ΠΟΛΥΤΕΧΝΕΙΟ ΚΡΗΤΗΣ

Τμήμα Μηχανικών Παραγωγής και Διοίκησης

Διπλωματική Εργασία

Παράγοντες επιρροής εσόδων σε διαδικτυακές πλατφόρμες βραχυχρόνιας ενοικίασης καταλυμάτων.

Τσόλης Γεώργιος

Επιβλέπων καθηγητής: Τσαφαράκης Στέλιος

Χανιά, Μάιος 2018

Ευχαριστίες

Αρχικά θα ήθελα να δώσω ιδιαίτερες ευχαριστίες στον επιβλέποντα καθηγητή μου κ. Τσαφαράκη Στέλιο που συνέβαλε με την βοήθειά του στην ολοκλήρωση αυτής της διπλωματικής εργασίας.
Έπειτα οφείλω ένα μεγάλο ευχαριστώ στην οικογένειά μου, που με στήριξε κατά τη διάρκεια των σπουδών μου και συνεχίζει να με στηρίζει αυτά τα χρόνια.
Τέλος θα ήθελα να ευχαριστήσω τους φίλους μου για τη στήριξη και τις όμορφες στιγμές που μου πρόσφεραν αυτά τα χρόνια.

Περίληψη

Η παρούσα διπλωματική εργασία έχει στόχο την εύρεση των καθοριστικών παραγόντων που συμβάλλουν στη διαμόρφωση των εσόδων στα καταλύματα της Αθήνας.
Στο βιβλιογραφικό μέρος έγινε συλλογή πληροφοριών από μελέτες και επιστημονικά άρθρα τα οποία αναφέρονται στους παράγοντες που επηρεάζουν την τιμή στις καταχωρήσεις της Airbnb και παρουσιάστηκαν τα αποτελέσματα των μελετών.
Στο ερευνητικό μέρος έγινε συλλογή δεδομένων από την ιστοσελίδα της Airbnb για την εύρεση του συνόλου των καταλυμάτων. Έπειτα έγινε ανάλυση και εφαρμογή της βηματικής παλινδρόμησης και εξετάστηκαν περίπου σαράντα διαφορετικοί παράγοντες οι οποίοι αναφέρονται στις παροχές του καταλύματος, στους κανόνες ακύρωσης, στην τοποθεσία του καταλύματος κ.α.
Τέλος έγινε εξαγωγή των συμπερασμάτων και σύγκριση των αποτελεσμάτων μας σε σχέση με τα αποτελέσματα των επιστημονικών ερευνών.

Abstract

This dissertation aims at finding the key determinants of the revenue structure in Airbnb accommodation in Athens.
As for the bibliography, a selection of information from studies and scientific articles was made , concerning the factors that affect entries' price in Airbnb accommodation establishments. Afterwards, the results of the studies were presented.
As far as the research project is concerned, a selection of data from Airbnb website was made in order to find the total number of the accommodation establishments. Then, regression analysis was applied and about forty different factors were examined, concerning the accommodation's facilities, the cancellation rules, the accommodation's location etc.
Finally , conclusions were drawn and we compared the results of our research to the ones of other scientific researches.

Περιεχόμενα

ΚΕΦΑΛΑΙΟ 1

1.Εισαγωγή.

1.1 Εισαγωγή.

Τα τελευταία χρόνια, αναδύθηκε ένα νέο επιχειρηματικό μοντέλο, γνωστό ως "οικονομία του διαμερισμού " στον τομέα της στέγασης του κλάδου του τουρισμού και της φιλοξενίας. Η Airbnb έχει πρωτοπορήσει στη χρήση αυτού του επιχειρηματικού μοντέλου για να συνδέσει ανθρώπους που διαθέτουν ιδιόκτητα καταλύματα (όπως κενά δωμάτια ή διαμερίσματα) με εκείνους που χρειάζονται προσωρινά καταλύματα (όπως οι τουρίστες) μέσω ψηφιακών αγορών. Η αύξηση των εσόδων είναι ένας πάρα πολύ σημαντικός παράγοντας που έχει ως αποτέλεσμα τη μακροπρόθεσμη επιτυχία της ενοικίασης καταλυμάτων. Μια έρευνα που θα είχε ως στόχο την διαμόρφωση εσόδων των καταλυμάτων στέγασης θα πρόσφερε σημαντικές ιδέες στους οικοδεσπότες ώστε να βελτιώσουν τα κέρδη και να αναπτύξουν την επιχείρηση τους. Ως εκ τούτου, έχει εντοπιστεί ότι τα χαρακτηριστικά του ξενιστή, οι ειδικές παροχές και ορισμένα διαφοροποιημένα χαρακτηριστικά διαμονής συνδέονται με τα έσοδα που θα επιφέρει το κατάλυμα. Έτσι η συγκεκριμένη έρευνα φιλοδοξεί να εξετάσει διάφορους παράγοντες ώστε να καταλήξει πόσοι και ποιοι τελικά είναι σημαντικοί, στην επίδραση των εσόδων του καταλύματος.

1.2 Σκοπός της μελέτης.

Σκοπός αυτής της μελέτης είναι να προσδιοριστούν οι καθοριστικοί παράγοντες που συμβάλλουν στη διαμόρφωση των εσόδων για την κατανομή των ευκαιριών διαμονής που βασίζονται στην οικονομία της ψηφιακής αγοράς (συγκεκριμένα στο Airbnb.com). Εξετάστηκε δείγμα 603 καταλυμάτων στη περιοχή της Αθήνας που είναι εγγεγραμμένα στο Airbnb.com. Η μέθοδος της βηματικής παλινδρόμησης χρησιμοποιήθηκε για τη διερεύνηση των καθοριστικών παραγόντων ώστε να εξεταστούν, χαρακτηριστικά τοποθεσίας, κανόνες ακύρωσης, παροχές και υπηρεσίες του καταλύματος, κριτικές κ.α. Επιπλέον η συγκεκριμένη έρευνα μπορεί να αποτελέσει κίνητρο για περαιτέρω μελέτη διότι τα αποτελέσματα έχουν σημαντικές συνέπειες για το σχεδιασμό συστημάτων υπολογισμού των εσόδων των καταλυμάτων της Airbnb.

1.3 Δομή της εργασίας.

Η συγκεκριμένη εργασία αποτελείται από εφτά κεφάλαια όπου η δομή της φαίνεται στη συνέχεια.

Πρώτο κεφάλαιο: Γίνεται μια γενική εισαγωγή στο θέμα και αναφέρονται οι στόχοι για τους οποίους πραγματοποιήθηκε.

Δεύτερο κεφάλαιο: Αναλύεται ο όρος τουρισμός, παρουσιάζονται διάφορα στατιστικά στοιχεία τα οποία αναφέρονται στον τουρισμό της Ελλάδας και γίνεται μια αναφορά στην εταιρεία της Airbnb.

Τρίτο κεφάλαιο: Παρουσιάζονται επιστημονικά άρθρα και μελέτες τα οποία αναφέρονται στους βασικούς παράγοντες που συμβάλλουν στη διαμόρφωση της τιμής στα καταλύματα της Airbnb.

Τέταρτο κεφάλαιο: Γίνεται παρουσίαση της μεθόδου της γραμμικής παλινδρόμησης, ανάλυση βασικών εννοιών και περιγράφεται το μοντέλο της βηματικής παλινδρόμησης.

Πέμπτο κεφάλαιο: Εφαρμόζεται η μέθοδος της βηματικής παλινδρόμησης, παρουσιάζονται αναλυτικά τα βήματα της μεθόδου και αναλύεται το τελικό μοντέλο.

Έκτο κεφάλαιο: Γίνεται εξέταση περίπου σαράντα διαφορετικών παραγόντων οι οποίοι αναφέρονται στα χαρακτηριστικά τοποθεσίας, κανόνες ακύρωσης, παροχές και υπηρεσίες του καταλύματος, κριτικές ώστε να δούμε ποιες μεταβλητές έχουν την πιο σημαντική επίδραση στα έσοδα των καταλυμάτων.

Έβδομο κεφάλαιο: Παρουσιάζονται τα συμπεράσματα και γίνεται σύγκριση των αποτελεσμάτων μας σε σχέση με τα αποτελέσματα των επιστημονικών ερευνών.

ΚΕΦΑΛΑΙΟ 2

2.Εισαγωγή στο φαινόμενο του τουρισμού.

2.1 Γενικά για τον τουρισμό.

Η λέξη τουρισμός προέρχεται από τη γαλλική λέξη “tour” και την αγγλική λέξη “touring”, που προέρχονται από τη λατινική λέξη “tornus” και σημαίνουν γύρος ή περιήγηση (βαρβαρέσος, 2000). Ο τουρισμός μπορεί να οριστεί ως μια δραστηριότητα του ανθρώπου η οποία περιλαμβάνει την μετακίνηση του από το τόπο μόνιμης κατοικίας του σε έναν άλλο, ώστε να ικανοποιήσει τις τουριστικές του επιθυμίες, τις φυσικές του ανάγκες, οι οποίες δεν είναι πάντα επιθυμίες ή ανάγκες ξεκούρασης αλλά είναι ανάγκες γνωριμίας του πολιτισμού του, της ιστορίας του και του περιβάλλοντος του.

Η διεθνής Ακαδημία του τουρισμού (A.I.T. – Monte Carlo) όρισε ότι « τουρισμός είναι το σύνολο των ανθρώπινων μετακινήσεων και των δραστηριοτήτων που προκύπτουν από αυτές. Αυτές υποκινούνται, σε διαφορετικό βαθμό και ένταση, από τον πόθο της απόδρασης που υπάρχει στα άτομα.»

Ο J. Gouer το 1929 ανέφερε ως τουρισμό « τον τρόπο του να ταξιδεύεις, συνδυάζοντας σε διαφορετικές αναλογίες, σύμφωνα με τα άτομα, την περιέργεια του πνεύματος μαζί με τον πόθο του να προσαρμοστεί κάποιος στα ξένα ήθη και έθιμα.»

Η Διεθνής Οργάνωση Ειδικών Επιστημόνων του Τουρισμού (International Association of Scientific Expert in Tourism) το 1981 όρισε ως τουρισμό « τις συγκεκριμένες δραστηριότητες που επιλέγονται κατά βούληση και λαμβάνουν χώρα εκτός του συνηθισμένου περιβάλλοντος »

Με την κατοχύρωση των «θεμελιωδών κοινωνικών δικαιωμάτων» δηλαδή συγκεκριμένο ωράριο εργασίας, πενθήμερη εργασία, δικαίωμα άδειας για διακοπές ο τουρισμός ενισχύθηκε ακόμα περισσότερο.

2.2 Ο τουρισμός στην αρχαιότητα

Ο τουρισμός, παρόλο που στηρίχτηκε από την αρχαιότητα έως σήμερα στον ελεύθερο χρόνο και τη μετακίνηση, εντούτοις παρουσίασε διαχρονικά μια μορφολογική εξέλιξη και μια διευρυνόμενη χωρικά διάρθρωση (Βαρβαρέσος, 2000).

Στην αρχαία Ελλάδα και στην αρχαία Ρώμη, η κοινωνική τάξη και η προέλευση των πολιτών καθοριζόταν εν μέρει από τον ελεύθερο χρόνο που διάθεταν. Οι πολίτες που ανήκαν σε υψηλότερη κοινωνική τάξη, άρα διέθεταν ελεύθερο χρόνο μπορούσαν να λαμβάνουν μέρος σε διάφορους τύπους ταξιδιωτικών μετακινήσεων, όπως για παράδειγμα τα διεθνή ταξίδια που έκαναν οι αρχαίοι Έλληνες και Ρωμαίοι στην Αίγυπτο για λόγους αναψυχής, τις παρακολουθήσεις θεατρικών παραστάσεων, αθλητικών αγώνων και εορτών.

Ο Ηρόδοτος (480-421 π.χ.) άρχισε να αναφέρει τα γεωγραφικά και πολιτιστικά χαρακτηριστικά των τόπων όπου ένας επισκέπτης μπορούσε να αναζητήσει σε μια ταξιδιωτική του μετακίνηση. Στο 19^{ο} αιώνα οι αναφορές του Ηρόδοτου χαρακτηρίστηκαν ως «μοντέλο τουρίστα». Εκείνη τη περίοδο άρχισαν να κατασκευάζονται τα πρώτα κρατικά πανδοχεία με δωρεάν υπηρεσίες στους ταξιδιώτες και τα πρώτα ιδιωτικά πανδοχεία με παροχές υπηρεσιών επί πληρωμή. Κατά το πέρας των χρόνων άρχισε να παρουσιάζεται αύξηση της κατασκευής των πανδοχείων και εμφανίστηκαν οι πρώτες τάσεις εμπορικοποίησης των ταξιδιωτικών μετακινήσεων.

2.3 Ανάλυση των ειδών τουρισμού.

Στη συνέχεια παρουσιάζονται οι κυριότερες ειδικές και εναλλακτικές μορφές τουρισμού οι οποίες μπορούν και αναπτύσσονται στην Ελλάδα. Τα είδη του τουρισμού κατηγοριοποιούνται σε (Σωτηριάδης και Φαρσάρη, 2010):

- Γενικός τουρισμός

Ο γενικός τουρισμός ή αλλιώς χαρακτηρίζεται και ως τουρισμός αναψυχής, περιλαμβάνει τις δραστηριότητες των ανθρώπων στον ελεύθερο χρόνο τους και ειδικότερα σχετίζεται με την επιθυμία του ατόμου για ανάπαυση και αναψυχή.

- Μορφωτικός τουρισμός

Ο μορφωτικός τουρισμός είναι ένα είδος τουρισμού, στον οποίο οι ταξιδιώτες έχουν σκοπό να παρακολουθήσουν πολιτιστικές εκδηλώσεις, να επισκεφτούν αρχαιολογικούς χώρους, μνημεία, μουσεία, θεατρικές παραστάσεις, ιστορικά μνημεία .

- Τουρισμός εκθέσεων

Ο τουρισμός εκθέσεων ουσιαστικά αποτελεί και αυτός ένα είδος τουρισμού. Σε αυτήν την κατηγορία ανήκουν κάθε είδους εκθέσεις , όπου οι ταξιδιώτες να μπορούν να συνδυάσουν το τουρισμό μαζί με την επαγγελματική τους ενημέρωση όπως για παράδειγμα τουρισμό πόλης, μορφωτικό τουρισμό κ.τ.λ.

- Τουρισμός υγείας

Ο τουρισμός υγείας ή αλλιώς θεραπευτικός τουρισμός, αναπτύσσεται δυναμικά ολοένα και περισσότερο. Οι άνθρωποι που λαμβάνουν μέρος σε αυτόν έχουν ως κύριο στόχο την αποκατάσταση και διατήρηση της υγείας τους, τη θεραπεία και ανάρρωση τους από ασθένειες κ.τ.λ.

- Τουρισμός άθλησης

Ο τουρισμός άθλησης περιλαμβάνει τα άτομα που έχουν ως στόχο την άσκηση ενός αθλήματος, όπως για παράδειγμα το ποδόσφαιρο, την κολύμβηση, το τρέξιμο και γενικά προγράμματα γυμναστικής.

- Τουρισμός πόλης

Ο τουρισμός πόλης αποτελεί μια μορφή τουρισμού όπου τα άτομα που συμμετέχουν σε αυτόν ταξιδεύουν σε μια πόλη για ένα μικρό χρονικό διάστημα ώστε να την περιηγηθούν.

- Συνεδριακός τουρισμός

Ο συνεδριακός τουρισμός είναι ένα είδος τουρισμού όπου τα άτομα συμμετέχουν σε κάθε είδους οργανωμένες εκδηλώσεις, συνέδρια ή συναντήσεις σε οποιοδήποτε επίπεδο.

- Τουρισμός περιπέτειας

Ο τουρισμός περιπέτειας έχει έντονο το στοιχείο του απροσδόκητου, του άγνωστου και της έκπληξης. Οι ταξιδιώτες λαμβάνουν μέρος σε σαφάρι, περιπλανήσεις σε άγνωστες περιοχές, κωπηλασία σε ορμητικούς ποταμούς κ.α.

- Οικογενειακός τουρισμός

Ο οικογενειακός τουρισμός είναι ένα είδος τουρισμού όπου για να πραγματοποιηθεί πρέπει να γίνει η επιλογή του μέρους και του χρονικού διαστήματος όπου θα πραγματοποιηθούν διότι υπάρχει η ύπαρξη παιδιών.

- Τουρισμός τρίτης ηλικίας

Μια άλλη μορφή τουρισμού είναι και ο τουρισμός τρίτης ηλικίας όπου σε αυτόν ανήκουν τα άτομα μεγαλύτερης ηλικίας όπου πρέπει συνήθως να διαθέτουν επιθυμία για ταξίδια, ελεύθερο χρόνο και σταθερό εισόδημα.

- Τουρισμός χειμερινών σπορ

Ο τουρισμός χειμερινών σπορ είναι μια μορφή του τουρισμού που λαμβάνει μεγάλο ποσοστό τα τελευταία χρόνια. Σε αυτόν λαμβάνουν μέρος άτομα με υψηλό εισόδημα διότι για να πραγματοποιηθεί έχει ακριβές παροχές υπηρεσιών, όπως για παράδειγμα το σκι σε χιονοδρομικά κέντρα

- Τουρισμός παραχείμασης

Ο τουρισμός παραχείμασης αναφέρεται σε άτομα όπου επιλέγουν να φύγουν από το τόπο μόνιμης κατοικίας τους το χειμώνα, για να αποφύγουν τις αντίξοες καιρικές συνθήκες, και επιλέγουν να κάνουν διακοπές σε χώρες με ήπια κλίματα.

- Ορεινός τουρισμός

Ο ορεινός τουρισμός αναφέρεται σε άτομα όπου παίρνουν μέρος σε δραστηριότητες υπαίθριας αναψυχής και τουρισμού που εκδηλώνεται αποκλειστικά και μόνο στις ορεινές περιοχές των χωρών υποδοχής τουριστών που επιθυμούν να αναπτύξουν αυτής της μορφής τον τουρισμό.

- Αγροτικός τουρισμός

Ο αγροτικός τουρισμός ή αλλιώς αγροτουρισμός, αναπτύσσεται στον αγροτικό χώρο, αναφέρεται στις δραστηριότητες υπαίθριας αναψυχής και εντάσσονται στο πλαίσιο του αγροτικού περιβάλλοντος.

- Θαλάσσιος τουρισμός

Ο θαλάσσιος τουρισμός αναφέρεται στα άτομα όπου λαμβάνουν μέρος σε τουριστικές δραστηριότητες που αναπτύσσονται στο θαλάσσιο χώρο μιας χώρας.

- Χρονομεριστικός τουρισμός

Ο συγκεκριμένος τουρισμός περιλαμβάνει άτομα με υψηλό εισόδημα και βασίζεται στην επί πληρωμή παροχών υπηρεσιών σε ξενοδοχεία ή συγκροτήματα διαμερισμάτων για συγκεκριμένες ημέρες κάθε χρόνο.

- Θρησκευτικός τουρισμός

Ο θρησκευτικός τουρισμός είναι μια άλλη μορφή τουρισμού όπου οι ταξιδιώτες επισκέπτονται θρησκευτικούς τόπους ώστε να παρακολουθήσουν θρησκευτικές εκδηλώσεις και ότι αναφέρεται σε θρησκευτική τελετή.

- Επιλεκτικός τουρισμός

Ο επιλεκτικός τουρισμός δεν μπορεί να πραγματοποιηθεί από όλα τα άτομα διότι πρέπει διαθέτουν υψηλό εισόδημα. Στη συγκεκριμένη κατηγορία οι ταξιδιώτες πραγματοποιούν τις διακοπές τους μεμονωμένα διαλέγοντας από μόνοι τους τη χώρα που θα επισκεφτούν.

- Τουρισμός κινήτρων

Το συγκεκριμένο είδος τουρισμού πραγματοποιείται από εταιρείες για τους εργαζόμενους της, ώστε να έχουν μια ολιγοήμερη ξεκούραση.

- Οικολογικός τουρισμός

Σε αυτό το είδος τουρισμού λαμβάνουν μέρος άτομα όπου σέβονται τη φύση και συνήθως οι ταξιδιώτες αυτοί προσπαθούν να μην μολύνουν το περιβάλλον.

- Λαϊκός τουρισμός

Ο λαϊκός τουρισμός είναι ένα είδος τουρισμού όπου πραγματοποιείται συνήθως από άτομα με χαμηλό εισόδημα. Ταξιδεύουν τις περισσότερες φορές οικογενειακά και με δικό τους μεταφορικό μέσο.

- Κοσμοπολίτικος τουρισμός

Ο κοσμοπολίτικος τουρισμός πραγματοποιείται από άτομα υψηλής εισοδηματικής τάξης, από αριστοκράτες όπου παραθερίζουν από το ένα κοσμοπολίτικο κέντρο στο άλλο.

- Τουρισμός αναπήρων

Στο συγκεκριμένο τουρισμό λαμβάνουν μέρος άτομα με ειδικές ανάγκες και πραγματοποιείται σε εγκαταστάσεις με εξοπλισμό για την εξυπηρέτηση τουριστών με κινητικά προβλήματα.

2.4 Ο τουρισμός στην Ελλάδα

Ο τουρισμός στην Ελλάδα είναι ένας ανταγωνιστικός και δυναμικός κλάδος της οικονομίας σε διεθνές επίπεδο. Η Ελλάδα έως και σήμερα είναι ένας από του δημοφιλέστερους προορισμούς των ταξιδιωτών, στο οποίο την βοηθάει η γεωγραφική της θέση, το κλίμα της, η ομορφιά της κ.α.

Στο πίνακα 1.1 βλέπουμε τις διεθνείς αφίξεις αλλοδαπών στα κυριότερα αεροδρόμια της Ελλάδας για το έτος 2017.

ΠΟΛΗ	ΑΦΙΞΕΙΣ
Αθήνα	4.797.365
Θεσσαλονίκη	1.929.916
Ρόδος	2.178.663
Κως	1.027.342
Κάρπαθος	81.812
Ηράκλειο	3.147.704
Χανιά	1.093.702
Κέρκυρα	1.300.189
Ζάκυνθος	781.093
Κεφαλονιά	266.549

Άκτιο	271.012
Μύκονος	346.458
Σαντορίνη	431.120
Άραξος	76.931
Καλαμάτα	121.609
Σάμος	115.603
Σκιάθος	182.459
Καβάλα	129.095
Μυτιλήνη	43.948
ΣΥΝΟΛΟ	18.322.570

Πίνακας 1.1: Διεθνείς αφίξεις αλλοδαπών στα κυριότερα αεροδρόμια της Ελλάδας 2017

Παρατηρούμε ότι το έτος 2017 έφτασαν στην Ελλάδα αεροπορικός 18.322.570 αλλοδαποί τουρίστες. Οι περισσότερες αφίξεις παρατηρήθηκαν στην Αθήνα, στο Ηράκλειο, στην Ρόδο, στην Θεσσαλονίκη, στην Κέρκυρα και στα Χανιά όπως φαίνεται στο παραπάνω πίνακα. Επίσης σύμφωνα με στατιστικές υπηρεσίες, είχαμε στη χώρα μας το έτος 2017, 8.788.122 οδικές αφίξεις ενώ ο συνολικός αριθμός τουριστών έφτασε τα 27,2 εκατομμύρια. Το όραμα του Συνδέσμου Ελληνικών Τουριστικών Επιχειρήσεων (ΣΕΤΕ) έγινε ακόμη πιο ορατό, δηλαδή η Ελλάδα να λάβει μία θέση στη δεκάδα.

Στη συνέχεια θα δούμε τις διεθνείς τουριστικές αφίξεις την πενταετία 2012-2016 σε διάφορες τουριστικές χώρες.

Διεθνείς τουριστικές αφίξεις την πενταετία 2012 - 2016 (σε εκατ.)

	Χώρα	2016	2015	2014	2013	2012
1.	Γαλλία	82,6	84,5	83,7	84,7	83,0
2.	ΗΠΑ	75,6	77,5	74,8	69,8	66,7
3.	Ισπανία	75,6	68,5	65,0	60,7	57,5
4.	Κίνα	59,3	56,9	55,6	55,7	57,7
5.	Ιταλία	52,4	50,7	48,6	47,7	46,4
6.	Ηνωμ. Βασίλειο	35,8	34,4	32,6	31,2	29,3
7.	Γερμανία	35,6	35,0	33,0	31,5	30,4
8.	Μεξικό	35,0	32,1	29,1	23,7	23,4
9.	Ταϊλάνδη	32,6	29,9	24,8	26,5	22,4
10.	Αυστρία	28,1	26,7	25,3	24,8	24,2
11.	Μαλαισία	26,7	25,7	27,4	25,7	25,0
12.	Χονγκ Κονγκ	26,5	26,7	27,8	25,7	23,8
13.	Τουρκία	25,3	39,5	39,8	37,8	35,7
14.	Ελλάδα	24,8	23,6	22,0	17,9	15,5
15.	Ρωσία	24,5	26,9	29,8	28,4	25,7

Εικόνα 1.1: Διεθνείς τουριστικές αφίξεις την πενταετία 2012-2016

Πηγή:[http://www.naftemporiki.gr]

Βλέπουμε ότι στην Ελλάδα από το 2012 έως το 2016 έχει αυξηθεί ο τουρισμός της σχεδόν 10 εκατομμύρια. Στην πρώτη τριάδα βρίσκεται η Γαλλία, οι Ηνωμένες Πολιτείες της Αμερικής και η Ισπανία όπου το 2016 είχαν γύρω στα 80 εκατομμύρια τουρίστες.

Είναι σημαντικό να αναφέρουμε τα στοιχεία του ελληνικού τουριστικού κυκλώματος τα οποία προσφέρουν υπηρεσίες σε σχεδόν 27 εκατομμύρια τουρίστες. Στην Ελλάδα σύμφωνα με το ξενοδοχειακό επιμελητήριο υπάρχουν 9.783 ξενοδοχεία τα οποία διαθέτουν 414.127 δωμάτια και 806.045 κλίνες. Εκτός από τα ξενοδοχειακά καταλύματα προστίθεται και η δυναμικότητα της παραξενοδοχείας όπου τα τελευταία χρόνια αυξάνεται ραγδαία. Τέτοιες επιχειρήσεις είναι η Airbnb κ.α. όπου προσφέρουν υπηρεσίες ενοικίασης καταλυμάτων.

Από τους παρακάτω πίνακες θα δούμε τον αριθμό των ενοικιαζόμενων δωματίων σε 13 περιφέρειες της Ελλάδας σύμφωνα με το Ινστιτούτο του Συνδέσμου Ελληνικών Τουριστικών Επιχειρήσεων.

ΠΕΡΙΦΕΡΕΙΑ ΑΝΑΤΟΛΙΚΗΣ ΜΑΚΕΔΟΝΙΑΣ ΚΑΙ ΘΡΑΚΗΣ / EASTERN MACEDONIA & THRACE REGION Ενοικιαζόμενα δωμάτια 2017 / ROOMS FOR RENT 2017								
Περιφερειακή Ενότητα	**Regional Area**			**4K**	**3K**	**2K**	**1K**	**Σύνολο**
Δράμας	**Drama**	Μονάδες	Units	2	3	7	0	**12**
		Δωμάτια	Rooms	15	20	47	0	**82**
		Κλίνες	Guest beds	38	43	110	0	**191**
Έβρου	**Evros**	Μονάδες	Units	2	18	25	45	**90**
		Δωμάτια	Rooms	11	120	186	270	**587**
		Κλίνες	Guest beds	27	281	388	559	**1.255**
Θάσου	**Thassos**	Μονάδες	Units	16	154	567	171	**908**
		Δωμάτια	Rooms	181	1.143	3.597	962	**5.883**
		Κλίνες	Guest beds	404	2.877	8.268	2.048	**13.597**
Καβάλας	**Kavala**	Μονάδες	Units	4	33	55	19	**111**
		Δωμάτια	Rooms	27	235	321	102	**685**
		Κλίνες	Guest beds	68	666	866	270	**1.870**
Ξάνθης	**Xanthi**	Μονάδες	Units	1	1	3	2	**7**
		Δωμάτια	Rooms	6	5	16	16	**43**
		Κλίνες	Guest beds	12	11	36	29	**88**
Ροδόπης	**Rodopi**	Μονάδες	Units	0	4	6	1	**11**
		Δωμάτια	Rooms	0	38	23	5	**66**
		Κλίνες	Guest beds	0	82	94	11	**187**
Σύνολο	**Total**	**Μονάδες**	**Units**	**25**	**213**	**663**	**238**	**1.139**
		Δωμάτια	**Rooms**	**240**	**1.561**	**4.190**	**1.355**	**7.346**
		Κλίνες	**Guest beds**	**549**	**3.960**	**9.762**	**2.917**	**17.188**

Πίνακας 1.2: Ενοικιαζόμενα δωμάτια στη περιφέρεια ανατολικής Μακεδονίας και Θράκης.

ΠΕΡΙΦΕΡΕΙΑ ΑΤΤΙΚΗΣ / ATTICA REGION Ενοικιαζόμενα δωμάτια 2017 / ROOMS FOR RENT 2017								
Περιφερειακή Ενότητα	**Regional Area**			**4K**	**3K**	**2K**	**1K**	**Σύνολο**
Ανατολική Αττική	**East Attica**	Μονάδες	Units	4	12	5	1	**22**
		Δωμάτια	Rooms	28	112	28	13	**181**
		Κλίνες	Guest beds	67	288	62	19	**436**
Βόρειος τομέας Αθηνών	**North Athens**	Μονάδες	Units	12	2	1	0	**15**
		Δωμάτια	Rooms	149	16	11	0	**176**
		Κλίνες	Guest beds	308	40	28	0	**376**
Δυτική Αττική	**West Attica**	Μονάδες	Units	1	2	1	0	**4**
		Δωμάτια	Rooms	5	15	3	0	**23**
		Κλίνες	Guest beds	20	30	7	0	**57**
Δυτικός τομέας Αθηνών	**West Athens**	Μονάδες	Units	1	0	0	0	**1**
		Δωμάτια	Rooms	9	0	0	0	**9**
		Κλίνες	Guest beds	17	0	0	0	**17**
Κεντρικός τομέας Αθηνών	**Central Athens**	Μονάδες	Units	43	1	4	3	**51**
		Δωμάτια	Rooms	472	15	21	22	**530**
		Κλίνες	Guest beds	1.168	19	70	44	**1.301**
Νήσων	**Islands**	Μονάδες	Units	53	113	303	128	**597**
		Δωμάτια	Rooms	471	1.053	2.227	725	**4.476**
		Κλίνες	Guest beds	1.092	2.359	4.884	1.615	**9.950**
Νότιος τομέας Αθηνών	**South Athens**	Μονάδες	Units	7	0	0	0	**7**
		Δωμάτια	Rooms	47	0	0	0	**47**
		Κλίνες	Guest beds	119	0	0	0	**119**
Πειραιώς	**Piraeus**	Μονάδες	Units	2	3	25	30	**60**
		Δωμάτια	Rooms	13	22	155	177	**367**
		Κλίνες	Guest beds	31	50	318	357	**756**
ΣΥΝΟΛΟ	**TOTAL**	**Μονάδες**	**Units**	**123**	**133**	**339**	**162**	**757**
		Δωμάτια	**Rooms**	**1.194**	**1.233**	**2.445**	**937**	**5.809**
		Κλίνες	**Guest beds**	**2.822**	**2.786**	**5.369**	**2.035**	**13.012**

Πίνακας 1.3: Ενοικιαζόμενα δωμάτια στη περιφέρεια Αττικής.

ΠΕΡΙΦΕΡΕΙΑ ΒΟΡΕΙΟΥ ΑΙΓΑΙΟΥ / NORTH AEGEAN REGION Ενοικιαζόμενα δωμάτια 2017 / ROOMS FOR RENT 2017								
Περιφερειακή Ενότητα	**Regional Area**			**4K**	**3K**	**2K**	**1K**	**Σύνολο**
Ικαρίας	**Icaria**	Μονάδες	Units	0	10	74	21	**105**
		Δωμάτια	Rooms	0	51	524	145	**720**
		Κλίνες	Guest beds	0	123	1.089	288	**1.500**
Λέσβου	**Lesvos**	Μονάδες	Units	4	56	251	119	**430**
		Δωμάτια	Rooms	22	406	1.676	618	**2.722**
		Κλίνες	Guest beds	65	988	3.670	1.315	**6.038**
Λήμνου	**Limnos**	Μονάδες	Units	3	35	42	3	**83**
		Δωμάτια	Rooms	8	264	276	20	**568**
		Κλίνες	Guest beds	28	660	624	49	**1.361**
Σάμου	**Samos**	Μονάδες	Units	3	29	278	11	**321**
		Δωμάτια	Rooms	27	293	1.877	75	**2.272**
		Κλίνες	Guest beds	70	679	4.260	163	**5.172**
Χίου	**Chios**	Μονάδες	Units	2	33	64	3	**102**
		Δωμάτια	Rooms	13	212	403	28	**656**
		Κλίνες	Guest beds	36	530	952	55	**1.573**
Σύνολο	**Total**	**Μονάδες**	**Units**	**12**	**163**	**709**	**157**	**1.041**
		Δωμάτια	**Rooms**	**70**	**1.226**	**4.756**	**886**	**6.938**
		Κλίνες	**Guest beds**	**199**	**2.980**	**10.595**	**1.870**	**15.644**

Πίνακας 1.4: Ενοικιαζόμενα δωμάτια στη περιφέρεια Βορείου Αιγαίου.

ΠΕΡΙΦΕΡΕΙΑ ΔΥΤΙΚΗΣ ΕΛΛΑΔΑΣ / WESTERN GREECE REGION Ενοικιαζόμενα δωμάτια 2017 / ROOMS FOR RENT 2017								
Περιφερειακή Ενότητα	**Regional Area**			**4K**	**3K**	**2K**	**1K**	**Σύνολο**
Αιτωλοακαρνανία	**Aitoloakarnania**	Μονάδες	Units	14	25	25	16	**80**
		Δωμάτια	Rooms	99	186	168	90	**543**
		Κλίνες	Guest beds	230	411	369	211	**1.221**
Αχαϊα	**Achaia**	Μονάδες	Units	6	37	37	16	**96**
		Δωμάτια	Rooms	35	277	249	87	**648**
		Κλίνες	Guest beds	103	641	602	205	**1.551**
Ηλεία	**Ilia**	Μονάδες	Units	1	40	44	12	**97**
		Δωμάτια	Rooms	7	268	227	53	**555**
		Κλίνες	Guest beds	12	622	532	122	**1.288**
Σύνολο	**Total**	**Μονάδες**	**Units**	**21**	**102**	**106**	**44**	**273**
		Δωμάτια	**Rooms**	**141**	**731**	**644**	**230**	**1.746**
		Κλίνες	**Guest beds**	**345**	**1.674**	**1.503**	**538**	**4.060**

Πίνακας 1.5: Ενοικιαζόμενα δωμάτια στη περιφέρεια Δυτικής Ελλάδας.

ΠΕΡΙΦΕΡΕΙΑ ΔΥΤΙΚΗΣ ΜΑΚΕΔΟΝΙΑΣ / WESTERN MACEDONIA REGION Ενοικιαζόμενα δωμάτια 2017 / ROOMS FOR RENT 2017								
Περιφερειακή Ενότητα	Regional Area			4K	3K	2K	1K	Σύνολο
Γρεβενά	Grevena	Μονάδες	Units	8	8	20	2	**38**
		Δωμάτια	Rooms	57	62	146	10	**275**
		Κλίνες	Guest beds	113	126	342	19	**600**
Καστοριά	Kastoria	Μονάδες	Units	8	19	8	0	**35**
		Δωμάτια	Rooms	46	133	48	0	**227**
		Κλίνες	Guest beds	102	304	107	0	**513**
Κοζάνη	Kozani	Μονάδες	Units	1	5	1	0	**7**
		Δωμάτια	Rooms	9	35	10	0	**54**
		Κλίνες	Guest beds	17	73	15	0	**105**
Φλώρινα	Florina	Μονάδες	Units	4	9	13	3	**29**
		Δωμάτια	Rooms	26	65	79	14	**184**
		Κλίνες	Guest beds	60	144	188	33	**425**
Σύνολο	Total	Μονάδες	Units	21	41	42	5	109
		Δωμάτια	Rooms	138	295	283	24	740
		Κλίνες	Guest beds	292	647	652	52	1.643

Πίνακας 1.6: Ενοικιαζόμενα δωμάτια στη περιφέρεια δυτικής Μακεδονίας.

ΠΕΡΙΦΕΡΕΙΑ ΗΠΕΙΡΟΥ / EPIRUS REGION Ενοικιαζόμενα δωμάτια 2017 / ROOMS FOR RENT 2017								
Περιφερειακή Ενότητα	Regional Area			4K	3K	2K	1K	Σύνολο
Άρτας	Arta	Μονάδες	Units	6	7	6	0	**19**
		Δωμάτια	Rooms	41	41	28	0	**110**
		Κλίνες	Guest beds	100	95	59	0	**254**
Θεσπρωτίας	Thesprotia	Μονάδες	Units	24	118	131	55	**328**
		Δωμάτια	Rooms	170	544	677	272	**1.663**
		Κλίνες	Guest beds	419	1.438	1.607	577	**4.041**
Ιωαννίνων	Ioannina	Μονάδες	Units	23	91	85	20	**219**
		Δωμάτια	Rooms	158	587	378	93	**1.216**
		Κλίνες	Guest beds	364	1.329	864	210	**2.767**
Πρέβεζας	Preveza	Μονάδες	Units	80	262	302	112	**756**
		Δωμάτια	Rooms	829	1.700	1.836	601	**4.966**
		Κλίνες	Guest beds	2.074	4.340	4.243	1.360	**12.017**
Σύνολο	Total	Μονάδες	Units	133	478	524	187	1.322
		Δωμάτια	Rooms	1.198	2.872	2.919	966	7.955
		Κλίνες	Guest beds	2.957	7.202	6.773	2.147	19.079

Πίνακας 1.7: Ενοικιαζόμενα δωμάτια στη περιφέρεια Ηπείρου.

ΠΕΡΙΦΕΡΕΙΑ ΘΕΣΣΑΛΙΑΣ / THESSALY REGION Ενοικιαζόμενα δωμάτια 2017 / ROOMS FOR RENT 2017								
Περιφερειακή Ενότητα	Regional Area			4K	3K	2K	1K	Σύνολο
Καρδίτσα	Karditsa	Μονάδες	Units	5	20	32	8	**65**
		Δωμάτια	Rooms	33	121	156	49	**359**
		Κλίνες	Guest beds	80	298	383	103	**864**
Λάρισα	Larissa	Μονάδες	Units	6	29	139	55	**229**
		Δωμάτια	Rooms	48	205	1.024	363	**1.640**
		Κλίνες	Guest beds	145	458	2.146	740	**3.489**
Μαγνησία	Magnesia	Μονάδες	Units	36	310	737	293	**1.376**
		Δωμάτια	Rooms	307	1.802	4.870	1.858	**8.837**
		Κλίνες	Guest beds	838	4.528	11.290	4.134	**20.790**
Τρίκαλα	Trikala	Μονάδες	Units	9	33	36	15	**93**
		Δωμάτια	Rooms	72	243	187	76	**578**
		Κλίνες	Guest beds	179	558	461	169	**1.367**
Σύνολο	**Total**	**Μονάδες**	**Units**	**56**	**392**	**944**	**371**	**1.763**
		Δωμάτια	**Rooms**	**460**	**2.371**	**6.237**	**2.346**	**11.414**
		Κλίνες	**Guest beds**	**1.242**	**5.842**	**14.280**	**5.146**	**26.510**

Πίνακας 1.8: Ενοικιαζόμενα δωμάτια στη περιφέρεια Θεσσαλίας.

ΠΕΡΙΦΕΡΕΙΑ ΙΟΝΙΩΝ ΝΗΣΩΝ / IONIAN ISLANDS REGION Ενοικιαζόμενα δωμάτια 2017 / ROOMS FOR RENT 2017								
Περιφερειακή Ενότητα	Regional Area			4K	3K	2K	1K	Σύνολο
Ζακύνθου	Zante	Μονάδες	Units	54	183	484	290	**1.011**
		Δωμάτια	Rooms	557	1.630	3.498	1.918	**7.603**
		Κλίνες	Guest beds	1.306	3.643	7.631	4.063	**16.643**
Ιθάκης	Ithaki	Μονάδες	Units	0	24	78	26	**128**
		Δωμάτια	Rooms	0	119	300	91	**510**
		Κλίνες	Guest beds	0	305	816	210	**1.331**
Κέρκυρας	Corfu	Μονάδες	Units	68	271	618	625	**1.582**
		Δωμάτια	Rooms	489	1.449	3.294	2.837	**8.069**
		Κλίνες	Guest beds	1.317	3.984	8.543	7.336	**21.180**
Κεφαλληνίας	Kefalonia	Μονάδες	Units	15	151	511	122	**799**
		Δωμάτια	Rooms	131	1.271	3.294	729	**5.425**
		Κλίνες	Guest beds	301	3.137	7.753	1.687	**12.878**
Λευκάδας	Lefkada	Μονάδες	Units	17	177	366	138	**698**
		Δωμάτια	Rooms	138	1.267	1.997	687	**4.089**
		Κλίνες	Guest beds	356	3.303	4.606	1.515	**9.780**
Σύνολο	**Total**	**Μονάδες**	**Units**	**154**	**806**	**2.057**	**1.201**	**4.218**
		Δωμάτια	**Rooms**	**1.315**	**5.736**	**12.383**	**6.262**	**25.696**
		Κλίνες	**Guest beds**	**3.280**	**14.372**	**29.349**	**14.811**	**61.812**

Πίνακας 1.9: Ενοικιαζόμενα δωμάτια στη περιφέρεια Ιόνιων Νήσων.

ΠΕΡΙΦΕΡΕΙΑ ΚΕΝΤΡΙΚΗΣ ΜΑΚΕΔΟΝΙΑΣ / CENTRAL MACEDONIA REGION Ενοικιαζόμενα δωμάτια 2017 / ROOMS FOR RENT 2017								
Περιφερειακή Ενότητα	**Regional Area**			**4K**	**3K**	**2K**	**1K**	**Σύνολο**
Ημαθίας	**Imathia**	Μονάδες	Units	12	8	2	0	**22**
		Δωμάτια	Rooms	89	45	14	0	**148**
		Κλίνες	Guest beds	197	106	33	0	**336**
Θεσσαλονίκης	**Thessaloniki**	Μονάδες	Units	12	106	282	110	**510**
		Δωμάτια	Rooms	127	1.064	2.144	717	**4.052**
		Κλίνες	Guest beds	324	2.272	5.141	1.830	**9.567**
Κιλκίς	**Kilkis**	Μονάδες	Units	1	1	2	0	**4**
		Δωμάτια	Rooms	8	7	8	0	**23**
		Κλίνες	Guest beds	16	14	31	0	**61**
Πέλλας	**Pellas**	Μονάδες	Units	8	55	32	1	**96**
		Δωμάτια	Rooms	51	324	161	3	**539**
		Κλίνες	Guest beds	119	733	351	6	**1.209**
Πιερίας	**Pierias**	Μονάδες	Units	6	184	457	436	**1.083**
		Δωμάτια	Rooms	56	1.875	3.853	3.118	**8.902**
		Κλίνες	Guest beds	119	3.785	8.858	7.171	**19.933**
Σερρών	**Serres**	Μονάδες	Units	2	9	1	0	**12**
		Δωμάτια	Rooms	17	62	4	0	**83**
		Κλίνες	Guest beds	32	138	8	0	**178**
Χαλκιδικής	**Xalkidiki**	Μονάδες	Units	61	616	1.077	526	**2.280**
		Δωμάτια	Rooms	516	4.318	7.032	2.954	**14.820**
		Κλίνες	Guest beds	1.325	10.164	17.154	6.933	**35.576**
Σύνολο	**Total**	**Μονάδες**	**Units**	**102**	**979**	**1.853**	**1.073**	**4.007**
		Δωμάτια	**Rooms**	**864**	**7.695**	**13.216**	**6.792**	**28.567**
		Κλίνες	**Guest beds**	**2.132**	**17.212**	**31.576**	**15.940**	**66.860**

Πίνακας 1.10: Ενοικιαζόμενα δωμάτια στη περιφέρεια κεντρικής Μακεδονίας.

ΠΕΡΙΦΕΡΕΙΑ ΚΡΗΤΗΣ / CRETE REGION Ενοικιαζόμενα δωμάτια 2017 / ROOMS FOR RENT 2017								
Περιφερειακή Ενότητα	**Regional Area**			**4K**	**3K**	**2K**	**1K**	**Σύνολο**
Ηρακλείου	**Heraklion**	Μονάδες	Units	66	349	360	101	**876**
		Δωμάτια	Rooms	663	3.975	3.310	719	**8.667**
		Κλίνες	Guest beds	1.605	8.441	6.943	1.503	**18.492**
Λασιθίου	**Lasithi**	Μονάδες	Units	21	115	222	130	**488**
		Δωμάτια	Rooms	187	837	1.361	656	**3.041**
		Κλίνες	Guest beds	431	2.018	3.142	1.468	**7.059**
Ρεθύμνου	**Rethymno**	Μονάδες	Units	17	220	330	113	**680**
		Δωμάτια	Rooms	200	1.415	1.942	639	**4.196**
		Κλίνες	Guest beds	459	3.662	4.494	1.439	**10.054**
Χανίων	**Chania**	Μονάδες	Units	38	521	613	153	**1.325**
		Δωμάτια	Rooms	172	3.122	3.872	951	**8.117**
		Κλίνες	Guest beds	466	8.163	9.167	2.162	**19.958**
Σύνολο	**Total**	**Μονάδες**	**Units**	**142**	**1.205**	**1.525**	**497**	**3.369**
		Δωμάτια	**Rooms**	**1.222**	**9.349**	**10.485**	**2.965**	**24.021**
		Κλίνες	**Guest beds**	**2.961**	**22.284**	**23.746**	**6.572**	**55.563**

Πίνακας 1.11: Ενοικιαζόμενα δωμάτια στη περιφέρεια Κρήτης.

Περιφερειακή Ενότητα	Regional Area			4K	3K	2K	1K	Σύνολο
ΠΕΡΙΦΕΡΕΙΑ ΝΟΤΙΟΥ ΑΙΓΑΙΟΥ / SOUTH AEGEAN REGION								
Ενοικιαζόμενα δωμάτια 2017 / ROOMS FOR RENT 2017								
Άνδρος	**Andros**	Μονάδες	Units	7	76	101	21	**205**
		Δωμάτια	Rooms	53	522	635	94	**1.304**
		Κλίνες	Guest beds	140	1.281	1.456	231	**3.108**
Θήρας	**Thira**	Μονάδες	Units	176	447	379	110	**1.112**
		Δωμάτια	Rooms	1.652	3.673	2.858	788	**8.971**
		Κλίνες	Guest beds	4.084	8.573	6.440	1.719	**20.816**
Καλύμνου	**Kalymnos**	Μονάδες	Units	6	76	191	78	**351**
		Δωμάτια	Rooms	41	492	1.104	385	**2.022**
		Κλίνες	Guest beds	117	1.100	2.333	819	**4.369**
Καρπάθου	**Karpathos**	Μονάδες	Units	0	22	46	28	**96**
		Δωμάτια	Rooms	0	161	344	181	**686**
		Κλίνες	Guest beds	0	357	726	361	**1.444**
Κέας - Κύθνου	**Kea - Kythnos**	Μονάδες	Units	4	53	72	26	**155**
		Δωμάτια	Rooms	27	311	433	137	**908**
		Κλίνες	Guest beds	75	808	1.087	323	**2.293**
Κω	**Kos**	Μονάδες	Units	8	54	92	116	**270**
		Δωμάτια	Rooms	74	710	1.022	983	**2.789**
		Κλίνες	Guest beds	162	1.456	2.156	2.033	**5.807**
Μήλου	**Milos**	Μονάδες	Units	16	177	403	116	**712**
		Δωμάτια	Rooms	126	1.199	2.233	467	**4.025**
		Κλίνες	Guest beds	325	2.861	4.931	990	**9.107**
Μυκόνου	**Mykonos**	Μονάδες	Units	67	170	172	35	**444**
		Δωμάτια	Rooms	606	1.512	1.094	159	**3.371**
		Κλίνες	Guest beds	1.500	3.642	2.433	342	**7.917**
Νάξου	**Naxos**	Μονάδες	Units	22	243	369	94	**728**
		Δωμάτια	Rooms	193	1.724	2.706	659	**5.282**
		Κλίνες	Guest beds	465	3.923	6.140	1.431	**11.959**
Πάρου	**Paros**	Μονάδες	Units	42	159	360	105	**666**
		Δωμάτια	Rooms	414	1.400	2.484	592	**4.890**
		Κλίνες	Guest beds	1.135	3.473	5.846	1.335	**11.789**
Ρόδου	**Rhodes**	Μονάδες	Units	61	205	368	237	**871**
		Δωμάτια	Rooms	509	1.972	2.793	1.466	**6.740**
		Κλίνες	Guest beds	1.134	4.318	5.970	3.166	**14.588**
Σύρου	**Syros**	Μονάδες	Units	5	80	170	39	**294**
		Δωμάτια	Rooms	51	552	940	162	**1.705**
		Κλίνες	Guest beds	131	1.306	2.228	364	**4.029**
Τήνου	**Tinos**	Μονάδες	Units	5	92	89	32	**218**
		Δωμάτια	Rooms	35	578	449	116	**1.178**
		Κλίνες	Guest beds	104	1.404	1.061	258	**2.827**
Σύνολο	**Total**	**Μονάδες**	**Units**	**419**	**1.854**	**2.812**	**1.037**	**6.122**
		Δωμάτια	**Rooms**	**3.781**	**14.806**	**19.095**	**6.189**	**43.871**
		Κλίνες	**Guest beds**	**9.372**	**34.502**	**42.807**	**13.372**	**100.053**

Πίνακας 1.12: Ενοικιαζόμενα δωμάτια στη περιφέρεια Νότιου Αιγαίου.

