Αθήνα, 2018

Polyzogopoulos Pavlos (bapt1730)

Dimitris Karlis

**STATISTICS II**

Project 1

Περιεχόμενα

[Εισαγωγή 2](#_Toc507105251)

[Μεθοδολογία 3](#_Toc507105252)

[Ανάλυση δεδομένων 4](#_Toc507105253)

[Αποτελέσματα Ανάλυσης 8](#_Toc507105254)

[Συμπεράσματα 12](#_Toc507105255)

# Εισαγωγή

Οι εταιρίες τηλεπικοινωνιών αντιμετωπίζουν παγκοσμίως μία μεγάλη πρόκληση, την απώλεια ενός όγκου πελατών τους. Το φαινόμενο αυτό είναι πιο έντονο στις εταιρίες της κινητής τηλεφωνίας όπου η αλλαγή παρόχου είναι πιο εύκολη καθώς πλέον ο καταναλωτής μπορεί να διατηρήσει τον αριθμό του αλλάζοντας εταιρία. Ως αποτέλεσμα αυτού και καθώς ο κλάδος των τηλεπικοινωνιών είναι ιδιαίτερα ανταγωνιστικός στις μέρες μας, οι καταναλωτές έχουν όλο και περισσότερες ευκαιρίες , τόσο από οικονομική άποψη όσο και από ποιοτική όσο αφορά τις επιλογές τους. Προφανώς το φαινόμενο αυτό , είναι ιδιαίτερα ανησυχητικό για τις εταιρίες καθώς συνδέεται άμεσα με την πορεία των κερδών τους. Για το λόγο αυτό όλο και περισσότερες τείνουν στην ανάλυση της συμπεριφοράς των συνδρομητών τους και πιο συγκεκριμένα στην ανάλυση της αποχώρησης τους, προκειμένου να βελτιώσουν και να ενισχύσουν τη σχέση τους με τους πελάτες τους. Σκοπός αυτής της εργασίας είναι η ανάλυση της αποχώρησης των συνδρομητών από την εταιρία. Το φαινόμενο αυτό είναι παγκοσμίως γνωστό και ως “churn”.[[1]](#footnote-1)

Προκειμένου να μελετηθεί το φαινόμενο της αποχώρησης καταναλωτών στην εταιρία μας, θα αναλυθούν ιστορικά δεδομένα τα οποία περιέχουν στοιχεία ενός μέρους από το χαρτοφυλάκιο των πελατών μας. Τα δεδομένα αυτά αφορούν την προηγούμενη χρονική περίοδο. Πιο συγκεκριμένα θα γίνει ανάλυση για 3333 διαφορετικούς πελάτες και τα στοιχεία που θα μπουν κάτω από το μικροσκόπιο αυτής της ανάλυσης αφορούν κυρίως την συμπεριφορά των συνδρομητών μας όσο αφορά τον χρόνο, το πλήθος αλλά και τις αντίστοιχες χρεώσεις για το πλήθος των κλήσεων τους. Είναι διαθέσιμα επίσης και μερικά δημογραφικά στοιχεία που αφορούν τους συγκεκριμένους πελάτες.

Πιο συγκεκριμένα τα στοιχεία/μεταβλητές είναι :

1. Το σύνολο των μηνών που ο καταναλωτής είναι πελάτης της εταίρας μας. Η μεταβλητή αυτή αναφέρεται ως **Account Length.**
2. Το σύνολο των ηχητικών μηνυμάτων για τα οποία χρεώθηκε ο καταναλωτής (**VMail Message)**
3. Ο συνολικός χρόνος (μετρημένος σε λεπτά) για τα οποία χρεώθηκε ο καταναλωτής κατά την διάρκεια της ημέρας (**Day Mins**).
4. Ο συνολικός χρόνος (μετρημένος σε λεπτά) για τα οποία χρεώθηκε ο καταναλωτής κατά την διάρκεια του απογεύματος (**Eve Mins**).
5. Ο συνολικός χρόνος (μετρημένος σε λεπτά) για τα οποία χρεώθηκε ο καταναλωτής κατά την διάρκεια της νύχτας (**Night Mins).**
6. Ο συνολικός χρόνος (μετρημένος σε λεπτά) για τα οποία χρεώθηκε ο καταναλωτής για διεθνείς κλήσεις(**Intl Mins**).
7. Το σύνολο των κλήσεων προς το τμήμα εξυπηρέτησης πελατών **(CustServ Calls).**
8. Η πληροφορία σχετικά με το αν ο καταναλωτής αποχώρησε εν τέλει από την εταιρία (**Churn**).
9. Αν ο καταναλωτής διαθέτει την υπηρεσία για διεθνείς κλήσεις **(Int'l Plan**).
10. Αν ο καταναλωτής διαθέτει την υπηρεσία για ηχητικά μηνύματα (**VMail Plan**).
11. Το σύνολο των κλήσεων που πραγματοποίησε κατά την διάρκεια της ημέρας **(Day Calls).**
12. Το σύνολο των κλήσεων που πραγματοποίησε κατά την διάρκεια του απογεύματος **(Eve Calls).**
13. Το σύνολο των κλήσεων που πραγματοποίησε κατά την διάρκεια της νύχτας (**Night Calls).**
14. Το σύνολο των κλήσεων που πραγματοποίησε για διεθνείς κλήσεις **(Intl Calls).**
15. H συνολική χρέωση για τις κλήσεις που πραγματοποίησε στην διάρκεια της ημέρας **(Day Charge).**
16. H συνολική χρέωση για τις κλήσεις που πραγματοποίησε στην διάρκεια του απογεύματος **(Eve Charge).**
17. H συνολική χρέωση για τις κλήσεις που πραγματοποίησε στην διάρκεια της νύχτας **(Night Charge).**
18. H συνολική χρέωση για τις διεθνείς κλήσεις που πραγματοποίησε **(Intl Charge).**
19. Την πολιτεία στην οποία κατοικεί ο καταναλωτής **(State).**
20. Το φύλλο του καταναλωτή **(Gender).**

