.43) δεν προκύπτει η ύπαρξη κανονικής κατανομής καθώς παρατηρούμε εξάρσεις κεντρικά και στήλες εκτός περιοχής στο διάγραμμα. Το ίδιο προκύπτει και από το P-P Plot (Εικόνα 5.44) καθώς η επιθυμία μας τα σημεία να είναι κοντά στην ευθεία, ώστε να αποδειχτεί η κανονικότητα, δεν συμβαίνει στο διάγραμμά μας.

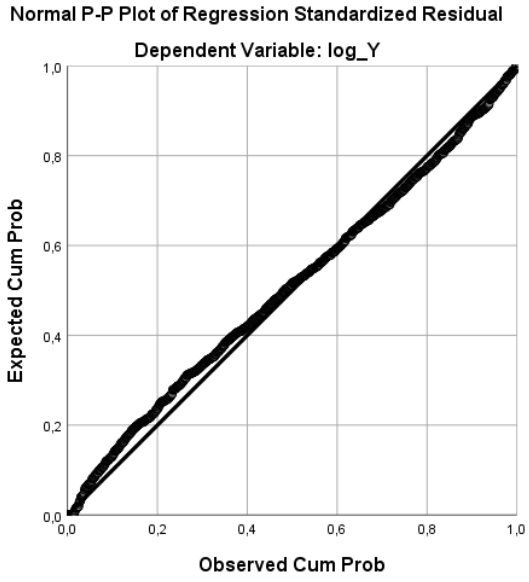
Εικόνα 5.43

Ιστόγραμμα υπολοίπων



Εικόνα 5.44

Διάγραμμα P-P_ Plot



Τέλος με το Kolmogorov-Smirnov (K-S) Test, η μηδενική υπόθεση είναι ότι τα κατάλοιπα ακολουθούν κανονική κατανομή, δεν φαίνεται να ικανοποιείται ως απαίτηση καθώς τα p-value<0.05 (Πίνακας 5.48).

Πίνακας 5.48

Kolmogorov-Smirnov (K-S) Test

Tests of Normality

	Kolmogorov-Smirnov[a]			Shapiro-Wilk		
	Statistic	df	Sig.	Statistic	df	Sig.
Standardized Residual	,045	964	,000	,955	964	,000

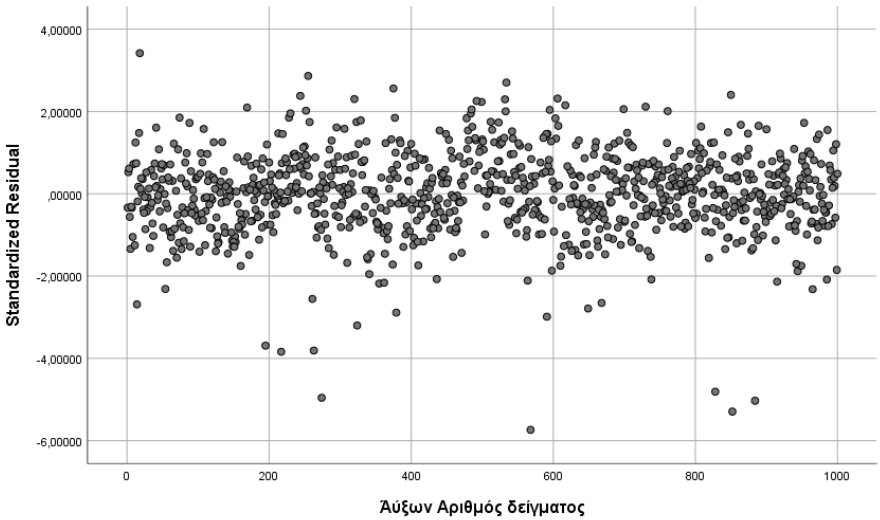
a. Lilliefors Significance Correction

Ανεξαρτησία υπολοίπων.

Παρατηρούμε ότι οι τιμές στο γράφημα σκέδασης scatter plot (Εικόνα 5.45) παρουσιάζουν μερικώς φαινόμενο συσσωρεύσεων σημείων (clustering) και κατ' επέκταση δεν μπορούμε να πούμε ότι τα κατάλοιπα είναι τυχαία.

Εικόνα 5.45

Scatter plot



5.4.3.6 Εξαγωγή Συμπερασμάτων Αποτελεσμάτων Τρίτου Μοντέλου μετά την επεξεργασία με το λογισμικό SPSS

Συμπερασματικά από την δημιουργία του τρίτου μοντέλου μέσα από το λογισμικό SPSS διαπιστώνουμε ότι πρόκειται για ένα μέτρια ικανοποιητικό μοντέλο. Τα κατάλοιπα της παλινδρόμησης δεν ακολουθούν την κανονική κατανομή. Τα

σφάλματα δεν φαίνεται να κατανέμονται με τυχαίο τρόπο, ενώ φαίνεται να έχουμε ετεροσκεδαστικότητα. Το δείγμα είναι χωρικά κατανεμημένο και αυτό θα πρέπει να ληφθεί υπ' όψιν. Στο επόμενο βήμα προχωράμε στην επεξεργασία μέσω του λογισμικού Arc Gis.

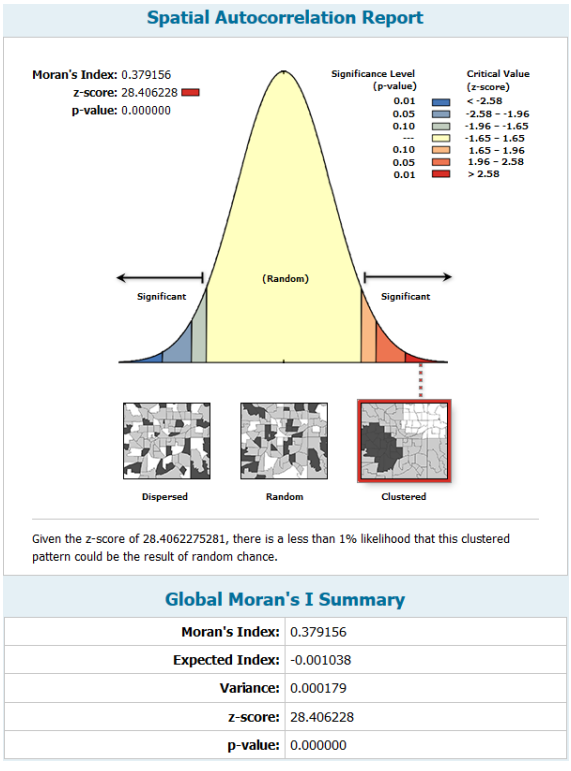
5.4.3.7 Εξέταση του Τελικού Τρίτου Μοντέλου Παλινδρόμησης σε Περιβάλλον Γεωγραφικού Συστήματος Πληροφοριών

5.4.3.7.1 Κλασσική παλινδρόμηση (Ordinary Least Squares Estimation-OLS) σε Περιβάλλον GIS.

Σε πρώτη φάση εφαρμόζουμε τον συντελεστή Moran's I για την εξαρτημένη μεταβλητή «log_ΥΑξία «. Τα αποτελέσματα παρουσιάζονται στην εικόνα 5.46 Παρατηρούμε ότι η εξαρτημένη μεταβλητή παρουσιάζει τιμή στο δείκτη ίση με 0,3791 και z-score ίσο με 28,4062, ενώ το p=0,000. Παράλληλα παρουσιάζεται φαινόμενο συγκέντρωσης (cluster), η σημαντικότητα δε του φαινομένου αυτού ενδυναμώνεται και από την υψηλή τιμή του z score, όπου δηλώνεται ότι το φαινόμενο cluster έχει πιθανότητα μικρότερη του 1% να οφείλεται σε τυχαίους παράγοντες.

Εικόνα 5.46

Εξέταση συντελεστή Moran's I για την εξαρτημένη μεταβλητή



Global Moran's I Summary

Moran's Index:	0.379156
Expected Index:	-0.001038
Variance:	0.000179
z-score:	28.406228
p-value:	0.000000

Συνεχίζουμε ακολουθώντας την διαδικασία όπως αυτή παρουσιάστηκε στο δεύτερο μοντέλο εισάγουμε μόνο εκείνες οι μεταβλητές που λειτούργησαν για την δημιουργία του τρίτου μοντέλου στο SPSS.

Τα αποτελέσματα της διαδικασίας παρουσιάζονται στον πίνακα 5.49 ενώ στον πίνακα 5.50 παρουσιάζονται τα διαγνωστικά της παλινδρόμησης. Παρατηρούμε ότι το adjusted R^2 =0,7361 το F=224,863 και το AIC είναι ίσο με -291,707 .

Τα αποτελέσματα ταυτίζονται απόλυτα με το μοντέλο όπως αυτό είχε δημιουργεί με το SPSS και παρουσιάζονται αναλυτικά στο παράρτημα «Γ7».

Πίνακας 5.49

Αποτελέσματα OLS διαδικασίας τρίτο μοντέλο

Summary of OLS Results - Model Variables

Variable	Coefficient [a]	StdError	t-Statistic	Probability [b]	Robust_SE	Robust_t	Robust_Pr [b]	VIF [c]
Intercept	3.423028	0.124824	27.422896	0.000000*	0.124156	27.570443	0.000000*	--------
X7_POSOSTO	-0.004093	0.002050	-1.996492	0.046157*	0.002051	-1.995468	0.046269*	1.374201
X9_POSOSTO	-0.012356	0.002675	-4.619942	0.000006*	0.002470	-5.003188	0.000001*	2.538353
X10_TM_PRA	0.004852	0.001615	3.004890	0.002737*	0.001525	3.182245	0.001523*	2.282392
X31_TIMH_Z	0.000140	0.000013	10.885268	0.000000*	0.000012	11.303914	0.000000*	1.425260
X37_YPARXH	0.047545	0.017437	2.726593	0.006516*	0.018335	2.593057	0.009651*	1.269080
X38_OROFOS	0.025474	0.003638	7.001587	0.000000*	0.003952	6.446112	0.000000*	1.146810
X41_EIDOS_	-0.100894	0.019435	-5.191376	0.000001*	0.021048	-4.793460	0.000003*	1.224165
X42_KATAST	-0.054073	0.006114	-8.844320	0.000000*	0.005887	-9.185811	0.000000*	1.367102
X48_EIDOS_	-0.065260	0.011770	-5.544721	0.000000*	0.012072	-5.405887	0.000000*	1.399704
X51_UEA_X5	0.058622	0.019066	3.074583	0.002180*	0.019673	2.979720	0.002968*	1.228324
LOG_26	0.063047	0.017612	3.579857	0.000375*	0.016687	3.778255	0.000179*	1.524409
LOG_32	0.999919	0.034970	28.593653	0.000000*	0.033320	30.009147	0.000000*	1.298708

Πίνακας 5.50

Διαγνωστικά παλινδρόμησης OLS διαδικασίας

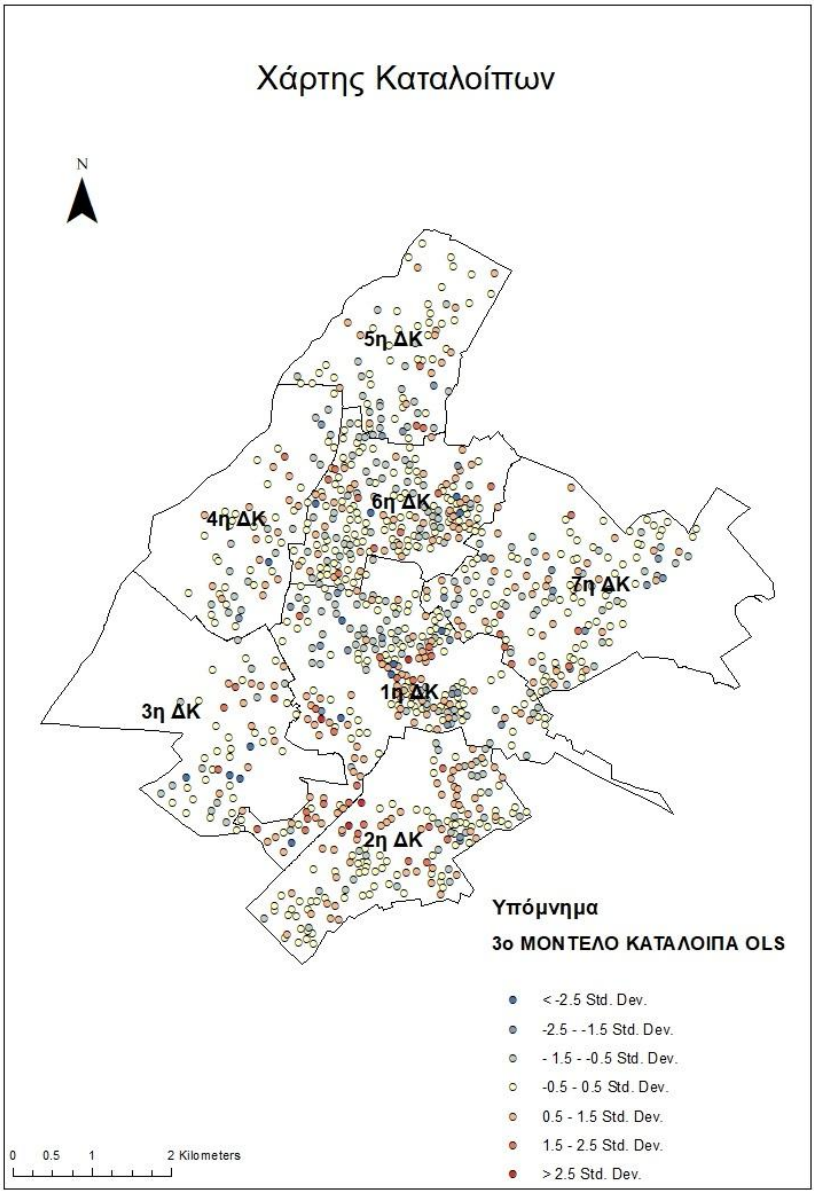
OLS Diagnostics

Input Features:	3_OLS_MONTELO_TELIKO	Dependent Variable:	LOG_Y
Number of Observations:	964	Akaike's Information Criterion (AICc) [d]:	-291.707811
Multiple R-Squared [d]:	0.739406	Adjusted R-Squared [d]:	0.736118
Joint F-Statistic [e]:	224.863198	Prob(>F), (12,951) degrees of freedom:	0.000000*
Joint Wald Statistic [e]:	3013.413096	Prob(>chi-squared), (12) degrees of freedom:	0.000000*
Koenker (BP) Statistic [f]:	19.503418	Prob(>chi-squared), (12) degrees of freedom:	0.077082
Jarque-Bera Statistic [g]:	719.738043	Prob(>chi-squared), (2) degrees of freedom:	0.000000*

Στο χάρτη 5.9 παρουσιάζεται η χαρτογράφηση των καταλοίπων ταξινομημένα σε επτά κλάσεις, ενώ στην εικόνα 5.47 παρουσιάζεται το ιστόγραμμα αυτών και στην εικόνα 5.48 παρουσιάζεται το διάγραμμα διασποράς τους. Υποδεικνύεται η μη κανονική κατανομή τους.

