

**ΣΧΟΛΗ ΟΙΚΟΝΟΜΙΚΩΝ
ΕΠΙΣΤΗΜΩΝ ΚΑΙ ΔΙΟΙΚΗΣΗΣ**



**ΑΝΟΙΚΤΟ
ΠΑΝΕΠΙΣΤΗΜΙΟ
ΚΥΠΡΟΥ**

**ΜΕΤΑΠΤΥΧΙΑΚΟ ΠΡΟΓΡΑΜΜΑ ΣΠΟΥΔΩΝ
«ΔΙΟΙΚΗΣΗ ΕΠΙΧΕΙΡΗΣΕΩΝ»**

ΔΙΑΤΡΙΒΗ ΕΠΙΠΕΔΟΥ ΜΑΣΤΕΡ

**ΠΑΡΑΓΟΝΤΕΣ ΠΟΥ ΕΠΗΡΕΑΖΟΥΝ ΤΟ
ΕΠΙΤΥΧΗΜΕΝΟ ΗΛΕΚΤΡΟΝΙΚΟ ΕΜΠΟΡΙΟ**

ΧΡΙΣΤΙΑΝΑ ΛΑΖΟΥ

**ΟΝΟΜΑ ΕΠΙΒΛΕΠΟΝΤΑ ΚΑΘΗΓΗΤΗ
ΓΙΩΡΓΟΣ ΧΑΤΖΗΝΙΚΟΛΑΣ**

ΜΑΪΟΣ 2013

Περιεχόμενα

Παράγοντες που επηρεάζουν το επιτυχημένο ηλεκτρονικό εμπόριο (Περίληψη)....	7
Factors influencing the electronic commerce (Summary)	9
1. Εισαγωγή	11
1.1 Προσδιορισμός του προβλήματος.....	11
1.2 Προέλευση των δεδομένων	12
1.3 Στόχος μελέτης.....	12
1.4 Εργαλεία που θα χρησιμοποιηθούν.....	13
1.5 Αναμενόμενα αποτελέσματα	13
2. Ηλεκτρονικό εμπόριο – Θεωρητική προσέγγιση	14
2.1 Συμπεριφορά του καταναλωτή – Η εμπειρία του στο διαδίκτυο.....	14
2.2 Τι οδηγεί τους καταναλωτές στις αγορές μέσω διαδικτύου. Μια θεωρητική προσέγγιση.....	16
2.3 Έρευνα στους παράγοντες που επηρεάζουν τον καταναλωτή. Αγοραστική συμπεριφορά στο διαδίκτυο.....	19
3. Συλλογή, ανάλυση και επιλογή δεδομένων	22
3.1 Προέλευση βάσης δεδομένων.....	22
3.2 Ανάλυση χαρακτηριστικών	23
3.2.1 Αριθμός εγγραφών κατά κατηγορία.....	23
3.2.2 Αριθμός εγγραφών κατά κατάσταση προϊόντος.....	24
3.2.3 Αριθμός εγγραφών κατά τύπο πωλητή.....	25
3.2.4 Αριθμός εγγραφών των προϊόντων που πωλήθηκαν	25
3.2.5 Αριθμός εγγραφών με βάση τον τρόπο πώλησης	26
3.3 Επιλογή κατηγοριών για ανάλυση	27
3.3.1 Collectables.....	27
3.3.2 Coins.....	28

3.3.3	Sports Memorabilia	28
3.3.4	Stamps	28
3.3.5	Computers.....	29
3.3.6	Antiques.....	29
3.3.7	Games and Toys.....	29
3.3.8	Clothes	30
3.3.9	Dvds	30
3.3.10	Books.....	30
3.3.11	Watches	31
3.3.12	Video Games.....	31
3.3.13	Magazines	31
3.3.14	Dolls	32
3.4	Επιλογή τελικών χαρακτηριστικών.....	33
4.	Ανάλυση δεδομένων	34
4.1	Έλεγχος υποθέσεων.....	34
4.1.1	Κατηγορία Collectables.....	34
4.1.1.1	Σχέση πώλησης με φυσική κατάσταση του προϊόντος	34
4.1.1.2	Σχέση πώλησης με είδος του πωλητή	36
4.1.1.3	Σχέση πώλησης με τρόπο πώλησης	36
4.1.1.4	Σχέση πώλησης με κόστος αποστολής προϊόντος	37
4.1.1.5	Ανάλυση διασποράς κόστους αποστολής	41
4.1.2	Κατηγορία Coins.....	43
4.1.2.1	Σχέση πώλησης με είδος του πωλητή	43
4.1.2.2	Σχέση πώλησης με τρόπο πώλησης	44
4.1.2.3	Σχέση πώλησης με κόστος αποστολής προϊόντος	45
4.1.2.4	Ανάλυση διασποράς κόστους αποστολής	48
4.1.3	Κατηγορία Antiques.....	50

4.1.3.1	Σχέση πώλησης με είδος του πωλητή	50
4.1.3.2	Σχέση πώλησης με τρόπο πώλησης	51
4.1.3.3	Σχέση πώλησης με κόστος αποστολής προϊόντος	52
4.1.3.4	Ανάλυση διασποράς κόστους αποστολής	55
4.1.4	Κατηγορία Books.....	57
4.1.4.1	Σχέση πώλησης με φυσική κατάσταση του προϊόντος	57
4.1.4.2	Σχέση πώλησης με είδος του πωλητή	58
4.1.4.3	Σχέση πώλησης με τρόπο πώλησης	59
4.1.4.4	Σχέση πώλησης με κόστος αποστολής προϊόντος	60
4.1.5	Κατηγορία Sports Memorabilia	62
4.1.5.1	Σχέση πώλησης με φυσική κατάσταση του προϊόντος	62
4.1.5.2	Σχέση πώλησης με είδος του πωλητή	63
4.1.5.3	Σχέση πώλησης με τρόπο πώλησης	63
4.1.5.4	Σχέση πώλησης με κόστος αποστολής προϊόντος	64
4.1.5.5	Ανάλυση διασποράς κόστους αποστολής	67
4.1.6	Κατηγορία Stamps	69
4.1.6.1	Σχέση πώλησης με είδος του πωλητή	69
4.1.6.2	Σχέση πώλησης με τρόπο πώλησης	70
4.1.6.3	Σχέση πώλησης με κόστος αποστολής προϊόντος	71
4.1.6.4	Ανάλυση διασποράς κόστους αποστολής	73
4.1.7	Κατηγορία Computers.....	76
4.1.7.1	Σχέση πώλησης με φυσική κατάσταση του προϊόντος	76
4.1.7.2	Σχέση πώλησης με είδος του πωλητή	76
4.1.7.3	Σχέση πώλησης με τρόπο πώλησης	77
4.1.7.4	Σχέση πώλησης με κόστος αποστολής προϊόντος	78
4.1.7.5	Ανάλυση διασποράς κόστους αποστολής	82
4.1.8	Κατηγορία Games-Toys	84

4.1.8.1	Σχέση πώλησης με φυσική κατάσταση του προϊόντος	84
4.1.8.2	Σχέση πώλησης με είδος του πωλητή	85
4.1.8.3	Σχέση πώλησης με τρόπο πώλησης	85
4.1.8.4	Σχέση πώλησης με κόστος αποστολής προϊόντος	86
4.1.8.5	Ανάλυση διασποράς κόστους αποστολής	89
4.1.9	Κατηγορία Clothes	91
4.1.9.1	Σχέση πώλησης με φυσική κατάσταση του προϊόντος	91
4.1.9.2	Σχέση πώλησης με είδος του πωλητή	92
4.1.9.3	Σχέση πώλησης με τρόπο πώλησης	93
4.1.9.4	Σχέση πώλησης με κόστος αποστολής προϊόντος	93
4.1.9.5	Ανάλυση διασποράς κόστους αποστολής	96
4.1.10	Κατηγορία DVDs	98
4.1.10.1	Σχέση πώλησης με φυσική κατάσταση του προϊόντος	98
4.1.10.2	Σχέση πώλησης με είδος του πωλητή	99
4.1.10.3	Σχέση πώλησης με τρόπο πώλησης	100
4.1.10.4	Σχέση πώλησης με κόστος αποστολής προϊόντος	100
4.2	Binary Logistic Regression	103
4.2.1	Κατηγορία Collectables	104
4.2.2	Κατηγορία Coins	105
4.2.3	Κατηγορία Antiques	107
4.2.4	Κατηγορία Books	109
4.2.5	Κατηγορία Sports Memorabilia	110
4.2.6	Κατηγορία Stamps	111
4.2.7	Κατηγορία Computers	113
4.2.8	Κατηγορία Games-Toys	115
4.2.9	Κατηγορία Clothes	117
4.2.10	Κατηγορία DVDs	118

5. Συμπεράσματα – Μελλοντική εργασία	120
5.1 Συμπεράσματα	120
5.1.1 Κατηγορία Collectables.....	120
5.1.2 Κατηγορία Coins.....	121
5.1.3 Κατηγορία Antiques.....	121
5.1.4 Κατηγορία Books.....	123
5.1.5 Κατηγορία Sports Memorabilia	123
5.1.6 Κατηγορία Stamps	123
5.1.7 Κατηγορία Computers.....	124
5.1.8 Κατηγορία Games – Toys	125
5.1.9 Κατηγορία Clothes.....	126
5.1.10 Κατηγορία DVDs.....	126
5.2 Μελλοντική εργασία	127
Βιβλιογραφία.....	129

Παράγοντες που επηρεάζουν το επιτυχημένο ηλεκτρονικό εμπόριο (Περίληψη)

Το ηλεκτρονικό εμπόριο στις μέρες μας φαίνεται να κερδίζει έδαφος έναντι των παραδοσιακών καταστημάτων. Οι λόγοι ποικίλουν και εξαρτώνται από το είδος των προϊόντων και τις κατηγορίες του αγοραστικού κοινού. Κάποια από τα πλεονεκτήματα που είναι κοινώς αποδεκτά είναι ότι ένα ηλεκτρονικό κατάστημα μπορεί να λειτουργεί χωρίς επιπλέον κόστος 24 ώρες την ημέρα, 7 ημέρες την εβδομάδα μειώνοντας έτσι τα λειτουργικά έξοδα μιας επιχείρησης. Κάποιοι άλλοι λόγοι που πιθανόν να παίζουν ρόλο, θετικά ή αρνητικά, στην ανάπτυξη του ηλεκτρονικού εμπορίου αναλύονται σε αυτή τη μελέτη.

Για το σκοπό της έρευνας χρησιμοποιήθηκαν δεδομένα από το ebay.co.uk. Τα δεδομένα προέρχονταν από πωλήσεις που έγιναν σε διάρκεια 90 ημερών με πωλητές από Κύπρο. Τα δεδομένα χωρίστηκαν σε κατηγορίες ανάλογα με το είδος τους.

Στη συνέχεια έγινε η επιλογή των κατηγοριών που είχαν μεγαλύτερη σημαντικότητα με σκοπό την ανάλυση τους και την εξαγωγή συμπερασμάτων ως προς τη συμπεριφορά του καταναλωτή. Στην ανάλυση, η μεταβλητή που εξετάστηκε ήταν ποιες παράμετροι και πως παίζουν ρόλο στην αγορά ή όχι ενός προϊόντος. Για το σκοπό αυτό χρησιμοποιήθηκαν οι μεταβλητές της κατάστασης του προϊόντος, το είδος του πωλητή, το είδος της πώλησης, η τιμή και το κόστος αποστολής (βλέπε κεφάλαιο 3).

Η ανάλυση χωρίστηκε σε δύο μέρη, τον έλεγχο υποθέσεων και την binary logistic regression. Και στις 2 περιπτώσεις η ανάλυση των δεδομένων έγινε ξεχωριστά για κάθε κατηγορία. Στον έλεγχο υποθέσεων εξετάστηκαν τα ακόλουθα:

- Κατά πόσο το ποσοστό των αντικειμένων που πωλούνται είναι μεγαλύτερο όταν είναι καινούρια ή χρησιμοποιημένα (ProductCondition).
- Κατά πόσο το ποσοστό των αντικειμένων που πωλούνται είναι μεγαλύτερο όταν προέρχονται από ιδιώτες ή επιχειρήσεις (SellerType).
- Κατά πόσο το ποσοστό των αντικειμένων που πωλούνται είναι μεγαλύτερο όταν η πώληση γίνεται μέσω δημοπρασίας ή απευθείας πώλησης (BuyMethod).

- Κατά πόσο πωλούνται τα αντικείμενα με χαμηλότερο κόστος αποστολής (ShippingCost).

Στο binary logistic regression χρησιμοποιήθηκαν σαν ανεξάρτητες μεταβλητές η κατάσταση του προϊόντος, το είδος του πωλητή, το είδος της πώλησης, η τιμή και το κόστος αποστολής και σαν εξαρτημένη η πώληση. Οι μεταβλητές κατάσταση του προϊόντος (ProductCondition), το είδος του πωλητή (SellerType), το είδος της πώλησης (BuyMethod) και πώληση (Sold) παίρνουν διακριτές τιμές και έγινε η αντίστοιχη κωδικοποίηση σε 0 και 1. Οι μεταβλητές της τιμής και του κόστους αποστολής είναι συνεχής και αφορούν αξία. Δημιουργήθηκε το μοντέλο κάθε κατηγορίας και όπου ήταν εφικτό η εξίσωση παλινδρόμησης όπου μπορεί να γίνει πρόβλεψη για το πότε ένα προϊόν πωλείται πιο εύκολα (βλέπε κεφάλαιο 4).

Τέλος, παρουσιάζονται τα αποτελέσματα της ανάλυσης των δεδομένων και τα συμπεράσματα που εξάγονται μέσα από αυτά. Είναι ξεκάθαρο ότι κάθε κατηγορία είναι ξεχωριστή και οι παράγοντες που επηρεάζουν διαφοροποιούνται αναλόγως. Ένα κοινό στοιχείο που φαίνεται από αυτή τη μελέτη είναι ότι το αγοραστικό κοινό στις περισσότερες κατηγορίες προτιμά να αγοράζει μέσω δημοπρασίας αφού θεωρεί ότι μπορεί να πετύχει καλύτερες τιμές. Παράλληλα, διαφάνηκε μέσα από τη μελέτη ότι ειδικά στις κατηγορίες που αφορούν συλλεκτικά ή αρχαία αντικείμενα ο αγοραστής εμπιστεύεται περισσότερο τις επιχειρήσεις απ' ότι τους ιδιώτες πιθανόν λόγω της αυθεντίας των προϊόντων (βλέπε κεφάλαιο 5).

Factors influencing the electronic commerce (Summary)

The internet revolution has led many companies to the creation of electronic shops and to the growth of electronic commerce. One of the e-commerce advantages is the accessibility that provides. An online shop can be accessed 24/7, which is very important for the consumers. This research, will focus in other factors influencing the electronic commerce.

For this purpose data from ebay.co.uk were used. Data came from sales completed in a period of 90 days and the seller was located in Cyprus. After the creation of the database there was a categorization based on product type.

Some of the categories had a significant amount of data and were used for further analysis. After the categorization, the final categories and the variables which were suitable for analysis were chosen. These selected variables were product condition, seller type, sales method, product price and shipping cost.

Data analysis was based in two analysis methods, hypothesis testing and binary logistic regression. Both methods were applied separately in each category. Looking at hypothesis testing for factors influencing the sale of a product were searched based on the following:

- Whether sold products percentage was higher when the products were new or used (ProductCondition).
- Whether sold products percentage was higher when the seller was business or private (SellerType).
- Whether sold products percentage was higher when the sale was listed as bid or «buy it now» (BuyMethod).
- Whether the consumers prefer buying products with low shipping cost (ShippingCost).

At binary logistic regression product condition, seller type, sale method, price and shipping cost were used as independent variables and sold (yes/no) was used as dependent variable. Product condition, seller type, sale method and sold were binary values so they were transformed into 1/0. Price and shipping cost were continuous variables and they were used as is. After the data preparation, the

binary regression model for each category separately was exported, and found the regression equation that predicts when an item can be sold and what factors are important in order to get sold.

Finally in the last chapter, the results of this research are presented. The results are different in each category and the conclusion is that there are different factors influencing sales in each category. Some of the factors though, are affecting most of the categories. One of these factors is the sales method (Bid/Buy it Now). Based on the research it can be said, that most consumers prefer buying from auction because of the lower prices they might get. Similarly, it can be said that, when it comes to antiques or high valued items, consumers prefer buying from business than private sellers due to the authentication of items.

Εισαγωγή

- 1.1 Προσδιορισμός του προβλήματος
 - 1.2 Προέλευση των δεδομένων
 - 1.3 Στόχος μελέτης
 - 1.4 Εργαλεία που θα χρησιμοποιηθούν
 - 1.5 Αναμενόμενα αποτελέσματα
-

1. Εισαγωγή

1.1 Προσδιορισμός του προβλήματος

Το ηλεκτρονικό εμπόριο έχει αναπτυχθεί στις μέρες μας και καθημερινά κερδίζει έδαφος έναντι των παραδοσιακών εμπορικών καταστημάτων. Οι καταναλωτές πλέον έχουν εύκολη πρόσβαση στο διαδίκτυο και μπορούν έτσι να πάρουν πληροφορίες για τα προϊόντα που τους ενδιαφέρουν, να συγκρίνουν τιμές και να κοιτάζουν για εναλλακτικές λύσεις.

Όλα αυτά σε συνδυασμό με την αύξηση της ασφάλειας των αγορών μέσω διαδικτύου συντελούν στην αύξηση των ηλεκτρονικών καταστημάτων. Τα πλεονεκτήματα ενός ηλεκτρονικού καταστήματος έναντι του παραδοσιακού είναι η μείωση των λειτουργικών εξόδων με αποτέλεσμα να μπορούν να πωλούν σε χαμηλότερες τιμές ή να μεγιστοποιούν το κέρδος, η εικοσιτετράωρη προσβασιμότητα και λειτουργία τους, η εύκολη και γρήγορη πρόσβαση των πελατών σε αυτά.

Η μελέτη αυτή επικεντρώνεται στους πωλητές και στις πωλήσεις που κάνουν μέσω ηλεκτρονικών καταστημάτων. Συγκεκριμένα, η πιο κάτω έρευνα θα χρησιμοποιήσει ως βάση της τον ιστόχωρο www.ebay.co.uk και το δείγμα που θα αναλυθεί αφορά τις πωλήσεις που έγιναν από την Κύπρο σε διάστημα 90 ημερών.

Το πρόβλημα που παρουσιάζεται είναι ποιοι παράγοντες επηρεάζουν την πώληση ενός προϊόντος. Για το σκοπό αυτό θα γίνει συλλογή των δεδομένων για διάφορες κατηγορίες προϊόντων. Οι παράγοντες που θα εξεταστούν θα είναι ο τύπος του πωλητή, η κατάσταση του προϊόντος, η μέθοδος της πώλησης, η τιμή του προϊόντος και το κόστος αποστολής.

1.2 Προέλευση των δεδομένων

Όλα τα δεδομένα προέρχονται από τον ιστόχωρο www.ebay.co.uk και αφορούν πωλήσεις που έγιναν από την Κύπρο το χρονικό διάστημα 27/09/2012 – 20/01/2013. Η συλλογή των δεδομένων αφορά τις πιο κάτω κατηγορίες.

Κατηγορία	Αριθμός εγγραφών
Collectables	1511
Coins	1226
Sports Memorabilia	441
Stamps	431
Computers	404
Antiques	238
Games	235
Clothes	148
Dvds	134
Books	115
Watches	99
Video Games	92
Magazines	68
Dolls	64
Toys	48

Πίνακας 1.1: Αριθμός εγγραφών κατά κατηγορία

1.3 Στόχος μελέτης

Στόχος της μελέτης είναι να καθοριστούν οι παράγοντες που επηρεάζουν τις πωλήσεις μέσω διαδικτύου. Για το σκοπό αυτό θα εξεταστούν οι παράγοντες όπως τύπος του πωλητή, η κατάσταση του προϊόντος, η μέθοδος της πώλησης, η τιμή του προϊόντος και το κόστος αποστολής για κάθε κατηγορία ξεχωριστά.

1.4 Εργαλεία που θα χρησιμοποιηθούν

Για τη δημιουργία της βάσης δεδομένων μετά τη συλλογή τους θα χρησιμοποιηθεί το λογισμικό Microsoft Access, ενώ για την ανάλυση των δεδομένων θα χρησιμοποιηθούν τα λογισμικά Microsoft Excel και Minitab.

1.5 Αναμενόμενα αποτελέσματα

Με την ολοκλήρωση της μελέτης αναμένεται να εξαχθούν συμπεράσματα που να αφορούν τους παράγοντες που επηρεάζουν τις πωλήσεις μέσω διαδικτύου όταν αυτές γίνονται από την Κύπρο, καθώς και τα προβλήματα που μπορεί να αντιμετωπίζουν οι Κύπριοι πωλητές.

Ηλεκτρονικό εμπόριο – Θεωρητική προσέγγιση

- 2.1 Συμπεριφορά του καταναλωτή – Η εμπειρία του στο διαδίκτυο
 - 2.2 Τι οδηγεί τους καταναλωτές στις αγορές μέσω διαδικτύου. Μια θεωρητική προσέγγιση
 - 2.3 Έρευνα στους παράγοντες που επηρεάζουν τον καταναλωτή. Αγοραστική συμπεριφορά στο διαδίκτυο
-

2. Ηλεκτρονικό εμπόριο – Θεωρητική προσέγγιση

Σε αυτό το κεφάλαιο αναφέρονται κάποιες μελέτες οι οποίες προσπαθούν να εντοπίσουν τις κατηγορίες προϊόντων που οι καταναλωτές αναζητούν στο διαδίκτυο, καθώς και τους διάφορους παράγοντες που επηρεάζουν και ωθούν τον καταναλωτή στο ηλεκτρονικό εμπόριο. Στις μελέτες αυτές, αναφέρονται παράλληλα οι δυσκολίες και τα προβλήματα που αντιμετωπίζουν τα ηλεκτρονικά καταστήματα.

2.1 Συμπεριφορά του καταναλωτή – Η εμπειρία του στο διαδίκτυο¹

Η συμπεριφορά του καταναλωτή και οι παράγοντες που τον επηρεάζουν για να αγοράσει ένα προϊόν είναι κάτι που αφορά το μάρκετινγκ όλων των εταιρειών. Η συμπεριφορά του καταναλωτή ως προς τις online αγορές έχει κάποιες ιδιαιτερότητες σε σχέση με τις «παραδοσιακές» αγορές. Γενικότερα ο καταναλωτής για να αγοράσει ένα προϊόν πρέπει πρώτα να εντοπίσει ότι υπάρχει ανάγκη για αγορά του, να συλλέξει πληροφορίες σχετικά με το προϊόν, να αξιολογήσει τις εναλλακτικές του, να πάρει την απόφαση και τέλος να επαναξιολογήσει την επιλογή που έκανε. Σημαντικό ρόλο στη διαδικασία λήψης αποφάσεων του καταναλωτή παίζει το ρίσκο που επιθυμεί να πάρει, το πόσο καινοτόμο άτομο είναι κτλ.

¹ Influencing the online consumer's behavior: the Web experience – Efthimios Constantinides: Internet Research (2004).

Με την πάροδο των χρόνων παρατηρείται ραγδαία αύξηση των αγορών που γίνονται μέσω διαδικτύου και συνεχώς τα ηλεκτρονικά καταστήματα κερδίζουν έδαφος έναντι των παραδοσιακών. Το μάρκετινγκ όσο αφορά τους παράγοντες που μπορούν να ελεγχθούν, όπως είναι η περιγραφή των χαρακτηριστικών των προϊόντων, παραμένει το ίδιο όπως στην προώθηση προϊόντων μέσω παραδοσιακών καταστημάτων.

Εκτός από παράγοντες όπως είναι τα χαρακτηριστικά, το σημαντικότερο ρόλο στην απόφαση του καταναλωτή παίζουν παράγοντες που θα αναλυθούν στη συνέχεια. Οι καταναλωτές θεωρούν πιο δύσκολη τη διαδικασία αγορά προϊόντων μέσω διαδικτύου, ειδικά στις περιπτώσεις που δεν υπάρχει εξοικείωση. Για την αγορά ο χρήστης πρέπει να αναζητήσει και να βρει το προϊόν, να το αξιολογήσει και να συγκρίνει με άλλες μάρκες μόνος του μέσω διαδικτύου. Άρα για την αγορά προϊόντων μέσω διαδικτύου απαιτείται εξοικείωση με αυτό.

Η μελέτη επικεντρώνεται στα στοιχεία που αποτελούν και καθορίζουν την εμπειρία στο διαδίκτυο. Ο αγοραστής που θα προχωρήσει σε αγορά μέσω διαδικτύου θεωρείται εκτός από αγοραστής και χρήστης ηλεκτρονικών υπολογιστών και ηλεκτρονικών συστημάτων. Η εμπειρία στο διαδίκτυο αφορά στοιχεία όπως αναζήτηση, πλοήγηση, αναζήτηση εναλλακτικών και αξιολόγηση των επιλογών. Αυτό έχει ως αποτέλεσμα τη διαδραστικότητα μεταξύ του χρήστη και της εταιρείας που διατηρεί κατάσταση στο διαδίκτυο.

Οι παράγοντες που επηρεάζουν την εμπειρία ενός πιθανού αγοραστή στο διαδίκτυο είναι λειτουργικοί, ψυχολογικοί και παράγοντες που αφορούν το περιεχόμενο. Οι λειτουργικοί παράγοντες έχουν να κάνουν με την ευκολία που μπορεί ένας χρήστης να κινηθεί σε ένα ιστόχωρο, με την ταχύτητα πλοήγησης και γενικότερα με την ευκολία στη χρήση των εργαλείων που προσφέρει. Παράλληλα, η αλληλεπίδραση με το ηλεκτρονικό κατάστημα παίζει σημαντικό ρόλο, όπως για παράδειγμα

η επικοινωνία με το προσωπικό, η μετά αγορά επικοινωνία, η τροποποίηση σε προϊόντα έτσι ώστε να ικανοποιούν τον πελάτη.

Οι ψυχολογικοί παράγοντες παίζουν ρόλο στην καλή ή κακή εμπειρία του αγοραστή με το διαδίκτυο όπως είναι η εμπιστοσύνη στις αγορές που κάνει, η εμπιστοσύνη στο να δίνει τα προσωπικά του στοιχεία και η εγγύηση που δίνει μία επιχείρηση ή η πολιτική για επιστροφή προϊόντων. Καταλυτικό ρόλο στη δημιουργία εμπιστοσύνης για τις αγορές στο διαδίκτυο παίζει αν ο ιστοχώρος που επισκέπτεται ένας πιθανός αγοραστής έχει κάποια σφραγίδα αξιοπιστίας από γνωστούς οργανισμούς.

Τέλος, οι παράγοντες που αφορούν το περιεχόμενο μπορεί να χωριστούν σε δύο επίπεδα. Το πρώτο είναι οι αισθητικοί παράγοντες που έχουν να κάνουν με τον σχεδιασμό, την παρουσίαση και την ποιότητα καθώς και το στυλ και την ατμόσφαιρα που δημιουργείται σε ένα χώρο. Αυτό μπορεί να παραλληλιστεί με ένα παραδοσιακό κατάστημα όπου ο σχεδιασμός και το στυλ του μαγαζιού προδιαθέτουν ή όχι τον αγοραστή στο να παραμείνει και να ψωνίσει. Ο σχεδιασμός και το στυλ ενός ιστόχωρου είναι η βιτρίνα και είναι υπεύθυνο για το αν θα παραμείνει ο χρήστης ή θα συνεχίσει την αναζήτηση του αλλού.

Άλλοι παράγοντες που παίζουν ρόλο είναι η επικοινωνία, η ικανοποίηση, το ίδιο το προϊόν, η τιμή του προϊόντος αλλά και η προώθηση που γίνεται σε αυτό.

2.2 Τι οδηγεί τους καταναλωτές στις αγορές μέσω διαδικτύου. Μια θεωρητική προσέγγιση²

Το ποσοστό των ατόμων που προτιμούν να κάνουν τις αγορές τους μέσω διαδικτύου αυξάνεται συνεχώς δραματικά τόσο στις ΗΠΑ όσο και στην Ευρώπη.

² What drives consumers to shop online? A literature view – Perea y Monsuwe, Bendict G.C Dellaert and Ko de Ruyter: International Journal of service industry management (2004)

Αυτή η μελέτη επικεντρώνεται σε ένα πλαίσιο (framework) για τις προθέσεις των καταναλωτών όσο αφορά τις αγορές στο διαδίκτυο. Οι διαδικτυακοί αγοραστές μπορούν να χωριστούν σε δύο κατηγορίες, στους «επιλυτές προβλημάτων» (problem solvers) και σε αυτούς που ψάχνουν για διασκέδαση, διέγερση της φαντασίας τους και απόλαυση. Η πρώτη ομάδα έχει ως στόχο την αγορά συγκεκριμένων προϊόντων και θέλει να το πετύχει όσο πιο αποδοτικά και γρήγορα γίνεται. Η δεύτερη ομάδα αρέσκεται στην εμπειρία η οποία μπορεί να οδηγήσει ή όχι σε αγορά. Οι παράγοντες όπως καθορίζονται και στο πλαίσιο και επηρεάζουν τόσο σε επίπεδο χρησιμότητας όσο και σε επίπεδο ευχαρίστησης αναφέρονται στη συνέχεια.

Πρώτος παράγοντας είναι η χρησιμότητα μίας νέας τεχνολογίας έτσι ώστε να βοηθήσει στην εκπλήρωση των στόχων του αγοραστή. Η χρησιμότητα συνδέεται άμεσα και επηρεάζεται από την ευκολία χρήσης αφού όσο πιο εύκολη είναι μία τεχνολογία να χρησιμοποιηθεί τόσο πιο χρήσιμη μπορεί να γίνει. Στο πλαίσιο φαίνεται ότι αν επενδύσει κάποιος στην αγορά ενός υπολογιστή και στη εκπαίδευση του ώστε να μπορεί να κάνει αγορές στο διαδίκτυο τότε αυτή η επένδυση μπορεί στο μέλλον να αποδώσει, αφού ο αγοραστής θα επωφελείται των πλεονεκτημάτων που έχουν οι αγορές στο διαδίκτυο.