Η μεταβλητή churn θα αποτελέσει την εξαρτημένη μεταβλητή στην ανάλυση που θα ακολουθήσει ενώ οι υπόλοιπες θα είναι οι εξαρτημένες.

Η ανάλυση αυτών των δεδομένων στοχεύει στο να απαντήσει με όσο μεγαλύτερη ακρίβεια στο ποιοι είναι οι λόγοι που οδηγούν ένα καταναλωτή να αποχωρήσει από την εταιρία. Ποιοι είναι δηλαδή οι συντελεστές που επιδρούν θετικά και ποιοι αρνητικά σε αυτήν του την απόφαση. Ένα τέτοιο μοντέλο είναι ιδιαίτερα σημαντικό, καθώς θα μπορούσε να απαντήσει σε αυτά τα ερωτήματα και να οριοθετήσει τις κατηγορίες πελατών οι οποίες έχουν αυξημένο κίνδυνο αποχώρησης.

# Μεθοδολογία

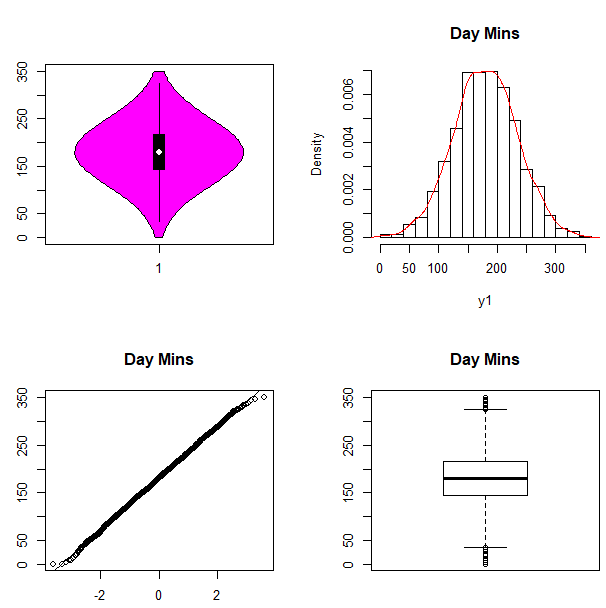
Στην πράξη θα επιδιώξουμε να φτιάξουμε ένα μοντέλο το οποίο θα μπορεί να μελετήσει την σχέση της κατηγορικής εξαρτημένης μεταβλητής (churn) με τις υπόλοιπες ποσοτικές και ποιοτικές ανεξάρτητες μεταβλητές (μεταβλητές που αναφέρθηκαν προηγουμένως). Για το λόγο αυτό είναι προτιμότερο να χρησιμοποιηθεί η μέθοδος της λογιστικής παλινδρόμησης (logistic regression). Η λογιστική παλινδρόμηση είναι στην ουσία ένα μοντέλο ταξινόμησης των τιμών μίας μεταβλητής απόκλισης Υ με βάση την θεωρία των πιθανοτήτων. Στο μοντέλο αυτό όπου η μεταβλητή έχει δυαδικό χαρακτήρα (παίρνει δύο τιμές ΝΑΙ και ΌΧΙ) στοχεύεται η πρόβλεψη της έκβασης αυτής από το πλήθος των υπόλοιπων μεταβλητών. Η λογιστική παλινδρόμηση χρησιμοποιείται για την πρόβλεψη της πιθανότητας εμφάνισης ενός γεγονότος (στην περίπτωσή μας το αν ένας συνδρομητής θα αποφασίσει να αποχωρήσει ή όχι από την εταιρία) προσαρμόζοντας τα δεδομένα της μελέτης. Η ανάλυση έχει γίνει με την βοήθεια της γλώσσας προγραμματισμού R και όλος ο κώδικάς υπάρχει σε ξεχωριστό αρχείο για περεταίρω μελέτη.