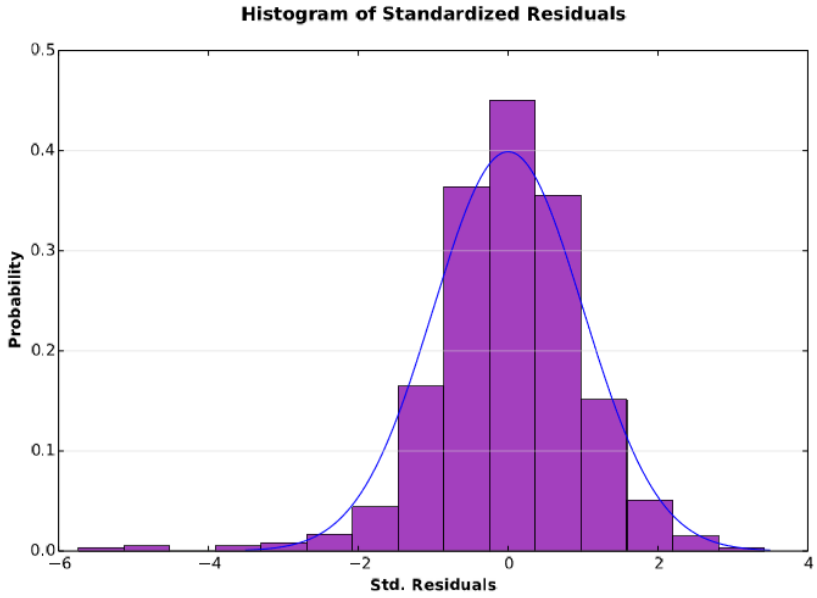
Χάρτης 5.9

Χαρτογράφηση Καταλοίπων τρίτου μοντέλου



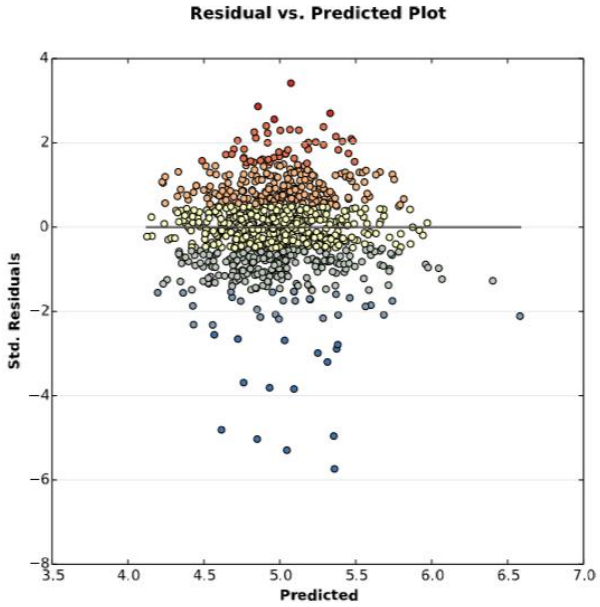
Εικόνα 5.47

Ιστόγραμμα καταλοίπων τρίτου μοντέλου



Εικόνα 5.48

Διάγραμμα διασποράς καταλοίπων τρίτου μοντέλου

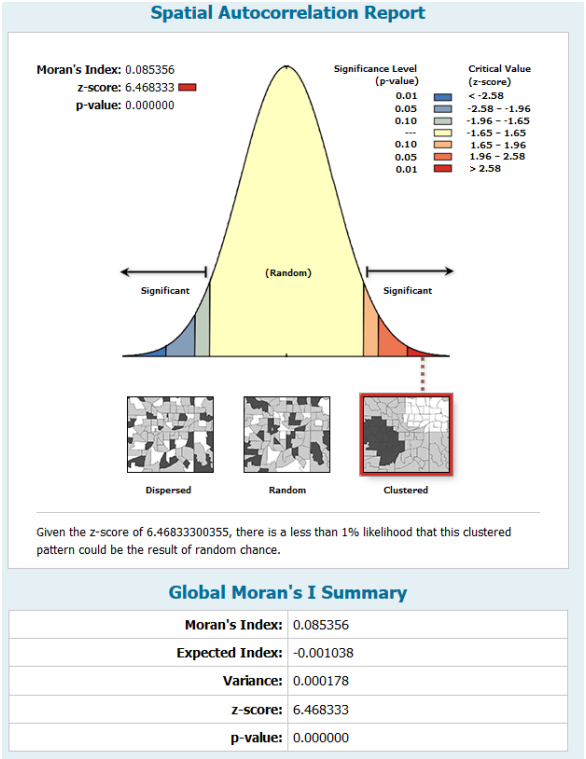


Για τον έλεγχο τη χωρικής αυτοσυσχέτισης των υπολοίπων χρησιμοποιούμε και τον δείκτη Morans I (Εικόνα 5.49). Παρατηρούμε ότι τα κατάλοιπα της παλινδρόμησης παρουσιάζουν τιμή στο δείκτη ίση με 0,0853 και z-score ίσο με

6,464, ενώ το p=0,0000. Παράλληλα παρουσιάζεται φαινόμενο συγκέντρωσης (cluster). Η σημαντικότητα του φαινομένου υποδεικνύεται από την τιμή του z score, όπου δηλώνεται ότι το φαινόμενο cluster έχει πιθανότητα μικρότερη του 1% να οφείλεται σε τυχαίους παράγοντες. Η τιμή Moran's I είναι θετική αλλά κοντά στο μηδέν και έτσι υποδεικνύεται ότι υπάρχει μικρή συγκέντρωση χωρικά παρόμοιων τιμών. Γενικά τιμές κοντά στο μηδέν υποδεικνύουν έλλειψη χωρικής εξάρτησης μεταξύ των γειτνιαζόντων παρατηρήσεων.

Εικόνα 5.49

Έλεγχος χωρικής αυτοσυσχέτισης καταλοίπων τρίτου μοντέλου



Ελέγχοντας τα κατάλοιπα της OLS διαδικασίας, με τον υπολογισμό του Moran I, διαπιστώνουμε την ύπαρξη χωρικής αυτοσυσχέτισης, δηλαδή clustered χωρικό πρότυπο και γι' αυτό τον λόγο κρίνεται λοιπόν σκόπιμο να προχωρήσουμε στη γεωγραφικά σταθμισμένη παλινδρόμηση.

5.4.3.7.2 Χωρική Παλινδρόμηση (Geographically Weighted Regression-GWR) σε περιβάλλον GIS.

Η εφαρμογή της GWR στο μοντέλο μας γίνεται και εδώ μέσω του open source λογισμικού GWR4. Τα αποτελέσματα φαίνονται στον πίνακα 5.51 και αναλυτικά παρουσιάζονται στο παράρτημα «Γ8» , ενώ αναλυτικά αποτελέσματα της διαδικασίας για κάθε ακίνητο παρουσιάζονται στο παράτημα «Γ9».
Παρατηρούμε ότι μέσω της διαδικασίας αυτής το Adjusted R square που στην OLS ήταν 0,736, τώρα είναι 0,7574 ενώ παράλληλα το AICc από -291.707 που ήταν στην OLS διαδικασία, υπολογίστηκε σε -352,666

Πίνακας 5.51
Αποτελέσματα GWR διαδικασίας

GWR (Geographically weighted regression) result

Bandwidth and geographic ranges
Bandwidth size: 554.470114

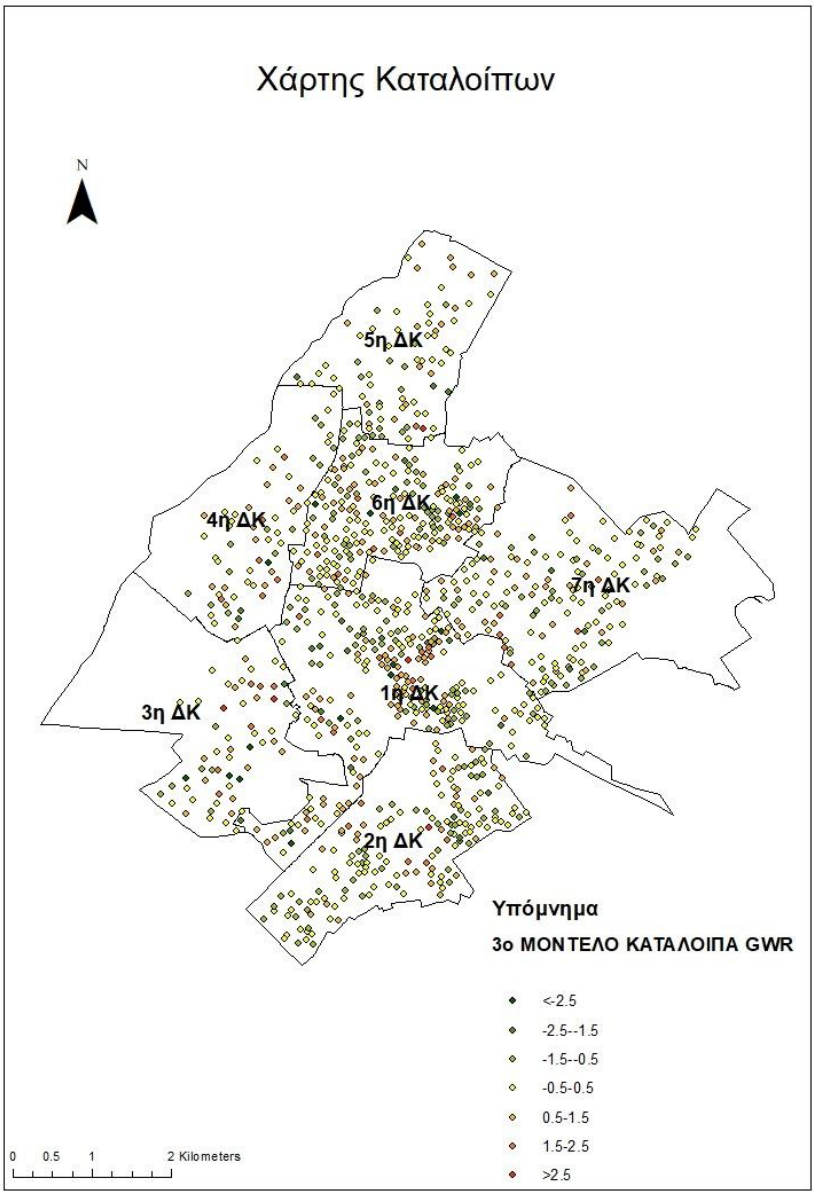
Coordinate	Min	Max	Range
X-coord	473725.000000	480487.000000	6762.000000
Y-coord	4200150.000000	4208998.000000	8848.000000

Diagnostic information
Residual sum of squares: 35.407901
Effective number of parameters (model: trace(S)): 45.050626
Effective number of parameters (variance: trace(S'S)): 31.038436
Degree of freedom (model: n - trace(S)): 918.949374
Degree of freedom (residual: n - 2trace(S) + trace(S'S)): 904.937184
ML based sigma estimate: 0.191651
Unbiased sigma estimate: 0.197807
-2 log-likelihood: -449.493195
Classic AIC: -357.391943
AICc: -352.666031
BIC/MDL: -133.075139
CV: 0.040470
R square: 0.772361
Adjusted R square: 0.757487

Στο χάρτη 5.10 φαίνονται τα τυποποιημένα κατάλοιπα (standardized) της διαδικασίας GWR. Από τον χάρτη αυτό σε σχέση με προηγούμενο δεν παρατηρείται μείωση του μεγέθους των καταλοίπων ενώ το ίδιο προκύπτει και από το ιστόγραμμα (Εικόνα 5.50)

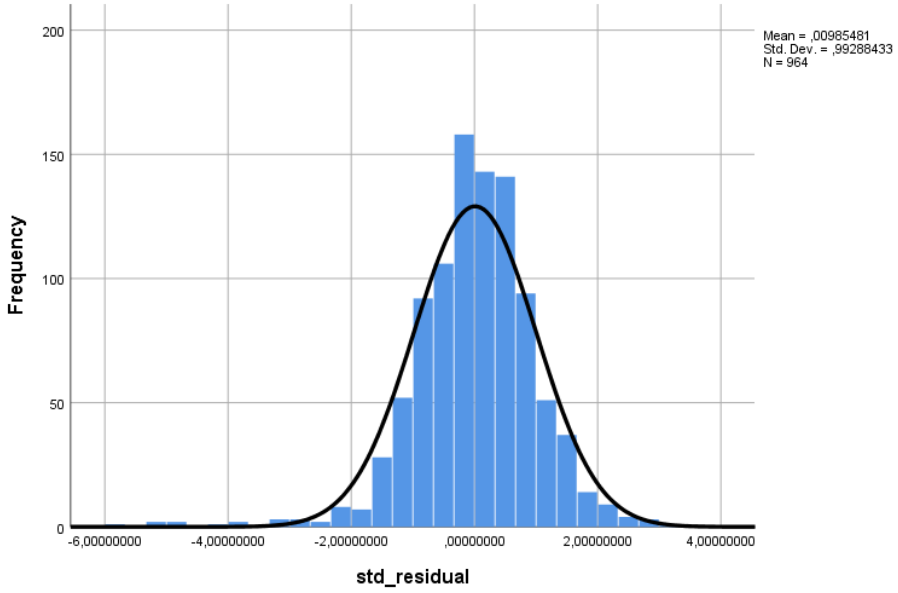
Χάρτης 5.10

Χάρτης καταλοίπων GWR διαδικασίας



Εικόνα 5.50

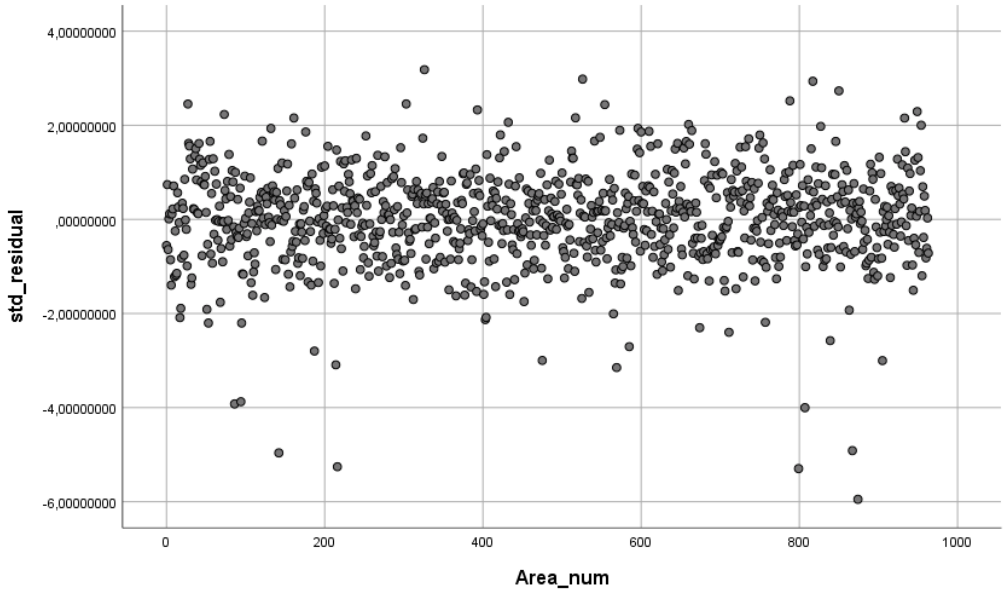
Ιστόγραμμα καταλοίπων



Απεικονίζοντας το scatter plot (Εικόνα 5.51) και υπολογίζοντας τον συντελεστή αυτοσυσχέτισης Moran I (Εικόνα 5.52) για τα κατάλοιπα της GWR δεν προκύπτει τυχαίο πρότυπο, δηλαδή τα κατάλοιπα φαίνεται να παρουσιάζουν ακόμα χωρική αυτοσυσχέτιση.

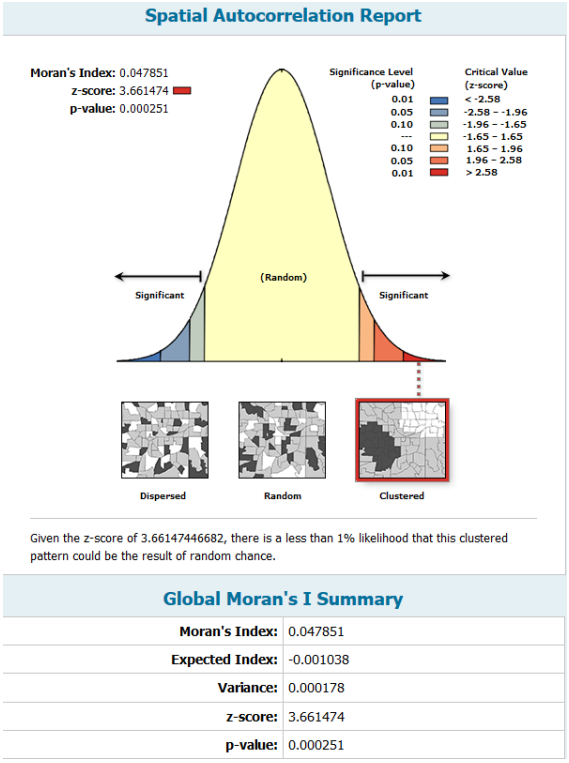
Εικόνα 5.51

Διάγραμμα διασποράς υπολοίπων.



Εικόνα 5.52

Υπολογισμός Morans I για τα κατάλοιπα



5.4.3.7.3 Εξαγωγή Συμπερασμάτων των Αποτελεσμάτων του Τρίτου Μοντέλου μέσα από την Χωρική παλινδρόμηση (Geographically Weighted Regression-GWR.)