Η ευκολία χρήσης καθορίζεται όταν η χρήση μίας νέας τεχνολογίας δεν απαιτεί ιδιαίτερη προσπάθεια. Η ευκολία χρήσης αναφέρεται στη διαδικασία κατά την οποία ο χρήστης οδηγείται στην αγορά ενός προϊόντος. Όσο πιο απλή και εύκολη είναι μία τεχνολογία τόσο πιο πολλοί καταναλωτές θα τη χρησιμοποιήσουν.

Άλλοι παράγοντες που επηρεάζουν είναι ο έλεγχος και η ανησυχία για τους υπολογιστές. Ο χρήστης πρέπει να νιώθει ότι έχει τον έλεγχο της διαδικασίας και να αποβάλλει τον φόβο που μπορεί να έχει για τους υπολογιστές.

Παράλληλα, χαρακτηριστικά του ηλεκτρονικού καταστήματος ελκύουν ή απωθούν τους χρήστες. Για παράδειγμα η εμφάνιση του υλικού ενός ηλεκτρονικού καταστήματος, η ταχύτητα με την οποία

μπορεί κάποιος να πλοηγηθεί, τα μέτρα ασφάλειας, η ευχαρίστηση που παίρνει κάποιος από τη χρήση, είναι κάποιιοι από τους παράγοντες που ένας χρήστης λαμβάνει υπόψη κατά τη διαδικασία αγοράς προϊόντων από ένα ηλεκτρονικό κατάστημα.

Την ίδια στιγμή υπάρχουν παράγοντες εξωγενείς που βοηθούν στην ύπαρξη ηλεκτρονικών καταστημάτων. Αυτοί μπορεί να είναι η ηλικία, το φύλο, η μόρφωση, το εισόδημα. Άτομα νεαρής ηλικίας φαίνεται να ψωνίζουν περισσότερο και άτομα με υψηλό μορφωτικό επίπεδο εμπιστεύονται περισσότερο το διαδίκτυο. Επιπλέον, άτομα τα οποία δε μπορούν να πάνε σε ένα παραδοσιακό κατάστημα λόγω χρόνου, χώρου διαμονής, κινητικών προβλημάτων, μπορούν να κάνουν τις αγορές τους όλες τις ώρες της ημέρας.

Εκτός αυτών υπάρχει και η κατηγορία των καταναλωτών που επιθυμούν προϊόντα ιδιαίτερα, προσαρμοσμένα στις ανάγκες τους, ή προϊόντα που πιθανόν δύσκολα να τα βρίσκει κάποιος σε ένα τοπικό παραδοσιακό κατάστημα, ιδιαίτερα αν η γεωγραφική τοποθεσία του καταναλωτή δεν είναι σε μία μεγαλούπολη.

Τα αρνητικά των αγορών μέσω διαδικτύου είναι ότι δεν υπάρχει η επαφή με τον πωλητή, άρα η επικοινωνία και ο τρόπος που δίνεται η βοήθεια πρέπει να είναι ξεκάθαρη και συγκεκριμένη έτσι ώστε να αποφεύγονται παρεξηγήσεις. Ακόμα, ο χρήστης πρέπει να αρκεστεί από την εικόνα και την περιγραφή που βλέπει για να αποφασίσει αν θα προχωρήσει στην αγορά. Αυτό δε μπορεί να γίνει πάντα αφού για την αγορά συγκεκριμένων κατηγοριών προϊόντων απαιτούνται και άλλες αισθήσεις (όσφρηση, αφή). Αυτός είναι και ο λόγος που τα καταστήματα στο διαδίκτυο προσφέρουν ως συνήθως προϊόντα που αφορούν υπολογιστές, βιβλία, DVDs κτλ, ενώ αποφεύγονται κατηγορίες που έχουν να κάνουν με προσωπική υγιεινή, αρώματα κτλ.

Τέλος, η μελέτη αναφέρει ότι σημαντικό ρόλο στην απόφαση του καταναλωτή να προτιμά αγορές μέσω διαδικτύου παίζει η προηγούμενη εμπειρία που είχε με τέτοιες αγορές. Αν ήταν θετική τότε

είναι πιο εύκολο να προχωρήσει σε περισσότερες αγορές, αν όχι τότε γίνεται πιο επιφυλακτικός και προτιμά τα παραδοσιακά καταστήματα.

2.3 Έρευνα στους παράγοντες που επηρεάζουν τον καταναλωτή. Αγοραστική συμπεριφορά στο διαδίκτυο³

Η μελέτη αυτή επικεντρώνεται στους παράγοντες που ωθούν τους αγοραστές να προχωρήσουν σε αγορές μέσω διαδικτύου. Η έρευνα έγινε με ερωτηματολόγια σε δείγμα 80 ατόμων και αφορούσε την κινέζικη αγορά.

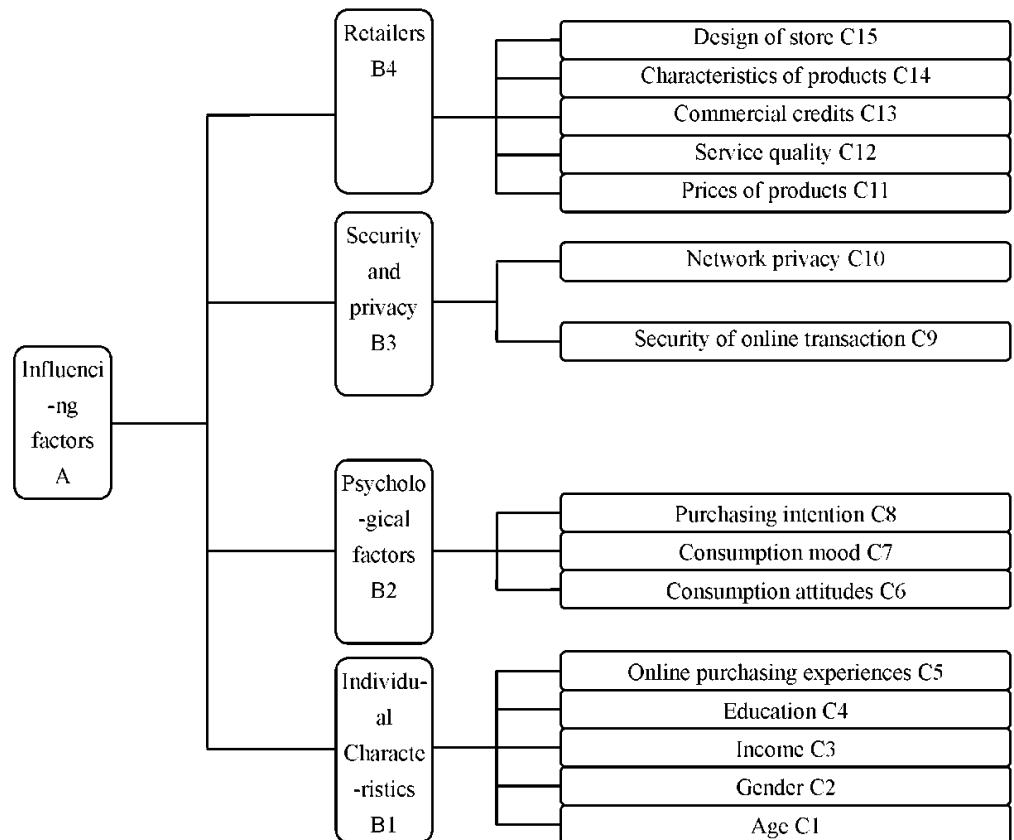
Το πρώτο μέρος της έρευνας αφορούσε πληροφορίες όπως φύλο, ηλικία, εισόδημα, μόρφωση και εμπειρία στις αγορές μέσω διαδικτύου. Το δεύτερο μέρος ήταν να εξηγήσει τους παράγοντες που επηρεάζουν έτσι ώστε να βοηθήσουν τους αγοραστές να τους κατανοήσουν. Το τρίτο μέρος αφορούσε την εξήγηση των παραγόντων τους οποίους κατηγοριοποίησαν και ζητήθηκε από τους αγοραστές να κρίνουν το βαθμό σημαντικότητας του κάθε ενός.

Με βάση τα αποτελέσματα από τα 80 ερωτηματολόγια το ποσοστό των ανδρών ήταν 42% σε αντίθεση με το ποσοστό των γυναικών που ήταν 58%. Να σημειωθεί εδώ ότι δόθηκαν ερωτηματολόγια αποκλειστικά σε άτομα που έχουν κάνει αγορές μέσω διαδικτύου. Η έρευνα έδειξε επίσης ότι το μεγαλύτερο ποσοστό ήταν ηλικία 18-35, με υψηλό επίπεδο μόρφωσης και ότι το 79% είχε κάνει αγορές πάνω από 10 φορές.

Στη συνέχεια ο αναλυτής δημιούργησε το ιεραρχικό μοντέλο με βάση τους παράγοντες που επηρεάζουν τη συμπεριφορά των αγοραστών στις αγορές μέσω διαδικτύου. Για τη δημιουργία του μοντέλου έγινε στατιστική ανάλυση των δεδομένων που συλλέχθηκαν δίνοντας αντίστοιχη βαρύτητα στο καθένα ανάλογα με τη σημαντικότητά τους. Το μοντέλο παρουσιάζεται στη συνέχεια στο Σχήμα 1.

³ A Research on Influencing Factors of Consumer. Purchasing Behaviors in Cyberspace - Li Guo: International Journal of Marketing Studies (2011)

Τα συμπεράσματα της έρευνας έδειξαν ότι οι πιθανοί αγοραστές επηρεάζονται από την ασφάλεια των συναλλαγών και του δικτύου, τις τιμές των προϊόντων, τις υπηρεσίες που παρέχει ένα ηλεκτρονικό κατάστημα και τη σημαντικότητα της αγοράς που θέλουν να κάνουν. Η έρευνα έδειξε επίσης ότι το φύλο, η αισθητική του ηλεκτρονικού καταστήματος και η εκπαίδευση δεν αποτελούν σημαντικούς παράγοντες που επηρεάζουν στις διαδικτυακές αγορές.



Σχήμα 2.1: Ιεραρχικό μοντέλο των παραγόντων που επηρεάζουν τους καταναλωτές στις αγορές μέσω διαδικτύου.

Οι εισηγήσεις που δίνονται στους πωλητές είναι αρχικά όπως δημιουργήσουν μία δυνατή σχέση με τον καταναλωτή. Αυτό μπορεί να γίνει με τη χρήση αξιόπιστων δικτύων, ασφαλών συναλλαγών και

σωστής επικοινωνίας με τον αγοραστή. Παράλληλα, ο πωλητής πρέπει να φροντίσει όπως οι τιμές του παραμείνουν σε ανταγωνιστικά επίπεδα χωρίς όμως να χάνει από την ποιότητα των προϊόντων και των υπηρεσιών του. Σημαντικό παράγοντα παίζουν η γρήγορη διανομή και το χαμηλό κόστος αποστολής. Άρα ο πωλητής πρέπει να βρει τρόπους ώστε να διανέμει τα προϊόντα του γρήγορα και φτηνά. Τέλος, ο πωλητής για να διατηρήσει την αξιοπιστία του πρέπει να παρέχει εγγυήσεις για τα προϊόντα του και να δίνει το δικαίωμα στον αγοραστή να επιστρέψει το προϊόν αν δεν τον ικανοποιεί.

Συλλογή, ανάλυση και επιλογή δεδομένων

- 3.1 Προέλευση βάσης δεδομένων
 - 3.2 Ανάλυση χαρακτηριστικών
 - 3.3 Επιλογή κατηγοριών για ανάλυση
 - 3.4 Επιλογή τελικών χαρακτηριστικών
-

3. Συλλογή, ανάλυση και επιλογή δεδομένων

3.1 Προέλευση βάσης δεδομένων

Για τη συλλογή των δεδομένων επιλέχθηκε το ebay.co.uk. Τα δεδομένα αφορούν προϊόντα που προσφέρθηκαν στην περίοδο 27/9/2012 – 20/1/2013, και έληξε η περίοδος προσφοράς τους. Η συλλογή των δεδομένων αφορά τις πιο κάτω κατηγορίες.

Κατηγορία	Αριθμός εγγραφών
Collectables	1511
Coins	1226
Sports Memorabilia	441
Stamps	431
Computers	404
Antiques	238
Games	235
Clothes	148
DVDs	134
Books	115
Watches	99
Video Games	92
Magazines	68
Dolls	64
Toys	48

Πίνακας 3.1: Αριθμός εγγραφών κατά κατηγορία

3.2 Ανάλυση χαρακτηριστικών

Για κάθε εγγραφή συλλέχτηκαν δεδομένα που αφορούν τα πιο κάτω:

Χαρακτηριστικό	Περιγραφή
SellerType	Τύπος πωλητής. Το πεδίο αυτό μπορεί να πάρει δύο τιμές, Business/Private.
ProductCategory	Η κατηγορία του προϊόντος
ProductCondition	Η κατάσταση του προϊόντος. Το πεδίο αυτό μπορεί να πάρει 2 τιμές, New/Used
Sold	Το πεδίο αυτό παίρνει δύο τιμές Yes/No και αναφέρεται στο κατά πόσο έχει πουληθεί το προϊόν ή όχι
BuyMethod	Η μέθοδος με την οποία έγινε η πώληση. Το πεδίο αυτό μπορεί να πάρει δύο τιμές Bid/Buy it Now
NoOfBids	Ο αριθμός των bids που έγιναν. Το πεδίο αυτό παίρνει τιμές μεγαλύτερες από το 0 μόνο στις περιπτώσεις εκείνες όπου το πεδίο BuyMethod=Bid
Price	Η τιμή που πωλήθηκε το προϊόν. Στις περιπτώσεις όπου δεν πουλήθηκε το προϊόν η τιμή αφορά την αρχική τιμή που έθεσε ο πωλητής
Currency	Το νόμισμα στο οποίο έκανε την προσφορά ο πωλητής
ShippingCost	Το κόστος αποστολής του προϊόντος
EbayCountry	Η χώρα στην οποία εδρεύει το ebay που επέλεξε ο πωλητής να απευθυνθεί, άρα προσφέρει το προϊόν
SellDate	Η ημερομηνία πώλησης του προϊόντος

Πίνακας 3.2: Περιγραφή χαρακτηριστικών

Στη συνέχεια δίνεται μία πρώτη κατηγοριοποίηση των δεδομένων βάση διαφόρων παραγόντων.

3.2.1 Αριθμός εγγραφών κατά κατηγορία

Στον πίνακα 3.3 φαίνονται οι σημαντικότερες κατηγορίες για τις οποίες έγινε καταγραφή των προϊόντων. Υπήρξαν

κατηγορίες οι οποίες φαίνονται στον πίνακα 3.4 όπου ο αριθμός των εγγραφών ήταν μικρός και δε θεωρήθηκαν σημαντικές να υπολογιστούν.

Product Category	Records	%
Collectables	1.511	28,76
Coins	1.226	23,33
Sports Memorabilia	441	8,39
Stamps	431	8,20
Computers	404	7,69
Antiques	238	4,53
Games	235	4,47
Clothes	148	2,82
Dvds	134	2,55
Books	115	2,19
Watches	99	1,88
Video Games	92	1,75
Magazines	68	1,29
Dolls	64	1,22
Toys	48	0,91
	5.254	100,00

Πίνακας 3.3: Αριθμός εγγραφών κατά κατηγορία

Product Category	Records
Mobile Phones	0
Sports Equipment	0
Jewelery	28
Electronics	6

Πίνακας 3.4: Αριθμός εγγραφών λιγότερο σημαντικών κατηγοριών

3.2.2 Αριθμός εγγραφών κατά κατάσταση προϊόντος

Στο ebay υπάρχει δυνατότητα να πουληθούν προϊόντα καινούργια ή χρησιμοποιημένα. Στον πίνακα 3.5 φαίνεται ο αριθμός καινούριων προϊόντων που τέθηκαν προς πώληση και ο αριθμός χρησιμοποιημένων προϊόντων.

Product Condition	Records	%
New	1.389	26,44
Used	3.865	73,56
	5.254	100,00

Πίνακας 3.5: Αριθμός εγγραφών κατά κατάσταση προϊόντος

3.2.3 Αριθμός εγγραφών κατά τύπο πωλητή

Οι πωλητές χωρίζονται σε Business και Private. Η πρώτη κατηγορία είναι πωλητές που εγγράφονται σαν επιχειρήσεις ενώ η δεύτερη αφορά ιδιώτες που πουλούν προϊόντα για προσωπικό τους όφελος. Στον πίνακα 3.6 φαίνεται ο αριθμός των προϊόντων που τέθηκαν προς πώληση από ιδιώτες και επιχειρήσεις.

Seller Type	Records	%
Business	2.758	52,49
Private	2.496	47,51
	5.254	100,00

Πίνακας 3.6: Αριθμός εγγραφών κατά τύπο πωλητή

3.2.4 Αριθμός εγγραφών των προϊόντων που πωλήθηκαν

Τα προϊόντα που οι πωλητές θέτουν προς πώληση μπορεί είτε να πουληθούν είτε όχι. Ο πίνακας 3.7 δείχνει πόσα από τα προϊόντα πουλήθηκαν.

Sold	Records	%
No	2.334	44,42
Yes	2.920	55,58

5.254	100,00
--------------	---------------

Πίνακας 3.7: Αριθμός εγγραφών
πωληθέντων προϊόντων

3.2.5 Αριθμός εγγραφών με βάση τον τρόπο πώλησης

Οι πωλητές μπορούν να χρησιμοποιήσουν δύο μεθόδους για να πουλήσουν τα προϊόντα τους. Η πρώτη μέθοδος είναι αυτή της δημοπρασίας (bid), όπου ο πωλητής θέτει το ελάχιστο όριο της τιμής και διάρκεια χρόνου της δημοπρασίας και πλειοδότης με την υψηλότερη προσφορά κερδίζει το προϊόν. Στην περίπτωση της απευθείας αγοράς (buy it now), ο πωλητής ορίζει την τιμή που επιθυμεί για το προϊόν του. Η τιμή είναι σταθερή, αν ο αγοραστής θέλει να αγοράσει το προϊόν τότε πρέπει να πληρώσει το προϊόν στην τιμή που όρισε ο πωλητής. Η προσφορά του πωλητή μπορεί να έχει διάρκεια χρόνου, όπως τις δημοπρασίες, ή όχι.

Buy Method	Records	%
Bid	3.935	74,90
Buy it Now	1.319	25,10
	5.254	100,00

Πίνακας 3.8: Αριθμός εγγραφών με βάση
τον τρόπο πώλησης

3.3 Επιλογή κατηγοριών για ανάλυση

Όπως έχει αναφερθεί κατηγορίες που έχουν αρχικά επιλεγεί προς ανάλυση έχουν ως ακολούθως:

Product Category	Records	%
Collectables	1.511	28,76
Coins	1.226	23,33
Sports Memorabilia	441	8,39
Stamps	431	8,20
Computers	404	7,69
Antiques	238	4,53
Games	235	4,47
Clothes	148	2,82
Dvds	134	2,55
Books	115	2,19
Watches	99	1,88
Video Games	92	1,75
Magazines	68	1,29
Dolls	64	1,22
Toys	48	0,91
	5.254	100,00

Πίνακας 3.9: Αριθμός εγγραφών κατά κατηγορία

Στη συνέχεια γίνεται επιλογή των κατηγοριών αφού κάποιες κατηγορίες έχουν είτε μικρό δείγμα είτε πολύ χαμηλό ποσοστό πωληθέντων προϊόντων. Η σχέση των προϊόντων που πωλήθηκαν με αυτών που δεν πωλήθηκαν κατά κατηγορία φαίνεται ακολούθως.

3.3.1 Collectables

Πωλήθηκαν	Αριθμός Εγγραφών	%
Ναι	616	41
Όχι	895	59

Σύνολο	1514	100
---------------	------	-----

Πίνακας 3.10: Αριθμός εγγραφών σε σχέση με την πώληση τους

3.3.2 Coins

Πωλήθηκαν	Αριθμός Εγγραφών	%
Ναι	1130	92
Όχι	96	8
Σύνολο	1226	100

Πίνακας 3.11: Αριθμός εγγραφών σε σχέση με την πώληση τους

3.3.3 Sports Memorabilia

Πωλήθηκαν	Αριθμός Εγγραφών	%
Ναι	358	81
Όχι	83	19
Σύνολο	441	100

Πίνακας 3.12: Αριθμός εγγραφών σε σχέση με την πώληση τους

3.3.4 Stamps

Πωλήθηκαν	Αριθμός Εγγραφών	%
Ναι	167	39
Όχι	264	61
Σύνολο	431	100

Πίνακας 3.13: Αριθμός εγγραφών σε σχέση με την πώληση τους

3.3.5 Computers

Πωλήθηκαν	Αριθμός Εγγραφών	%
Ναι	45	89
Όχι	359	11
Σύνολο	404	100

Πίνακας 3.14: Αριθμός εγγραφών σε σχέση με την πώληση τους

3.3.6 Antiques

Πωλήθηκαν	Αριθμός Εγγραφών	%
Ναι	195	82
Όχι	43	18
Σύνολο	238	100

Πίνακας 3.15: Αριθμός εγγραφών σε σχέση με την πώληση τους

3.3.7 Games and Toys

Έχει γίνει ενοποίηση των κατηγοριών Games και Toys αφού έχουν κοινά χαρακτηριστικά και απευθύνονται στο ίδιο κοινό.

Πωλήθηκαν	Αριθμός Εγγραφών	%
Ναι	205	72
Όχι	78	28
Σύνολο	283	100

Πίνακας 3.16: Αριθμός εγγραφών σε σχέση με την πώληση τους

3.3.8 Clothes

Πωλήθηκαν	Αριθμός Εγγραφών	%
Ναι	30	20
Όχι	118	80
Σύνολο	148	100

Πίνακας 3.17: Αριθμός εγγραφών σε σχέση με την πώληση τους

3.3.9 Dvds

Πωλήθηκαν	Αριθμός Εγγραφών	%
Ναι	25	19
Όχι	109	81
Σύνολο	134	100

Πίνακας 3.18: Αριθμός εγγραφών σε σχέση με την πώληση τους

3.3.10 Books

Πωλήθηκαν	Αριθμός Εγγραφών	%
Ναι	25	21
Όχι	90	79
Σύνολο	115	100

Πίνακας 3.19: Αριθμός εγγραφών σε σχέση με την πώληση τους

3.3.11 Watches

Πωλήθηκαν	Αριθμός Εγγραφών	%
Ναι	35	35
Όχι	64	65
Σύνολο	99	100

Πίνακας 3.20: Αριθμός εγγραφών σε σχέση με την πώληση τους

3.3.12 Video Games

Πωλήθηκαν	Αριθμός Εγγραφών	%
Ναι	20	22
Όχι	72	78
Σύνολο	92	100

Πίνακας 3.21: Αριθμός εγγραφών σε σχέση με την πώληση τους

3.3.13 Magazines

Πωλήθηκαν	Αριθμός Εγγραφών	%
Ναι	18	26
Όχι	50	74
Σύνολο	68	100

Πίνακας 3.22: Αριθμός εγγραφών σε σχέση με την πώληση τους

3.3.14 Dolls

Πωλήθηκαν	Αριθμός Εγγραφών	%
Ναι	51	80
Όχι	13	20
Σύνολο	64	100

Πίνακας 3.23: Αριθμός εγγραφών σε σχέση με την πώληση τους

Με βάση τα πιο πάνω επιλέγονται τελικά προς ανάλυση οι κατηγορίες Collectables, Coins, Sports Memorabilia, Stamps, Computers, Antiques, Games and Toys, Clothes, Dvds και Books.

Οι κατηγορίες Watches, Video Games, Magazines και Dolls δεν παρουσίασαν μεγάλο αριθμό εγγραφών κατά τη συλλογή δεδομένων με αποτέλεσμα μία περαιτέρω ανάλυση να μη δώσει ορθή εικόνα για τους παράγοντες που επηρεάζουν την πώληση τους.

3.4 Επιλογή τελικών χαρακτηριστικών

Στην ενότητα 3.2 έγινε περιγραφή των χαρακτηριστικών που καταχωρήθηκαν στη βάση δεδομένων. Από αυτά τα χαρακτηριστικά δε θα ληφθούν υπόψη το νόμισμα (Currency) αφού όλες οι εγγραφές είναι σε αγγλικές λίρες UKP, η χώρα που λειτουργεί το ebay που επιλέχθηκε (EbayCountry) αφού η συλλογή δεδομένο έχει γίνει αποκλειστικά από το ebay.co.uk ,η ημερομηνία πώλησης (SellDate) και ο αριθμός των bids (NoOfBids). Ο αριθμός των bids δε μπορεί να χρησιμοποιηθεί αφού δεν ήταν εφικτό να γίνει συλλογή της αρχικής τιμής (starting bid) έτσι ώστε να διαφανεί κατά πόσο υπάρχει θετική σχέση της τιμής και του αριθμού των bids.

Τα υπόλοιπα χαρακτηριστικά θα χρησιμοποιηθούν και θα εξεταστούν σε σχέση με το χαρακτηριστικό Sold. Τα χαρακτηριστικά SellerType, ProductCondition, BuyMethod και ShippingCost θα χρησιμοποιηθούν ως ανεξάρτητες τιμές για να διαπιστωθεί κατά πόσο και με ποιο τρόπο επηρεάζεται η πώληση ενός προϊόντος. Η εξαρτημένη τιμή θα είναι το χαρακτηριστικό Sold. Το κόστος αποστολής (ShippingCost) θα χρησιμοποιηθεί σε σχέση με το κόστος του προϊόντος (Price) έτσι ώστε να γίνει κατηγοριοποίηση των προϊόντων με βάση την αξία τους.

Τέλος, κάθε κατηγορία προϊόντων (ProductCategory) θα εξεταστεί ξεχωριστά αφού τα χαρακτηριστικά της είναι διαφορετικά και απευθύνεται σε διαφορετικές ομάδες των καταναλωτών με αποτέλεσμα οι παράγοντες που επηρεάζουν σε κάθε κατηγορία προϊόντων να είναι διαφορετικοί.

Ανάλυση δεδομένων

4.1 Έλεγχος υποθέσεων

4.2 Binary Logistic Regression

4. Ανάλυση δεδομένων

Στο κεφάλαιο αυτό θα ασχοληθούμε με την ανάλυση των δεδομένων και την παρατήρηση των ανεξάρτητων μεταβλητών σε σχέση με την εξαρτημένη. Η εξαρτημένη τιμή θα είναι σε όλες η μεταβλητή Sold που αντιπροσωπεύει το προϊόν αν έχει πουληθεί ή όχι.

4.1 Έλεγχος υποθέσεων

Με τον έλεγχο υποθέσεων θα προσπαθήσουμε να δούμε ποιοι παράγοντες και πως επηρεάζουν την απόφαση του καταναλωτή για το αν θα προχωρήσει με την αγορά ή όχι ενός προϊόντος. Σαν εξαρτημένη μεταβλητή θεωρούμε την πώληση σε όλες τις περιπτώσεις. Η ανάλυση των δεδομένων γίνεται ξεχωριστά για κάθε κατηγορία αφού απευθύνονται σε διαφορετικό κοινό, άρα οι λόγοι που επηρεάζουν την απόφαση του καταναλωτή πιθανόν να είναι διαφορετικοί. Σε όλες τις κατηγορίες θα εξεταστεί η μεταβλητή της πώλησης (Sold) με τις μεταβλητές της φυσικής κατάστασης του προϊόντος, του είδους του πωλητή, του τρόπου πώλησης και του κόστους αποστολής. Το κόστος αποστολής θα εξεταστεί σε σχέση με τη τιμή πώλησης του προϊόντος.

4.1.1 Κατηγορία Collectables

4.1.1.1 Σχέση πώλησης με φυσική κατάσταση του προϊόντος

Υποθέτουμε ότι οι πιθανοί αγοραστές θα προτιμούν τα χρησιμοποιημένα προϊόντα της κατηγορίας αφού θα πωλούνται σε χαμηλότερες τιμές. Γι' αυτό το λόγο σαν αρχική υπόθεση θεωρούμε το ποσοστό των καινούριων αντικειμένων που πωλούνται είναι

μικρότερο από το ποσοστό των χρησιμοποιημένων αντικειμένων που πωλούνται.

	New	%	Used	%	
Yes	293	0,46	323	0,37	616
No	348	0,54	547	0,63	895
	641	1,00	870	1,00	1511

Πρόκειται για one-tailed έλεγχο. Για τον έλεγχο θα χρησιμοποιηθεί η μέθοδος z-statistics με επίπεδο σημαντικότητας 5% ($\alpha=0,05$).

$$H_0 : p_1 \leq p_2$$

$$H_1 : p_1 > p_2$$

$$p_c = \frac{\bar{x}_1 + \bar{x}_2}{n_1 + n_2} = 0,41$$

$$z = \frac{p_1 - p_2}{\sqrt{\frac{p_c(1-p_c)}{n_1} + \frac{p_c(1-p_c)}{n_2}}} = 3,36$$

Αφού το $z = 3,36 > z_{0,05} = 1,645$ απορρίπτεται η μηδενική υπόθεση και ισχύει η εναλλακτική. Μπορούμε να πούμε δηλαδή ότι το ποσοστό των καινούριων προϊόντων στην κατηγορία Collectables πωλούνται περισσότερο από το ποσοστό των χρησιμοποιημένων.

Μετά από μελέτη των δεδομένων, η απόρριψη της μηδενικής υπόθεσης φαίνεται να οφείλεται στο γεγονός ότι στην κατηγορία αυτή πωλούνται συλλεκτικά αντικείμενα και ενθύμια (π.χ. καρφίτσες των Ολυμπιακών Αγώνων), όπου οι συλλέκτες τείνουν να προτιμούν τα καινούρια που είναι σε καλύτερη κατάσταση. Παράλληλα, το κόστος των προϊόντων αυτής της κατηγορίας δεν είναι μεγάλο με αποτέλεσμα ο αγοραστής να προτιμά τα καινούρια.

4.1.1.2 Σχέση πώλησης με είδος του πωλητή

Υποθέτουμε ότι οι πιθανοί αγοραστές θα προτιμούν τα προϊόντα της κατηγορίας που πωλούνται από επιχειρήσεις αφού ίσως να εμπιστεύονται περισσότερο τους πωλητές. Γι' αυτό το λόγο σαν αρχική υπόθεση θεωρούμε το ποσοστό των αντικειμένων που πωλούνται από επιχειρήσεις είναι μικρότερο από το ποσοστό των αντικειμένων που πωλούνται από ιδιώτες.