# Ανάλυση δεδομένων

Το πρώτο βήμα είναι η ανάλυση των δεδομένων . Όπως αναφέρθηκε παραπάνω τα δεδομένα που είναι προς μελέτη αφορούν 3.333 διαφορετικούς πελάτες της εταιρίας της προηγούμενης χρονικής περιόδου.

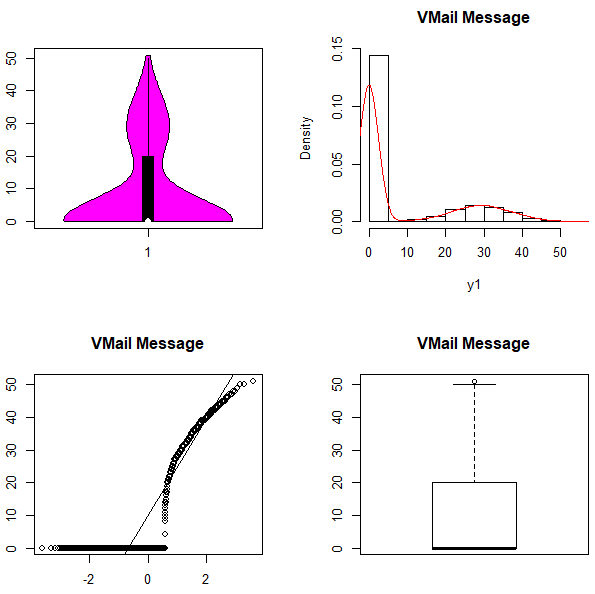
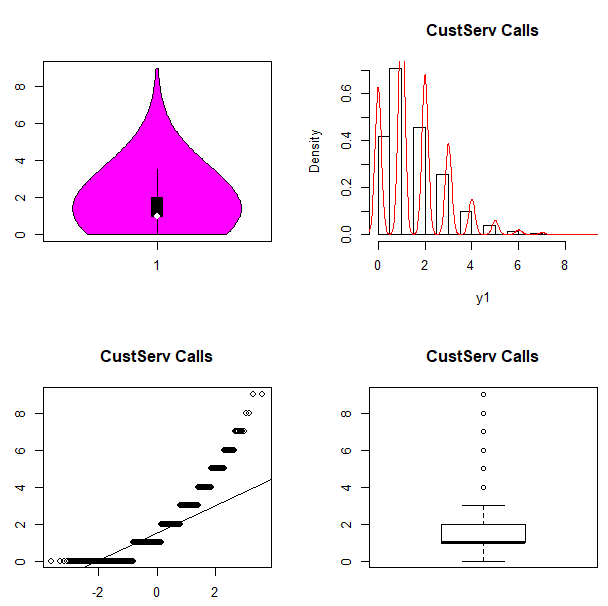
Αρχικά τα δεδομένα εισάγονται στην R. Το πρώτο πράγμα που θα ελέγξουμε είναι εάν υπάρχουν κενά ή λάθη στα δεδομένα (NAs , NANs , missing values). Τα δεδομένα είναι «καθαρά» πράγμα που επιτρέπει την συνέχεια της ανάλυσης. Προκειμένου να γίνουν πιο κατανοητά τα δεδομένα θα ακολουθήσει περιγραφή ενός μέσου συνδρομητή. Ο μέσος συνδρομητής λοιπόν είναι πελάτης της εταιρίας για περίπου 101 μήνες (8,4 χρόνια). Κατά την διάρκεια της τελευταίας χρονικής περιόδου, πραγματοποίησε περίπου ίδιο αριθμό κλήσεων (περίπου 100) στις διαφορετικές «ζώνες» της ημέρας (ημέρα , απόγευμα ,νύχτα). Την ημέρα μίλησε κατά μέσο όρο τον λιγότερο χρόνο αθροιστικά (180 λεπτά) σε σύγκριση με το απόγευμα και την νύχτα όπου χρεώθηκε για περίπου τον ίδιο χρόνο δηλαδή 201 λεπτά. Αντιθέτως , λόγω των διαφορετικών χρεώσεων που ισχύουν για τις διαφορετικές χρονικές περιόδους η μέση χρέωση για τις πραγματοποιθείσες κλήσεις ήταν 31€ για την ημέρα , 17€ για το απόγευμα και 9€ για το βράδυ. Συνολικά δηλαδή ένας μέσος πελάτης πραγματοποίησε 300 κλήσεις στις οποίες μίλησε για 983 λεπτά και χρεώθηκε συνολικά περίπου 57€. Όσο αφορά τις διεθνής κλήσεις κάθε καταναλωτής πραγματοποίησε περίπου 4 κλήσεις ενώ μίλησε για περίπου 10,2 λεπτά γεγονός που χρέωσε τον λογαριασμό του κατά 2,7€. Ένα σημείο που πρέπει να τονιστεί είναι ότι κάθε συνδρομητής κάλεσε κατά μέσο όρο 2 φορές το τμήμα εξυπηρέτησης πελατών. Επιπλέον μόνο το 10% (323 στους 3.333) των συνδρομητών διαθέτουν το πρόγραμμα για διεθνείς κλήσεις. Αντιθέτως το 28% (922 στους 3.333) έχουν ενεργοποιημένη την υπηρεσία για φωνητικά μηνύματα με τον κάθε συνδρομητή να πραγματοποιεί 8 φωνητικά μνήματα κατά μέσο όρο. Τέλος το ποσοστό ανδρών και γυναικών είναι ίδιο και οι πελάτες προέρχονται από 51 διαφορετικές πολιτείες των Ηνωμένων Πολιτειών της Αμερικής.