Συμπερασματικά από την επεξεργασία του τρίτου μοντέλου μέσα από την χρήση της χωρικής παλινδρόμησης διαπιστώνουμε ότι αυτό δεν βελτιώθηκε (βλέπε πίνακα 5.52) δεν παρατηρείται ουσιαστική μείωση του μεγέθους των καταλοίπων, ενώ αυτά φαίνεται ακόμα να παρουσιάζουν χωρική αυτοσυσχέτιση (clustered φαινόμενο).

Πίνακας 5.52

Σύγκριση μοντέλων OLS & GWR διαδικασίας.

Συντελεστές	OLS	GWR
Adjusted R Square	0.736	0.7574
Morans I residuals	0.085	0.047
AICc	-291.700	-352.666

Είναι φανερό λοιπόν ότι από το συγκεκριμένο μοντέλο δεν προέκυψε καλύτερη ερμηνεία της εξαρτημένης μεταβλητής και ακριβέστερος προσδιορισμός του μοντέλου της παλινδρόμησης.

5.4.4 Δημιουργία Τέταρτου Μοντέλου

Στο τέταρτο μοντέλο ακολουθείται η ίδια διαδικασία δημιουργίας εξίσωσης παλινδρόμησης με την χρήση μετασχηματισμών με την βοήθεια λογαρίθμων και ρίζας, στο εναπομένων δείγμα, όπως αυτό προέκυψε από το τρίτο μοντέλο με την αφαίρεση των ακραίων οντοτήτων. Η διάφορα μεταξύ των δύο αυτών μοντέλων είναι ότι στη συγκεκριμένη περίπτωση η εισαγωγή των επιλεγμένων μεταβλητών γίνεται μαζικά και όχι επιλεκτικά ως αποτέλεσμα ομαδοποίησης. Η τελική εξίσωση παλινδρόμησης που παράχθηκε από το λογισμικό SPSS επεξεργάστηκε και με την GWR διαδικασία.

5.4.4.1 Προσαρμογή Μοντέλου Γραμμικής Παλινδρόμησης

Οι κατάλληλοι μετασχηματισμοί των μεταβλητών έγιναν σύμφωνα με τα αναφερόμενα στο τρίτο μοντέλο και οι μεταβλητές που λαμβάνουν τελικά μέρος στο μοντέλο αυτό είναι οι : log_ εμβαδόν32, Αριθμός υπνοδωματίων_X34, Αριθμός μπάνιων_X35, Ύπαρξη WC_X36, Ύπαρξη αποθήκης_X37, Όροφος_X38, Ηλικία κατασκευής_X40, Είδος κατοικίας_X41, Κατάσταση κατασκευής_X42, Ύπαρξη ανακαίνισης_X43, Ύπαρξη βεράντας_X44, Ύπαρξη ασανσέρ_X46, Ύπαρξη θέσης στάθμευσης_X47, Είδος θέρμανσης_X48, Είδος καυσίμου θέρμανσης_X49, Ύπαρξη κλιματισμού_X50, Ύπαρξη τζακιού_X52, sqrt_Απόσταση από αστυνομικό τμήμα11, Απόσταση από βιομηχανία_X12, sqrt_Απόσταση από βρεφονηπιακό σταθμό13, sqrt_Απόσταση από εμπόριο14, log_Απόσταση από ιστορικό κέντρο15, log_Απόσταση από νεκροταφείο16, sqrt_Απόσταση από νοσοκομείο17, Είδος πλησιέστερης συγκοινωνίας_X18, log_Απόσταση από πλησιέστερη συγκοινωνία19, log_Απόσταση από ναό20, sqrt_ Απόσταση από σχολείο21, log_ Απόσταση από αθλητισμό22, log_Απόσταση από πράσινο23, sqrt_Απόσταση από Πλάκα24, sqrt_Απόσταση από πεζόδρομο25,log_Απόσταση από ΟΣΕ26, log_Απόσταση από Μουσείο27, log_Απόσταση από πρωτεύον δίκτυο28, log_Απόσταση από βασικό δίκτυο29,Γωνιακό ή διαμπερές_X45, Ύπαρξη θέας_X51, Ποσοστό ελλήνων_X4, Ποσοστό ανεργίας_X6, Ποσοστό πρασίνου_X8, Ποσοστό κατοίκων ανά ΔΚ_ X9, τμ

πρασίνου ανά κάτοικο_X10, Μέσο εισόδημα κατοίκων_X30, Ποσοστό κενών κατοικιών_X7, Τιμή ζώνης_X31, log_Αριθμός ημερών προς διάθεση53 .

Η επεξεργασία του τέταρτου μοντέλου έδωσε τα παρακάτω αποτελέσματα για την εξαρτημένη μεταβλητή log_Y (παράρτημα Δ1). Στον πίνακα 5.53 φαίνεται το R να είναι ίσο με 0,863 και το Adjusted R^2 να είναι ίσο με 0,738. Η εκτίμηση σφάλματος είναι 0,17508 , ενώ το F είναι ίσο με 123,414 όπως φαίνεται στον πίνακα 5.54

Πίνακας 5.53

Γραμμική παλινδρόμηση: Συντελεστές προσδιορισμού

Model Summary[o]

Model	R	R Square	Adjusted R Square	Std. Error of the Estimate
14	,863[n]	,744	,738	,17508

n. Predictors: (Constant), log_32, X40_Ηλικία Κατασκευής, X8_Ποσοστό Πρασίνου, X48_Είδος Θέρμανσης, log_16, X47_Ύπαρξη Θέσης στάθμευσης, X38_Όροφος, X42_Κατασταση Κατασκευής, X41_Είδος Κατοικίας, X9_Ποσοστό Κατοίκων ανά ΔΚ, log_26, X31_Τιμή ζώνης, SQRT_17, log_22

o. Dependent Variable: log_Y

Πίνακας 5.54

Γραμμική παλινδρόμηση: Ανάλυση διασποράς

ANOVA[a]

Model		Sum of Squares	df	Mean Square	F	Sig.
14	Regression	52,961	14	3,783	123,414	,000[o]
	Residual	18,177	593	,031		
	Total	71,137	607			

Στον πίνακα 5.55 παρουσιάζονται οι συμμετέχουσες μεταβλητές οι οποίες είναι οι : log_32, X40, X8, X48, log_16, X47, X38,X42, X41, X9, log_26, X31, sqrt_17, log_22.

Πίνακας 5.55

Αποτελέσματα παλινδρόμησης: Συντελεστές παλινδρόμησης

Coefficients[a]

Model		Unstandardized Coefficients		Standardized Coefficients	t	Sig.
		B	Std. Error	Beta		
14	(Constant)	3,375	,168		20,122	,000

log_32_Εμβαδόν	1,047	,040	,599	26,258	,000
X40_Ηλικία Κατασκευής	-,002	,000	-,133	-4,528	,000
X8_Ποσοστό Πρασίνου	,005	,001	,147	4,205	,000
X48_Είδος Θέρμανσης	-,074	,014	-,138	-5,390	,000
log_16_Απόσταση από νεκροταφείο	-,056	,026	-,053	-2,146	,032
X47_Ύπαρξη Θέσης στάθμευσης	,076	,024	,084	3,252	,001
X38_Όροφος	,022	,004	,126	5,641	,000
X42_Κατασταση Κατασκευής	-,031	,007	-,111	-4,402	,000
X41_Είδος Κατοικίας	-,076	,022	-,080	-3,550	,000
X9_Ποσοστό Κατοίκων ανά ΔΚ	-,010	,002	-,103	-4,257	,000
log_26_ Απόσταση από ΟΣΕ	,041	,022	,060	1,914	,056
X31_Τιμή ζώνης	,00017	,000	,088	3,083	,002
SQRT_17_ Νοσοκομείο	,003	,001	,075	2,670	,008
log_22_ Απόσταση από αθλητισμό	-,056	,027	-,045	-2,099	,036

a. Dependent Variable: log_Y

Από τους συντελεστές της παλινδρόμησης και αναφορικά με τις λογαριθμήσεις παρατηρούμε ότι:

i. η αύξηση κατά 1% στο εμβαδόν αυξάνει την τιμή του ακινήτου κατά 0.453% (όπως προκύπτει από $e^{1,047 \cdot \log(1,01)}$ =1,00453→ 0,453%,

ii. αύξηση κατά 1% στην απόσταση από ΟΣΕ αυξάνει την τιμή του ακινήτου κατά 0,017% (όπως προκύπτει από $e^{0,041 \cdot \log(1,01)}$ =1,000177→ 0,0177%

iii. αύξηση κατά μία μονάδα στην κατάσταση κατασκευής μειώνει την τιμή του ακινήτου κατά 5,26% (όπως προκύπτει από $e^{-0.031)}$=0,969→(1-0,969=0,03)→ -3,05%

iv. αύξηση κατά 1% στην απόσταση από αθλητισμό μειώνει την τιμή του ακινήτου κατά 0,024%

v. αύξηση κατά 1% στην απόσταση από νεκροταφείο μειώνει την τιμή του ακινήτου κατά 0,024%

vi. αύξηση κατά μία μονάδα στην τιμή ζώνης αυξάνει την τιμή του ακινήτου κατά 0,017%

vii. αύξηση κατά μία μονάδα στο ποσοστό κατοίκων ανά ΔΚ μειώνει την τιμή του ακινήτου κατά 0,995%

viii. αύξηση κατά μία μονάδα στην τιμή του ορόφου (επιπλέον όροφος) αυξάνει την τιμή του ακινήτου κατά 2,225%

ix. αύξηση κατά μία μονάδα στο είδος θέρμανσης μειώνει την τιμή του ακινήτου κατά 7,137%, (υπενθυμίζουμε ότι η αντιστοίχηση ήταν 1=φυσικό αέριο, 2= πετρέλαιο, 3=χωρίς θέρμανση),

x. αύξηση κατά μία μονάδα στο είδος κατοικίας μειώνει την τιμή του ακινήτου κατά 7,322% (υπενθυμίζουμε ότι η αντιστοίχηση ήταν 1= μονοκατοικία, 2= μεζονέτα, 3= διαμέρισμα),

xi. αύξηση κατά μία μονάδα στην ύπαρξη θέσης στάθμευσης αυξάνει την τιμή του ακινήτου κατά 7,901%, (υπενθυμίζουμε ότι η αντιστοίχηση ήταν 0= χωρίς, 1=με θέση στάθμευσης),

xii. αύξηση κατά μία μονάδα στην τιμή του ποσοστό πράσινο αυξάνει την τιμή του ακινήτου κατά 0,501%

xiii. αύξηση κατά μία μονάδα στην ηλικία κατασκευής μειώνει την τιμή της κατοικίας κατά 0,199%

xiv. αύξηση κατά μία μονάδα στην κατάσταση κατασκευής μειώνει την τιμή του ακινήτου κατά 3,054% (υπενθυμίζουμε ότι η αντιστοίχηση ήταν 1= άριστη…….5= κακή).

xv. Αύξηση κατά μία μονάδα της SQRT(απόσταση από νοσοκομείο) αυξάνει την τιμή του ακινήτου κατά 0,300%.

Το μοντέλο μας εξηγεί τον λογάριθμο της μεταβλητής « Αξία Υ» και σε αυτό λαμβάνει μέρος ο λογάριθμος της μεταβλητής « Εμβαδόν Χ32» με θετικό πρόσημο καθώς η αύξηση του εμβαδού αυξάνει και την τιμή του ακινήτου. Η ιεραρχική μεταβλητή « Κατάσταση κατασκευής Χ42» συμβάλει με αρνητικό πρόσημο καθώς η ιεράρχηση της μεταβλητής ήταν από άριστη= 1 έως κακή = 5 υποδεικνύοντας ότι όσο αυξάνεται ο αριθμός τόσο αρνητικότερη συνεισφορά έχει στην τιμή της εξαρτημένης μεταβλητής. Η αύξηση των τιμών των μεταβλητών «Τιμή ζώνης Χ31», « Όροφος Χ38», « Ύπαρξη θέσης στάθμευσης Χ47», « Ποσοστό πρασίνου Χ8» συμβάλουν ορθώς θετικά στην εξαρτημένη μεταβλητή. Η μεταβλητή « Ηλικία κατασκευής Χ40» συμβάλει αρνητικά στην εξαρτημένη μεταβλητή αφού υποδεικνύει μειωμένες τιμές σε παλαιές κατασκευές. Η μεταβλητή « Ποσοστό κατοίκων ανά ΔΚ Χ9» συμβάλει αρνητικά στην εξαρτημένη μεταβλητή υποδεικνύοντας ότι όσο πιο

πυκνοκατοικημένη είναι η περιοχή τόσο περισσότερο μειώνονται είναι οι τιμές των κατοικιών. Ο λογάριθμος της μεταβλητής « Απόσταση από ΟΣΕ Χ26» συμβάλει θετικά καθώς υποδεικνύει ότι όσο μεγαλύτερη απόσταση έχουν τα ακίνητα από τον σταθμό του ΟΣΕ τόσο μεγαλύτερη τιμή έχουν. Η ιεραρχική μεταβλητή « Είδος κατοικίας Χ41» συμβάλει αρνητικά. Υπενθυμίζουμε ότι η ιεράρχηση ήταν από μονοκατοικία= 1 μεζονέτα =2 και διαμέρισμα =3 . Η ιεραρχική μεταβλητή « Είδος θέρμανσης Χ48» συμβάλει αρνητικά καθώς η ιεράρχηση ήταν Αυτόνομη=1, Κεντρική= 2 και Χωρίς θέρμανση =3. Έτσι για παράδειγμα εάν έχουμε ακίνητο χωρίς θέρμανση η τιμή του είναι μειωμένη εν σχέση με ένα ακίνητο με θέρμανση . Οι μεταβλητές log « Απόσταση από Νεκροταφείο Χ16» και log « απόσταση από αθλητισμό Χ22» συμβάλουν αρνητικά στην εξαρτημένη υποδεικνύοντας ότι όσο μεγαλύτερη είναι η απόσταση τόσο περισσότερο μειώνεται η τιμή του ακινήτου. Αυτό βέβαια δεν είναι λογικό για την μεταβλητή της απόστασης από νεκροταφείο καθώς κάποιος θα περίμενε το ανάποδο. Η λογική εξήγηση είναι ότι το βάρος λογικά θα πρέπει να έχει πέσει στα ακίνητα που βρίσκονται γύρω από το Α’ νεκροταφείο Αθηνών που ουσιαστικά βρίσκονται σε μια από τις καλές περιοχές της πόλης. Τέλος η μεταβλητή sqrt « Απόσταση από νοσοκομείο Χ17» συμβάλει θετικά στο μοντέλο υποδεικνύοντας την αυξημένη τιμή του ακινήτου όσο η απόσταση από νοσοκομείο μεγαλώνει. Φυσιολογικά θα περίμενε κανείς αυξημένες τιμές πέριξ υγειονομικών εξυπηρετήσεων όμως θα πρέπει να δούμε το θέμα και μέσα από το πρίσμα της όχλησης που προκαλείται από την έντονη κυκλοφορία πέριξ ενός νοσοκομείου.

5.4.4.2 Έλεγχος Ορθότητας Μοντέλου

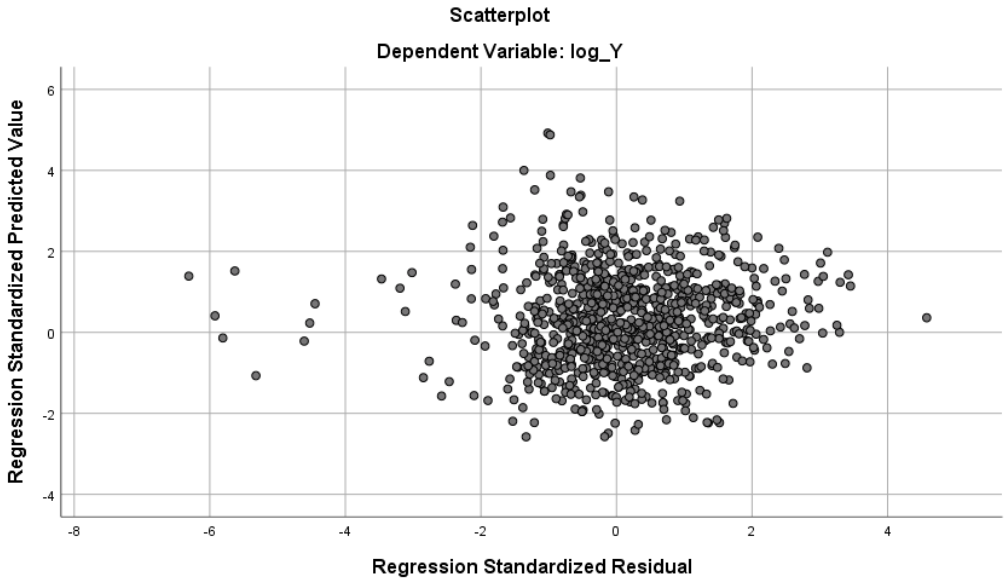
Έπεται ο έλεγχος της ορθότητας του μοντέλου, αναφορικά με την αξιοπιστία των αποτελεσμάτων που δίνει .και την αξιόπιστη προσαρμογή των παρατηρήσεων σε αυτό. Οι υποθέσεις γίνονται στα κατάλοιπα.