	Business	%	Private	%	
Yes	269	0,4119449	347	0,404429	616
No	384	0,5880551	511	0,595571	895
	653	1	858	1	1511

Πρόκειται για one-tailed έλεγχο. Για τον έλεγχο θα χρησιμοποιηθεί η μέθοδος z-statistics με επίπεδο σημαντικότητας 5% ($\alpha=0,05$).

$$H_0 : p_1 \leq p_2$$

$$H_1 : p_1 > p_2$$

$$p_c = \frac{\bar{x}_1 + \bar{x}_2}{n_1 + n_2} = 0,41$$

$$z = \frac{p_1 - p_2}{\sqrt{\frac{p_c(1-p_c)}{n_1} + \frac{p_c(1-p_c)}{n_2}}} = 0,3$$

Αφού το $z = 0,3 < z_{0,05} = 1,645$ δεν απορρίπτεται η μηδενική υπόθεση με αποτέλεσμα να μη μπορούμε να αποδείξουμε την εναλλακτική ότι δηλαδή το ποσοστό των προϊόντων που πωλούνται από επιχειρήσεις είναι μεγαλύτερο από το ποσοστό προϊόντων που πωλούνται από ιδιώτες.

4.1.1.3 Σχέση πώλησης με τρόπο πώλησης

Υποθέτουμε ότι οι πιθανοί αγοραστές θα προτιμούν τα προϊόντα της κατηγορίας που πωλούνται μέσω δημοπρασίας αφού η αρχική τιμή είναι πάντα χαμηλότερη και υπάρχει δυνατότητα ευκαιριακής

αγοράς. Γι' αυτό το λόγο σαν αρχική υπόθεση θεωρούμε το ποσοστό των αντικειμένων που πωλούνται μέσω δημοπρασίας είναι μικρότερο από το ποσοστό των αντικειμένων που πωλούνται μέσω άμεσης πώλησης (buy it now).

	Bid	%	Buy it Now	%	
Yes	435	0,418269	181	0,384289	616
No	605	0,581731	290	0,615711	895
	1040	1	471	1	1511

Πρόκειται για one-tailed έλεγχο. Για τον έλεγχο θα χρησιμοποιηθεί η μέθοδος z-statistics με επίπεδο σημαντικότητας 5% ($\alpha=0,05$).

$$H_0 : p_1 \leq p_2$$

$$H_1 : p_1 > p_2$$

$$p_c = \frac{\bar{x}_1 + \bar{x}_2}{n_1 + n_2} = 0,41$$

$$z = \frac{p_1 - p_2}{\sqrt{\frac{p_c(1-p_c)}{n_1} + \frac{p_c(1-p_c)}{n_2}}} = 1,25$$

Αφού το $z = 1,25 < z_{0,05} = 1,645$ δεν απορρίπτεται η μηδενική υπόθεση με αποτέλεσμα να μη μπορούμε να αποδείξουμε την εναλλακτική ότι δηλαδή το ποσοστό των προϊόντων που πωλούνται με δημοπρασία είναι μεγαλύτερο από το ποσοστό προϊόντων που υπάρχουν διαθέσιμα προς άμεση πώληση.

4.1.1.4 Σχέση πώλησης με κόστος αποστολής προϊόντος

Το εύρος τιμών της κατηγορίας collectable είναι μεγάλο, με αποτέλεσμα αν γίνει έλεγχος υπόθεσης στο κόστος αποστολής συνολικά για την κατηγορία τα αποτελέσματα δε θα είναι αντιπροσωπευτικά. Γι' αυτό το σκοπό έγινε κατηγοριοποίηση των προϊόντων με βάση το κόστος πώλησης τους. Για κάθε κατηγορία

θα γίνει έλεγχος υπόθεσης σχετικά με το κόστος αποστολής. Η κατηγοριοποίηση έγινε ως ακολούθως:

Κόστος προϊόντος (UKP)	Αριθμός εγγραφών
0-10	688
10-50	427
50-100	63
>100	49

Σημ. Οι τιμές που εμπίπτουν στα όρια των κλάσεων συμπεριλαμβάνονται στην χαμηλότερη κλάση.

Στην αρχική υπόθεση θεωρούμε ότι το κόστος αποστολής των προϊόντων που πωλούνται είναι μεγαλύτερο από το κόστος αποστολής των προϊόντων που δεν πωλούνται.

Κατηγορία 1 (0 - 10):

Πρόκειται για one-tailed έλεγχο. Για τον έλεγχο θα χρησιμοποιηθεί η μέθοδος z-statistics με επίπεδο σημαντικότητας 5% ($\alpha=0,05$).

$$H_0 : \mu_1 \geq \mu_2$$

$$H_1 : \mu_1 < \mu_2$$

$$s_{\bar{x}_1, \bar{x}_2} = \sqrt{\frac{s_1^2}{n_1} + \frac{s_2^2}{n_2}} = 0,16$$

$$z = \frac{(\bar{x}_1 - \bar{x}_2)}{s_{\bar{x}_1, \bar{x}_2}} = 0,11$$

Αφού το $z = 0,11 > z_{0,05} = -1,645$ δεν απορρίπτεται η μηδενική υπόθεση με αποτέλεσμα να μη μπορούμε να αποδείξουμε την εναλλακτική.

Κατηγορία 2 (11 - 50):

Πρόκειται για one-tailed έλεγχο. Για τον έλεγχο θα χρησιμοποιηθεί η μέθοδος z-statistics με επίπεδο σημαντικότητας 5% ($\alpha=0,05$).

$$H_0 : \mu_1 \geq \mu_2$$

$$H_1 : \mu_1 < \mu_2$$

$$s_{\bar{x}_1, \bar{x}_2} = \sqrt{\frac{s_1^2}{n_1} + \frac{s_2^2}{n_2}} = 0,75$$

$$z = \frac{(\bar{x}_1 - \bar{x}_2)}{s_{\bar{x}_1, \bar{x}_2}} = -0,9$$

Αφού το $z = -0,9 > z_{0,05} = -1,645$ δεν απορρίπτεται η μηδενική υπόθεση.

Κατηγορία 3 (51 - 100):

Πρόκειται για one-tailed έλεγχο. Για τον έλεγχο θα χρησιμοποιηθεί η μέθοδος t-statistics αφού πρόκειται για δείγμα μικρότερο του 30, με επίπεδο σημαντικότητας 5% ($\alpha=0,05$).

$$H_0 : \mu_1 \geq \mu_2$$

$$H_1 : \mu_1 < \mu_2$$

$$N_1 = 25$$

$$N_2 = 24$$

Για να βρεθεί το $t_{0,05}$ θα χρησιμοποιηθούν οι βαθμοί ελευθερίας του μικρότερου δείγματος δηλαδή $N_2 - 1 = 23$

$$s_{\bar{x}_1, \bar{x}_2} = \sqrt{\frac{s_1^2}{n_1} + \frac{s_2^2}{n_2}} = 3,01$$

$$t = \frac{(\bar{x}_1 - \bar{x}_2)}{s_{\bar{x}_1, \bar{x}_2}} = -1,06$$

Αφού το $t = -1,06 > t_{0,05,23} = -1,714$ δεν απορρίπτεται η μηδενική υπόθεση.

Κατηγορία 4 (>100):

Πρόκειται για one-tailed έλεγχο. Για τον έλεγχο θα χρησιμοποιηθεί η μέθοδος t-statistics αφού πρόκειται για δείγμα μικρότερο του 30, με επίπεδο σημαντικότητας 5% ($\alpha=0,05$).

$$H_0 : \mu_1 \geq \mu_2$$

$$H_1 : \mu_1 < \mu_2$$

$$N_1 = 23$$

$$N_2 = 33$$

Για να βρεθεί το $t_{0,05}$ θα χρησιμοποιηθούν οι βαθμοί ελευθερίας του μικρότερου δείγματος δηλαδή $N_1 - 1 = 22$

$$s_{\bar{x}_1, \bar{x}_2} = \sqrt{\frac{s_1^2}{n_1} + \frac{s_2^2}{n_2}} = 9,79$$
$$t = \frac{(\bar{x}_1 - \bar{x}_2)}{s_{\bar{x}_1, \bar{x}_2}} = -0,59$$

Αφού το $t = -0,59 > t_{0,05,22} = -1,717$ δεν απορρίπτεται η μηδενική υπόθεση.

Παρατηρούμε ότι και στις τέσσερις κατηγορίες δε μπορέσαμε τα αποδείξουμε ότι το κόστος αποστολής είναι μικρότερο στα αντικείμενα που πωλούνται σε σχέση με αυτά που δεν πωλούνται. Πολύ πιθανόν ο παράγοντας κόστος αποστολής να μην επηρεάζει την αγοραστική απόφαση των καταναλωτών σε αυτή την κατηγορία λόγω του ότι πρόκειται για κυρίως μικρά αντικείμενα με χαμηλό κόστος αποστολής. Επιπλέον, ο καταναλωτής πιθανόν να μην επηρεάζεται από το κόστος αποστολής αφού πρόκειται για συλλεκτικά αντικείμενα όπου δε τα βρίσκει εύκολα στην αγορά.

4.1.1.5 Ανάλυση διασποράς κόστους αποστολής

Με την ανάλυση της διασποράς θα παρατηρήσουμε αν υπάρχει διαφοροποίηση των δεδομένων μεταξύ των κατηγοριών, δηλαδή κατά πόσο αλλάζει το κόστος αποστολής όταν αλλάζει η αξία του προϊόντος. Συγκεκριμένα, τα δεδομένα έχουν χωριστεί σε 4 κατηγορίες βάση της αξίας των προϊόντων της όπως φαίνεται πιο κάτω:

Κόστος προϊόντος (UKP)	Αριθμός εγγραφών
0-10	688
10-50	427
50-100	63
>100	49

Σημ. Στον αριθμό των εγγραφών συμπεριλαμβάνονται και αυτές που δεν αναφέρουν το κόστος αποστολής.

Anova: Single Factor

SUMMARY

Groups	Count	Sum	Average	Variance
CAT1 (0-10)	355	1122,02	3,16062	6,56659
CAT2 (10-50)	169	1147,91	6,792367	25,82577
CAT3 (50-100)	21	112,22	5,34381	80,35403
CAT4 (>100)	17	231,99	13,64647	107,4722

ANOVA

Source of Variation	SS	df	MS	F	P-value	F crit
Between Groups	2949,173	3	983,0577	54,90987	4,11169E-31	2,62087509
Within Groups	9989,938	558	17,90311			
Total	12939,11	561				

Αναλύοντας τη διασπορά των δεδομένων μεταξύ των κατηγοριών παρατηρείται ότι υπάρχει διαφοροποίηση μεταξύ των μέσων όρων των προϊόντων που πωλήθηκαν, αφού το p-value είναι πολύ μικρότερο από 0,001 υποδεικνύοντας την ύπαρξη του μοντέλου. Αυτό σημαίνει ότι ανάλογα με το πραγματικό κόστος του προϊόντος

διαφοροποιείται και το κόστος αποστολής του προϊόντος έτσι ώστε να μπορέσει να πουληθεί.

Αν συγκρίνουμε τους μέσους όσου της κατηγορίας 1 (προϊόντα μέχρι 10 UKP) με την κατηγορία 4 (προϊόντα πάνω από 100 UKP) με έλεγχο υπόθεσης θα έχουμε:

$$H_0 : \mu_2 \leq \mu_1$$

$$H_1 : \mu_2 > \mu_1$$

$$s_{\bar{x}_1, \bar{x}_2} = \sqrt{\frac{s_1^2}{n_1} + \frac{s_2^2}{n_2}} = 2,52$$

$$z = \frac{(\bar{x}_2 - \bar{x}_1)}{s_{\bar{x}_1, \bar{x}_2}} = 4,16$$

Αφού το $z = 4,16 > z_{0,05} = 1,645$ απορρίπτεται η μηδενική υπόθεση. Φαίνεται ότι οι αγοραστές είναι διατεθειμένοι να πληρώσουν μεγαλύτερο κόστος αποστολής όσο αυξάνεται η αξία του προϊόντος αφού η αναλογία της αξίας και του κόστους αποστολής δεν είναι τόσο μεγάλη. Σε προϊόντα μικρής αξίας αν το κόστος αποστολής είναι υψηλό, πιθανόν να μην αξίζει να προχωρήσει ο καταναλωτής σε αγορά του προϊόντος.

4.1.2 Κατηγορία Coins

4.1.2.1 Σχέση πώλησης με είδος του πωλητή

Υποθέτουμε ότι το ποσοστό των νομισμάτων που θα πωλούνται από επιχειρήσεις είναι μικρότερο από το ποσοστό που θα πωλούνται από ιδιώτες.

	Business	%	Private	%	
Yes	832	0,969697	298	0,809783	1130
No	26	0,030303	70	0,190217	96
	858	1	368	1	1226

Πρόκειται για one-tailed έλεγχο. Για τον έλεγχο θα χρησιμοποιηθεί η μέθοδος z-statistics με επίπεδο σημαντικότητας 5% ($\alpha=0,05$).

$$H_0 : p_1 \leq p_2$$

$$H_1 : p_1 > p_2$$

$$p_c = \frac{\bar{x}_1 + \bar{x}_2}{n_1 + n_2} = 0,92$$

$$z = \frac{p_1 - p_2}{\sqrt{\frac{p_c(1-p_c)}{n_1} + \frac{p_c(1-p_c)}{n_2}}} = 9,55$$

Αφού το $z = 9,55 > z_{0,05} = 1,645$ απορρίπτεται η μηδενική υπόθεση και ισχύει η εναλλακτική. Κατ' επέκταση μπορούμε να πούμε ότι το ποσοστό των νομισμάτων που πωλούνται από επιχειρήσεις είναι μεγαλύτερο από το ποσοστό που πωλούνται από ιδιώτες.

Λόγω ότι τα νομίσματα είναι πολλές φορές αρχαία φαίνεται ότι οι αγοραστές προτιμούν να αγοράζουν από επιχείρηση, ίσως επειδή εμπιστεύονται περισσότερο τις αγορές αυτές. Όταν η αγορά γίνεται από επιχείρηση τότε μειώνεται ο κίνδυνος των κλοπιμαίων και των πλαστών νομισμάτων.

4.1.2.2 Σχέση πώλησης με τρόπο πώλησης

Για αρχική υπόθεση θεωρούμε ότι το ποσοστό των νομισμάτων που πωλούνται με δημοπρασία είναι μικρότερο από το ποσοστό που πωλούνται με άμεση πώληση.

	Bid	%	Buy it Now	%	
Yes	1033	0,95	97	0,69	1130
No	52	0,05	44	0,31	96
	1085	1	141	1	1226

Πρόκειται για one-tailed έλεγχο. Για τον έλεγχο θα χρησιμοποιηθεί η μέθοδος z-statistics με επίπεδο σημαντικότητας 5% ($\alpha=0,05$).

$$H_0 : p_1 \leq p_2$$

$$H_1 : p_1 > p_2$$

$$p_c = \frac{\bar{x}_1 + \bar{x}_2}{n_1 + n_2} = 0,92$$

$$z = \frac{p_1 - p_2}{\sqrt{\frac{p_c(1-p_c)}{n_1} + \frac{p_c(1-p_c)}{n_2}}} = 10,98$$

Αφού το $z = 10,98 > z_{0,05} = 1,645$ απορρίπτεται η μηδενική υπόθεση και ισχύει η εναλλακτική. Δηλαδή, μπορούμε να πούμε ότι το ποσοστό των νομισμάτων που πωλούνται με δημοπρασία είναι μεγαλύτερο από το ποσοστό που πωλούνται με τη μέθοδο της απευθείας αγοράς.

Μία εξήγηση που μπορεί να δοθεί είναι η ευκαιριακή αγορά των νομισμάτων. Με τις δημοπρασίες η αρχική τιμή των προϊόντων είναι χαμηλή και ο πωλητής απευθύνεται σε συγκεκριμένη ομάδα αγοραστών τους συλλέκτες νομισμάτων. Ο κάθε αγοραστής ψάχνει συγκεκριμένα εξειδικευμένα αντικείμενα να αγοράσει μειώνοντας έτσι τον ανταγωνισμό με άλλους πιθανούς αγοραστές.

Έτσι δεν υπάρχουν πολλές προσφορές (bids) για κάθε νόμισμα με αποτέλεσμα οι αγοραστές να επιτυγχάνουν καλύτερες τιμές.

4.1.2.3 Σχέση πώλησης με κόστος αποστολής προϊόντος

Το εύρος τιμών της κατηγορίας coins είναι μεγάλο, με αποτέλεσμα αν γίνει έλεγχος υπόθεσης στο κόστος αποστολής συνολικά για την κατηγορία τα αποτελέσματα δε θα είναι αντιπροσωπευτικά. Γι' αυτό το σκοπό έγινε κατηγοριοποίηση των προϊόντων με βάση το κόστος πώλησης τους. Για κάθε κατηγορία θα γίνει έλεγχος υπόθεσης σχετικά με το κόστος αποστολής. Η κατηγοριοποίηση έγινε ως ακολούθως:

Κόστος προϊόντος (UKP)	Αριθμός εγγραφών
0-10	689
10-50	426
50-100	62
>100	50

Σημ. Οι τιμές που εμπίπτουν στα όρια των κλάσεων συμπεριλαμβάνονται στην χαμηλότερη κλάση.

Στην αρχική υπόθεση θεωρούμε ότι το κόστος αποστολής των νομισμάτων που πωλούνται είναι μεγαλύτερο από το κόστος αποστολής αυτών που δεν πωλούνται.

Κατηγορία 1 (0 - 10):

Πρόκειται για one-tailed έλεγχο. Για τον έλεγχο θα χρησιμοποιηθεί η μέθοδος z-statistics με επίπεδο σημαντικότητας 5% ($\alpha=0,05$).

$$H_0 : \mu_1 \geq \mu_2$$

$$H_1 : \mu_1 < \mu_2$$

$$s_{\bar{x}_1, \bar{x}_2} = \sqrt{\frac{s_1^2}{n_1} + \frac{s_2^2}{n_2}} = 0,24$$

$$z = \frac{(\bar{x}_1 - \bar{x}_2)}{s_{\bar{x}_1, \bar{x}_2}} = 4,86$$

Αφού το $z = 4,86 > z_{0,05} = -1,645$ δεν απορρίπτεται η μηδενική υπόθεση.

Κατηγορία 2 (11 - 50):

Πρόκειται για one-tailed έλεγχο. Για τον έλεγχο θα χρησιμοποιηθεί η μέθοδος z-statistics με επίπεδο σημαντικότητας 5% ($\alpha=0,05$).

$$H_0 : \mu_1 \geq \mu_2$$

$$H_1 : \mu_1 < \mu_2$$

$$s_{\bar{x}_1, \bar{x}_2} = \sqrt{\frac{s_1^2}{n_1} + \frac{s_2^2}{n_2}} = 0,73$$

$$z = \frac{(\bar{x}_1 - \bar{x}_2)}{s_{\bar{x}_1, \bar{x}_2}} = 0,58$$

Αφού το $z = 0,58 > z_{0,05} = -1,645$ δεν απορρίπτεται η μηδενική υπόθεση.

Κατηγορία 3 (51 - 100):

Πρόκειται για one-tailed έλεγχο. Για τον έλεγχο θα χρησιμοποιηθεί η μέθοδος t-statistics αφού το μικρότερο δείγμα είναι μικρότερο του 30, με επίπεδο σημαντικότητας 5% ($\alpha=0,05$).

$$H_0 : \mu_1 \geq \mu_2$$

$$H_1 : \mu_1 < \mu_2$$

$$N_1 = 56$$

$$N_2 = 6$$

Για να βρεθεί το $t_{0,05}$ θα χρησιμοποιηθούν οι βαθμοί ελευθερίας του μικρότερου δείγματος δηλαδή $N_2 - 1 = 5$

$$s_{\bar{x}_1, \bar{x}_2} = \sqrt{\frac{s_1^2}{n_1} + \frac{s_2^2}{n_2}} = 0,76$$
$$t = \frac{(\bar{x}_1 - \bar{x}_2)}{s_{\bar{x}_1, \bar{x}_2}} = 5,05$$

Αφού το $t = 5,04 > t_{0,05,5} = -2,015$ δεν απορρίπτεται η μηδενική υπόθεση.

Κατηγορία 4 (>100):

Πρόκειται για one-tailed έλεγχο. Για τον έλεγχο θα χρησιμοποιηθεί η μέθοδος t-statistics αφού το μικρότερο δείγμα είναι μικρότερο του 30, με επίπεδο σημαντικότητας 5% ($\alpha=0,05$).

$$H_0 : \mu_1 \geq \mu_2$$

$$H_1 : \mu_1 < \mu_2$$

$$N_1 = 45$$

$$N_2 = 5$$

Για να βρεθεί το $t_{0,05}$ θα χρησιμοποιηθούν οι βαθμοί ελευθερίας του μικρότερου δείγματος δηλαδή $N_2 - 1 = 4$

$$s_{\bar{x}_1, \bar{x}_2} = \sqrt{\frac{s_1^2}{n_1} + \frac{s_2^2}{n_2}} = 1,53$$
$$t = \frac{(\bar{x}_1 - \bar{x}_2)}{s_{\bar{x}_1, \bar{x}_2}} = 2,056$$

Αφού το $t = 2,056 > t_{0,05,4} = -2,132$ δεν απορρίπτεται η μηδενική υπόθεση.

Παρατηρούμε ότι το κόστος αποστολής δε φαίνεται να αποτελεί σημαντικό παράγοντα στην απόφαση των αγοραστών, πολύ πιθανόν λόγω της φύσης του προϊόντος, όπου αν κάποιος συλλέκτης

επιθυμεί να αγοράσει ένα νόμισμα δε θα λάβει υπόψη στην απόφαση του το κόστος αποστολής.

4.1.2.4 Ανάλυση διασποράς κόστους αποστολής

Με την ανάλυση της διασποράς θα παρατηρήσουμε αν υπάρχει διαφοροποίηση των δεδομένων μεταξύ των κατηγοριών, δηλαδή κατά πόσο αλλάζει το κόστος αποστολής όταν αλλάζει η αξία του προϊόντος. Συγκεκριμένα, τα δεδομένα έχουν χωριστεί σε 4 κατηγορίες βάση της αξίας των προϊόντων της όπως φαίνεται πιο κάτω:

Κόστος προϊόντος (UKP)	Αριθμός εγγραφών
0-10	689
10-50	426
50-100	62
>100	50

Σημ. Στον αριθμό των εγγραφών συμπεριλαμβάνονται και αυτές που δεν αναφέρουν το κόστος αποστολής.

Anova: Single Factor

SUMMARY

<i>Groups</i>	<i>Count</i>	<i>Sum</i>	<i>Average</i>	<i>Variance</i>
CAT1	625	2213,9	3,54224	1,629609
CAT2	391	1472,57	3,766164	1,707425
CAT3	56	217,79	3,889107	1,875517
CAT4	45	240,69	5,348667	22,66306

ANOVA

<i>Source of Variation</i>	<i>SS</i>	<i>df</i>	<i>MS</i>	<i>F</i>	<i>P-value</i>	<i>F crit</i>
Between Groups	141,454	3	47,15134	18,85647	6,21E-12	2,612899
Within Groups	2783,099	1113	2,500539			
Total	2924,554	1116				

Αναλύοντας τη διασπορά των δεδομένων μεταξύ των κατηγοριών παρατηρείται ότι υπάρχει διαφοροποίηση μεταξύ των μέσων όρων των προϊόντων που πωλήθηκαν, αφού το p-value είναι πολύ μικρότερο από 0,001 υποδεικνύοντας την ύπαρξη του μοντέλου. Αυτό σημαίνει ότι ανάλογα με το πραγματικό κόστος του προϊόντος διαφοροποιείται και το κόστος αποστολής του προϊόντος έτσι ώστε να μπορέσει να πουληθεί.

Αν συγκρίνουμε τους μέσους όρους της κατηγορίας 1 (προϊόντα μέχρι 10 UKP) με την κατηγορία 4 (προϊόντα πάνω από 100 UKP) με έλεγχο υπόθεσης θα έχουμε:

$$H_0 : \mu_2 \leq \mu_1$$

$$H_1 : \mu_2 > \mu_1$$

$$s_{\bar{x}_1, \bar{x}_2} = \sqrt{\frac{s_1^2}{n_1} + \frac{s_2^2}{n_2}} = 0,55$$

$$z = \frac{(\bar{x}_2 - \bar{x}_1)}{s_{\bar{x}_1, \bar{x}_2}} = 3,26$$

Αφού το $z = 3,26 > z_{0,05} = 1,645$ απορρίπτεται η μηδενική υπόθεση. Φαίνεται ότι οι αγοραστές είναι διατεθειμένοι να πληρώσουν μεγαλύτερο κόστος αποστολής όσο αυξάνεται η αξία του προϊόντος αφού η αναλογία της αξίας και του κόστους αποστολής δεν είναι τόσο μεγάλη. Σε προϊόντα μικρής αξίας αν το κόστος αποστολής είναι υψηλό, πιθανόν να μην αξίζει να προχωρήσει ο καταναλωτής σε αγορά του προϊόντος.

4.1.3 Κατηγορία Antiques

4.1.3.1 Σχέση πώλησης με είδος του πωλητή

Για τον έλεγχο αυτό θεωρούμε ότι το ποσοστό των αντικών που πωλούνται μέσω επιχειρήσεων είναι μικρότερο από το ποσοστό που πωλούνται μέσω ιδιωτών.

	Business	%	Private	%	
Yes	179	0,94	16	0,33	195
No	11	0,06	32	0,67	43
	190	1	48	1	238

Πρόκειται για one-tailed έλεγχο. Για τον έλεγχο θα χρησιμοποιηθεί η μέθοδος z-statistics με επίπεδο σημαντικότητας 5% ($\alpha=0,05$).

$$H_0 : p_1 \leq p_2$$

$$H_1 : p_1 > p_2$$

$$p_c = \frac{\bar{x}_1 + \bar{x}_2}{n_1 + n_2} = 0,82$$

$$z = \frac{p_1 - p_2}{\sqrt{\frac{p_c(1-p_c)}{n_1} + \frac{p_c(1-p_c)}{n_2}}} = 9,79$$

Αφού το $z = 9,79 > z_{0,05} = 1,645$ απορρίπτεται η μηδενική υπόθεση και ισχύει η εναλλακτική. Κατ' επέκταση μπορούμε να πούμε ότι το ποσοστό των αντικών που πωλούνται από επιχειρήσεις είναι μεγαλύτερο από το ποσοστό που πωλούνται από ιδιώτες.

Λόγω ότι οι αντίκες είναι ιδιαίτερη κατηγορία φαίνεται και από τον έλεγχο υπόθεσης ότι οι αγοραστές προτιμούν να αγοράζουν από επιχείρηση. Ένας λόγος είναι ότι εμπιστεύονται περισσότερο τις αγορές αυτές. Όταν η αγορά γίνεται από επιχείρηση τότε μειώνεται ο κίνδυνος των κλοπιμαίων και των πλαστών αντικειμένων και ο αγοραστής νιώθει πιο σίγουρος για την αγορά του.

4.1.3.2 Σχέση πώλησης με τρόπο πώλησης

Για αρχική υπόθεση θεωρούμε ότι το ποσοστό των αντικών που πωλούνται με δημοπρασία είναι μικρότερο από το ποσοστό που πωλούνται με άμεση πώληση.

	Bid	%	Buy it Now	%	
Yes	192	0,88	3	0,15	195
No	26	0,12	17	0,85	43
	218	1	20	1	238

Πρόκειται για one-tailed έλεγχο. Για τον έλεγχο θα χρησιμοποιηθεί η μέθοδος z-statistics με επίπεδο σημαντικότητας 5% ($\alpha=0,05$).

$$H_0 : p_1 \leq p_2$$

$$H_1 : p_1 > p_2$$

$$p_c = \frac{\bar{x}_1 + \bar{x}_2}{n_1 + n_2} = 0,82$$

$$z = \frac{p_1 - p_2}{\sqrt{\frac{p_c(1-p_c)}{n_1} + \frac{p_c(1-p_c)}{n_2}}} = 8,13$$

Αφού το $z = 8,13 > z_{0,05} = 1,645$ απορρίπτεται η μηδενική υπόθεση και ισχύει η εναλλακτική. Δηλαδή, μπορούμε να πούμε ότι το ποσοστό των αντικών που πωλούνται με δημοπρασία είναι μεγαλύτερο από το ποσοστό που πωλούνται με τη μέθοδο της απευθείας αγοράς.

Μία εξήγηση που μπορεί να δοθεί είναι η αγορά των αντικών σε χαμηλότερη τιμή. Με τις δημοπρασίες η αρχική τιμή των προϊόντων είναι χαμηλή και ο πωλητής απευθύνεται σε συγκεκριμένη ομάδα αγοραστών τους συλλέκτες αντικών. Ο κάθε αγοραστής ψάχνει συγκεκριμένα εξειδικευμένα αντικείμενα να αγοράσει μειώνοντας έτσι τον ανταγωνισμό με άλλους πιθανούς αγοραστές. Έτσι δεν

υπάρχουν πολλές προσφορές (bids) για κάθε αντικα που πωλείται, με αποτέλεσμα οι αγοραστές να επιτυγχάνουν καλύτερες τιμές.

4.1.3.3 Σχέση πώλησης με κόστος αποστολής προϊόντος

Το εύρος τιμών της κατηγορίας αντικών είναι μεγάλο, με αποτέλεσμα αν γίνει έλεγχος υπόθεσης στο κόστος αποστολής συνολικά για την κατηγορία τα αποτελέσματα δε θα είναι αντιπροσωπευτικά. Γι' αυτό το σκοπό έγινε κατηγοριοποίηση των προϊόντων με βάση το κόστος πώλησης τους. Για κάθε κατηγορία θα γίνει έλεγχος υπόθεσης σχετικά με το κόστος αποστολής. Η κατηγοριοποίηση έγινε ως ακολούθως:

Κόστος προϊόντος (UKP)	Αριθμός εγγραφών
0-10	120
10-50	92
50-100	9
>100	17

Σημ. Οι τιμές που εμπίπτουν στα όρια των κλάσεων συμπεριλαμβάνονται στην χαμηλότερη κλάση.