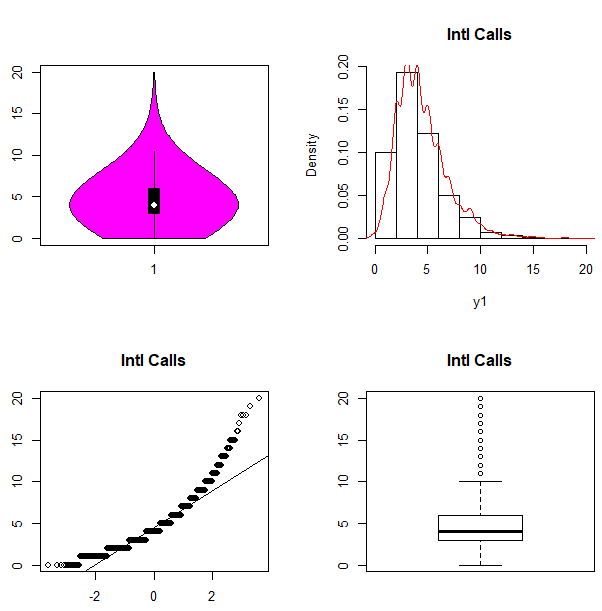
Ένα πολύ ενδιαφέρον στοιχεί είναι ότι για τις περισσότερες από τις παραπάνω ποσοτικές μεταβλητές μπορούμε να πούμε ότι ακολουθούν την κανονική κατανομή. Αυτή η πληροφορία μπορεί να αποδειχτεί ιδιαίτερα χρήσιμη για την καλύτερη κατανόηση της συμπεριφοράς των συνδρομητών της εταιρίας. Τα παρακάτω διαγράμματα αφορούν την μεταβλητή Day Mins (χρόνος ομιλίας κατά την διάρκεια της ημέρας) και αποδεικνύουν την εν λόγω παρατήρηση. Όμοια αποτελέσματα έχουν και τα διαγράμματα για τις υπόλοιπες αριθμητικές μεταβλητές.



Διάγραμμα 1: Πάνω αριστερά : διάγραμμα βιολιού (violin plots ή vioplots) . Πάνω δεξιά ιστόγραμμα (histogram). Κάτω αριστερά Q-Q Plot. Κάτω δεξιά: Γράφημα πλαισίου απολήξεων. Και τα τέσσερα διαγράμματα δείχνουν ότι ακολουθείται η κανονική κατανομή.

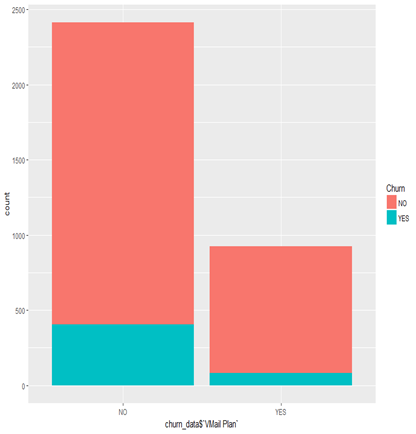
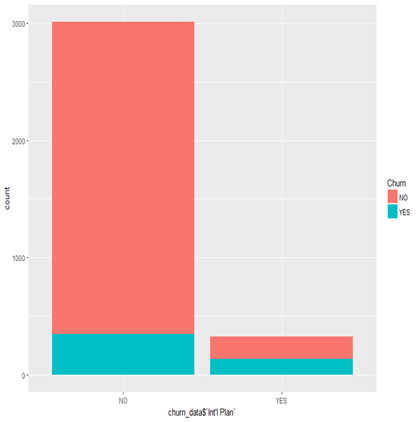
Αντίθετη εικόνα παρουσιάζουν οι μεταβλητές για τον αριθμό των ηχητικών μηνυμάτων (Vmail Message) , τον αριθμό των κλήσεων προς το τμήμα εξυπηρέτησης πελατών (CustSerc Calls) και για τον αριθμό των διεθνών κλήσεων (Intl Calls). Όπως δείνουν και τα διαγράμματα παρακάτω και οι τρεις μεταβλητές παρουσιάζουν έντονη δεξιά συμμετρία γεγονός που μας υποδηλώνει ότι οι συνδρομητές δεν χρησιμοποιούν με την ίδια συχνότητα τις συγκεκριμένες υπηρεσίες.