ΠΕΡΙΦΕΡΕΙΑ ΠΕΛΟΠΟΝΝΗΣΟΥ / PELOPONESSE REGION Ενοικιαζόμενα δωμάτια 2017 / ROOMS FOR RENT 2017								
Περιφερειακή Ενότητα	**Regional Area**			**4K**	**3K**	**2K**	**1K**	**Σύνολο**
Αργολίδας	**Argolida**	Μονάδες	Units	28	141	121	9	**299**
		Δωμάτια	Rooms	197	1.088	843	44	**2.172**
		Κλίνες	Guest beds	505	2.594	1.830	106	**5.035**
Αρκαδίας	**Arkadia**	Μονάδες	Units	5	49	132	29	**215**
		Δωμάτια	Rooms	31	315	746	131	**1.223**
		Κλίνες	Guest beds	71	787	1.724	319	**2.901**
Κορίνθου	**Korinthos**	Μονάδες	Units	9	31	44	2	**86**
		Δωμάτια	Rooms	66	203	192	14	**475**
		Κλίνες	Guest beds	194	499	474	30	**1.197**
Λακωνίας	**Lakonia**	Μονάδες	Units	15	115	116	22	**268**
		Δωμάτια	Rooms	154	773	740	104	**1.771**
		Κλίνες	Guest beds	400	1.844	1.794	221	**4.259**
Μεσσηνίας	**Messinia**	Μονάδες	Units	8	81	220	125	**434**
		Δωμάτια	Rooms	57	468	1.126	573	**2.224**
		Κλίνες	Guest beds	156	1.230	3.014	1.427	**5.827**
Σύνολο	**Total**	**Μονάδες**	**Units**	**65**	**417**	**633**	**187**	**1.302**
		Δωμάτια	**Rooms**	**505**	**2.847**	**3.647**	**866**	**7.865**
		Κλίνες	**Guest beds**	**1.326**	**6.954**	**8.836**	**2.103**	**19.219**

Πίνακας 1.13: Ενοικιαζόμενα δωμάτια στη περιφέρεια Πελοποννήσου.

ΠΕΡΙΦΕΡΕΙΑ ΣΤΕΡΕΑΣ ΕΛΛΑΔΑΣ / CENTRAL GREECE REGION Ενοικιαζόμενα δωμάτια 2017 / ROOMS FOR RENT 2017								
Περιφερειακή Ενότητα	**Regional Area**			**4K**	**3K**	**2K**	**1K**	**Σύνολο**
Βοιωτίας	**Voiotia**	Μονάδες	Units	9	26	19	2	**56**
		Δωμάτια	Rooms	49	184	87	18	**338**
		Κλίνες	Guest beds	157	411	209	30	**807**
Έυβοιας	**Evia**	Μονάδες	Units	15	300	388	167	**870**
		Δωμάτια	Rooms	122	2.521	2.689	917	**6.249**
		Κλίνες	Guest beds	298	5.462	5.829	1.966	**13.555**
Ευρυτανίας	**Evritania**	Μονάδες	Units	8	42	53	6	**109**
		Δωμάτια	Rooms	43	282	247	25	**597**
		Κλίνες	Guest beds	91	593	539	56	**1.279**
Φθιώτιδας	**Fthiotida**	Μονάδες	Units	8	34	33	64	**139**
		Δωμάτια	Rooms	62	278	242	333	**915**
		Κλίνες	Guest beds	124	654	515	673	**1.966**
Φωκίδας	**Fokida**	Μονάδες	Units	9	49	42	3	**103**
		Δωμάτια	Rooms	71	329	254	13	**667**
		Κλίνες	Guest beds	159	736	577	29	**1.501**
Σύνολο	**Total**	**Μονάδες**	**Units**	**49**	**451**	**535**	**242**	**1.277**
		Δωμάτια	**Rooms**	**347**	**3.594**	**3.519**	**1.306**	**8.766**
		Κλίνες	**Guest beds**	**829**	**7.856**	**7.669**	**2.754**	**19.108**

Πίνακας 1.14: Ενοικιαζόμενα δωμάτια στη περιφέρεια Στερεά Ελλάδας.

[***Πηγή:*** *ΜΗΤΕ - Επεξεργασία SETE Intelligence/* ***Source:*** *Tourism Enterprise Register - Processing SETE Intelligence*].

2.5 Η εταιρεία Airbnb

Στον τομέα της στέγασης του κλάδου του τουρισμού και της φιλοξενίας τα τελευταία χρόνια, αναδύθηκε ένα νέο επιχειρηματικό μοντέλο, γνωστό ως "οικονομία του διαμερισμού", (Gansky, 2010, Sundararajan, 2013). Τα τελευταία χρόνια, η οικονομία του διαμερισμού έχει κερδίσει δημοτικότητα σε πολλούς επιχειρηματικούς τομείς όπως και στη διαμονή (π.χ. Airbnb). Η ιστοσελίδα της Airbnb έχει πρωτοπορήσει στη χρήση αυτού του επιχειρηματικού μοντέλου ώστε να μπορέσει να φέρει σε επαφή ανθρώπους που διαθέτουν ιδιόκτητα καταλύματα, όπως διαμερίσματα, ολόκληρα δωμάτια και κοινόχρηστα δωμάτια με εκείνους που χρειάζονται να ενοικιάσουν κάποιο είδος καταλύματος για συγκεκριμένο χρονικό διάστημα μέσω ψηφιακών αγορών. (Botsman and Rogers, 2011, Zervas et al., 2016).
Η εταιρεία Airbnb, που ιδρύθηκε το 2008, αναπτύσσεται ραγδαία και έχει εξυπηρετήσει περισσότερους από 150 εκατομμύρια επισκέπτες. Την τελευταία δεκαετία διαθέτει περισσότερες από 3 εκατομμύρια καταχωρήσεις σε 191 χώρες. Είναι μια νεοϊδρυθείσα εταιρεία που διαφέρει από τα παραδοσιακά ξενοδοχεία όσον αφορά τα συστήματα κρατήσεων, τις εγκαταστάσεις, τα σχέδια ιστοσελίδων και την εξυπηρέτηση πελατών και προσφέρει ταξιδιωτικές υπηρεσίες με ένα καινοτόμο τρόπο. Όλα αυτά συγκεντρώνονται σε μια σωστά δομημένη και καλαίσθητη ιστοσελίδα έτοιμη για εφαρμογή.
Τα κλειδιά για την επιτυχία της Airbnb περιλαμβάνουν χαμηλό κόστος και άμεσες αλληλεπιδράσεις με την τοπική κοινότητα, που προσφέρει στους επισκέπτες μοναδικές εμπειρίες διαμονής.
Η εταιρεία έκανε έσοδα ύψους 900 εκατομμυρίων δολαρίων το 2015 και έχει αγοραία αξία 24 δισεκατομμυρίων δολαρίων.

Πηγή[airbnb.commons.wikipedia]

ΚΕΦΑΛΑΙΟ 3

3. Βασικοί παράγοντες που επηρεάζουν την τιμή στις καταχωρήσεις της Airbnb.

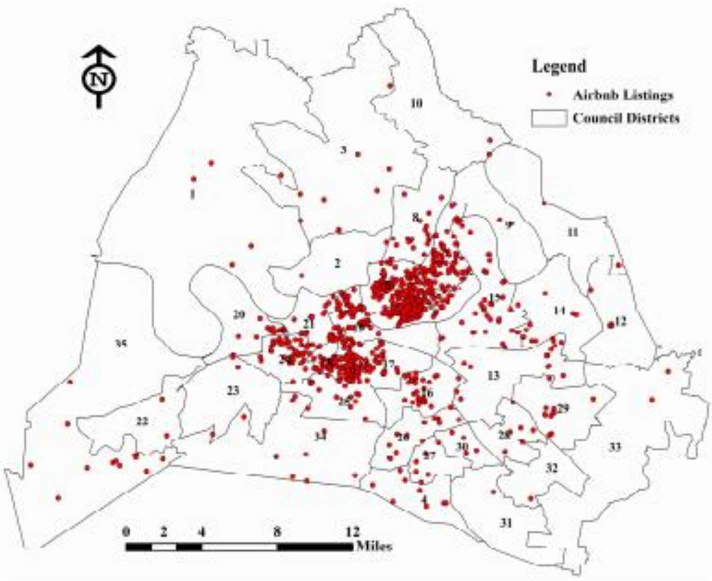
3.1 Εφαρμογή στη πόλη της Nashville.

3.1.1 Περιοχή μελέτης και δεδομένα

Στην συγκεκριμένη έρευνα των (Zhihua Zhang, Rachel JC Che, Lee D. Han και Lu Yang,2017) μελετήθηκε η πόλη της Nashville γνωστή ως Metro Nashville και Davidson County που βρίσκεται στην πολιτεία του Tenesse στις Ηνωμένες Πολιτείες και έγινε ένας προσδιορισμός των βασικών παραγόντων που επηρεάζουν τις τιμές εισαγωγής στο Airbnb. Η κυβέρνηση της πόλης Nashville και Davidson County συγχωνεύθηκε το 1963 και χωρίστηκε σε 35 συνοικίες.
Τα δεδομένα των καταχωρήσεων της Airbnb της πόλης του Metro Nashville συλλέχθηκαν στις 7 Αυγούστου του 2017 από την ιστοσελίδα της Airbnb (https://www.airbnb.com/) η οποία έχει πληροφορίες για τα καταλύματα που είναι καταχωρημένα.

Στη συλλογή των δεδομένων συλλέχθηκαν καταχωρήσεις στις οποίες ο οικοδεσπότης ήταν super host δηλαδή είναι ένας έμπειρος οικοδεσπότης διότι υπήρχαν πολλές καταχωρήσεις που ήταν απλά εισηγμένες και δεν είχαν καμία κράτηση. Έτσι συγκεντρώθηκαν 974 διαφορετικά καταλύματα από τη πόλη του Metro Nashville όπου έγινε και η ανάλυση τους.

Στην εικόνα 3.1 παρουσιάστηκε πως είναι κατανεμημένα τα καταλύματα στη πόλη του Metro Nashville. Παρατηρούμε ότι τα περισσότερα καταλύματα βρίσκονται στο κέντρο της πόλης.



Εικόνα 3.1: Χωρική κατανομή των καταχωρίσεων της Airbnb στο Metro Nashville.

Στην συγκεκριμένη έρευνα επιλέχθηκαν έξι διαφορετικές μεταβλητές για να μελετηθούν από τις οποίες οι πέντε είναι ανεξάρτητες και μια είναι εξαρτημένη. Συγκεκριμένα η Τιμή ορίστηκε ως εξαρτημένη και ως ανεξάρτητες ορίστηκαν οι Κριτικές, η Η-απόσταση, η C-απόσταση, η Ηλικία και η Βαθμολογία όπου φαίνονται στο παρακάτω πινάκα και θα αναλύθηκαν στη συνέχεια.

ΜΕΤΑΒΛΗΤΕΣ	ΠΕΡΙΓΡΑΦΗ
Τιμή	Τιμή ανά διανυκτέρευση σε σχέση με τους μέγιστους φιλοξενούμενους.
Κριτικές	Αριθμός κριτικών.
Η-απόσταση	Η απόσταση από τον πλησιέστερο αυτοκινητόδρομο (km).
C-απόσταση	Η απόσταση από το κέντρο του Nashville (km).
Ηλικία	Μήνες από τη καταχώρηση στο Airbnb
Βαθμολογία	Συνολικό βαθμολογία από τους φιλοξενούμενους.

Πίνακας 3.1: Μεταβλητές και σύντομες περιγραφές

Όπως αναφέρθηκε και παραπάνω ως Τιμή ορίστηκε η τιμή ανά διανυκτέρευση σε σχέση με τους μέγιστους φιλοξενούμενους όπου θα εξεταστεί πώς επηρεάζεται σε σχέση με τις υπόλοιπες ανεξάρτητες μεταβλητές.

Οι Κριτικές είναι ένας πάρα πολύ σημαντικός παράγοντας της τιμής διότι όσες περισσότερες καλές κριτικές έχει ένα κατάλυμα μπορεί να το οδηγήσει σε υψηλότερη τιμή.

Εκτός από τις κριτικές η τοποθεσία είναι ένας ακόμα παράγοντας που επηρεάζει τη τιμή του καταλύματος. Οπότε ορίστηκαν η H-απόσταση και η C-απόσταση όπου αναφέρονται στην απόσταση από το πλησιέστερο αυτοκινητόδρομο και στην απόσταση από το κέντρο του Nashville αντίστοιχα.

Η ηλικία είναι μία μεταβλητή η οποία μπορεί να συμβάλει στο καθορισμό της τιμής, δηλαδή η χρονική περίοδος από την εισαγωγή της στο Airbnb, έτσι αποφασίστηκε να επιλεχθεί η συγκεκριμένη μεταβλητή.

Η βαθμολογία είναι και αυτή μια μεταβλητή όπου επηρεάζει τη τιμή, διότι αν το κατάλυμα έχει αποκτήσει υψηλή βαθμολογία από τους φιλοξενούμενους μπορεί να οδηγήσει σε αύξηση της τιμής.

3.2 Μοντέλα της διαδικασία

Τα μοντέλα που χρησιμοποιήθηκαν για να βρεθεί η σχέση της τιμής με τις μεταβλητές Κριτικές, H-απόσταση, C-απόσταση, Ηλικία και Βαθμολογία είναι το γενικό γραμμικό μοντέλο (GLM) και η γεωγραφικά σταθμισμένη παλινδρόμηση (GWR).

3.2.1 Γενικό γραμμικό μοντέλο (GLM)

Το γενικό γραμμικό μοντέλο είναι μια γενίκευση μοντέλου πολλαπλής γραμμικής παλινδρόμησης στην περίπτωση που υπάρχουν περισσότερες από μία ανεξάρτητες μεταβλητές. Σε ένα γραμμικό μοντέλο η εξαρτημένη μεταβλητή εκφράζεται με ένα σύνολο ανεξάρτητων μεταβλητών. Η μορφή του μοντέλου εμφανίζεται στη συνέχεια.

$$y_i = \beta_0 + \sum_{k=1}^{p} \beta_k X_{ik} + \varepsilon_i$$

Όπου:

y_i δηλώνει την i παρατήρηση της εξαρτημένης μεταβλητής

β_i είναι οι συντελεστές των προγνωστικών που υπολογίζονται στο μοντέλο και

e_i είναι το σφάλμα.

3.2.2 Γεωγραφικά σταθμισμένη παλινδρόμηση (GWR)

Η Γεωγραφικά Σταθμισμένη Παλινδρόμηση (GWR) είναι μια σύγχρονη μέθοδος χωρικής ανάλυσης, η οποία επιτρέπει την εξέταση τοπικών διακυμάνσεων σε χωρικές διεργασίες. Επίσης η GWR αποτελεί ένα ισχυρό εργαλείο ανίχνευσης χωρικά μεταβαλλόμενων σχέσεων στα χωρικά δεδομένα. Η τεχνική επιτρέπει τη βαθμονόμηση ενός τοπικού μοντέλου εκτιμώντας τοπικές παραμέτρους αντί μιας ολικής παραμέτρου για κάθε μεταβλητή. Αυτό είναι δυνατό ορίζοντας και βαθμονομώντας ένα ξεχωριστό μοντέλο γύρω από κάθε παρατήρηση i με γεωγραφικές συντεταγμένες (u_i, v_i) στο οποίο περιλαμβάνονται όλες ή μερικές από τις n παρατηρήσεις του ολικού μοντέλου σταθμισμένες με μια μέθοδο στάθμισης που συνήθως είναι συνάρτηση της απόστασης από το i. Ως τοπικό μοντέλο ορίζεται το σύνολο αυτών των ξεχωριστών μοντέλων που συνήθως είναι τόσα, όσες και οι παρατηρήσεις του ολικού μοντέλου.

Το μοντέλο της GWR παρουσιάζεται στη συνέχεια.

$$y_i = \beta_0(u_i, v_i) + \sum_{k=1}^{m} \beta_\kappa(u_i, v_i) X_{ik} + \varepsilon_i$$

Το y_i δηλώνει την εξαρτημένη μεταβλητή στη θέση i.
Το X_{ik} δηλώνει την τιμή της k ερμηνευτικής μεταβλητής σε θέση i.
Το β = (β 0 , β 1 , β 2 , ... β m) είναι οι συντελεστές των προβλέψεων που υπολογίζονται στο μοντέλο.
Το (u i , v i) δηλώνει τις συντεταγμένες της θέσης i και ε_i είναι το τυχαίο σφάλμα στη θέση i.

Το β στη θέση i εκτιμάται χρησιμοποιώντας το σταθμισμένο ελάχιστο τετράγωνο και μπορεί να εκφραστεί ως

$$\beta_i = [X^T W_i X]^{-1} X^T W_i Y$$

Όπου

$$W_i = \begin{bmatrix} w_{i1} & 0 & 0 \\ 0 & \ddots & 0 \\ 0 & 0 & w_{in} \end{bmatrix}$$

Ο πίνακας W_i είναι ένας πίνακας μήτρας (n x n) όπου τα στοιχεία της διαγωνίου είναι τα βάρη των παρατηρήσεων για το σημείο i και τα στοιχεία εκτός της διαγωνίου είναι μηδέν. Το w_{ij} είναι το βάρος της παρατήρησης j στη βαθμονόμηση του μοντέλου στο σημείο i.

Ο πίνακας X των ανεξάρτητων μεταβλητών φαίνεται στη συνέχεια.

$$X = \begin{bmatrix} 1 & x_{11} \cdots & x_{m1} \\ \vdots & \vdots \cdots & \vdots \\ 1 & x_{1n} \cdots & x_{mn} \end{bmatrix}$$

Το διάνυσμα Y αναπαριστά τις παρατηρήσεις της εξαρτημένης μεταβλητής.

$$Y = \begin{bmatrix} y_1 \\ \vdots \\ yn \end{bmatrix}$$

Για το πίνακα μήτρας W_i μπορούν να εφαρμοστούν διάφοροι μέθοδοι. Όπως η μέθοδος Gauss και η μέθοδος bi-square όμως στη συγκεκριμένη περίπτωση εφαρμόστηκε η μέθοδος Gauss η οποία εκφράζεται ως

$$W_{j(i)} = \exp(-0.5\left(\frac{d_{ij}}{b}\right)^2)$$

Όπου

d_{ij} δηλώνει την απόσταση μεταξύ της θέσης i και του παρατηρούμενου σημείου j.
b δηλώνει την απόσταση εύρου ζώνης.

Η επιλογή του εύρους ζώνης είναι ιδιαίτερα σημαντική διότι επηρεάζει σε μεγάλο βαθμό τα αποτελέσματα της μεθόδου. Ένας τρόπος είναι να οριστεί από το χρήστη ο οποίος μετά από δοκιμασίες διαφορετικών εύρων ζώνης να μελετήσει την ευαισθησία των αποτελεσμάτων σε αυτά. Όμως, και η μέθοδος GWR μπορεί να υπολογίσει το ιδανικότερο εύρος ζώνης με τη βοήθεια ενός αλγορίθμου που αποτελεί αναπόσπαστη λειτουργία της μεθόδου σε κάθε λογισμικό. Για τον υπολογισμό του ιδανικού εύρους ζώνης, ο αλγόριθμος δοκιμάζει επαναληπτικά διαφορετικές τιμές εύρους ζώνης έως ότου συγκλίνει σε μια τιμή ελαχιστοποιώντας ένα στατιστικό μέτρο. Τα πιο διαδεδομένα στατιστικά μέτρα είναι το Cross-Validation σκορ (CV) και το κριτήριο Akaike Information Criterion (AIC).

3.3 Αποτελέσματα

3.3.1 Εκτίμηση της (GLM)

Με τη συγκεκριμένη μέθοδο έγινε διερεύνηση στις σχέσεις μεταξύ της εξαρτημένης μεταβλητής (Τιμή) και των ανεξάρτητων μεταβλητών (Κριτικές, H-απόσταση, C-απόσταση, Ηλικία και Βαθμολογία). Πρέπει να αναφερθεί ότι στη συγκεκριμένη μελέτη δεν έγινε σύγκριση μεταξύ των τύπων των καταλυμάτων δηλαδή αν είναι ολόκληρο σπίτι/ διαμέρισμα, ιδιωτικό δωμάτιο ή κοινόχρηστο δωμάτιο για αυτό ορίστηκε η εξαρτημένη μεταβλητή Τιμή ως την τιμή με το μέγιστο αριθμό επισκεπτών ανά διανυκτέρευση.

Ο πίνακας 3.2 παρουσιάζει τα αποτελέσματα που διαμορφώθηκαν από το μοντέλο GLM.

Variable	Coefficient	St. Error	t-Value	p-Value	95% Confidence limits	
Σταθερός συντελεστής	56,32561**	3,78106	14,90	<0,0001	48,90558	63,74563
H-απόσταση	-0,36897	0,45514	-0,81	0,4178	-1,26215	0,52421
C-απόσταση	-1,52435**	0,17030	-8,95	<0,0001	-1,85855	-1,19016
Κριτικές	-0,05080**	0,01303	-3,90	0,0001	-0,07637	-0,02524
Ηλικία	0,02233	0,03803	0,59	0,5572	-0,05229	0,09695
Βαθμολογία	-0,07317*	0,03662	-2,00	0,0460	-0,14502	-0,00131
Σημείωση: **= p<0.01, *= p<0.05						

Πίνακας 3.2: Αποτελέσματα μεθόδου (GLM).

Από τη στήλη Coefficient βλέπουμε πως επηρεάζει κάθε μεταβλητή την εξαρτημένη μεταβλητή (Τιμή).

- Η μεταβλητή (H- απόσταση) δηλαδή η απόσταση από το πλησιέστερο αυτοκινητόδρομο έχει αρνητική επίδραση στη τιμή. Βλέπουμε ότι την επηρεάζει αρνητικά κατά 36,9%. Συγκεκριμένα μπορούμε να πούμε ότι μια αύξηση της μονάδας σε απόσταση θα επιφέρει μείωση της τιμής κατά 0,369 δολάρια.

- Η μεταβλητή (C- απόσταση) δηλαδή η απόσταση από το κέντρο του Nashville έχει και αυτή αρνητική επίδραση στη τιμή και μπορούμε να πούμε ότι είναι πολύ μεγαλύτερη από την επίδραση της απόστασης από το πλησιέστερο αυτοκινητόδρομο . Παρατηρούμε ότι επηρεάζεται αρνητική κατά 152,44%. Ειδικότερα μια αύξηση της μονάδας σε απόσταση θα επιφέρει μείωση της τιμής κατά 1,52 δολάρια. Αυτό το αποτέλεσμα φαίνεται λογικό αφού οι επισκέπτες προτιμούν το κατάλυμα όπου θα μείνουν να βρίσκεται στο κέντρο της πόλης.

- Η μεταβλητή (Κριτικές) βλέπουμε ότι επηρεάζει και αυτή αρνητικά την εξαρτημένη μεταβλητή (Τιμή). Το ποσοστό της αρνητικής επίδρασης είναι 5,08%. Οπότε καταλήγουμε ότι μια μονάδα αύξησης του αριθμού των κριτικών θα επιφέρει μείωση ίση με 0,0508 δολάρια στη τιμή. Πρέπει να αναφέρουμε ότι αυτό θα ήταν λογικό αν η κριτική δεν ήταν καλή.

- Η μεταβλητή (Ηλικία) δηλαδή οι μήνες από την καταχώρηση του καταλύματος στο Airbnb βλέπουμε ότι επηρεάζει θετικά την Τιμή. Παρατηρούμε ότι το ποσοστό θετικής επίδρασης είναι 2,23%. Δηλαδή μια μονάδα αύξησης της Ηλικίας προκαλεί αύξηση στη Τιμή κατά 0,022 δολάρια.

- Τέλος η μεταβλητή (Βαθμολογία) παρατηρούμε ότι επιδρά αρνητικά στην εξαρτημένη μεταβλητή (Τιμή). Βλέπουμε ότι την επηρεάζει αρνητικά κατά 7,32%. Ειδικότερα μια αύξηση της μονάδας σε βαθμολογία θα επιφέρει μείωση της τιμής κατά 0,0732 δολάρια.

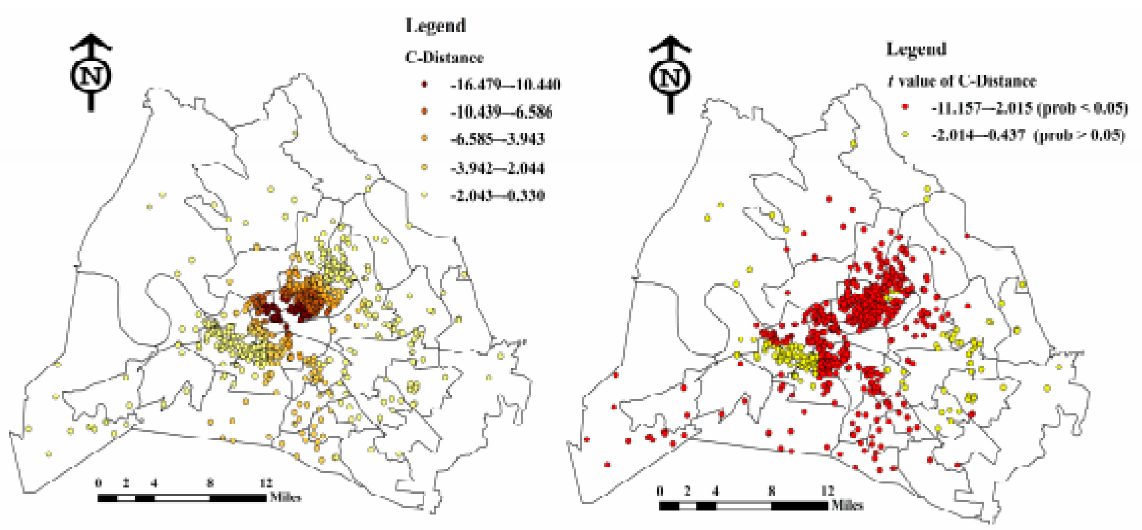
3.3.2 Εκτίμηση της (GWR)

Με το μοντέλο GWR έγινε ανάλυση στη σχέση μεταξύ της τιμής και των περιοχών όπου είναι καταχωρημένα τα καταλύματα. Αναφέρθηκε ότι περιλήφθηκαν επιλεκτικά μερικές περιοχές διότι κάποιες δεν συνδέονταν πλήρως με τη διαδικασία της μοντελοποίησης GWR. Τα δεδομένα που χρησιμοποιήθηκαν στο συγκεκριμένο μοντέλο είναι ίδια με αυτά του μοντέλου GLM. Ο πίνακας 3.3 εμφανίζει κάποιους συντελεστές εκτιμώμενης γεωγραφικής σταθμισμένης παλινδρόμησης

Variable	Mean	Min	First Quartile	Median	Third Quartile	Max
Σταθερός όρος	59.794	26.863	47.879	57.155	71.177	96.336
Κριτικές	-0.063	-0.184	-0.128	-0.016	-0.003	0.021
Βαθμολογία	-0.096	-0.345	-0.167	-0.051	-0.005	0.054
Ηλικία	0.064	-0.118	-0.039	0.013	0.168	0.329
Η-απόσταση	5.580	-1.079	2.292	4.964	7.425	18.403
C-απόσταση	-4.542	-16.479	-7.599	-2.536	-1.487	-0.330

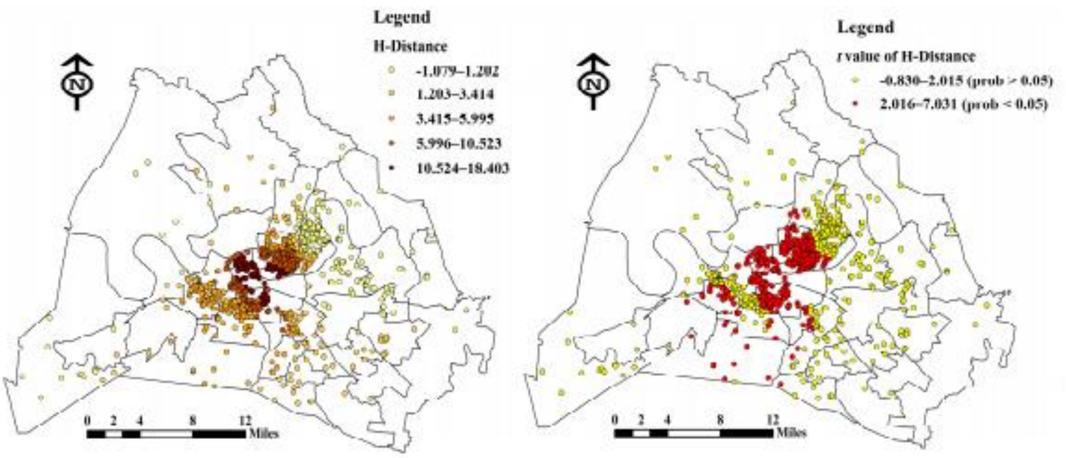
Πίνακας 3.3: Συντελεστές εκτιμώμενης γεωγραφικής σταθμισμένης παλινδρόμησης (GWR).

Στη συνέχεια από τις παρακάτω εικόνες θα γίνει παρουσίαση της επίδρασης των ανεξάρτητων μεταβλητών στην εξαρτημένη μεταβλητή Τιμή. Ωστόσο αναφέρθηκε ότι το μοντέλο GWR είναι προτιμότερο διότι εμφανίζει μια σημαντική χωρική ετερογένεια που υπάρχει στις ανεξάρτητες μεταβλητές. Οι κριτικές, η βαθμολογία και η απόσταση από το κέντρο του Nashville επιδρούν κατά μέσο όρο αρνητικά στη τιμή ενώ η απόσταση από τον κοντινότερο αυτοκινητόδρομο και η ηλικία επιδρούν κατά μέσο όρο θετικά. Στη συνέχεια παρουσιάστηκε πως επηρέασε κάθε ανεξάρτητη μεταβλητή την τιμή αναλόγως την περιοχή.



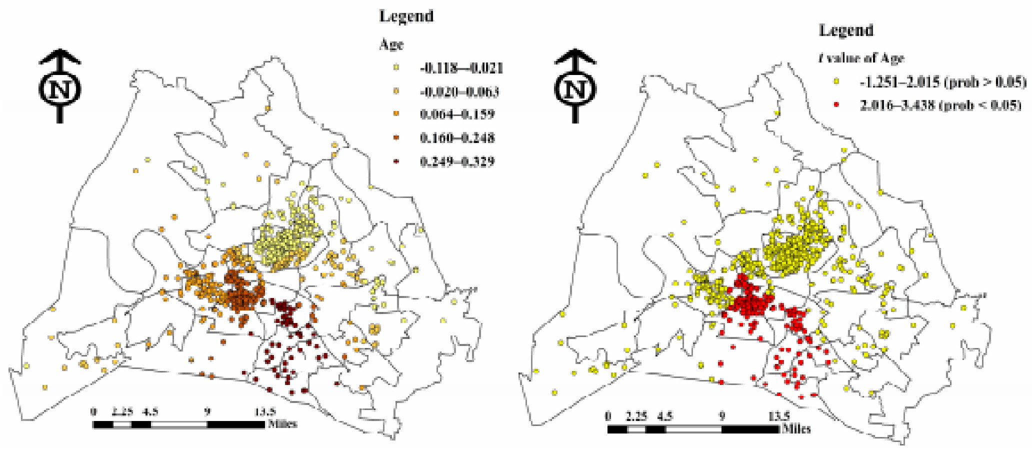
Εικόνα 3.2: (Αριστερά)Χωρική κατανομή των εκτιμώμενων συντελεστών της απόστασης και (δεξιά) των t τιμών της απόστασης από το κέντρο του Nashville.

Στην εικόνα 3.2 εμφανίζεται η σχέση μεταξύ της εξαρτημένης μεταβλητής τη Τιμή και της απόστασης από το κέντρο του Nashville. Οι αρνητικές συσχετίσεις παρατηρούνται στις περισσότερες περιοχές εκτός από τις δυτικές περιοχές 18, 24. Είναι λογικό ότι στις περιοχές κοντά στο κέντρο της πόλης του Nashville όπου προσφέρει άνεση στους ταξιδιώτες ως προς τη πρόσβαση σε αξιοθέατα, συγκοινωνίες και εγκαταστάσεις ψυχαγωγίας να εμφανίζεται υψηλότερη τιμή. Όμως οι συντελεστές δείχνουν μια πτωτική τάση προς το κέντρο της πόλης. Ίσως στις καταχωρήσεις στο κέντρο της πόλης να υπάρχει μια στρατηγική ηγεσία κόστους, όπου να προσφέρουν χαμηλότερη τιμή ώστε να προσελκύσουν περισσότερους πελάτες. Στρατηγική Ηγεσίας Κόστους λοιπόν ακολουθεί μία εταιρεία όταν θέτει ως βασικό στόχο να πετύχει ανταγωνιστικό πλεονέκτημα με την επίτευξη χαμηλότερου κόστους στην παροχή υπηρεσιών. Το πλεονέκτημα της χαμηλότερης τιμής είναι η αύξηση των κρατήσεων που συνεπάγεται με αύξηση των εσόδων. Γενικότερα μια επιχείρηση που επιλέγει να ακολουθήσει αυτή την ανταγωνιστική στρατηγική απευθύνεται στο σύνολο της αγοράς και επιδιώκει επίτευξη μεγάλου όγκου παροχής υπηρεσιών.



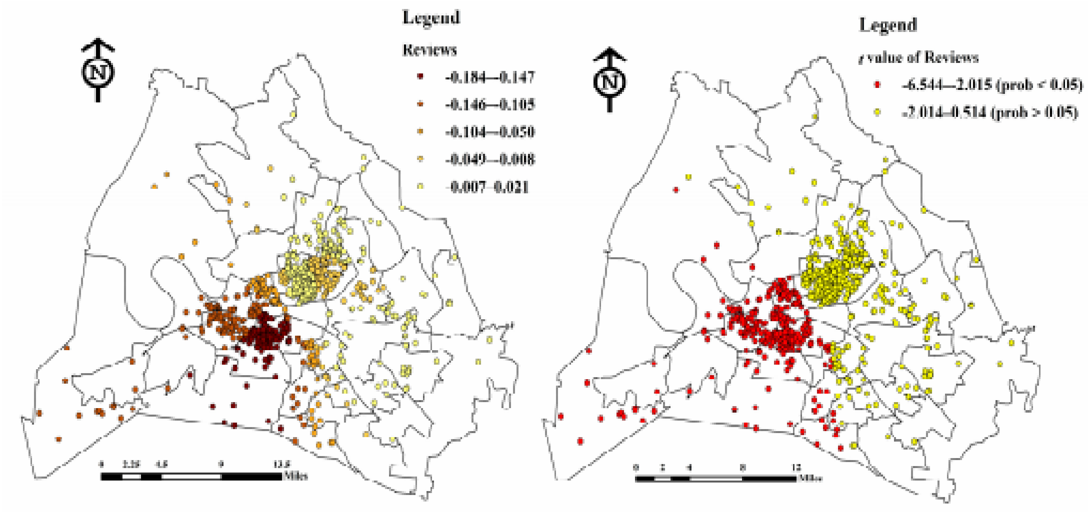
Εικόνα 3.3: (Αριστερά)Χωρική κατανομή των εκτιμώμενων συντελεστών και (δεξιά) των t τιμών της απόστασης από το πλησιέστερο αυτοκινητόδρομο του Nashville.

Στην εικόνα 3.3 εμφανίζεται η σχέση μεταξύ της εξαρτημένης μεταβλητής, τη Τιμή και της απόστασης από το πλησιέστερο αυτοκινητόδρομο του Nashville. Στο κέντρο της πόλης δηλαδή στις περιοχές 5,6,18,19 υπάρχει πολύ θετική επίδραση στη τιμή. Αυτό είναι κατανοητό ότι στις περιοχές μακριά από τους αυτοκινητόδρομους δηλαδή τις περιοχές του κέντρου όπου δεν υπάρχει ηχορύπανση θα εμφανίζεται υψηλότερη τιμή. Άρα στις περιοχές μακριά από το κέντρο της πόλης παρατηρούμε πτωτική τάση στις τιμές σε αντίθεση με τις περιοχές που βρίσκονται στο κέντρο.



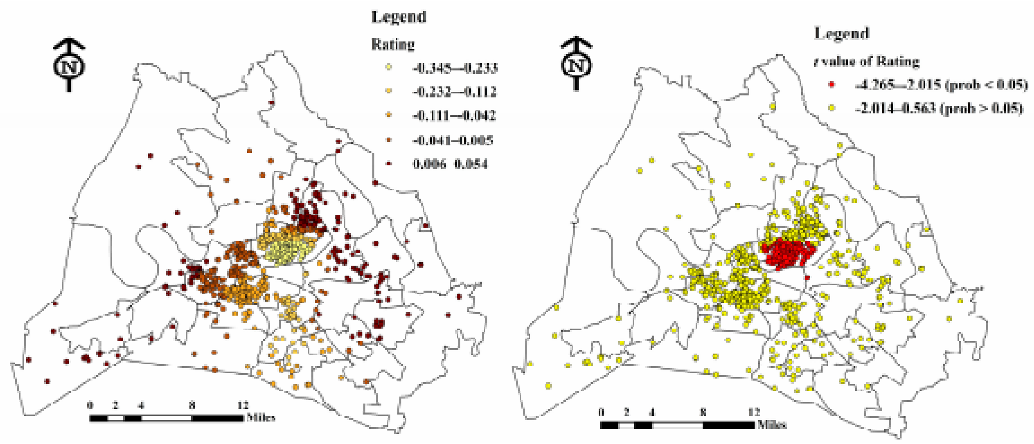
Εικόνα 3.4: (Αριστερά)Χωρική κατανομή των εκτιμώμενων συντελεστών και (δεξιά) των τιμών t της ηλικίας.

Όπως παρατηρείται στην εικόνα 3.4 εμφανίζεται η σχέση μεταξύ της εξαρτημένης μεταβλητής, τη Τιμή και της ηλικίας του καταλύματος. Βλέπουμε ότι η θετική επίδραση στις τιμές υπάρχει στις περιοχές 16-18,25-27,30 δηλαδή στο νότιο Nashville. Άρα μπορούμε να πούμε ότι σε αυτές τις περιοχές όπου τα καταλύματα είχαν εισαχθεί νωρίτερα στο Airbnb εμφανίζονται και υψηλότερες τιμές διότι οι οικοδεσπότες είναι πιο έμπειροι και οι ταξιδιώτες νιώθουν μια ασφάλεια ως προς την ποιότητα των υπηρεσιών.



Εικόνα 3.5: (Αριστερά)Χωρική κατανομή των εκτιμώμενων συντελεστών και (δεξιά) των τιμών t των κριτικών.

Στην εικόνα 3.5 εμφανίζεται η σχέση μεταξύ της εξαρτημένης μεταβλητής, τη Τιμή και των κριτικών. Παρατηρούμε ότι υπάρχει αρνητική επίδραση στη τιμή στις περιοχές 18,20-25,34 δηλαδή στο κέντρο και δυτικά του Nashville. Δηλαδή όσες περισσότερες κριτικές έχει ένα κατάλυμα τόσο χαμηλότερη θα είναι και η τιμή του, όπου έρχεται σε αντίθεση με την έρευνα του (Wang D, Nicolau J.L,2017). Μια εξήγηση που μπορούμε να δώσουμε είναι ότι επειδή οι ταξιδιώτες επιλέγουν χαμηλές τιμές καταλυμάτων έχει ως αποτέλεσμα και οι οικοδεσπότες να μειώνουν τις τιμές. Για αυτό δεν έχει αποδειχθεί ότι οι κριτικές σχετίζονται θετικά με τη τιμή.



Εικόνα 3.6: (Αριστερά)Χωρική κατανομή των εκτιμώμενων συντελεστών και (δεξιά) των τιμών t της βαθμολογίας.

Στην εικόνα 3.6 εμφανίζεται η σχέση μεταξύ της εξαρτημένης μεταβλητής, τη Τιμή και της βαθμολογίας. Βλέπουμε ότι υπάρχει αρνητική επίδραση στη τιμή στις περιοχές 5,6,19 όπου είναι κάποιες μικρές περιοχές στο κεντρικό και βορειοανατολικό Nashville. Αυτό σημαίνει ότι όσο υψηλότερες είναι οι βαθμολογίες τόσο χαμηλότερη είναι και η τιμή όπου έρχεται σε αντίθεση με άλλες μελέτες. Έχει αναφερθεί ότι αυτό μπορεί να οφείλεται στο φαινόμενο ότι οι μέσες βαθμολογίες στην Airbnb να είναι πιο υψηλές από αυτές σε άλλες πλατφόρμες. (Zervas G. , Proserpio D. , Byers, J.W. ,2017). Επί του παρόντος, στη μελέτη αυτή το αποτέλεσμα δεν είναι λογικά διότι τα αποτελέσματα των πορισμάτων των προηγούμενων μελετών δείχνουν ότι η υψηλότερη τιμή οδηγεί σε υψηλότερες βαθμολογίες ικανοποίησης. Όμως μια άλλη πιθανή εξήγηση που μπορούμε να αναφέρουμε είναι ότι η αυξημένη τιμή του καταλύματος δημιουργεί μεγαλύτερες προσδοκίες στο πελάτη με αποτέλεσμα να κρίνει τις υπηρεσίες του καταλύματος πιο αυστηρά. Έτσι μπορεί να δημιουργηθεί χαμηλή βαθμολογία σε καταλύματα με αυξημένη τιμή.

3.4 Συμπεράσματα

Συμπεραίνουμε βάση των αποτελεσμάτων ότι η σχέση μεταξύ τιμής και των υπόλοιπων ανεξάρτητων μεταβλητών είναι στατιστικά σημαντική στις περιοχές 5, 6 και 19 ενώ δεν υπάρχει σημαντική σχέση μεταξύ του χωρικού παράγοντα και των καταχωρήσεων Airbnb. Είναι εύλογο το γεγονός ότι η Airbnb στο Nashville λαμβάνει μέση ή υψηλότερη βαθμολογία λόγω της συγκριτικής χαμηλής τιμής που προσφέρεται στην αγορά, η οποία εκτιμάται από την πλειοψηφία των πελατών που διανυκτέρευσαν σε αυτήν την περιοχή της μελέτης.

3.5 Εφαρμογή σε δείγμα 33 διαφορετικών πόλεων.

3.5.1 Σκοπός.

Σκοπός αυτής της μελέτης είναι να προσδιοριστούν οι παράγοντες που επηρεάζουν την τιμή των καταλυμάτων προς την ενοικίαση στην ψηφιακή αγορά και συγκεκριμένα στο Airbnb.com. Στην παρακάτω έρευνα εξετάζεται δείγμα 180.533 κατοικιών προς ενοικίαση από 33 διαφορετικές πόλεις που είναι καταχωρημένα στο Airbnb.com. Οι συνήθεις αναλύσεις των ελαχίστων τετραγώνων (OLS) και η ποσοτική παλινδρόμηση (QR) χρησιμοποιούνται για τη διερεύνηση των καθοριστικών παραγόντων της τιμής που χωρίζονται στις επόμενες πέντε κατηγορίες:

- ιδιότητες υποδοχής,
- χαρακτηριστικά τοποθεσίας και ιδιοκτησίας,
- παροχές και υπηρεσίες,
- κανόνες ενοικίασης και αξιολογήσεις μέσω διαδικτύου.

Τα αποτελέσματα των ερευνών αυτών έχουν σημαντικές συνέπειες για τη σχεδίαση του υπολογισμού της τιμής ανάλογα τις παροχές που διαθέτει κάθε κατάλυμα, όπως το εργαλείο σύστασης τιμών που πρόσφατα δρομολόγησε η Airbnb.

3.6 Δεδομένα και περιγραφικά μέτρα μεταβλητών.

Στην παρακάτω έρευνα εξετάστηκαν τα αποτελέσματα των 25 μεταβλητών στις ακόλουθες 5 κατηγορίες:

- ✓ ιδιότητες υποδοχής,
- ✓ χαρακτηριστικά τοποθεσίας και ιδιοκτησίας,
- ✓ παροχές και υπηρεσίες,
- ✓ κανόνες ενοικίασης και
- ✓ αξιολογήσεις μέσω διαδικτύου.

Αυτές οι μεταβλητές εμφανίζονται και ορίζονται στον πίνακα 3.4.

Όνομα μεταβλητής	**Μέση τιμή / αναλογία**	**Τυπική απόκλιση**	**Ορισμός**
Τιμή	117.18	127.87	Τιμή ανά διανυκτέρευση στο Airbnb.com.(Μετρήθηκε σε US δολάρια).
Super host	0.09	0.29	Έχει γίνει οικοδεσπότης τουλάχιστον 10 φορές το χρόνο. Διατηρεί ποσοστό ανταπόκρισης 90% ή υψηλότερο. Έλαβε συνολική αξιολόγηση 4,8+. Ολοκλήρωσαν τις επιβεβαιωμένες κρατήσεις τους χωρίς ακύρωση.
Καταγραφές κεντρικών υπολογιστών	5.66	27.48	Ο αριθμός των ενοικιάσεων του καταλύματος που αναγράφονται στην Airbnb.com
Εικόνα προφίλ οικοδεσπότη	0.99	0.04	Ο κεντρικός υπολογιστής διαθέτει μια φωτογραφία προφίλ

Επαλήθευση ταυτότητας κεντρικού υπολογιστή	0.64	0.47	Ο κεντρικός υπολογιστής έχει ολοκληρώσει τις διαδικασίες Verified ID στο Airbnb.com
Απόσταση	6.28	18.83	Η απόσταση (Km) μεταξύ της τοποθεσίας μιας καταχωρημένης μίσθωσης και του κέντρου της πόλης, υπολογιζόμενη με τη χρήση του τύπου "Harversine".
Τύπος διαμονής 1	0.77	0.41	Συνδυασμένοι τύποι καταλυμάτων: Διαμέρισμα, και σοφίτα.
Τύπος διαμονής 2	0.02	0.14	Συνδυασμένοι τύποι καταλυμάτων: διαμονή με πρωινό και κοιτώνες.
Τύπος διαμονής 3	0.20	0.40	Συνδυασμένοι τύποι: σπίτι, σουίτα, αρχοντικό, βίλα, καμπίνα.
Ολόκληρο σπίτι / διαμέρισμα	0.65	0.47	Ολόκληρο σπίτι / διαμέρισμα
Ιδιωτικό δωμάτιο	0.32	0.46	Ένα ιδιωτικό δωμάτιο
Κοινόχρηστος χώρος	0.01	0.13	Κοινόχρηστο δωμάτιο
Διαμονή	3.25	1.97	Ο αριθμός των ατόμων που μπορούν να φιλοξενήσει.
Μπάνια	1.20	0.54	Ο αριθμός των λουτρών
Υπνοδωμάτια	1.30	0.85	Ο αριθμός των υπνοδωματίων
Πραγματικό κρεβάτι	0.95	0.21	Προσφέρετε ένα πραγματικό κρεβάτι (σε σχέση με άλλους τύπους κλινών όπως αερόθερμα).
Ασύρματο Internet	0.94	0.23	Προσφέρετε ασύρματη πρόσβαση στο Internet.
Πρωινό γεύμα	0.09	0.287	Προσφέρετε πρωινό.
Δωρεάν στάθμευση	0.26	0.44	Προσφέρετε δωρεάν χώρος στάθμευσης.
Άμεση κράτηση	0.16	0.37	Δυνατότητα στιγμιαίας κράτησης.
Πολιτική ακύρωσης (Μέτρια και	0.73	0.43	Δεν ισχύουν ακυρώσεις.

αυστηρή)			
Κάπνισμα	0.12	0.33	Το κάπνισμα επιτρέπεται.
Απαιτούμενη φωτογραφία προφίλ επισκέπτη	0.04	0.20	Απαιτείτε εικόνα προφίλ του επισκέπτη για έγκριση κράτησης.
Απαιτείται επαλήθευση τηλεφώνου από τον επισκέπτη	0.06	0.24	Απαιτείται αριθμός τηλεφώνου επισκέπτη για έγκριση κράτησης.
Κριτικές ανά έτος	9.65	14.41	Ο συνολικός αριθμός κριτικών κατά τα έτη κατά τα οποία καταχωρήθηκε η μίσθωση
Βαθμολογία	92.08	9.09	Συνολική βαθμολογία του καταλύματος.

Πίνακας 3.4: Περιγραφικά μέτρα και ορισμός μεταβλητών.