Σαν αρχική υπόθεση θεωρούμε ότι πωλούνται περισσότερο οι αντικες που έχουν μεγαλύτερο κόστος αποστολής.

Κατηγορία 1 (0 - 10):

Πρόκειται για one-tailed έλεγχο. Για τον έλεγχο θα χρησιμοποιηθεί η μέθοδος z-statistics με επίπεδο σημαντικότητας 5% ($\alpha=0,05$).

$$H_0 : \mu_1 \geq \mu_2$$

$$H_1 : \mu_1 < \mu_2$$

$$s_{\bar{x}_1, \bar{x}_2} = \sqrt{\frac{s_1^2}{n_1} + \frac{s_2^2}{n_2}} = 0,73$$

$$z = \frac{(\bar{x}_1 - \bar{x}_2)}{s_{\bar{x}_1, \bar{x}_2}} = -8,23$$

Αφού το $z = -8,23 < z_{0,05} = -1,645$ απορρίπτεται η μηδενική υπόθεση και ισχύει η εναλλακτική ότι στην κατηγορία με αντίκες αξίας 0-10 UKP ο αγοραστής προτιμά αυτά με το χαμηλότερο κόστος αποστολής.

Κατηγορία 2 (11 - 50):

Πρόκειται για one-tailed έλεγχο. Για τον έλεγχο θα χρησιμοποιηθεί η μέθοδος z-statistics με επίπεδο σημαντικότητας 5% ($\alpha=0,05$).

$$H_0 : \mu_1 \geq \mu_2$$

$$H_1 : \mu_1 < \mu_2$$

$$s_{\bar{x}_1, \bar{x}_2} = \sqrt{\frac{s_1^2}{n_1} + \frac{s_2^2}{n_2}} = 1,54$$

$$z = \frac{(\bar{x}_1 - \bar{x}_2)}{s_{\bar{x}_1, \bar{x}_2}} = -2,02$$

Αφού το $z = -2,02 < z_{0,05} = -1,645$ απορρίπτεται η μηδενική υπόθεση και ισχύει η εναλλακτική ότι στην κατηγορία στην κατηγορία με αντίκες αξίας 10-50 UKP ο αγοραστής προτιμά αυτά με το χαμηλότερο κόστος αποστολής.

Κατηγορία 3 (51 - 100):

Πρόκειται για one-tailed έλεγχο. Για τον έλεγχο θα χρησιμοποιηθεί η μέθοδος t-statistics αφού το μικρότερο δείγμα είναι μικρότερο του 30, με επίπεδο σημαντικότητας 5% ($\alpha=0,05$).

$$H_0 : \mu_1 \geq \mu_2$$

$$H_1 : \mu_1 < \mu_2$$

$$N_1 = 6$$

$$N_2 = 3$$

Για να βρεθεί το $t_{0,05}$ θα χρησιμοποιηθούν οι βαθμοί ελευθερίας του μικρότερου δείγματος δηλαδή $N_2 - 1 = 2$

$$s_{\bar{x}_1, \bar{x}_2} = \sqrt{\frac{s_1^2}{n_1} + \frac{s_2^2}{n_2}} = 1,40$$

$$t = \frac{(\bar{x}_1 - \bar{x}_2)}{s_{\bar{x}_1, \bar{x}_2}} = -5,9$$

Αφού το $t = -5,9 < t_{0,05,2} = -2,92$ απορρίπτεται η μηδενική υπόθεση και ισχύει η εναλλακτική ότι στην κατηγορία με αντίκες αξίας 50-100 UKP ο αγοραστής προτιμά αυτά με το χαμηλότερο κόστος αποστολής.

Κατηγορία 4 (>100):

Πρόκειται για one-tailed έλεγχο. Για τον έλεγχο θα χρησιμοποιηθεί η μέθοδος t-statistics αφού το μικρότερο δείγμα είναι μικρότερο του 30, με επίπεδο σημαντικότητας 5% ($\alpha=0,05$).

$$H_0 : \mu_1 \geq \mu_2$$

$$H_1 : \mu_1 < \mu_2$$

$$N_1 = 7$$

$$N_2 = 10$$

Για να βρεθεί το $t_{0,05}$ θα χρησιμοποιηθούν οι βαθμοί ελευθερίας του μικρότερου δείγματος δηλαδή $N_1 - 1 = 6$

$$s_{\bar{x}_1, \bar{x}_2} = \sqrt{\frac{s_1^2}{n_1} + \frac{s_2^2}{n_2}} = 1,97$$

$$t = \frac{(\bar{x}_1 - \bar{x}_2)}{s_{\bar{x}_1, \bar{x}_2}} = 0,999$$

Αφού το $t = 0,999 > t_{0,05,6} = -1,943$ δεν απορρίπτεται η μηδενική υπόθεση.

Από το αποτέλεσμα και την κατάρριψη της μηδενικής υπόθεσης στις πρώτες κατηγορίες φαίνεται ότι για αντίκες μικρής αξίας ο αγοραστής λαμβάνει υπόψη το κόστος αποστολής και προτιμά αυτά με το χαμηλότερο αφού δεν είναι σπάνια και μοναδικά. Όσο αφορά την κατηγορία με αντίκες πάνω των 100 UKP δεν έχουμε καταφέρει να καταρρίψουμε την αρχική υπόθεση και φαίνεται ότι το κόστος αποστολής δεν αποτελεί παράγοντα που να επηρεάζει τον αγοραστή στη λήψη απόφασης σχετικά με το κατά πόσο θα αγοράσει ή όχι μία αντίκα. Αυτό πιθανόν να οφείλεται στο γεγονός ότι ο αγοραστής που ασχολείται με αγορά αντικών είναι διατεθειμένος να πληρώσει περισσότερο για να αποκτήσει κάποιο κομμάτι που θεωρεί μοναδικό.

4.1.3.4 Ανάλυση διασποράς κόστους αποστολής

Με την ανάλυση της διασποράς θα παρατηρήσουμε αν υπάρχει διαφοροποίηση των δεδομένων μεταξύ των κατηγοριών, δηλαδή κατά πόσο αλλάζει το κόστος αποστολής όταν αλλάζει η αξία του προϊόντος. Συγκεκριμένα, τα δεδομένα έχουν χωριστεί σε 4 κατηγορίες βάση της αξίας των προϊόντων της όπως φαίνεται πιο κάτω:

Κόστος προϊόντος (UKP)	Αριθμός εγγραφών
0-10	120
10-50	92
50-100	9
>100	17

Σημ. Στον αριθμό των εγγραφών συμπεριλαμβάνονται και αυτές που δεν αναφέρουν το κόστος αποστολής. Οι τιμές που εμπίπτουν στα όρια των κλάσεων συμπεριλαμβάνονται στην χαμηλότερη κλάση.

SUMMARY

<i>Groups</i>	<i>Count</i>	<i>Sum</i>	<i>Average</i>	<i>Variance</i>
CAT1	95	368,22	3,876	1,229852
CAT2	86	373,7	4,345349	1,60166
CAT3	5	27,16	5,432	9,87582
CAT4	7	68,36	9,765714	20,9547

ANOVA

<i>Source of Variation</i>	<i>SS</i>	<i>df</i>	<i>MS</i>	<i>F</i>	<i>P-value</i>	<i>F crit</i>
Between Groups	232,4864	3	77,49545	35,12564	4,34E-18	2,65239
Within Groups	416,9787	189	2,206236			
Total	649,465	192				

Αναλύοντας τη διασπορά των δεδομένων μεταξύ των κατηγοριών παρατηρείται ότι υπάρχει διαφοροποίηση μεταξύ των μέσων όρων των προϊόντων που πωλήθηκαν, αφού το p-value είναι πολύ μικρότερο από 0,001 υποδεικνύοντας την ύπαρξη του μοντέλου. Αυτό σημαίνει ότι ανάλογα με το πραγματικό κόστος του προϊόντος διαφοροποιείται και το κόστος αποστολής του προϊόντος έτσι ώστε να μπορέσει να πουληθεί.

Αν συγκρίνουμε τους μέσους όσους της κατηγορίας 1 (προϊόντα κάτω των 10 UKP) με την κατηγορία 4 (προϊόντα πάνω από 100 UKP) με έλεγχο υπόθεσης θα έχουμε:

$$H_0 : \mu_2 \leq \mu_1$$

$$H_1 : \mu_2 > \mu_1$$

$$s_{\bar{x}_1, \bar{x}_2} = \sqrt{\frac{s_1^2}{n_1} + \frac{s_2^2}{n_2}} = 1,74$$

$$z = \frac{(\bar{x}_2 - \bar{x}_1)}{s_{\bar{x}_1, \bar{x}_2}} = 3,40$$

Αφού το $z = 3,40 > z_{0,05} = 1,645$ απορρίπτεται η μηδενική υπόθεση. Φαίνεται ότι οι αγοραστές είναι διατεθειμένοι να πληρώσουν μεγαλύτερο κόστος αποστολής όσο αυξάνεται η αξία

του προϊόντος αφού η αναλογία της αξίας και του κόστους αποστολής δεν είναι τόσο μεγάλη. Σε προϊόντα μικρής αξίας αν το κόστος αποστολής είναι υψηλό, πιθανόν να μην αξίζει να προχωρήσει ο καταναλωτής σε αγορά του προϊόντος.

4.1.4 Κατηγορία Books

4.1.4.1 Σχέση πώλησης με φυσική κατάσταση του προϊόντος

Σαν αρχική υπόθεση θεωρούμε ότι το ποσοστό των καινούριων βιβλίων που πωλούνται είναι μικρότερο από το ποσοστό των χρησιμοποιημένων βιβλίων που πωλούνται, πιθανό λόγω της χαμηλότερης τιμής τους.

	New	%	Used	%	
Yes	2	0,17	23	0,22	25
No	10	0,83	80	0,78	90
	12	1,00	103	1,00	115

Πρόκειται για one-tailed έλεγχο. Για τον έλεγχο θα χρησιμοποιηθεί η μέθοδος z-statistics με επίπεδο σημαντικότητας 5% ($\alpha=0,05$).

$$H_0 : p_1 \leq p_2$$

$$H_1 : p_1 > p_2$$

$$p_c = \frac{\bar{x}_1 + \bar{x}_2}{n_1 + n_2} = 0,21$$

$$z = \frac{p_1 - p_2}{\sqrt{\frac{p_c(1-p_c)}{n_1} + \frac{p_c(1-p_c)}{n_2}}} = -0,45$$

Αφού το $z = -0,45 < z_{0,05} = 1,645$ δεν απορρίπτεται η μηδενική υπόθεση και έτσι δε μπορούμε να προβούμε σε συμπεράσματα με τον έλεγχο υπόθεσης.

Φαίνεται ότι η φυσική κατάσταση του βιβλίου δεν παίζει ρόλο στην πώληση του και οι αγοραστές λαμβάνουν υπόψη άλλους παράγοντες την αγορά ενός βιβλίου.

4.1.4.2 Σχέση πώλησης με είδος του πωλητή

Θεωρούμε σαν αρχική υπόθεση ότι το ποσοστό βιβλίων που πωλούνται από επιχειρήσεις είναι μικρότερο από αυτό των βιβλίων που πωλούνται από ιδιώτες.

	Business	%	Private	%	
Yes	23	0,23	2	0,13	25
No	76	0,77	14	0,88	90
	99	1,00	16	1,00	115

Πρόκειται για one-tailed έλεγχο. Για τον έλεγχο θα χρησιμοποιηθεί η μέθοδος z-statistics με επίπεδο σημαντικότητας 5% ($\alpha=0,05$).

$$H_0 : p_1 \leq p_2$$

$$H_1 : p_1 > p_2$$

$$p_c = \frac{\bar{x}_1 + \bar{x}_2}{n_1 + n_2} = 0,22$$

$$z = \frac{p_1 - p_2}{\sqrt{\frac{p_c(1 - p_c)}{n_1} + \frac{p_c(1 - p_c)}{n_2}}} = 0,97$$

Αφού το $z = 0,97 < z_{0,05} = 1,645$ δεν απορρίπτεται η μηδενική υπόθεση με αποτέλεσμα να μη μπορούμε να αποδείξουμε την εναλλακτική ότι δηλαδή το ποσοστό των προϊόντων που πωλούνται από επιχειρήσεις είναι μεγαλύτερο από το ποσοστό προϊόντων που πωλούνται από ιδιώτες.

Φαίνεται ότι ο τύπος του πωλητή (επιχείρηση / ιδιώτης) δεν παίζει ρόλο στην πώληση του και οι αγοραστές λαμβάνουν υπόψη άλλους παράγοντες την αγορά ενός βιβλίου.

4.1.4.3 Σχέση πώλησης με τρόπο πώλησης

Θεωρούμε σαν αρχική υπόθεση ότι πωλούνται λιγότερα βιβλία με τη μέθοδο της δημοπρασίας απ' ότι με τη μέθοδο της απευθείας αγοράς, πιθανόν για το λόγο ότι ο αγοραστής θέλει να αγοράσει άμεσα ένα βιβλίο και έτσι προχωρά στην απευθείας αγορά.

	Bid	%	Buy it Now	%	
Yes	23	0,30	2	0,05	25
No	53	0,70	37	0,95	90
	76	1,00	39	1,00	115

Πρόκειται για one-tailed έλεγχο. Για τον έλεγχο θα χρησιμοποιηθεί η μέθοδος z-statistics με επίπεδο σημαντικότητας 5% ($\alpha=0,05$).

$$H_0 : p_1 \leq p_2$$

$$H_1 : p_1 > p_2$$

$$p_c = \frac{\bar{x}_1 + \bar{x}_2}{n_1 + n_2} = 0,21$$

$$z = \frac{p_1 - p_2}{\sqrt{\frac{p_c(1-p_c)}{n_1} + \frac{p_c(1-p_c)}{n_2}}} = 3,01$$

Αφού το $z = 3,01 > z_{0,05} = 1,645$ απορρίπτεται η μηδενική υπόθεση με αποτέλεσμα και αποδεικνύεται η εναλλακτική. Μπορούμε δηλαδή να πούμε ότι το ποσοστό των προϊόντων που πωλούνται με δημοπρασία είναι μεγαλύτερο από το ποσοστό προϊόντων που υπάρχουν διαθέσιμα προς άμεση πώληση.

Αυτό οφείλεται στο γεγονός ότι ο αγοραστής ψάχνει την καλύτερη ευκαιρία και τη χαμηλότερη τιμή. Το ίδιο βιβλίο μπορεί να το βρει πιθανόν από πολλούς πωλητές και έτσι έχει τη δυνατότητα να κάνει μία χαμηλή προσφορά (bid) και να θέσει αυτός τα όρια του ποσού που μπορεί να διαθέσει. Σε αντίθεση στην απευθείας αγορά η τιμή είναι σταθερή και τις περισσότερες φορές υψηλότερη από αυτή της δημοπρασίας.

4.1.4.4 Σχέση πώλησης με κόστος αποστολής προϊόντος

Το εύρος τιμών της κατηγορίας books είναι μεγάλο, με αποτέλεσμα αν γίνει έλεγχος υπόθεσης στο κόστος αποστολής συνολικά για την κατηγορία τα αποτελέσματα δε θα είναι αντιπροσωπευτικά. Γι' αυτό το σκοπό έγινε κατηγοριοποίηση των προϊόντων με βάση το κόστος πώλησης τους. Για κάθε κατηγορία θα γίνει έλεγχος υπόθεσης σχετικά με το κόστος αποστολής. Η κατηγοριοποίηση έγινε ως ακολούθως:

Κόστος προϊόντος (UKP)	Αριθμός εγγραφών
0-10	76
10-50	10
50-100	3
>100	25

Σημ. Οι τιμές που εμπίπτουν στα όρια των κλάσεων συμπεριλαμβάνονται στην χαμηλότερη κλάση. Παράλληλα, μία εγγραφή δεν έχει τιμή.

Θεωρούμε σαν αρχική υπόθεση ότι το κόστος αποστολής των προϊόντων που πωλούνται είναι μεγαλύτερο από το κόστος αποστολής των προϊόντων που δεν πωλούνται.

Κατηγορία 1 (0 - 10):

Πρόκειται για one-tailed έλεγχο. Για τον έλεγχο θα χρησιμοποιηθεί η μέθοδος z-statistics με επίπεδο σημαντικότητας 5% ($\alpha=0,05$).

$$H_0 : \mu_1 \geq \mu_2$$

$$H_1 : \mu_1 < \mu_2$$

$$s_{\bar{x}_1, \bar{x}_2} = \sqrt{\frac{s_1^2}{n_1} + \frac{s_2^2}{n_2}} = 0,45$$

$$z = \frac{(\bar{x}_1 - \bar{x}_2)}{s_{\bar{x}_1, \bar{x}_2}} = -0,15$$

Αφού το $z = -0,15 > z_{0,05} = -1,645$ δεν απορρίπτεται η μηδενική υπόθεση και έτσι δεν μπορούμε να αποδείξουμε την εναλλακτική.

Κατηγορία 2 (11 - 50):

Δεν υπάρχουν αρκετά δεδομένα έτσι ώστε να γίνει έλεγχος υπόθεσης αφού μόνο 1 από τα 10 προϊόντα της κατηγορίας έχει πουληθεί.

Κατηγορία 3 (51 - 100):

Δεν υπάρχουν αρκετά δεδομένα έτσι ώστε να γίνει έλεγχος υπόθεσης αφού το δείγμα μας αποτελείται από 3 μόνο εγγραφές.

Κατηγορία 4 (>100):

Δεν υπάρχουν αρκετά δεδομένα έτσι ώστε να γίνει έλεγχος υπόθεσης αφού κανένα προϊόν της κατηγορίας δεν έχει πουληθεί.

Φαίνεται ότι το κόστος αποστολής δεν παίζει ρόλο στην πώληση του και οι αγοραστές λαμβάνουν υπόψη άλλους παράγοντες την αγορά ενός βιβλίου.

4.1.5 Κατηγορία Sports Memorabilia

4.1.5.1 Σχέση πώλησης με φυσική κατάσταση του προϊόντος

Σαν αρχική υπόθεση θεωρούμε ότι πωλούνται περισσότερα χρησιμοποιημένα αντικείμενα απ' ότι καινούρια.

	New	%	Used	%	
Yes	16	1,00	342	0,80	358
No	0	0,00	83	0,20	83
	16	1,00	425	1,00	441

Πρόκειται για one-tailed έλεγχο. Για τον έλεγχο θα χρησιμοποιηθεί η μέθοδος z-statistics με επίπεδο σημαντικότητας 5% ($\alpha=0,05$).

$$H_0 : p_1 \leq p_2$$

$$H_1 : p_1 > p_2$$

$$p_c = \frac{\bar{x}_1 + \bar{x}_2}{n_1 + n_2} = 0,81$$

$$z = \frac{p_1 - p_2}{\sqrt{\frac{p_c(1-p_c)}{n_1} + \frac{p_c(1-p_c)}{n_2}}} = 1,96$$

Αφού το $z = 1,96 > z_{0,05} = 1,645$ απορρίπτεται η μηδενική υπόθεση και ισχύει η εναλλακτική. Μπορούμε να πούμε δηλαδή ότι το ποσοστό των καινούριων προϊόντων στην κατηγορία Sports Memorabilia πωλούνται περισσότερο από το ποσοστό των χρησιμοποιημένων.

Φαίνεται ότι οι αγοραστές προτιμούν να αγοράζουν καινούρια αντικείμενα παρά χρησιμοποιημένα. Ένας λόγος πιθανόν να είναι η καλύτερη φυσική κατάσταση των καινούριων αντικειμένων, έναντι των χρησιμοποιημένων. Παράλληλα, επειδή πρόκειται για αναμνηστικά αθλημάτων οι τιμές δεν είναι απαγορευτικές και έτσι επιτρέπεται στους αγοραστές να προτιμούν τα καινούρια.

4.1.5.2 Σχέση πώλησης με είδος του πωλητή

Σαν μηδενική υπόθεση θεωρούμε ότι πωλούνται λιγότερο τα αντικείμενα που προέρχονται από επιχειρήσεις.

	Business	%	Private	%	
Yes	274	0,78	84	0,92	358
No	76	0,22	7	0,08	83
	350	1,00	91	1,00	441

Πρόκειται για one-tailed έλεγχο. Για τον έλεγχο θα χρησιμοποιηθεί η μέθοδος z-statistics με επίπεδο σημαντικότητας 5% ($\alpha=0,05$).

$$H_0 : p_1 \leq p_2$$

$$H_1 : p_1 > p_2$$

$$p_c = \frac{\bar{x}_1 + \bar{x}_2}{n_1 + n_2} = 0,81$$

$$z = \frac{p_1 - p_2}{\sqrt{\frac{p_c(1-p_c)}{n_1} + \frac{p_c(1-p_c)}{n_2}}} = -3,05$$

Αφού το $z = -3,05 < z_{0,05} = 1,645$ δεν απορρίπτεται η μηδενική υπόθεση με αποτέλεσμα να μη μπορούμε να αποδείξουμε την εναλλακτική ότι δηλαδή το ποσοστό των προϊόντων που πωλούνται από επιχειρήσεις είναι μεγαλύτερο από το ποσοστό προϊόντων που πωλούνται από ιδιώτες.

Φαίνεται ότι ο παράγοντας του τύπου του πωλητή δεν επηρεάζει τις αγοραστικές αποφάσεις του καταναλωτή και ο καταναλωτής λαμβάνει άλλους παράγοντες υπόψη για το αν θα αγοράσει ή όχι ένα αντικείμενο αυτής της κατηγορίας.

4.1.5.3 Σχέση πώλησης με τρόπο πώλησης

Θεωρούμε σαν αρχική υπόθεση ότι το ποσοστό των αντικειμένων που αγοράστηκαν με τη μέθοδο της απευθείας αγοράς είναι μεγαλύτερο από το ποσοστό που αγοράστηκαν μέσω δημοπρασίας.

	Bid	%	Buy it Now	%	
Yes	353	0,82	5	0,63	358
No	80	0,18	3	0,38	83
	433	1,00	8	1,00	441

Πρόκειται για one-tailed έλεγχο. Για τον έλεγχο θα χρησιμοποιηθεί η μέθοδος z-statistics με επίπεδο σημαντικότητας 5% ($\alpha=0,05$).

$$H_0 : p_1 \leq p_2$$

$$H_1 : p_1 > p_2$$

$$p_c = \frac{\bar{x}_1 + \bar{x}_2}{n_1 + n_2} = 0,81$$

$$z = \frac{p_1 - p_2}{\sqrt{\frac{p_c(1-p_c)}{n_1} + \frac{p_c(1-p_c)}{n_2}}} = 1,36$$

Αφού το $z = 1,36 < z_{0,05} = 1,645$ δεν απορρίπτεται η μηδενική υπόθεση με αποτέλεσμα να μη μπορούμε να αποδείξουμε την εναλλακτική ότι δηλαδή το ποσοστό των προϊόντων που πωλούνται με δημοπρασία είναι μεγαλύτερο από το ποσοστό προϊόντων που υπάρχουν διαθέσιμα προς άμεση πώληση.

Φαίνεται ότι η μέθοδος πώλησης ενός αντικειμένου αυτής της κατηγορίας δεν παίζει ρόλο στην επιλογή του αγοραστή. Αυτό πιθανό να οφείλεται στο γεγονός ότι η κατηγορία αφορά αναμνηστικά που έχουν σχέση με τον αθλητισμό, χαμηλό κόστος και ο αγοραστής δεν ενδιαφέρεται κατά πόσο είναι δημοπρασία ή άμεση πώληση.

4.1.5.4 Σχέση πώλησης με κόστος αποστολής προϊόντος

Το εύρος τιμών της κατηγορίας sports memorabilia είναι μεγάλο, με αποτέλεσμα αν γίνει έλεγχος υπόθεσης στο κόστος αποστολής συνολικά για την κατηγορία τα αποτελέσματα δε θα είναι αντιπροσωπευτικά. Γι' αυτό το σκοπό έγινε κατηγοριοποίηση των

προϊόντων με βάση το κόστος πώλησης τους. Για κάθε κατηγορία θα γίνει έλεγχος υπόθεσης σχετικά με το κόστος αποστολής. Η κατηγοριοποίηση έγινε ως ακολούθως:

Κόστος προϊόντος (UKP)	Αριθμός εγγραφών
0-10	393
10-50	45
50-100	3

Σαν μηδενική υπόθεση θεωρούμε ότι το κόστος αποστολής των προϊόντων που πωλούνται είναι μεγαλύτερο από το κόστος αποστολής αυτών που δεν πωλούνται.

Κατηγορία 1 (0 - 10):

Πρόκειται για one-tailed έλεγχο. Για τον έλεγχο θα χρησιμοποιηθεί η μέθοδος z-statistics με επίπεδο σημαντικότητας 5% ($\alpha=0,05$).

$$H_0 : \mu_1 \geq \mu_2$$

$$H_1 : \mu_1 < \mu_2$$

$$s_{\bar{x}_1, \bar{x}_2} = \sqrt{\frac{s_1^2}{n_1} + \frac{s_2^2}{n_2}} = 0,28$$

$$z = \frac{(\bar{x}_1 - \bar{x}_2)}{s_{\bar{x}_1, \bar{x}_2}} = 1,53$$

Αφού το $z = 1,53 > z_{0,05} = -1,645$ δεν απορρίπτεται η μηδενική υπόθεση και δεν αποδεικνύεται η εναλλακτική ότι δηλαδή για προϊόντα με τιμή πώλησης από 0-10 UKP πωλούνται αυτά με χαμηλότερο κόστος αποστολής.

Κατηγορία 2 (11 - 50):

Πρόκειται για one-tailed έλεγχο. Για τον έλεγχο θα χρησιμοποιηθεί η μέθοδος t-statistics αφού πρόκειται για δείγμα μικρότερο του 30, με επίπεδο σημαντικότητας 5% ($\alpha=0,05$).

$$H_0 : \mu_1 \geq \mu_2$$

$$H_1 : \mu_1 < \mu_2$$

$$N_1 = 36$$

$$N_2 = 9$$

Για να βρεθεί το $t_{0,05}$ θα χρησιμοποιηθούν οι βαθμοί ελευθερίας του μικρότερου δείγματος δηλαδή $N_2 - 1 = 8$

$$s_{\bar{x}_1, \bar{x}_2} = \sqrt{\frac{s_1^2}{n_1} + \frac{s_2^2}{n_2}} = 0,63$$

$$t = \frac{(\bar{x}_1 - \bar{x}_2)}{s_{\bar{x}_1, \bar{x}_2}} = 1,16$$

Αφού το $t = 1,16 > t_{0,05,8} = -1,860$ δεν απορρίπτεται η μηδενική υπόθεση.

Κατηγορία 3 (51 - 100):

Δεν υπάρχουν αρκετά δεδομένα έτσι ώστε να γίνει έλεγχος υπόθεσης αφού το δείγμα μας αποτελείται από 3 μόνο εγγραφές.

Στις κατηγορίες όπου υπήρχαν δεδομένα για ανάλυση φαίνεται ότι το κόστος αποστολής δεν επηρεάζει την απόφαση των αγοραστών για αγορά ή όχι ενός αντικειμένου της κατηγορίας αυτής. Αυτό πιθανόν να οφείλεται στο γεγονός ότι τα αθλητικά ενθύμια έχουν μικρές διαστάσεις και βάρος με αποτέλεσμα να έχουν χαμηλό κόστος αποστολής και γι' αυτό να μην αποτελεί παράγοντα που επηρεάζει. Την ίδια στιγμή, επειδή πρόκειται για αναμνηστικά ο αγοραστής θα προχωρήσει στην αγορά ανεξάρτητα από το κόστος αποστολής αφού είναι προϊόν που επιθυμεί να αγοράσει.

4.1.5.5 Ανάλυση διασποράς κόστους αποστολής

Με την ανάλυση της διασποράς θα παρατηρήσουμε αν υπάρχει διαφοροποίηση των δεδομένων μεταξύ των κατηγοριών, δηλαδή κατά πόσο αλλάζει το κόστος αποστολής όταν αλλάζει η αξία του προϊόντος. Συγκεκριμένα, τα δεδομένα έχουν χωριστεί σε 3 κατηγορίες βάση της αξίας των προϊόντων της όπως φαίνεται πιο κάτω:

Κόστος προϊόντος (UKP)	Αριθμός εγγραφών
0-10	393
10-50	45
50-100	3

Σημ: Στον αριθμό των εγγραφών συμπεριλαμβάνονται και αυτές που δεν αναφέρουν το κόστος αποστολής.

Anova: Single Factor

SUMMARY

Groups	Count	Sum	Average	Variance
CAT1	316	1361,96	4,31	5,019924
CAT2	28	191,41	6,836071	6,969647

ANOVA

Source of Variation	SS	df	MS	F	P-value	F crit
Between Groups	164,1262	1	164,1262	31,72226	3,72E-08	3,868792
Within Groups	1769,456	342	5,173849			
Total	1933,583	343				

Αναλύοντας τη διασπορά των δεδομένων μεταξύ των κατηγοριών παρατηρείται ότι υπάρχει διαφοροποίηση μεταξύ των μέσων όρων των προϊόντων που πωλήθηκαν, αφού το p-value είναι πολύ μικρότερο από 0,001 υποδεικνύοντας την ύπαρξη του μοντέλου.

Αυτό σημαίνει ότι ανάλογα με το πραγματικό κόστος του προϊόντος διαφοροποιείται και το κόστος αποστολής του προϊόντος έτσι ώστε να μπορέσει να πουληθεί.

Αν συγκρίνουμε τους μέσους όρους της κατηγορίας 1 (προϊόντα μέχρι 10 UKP) με την κατηγορία 2 (προϊόντα από 10 μέχρι 50) με έλεγχο υπόθεσης θα έχουμε:

$$H_0 : \mu_2 \leq \mu_1$$

$$H_1 : \mu_2 > \mu_1$$

$$s_{\bar{x}_1, \bar{x}_2} = \sqrt{\frac{s_1^2}{n_1} + \frac{s_2^2}{n_2}} = 0,51$$

$$z = \frac{(\bar{x}_2 - \bar{x}_1)}{s_{\bar{x}_1, \bar{x}_2}} = 4,90$$

Αφού το $z = 4,90 > z_{0,05} = 1,645$ απορρίπτεται η μηδενική υπόθεση. Φαίνεται ότι οι αγοραστές είναι διατεθειμένοι να πληρώσουν μεγαλύτερο κόστος αποστολής όσο αυξάνεται η αξία του προϊόντος αφού η αναλογία της αξίας και του κόστους αποστολής δεν είναι τόσο μεγάλη. Σε προϊόντα μικρής αξίας αν το κόστος αποστολής είναι υψηλό, πιθανόν να μην αξίζει να προχωρήσει ο καταναλωτής σε αγορά του προϊόντος.