Ομοσκεδαστικότητα υπολοίπων.

Παρατηρούμε στο διάγραμμα σκεδασμού, όπως αυτό έχει προκύψει από την επεξεργασία στην εικόνα.5.53, ότι δεν υπάρχει σταθερή διασπορά και τα σφάλματα δεν φαίνεται να κατανέμονται με τυχαίο τρόπο, γεγονός που φανερώνει ετεροσκεδαστικότητα .

Εικόνα 5.53

Διάγραμμα σκεδασμού υπολοίπων

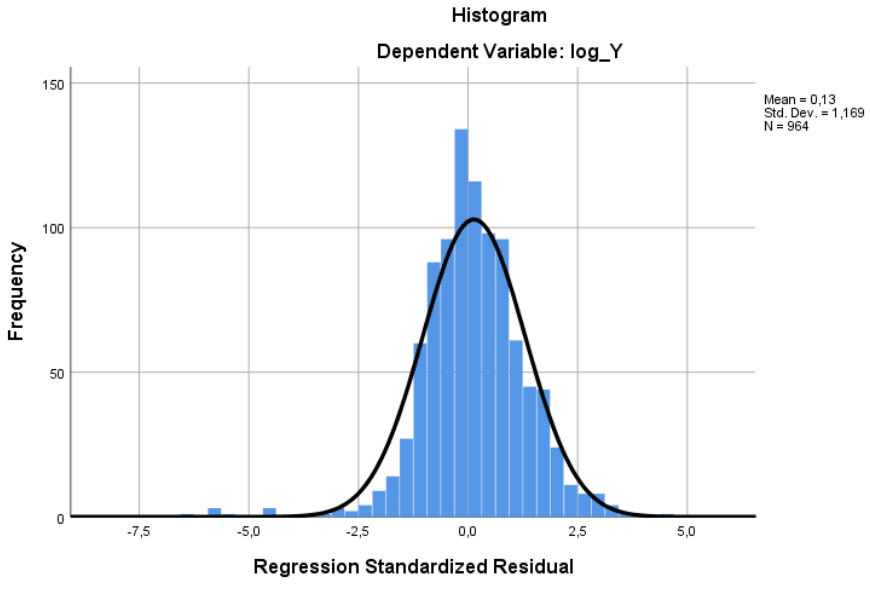


Κανονικότητα υπολοίπων

Από το ιστόγραμμα (Εικόνα 5.54) δεν προκύπτει η ύπαρξη κανονικής κατανομής καθώς παρατηρούμε εξάρσεις κεντρικά και στήλες εκτός περιοχής στο διάγραμμα. Το ίδιο προκύπτει και από το P-P Plot (Εικόνα 5.55) καθώς η επιθυμία μας τα σημεία να είναι κοντά στην ευθεία, ώστε να αποδειχτεί η κανονικότητα, δεν συμβαίνει στο διάγραμμά μας.

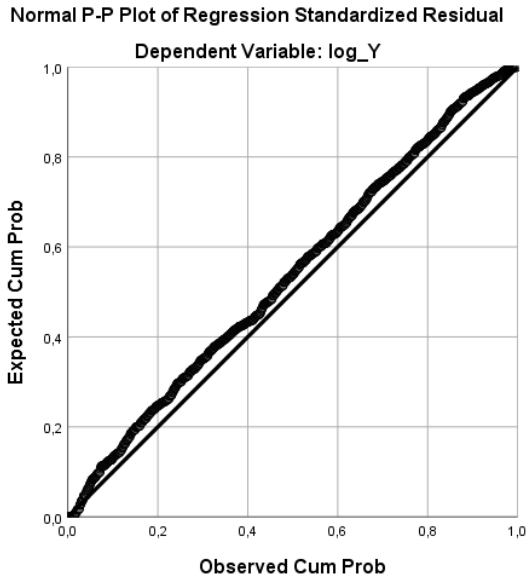
Εικόνα 5.54

Ιστόγραμμα υπολοίπων



Εικόνα 5.55

Διάγραμμα P-P_ Plot



Τέλος με το Kolmogorov-Smirnov (K-S) Test, η μηδενική υπόθεση ότι τα κατάλοιπα ακολουθούν κανονική κατανομή δεν φαίνεται να ικανοποιείται ως απαίτηση καθώς τα p-value=«0.05 (Πίνακας 5.56).

Πίνακας 5.56

Kolmogorov-Smirnov (K-S) Test

Tests of Normality

	Kolmogorov-Smirnov[a]			Shapiro-Wilk		
	Statistic	df	Sig.	Statistic	df	Sig.
Standardized Residual	,051	964	,000	,954	964	,000

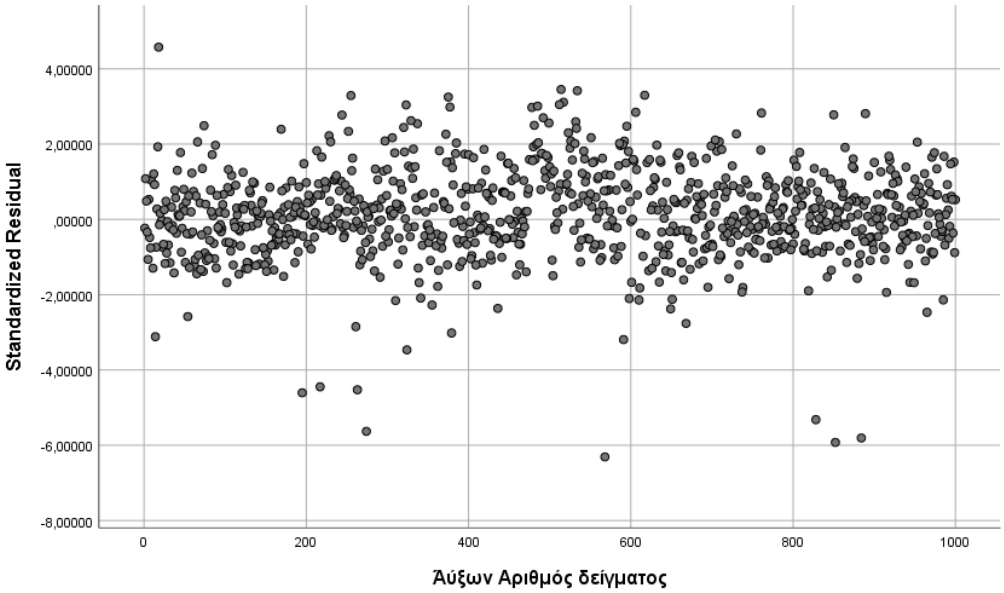
a. Lilliefors Significance Correction

Ανεξαρτησία υπολοίπων

Παρατηρούμε ότι οι τιμές στο γράφημα σκέδασης scatter plot (Εικόνα 5.56) παρουσιάζουν μερικώς φαινόμενο συσσωρεύσεων σημείων (clustering) άρα δεν μπορούμε να πούμε ότι τα κατάλοιπα είναι τυχαία.

Εικόνα 5.56

Scatter plot residuals / predicted



5.4.4.3 Εξαγωγή Συμπερασμάτων Αποτελεσμάτων Τέταρτου Μοντέλου μετά την επεξεργασία με το λογισμικό SPSS

Συμπερασματικά από την δημιουργία του τέταρτου μοντέλου μέσα από το λογισμικό SPSS διαπιστώνουμε ότι δεν πρόκειται για ένα ικανοποιητικό μοντέλο. Τα κατάλοιπα της παλινδρόμησης δεν ακολουθούν την κανονική κατανομή. Τα σφάλματα δεν φαίνεται να κατανέμονται με τυχαίο τρόπο, ενώ φαίνεται να έχουμε ετεροσκεδαστικότητα. Το δείγμα είναι χωρικά κατανεμημένο και αυτό θα πρέπει να ληφθεί υπ' όψιν. Στο επόμενο βήμα προχωράμε στην επεξεργασία μέσω του λογισμικού Arc Gis.

5.4.4.4 Εξέταση του Τέταρτου Μοντέλου Παλινδρόμησης σε Περιβάλλον Συστημάτων Γεωγραφικών Πληροφοριών

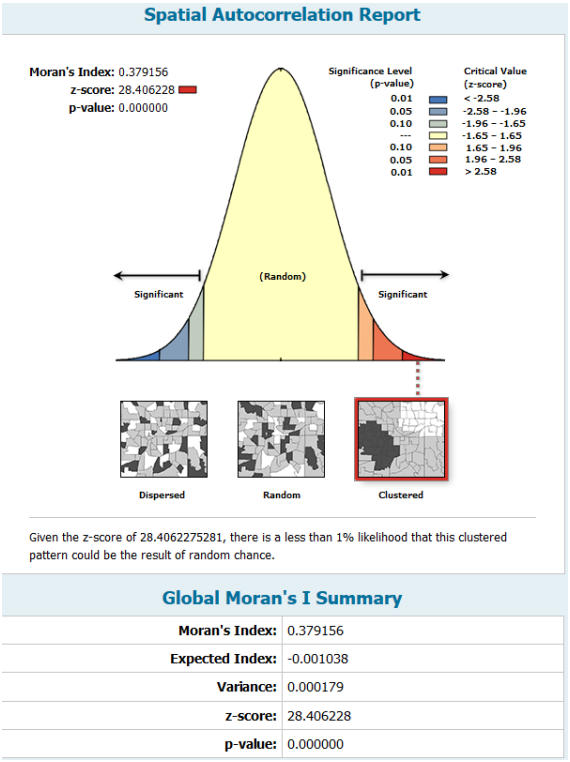
5.4.4.4.1 Χωρική Στατιστική Ανάλυση.

Σε πρώτη φάση εφαρμόζουμε τον συντελεστή Moran's I για την εξαρτημένη μεταβλητή «log_ΥΑξία «. Τα αποτελέσματα παρουσιάζονται στην εικόνα 5.57 Παρατηρούμε ότι η εξαρτημένη μεταβλητή παρουσιάζει τιμή στο δείκτη ίση με 0,3791 και z-score ίσο με 28,4062, ενώ το p=0,000. Παράλληλα παρουσιάζεται φαινόμενο συγκέντρωσης (cluster), η σημαντικότητα δε του φαινομένου αυτού ενδυναμώνεται και από την υψηλή τιμή του z score, όπου δηλώνεται ότι το

φαινόμενο cluster έχει πιθανότητα μικρότερη του 1% να οφείλεται σε τυχαίους παράγοντες.

Εικόνα 5.57

Εξέταση συντελεστή Moran's I για την εξαρτημένη μεταβλητή



5.4.4.4.2 Κλασσική παλινδρόμηση (Ordinary Least Squares Estimation-OLS).

Ακολουθώντας την διαδικασία όπως αυτή παρουσιάστηκε στο δεύτερο και τρίτο μοντέλο εισάγουμε μόνο εκείνες οι μεταβλητές που λειτούργησαν για την δημιουργία του τέταρτου μοντέλου στο SPSS για να επιτευχθεί η σύγκριση των αποτελεσμάτων ώστε να φανούν οι ομοιότητες ή οι διαφορές στις μαθηματικές εξισώσεις των δύο λογισμικών πακέτων.

Τα αποτελέσματα της διαδικασίας παρουσιάζονται στον πίνακα 5.57 ενώ στον πίνακα 5.58 παρουσιάζονται τα διαγνωστικά της παλινδρόμησης. Παρατηρούμε ότι το adjusted R^2 =0,74864 το F=205,877 και το AIC είναι ίσο με -336,5031 .

Τα αποτελέσματα ταυτίζονται, με μικρές διαφοροποιήσεις, με το μοντέλο όπως αυτό είχε δημιουργεί με το SPSS και παρουσιάζονται αναλυτικά στο παράρτημα «Δ2».

Πίνακας 5.57

Αποτελέσματα OLS διαδικασίας

Summary of OLS Results - Model Variables

Variable	Coefficient [a]	StdError	t-Statistic	Probability [b]	Robust_SE	Robust_t	Robust_Pr [b]	VIF [c]
Intercept	3.667550	0.155474	23.589457	0.000000*	0.141057	26.000559	0.000000*	--------
X8_POSOSTO	0.007171	0.001146	6.256658	0.000000*	0.001067	6.721362	0.000000*	2.741436
X9_POSOSTO	-0.009539	0.002163	-4.409493	0.000015*	0.002192	-4.351656	0.000018*	1.743562
X31_TIMH_Z	0.000148	0.000013	11.796087	0.000000*	0.000011	13.252189	0.000000*	1.424918
X38_OROFOS	0.024918	0.003497	7.125164	0.000000*	0.003754	6.638528	0.000000*	1.112405
X40_HLIKIA	-0.001166	0.000393	-2.963031	0.003131*	0.000418	-2.788819	0.005398*	1.795021
X41_EIDOS_	-0.099587	0.019044	-5.229207	0.000000*	0.021398	-4.653914	0.000005*	1.234044
X42_KATAST	-0.047196	0.006204	-7.607731	0.000000*	0.006018	-7.842015	0.000000*	1.477773
X47_YPARXH	0.059569	0.020746	2.871345	0.004184*	0.019197	3.102994	0.001985*	1.394762
X48_EIDOS_	-0.056685	0.011892	-4.766757	0.000003*	0.012166	-4.659342	0.000005*	1.500127
LOG_16	-0.129777	0.022737	-5.707753	0.000000*	0.022069	-5.880408	0.000000*	1.337248
LOG_22	-0.025222	0.024728	-1.019959	0.307997	0.021605	-1.167417	0.243334	1.162891
LOG_26	0.011338	0.019511	0.581083	0.561328	0.017900	0.633375	0.526642	1.964262
LOG_32	1.059469	0.034438	30.764464	0.000000*	0.033205	31.907265	0.000000*	1.322290
SQRT_17	0.001967	0.000950	2.070242	0.038688*	0.000824	2.387648	0.017137*	1.748756