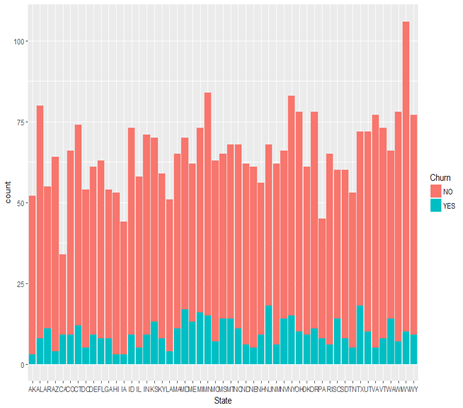
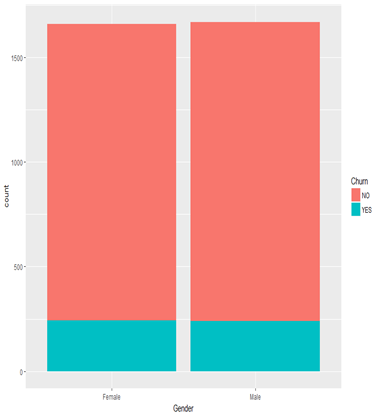
 



Διάγραμμα 2: Πάνω αριστερά : διάγραμμα βιολιού (violin plots ή vioplots) . Πάνω δεξιά ιστόγραμμα (histogram). Κάτω αριστερά Q-Q Plot. Κάτω δεξιά: Γράφημα πλαισίου απολήξεων. Υπάρχει έντονη δεξιά συμμετρία και αρκετές ακραίες τιμές.

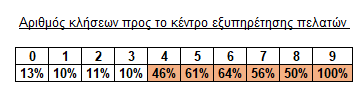
Ιδιαίτερο ενδιαφέρον παρουσιάζουν και οι σχέσεις μεταξύ ορισμένων μεταβλητών με το ζητούμενο αυτής της ανάλυσης που είναι το αν ο πελάτης αποχώρησε από την εταιρία η όχι. Το 42% των πελατών που έχει ενεργοποιημένη την υπηρεσία για τις διεθνείς κλήσεις επέλεξε τελικά να αλλάξει πάροχο κινητής τηλεφωνίας. Το ποσοστό αυτό είναι αισθητά πιο μικρό για την υπηρεσία των φωνητικών μηνυμάτων όπου μόλις το 9%. Επίσης οι πολιτείες Αριζόνα , Καλιφόρνια , Μέιν , Μέριλανντ , Μίσιγκαν ,Μισισίπι , Μοντάνα , Νεβάδα Νότια Καρολίνα και Τέξας παρουσιάζουν ποσοστά αποχώρησης μεγαλύτερα του 20% ενώ αντιθέτως οι πολιτείες Αλάσκα , Αριζόνα , Ουάσιγκτον , Χαβάη , Αϊόβα ,Ιλινόις , Ρόουντ Άιλαντ , Τένεσυ , Γουισκόζιν και Γουομίν να έχουν ποσοστά αποχώρησης κάτω του 10%. Τέλος δεν παρατηρείται κάποια διαφορά στα ποσοστά μεταξύ των δύο φύλλων. Όλα τα παραπάνω συμπεράσματα παρουσιάζονται στο διάγραμμα (Διάγραμμα 3).





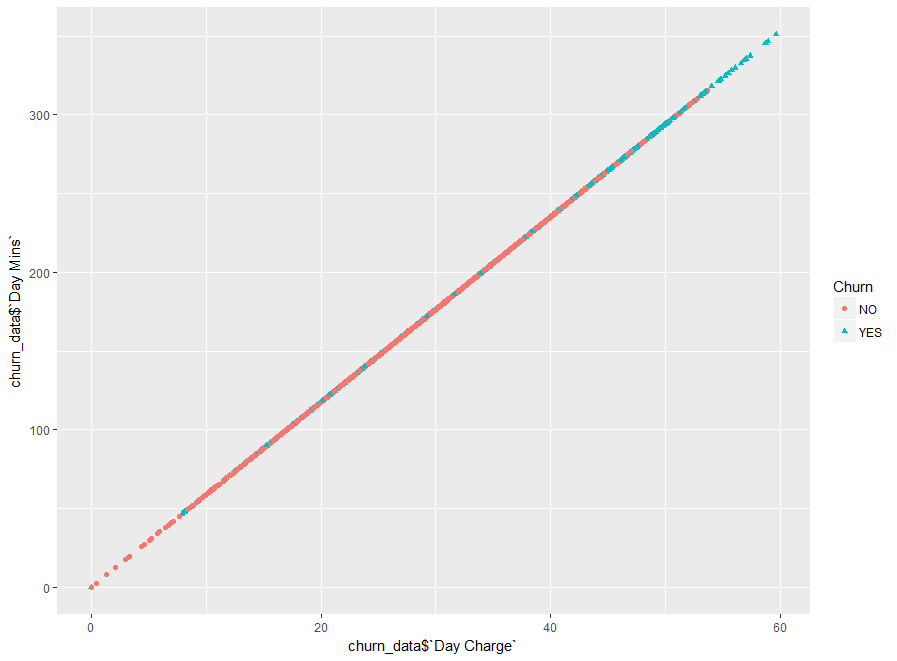
Διάγραμμα 3: Bar plots.