Από το πίνακα 3.4 βλέπουμε ότι τα καταλύματα έχουν μια μέση τιμή 117,18 δολάρια US. Από τα κύρια χαρακτηριστικά του δείγματος μπορούμε να αναφέρουμε ότι super hosts ιδιοκτήτες είναι μόνο το 9%. Για την εικόνα προφίλ του οικοδεσπότη έχει το 99% των καταχωρήσεων. Στο 64% των καταχωρήσεων έχει ολοκληρώσει τις διαδικασίες Verified ID στο Airbnb.com. Η μέση απόσταση των καταλυμάτων από το κέντρο της κάθε πόλης είναι στα 6,28 χιλιόμετρα. Για τους τύπους των καταλυμάτων βλέπουμε ότι σε μεγαλύτερο ποσοστό με 65% είναι Ολόκληρο σπίτι/ διαμέρισμα, στη συνέχεια με 32% είναι τα ιδιωτικά δωμάτια και τέλος με 1% οι κοινόχρηστοι χώροι. Ο αριθμός των ατόμων που μπορεί να μείνει στο κατάλυμα έχει μέση τιμή 3,25 και διαθέτουν με μέση τιμή 1,2 μπάνια.
Επίσης για τα υπνοδωμάτια βλέπουμε ότι τα καταλύματα έχουν με μέση τιμή 1,3 υπνοδωμάτια και το 95% των καταλυμάτων διαθέτει πραγματικά κρεβάτια. Το 94% διαθέτει ασύρματο ίντερνετ. Ένα πολύ μικρό ποσοστό 9% διαθέτει στους φιλοξενούμενους πρωινό γεύμα. Δωρεάν χώρο στάθμευσης διαθέτει μόνο το 26%. Ακόμα παρατηρούμε ότι μόνο στο 16% των καταλυμάτων μπορεί να γίνει άμεση κράτηση και μόνο στο 12% επιτρέπεται το κάπνισμα. Για τα καταλύματα που απαιτούν φωτογραφία προφίλ επισκέπτη και επαλήθευση μέσω τηλεφώνου τα ποσοστά είναι αντίστοιχα 4% και 6%, πολύ μικρά ποσοστά.
Τέλος τα καταλύματα έχουν 9,65 κριτικές ανά έτος και η μέση βαθμολογία είναι των καταλυμάτων είναι στο 92,08 στα εκατό.

3.7 Περιγραφικά μέτρα των πόλεων του δείγματος.

Οι καταχωρήσεις των καταλυμάτων της Airbnb με τα αντίστοιχα χαρακτηριστικά τους έχουν συλλεχθεί από την ιστοσελίδα www.insideairbnb.com για την κάθε πόλη ξεχωριστά. Στον πίνακα 3.5 παρέχονται όλες οι λεπτομέρειες των καταχωρήσεων για κάθε πόλη. Καθώς ορισμένα καταλύματα είναι εισηγμένα αλλά δεν συνδέονται με πραγματικές συναλλαγές, χρησιμοποιήθηκαν μόνο καταχωρήσεις με τουλάχιστον μία επιθεώρηση πελατών σε απευθείας σύνδεση για να διασφαλίσουμε ότι η τιμή των καταλυμάτων που αναγράφονται αντικατοπτρίζει την ισορροπία της αγοράς σε κάποιο βαθμό.

Πόλη	**Επιλεγμένες καταχωρίσεις (με τουλάχιστον μία αναθεώρηση ανά έτος)**
Άμστερνταμ	7525
Αμβέρσα	522
Αθήνα	1373
Ώστιν	2926
Βαρκελώνη	9392
Βερολίνο	9910
Βοστώνη	1759
Βρυξέλες	3085
Σικάγο	3632
Δουβλίνο	2690
Λονδίνο	19682
Λος Άντζελες	12404
Μαδρίτη	4743
Μαγιόρκα	4524
Μελβούρνη	4883
Μόντρεαλ	5074
Νάσβιλ	1504
Νέα Ορλεάνη	2387
Νέα Υόρκη	21733
Όκλαντ	713
Παρίσι	26472
Πόρτλαντ	2206
Σαν Ντιέγκο	1977
Σαν Φρανσίσκο	4650
Επαρχία Σάντα Κρουζ	630
Σηάτλ	2778
Σύδνεϋ	6633
Τορόντο	3705
Τρεντίνο	712
Βανκούβερ	3058
Βενετία	2275
Βιέννη	2601
Ουάσιγκτον	2375

Πίνακας 3.5: Περιγραφικά μέτρα πόλεων.

Το σύνολο των καταχωρήσεων με τουλάχιστον μια αναθεώρηση ανά έτος είναι 180533.

3.8 Μοντέλα της διαδικασίας.

Τα μοντέλα που χρησιμοποιήθηκαν για να βρεθεί η σχέση μεταξύ της ανεξάρτητης μεταβλητής και ενός συνόλου ανεξάρτητων μεταβλητών είναι το γραμμικό μοντέλο QR και το γραμμικό μοντέλο παλινδρόμησης OLS.

3.8.1 Μοντέλο OLS (Ordinary Least Square).

Η μέθοδος ελαχίστων τετραγώνων είναι μια μέθοδος που χρησιμοποιείται για την εύρεση της εξίσωσης, της καλύτερης ευθείας για την εκτίμηση των άγνωστων παραμέτρων σε ένα μοντέλο γραμμικής παλινδρόμησης.

Οπότε με τη μέθοδο των ελαχίστων τετραγώνων η ευθεία που θα προσαρμόζεται καλύτερα στα δεδομένα θα ελαχιστοποιεί το άθροισμα των τετραγώνων των υπολοίπων ε_ι.

Επίσης έχει αναφερθεί ότι η μέθοδος OLS επιλέγει τις παραμέτρους μιας γραμμικής συνάρτησης ενός συνόλου επεξηγηματικών μεταβλητών, ελαχιστοποιώντας το άθροισμα των τετραγώνων των υπολοίπων, μεταξύ της εξαρτημένης μεταβλητής στο σύνολο του δείγματος.

$$\sum_{i=1}^{n} e_i^2 = \sum_{i=1}^{n} (y_i - (\beta_0 + \beta_1 x_i))^2$$

Οι εκτιμήτριες ελαχίστων τετραγώνων για τις παραμέτρους β_0 και β_1της ευθείας $y = \beta_0 + \beta_1 x_i$ φαίνονται στη συνέχεια.

$$\beta_1 = \frac{\nu \sum_{i=1}^{\nu} x_i y_i - (\sum_{i=1}^{n} x_i)(\sum_{i=1}^{n} y_i)}{\nu \sum_{i=1}^{n} {x_i}^2 - (\sum_{i=1}^{n} x_i)^2}$$

$$\beta_0 = \frac{1}{\nu}\sum_{i=1}^{n} y_i - \beta_1 * \frac{1}{\nu}\sum_{i=1}^{n} x_i$$

Η ευθεία $y = \beta_0 + \beta_1 x_i$ καλείται ευθεία ελαχίστων τετραγώνων της Υ πάνω στη Χ.

3.8.2 Μοντέλο QR (Quantile Regression).

Ενώ η μέθοδος των ελαχίστων τετραγώνων οδηγεί σε εκτιμήσεις του μέσου όρου της μεταβλητής απόκρισης, η ποσοτική παλινδρόμηση στοχεύει στην εκτίμηση είτε του μέσου όρου υπό όρους είτε άλλων ποσοτήτων της μεταβλητής απόκρισης.

Τα γραμμικά μοντέλα QR μοντελοποιούν με βάση την υπό όρους τ $\tau \in (0, 1)$ του ποσοστού της εξαρτημένης μεταβλητής. Επομένως, το QR υπερβαίνει την ανάλυση του μέσου όρου εξαρτώμενης μεταβλητής, παρέχοντας μια πληρέστερη περιγραφή της κατανομής υπό όρους. Δηλαδή, αντί να υπολογίσει τη μέση απόκριση της εξαρτώμενης μεταβλητής στις αλλαγές στις επεξηγηματικές μεταβλητές, το QR μετρά τα αποτελέσματα των μεμονωμένων επεξηγηματικών μεταβλητών σε ολόκληρη τη διανομή της εξαρτώμενης μεταβλητής.

Το QR καθορίζεται ως εξής (Koenker and Bassett, 1978).

Έστω μια τυχαία μεταβλητή Y με μια συνάρτηση κατανομής πιθανοτήτων $F (y)$ = Prob ($Y \le y$),

Η Y μπορεί να οριστεί ως η μικρότερη τιμή *y που* ικανοποιεί την $F_y^{-1}(\tau) = Q_y (\tau) = \inf \{ y : F (y) \ge \tau \}$, όπου $0 < \tau < 1$.

Για n παρατηρήσεις της Y, η συνάρτηση της εμπειρικής κατανομής δίνεται ως $F_n (y) = \Sigma\, 1 (Y_i \le y)$, όπου $1 (z)$ είναι συνάρτηση δείκτη που παίρνει την τιμή 1 εάν το όρισμα z είναι αληθές και 0 αλλιώς.

Συνεπώς, οι εμπειρικές ποσοστώσεις ορίζονται ως εξής:

$Q_n (\tau) = \inf \{ y : F (y) \ge \tau \}$.

Αυτή η έκφραση δίνεται ως πρόβλημα βελτιστοποίησης παρακάτω:

$$Q_{n(\tau)} = argmin(\xi)\{ \sum_{i:Y_i \ge \xi} \tau|Y_i - \xi| + \sum_{i:Y_i < \xi} (1-\tau)|Y_i - \xi|\}$$
$$= argmin(\xi)\{ \sum_{i:Y_i < \xi} \rho_\tau |Y_i - \xi|\}$$

όπου $r_t(u) = u(\tau - 1(u < 0))$ είναι η λεγόμενη *λειτουργία ελέγχου* , η οποία ζυγίζει ασύμμετρα θετικές και αρνητικές τιμές. Μια γραμμική προδιαγραφή του υπό όρους ποσοστημόριο της εξαρτημένης μεταβλητής δίνει $Q(\tau \mid X_i, \beta(\tau)) = X_i' \beta(\tau)$, όπου το X_i είναι το διάνυσμα των επεξηγηματικών μεταβλητών και $\beta(\tau)$ είναι ο φορέας των συντελεστών που συνδέονται με την τ ποσοτική. Υπό τις συνθήκες αυτές, το προηγούμενο πρόβλημα βελτιστοποίησης έχει ως εξής:

$$\beta\hat{}n(\tau), = argmin_{\beta(\tau)}\{\sum_{\iota} \rho_\tau\,(Y_i - X'_i\beta(\tau))\}$$

Διαισθητικά, οι παράμετροι του QR υπολογίζονται λαμβάνοντας υπόψη τα διαφορετικά βάρη των απόλυτων υπολειμμάτων. Για την ανάλυση των τιμών εισαγωγής, επιλέγεται η μεταβλητή της τιμής ανά άτομο ανά διανυκτέρευση (σε λογαριθμική μορφή) ως εξαρτημένη μεταβλητή. Δεδομένου ότι η προκύπτουσα έκφραση είναι λογαριθμική προδιαγραφή, οι τιμές των συντελεστών αντιπροσωπεύουν ημιελαστικότητα, δηλαδή την ποσοστιαία μεταβολή της τιμής όταν μια ερμηνευτική μεταβλητή ποικίλει κατά 1, έχοντας υπόψη ότι η επίδραση μιας εικονικής ανεξάρτητης μεταβλητής σε μια λογαριθμική εξαρτώμενη μεταβλητή είναι μετρούμενη με e^{α} -1.

3.9 Αποτελέσματα.

Στο πίνακα 3.6 βλέπουμε τα αποτελέσματα των μεθόδων. Συγκεκριμένα μας δίνει τους συντελεστές των παρακάτω ανεξάρτητων μεταβλητών ως προς την τιμή ανά άτομο ανά διανυκτέρευση από τη μέθοδο OLS. Εκτός από τους συντελεστές από την OLS μας δίνει και κάποια ποσοστά (το 10^{ου}, 25^{ου}, 50^{ου}, 75^{ου}, και 90^{ου}) της εξαρτημένης μεταβλητής από τη μέθοδο QR για να αποδειχθούν τα αποτελέσματα των επεξηγηματικών μεταβλητών για κάθε ποσό. Αυτό επιτρέπει τη διάκριση των επιδράσεων μιας συγκεκριμένης μεταβλητής ανάλογα με το όριο της εξαρτημένης μεταβλητής.

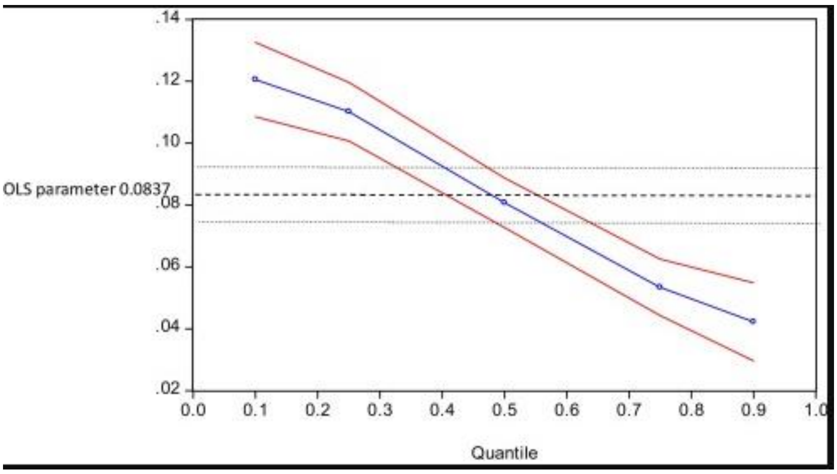
Μεταβλητές	OLS	0,1	0,25	0,5	0,75	0,9
Constant	2.4083^{a} (0.0296)	1.8154^{a} (0.0443)	2.0438^{a} (0.0426)	2.3263^{a} (0.0451)	2.7156^{a} (0.0484)	3.0724^{a} (0.0447)
Super host	0.0837^{a} (0.0039)	0.1206^{a} (0.0061)	0.1102^{a} (0.0048)	0.0807^{a} (0.0041)	0.0534^{a} (0.0046)	0.0422^{a} (0.0064)
Καταγραφές κεντρικών υπολογιστών	0.0006^{a} (0.0001)	0.0007^{a} (0.0002)	0.0006^{a} (0.0001)	0.0008^{a} (0.0001)	0.0008^{a} (0.0001)	0.0005^{a} (0.0001)
Εικόνα προφίλ	-0.1154^{a} (0.0241)	-0.0856^{a} (0.0267)	-0.0994^{a} (0.0330)	-0.1142^{a} (0.0392)	-0.1313^{a} (0.0414)	-0.1473^{a} (0.0357)

οικοδεσπότη						
Επαλήθευση ταυτότητας κεντρικού υπολογιστή	0.0856[a] (0.0025)	0.0679[a] (0.0039)	0.0843[a] (0.0031)	0.0948[a] (0.0029)	0.0861[a] (0.0033)	0.0767[a] (0.0042)
Απόσταση (km)	−0.0059[a] (0.0002)	−0.0053[a] (0.0003)	−0.0054[a] (0.0002)	−0.0058[a] (0.0002)	−0.0062[a] (0.0003)	−0.0063[a] (0.0003)
Τύπος διαμονής 1	−0.0827[a] (0.0033)	−0.1390[a] (0.0056)	−0.1073[a] (0.0043)	−0.0619[a] (0.0038)	−0.0422[a] (0.0042)	−0.0329[a] (0.0058)
Τύπος διαμονής 2	−0.0890[a] (0.0088)	−0.2483[a] (0.0193)	−0.1175[a] (0.0147)	−0.0303[b] (0.0119)	0.0345[a] (0.0113)	0.0556[a] (0.0162)
Ολόκληρο σπίτι / διαμέρισμα	0.8948[a] (0.0090)	1.0371[a] (0.0151)	0.9555[a] (0.0127)	0.8419[a] (0.0128)	0.7535[a] (0.0171)	0.6667[a] (0.0187)
Ιδιωτικό δωμάτιο	0.3419[a] (0.0090)	0.4097[a] (0.0148)	0.3727[a] (0.0128)	0.3162[a] (0.0128)	0.2582[a] (0.0169)	0.1735[a] (0.0183)
Διαμονή	0.0616[a] (0.0010)	0.0554[a] (0.0015)	0.0589[a] (0.0012)	0.0660[a] (0.0014)	0.0711[a] (0.0016)	0.0780[a] (0.0021)
Μπάνια	0.1085[a] (0.0027)	0.0515[a] (0.0039)	0.0798[a] (0.0040)	0.1237[a] (0.0042)	0.1705[a] (0.0055)	0.2092[a] (0.0067)
Υπνοδωμάτια	0.1249[a] (0.0021)	0.0976[a] (0.0031)	0.1111[a] (0.0027)	0.1216[a] (0.0029)	0.1318[a] (0.0034)	0.1275[a] (0.0043)
Πραγματικό κρεβάτι	0.1555[a] (0.0055)	0.0819[a] (0.0070)	0.1244[a] (0.0060)	0.1831[a] (0.0055)	0.2106[a] (0.0083)	0.2336[a] (0.0085)
Ασύρματο Internet	0.0951[a] (0.0052)	0.1331[a] (0.0089)	0.1262[a] (0.0073)	0.0886[a] (0.0068)	0.0646[a] (0.0076)	0.0751[a] (0.0099)
Πρωινό Γεύμα	−0.0106[a] (0.0042)	0.0108 (0.0068)	−0.0031 (0.0054)	−0.0103[b] (0.0048)	−0.0253[a] (0.0055)	−0.0210[a] (0.0099)
Δωρεάν στάθμευση	0.0811[a] (0.0029)	0.1184[a] (0.0049)	0.1103[a] (0.0037)	0.0891[a] (0.0033)	0.0433[a] (0.0036)	0.0084 (0.0049)
Άμεση κράτηση	−0.0665[a] (0.0031)	−0.0606[a] (0.0048)	−0.0614[a] (0.0040)	−0.0607[a] (0.0036)	−0.0680[a] (0.0041)	−0.0756[a] (0.0051)
Πολιτική ακύρωσης (Μέτρια και αυστηρή)	0.0448[a] (0.0028)	0.0446[a] (0.0043)	0.0431[a] (0.0035)	0.0478[a] (0.0034)	0.0490[a] (0.0039)	0.0406[a] (0.0049)
Κάπνισμα	−0.2654[a] (0.0035)	−0.2253[a] (0.0053)	−0.2536[a] (0.0042)	−0.2804[a] (0.0040)	−0.2876[a] (0.0050)	−0.2588[a] (0.0063)
Απαιτούμενη φωτογραφία προφίλ επισκέπτη	0.0102 (0.0082)	0.0096 (0.0126)	−0.0008 (0.0101)	0.0178 (0.0103)	0.0096 (0.0095)	0.0197 (0.0129)
Απαιτείται επαλήθευση τηλεφώνου από τον	0.0220[a] (0.0071)	0.0217 (0.0115)	0.0303[a] (0.0085)	0.0172 (0.0090)	0.0272 (0.0088)	0.0141[a] (0.0105)

επισκέπτη						
Κριτικές ανά έτος	-0.0010^a (0.0001)	0.00004 (0.0001)	-0.0002^b (0.0001)	-0.0007^a (0.0001)	-0.0015^a (0.0001)	-0.0022^a (0.0001)
Βαθμολογία	0.0087^a (0.0001)	0.0091^a (0.0003)	0.0091^a (0.0002)	0.0091^a (0.0002)	0.0086^a (0.0002)	0.0082^a (0.0002)
Notes: ^{a} prob <1%; ^{b} prob <5%. Στις παρενθέσεις είναι τα τυπικά σφάλματα.						

Πίνακας 3.6: καθοριστικοί παράγοντες της τιμής ανά διανυκτέρευση (OLS και QR).

Η μεταβλητή super host παρατηρούμε ότι προκαλεί αύξηση της τιμής. Υπολογίζοντας το (e^{α}-1) θα βρούμε το ποσοστό επίδρασης στη τιμή. Οπότε από το μοντέλο OLS το $e^{0,0837}$ -1 = 0,0873 άρα η τιμή αυξάνεται κατά 8,73%. Στη συνέχεια θα ελέγξουμε αν ισχύει και το ίδιο από το μοντέλο QR και συγκεκριμένα από τη γραμμική εξίσωση των συντελεστών των ποσοστών που παρουσιάστηκε στην εικόνα 3.7. Όπως παρατηρείται στο παρακάτω σχήμα δημιουργείται φθίνουσα εικόνα των ποσοστών. Έτσι μπορούμε να πούμε ότι και αν οι παράμετροι που λαμβάνονται με την παλινδρόμηση OLS και με το μοντέλο QR έχουν το ίδιο πρόσημο, μπορεί το αποτέλεσμα της παλινδρόμησης OLS να μην είναι σταθερό μεταξύ των ποσοτήτων, όπως παρατηρείται στο γράφημα ότι ένα μεγάλο μέρος του διαστήματος εμπιστοσύνης του QR δεν εμπίπτει στο σταθερό διάστημα εμπιστοσύνης OLS.

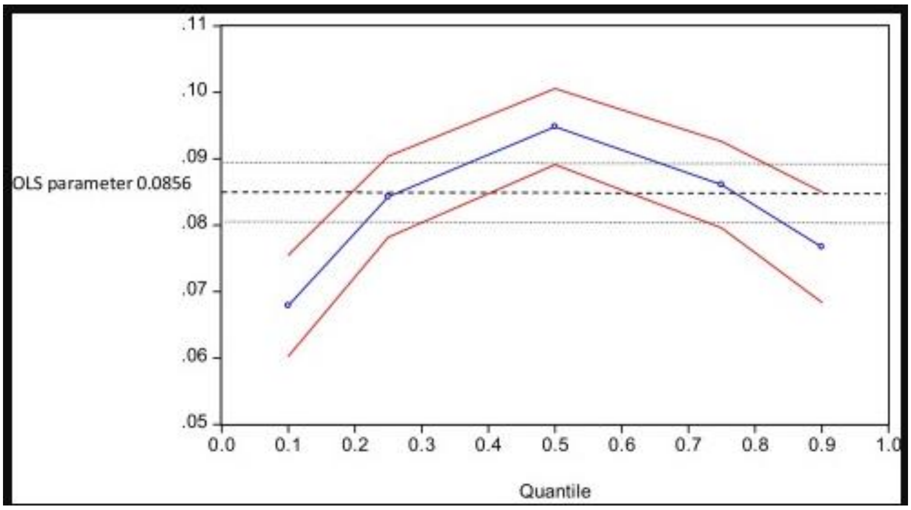


Εικόνα 3.7:Διάγραμμα συντελεστών super host από OLS και QR.

- Για τον αριθμό των καταλόγων υποδοχής παρατηρούμε θετικούς συντελεστές από τη παλινδρόμηση OLS και από το μοντέλο QR. Συγκεκριμένα η τιμή αυξάνεται κατά $e^{0,0006}$ -1=0,0006 ή 0,06% για κάθε καταμέτρηση συμφώνα με την OLS όμως παρατηρούμε ότι από το μοντέλο QR οι συντελεστές του $50^{ου}$ και του $75^{ου}$ είναι υψηλότεροι από το συντελεστή της OLS και του $90^{ου}$ είναι σημαντικά χαμηλότεροι από τους υπόλοιπους. Οπότε μπορούμε να πούμε για τη συγκεκριμένη μεταβλητή ότι έχει μεγαλύτερη θετική επίδραση στις καταχωρήσεις χαμηλότερων τιμών και μικρότερη θετική επίδραση στις καταχωρήσεις υψηλότερων τιμών.

- Για την εικόνα προφίλ του οικοδεσπότη βλέπουμε ότι επηρεάζει αρνητικά τη τιμή και συγκεκριμένα από το μοντέλο OLS παρουσιάζει μείωση $e^{-0,1554}$-1= -0,1089 ή -10,89%. Όμως από το μοντέλο QR παρατηρούμε ότι οι συντελεστές των ποσοστών αυξάνονται καθώς αυξάνονται και αυτοί. Άρα η συγκεκριμένη μεταβλητή έχει μεγαλύτερη επίδραση στις υψηλές τιμές σύμφωνα με το μοντέλο QR.

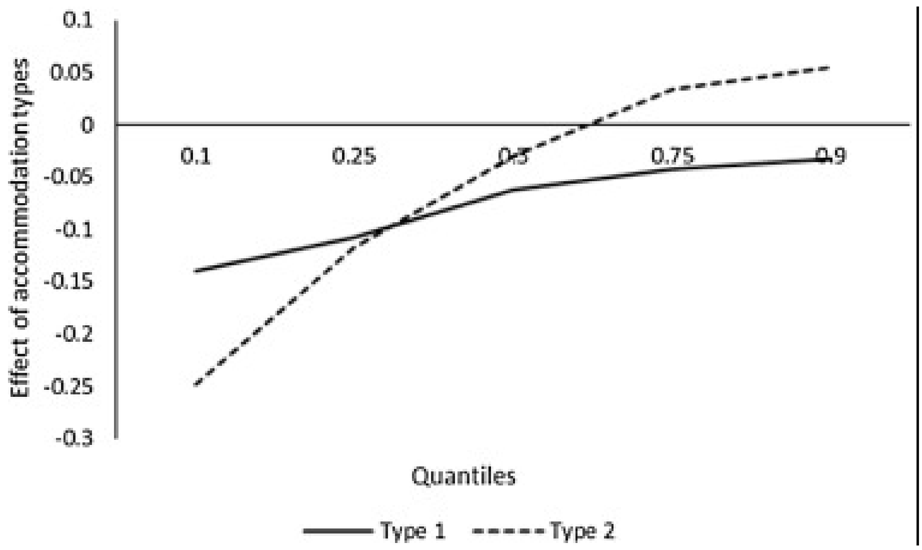
- Η επαλήθευση ταυτότητας του ξενιστή επηρεάζει θετικά την τιμή όπως βλέπουμε και υπολογίζοντας από το μοντέλο OLS το $e^{0,0856}$-1=0,0894 θα υπάρχει αύξηση της τιμής κατά 8,94%.Από το μοντέλο QR οι επιδράσεις των ποσοτικών παραμέτρων που φαίνονται στην εικόνα 3.8 έχουν τη μορφή ενός ανεστραμμένου U. Παρατηρούμε ότι η θετική επίδραση στη τιμή είναι χαμηλότερη στις ουρές της καμπύλη και συγκεκριμένα στο $10^{ο}$ και στο $90^{ο}$ ενώ την υψηλότερη θετική επίδραση την έχει στο $50^{ο}$. Όλα αυτά τα ευρήματα υποστηρίζουν το συμπέρασμα που έχει παρατηρηθεί σε μελέτες ότι οι οικοδεσπότες αξιοποιούν συνήθως καλή φήμη και επαγγελματική κατάσταση. (Gutt and Hermann, 2015, Ikkala and Lampien, 2014, Li et al., 2015)



Εικόνα 3.8:Διάγραμμα συντελεστών επαλήθευσης ταυτότητας του ξενιστή από OLS και QR.

- Η μεταβλητή απόσταση έχει σημαντική αρνητική επίδραση στη τιμή και μπορούμε να πούμε ότι όσο πιο μακριά είναι το κατάλυμα από το κέντρο της πόλης τόσο η τιμή θα είναι χαμηλότερη. Ειδικότερα από το μοντέλο OLS υπολογίζοντας το $e^{-0,0059}$-1=0,0059 ή 0,59% παρατηρούμε μείωση κατά 0,59 ανά χιλιόμετρο. Από το μοντέλο QR βλέπουμε ότι στις υψηλές τιμές (50^{η} ,75^{η},90^{η}) υπάρχει μεγαλύτερη αρνητική επίδραση.

- Για τον τύπο καταλυμάτων 1 και 2 παρατηρούμε μια αρνητική επίδραση στη τιμή. Συγκεκριμένα για το τύπο καταλύματος 1 από το μοντέλο OLS υπολογίζοντας το $e^{-0,0851}$-1= -0,0794 ή 7,94% δηλαδή η τιμή μειώνεται κατά 7,94% και για το τύπο καταλύματος 2 υπολογίζοντας το $e^{-0,0890}$-1= - 0,0851 ή 8,51% δηλαδή η τιμή μειώνεται κατά 8,51%. Από το μοντέλο QR επίσης παρατηρούμε αρνητικές επιδράσεις των τύπων 1 και 2 ενώ για τις υψηλές τιμές η αρνητική επίδραση μειώνεται και η μόνη κοινή τιμή παρατηρείται στην τιμή του 25^{ου} ποσού. Στην εικόνα 3.9 φαίνονται οι καμπύλες των τύπων 1 και 2.



Εικόνα 3.9:Διάγραμα συντελεστών τύπου 1 και 2 καταλυμάτων από OLS και QR.

- Για τη μεταβλητή ολόκληρο σπίτι/ διαμέρισμα παρατηρούμε μια θετική επίδραση στη τιμή, ειδικότερα από το μοντέλο OLS υπολογίζοντας το $e^{0,894}$-1= 1,4468 ή 144,68% βλέπουμε ότι η τιμή αυξάνεται κατά 144,68% αν ο τύπος σπιτιού είναι ολόκληρο σπίτι/ διαμέρισμα. Από το μοντέλο QR βλέπουμε ότι υπάρχει και εκεί αύξηση της τιμής και συγκεκριμένα αυτή η αύξηση είναι μεγαλύτερη για καταχωρίσεις σε χαμηλές τιμές και μικρότερη για καταχωρίσεις υψηλού κόστους.

- Η μεταβλητή ιδιωτικό δωμάτιο παρουσιάζει και αυτή επίσης μια θετική επίδραση στη τιμή και υπολογίζοντας από το μοντέλο OLS το $e^{0,3419}$-1= 0,4076 ή 40,76% βλέπουμε ότι η τιμή αυξάνεται κατά 40,76% αν ο τύπος σπιτιού είναι ιδιωτικό δωμάτιο. Όπως και στη προηγούμενη μεταβλητή από το μοντέλο QR βλέπουμε ότι η αύξηση είναι μεγαλύτερη για καταχωρίσεις σε χαμηλές τιμές και μικρότερη για καταχωρίσεις υψηλού κόστους.

- Η μεταβλητή διαμονή που εκφράζει τον αριθμό των ατόμων που μπορεί να φιλοξενήσει το κατάλυμα παρατηρούμε ότι έχει θετική επίδραση στη τιμή και ειδικότερα από το μοντέλο OLS υπολογίζοντας το $e^{0,0616}$-1=0,0635 ή 6,35% όπου ισχύει ότι η τιμή αυξάνεται κατά 6,35% όσο αυξάνεται ο αριθμός των ατόμων. Από το μοντέλο QR βλέπουμε ότι η αύξηση είναι μεγαλύτερη για καταχωρίσεις μεγάλου κόστους.

- Η μεταβλητή μπάνια έχει θετική επίδραση στη τιμή και ειδικότερα υπολογίζοντας από το μοντέλο OLS το $e^{0,1085}$-1=0,1146 ή 11,46% όπου ισχύει ότι η τιμή αυξάνεται κατά 11,46% όσο αυξάνονται τα μπάνια. Όπως και στη προηγούμενη μεταβλητή από το μοντέλο QR βλέπουμε ότι η αύξηση είναι μεγαλύτερη για καταχωρίσεις σε υψηλές τιμές και μικρότερη για καταχωρίσεις χαμηλού κόστους.

- Η μεταβλητή υπνοδωμάτια έχει θετική επίδραση στη τιμή και συγκεκριμένα από το μοντέλο OLS υπολογίζοντας το $e^{0,1249}$-1=0,1330 ή 13,3% όπου ισχύει ότι η τιμή αυξάνεται κατά 13,3% όσο αυξάνονται τα υπνοδωμάτια. Από το μοντέλο QR βλέπουμε ότι η αύξηση είναι μεγαλύτερη για καταχωρίσεις μεγάλου κόστους και συγκεκριμένα μεγιστοποιείται στη τιμή του 75ου ποσού.

- Η μεταβλητή πραγματικά κρεβάτια παρουσιάζει και αυτή επίσης μια θετική επίδραση στη τιμή και ειδικότερα υπολογίζοντας από το μοντέλο OLS το

$e^{0,1555}$-1= 0,1682 ή 16,82% βλέπουμε ότι η τιμή αυξάνεται κατά 16,82% αν υπάρχουν πραγματικά κρεβάτια στο κατάλυμα. Όπως και στη προηγούμενη μεταβλητή από το μοντέλο QR βλέπουμε ότι η αύξηση είναι μεγαλύτερη για καταχωρίσεις σε χαμηλές τιμές και μικρότερη για καταχωρίσεις υψηλού κόστους. Από το μοντέλο QR βλέπουμε ότι η αύξηση είναι μεγαλύτερη για καταχωρίσεις μεγάλου κόστους.

- Η μεταβλητή ασύρματου ίντερνετ παρατηρούμε ότι έχει μια θετική επίδραση στη τιμή. Υπολογίζοντας από το μοντέλο OLS το $e^{0,0951}$-1=0,0997 ή 9,97% βλέπουμε ότι αν το κατάλυμα παρέχει πρόσβαση ίντερνετ η τιμή θα αυξάνεται κατά 9,97% όπως παρατηρούμε αύξηση και από το μοντέλο QR αλλά οι ποσοτικές εκτιμήσεις της παρουσιάζουν ένα φθίνουσα μοτίβο. Ως εκ τούτου η συγκεκριμένη παροχή είναι πιο σημαντική για τις χαμηλές τιμές. Αυτό το συμπέρασμα είναι ασυμβίβαστο με την παρατήρηση των επαγγελματιών της ξενοδοχειακής βιομηχανίας ότι τα ξενοδοχεία και οι ξενώνες χαμηλού τιμολογίου είναι πιθανότερο να παρέχουν δωρεάν πρόσβαση στο Διαδίκτυο. (Rent et al., 2016).

- Για τη μεταβλητή πρωινό γεύμα μια αρνητική επίδραση στη τιμή από τη παλινδρόμηση OLS ίση με -1,05% υπολογίζοντας το $e^{-0,0106}$-1=-0,0105. Από το μοντέλο QR η επίδραση αυτής της μεταβλητής είναι μηδενική για τη 10^{o} και 25^{o} και αρνητική για τα 50^{o} , 75^{o} και 90^{o} ποσά. Συμπεραίνουμε ότι η παλινδρόμηση OLS δείχνει μια σημαντική αρνητική επίδραση σε όλη την κατανομή των τιμών, ενώ η QR αποκαλύπτει ότι μόνο εκείνες οι τιμές που υπερβαίνουν το μέσο όρο έχουν μια σημαντικά αρνητική παράμετρο. Επίσης πρέπει να αναφέρουμε ότι μόνο στο 9% των καταλυμάτων παρέχεται πρωινό γεύμα. Αυτό το εύρημα είναι ασυμβίβαστο με τα ευρήματα των επαγγελματιών της ξενοδοχειακής βιομηχανίας. (Masiero et al., 2015, Schamel, 2012).

- Η μεταβλητή δωρεάν χώρου στάθμευσης στις εγκαταστάσεις παρουσιάζει θετική επίδραση στη τιμή και ειδικότερα υπολογίζοντας από το μοντέλο OLS το $e^{0,0811}$-1= 0,0844 ή 8,44% βλέπουμε ότι η τιμή αυξάνεται κατά 8,44% αν υπάρχει χώρος στάθμευσης. Η θετική επίδραση παρατηρείται και από τα δύο μοντέλα OLS και QR. Συγκεκριμένα από το μοντέλο QR ότι στο 10^{o} και 25^{o} ποσό εμφανίζεται μεγαλύτερη θετική επίδραση από το 50^{o} που είναι παρόμοια με την εκτίμηση που μας έδωσε το μοντέλο OLS ενώ στο 75^{o} και στο 90^{o} υπάρχει χαμηλότερη επίδραση. Άρα μπορούμε να πούμε ότι η θετική επίδραση του δωρεάν χώρου στάθμευσης είναι πιο μεγάλη για τις χαμηλού κόστους τιμές σε σχέση με τις τιμές σε υψηλού κόστους. Το εύρημα αυτό είναι σύμφωνο με τα αποτελέσματα που αναφέρθηκαν από τους επαγγελματίες της ξενοδοχειακής βιομηχανίας. (Espinet et al., 2003, Lee and Jang, 2012).

- Η παροχή της άμεσης κράτησης βλέπουμε ότι έχει αρνητική επίδραση στη τιμή από τη παλινδρόμηση OLS ίση με -6,43% υπολογίζοντας το $e^{-0,0665}$-1= -0,0643. Αυτή η επίδραση δεν είναι λογική διότι αφού βοηθά τους ενδιαφερόμενους στην κράτηση του καταλύματος συνδέεται με μείωση της τιμής. Από το μοντέλο QR βλέπουμε ότι η συγκεκριμένη επίδραση είναι πιο αισθητή στις καταχωρήσεις με υψηλές τιμές.

- Για τη μεταβλητή πολιτική ακύρωσης (Μέτρια και Αυστηρή) παρατηρούμε ότι έχει μια θετική επίδραση στη τιμή. Υπολογίζοντας από το μοντέλο OLS το $e^{0,0448}$-1=0,0458 ή 4,58% καταλήγουμε ότι η τιμή θα αυξάνεται κατά 4,58%. Και από το μοντέλο QR παρατηρούμε σχεδόν σταθερή αύξηση για όλο το εύρος των τιμών.

- Για τη μεταβλητή κάπνισμα η οποία αναφέρεται στην άδεια καπνίσματος μέσα στο σπίτι παρατηρούμε ότι προκαλεί μια αρνητική επίδραση στη τιμή. Υπολογίζοντας από το μοντέλο OLS το $e^{-0,2654}$-1= -0,2331 ή -23,31% βλέπουμε ότι η συγκεκριμένη μεταβλητή οδηγεί σε μείωση της τιμής κατά 23,31%. Από το μοντέλο QR παίρνουμε την πληροφορία ότι η συγκεκριμένη μεταβλητή επιδρά αρνητικά πιο πολύ στα καταλύματα χαμηλού κόστους. Συμπεραίνουμε ότι οι οικοδεσπότες που επιτρέπουν το κάπνισμα μέσα στο κατάλυμα έχουν σαν αποτέλεσμα τη μείωση των εσόδων τους.

- Η μεταβλητή όπου απαιτεί να έχει εικόνα προφίλ ο επισκέπτης παρατηρούμε ότι δεν παίζει σημαντικό ρόλο στη τιμή με βάση τα αποτελέσματα των παλινδρομήσεων OLS και QR. Από τη παλινδρόμηση OLS βλέπουμε ότι επηρεάζει κατά $e^{0,0102}$-1=0,0102 ή 1,02% που είναι πάρα πολύ μικρό ποσοστό και δεν το λαμβάνουμε υπόψη.

- Για τη μεταβλητή η οποία απαιτεί επαλήθευση μέσω τηλεφώνου βλέπουμε ότι επηρεάζει θετικά τη τιμή και ειδικότερα υπολογίζοντας από το μοντέλο OLS το $e^{0,0220}$-1=0,0222 βλέπουμε ότι η συγκεκριμένη μεταβλητή προκαλεί αύξηση στη τιμή κατά 2,22%. Από το μοντέλο QR παρατηρούμε ότι αυτή η θετική επίδραση είναι υψηλότερη στο 25^{o} και στο 90^{o} ποσό ενώ στα υπόλοιπα υπάρχει χαμηλότερη θετική επίδραση. Πρέπει να αναφέρουμε ότι οι οικοδεσπότες που απαιτούν επαλήθευση μέσω τηλεφώνου ανήκουν στο 6% του δείγματος και θέλουν να επιβεβαιωθούν ότι ο επισκέπτης θα έρθει.

- Για τη μεταβλητή κριτικές ανά έτος παρατηρούμε ότι προκαλεί μια πολύ μικρή αρνητική επίδραση στη τιμή όπως φαίνεται και από τα δυο μοντέλα. Υπολογίζοντας από το μοντέλο OLS το $e^{-0,0010}$-1= -0,00099 ή -0,01% βλέπουμε ότι οι κριτικές επηρεάζουν αρνητικά κατά 0,01%. Επιπροσθέτως μπορούμε να πούμε ότι τα καταλύματα με μειωμένη τιμή θα έχουν περισσότερες κρατήσεις άρα και περισσότερες κριτικές. Οι προηγούμενοι ερευνητές ανέφεραν ότι οι περισσότεροι τουρίστες επιλέγουν να νοικιάσουν καταλύματα που βασίζονται στην οικονομία για να μειώσουν το κόστος (Balck and Cracau, 2015, Guttentag, 2015, Quinby και Gasdia, 2014).

- Η μεταβλητή βαθμολογία βλέπουμε ότι επηρεάζει θετικά τη τιμή και με βάση το μοντέλο OLS υπολογίζοντας το $e^{0,0087}$-1=0,0087 παρατηρούμε ότι η τιμή επηρεάζεται κατά 0,87% από τη βαθμολογία. Από το μοντέλο QR παρατηρούμε επίσης θετική επίδραση με μεγαλύτερη επίδραση στα καταλύματα χαμηλού κόστους.

ΚΕΦΑΛΑΙΟ 4

4. Η μέθοδος της βηματικής παλινδρόμησης.

4.1 Απλή γραμμική παλινδρόμηση.

Η απλή γραμμική παλινδρόμηση είναι ένα μοντέλο γραμμικής παλινδρόμησης με μία μόνο εξαρτημένη μεταβλητή. Δηλαδή, πρόκειται για δισδιάστατα σημεία δειγματοληψίας με μια ανεξάρτητη μεταβλητή και μια εξαρτημένη και βρίσκει γραμμική συνάρτηση (μια μη κατακόρυφη ευθεία γραμμή) που, όσο το δυνατόν ακριβέστερα, προβλέπει τις εξαρτώμενες τιμές μεταβλητών ως συνάρτηση των ανεξάρτητων μεταβλητών. Είναι συνηθισμένο να κάνουμε την πρόσθετη υπόθεση ότι πρέπει να χρησιμοποιείται η μέθοδος ελαχίστων τετραγώνων για την ελαχιστοποίηση των υπολειμμάτων (κάθετες αποστάσεις μεταξύ των σημείων του συνόλου δεδομένων και της προσαρμοσμένης γραμμής). Σύμφωνα με αυτή την υπόθεση, η ακρίβεια μιας γραμμής μέσω των σημείων δειγματοληψίας μετράται με το άθροισμα τετραγωνικών υπολειμμάτων και ο στόχος είναι να γίνει αυτό το άθροισμα όσο το δυνατόν μικρότερο.

Η κλίση της προσαρμοσμένης γραμμής είναι ίση με τη συσχέτιση μεταξύ y και x διορθωμένη με την αναλογία τυπικών αποκλίσεων αυτών των μεταβλητών. Η τομή της προσαρμοσμένης γραμμής είναι τέτοια ώστε να διέρχεται από το κέντρο της μάζας ($\bar{x}$, $\bar{y}$) των σημείων των δεδομένων.

4.1.1 Μέθοδος ελαχίστων τετραγώνων.

Η μέθοδος των ελαχίστων τετραγώνων είναι μια τυπική προσέγγιση στην ανάλυση παλινδρόμησης για την προσέγγιση της λύσης των υπέρ προσδιορισμένων συστημάτων , δηλαδή σε ομάδες εξισώσεων στις οποίες υπάρχουν περισσότερες εξισώσεις από άγνωστες."Λιγότερα τετράγωνα" σημαίνει ότι η συνολική λύση ελαχιστοποιεί το άθροισμα των τετραγώνων των υπολειμμάτων που γίνονται στα αποτελέσματα κάθε μοναδικής εξίσωσης. Δηλαδή το άθροισμα των u_i^2 όπου είναι το $\sum_{i=0}^{n} u_i^2$ να είναι ελάχιστο.

Η μορφή της ευθείας των ελαχίστων τετραγώνων: y=a+bx

Ισχύει ότι: F(a,b)= $\sum_{i=1}^{n}(a + bx_i - y_i)^2$

Για να ελαχιστοποιηθεί το $\sum_{i=0}^{n} u_i^2$ πρέπει το $\frac{dF(a,b)}{da} = 0$

Για τους συντελεστές a και b ισχύει

$$a = \frac{\sum_{i=1}^{n} x_i^2 * \sum_{i=1}^{n} y_i - \sum_{i=1}^{n} x_i \sum_{i=0}^{n} x_i y_i}{N \sum_{i=1}^{n} x_i^2 - [\sum_{i=1}^{n} x_i]^2}$$

$$b = \frac{N \sum_{i=1}^{n} x_i y_i - \sum_{i=1}^{n} x_i * \sum_{i=1}^{n} y_i}{N \sum_{i=1}^{n} x_i^2 - [\sum_{i=1}^{n} x_i]^2}$$

Τα σφάλματα των συντελεστών a και b φαίνονται στη συνέχεια.

$$\delta a = \sqrt{\frac{1}{N-2}} * \sqrt{\frac{\sum_{i=1}^{n} x_i^2 * \sum_{i=1}^{n} d_i^2}{N \sum_{i=1}^{n} x_i^2 - [\sum_{i=1}^{n} x_i]^2}}$$

$$\delta b = \sqrt{\frac{N}{N-2}} \sqrt{\frac{\sum_{i=1}^{n} d_i^2}{N \sum_{i=1}^{n} x_i^2 - [\sum_{i=1}^{n} x_i]^2}}$$

4.1.2 Γραμμική συσχέτιση.

Η γραμμική συσχέτιση δύο μεταβλητών X και Y με διασπορά σ^2X και σ^2Y αντίστοιχα και συνδιασπορά σ_{xy} = Cov(X, Y) = E(X, Y) − E(X)E(Y), μπορεί να εκφραστεί με τον συντελεστή συσχέτισης r.

$$r = \frac{\sigma_{xy}}{\sigma_x * \sigma_y}$$

Η τιμή του *r* παίρνει τιμές -1 $\leq r \leq$ +1. Τα σύμβολα + και - χρησιμοποιούνται για θετικούς γραμμικούς συσχετισμούς και αρνητικούς γραμμικούς συσχετισμούς, αντίστοιχα.

- Θετική συσχέτιση: Εάν τα *x* και *y* έχουν μια ισχυρή θετική γραμμική συσχέτιση, το *r* είναι κοντά στο +1. Μια τιμή *r* ακριβώς +1 δείχνει μια τέλεια θετική εφαρμογή. Οι θετικές τιμές υποδεικνύουν μια σχέση μεταξύ των μεταβλητών *x* και *y* έτσι ώστε, καθώς οι τιμές για το *x* αυξάνονται, οι τιμές για το *y* επίσης να αυξηθούν.

- Αρνητικός συσχετισμός: Εάν τα *x* και *y* έχουν μια ισχυρή αρνητική γραμμική συσχέτιση, το *r* είναι κοντά στο -1. Μια τιμή *r* ακριβώς -1 δηλώνει μια τέλεια αρνητική εφαρμογή. Οι αρνητικές τιμές υποδηλώνουν μια σχέση μεταξύ x και y έτσι ώστε, καθώς οι τιμές για το *x* αυξάνονται, οι τιμές για το *y* μειώνονται.

- Καμία συσχέτιση: Εάν δεν υπάρχει γραμμική συσχέτιση ή ασθενής γραμμική συσχέτιση, το *r* είναι κοντά στο 0.
  Μια τιμή κοντά στο μηδέν σημαίνει ότι υπάρχει μια τυχαία, μη γραμμική σχέση μεταξύ των δύο μεταβλητών.

Μια *τέλεια* συσχέτιση ± 1 εμφανίζεται μόνο όταν όλα τα σημεία των δεδομένων βρίσκονται ακριβώς σε ευθεία γραμμή. Εάν *r* = +1, η κλίση αυτής της γραμμής είναι θετική ενώ όταν *r* = -1, η κλίση αυτής της γραμμής είναι αρνητική.

Οι αμερόληπτες εκτιμήτριες s_{xy} και s_x^2 δίνονται ως

$$s_{xy} = \frac{1}{n-1}\left(\sum_{i=1}^{n} x_i * y_i - n\bar{x}\bar{y}\right)$$

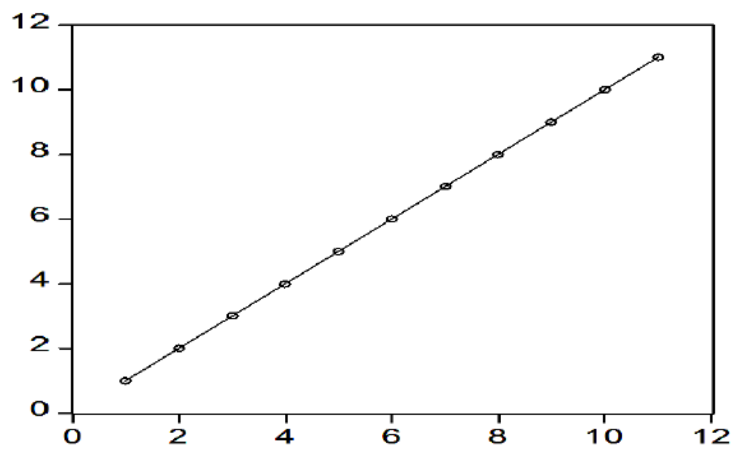
$$s_x^2 = \frac{1}{n-1}\left(\sum_{i=1}^{n} x_i^2 - n\bar{x^2}\right)$$

όπου $\bar{x}$ και $\bar{y}$ είναι οι δειγματικές μέσες τιμές των X και Y .