4.1.6 Κατηγορία Stamps

4.1.6.1 Σχέση πώλησης με είδος του πωλητή

Σαν μηδενική υπόθεση θεωρούμε ότι το ποσοστό των γραμματοσήμων που πωλούνται από ιδιώτες είναι μεγαλύτερο από το ποσοστό που πωλούνται από επιχειρήσεις.

	Business	%	Private	%	
Yes	24	0,77	143	0,36	167
No	7	0,23	257	0,64	264
	31	1,00	400	1,00	431

Πρόκειται για one-tailed έλεγχο. Για τον έλεγχο θα χρησιμοποιηθεί η μέθοδος z-statistics με επίπεδο σημαντικότητας 5% ($\alpha=0,05$).

$$H_0 : p_1 \leq p_2$$

$$H_1 : p_1 > p_2$$

$$p_c = \frac{\bar{x}_1 + \bar{x}_2}{n_1 + n_2} = 0,39$$

$$z = \frac{p_1 - p_2}{\sqrt{\frac{p_c(1-p_c)}{n_1} + \frac{p_c(1-p_c)}{n_2}}} = 4,59$$

Αφού το $z = 4,59 > z_{0,05} = 1,645$ απορρίπτεται η μηδενική υπόθεση και ισχύει η εναλλακτική. Κατ' επέκταση μπορούμε να πούμε ότι το ποσοστό των γραμματόσημων που πωλούνται από επιχειρήσεις είναι μεγαλύτερο από το ποσοστό που πωλούνται από ιδιώτες.

Λόγω ότι τα γραμματόσημα είναι ιδιαίτερη κατηγορία φαίνεται και από τον έλεγχο υπόθεσης ότι οι αγοραστές προτιμούν να αγοράζουν από επιχειρήσεις. Ένας λόγος είναι ότι εμπιστεύονται περισσότερο τις αγορές αυτές. Όταν η αγορά γίνεται από επιχείρηση τότε μειώνεται ο κίνδυνος των πλαστών γραμματοσήμων και ο

αγοραστής νιώθει πιο σίγουρος για την αγορά του. Ο αγοραστής δεν έχει φυσική επαφή με το αντικείμενο που θα αγοράσει έτσι δίνει μεγάλη σημασία στην αξιοπιστία του πωλητή.

4.1.6.2 Σχέση πώλησης με τρόπο πώλησης

Σαν μηδενική υπόθεση θεωρούμε ότι το ποσοστό των γραμματοσήμων που πωλούνται μέσω απευθείας πώλησης είναι μεγαλύτερο από αυτό που πωλούνται μέσω δημοπρασίας.

	Bid	%	Buy it Now	%	
Yes	163	0,39	4	0,50	167
No	260	0,61	4	0,50	264
	423	1,00	8	1,00	431

Πρόκειται για one-tailed έλεγχο. Για τον έλεγχο θα χρησιμοποιηθεί η μέθοδος z-statistics με επίπεδο σημαντικότητας 5% ($\alpha=0,05$).

$$H_0 : p_1 \leq p_2$$

$$H_1 : p_1 > p_2$$

$$p_c = \frac{\bar{x}_1 + \bar{x}_2}{n_1 + n_2} = 0,39$$

$$z = \frac{p_1 - p_2}{\sqrt{\frac{p_c(1-p_c)}{n_1} + \frac{p_c(1-p_c)}{n_2}}} = -0,66$$

Αφού το $z = -0,66 < z_{0,05} = 1,645$ δεν απορρίπτεται η μηδενική υπόθεση με αποτέλεσμα να μη μπορούμε να αποδείξουμε την εναλλακτική. Ότι δηλαδή, το ποσοστό των γραμματοσήμων που πωλούνται με δημοπρασία είναι μεγαλύτερο από το ποσοστό που πωλούνται με τη μέθοδο της απευθείας αγοράς.

Φαίνεται ότι ο τρόπος πώλησης δεν είναι παράγοντας που να καθορίζει κατά πόσο θα αγοράσει κάποιος ένα γραμματόσημο ή όχι. Η κατηγορία αυτή απευθύνεται σε συλλέκτες έτσι αν βρεθεί το

κατάλληλο αντικείμενο ο αγοραστής θα προχωρήσει στην αγορά ανεξαρτήτως του τρόπου πώλησης του.

4.1.6.3 Σχέση πώλησης με κόστος αποστολής προϊόντος

Το εύρος τιμών της κατηγορίας coins είναι μεγάλο, με αποτέλεσμα αν γίνει έλεγχος υπόθεσης στο κόστος αποστολής συνολικά για την κατηγορία τα αποτελέσματα δε θα είναι αντιπροσωπευτικά. Γι' αυτό το σκοπό έγινε κατηγοριοποίηση των προϊόντων με βάση το κόστος πώλησης τους. Για κάθε κατηγορία θα γίνει έλεγχος υπόθεσης σχετικά με το κόστος αποστολής. Η κατηγοριοποίηση έγινε ως ακολούθως:

Κόστος προϊόντος (UKP)	Αριθμός εγγραφών
0-10	396
10-50	30
50-100	2
>100	3

Σημ. Οι τιμές που emπίπτουν στα όρια των κλάσεων συμπεριλαμβάνονται στην χαμηλότερη κλάση.

Σαν μηδενική υπόθεση υποθέτουμε ότι πωλούνται περισσότερο τα γραμματόσημα με μεγαλύτερο κόστος αποστολής.

Κατηγορία 1 (0 - 10):

Πρόκειται για one-tailed έλεγχο. Για τον έλεγχο θα χρησιμοποιηθεί η μέθοδος z-statistics με επίπεδο σημαντικότητας 5% ($\alpha=0,05$).

$$H_0 : \mu_1 \geq \mu_2$$

$$H_1 : \mu_1 < \mu_2$$

$$s_{\bar{x}_1, \bar{x}_2} = \sqrt{\frac{s_1^2}{n_1} + \frac{s_2^2}{n_2}} = 0,06$$

$$z = \frac{(\bar{x}_1 - \bar{x}_2)}{s_{\bar{x}_1, \bar{x}_2}} = 4,48$$

Αφού το $z = 4,48 > z_{0,05} = -1,645$ δεν απορρίπτεται η μηδενική υπόθεση.

Κατηγορία 2 (11 - 50):

Πρόκειται για one-tailed έλεγχο. Για τον έλεγχο θα χρησιμοποιηθεί η μέθοδος t-statistics αφού το μικρότερο δείγμα είναι μικρότερο του 30, με επίπεδο σημαντικότητας 5% ($\alpha=0,05$).

$$H_0 : \mu_1 \geq \mu_2$$

$$H_1 : \mu_1 < \mu_2$$

$$N_1 = 24$$

$$N_2 = 6$$

Για να βρεθεί το $t_{0,05}$ θα χρησιμοποιηθούν οι βαθμοί ελευθερίας του μικρότερου δείγματος δηλαδή $N_2 - 1 = 5$

$$s_{\bar{x}_1, \bar{x}_2} = \sqrt{\frac{s_1^2}{n_1} + \frac{s_2^2}{n_2}} = 0,49$$

$$t = \frac{(\bar{x}_1 - \bar{x}_2)}{s_{\bar{x}_1, \bar{x}_2}} = 3,58$$

Αφού το $t = 3,58 > t_{0,05,5} = -2,015$ δεν απορρίπτεται η μηδενική υπόθεση.

Κατηγορία 3 (51 - 100):

Δεν υπάρχουν αρκετά δεδομένα έτσι ώστε να γίνει έλεγχος υπόθεσης αφού το δείγμα μας αποτελείται από 2 μόνο εγγραφές.

Κατηγορία 4 (>100):

Δεν υπάρχουν αρκετά δεδομένα έτσι ώστε να γίνει έλεγχος υπόθεσης αφού το δείγμα μας αποτελείται από 3 μόνο εγγραφές.

Στις κατηγορίες όπου υπάρχουν διαθέσιμα δεδομένα για ανάλυση παρατηρούμε ότι οι αγοραστές δε φαίνεται να επηρεάζονται από το κόστος αποστολής και να λαμβάνουν άλλους παράγοντες κατά την απόφαση τους κατά πόσο θα αγοράσουν ή όχι ένα γραμματόσημο (ή μία σειρά γραμματοσήμων). Αυτό συμβαίνει, πιθανόν γιατί οι αγοραστές είναι συλλέκτες και έτσι είναι διατεθειμένοι να πληρώσουν μεγαλύτερο κόστος αποστολής αν το αντικείμενο που είναι προς πώληση τους ικανοποιεί.

4.1.6.4 Ανάλυση διασποράς κόστους αποστολής

Με την ανάλυση της διασποράς θα παρατηρήσουμε αν υπάρχει διαφοροποίηση των δεδομένων μεταξύ των κατηγοριών, δηλαδή κατά πόσο αλλάζει το κόστος αποστολής όταν αλλάζει η αξία του προϊόντος. Συγκεκριμένα, τα δεδομένα έχουν χωριστεί σε 4 κατηγορίες βάση της αξίας των προϊόντων της όπως φαίνεται πιο κάτω:

Κόστος προϊόντος (UKP)	Αριθμός εγγραφών
0-10	396
10-50	30
50-100	2
>100	3

Σημ. Στον αριθμό των εγγραφών συμπεριλαμβάνονται και αυτές που δεν αναφέρουν το κόστος αποστολής. Οι τιμές που εμπίπτουν στα όρια των κλάσεων συμπεριλαμβάνονται στην χαμηλότερη κλάση.

Anova: Single Factor

SUMMARY

Groups	Count	Sum	Average	Variance
CAT1	135	198,97	1,473852	0,329596
CAT2	23	82,07	3,568261	5,426415

ANOVA

Source of Variation	SS	df	MS	F	P-value	F crit
Between Groups	86,20402	1	86,20402	82,22611	4,89E-16	3,901761
Within Groups	163,5469	156	1,048378			
Total	249,7509	157				

Αναλύοντας τη διασπορά των δεδομένων μεταξύ των κατηγοριών παρατηρείται ότι υπάρχει διαφοροποίηση μεταξύ των μέσων όρων των προϊόντων που πωλήθηκαν, αφού το p-value είναι πολύ μικρότερο από 0,001 υποδεικνύοντας την ύπαρξη του μοντέλου. Αυτό σημαίνει ότι ανάλογα με το πραγματικό κόστος του προϊόντος διαφοροποιείται και το κόστος αποστολής του προϊόντος έτσι ώστε να μπορέσει να πουληθεί.

Αν συγκρίνουμε τους μέσους όρους της κατηγορίας 1 (προϊόντα μέχρι 10 UKP) με την κατηγορία 2 (προϊόντα από 10 μέχρι 50) με έλεγχο υπόθεσης θα έχουμε:

$$H_0 : \mu_2 \leq \mu_1$$

$$H_1 : \mu_2 > \mu_1$$

$$s_{\bar{x}_1, \bar{x}_2} = \sqrt{\frac{s_1^2}{n_1} + \frac{s_2^2}{n_2}} = 0,49$$

$$z = \frac{(\bar{x}_2 - \bar{x}_1)}{s_{\bar{x}_1, \bar{x}_2}} = 4,29$$

Αφού το $z = 4,29 > z_{0,05} = 1,645$ απορρίπτεται η μηδενική υπόθεση. Φαίνεται ότι οι αγοραστές είναι διατεθειμένοι να πληρώσουν μεγαλύτερο κόστος αποστολής όσο αυξάνεται η αξία του προϊόντος αφού η αναλογία της αξίας και του κόστους

αποστολής δεν είναι τόσο μεγάλη. Σε προϊόντα μικρής αξίας αν το κόστος αποστολής είναι υψηλό, πιθανόν να μην αξίζει να προχωρήσει ο καταναλωτής σε αγορά του προϊόντος.

4.1.7 Κατηγορία Computers

4.1.7.1 Σχέση πώλησης με φυσική κατάσταση του προϊόντος

Σαν μηδενική υπόθεση θεωρούμε ότι το ποσοστό χρησιμοποιημένων των προϊόντων της κατηγορίας υπολογιστές που πωλούνται είναι μεγαλύτερο από αυτό των καινούριων.

	New	%	Used	%	
Yes	27	0,11	18	0,12	45
No	229	0,89	130	0,88	359
	256	1,00	148	1,00	404

Πρόκειται για one-tailed έλεγχο. Για τον έλεγχο θα χρησιμοποιηθεί η μέθοδος z-statistics με επίπεδο σημαντικότητας 5% ($\alpha=0,05$).

$$H_0 : p_1 \leq p_2$$

$$H_1 : p_1 > p_2$$

$$p_c = \frac{\bar{x}_1 + \bar{x}_2}{n_1 + n_2} = 0,11$$

$$z = \frac{p_1 - p_2}{\sqrt{\frac{p_c(1-p_c)}{n_1} + \frac{p_c(1-p_c)}{n_2}}} = -0,50$$

Αφού το $z = -0,50 < z_{0,05} = 1,645$ δεν απορρίπτεται η μηδενική υπόθεση με αποτέλεσμα να μην αποδεικνύεται η εναλλακτική.

Η φυσική κατάσταση των προϊόντων δε φαίνεται να είναι ένας από τους παράγοντες που παίζουν ρόλο στην απόφαση του αγοραστή κατά πόσο θα αγοράσει ή όχι ένα προϊόν αυτής της κατηγορίας.

4.1.7.2 Σχέση πώλησης με είδος του πωλητή

Σαν μηδενική υπόθεση θεωρούμε ότι πωλούνται περισσότερο προϊόντα από ιδιώτες παρά από επιχειρήσεις.

	Business	%	Private	%	
Yes	15	0,05	30	0,39	45
No	312	0,95	47	0,61	359
	327	1	77	1	404

Πρόκειται για one-tailed έλεγχο. Για τον έλεγχο θα χρησιμοποιηθεί η μέθοδος z-statistics με επίπεδο σημαντικότητας 5% ($\alpha=0,05$).

$$H_0 : p_1 \leq p_2$$

$$H_1 : p_1 > p_2$$

$$p_c = \frac{\bar{x}_1 + \bar{x}_2}{n_1 + n_2} = 0,11$$

$$z = \frac{p_1 - p_2}{\sqrt{\frac{p_c(1-p_c)}{n_1} + \frac{p_c(1-p_c)}{n_2}}} = -8,63$$

Αφού το $z = -8,63 < z_{0,05} = 1,645$ δεν απορρίπτεται η μηδενική υπόθεση με αποτέλεσμα να μη μπορούμε να αποδείξουμε την εναλλακτική ότι δηλαδή το ποσοστό των προϊόντων που πωλούνται από επιχειρήσεις είναι μεγαλύτερο από το ποσοστό προϊόντων που πωλούνται από ιδιώτες.

Το είδος του πωλητή δε φαίνεται να επηρεάζει την απόφαση του αγοραστή κατά πόσο θα αγοράσει ή όχι ένα προϊόν. Ο αγοραστής, αν βρει το προϊόν που ικανοποιεί τις ανάγκες του θα το αγοράσει ανεξαρτήτως του πωλητή.

4.1.7.3 Σχέση πώλησης με τρόπο πώλησης

Σαν μηδενική υπόθεση θεωρούμε ότι το ποσοστό των αντικειμένων που πωλούνται με τη μέθοδο της απευθείας πώλησης είναι μεγαλύτερος από το ποσοστό των αντικειμένων που πωλούνται με δημοπρασίες.

	Bid	%	Buy it Now	%	
Yes	23	0,45	22	0,06	45

No	28	0,55	331	0,94	359
	51	1	353	1	404

Πρόκειται για one-tailed έλεγχο. Για τον έλεγχο θα χρησιμοποιηθεί η μέθοδος z-statistics με επίπεδο σημαντικότητας 5% ($\alpha=0,05$).

$$H_0 : p_1 \leq p_2$$

$$H_1 : p_1 > p_2$$

$$p_c = \frac{\bar{x}_1 + \bar{x}_2}{n_1 + n_2} = 0,11$$

$$z = \frac{p_1 - p_2}{\sqrt{\frac{p_c(1-p_c)}{n_1} + \frac{p_c(1-p_c)}{n_2}}} = 8,25$$

Αφού το $z = 8,25 > z_{0,05} = 1,645$ απορρίπτεται η μηδενική υπόθεση αποδεικνύεται η εναλλακτική. Μπορούμε δηλαδή να πούμε ότι το ποσοστό των προϊόντων που πωλούνται με δημοπρασία είναι μεγαλύτερο από το ποσοστό προϊόντων που υπάρχουν διαθέσιμα προς άμεση πώληση.

Φαίνεται ότι οι αγοραστές προτιμούν να αγοράζουν τα αντικείμενα αυτής της κατηγορίας μέσω δημοπρασίας. Ένας πιθανός λόγος που να εξηγεί τη συμπεριφορά αυτή είναι ότι με τις δημοπρασίες υπάρχει η δυνατότητα να αγοραστεί ένα προϊόν σε χαμηλή τιμή, αν αυτό δεν έχει μεγάλη ζήτηση.

4.1.7.4 Σχέση πώλησης με κόστος αποστολής προϊόντος

Το εύρος τιμών της κατηγορίας computers είναι μεγάλο, με αποτέλεσμα αν γίνει έλεγχος υπόθεσης στο κόστος αποστολής συνολικά για την κατηγορία τα αποτελέσματα δε θα είναι αντιπροσωπευτικά. Γι' αυτό το σκοπό έγινε κατηγοριοποίηση των προϊόντων με βάση το κόστος πώλησης τους. Για κάθε κατηγορία

θα γίνει έλεγχος υπόθεσης σχετικά με το κόστος αποστολής. Η κατηγοριοποίηση έγινε ως ακολούθως:

Κόστος προϊόντος (UKP)	Αριθμός εγγραφών
0-10	175
10-50	169
50-100	30
>100	30

Σημ. Οι τιμές που εμπίπτουν στα όρια των κλάσεων συμπεριλαμβάνονται στην χαμηλότερη κλάση.

Σαν μηδενική υπόθεση θεωρούμε ότι πωλούνται περισσότερο τα αντικείμενα με μεγαλύτερο κόστος αποστολής.

Κατηγορία 1 (0 - 10):

Πρόκειται για one-tailed έλεγχο. Για τον έλεγχο θα χρησιμοποιηθεί η μέθοδος z-statistics με επίπεδο σημαντικότητας 5% ($\alpha=0,05$).

$$H_0 : \mu_1 \geq \mu_2$$

$$H_1 : \mu_1 < \mu_2$$

$$s_{\bar{x}_1, \bar{x}_2} = \sqrt{\frac{s_1^2}{n_1} + \frac{s_2^2}{n_2}} = 0,93$$

$$z = \frac{(\bar{x}_1 - \bar{x}_2)}{s_{\bar{x}_1, \bar{x}_2}} = 0,019$$

Αφού το $z = 0,019 > z_{0,05} = -1,645$ δεν απορρίπτεται η μηδενική υπόθεση με αποτέλεσμα να μη μπορούμε να αποδείξουμε την εναλλακτική.

Κατηγορία 2 (11 - 50):

Πρόκειται για one-tailed έλεγχο. Για τον έλεγχο θα χρησιμοποιηθεί η μέθοδος z-statistics με επίπεδο σημαντικότητας 5% ($\alpha=0,05$).

$$H_0 : \mu_1 \geq \mu_2$$

$$H_1 : \mu_1 < \mu_2$$

$$s_{\bar{x}_1, \bar{x}_2} = \sqrt{\frac{s_1^2}{n_1} + \frac{s_2^2}{n_2}} = 1,31$$

$$z = \frac{(\bar{x}_1 - \bar{x}_2)}{s_{\bar{x}_1, \bar{x}_2}} = 2,31$$

Αφού το $z = 2,31 > z_{0,05} = -1,645$ δεν απορρίπτεται η μηδενική υπόθεση.

Κατηγορία 3 (51 - 100):

Πρόκειται για one-tailed έλεγχο. Για τον έλεγχο θα χρησιμοποιηθεί η μέθοδος t-statistics αφού πρόκειται για δείγμα μικρότερο του 30, με επίπεδο σημαντικότητας 5% ($\alpha=0,05$).

$$H_0 : \mu_1 \geq \mu_2$$

$$H_1 : \mu_1 < \mu_2$$

$$N_1 = 7$$

$$N_2 = 23$$

Για να βρεθεί το $t_{0,05}$ θα χρησιμοποιηθούν οι βαθμοί ελευθερίας του μικρότερου δείγματος δηλαδή $N_1 - 1 = 6$

$$s_{\bar{x}_1, \bar{x}_2} = \sqrt{\frac{s_1^2}{n_1} + \frac{s_2^2}{n_2}} = 3,18$$

$$t = \frac{(\bar{x}_1 - \bar{x}_2)}{s_{\bar{x}_1, \bar{x}_2}} = 0,95$$

Αφού το $t = 0,95 > t_{0,05,6} = -1,943$ δεν απορρίπτεται η μηδενική υπόθεση.

Κατηγορία 4 (>100):

Πρόκειται για one-tailed έλεγχο. Για τον έλεγχο θα χρησιμοποιηθεί η μέθοδος t-statistics αφού πρόκειται για δείγμα μικρότερο του 30, με επίπεδο σημαντικότητας 5% ($\alpha=0,05$).

$$H_0 : \mu_1 \geq \mu_2$$

$$H_1 : \mu_1 < \mu_2$$

$$N_1 = 12$$

$$N_2 = 18$$

Για να βρεθεί το $t_{0,05}$ θα χρησιμοποιηθούν οι βαθμοί ελευθερίας του μικρότερου δείγματος δηλαδή $N_1 - 1 = 11$

$$s_{\bar{x}_1, \bar{x}_2} = \sqrt{\frac{s_1^2}{n_1} + \frac{s_2^2}{n_2}} = 7,68$$
$$t = \frac{(\bar{x}_1 - \bar{x}_2)}{s_{\bar{x}_1, \bar{x}_2}} = 0,39$$

Αφού το $t = 0,39 > t_{0,05,11} = -1,796$ δεν απορρίπτεται η μηδενική υπόθεση.

Στην πρώτες κατηγορίες όπου οι τιμές είναι μέχρι 50 UKP, παρατηρείται ότι το κόστος αποστολής δεν παίζει ρόλο στην απόφαση του αγοραστή πιθανόν γιατί πρόκειται για αντικείμενα μικρά (computer accessories) με χαμηλό κόστος αποστολής και έτσι ο αγοραστής δε λαμβάνει υπόψη το κόστος αποστολής. Στις κατηγορίες άνω των 50 UKP φαίνεται ότι το κόστος αποστολής δεν παίζει ρόλο αφού πρόκειται για αντικείμενα με μεγαλύτερη αξία και αναλογικά το κόστος αποστολής είναι μικρό και δεν επηρεάζει την αγοραστική απόφαση.

4.1.7.5 Ανάλυση διασποράς κόστους αποστολής

Με την ανάλυση της διασποράς θα παρατηρήσουμε αν υπάρχει διαφοροποίηση των δεδομένων μεταξύ των κατηγοριών, δηλαδή κατά πόσο αλλάζει το κόστος αποστολής όταν αλλάζει η αξία του προϊόντος. Συγκεκριμένα, τα δεδομένα έχουν χωριστεί σε 4 κατηγορίες βάση της αξίας των προϊόντων της όπως φαίνεται πιο κάτω:

Κόστος προϊόντος (UKP)	Αριθμός εγγραφών
0-10	175
10-50	169
50-100	30
>100	30

Σημ. Στον αριθμό των εγγραφών συμπεριλαμβάνονται και αυτές που δεν αναφέρουν το κόστος αποστολής.

Anova: Single Factor

SUMMARY

Groups	Count	Sum	Average	Variance
CAT1	8	28,95	3,61875	6,054241
CAT2	18	169,48	9,415556	27,70305
CAT3	7	82,94	11,84857	34,13001
CAT4	10	175,46	17,546	90,1588

ANOVA

Source of Variation	SS	df	MS	F	P-value	F crit
Between Groups	908,5122	3	302,8374	7,721702	0,000363	2,845068
Within Groups	1529,541	39	39,219			
Total	2438,053	42				

Αναλύοντας τη διασπορά των δεδομένων μεταξύ των κατηγοριών παρατηρείται ότι υπάρχει διαφοροποίηση μεταξύ των μέσων όρων των προϊόντων που πωλήθηκαν, αφού το p-value είναι πολύ μικρότερο από 0,001 υποδεικνύοντας την ύπαρξη του μοντέλου.

Αυτό σημαίνει ότι ανάλογα με το πραγματικό κόστος του προϊόντος διαφοροποιείται και το κόστος αποστολής του προϊόντος έτσι ώστε να μπορέσει να πουληθεί.

Αν συγκρίνουμε τους μέσους όρους της κατηγορίας 1 (προϊόντα μέχρι 10 UKP) με την κατηγορία 4 (προϊόντα πάνω από 100 UKP) με έλεγχο υπόθεσης θα έχουμε:

$$H_0 : \mu_2 \leq \mu_1$$

$$H_1 : \mu_2 > \mu_1$$

$$s_{\bar{x}_1, \bar{x}_2} = \sqrt{\frac{s_1^2}{n_1} + \frac{s_2^2}{n_2}} = 3,13$$

$$z = \frac{(\bar{x}_2 - \bar{x}_1)}{s_{\bar{x}_1, \bar{x}_2}} = 4,46$$

Αφού το $z = 4,46 > z_{0,05} = 1,645$ απορρίπτεται η μηδενική υπόθεση. Φαίνεται ότι οι αγοραστές είναι διατεθειμένοι να πληρώσουν μεγαλύτερο κόστος αποστολής όσο αυξάνεται η αξία του προϊόντος αφού η αναλογία της αξίας και του κόστους αποστολής δεν είναι τόσο μεγάλη. Σε προϊόντα μικρής αξίας αν το κόστος αποστολής είναι υψηλό, πιθανόν να μην αξίζει να προχωρήσει ο καταναλωτής σε αγορά του προϊόντος.

4.1.8 Κατηγορία Games-Toys

4.1.8.1 Σχέση πώλησης με φυσική κατάσταση του προϊόντος

Σαν μηδενική υπόθεση θεωρούμε ότι πωλούνται περισσότερο τα χρησιμοποιημένα προϊόντα παρά τα καινούρια.

	New	%	Used	%	
Yes	166	0,75	39	0,62	205
No	54	0,25	24	0,38	78
	220	1,00	63	1,00	283

Πρόκειται για one-tailed έλεγχο. Για τον έλεγχο θα χρησιμοποιηθεί η μέθοδος z-statistics με επίπεδο σημαντικότητας 5% ($\alpha=0,05$).

$$H_0 : p_1 \leq p_2$$

$$H_1 : p_1 > p_2$$

$$p_c = \frac{\bar{x}_1 + \bar{x}_2}{n_1 + n_2} = 0,72$$

$$z = \frac{p_1 - p_2}{\sqrt{\frac{p_c(1-p_c)}{n_1} + \frac{p_c(1-p_c)}{n_2}}} = 2,12$$

Αφού το $z = 2,12 > z_{0,05} = 1,645$ απορρίπτεται η μηδενική υπόθεση και ισχύει η εναλλακτική. Μπορούμε να πούμε δηλαδή ότι το ποσοστό των καινούριων προϊόντων στην κατηγορία Games-Toys πωλούνται περισσότερο από το ποσοστό των χρησιμοποιημένων.

Παρατηρείται ότι οι αγοραστές προτιμούν τα καινούρια προϊόντα. Ένας πιθανός λόγος είναι ότι τα παιχνίδια αφορούν παιδιά άρα ο αγοραστής δε διακινδυνεύει την ασφάλεια του παιδιού και την ομαλή λειτουργία των παιχνιδιών με αγορά πιο φτηνών χρησιμοποιημένων παιχνιδιών. Αποτέλεσμα αυτού η αγορά καινούριων προϊόντων έστω και σε υψηλότερη τιμή.

4.1.8.2 Σχέση πώλησης με είδος του πωλητή

Σαν μηδενική υπόθεση θεωρούμε ότι πωλούνται περισσότερο τα αντικείμενα από ιδιώτες απ' ό,τι τα αντικείμενα από επιχειρήσεις.

	Business	%	Private	%	
Yes	18	0,62	187	0,74	205
No	11	0,38	67	0,26	78
	29	1	254	1	283

Πρόκειται για one-tailed έλεγχο. Για τον έλεγχο θα χρησιμοποιηθεί η μέθοδος z-statistics με επίπεδο σημαντικότητας 5% ($\alpha=0,05$).

$$H_0 : p_1 \leq p_2$$

$$H_1 : p_1 > p_2$$

$$p_c = \frac{\bar{x}_1 + \bar{x}_2}{n_1 + n_2} = 0,72$$

$$z = \frac{p_1 - p_2}{\sqrt{\frac{p_c(1 - p_c)}{n_1} + \frac{p_c(1 - p_c)}{n_2}}} = -1,32$$

Αφού το $z = -1,32 < z_{0,05} = 1,645$ δεν απορρίπτεται η μηδενική υπόθεση με αποτέλεσμα να μη μπορούμε να αποδείξουμε την εναλλακτική ότι δηλαδή το ποσοστό των προϊόντων που πωλούνται από επιχειρήσεις είναι μεγαλύτερο από το ποσοστό προϊόντων που πωλούνται από ιδιώτες.

Φαίνεται ότι ο παράγοντας τύπος του πωλητή δεν αποτελεί σημαντικό παράγοντα στην απόφαση ενός αγοραστή κατά πόσο θα αγοράσει ένα παιχνίδι ή όχι.

4.1.8.3 Σχέση πώλησης με τρόπο πώλησης

Σαν μηδενική υπόθεση θεωρούμε ότι πωλούνται περισσότερα παιχνίδια με την απευθείας πώληση παρά με τη μέθοδο της δημοπρασίας.