Συνεχίζοντας την ανάλυση φαίνεται ότι (όπως είναι και άλλωστε λογικό) οι κλήσεις προς το κέντρο εξυπηρέτησης πελατών έχει επηρεάζει έντονα αρνητικά την απόφαση των συνδρομητών για να αποχωρήσουν. Πιο συγκεκριμένα όπως φαίνεται και στον Πίνακα 1 τα ποσοστά αποχώρησης εκτοξεύονται πάνω από το 40% όταν οι κλήσεις ξεπερνούν τις 4.



Πίνακας 1: Ποσοστά αποχώρησης συνδρομητών

Τέλος συγκρίνοντας τις τιμές του συνόλου των κλήσεων , λεπτών και χρεώσεων σε σχέση με την αποχώρηση των πελατών το μόνο που θα μπορούσε να συμπεράνει κάνεις είναι ότι κατά κανόνα η πλειοψηφία των πελατών που επέλεξαν να αλλάξουν πάροχο φαίνεται να είναι αυτοί που μιλούν περισσότερο κατά την διάρκεια της ημέρας. Το γεγονός αυτό οφείλεται προφανώς στο ότι οι χρεώσεις αυτήν την χρονική περίοδο είναι ψηλότερες σε σχέση με αυτές που ισχύουν το απόγευμα ή την νύχτα (0,17€ το λεπτό). Η σχέση αυτή φαίνεται ξεκάθαρα στο παρακάτω διάγραμμα (Διάγραμμα 4).



Διάγραμμα 4: Στον άξονα Χ η χρέωση ανά λεπτό για την ημέρα, στον άξονα Υ τα λεπτά χρέωσης για τις κλήσεις στην διάρκεια της ημέρας. Οι δύο τιμές είναι πάνω σε μία ευθεία καθώς η χρέωση ανά λεπτό είναι σταθερή και ίση με 0,17€ ανά λεπτό. Με μπλε χρώμα είναι οι πελάτες που επέλεξαν να αλλάξουν πάροχο.

# Αποτελέσματα Ανάλυσης

Το τελικό μοντέλο είναι της μορφής

logit=exp(α + βx1 + γx2) / 1+exp(α + βx1 + γx2)

Όπου P είναι η πιθανότητα ένας καταναλωτής να αποχωρήσει από την εταιρία και α + βx1 + γx2 είναι οι γραμμικοί συντελεστές (με α, β, γ να πρέπει να βρεθούν).

Οι προϋποθέσεις που θα πρέπει να τηρούνται ώστε το μοντέλο να είναι σωστό είναι :

Η ανεξαρτησία των παρατηρήσεων . Λόγω ότι τα δεδομένα απευθύνονται σε μοναδικούς συνδρομητές και καθώς αφορούν μία μόνο χρονική περίοδο μπορεί να θεωρηθεί ότι αυτή η προϋπόθεση ισχύει.

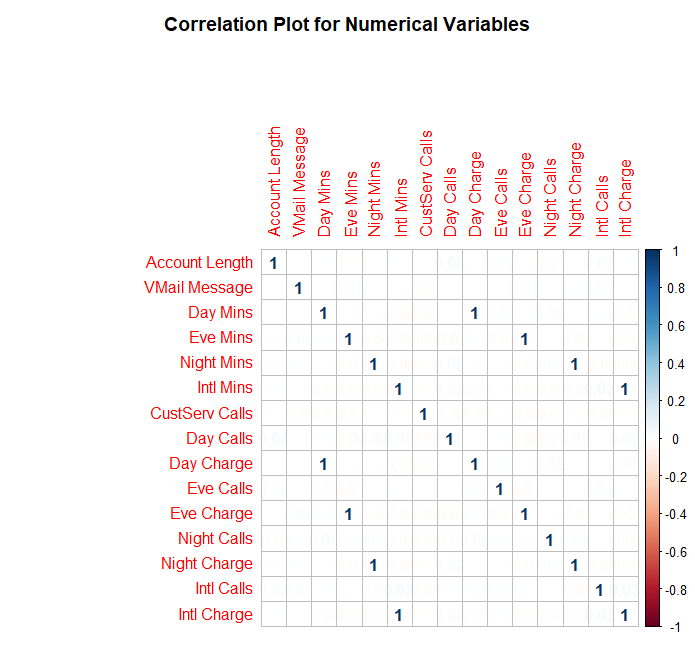
Δεν υπάρχει πολυγραμμικότητα (multicollinearity) μεταξύ των μεταβλητών. Αυτή η προϋπόθεση θα εξεταστεί ξεχωριστά παρακάτω.