Από τα παραπάνω προκύπτει η έκφραση της εκτιμήτρια r

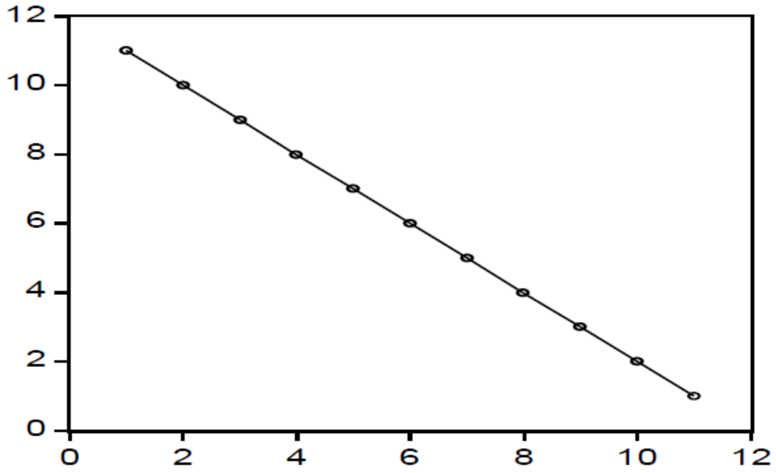
$$r = \frac{\sum_{i=1}^{n} x_i * y_i - n\bar{x}\bar{y}}{\sqrt{(\sum_{i=1}^{n} x_i^2 - n\bar{x}^2)(\sum_{i=1}^{n} y_i^2 - n\bar{y}^2)}}$$

Μια τέλεια θετική συσχέτιση φαίνεται στην εικόνα 4.1.



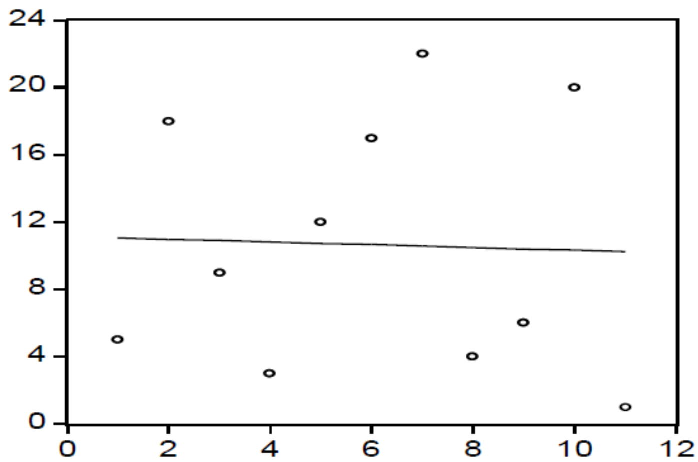
Εικόνα 4.1: Τέλεια θετική συσχέτιση.

Μια τέλεια αρνητική συσχέτιση φαίνεται στην εικόνα 4.2.



Εικόνα 4.2: Τέλεια αρνητική συσχέτιση.

Όταν δεν υπάρχει συσχέτιση των σημείων φαίνεται στην εικόνα 4.3.



Εικόνα 4.3: Καμία συσχέτιση σημείων.

4.1.3 Συνδιακύμανση.

Η συνδιακύμανση είναι ένα μέτρο της κοινής μεταβλητότητας δύο τυχαίων μεταβλητών. Αν οι μεγαλύτερες τιμές μιας μεταβλητής αντιστοιχούν κυρίως στις μεγαλύτερες τιμές της άλλης μεταβλητής και το ίδιο ισχύει και για τις μικρότερες τιμές, τότε οι μεταβλητές τείνουν να εμφανίζουν παρόμοια συμπεριφορά και η συνδιακύμανση είναι θετική. Στην αντίθετη περίπτωση, όταν οι μεγαλύτερες τιμές μιας μεταβλητής αντιστοιχούν κυρίως στις μικρότερες τιμές του άλλου, τότε οι μεταβλητές τείνουν να δείχνουν αντίθετη συμπεριφορά και η συνδιακύμανση είναι αρνητική. Το σημάδι της συνδιακύμανσης δείχνει επομένως την τάση στην γραμμική σχέση μεταξύ των μεταβλητών.

Ο τύπος της συνδιακύμανσης παρουσιάζεται στη συνέχεια.

$$Cov(X,Y) = \frac{1}{n}\sum_{i=1}^{n}\big(x_i - E(X)\big)(y_i - E(Y))$$

4.1.4 Συντελεστής προσδιορισμού R^2.

Ο συντελεστής προσδιορισμού R^2, είναι η αναλογία της διακύμανσης στην εξαρτώμενη μεταβλητή που είναι προβλέψιμη από την ανεξάρτητη μεταβλητή. Πρόκειται για στατιστικό στοιχείο που χρησιμοποιείται στο πλαίσιο στατιστικών μοντέλων που έχουν ως κύριο σκοπό είτε την πρόβλεψη μελλοντικών αποτελεσμάτων είτε τη δοκιμή υποθέσεων, με βάση άλλες σχετικές πληροφορίες. Παρέχει ένα μέτρο για το πόσο καλά παρατηρούνται τα αποτελέσματα από το μοντέλο, με βάση το ποσοστό της συνολικής διακύμανσης των αποτελεσμάτων που εξηγείται από το μοντέλο.

$$R^2 = \frac{SSR}{SST}$$

Παίρνει τιμές $0 \leq R^2 \leq 1$.

$R^2 = 0$ σημαίνει ότι η ερμηνευτική μεταβλητή X εξηγεί μηδενικό ποσοστό της διακύμανσης του Y.

$R^2 = 1$ σημαίνει ότι η X εξηγεί όλη τη διακύμανση της Y.

Ο συντελεστής προσδιορισμού αυξάνεται καθώς οι μεταβλητές εισχωρούν στο μοντέλο. Όταν η μεταβλητή που θα εισχωρήσει δεν θα παρουσιάσει αύξηση του συντελεστή αλλά σταθεροποιείται στη μέγιστη τιμή, η διαδικασία σταματάει. Έτσι βοηθάει ώστε να βρεθούν πόσες μεταβλητές θα εισχωρήσουν στο μοντέλο. Όμως πρέπει να πούμε ότι αυτή η τιμή του συντελεστή δεν είναι η βέλτιστη αλλά ανήκει σε ένα ικανοποιητικό επίπεδο.

4.1.5 Προσαρμοσμένος συντελεστής προσδιορισμού $\bar{R}^2$.

Ο προσαρμοσμένος συντελεστής προσδιορισμού $\bar{R}^2$ είναι μία παράμετρος όπου καθόρισαν οι αναλυτές ώστε να αντιμετωπίσει το πρόβλημα της ερμηνείας του συντελεστή προσδιορισμού R^2. Ο προσαρμοσμένος συντελεστής προσδιορισμού σε σχέση με τον συντελεστή προσδιορισμού δεν αυξάνεται καθώς εισάγονται οι μεταβλητές στο μοντέλο. Αντίθετα, αποδεικνύεται ότι εάν s μεταβλητές εισαχθούν σε ένα μοντέλο με αρχικό $\bar{R}^2$ ίσο με$\bar{R}_p^2$, τότε το $\bar{R}_{p+s}^2$ θα ξεπεράσει σε τιμή το $\bar{R}_p^2$, εάν και μόνο εάν οι συντελεστές μερικής συσχέτισης των εισαχθέντων s μεταβλητών ξεπερνούν τη μονάδα.(Σταμπολίδης,2010). Τέλος μπορούμε να πούμε ότι το καλύτερο μοντέλο παλινδρόμησης είναι αυτό με το μεγαλύτερο $\bar{R}_p^2$.

Προσαρμοσμένος συντελεστής προσδιορισμού για μοντέλο παλινδρόμησης με p όρους:

$$R_p^2 = 1 - \left(\frac{n-1}{n-p}\right) * (1 - R_p^2)$$

Όπου n το σύνολο των παρατηρήσεων των ανεξάρτητων μεταβλητών.

4.2 Πολλαπλή γραμμική παλινδρόμηση.

Ένα μοντέλο γραμμικής παλινδρόμησης που περιέχει δύο ή περισσότερες ανεξάρτητες μεταβλητές ονομάζεται μοντέλο πολλαπλής γραμμικής παλινδρόμησης. Στην συγκεκριμένη περίπτωση εξετάζουμε τη γραμμική σχέση μεταξύ μιας εξαρτημένης μεταβλητής και k ανεξάρτητων μεταβλητών.

Στη συνέχεια φαίνεται το μοντέλο της πολλαπλής γραμμικής παλινδρόμησης με k ανεξάρτητες μεταβλητές.

$$Y_i = a + \beta_1 X_{i1} + \beta_2 X_{i2} + .. + \beta_\kappa X_{ik} + \varepsilon_i$$

Όπου:

i=1, 2, …, n.

Y_i είναι η τιμή της εξαρτημένης μεταβλητής

X_{ik} είναι οι τιμές των ανεξάρτητων μεταβλητών

ε_i είναι ανεξάρτητες τυχαίες μεταβλητές που ακολουθούν N(0,σ^2) κατανομές.

4.2.1 Πολυσυγγραμικότητα.

Η πολυσυγγραμμικότητα αναφέρεται σε μια κατάσταση όπου ένας αριθμός ανεξάρτητων μεταβλητών σε ένα μοντέλο πολλαπλής παλινδρόμησης συσχετίζονται στενά μεταξύ τους. Η πολυσυγγραμμικότητα μπορεί να οδηγήσει σε αντικανονικά ή παραπλανητικά αποτελέσματα όταν ένας ερευνητής ή αναλυτής επιχειρεί να προσδιορίσει πόσο καλά μπορεί να χρησιμοποιηθεί η καθεμία από μια σειρά μεμονωμένων ανεξάρτητων μεταβλητών για την πρόβλεψη ή κατανόηση της εξαρτημένης μεταβλητής σε ένα στατιστικό μοντέλο.

Γενικά, η πολυσυγγραμμικότητα μπορεί να οδηγήσει σε μεγαλύτερα διαστήματα εμπιστοσύνης και λιγότερο αξιόπιστες τιμές πιθανοτήτων (τιμές P) για τις ανεξάρτητες μεταβλητές. Στην περίπτωση της πολλαπλής παλινδρόμησης μπορεί κάποιες ανεξάρτητες μεταβλητές να είναι γραμμικά εξαρτημένες οπότε ο πίνακας πληροφορίας $X^T X$ να έχει ορίζουσα μηδέν δηλαδή να μην αντιστρέφεται. Ένας τρόπος αντιμετώπισης του προβλήματος είναι η αφαίρεση κάποιων ανεξάρτητων μεταβλητών. Ωστόσο η αντιμετώπιση του παραπάνω προβλήματος από το SPSS γίνεται τυποποιώντας όλες τις μεταβλητές με τιμές στο (-1,1).(Δημητριάδης,2014). Ο έλεγχος της πολυσυγγραμμικότητας μπορεί να γίνει κοιτώντας το δείκτη VIF όπου $VIF = \frac{1}{1-R^2}$. Αν η τιμή του δείκτη είναι κοντά στο 2 τότε δεν υπάρχει το πρόβλημα της πολυσυγγραμμικότητας.

4.3 Μεταβλητές

Η μεταβλητή είναι ένας ειδικός τύπος που αναφέρεται σε ποσότητα ή ποιότητα μιας άγνωστης τιμής. Χωρίζονται σε ποσοτικές και ποιοτικές μεταβλητές.

4.3.1 Ποσοτικές μεταβλητές

Οι ποσοτικές μεταβλητές είναι μία από τις δύο κατηγορίες των μεταβλητών χρησιμοποιούνται σε μεγέθη που μπορούν να μετρηθούν(χιλιόμετρα, τετραγωνικά, χρόνος).

Οι ποσοτικές μεταβλητές διακρίνονται σε συνεχείς και ασυνεχείς.

- Συνεχείς μεταβλητές: παίρνουν οποιαδήποτε πραγματική τιμή με απροσδιόριστες ενδιάμεσες διαφορές.(π.χ. χρονομέτρηση).
- Ασυνεχείς μεταβλητές: παίρνουν συγκεκριμένες αριθμητικές τιμές οι οποίες έχουν καθοριστεί.(π.χ. μέτρηση παιδιών).

4.3.2 Ποιοτικές μεταβλητές

Οι ποιοτικές μεταβλητές χρησιμοποιούνται για να προσδιορίσουν παράγοντες που μεταβάλλονται κατά είδος.

Οι ποιοτικές μεταβλητές διακρίνονται σε κατηγορικές και διατάξιμες.

- Κατηγορικές μεταβλητές: παίρνουν τιμές για να προσδιορίσουν μεταβλητές που δεν υπάρχει σχέση μεταξύ τους.(π.χ. περιοχές της Αθήνας).

- Διατάξιμες μεταβλητές: παίρνουν τιμές για να προσδιορίσουν μια σειριακή σχέση.(π.χ. ευέλικτη ακύρωση, μέτρια ακύρωση, αυστηρή ακύρωση).

4.3.3 Ψευδομεταβλητές

Οι συγκεκριμένες μεταβλητές χρησιμοποιούνται για να ταξινομήσουν τα δεδομένα σε κατηγορίες. Παίρνουν τιμές 0 και 1 ώστε να δείξουν που ανήκει κάποια παρατήρηση. Παράδειγμα: Αν ένα κατάλυμα δεν διαθέτει είσοδο για αναπηρικό αμαξίδιο παίρνει τη τιμή 0 ενώ άμα διαθέτει θα πάρει τη τιμή 1.

4.4 Σύντομη περιγραφή της βηματικής παλινδρόμησης.

Στα πλαίσια του μοντέλου της βηματικής παλινδρόμησης επιτυγχάνεται η "βήμα – βήμα'' επαναληπτική διαδικασία κατασκευής ενός μοντέλου παλινδρόμησης στην οποία γίνεται αυτόματη εισχώρηση των ανεξάρτητων μεταβλητών. Η βηματική παλινδρόμηση μπορεί να πραγματοποιηθεί ή με τον σταδιακό έλεγχο κάθε ανεξάρτητης μεταβλητής και την επιλογή της στο μοντέλο ή με την επιλογή όλων των πιθανών ανεξάρτητων μεταβλητών και τη σταδιακή αφαίρεση των μη σημαντικών. Ο συντελεστής της μερικής συσχέτισης της ανεξάρτητης μεταβλητής σε σχέση με την εξαρτημένη μας καθορίζει τη σειρά εξέτασης των μεταβλητών για την τυχόν εισαγωγή τους στο μοντέλο. Οπότε η βηματική παλινδρόμηση είναι ένας συνδυασμός των μεθόδων προς τα εμπρός επιλογή (Forward Selection) και της μεθόδου προς τα πίσω αφαίρεση (Backward Elimination).

4.5 Η βηματική παλινδρόμηση (Stepwise Regression)

Ο Efroymson το (1960) ήταν ο πρώτος που παρουσίασε τον αλγόριθμο της βηματικής παλινδρόμησης . Είναι ένα συνδυασμός των μεθόδων, προσθήκης μεταβλητών και αφαίρεσης μεταβλητών. Στη διαδικασία αυτή γίνεται επιλογή ενός συνόλου Ω_i έστω το Ω_1, στο οποίο παρατηρείται η μεγαλύτερη συσχέτιση με την εξαρτημένη μεταβλητή Y, και υπολογίζεται η πρώτης τάξεως γραμμική συνάρτηση παλινδρόμησης $Y=f(\Omega_1)$. Έπειτα ελέγχεται αν η μεταβλητή αυτή είναι σημαντική. Αν δεν είναι, η διαδικασία τελειώνει και το καλύτερο μοντέλο της παλινδρόμησης είναι το $Y=\bar{Y}$. Αν είναι τότε συνεχίζεται η διαδικασία για την επιλογή της επόμενης μεταβλητής. Οπότε πρέπει να γίνεται έλεγχος των συντελεστών μερικών συσχετίσεων για όλες τις μεταβλητές, έστω (Ω_i, $i\neq1$) που δεν έχουν λάβει μέρος, έως αυτό το στάδιο, σε σχέση με την εξαρτημένη μεταβλητή. Αυτό σημαίνει ότι, οι Y και Ω_i έχουν προσαρμοστεί σε σχέση με την Ω_1 ως προς τη γραμμικότητα τους και στη συνέχεια υπολογίζεται η συσχέτιση για όλα τα $i\neq1$. Από μαθηματικής άποψης είναι ισοδύναμο με τον υπολογισμό των συσχετίσεων μεταξύ των υπολοίπων από τη συνάρτηση παλινδρόμησης και των υπολοίπων από τις κάθε j συναρτήσεις

παλινδρόμησης οι οποίες, δεν έχουν υπολογιστεί. Άρα επιλέγεται το Ω_i έστω το Ω_2 με το μεγαλύτερο συντελεστή μερικής συσχέτισης και υπολογίζεται μια καινούργια συνάρτηση παλινδρόμησης Y=f(Ω_1, Ω_2). Η συνάρτηση αυτή ελέγχεται για τη σημαντικότητα της και παρατηρείται η πιθανή αύξηση του συντελεστή προσδιορισμού R^2. Αφού έχει οριστεί ένα κρίσιμο επίπεδο για την επιλογή και την απόρριψη της μεταβλητής γίνεται σύγκριση των τιμών των μερικών F των συγκεκριμένων μεταβλητών και ακολουθεί η επιλογή ή η απόρριψη της.
Η συγκεκριμένη διαδικασία επαναλαμβάνεται σε κάθε στάδιο και μπορεί να παρατηρηθεί αφαίρεση μια μεταβλητής όπου στο προηγούμενο στάδιο είχε επιλεχθεί και αυτό γίνεται λόγο των συσχετίσεων των νέων μεταβλητών που μπήκαν στο μοντέλο. Συνεχίζεται η διαδικασία μέχρι όταν δεν μπορεί να εισαχθεί άλλη μεταβλητή και υπολογίζεται η κατάλληλη συνάρτηση βάση των μεταβλητών που έχουν εισαχθεί στο τελευταίο στάδιο.

4.5.1 Σταδιακή προσθήκη μεταβλητών (Forward Selection).

Στη διαδικασία της σταδιακής προσθήκης μεταβλητών, αρχικά στο μοντέλο δεν υπάρχει καμία μεταβλητή ώστε να γίνει η σωστή και σταδιακή επιλογή των μεταβλητών. Πρώτα γίνεται ο έλεγχος για να βρεθεί η μεταβλητή που παρουσιάζει τη μεγαλύτερη τιμή συσχέτισης με την εξαρτημένη μεταβλητή. Αφού έχει οριστεί ένα κρίσιμο επίπεδο για την επιλογή των μεταβλητών συγκρίνεται με αυτή, και επιλέγετε αν τηρεί την προϋπόθεση. Έτσι ελέγχεται η στατιστική σημαντικότητα της μεταβλητής που θέλουμε να εισχωρήσει στο μοντέλο. Στη συνέχεια επιλέγεται η αμέσως επόμενη μεταβλητή με την μεγαλύτερη συσχέτιση και αφού γίνει πάλι ο έλεγχος προϋπόθεση μπαίνει και αυτή στο μοντέλο. Η παραπάνω διαδικασία τελειώνει όταν η στατιστική σημαντικότητα της μεταβλητής που είναι να εισχωρήσει παραβιάζει την προϋπόθεση σημαντικότητας που έχουμε επιλέξει.

4.5.2 Σταδιακή αφαίρεση μεταβλητών (Backward Elimination).

Στη διαδικασία της σταδιακής αφαίρεσης μεταβλητών, αρχικά στο μοντέλο υπάρχει όλο το σύνολο των ανεξάρτητων μεταβλητών. Στη συνέχεια υπολογίζονται οι τιμές p των μεταβλητών και αφαιρούμε τη μεταβλητή που έχει p μεγαλύτερο από το acrit. Αφού αφαιρεθεί η συγκεκριμένη μεταβλητή ξανά υπολογίζονται οι τιμές των p και πραγματοποιείτε η παραπάνω διαδικασία. Αυτό γίνεται μέχρι καμία μεταβλητή να μην παραβιάζει την προϋπόθεση όπου έχουμε και το τελικό μοντέλο.

4.6 Κριτήρια εισαγωγής - εξαγωγής μεταβλητών.

4.6.1 Η στατιστική F.

Για τον υπολογισμό της τιμής αυτής πρέπει να γίνει σύγκριση των διακυμάνσεων των δυο δειγμάτων των οποίων θέλουμε να ελέγξουμε. Τα συνήθη παραδείγματα χρήσης των δοκιμών *F* περιλαμβάνουν τη μελέτη των ακόλουθων περιπτώσεων:

i) οι μέσοι όροι των δύο δειγμάτων είναι ίσοι.
ii) δεν υπάρχει καμία σχέση μεταξύ της εξαρτημένης μεταβλητής και κάποιων από τις ανεξάρτητες μεταβλητές.
iii) οι διακυμάνσεις των δειγμάτων που συγκρίνονται προέρχονται από δείγματα με κανονική κατανομή, με κοινή διακύμανση.

Για την περίπτωση της παλινδρόμησης ισχύει ότι F= MSR/MSE

Όπου:

MSR είναι το μέσο τετραγωνικό άθροισμα των σφαλμάτων (δηλαδή το τετραγωνικό άθροισμα διαιρεμένο με τους αντίστοιχους βαθμούς ελευθερίας.

MSE είναι το τετραγωνικό άθροισμα των υπολοίπων

Ο παρακάτω πίνακας δείχνει τα αθροίσματα αυτά σε συνάρτηση με τους βαθμούς ελευθερίας που τους αντιστοιχούν (Montgomery and Peck, 1992).

Ανάλυση διακύμανσης για τον έλεγχο σημαντικότητας στην παλινδρόμηση				
Πηγή διακύμανσης λόγω:	Τετραγωνικό Άθροισμα	Βαθμοί Ελευθερίας	Μέσο Τετραγωνικό Άθροισμα	Στατιστική Μηδενικής Υπόθεσης, F0
Παλινδρόμησης Υπολοίπων Συνολικά	SSR SSE Syy	K n-k-1 n-1	MSR MSE	MSR/MSE
n: Συνολικές διαθέσιμες μεταβλητές		k: Μεταβλητές που συμμετέχουν στο μοντέλο		

Ο έλεγχος σημαντικότητας σε μονόδρομη ανάλυση της διακύμανσης χρησιμοποιείται για να εκτιμηθεί εάν οι αναμενόμενες τιμές μιας ποσοτικής μεταβλητής μέσα σε αρκετές προκαθορισμένες ομάδες διαφέρουν μεταξύ τους. Άρα μπορούμε να πούμε ότι μια ανεξάρτητη μεταβλητή είναι σημαντική αν απορρίψουμε την μηδενική υπόθεση δηλαδή ότι υπάρχει ανεξάρτητη μεταβλητή με μηδενικό συντελεστή. Πιο αναλυτικά μπορούμε να πούμε, ότι θα γίνει σύγκριση της F_0 με την στατιστική F όπου εάν η τιμή της F_0 είναι μεγαλύτερη από την F απορρίπτουμε την μηδενική υπόθεση.

4.6.2 Η τιμή p.

Η στατιστική p (p-value) είναι ένας άλλος τρόπος στατιστικού ελέγχου της σημαντικότητας των ανεξάρτητων μεταβλητών. Στο συγκεκριμένο έλεγχο, ορίζεται ένα κατώφλι της p όπου συνήθως παίρνει την τιμή 0,05. Με τον ορισμό της συγκεκριμένης τιμής δεχόμαστε ότι οι μεταβλητές οι οποίες εισήχθησαν στο μοντέλο έχουν τυχαία κατανομή σε σύγκριση με τις υπόλοιπες με πιθανότητα 5%, οπότε μπορούμε να δεχτούμε ότι με πιθανότητα 95% ότι οι μεταβλητές οι οποίες εισήχθησαν στο μοντέλο είναι σημαντικές για αυτό. Τέλος πρέπει να αναφέρουμε ότι όσο μικρότερες τιμές της στατιστικής p παρατηρούμε στο μοντέλο μας, αυτό συνεπάγεται με περισσότερες ενδείξεις ότι δεν ισχύει η μηδενική υπόθεση δηλαδή ότι η ανεξάρτητη μεταβλητή δεν είναι σημαντική για το μοντέλο.

4.7 Κριτική στην επιλογή μοντέλου παλινδρόμησης.

Παρακάτω φαίνονται κάποια σημαντικά στοιχεία τα οποία οι ερευνητές θα πρέπει να λάβουν υπόψη ώστε να μπορούν να αποφασίσουν αν είναι σωστή η επιλογή των μεταβλητών από τη διαδικασία που ακολούθησαν.

1. Είναι η τελική συνάρτηση λογική;(Σταμπολίδης,2010) Δηλαδή, οι μεταβλητές που έχουν επιλεχθεί για το μοντέλο έχουν κάποιο νόημα σχετικά με το πρόβλημα;

2. Το μοντέλο που δημιουργήθηκε είναι χρήσιμο για τον σκοπό για τον οποίο αναπτύχθηκε; (Σταμπολίδης,2010)

3. Είναι οι συντελεστές παλινδρόμησης στο μοντέλο λογικοί; (Σταμπολίδης,2010)

4. Είναι οι διαγνωστικοί έλεγχοι για την καταλληλότητα και την επάρκεια του μοντέλου ικανοποιητικοί; (Σταμπολίδης,2010)

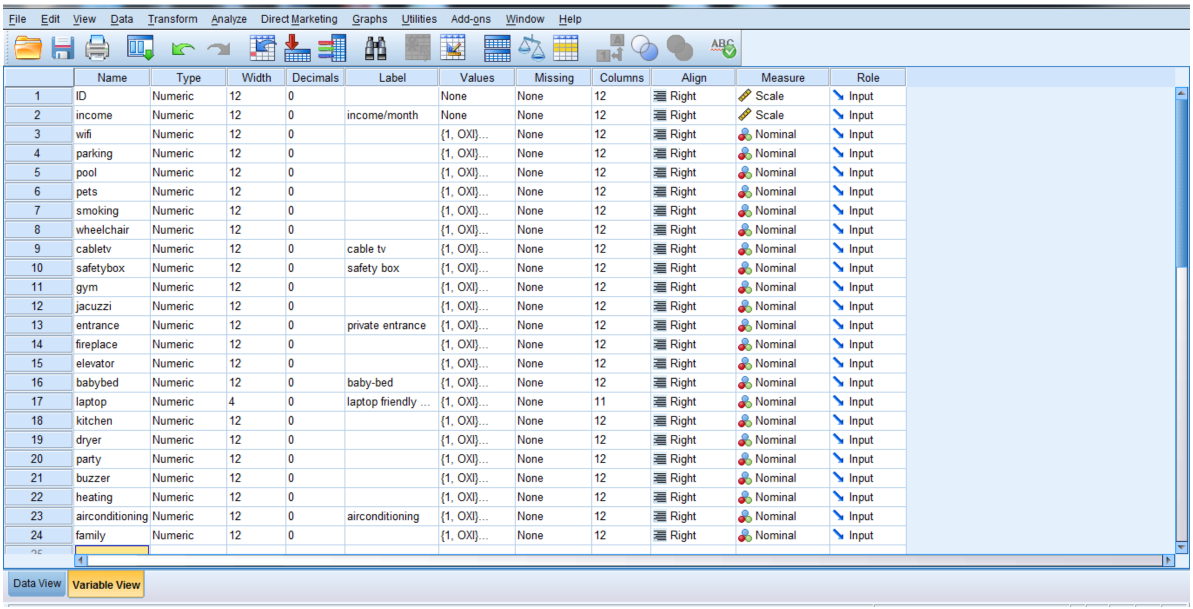
Καταλήγουμε ότι αν ο ερευνητής λάβει υπόψη σοβαρά τις παραπάνω ερωτήσεις τότε θα έχει μεγάλες πιθανότητες το μοντέλο του να είναι ικανοποιητικό. Αναφερόμαστε σε μεγάλες πιθανότητες διότι και αν ένα μοντέλο πληρεί τους διαγνωστικούς ελέγχους δεν είναι σίγουρο ότι θα εμφανιστούν καλά αποτελέσματα.

ΚΕΦΑΛΑΙΟ 5

5. Ανάλυση εφαρμογής της μεθόδου βηματικής παλινδρόμησης.

5.1 Ανάλυση των δεδομένων του παραδείγματος.

Η συγκεκριμένη μέθοδος θα εφαρμοστεί σε μια συλλογή δεδομένων η οποία περιέχει τις παροχές, των κατοικιών που υπάρχουν στην ιστοσελίδα της εταιρείας Airbnb. Σκοπός μας είναι να εξετάσουμε το εξαρτημένο μέγεθος μας δηλαδή τα έσοδα ανά μήνα με βάση τις ανεξάρτητες μεταβλητές όπου αντιστοιχούν στις παροχές των κατοικιών. Άρα θα δούμε πως επιδρά η κάθε παροχή των κατοικιών στα έσοδα. Από την ιστοσελίδα www.insideairbnb.com/athens έγινε η συλλογή δεδομένων για τα εκτιμώμενα έσοδα ανά μήνα που θα έχει ο ιδιοκτήτης, τις εκτιμώμενες ημέρες που θα βρίσκεται υπό κράτηση και για την διαθεσιμότητα του. Από το id της καταχώρησης συνδεόμασταν με την ιστοσελίδα του airbnb για να δούμε τις παροχές του κάθε καταλύματος, όπου φαίνονται στη συνεχεία και θα αναλύσουμε περισσότερο στο επόμενο κεφάλαιο. Η καταγραφή των δεδομένων έγινε σε ένα αρχείο excel όπου από τη περιοχή Data View πατώντας File → Open → Data έγινε η είσοδος των 603 καταλυμάτων στο SPSS. Στο Variable view έγινε ο ορισμός των μεταβλητών όπως φαίνεται στην εικόνα 5.1.



	Name	Type	Width	Decimals	Label	Values	Missing	Columns	Align	Measure	Role
1	ID	Numeric	12	0		None	None	12	Right	Scale	Input
2	income	Numeric	12	0	income/month	None	None	12	Right	Scale	Input
3	wifi	Numeric	12	0		{1, OXI}...	None	12	Right	Nominal	Input
4	parking	Numeric	12	0		{1, OXI}...	None	12	Right	Nominal	Input
5	pool	Numeric	12	0		{1, OXI}...	None	12	Right	Nominal	Input
6	pets	Numeric	12	0		{1, OXI}...	None	12	Right	Nominal	Input
7	smoking	Numeric	12	0		{1, OXI}...	None	12	Right	Nominal	Input
8	wheelchair	Numeric	12	0		{1, OXI}...	None	12	Right	Nominal	Input
9	cabletv	Numeric	12	0	cable tv	{1, OXI}...	None	12	Right	Nominal	Input
10	safetybox	Numeric	12	0	safety box	{1, OXI}...	None	12	Right	Nominal	Input
11	gym	Numeric	12	0		{1, OXI}...	None	12	Right	Nominal	Input
12	jacuzzi	Numeric	12	0		{1, OXI}...	None	12	Right	Nominal	Input
13	entrance	Numeric	12	0	private entrance	{1, OXI}...	None	12	Right	Nominal	Input
14	fireplace	Numeric	12	0		{1, OXI}...	None	12	Right	Nominal	Input
15	elevator	Numeric	12	0		{1, OXI}...	None	12	Right	Nominal	Input
16	babybed	Numeric	12	0	baby-bed	{1, OXI}...	None	12	Right	Nominal	Input
17	laptop	Numeric	4	0	laptop friendly ...	{1, OXI}...	None	11	Right	Nominal	Input
18	kitchen	Numeric	12	0		{1, OXI}...	None	12	Right	Nominal	Input
19	dryer	Numeric	12	0		{1, OXI}...	None	12	Right	Nominal	Input
20	party	Numeric	12	0		{1, OXI}...	None	12	Right	Nominal	Input
21	buzzer	Numeric	12	0		{1, OXI}...	None	12	Right	Nominal	Input
22	heating	Numeric	12	0		{1, OXI}...	None	12	Right	Nominal	Input
23	airconditioning	Numeric	12	0	airconditioning	{1, OXI}...	None	12	Right	Nominal	Input
24	family	Numeric	12	0		{1, OXI}...	None	12	Right	Nominal	Input

Εικόνα 5.1: Περιοχή δήλωσης μεταβλητών στο SPSS.

Στη συνέχεια θα αναλύσουμε τις ανεξάρτητες μεταβλητές.

Οι μεταβλητές στη συγκεκριμένη ανάλυση χωρίζονται σε δυο κατηγορίες, τις ποσοτικές και τις ποιοτικές. Η εξαρτημένη μεταβλητή τα έσοδα ανά μήνα (income/ month) είναι μια ποσοτική μεταβλητή αφού μας δείχνει ένα ποσοτικό μέγεθος. Οι ανεξάρτητες μεταβλητές όπου φαίνονται παρακάτω είναι ποιοτικές μεταβλητές διότι αναφέρονται σε ένα διακριτό είδος.

1) WIFI
2) PARKING
3) POOL
4) PETS
5) SMOKING
6) WHEELCHAIR
7) CABLE TV
8) GYM
9) JACUZZI
10) PRIVATE ENTRANCE
11) FIREPLACE
12) ELEVATOR
13) BABY BED
14) LAPTOP FRIENDLY WORKSPACE
15) KITCHEN
16) DRYER
17) PARTY
18) BUZZER
19) HEATING

20) AIR CONDITIONING
21) FAMILY

Δηλαδή οι ανεξάρτητες μεταβλητές αναφέρονται στο αν το κατάλυμα διαθέτει 1) ασύρματο ίντερνετ 2) χώρο παρκαρίσματος 3) πισίνα 4) αν επιτρέπονται τα κατοικίδια 5) επιτρέπετε το κάπνισμα, 6) αν έχει πρόσβαση αναπηρικό αμαξίδιο, 7) αν διαθέτει δορυφορική τηλεόραση 8) γυμναστήριο 9) τζακούζι 10) ιδιωτική είσοδο 11) τζάκι 12) ασανσέρ 13) κρεβάτι μωρού 14) επιφάνεια για τοποθέτηση του laptop 15) κουζίνα 16) στεγνωτήριο 17) αν επιτρέπονται πάρτι/εκδηλώσεις 18) αν διαθέτει κουδούνι/ασύρματη ενδοεπικοινωνία 19) θέρμανση 20) κλιματισμό και τέλος 20) αν είναι κατάλληλη προς ενοικίαση από μια οικογένεια.

Για να μπορέσει το SPSS να αναγνωρίζει ποιες παροχές διαθέτει κάθε κατάλυμα θα πρέπει να γίνει μια κωδικοποίηση των μεταβλητών. Δηλαδή για μια κατοικία η οποία δεν θα έχει κάποια παροχή θα κωδικοποιηθεί ¨ΟΧΙ¨=0 και για μια που θα έχει θα κωδικοποιηθεί ¨ΝΑΙ¨=1.

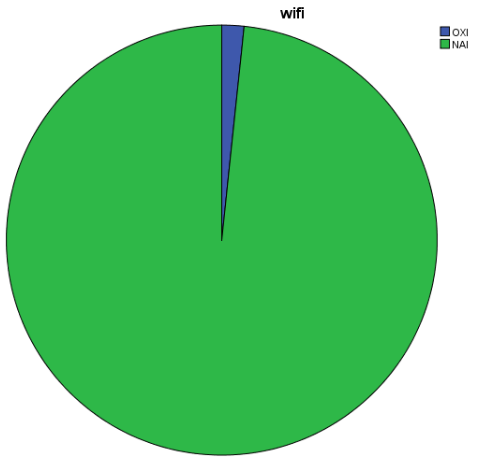
5.2 Περιγραφικά μέτρα ποιοτικών και ποσοτικών μεταβλητών.

Ένα ενδιαφέρον κομμάτι είναι να δούμε τα περιγραφικά μέτρα των ποιοτικών μεταβλητών, όπου θα δούμε από τις 603 καταχωρήσεις πόσες διαθέτουν ή όχι κάθε παροχή που αναφέραμε πριν. Από τη περιοχή Data View πατώντας analyze → descriptive statistics→frequencies επιλέγουμε τις ποιοτικές μεταβλητές που θέλουμε να αναλύσουμε και τις τοποθετούμε στο variable(s). Παρακάτω βλέπουμε τους πίνακες συχνοτήτων για κάθε μια ανεξάρτητη μεταβλητή.

wifi

		Frequency	Percent	Valid Percent	Cumulative Percent
Valid	ΟΧΙ	10	1,7	1,7	1,7
	ΝΑΙ	593	98,3	98,3	100,0
	Total	603	100,0	100,0	

Πίνακας 5.1: Συχνότητες μεταβλητής wifi.



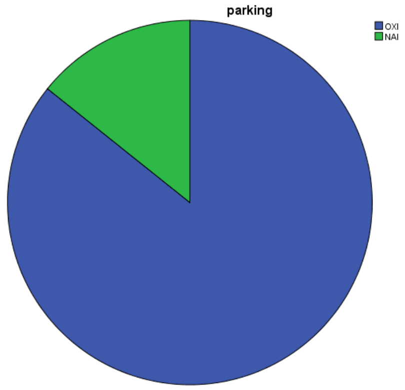
Εικόνα 5.2: Πίτα συχνοτήτων μεταβλητής wifi.

Βλέπουμε ότι 98,3% διαθέτει ασύρματο ίντερνετ ενώ το 1,7% δεν διαθέτει

parking

		Frequency	Percent	Valid Percent	Cumulative Percent
Valid	OXI	517	85,7	85,7	85,7
	NAI	86	14,3	14,3	100,0
	Total	603	100,0	100,0	

Πίνακας 5.2: Συχνότητες μεταβλητής parking.



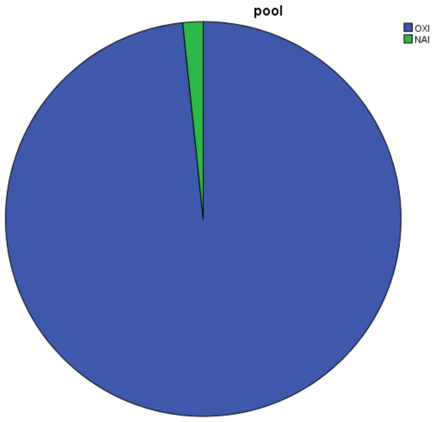
Εικόνα 5.3: Πίτα συχνοτήτων μεταβλητής parking.

Το 85,7% των κατοικιών διαθέτει χώρο παρκαρίσματος ενώ το 14,3% δεν διαθέτει.

pool

		Frequency	Percent	Valid Percent	Cumulative Percent
Valid	OXI	593	98,3	98,3	98,3
	NAI	10	1,7	1,7	100,0
	Total	603	100,0	100,0	

Πίνακας 5.3: Συχνότητες μεταβλητής pool.



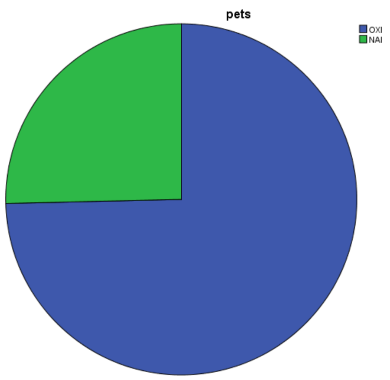
Εικόνα 5.4: Πίτα συχνοτήτων μεταβλητής pool.

Εδώ βλέπουμε ότι μόνο το 1,7% των κατοικιών έχει πισίνα ενώ το 98,3% δεν έχει.

pets

		Frequency	Percent	Valid Percent	Cumulative Percent
Valid	ΟΧΙ	450	74,6	74,6	74,6
	ΝΑΙ	153	25,4	25,4	100,0
	Total	603	100,0	100,0	

Πίνακας 5.4: Συχνότητες μεταβλητής pets.



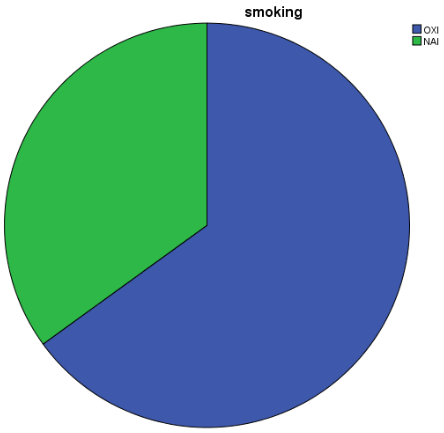
Εικόνα 5.5: Πίτα συχνοτήτων μεταβλητής pets.

Παρατηρούμε ότι το 25,4% των ιδιοκτητών επιτρέπουν τα κατοικίδια ενώ το 74,6 % δεν τα επιτρέπει.

smoking

		Frequency	Percent	Valid Percent	Cumulative Percent
Valid	ΟΧΙ	392	65,0	65,0	65,0
	ΝΑΙ	211	35,0	35,0	100,0
	Total	603	100,0	100,0	

Πίνακας 5.5: Συχνότητες μεταβλητής smoking.



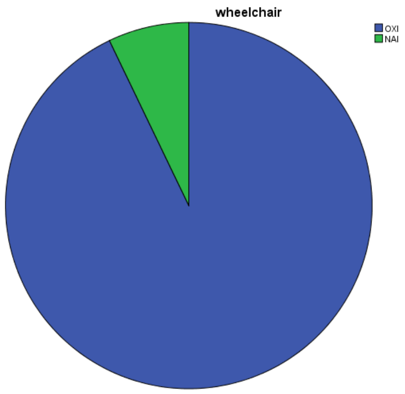
Εικόνα 5.6: Πίτα συχνοτήτων μεταβλητής smoking.

Για το κάπνισμα βλέπουμε ότι στο 35% των καταλυμάτων επιτρέπετε το κάπνισμα ενώ στο 65% απαγορεύετε.

wheelchair

		Frequency	Percent	Valid Percent	Cumulative Percent
Valid	OXI	560	92,9	92,9	92,9
	NAI	43	7,1	7,1	100,0
	Total	603	100,0	100,0	

Πίνακας 5.6: Συχνότητες μεταβλητής wheelchair.



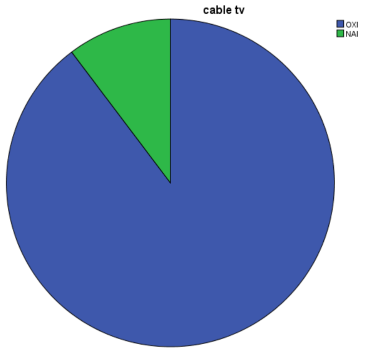
Εικόνα 5.7: Πίτα συχνοτήτων μεταβλητής wheelchair.

Εδώ βλέπουμε ότι μόνο στο 7,1% των καταλυμάτων έχει πρόσβαση αναπηρικό αμαξίδιο ενώ στο 92,9% δεν υπάρχει πρόσβαση.

cable tv

		Frequency	Percent	Valid Percent	Cumulative Percent
Valid	OXI	541	89,7	89,7	89,7
	NAI	62	10,3	10,3	100,0
	Total	603	100,0	100,0	

Πίνακας 5.7: Συχνότητες μεταβλητής cable tv.



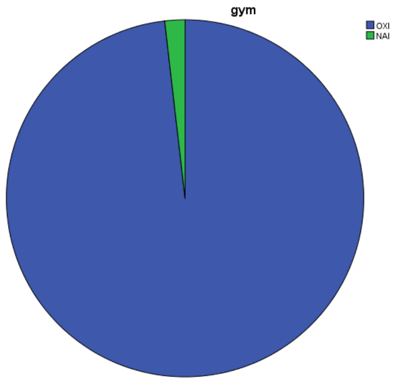
Εικόνα 5.8: Πίτα συχνοτήτων μεταβλητής cable tv.

Δορυφορική τηλεόραση έχει μόνο το 10,3% των καταλυμάτων ενώ το 89,7% δεν διαθέτει.

gym

		Frequency	Percent	Valid Percent	Cumulative Percent
Valid	OXI	592	98,2	98,2	98,2
	NAI	11	1,8	1,8	100,0
	Total	603	100,0	100,0	

Πίνακας 5.8: Συχνότητες μεταβλητής gym.



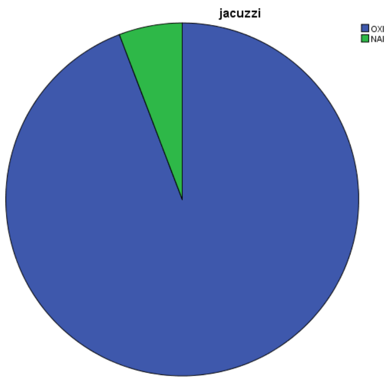
Εικόνα 5.9: Πίτα συχνοτήτων μεταβλητής gym.

Προφανές ήταν ότι τα καταλύματα που διαθέτουν γυμναστήριο θα κατείχαν μικρό ποσοστό 1,8% και με 98,2% αυτές που δεν έχουν.

jacuzzi

		Frequency	Percent	Valid Percent	Cumulative Percent
Valid	OXI	568	94,2	94,2	94,2
	NAI	35	5,8	5,8	100,0
	Total	603	100,0	100,0	

Πίνακας 5.9: Συχνότητες μεταβλητής Jacuzzi.



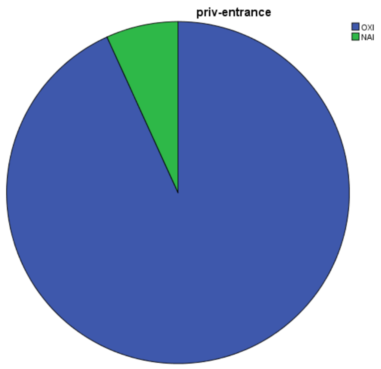
Εικόνα 5.10: Πίτα συχνοτήτων μεταβλητής Jacuzzi.

Jacuzzi διαθέτει μόνο το 5,8% ενώ το 94,2% δεν έχει.

priv-entrance

		Frequency	Percent	Valid Percent	Cumulative Percent
Valid	ΟΧΙ	562	93,2	93,2	93,2
	ΝΑΙ	41	6,8	6,8	100,0
	Total	603	100,0	100,0	

Πίνακας 5.10: Συχνότητες μεταβλητής private entrance.



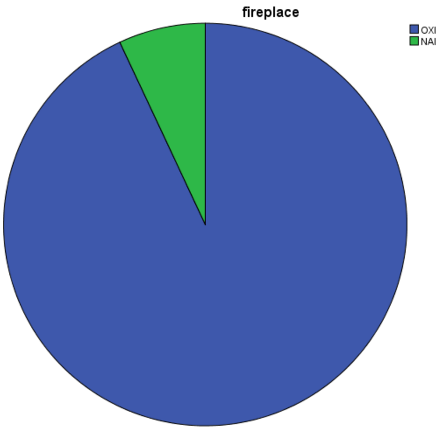
Εικόνα 5.11: Πίτα συχνοτήτων μεταβλητής private entrance.

Εδώ βλέπουμε ότι το 6,8% έχουν ιδιωτική είσοδο για τα καταλύματα ενώ το 93,2% δεν έχει.

fireplace

		Frequency	Percent	Valid Percent	Cumulative Percent
Valid	ΟΧΙ	561	93,0	93,0	93,0
	ΝΑΙ	42	7,0	7,0	100,0
	Total	603	100,0	100,0	

Πίνακας 5.11: Συχνότητες μεταβλητής fireplace.



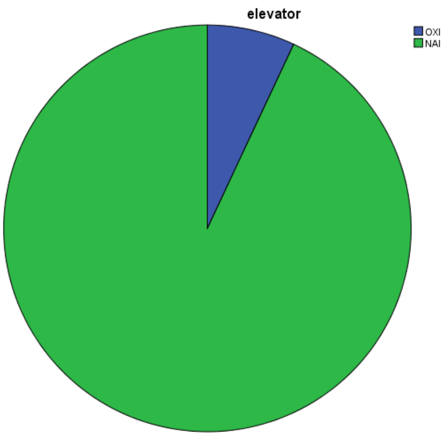
Εικόνα 5.12: Πίτα συχνοτήτων μεταβλητής fireplace.

Επίσης βλέπουμε ότι τζάκι διαθέτει μόνο το 7% ενώ το υπόλοιπο 93% των καταλυμάτων δεν διαθέτει.

elevator

		Frequency	Percent	Valid Percent	Cumulative Percent
Valid	OXI	42	7,0	7,0	7,0
	NAI	561	93,0	93,0	100,0
	Total	603	100,0	100,0	

Πίνακας 5.12: Συχνότητες μεταβλητής elevator.



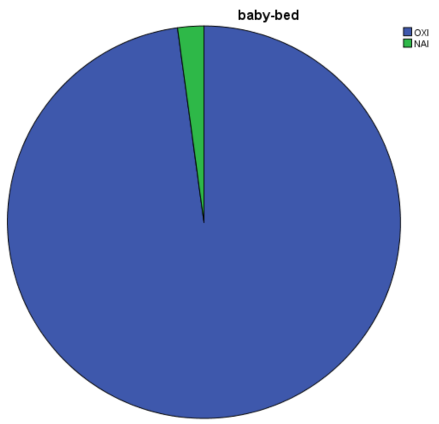
Εικόνα 5.13: Πίτα συχνοτήτων μεταβλητής elevator.

Ασανσέρ έχει το 93% των καταλυμάτων ενώ το 7% δεν έχει.

baby-bed

		Frequency	Percent	Valid Percent	Cumulative Percent
Valid	ΟΧΙ	590	97,8	97,8	97,8
	ΝΑΙ	13	2,2	2,2	100,0
	Total	603	100,0	100,0	

Πίνακας 5.13: Συχνότητες μεταβλητής baby bed.