	Bid	%	Buy it Now	%	
Yes	187	0,87	18	0,26	205
No	27	0,13	51	0,74	78
	214	1	69	1	283

Πρόκειται για one-tailed έλεγχο. Για τον έλεγχο θα χρησιμοποιηθεί η μέθοδος z-statistics με επίπεδο σημαντικότητας 5% ($\alpha=0,05$).

$$H_0 : p_1 \leq p_2$$

$$H_1 : p_1 > p_2$$

$$p_c = \frac{\bar{x}_1 + \bar{x}_2}{n_1 + n_2} = 0,72$$

$$z = \frac{p_1 - p_2}{\sqrt{\frac{p_c(1 - p_c)}{n_1} + \frac{p_c(1 - p_c)}{n_2}}} = 9,91$$

Αφού το $z = 9,91 > z_{0,05} = 1,645$ απορρίπτεται η μηδενική υπόθεση και ισχύει η εναλλακτική. Μπορούμε δηλαδή να πούμε ότι το ποσοστό των προϊόντων που πωλούνται με δημοπρασία είναι μεγαλύτερο από το ποσοστό προϊόντων που υπάρχουν διαθέσιμα προς άμεση πώληση.

Φαίνεται ότι οι αγοραστές προτιμούν να αγοράζουν παιχνίδια μέσω δημοπρασίας παρά μέσω απευθείας αγοράς. Η προτίμηση αυτή οφείλεται στο γεγονός ότι σε μία δημοπρασία μπορεί ο αγοραστής να πετύχει καλύτερη τιμή αν δεν έχει μεγάλη ζήτηση το συγκεκριμένο αντικείμενο. Παράλληλα, επειδή είναι αντικείμενα που ο αγοραστής μπορεί να βρει εύκολα είναι διατεθειμένος να περιμένει την κατάλληλη ευκαιρία μέσω δημοπρασίας.

4.1.8.4 Σχέση πώλησης με κόστος αποστολής προϊόντος

Το εύρος τιμών της κατηγορίας Games-Toys είναι μεγάλο, με αποτέλεσμα αν γίνει έλεγχος υπόθεσης στο κόστος αποστολής

συνολικά για την κατηγορία τα αποτελέσματα δε θα είναι αντιπροσωπευτικά. Γι' αυτό το σκοπό έγινε κατηγοριοποίηση των προϊόντων με βάση το κόστος πώλησης τους. Για κάθε κατηγορία θα γίνει έλεγχος υπόθεσης σχετικά με το κόστος αποστολής. Η κατηγοριοποίηση έγινε ως ακολούθως:

Κόστος προϊόντος (UKP)	Αριθμός εγγραφών
0-10	208
10-50	57
50-100	8
>100	10

Σημ. Οι τιμές που εμπίπτουν στα όρια των κλάσεων συμπεριλαμβάνονται στην χαμηλότερη κλάση.

Σαν μηδενική υπόθεση θεωρούμε ότι πωλούνται τα παιχνίδια με μεγαλύτερο κόστος αποστολής.

Κατηγορία 1 (0 - 10):

Πρόκειται για one-tailed έλεγχο. Για τον έλεγχο θα χρησιμοποιηθεί η μέθοδος z-statistics με επίπεδο σημαντικότητας 5% ($\alpha=0,05$).

$$H_0 : \mu_1 \geq \mu_2$$

$$H_1 : \mu_1 < \mu_2$$

$$s_{\bar{x}_1, \bar{x}_2} = \sqrt{\frac{s_1^2}{n_1} + \frac{s_2^2}{n_2}} = 0,49$$

$$z = \frac{(\bar{x}_1 - \bar{x}_2)}{s_{\bar{x}_1, \bar{x}_2}} = -1,52$$

Αφού το $z = -1,52 > z_{0,05} = -1,645$ δεν απορρίπτεται η μηδενική υπόθεση με αποτέλεσμα να μη μπορούμε να αποδείξουμε την εναλλακτική.

Κατηγορία 2 (11 - 50):

Πρόκειται για one-tailed έλεγχο. Για τον έλεγχο θα χρησιμοποιηθεί η μέθοδος t-statistics αφού πρόκειται για δείγμα μικρότερο του 30, με επίπεδο σημαντικότητας 5% ($\alpha=0,05$).

$$H_0 : \mu_1 \geq \mu_2$$

$$H_1 : \mu_1 < \mu_2$$

$$N_1 = 29$$

$$N_2 = 28$$

Για να βρεθεί το $t_{0,05}$ θα χρησιμοποιηθούν οι βαθμοί ελευθερίας του μικρότερου δείγματος δηλαδή $N_2 - 1 = 27$

$$s_{\bar{x}_1, \bar{x}_2} = \sqrt{\frac{s_1^2}{n_1} + \frac{s_2^2}{n_2}} = 0,64$$

$$t = \frac{(\bar{x}_1 - \bar{x}_2)}{s_{\bar{x}_1, \bar{x}_2}} = 3,49$$

Αφού το $t = 3,49 > t_{0,05,27} = -1,701$ δεν απορρίπτεται η μηδενική υπόθεση.

Κατηγορία 3 (51 - 100):

Δεν υπάρχουν αρκετά δεδομένα έτσι ώστε να γίνει έλεγχος υπόθεσης αφού το δείγμα μας αποτελείται από 8 μόνο εγγραφές, εκ των οποίων 4 αφορούν τα προϊόντα που πωλήθηκαν και 4 αυτά που δεν πωλήθηκαν.

Κατηγορία 4 (>100):

Δεν υπάρχουν αρκετά δεδομένα έτσι ώστε να γίνει έλεγχος υπόθεσης αφού το δείγμα των προϊόντων που πωλήθηκαν αποτελείται μόνο 2 εγγραφές.

Για τις κατηγορίες που υπήρχε επάρκεια δεδομένων, παρατηρείται ότι ο αγοραστής δεν επηρεάζεται από το κόστος αποστολής και λαμβάνει υπόψη άλλους παράγοντες. Αυτό πιθανό να οφείλεται στο γεγονός ότι υπάρχει ανομοιογένεια στα προϊόντα της κατηγορίας και ο αγοραστής δεν έχει πολλές εναλλακτικές για το ίδιο αντικείμενο. Έτσι αν θελήσει να αγοράσει ένα συγκεκριμένο αντικείμενο της κατηγορίας πρέπει να πληρώσει το όποιο κόστος αποστολής.

4.1.8.5 Ανάλυση διασποράς κόστους αποστολής

Με την ανάλυση της διασποράς θα παρατηρήσουμε αν υπάρχει διαφοροποίηση των δεδομένων μεταξύ των κατηγοριών, δηλαδή κατά πόσο αλλάζει το κόστος αποστολής όταν αλλάζει η αξία του προϊόντος. Συγκεκριμένα, τα δεδομένα έχουν χωριστεί σε 4 κατηγορίες βάση της αξίας των προϊόντων της όπως φαίνεται πιο κάτω:

Κόστος προϊόντος (UKP)	Αριθμός εγγραφών
0-10	208
10-50	57
50-100	8
>100	10

Σημ. Στον αριθμό των εγγραφών συμπεριλαμβάνονται και αυτές που δεν αναφέρουν το κόστος αποστολής.

Anova: Single Factor

SUMMARY

Groups	Count	Sum	Average	Variance
CAT1	170	536,43	3,155471	1,81327
CAT2	29	151,05	5,208621	9,594512

ANOVA

Source of Variation	SS	df	MS	F	P-value	F crit
---------------------	----	----	----	---	---------	--------

Between Groups	104,4324	1	104,4324	35,77391	1,03E-08	3,889096
Within Groups	575,089	197	2,919233			
Total	679,5214	198				

Αναλύοντας τη διασπορά των δεδομένων μεταξύ των κατηγοριών παρατηρείται ότι υπάρχει διαφοροποίηση μεταξύ των μέσων όρων των προϊόντων που πωλήθηκαν, αφού το p-value είναι πολύ μικρότερο από 0,001 υποδεικνύοντας την ύπαρξη του μοντέλου. Αυτό σημαίνει ότι ανάλογα με το πραγματικό κόστος του προϊόντος διαφοροποιείται και το κόστος αποστολής του προϊόντος έτσι ώστε να μπορέσει να πουληθεί.

Αν συγκρίνουμε τους μέσους όρους της κατηγορίας 1 (προϊόντα κάτω των 10 UKP) με την κατηγορία 2 (προϊόντα από 11 μέχρι 50) με έλεγχο υπόθεσης θα έχουμε:

$$H_0 : \mu_2 \leq \mu_1$$

$$H_1 : \mu_2 > \mu_1$$

$$s_{\bar{x}_1, \bar{x}_2} = \sqrt{\frac{s_1^2}{n_1} + \frac{s_2^2}{n_2}} = 0,58$$

$$z = \frac{(\bar{x}_2 - \bar{x}_1)}{s_{\bar{x}_1, \bar{x}_2}} = 3,51$$

Αφού το $z = 3,51 > z_{0,05} = 1,645$ απορρίπτεται η μηδενική υπόθεση. Φαίνεται ότι οι αγοραστές είναι διατεθειμένοι να πληρώσουν μεγαλύτερο κόστος αποστολής όσο αυξάνεται η αξία του προϊόντος αφού η αναλογία της αξίας και του κόστους αποστολής δεν είναι τόσο μεγάλη. Σε προϊόντα μικρής αξίας αν το κόστος αποστολής είναι υψηλό, πιθανόν να μην αξίζει να προχωρήσει ο καταναλωτής σε αγορά του προϊόντος.

4.1.9 Κατηγορία Clothes

4.1.9.1 Σχέση πώλησης με φυσική κατάσταση του προϊόντος

Σαν μηδενική υπόθεση θεωρούμε ότι πωλούνται περισσότερα τα χρησιμοποιημένα ρούχα απ' ότι τα καινούρια.

	New	%	Used	%	
Yes	20	0,41	10	0,10	30
No	29	0,59	89	0,90	118
	49	1,00	99	1,00	148

Πρόκειται για one-tailed έλεγχο. Για τον έλεγχο θα χρησιμοποιηθεί η μέθοδος z-statistics με επίπεδο σημαντικότητας 5% ($\alpha=0,05$).

$$H_0 : p_1 \leq p_2$$

$$H_1 : p_1 > p_2$$

$$p_c = \frac{\bar{x}_1 + \bar{x}_2}{n_1 + n_2} = 0,20$$

$$z = \frac{p_1 - p_2}{\sqrt{\frac{p_c(1 - p_c)}{n_1} + \frac{p_c(1 - p_c)}{n_2}}} = 4,37$$

Αφού το $z = 4,37 > z_{0,05} = 1,645$ απορρίπτεται η μηδενική υπόθεση και ισχύει η εναλλακτική. Μπορούμε να πούμε δηλαδή ότι το ποσοστό των καινούριων προϊόντων στην κατηγορία Clothes πωλούνται περισσότερο από το ποσοστό των χρησιμοποιημένων.

Φαίνεται ότι οι αγοραστές προτιμούν να αγοράζουν καινούρια ρούχα, πιθανόν για λόγους υγιεινής αλλά και φθοράς. Τα ρούχα είναι αντικείμενα που φθείρονται εύκολα και μπορεί οι φθορές να μη φαίνονται οπτικά (μέσω των φωτογραφιών που παρέχονται). Αυτό έχει ως αποτέλεσμα οι αγοραστές να προτιμούν να αγοράζουν καινούρια προϊόντα έστω και σε υψηλότερη τιμή.

4.1.9.2 Σχέση πώλησης με είδος του πωλητή

Σαν μηδενική υπόθεση θεωρούμε ότι το ποσοστό των ρούχων που πωλούνται από ιδιώτες είναι μεγαλύτερο από το ποσοστό ρούχων που πωλούνται από επιχειρήσεις.

	Business	%	Private	%	
Yes	27	0,55	3	0,03	30
No	22	0,45	96	0,97	118
	49	1	99	1	148

Πρόκειται για one-tailed έλεγχο. Για τον έλεγχο θα χρησιμοποιηθεί η μέθοδος z-statistics με επίπεδο σημαντικότητας 5% ($\alpha=0,05$).

$$H_0 : p_1 \leq p_2$$

$$H_1 : p_1 > p_2$$

$$p_c = \frac{\bar{x}_1 + \bar{x}_2}{n_1 + n_2} = 0,20$$

$$z = \frac{p_1 - p_2}{\sqrt{\frac{p_c(1-p_c)}{n_1} + \frac{p_c(1-p_c)}{n_2}}} = 7,41$$

Αφού το $z7,41 < z_{0,05} = 1,645$ απορρίπτεται η μηδενική υπόθεση με αποτέλεσμα να ισχύει η εναλλακτική. Μπορούμε δηλαδή να πούμε ότι το ποσοστό των προϊόντων που πωλούνται από επιχειρήσεις είναι μεγαλύτερο από το ποσοστό προϊόντων που πωλούνται από ιδιώτες.

Ο τύπος του πωλητή αποτελεί παράγοντα που επηρεάζει την αγοραστική απόφαση. Οι αγοραστές προτιμούν να αγοράζουν από επιχειρήσεις αφού έτσι κερδίζουν καλύτερη ποιότητα και αποφεύγουν τα μη αυθεντικά προϊόντα.

4.1.9.3 Σχέση πώλησης με τρόπο πώλησης

Σαν μηδενική υπόθεση θεωρούμε ότι το ποσοστό των ρούχων που πωλείται μέσω απευθείας πώλησης είναι μεγαλύτερο από το ποσοστό που πωλείται μέσω δημοπρασίας.

	Bid	%	Buy it Now	%	
Yes	14	0,13	16	0,38	30
No	92	0,87	26	0,62	118
	106	1	42	1	148

Πρόκειται για one-tailed έλεγχο. Για τον έλεγχο θα χρησιμοποιηθεί η μέθοδος z-statistics με επίπεδο σημαντικότητας 5% ($\alpha=0,05$).

$$H_0 : p_1 \leq p_2$$

$$H_1 : p_1 > p_2$$

$$p_c = \frac{\bar{x}_1 + \bar{x}_2}{n_1 + n_2} = 0,20$$

$$z = \frac{p_1 - p_2}{\sqrt{\frac{p_c(1-p_c)}{n_1} + \frac{p_c(1-p_c)}{n_2}}} = -3,40$$

Αφού το $z = -3,40 < z_{0,05} = 1,645$ δεν απορρίπτεται η μηδενική υπόθεση με αποτέλεσμα να μη μπορούμε να αποδείξουμε την εναλλακτική.

Ο τρόπος πώλησης δεν φαίνεται να αποτελεί παράγοντα που επηρεάζει τον αγοραστή κατά πόσο θα αγοράσει ή όχι ένα ρούχο.

4.1.9.4 Σχέση πώλησης με κόστος αποστολής προϊόντος

Το εύρος τιμών της κατηγορίας Clothes είναι μεγάλο, με αποτέλεσμα αν γίνει έλεγχος υπόθεσης στο κόστος αποστολής συνολικά για την κατηγορία τα αποτελέσματα δε θα είναι αντιπροσωπευτικά. Γι' αυτό το σκοπό έγινε κατηγοριοποίηση των

προϊόντων με βάση τη τιμή πώλησης τους. Για κάθε κατηγορία θα γίνει έλεγχος υπόθεσης σχετικά με το κόστος αποστολής. Η κατηγοριοποίηση έγινε ως ακολούθως:

Κόστος προϊόντος (UKP)	Αριθμός εγγραφών
0-10	104
10-50	39
50-100	5

Σημ. Οι τιμές που εμπίπτουν στα όρια των κλάσεων συμπεριλαμβάνονται στην χαμηλότερη κλάση.

Σαν μηδενική υπόθεση θεωρούμε ότι πωλούνται τα ρούχα με μεγαλύτερο κόστος αποστολής.

Κατηγορία 1 (0 - 10):

Πρόκειται για one-tailed έλεγχο. Για τον έλεγχο θα χρησιμοποιηθεί η μέθοδος z-statistics με επίπεδο σημαντικότητας 5% ($\alpha=0,05$).

$$H_0 : \mu_1 \geq \mu_2$$

$$H_1 : \mu_1 < \mu_2$$

$$s_{\bar{x}_1, \bar{x}_2} = \sqrt{\frac{s_1^2}{n_1} + \frac{s_2^2}{n_2}} = 0,55$$

$$z = \frac{(\bar{x}_1 - \bar{x}_2)}{s_{\bar{x}_1, \bar{x}_2}} = -2,67$$

Αφού το $z = -2,67 < z_{0,05} = -1,645$ απορρίπτεται η μηδενική υπόθεση και ισχύει η εναλλακτική ότι δηλαδή στην κατηγορία με ρούχα αξία μέχρι 10 UKP πωλούνται αυτά με το χαμηλότερο κόστος.

Κατηγορία 2 (11 - 50):

Πρόκειται για one-tailed έλεγχο. Για τον έλεγχο θα χρησιμοποιηθεί η μέθοδος t-statistics αφού πρόκειται για δείγμα μικρότερο του 30, με επίπεδο σημαντικότητας 5% ($\alpha=0,05$).

$$H_0 : \mu_1 \geq \mu_2$$

$$H_1 : \mu_1 < \mu_2$$

$$N_1 = 20$$

$$N_2 = 19$$

Για να βρεθεί το $t_{0,05}$ θα χρησιμοποιηθούν οι βαθμοί ελευθερίας του μικρότερου δείγματος δηλαδή $N_2 - 1 = 18$

$$s_{\bar{x}_1, \bar{x}_2} = \sqrt{\frac{s_1^2}{n_1} + \frac{s_2^2}{n_2}} = 1,06$$

$$t = \frac{(\bar{x}_1 - \bar{x}_2)}{s_{\bar{x}_1, \bar{x}_2}} = 5,23$$

Αφού το $t = 5,23 > t_{0,05,18} = -1,734$ δεν απορρίπτεται η μηδενική υπόθεση.

Κατηγορία 3 (51 - 100):

Δεν υπάρχουν αρκετά δεδομένα έτσι ώστε να γίνει έλεγχος υπόθεσης αφού στο δείγμα μας δεν υπάρχει κανένα προϊόν που να έχει πωληθεί.

Για τις κατηγορίες που υπήρχε επάρκεια δεδομένων, παρατηρείται ότι στην κατηγορία μέχρι 10 UKP ο αγοραστής επηρεάζεται από το κόστος αποστολής αφού αν αυτό γίνει δυσανάλογο της αξίας δεν αξίζει να αγοραστεί. Αυτό οφείλεται στο γεγονός ότι υπάρχουν πολλές εναλλακτικές επιλογές στην κατηγορία των ρούχων και έτσι ο αγοραστής μπορεί να επιλέξει αυτές με χαμηλό κόστος αποστολής. Στην κατηγορία 10-50 UKP ο αγοραστής δε φαίνεται

να επηρεάζεται από το κόστος αποστολής αφού όσο ανεβαίνει η αξία των προϊόντος τόσο πιο εξειδικευμένη γίνεται η αγορά και ο αγοραστής ψάχνει κάτι συγκεκριμένο.

4.1.9.5 Ανάλυση διασποράς κόστους αποστολής

Με την ανάλυση της διασποράς θα παρατηρήσουμε αν υπάρχει διαφοροποίηση των δεδομένων μεταξύ των κατηγοριών, δηλαδή κατά πόσο αλλάζει το κόστος αποστολής όταν αλλάζει η αξία του προϊόντος. Συγκεκριμένα, τα δεδομένα έχουν χωριστεί σε 3 κατηγορίες βάση της αξίας των προϊόντων της όπως φαίνεται πιο κάτω:

Κόστος προϊόντος (UKP)	Αριθμός εγγραφών
0-10	104
10-50	39
50-100	5

Σημ. Στον αριθμό των εγγραφών συμπεριλαμβάνονται και αυτές που δεν αναφέρουν το κόστος αποστολής. Οι τιμές που εμπίπτουν στα όρια των κλάσεων συμπεριλαμβάνονται στην χαμηλότερη κλάση.

Anova: Single Factor

SUMMARY

Groups	Count	Sum	Average	Variance
CAT1	10	35,87	3,587	2,913846
CAT2	20	233,92	11,696	15,69645

ANOVA

Source of Variation	SS	df	MS	F	P-value	F crit
Between Groups	438,3725	1	438,3725	37,83068	1,22E-06	4,195972
Within Groups	324,4571	28	11,58775			
Total	762,8296	29				

Αναλύοντας τη διασπορά των δεδομένων μεταξύ των κατηγοριών παρατηρείται ότι υπάρχει διαφοροποίηση μεταξύ των μέσων όρων των προϊόντων που πωλήθηκαν, αφού το p-value είναι πολύ μικρότερο από 0,001 υποδεικνύοντας την ύπαρξη του μοντέλου. Αυτό σημαίνει ότι ανάλογα με το πραγματικό κόστος του προϊόντος διαφοροποιείται και το κόστος αποστολής του προϊόντος έτσι ώστε να μπορέσει να πουληθεί.

Αν συγκρίνουμε τους μέσους όρους της κατηγορίας 1 (προϊόντα κάτω των 10 UKP) με την κατηγορία 2 (προϊόντα από 11 μέχρι 50) με έλεγχο υπόθεσης θα έχουμε:

$$H_0 : \mu_2 \leq \mu_1$$

$$H_1 : \mu_2 > \mu_1$$

$$s_{\bar{x}_1, \bar{x}_2} = \sqrt{\frac{s_1^2}{n_1} + \frac{s_2^2}{n_2}} = 1,04$$

$$z = \frac{(\bar{x}_2 - \bar{x}_1)}{s_{\bar{x}_1, \bar{x}_2}} = 7,81$$

Αφού το $z = 7,81 > z_{0,05} = 1,645$ απορρίπτεται η μηδενική υπόθεση. Φαίνεται ότι οι αγοραστές είναι διατεθειμένοι να πληρώσουν μεγαλύτερο κόστος αποστολής όσο αυξάνεται η αξία του προϊόντος αφού η αναλογία της αξίας και του κόστους αποστολής δεν είναι τόσο μεγάλη. Σε προϊόντα μικρής αξίας αν το κόστος αποστολής είναι υψηλό, πιθανόν να μην αξίζει να προχωρήσει ο καταναλωτής σε αγορά του προϊόντος.

4.1.10 Κατηγορία DVDs

4.1.10.1 Σχέση πώλησης με φυσική κατάσταση του προϊόντος

Σαν μηδενική υπόθεση θεωρούμε ότι το ποσοστό των χρησιμοποιημένων DVDs που χρησιμοποιούνται είναι μεγαλύτερο από το ποσοστό των καινούριων.

	New	%	Used	%	
Yes	11	0,41	14	0,13	25
No	16	0,59	93	0,87	109
	27	1,00	107	1,00	134

Πρόκειται για one-tailed έλεγχο. Για τον έλεγχο θα χρησιμοποιηθεί η μέθοδος z-statistics με επίπεδο σημαντικότητας 5% ($\alpha=0,05$).

$$H_0 : p_1 \leq p_2$$

$$H_1 : p_1 > p_2$$

$$p_c = \frac{\bar{x}_1 + \bar{x}_2}{n_1 + n_2} = 0,19$$

$$z = \frac{p_1 - p_2}{\sqrt{\frac{p_c(1-p_c)}{n_1} + \frac{p_c(1-p_c)}{n_2}}} = 3,30$$

Αφού το $z = 3,30 > z_{0,05} = 1,645$ απορρίπτεται η μηδενική υπόθεση και ισχύει η εναλλακτική. Μπορούμε να πούμε δηλαδή ότι το ποσοστό των καινούριων προϊόντων στην κατηγορία DVDs πωλούνται περισσότερο από το ποσοστό των χρησιμοποιημένων.

Φαίνεται ότι η φυσική κατάσταση ενός DVD αποτελεί παράγοντα στην απόφαση του αγοραστή. Τα DVDs είναι προϊόντα χαμηλής αξίας και επιπλέον μπορεί να έχουν φθορές που να μη φαίνονται οπτικά μέσω φωτογραφιών. Αυτοί οι δύο παράγοντες οδηγούν στην αγορά καινούριων DVD αφού η αγορά καινούριων DVDs δεν έχει

μεγάλη διάφορα σε σχέση με τα χρησιμοποιημένα ενώ παράλληλα εξασφαλίζεται η ποιότητα του.

4.1.10.2 Σχέση πώλησης με είδος του πωλητή

Σαν μηδενική υπόθεση θεωρούμε ότι το ποσοστό των DVDs που πωλούνται από ιδιώτες είναι μεγαλύτερο από αυτό που πωλούνται από επιχειρήσεις.

	Business	%	Private	%	
Yes	10	0,48	15	0,13	25
No	11	0,52	98	0,87	109
	21	1	113	1	134

Πρόκειται για one-tailed έλεγχο. Για τον έλεγχο θα χρησιμοποιηθεί η μέθοδος z-statistics με επίπεδο σημαντικότητας 5% ($\alpha=0,05$).

$$H_0 : p_1 \leq p_2$$

$$H_1 : p_1 > p_2$$

$$p_c = \frac{\bar{x}_1 + \bar{x}_2}{n_1 + n_2} = 0,19$$

$$z = \frac{p_1 - p_2}{\sqrt{\frac{p_c(1 - p_c)}{n_1} + \frac{p_c(1 - p_c)}{n_2}}} = 3,71$$

Αφού το $z = 3,71 > z_{0,05} = 1,645$ απορρίπτεται η μηδενική υπόθεση με αποτέλεσμα να ισχύει η εναλλακτική. Μπορούμε δηλαδή να πούμε ότι το ποσοστό των προϊόντων που πωλούνται από επιχειρήσεις είναι μεγαλύτερο από το ποσοστό προϊόντων που πωλούνται από ιδιώτες.

Ο τύπος του πωλητή αποτελεί παράγοντα που επηρεάζει την αγοραστική απόφαση. Οι αγοραστές προτιμούν να αγοράζουν από επιχειρήσεις αφού έτσι νιώθουν μεγαλύτερη ασφάλεια για την ποιότητα των DVDs και αποφεύγουν τα μη αυθεντικά προϊόντα.

4.1.10.3 Σχέση πώλησης με τρόπο πώλησης

Σαν μηδενική υπόθεση θεωρούμε ότι το ποσοστό των DVDs που πωλούνται μέσω απευθείας πώλησης είναι μεγαλύτερο από αυτό που πωλούνται μέσω δημοπρασίας.

	Bid	%	Buy it Now	%	
Yes	22	0,18	3	0,33	25
No	103	0,82	6	0,67	109
	125	1	9	1	134

Πρόκειται για one-tailed έλεγχο. Για τον έλεγχο θα χρησιμοποιηθεί η μέθοδος z-statistics με επίπεδο σημαντικότητας 5% ($\alpha=0,05$).

$$H_0 : p_1 \leq p_2$$

$$H_1 : p_1 > p_2$$

$$p_c = \frac{\bar{x}_1 + \bar{x}_2}{n_1 + n_2} = 0,19$$

$$z = \frac{p_1 - p_2}{\sqrt{\frac{p_c(1-p_c)}{n_1} + \frac{p_c(1-p_c)}{n_2}}} = -1,17$$

Αφού το $z = -1,17 < z_{0,05} = 1,645$ δεν απορρίπτεται η μηδενική υπόθεση με αποτέλεσμα να μη μπορούμε να αποδείξουμε την εναλλακτική.

Ο τρόπος πώλησης δεν φαίνεται να αποτελεί παράγοντα που επηρεάζει τον αγοραστή κατά πόσο θα αγοράσει ή όχι ένα DVD.

4.1.10.4 Σχέση πώλησης με κόστος αποστολής προϊόντος

Το εύρος τιμών της κατηγορίας DVDs διαφέρει, με αποτέλεσμα αν γίνει έλεγχος υπόθεσης στο κόστος αποστολής συνολικά για την κατηγορία τα αποτελέσματα δε θα είναι αντιπροσωπευτικά. Γι' αυτό το σκοπό έγινε κατηγοριοποίηση των προϊόντων με βάση το κόστος

πώλησης τους. Για κάθε κατηγορία θα γίνει έλεγχος υπόθεσης σχετικά με το κόστος αποστολής. Η κατηγοριοποίηση έγινε ως ακολούθως:

Κόστος προϊόντος (UKP)	Αριθμός εγγραφών
0-10	129
10-50	5

Σαν μηδενική υπόθεση θεωρούμε ότι πωλούνται τα DVDs με μεγαλύτερο κόστος αποστολής.

Κατηγορία 1 (0 - 10):

Πρόκειται για one-tailed έλεγχο. Για τον έλεγχο θα χρησιμοποιηθεί η μέθοδος z-statistics με επίπεδο σημαντικότητας 5% ($\alpha=0,05$).

$$H_0 : \mu_1 \geq \mu_2$$

$$H_1 : \mu_1 < \mu_2$$

$$s_{\bar{x}_1, \bar{x}_2} = \sqrt{\frac{s_1^2}{n_1} + \frac{s_2^2}{n_2}} = 0,50$$

$$z = \frac{(\bar{x}_1 - \bar{x}_2)}{s_{\bar{x}_1, \bar{x}_2}} = -4,70$$

Αφού το $z = -4,70 < z_{0,05} = -1,645$ απορρίπτεται η μηδενική υπόθεση και ισχύει η εναλλακτική.

Κατηγορία 2 (11 - 50):

Δεν υπάρχουν αρκετά δεδομένα έτσι ώστε να γίνει έλεγχος υπόθεσης αφού στο δείγμα μας δεν υπάρχει κανένα προϊόν που να έχει πωληθεί.

Το κόστος αποστολής φαίνεται να αποτελεί παράγοντα στην απόφαση του αγοραστή κατά πόσο θα αγοράσει ένα DVD ή όχι

αφού το μέγεθος και το βάρος είναι σταθερό με αποτέλεσμα ο αγοραστής να προτιμά τα προϊόντα αυτά που έχουν χαμηλότερο το κόστος αποστολής.