Πίνακας 5.58

Διαγνωστικά παλινδρόμησης OLS διαδικασίας

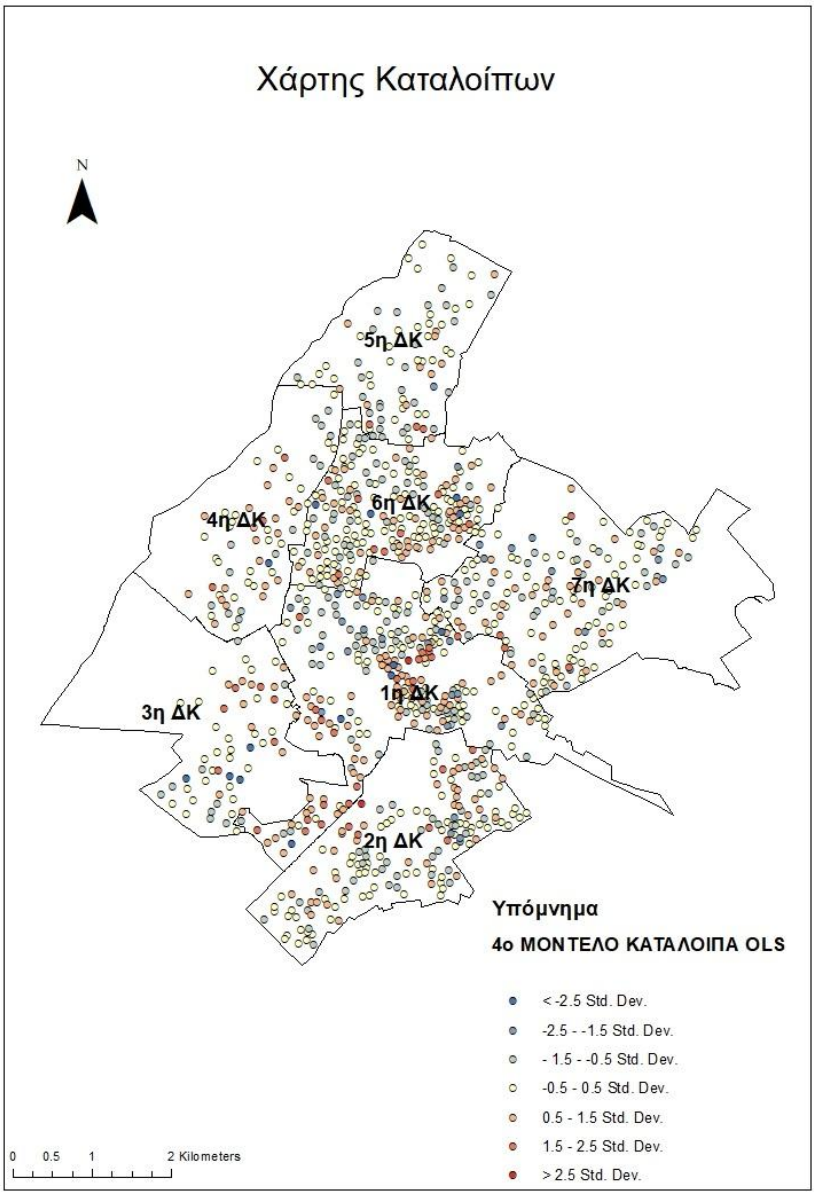
OLS Diagnostics

Input Features:	4_OLS_MONTELO_TELIKO	Dependent Variable:	LOG_Y
Number of Observations:	964	Akaike's Information Criterion (AICc) [d]:	-336.503167
Multiple R-Squared [d]:	0.752302	Adjusted R-Squared [d]:	0.748648
Joint F-Statistic [e]:	205.877565	Prob(>F), (14,949) degrees of freedom:	0.000000*
Joint Wald Statistic [e]:	3349.764698	Prob(>chi-squared), (14) degrees of freedom:	0.000000*
Koenker (BP) Statistic [f]:	19.916748	Prob(>chi-squared), (14) degrees of freedom:	0.132788
Jarque-Bera Statistic [g]:	678.859657	Prob(>chi-squared), (2) degrees of freedom:	0.000000*

Στον χάρτη 5.11 παρουσιάζεται η χαρτογράφηση των καταλοίπων ταξινομημένα σε επτά κλάσεις, ενώ στην εικόνα 5.58 παρουσιάζεται το ιστόγραμμα αυτών και στην εικόνα 5.59 παρουσιάζεται το διάγραμμα διασποράς τους. Υποδεικνύεται η μη κανονική κατανομή τους, η ετεροσκεδαστικότητα και η μη ανεξαρτησία.

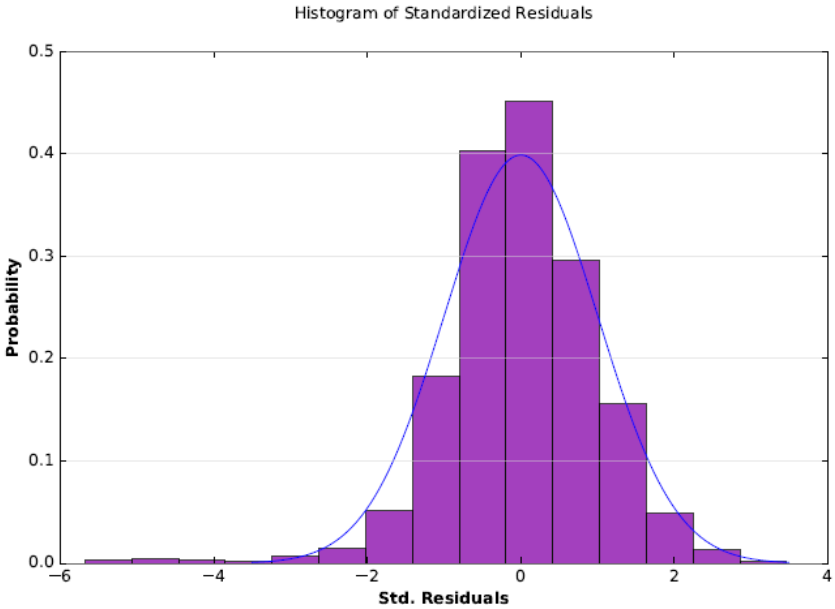
Χάρτης 5.11

Χαρτογράφηση Καταλοίπων



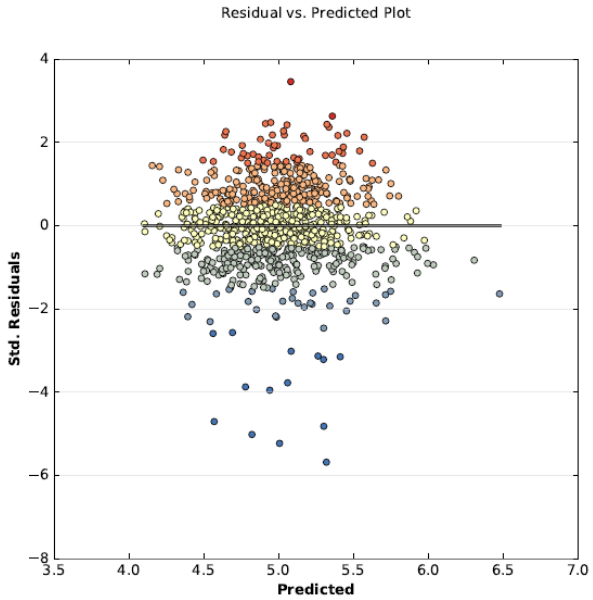
Εικόνα 5.58

Ιστόγραμμα καταλοίπων



Εικόνα 5.59

Διάγραμμα διασποράς καταλοίπων

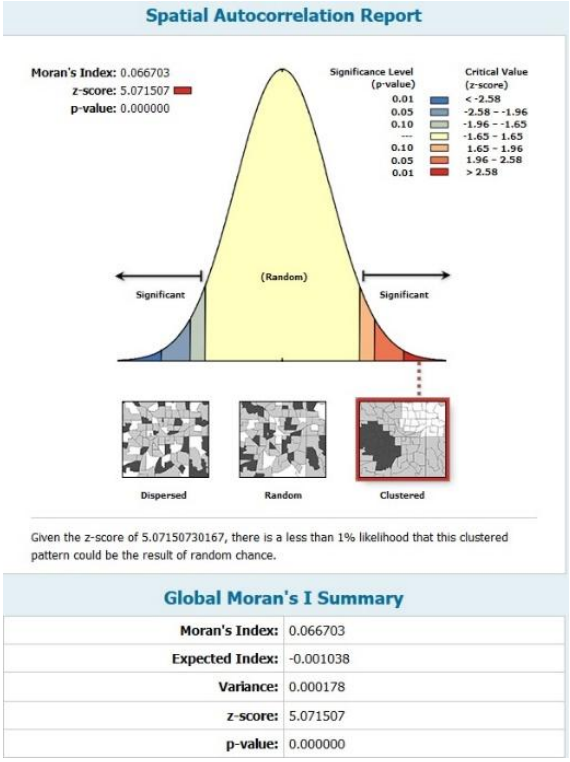


Για τον έλεγχο τη χωρικής αυτοσυσχέτισης των υπολοίπων χρησιμοποιούμε τον δείκτη Morans I (Εικόνα 5.60). Παρατηρούμε ότι τα κατάλοιπα της παλινδρόμησης παρουσιάζουν τιμή στο δείκτη ίση με 0,066 και z-score ίσο με 5,071, ενώ το p=0,0000. Παράλληλα παρουσιάζεται φαινόμενο συγκέντρωσης (clustered). Η σημαντικότητα του φαινομένου υποδεικνύεται από την τιμή του z score, όπου

δηλώνεται ότι το φαινόμενο cluster έχει πιθανότητα μικρότερη του 1% να οφείλεται σε τυχαίους παράγοντες.

Εικόνα 5.60

Έλεγχος χωρικής αυτοσυσχέτισης καταλοίπων



Global Moran's I Summary

Moran's Index:	0.066703
Expected Index:	-0.001038
Variance:	0.000178
z-score:	5.071507
p-value:	0.000000

Ελέγχοντας τα κατάλοιπα της OLS διαδικασίας, με τον υπολογισμό του Moran I, διαπιστώνουμε την ύπαρξη χωρικής αυτοσυσχέτισης, δηλαδή cluster χωρικό πρότυπο και γι' αυτό τον λόγο κρίνεται λοιπόν σκόπιμο να προχωρήσουμε στη γεωγραφικά σταθμισμένη παλινδρόμηση.

5.4.4.4.3 Χωρική Παλινδρόμηση (Geographically Weighted Regression-GWR) σε περιβάλλον GIS.

Η εφαρμογή της GWR στο μοντέλο μας γίνεται και εδώ μέσω του open source λογισμικού GWR4. Τα αποτελέσματα φαίνονται στον πίνακα 5.59 και αναλυτικά παρουσιάζονται στο παράρτημα «Δ3» , ενώ αναλυτικά αποτελέσματα της διαδικασίας για κάθε ακίνητο παρουσιάζονται στο παράτημα «Δ4». Παρατηρούμε ότι μέσω της διαδικασίας αυτής το Adjusted R square που στην OLS ήταν 0,7486, τώρα είναι 0,7712, ενώ παράλληλα το AICc από -336,50 που ήταν στην OLS διαδικασία,

υπολογίστηκε σε -395.266. Συνολικά και σε γενικές γραμμές το μοντέλο δείχνει να έχει βελτιωθεί αισθητά.

Πίνακας 5.59

Αποτελέσματα GWR διαδικασίας

GWR (Geographically weighted regression) result

Bandwidth and geographic ranges

Bandwidth size:	510.364541		
Coordinate	Min	Max	Range
X-coord	473725.000000	480487.000000	6762.000000
Y-coord	4200150.000000	4208998.000000	8848.000000

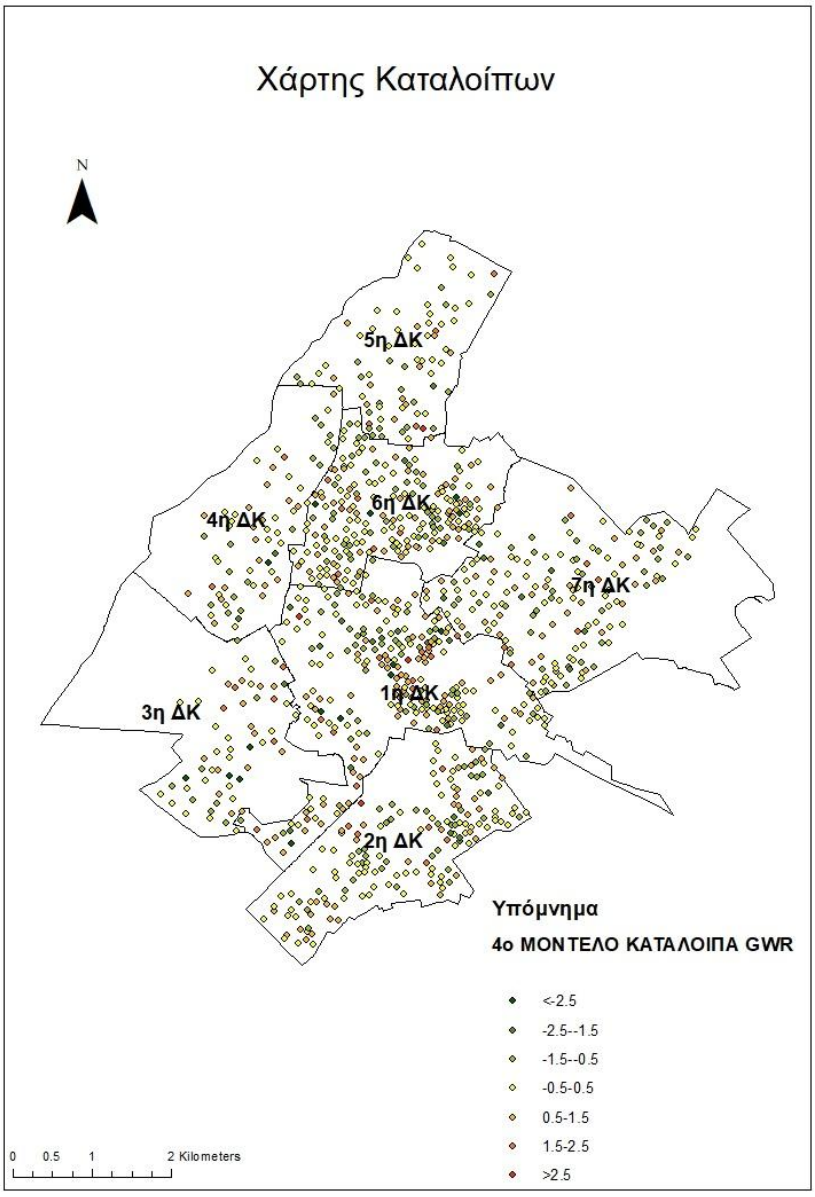
Diagnostic information

Residual sum of squares:	32.673154
Effective number of parameters (model: trace(S)):	60.581866
Effective number of parameters (variance: trace(S'S)):	42.568634
Degree of freedom (model: n - trace(S)):	903.418134
Degree of freedom (residual: n - 2trace(S) + trace(S'S)):	885.404901
ML based sigma estimate:	0.184101
Unbiased sigma estimate:	0.192099
-2 log-likelihood:	-526.980683
Classic AIC:	-403.816951
AICc:	-395.266184
BIC/MDL:	-103.846058
CV:	0.038437
R square:	0.789943
Adjusted R square:	0.771276

Στο χάρτη 5.12 φαίνονται τα τυποποιημένα κατάλοιπα (standardized) της διαδικασίας GWR. Από τον χάρτη αυτό φαίνεται και σε σχέση με τον προηγούμενο χάρτη μείωση του μεγέθους των καταλοίπων. Το ίδιο προκύπτει και από το ιστόγραμμα (Εικόνα 5.61).

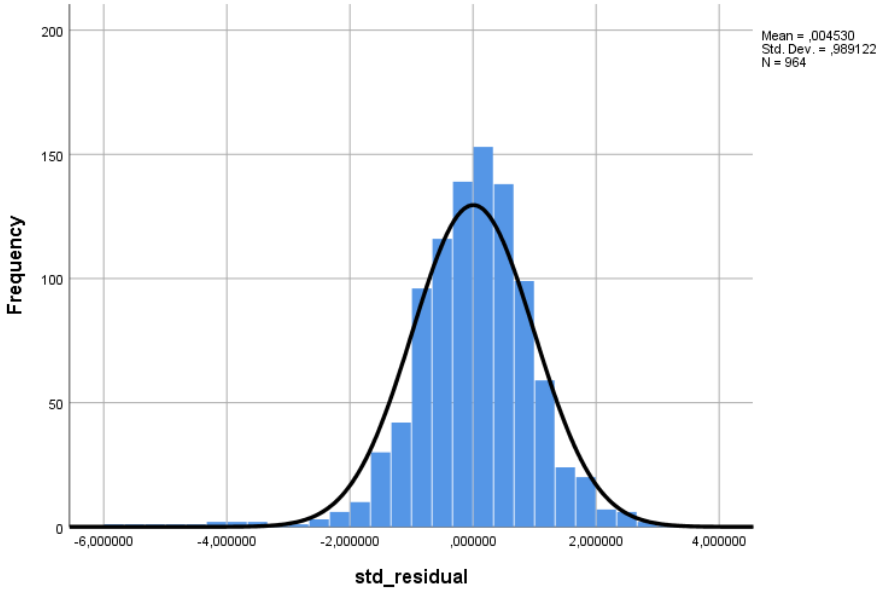
Χάρτης 5.12

Χάρτης καταλοίπων GWR διαδικασίας



Εικόνα 5.61

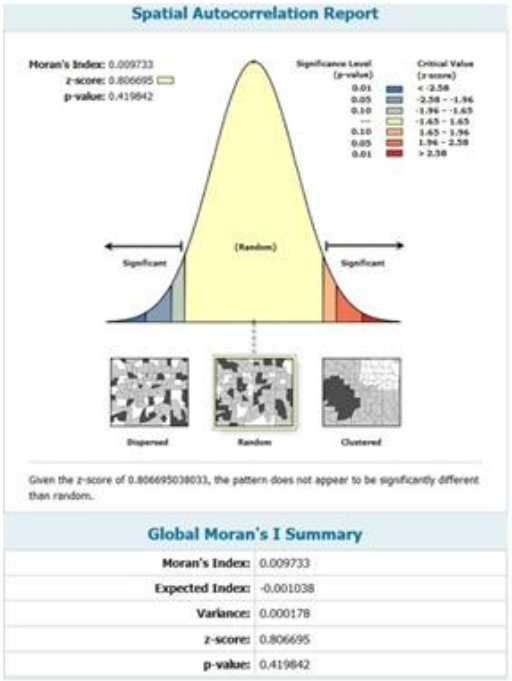
Ιστόγραμμα καταλοίπων



Υπολογίζουμε τον συντελεστή αυτοσυσχέτισης Moran I (Εικόνα 5.62) για τα κατάλοιπα της GWR και βλέπουμε ότι προκύπτει τυχαίο πρότυπο, δηλαδή τα κατάλοιπα δεν παρουσιάζουν πλέον χωρική αυτοσυσχέτιση και οφείλονται σε τυχαίους παράγοντες.