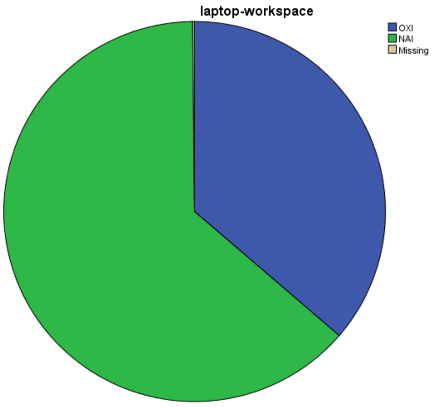
Εικόνα 5.14: Πίτα συχνοτήτων μεταβλητής baby bed.

Παιδικό κρεβάτι έχει μόνο το 2,2% των καταλυμάτων ενώ το υπόλοιπο 97,8% δεν έχει

laptop-workspace

		Frequency	Percent	Valid Percent	Cumulative Percent
Valid	ΟΧΙ	219	36,3	36,3	36,3
	ΝΑΙ	384	63,7	63,7	100,0
	Total	603	100,0	100,0	

Πίνακας 5.14: Συχνότητες μεταβλητής laptop workspace.



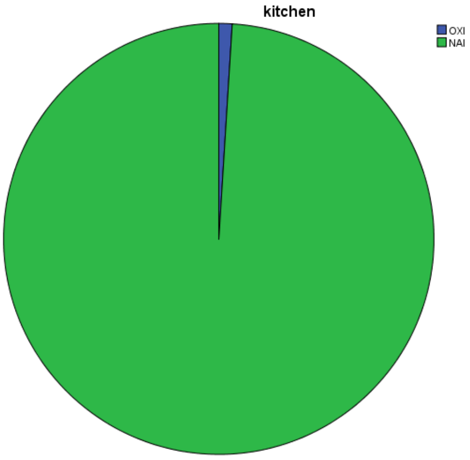
Εικόνα 5.14: Πίτα συχνοτήτων μεταβλητής laptop workspace.

Επιφάνεια εργασίας για φορητό υπολογιστή διαθέτει το 63,7% ενώ το υπόλοιπο 36,3% δεν διαθέτει.

kitchen

		Frequency	Percent	Valid Percent	Cumulative Percent
Valid	OXI	6	1,0	1,0	1,0
	NAI	597	99,0	99,0	100,0
	Total	603	100,0	100,0	

Πίνακας 5.15: Συχνότητες μεταβλητής kitchen.



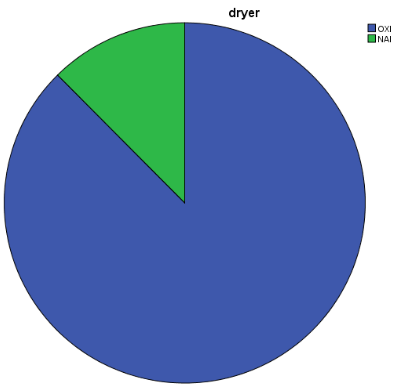
Εικόνα 5.16: Πίτα συχνοτήτων μεταβλητής kitchen.

Προφανές είναι ότι το 99% των καταλυμάτων θα είχε κουζίνα ενώ το υπόλοιπο 1% δεν θα είχε.

dryer

		Frequency	Percent	Valid Percent	Cumulative Percent
Valid	ΟΧΙ	528	87,6	87,6	87,6
	ΝΑΙ	75	12,4	12,4	100,0
	Total	603	100,0	100,0	

Πίνακας 5.16: Συχνότητες μεταβλητής dryer.



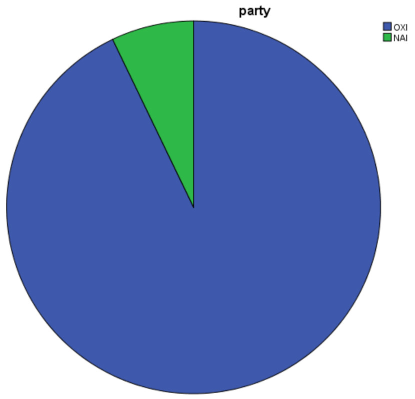
Εικόνα 5.17: Πίτα συχνοτήτων μεταβλητής dryer.

Ακόμα βλέπουμε ότι το 12,4% των καταλυμάτων διαθέτει στεγνωτήριο ενώ το 87,6% δεν συμπεριλαμβάνει.

party

		Frequency	Percent	Valid Percent	Cumulative Percent
Valid	ΟΧΙ	560	92,9	92,9	92,9
	ΝΑΙ	43	7,1	7,1	100,0
	Total	603	100,0	100,0	

Πίνακας 5.17: Συχνότητες μεταβλητής party.



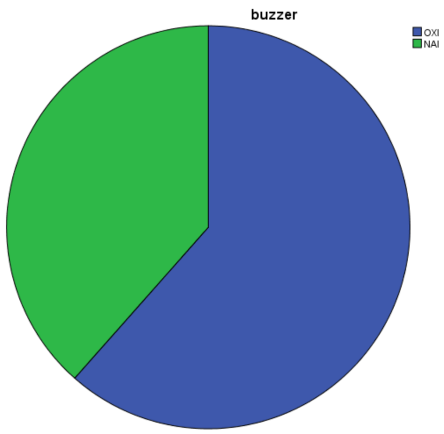
Εικόνα 5.18: Πίτα συχνοτήτων μεταβλητής party.

Εδώ βλέπουμε ότι τα πάρτι/εκδηλώσεις επιτρέπονται μόνο στο 7,1% ενώ στο υπόλοιπο 92,9% απαγορεύονται.

buzzer

		Frequency	Percent	Valid Percent	Cumulative Percent
Valid	OXI	371	61,5	61,5	61,5
	NAI	232	38,5	38,5	100,0
	Total	603	100,0	100,0	

Πίνακας 5.18: Συχνότητες μεταβλητής buzzer.



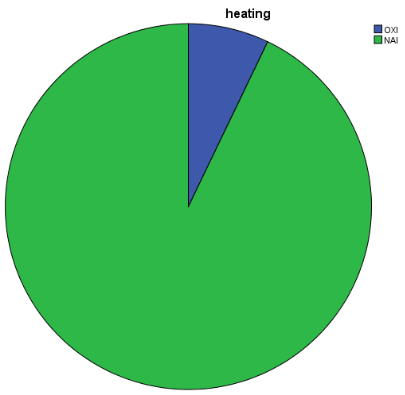
Εικόνα 5.19: Πίτα συχνοτήτων μεταβλητής buzzer.

Επίσης βλέπουμε ότι το 38,5% των καταλυμάτων έχουν κουδούνι/ασύρματη ενδοεπικοινωνία ενώ το υπόλοιπο 61,5% δεν έχει.

heating

		Frequency	Percent	Valid Percent	Cumulative Percent
Valid	OXI	43	7,1	7,1	7,1
	NAI	560	92,9	92,9	100,0
	Total	603	100,0	100,0	

Πίνακας 5.19: Συχνότητες μεταβλητής heating.



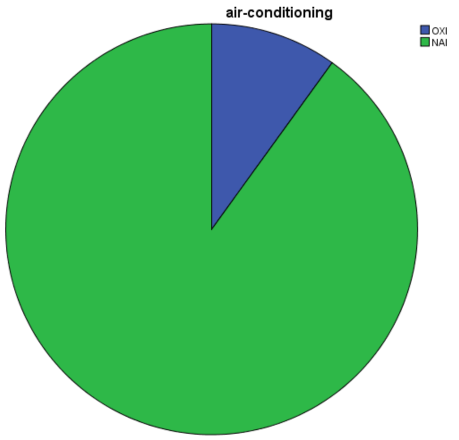
Εικόνα 5.20: Πίτα συχνοτήτων μεταβλητής heating

Θέρμανση παρατηρούμε ότι διαθέτει ένα μεγάλο ποσοστό, το 92,9% ενώ το 7,1% των καταλυμάτων δεν διαθέτει.

air-conditioning

		Frequency	Percent	Valid Percent	Cumulative Percent
Valid	OXI	60	10,0	10,0	10,0
	NAI	543	90,0	90,0	100,0
	Total	603	100,0	100,0	

Πίνακας 5.20: Συχνότητες μεταβλητής air conditioning.



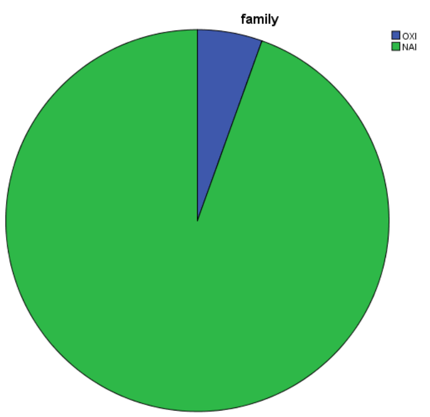
Εικόνα 5.21: Πίτα συχνοτήτων μεταβλητής air conditioning.

Εδώ βλέπουμε ότι κλιματισμό στα καταλύματα διαθέτει το 90% ενώ το υπόλοιπο 10% δεν παρέχει.

family

		Frequency	Percent	Valid Percent	Cumulative Percent
Valid	OXI	33	5,5	5,5	5,5
	NAI	570	94,5	94,5	100,0
	Total	603	100,0	100,0	

Πίνακας 5.21: Συχνότητες μεταβλητής family.



Εικόνα 5.22: Πίτα συχνοτήτων μεταβλητής family.

Τέλος παρατηρούμε ότι το 94,5% είναι κατάλληλο για οικογένειες ενώ το 5,5% δεν απευθύνεται προς αυτές.

Τώρα θα εξάγουμε τα περιγραφικά μέτρα της ποσοτικής μας μεταβλητής. Όπως και πριν από τη περιοχή Data View πατώντας analyze → descriptive statistics→frequencies επιλέγουμε την ποσοτική μεταβλητή που θέλουμε να αναλύσουμε και την τοποθετούμε στο variable(s). Ακόμα από το statistics επιλέγουμε τα *quartiles, minimum ,maximum* και *range.*

Από το Output παίρνουμε το πίνακα 5.22.

Statistics

income/month

N	Valid	603
	Missing	0
Range		3636
Minimum		0
Maximum		3636
Percentiles	25	97,00
	50	278,00
	75	585,00

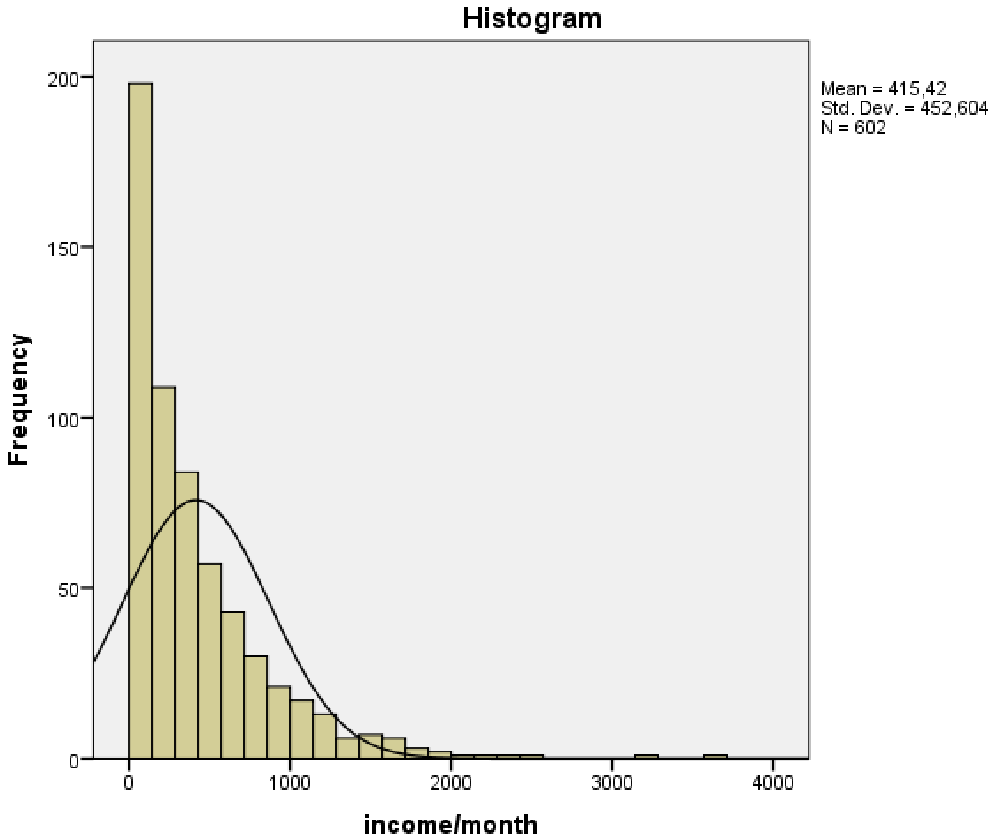
Πίνακας 5.22: Συχνότητες μεταβλητής income/ month.

Όπως βλέπουμε από το πίνακα το εύρος είναι 3.636, το ελάχιστο εισόδημα είναι 0 ευρώ και το μέγιστο 3.636 ευρώ. Επίσης παρατηρούμε ότι το 25% του δείγματος διαθέτει εισόδημα έως 97 ευρώ, το 50% διαθέτει έως 278 ευρώ και το 75% του δείγματος διαθέτει εισόδημα έως 585 ευρώ.

5.3 Έλεγχος προϋποθέσεων βηματικής παλινδρόμησης.

Αφού έχει γίνει η κατανόηση των δεδομένων τώρα θα γίνει ο έλεγχος των προϋποθέσεων. Αρχικά ελέγχουμε την κατανομή της Εξαρτημένης μεταβλητής, αν παρουσιάζει κάποια ασυμμετρία όπου είναι μια από τις βασικότερες προϋποθέσεις. Η κανονικότητα ενός εξεταζόμενου μεγέθους μπορεί να απεικονιστεί γραφικά μέσω του Ιστογράμματος και να επιβεβαιωθεί στατιστικά μέσω του μη παραμετρικού ελέγχου Kolmogorov-Smirnof.

Παρακάτω φαίνεται το ιστόγραμμα.



Εικόνα 5.23: Ιστόγραμμα εμφάνισης κανονικότητας εξαρτημένης μεταβλητής.

Παρατηρούμε ότι υπάρχει μια θετική ασυμμετρία και συγκεκριμένα βλέπουμε να εμφανίζονται κάποιες ακραίες τιμές, πράγμα που δεν θέλουμε ώστε να πληρούμε την προϋπόθεση της κανονικής κατανομής.

Αυτό που υποθέτουμε από το Ιστόγραμμα το επιβεβαιώνουμε και στατιστικά μέσω του μη παραμετρικού ελέγχου Kolmogorov-Smirnof δηλαδή θα απορρίψουμε ή θα δεχτούμε την αρχική υπόθεση αν η κατανομή είναι κανονική. Πατώντας Analyze -> Nonparametric test ->legacy dialogs -> 1 sample K-S. Στη συνέχεια βάζουμε τη μεταβλητή που θέλουμε να εξετάσουμε στη λίστα Test Variable List όπου στη περίπτωση μας είναι η income/month και πραγματοποιείται ο έλεγχος Kolmogorov-Smirnof.

Βάση του ελέγχου αυτού μπορούμε να δεχτούμε ή να απορρίψουμε την αρχική υπόθεση ότι η κατανομή είναι κανονική. Από την τιμή του sig. στον πίνακα που ακολουθεί απορρίπτουμε ή όχι την αρχική υπόθεση. Αν η τιμή του sig. είναι μεγαλύτερη του 0,05 η κατανομή είναι κανονική. Παρακάτω φαίνονται τα αποτελέσματα από τον έλεγχο που πήραμε από το Output.

One-Sample Kolmogorov-Smirnov Test

		income/month
N		602
Normal Parameters[a,b]	Mean	415,42
	Std. Deviation	452,604
Most Extreme Differences	Absolute	,179
	Positive	,142
	Negative	-,179
Kolmogorov-Smirnov Z		4,400
Asymp. Sig. (2-tailed)		,000

a. Test distribution is Normal.

b. Calculated from data.

Πίνακας 5.23: One-Sample Kolmogorov-Smirnov Test.

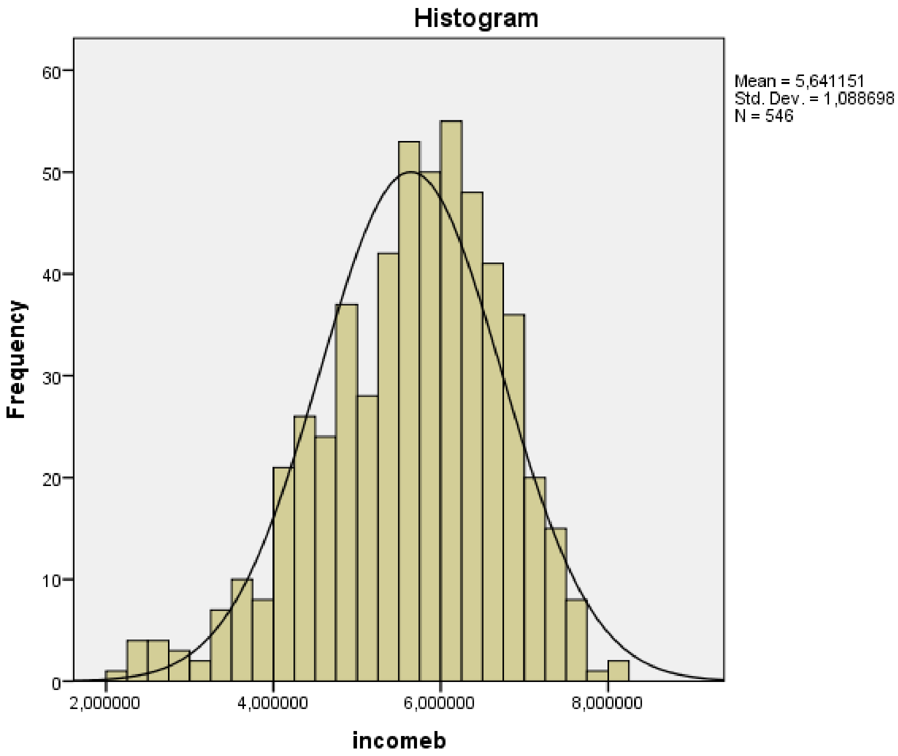
Βλέπουμε ότι η τιμή του sig είναι μικρότερη του 0,05 άρα η κατανομή δεν είναι κανονική, πράγμα που μας εμποδίζει να συνεχίσουμε την παλινδρόμηση μας. Για να αντιμετωπίσουμε το πρόβλημα πρέπει να τροποποιήσουμε τη μεταβλητή μας.

Οι πιο συνηθισμένες τροποποιήσεις είναι:

1) Ο υπολογισμός του Τετραγώνου,

2) της ρίζας ή του λογαρίθμου της εξαρτημένης μεταβλητής.

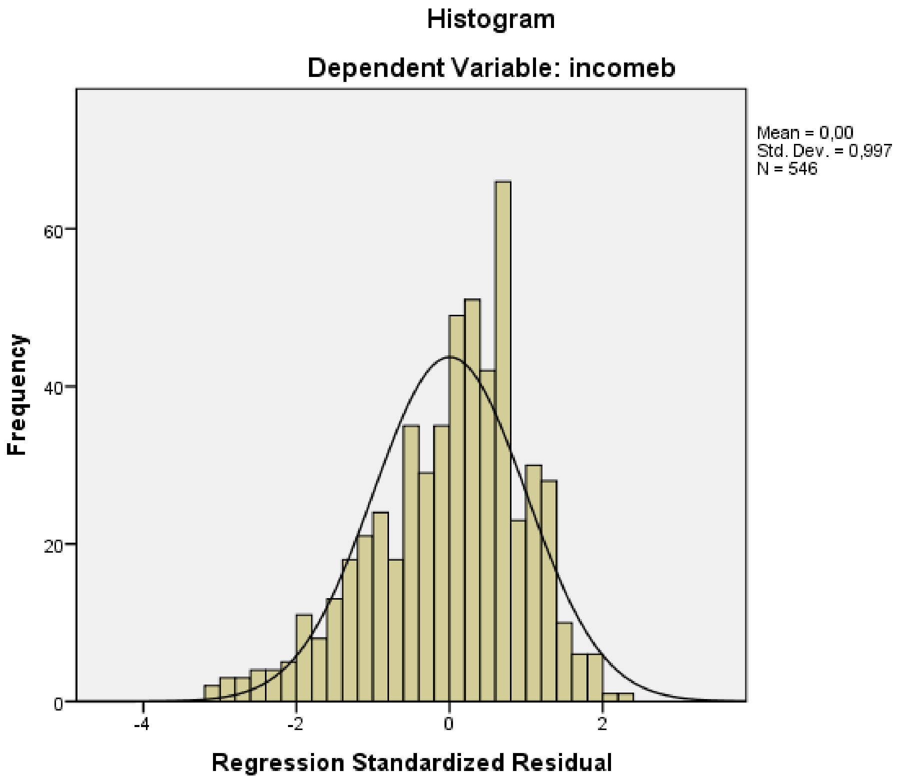
Στο συγκεκριμένο παράδειγμα θα επιλέξουμε να λογαριθμίσουμε την εξαρτημένη μεταβλητή από το Transform->Compute Variable. Αφού πραγματοποιήσουμε αυτή τη διαδικασία δηλαδή δημιουργήσαμε μια νέα μεταβλητή με στοιχεία τους λογαρίθμους των αρχικών τιμών, κάνουμε πάλι τον έλεγχο που πραγματοποιήσαμε προηγουμένως. Παρακάτω εμφανίζουμε το ιστόγραμμα των λογαριθμισμένων τιμών.



Εικόνα 5.24: Ιστόγραμμα εμφάνισης κανονικότητας εξαρτημένης μεταβλητής.

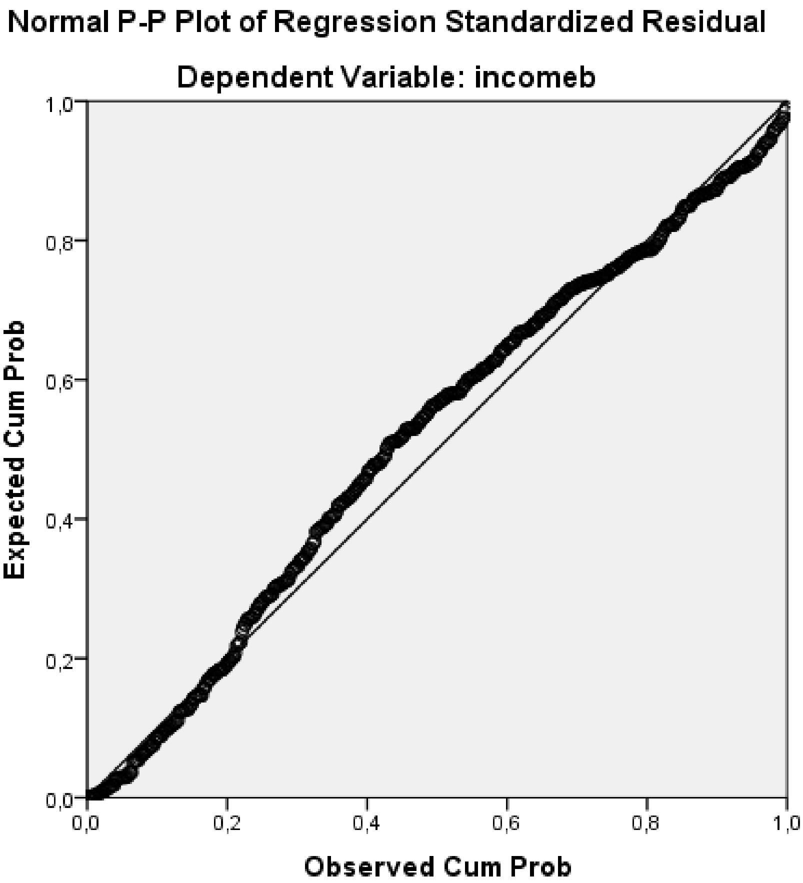
Από το ιστόγραμμα δεν απορρίπτουμε την υπόθεση της κανονικότητας διότι η εξαρτημένη μεταβλητή πλησιάζει πάρα πολύ κοντά στην κανονική κατανομή χωρίς να παρουσιάζει ακραίες τιμές, όπου μπορούμε να ισχυριστούμε ότι την ακολουθεί. Τέλος πρέπει να αναφέρουμε ότι είναι πάρα πολύ σημαντική προϋπόθεση διότι από αυτή γνωρίζουμε την εγκυρότητα της μεθόδου.

- Κανονικότητα καταλοίπων: γίνεται έλεγχος αν τα κατάλοιπα ακολουθούν κανονική κατανομή. Η ύπαρξη κανονικής κατανομής των καταλοίπων δημιουργεί μη ύπαρξη συσχέτισης τους. Αυτό φαίνεται από το ιστόγραμμα που παρουσιάζουμε παρακάτω και εξασφαλίζουμε αυτή την προϋπόθεση.



Εικόνα 5.25: Ιστόγραμμα εμφάνισης κανονικότητας καταλοίπων.

Σε ένα επόμενο βήμα πρέπει να ελέγξουμε αν οι ανεξάρτητες μεταβλητές που θα χρησιμοποιηθούν για την επεξήγηση της συμπεριφοράς της εξαρτημένης μεταβλητής συνδέονται γραμμικά με το εξαρτημένο μέγεθος. Δηλαδή θα έπρεπε στην βέλτιστή περίπτωση να είναι μια ευθεία γραμμή. Ωστόσο, αυτό είναι πολύ δύσκολο να συμβεί σε πραγματικά δεδομένα. Παρακάτω βλέπουμε στη δικιά μας περίπτωση από το P-P Plot ότι πλησιάζει αρκετά στην ευθεία γραμμή.



Εικόνα 5.26: Normal P-P Plot Regression Standardized Residual.

- Ανεξαρτησία καταλοίπων: πρέπει τα κατάλοιπα να παρουσιάζουν ανεξαρτησία μεταξύ τους διότι αλλιώς θα υπάρχει αυτοσυσχέτιση μεταξύ τους και μπορεί να παραληφθεί κάποια σημαντική ανεξάρτητη μεταβλητή. Αυτό μπορούμε να το δούμε από το πίνακα model summary κοιτώντας το δείκτη Durbin-Watson όπου η τιμή του θέλουμε να είναι κοντά στη τιμή 2.Εξασφαλιζουμε και αυτή τη προϋπόθεση αφού ο δείκτης μας έχει τη τιμή 1,807 όπως φαίνεται και παρακάτω άρα υπάρχει ανεξαρτησία μεταξύ των καταλοίπων.

Model Summary[e]

Model	R	R Square	Adjusted R Square	Std. Error of the Estimate	Change Statistics: R Square Change	F Change	df1	df2	Sig. F Change	Durbin-Watson
1	,231[a]	,053	,052	1,060432169	,053	30,622	1	543	,000	
2	,273[b]	,075	,071	1,049468905	,021	12,404	1	542	,000	
3	,294[c]	,087	,082	1,043568426	,012	7,146	1	541	,008	
4	,308[d]	,095	,088	1,039812125	,008	4,916	1	540	,027	1,807

a. Predictors: (Constant), airconditioning

b. Predictors: (Constant), airconditioning, smoking

c. Predictors: (Constant), airconditioning, smoking, pool

d. Predictors: (Constant), airconditioning, smoking, pool, wheelchair

e. Dependent Variable: incomeb

Πίνακας 5.24: Model summary.

- πολυσυγραμμικότητα: θα ελέγξουμε αν υπάρχει γραμμική συσχέτιση μεταξύ των ανεξάρτητων μεταβλητών, πράγμα που δεν θέλουμε. Ο έλεγχος θα γίνει κοιτώντας από το πίνακα coefficients τις τιμές των στηλών Tolerance και VIF.Πρέπει οι τιμές να είναι κοντά στη μονάδα για να ικανοποιείτε και αυτή η προϋπόθεση. Όπως φαίνεται παρακάτω ισχύει και στη δική μας περίπτωση.

Coefficients[a]

Model		Unstandardized Coefficients		Standardized Coefficients	t	Sig.	Correlations		
		B	Std. Error	Beta			Zero-order	Partial	Part
1	(Constant)	4,840	,151		31,976	,000			
	air-conditioning	,882	,159	,232	5,558	,000	,232	,232	,232
2	(Constant)	5,037	,160		31,435	,000			
	air-conditioning	,789	,159	,207	4,956	,000	,232	,208	,204
	smoking	-,334	,096	-,146	-3,479	,001	-,180	-,148	-,144
3	(Constant)	5,029	,159		31,552	,000			
	air-conditioning	,776	,158	,204	4,900	,000	,232	,206	,201
	smoking	-,320	,096	-,139	-3,342	,001	-,180	-,142	-,137
	pool	,938	,351	,110	2,668	,008	,127	,114	,109
4	(Constant)	5,019	,159		31,593	,000			
	air-conditioning	,756	,158	,199	4,779	,000	,232	,201	,195
	smoking	-,317	,095	-,138	-3,323	,001	-,180	-,141	-,136
	pool	,884	,351	,103	2,520	,012	,127	,108	,103
	wheelchair	,373	,170	,090	2,199	,028	,115	,094	,090

a. Dependent Variable: incomeb

Πίνακας 5.25: Coefficients.

Επίσης μπορούμε να κάνουμε έλεγχο της πολυσυγραμμικότητας και από τον πίνακα correlations κοιτώντας τους συντελεστές συσχέτισης των ανεξάρτητων μεταβλητών οι οποίοι πρέπει να μην υπερβαίνουν το 0,8 κάτι που ισχύει σε εμάς.

Correlations

		incomeb	wifi	parking	pool	pets	smoking	wheelchair	cable tv	gym	jacuzzi	priv-entrance	fireplace
Pearson Correlation	incomeb	1,000	,069	,052	,127	-,107	-,180	,115	,065	,026	,029	,068	,017
	wifi	,069	1,000	,042	,014	,019	,001	,030	,036	,014	,026	,028	,028
	parking	,052	,042	1,000	,239	-,020	-,043	,167	,068	-,055	,058	,102	,197
	pool	,127	,014	,239	1,000	-,040	-,063	,072	,049	,090	,029	,022	,025
	pets	-,107	,019	-,020	-,040	1,000	,372	,096	-,030	,017	-,033	-,052	-,045
	smoking	-,180	,001	-,043	-,063	,372	1,000	-,028	-,097	,046	-,064	-,040	-,046
	wheelchair	,115	,030	,167	,072	,096	-,028	1,000	,195	,117	,107	,062	,181
	cable tv	,065	,036	,068	,049	-,030	-,097	,195	1,000	,086	,066	,002	-,017
	gym	,026	,014	-,055	,090	,017	,046	,117	,086	1,000	,082	,018	-,036
	jacuzzi	,029	,026	,058	,029	-,033	-,064	,107	,066	,082	1,000	-,067	,221
	priv-entrance	,068	,028	,102	,022	-,052	-,040	,062	,002	,018	-,067	1,000	-,011
	fireplace	,017	,028	,197	,025	-,045	-,046	,181	-,017	-,036	,221	-,011	1,000
	elevator	,090	,047	,057	,033	,037	,004	,043	-,012	-,023	,030	-,085	,035
	baby-bed	,066	,016	,048	,079	,002	,024	,005	-,011	-,020	-,037	,158	,063
	laptop-workspace	,059	,144	,060	,034	-,084	-,165	,135	,102	,014	,100	,151	,018
	kitchen	,021	-,010	,039	,012	-,035	-,053	,027	,033	,013	,024	,026	,025
	dryer	,034	,040	,025	-,049	-,034	-,048	-,002	,140	,031	,166	,009	,037
	party	-,102	,028	,043	-,034	,192	,199	,036	,028	,019	-,003	,017	,051
	buzzer	,016	,083	,141	,017	-,019	-,048	,176	,218	,005	,077	,013	-,006
	heating	,056	,251	,045	,035	,052	,024	,077	,069	,037	,005	,073	,041
	air-conditioning	,232	,090	-,004	,041	-,059	-,167	,065	,066	-,005	,024	,008	,030
	family	,086	,149	,034	,028	-,003	-,035	,061	,016	-,037	,015	,022	,056

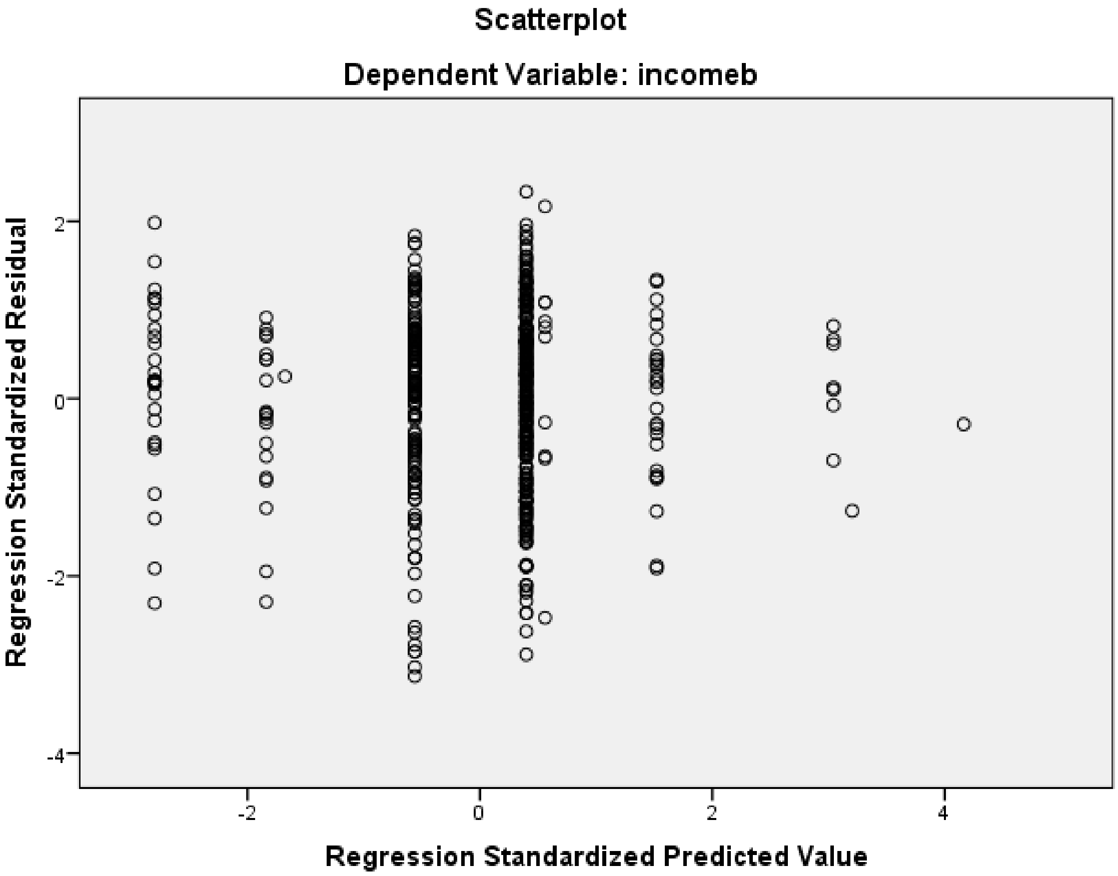
Πίνακας 5.26: Correlations.

Correlations

jacuzzi	priv-entrance	fireplace	elevator	baby-bed	laptop-workspace	kitchen	dryer	party	buzzer	heating	air-conditioning	family
,029	,068	,017	,090	,066	,059	,021	,034	-,102	,016	,056	,232	,086
,026	,028	,028	,047	,016	,144	-,010	,040	,028	,083	,251	,090	,149
,058	,102	,197	,057	,048	,060	,039	,025	,043	,141	,045	-,004	,034
,029	,022	,025	,033	,079	,034	,012	-,049	-,034	,017	,035	,041	,028
-,033	-,052	-,045	,037	,002	-,084	-,035	-,034	,192	-,019	,052	-,059	-,003
-,064	-,040	-,046	,004	,024	-,165	-,053	-,048	,199	-,048	,024	-,167	-,035
,107	,062	,181	,043	,005	,135	,027	-,002	,036	,176	,077	,065	,061
,066	,002	-,017	-,012	-,011	,102	,033	,140	,028	,218	,069	,066	,016
,082	,018	-,036	-,023	-,020	,014	,013	,031	,019	,005	,037	-,005	-,037
1,000	-,067	,221	,030	-,037	,100	,024	,166	-,003	,077	,005	,024	,015
-,067	1,000	-,011	-,085	,158	,151	,026	,009	,017	,013	,073	,008	,022
,221	-,011	1,000	,035	,063	,018	,025	,037	,051	-,006	,041	,030	,056
,030	-,085	,035	1,000	-,014	-,024	,218	,002	,067	,136	,023	,162	,358
-,037	,158	,063	-,014	1,000	,031	,014	,019	-,040	,036	,040	,047	,032
,100	,151	,018	-,024	,031	1,000	,091	,019	,008	,273	,186	,148	,031
,024	,026	,025	,218	,014	,091	1,000	-,022	,025	,076	-,026	,104	,261
,166	,009	,037	,002	,019	,019	-,022	1,000	,012	,057	,035	,041	,054
-,003	,017	,051	,067	-,040	,008	,025	,012	1,000	-,027	,013	-,046	,021
,077	,013	-,006	,136	,036	,273	,076	,057	-,027	1,000	,107	,128	,077
,005	,073	,041	,023	,040	,186	-,026	,035	,013	,107	1,000	,094	,120
,024	,008	,030	,162	,047	,148	,104	,041	-,046	,128	,094	1,000	,183
,015	,022	,056	,358	,032	,031	,261	,054	,021	,077	,120	,183	1,000

Πίνακας 5.27: Correlations.

- Ομοσκεδαστικότητα: Πρέπει να υπάρχει σταθερή διακύμανση καταλοίπων. Ο έλεγχος θα γίνει μέσω του διαγράμματος διασποράς. Στην περίπτωση μας που έχουμε κατηγορικές μεταβλητές αναμένουμε τα σημεία στο διάγραμμα να είναι κατανεμημένα γύρω από το μηδέν όπου ισχύει όπως βλέπουμε παρακάτω.



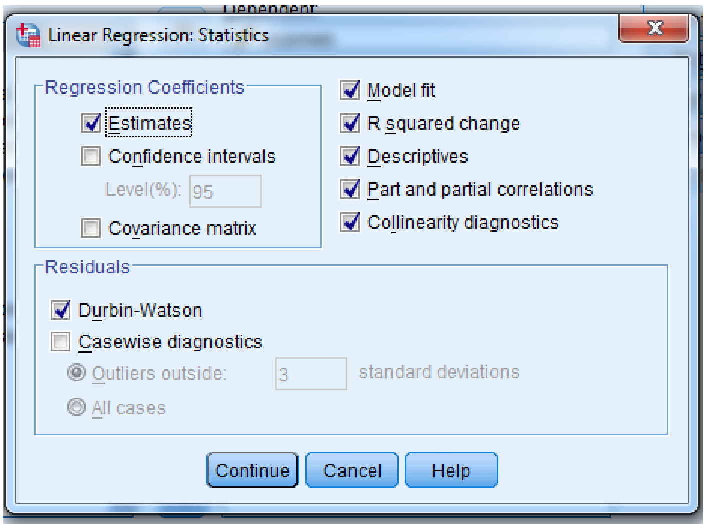
Εικόνα 5.27:Διάγραμμα διασποράς .

Έπειτα από τη τροποποίηση της εξαρτημένης μεταβλητής μας είδαμε ότι πληρούνται όλες οι προϋποθέσεις για την ανάλυση και συνεχίζουμε με την ερμηνεία του μοντέλου, όπου θα αναλύσουμε τα βήματα στη συνέχεια.

5.4 Επιλογές και ρυθμίσεις στο SPSS.

Μπορούμε να πάρουμε χρήσιμες πληροφορίες για τη στατιστική σημαντικότητα του παραγόμενου μοντέλου, για την προσαρμογή του στα εξεταζόμενα δεδομένα ή για την γενικότερη επιτυχία εφαρμογής της τεχνικής από τις διάφορες επιλογές του SPSS.

Από την επιλογή statistics.



Εικόνα 5.28: Επιλογή statistics στο SPSS.

Estimates: Οι συντελεστές του μοντέλου της παλινδρόμησης και σχετικά με αυτούς μέτρα όπως το τυπικό τους σφάλμα.

Confidence Interval: Η εκτίμηση της τιμής των συντελεστών της παλινδρόμησης έχει 95% διάστημα εμπιστοσύνης.

Covariance Matrix: Μας εμφανίζει το πίνακα συνδιακύμανσης και συσχέτισης.

Model fit: Εμφανίζεται ο πίνακας της **ANOVA .**

R squared change: Παίρνουμε πληροφορίες για την αλλαγή της τιμής του R^2 με την προσθήκη ή την διαγραφή μίας ανεξάρτητης μεταβλητής.

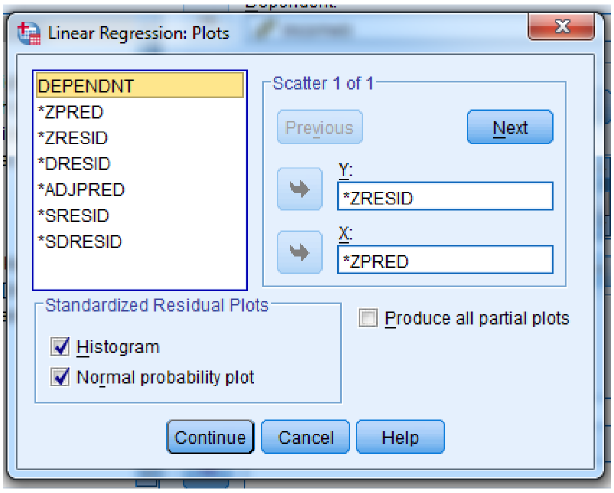
Part and Partial correlation: Μας εμφανίζει τους δείκτη αυτοσυσχέτισης των μεταβλητών.

Collinearity diagnostics: Παραγωγή ενός πλήθους στατιστικών μέτρων για την διάγνωση της συγραμμικότητας.

Durbin- Watson:Στατιστικός έλεγχος για την σειριακή συσχέτιση των καταλοίπων. Με την επιλογή αυτή εμφανίζονται τα περιγραφικά στατιστικά μέτρα των καταλοίπων και των προβλεπόμενων από το μοντέλο τιμών.

Casewise diagnostics: Πίνακας περιγραφής της συμπεριφοράς των ακραίων τιμών.

Από την επιλογή Plots παίρνουμε τα ιστογράμματα.



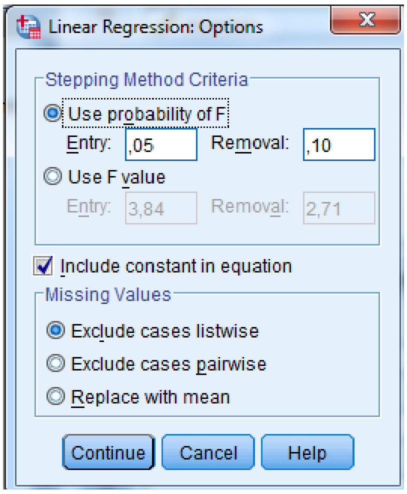
Εικόνα 5.29: Επιλογή Plots στο SPSS.

Normal probability plot: Μας εμφανίζει το διάγραμμα για τον έλεγχο της κανονικότητας των καταλοίπων.

Produce all partial plots: Μας εμφανίζει τα **scatterplots**

Στην επιλογή Y επιλέγουμε ZRESID και στη X επιλέγουμε ZPRED, έτσι επιλέγουμε το ζευγάρι που θα μας εμφανίσει στο **scatterplots.**

Από την επιλογή Options



Εικόνα 5.30: Επιλογή Options στο SPSS.

Ορίζουμε το επίπεδο σημαντικότητας για την απόρριψη ή την αποδοχή του μοντέλου.

5.5 Ανάλυση αποτελεσμάτων.

Από το παρακάτω πίνακα **Variables Entered/Removed[a]** βλέπουμε τις ανεξάρτητες μεταβλητές όπου μπήκαν στο μοντέλο μας και είναι οι 1) air conditioning 2) smoking 3) pool 4) wheelchair .Στην επιλογή των μεταβλητών, γίνεται σε κάθε βήμα έλεγχος στατιστικής σημαντικότητας για την εισαγωγή της στην εξίσωση παλινδρόμησης όπως φαίνεται ο έλεγχος από κάτω.

Variables Entered/Removed[a]

Model	Variables Entered	Variables Removed	Method
1	air-conditioning	.	Stepwise (Criteria: Probability-of-F-to-enter <= ,050, Probability-of-F-to-remove >= ,100).
2	smoking	.	Stepwise (Criteria: Probability-of-F-to-enter <= ,050, Probability-of-F-to-remove >= ,100).
3	pool	.	Stepwise (Criteria: Probability-of-F-to-enter <= ,050, Probability-of-F-to-remove >= ,100).
4	wheelchair	.	Stepwise (Criteria: Probability-of-F-to-enter <= ,050, Probability-of-F-to-remove >= ,100).

a. Dependent Variable: incomeb

Πίνακας 5.28: Variables Entered/Removed[a].

Τώρα θα αναλύσουμε για το πώς έγινε η επιλογή των τεσσάρων ανεξάρτητων μεταβλητών. Από το πίνακα correlations που φαίνεται παρακάτω θα δούμε πως έγινε η επιλογή της πρώτης ανεξάρτητης μεταβλητής. Ο πίνακας εμφανίζει όλες τις συσχετίσεις μεταξύ των μεταβλητών. Η επιλογή γίνεται με βάση τη συσχέτιση δηλαδή όποια ανεξάρτητη μεταβλητή έχει την μεγαλύτερη συσχέτιση σε σχέση με την εξαρτημένη τοποθετείται πρώτη στο μοντέλο. Βλέπουμε ότι τη μεγαλύτερη συσχέτιση την έχει το airconditioning με R=0,232, στατιστική σημαντικότητα sig=0,00 και από το πίνακα model summary όπου θα αναλύσουμε πιο μετά, βλέπουμε το συντελεστή προσδιορισμού R^2=0,054.

Correlations

		incomeb	wifi	parking	pool	pets	smoking	wheelchair	cable tv	gym	jacuzzi	priv-entrance	fireplace
Pearson Correlation	incomeb	1,000	,069	,052	,127	-,107	-,180	,115	,065	,026	,029	,068	,017
	wifi	,069	1,000	,042	,014	,019	,001	,030	,036	,014	,026	,028	,028
	parking	,052	,042	1,000	,239	-,020	-,043	,167	,068	-,055	,058	,102	,197
	pool	,127	,014	,239	1,000	-,040	-,063	,072	,049	,090	,029	,022	,025
	pets	-,107	,019	-,020	-,040	1,000	,372	,096	-,030	,017	-,033	-,052	-,045
	smoking	-,180	,001	-,043	-,063	,372	1,000	-,028	-,097	,046	-,064	-,040	-,046
	wheelchair	,115	,030	,167	,072	,096	-,028	1,000	,195	,117	,107	,062	,181
	cable tv	,065	,036	,068	,049	-,030	-,097	,195	1,000	,086	,066	,002	-,017
	gym	,026	,014	-,055	,090	,017	,046	,117	,086	1,000	,082	,018	-,036
	jacuzzi	,029	,026	,058	,029	-,033	-,064	,107	,066	,082	1,000	-,067	,221
	priv-entrance	,068	,028	,102	,022	-,052	-,040	,062	,002	,018	-,067	1,000	-,011
	fireplace	,017	,028	,197	,025	-,045	-,046	,181	-,017	-,036	,221	-,011	1,000
	elevator	,090	,047	,057	,033	,037	,004	,043	-,012	-,023	,030	-,085	,035
	baby-bed	,066	,016	,048	,079	,002	,024	,005	-,011	-,020	-,037	,158	,063
	laptop-workspace	,059	,144	,060	,034	-,084	-,165	,135	,102	,014	,100	,151	,018
	kitchen	,021	-,010	,039	,012	-,035	-,053	,027	,033	,013	,024	,026	,025
	dryer	,034	,040	,025	-,049	-,034	-,048	-,002	,140	,031	,166	,009	,037
	party	-,102	,028	,043	-,034	,192	,199	,036	,028	,019	-,003	,017	,051
	buzzer	,016	,083	,141	,017	-,019	-,048	,176	,218	,005	,077	,013	-,006
	heating	,056	,251	,045	,035	,052	,024	,077	,069	,037	,005	,073	,041
	air-conditioning	,232	,090	-,004	,041	-,059	-,167	,065	,066	-,005	,024	,008	,030
	family	,086	,149	,034	,028	-,003	-,035	,061	,016	-,037	,015	,022	,056

Πίνακας 5.29: Correlations.

Άρα η μεταβλητή air conditioning μπαίνει πρώτη στο μοντέλο1.