4.2 Binary Logistic Regression

Τα περισσότερα από τα δεδομένα που συλλέχτηκαν αφορούν χαρακτηριστικά με δυαδικές, διακριτές τιμές και όχι συνεχείς. Συγκεκριμένα τα χαρακτηριστικά Sold, ProductCondition, SellerType, BuyMethod παίρνουν διακριτές τιμές και τα χαρακτηριστικά Price και ShippingCost συνεχείς. Για το σκοπό της παλινδρόμησης και δημιουργίας μοντέλου, τα χαρακτηριστικά με διακριτές τιμές κωδικοποιούνται ως ακολούθως:

	VALUES	
	0	1
Sold	No	Yes
ProductCondition	Used	New
SellerType	Private	Business
BuyMethod	Buy it Now	Bid

Δεδομένου ότι τα περισσότερα δεδομένα παίρνουν διακριτές τιμές θα χρησιμοποιηθεί binary logistic regression αντί της απλής γραμμικής παλινδρόμησης, έτσι ώστε το μοντέλο να δίνει αποτελέσματα και προβλέψεις πιο κοντά στην πραγματικότητα.

Η ανάλυση των δεδομένων γίνεται ξεχωριστά για κάθε κατηγορία αφού απευθύνονται σε διαφορετικό κοινό, άρα οι λόγοι που επηρεάζουν την απόφαση του καταναλωτή πιθανόν να είναι διαφορετικοί. Σε όλες τις κατηγορίες θα εξεταστεί η εξαρτημένη μεταβλητή της πώλησης (Sold) σε σχέση με τις ανεξάρτητες μεταβλητές της φυσικής κατάστασης του προϊόντος, του είδους του πωλητή, του τρόπου πώλησης, της τιμής πώλησης του προϊόντος και του κόστους αποστολής.

4.2.1 Κατηγορία Collectables

Link Function: Logit

Response Information

Variable	Value	Count	
SoldB	1	562	(Event)
	0	817	
	Total	1379	

Logistic Regression Table

Predictor	Coef	SE Coef	Z	P	Odds Ratio
Constant	-1,51174	0,221112	-6,84	0,000	
SellerTypeB					
1	0,180287	0,134409	1,34	0,180	1,20
ProductConditionB					
1	0,898318	0,150390	5,97	0,000	2,46
BuyMethodB					
1	0,901871	0,186844	4,83	0,000	2,46
ShippingCost	-0,0029827	0,0071039	-0,42	0,675	1,00
Price	0,0038814	0,0017466	2,22	0,026	1,00

Log-Likelihood = -911,574

Test that all slopes are zero: G = 41,125, DF = 5, P-Value = 0,000

Goodness-of-Fit Tests

Method	Chi-Square	DF	P
Pearson	1000,16	599	0,000
Deviance	1289,62	599	0,000
Hosmer-Lemeshow	101,71	8	0,000

Measures of Association:

(Between the Response Variable and Predicted Probabilities)

Pairs	Number	Percent	Summary Measures
Concordant	272262	59,3	Somers' D 0,23
Discordant	166537	36,3	Goodman-Kruskal Gamma 0,24
Ties	20355	4,4	Kendall's Tau-a 0,11
Total	459154	100,0	

Σημ. Στο λογισμικό πρόγραμμα Minitab αν το p-value είναι μικρότερο από 0,001 τότε παρουσιάζεται σαν 0,000.

Αφού τρέξαμε το μοντέλο με όλες τις ανεξάρτητες μεταβλητές παρατηρούμε ότι αυτές που επηρεάζουν την εξαρτημένη είναι η κατάσταση του προϊόντος, η μέθοδος με την οποία γίνεται η πώληση και η τιμή πώλησης δεδομένου του πολύ χαμηλού p-value που έχουν οι συγκεκριμένες μεταβλητές. Το πολύ χαμηλό p-value υποδεικνύει ύπαρξη του μοντέλου, ύπαρξη δηλαδή σχέση μεταξύ της εξαρτημένης και της ανεξάρτητης μεταβλητής. Το p-value του μοντέλου είναι επίσης μικρότερο από 0,001 δείχνοντας έτσι την ύπαρξη μοντέλου.

Στον συντελεστή προσδιορισμού παρατηρούμε ότι το ποσοστό είναι στο 60% που υποδηλώνει την ύπαρξη ενός καλού μοντέλου. Αυτό σαν γεγονός μπορεί να ερμηνευθεί ότι η πώληση ενός προϊόντος της κατηγορίας Collectable εξαρτάται από τις μεταβλητές κατάσταση του προϊόντος, μέθοδος με την οποία γίνεται η πώληση και τιμή πώλησης.

Η εξίσωση παλινδρόμησης, έχει ως ακολούθως:

$$p(x) = \frac{1}{1 + e^{-(a + \beta_1 x_1 + \beta_2 x_2 + \dots + \beta_n x_n)}}$$

Με βάση το μοντέλο μας και λαμβάνοντας υπόψη μόνο τις ανεξάρτητες μεταβλητές με p-value μικρότερο από 0,01 έχουμε την εξής εξίσωση:

$$Sold = \frac{1}{1 + e^{-(-1,51 + 0,90(ProductCondition) + 0,90(BuyMethod) + 0,004(Price))}}$$

Αν η μέθοδος πώλησης είναι με δημοπρασία (bid) και έχει τη τιμή 1 τότε η τιμή του Sold θα τείνει προς το 1 δηλαδή έχει περισσότερες πιθανότητες να πουληθεί. Αντίστοιχα, αν το προϊόν είναι καινούριο, το Sold θα τείνει προς το 1 και η πιθανότητα να πουληθεί αυξάνεται. Όσο αφορά τη τιμή παρατηρούμε ότι ο συντελεστής β είναι πολύ μικρός με αποτέλεσμα να μη διαφοροποιείται αισθητά το αποτέλεσμα αν αυξηθεί ή μειωθεί η τιμή του προϊόντος.

Με βάση τα πιο πάνω μπορούμε να κάνουμε πρόβλεψη ότι τα προϊόντα της κατηγορίας Collectables πωλούνται με μεγαλύτερη ευκολία αν είναι καινούρια και πωλούνται μέσω δημοπρασίας.

4.2.2 Κατηγορία Coins

Link Function: Logit

Response Information

Variable	Value	Count	
SoldB	1	1117	(Event)
	0	88	
	Total	1205	

* NOTE * 1205 cases were used
 * NOTE * 21 cases contained missing values

Logistic Regression Table

Predictor	Coef	SE Coef	Z	P	Odds Ratio
Constant	-0,291492	0,261265	-1,12	0,265	
SellerTypeB					
1	1,95781	0,268393	7,29	0,000	7,08
BuyMethodB					
1	2,41770	0,289075	8,36	0,000	11,22
ShippingCost	-0,0001433	0,0568523	-0,00	0,998	1,00
Price	0,0000781	0,0013648	0,06	0,954	1,00

Log-Likelihood = -237,826
 Test that all slopes are zero: G = 154,333, DF = 4, P-Value = 0,000

Goodness-of-Fit Tests

Method	Chi-Square	DF	P
Pearson	808,014	686	0,001
Deviance	343,964	686	1,000
Hosmer-Lemeshow	66,442	8	0,000

Measures of Association:

(Between the Response Variable and Predicted Probabilities)

Pairs	Number	Percent	Summary Measures
Concordant	73900	75,2	Somers' D 0,68
Discordant	7014	7,1	Goodman-Kruskal Gamma 0,83
Ties	17382	17,7	Kendall's Tau-a 0,09
Total	98296	100,0	

Σημ. Στο λογισμικό πρόγραμμα Minitab αν το p-value είναι μικρότερο από 0,001 τότε παρουσιάζεται σαν 0,000.

Τρέξαμε το μοντέλο με όλες τις ανεξάρτητες μεταβλητές εκτός της κατάστασης του προϊόντος (ProductCondition) αφού σε όλες τις περιπτώσεις παίρνει τη τιμή 0. Παρατηρούμε ότι αυτές που επηρεάζουν την εξαρτημένη είναι ο τύπος του πωλητή και η μέθοδος με την οποία γίνεται η πώληση, δεδομένου του πολύ χαμηλού p-value που έχουν οι συγκεκριμένες μεταβλητές. Το πολύ χαμηλό p-value υποδεικνύει ύπαρξη του μοντέλου, ύπαρξη δηλαδή σχέση μεταξύ της εξαρτημένης και της ανεξάρτητης μεταβλητής. Το p-value του μοντέλου είναι επίσης μικρότερο από 0,001 δείχνοντας έτσι την ύπαρξη μοντέλου.

Στον συντελεστή προσδιορισμού παρατηρούμε ότι το ποσοστό είναι στο 75,2% που υποδηλώνει την ύπαρξη ενός καλού μοντέλου. Αυτό σαν γεγονός μπορεί να ερμηνευθεί ότι η πώληση ενός προϊόντος της

κατηγορίας Coins εξαρτάται από τον τύπο του πωλητή και τη μέθοδο της πώλησης. Η εξίσωση παλινδρόμησης, έχει ως ακολούθως:

$$y = \frac{1}{1 + e^{-(a + \beta_1 x_1 + \beta_2 x_2 + \dots + \beta_n x_n)}}$$

Με βάση το μοντέλο μας και λαμβάνοντας υπόψη μόνο τις ανεξάρτητες μεταβλητές με p-value μικρότερο από 0,01 έχουμε την εξής εξίσωση:

$$Sold = \frac{1}{1 + e^{-(-0,29 + 1,96(SellerType) + 2,41(BuyMethod))}}$$

Αν ο τύπος του πωλητή είναι επιχείρηση και έχει τη τιμή 1 τότε η τιμή του Sold θα τείνει προς το 1 δηλαδή έχει περισσότερες πιθανότητες να πουληθεί. Αντίστοιχα, αν το προϊόν έχει πουληθεί μέσω δημοπρασίας, το Sold θα τείνει προς το 1 και η πιθανότητα να πουληθεί αυξάνεται.

Με βάση τα πιο πάνω μπορούμε να κάνουμε πρόβλεψη ότι τα προϊόντα της κατηγορίας Coins πωλούνται με μεγαλύτερη ευκολία αν προέρχονται από επιχειρήσεις και πωλούνται μέσω δημοπρασίας.

4.2.3 Κατηγορία Antiques

Link Function: Logit

Response Information

Variable	Value	Count
SoldB	1	193 (Event)
	0	42
Total		235

* NOTE * 235 cases were used

* NOTE * 3 cases contained missing values

Logistic Regression Table

Predictor	Coef	SE Coef	Z	P	Odds Ratio
Constant	-4,49560	1,63856	-2,74	0,006	
SellerTypeB					
1	3,09609	0,776927	3,99	0,000	22,11
BuyMethodB					
1	6,13390	1,30453	4,70	0,000	461,23
ShippingCost	-0,307102	0,109346	-2,81	0,005	0,74
Price	0,0339128	0,0095426	3,55	0,000	1,03

Log-Likelihood = -39,595
 Test that all slopes are zero: G = 141,453, DF = 4, P-Value = 0,000

Goodness-of-Fit Tests

Method	Chi-Square	DF	P
Pearson	308,977	157	0,000
Deviance	53,637	157	1,000
Hosmer-Lemeshow	17,986	8	0,021

Measures of Association:

(Between the Response Variable and Predicted Probabilities)

Pairs	Number	Percent	Summary Measures
Concordant	7686	94,8	Somers' D 0,91
Discordant	342	4,2	Goodman-Kruskal Gamma 0,91
Ties	78	1,0	Kendall's Tau-a 0,27
Total	8106	100,0	

Σημ. Στο λογισμικό πρόγραμμα Minitab αν το p-value είναι μικρότερο από 0,001 τότε παρουσιάζεται σαν 0,000.

Τρέξαμε το μοντέλο με όλες τις ανεξάρτητες μεταβλητές εκτός της κατάστασης του προϊόντος (ProductCondition) αφού σε όλες τις περιπτώσεις παίρνει τη τιμή 0. Παρατηρούμε ότι αυτές που επηρεάζουν την εξαρτημένη είναι ο τύπος του πωλητή, η μέθοδος με την οποία γίνεται η πώληση, το κόστος αποστολής και η τιμή του προϊόντος δεδομένου του πολύ χαμηλού p-value που έχουν οι συγκεκριμένες μεταβλητές. Το πολύ χαμηλό p-value υποδεικνύει ύπαρξη του μοντέλου, ύπαρξη δηλαδή σχέση μεταξύ της εξαρτημένης και της ανεξάρτητης μεταβλητής. Το p-value του μοντέλου είναι επίσης μικρότερο από 0,001 δείχνοντας έτσι την ύπαρξη μοντέλου.

Στον συντελεστή προσδιορισμού παρατηρούμε ότι το ποσοστό είναι στο 94,8% είναι εξαιρετικά που δείχνει την ύπαρξη ενός πολύ δυνατού μοντέλου. Αυτό σαν γεγονός μπορεί να ερμηνευθεί ότι η πώληση ενός προϊόντος της κατηγορίας Coins εξαρτάται από τον τύπο του πωλητή, τη μέθοδο της πώλησης το κόστος αποστολής και την τιμή του προϊόντος. Η εξίσωση παλινδρόμησης, έχει ως ακολούθως:

$$y = \frac{1}{1 + e^{-(a + \beta_1 x_1 + \beta_2 x_2 + \dots + \beta_n x_n)}}$$

Με βάση το μοντέλο μας και λαμβάνοντας υπόψη μόνο τις ανεξάρτητες μεταβλητές με p-value μικρότερο από 0,01 έχουμε την εξής εξίσωση:

$$Sold = \frac{1}{1 + e^{-(-4,5+3,10(SellerType)+6,13(BuyMethod)-0,3(ShippingCost)+0,03(Price))}}$$

Αν ο τύπος του πωλητή είναι επιχείρηση και έχει τη τιμή 1 τότε η τιμή του Sold θα τείνει προς το 1 δηλαδή έχει περισσότερες πιθανότητες να πουληθεί. Αντίστοιχα, αν το προϊόν έχει πουληθεί μέσω δημοπρασίας, το Sold θα τείνει προς το 1 και η πιθανότητα να πουληθεί αυξάνεται. Παράλληλα, η πιθανότητα να πωληθεί αυξάνεται αν το κόστος αποστολής γίνει πιο μικρό.

Με βάση τα πιο πάνω μπορούμε να κάνουμε πρόβλεψη ότι τα προϊόντα της κατηγορίας Antiques πωλούνται με μεγαλύτερη ευκολία αν προέρχονται από επιχειρήσεις, πωλούνται μέσω δημοπρασίας και έχουν χαμηλό κόστος αποστολής. Την ίδια στιγμή μπορούμε να πούμε ότι αν αυξηθεί η τιμή μίας αντίκας τότε υπάρχει μεγαλύτερη πιθανότητα να πωληθεί, λόγω της ιδιαιτερότητας της κατηγορίας.

4.2.4 Κατηγορία Books

Link Function: Logit

Response Information

Variable	Value	Count	
SoldB	1	25	(Event)
	0	87	
	Total	112	

* NOTE * 112 cases were used

* NOTE * 2 cases contained missing values

Logistic Regression Table

Predictor	Coef	SE Coef	Z	P	Odds Ratio
Constant	-21,3110	4760,97	-0,00	0,996	
SellerTypeB					
1	2,25287	2,87474	0,78	0,433	9,51
ProductConditionB					
1	21,3071	4760,97	0,00	0,996	1,79E+09
BuyMethodB					
1	18,6620	4760,97	0,00	0,997	1,27E+08
ShippingCost	-0,0777258	0,232034	-0,33	0,738	0,93
Price	-0,0216937	0,0226637	-0,96	0,338	0,98

Log-Likelihood = -48,742

Test that all slopes are zero: G = 21,449, DF = 5, P-Value = 0,001

Goodness-of-Fit Tests

Method	Chi-Square	DF	P
Pearson	35,2087	33	0,364
Deviance	40,4610	33	0,174
Hosmer-Lemeshow	2,8937	5	0,716

Measures of Association:

(Between the Response Variable and Predicted Probabilities)

Pairs	Number	Percent	Summary Measures
Concordant	1391	64,0	Somers' D 0,43
Discordant	446	20,5	Goodman-Kruskal Gamma 0,51
Ties	338	15,5	Kendall's Tau-a 0,15
Total	2175	100,0	

Σημ. Στο λογισμικό πρόγραμμα Minitab αν το p-value είναι μικρότερο από 0,001 τότε παρουσιάζεται σαν 0,000.

Το p-value του μοντέλου 0,364 που δείχνει τη μη ύπαρξη μοντέλου. Αυτό άλλωστε αποδεικνύεται και από το υψηλό p-value που έχουν όλες οι ανεξάρτητες μεταβλητές.

Συμπερασματικά, μπορούμε να πούμε ότι η πώληση ενός βιβλίου εξαρτάται από παράγοντες οι οποίοι δεν έχουν καταγραφεί.

4.2.5 Κατηγορία Sports Memorabilia

Link Function: Logit

Response Information

Variable	Value	Count
SoldB	1	345 (Event)
	0	81
Total		426

* NOTE * 426 cases were used

* NOTE * 15 cases contained missing values

Logistic Regression Table

Predictor	Coef	SE Coef	Z	P	Odds Ratio
Constant	0,319238	1,12944	0,28	0,777	
SellerTypeB					
1	-2,01812	0,733231	-2,75	0,006	0,13
ProductConditionB					
1	19,4430	7529,78	0,00	0,998	2,78E+08
BuyMethodB					
1	2,55158	1,25823	2,03	0,043	12,83
ShippingCost	0,131424	0,0625348	2,10	0,036	1,14
Price	-0,0389070	0,0216619	-1,80	0,072	0,96

Log-Likelihood = -193,150
 Test that all slopes are zero: G = 28,136, DF = 5, P-Value = 0,000

Goodness-of-Fit Tests

Method	Chi-Square	DF	P
Pearson	181,465	139	0,009
Deviance	176,523	139	0,017
Hosmer-Lemeshow	29,201	7	0,000

Measures of Association:

(Between the Response Variable and Predicted Probabilities)

Pairs	Number	Percent	Summary Measures
Concordant	18575	66,5	Somers' D 0,37
Discordant	8096	29,0	Goodman-Kruskal Gamma 0,39
Ties	1274	4,6	Kendall's Tau-a 0,12
Total	27945	100,0	

Το p-value του μοντέλου είναι 0,009 που δείχνει την ύπαρξη κάποιου μοντέλου, χωρίς όμως δυνατή σχέση μεταξύ των ανεξάρτητων μεταβλητών και της εξαρτημένης. Η μόνη μεταβλητή που φαίνεται να επηρεάζει σε κάποιο βαθμό είναι ο τύπος του πωλητή.

Συμπερασματικά, μπορούμε να πούμε ότι η πώληση αντικειμένων της κατηγορίας Sports Memorabilia επηρεάζεται από άλλους παράγοντες οι οποίοι δεν αναλύονται.

4.2.6 Κατηγορία Stamps

Link Function: Logit

Response Information

Variable	Value	Count
SoldB	1	159 (Event)
	0	235
Total		394

* NOTE * 394 cases were used
 * NOTE * 37 cases contained missing values

Logistic Regression Table

Predictor	Coef	SE Coef	Z	P	Odds Ratio
Constant	-3,52840	1,29336	-2,73	0,006	
SellerTypeB					
1	2,14921	0,482204	4,46	0,000	8,58
BuyMethodB					
1	1,38416	1,26181	1,10	0,273	3,99
ShippingCost	1,00827	0,223387	4,51	0,000	2,74
Price	0,0538051	0,0288839	1,86	0,062	1,06

Log-Likelihood = -228,847
 Test that all slopes are zero: G = 73,754, DF = 4, P-Value = 0,000

Goodness-of-Fit Tests

Method	Chi-Square	DF	P
Pearson	177,651	113	0,000
Deviance	192,101	113	0,000
Hosmer-Lemeshow	26,459	6	0,000

Measures of Association:

(Between the Response Variable and Predicted Probabilities)

Pairs	Number	Percent	Summary Measures
Concordant	26391	70,6	Somers' D 0,45
Discordant	9426	25,2	Goodman-Kruskal Gamma 0,47
Ties	1548	4,1	Kendall's Tau-a 0,22
Total	37365	100,0	

Σημ. Στο λογισμικό πρόγραμμα Minitab αν το p-value είναι μικρότερο από 0,001 τότε παρουσιάζεται σαν 0,000.

Τρέξαμε το μοντέλο με όλες τις ανεξάρτητες μεταβλητές εκτός της κατάστασης του προϊόντος (ProductCondition) αφού σε όλες τις περιπτώσεις παίρνει τη τιμή 0. Παρατηρούμε από το p-value των μεταβλητών ότι αυτές που επηρεάζουν την εξαρτημένη είναι ο τύπος του πωλητή, το κόστος αποστολής και σε μικρότερο βαθμό η τιμή του προϊόντος. Το πολύ χαμηλό p-value υποδεικνύει ύπαρξη του μοντέλου, ύπαρξη δηλαδή σχέση μεταξύ της εξαρτημένης και των ανεξάρτητων μεταβλητών. Το p-value του μοντέλου είναι επίσης μικρότερο από 0,001 δείχνοντας έτσι την ύπαρξη μοντέλου.

Στον συντελεστή προσδιορισμού παρατηρούμε ότι το ποσοστό είναι στο 70,6% που δείχνει την ύπαρξη ενός καλού μοντέλου. Αυτό σαν γεγονός μπορεί να ερμηνευθεί ότι η πώληση ενός προϊόντος της κατηγορίας Stamps εξαρτάται από τον τύπο του πωλητή, το κόστος αποστολής και την τιμή του προϊόντος. Η εξίσωση παλινδρόμησης, έχει ως ακολούθως:

$$y = \frac{1}{1 + e^{-(a + \beta_1 x_1 + \beta_2 x_2 + \dots + \beta_n x_n)}}$$

Με βάση το μοντέλο μας και λαμβάνοντας υπόψη μόνο τις ανεξάρτητες μεταβλητές με p-value μικρότερο από 0,01 έχουμε την εξής εξίσωση:

$$Sold = \frac{1}{1 + e^{-(-3,53+2,15(SellerType)+1,00(ShippingCost)+0,05(Price))}}$$

Αν ο τύπος του πωλητή είναι επιχείρηση και έχει τη τιμή 1 τότε η τιμή του Sold θα τείνει προς το 1 δηλαδή έχει περισσότερες πιθανότητες να πουληθεί. Παράλληλα, η πιθανότητα να πουληθεί αυξάνεται αν το κόστος αποστολής αυξηθεί δεδομένου ότι θα αυξηθεί η τιμή.

Με βάση τα πιο πάνω μπορούμε να κάνουμε πρόβλεψη ότι τα προϊόντα της κατηγορίας Stamps πωλούνται με μεγαλύτερη ευκολία αν προέρχονται από επιχειρήσεις. Την ίδια στιγμή μπορούμε να πούμε ότι αν αυξηθεί η τιμή ενός γραμματοσήμου τότε υπάρχει μεγαλύτερη πιθανότητα να πουληθεί, λόγω της ιδιαιτερότητας της κατηγορίας.

4.2.7 Κατηγορία Computers

Link Function: Logit

Response Information

Variable	Value	Count	
SoldB	1	43	(Event)
	0	351	
Total		394	

* NOTE * 394 cases were used

* NOTE * 10 cases contained missing values

Logistic Regression Table

Predictor	Coef	SE Coef	Z	P	Odds Ratio
Constant	-0,861351	0,642820	-1,34	0,180	
SellerTypeB					
1	-1,90189	0,565408	-3,36	0,001	0,15
ProductConditionB					
1	-0,289300	0,455539	-0,64	0,525	0,75
BuyMethodB					
1	1,28893	0,578081	2,23	0,026	3,63
ShippingCost	-0,0276595	0,0291691	-0,95	0,343	0,97
Price	0,0013275	0,0009646	1,38	0,169	1,00

Log-Likelihood = -103,964

Test that all slopes are zero: G = 63,702, DF = 5, P-Value = 0,000

Goodness-of-Fit Tests

Method	Chi-Square	DF	P
Pearson	290,239	219	0,001
Deviance	178,511	219	0,979
Hosmer-Lemeshow	3,227	8	0,919

Measures of Association:

(Between the Response Variable and Predicted Probabilities)

Pairs	Number	Percent	Summary Measures
Concordant	11760	77,9	Somers' D 0,61
Discordant	2563	17,0	Goodman-Kruskal Gamma 0,64
Ties	770	5,1	Kendall's Tau-a 0,12
Total	15093	100,0	

Σημ. Στο λογισμικό πρόγραμμα Minitab αν το p-value είναι μικρότερο από 0,001 τότε παρουσιάζεται σαν 0,000.

Τρέξαμε το μοντέλο με όλες τις ανεξάρτητες μεταβλητές. Παρατηρούμε από το p-value των μεταβλητών ότι αυτές που επηρεάζουν την εξαρτημένη είναι ο τύπος του πωλητή και η μέθοδος που γίνεται η πώληση. Το πολύ χαμηλό p-value υποδεικνύει ύπαρξη του μοντέλου, ύπαρξη δηλαδή σχέση μεταξύ της εξαρτημένης και των ανεξάρτητων μεταβλητών. Το p-value του μοντέλου είναι 0,001 δείχνοντας έτσι την ύπαρξη μοντέλου.

Στον συντελεστή προσδιορισμού παρατηρούμε ότι το ποσοστό είναι στο 77,9% που δείχνει την ύπαρξη ενός καλού μοντέλου. Αυτό σαν γεγονός μπορεί να ερμηνευθεί ότι η πώληση ενός προϊόντος της κατηγορίας Computers εξαρτάται από τον τύπο του πωλητή και τη μέθοδο πώλησης του προϊόντος. Η εξίσωση παλινδρόμησης, έχει ως ακολούθως:

$$y = \frac{1}{1 + e^{-(a + \beta_1 x_1 + \beta_2 x_2 + \dots + \beta_n x_n)}}$$

Με βάση το μοντέλο μας και λαμβάνοντας υπόψη μόνο τις ανεξάρτητες μεταβλητές με p-value μικρότερο από 0,01 έχουμε την εξής εξίσωση:

$$Sold = \frac{1}{1 + e^{-(-0,86 - 1,9(SellerType) + 1,3(BuyMethod))}}$$

Αν ο τύπος του πωλητή είναι επιχείρηση και έχει τη τιμή 1 τότε η τιμή του Sold θα τείνει προς το 0 δηλαδή έχει λιγότερες πιθανότητες να πουληθεί. Παράλληλα, η πιθανότητα να πουληθεί αυξάνεται αν η πώληση γίνει μέσω δημοπρασίας.

Με βάση τα πιο πάνω μπορούμε να κάνουμε πρόβλεψη ότι τα προϊόντα της κατηγορίας Computers πωλούνται με μεγαλύτερη ευκολία αν προέρχονται από ιδιώτες και πωλούνται μέσω δημοπρασίας.

4.2.8 Κατηγορία Games-Toys

Link Function: Logit

Response Information

Variable	Value	Count	
SoldB	1	205	(Event)
	0	71	
	Total	276	

* NOTE * 276 cases were used

* NOTE * 7 cases contained missing values

Logistic Regression Table

Predictor	Coef	SE Coef	Z	P	Odds Ratio
Constant	-3,33981	0,715294	-4,67	0,000	
SellerTypeB					
1	1,97882	0,642636	3,08	0,002	7,23
ProductConditionB					
1	1,43710	0,451894	3,18	0,001	4,21
BuyMethodB					
1	4,06560	0,563915	7,21	0,000	58,30
ShippingCost	0,150546	0,0611616	2,46	0,014	1,16
Price	-0,0167939	0,0086549	-1,94	0,052	0,98

Log-Likelihood = -92,992

Test that all slopes are zero: G = 128,742, DF = 5, P-Value = 0,000

Goodness-of-Fit Tests

Method	Chi-Square	DF	P
Pearson	209,310	142	0,000
Deviance	137,862	142	0,583
Hosmer-Lemeshow	54,587	7	0,000

Measures of Association:

(Between the Response Variable and Predicted Probabilities)

Pairs	Number	Percent	Summary Measures	
Concordant	11841	81,4	Somers' D	0,66
Discordant	2302	15,8	Goodman-Kruskal Gamma	0,67
Ties	412	2,8	Kendall's Tau-a	0,25
Total	14555	100,0		

Σημ. Στο λογισμικό πρόγραμμα Minitab αν το p-value είναι μικρότερο από 0,001 τότε παρουσιάζεται σαν 0,000.

Τρέξαμε το μοντέλο με όλες τις ανεξάρτητες μεταβλητές. Παρατηρούμε, από το p-value των μεταβλητών, ότι αυτές που επηρεάζουν την εξαρτημένη είναι ο τύπος του πωλητή, η κατάσταση του προϊόντος, η μέθοδος με την οποία γίνεται η πώληση και η τιμή του προϊόντος. Το πολύ χαμηλό p-value υποδεικνύει ύπαρξη του μοντέλου, ύπαρξη δηλαδή σχέση μεταξύ της εξαρτημένης και των ανεξάρτητων μεταβλητών. Το p-value του μοντέλου είναι επίσης μικρότερο από 0,001 δείχνοντας έτσι την ύπαρξη μοντέλου.

Στον συντελεστή προσδιορισμού παρατηρούμε ότι το ποσοστό είναι στο 81,4% είναι εξαιρετικά που δείχνει την ύπαρξη ενός δυνατού μοντέλου. Αυτό σαν γεγονός μπορεί να ερμηνευθεί ότι η πώληση ενός προϊόντος της κατηγορίας Games-Toys εξαρτάται από τον τύπο του πωλητή, την κατάσταση του προϊόντος, τη μέθοδο της πώλησης και την τιμή του προϊόντος. Η εξίσωση παλινδρόμησης, έχει ως ακολούθως:

$$y = \frac{1}{1 + e^{-(a + \beta_1 x_1 + \beta_2 x_2 + \dots + \beta_n x_n)}}$$

Με βάση το μοντέλο μας και λαμβάνοντας υπόψη μόνο τις ανεξάρτητες μεταβλητές με p-value μικρότερο από 0,01 έχουμε την εξής εξίσωση:

$$Sold = \frac{1}{1 + e^{-(-3,34 + 1,98(SellerType) + 4,1(BuyMethod) + 1,44(ProductCondition) - 0,02(Price))}}$$

Αν ο τύπος του πωλητή είναι επιχείρηση και έχει τη τιμή 1 τότε η τιμή του Sold θα τείνει προς το 1 δηλαδή έχει περισσότερες πιθανότητες να πουληθεί. Αν το προϊόν έχει πουληθεί μέσω δημοπρασίας, το Sold θα τείνει προς το 1 και η πιθανότητα να πουληθεί αυξάνεται. Αν το προϊόν είναι καινούριο το Sold θα τείνει προς το 1 και έτσι έχει μεγαλύτερη πιθανότητα να πουληθεί. Παράλληλα, η πιθανότητα να πωληθεί αυξάνεται αν το κόστος του προϊόντος μειωθεί.