Υπάρχει γραμμική συσχέτιση μεταξύ των των ανεξάρτητων μεταβλητών και της πιθανότητας ενός καταναλωτή να αποχωρήσει από την εταιρία. Από την μαθηματική εξίσωση υποθέτουμε αυτήν την σχέση η οποία θα εξεταστεί παρακάτω.

Αρχικά όπως έχει αναφερθεί και παραπάνω παρατηρούμε ότι πολιτική της εταιρίας είναι να έχει σταθερές χρεώσεις ανά χρονική ζώνη. Πιο συγκεκριμένα οι χρεώσεις αυτές είναι

* 0,17€ / λεπτό για την ημέρα
* 0,09€ / λεπτό για το απόγευμα
* 0,04€ / λεπτό για το βράδυ
* 0,27€ / λεπτό για διεθνής κλήσεις

Χρησιμοποιώντας την εντολή corrplot από το ομώνυμο πακέτο της R για τις αριθμιτικές μεταβλητές φαίνεται ότι, όπως ήταν και αναμενόμενο, υπάρχει απόλυτη συσχέτιση μεταξύ των μεταβλητών λεπτά και χρέωση για τις αντίστοιχες χρονικές ζώνες της ημέρας. (Πίνακας 2)



Πίνακας 2: Core plot.

Για το λόγο αυτό οι μεταβλητές που αφορούν τις χρεώσεις θα πρέπει να βγουν εκτός ανάλυσης καθώς σε διαφορετική περίπτωση θα παραβιαστεί μία από τις προϋποθέσεις του μοντέλου. Αντί αυτών θα προστεθούν οι αντίστοιχες χρεώσεις στα δεδομένα μας.

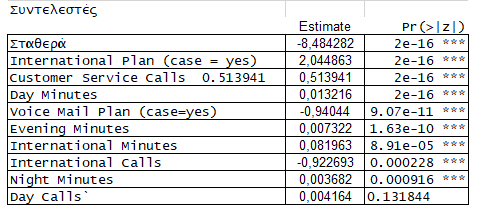
Για να καταλήξουμε σε αυτήν την απόφαση πρώτα εφαρμόσαμε ένα μοντέλο που περιείχε όλες τις μεταβλητές και ύστερα έγινε η ανάλυση των αποκλίσεων (εντολή anova) της οποίας τα αποτελέσματα έδειξαν ότι οι μεταβλητές που αφορούν τα λεπτά ομιλίας είανι πιο σημαντικές από τις αντίστοιχές των χρεώσεων στην πρόβλεψη της πιθανότητας αποχώρησης ενός πελάτη.

Ύστερα από πολλές δοκιμές και προσαρμόζοντας το μοντέλο στα δικά μας δεδομένα παίρνουμε το μοντέλο :

P=P(churn=1) = exp(0.51\*CustumerServiceCalls + 0.01\*DayMinutes + 0.007\*EveningMinutes + 0.08\*InternationalMinutes – 0.09\*InternationalCalls + 0.003\*NightMinutes + 0.004DayCalls + 2.04\*InternationalPlan (case plan = yes) – 0.94\*VoiceMailPlan(case Voice Mail Plan=yes + ε) /1+ exp(0.51\*CustumerServiceCalls + 0.01\*DayMinutes + 0.007\*EveningMinutes + 0.08\*InternationalMinutes – 0.09\*InternationalCalls + 0.003\*NightMinutes + 0.004DayCalls + 2.04\*InternationalPlan (case plan = yes) – 0.94\*VoiceMailPlan(case Voice Mail Plan=yes + ε)

Το συγκεκριμένο μοντέλο προέκυψε με την χρήση της μεθόδου stepwise (direction = forward).

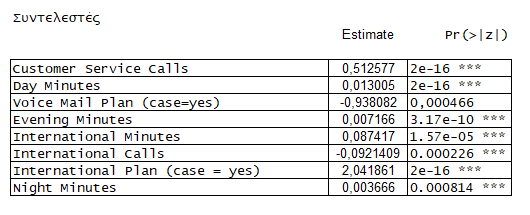
Τα στατιστικά του μοντέλου αυτού φαίνονται στον παρακάτω πίνακα (Πίνακας3)



Πίνακας 3: Πίνακας εκτιμήσεων των συντελεστών των παραμέτρων του αρχικού μοντέλου.

Με μία πρώτη ματιά φαίνεται ξεκάθαρα πως μόνο οι μεταβλητές Voice Mai Plan και International Calls μειώνουν την πιθανότητα ενός καταναλωτή να αποχωρήσει από την εταιρία με όλες τις υπόλοιπες να την αυξάνουν.