Για την επιλογή της δεύτερης ανεξάρτητης μεταβλητής θα γίνει έλεγχος συσχετίσεων μεταξύ των ανεξάρτητων μεταβλητών ώστε να λάβει υπόψη τη συσχέτιση της ήδη επιλεγμένης, οπότε ο έλεγχος θα γίνει από το πίνακα Excluded Variables κοιτάζοντας το μοντέλο 1. Ο συγκεκριμένος πίνακας μας δείχνει τις μερικές συσχετίσεις των ανεξάρτητων μεταβλητών σε σχέση με την εξαρτημένη μεταβλητή ώστε να αποφασίσουμε πια θα επιλεχθεί. Παρακάτω φαίνεται ο πίνακας Excluded Variables.

Excluded Variables[a]

Model		Beta In	t	Sig.	Partial Correlation	Collinearity Statistics Tolerance	VIF	Minimum Tolerance
1	wifi	,048[b]	1,158	,248	,050	,992	1,008	,992
	parking	,053[b]	1,280	,201	,055	1,000	1,000	1,000
	pool	,118[b]	2,836	,005	,121	,998	1,002	,998
	pets	-,094[b]	-2,253	,025	-,096	,996	1,004	,996
	smoking	-,146[b]	-3,479	,001	-,148	,972	1,029	,972
	wheelchair	,100[b]	2,405	,017	,103	,996	1,004	,996
	cable tv	,050[b]	1,200	,231	,051	,996	1,004	,996
	gym	,027[b]	,654	,513	,028	1,000	1,000	1,000
	jacuzzi	,024[b]	,564	,573	,024	,999	1,001	,999
	priv-entrance	,066[b]	1,597	,111	,068	1,000	1,000	1,000
	fireplace	,010[b]	,233	,816	,010	,999	1,001	,999
	elevator	,054[b]	1,269	,205	,054	,974	1,027	,974
	baby-bed	,055[b]	1,319	,188	,056	,998	1,002	,998
	laptop-workspace	,026[b]	,611	,541	,026	,978	1,022	,978
	kitchen	-,003[b]	-,079	,937	-,003	,989	1,011	,989
	dryer	,025[b]	,600	,549	,026	,998	1,002	,998
	party	-,091[b]	-2,191	,029	-,094	,998	1,002	,998
	buzzer	-,014[b]	-,339	,734	-,015	,984	1,017	,984
	heating	,035[b]	,826	,409	,035	,991	1,009	,991
	family	,045[b]	1,058	,290	,045	,967	1,035	,967

Πίνακας 5.30: Excluded Variables.

Τώρα για να επιλέξουμε ποια ανεξάρτητη μεταβλητή έχει τη μεγαλύτερη συσχέτιση με την εξαρτημένη θα κοιτάξουμε τη στήλη Partial Correlation. Παρατηρούμε ότι η μεταβλητή smoking έχει τη μεγαλύτερη συσχέτιση που ισούται με –0,148 και με στατιστική σημαντικότητα sig=0,00. Οι θετικές τιμές της συσχέτισης δεν υποδηλώνουν μεγαλύτερο βαθμό γραμμικής συσχέτισης από το βαθμό γραμμικής συσχέτισης που υποδηλώνουν αρνητικές τιμές. Ο βαθμός γραμμικής συσχέτισης καθορίζεται από την απόλυτη τιμή του R και όχι από το πρόσημο του R .
Ο συντελεστής προσδιορισμού έγινε τώρα R^2=0,074

Για την επιλογή της τρίτης ανεξάρτητης μεταβλητής θα γίνει έλεγχος από το πίνακα Excluded Variables κοιτάζοντας το μοντέλο 2. Ο πίνακας φαίνεται παρακάτω

Excluded Variables

2	wifi	,051[c]	1,228	,220	,053	,992	1,008	,964
	parking	,047[c]	1,140	,255	,049	,998	1,002	,970
	pool	,110[c]	2,668	,008	,114	,995	1,005	,969
	pets	-,047[c]	-1,063	,288	-,046	,862	1,160	,841
	wheelchair	,097[c]	2,367	,018	,101	,995	1,005	,969
	cable tv	,038[c]	,912	,362	,039	,988	1,012	,965
	gym	,034[c]	,821	,412	,035	,998	1,002	,970
	jacuzzi	,015[c]	,359	,720	,015	,996	1,004	,969
	priv-entrance	,061[c]	1,478	,140	,063	,998	1,002	,971
	fireplace	,004[c]	,092	,927	,004	,997	1,003	,971
	elevator	,058[c]	1,393	,164	,060	,973	1,028	,946
	baby-bed	,060[c]	1,447	,148	,062	,997	1,003	,970
	laptop-workspace	,005[c]	,117	,907	,005	,958	1,044	,952
	kitchen	-,009[c]	-,205	,838	-,009	,988	1,012	,963
	dryer	,019[c]	,460	,646	,020	,997	1,003	,970
	party	-,066[c]	-1,563	,119	-,067	,960	1,041	,936
	buzzer	-,018[c]	-,438	,661	-,019	,983	1,017	,958
	heating	,041[c]	,977	,329	,042	,990	1,011	,963
	family	,044[c]	1,054	,292	,045	,967	1,035	,941

Πίνακας 5.31: Excluded Variables.

Οπότε για να επιλέξουμε ποια ανεξάρτητη μεταβλητή έχει τη μεγαλύτερη συσχέτιση με την εξαρτημένη θα κοιτάξουμε πάλι τη στήλη Partial Correlation. Παρατηρούμε ότι η μεταβλητή pool έχει τη μεγαλύτερη συσχέτιση που ισούται με 0,114 και με στατιστική σημαντικότητα sig=0,008. Ο συντελεστής προσδιορισμού έγινε τώρα R^2=0,086.

Για την επιλογή της τέταρτης ανεξάρτητης μεταβλητής θα γίνει έλεγχος από το πίνακα Excluded Variables κοιτάζοντας το μοντέλο 3. Ο πίνακας φαίνεται παρακάτω.

Excluded Variables

3	wifi	,050ᵈ	1,206	,228	,052	,992	1,009	,963
	parking	,022ᵈ	,528	,598	,023	,942	1,062	,939
	pets	-,045ᵈ	-1,020	,308	-,044	,861	1,161	,839
	wheelchair	,090ᵈ	2,199	,028	,094	,991	1,009	,968
	cable tv	,033ᵈ	,806	,420	,035	,986	1,014	,962
	gym	,024ᵈ	,581	,562	,025	,989	1,011	,967
	jacuzzi	,012ᵈ	,295	,768	,013	,995	1,005	,966
	priv-entrance	,059ᵈ	1,434	,152	,061	,998	1,002	,968
	fireplace	,001ᵈ	,035	,972	,002	,997	1,003	,968
	elevator	,055ᵈ	1,325	,186	,057	,972	1,029	,945
	baby-bed	,051ᵈ	1,248	,213	,054	,991	1,010	,968
	laptop-workspace	,003ᵈ	,063	,950	,003	,957	1,044	,949
	kitchen	-,009ᵈ	-,222	,824	-,010	,988	1,012	,962
	dryer	,025ᵈ	,605	,546	,026	,994	1,006	,967
	party	-,063ᵈ	-1,513	,131	-,065	,960	1,042	,933
	buzzer	-,019ᵈ	-,468	,640	-,020	,983	1,018	,957
	heating	,037ᵈ	,893	,372	,038	,988	1,012	,962
	family	,042ᵈ	1,006	,315	,043	,966	1,035	,940

Πίνακας 5.31: Excluded Variables.

Παρατηρούμε ότι η μεταβλητή wheelchair έχει τη μεγαλύτερη συσχέτιση που ισούται με 0,094 και με στατιστική σημαντικότητα sig=0,028. O συντελεστής προσδιορισμού έγινε τώρα R^2=0,094.

Για την επιλογή της πέμπτης ανεξάρτητης μεταβλητής θα γίνει έλεγχος από το πίνακα Excluded Variables κοιτάζοντας το μοντέλο 4. Ο πίνακας φαίνεται παρακάτω

Excluded Variables

4	wifi	,048ᵉ	1,158	,248	,050	,991	1,009	,960
	parking	,008ᵉ	,191	,848	,008	,919	1,088	,919
	pets	-,057ᵉ	-1,289	,198	-,055	,850	1,177	,836
	cable tv	,017ᵉ	,402	,688	,017	,951	1,051	,951
	gym	,014ᵉ	,338	,736	,015	,977	1,024	,966
	jacuzzi	,003ᵉ	,071	,944	,003	,985	1,016	,966
	priv-entrance	,054ᵉ	1,310	,191	,056	,995	1,006	,968
	fireplace	-,015ᵉ	-,361	,718	-,016	,965	1,036	,959
	elevator	,052ᵉ	1,261	,208	,054	,971	1,030	,942
	baby-bed	,052ᵉ	1,259	,209	,054	,991	1,010	,966
	laptop-workspace	-,009ᵉ	-,212	,832	-,009	,943	1,061	,943
	kitchen	-,011ᵉ	-,267	,790	-,011	,987	1,013	,959
	dryer	,025ᵉ	,611	,542	,026	,994	1,006	,967
	party	-,068ᵉ	-1,621	,106	-,070	,958	1,044	,933
	buzzer	-,036ᵉ	-,853	,394	-,037	,955	1,047	,955
	heating	,031ᵉ	,745	,457	,032	,984	1,017	,959
	family	,038ᵉ	,903	,367	,039	,964	1,037	,938

Πίνακας 5.32: Excluded Variables.

Παρατηρούμε ότι η μεταβλητή party έχει τη μεγαλύτερη συσχέτιση που ισούται με -0,7 όμως η στατιστική σημαντικότητα sig=0.106 > 0.05 άρα δεν δεχόμαστε την είσοδο άλλων μεταβλητών στο μοντέλο.

5.5.1 Παρουσίαση του πίνακα model summary.

Model Summary[e]

Model	R	R Square	Adjusted R Square	Std. Error of the Estimate	Change Statistics					Durbin-Watson
					R Square Change	F Change	df1	df2	Sig. F Change	
1	,232[a]	,054	,052	1,059457131	,054	30,897	1	545	,000	
2	,272[b]	,074	,071	1,048824964	,021	12,106	1	544	,001	
3	,294[c]	,086	,081	1,042974243	,012	7,120	1	543	,008	
4	,307[d]	,094	,088	1,039310980	,008	4,835	1	542	,028	1,806

a. Predictors: (Constant), air-conditioning

b. Predictors: (Constant), air-conditioning, smoking

c. Predictors: (Constant), air-conditioning, smoking, pool

d. Predictors: (Constant), air-conditioning, smoking, pool, wheelchair

e. Dependent Variable: incomeb

Πίνακας 5.33: Model Summary.

Ο πίνακας model summary μας εμφανίζει κάποιους δείκτες καλής προσαρμογής. Παρατηρούμε ότι με την ένταξη των ανεξάρτητων μεταβλητών στο μοντέλο μας ο συντελεστής προσδιορισμού αυξήθηκε. Με την εισαγωγή της πρώτης μεταβλητή η τιμή ήταν 0,052 και μετά την ένταξη της τέταρτης έφτασε η τιμή του στο 0,88. Η στήλη Std. Error of the Estimate μας δείχνει το σφάλμα μεταξύ της μετρούμενης και της πραγματικής τιμής. Τέλος από τη στήλη Sig. F Change βλέπουμε την επίδραση του κάθε μοντέλου ξεχωριστά.

5.5.2 Παρουσίαση του πίνακα ANOVA.

Ο πίνακας ανάλυσης διακύμανσης ANOVA παρουσιάζεται παρακάτω.

ANOVA[a]

Model		Sum of Squares	df	Mean Square	F	Sig.
1	Regression	34,680	1	34,680	30,897	,000[b]
	Residual	611,735	545	1,122		
	Total	646,415	546			
2	Regression	47,997	2	23,998	21,816	,000[c]
	Residual	598,418	544	1,100		
	Total	646,415	546			
3	Regression	55,742	3	18,581	17,081	,000[d]
	Residual	590,673	543	1,088		
	Total	646,415	546			
4	Regression	60,964	4	15,241	14,110	,000[e]
	Residual	585,451	542	1,080		
	Total	646,415	546			

a. Dependent Variable: incomeb

b. Predictors: (Constant), air-conditioning

c. Predictors: (Constant), air-conditioning, smoking

d. Predictors: (Constant), air-conditioning, smoking, pool

e. Predictors: (Constant), air-conditioning, smoking, pool, wheelchair

Πίνακας 5.34: ANOVA.

Από τον παραπάνω πίνακα γίνεται ένας έλεγχος για τη σημαντικότητα του μοντέλου. Η συνάρτηση F μας βοηθάει να ελέγξουμε την υπόθεση ότι οι συντελεστές των ανεξάρτητων μεταβλητών που συμμετέχουν είναι και ταυτόχρονα μηδέν. Ωστόσο ελέγχουμε τη τιμή sig που δηλώνει τη στατιστική σημαντικότητα του συνολικού μοντέλου. Βλέπουμε ότι οι τιμές των F>1 και οι τιμές των sig για τα τέσσερα μοντέλα είναι ίση με sig=0,000 άρα είναι μικρότερη του 0,05 οπότε μπορούμε να πούμε ότι και τα τέσσερα μοντέλα μας είναι στατιστικά σημαντικά. Επίσης από το πίνακα μπορούμε να δούμε τη διακύμανση της παλινδρόμησης για κάθε μοντέλο που είναι η γραμμή Regression, τη διακύμανση του λάθους που είναι η γραμμή Residual και από τη γραμμή Total είναι το άθροισμα των δυο παραπάνω. Παρατηρούμε ότι με την εισαγωγή της ανεξάρτητης μεταβλητής σε κάθε μοντέλο μειώνεται η διακύμανση λάθους και αυξάνεται η διακύμανση του μοντέλου.

5.5.3 Παρουσίαση του πίνακα Coefficients.

Στη συνέχεια βλέπουμε το πίνακα coefficients όπου θα πάρουμε τους συντελεστές της παλινδρόμησης.

Coefficients[a]

Model		Unstandardized Coefficients B	Unstandardized Coefficients Std. Error	Standardized Coefficients Beta	t	Sig.	Correlations Zero-order	Correlations Partial	Correlations Part	Collinearity Statistics Tolerance	Collinearity Statistics VIF
1	(Constant)	4,840	,151		31,976	,000					
	air-conditioning	,882	,159	,232	5,558	,000	,232	,232	,232	1,000	1,000
2	(Constant)	5,037	,160		31,435	,000					
	air-conditioning	,789	,159	,207	4,956	,000	,232	,208	,204	,972	1,029
	smoking	-,334	,096	-,146	-3,479	,001	-,180	-,148	-,144	,972	1,029
3	(Constant)	5,029	,159		31,552	,000					
	air-conditioning	,776	,158	,204	4,900	,000	,232	,206	,201	,971	1,030
	smoking	-,320	,096	-,139	-3,342	,001	-,180	-,142	-,137	,969	1,032
	pool	,938	,351	,110	2,668	,008	,127	,114	,109	,995	1,005
4	(Constant)	5,019	,159		31,593	,000					
	air-conditioning	,756	,158	,199	4,779	,000	,232	,201	,195	,968	1,033
	smoking	-,317	,095	-,138	-3,323	,001	-,180	-,141	-,136	,969	1,032
	pool	,884	,351	,103	2,520	,012	,127	,108	,103	,990	1,010
	wheelchair	,373	,170	,090	2,199	,028	,115	,094	,090	,991	1,009

a. Dependent Variable: incomeb

Πίνακας 5.35: Coefficients.

Όπως είπαμε στο συγκεκριμένο πίνακα από τη στήλη B παίρνουμε το σταθερό όρο και τους συντελεστές παλινδρόμησης κάθε μιας ανεξάρτητης μεταβλητής για κάθε μοντέλο ξεχωριστά, όπου βλέπουμε την επίδραση τους προς την εξαρτημένη μεταβλητή. Μέσω αυτών θα βγάλουμε και την τελική εξίσωση μετέπειτα.

Από τη στήλη Std.Error παίρνουμε τις τιμές του τυπικού σφάλματος της εκτίμησης των συντελεστών B. Στην στήλη t περιέχονται οι τιμές τις στατιστικής συνάρτησης βάση της οποίας γίνεται ο έλεγχος για την σημαντικότητα ύπαρξης των συντελεστών στην συνάρτηση. Από την τελευταία στήλη του πίνακα, την στήλη sig., δίνεται η τιμή σημαντικότητας βάση της οποίας επιβεβαιώνεται η όχι η διατήρηση της μεταβλητής στο μοντέλο.

Από τη στήλη Beta μπορούμε να ταξινομήσουμε τις ανεξάρτητες μεταβλητές ως προς την ερμηνευτική τους ικανότητα. Από το μοντέλο 4 που έχουν εισχωρήσει και οι τέσσερεις μεταβλητές βλέπουμε ότι το air conditioning επηρεάζει πιο πολύ με συντελεστή Beta 0,199, έπειτα το smoking με συντελεστή Beta –0,138 μετά η pool με συντελεστή Beta 0,103 και τέλος η wheelchair με συντελεστή Beta 0,09. Επίσης από τη στήλη sig βλέπουμε ότι υπάρχει στατιστική σημαντικότητα κάθε μεταβλητής σε κάθε μοντέλο αφού οι τιμές των sig<0,05.

5.5.4 Μορφή της εξίσωσης και ερμηνεία τελικού μοντέλου.

Από το πίνακα coefficients που εμφανίσαμε παραπάνω και συγκεκριμένα από το μοντέλο 4 θα πάρουμε το σταθερό όρο και τους συντελεστές παλινδρόμησης κάθε μιας ανεξάρτητης μεταβλητής για να δημιουργήσουμε την τελική εξίσωση όπως θα δούμε στη συνέχεια.

Η εξίσωση έχει τη μορφή:

$$Ln(Y) = a_0 + a_1 * x_1 + a_2 * x_2 + a_3 * x_3 + a_4 * x_4$$

Όπου

a_0=5,019 a_1=0,756 a_2=-0,317 a_3=0,884 a_4=0,373

Άρα έχει τη μορφή

$$\mathrm{Ln(Y)} = 5.019 + 0.756 * \mathrm{x}_1 - 0.317 * \mathrm{x}_2 + 0.884 * \mathrm{x}_3 + 0.373 * \mathrm{x}_4$$

Οι μεταβλητές $\mathrm{x}_1, \mathrm{x}_2, \mathrm{x}_3, \mathrm{x}_4$ αναφέρονται στις μεταβλητές air conditioning, smoking, pool και wheelchair.

Αρχικά θα βγάλουμε τη μέση τιμή των εσόδων/ μήνα για τα καταλύματα όπου δεν έχουν καμία από τις παραπάνω παροχές. Δηλαδή οι τιμές των $\mathrm{x}_1, \mathrm{x}_2, \mathrm{x}_3, \mathrm{x}_4$ είναι ίσες με το μηδέν.

Ισχύει ότι:

$$Ln(Y) = 5.019 \Leftrightarrow$$

$$e^{Ln(Y)} = e^{5.019} \Leftrightarrow$$

$$Y = 151.24$$

Συμπεραίνουμε ότι η μέση τιμή των εσόδων/μήνα των καταχωρήσεων κυμαίνεται στα 151,24 ευρώ.

Τώρα θα δούμε πως επηρεάζει κάθε μεταβλητή ξεχωριστά την τελική τιμή. Επειδή η εξίσωση μου έχει λογαριθμική μορφή θα πρέπει να υπολογίζω κάθε φορά το e^{α}-1 ώστε να βρω τη ποσοστιαία μεταβολή των εσόδων/μήνα όταν μια ερμηνευτική μεταβλητή ποικίλει κατά 1.

Για τη μεταβλητή air conditioning υπολογίζω το $e^{0.756}$ -1=1.1297.Οπότε για τα καταλύματα που έχουν κλιματισμό θα αυξάνονται τα έσοδα/μήνα κατά 113% σε σχέση με τα καταλύματα που δεν διαθέτουν.

Για τη μεταβλητή smoking υπολογίζω το $e^{-0.317}$-1=-0.2716.Οπότε για τα καταλύματα που επιτρέπεται το κάπνισμα θα μειώνονται τα έσοδα/μήνα κατά 27% σε σχέση με τα καταλύματα που επιτρέπεται.

Για τη μεταβλητή pool υπολογίζω το $e^{0.884}$ -1=1.4205.Οπότε για τα καταλύματα που έχουν πισίνα θα αυξάνονται τα έσοδα/μήνα κατά 142% σε σχέση με τα καταλύματα που δεν διαθέτουν.

Για τη μεταβλητή wheelchair υπολογίζω το$e^{0.373}$ -1=0.4520.Οπότε για τα καταλύματα που έχουν είσοδο για αναπηρικό αμαξίδιο θα αυξάνονται τα έσοδα/μήνα κατά 45,2% σε σχέση με τα καταλύματα που δεν διαθέτουν.

5.6 Κατανόηση μοντέλου.

Για να γίνει σύγκριση των εσόδων μεταξύ δυο καταλυμάτων όπου το ένα από τα δυο διαθέτει μια παροχή θα πρέπει όλες οι υπόλοιπες παροχές να είναι σταθερές και στα δύο, διότι στην παλινδρόμηση έλαβαν μέρος όλες οι μεταβλητές.

Έστω ότι έχουμε ένα κατάλυμα που διαθέτει κλιματισμό και ένα άλλο που δεν διαθέτει τη συγκεκριμένη παροχή ενώ όλες οι υπόλοιπες παροχές είναι σταθερές . Οι συγκεκριμένες μεταβλητές είναι ποιοτικές και παίρνουν τιμές 0 και 1 αν διαθέτουν ή όχι αντίστοιχα τη συγκεκριμένη παροχή. Για τον υπολογισμό των εσόδων/μήνα του πρώτου καταλύματος θα χρησιμοποιήσουμε την παραπάνω εξίσωση. x_1, x_2, x_3, x_4.

$$Ln(Y_1) = 5.019 + 0.756 * x_1 - 0.317 * x_2 + 0.884 * x_3 + 0.373 * x_4 \Leftrightarrow$$

$$Ln(Y_1) = 5.019 + 0.756 * 1 - 0.317 * 0 + 0.884 * 0 + 0.373 * 0 \Leftrightarrow$$

$$Ln(Y_1) = 5{,}775 \Rightarrow$$

$$e^{Ln(Y_1)} = e^{5{,}775} \Leftrightarrow$$

$$Y_1 = 322{,}14$$

Καταλήγουμε ότι τα έσοδα/ μήνα του πρώτου καταλύματος θα είναι 322,14 ευρώ.

Ενώ για το δεύτερο ισχύει

$$Ln(Y_2) = 5.019 + 0.756 * x_1 - 0.317 * x_2 + 0.884 * x_3 + 0.373 * x_4 \Leftrightarrow$$

$$Ln(Y_2) = 5.019 + 0.756 * 0 - 0.317 * 0 + 0.884 * 0 + 0.373 * 0 \Leftrightarrow$$

$$Ln(Y_2) = 5{,}019 \Rightarrow$$

$$e^{Ln(Y_2)} = e^{5{,}019} \Leftrightarrow$$

$$Y_2 = 151{,}24$$

Οπότε τα έσοδα/ μήνα του δεύτερου καταλύματος θα είναι 151,24 ευρώ.

ΚΕΦΑΛΑΙΟ 6

6. Παράγοντες διαμόρφωσης εσόδων των καταλυμάτων της Αθήνας.

6.1 Περιγραφή διαδικασίας συλλογής δεδομένων.

Για τη συλλογή των δεδομένων αρχικά χρησιμοποιήσαμε την ιστοσελίδα www.insideairbnb.com/athens. Μέσω αυτής μπορούμε να δούμε τις καταχωρήσεις των καταλυμάτων που έχει κάθε περιοχή. Επιλέγοντας μια κατοικία συλλέγαμε δεδομένα για τα εκτιμώμενα έσοδα ανά μήνα που θα έχει ο ιδιοκτήτης, τις εκτιμώμενες ημέρες που θα βρίσκεται υπό κράτηση και για την διαθεσιμότητα του. Έπειτα από το id της καταχώρησης συνδεόμασταν με την ιστοσελίδα του airbnb για να δούμε τις υπόλοιπες παροχές όπου θα αναλύσουμε στη συνέχεια. Από τα διαθέσιμα δεδομένα της ιστοσελίδας www.insideairbnb.com/athens υπήρχαν πληροφορίες για το γεωγραφικό πλάτος και μήκος κάθε καταχώρησης, όπου μέσω αυτών βρήκαμε τις αποστάσεις των καταλυμάτων από το μουσείο της Ακρόπολης, από τη Πλάκα και από την Ομόνοια με τη βοήθεια του Google maps. Καταφέραμε να συλλέξουμε δεδομένα για 603 καταχωρήσεις την περίοδο από 1 Νοέμβρη έως 15 Δεκέμβρη.

6.2 Ανάλυση μεταβλητών.

Στη διαδικασία της βηματικής παλινδρόμησης πήραν μέρος 40 διαφορετικές μεταβλητές, κάποιες από τις οποίες ήταν ποσοτικές δηλαδή αντιστοιχούν σε μεγέθη που μπορούν να μετρηθούν και κάποιες ποιοτικές όπου μας δίνουν μια μη ποσοτική πληροφορία όπως θα δούμε στη συνέχεια. Τώρα θα γίνει η ανάλυση των ποσοτικών μεταβλητών που πήραν μέρος στη διαδικασία.

Η μεταβλητή listings μας δίνει πληροφορίες για τις υπόλοιπες καταχωρήσεις που έχει ο ιδιοκτήτης στην επίβλεψη του.

Η μεταβλητή min nights μας δίνει πληροφορίες για τις ελάχιστες διανυκτερεύσεις που έχει ορίσει ο ιδιοκτήτης.

Η μεταβλητή income μας δίνει πληροφορίες για τα έσοδα ανά μήνα που θα έχει ο ιδιοκτήτης. Η συγκεκριμένη μεταβλητή θα οριστεί ως εξαρτημένη στη διαδικασία της παλινδρόμησης.

Η μεταβλητή price μας δίνει πληροφορίες για τη τιμή που έχει ορίσει ο ιδιοκτήτης στο κατάλυμα.

Η μεταβλητή reviews μας δίνει πληροφορίες για τον αριθμό των αξιολογήσεων που έχει μαζέψει από τους ενοικιαστές του το κάθε κατάλυμα.

Η μεταβλητή plaka μας δίνει πληροφορίες για την απόσταση του καταλύματος από την περιοχή της Πλάκας.

Η μεταβλητή omonoia μας δίνει πληροφορίες για την απόσταση του καταλύματος από την περιοχή της Ομόνοιας.

Η μεταβλητή akropoli μας δίνει πληροφορίες για την απόσταση του καταλύματος από το μουσείο της ακρόπολης.

Η μεταβλητή photos μας δίνει πληροφορίες για των αριθμό των φωτογραφιών του καταλύματος που έχει ανεβάσει ο ιδιοκτήτης.

Η μεταβλητή guests μας δίνει πληροφορίες για το πόσους ανθρώπους μπορεί να φιλοξενήσει το κατάλυμα.

Η μεταβλητή bedroomsμας δίνει πληροφορίες για τον αριθμό των υπνοδωματίων.

Η μεταβλητή beds μας δίνει πληροφορίες για τον αριθμό των κρεβατιών.

Η μεταβλητή guests in bedroom μας δίνει πληροφορίες για τον αριθμό των ανθρώπων όπου θα κοιμούνται σε υπνοδωμάτιο και όχι σε κάποιο κοινόχρηστο χώρο.

Τώρα θα γίνει η ανάλυση και ο ορισμός των ποιοτικών μεταβλητών. Για να μπορέσουμε να εισάγουμε τις συγκεκριμένες μεταβλητές στο SPSS θα πρέπει να ορίσουμε κάποιες ψευδομεταβλητές όπου θα τις εκφράζουν.

Παρακάτω φαίνονται όλες οι ποιοτικές μεταβλητές και η κωδικοποίηση τους.

- Η μεταβλητή Wifi αναφέρεται στο αν το κατάλυμα διαθέτει δωρεάν ασύρματο ίντερνετ.

0= Το κατάλυμα δεν διαθέτει δωρεάν ασύρματο ίντερνετ.

1= Το κατάλυμα διαθέτει δωρεάν ασύρματο ίντερνετ.

- Η μεταβλητή parking αναφέρεται στην ύπαρξη δωρεάν χώρου στάθμευσης στις εγκαταστάσεις.

0= Το κατάλυμα δεν διαθέτει χώρο στάθμευσης.

1= Το κατάλυμα διαθέτει χώρο στάθμευσης.

- Η μεταβλητή pool αναφέρεται στο εάν το κατάλυμα διαθέτει πισίνα.

0= Το κατάλυμα δεν διαθέτει πισίνα.

1= Το κατάλυμα διαθέτει πισίνα.

- Η μεταβλητή pets αναφέρεται στην είσοδο των κατοικίδιων στο σπίτι.

0= Το κατάλυμα δεν επιτρέπει την είσοδο των κατοικίδιων.

1= Το κατάλυμα επιτρέπει την είσοδο των κατοικίδιων.

- Η μεταβλητή smoking αναφέρεται στο αν επιτρέπεται το κάπνισμα μέσα στο χώρο.

0= Το κατάλυμα απαγορεύει το κάπνισμα μέσα στο χώρο.

1= Το κατάλυμα επιτρέπει το κάπνισμα μέσα στο χώρο.

- Η μεταβλητή wheelchair αναφέρεται στην ύπαρξη εισόδου για αναπηρικό αμαξίδιο.

0= Το κατάλυμα δεν διαθέτει είσοδο για αναπηρικό αμαξίδιο.

. 1= Το κατάλυμα διαθέτει είσοδο για αναπηρικό αμαξίδιο.

- Η μεταβλητή cable TV αναφέρεται στην ύπαρξη δορυφορικής τηλεόρασης

0= Το κατάλυμα δεν διαθέτει δορυφορική τηλεόραση.

. 1= Το κατάλυμα διαθέτει δορυφορική τηλεόραση.

- Η μεταβλητή gym αναφέρεται στην ύπαρξη εισόδου γυμναστηρίου.

0= Το κατάλυμα δεν διαθέτει γυμναστήριο.

. 1= Το κατάλυμα διαθέτει γυμναστήριο.

- Η μεταβλητή Jacuzzi αναφέρεται στην ύπαρξη Jacuzzi.

0= Το κατάλυμα δεν διαθέτει Jacuzzi.

. 1= Το κατάλυμα διαθέτει Jacuzzi.

- Η μεταβλητή private entrance αναφέρεται στην ύπαρξη ιδιωτικής εισόδου.

0= Το κατάλυμα δεν διαθέτει ιδιωτική είσοδο.

. 1= Το κατάλυμα διαθέτει ιδιωτική είσοδο.

- Η μεταβλητή fireplace αναφέρεται στην ύπαρξη τζακιού στο κατάλυμα.

0= Το κατάλυμα δεν διαθέτει τζάκι.

. 1= Το κατάλυμα διαθέτει τζάκι.

- Η μεταβλητή elevator αναφέρεται στην ύπαρξη ασανσέρ.

0= Το κατάλυμα δεν διαθέτει ασανσέρ.

. 1= Το κατάλυμα διαθέτει ασανσέρ.

- Η μεταβλητή baby bed αναφέρεται στην ύπαρξη παιδικού κρεβατιού.

0= Το κατάλυμα δεν διαθέτει παιδικό κρεβάτι.

. 1= Το κατάλυμα διαθέτει παιδικό κρεβάτι.

- Η μεταβλητή laptop αναφέρεται στην ύπαρξη επιφάνειας για τοποθέτηση φορητού υπολογιστή.

0= Το κατάλυμα δεν διαθέτει επιφάνεια.

. 1= Το κατάλυμα διαθέτει επιφάνεια.

- Η μεταβλητή kitchen αναφέρεται στην ύπαρξη κουζίνας στο κατάλυμα.

0= Το κατάλυμα δεν διαθέτει κουζίνα.

. 1= Το κατάλυμα διαθέτει κουζίνα.

- Η μεταβλητή dryer αναφέρεται στην ύπαρξη στεγνωτηρίου.

0= Το κατάλυμα δεν διαθέτει στεγνωτήριο.

. 1= Το κατάλυμα διαθέτει στεγνωτήριο.

- Η μεταβλητή party αναφέρεται στο εάν επιτρέπονται πάρτι και εκδηλώσεις στο κατάλυμα.

0= Το κατάλυμα δεν επιτρέπει.

. 1= Το κατάλυμα επιτρέπει.

- Η μεταβλητή buzzer αναφέρεται στην ύπαρξη κουδουνιού/ ασύρματης ενδοεπικοινωνίας .

0= Το κατάλυμα δεν διαθέτει ασύρματη ενδοεπικοινωνία.

. 1= Το κατάλυμα διαθέτει ασύρματη ενδοεπικοινωνία.

- Η μεταβλητή heating αναφέρεται στην ύπαρξη θέρμανσης του χώρου.

0= Το κατάλυμα δεν διαθέτει θέρμανση.

. 1= Το κατάλυμα διαθέτει θέρμανση.

- Η μεταβλητή air conditioning αναφέρεται στην ύπαρξη κλιματισμού του χώρου.

0= Το κατάλυμα δεν διαθέτει θέρμανση.

. 1= Το κατάλυμα διαθέτει θέρμανση.

- Η μεταβλητή family αναφέρεται στο εάν το κατάλυμα είναι κατάλληλο για οικογένεια.

0= Το κατάλυμα δεν είναι κατάλληλο για οικογένεια.

. 1= Το κατάλυμα είναι κατάλληλο για οικογένεια

- Η μεταβλητή cancel αναφέρεται στην ακύρωση που μπορεί να κάνει ο ενοικιαστής.

1= Ευέλικτη. Πλήρη επιστροφή χρημάτων, εάν η ακύρωση γίνει 24 ώρες πριν την άφιξη.

2= Μέτρια. Πλήρη επιστροφή χρημάτων, εάν η ακύρωση γίνει 5 ημέρες πριν την άφιξη.

3= Αυστηρή. Πλήρη επιστροφή χρημάτων, εάν η ακύρωση γίνει 7 ημέρες πριν την άφιξη.

- Η μεταβλητή type αναφέρεται στο τύπο του καταλύματος.

1= Entire home /apt

2= Private room

3= Shared room

- Η μεταβλητή neighborhood προσδιορίζει τη περιοχή που βρίσκεται το κατάλυμα. Τα καταλύματα που καταχωρήσαμε είναι από 20 διαφορετικές περιοχές όπως φαίνονται παρακάτω.

1= 1^{ο} νεκροταφείο

2= Άνω Κυψέλη

3= Νέα Κυψέλη

4= Άνω Πατήσια

5=Νεος Κόσμος

6= Αμπελόκηποι

7= Γκάζι

8= Γκύζη

9= Γούβα

10= Γουδί

11= Πεδίο Άρεως

12= Άγιος Νικόλαος

13= Άγιος Ελευθέριος

14= Άγιος Κων/νος

15= Θησείο

16= Ιλίσια

17= Κυψέλη

18= Κουκάκι

19= Κολωνάκι

20= Εξάρχεια

6.3 Περιγραφικά μέτρα ποσοτικών και ποιοτικών μεταβλητών.

Τώρα θα βγάλουμε τα περιγραφικά μέτρα των ποσοτικών μας μεταβλητών. Από τη περιοχή Data View πατώντας analyze → descriptive statistics→frequencies επιλέγουμε τις ποσοτικές μεταβλητές που θέλουμε να αναλύσουμε και τις τοποθετούμε στο variable(s). Ακόμα από το statistics επιλέγουμε τα quartiles, min, max και range.

Από το Output παίρνουμε τους παρακάτω πίνακες.

listings

		Frequency	Percent	Valid Percent	Cumulative Percent
Valid	0	74	12,3	12,3	12,3
	1	293	48,6	48,6	60,9
	2	86	14,3	14,3	75,1

3	27	4,5	4,5	79,6
4	19	3,2	3,2	82,8
5	24	4,0	4,0	86,7
6	21	3,5	3,5	90,2
7	10	1,7	1,7	91,9
8	2	,3	,3	92,2
9	2	,3	,3	92,5
10	3	,5	,5	93,0
11	8	1,3	1,3	94,4
12	3	,5	,5	94,9
14	1	,2	,2	95,0
16	2	,3	,3	95,4
17	1	,2	,2	95,5
19	1	,2	,2	95,7
20	4	,7	,7	96,4
25	6	1,0	1,0	97,3
43	1	,2	,2	97,5
46	2	,3	,3	97,8
47	11	1,8	1,8	99,7
58	2	,3	,3	100,0
Total	603	100,0	100,0	

Πίνακας 6.1: Συχνότητες μεταβλητής listings.

Από το παραπάνω πίνακα βλέπουμε πόσες ακόμα καταχωρήσεις έχει ο ιδιοκτήτης του κάθε καταλύματος. Καμία άλλη καταχώρηση έχουν 74 από τους 603 με ποσοστό 12,3%. Άλλη μια καταχώρηση έχουν 293 με ποσοστό 48,6%. Άλλες δύο καταχωρήσεις έχουν 86 με ποσοστό 14,3%. Άλλες 3 καταχωρήσεις έχουν 27 με ποσοστό 4,5%. Άλλες 4 καταχωρήσεις έχουν 19 με ποσοστό 3,2%. Άλλες 5 καταχωρήσεις έχουν 24 με ποσοστό 4%. Άλλες 6 καταχωρήσεις έχουν 21 με ποσοστό 3,5%. Άλλες 7 καταχωρήσεις έχουν 10 με ποσοστό 1,7%.

Άλλες 8 καταχωρήσεις έχουν 2 με ποσοστό 0,3%. Άλλες 9 καταχωρήσεις έχουν 2 με ποσοστό 0,3%. Άλλες 10 καταχωρήσεις έχουν 3 με ποσοστό 0,5%. Άλλες 11 καταχωρήσεις έχουν 8 με ποσοστό 1,3%. Άλλες 12 καταχωρήσεις έχουν 3 με ποσοστό 0,5%. Άλλες 14 καταχωρήσεις έχουν 1 με ποσοστό 0,2%. Άλλες 16 καταχωρήσεις έχουν 2 με ποσοστό 0,3%. Άλλες 17 καταχωρήσεις έχουν 1 με ποσοστό 0,2%. Άλλες 19 καταχωρήσεις έχουν 1 με ποσοστό 0,2%. Άλλες 20 καταχωρήσεις έχουν 4 με ποσοστό 0,7%. Άλλες 25 καταχωρήσεις έχουν 6 με ποσοστό 1%. Άλλες 43 καταχωρήσεις έχουν 1 με ποσοστό 0,2%. Άλλες 46 καταχωρήσεις έχουν 2 με ποσοστό 0,3%. Άλλες 47 καταχωρήσεις έχουν 11 με ποσοστό 1,8%. Άλλες 58 καταχωρήσεις έχουν 2 με ποσοστό 0,3%.

Ο παρακάτω πίνακας αναφέρεται στις ελάχιστες διανυκτερεύσεις.

Statistics

minimum nights

N	Valid	603
	Missing	0
Mean		2,16
Median		2,00
Range		27
Minimum		1
Maximum		28
Percentiles	25	1,00
	50	2,00
	75	2,00

Πίνακας 6.2: Συχνότητες μεταβλητής minimum nights.

Βλέπουμε ότι η ελάχιστη διανυκτέρευση είναι μια ημέρα και οι μέγιστες διανυκτερεύσεις είναι 28 ημέρες. Ακόμα βλέπουμε ότι η μέση τιμή στις ελάχιστες διανυκτερεύσεις είναι στις 2 ημέρες.

Ο πίνακας με τα έσοδα ανά μήνα φαίνεται παρακάτω

Statistics

income/month

N	Valid	603
	Missing	0
Range		3636
Minimum		0
Maximum		3636
Percentiles	25	97,00
	50	278,00
	75	585,00

Πίνακας 6.3: Συχνότητες μεταβλητής income/ month.

Όπως βλέπουμε από το πίνακα το εύρος είναι 3.636, το ελάχιστο εισόδημα είναι 0 ευρώ και το μέγιστο 3.636 ευρώ. Επίσης παρατηρούμε ότι το 25% του δείγματος διαθέτει εισόδημα έως 97 ευρώ, το 50% διαθέτει έως 278 ευρώ και το 75% του δείγματος διαθέτει εισόδημα έως 585 ευρώ.

Ο πίνακας με τις τιμές των καταλυμάτων παρουσιάζεται στη συνέχεια.

Statistics

price

N	Valid	603
	Missing	0
Mean		46,83
Median		39,00
Range		396
Minimum		8
Maximum		404
Percentiles	25	27,00
	50	39,00
	75	55,00

Πίνακας 6.4: Συχνότητες μεταβλητής price.

Βλέπουμε ότι η ελάχιστη τιμή είναι 8 ευρώ ενώ η μέγιστη είναι στα 404 ευρώ. Η μέση τιμή των τιμών είναι στα 46,83 ευρώ ανά βράδυ.

Ο πίνακας με τις κριτικές των καταλυμάτων φαίνεται παρακάτω.

Statistics

reviews

N	Valid	603
	Missing	0
Mean		30,29
Median		13,00
Range		269
Minimum		0
Maximum		269
Percentiles	25	4,00
	50	13,00
	75	41,00

Πίνακας 6.5: Συχνότητες μεταβλητής reviews.

Βλέπουμε ότι οι ελάχιστες κριτικές από τις καταχωρήσεις μας είναι μηδέν και οι μέγιστες 269. Ο μέσος όρος κριτικών είναι στις 30.

Ο πίνακας με τα τετραγωνικά μέτρα των καταλυμάτων παρουσιάζεται στη συνέχεια.

Statistics

m^2

N	Valid	170
	Missing	433
Mean		77,88
Median		72,00
Range		320
Minimum		14
Maximum		334
Percentiles	25	50,00
	50	72,00
	75	100,00

Πίνακας 6.6: Συχνότητες μεταβλητής m^2.

Βλέπουμε ότι το κατάλυμα με τα λιγότερα τετραγωνικά μέτρα είναι 14 και το μεγαλύτερο 334. Ο μέσος όρος τετραγωνικών είναι στα 78. Για τη συγκεκριμένη μεταβλητή δεν μπορούμε να βγάλουμε πληροφορίες με ακρίβεια γιατί έχουμε πάρα πολλά κενά.

Ο πίνακας με τις αποστάσεις φαίνεται στη συνέχεια.

Statistics

		km-plaka	km-omonoia	km-archaiol
N	Valid	603	603	603
	Missing	0	0	0
Mean		2,4947	2,1202	2,5471
Median		2,3400	2,1000	2,2000
Range		5,60	3,96	6,34
Minimum		,26	,31	,06
Maximum		5,86	4,27	6,40
Percentiles	25	1,5000	1,5100	1,4500
	50	2,3400	2,1000	2,2000
	75	3,3700	2,7100	3,3600

Πίνακας 6.7: Συχνότητες μεταβλητών km-plaka, km-omonoia, km-archaiol.

Από το παραπάνω πίνακα βλέπουμε για τα καταλύματα ότι, η μικρότερη απόσταση από τη Πλάκα είναι 0,26 χιλιόμετρα ενώ η μέγιστη 5,86. Η μέση τιμή των αποστάσεων από τη Πλάκα είναι στα 2,5 χιλιόμετρα. Επίσης η μικρότερη απόσταση από την Ομόνοια είναι 0,31 χιλιόμετρα ενώ η μέγιστη 4,27. Η μέση τιμή των αποστάσεων από την Ομόνοια είναι 2,1 χιλιόμετρα. Τέλος βλέπουμε ότι η μικρότερη απόσταση από το μουσείο της ακρόπολης είναι 0,06 χιλιόμετρα ενώ η μέγιστη 6,4. Η μέση τιμή των αποστάσεων από το μουσείο είναι 2,5 χιλιόμετρα.

Ο πίνακας μας δίνει πληροφορίες για τις φωτογραφίες.

Statistics

photos

N	Valid	603
	Missing	0
Mean		24,56
Median		22,00
Range		97
Minimum		2
Maximum		99
Percentiles	25	14,00
	50	22,00
	75	30,00

Πίνακας 6.8: Συχνότητες μεταβλητής photos.

Βλέπουμε ότι οι ελάχιστες φωτογραφίες που είχε μια καταχώρηση είναι 2 ενώ οι περισσότερες είναι 99. Η μέση τιμή των φωτογραφιών που έχει μια καταχώρηση είναι 25 φωτογραφίες

Statistics

		guests	bedrooms	beds	guests-in-bedroom
N	Valid	603	603	603	603
	Missing	0	0	0	0
Mean		3,74	1,41	2,18	2,77
Median		4,00	1,00	2,00	2,00
Range		15	10	15	16
Minimum		1	0	1	0
Maximum		16	10	16	16
Percentiles	25	2,00	1,00	1,00	2,00
	50	4,00	1,00	2,00	2,00
	75	4,00	2,00	3,00	4,00

Πίνακας 6.9: Συχνότητες μεταβλητών guests, bedrooms, beds, guests in bedroom.

Από το παραπάνω πίνακα παίρνουμε ενδιαφέρουσες πληροφορίες σχετικά με τη χωρητικότητα, τα υπνοδωμάτια, τα κρεβάτια και τη χωρητικότητα των φιλοξενούμενων σε υπνοδωμάτια για τις 603 καταχωρήσεις. Βλέπουμε ότι η ελάχιστη χωρητικότητα σε κατάλυμα είναι ενός ατόμου ενώ η μέγιστη στα 16 άτομα. Αυτό

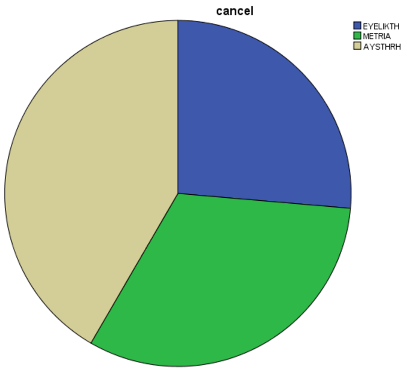
συμβαίνει διότι η καταχώρηση μπορεί να είναι ένα ολόκληρο σπίτι ή ένα κρεβάτι. Για τα υπνοδωμάτια βλέπουμε ότι τα ελάχιστα είναι μηδέν όπου αναφερόμαστε σε στούντιο ενώ η καταχώρηση με τα μέγιστα υπνοδωμάτια είναι 10.
Ο μέσος όρος υπνοδωματίων στις καταχωρήσεις είναι 1. Για τα κρεβάτια μπορούμε να πούμε ότι ο ελάχιστος αριθμός είναι 1 ενώ ο μέγιστος 16 και ο μέσος όρος είναι στα 2 κρεβάτια. Μια σημαντική πληροφορία είναι η χωρητικότητα σε υπνοδωμάτια, διότι μπορεί να έχει υψηλή χωρητικότητα ένα κατάλυμα αλλά οι φιλοξενούμενοι να μην κοιμούνται σε υπνοδωμάτια. Παρατηρούμε ότι υπάρχει κατάλυμα όπου να μην μπορεί να κοιμηθεί ο φιλοξενούμενος σε υπνοδωμάτιο δηλαδή να είναι ένα στούντιο αλλά και ο μέγιστος αριθμός φιλοξενίας σε υπνοδωμάτια είναι στα 16 άτομα. Ο μέσος όρος είναι να κοιμούνται 3 άτομα σε υπνοδωμάτιο.

Ένα ενδιαφέρον κομμάτι είναι να δούμε τα περιγραφικά μέτρα των ποιοτικών μεταβλητών που δεν αναλύσαμε στο προηγούμενο κεφάλαιο. Από τη περιοχή Data View πατώντας analyze → descriptive statistics→frequencies επιλέγουμε τις ποιοτικές μεταβλητές που θέλουμε να αναλύσουμε και τις τοποθετούμε στο variable(s). Παρακάτω βλέπουμε τους πίνακες συχνοτήτων για κάθε μια ανεξάρτητη μεταβλητή.

cancel

		Frequency	Percent	Valid Percent	Cumulative Percent
Valid	EYELIKTH	159	26,4	26,4	26,4
	METRIA	193	32,0	32,0	58,4
	AYSTHRH	251	41,6	41,6	100,0
	Total	603	100,0	100,0	

Πίνακας 6.10: Συχνότητες μεταβλητής cancel.



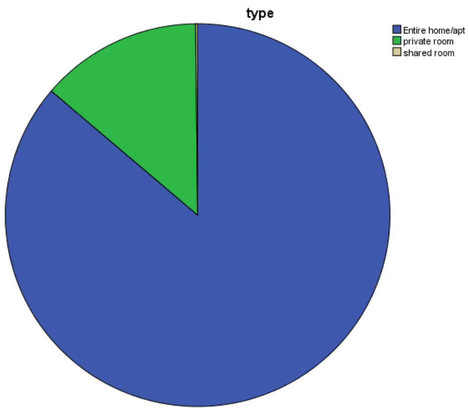
Εικόνα 6.1:Πίτα συχνοτήτων μεταβλητής cancel.

Ο παραπάνω πίνακας μας δίνει πληροφορίες σχετικά με την ακύρωση. Βλέπουμε ότι το 26,4% των καταλυμάτων έχει ευέλικτη ακύρωση, το 32% έχει μέτρια και τέλος το 41,6% έχει αυστηρή ακύρωση.

type

		Frequency	Percent	Valid Percent	Cumulative Percent
Valid	Entire home/apt	520	86,2	86,2	86,2
	private room	82	13,6	13,6	99,8
	shared room	1	,2	,2	100,0
	Total	603	100,0	100,0	

Πίνακας 6.11: Συχνότητες μεταβλητής type.