Με βάση τα πιο πάνω μπορούμε να κάνουμε πρόβλεψη ότι τα προϊόντα της κατηγορίας Games-Toys πωλούνται με μεγαλύτερη ευκολία αν προέρχονται από επιχειρήσεις, πωλούνται μέσω δημοπρασίας, είναι καινούρια και έχουν χαμηλό κόστος.

4.2.9 Κατηγορία Clothes

Link Function: Logit

Response Information

Variable	Value	Count	
SoldB	1	30	(Event)
	0	118	
Total		148	

Logistic Regression Table

Predictor	Coef	SE Coef	Z	P	Odds Ratio
Constant	-5,22411	1,33600	-3,91	0,000	
SellerTypeB					
1	4,85573	1,13902	4,26	0,000	128,47
ProductConditionB					
1	0,817763	1,26133	0,65	0,517	2,27
BuyMethodB					
1	-0,169629	1,15169	-0,15	0,883	0,84
ShippingCost	0,379280	0,134300	2,82	0,005	1,46
Price	-0,121118	0,0483215	-2,51	0,012	0,89

Log-Likelihood = -35,343

Test that all slopes are zero: G = 78,535, DF = 5, P-Value = 0,000

Goodness-of-Fit Tests

Method	Chi-Square	DF	P
Pearson	36,7944	51	0,933
Deviance	39,2924	51	0,884
Hosmer-Lemeshow	4,0437	6	0,671

Measures of Association:

(Between the Response Variable and Predicted Probabilities)

Pairs	Number	Percent	Summary Measures	
Concordant	3227	91,2	Somers' D	0,84
Discordant	249	7,0	Goodman-Kruskal Gamma	0,86
Ties	64	1,8	Kendall's Tau-a	0,27
Total	3540	100,0		

Σημ. Στο λογισμικό πρόγραμμα Minitab αν το p-value είναι μικρότερο από 0,001 τότε παρουσιάζεται σαν 0,000.

Το p-value του μοντέλου είναι πολύ μεγάλο που δείχνει ότι δεν υπάρχει κάποιο μοντέλο, δηλαδή η εξαρτημένη τιμή δεν εξηγείται από τις ανεξάρτητες και υπάρχουν άλλες παράμετροι που επηρεάζουν.

Συμπερασματικά, δε μπορεί να γίνει κάποια πρόβλεψη με βάση αυτά τα δεδομένα και μπορούμε να πούμε ότι η πώληση αντικειμένων της κατηγορίας Clothes επηρεάζεται από άλλους παράγοντες οι οποίοι δεν αναλύονται.

4.2.10 Κατηγορία DVDs

Link Function: Logit

Response Information

Variable	Value	Count	
SoldB	1	25	(Event)
	0	109	
	Total	134	

Logistic Regression Table

Predictor	Coef	SE Coef	Z	P	Odds Ratio
Constant	-0,181316	1,20279	-0,15	0,880	
SellerTypeB					
1	1,72831	1,34892	1,28	0,200	5,63
ProductConditionB					
1	-0,0981113	1,41130	-0,07	0,945	0,91
BuyMethodB					
1	-0,288619	1,08672	-0,27	0,791	0,75
ShippingCost	-0,387935	0,126713	-3,06	0,002	0,68
Price	0,0824640	0,105104	0,78	0,433	1,09

Log-Likelihood = -51,295

Test that all slopes are zero: G = 26,374, DF = 5, P-Value = 0,000

Goodness-of-Fit Tests

Method	Chi-Square	DF	P
Pearson	27,8728	22	0,180
Deviance	33,5837	22	0,054
Hosmer-Lemeshow	5,3972	4	0,249

Measures of Association:

(Between the Response Variable and Predicted Probabilities)

Pairs	Number	Percent	Summary Measures
Concordant	2026	74,3	Somers' D 0,57
Discordant	467	17,1	Goodman-Kruskal Gamma 0,63
Ties	232	8,5	Kendall's Tau-a 0,17
Total	2725	100,0	

Σημ. Στο λογισμικό πρόγραμμα Minitab αν το p-value είναι μικρότερο από 0,001 τότε παρουσιάζεται σαν 0,000.

Το p-value του μοντέλου είναι πολύ μεγάλο που δείχνει ότι δεν υπάρχει κάποιο μοντέλο, δηλαδή η εξαρτημένη τιμή δεν εξηγείται από τις ανεξάρτητες και υπάρχουν άλλες παράμετροι που επηρεάζουν.

Συμπερασματικά, δε μπορεί να γίνει κάποια πρόβλεψη με βάση αυτά τα δεδομένα και μπορούμε να πούμε ότι η πώληση αντικειμένων της κατηγορίας DVDs επηρεάζεται από άλλους παράγοντες οι οποίοι δεν αναλύονται.

Συμπεράσματα – Μελλοντική εργασία

5.1 Συμπεράσματα

5.2 Μελλοντική εργασία

5. Συμπεράσματα – Μελλοντική εργασία

Σε αυτό το κεφάλαιο θα παρουσιάσουμε τα αποτελέσματα της μελέτης αυτής και τα σημαντικότερα ευρήματα μετά από την ανάλυση των δεδομένων. Στη ενότητα 5.2 θα γίνουν εισηγήσεις για μελλοντική εργασία επί του θέματος.

5.1 Συμπεράσματα

Από την ανάλυση των δεδομένων διαφάνηκε ότι κάθε κατηγορία είναι ξεχωριστοί και επηρεάζεται από διαφορετικούς παράγοντες. Στη συνέχεια θα γίνει παρουσίαση των αποτελεσμάτων – συμπερασμάτων που προέκυψαν από την ανάλυση για κάθε κατηγορία ξεχωριστά.

5.1.1 Κατηγορία Collectables

Στην κατηγορία Collectables πωλούνται συλλεκτικά αντικείμενα και ενθύμια, όποτε οι συλλέκτες τείνουν να προτιμούν τα καινούρια που είναι σε καλύτερη κατάσταση. Παράλληλα, το κόστος των προϊόντων αυτής της κατηγορίας δεν είναι μεγάλο με αποτέλεσμα ο αγοραστής να προτιμά τα καινούρια. Την ίδια στιγμή ο καταναλωτής λαμβάνει υπόψη τη σχέση της τιμής του προϊόντος με το κόστος αποστολής δηλαδή σε αντικείμενα μεγάλης αξίας είναι διατεθειμένος να πληρώσει περισσότερα σαν κόστος αποστολής.

Στο μοντέλο πρόβλεψης έχουμε την εξής εξίσωση παλινδρόμησης:

$$Sold = \frac{1}{1 + e^{-(-1,51+0,90(ProductCondition)+0,90(BuyMethod)+0,004(Price))}}$$

Με βάση τα πιο πάνω μπορούμε να κάνουμε πρόβλεψη ότι τα προϊόντα της κατηγορίας Collectables πωλούνται με μεγαλύτερη ευκολία αν είναι καινούρια και πωλούνται μέσω δημοπρασίας.

Άρα με βάση αυτά τα δεδομένα ένας πωλητής μπορεί να χαράξει την πολιτική του βάση των προτιμήσεων του καταναλωτή.

5.1.2 Κατηγορία Coins

Στην κατηγορία Coins παρατηρήθηκε ότι η αγοραστική απόφαση του καταναλωτή εξαρτάται από τη μέθοδο της πώλησης. Μία εξήγηση που μπορεί να δοθεί είναι η ευκαιριακή αγορά των νομισμάτων. Με τις δημοπρασίες η αρχική τιμή των προϊόντων είναι χαμηλή και ο πωλητής απευθύνεται σε συγκεκριμένη ομάδα αγοραστών τους συλλέκτες νομισμάτων. Ο κάθε αγοραστής ψάχνει συγκεκριμένα εξειδικευμένα αντικείμενα να αγοράσει μειώνοντας έτσι τον ανταγωνισμό με άλλους πιθανούς αγοραστές. Έτσι δεν υπάρχουν πολλές προσφορές (bids) για κάθε νόμισμα με αποτέλεσμα οι αγοραστές να επιτυγχάνουν καλύτερες τιμές. Την ίδια στιγμή ο καταναλωτής λαμβάνει υπόψη τη σχέση της τιμής του προϊόντος με το κόστος αποστολής δηλαδή σε αντικείμενα μεγάλης αξίας είναι διατεθειμένος να πληρώσει περισσότερα σαν κόστος αποστολής.

Στο μοντέλο πρόβλεψης έχουμε την εξής εξίσωση παλινδρόμησης:

$$Sold = \frac{1}{1 + e^{-(-0,29+1,96(SellerType)+2,41(BuyMethod))}}$$

Με βάση τα πιο πάνω μπορούμε να κάνουμε πρόβλεψη ότι τα προϊόντα της κατηγορίας Coins πωλούνται με μεγαλύτερη ευκολία αν προέρχονται από επιχειρήσεις και πωλούνται μέσω δημοπρασίας.

Άρα με βάση αυτά τα δεδομένα ένας πωλητής μπορεί να χαράξει την πολιτική του βάση των προτιμήσεων του καταναλωτή.

5.1.3 Κατηγορία Antiques

Στην κατηγορία Antiques παρατηρούμε ότι οι καταναλωτές προτιμούν να αγοράζουν από επιχειρήσεις παρά από ιδιώτες. Ένας λόγος είναι ότι εμπιστεύονται περισσότερο τις αγορές αυτές. Όταν η αγορά γίνεται από επιχείρηση τότε μειώνεται ο κίνδυνος των κλοπιμαίων και των

πλαστών αντικειμένων και ο αγοραστής νιώθει πιο σίγουρος για την αγορά του.

Ακόμα ένας παράγοντας που επηρεάζει είναι ο τρόπος με τον οποίο γίνονται οι πωλήσεις (δημοπρασίες/ απευθείας αγορές). Με τις δημοπρασίες η αρχική τιμή των προϊόντων είναι χαμηλή και ο πωλητής απευθύνεται σε συγκεκριμένη ομάδα αγοραστών τους συλλέκτες αντικών. Ο κάθε αγοραστής ψάχνει συγκεκριμένα εξειδικευμένα αντικείμενα να αγοράσει μειώνοντας έτσι τον ανταγωνισμό με άλλους πιθανούς αγοραστές. Έτσι δεν υπάρχουν πολλές προσφορές (bids) για κάθε αντικείμενο που πωλείται, με αποτέλεσμα οι αγοραστές να επιτυγχάνουν καλύτερες τιμές.

Την ίδια στιγμή η απόφαση του καταναλωτή επηρεάζεται από το κόστος αποστολής. Όσο το προϊόν είναι χαμηλής αξίας ο αγοραστής λαμβάνει υπόψη το κόστος αποστολής και δεν είναι διατεθειμένος να πληρώσει πολλά ενώ όσο αυξάνεται η αξία του προϊόντος το κόστος αποστολής δεν έχει ιδιαίτερη σημασία. Αυτό πιθανόν να οφείλεται στο γεγονός ότι ο αγοραστής που ασχολείται με αγορά αντικών είναι διατεθειμένος να πληρώσει περισσότερο για να αποκτήσει κάποιο κομμάτι που θεωρεί μοναδικό.

Στο μοντέλο πρόβλεψης έχουμε την εξής εξίσωση παλινδρόμησης:

$$Sold = \frac{1}{1 + e^{-(-4,5+3,10(SellerType)+6,13(BuyMethod)-0,3(ShippingCost)+0,03(Price))}}$$

Με βάση τα πιο πάνω μπορούμε να κάνουμε πρόβλεψη ότι τα προϊόντα της κατηγορίας Antiques πωλούνται με μεγαλύτερη ευκολία αν προέρχονται από επιχειρήσεις, πωλούνται μέσω δημοπρασίας και έχουν χαμηλό κόστος αποστολής, ενώ αν αυξηθεί η τιμή μίας αντικας τότε υπάρχει μεγαλύτερη πιθανότητα να πωληθεί, λόγω της ιδιαιτερότητας της κατηγορίας.

Άρα με βάση αυτά τα δεδομένα ένας πωλητής μπορεί να χαράξει την πολιτική του βάση των προτιμήσεων του καταναλωτή.

5.1.4 Κατηγορία Books

Στην κατηγορία Books ο αγοραστής προτιμά να αγοράζει μέσω δημοπρασίας αντί με απευθείας αγορά. Αυτό οφείλεται στο γεγονός ότι ο αγοραστής ψάχνει την καλύτερη ευκαιρία και τη χαμηλότερη τιμή. Το ίδιο βιβλίο μπορεί να το βρει πιθανόν από πολλούς πωλητές και έτσι έχει τη δυνατότητα να κάνει μία χαμηλή προσφορά (bid) και να θέσει αυτός τα όρια του ποσού που μπορεί να διαθέσει. Σε αντίθεση στην απευθείας αγορά η τιμή είναι σταθερή και τις περισσότερες φορές υψηλότερη από αυτή της δημοπρασίας.

Άρα με βάση αυτά τα δεδομένα ένας πωλητής μπορεί να χαράξει την πολιτική του βάση των προτιμήσεων του καταναλωτή.

5.1.5 Κατηγορία Sports Memorabilia

Στην κατηγορία Sports Memorabilia φαίνεται ότι οι αγοραστές προτιμούν να αγοράζουν καινούρια αντικείμενα παρά χρησιμοποιημένα. Ένας λόγος πιθανόν να είναι η καλύτερη φυσική κατάσταση των καινούριων αντικειμένων, έναντι των χρησιμοποιημένων. Παράλληλα, επειδή πρόκειται για αναμνηστικά αθλημάτων οι τιμές δεν είναι απαγορευτικές και έτσι επιτρέπεται στους αγοραστές να προτιμούν τα καινούρια. Την ίδια στιγμή ο καταναλωτής λαμβάνει υπόψη τη σχέση της τιμής του προϊόντος με το κόστος αποστολής δηλαδή σε αντικείμενα μεγάλης αξίας είναι διατεθειμένος να πληρώσει περισσότερα σαν κόστος αποστολής.

Άρα με βάση αυτά τα δεδομένα ένας πωλητής μπορεί να χαράξει την πολιτική του βάση των προτιμήσεων του καταναλωτή.

5.1.6 Κατηγορία Stamps

Στην κατηγορία Stamps φαίνεται ότι οι αγοραστές προτιμούν να αγοράζουν από επιχειρήσεις. Ένας λόγος είναι ότι εμπιστεύονται περισσότερο τις αγορές αυτές. Όταν η αγορά γίνεται από επιχείρηση τότε μειώνεται ο κίνδυνος των πλαστών γραμματοσήμων και ο αγοραστής νιώθει πιο σίγουρος για την αγορά του. Ο αγοραστής δεν

έχει φυσική επαφή με το αντικείμενο που θα αγοράσει έτσι δίνει μεγάλη σημασία στην αξιοπιστία του πωλητή.

Την ίδια στιγμή ο καταναλωτής λαμβάνει υπόψη τη σχέση της τιμής του προϊόντος με το κόστος αποστολής δηλαδή σε αντικείμενα μεγάλης αξίας είναι διατεθειμένος να πληρώσει περισσότερο σαν κόστος αποστολής.

Στο μοντέλο πρόβλεψης έχουμε την εξής εξίσωση παλινδρόμησης:

$$Sold = \frac{1}{1 + e^{-(-3,53+2,15(SellerType)+1,00(ShippingCost)+0,05(Price))}}$$

Με βάση τα πιο πάνω μπορούμε να κάνουμε πρόβλεψη ότι τα προϊόντα της κατηγορίας Stamps πωλούνται με μεγαλύτερη ευκολία αν προέρχονται από επιχειρήσεις. Την ίδια στιγμή μπορούμε να πούμε ότι αν αυξηθεί η τιμή ενός γραμματοσήμου τότε υπάρχει μεγαλύτερη πιθανότητα να πωληθεί, λόγω της ιδιαιτερότητας της κατηγορίας.

Άρα με βάση αυτά τα δεδομένα ένας πωλητής μπορεί να χαράξει την πολιτική του βάση των προτιμήσεων του καταναλωτή.

5.1.7 Κατηγορία Computers

Στην κατηγορία Computers φαίνεται ότι οι αγοραστές προτιμούν να αγοράζουν τα αντικείμενα αυτής της κατηγορίας μέσω δημοπρασίας. Ένας πιθανός λόγος που να εξηγεί τη συμπεριφορά αυτή είναι ότι με τις δημοπρασίες υπάρχει η δυνατότητα να αγοραστεί ένα προϊόν σε χαμηλή τιμή, αν αυτό δεν έχει μεγάλη ζήτηση.

Την ίδια στιγμή ο καταναλωτής λαμβάνει υπόψη τη σχέση της τιμής του προϊόντος με το κόστος αποστολής δηλαδή σε αντικείμενα μεγάλης αξίας είναι διατεθειμένος να πληρώσει περισσότερο σαν κόστος αποστολής.

Στο μοντέλο πρόβλεψης έχουμε την εξής εξίσωση παλινδρόμησης:

$$Sold = \frac{1}{1 + e^{-(-0,86 - 1,9(SellerType)+1,3(BuyMethod))}}$$

Με βάση τα πιο πάνω μπορούμε να κάνουμε πρόβλεψη ότι τα προϊόντα της κατηγορίας Computers πωλούνται με μεγαλύτερη ευκολία αν προέρχονται από ιδιώτες και πωλούνται μέσω δημοπρασίας.

Άρα με βάση αυτά τα δεδομένα ένας πωλητής μπορεί να χαράξει την πολιτική του βάση των προτιμήσεων του καταναλωτή.

5.1.8 Κατηγορία Games – Toys

Στην κατηγορία Games-Toys παρατηρείται ότι οι αγοραστές προτιμούν τα καινούρια προϊόντα. Ένας πιθανός λόγος είναι ότι τα παιχνίδια αφορούν παιδιά άρα ο αγοραστής δε διακινδυνεύει την ασφάλεια του παιδιού και την ομαλή λειτουργία των παιχνιδιών με αγορά πιο φτηνών χρησιμοποιημένων παιχνιδιών. Αποτέλεσμα αυτού η αγορά καινούριων προϊόντων έστω και σε υψηλότερη τιμή.

Άλλος παράγοντας που επηρεάζει φαίνεται να είναι ότι οι αγοραστές προτιμούν να αγοράζουν παιχνίδια μέσω δημοπρασίας παρά μέσω απευθείας αγοράς. Η προτίμηση αυτή οφείλεται στο γεγονός ότι σε μία δημοπρασία μπορεί ο αγοραστής να πετύχει καλύτερη τιμή αν δεν έχει μεγάλη ζήτηση το συγκεκριμένο αντικείμενο. Παράλληλα, επειδή είναι αντικείμενα που ο αγοραστής μπορεί να βρει εύκολα είναι διατεθειμένος να περιμένει την κατάλληλη ευκαιρία μέσω δημοπρασίας.

Στο μοντέλο πρόβλεψης έχουμε την εξής εξίσωση παλινδρόμησης:

$$Sold = \frac{1}{1 + e^{-(-3,34+1,98(SellerType)+4,1(BuyMethod)+1,44(ProductCondition)-0,02(Price))}}$$

Με βάση τα πιο πάνω μπορούμε να κάνουμε πρόβλεψη ότι τα προϊόντα της κατηγορίας Games-Toys πωλούνται με μεγαλύτερη ευκολία αν προέρχονται από επιχειρήσεις, πωλούνται μέσω δημοπρασίας, είναι καινούρια και έχουν χαμηλό κόστος.

Άρα με βάση αυτά τα δεδομένα ένας πωλητής μπορεί να χαράξει την πολιτική του βάση των προτιμήσεων του καταναλωτή.

5.1.9 Κατηγορία Clothes

Στην κατηγορία Clothes φαίνεται ότι οι αγοραστές προτιμούν να αγοράζουν καινούρια ρούχα, πιθανόν για λόγους υγιεινής αλλά και φθοράς. Τα ρούχα είναι αντικείμενα που φθείρονται εύκολα και μπορεί οι φθορές να μη φαίνονται οπτικά (μέσω των φωτογραφιών που παρέχονται). Αυτό έχει ως αποτέλεσμα οι αγοραστές να προτιμούν να αγοράζουν καινούρια προϊόντα έστω και σε υψηλότερη τιμή.

Παράλληλα, ο τύπος του πωλητή αποτελεί παράγοντα που επηρεάζει την αγοραστική απόφαση. Οι αγοραστές προτιμούν να αγοράζουν από επιχειρήσεις αφού έτσι κερδίζουν καλύτερη ποιότητα και αποφεύγουν τα προϊόντα μαϊμού.

Την ίδια στιγμή ο καταναλωτής λαμβάνει υπόψη τη σχέση της τιμής του προϊόντος με το κόστος αποστολής δηλαδή σε αντικείμενα μεγάλης αξίας είναι διατεθειμένος να πληρώσει περισσότερα σαν κόστος αποστολής.

Αρα με βάση αυτά τα δεδομένα ένας πωλητής μπορεί να χαράξει την πολιτική του βάση των προτιμήσεων του καταναλωτή.

5.1.10 Κατηγορία DVDs

Στην κατηγορία DVDs φαίνεται ότι η φυσική κατάσταση ενός DVD αποτελεί παράγοντα στην απόφαση του αγοραστή. Τα DVDs είναι προϊόντα χαμηλής αξίας και επιπλέον μπορεί να έχουν φθορές που να μη φαίνονται οπτικά μέσω φωτογραφιών. Αυτοί οι δύο παράγοντες οδηγούν στην αγορά καινούριων DVD αφού η αγορά καινούριων DVDs δεν έχει μεγάλη διάφορα σε σχέση με τα χρησιμοποιημένα ενώ παράλληλα εξασφαλίζεται η ποιότητα του.

Παράλληλα φαίνεται ότι ο τύπος του πωλητή αποτελεί παράγοντα που επηρεάζει την αγοραστική απόφαση. Οι αγοραστές προτιμούν να αγοράζουν από επιχειρήσεις αφού έτσι νιώθουν μεγαλύτερη ασφάλεια για την ποιότητα των DVDs και αποφεύγουν τα μη αυθεντικά προϊόντα.

Το κόστος αποστολής φαίνεται να αποτελεί παράγοντα στην απόφαση του αγοραστή κατά πόσο θα αγοράσει ένα DVD ή όχι αφού το μέγεθος και το βάρος είναι σταθερό με αποτέλεσμα ο αγοραστής να προτιμά τα προϊόντα αυτά που έχουν χαμηλότερο το κόστος αποστολής.

Άρα με βάση αυτά τα δεδομένα ένας πωλητής μπορεί να χαράξει την πολιτική του βάση των προτιμήσεων του καταναλωτή.

5.2 Μελλοντική εργασία

Η μελέτη αυτή είχε ως στόχο να δείξει ποιοι παράγοντες είναι αυτοί που επηρεάζουν ή όχι την πώληση ενός προϊόντος. Τα αποτελέσματα και τα συμπεράσματα της έρευνας φαίνονται στην ενότητα 5.1.

Κατά τη διάρκεια της μελέτης παρατηρήθηκε ότι παρά το γεγονός ότι υπήρχε η πληροφόρηση για κάποιες μεταβλητές αυτές δε μπορούσαν να αναλυθούν αφού συνδέονταν με άλλες μεταβλητές που δεν ήταν διαθέσιμες.

Για παράδειγμα, κατά τη συλλογή των δεδομένων έγινε καταχώρηση στη βάση δεδομένων του πεδίου NoOfBids. Το πεδίο αυτό χαρακτηρίζει τον αριθμό των bids που έγιναν σε ένα αντικείμενο δεδομένου ότι αυτό πωλήθηκε μέσω δημοπρασίας. Με το στοιχείο αυτό θα μπορούσε να αποδειχτεί κατά πόσο υπάρχει θετική σχέση μεταξύ του αριθμού των bids και της τιμής. Το στοιχείο αυτό παρά το γεγονός ότι είναι σημαντικό δε μπορούσε να χρησιμοποιηθεί αφού δεν ήταν διαθέσιμη η αρχική τιμή του αντικειμένου.

Ένα άλλο σημαντικό στοιχείο που δεν έχει εξεταστεί είναι κατά πόσο επηρεάζει την πώληση η διάρκεια που είναι «ενεργό προς πώληση» ένα αντικείμενο. Δηλαδή, αν οι καταναλωτές προτιμούν να αγοράζουν και να κάνουν bid σε προϊόντα που η διάρκεια που βρίσκονται προς πώληση είναι πιο μικρή ή πιο μεγάλη. Στη μελέτη αυτή δεν ήταν δυνατό να εξεταστεί αυτός ο παράγοντας αφού δεν υπήρχε διαθέσιμη η

πληροφορία της ημερομηνίας καταχώρησης του προϊόντος (Starting Date).

Τέλος, μπορεί να γίνει η επέκταση της μελέτης και σε άλλα ηλεκτρονικά καταστήματα με στόχο την πιο ολοκληρωμένη εικόνα της συμπεριφοράς του καταναλωτή ως προς τις ηλεκτρονικές αγορές. Έτσι θα διαφανεί κατά πόσο ο πιθανός αγοραστής συμπεριφέρεται το ίδιο σε κάθε ηλεκτρονικό κατάστημα ή ανάλογα με το κατάστημα αλλάζουν οι παράγοντες που λαμβάνει υπόψη.

Βιβλιογραφία

- [1] **Gerald Keller**, (2009), “Managerial Statistics”, 8th edition.
- [2] **Douglas Downing and Jeffrey Clark** (2003),”Business Statistics”, 4th edition, USA.
- [3] **Gelman, A & Hill, J.** (2006). *Data Analysis Using Regression and Multilevel/Hierarchical Models*. Cambridge,UK: Cambridge University Press.
- [4] **David W. Hosmer and Stanley Lemeshow** (2000), “Applied Logistic Regression”, 2nd edition, John Wiley & Sons Inc, USA.
- [5] **Menard, S.** (2002). *Applied Logistic Regression Analysis* (2nd ed.). Thousand Oaks, CA: Sage Publications.
- [6] **Cohen Jacob** (2003), “Applied multiple regression”, 3rd edition, Mahwah, N.J. : L. Erlbaum Associates
- [7] **Ken Black** (2012), “Business Statistics: For Contemporary Deciesion Making”, 7th edition, John Wiley & Sons Inc. USA.
- [8] **Ning-Zhong Shi and Jian Tao** (2008), “Statistical Hypothesis Testing: Theory and methods”,World Scientific Publishing Co. Pte. Ltd. Singapore.
- [9] **Efthimios Constantinides** (2004), “Influencing the online consumer’s behavior: the Web experience”, *Internet Research*, Vol.14, No.2, pp.111-126.
- [10] **Perea y Monsuwe, Bendict G.C Dellaert and Ko de Ruyter** (2004), “What drives consumers to shop online? A literature view”,*International Journal of service industry management*, Vol.15, No.1 pp.102-121.
- [11] **Li Guo** (2011), “A Research on Influencing Factors of Consumer. Purchasing Behaviors in Cyberspace”, *International Journal of Marketing Studies*, Vol.3, No.3.
- [12] **Brendan Hannah & Kristina M. Lybecker** (2010), “Determinants of Recent Online Purchasing and the Percentage of Income Spent Online”, *International Business Research*, Vol.3 No.4, pp. 70-61.
- [13] **Inma Rodriguez-Adura, Antoni Meseguer-Artola and Jordi Vilaseca-Requena** (2008), “Factors Influencing the Evolution of Electronic Commerce: An Empirical Analysis in a Developed Market Economy”, *Journal of Theoretical and Applied Electronic Commerce Research*, Vol.3, No.1, pp.18-29.

[14] **Yanfen Shang, Fugee Tsung and Changliang Zou** (2011), “Statistical Process Control for Multistage Processes with BinaryOutputs”, ProQuest Dissertations and Theses.

[15] **Wojciech Peter Latusek**, (2010) "B2B relationship marketing analytical support with GBC modeling", *Journal of Business & Industrial Marketing*, Vol. 25 Iss: 3, pp.209 – 219.

[16] **Junhong Chu, Marta Arce-Urriza, José-Javier Cebollada-Calvo and Pradeep K. Chintagunta** (2010), “An Empirical Analysis of Shopping Behavior Across Online and Offline Channels for Grocery Products: The Moderating Effects of Household and Product Characteristics”, *Journal of Interactive Marketing*, Vol.14, pp.251-268.

[17] **Teo, T.S.H., Wang, P. and Leong, H.C.** (2004) ‘Understanding online shopping behavior using a transaction cost economics approach’, *Int. J. Internet Marketing and Advertising*, Vol. 1, No. 1, pp.62–84.

[18] **Jee Young K. Kim** (2005), “Factors influencing consumers’ apparel purchasing intention in the C2C e-commerce market”, University of Nebraska.

[19] **Fang He** (2009), “Decisions factors for the adoption of e-finance and other e-commerce activities”, Southern Illinois University Carbondale.

[20] **Sangeeta Sahney** (2008), “Critical Success Factors in Online Retail – An Application of Quality Function Deployment And Interpretive Structural Modeling”, *International Journal of Business and Information*, Vol.3, No.1.

[21] **Mohammad Hossein Moshref Javadi, Hossein Rezaei Dolatabadi, Mojtaba Nourbakhsh, Amir Poursaedi & Ahmad Reza Asadollahi** (2012), “An Analysis of Factors Affecting on Online Shopping Behavior of Consumers”, *International Journal of Marketing Studies*; Vol. 4, No. 5.

[22] **Peng Huang, Nicholas H. Lurie, & Sabyasachi Mitra** (2009). “Searching for Experience on the Web: An Empirical Examination of Consumer Behavior for Search and Experience Goods” *Journal of Marketing*, Vol. 73, pp. 55–69.

[23] **George J. Avlonitis and Despina A. Karayanni** (2000), “The Impact of Internet Use on Business-to-Business Marketing”, *Industrial Marketing Management* Vo.l.29, pp. 441–459.

[24] **Jae-II Kim and Hee Chun Lee** (2004), “Factors Affecting Online Search Intention and Online Purchase Intention”, *Seoul Journal of Business*, Vol. 10, No.2.