Προκειμένου να γίνει πιο κατανοητή η εξήγηση του μοντέλου θα θέσουμε όλες τις μεταβλητές στο Μ.Ο των δεδομένων ώστε η σύγκριση να γίνεται πλέον με βάση τον μέσο πελάτη που περιεγράφηκε παραπάνω. Μετά τις αλλαγές οι νέοι συντελεστές φαίνονται στον παρακάτω πίνακα (Πίνακας 4)



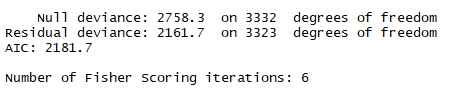
Πίνακας 4: Πίνακας εκτιμήσεων των συντελεστών των παραμέτρων του μοντέλου με κεντρικοποιημένες τιμές.

Η εξήγηση των συντελεστών των παραμέτρων για το μοντέλο είναι η εξής:

Για κάθε ένα παραπάνω λεπτό (αύξηση κατά μία μονάδα) που μιλάει ένας πελάτης σε σχέση με τον μέσο όρο , στην διάρκεια της ημέρας και μονάδες διατηρώντας τις υπόλοιπες μεταβλητές σταθερές, αυξάνεται ο λογιστικός μετασχηματισμός (logit) κατά 0,013005 μονάδες.

Αντίστοιχα ένας πελάτης που έχει ενεργοποιημένη την υπηρεσία φωνητικών μηνυμάτων έχει μειωμένο λογιστικό μετασχηματισμό κατά -0.93808 μονάδες σε σχέση με έναν πελάτη που δεν την έχει ενεργοποιήσει διατηρώντας πάντα τις υπόλοιπες μεταβλητές σταθερές.

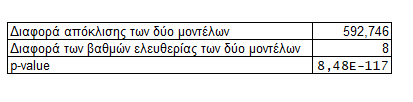
Η απόκλιση (deviance) του μηδενικού μοντέλου και των υπολοίπων, μαζί µε τους βαθμούς ελευθερίας τους αλλά και την τιμή του κριτηρίου AIC την βλέπουμε στον πίνακα (Πίνακας 5)



Πίνακας 5: Απόκλιση μηδενικού μοντέλου και τιμή κριτηρίου AIC

Tέλος για να δούμε εάν το μοντέλο μας έχει καλή εφαρμογή στα δεδομένα θα το συγκρίνουμε με το μηδενικό μοντέλο δηλαδή αυτό που περιλαμβάνει μόνο την σταθερά.

Για να γίνει αυτό θα χρησιμοποιήσουμε chi-square τεστ το οποίο μας δίνει τα εξής αποτελέσματα (ο κώδικας για το chi-square τεστ όπως και για όλα τα παραπάνω βρίσκεται στο αρχείο R)



Πίνακας 6: Αποτελέσματα chi-square τεστ.

Από τα παραπάνω συμπεραίνουμε ότι το μοντέλο μας έχει καλύτερη εφαρμογή στα δεδομένα μας σε σχέση με το μηδενικό πράγμα που σημαίνει ότι η προσθήκη μεταβλητών βοήθησε το μοντέλο μας να εξηγήσει καλύτερα την πιθανότητα αποχώρησης ενός μέσου συνδρομητή.

# Συμπεράσματα

Τα συμπεράσματα αυτής της έρευνας είναι ιδιαίτερα ανησυχητικά καθώς βρέθηκε ότι οι χρεώσεις για όλες τις χρονικές περιόδους της ημέρας αλλά και για τις διεθνείς κλήσεις επηρεάζουν αρνητικά την πιθανότητά ενός συνδρομητή να παραμείνει στην εταιρεία. Επίσης η υπηρεσία για τις διεθνείς κλήσεις έχει επίσης αρνητική επίδραση και μάλιστα ιδιαίτερα ισχυρή δεδομένου του γεγονότος ότι ο κλάδος των τηλεπικοινωνιών είναι ιδιαίτερα ανταγωνιστικός ίσως θα πρέπει η εταιρεία να εφαρμόσει διαφορετική πολιτική χρέωσης αλλά και να δώσει ιδιαίτερη έμφαση στις υπηρεσίες που προσφέρει προς τους πελάτες της, παίρνοντας παράδειγμα από την υπηρεσία ηχητικού μηνύματος η οποία δρα ευεργετικά. Τέλος, θα πρέπει αν μπουν στο μικροσκόπιο της εταιρείας (retention division) οι πελάτες οι οποίοι τηλεφωνούν συχνά στο κέντρο εξυπηρέτησης πελατών.

1. Σε αυτό το σημείο θα πρέπει να σημειωθεί πως δεν υπάρχει ακριβής μετάφραση στα Ελληνικά αυτού του όρου και για τους σκοπούς αυτής της εργασίας θα αναφέρεται ως “αποχώρηση καταναλωτών”. [↑](#footnote-ref-1)