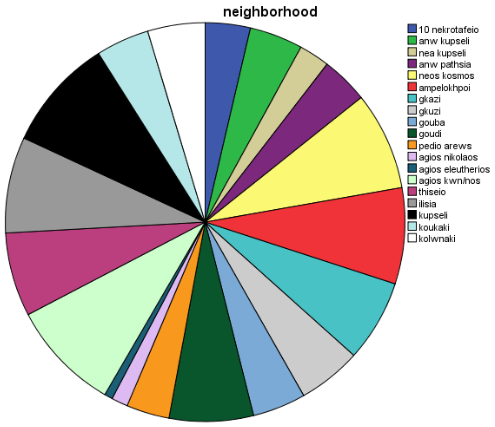
Εικόνα 6.2:Πίτα συχνοτήτων μεταβλητής type.

Ο παραπάνω πίνακας μας δίνει πληροφορίες σχετικά με το τύπο των καταλυμάτων. Βλέπουμε ότι από τα 603 καταλύματα το 86,2% αναφέρεται σε ολόκληρα διαμερίσματα, το 13,6 % σε ιδιωτικά δωμάτια και το 0,2% σε κοινόχρηστο δωμάτιο. Παρατηρούμε ότι πολύ μεγάλο ποσοστό καταλαμβάνουν τα ολόκληρα διαμερίσματα, όπου στο τέλος μπορεί να παρατηρήσουμε επίδραση στα έσοδα ανά μήνα αν το κατάλυμα είναι του τύπου αυτού.

neighborhood

		Frequency	Percent	Valid Percent	Cumulative Percent
Valid	10 nekrotafeio	22	3,6	3,6	3,6
	anw kupseli	26	4,3	4,3	8,0
	nea kupseli	15	2,5	2,5	10,4
	anw pathsia	23	3,8	3,8	14,3
	neos kosmos	48	8,0	8,0	22,2
	ampelokhpoi	47	7,8	7,8	30,0
	gkazi	40	6,6	6,6	36,7
	gkuzi	31	5,1	5,1	41,8
	gouba	26	4,3	4,3	46,1
	goudi	41	6,8	6,8	52,9
	pedio arews	21	3,5	3,5	56,4
	agios nikolaos	8	1,3	1,3	57,7
	agios eleutherios	4	,7	,7	58,4
	agios kwn/nos	54	9,0	9,0	67,3
	thiseio	41	6,8	6,8	74,1
	ilisia	47	7,8	7,8	81,9
	kupseli	55	9,1	9,1	91,0
	koukaki	26	4,3	4,3	95,4
	kolwnaki	28	4,6	4,6	100,0
	Total	603	100,0	100,0	

Πίνακας 6.12: Συχνότητες μεταβλητής neighborhood.



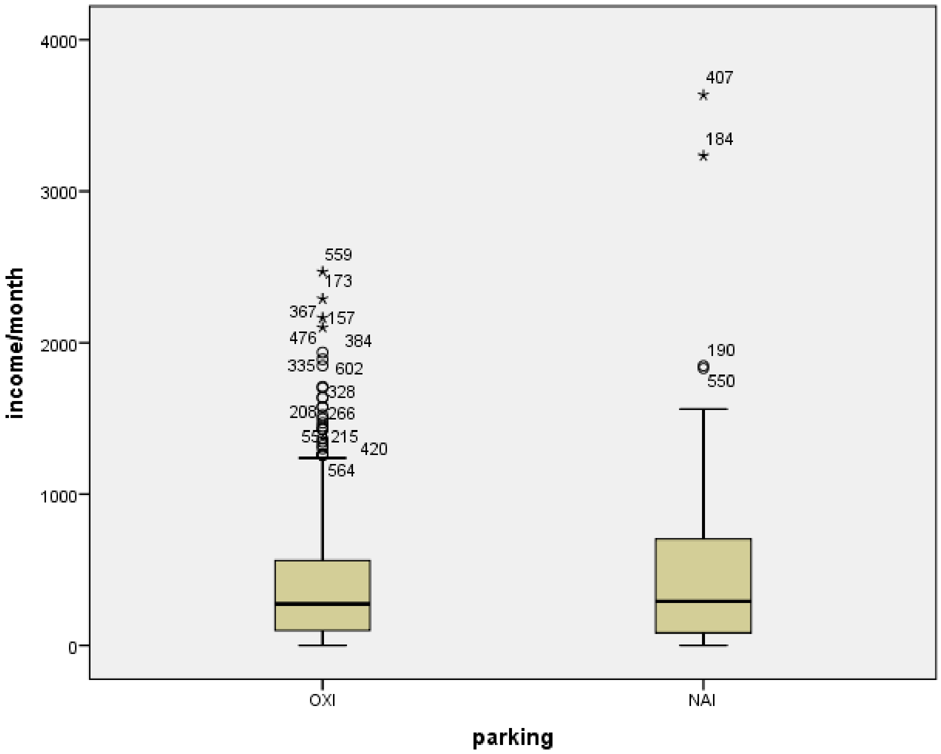
Εικόνα 6.1:Πίτα συχνοτήτων μεταβλητής neighborhood.

Ο παραπάνω πίνακας μας δίνει πληροφορίες για την τοποθεσία των καταλυμάτων. Παρατηρούμε ότι από τις 603 καταχωρήσεις το 3,6% των καταλυμάτων βρίσκεται στο 1^{ο} νεκροταφείο, το 4,3% στην Άνω Κυψέλη, το 2,5% στη Νέα Κυψέλη, το 3,8% στα Άνω Πατήσια, το 8% στο Νέο Κόσμο, το 7,8% στους Αμπελόκηπους, το 6,6% στο Γκάζι, το 5,1% στο Γκύζη, το 4,3% στη Γούβα, το 6,8% στο Γουδί, το 3,5% στο Πεδίο του Άρεως, το 1,3% στον Άγιο Νικόλαο, το 0,7% στον Άγιο Ελευθέριο, το 9% στον Άγιο Κων/νο, το 6,8% στο Θησείο, το 7,8% στα Ιλίσια, το 9,1% στη Κυψέλη, το 4,3% στο Κουκάκι και τέλος το 4,6% στο Κολωνάκι.

6.4 Σύγκριση εσόδων με συγκεκριμένες μεταβλητές.

Από τα περιγραφικά μέτρα που εμφανίσαμε στο προηγούμενο κεφάλαιο, παρατηρήσαμε ότι ένα μικρό ποσοστό καταλυμάτων διαθέτει αρκετές και σημαντικές παροχές. Ενδιαφέρον θα ήταν να δούμε πως κυμαίνονται τα έσοδα ανά μήνα σε σύγκριση με το αν διαθέτει κάποια παροχή το κατάλυμα και ανάλογα τη περιοχή που βρίσκεται. Από τη περιοχή Data View πατώντας analyze → descriptive statistics→ explore επιλέγουμε τη μεταβλητή εισόδημα ανά μήνα στη dependent list όπου θέλουμε να εξετάσουμε και στη factor list τη μεταβλητή που θέλουμε να συγκρίνουμε κάθε φορά.

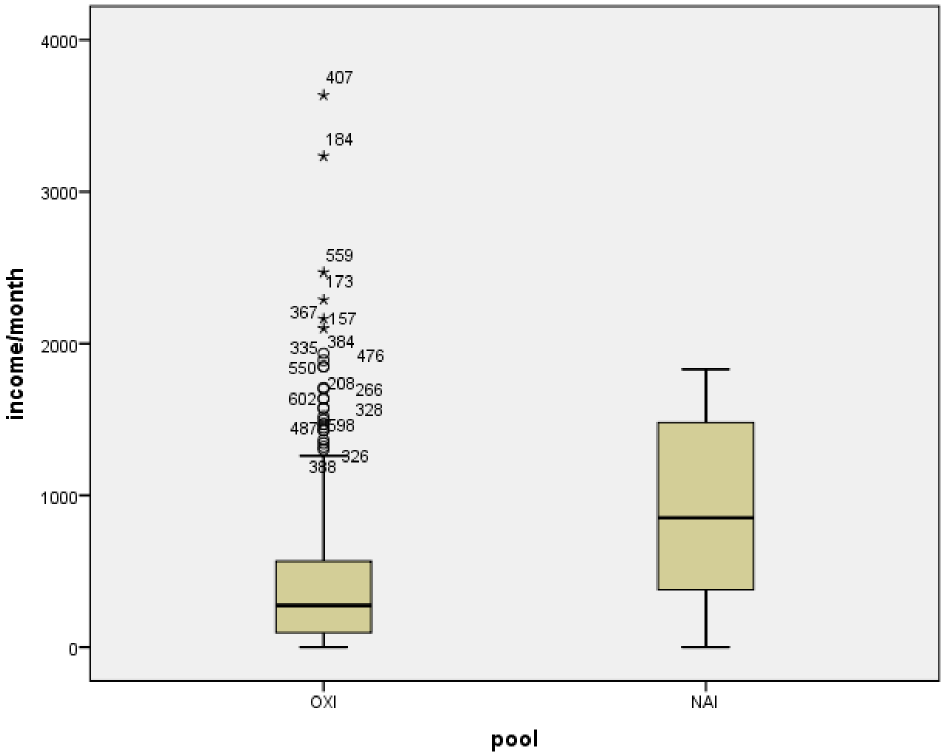
Το παρακάτω διάγραμμα εμφανίζει το μέσο όρο των εσόδων σε σχέση με τη μεταβλητή parking.



Εικόνα 6.2:Σύγκριση εσόδων με τη μεταβλητή parking.

Από τα περιγραφικά μέτρα παρατηρήσαμε ότι παρκινγκ διαθέτουν μόνο 86 καταλύματα από τα 603 και είναι ενδιαφέρον να δούμε πως κυμαίνεται το εισόδημα. Τα έσοδα ανά μήνα στα καταλύματα που δεν διαθέτουν παρκινγκ κυμαίνεται στα 401 ευρώ ενώ σε αυτά που διαθέτουν κυμαίνεται στα 498 ευρώ. Μπορούμε να πούμε ότι δεν επηρεάζει σημαντικά η παροχή παρκινγκ το εισόδημα ανά μήνα του ιδιοκτήτη. Επίσης βλέπουμε ότι υπάρχουν αρκετά καταλύματα που δεν διαθέτουν παρκινγκ αλλά εμφανίζουν αρκετά υψηλά έσοδα με μέγιστο το κατάλυμα με ID 559 που εμφανίζει 2468 ευρώ και ελάχιστο το κατάλυμα με ID 564 που εμφανίζει 1260 ευρώ. Για όλα αυτά τα καταλύματα που δεν διαθέτουν παρκινγκ αλλά έχουν υψηλά έσοδα, μπορούμε να πούμε ότι ίσως διαθέτουν κάποια άλλη παροχή πιο σημαντική και για αυτό έχουν υψηλά έσοδα ανά μήνα.

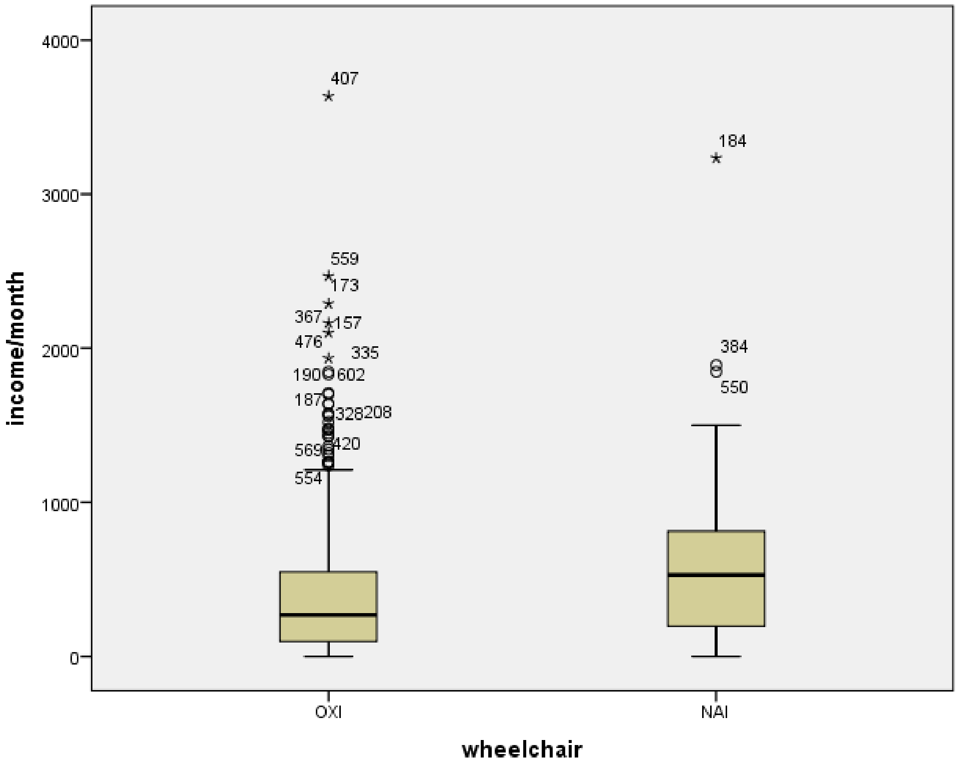
Τώρα θα δούμε το διάγραμμα που εμφανίζει το μέσο όρο των εσόδων σε σχέση με τη μεταβλητή pool.



Εικόνα 6.3:Σύγκριση εσόδων με τη μεταβλητή pool.

Είχαμε παρατηρήσει ότι πισίνα διαθέτουν μόνο 10 καταλύματα από τα 603 και τώρα ήρθε η ώρα να δούμε πως κυμαίνεται το εισόδημα για τα καταλύματα που διαθέτουν και για αυτά που δεν διαθέτουν. Τα έσοδα ανά μήνα στα καταλύματα που δεν διαθέτουν πισίνα κυμαίνεται στα 407 ευρώ ενώ αυτά που διαθέτουν εμφανίζουν εισόδημα στα 878 ευρώ. Βλέπουμε ότι υπάρχει σημαντική διαφορά και αυτό μπορούμε να το αιτιολογήσουμε ότι για να διαθέτει πισίνα δεν θα είναι ένα απλό κατάλυμα. Ωστόσο παρατηρούμε ότι υπάρχουν και πολλά άλλα καταλύματα που δεν διαθέτουν πισίνα άλλα εμφανίζουν αρκετά υψηλά έσοδα, όπως το κατάλυμα με ID 407 που εμφανίζει 3636 ευρώ ανά μήνα. Για τα συγκεκριμένα μπορούμε να πούμε ότι θα διαθέτουν κάποιες άλλες παροχές για να εμφανίζει τόσο υψηλά έσοδα.

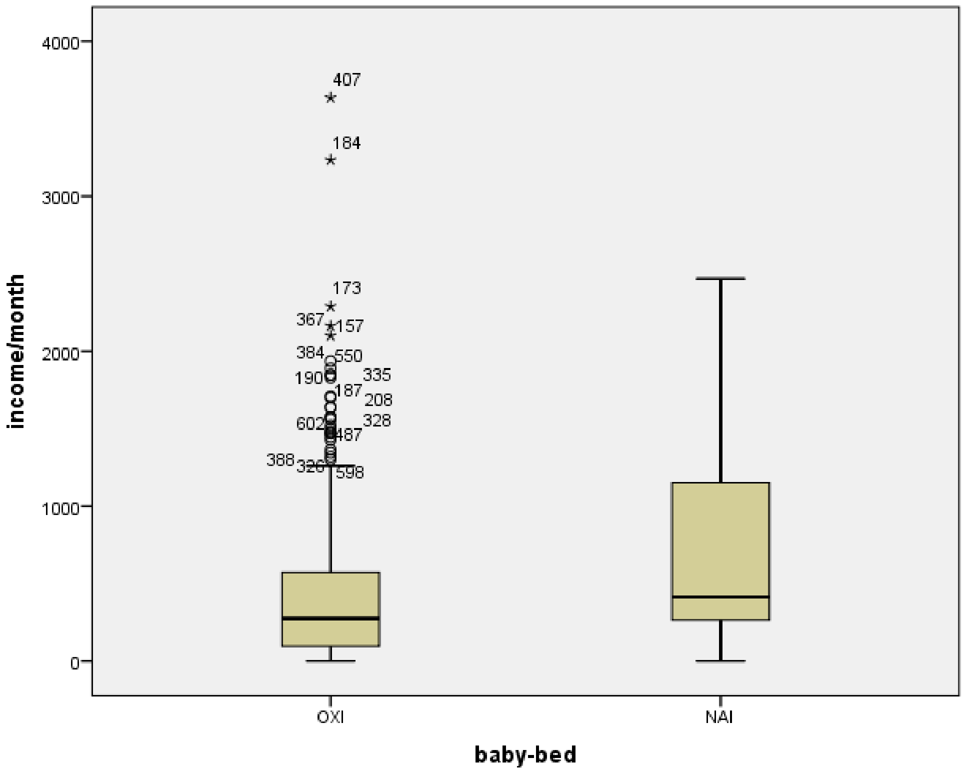
Το διάγραμμα εμφανίζει το μέσο όρο των εσόδων σε σχέση με τη μεταβλητή wheelchair.



Εικόνα 6.4:Σύγκριση εσόδων με τη μεταβλητή wheelchair.

Σημαντική ήταν και η πληροφορία που είχαμε πάρει, ότι από τα 603 καταλύματα μόνο τα 43 είχαν είσοδο για αναπηρικό αμαξίδιο. Τώρα θα δούμε πως κυμαίνονται τα έσοδα αναλόγως τη συγκεκριμένη παροχή. Τα έσοδα ανά μήνα στα καταλύματα που δεν διαθέτουν είσοδο για αναπηρικό αμαξίδιο κυμαίνεται στα 398 ευρώ ενώ αυτά που διαθέτουν είσοδο για αναπηρικό αμαξίδιο στα 630. Παρατηρούμε αισθητή διαφορά εσόδων και θα μπορούσαμε να πούμε ότι η συγκεκριμένη παροχή είναι πολύ σημαντική για την αύξηση των εσόδων ανά μήνα. Επίσης βλέπουμε ότι υπάρχουν καταλύματα που δεν διαθέτουν τη συγκεκριμένη παροχή αλλά εμφανίζουν υψηλότερα έσοδα από το μέσο όρο με υψηλότερο το κατάλυμα με ID 407 που εμφανίζει 3636 ευρώ. Τα συγκεκριμένα ίσως να διαθέτουν κάποιες άλλες παροχές και ίσως να βρίσκονται σε καλύτερη περιοχή.

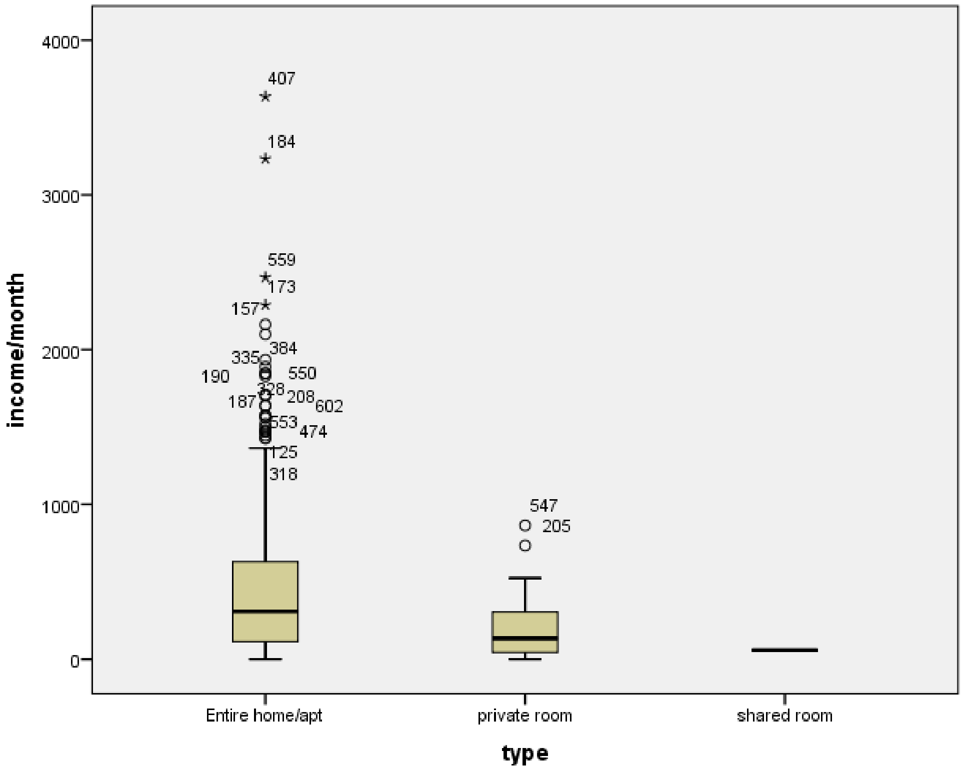
Στη συνέχεια θα δούμε το διάγραμμα που εμφανίζει το μέσο όρο των εσόδων σε σχέση με τη μεταβλητή baby bed.



Εικόνα 6.5:Σύγκριση εσόδων με τη μεταβλητή baby bed.

Από τα περιγραφικά μέτρα παρατηρήσαμε ότι κρεβάτι μωρού διαθέτουν μόνο 13 καταλύματα από τα 603 και είναι ενδιαφέρον να δούμε πως κυμαίνονται τα έσοδα. Τα έσοδα ανά μήνα στα καταλύματα που δεν διαθέτουν κρεβάτι μωρού κυμαίνεται στα 408 ευρώ ενώ σε αυτά που διαθέτουν κυμαίνεται στα 735 ευρώ.
Επίσης μπορούμε να πούμε ότι η συγκεκριμένη μεταβλητή αυξάνει σημαντικά τα έσοδα ανά μήνα. Όπως και στις προηγούμενες συγκρίσεις που κάναμε, έτσι και εδώ παρατηρούμε ότι υπάρχουν καταλύματα που δεν διαθέτουν τη συγκεκριμένη μεταβλητή αλλά εμφανίζουν υψηλά έσοδα, και αυτό γίνεται διότι θα έχουν κάποιες άλλες παροχές.

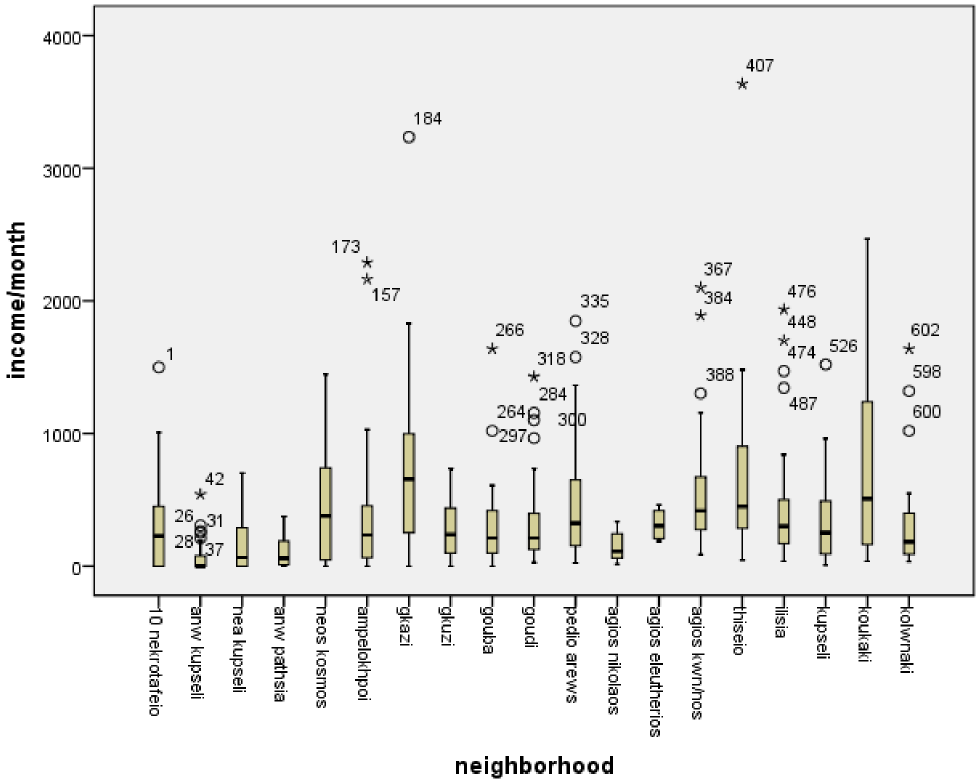
Το διάγραμμα εμφανίζει το μέσο όρο των εσόδων σε σχέση με το τύπο του καταλύματος.



Εικόνα 6.6:Σύγκριση εσόδων με τη μεταβλητή type.

Στη συγκεκριμένη σύγκριση βλέπουμε πώς κυμαίνονται τα έσοδα σε σχέση με το τύπο του καταλύματος. Παρατηρούμε ότι τα καταλύματα τύπου ολόκληρου διαμερίσματος εμφανίζουν υψηλότερα έσοδα με σχέση με τα ιδιωτικά δωμάτια και συγκεκριμένα τα πρώτα εμφανίζουν 450 ευρώ ανά μήνα ενώ ο δεύτερος τύπος καταλυμάτων εμφανίζουν 192 ευρώ ανά μήνα. Αυτό είναι λογικό διότι στα ολόκληρα διαμερίσματα οι ενοικιαστές έχουν ολόκληρο το κατάλυμα στη διάθεση τους και δεν συγκατοικούν με άλλους. Επίσης υπάρχει και μεγαλύτερη χωρητικότητα φιλοξενούμενων σε σχέση με τα ιδιωτικά δωμάτια που συνεπάγεται με υψηλότερη τιμή. Βλέπουμε ότι υπάρχουν και δυο καταλύματα τύπου ιδιωτικού δωματίου που παρουσιάζουν υψηλότερα έσοδα από το μέσο όρο τους, με υψηλότερο το κατάλυμα με ID 547 που εμφανίζει 865 ευρώ ανά μήνα. Αυτό συμβαίνει διότι θα έχουν περισσότερες παροχές. Για τα καταλύματα τύπου κοινόχρηστου δωματίου δεν μπορούμε να βγάλουμε κάποιο συμπέρασμα διότι κατείχαν πάρα πολύ μικρό ποσοστό στο δείγμα μας.

Το συγκεκριμένο διάγραμμα εμφανίζει το μέσο όρο των εσόδων σε σχέση με τη περιοχή.



Εικόνα 6.7:Σύγκριση εσόδων με τη μεταβλητή neighborhood.

Πολύ ενδιαφέρων πληροφορίες θα πάρουμε από το παραπάνω πίνακα και συγκεκριμένα θα δούμε που κυμαίνονται τα έσοδα ανά μήνα αναλόγως τη περιοχή που βρίσκεται το κατάλυμα.

Για τη περιοχή του 1^{ο} Νεκροταφείου τα έσοδα κυμαίνονται στα 327 ευρώ ανά μήνα εκτός από ένα κατάλυμα που εμφανίζει αυξημένο διότι θα έχει περισσότερες παροχές σε σχέση με τα άλλα. Η συγκεκριμένη περιοχή δεν είναι και από τις πιο δημοφιλείς στην Αθήνα για αυτό μπορεί να μην υπάρχει η μέγιστη ζήτηση με αποτέλεσμα να μην είναι από τις περιοχές με το πιο αυξημένα έσοδα ανά μήνα. Ακόμα μπορούμε να πούμε ότι υπάρχει σημαντική απόσταση από περιοχές με μεγάλη ζήτηση.

Για τη περιοχή της Άνω Κυψέλης τα έσοδα κυμαίνονται στα 77 ευρώ ανά μήνα ωστόσο υπάρχουν και κάποια υψηλά έσοδα όπως στα καταλύματα με ID 37,28,31,26 και 42 με έσοδα 251,306,210,264 και 540 ευρώ αντίστοιχα. Επίσης είναι από τις περιοχές με τα χαμηλότερα έσοδα. Η περιοχή της Άνω Κυψέλης είναι μακριά από το κέντρο της Αθήνας και ίσως για αυτό να παρουσιάζει χαμηλά έσοδα.

Για τη περιοχή της Νέας Κυψέλης τα έσοδα κυμαίνονται στα 180 ευρώ ανά μήνα. Παρατηρούμε ότι τα έσοδα όλων των καταλυμάτων κυμαίνονται στο μέσο όρο δηλαδή τα 180 ευρώ. Παρομοίως με τις παραπάνω περιοχές μπορούμε να πούμε ότι δεν είναι κοντά στο κέντρο της Αθήνας και για αυτό δεν υπάρχει υψηλή ζήτηση ώστε να αυξάνεται το εισόδημα.

Για τη περιοχή των Άνω Πατησίων τα έσοδα κυμαίνονται στα 105 ευρώ ανά μήνα και μπορούμε ν πούμε ότι είναι χαμηλό. Επίσης και σε αυτή τη περιοχή όλα τα έσοδα των καταλυμάτων κυμαίνονται στα 105 ευρώ.

Για τη περιοχή του Νέου Κόσμου τα έσοδα κυμαίνονται στα 452 ευρώ ανά μήνα. Έχει υψηλότερα έσοδα σε σχέση με τις προηγούμενες περιοχές. Η περιοχή του Νέου Κόσμου είναι κοντά στο κέντρο της Αθήνας και τέλος βλέπουμε ότι τα έσοδα όλων των καταλυμάτων κυμαίνεται στα 452 ευρώ.

Για τη περιοχή του Αμπελόκηπου τα έσοδα κυμαίνονται στα 364 ευρώ ανά μήνα όπου μπορούμε να πούμε ότι μοιάζει με τη περιοχή του Νέου Κόσμου. Παρατηρούμε ότι υπάρχουν τα καταλύματα με ID 157 και 173 όπου τα έσοδα τους είναι υψηλότερα από το μέσο όρο και συγκεκριμένα είναι 2163 και 2289 ευρώ αντίστοιχα. Τα συγκεκριμένα θα έχουν κάποια σημαντική παροχή για να αιτιολογήσει τη μεγάλη αύξηση των εσόδων.

Για τη περιοχή του Γκάζι τα έσοδα κυμαίνονται στα 748 ευρώ ανά μήνα και είναι από τις περιοχές με το μεγαλύτερο εισόδημα ανά μήνα. Η συγκεκριμένη περιοχή είναι πολύ γνωστή στην Αθήνα οπότε θα υπάρχει και μεγάλη ζήτηση με αποτέλεσμα να εμφανίζεται και αυτή τη μεγάλη τιμή των εσόδων. Βλέπουμε ότι μόνο το κατάλυμα με ID 184 εμφανίζει υψηλότερα έσοδα από το μέσο όρο και συγκεκριμένα 3234 ευρώ.

Για τη περιοχή του Γκύζη τα έσοδα κυμαίνονται στα 275 ευρώ ανά μήνα. Μπορούμε να πούμε ότι επειδή δεν είναι και πολύ κοντά στο κέντρο της Αθήνας να μην υπάρχει τόση πολύ ζήτηση με αποτέλεσμα να παρουσιάζει και ένα μέτριο εισόδημα. Όλα τα έσοδα των καταλυμάτων κυμαίνονται στο μέσο όρο της περιοχής.

Για τη περιοχή της Γούβας και τη περιοχή του Γουδί τα έσοδα κυμαίνονται στα 316 και 345 ευρώ ανά μήνα αντίστοιχα. Είναι δυο περιοχές όπου τα μέσα έσοδα τους είναι πολύ κοντά ωστόσο και στις δυο περιοχές υπάρχουν καταλύματα που εμφανίζουν έσοδα μεγαλύτερα από το μέσο όρο της κάθε περιοχής, όπου φαίνονται στο παραπάνω πίνακα.

Για τη περιοχή του Πεδίου του Άρεως τα έσοδα κυμαίνονται στα 540 ευρώ ανά μήνα. Είναι μια περιοχή πολύ γνωστή με αποτέλεσμα να υπάρχει υψηλή ζήτηση και αυξημένο εισόδημα. Στη συγκεκριμένη περιοχή υπάρχουν καταλύματα που έχουν υψηλότερα έσοδα από το μέσο όρο της περιοχής και συγκεκριμένα με ID 328 και 335 με εισοδήματα 1575 ευρώ και 1848 ευρώ ανά μήνα αντίστοιχα.

Για τη περιοχή του Αγίου Νικολάου τα έσοδα κυμαίνονται στα 148 ευρώ ανά μήνα. Μπορούμε να πούμε ότι και αυτή η περιοχή δεν έχει μεγάλη ζήτηση έτσι εμφανίζει χαμηλό εισόδημα. Όλα τα έσοδα των καταλυμάτων κυμαίνονται στο μέσο όρο της περιοχής.

Για τη περιοχή του Αγίου Ελευθερίου, τη περιοχή της Κυψέλης και τη περιοχή του Κολωνακίου τα έσοδα είναι αντίστοιχα 313,340 και 333 ευρώ ανά μήνα. Μπορούμε να αναφέρουμε ότι και οι τρεις περιοχές είναι παρόμοιες ως προς το μέσο εισόδημα ανά μήνα. Παρατηρούμε ότι στη περιοχή του Αγίου Ελευθερίου τα εισοδήματα των καταλυμάτων κυμαίνονται στο μέσο όρο της περιοχής ενώ στη περιοχή της Κυψέλης και τη περιοχή του Κολωνακίου υπάρχουν καταλύματα με έσοδα μεγαλύτερα από το μέσο όρο κάθε περιοχής όπως φαίνονται στο παραπάνω πίνακα.

Για τη περιοχή του Αγίου Κων/νου και τη περιοχή των Ιλισίων τα έσοδα είναι 538 ευρώ και 419 ευρώ αντίστοιχα. Είναι δυο περιοχές που μοιάζουν ως προς τα έσοδα και παρατηρούμε ότι και στις δυο περιοχές υπάρχουν καταλύματα με μεγαλύτερα έσοδα σε σχέση με το μέσο όρο της κάθε περιοχής.

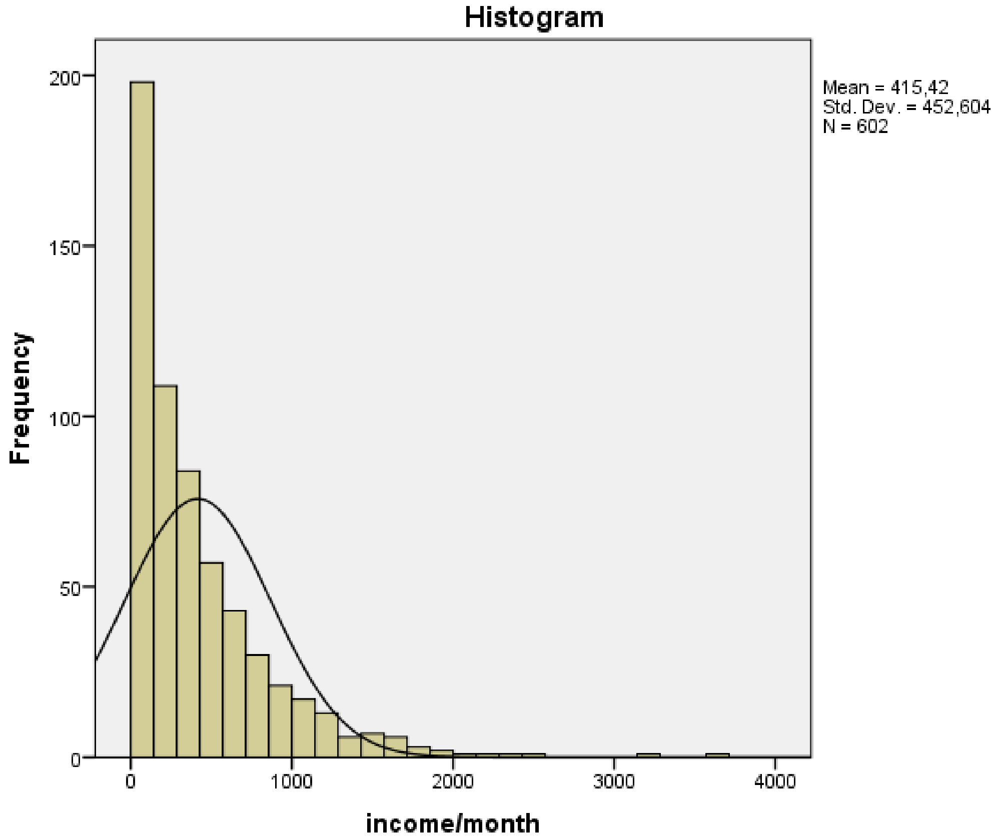
Για τη περιοχή του Θησείου και τη περιοχή του Κουκάκι τα έσοδα είναι 657 ευρώ και 725 ευρώ ανά μήνα αντίστοιχα. Βλέπουμε ότι μοιάζουν με τη περιοχή του Γκάζι αφού είναι και αυτές πολύ δημοφιλείς περιοχές. Και στις δυο περιοχές τα έσοδα ανά μήνα των καταλυμάτων κυμαίνονται στα μέσα έσοδα της κάθε περιοχής.

Στη συνέχεια θα γίνει ο έλεγχος των προϋποθέσεων ώστε να εφαρμόσουμε τη μέθοδο της βηματικής παλινδρόμησης. Ενδιαφέρον θα είναι να δούμε το αποτέλεσμα της διαδικασίας και να καταλήξουμε τελικά ποιες είναι οι πιο σημαντικές μεταβλητές.

6.5 Έλεγχος προϋποθέσεων της βηματικής παλινδρόμησης.

Αρχικά ελέγχουμε την κατανομή της Εξαρτημένης μεταβλητής, αν παρουσιάζει κάποια ασυμμετρία, όπου είναι μια από τις βασικότερες προϋποθέσεις.

Παρακάτω φαίνεται το ιστόγραμμα.



Εικόνα 6.8 Ιστόγραμμα εμφάνισης κανονικότητας εξαρτημένης μεταβλητής.

Παρατηρούμε ότι υπάρχει μια θετική ασυμμετρία και συγκεκριμένα βλέπουμε να εμφανίζονται κάποιες ακραίες τιμές, πράγμα που δεν θέλουμε ώστε να πληρούμε την προϋπόθεση της κανονικής κατανομής.

Αυτό που υποθέτουμε από το Ιστόγραμμα το επιβεβαιώνουμε και στατιστικά μέσω του μη παραμετρικού ελέγχου Kolmogorov-Smirnof δηλαδή θα απορρίψουμε ή θα δεχτούμε την αρχική υπόθεση αν η κατανομή είναι κανονική.

Παρακάτω φαίνονται τα αποτελέσματα από τον έλεγχο που πήραμε από το Output.

One-Sample Kolmogorov-Smirnov Test

		income/month
N		602
Normal Parameters[a,b]	Mean	415,42
	Std. Deviation	452,604
Most Extreme Differences	Absolute	,179
	Positive	,142
	Negative	-,179
Kolmogorov-Smirnov Z		4,400
Asymp. Sig. (2-tailed)		,000

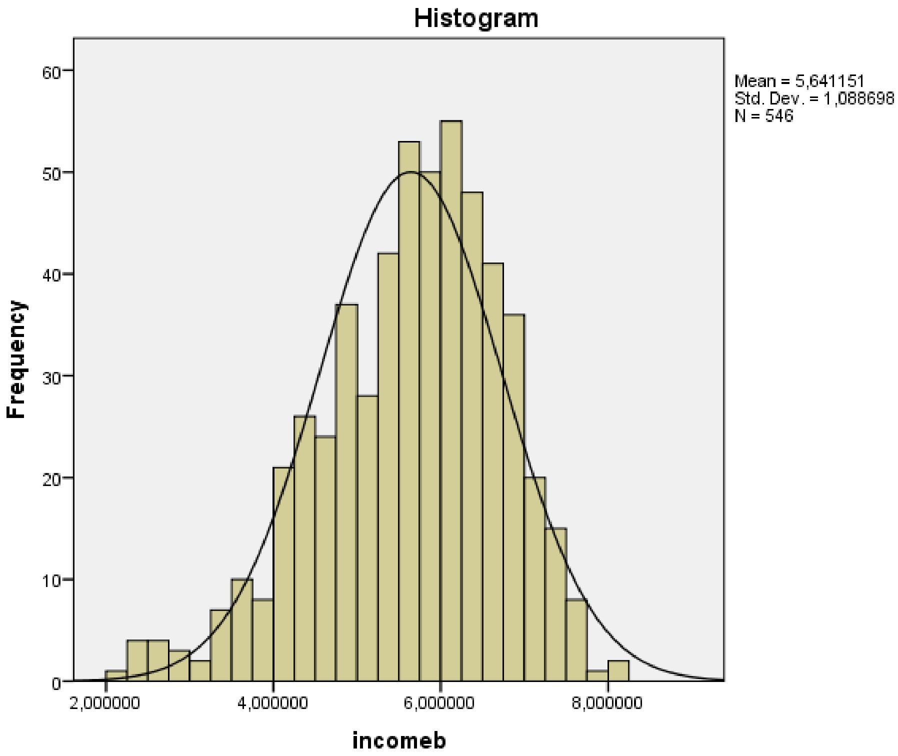
a. Test distribution is Normal.

b. Calculated from data.

Πίνακας 6.13: One-Sample Kolmogorov-Smirnov Test.

Βλέπουμε ότι η τιμή του sig είναι μικρότερη του 0,05 άρα η κατανομή δεν είναι κανονική, πράγμα που μας εμποδίζει να συνεχίσουμε την παλινδρόμηση μας. Για να αντιμετωπίσουμε το πρόβλημα πρέπει να τροποποιήσουμε τη μεταβλητή μας. Στο συγκεκριμένο παράδειγμα θα επιλέξουμε να λογαριθμίσουμε την εξαρτημένη μεταβλητή μας.

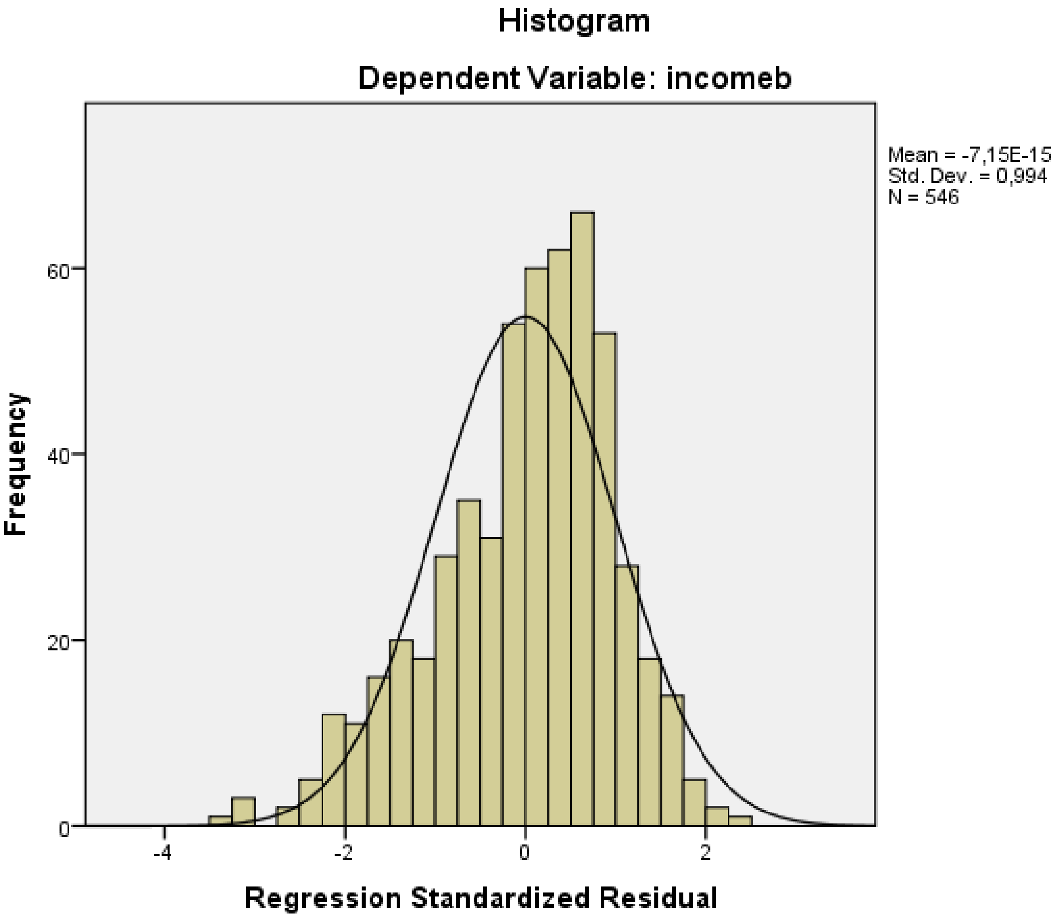
Παρακάτω εμφανίζουμε το νέο ιστόγραμμα των λογαριθμισμένων τιμών.



Εικόνα 6.9 Ιστόγραμμα εμφάνισης κανονικότητας εξαρτημένης μεταβλητής.

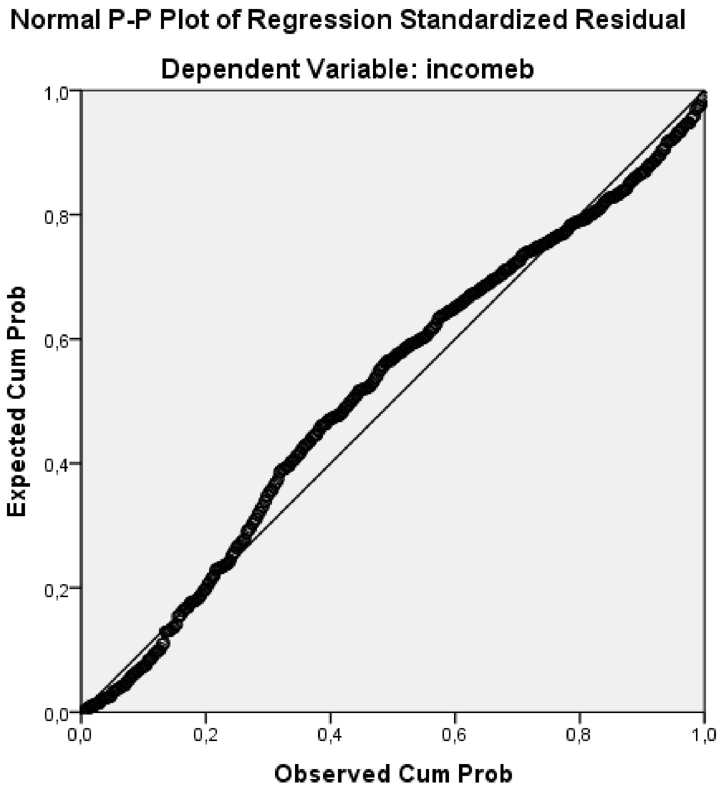
Από το ιστόγραμμα δεν απορρίπτουμε την υπόθεση της κανονικότητας διότι η εξαρτημένη μεταβλητή πλησιάζει πάρα πολύ κοντά στην κανονική κατανομή χωρίς να παρουσιάζει ακραίες τιμές.

- Κανονικότητα καταλοίπων: γίνεται έλεγχος αν τα κατάλοιπα ακολουθούν κανονική κατανομή. Η ύπαρξη κανονικής κατανομής των καταλοίπων δημιουργεί μη ύπαρξη συσχέτισης τους. Αυτό φαίνεται από το ιστόγραμμα που θα παρουσιάσαμε παρακάτω.



Εικόνα 6.10: Ιστόγραμμα εμφάνισης κανονικότητας καταλοίπων.

Σε ένα επόμενο βήμα πρέπει να ελέγξουμε αν οι ανεξάρτητες μεταβλητές που θα χρησιμοποιηθούν για την επεξήγηση της συμπεριφοράς της εξαρτημένης μεταβλητής συνδέονται γραμμικά με το εξαρτημένο μέγεθος. Δηλαδή θα έπρεπε στην βέλτιστή περίπτωση να είναι μια ευθεία γραμμή. Ωστόσο, αυτό είναι πολύ δύσκολο να συμβεί σε πραγματικά δεδομένα. Παρακάτω βλέπουμε στη δικιά μας περίπτωση από το P-P Plot ότι πλησιάζει αρκετά στην ευθεία γραμμή.



Εικόνα 6.11: Normal P-P Plot Regression Standardized Residual.

- Ανεξαρτησία καταλοίπων: πρέπει τα κατάλοιπα να παρουσιάζουν ανεξαρτησία μεταξύ τους διότι αλλιώς θα υπάρχει αυτοσυσχέτιση μεταξύ τους και μπορεί να παραληφθεί κάποια σημαντική ανεξάρτητη μεταβλητή. Αυτό μπορούμε να το δούμε από το πίνακα model summary κοιτώντας το δείκτη Durbin-Watson όπου η τιμή του θέλουμε να είναι κοντά στη τιμή 2.Εξασφαλιζουμε και αυτή τη προϋπόθεση αφού ο δείκτης μας έχει τη τιμή 1,903 όπως φαίνεται και παρακάτω άρα υπάρχει ανεξαρτησία μεταξύ των καταλοίπων.