Εικόνα 5.62

Υπολογισμός Morans I για τα κατάλοιπα



5.4.4.4.4 Εξαγωγή Συμπερασμάτων των Αποτελεσμάτων του Τέταρτου Μοντέλου μέσα από την Χωρική παλινδρόμηση (Geographically Weighted Regression-GWR).

Συμπερασματικά από την επεξεργασία του τέταρτου μοντέλου μέσα από την χρήση της χωρικής παλινδρόμησης διαπιστώνουμε ότι αυτό βελτιώθηκε αισθητά (βλέπε πίνακα 5.60). Ο συντελεστής συσχέτισης βελτιώθηκε αισθητά. Παρατηρείται μείωση του μεγέθους των καταλοίπων ενώ αυτά δεν παρουσιάζουν πλέον χωρική αυτοσυσχέτιση και οφείλονται σε τυχαίους παράγοντες.

Πίνακας 5.60

Σύγκριση μοντέλων OLS & GWR διαδικασίας τέταρτου μοντέλου μέσα από την χωρική παλινδρόμηση .

Συντελεστές	OLS	GWR
Adjusted R Square	0,748	0,771
Morans I residuals	0,066	0,010
AICc	-336.50	-395.266

Είναι φανερό πλέον ότι από το μοντέλο της γεωγραφικά σταθμισμένης παλινδρόμησης προκύπτει καλύτερη ερμηνεία της εξαρτημένης μεταβλητής και ακριβέστερος προσδιορισμός του μοντέλου της παλινδρόμησης.

5.4.5 Σύγκριση Μοντέλων

Στοn πίνακα 5.61 παρουσιάζονται συγκεντρωμένα τα αποτελέσματα της διαδικασίας και για τα τέσσερα μοντέλα .

Το τέταρτο και δεύτερο μοντέλο είναι αυτά που παρήγαγαν τα καλύτερα αποτελέσματα. Παρατηρώντας και τα τέσσερα μοντέλα διαπιστώνουμε ότι οι μεταβλητές «Εμβαδόν», «Κατάσταση κατασκευής», «Όροφος», «Είδος θέρμανσης» , «Τιμή ζώνης» συμμετέχουν σε όλα, αποδεικνύοντας την σημαντικότητά τους.

Παρατηρώντας τα δύο καλύτερα μοντέλα διαπιστώνουμε ότι έχουν κοινές τις μεταβλητές: «Εμβαδόν κατοικίας», «Κατάσταση κατασκευής», «Τιμή ζώνης», «Ύπαρξη θέσης στάθμευσης», «Είδος κατοικίας», «Όροφος», «Είδος θέρμανσης», «Απόσταση από νεκροταφείο», «Ηλικία κατασκευής», «Ποσοστό πρασίνου», «Απόσταση από νοσοκομείο».

Η λογαρίθμηση των ανεξάρτητων μεταβλητών: «Εμβαδόν», «Απόσταση από Νεκροταφείο»«, « Απόσταση από ΟΣΕ» και «Απόσταση από αθλητικές

εγκαταστάσεις» στο τέταρτο μοντέλο, διαφοροποιούν τον τρόπο επίδρασης τους στην εξαρτημένη μεταβλητή, η οποία είναι και αυτή λογαριθμισμένη, καθώς η ερμηνεία των παραμέτρων της εξίσωσης της παλινδρόμησης αφορά πλέον ελαστικότητες αντικατοπτρίζοντας την ελαστικότητα της κάθε ανεξάρτητης μεταβλητής ως προς την εξαρτημένη, υπολογίζοντας την ποσοστιαία μεταβολή της για κάθε 1% μεταβολή της ανεξάρτητης.

Πίνακας 5.61 Σύγκριση μοντέλων

	Πρώτο μοντέλο					Δεύτερο μοντέλο				Τρίτο μοντέλο				Τέταρτο μοντέλο			
	OLS Διαδικασία		GWR Διαδικασία			OLS Διαδικασία		GWR Διαδικασία		OLS Διαδικασία		GWR Διαδικασία		OLS Διαδικασία		GWR Διαδικασία	
	μέσω SPSS	μέσω ArcGis	μέσω ArcGis	μέσω GWR4	ΣΥΜΒΟΛΗ ΣΤΟ ΜΟΝΤΕΛΟ	μέσω SPSS	μέσω ArcGis	μέσω GWR4	ΣΥΜΒΟΛΗ ΣΤΟ ΜΟΝΤΕΛΟ	μέσω SPSS	μέσω ArcGis	μέσω GWR4	ΣΥΜΒΟΛΗ ΣΤΟ ΜΟΝΤΕΛΟ	μέσω SPSS	μέσω ArcGis	μέσω GWR4	ΣΥΜΒΟΛΗ ΣΤΟ ΜΟΝΤΕΛΟ
Εξαρτημένη μεταβλητή	Αξία Υ					Αξία Υ				log_Αξία Υ				log_Αξία Υ			
R	0,630					0,810				0,860				0.8630			
R Square	0,397	0,3969	0,4563	0.6006		0,656	0,6562	0,72184		0,739	0.7394	0.7723		0.7440	0.7523	0.7899	
Adjusted R Square	0,393	0.3926	0.4416	0.5339		0,650	0.6504	0,6891		0.7360	0.7361	0.7574		0.7380	0.7486	0.7712	
Std. Error of the Estimate	235022.2970		225362.3180	205769.2700		44981.8200		42394.4630		0.2064		0.1978		0.1751		0.1920	
F	93.2720	93.2710				114.5320	114.5310			224.8630	224.8630			123.4140	205.8770		
AICc		27582.8980	27503.9310	27398.2200			20764.8600	20703.2800			-291.7078	-352.6660			-336.5030	-395.2660	
Συμμετέχουσες ανεξάρτητες μεταβλητές	X23 -Απόσταση από πράσινο				-	X32 -Εμβαδόν κατοικίας			+	log_32_Εμβαδόν			+	log_32-Εμβαδόν			+
	X31- Τιμή ζώνης				+	X42- Κατάσταση κατασκευής			-	X42- Κατάσταση κατασκευής			-	X40- Ηλικία κατασκευής			-
	X32 -Εμβαδόν κατοικίας				+	X31- Τιμή ζώνης			+	X31- Τιμή ζώνης			+	X8- Ποσοστό πρασίνου			+
	X38- Όροφος				+	X47- Ύπαρξη θέσης στάθμευσης			+	X9- Ποσοστό κατοίκων ανα ΔΚ			-	X48- Είδος θέρμανσης			-
	X40- Ηλικία κατασκευής				+	X41- Είδος κατοικίας			-	X38- Όροφος			+	log_16_Απόσταση από νεκροταφείο			-
	X42- Κατάσταση κατασκευής				-	X38- Όροφος			+	X48_ Είδος θέρμανσης			-	X47- Ύπαρξη θέσης σταθμευσης			+
	X48- Είδος θέρμανσης				-	X48- Είδος θέρμανσης			-	X41_ Είδος κατοικίας			-	X38- Όροφος			+
						X51- Ύπαρξη θέας			+	log_26_Απόσταση από ΟΣΕ			+	X42- Κατάσταση κατασκευής			-
						X35- Αριθμός μπάνιων			+	X51- Ύπαρξη θέας			+	X41- Είδος κατοικίας			-
						X16- Απόσταση από νεκροταφείο			-	X37- Ύπαρξη αποθήκης			+	X9- Ποσοστό κατοίκων ανα ΔΚ			-
						X23- Απόσταση από πράσινο			-	X10- τμ πρασίνου ανα κάτοικο			+	log_26-Απόσταση από ΟΣΕ			+
						X40- Ηλικία κατασκευής			-	X7- Ποσοστό κενών κατοικιών			-	X31- Τιμή ζώνης			+
						X8- Ποσοστό πρασίνου			+					sqrt_X17- Απόσταση από νοσοκομείο			+
						X17- Απόσταση από νοσοκομείο			+					log_22- Απόσταση από αθλητισμό			-
Ομοσκεδαστικότητα υπολοίπων	ΟΧΙ		ΟΧΙ			ΟΧΙ		ΝΑΙ		ΟΧΙ		ΟΧΙ		ΟΧΙ		ΝΑΙ	
Κανονικότητα υπολοίπων	ΟΧΙ		ΟΧΙ			ΟΧΙ		ΝΑΙ		ΟΧΙ		ΟΧΙ		ΟΧΙ		ΝΑΙ	
Ανεξαρτησία υπολοίπων	ΟΧΙ		ΟΧΙ			ΟΧΙ		ΝΑΙ		ΟΧΙ		ΟΧΙ		ΟΧΙ		ΝΑΙ	
Moran's I Υπολοίπων		0.0853	0.0522				0.0363	-0.0187			0.0850	0.0478			0.0660	0.0100	
z- score Υπολοίπων		6.9975	4.3440				2.3093	-1.0813			6.4680	3.6614			5.0170	0.8066	
Παρατηρούμενο φαινόμενο		cluster	cluster				cluster	Random			cluster	cluster			cluster	Random	

Κεφάλαιο 6^{ο} - Συμπεράσματα

Σε αυτό το κεφάλαιο παρουσιάζονται όλα τα συμπεράσματα που εξάχθηκαν από τη μεταπτυχιακή διπλωματική εργασία. Εστιάζοντας στα αποτελέσματα που προκύπτουν από την εφαρμογή της στατιστικής ανάλυσης μέσω της διαδικασίας των ελαχίστων τετραγώνων και της γεωγραφικά σταθμισμένης παλινδρόμησης εντοπίζονται τα πλεονεκτήματα και μειονεκτήματα της μεθοδολογίας.

Ένα βασικό ερώτημα που καλείται να απαντηθεί, είναι εάν δύναται να προσεγγιστεί η αγοραία αξία ενός ακινήτου με την χρήση γραμμικού μοντέλου παλινδρόμησης. Η κλασική μέθοδος της γραμμικής παλινδρόμησης έχει την δυνατότητα να αποδώσει πολύπλοκες σχέσεις με απλό τρόπο και αυτό γίνεται ακόμα πιο εύκολο με την χρήση των κατάλληλων λογισμικών. Στην παρούσα περίπτωση η γραμμική παλινδρόμηση κλήθηκε να αναλύσει χωρικά φαινόμενα και αυτό έγινε εφικτό με την χρήση της γεωγραφικά σταθμισμένης παλινδρόμησης.

Το εξεταζόμενο μοντέλο μέσα από την πολλαπλή παλινδρόμηση έχει μεταβλητές που μένουν σταθερές στον χώρο, ενώ όταν αυτό εξετάζεται με την γεωγραφικά σταθμισμένη παλινδρόμηση ενσωματώνεται η γεωγραφική θέση της κάθε μιας παρατήρησης στην ερμηνεία των αποτελεσμάτων. Έχουμε λοιπόν πρακτικά μια ανάλυση σε τοπική κλίμακα και το μοντέλο εφαρμόζεται πλέον σε επίπεδο ακινήτου.

Η συγκέντρωση ενός αξιόπιστου δείγματος με την βοήθεια χρήσιμων πρωτογενών δεδομένων και με την βοήθεια της στατιστικής ανάλυσης έχουν ως στόχο την απόδοση εκτιμώμενων τιμών αξίας ακινήτων που ανταποκρίνονται στην πραγματικότητα. Σ' ένα ευρύτερο πλαίσιο το βασικό λοιπόν ερώτημα μετακυλίεται στο εάν μέσα από ένα μαθηματικό μοντέλο, μέσω μιας αυτοματοποιημένης διαδικασίας, μπορεί να λυθεί το πρόβλημα της μαζικής εκτίμησης τιμών ακινήτων. Το πρόβλημα είναι πολύπλοκο καθώς το φαινόμενο έχει γεωγραφικό και οικονομικό χαρακτήρα και για την λύση του απαιτείται προσεκτική επιλογή της πρωτογενούς πληροφορίας.

Στην παρούσα εργασία συλλέχθηκε δείγμα 1000 κατοικιών από γνωστή ιστοσελίδα του διαδικτύου. Η πληροφορία που μεταφέρθηκε προς επεξεργασία αφορούσε μεταβλητές κατασκευαστικών χαρακτηριστικών, θέσης, γειτονιάς καθώς και λοιπών μεταβλητών. Δημιουργήθηκαν τέσσερα μοντέλα τα οποία

επεξεργάστηκαν με το στατιστικό πακέτο SPSS και με τις εφαρμογές που παρέχει το πακέτο ArcGis και το πακέτο GWR4.

Η λογική επεξεργασίας του πρώτου μοντέλου ήταν η χρήση όλου του δείγματος χωρίς να αφαιρεθούν οι ακραίες οντότητες, ενώ απομακρύνθηκαν από τα ζεύγη των μεταβλητών που εμφάνιζαν συσχέτιση μεταξύ τους, αυτές που εμφάνιζαν την μικρότερη συσχέτιση με την εξαρτημένη μεταβλητή. Η διαδικασία εισαγωγής στο στατιστικό πακέτο ήταν η μαζική εισαγωγή (enter) των μεταβλητών, ενώ η διαδικασία συνεχίστηκε με την αφαίρεση των μεταβλητών αυτών που δεν μετέβαλαν σημαντικά τον διορθωμένο συντελεστή προσδιορισμού R^2. Τα αποτελέσματα που εξήχθησαν μέσα από την διαδικασία των ελαχίστων τετραγώνων δεν βελτιώθηκαν με την επεξεργασία μέσα από την γεωγραφικά σταθμισμένη παλινδρόμηση κρατώντας τον Adjusted R^2 σε χαμηλά επίπεδα, ενώ από τον έλεγχο των υπολοίπων της παλινδρόμησης συνεχίστηκε να διαπιστώνεται φαινόμενο cluster.

Η επεξεργασία του δευτέρου μοντέλου έγινε μετά την αφαίρεση των ακραίων οντοτήτων, ενώ απομακρύνθηκαν οι μεταβλητές που επαναλάμβαναν την πληροφορία μιας άλλης μεταβλητής. Η διαδικασία εισαγωγής των μεταβλητών έγινε με την μέθοδο stepwise, ενώ το αποτέλεσμα βελτιώθηκε αισθητά με την χρήση της γεωγραφικά σταθμισμένης παλινδρόμησης με τον δείκτη Adjusted R^2 να φτάνει σε επίπεδα 0,6891 χωρίς να παρουσιάζεται πλέον φαινόμενο cluster.

Η επεξεργασία του τρίτου μοντέλου έγινε αφού εφαρμόστηκαν οι κατάλληλοι μετασχηματισμοί στις μεταβλητές και αφού αφαιρέθηκαν οι ακραίες οντότητες κάνοντας χρήση των δεικτών DfFITS και Cook's D. Βασικό σημείο στην διαδικασία του τρίτου μοντέλου ήταν η ομαδοποίηση των μεταβλητών σε τέσσερις κατηγορίες: των κατασκευαστικών χαρακτηριστικών, της θέσης, της γειτονιάς και των λοιπών μεταβλητών. Ένα ακόμα βασικό σημείο ήταν η λογαρίθμηση της εξαρτημένης μεταβλητής. Το τελικό μοντέλο που παρήχθη ήταν της μορφής log- log, ενώ τα αποτελέσματα μέσω της OLS διαδικασίας δεν βελτιώθηκαν μέσω της GWR διαδικασίας, με υψηλό μεν διορθωμένο συντελεστή προσδιορισμού R^2, αλλά με cluster φαινόμενο.