Model Summary[h]

					Change Statistics					
Model	R	R Square	Adjusted R Square	Std. Error of the Estimate	R Square Change	F Change	df1	df2	Sig. F Change	Durbin-Watson
1	,460[a]	,212	,210	,967429428	,212	146,092	1	544	,000	
2	,545[b]	,297	,294	,914677535	,085	65,557	1	543	,000	
3	,591[c]	,350	,346	,880195192	,053	44,378	1	542	,000	
4	,620[d]	,385	,380	,856946124	,035	30,808	1	541	,000	
5	,626[e]	,392	,386	,852947987	,007	6,084	1	540	,014	
6	,632[f]	,399	,392	,848717698	,007	6,396	1	539	,012	
7	,637[g]	,406	,398	,844738312	,007	6,090	1	538	,014	1,903

a. Predictors: (Constant), reviews
b. Predictors: (Constant), reviews, km-archaiol
c. Predictors: (Constant), reviews, km-archaiol, photos
d. Predictors: (Constant), reviews, km-archaiol, photos, price
e. Predictors: (Constant), reviews, km-archaiol, photos, price, air-conditioning
f. Predictors: (Constant), reviews, km-archaiol, photos, price, air-conditioning, buzzer
g. Predictors: (Constant), reviews, km-archaiol, photos, price, air-conditioning, buzzer, priv-entrance
h. Dependent Variable: incomeb

Πίνακας 6.14: Model summary.

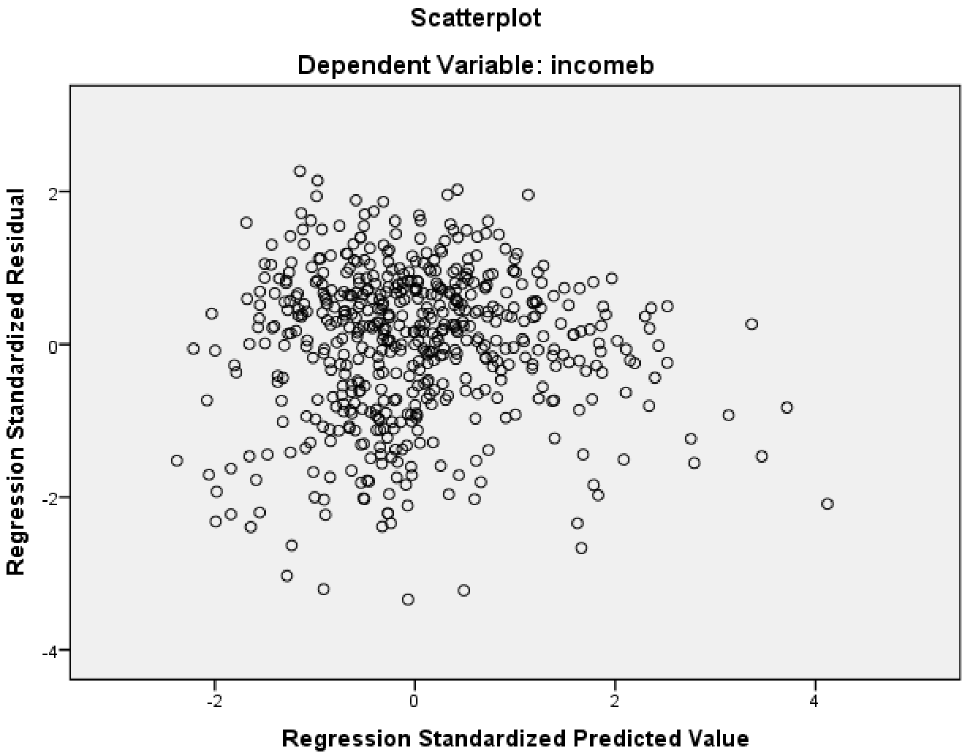
- Πολυσυγραμμικότητα: θα ελέγξουμε αν υπάρχει γραμμική συσχέτιση μεταξύ των ανεξάρτητων μεταβλητών, πράγμα που δεν θέλουμε. Ο έλεγχος θα γίνει κοιτώντας από το πίνακα coefficients τις τιμές των στηλών Tolerance και VIF για το μοντέλο 7.Πρέπει οι τιμές να είναι κοντά στη μονάδα για να ικανοποιείτε και αυτή η προϋπόθεση. Όπως φαίνεται παρακάτω ισχύει και στη δική μας περίπτωση.

Coefficients[a]

Model		Unstandardized Coefficients		Standardized Coefficients	t	Sig.	Correlations			Collinearity Statistics	
		B	Std. Error	Beta			Zero-order	Partial	Part	Tolerance	VIF
	reviews	,010	,001	,386	10,888	,000	,460	,424	,377	,953	1,049
	km-archaiol	-,230	,028	-,287	-8,243	,000	-,334	-,334	-,285	,990	1,010
	photos	,016	,002	,235	6,662	,000	,324	,275	,231	,960	1,041
4	(Constant)	5,188	,113		45,776	,000					
	reviews	,010	,001	,390	11,278	,000	,460	,436	,380	,953	1,049
	km-archaiol	-,197	,028	-,246	-7,099	,000	-,334	-,292	-,239	,946	1,057
	photos	,014	,002	,200	5,713	,000	,324	,239	,193	,928	1,077
	price	,006	,001	,195	5,550	,000	,301	,232	,187	,921	1,085
5	(Constant)	4,902	,162		30,324	,000					
	reviews	,010	,001	,379	10,944	,000	,460	,426	,367	,939	1,065
	km-archaiol	-,189	,028	-,236	-6,804	,000	-,334	-,281	-,228	,933	1,071
	photos	,013	,002	,188	5,338	,000	,324	,224	,179	,910	1,099
	price	,006	,001	,192	5,498	,000	,301	,230	,185	,921	1,086
	air-conditioning	,329	,133	,086	2,467	,014	,230	,106	,083	,936	1,068
6	(Constant)	4,907	,161		30,500	,000					
	reviews	,010	,001	,385	11,147	,000	,460	,433	,372	,934	1,070
	km-archaiol	-,186	,028	-,232	-6,691	,000	-,334	-,277	-,223	,931	1,074
	photos	,014	,002	,198	5,617	,000	,324	,235	,188	,898	1,113
	price	,006	,001	,199	5,706	,000	,301	,239	,191	,915	1,093
	air-conditioning	,360	,133	,094	2,701	,007	,230	,116	,090	,928	1,077
	buzzer	-,194	,077	-,086	-2,529	,012	,015	-,108	-,084	,954	1,049
7	(Constant)	4,891	,160		30,520	,000					
	reviews	,010	,001	,392	11,354	,000	,460	,440	,377	,929	1,076
	km-archaiol	-,186	,028	-,232	-6,738	,000	-,334	-,279	-,224	,931	1,074
	photos	,013	,002	,191	5,426	,000	,324	,228	,180	,893	1,120
	price	,006	,001	,200	5,766	,000	,301	,241	,192	,915	1,093
	air-conditioning	,358	,132	,093	2,701	,007	,230	,116	,090	,928	1,077
	buzzer	-,195	,076	-,087	-2,560	,011	,015	-,110	-,085	,954	1,049
	priv-entrance	,357	,145	,082	2,468	,014	,068	,106	,082	,989	1,011

Πίνακας 6.15: Coefficients.

- Ομοσκεδαστικότητα: Πρέπει να υπάρχει σταθερή διακύμανση καταλοίπων. Ο έλεγχος θα γίνει μέσω του διαγράμματος διασποράς. Παρατηρούμε ότι δεν εμφανίζεται κάποιο γραμμικό πρότυπο και τα σημεία στο διάγραμμα είναι κατανεμημένα γύρω από το μηδέν άρα μπορούμε να ισχυριστούμε ότι πληρούμε και αυτή την προϋπόθεση.



Εικόνα 6.12:Διάγραμμα διασποράς .

6.6 Παρουσίαση αποτελεσμάτων.

Αφού έγινε ο έλεγχος των προϋποθέσεων, μπορέσαμε και εφαρμόσαμε τη μέθοδο της βηματικής παλινδρόμησης ώστε να δούμε ποιες μεταβλητές είναι στατιστικά σημαντικές προς την εξαρτημένη μεταβλητή μας. Στη συνέχεια θα γίνει μια σύντομη περιγραφή της διαδικασίας διότι η πλήρης ανάλυση έγινε στο προηγούμενο κεφάλαιο. Από το πίνακα Variables Entered/Removed βλέπουμε ποιες μεταβλητές εισχωρούν στο μοντέλο μας σε κάθε βήμα. Ο συγκεκριμένος πίνακας παρουσιάζεται στη συνέχεια.

Variables Entered/Removed[a]

Model	Variables Entered	Variables Removed	Method
1	reviews	.	Stepwise (Criteria: Probability-of-F-to-enter <= ,050, Probability-of-F-to-remove >= ,100).
2	km-archaiol	.	Stepwise (Criteria: Probability-of-F-to-enter <= ,050, Probability-of-F-to-remove >= ,100).
3	photos	.	Stepwise (Criteria: Probability-of-F-to-enter <= ,050, Probability-of-F-to-remove >= ,100).
4	price	.	Stepwise (Criteria: Probability-of-F-to-enter <= ,050, Probability-of-F-to-remove >= ,100).
5	air-conditioning	.	Stepwise (Criteria: Probability-of-F-to-enter <= ,050, Probability-of-F-to-remove >= ,100).
6	buzzer	.	Stepwise (Criteria: Probability-of-F-to-enter <= ,050, Probability-of-F-to-remove >= ,100).
7	priv-entrance	.	Stepwise (Criteria: Probability-of-F-to-enter <= ,050, Probability-of-F-to-remove >= ,100).

a. Dependent Variable: incomeb

Πίνακας 6.16: Variables Entered/Removed[a].

Παρατηρούμε ότι η πρώτη ανεξάρτητη μεταβλητή που μπαίνει στο μοντέλο είναι η reviews διότι έχει τη μεγαλύτερη συσχέτιση με την εξαρτημένη μεταβλητή με R=0,46 και στατιστική σημαντικότητα sig=0,00, όπως θα δούμε στη συνέχεια από το πίνακα correlations. Επίσης όλες οι μεταβλητές που εισχώρησαν στο μοντέλο μας είναι στατιστικά σημαντικές με $p<0,05$.

Correlations

	incomeb
Listings	,073
min-nights	-,080
Price	,301
Reviews	,460
km-plaka	-,321
km-omonoia	-,190
km-archaiol	-,334
Wifi	,069
Parking	,052
Pool	,127
Pets	-,105
Smoking	-,179
Wheelchair	,114
cable tv	,065
Gym	,026
Jacuzzi	,029
Priv-entrance	,068
Fireplace	,016
Elevator	,090
Baby-bed	,066
Laptop-workspace	,058
Kitchen	,021
Dryer	,034
Party	-,102
Buzzer	,015
Heating	,056
Air-conditioning	,230
Family	,086
Cancel	,183
Photos	,324
Guests	,221

Bedrooms	,073
Beds	,102
Guests in bedroom	,118
Type	-,238
Neighborhood	,058

Πίνακας 6.17: Correlations.

Για την επιλογή της δεύτερης ανεξάρτητης μεταβλητής που θα εισχωρήσει στο μοντέλο θα κοιτάξουμε το πίνακα Excluded Variables και συγκεκριμένα τη στήλη Partial Correlation του μοντέλου 1 ο οποίος μας δείχνει τις μερικές συσχετίσεις των ανεξάρτητων μεταβλητών σε σχέση με την εξαρτημένη μεταβλητή και λαμβάνει υπόψη τη συσχέτιση της μεταβλητής που έχει ήδη μπει στο μοντέλο. Παρακάτω φαίνεται ο πίνακας.

Excluded Variables

Model 1	sig	Partial Correlation
Listings	,095	,072
min-nights	,150	-,062
Price	,000	,319
km-plaka	,000	-,311
km-omonoia	,000	-,194
km-archaiol	,000	-,328
Wifi	,270	,047
Parking	,446	,033
Pool	,001	,145
Pets	,041	-,088
Smoking	,001	-,147
Wheelchair	,040	,088
cable tv	,972	-,002
Gym	,885	-,006
Jacuzzi	,926	,004
Priv-entrance	,011	,108
Fireplace	,641	,020
Elevator	,062	,080

Baby-bed	,074	,077
Laptop-workspace	,158	,061
Kitchen	,396	,036
Dryer	,802	-,011
Party	,016	-,103
Buzzer	,369	-,039
Heating	,121	,066
Air-conditioning	,000	,177
Family	,015	,104
Cancel	,014	,105
Photos	,000	,268
Guests	,000	,196
Bedrooms	,030	,093
Beds	,031	,093
Guests in bedroom	,004	,124
Type	,000	-,218
Neighborhood	,080	,075

Πίνακας 6.18: Excluded Variables.

Βλέπουμε ότι η δεύτερη μεταβλητή που εισχωρεί στο μοντέλο είναι η km-archaiol και έχει τη μεγαλύτερη συσχέτιση με τιμή -0,328 και στατιστική σημαντικότητα sig=0,00

Για την επιλογή της τρίτης ανεξάρτητης μεταβλητής θα γίνει έλεγχος όπως και πριν από το πίνακα Excluded Variables κοιτάζοντας το μοντέλο 2. Ο πίνακας φαίνεται παρακάτω

Excluded Variables

Model 2	sig	Partial Correlation
Listings	,486	,030
min-nights	,592	-,023

Price	,000	,270
km-plaka	,889	-,006
km-omonoia	,561	-,025
Wifi	,813	-,010
Parking	,070	,078
Pool	,002	,136
Pets	,014	-,105
Smoking	,001	-,140
Wheelchair	,017	,102
cable tv	,887	,006
Gym	,857	-,008
Jacuzzi	,691	,017
Priv-entrance	,006	,117
Fireplace	,596	,023
Elevator	,264	,048
Baby-bed	,050	,084
Laptop-workspace	,104	,070
Kitchen	,627	,021
Dryer	,637	-,020
Party	,037	-,089
Buzzer	,445	-,033
Heating	,320	,043
Air-conditioning	,001	,145
Family	,073	,077
Cancel	,009	,112
Photos	,000	,275
Guests	,000	,201
Bedrooms	,020	,100
Beds	,005	,121
Guests in bedroom	,002	,133
Type	,000	-,205
Neighborhood	,389	-,037

Πίνακας 6.19: Excluded Variables.

Βλέπουμε ότι η μεταβλητή photos έχει τη μεγαλύτερη συσχέτιση που ισούται με 0,275 και με στατιστική σημαντικότητα sig=0,00.

Για την επιλογή της τέταρτης ανεξάρτητης μεταβλητής θα γίνει έλεγχος από το πίνακα Excluded Variables κοιτάζοντας το μοντέλο 3. Ο πίνακας φαίνεται παρακάτω

Excluded Variables

Model 3	sig	Partial Correlation
Listings	,656	,019
min-nights	,828	-,009
Price	,000	,232
km-plaka	,454	,032
km-omonoia	,658	-,019
Wifi	,369	-,039
Parking	,103	,070
Pool	,004	,125
Pets	,039	-,089
Smoking	,013	-,107
Wheelchair	,107	,069
cable tv	,825	-,010
Gym	,738	-,014
Jacuzzi	,838	-,009
Priv-entrance	,022	,098
Fireplace	,808	-,010
Elevator	,310	,044
Baby-bed	,289	,046
Laptop-workspace	,708	,016
Kitchen	,741	,014
Dryer	,369	-,039
Party	,052	-,084

Buzzer	,077	-,076
Heating	,699	,017
Air-conditioning	,011	,110
Family	,112	,068
Cancel	,174	,058
Guests	,002	,130
Bedrooms	,131	,065
Beds	,100	,071
Guests in bedroom	,030	,093
Type	,001	-,145
Neighborhood	,594	-,023

Πίνακας 6.20: Excluded Variables.

Παρατηρούμε ότι η μεταβλητή price έχει τη μεγαλύτερη συσχέτιση που ισούται με 0,232 και με στατιστική σημαντικότητα sig=0,00.

Για την επιλογή της πέμπτης ανεξάρτητης μεταβλητής θα γίνει έλεγχος από το πίνακα Excluded Variables κοιτάζοντας το μοντέλο 4. Ο πίνακας φαίνεται παρακάτω.

Excluded Variables

Model 4	sig	Partial Correlation
Listings	,190	-,056
min-nights	,615	-,022
km-plaka	,642	,020
km-omonoia	,850	-,008
Wifi	,354	-,040
Parking	,494	,029
Pool	,085	,074
Pets	,063	-,080
Smoking	,092	-,072
Wheelchair	,206	,054
cable tv	,395	-,037

Gym	,848	-,008
Jacuzzi	,589	-,023
Priv-entrance	,015	,104
Fireplace	,431	-,034
Elevator	,505	,029
Baby-bed	,514	,028
Laptop-workspace	,689	-,017
Kitchen	,785	,012
Dryer	,086	-,074
Party	,082	-,075
Buzzer	,023	-,098
Heating	,547	,026
Air-conditioning	,014	,106
Family	,204	,055
Cancel	,451	,032
Guests	,497	,029
Bedrooms	,588	-,023
Beds	,444	-,033
Guests in bedroom	,952	-,003
Type	-,105	,886
Neighborhood	-,050	,878

Πίνακας 6.21: Excluded Variables.

Παρατηρούμε ότι η μεταβλητή air conditioning έχει τη μεγαλύτερη συσχέτιση που ισούται με 0,106 και με στατιστική σημαντικότητα sig=0.014

Για την επιλογή της έκτης ανεξάρτητης μεταβλητής θα γίνει έλεγχος από το πίνακα Excluded Variables κοιτάζοντας το μοντέλο 5. Ο πίνακας φαίνεται παρακάτω

Excluded Variables

Model 5	sig	Partial Correlation
Listings	,153	-,062
min-nights	,650	-,020
km-plaka	,697	,017
km-omonoia	,983	-,001
Wifi	,292	-,045
Parking	,477	,031
Pool	,093	,072
Pets	,074	-,077
Smoking	,160	-,060
Wheelchair	,235	,051
cable tv	,347	-,041
Gym	,887	-,006
Jacuzzi	,582	-,024
Priv-entrance	,015	,104
Fireplace	,415	-,035
Elevator	,754	,013
Baby-bed	,553	,026
Laptop-workspace	,489	-,030
Kitchen	,984	,001
Dryer	,081	-,075
Party	,095	-,072
Buzzer	,012	-,108
Heating	,678	,018
Family	,391	,037
Cancel	,573	,024
Guests	,651	,019
Bedrooms	,623	-,021
Beds	,410	-,036
Guests in bedroom	,938	-,003
Type	,053	-,083
Neighborhood	,218	-,053

Πίνακας 6.22: Excluded Variables.

Βλέπουμε ότι η μεταβλητή buzzer έχει τη μεγαλύτερη συσχέτιση που ισούται με -0,108 και με στατιστική σημαντικότητα sig=0,012.

Για την επιλογή της έβδομης ανεξάρτητης μεταβλητής θα γίνει έλεγχος από το πίνακα Excluded Variables κοιτάζοντας το μοντέλο 6. Ο πίνακας φαίνεται παρακάτω

Excluded Variables

Model 6	sig	Partial Correlation
Listings	,129	-,065
min-nights	,700	-,017
km-plaka	,910	,005
km-omonoia	,981	,001
Wifi	,371	-,039
Parking	,305	,044
Pool	,097	,072
Pets	,076	-,076
Smoking	,163	-,060
Wheelchair	,115	,068
cable tv	,636	-,020
Gym	,883	-,006
Jacuzzi	,675	-,018
Priv-entrance	,014	,106
Fireplace	,367	-,039
Elevator	,537	,027
Baby-bed	,539	,026
Laptop-workspace	,917	-,005

Kitchen	,848	,008
Dryer	,093	-,072
Party	,086	-,074
Heating	,509	,028
Family	,314	,043
Cancel	,526	,027
Guests	,669	,018
Bedrooms	,687	-,017
Beds	,378	-,038
Guests in bedroom	,998	,000
Type	,044	-,087
Neighborhood	,223	-,053

Πίνακας 6.23: Excluded Variables.

Παρατηρούμε ότι η μεταβλητή private entrance έχει τη μεγαλύτερη συσχέτιση που ισούται με 0,106 και η στατιστική σημαντικότητα είναι sig=0.014

Για την επιλογή της όγδοης ανεξάρτητης μεταβλητής θα γίνει έλεγχος από το πίνακα Excluded Variables κοιτάζοντας το μοντέλο 7. Ο πίνακας φαίνεται παρακάτω

Excluded Variables

Model 7	sig	Partial Correlation
Listings	,144	-,063
min-nights	,716	-,016
km-plaka	,885	,006
km-omonoia	,971	,002
Wifi	,336	-,041
Parking	,441	,033
Pool	,106	,070

Pets	,096	-,072
Smoking	,190	-,057
Wheelchair	,148	,062
cable tv	,624	-,021
Gym	,844	-,008
Jacuzzi	,808	-,011
Fireplace	,392	-,037
Elevator	,403	,036
Baby-bed	,802	,011
Laptop-workspace	,643	-,020
Kitchen	,890	,006
Dryer	,086	-,074
Party	,077	-,076
Heating	,613	,022
Family	,334	,042
Cancel	,578	,024
Guests	,654	,019
Bedrooms	,727	-,015
Beds	,342	-,041
Guests in bedroom	,975	,001
Type	,065	-,079
Neighborhood	,251	-,050

Πίνακας 6.24: Excluded Variables.

Παρατηρούμε ότι η μεταβλητή type έχει τη μεγαλύτερη συσχέτιση που ισούται με -0,079 όμως η στατιστική σημαντικότητα sig=0.065 > 0.05 άρα δεν δεχόμαστε την είσοδο άλλων μεταβλητών στο μοντέλο.

Έπειτα βλέπουμε το πίνακα model summary.

Model Summary[h]

Model	R	R Square	Adjusted R Square	Std. Error of the Estimate	Change Statistics					Durbin-Watson
					R Square Change	F Change	df1	df2	Sig. F Change	
1	,460[a]	,212	,210	,967429428	,212	146,092	1	544	,000	
2	,545[b]	,297	,294	,914677535	,085	65,557	1	543	,000	
3	,591[c]	,350	,346	,880195192	,053	44,378	1	542	,000	
4	,620[d]	,385	,380	,856946124	,035	30,808	1	541	,000	
5	,626[e]	,392	,386	,852947987	,007	6,084	1	540	,014	
6	,632[f]	,399	,392	,848717698	,007	6,396	1	539	,012	
7	,637[g]	,406	,398	,844738312	,007	6,090	1	538	,014	1,903

a. Predictors: (Constant), reviews

b. Predictors: (Constant), reviews, km-archaiol

c. Predictors: (Constant), reviews, km-archaiol, photos

d. Predictors: (Constant), reviews, km-archaiol, photos, price

e. Predictors: (Constant), reviews, km-archaiol, photos, price, air-conditioning

f. Predictors: (Constant), reviews, km-archaiol, photos, price, air-conditioning, buzzer

g. Predictors: (Constant), reviews, km-archaiol, photos, price, air-conditioning, buzzer, priv-entrance

h. Dependent Variable: incomeb

Πίνακας 6.25: Model Summary.

Παρατηρούμε ότι με την ένταξη των ανεξάρτητων μεταβλητών στο μοντέλο μας ο συντελεστής προσδιορισμού αυξήθηκε. Με την εισαγωγή της πρώτης μεταβλητή η τιμή ήταν 0,212 και μετά την ένταξη της έβδομης έφτασε η τιμή του στο 0,406.

Στην συνέχεια φαίνεται ο πίνακας ANOVA

ANOVA[a]

Model		Sum of Squares	df	Mean Square	F	Sig.
1	Regression	136,730	1	136,730	146,092	,000[b]
	Residual	509,140	544	,936		
	Total	645,871	545			
2	Regression	191,578	2	95,789	114,493	,000[c]
	Residual	454,293	543	,837		
	Total	645,871	545			
3	Regression	225,960	3	75,320	97,219	,000[d]
	Residual	419,911	542	,775		
	Total	645,871	545			
4	Regression	248,584	4	62,146	84,626	,000[e]
	Residual	397,287	541	,734		
	Total	645,871	545			
5	Regression	253,010	5	50,602	69,554	,000[f]
	Residual	392,861	540	,728		
	Total	645,871	545			
6	Regression	257,617	6	42,936	59,607	,000[g]
	Residual	388,253	539	,720		
	Total	645,871	545			
7	Regression	261,963	7	37,423	52,444	,000[h]
	Residual	383,908	538	,714		
	Total	645,871	545			

a. Dependent Variable: incomeb

b. Predictors: (Constant), reviews

c. Predictors: (Constant), reviews, km-archaiol

d. Predictors: (Constant), reviews, km-archaiol, photos

e. Predictors: (Constant), reviews, km-archaiol, photos, price

f. Predictors: (Constant), reviews, km-archaiol, photos, price, air-conditioning

g. Predictors: (Constant), reviews, km-archaiol, photos, price, air-conditioning, buzzer

h. Predictors: (Constant), reviews, km-archaiol, photos, price, air-conditioning, buzzer, priv-entrance

Πίνακας 6.26: ANOVA.

Από τη στήλη sig βλέπουμε τη στατιστική σημαντικότητα των μοντέλων και παρατηρούμε ότι είναι μικρότερη από 0,05 άρα μπορούμε να πούμε ότι είναι στατιστικά σημαντικά. Για το μοντέλο 7 ισχύει ότι είναι το καλύτερο αφού είναι στατιστικά σημαντικό και έχει το μεγαλύτερο συντελεστή προσδιορισμού.

Στη συνέχεια βλέπουμε το πίνακα coefficients όπου θα πάρουμε τους συντελεστές της παλινδρόμησης.

Coefficients

Model 7	B	Std. Error	Beta	t	Sig
(Constant)	4,891	,160		30,520	,000
Reviews	,010	,001	,392	11,354	,000
Km–archaiol	-,186	,028	-,232	-6,738	,000
Photos	,013	,002	,191	5,426	,000
Price	,006	,001	,200	5,766	,000
Air-conditioning	,358	,132	,093	2,701	,007
Buzzer	-,195	,076	-,087	-2,560	,011
Private entrance	,357	,145	,082	2,468	,014

Πίνακας 6.27: Coefficients.

Όπως είχαμε αναφέρει από πίνακα Coefficients και τη στήλη B παίρνουμε το σταθερό όρο και τους συντελεστές παλινδρόμησης κάθε μιας ανεξάρτητης μεταβλητής για το μοντέλο 7 που μας ενδιαφέρει, όπου βλέπουμε την επίδραση τους προς την εξαρτημένη μεταβλητή.

Από τη στήλη Std. Error παίρνουμε τις τιμές του τυπικού σφάλματος της εκτίμησης των συντελεστών B.

Στην στήλη t περιέχονται οι τιμές τις στατιστικής συνάρτησης βάση της οποίας γίνεται ο έλεγχος για την σημαντικότητα ύπαρξης των συντελεστών στην συνάρτηση. Από την τελευταία στήλη του πίνακα, την στήλη sig., δίνεται η τιμή σημαντικότητας βάση της οποίας επιβεβαιώνεται η όχι η διατήρηση της μεταβλητής στο μοντέλο. Τέλος η στήλη Beta μας εμφανίζει την ερμηνευτική ικανότητα κάθε ανεξάρτητης μεταβλητής.

Στη συνέχεια από το πίνακα Coefficients θα δούμε και τα διαστήματα εμπιστοσύνης για κάθε ανεξάρτητη μεταβλητή.

Coefficients

Model 7	B	95,0% Confidence Interval for B	
		Lower Bound	Upper Bound
(Constant)	4,891	4,576	5,205
Reviews	,010	,008	,012
Km–archaiol	-,186	-,240	-,132
Photos	,013	,008	,018
Price	,006	,004	,008
Air-conditioning	,358	,098	,618
Buzzer	-,195	-,345	-,045
Private entrance	,357	,073	,641

Πίνακας 6.28: Coefficients.

Βλέπουμε ότι όλοι οι συντελεστές παλινδρόμησης είναι μέσα στα όρια του διαστήματος εμπιστοσύνης.

6.7 Ερμηνεία τελικού μοντέλου.

Από το πίνακα coefficients του μοντέλου 7 που εμφανίσαμε παραπάνω θα πάρουμε τους συντελεστές παλινδρόμησης κάθε μιας ανεξάρτητης μεταβλητής και το σταθερό όρο για να δημιουργήσουμε την τελική εξίσωση όπως θα δούμε στη συνέχεια.

Η εξίσωση είναι της μορφής:

$$Ln(Y) = a_0 + a_1 * x_1 + a_2 * x_2 + a_3 * x_3 + a_4 * x_4 + a_5 * \chi_5 + \alpha_6 * \chi_6 + \alpha_7 * \chi_7$$

Όπου

a_0=4,891 a_1=0,010 a_2=-0,186 a_3=0,13 a_4=0,006 a_5=0,358

α_6= -0,195 α_7=0,357

Η x_1αναφέρεται στη μεταβλητή Reviews.

Η x_2αναφέρεται στη μεταβλητή Km–archaiol.

Η x_3αναφέρεται στη μεταβλητή Photos.

Η x_4αναφέρεται στη μεταβλητή Price.

Η x_5αναφέρεται στη μεταβλητή Air-conditioning.

Η x_6αναφέρεται στη μεταβλητή Buzzer.

Η x_7αναφέρεται στη μεταβλητή Private entrance.

Άρα η τελική μορφή της εξίσωσης είναι

$$\mathrm{Ln(Y)} = 4.891 + 0.010 * x_1 - 0.186 * x_2 + 0.013 * x_3 + 0.006 * x_4 + 0.358 * x_5 - 0.195 * x_6 + 0.357 * x_7$$

Τώρα θα δούμε πως επηρεάζει κάθε μεταβλητή ξεχωριστά τα έσοδα/ μήνα. Όπως και στο προηγούμενο κεφάλαιο, επειδή η εξίσωση μου έχει λογαριθμική μορφή θα πρέπει να υπολογίζω κάθε φορά το e^{α}-1 ώστε να βρω τη ποσοστιαία μεταβολή των εσόδων/μήνα όταν μια ερμηνευτική μεταβλητή ποικίλει κατά 1.

Για τη μεταβλητή Reviews υπολογίζω το $e^{0.010}$ -1=0.010.Οπότε μπορούμε να πούμε ότι μια μονάδα αύξησης των κριτικών θα επιφέρει αύξηση στα έσοδα/ μήνα κατά 1% σε σχέση με ένα κατάλυμα όπου δεν θα παρουσιάσει τη συγκεκριμένη αύξηση.

Για τη μεταβλητή Km–archaiol υπολογίζω το $e^{-0.186}$ -1= -0,1697. Παρατηρούμε ότι η συγκεκριμένη μεταβλητή έχει αρνητική επίδραση στα έσοδα. Συγκεκριμένα μια μονάδα αύξησης των χιλιομέτρων του καταλύματος από το Μουσείο της Ακρόπολης θα υπάρχει μια μείωση των εσόδων/ μήνα κατά 16,97% σε σχέση με ένα κατάλυμα είναι σε κοντινότερη απόσταση.

Για τη μεταβλητή Photos υπολογίζω το $e^{0.013}$ -1= 0,013. Βλέπουμε ότι η μεταβλητή φωτογραφίες έχει θετική επίδραση στα έσοδα. Ειδικότερα μια μονάδα αύξηση των φωτογραφιών του καταλύματος θα επιφέρει αύξηση κατά 1,3% σε σχέση με ένα κατάλυμα που έχει λιγότερες φωτογραφίες.

Για τη μεταβλητή Price υπολογίζω το $e^{0.006}$ -1= 0,006. Παρατηρούμε ότι η συγκεκριμένη μεταβλητή έχει θετική επίδραση στα έσοδα. Συγκεκριμένα μια μονάδα αύξηση της τιμής θα έχει ως αποτέλεσμα αύξηση κατά 0,6% στα έσοδα σε σχέση με ένα κατάλυμα με χαμηλότερη τιμή.

Για τη μεταβλητή Air-conditioning υπολογίζω το $e^{0.358}$ -1= 0,4304. Ειδικότερα για τα καταλύματα που διαθέτουν κλιματισμό θα υπάρχει μια αύξηση κατά 43,04% των εσόδων σε σχέση με τα καταλύματα που δεν διαθέτουν.

Για τη μεταβλητή Buzzer υπολογίζω το $e^{-0.195}$ -1= -0,1771. Άρα για τα καταλύματα που διαθέτουν κουδούνι/ ασύρματη ενδοεπικοινωνία θα προβλέπεται μια μείωση στα έσοδα κατά 17,71% σε σχέση με τα καταλύματα που δεν διαθέτουν. Πρέπει να αναφέρουμε ότι το συγκεκριμένο αποτέλεσμα δεν είναι λογικό.

Για τη μεταβλητή Private entrance υπολογίζω το $e^{0.357}$ -1= 0,429 . Οπότε για τα καταλύματα που έχουν ιδιωτική είσοδο θα αυξάνονται τα έσοδα/μήνα κατά 42,9% σε σχέση με τα καταλύματα που δεν διαθέτουν.

Θα βγάλουμε τη μέση τιμή των εσόδων/ μήνα για τα καταλύματα όπου δεν έχουν καμία από τις παραπάνω παροχές. Δηλαδή οι τιμές των $x_1, x_2, x_3, x_4, x_5, x_6, x_7$ είναι ίσες με το μηδέν.

Ισχύει ότι:

$$Ln(Y) = 4.891 \Leftrightarrow$$

$$e^{Ln(Y)} = e^{4.891} \Leftrightarrow$$

$$Y = 133.08$$

Συμπεραίνουμε ότι η μέση τιμή των εσόδων/ μήνα των καταχωρήσεων κυμαίνεται στα 133,08 ευρώ.

ΚΕΦΑΛΑΙΟ 7

7. Συμπεράσματα.

Στην έρευνα μας αναλύσαμε τους παράγοντες οι οποίοι συμβάλουν στη διαμόρφωση των εσόδων στα καταλύματα της Airbnb. Εφαρμόσαμε τη μέθοδο της βηματικής παλινδρόμησης και εντοπίσαμε τις ανεξάρτητες μεταβλητές οι οποίες επηρέασαν την εξαρτημένη μεταβλητή μας, έσοδα ανά μήνα. Η μέθοδος εφαρμόστηκε σε δείγμα 603 καταλυμάτων και εξετάστηκαν τελικά 36 ανεξάρτητες μεταβλητές όπου στο τέλος παρουσιάστηκε ότι μόνο 7 είναι σημαντικές ως προς την εξαρτημένη μεταβλητή μας.

Αφού εφαρμόσαμε τη μέθοδο της βηματικής παλινδρόμησης παρατηρήσαμε ότι η μεταβλητή reviews έχει θετική επίδραση κατά 1% στη διαμόρφωση των εσόδων ανά μήνα του καταλύματος. Ωστόσο είδαμε ότι η συγκεκριμένη μεταβλητή έχει το μεγαλύτερο συντελεστή Beta όπου από αυτόν βλέπουμε τη βαρύτητα της μεταβλητής και είναι στατιστικά σημαντική. Αυτή η θετική επίδραση είναι λογική διότι όλοι οι ενδιαφερόμενοι ενοικιαστές νιώθουν περισσότερη ασφάλεια όσο αυξάνεται ο αριθμός των κριτικών.

Επίσης είδαμε ότι η μεταβλητή km- archaiol έχει αρνητική επίδραση κατά 16,97% στη διαμόρφωση των εσόδων ανά μήνα του καταλύματος. Από το συντελεστή Beta βλέπουμε ότι και αυτή η μεταβλητή έχει από τις μεγαλύτερες βαρύτητες στη διαμόρφωση των εσόδων ανά μήνα και είναι στατιστικά σημαντική. Όσο για τον αρνητικό συντελεστή ο οποίος θα προκαλέσει μείωση των εσόδων φαίνεται αναμενόμενο διότι οι περισσότεροι ενοικιαστές που στην πλειοψηφία είναι τουρίστες επιθυμούν να είναι κοντά στους αρχαιολογικούς χώρους.

Ακόμα παρατηρήσαμε ότι η μεταβλητή photos έχει θετική επίδραση κατά 1,3% στη διαμόρφωση των εσόδων ανά μήνα του καταλύματος. Η συγκεκριμένη μεταβλητή έχει και αυτή από τις μεγαλύτερες βαρύτητες στη διαμόρφωση των εσόδων ανά μήνα και είναι στατιστικά σημαντική. Αυτή η θετική επίδραση φαίνεται λογική αφού όσες περισσότερες φωτογραφίες έχει το κατάλυμα τόσο περισσότερη σιγουριά δέχεται ο ενδιαφερόμενος ενοικιαστής με αποτέλεσμα να έχει και περισσότερες κρατήσεις.

Επιπρόσθετα είδαμε ότι η μεταβλητή price έχει θετική επίδραση κατά 0,6% στη διαμόρφωση των εσόδων ανά μήνα του καταλύματος. Από το συντελεστή Beta παρατηρούμε ότι η συγκεκριμένη μεταβλητή έχει και αυτή από τις μεγαλύτερες βαρύτητες στη διαμόρφωση των εσόδων ανά μήνα και είναι στατιστικά σημαντική. Η θετική επίδραση που εμφανίστηκε είναι αναμενόμενη αφού ισχύει ότι όσο υψηλότερη είναι η τιμή του καταλύματος τόσο θα αυξάνονται και τα έσοδα ανά μήνα.

Επιπλέον παρατηρήσαμε ότι η μεταβλητή air conditioning έχει τη μεγαλύτερη θετική επίδραση κατά 43,04% στη διαμόρφωση των εσόδων ανά μήνα του καταλύματος. Η συγκεκριμένη μεταβλητή έχει συντελεστή Beta με τιμή 0,093 και είναι στατιστικά σημαντική. Η συγκεκριμένη παροχή είναι λογικό ότι θα έχει θετική επίδραση στα έσοδα διότι το μεγαλύτερο ποσοστό τουρισμού στην Ελλάδα παρατηρείται το καλοκαίρι και οι ενοικιαστές επιθυμούν να παρέχει το κατάλυμα κλιματισμό. Οπότε αν διαθέτει τη συγκεκριμένη παροχή θα έχει και μεγαλύτερες πιθανότητες για κράτηση με αποτέλεσμα και περισσότερα έσοδα ανά μήνα.

Για τη μεταβλητή buzzer παρατηρήσαμε αρνητική επίδραση κατά 17,71% στη διαμόρφωση των εσόδων ανά μήνα του καταλύματος. Από το συντελεστή Beta παρατηρούμε ότι η συγκεκριμένη μεταβλητή έχει μικρότερη βαρύτητα και από το συντελεστή sig μικρότερο βαθμό σημαντικότητας, στη διαμόρφωση των εσόδων ανά μήνα. Η συγκεκριμένη επίδραση δεν ήταν αναμενόμενη διότι όταν το κατάλυμα διαθέτει μια παροχή η οποία βοηθάει τον φιλοξενούμενο θα είχε σαν αποτέλεσμα θετική επίδραση στα έσοδα.

Τέλος για τη μεταβλητή private entrance είδαμε ότι προκαλεί μεγάλη θετική επίδραση κατά 42,9% στη διαμόρφωση των εσόδων ανά μήνα του καταλύματος. Η συγκεκριμένη μεταβλητή έχει συντελεστή Beta με τιμή 0,082 και μικρότερο βαθμό σημαντικότητας, στη διαμόρφωση των εσόδων ανά μήνα. . Η θετική επίδραση που εμφανίστηκε είναι αναμενόμενη αφού η ιδιωτική είσοδος είναι μια παροχή όπου εξυπηρετεί τον ενοικιαστή, όπου στην επιλογή του καταλύματος θα το λάβει υπόψη του. Άρα η συγκεκριμένη παροχή θα επιφέρει περισσότερες κρατήσεις με αποτέλεσμα και υψηλότερα έσοδα ανά μήνα.

Καταλήγουμε ότι οι παροχές που αναφέραμε εκτός από τη παροχή της ασύρματης ενδοεπικοινωνίας θα επιφέρουν στο κατάλυμα περισσότερα έσοδα ανά μήνα σε σχέση με ένα άλλο κατάλυμα που δεν διαθέτει τις συγκεκριμένες παροχές.

Για να δούμε αν όντως είναι λογικά τα αποτελέσματα που εμφανίσαμε θα τα συγκρίνουμε με τα αποτελέσματα των ερευνών που παρουσιάσαμε σε προηγούμενο κεφάλαιο. Σύμφωνα με τη μελέτη που έγινε για τη πόλη του Nashville και τη μελέτη των (Wang D, Nicolau J.L,2017) για τους καθοριστικούς παράγοντες που συμβάλλουν στη διαμόρφωση της τιμής παρατηρήθηκε αρνητική επίδραση της μεταβλητής των κριτικών στη διαμόρφωση της τιμής όπου εμείς ερχόμαστε σε αντίθεση με το αποτέλεσμα που εμφανίσαμε. Όμως και στις δύο μελέτες οι ερευνητές ανέφεραν ότι τα αποτελέσματα τους έρχονται σε αντίθεση με τη πραγματικότητα, οπότε μπορούμε να δεχθούμε το δικό μας αποτέλεσμα. Η αρνητική επίδραση της μεταβλητής χιλιόμετρα από αρχαιολογικό χώρο (Μουσείο Ακρόπολη) έχει παρόμοιο αποτέλεσμα με των μελετών όπου και σε αυτές παρουσιάζεται αρνητική επίδραση της απόστασης στη τιμή. Για τη θετική επίδραση της μεταβλητής φωτογραφίες δεν μπορούμε να τη συγκρίνουμε με τις προηγούμενες μελέτες διότι δεν έχουν λάβει μέρος τη συγκεκριμένη μεταβλητή, μόνο μπορούμε να κάνουμε μια σύγκριση με τη μεταβλητή που αναφερόταν στο αν διαθέτει ο οικοδεσπότης φωτογραφία προφίλ. Σε σχέση με τη μεταβλητή που αναφέραμε εμφανίζουμε παρόμοιο αποτέλεσμα άρα μπορούμε να δεχθούμε ότι είναι λογικό το αποτέλεσμα μας. Όσο αφορά τις θετικές επιδράσεις των μεταβλητών κλιματισμός και ιδιωτική είσοδος μπορούμε να τις συγκρίνουμε με τις παροχές πραγματικό κρεβάτι και ασύρματο ίντερνετ που μελετήθηκαν στις παραπάνω έρευνες, όπου και οι δύο ομάδες αναφέρονται σε καλύτερη εξυπηρέτηση του ενοικιαστή. Πράγματι παρατηρούμε θετική επίδραση στις παροχές πραγματικό κρεβάτι και ασύρματο ίντερνετ άρα μπορούμε να πούμε ότι είναι λογικά τα αποτελέσματα μας. Τέλος την αρνητική επίδραση της παροχής ασύρματης ενδοεπικοινωνίας δεν μπορούμε να τη συγκρίνουμε με τις προηγούμενες μελέτες διότι δεν έχουν λάβει μέρος τη συγκεκριμένη μεταβλητή, το μόνο που μπορούμε να πούμε είναι ότι ίσως οι καταχωρήσεις που διέθεταν ασύρματη ενδοεπικοινωνία να εμφάνιζαν χαμηλό εισόδημα με αποτέλεσμα και εμάς να μας εμφανιστεί αυτός ο αρνητικός συντελεστής.

Ακόμα θα πρέπει να απαντήσουμε σε κάποιες σημαντικές ερωτήσεις που ανάφερε ο (Σταμπολίδης,2010) και παρουσιάσαμε σε προηγούμενο κεφάλαιο, από τις οποίες καταλήγουμε στο αν το μοντέλο μας είναι ικανοποιητικό. Για την δικιά μας μελέτη μπορούμε να πούμε ότι οι μεταβλητές που έχουν επιλεχθεί για το μοντέλο έχουν κάποιο νόημα σχετικά με το πρόβλημα, το μοντέλο που δημιουργήθηκε είναι χρήσιμο για τον σκοπό για τον οποίο αναπτύχθηκε, οι διαγνωστικοί έλεγχοι είναι ικανοποιητικοί για την καταλληλότητα και την επάρκεια του μοντέλου και τέλος οι συντελεστές παλινδρόμησης στο μοντέλο μας είναι λογικοί εκτός από το συντελεστή της μεταβλητής buzzer. Καταλήγουμε ότι το μοντέλο μας είναι αρκετά ικανοποιητικό.

Βιβλιογραφία.

Ikkala, T., Lampinen, A.(2014) pp. 173–176 Defining the price of hospitality: networked hospitality exchange via Airbnb. In Proceedings of the Companion Publication of the 17th ACM Conference on Computer Supported Cooperative Work & Social Computing, Baltimore, MD, USA.

Gutt, D., Herrmann P. (2017) Sharing Means Caring? Hosts' Price Reaction to Rating Visibility.

Zhang H., Zhang J., Lu S., Cheng S., Zhang J. (2011) Modeling hotel room price with geographically weighted regression.

Zervas G., Proserpio D., Byers J.W. (2017) The rise of the sharing economy: Estimating the impact of Airbnb on the hotel industry.

Dan wang , Juan L. Nicolau. (2017) Price determinants of sharing economy based accommodation rental: A study of listings from 33 cities on Airbnb.com.

Zhihua Zhang , Rachel J. C. Chen , Lee D. Han and Lu Yang (2017) Key Factors Affecting the Price of Airbnb Listings: A Geographically Weighted Approach

Κρασαδάκη Ευαγγελία (2016) Descriptive Statistics with SPSS.

Μαρία Τζαβάρα και Ιωάννα Γιανούτσου (2008) Πτυχιακή Εργασία : Εναλλακτικές μορφές τουρισμού: Θαλάσσιος τουρισμός και η ανάπτυξη του στο νομό Χανίων.

Βασιλείου Λ. Σταμπολίδη (2010) Βηματική ποιοτική παλινδρόμηση: Ανάπτυξη και διερεύνηση εναλλακτικών μοντέλων.

Τσόκα Ιωάννα (2008) Στρατηγική και ιατρικός τουρισμός: η περίπτωση της Ελλάδας.

Montgomery D.C. and Peck E.A. (1992), Introduction to linear regression analysis, , John Wiley and Sons inc, New York.

Grechanovsky E. (1987), Stepwise Regression procedures: overview, problems, results, and suggestions, Annals of New York Academy of Sciences, 491, pp. 197-232.

Jan Van der Borg, NicolaCamatti, Dario Bertocchi,and Andrea Albarea (2017) The Rise of the Sharing Economy in Tourism: Exploring Airbnb.

Daniel Guttentag, Stephen Smith, Luke Potwarka and Mark Havitz (2017) Why Tourists Choose Airbnb: A Motivation-Based Segmentation Study.

Daniel Adams Guttentag (2016) Why tourists choose Airbnb: A motivation-based segmentation study underpinned by innovation concepts.

Δημητριάδης Ιωάννης (2014) Ανάλυση των παραγόντων που συμβάλλουν στη διαμόρφωση της τιμής στα ξενοδοχεία του νομού Χανίων με χρήση ηδονικής μεθόδου.

Montgomery D.C. and Peck E.A. (1992), Introduction to linear regression analysis, , John Wiley and Sons inc, New York.

Salt W.D., Subhash A., Crichton R, and Livingstone J.D. (2007), An improved approximation to the estimation of the critical F values in best subset regression, 43, pp. 143-149.

Lampinen, A.; Cheshire, C. Hosting via Airbnb: Motivations and financial assurances in monetized network hospitality. In Proceedings of the 2016 CHI Conference on Human Factors in Computing Systems, San Jose, CA, USA, 7–12 May 2016; pp. 1669–1680.

Meleo, L.; Romolini, A.; de Marco, M. The Sharing Economy Revolution and Peer-to-peer Online Platforms. The Case of Airbnb. In Proceedings of the International Conference on Exploring Services Science, Bucharest, Romania, 25–27 May 2016; Volume 247, pp. 561–570.

Fotheringham, A.S.; Charlton, M.E.; Brunsdon, C. Geographically weighted regression: A natural evolution of the expansion method for spatial data analysis. *Environ. Plan.* 1998, *30*, 1905–1927.

Σύνδεσμοι.

https://press.atairbnb.com/about-us/

www.insideairbnb.com

http://www.naftemporiki.gr

https://en.wikipedia.org/wiki/F-test

https://el.wikipedia.org/wiki/%CE%91%CF%80%CE%BB%CE%AE_%CE%B3%CF%81%CE%B1%CE%BC%CE%BC%CE%B9%CE%BA%CE%AE_%CF%80%CE%B1%CE%BB%CE%B9%CE%BD%CE%B4%CF%81%CF%8C%CE%BC%CE%B7%CF%83%CE%B7

https://www.materials.uoc.gr/el/undergrad/courses/ETY204/MET_new.pdf

http://mathbits.com/MathBits/TISection/Statistics2/correlation.htm

http://mathworld.wolfram.com/Covariance.html

http://www.statisticshowto.com/what-is-a-coefficient-of-determination/#COD

https://en.wikipedia.org/wiki/Multicollinearity

https://www.youtube.com/watch?v=Ii5iNAtSQtI

https://www.youtube.com/watch?v=alvj86XnEbY

http://www.unipi.gr/faculty/mkoutras/regres/regres1_1.pdf

http://www2.stat-athens.aueb.gr/~jpan/grammika-montela/chapter-III-1.pdf

https://en.wikibooks.org/wiki/Statistics/Numerical_Methods/Quantile_Regression

http://data.library.virginia.edu/getting-started-with-quantile-regression/