Τέλος στο τέταρτο μοντέλο χρησιμοποιήθηκε το μέγεθος δείγματος του τρίτου, όμως η εισαγωγή των μεταβλητών έγινε μαζικά και με την μέθοδο stepwise. Το μοντέλο ήταν της μορφής log- log. Τα αποτελέσματα ήταν αρκετά ικανοποιητικά με το Adjusted R^2 να φτάνει στα επίπεδα των 0,7899 και με τυχαία κατάλοιπα.

Τα μοντέλα που έδωσαν τα καλύτερα αποτελέσματα ήταν το τέταρτο και δεύτερο, ενώ το βέλτιστο adjusted R^2 και στην OLS αλλά και στην GWR διαδικασία ήταν του τέταρτου μοντέλου. Σημαντικό είναι να αναφερεθεί ότι και στα τέσσερα μοντέλα διαπιστώνουμε ότι οι μεταβλητές «Εμβαδόν», «Κατάσταση κατασκευής», «Όροφος», «Είδος θέρμανσης», «Τιμή ζώνης» συμμετέχουν σε όλα, γεγονός που αποδεικνύει την σημαντικότητά τους.

Από την μελέτη προέκυψε ότι η χρήση της γεωγραφικά σταθμισμένης παλινδρόμησης έδωσε λύση στο πρόβλημα της χωρικής αυτοσυσχέτισης των ακινήτων. Το χωρικό φαινόμενο αναλύθηκε και αυτό έγινε εφικτό με την ενσωμάτωση της γεωγραφικής θέσης της κάθε μιας μεταβλητής στην ερμηνεία των αποτελεσμάτων.

Η εφαρμογή της ανάλυσης σε μια τοπική κλίμακα σε επίπεδο ακίνητου φαίνεται ότι αποτελεί την κατάλληλη μέθοδο εκτίμησης σε σχέση με την παραδοσιακή μέθοδο της παλινδρόμησης, αποδεικνύοντας ότι όταν δεν λαμβάνεται υπόψη η χωρική διάσταση των τιμών των ακινήτων τότε τα αποτελέσματα παρουσιάζονται λανθασμένα.

Τα μοντέλα που προκρίθηκαν θα μπορούσαν να αποτελέσουν βάση και σε μαζικές εκτιμήσεις ακινήτων.

ΒΙΒΛΙΟΓΡΑΦΙΑ

Αβούρης, Ν., Κατσάνος, Χ., Τσέλιος Ν., Μουστάκας , Κ. (2015). *Εισαγωγή στην αλληλεπίδραση ανθρώπου – υπολογιστή*. Αθήνα: Ελληνικά ακαδημαϊκά συγγράμματα και βοηθήματα.

Anselin, L. and Gracia, N.L. (2008). *Errors in Variables and Spatial Effects in Hedonic House Price Models of Ambient Air Quality*. Empirical Economics, 34: 5–34.

Αποστολίδης, Μ., Μανομενίδης, Ν. (2014). *Εκτιμήσεις ακινήτων & νέα νομοθεσία για τους μεσίτες ακινήτων*. Αθήνα: ΙΜΕ ΓΣΕΒΕΕ.

Αρβανίτης, Α. (2009). *Διαχείριση γης και αξιοποίηση ακινήτων*. Ημερίδα Σ.Α.Τ.Μ.Β.Ε-ΤΕΕ- ΤΑΤΜ, ΑΠΘ, Θεσσαλονίκη.

Αρβανίτης, Α., Λαφαζάνη, Π., Μισίρογλου, Σ. Επιμέλεια: Παππάς, Β.,Κοτζινός, Δ., Σταθάκης, Δ. (2010). Μοντέλο διαχείρισης δημοτικού κτηματολογίου σε περιβάλλον Γεωγραφικού Συστήματος Πληροφοριών με την βοήθεια της στατιστικής ανάλυσης. Κείμενα πολεοδομίας, χωροταξίας και ανάπτυξης. Επιστημονικό περιοδικό: Αειχώρος . Τεύχος 14. Πανεπιστημιακές εκδόσεις Θεσσαλίας.

Benjamin,J., Guttery, R., Sirmans, C.(2004). *Mass appraisal: an introduction to multiple regression analysis for real state valuation*. Journal of real estate practice and education. Vol. 7, No. 1 , pp. 65-78. American Real Estate society.

Benoit, K. (2011). *Linear regression models with logarithmic transformation*. London school of economics.

Box, G., Cox, D. (1964). *An analysis of transformations*. *Journal of the Royal Statistical Society*. Series B (Methodological) Vol. 26, No. 2 (1964), pp. 211-252

Breiman, L., Friedman, J., Olshen, R., Stone, C (1984). *Classification and regression trees*. Chapman & Hall.

Brukstaitiene, D. (2000). *Application of mass appraisal models*, Vilnius Gediminas University, Lithuania.

Brunsdon, C., Fotheringham , A.S. & Charlton, M. (1998). *Geographically weighted regression – modelling spatial nonstationarity*, The Statistician, 47

(3) 431- 443.

Γανωτάκης, Γ. (2014). *Αποτίμηση ακινήτων λογιστικός χειρισμός αυτών στα πλαίσια εύλογης αξίας και ανώνυμες εταιρίες επενδύσεων σε ακίνητη περιουσία*. Διπλωματική εργασία, Επιβλέπων καθηγητής: Παπαναστασίου Ι. Τμήμα Λογιστικής και Χρηματοοικονομικής, Πανεπιστήμιο Μακεδονίας, Θεσσαλονίκη.

Can, A. (1990). *The Measurement of Neighborhood Dynamics in Urban House Prices. Economic Geography*, 66, 254–72.

Can, A. (1992). *Specification and estimation of hedonic housing price models*, Regional science and urban economics 22, 453-474. North Holland

Chrostek, K., Kopczewska, K.(2013). *Spatial Prediction Models for Real Estate Market Analysis . Εψονομιανρ35/2013*

Chin, T., Chau, K. (2002). *A critical review of literature on the hedonic price model.* International Journal for Housing Science and Its Applications 27(2):145-165

Cho, S., Poudyalb, N., Roland, K. (2007). *Spatial analysis of the amenity value of green open space.* Science direct. Ecological economics 66(2008) 403-406

Chatterjee, S. Simonoff, J. (2013). *Handbook of regression analysis. Wiley Publication*. New Jersey.

Γκαγκάτσης, Φ. (2016). *Τα ευδαιμονικά μοντέλα και οι εφαρμογές τους στις αξίες των ακινήτων.* Ερευνητική εργασία. Επιβλέπων: Καρανικόλας.ΑΠΘ. Θεσσαλονίκη.

Γκαρούτσος, Γ. (2008). *Μαθήματα στατιστικής.* Εκδόσεις: SPIN. Αθήνα.

Clauretie, T.M. and Neill, H.R. (2000). *Year-round school schedules and residential property values*. Journal of Real Estate Finance and Economics, 20(3): 311-322.

Δαμιανού, Χ., Παπαδάτος, Ν., Χαραλαμπίδης, Χ. (2010) . *Εισαγωγή στις πιθανότητες και τη στατιστική. Εκδόσεις: Συμμετρία. Αθήνα.*

Δαμιανού, Χ., Παπαδάτος, Ν., Χαραλαμπίδης, Χ. (2003) . *Εισαγωγή στις πιθανότητες και τη στατιστική. Εκδόσεις: Συμμετρία. Αθήνα.*

Daniel , WW. (1999). *Biostatistics: A Foundation for Analysis in the Health Sciences. 7th edition.* New York: John Wiley & Sons.

Δαφέρμος Β. (2005), *Κοινωνική Στατιστική με το SPSS*, Εκδόσεις Ζήτη, Θεσσαλονίκη.

Δαφέρμος, Β. (2002), Επαναληπτικές στατιστικές μετρήσεις στις κοινωνικές επιστήμες.Εκδόσεις Leader books. Αθήνα.

Dubin, R. (1988). *Estimation coefficients in the presence of spatially autocorrelated error terms.* The Review of Economics and Statistics.Vol. 70, No.3 pp 466-474. The MIT Press

Dubin, R.A.(1988). *Spatial Autocorrelation: A Primer*. Journal of Housing Economics, 7, 304–27.

Dubin, R.A., J.K. Pace, and T.G. (1999) Thibodeau. *Spatial Autoregression Techniques for Real Estate Data*. Journal of Real Estate Literature, 7, 79–95.

Εμβαλωτής, Α., Κάτης, Α., Σιδερίδης, Γ. (2006). *Στατιστική μεθοδολογία εκπαιδευτικής έρευνας*. Ιωάννινα.

Fan, G.,Ong, S.,Koh, H. (2006) *Determinants of housing price: A decision tree approach. Urban studies* , Vol.43, No12,2301-2315.

Field, A. *Discovering Statistics Using IBM SPSS Statistics.* SAGE Publications (2013).

Ζαφειρόπουλος, Κ. (2015). *Πως γίνεται μια επιστημονική εργασία*. Εκδόσεις: Κριτική Α.Ε. Αθήνα.

Ζεντέλης, Π. (2015). *Real Estate. Αξία, εκτίμηση, ανάπτυξη, επενδύσεις*. Αθήνα: Ελληνικά ακαδημαϊκά συγγράμματα και βοηθήματα.

Zietz, J., Zietz, E. Simans, G. (2008). *Determinants of house process: A quantile regression approach.* Department of economics and finance working series.

Gatzlaff, D. & Smith, M., (1993).*Uncertainty, growth controls and the efficiency of development patterns*. Journal of real estate and economics, 6:147-155.

Galton, F. *(1889). Natural Inheritance*. London: Macmillan.

Galton, F. *(1886). Regression Towards Mediocrity in Hereditary Stature.* Anthropological Miscellanea. The Journal of the Anthropological Institute of Great Britain and Ireland .Vol. 15 (1886), pp. 246-263

Giudice, V.,Paola, P.,Cantisani, G. (2017). *Valuation of Real Estate Investments through fuzzy logic*. Naples.

Ηλιοπούλου, Π. (2015). *Γεωγραφική ανάλυση.* Αθήνα: Ελληνικά ακαδημαϊκά συγγράμματα και βοηθήματα.

Ηλιοπούλου, Π. (2017). *Γεωγραφική ανάλυση (Αναθεώρηση).* Αθήνα: Ελληνικά ακαδημαϊκά συγγράμματα και βοηθήματα.

Ηλιοπούλου , Π., Στρατάκης , Π.(2018). *Spatial analysis of housing prices in the Athens Region, Greece.* International journal of real estate and land planning. Vol.1

Harrell, F. (2001). *Regression modeling strategies.* Springer. New York

Hayrullahoglu, G., Aliefendioglu, Y., Tanrivermis, H., Hayrullahoglu, A., (2018) . *Estimation of the hedonic valuation model in housing markets: The case of Cukurambar region in Cankaya district of Ankara province.* Egoforum, Vol. 7 Issue1(14).

Irwin, E., Bockstael, N. (2001). *The problem of identifying land use spillovers Qn Measuring the effects of open space on residential property values*.

Johnston, J., Dinardo, J. (2005). *Οικονομετρικές μέθοδοι.* Αθήνα. Εκδόσεις : Κλειδάριθμος.

Jahanshiri, E., Buyong, T., Shariff, A. (2011). *A review of property mass valuation models.* Pertanika J. Sci. & Technol. 19 (S): 23 – 30. University of Malaysia.

Καλαματιανού, Α. (2000). Κοινωνική στατιστική . Μέθοδοι μονοδιάστατης ανάλυσης. Αθήνα. Το οικονομικό.

Καρανικόλας, Ν. (2010). *Η εκτίμηση των ακινήτων*. Αθήνα. Εκδόσεις : Δίσιγμα.

Κουγιουμτζής, Δ. (2011). *Ανάλυση δεδομένων*, Σημειώσεις μαθήματος. ΑΠΘ. Θεσσαλονίκη.

Koenker, R. & Bassett, G. Jr. (1978). *Regression Quantiles, Econometrica*, 46(1), pp.33-50.

Kolbe, J., Wustemann, H. (2015) *Estimating the value of urban green space: A hedonic pricing analysis of the housing market in Cologne, Germany*. SFB, 649, Discussion Paper 2015-002

Kong F., Yin H., Nakagoshi N. (2007). *Using Gis and landscape metrics in the hedonic price modeling of the amenity value of urban green space: a case*

study inJinan City, Chin. Landscape and Urban Planning, 79, 240-252

Κουγιουμτζής, Δ. (2011). *Συσχέτιση και παλινδρόμηση*. Σημειώσεις μαθήματος . ΑΠΘ. Θεσσαλονίκη .

Κούτρας, Μ. (2011). Ανάλυση παλινδρόμησης. Σημειώσεις μαθημάτων. Πανεπιστήμιο Πειραιά .

Κύρκος, Ε.(2015). *Επιχειρηματική Ευφυΐα και εξόρυξη δεδομένων*. Αθήνα: Ελληνικά ακαδημαϊκά συγγράμματα και βοηθήματα.

Kutner,M., Nachtsheim, C., Neter, J., Li, W. (2005). Applied statistical linear models. McGraw-Hill. Boston

Lake, I., Lovett, A., Bateman, I., Day, B. (1999). *Using GIS and large scale digital data to implement hedonic pricing studies.* Research article . Geographical Information Science, 2000,Vol.14,No6,521-541

Λαμπρόπουλος, Α.(2013). *Αυτοματοποίηση μαζικών εκτιμήσεων αξιών ακινήτων για τον ελληνικό χώρο*. Διδακτορική Διατριβή. Επιβλέπων Καθηγητής: Ζεντέλης, Π. Σχολή Αγρονόμων και Τοπογράφων μηχανικών, Εθνικό Μετσόβιο Πολυτεχνείο, Αθήνα.

Λιανός, Θ., Παπαβασιλείου, Α., Χατζηανδρέου, Α. (2015). *Αρχές οικονομικής θεωρίας*. Ινστιτούτο τεχνολογίας υπολογιστών και εκδόσεων, Διοφαντος.

Li, M.M. and Brown, H.J. (1980) *Micro-neighbourhood externalities and hedonic housing prices, Land Economics*, 56(2): 125-141.

Li, P. (2005). *Box- Cox transformations: An overview. Department of statistics*, University of Connecticut.

Li, D., Zhang, X ,(2014) . *Analysis of shanghai housing price based on the hedonic price method.* Bio Technology, an Indian journal.

Μαντούβαλου, Μ. (1996) *Αστική γαιοπρόσοδος, τιμές γης και διαδικασίες ανάπτυξης του αστικού χώρου*. Επιθεώρηση Κοινωνικών Ερευνών, 89-90,1996,53-80.

Μαστοροκώστας, Π. (2015). *Εισαγωγή στην ασαφή λογική- Ασαφή σύνολα-συναρτήσεις συμμετοχής.* Αθήνα: Ελληνικά ακαδημαϊκά συγγράμματα και βοηθήματα.

Mach, L. (2017). *The application of classical neural regression models for the valuation of residential real estate*. Folia Oeconomica stetinensia, Opole

University of technology, Opole, Poland.

Mak, S., Choy, L., & Ho, W.,(2009). *Quantile Regression Estimates of Hong Kong Real Estate Prices.*

Mayor, K.,Lyons, S., Duffy, D.(2009). *A hedonic analysis of the value of parks and green spaces in thw Dublin area*. Working paper No331

Mankat, M., (2018). *Analysis of Impact of Accessibility on Residential Property Values in Gotri Area of Vadodara City, India Using OLS and GWR.* IJSRSET, Volume 4 Issue 1.

Μηλάκα, Κ. (2010). *Χωρική ανάλυση της εξέλιξης αστικών περιοχών μέσω πολυμεταβλητής επεξεργασίας εναλλακτικών σεναρίων*. Διδακτορική διατριβή, Επιβλέπων Καθηγητής: Φώτης Γ., Τμήμα μηχανικών χωροταξίας πολεοδομίας και περιφερειακής ανάπτυξης, Πανεπιστήμιο Θεσσαλίας, Βόλος.

Μίμης, Α., Ροβολής, Α.,, Στάμου, Μ. (2012). *Η χρήση των τεχνητών νευρωνικών δικτύων στην ανάλυση των αξιών ακινήτων*. Άρθρο που παρουσιάστηκε στα πρακτικά Τακτικού επιστημονικού συνεδρίου 2012.

Μπακιρτζόγλου, Χ.(2012). *Εισαγωγή στις εκτιμήσεις ακινήτων*. Πρακτικά συνεδρίου ΤΕΕ. Θεσσαλονίκη.

Μπούτσικας, Μ.(2004). *Σημειώσεις μαθήματος «« Στατιστικά προγράμματα»«*. Πανεπιστήμιο Πειραιώς.

Montgomery, D., Peck, E.,Vining (2012). *Introduction to linear regression analysis*. Fifth edition. John Wiley & Sons INC.USA

Morgan, G.,Leech, N., Gloecner, G.,Berrett, K. (2004). *SPSS for introductory statistics , use and interpretation.* Second edition. Lawrence Erlbaum associates, Publishers. London.

Neuman, W. L.(2013). *Social Research Methods_ Qualitative and Quantitative Approaches*-Pearson Education Limited

Νικήτα, Ε. (2012). *Έννοιες στατιστικής και εφαρμογές με το SPSS*. Θεσσαλονίκη.

Ozgur, C., Hughes, Z., Rogers, G., Parvee, S. (2016). *Multiple linear regression applications in real estate pricing.* international journal of mathematics and statistics invention (IJMSI). E-ISSN: 2321 – 4767 P-ISSN: 2321 – 4759

Osborne, J. (2010). *Improving your data transformations: Applying the |Cox Box*

transformation. Practical assessment research & evaluation. Vol 15 No12

Παγουρτζή, Ε. (2003). *Μεθοδολογία και ανάπτυξη συστήματος χωροχρονικών δεδομένων για την εκτίμηση τιμών ακινήτων.* Διδακτορική διατριβή. ΕΜΠ. Αθήνα.

Παπαγεωργίου, Ι. (2015). *Θεωρία δειγματοληψίας*. Αθήνα: Ελληνικά ακαδημαϊκά συγγράμματα και βοηθήματα.

Παγουρτζή, Ε.,Ασσημακόπουλος, Β.,Χατζηχρήστος, Θ., French, N. (2003). *Real estate appraisal*. A review of valuation methods. Journal of property investment and finance, 383-401

Πανάρετος, Ι.(1994). Γραμμικά μοντέλα με έμφαση στις εφαρμογές. Οικονομικό Πανεπιστήμιο Αθηνών.

Παπαδόπουλος, Γ. (χ.η). *Σημειώσεις εργαστηρίου μαθηματικών και στατιστικής*. Γεωπονικό πανεπιστήμιο Αθηνών.

Παπαευθυμίου, Ι. (2013). *Εκτίμηση αξίας ακινήτων. Αξιοποίηση τοπικών μοντέλων παλινδρόμησης.* Άρθρο που παρουσιάστηκε στο 1ο Συνέδριο Χωρικής Ανάλυσης, 2013.

Παυλόπουλος, Γ. (2006). Εισαγωγή στην στατιστική επεξεργασία δεδομένων με το SPSS for Windows. Εθνικό και Καποδιστριακό Πανεπιστήμιο Αθηνών, τμήμα φιλοσοφίας, ψυχολογίας, παιδαγωγικής και ψυχολογίας τομέας ψυχολογίας.

Πετρίδης, Δ. (2015) *Ανάλυση πολυμεταβλητών τεχνικών, Εφαρμογές περιπτώσεων*. Αθήνα: Ελληνικά ακαδημαϊκά συγγράμματα και βοηθήματα.

Pek, J.,Wong, C.(2017). *Data transformations for inference with linear regression: Clarifications and recommendations*. Practical assessment research & evaluation. Vol 22 No9

Peresetsky, A., (2014). *The price of Mpscow apartments*. National Research University Higher School of Economics.

Πομώνης, Γ.(2009). *Διδακτικές στρατηγικές μαθήτυποι και δεξιότητες κριτικής σκέψης στην διδασκαλία της χρηματοοικονομικής προσέγγισημε την quantile regression* . Διδακτορική Διατριβή. Πανεπιστήμιο Πατρών. Τμήμα διοίκησης επιχειρήσεων. Πάτρα.

Πραμαγγιούλης, Π. (2008). *Οδηγός ανάλυσης δεδομένων με την χρήση του SPSS*. Οικονομικό Πανεπιστήμιο Αθηνών.

Rahman, K., Sathik,M., Kannan. S,(2012) , *Multiple linear regression models in outlier detection*, International Journal of Research in Computer Science, eISSN 2249-8265 Volume 2 Issue 2 (2012) pp. 23-28, White Globe Publications, doi:10.7815/ijorcs.22.2012.018.

Raslanas, S., Tupenaite, L., Steinbergas, T. (2005). *Research on the prices of flats in the south east London and Vilnius*. International Journal of Strategic management (2006) 10, 51-63.

Rawlings, J., Pantula, S., Dickey, D.(1998). *Applied regression analysis: A research tool* . Second edition. Springer.USA.

Rodriguez, M., Sirmans, C.F (1994). *Quantifying the value of a view in single family housing markets*. 600-603

So, H.M., Tse, R.Y.C. and Ganesan, S. (1996). *Estimating the influence of transport on house prices: Evidence from Hong Kong, journal of property valuation and investment*. 15(1): 40-47.

Στάμου, Μ. (2016). *Διερεύνηση των αξιών ακινήτων στην μητροπολιτική περιοχή εφαρμογή σύγχρονων μεθόδων χωρικής ανάλυσης.* Διδακτορική διατριβή, Επιβλέπων Καθηγητής: Καραγάνης , Α., Τμήμα Οικονομικής και Περιφερειακής ανάπτυξης, Πάντειο Πανεπιστήμιο, Αθήνα.

Sirmans, G., Macpherson, D. (2003). *The composition of hedonic pricing models: A review of the literature*. National association of realtors.

Sopranzetti, B.,(χ.η). *Hedonic regression models*. Rutgers University, USA

Stevens, S. (1946). *On the theory of scales of measurement*. Science, Vol.103, No2694, pp 677-680.

Τασιόπουλος, Κ.(2012). *Οι αξίες των ακινήτων υπό το πρίσμα ευρύτερων παραμέτρων οικονομικής κρίσης*, Ερευνητική εργασία, Επιβλέπων Καθηγητής: Καρανικόλας, Ν., Τμήμα μηχανικών χωροταξίας και ανάπτυξης, ΑΠΘ, Θεσσαλονίκη.

Ταγρής, Μ. (2014). *Στατιστική με τη χρήση του πακέτου IBM SPSS 22. Statistical package for the social sciences.* Αθήνα και Nottingham.

Τζαβέλας, Γ.(2011) . *Εισαγωγή στη Λογιστική παλινδρόμηση.*

Tobler, W. R. (1970). *A computer movie simulating urban growth in the Detroit*

region. Economic Geography, 46(2), 234-240.

Farmer, M., Lipscomb. (2010) C. *Using quandile regression in hedonic analysis to reveal submarket competition*. JRER, Vol.32, No.4-2010 pp, 436,459.

Χαλικιάς, Μ. Μανωλέσσου, Α. Λάλου, Π. (2015). *Μεθοδολογία έρευνας και εισαγωγή στη στατιστική ανάλυση δεδομένων με το IBM SPSS STATISTICS.* Αθήνα: Ελληνικά ακαδημαϊκά συγγράμματα και βοηθήματα.

Χατζησπύρος, Σ. (2013). *Πιθανότητες.* Πανεπιστήμιο Αιγαίου. Σάμος.

Ιστοσελίδες Διαδικτύου

Αγγελίδης, Β. (2011). Ανάλυση δεδομένων διαφάνειες μαθήματος. Δημοκρίτειο Πανεπιστήμιο Θράκης.
Ανακτήθηκε 20 Μαΐου 2018, από
https://eclass.duth.gr/modules/document/file.php/TME179/%CE%A0%CE%B1%CF%81%CE%BF%CE%B8%CF%83%CE%B9%CE%AC%CF%83%CE%B5%CE%B9%CF%82%20draft/lecture6_regretion.pdf

Αστικός Κώδικας , Εμπράγματο Δίκαιο.
Ανακτήθηκε 1 Μαΐου 2018 , από
http://www.fa3.gr/nomothesia_2/nomoth_gen/19-Dikaio-embragmato.htm

Αθήνα.
Ανακτήθηκε 15 Μαΐου 2018, από
https://el.wikipedia.org/wiki/%CE%91%CE%B8%CE%AE%CE%BD%CE%B1

Curtis, K. (2013). Econometrics Notes & R code.
Ανακτήθηκε 1 Ιουνίου 2018, από
http://www.cazaar.com/ta/econ113/interpreting-beta

Δημητροπουλάκης, Π. (χ.η). Εισαγωγή στην χρήση του SPSS για Windows.
Ανακτήθηκε 20 Μαΐου 2018 από
http://www.lib.teiher.gr/webnotes/seyp/SPSS/

Δημοτικά διαμερίσματα της Αθήνας.
Ανακτήθηκε 15 Μαΐου 2018, από
http://www.kallikratis.org/dimos-athinaion-kentrikoy-tomea/

Ελληνική Στατιστική Αρχή. (2016) . *Συμβολαιογραφικές πράξεις ετών 2010 έως 2016*.
Ανακτήθηκε 15 Απριλίου 2018 , από

http://www.statistics.gr/el/statistics/-/publication/SJU24/2015

Ευρωπαϊκή Στατιστική Υπηρεσία Eurostat.(2015) *Στατιστικές στέγασης* .
Ανακτήθηκε 18, Απριλίου 2018, από
http://ec.europa.eu/eurostat/statistics-explained/index.php/Housing_statistics/el

Ευρωπαϊκή Στατιστική Υπηρεσία Eurostat Ποσοστό δαπάνης επί του διαθέσιμου εισοδήματος για το έτος 2014-2015.
Ανακτήθηκε 13 Απριλίου 2018, από
http://ec.europa.eu/eurostat/statistics-explained/index.php?title=File:Housing_cost_overburden_rate,_2014%E2%80%9315.png

Ιστορία της πόλης της Αθήνας.
Ανακτήθηκε 15 Μαΐου 2018, από
https://www.cityofathens.gr/episkeptes/eksereyniste-tin-athina/istoria-tis-polis

Κανονισμός (ΕΚ) αριθ. 1126/2008της Επιτροπης της 3ης Νοεμβρίου 2008 για την υιοθέτηση ορισμένων διεθνών λογιστικών προτύπων σύμφωνα με τον κανονισμό (ΕΚ) αριθ. 1606/2002 του Ευρωπαϊκού Κοινοβουλίου και του Συμβουλίου.
Ανακτήθηκε 20 Απριλίου 2018, από
https://www.taxheaven.gr/laws/circular/view/id/8142

Καρανικόλας, Ν.(2010). Μέθοδος κεφαλαιοποίησης εισοδήματος. Κύκλος σεμιναρίων μικρής διάρκειας. ΤΕΕ Κεντρικής Μακεδονίας.
Ανακτήθηκε 25 Μαΐου 2018, από
http://library.tee.gr/digital/kma/kma_m1473/kma_m1473_karanikolas.pdf

Κείμενο διαβούλευσης για το έργο : Αθήνα 2020 του Δήμου Αθηναίων.
Ανακτήθηκε 1 Μάιου 2018, από
http://developathens.gr/athina2020/images/keimeno_diavoulefsis.pdf

Μηλιώνης, Α. (2017). Οικονομική χρηματοοικονομική στατιστική . Σημειώσεις παραδόσεων . Οικονομετρικά υποδείγματα χρονικών σειρών. Πανεπιστήμιο Αιγαίου.
Ανακτήθηκε 20 Μαΐου 2018, από
http://www.samos.aegean.gr/actuar/amilionis/notes/%CE%A4%CE%99%CE%9C%CE%95%20-%CE%9A%CE%95%CE%A6%CE%91%CE%9B%CE%91%CE%99%CE%9F%202-2015-BETA%20VERSION-30.10.15-SENT.pdf

Οικονομική και Κοινωνική Επιτροπή της Ελλάδος (ΟΚΕ) (2009), Η κατασκευαστική και οικοδομική δραστηριότητα ως συνιστώσες της οικονομικής ανάπτυξης και οι επιπτώσεις της οικονομικής κρίσης σε αυτές, Γνώμη Πρωτοβουλίας.
Ανακτήθηκε 5 Μαΐου 2018, από

http://www.oke.gr/opinion/op_225.pdf

Open source λογισμικό GWR4.
Ανακτήθηκε 28 Οκτωβρίου 2018, από
https://gwrtools.github.io/gwr4-downloads.html

Open source λογισμικό MGWR4.
Ανακτήθηκε 24 Νοεμβρίου 2018, από
https://sgsup.asu.edu/sparc/mgwr

Παπάνα, Α. (2018). Οικονομετρία. Οικονομετρικά προβλήματα. Ενότητα 10. Διαφάνειες μαθήματος οικονομετρίας. ΑΠΘ.
Ανακτήθηκε 1 Ιουνίου 2018, από
http://users.auth.gr/agpapana/Econometrics_Lectures/Lecture10_econometrics.pdf

REMAX Ελλάς.(2016). *Πανελλαδική έρευνα της για τις πωλήσεις ακινήτων έτους 2016*.
Ανακτήθηκε 20 Απριλίου 2018, από
https://www.remax.gr/media/files/%CE%A0%CE%B1%CE%BD%CE%B5%CE%BB%CE%BB%CE%B1%CE%B4%CE%B9%CE%BA%CE%AE%20%CE%95%CF%84%CE%AE%CF%83%CE%B9%CE%B1%20%CE%88%CF%81%CE%B5%CF%85%CE%BD%CE%B1%20%CE%91%CE%BA%CE%B9%CE%BD%CE%AE%CF%84%CF%89%CE%BD%20REMAX%202016(1).pdf

Σταμούλης, Α. (2011). Γραμμική παλινδρόμηση, διαφάνειες μαθήματος.
Ανακτήθηκε 20 Μαΐου 2018, από
http://www.icbs.gr/innet/UsersFiles/students/tutor_notes/postgraduate/Stage_1/Stamoulis/2011-2012/Statistics/Grammiki%20Palindromisi.pdf

Standard of mass appraisal of real property.(2012- 2013)
Ανακτήθηκε 1 Μαΐου 2018, από
https://www.iaao.org/media/standards/MARP_2013.pdf

Spiti24
Ανακτήθηκε την 1 Μαίου 2018 από
www.spiti24.gr

Spatial structures in the social sciences. Spatial regression with GeoDa . Geoda: Spatial regression. S4 Training modules. Brown University.
Ανακτήθηκε 12 Ιουνίου 2018, από
https://www.brown.edu/academics/spatial-structures-in-social-sciences/

Σύνολα δεδομένων *data.gov*

Ανακτήθηκε 1 Ιουλίου 2018, από

http://www.data.gov.gr/dataset?_organization_limit=0&organization=cityofathens

The survey system.
Ανακτήθηκε 10 Μαΐου 2018, από
https://www.surveysystem.com/sscalc.htm

Τράπεζας της Ελλάδος.(2016) *Έκθεση Διοικητή Στουρνάρα για το έτος 2016 (Φεβρουάριος 2017, Κεφάλαιο V, Ενότητα 3).*
Ανακτήθηκε 20 Απριλίου 2018, από
https://www.bankofgreece.gr/BoGDocuments/%CE%95%CE%9A%CE%98%CE%95%CE%A3%CE%97_%CE%94%CE%99%CE%9F%CE%99%CE%9A%CE%97%CE%A4%CE%97_2016_%CE%A6%CE%95%CE%92_2017_%CE%9A%CE%95%CE%A6_V_%CE%95%CE%9D%CE%9F%CE%A4_3.pdf

Uniform Standards of professional appraisal practice (σελ.U-3)
Ανακτήθηκε 1 Μαΐου 2018, από
http://www.brienzaappraisal.com/images/2012-2013USPAP.pdf

Υπουργείο Οικονομικών
Ανακτήθηκε την 1 Μαίου 2018
https://www.minfin.gr/-/e-appa

Χαρδούβελης, Γ. *Η σπουδαιότητα της αγοράς κατοικίας στην οικονομία.*
Ανακτήθηκε 15 Απριλίου 2018, από
https://www.bankofgreece.gr/BoGDocuments/%CE%91%CE%93%CE%9F%CE%A1%CE%91_%CE%91%CE%9A%CE%99%CE%9D%CE%97%CE%A4%CE%A9%CE%9D1.pdf