# Βιβλιογραφία

## **Τμηματοποίηση πελατών**

### 1.1 Προέλευση και πρώτες έρευνες για την τμηματοποίηση πελατών

### 1.2 Τύποι τμηματοποίησης πελατών

## **Ανάλυση RFM**

### 2.1 Προέλευση και υλοποίηση

### 2.2 Παραλλαγές/Επεκτάσεις της ανάλυσης RFM

### 2.3 RFM segments: Χαρακτηριστικά & προτεινόμενες στρατηγικές μάρκετινγκ

## **3. Μηχανική Μάθηση**

### 3.1 Κατηγορίες Μηχανικής Μάθησης

### 3.2 Αλγόριθμοι Συσταδοποίησης

### 3.3 Συσταδοποίηση K-means

## **4. Συνδυασμός RFM και K-means**

## **Συστήματα επιχειρηματικής ευφυΐας και οπτικοποίηση δεδομένων**

~~~~~~~Η παράγραφος που ακολουθεί ΊΣΩΣ ΣΤΗΝ ΕΙΣΑΓΩΓΗ ή κάτι τέτοιο ~~~~~~~~~~

Σύμφωνα με τους Verhoef κ.ά. (2003), η ανάλυση RFM είναι μία από τις πιο αποτελεσματικές μεθόδους για την κατανόηση της αξίας και της αφοσίωσης των πελατών, ιδίως σε κλάδους όπου οι επαναλαμβανόμενες αγορές αποτελούν βασικό παράγοντα επιτυχίας, όπως το λιανικό εμπόριο, το ηλεκτρονικό εμπόριο και οι χρηματοπιστωτικές υπηρεσίες. Η δυνατότητα κατηγοριοποίησης των πελατών σε εφαρμόσιμα τμήματα, όπως «υψηλής αξίας» ή «υψηλού κινδύνου», επιτρέπει στις επιχειρήσεις να προσαρμόζουν αποτελεσματικότερα τις προσπάθειες μάρκετινγκ, οδηγώντας σε βελτιωμένη διατήρηση των πελατών και αύξηση των πωλήσεων (Verhoef et al., 2003).

## Τμηματοποίηση πελατών

Η τμηματοποίηση των πελατών παραμένει ακρογωνιαίος λίθος της στρατηγικής μάρκετινγκ, διότι επιτρέπει στις επιχειρήσεις να στοχεύουν αποτελεσματικότερα σε συγκεκριμένες ομάδες πελατών, αυξάνοντας τη συνάφεια και τον αντίκτυπο των εκστρατειών μάρκετινγκ. Εστιάζοντας στους σωστούς πελάτες, οι επιχειρήσεις μπορούν να βελτιώσουν την ικανοποίηση των πελατών, την αφοσίωση και τη συνολική κερδοφορία (Wind & Bell, 2007).

Οι Rust και Verhoef (2005) υπογραμμίζουν τη σημασία των δεδομένων των πελατών στην καθοδήγηση των αποφάσεων μάρκετινγκ, βοηθώντας τις επιχειρήσεις να κατανέμουν καλύτερα τους πόρους και να βελτιστοποιούν τις παρεμβάσεις μάρκετινγκ, θέτοντας τις βάσεις για το μέλλον των στρατηγικών μάρκετινγκ με βάση τα δεδομένα και την εξατομίκευση.

Η βέλτιστη τμηματοποίηση μπορεί να βελτιώσει σημαντικά τις λύσεις εξατομίκευσης, με τις μεθόδους άμεσης ομαδοποίησης να υπερτερούν συχνά των παραδοσιακών προσεγγίσεων που βασίζονται στη στατιστική (Jiang & Tuzhilin, 2006). Συνολικά, η τμηματοποίηση πελατών διαδραματίζει ζωτικό ρόλο στην κατανόηση των αναγκών των πελατών, στη βελτίωση της διατήρησης και στην αύξηση των πωλήσεων στο σημερινό ανταγωνιστικό επιχειρηματικό περιβάλλον (Shinde, 2023).

### 1.1 Προέλευση και πρώτες έρευνες για την τμηματοποίηση πελατών

Η έννοια της τμηματοποίησης πελατών προήλθε από τις πρώτες εργασίες για την τμηματοποίηση της αγοράς στην έρευνα μάρκετινγκ. Η πρώτη σημαντική ακαδημαϊκή συζήτηση εισήχθη το 1956 από τον Wendell R. Smith στην επιδραστική εργασία του με τίτλο «Διαφοροποίηση προϊόντος και τμηματοποίηση αγοράς ως εναλλακτικές στρατηγικές μάρκετινγκ» (“Product Differentiation and Market Segmentation as Alternative Marketing Strategies”). Ο Smith υποστήριξε ότι η μαζική αγορά θα μπορούσε να χωριστεί σε διακριτές ομάδες με βάση τα κοινά σημεία μεταξύ των καταναλωτών, οδηγώντας σε πιο στοχευμένες στρατηγικές μάρκετινγκ. Τόνισε ότι η τμηματοποίηση επιτρέπει στις εταιρείες να ικανοποιούν τις ανάγκες διαφορετικών ομάδων καταναλωτών πιο αποτελεσματικά, προσαρμόζοντας τα προϊόντα και τις προσεγγίσεις μάρκετινγκ (Smith, 1956).

Καθώς η ιδέα της τμηματοποίησης της αγοράς κέρδισε δημοτικότητα, αρκετοί μελετητές συνέβαλαν στην εξέλιξή της. Για παράδειγμα, ο Frank Bass (1969) ανέπτυξε το μοντέλο διάχυσης Bass, το οποίο χρησιμοποιείται για την ανάλυση του τρόπου με τον οποίο τα νέα προϊόντα υιοθετούνται με την πάροδο του χρόνου, με τους πρώιμους και τους όψιμους υιοθετούντες να σχηματίζουν διακριτά τμήματα της αγοράς. Αυτό το μοντέλο έθεσε τα θεμέλια για την τμηματοποίηση των πελατών με βάση τα πρότυπα υιοθέτησης καινοτομιών, το οποίο έκτοτε επηρέασε πολυάριθμους κλάδους, ιδίως στις αγορές υψηλής τεχνολογίας (Bass, 1969).

Περαιτέρω εξελίξεις στην τμηματοποίηση προήλθαν από την εργασία των Valentine και Powers (1972) για τη δημογραφική τμηματοποίηση, όπου οι ερευνητές διερεύνησαν πώς μεταβλητές όπως το εισόδημα, η εκπαίδευση και η ηλικία επηρεάζουν τη συμπεριφορά και τις προτιμήσεις των καταναλωτών. Ο Philip Kotler (1980) έπαιξε επίσης σημαντικό ρόλο στην επισημοποίηση της πρακτικής της τμηματοποίησης στη βιβλιογραφία του μάρκετινγκ, προσδιορίζοντας διάφορες προσεγγίσεις για την τμηματοποίηση των αγορών, συμπεριλαμβανομένης της δημογραφικής, γεωγραφικής, ψυχογραφικής και συμπεριφορικής τμηματοποίησης (Kotler, 1980).

1.2 Τύποι τμηματοποίησης πελατών

* ***Δημογραφική τμηματοποίηση***

Η δημογραφική τμηματοποίηση είναι η πιο συχνά χρησιμοποιούμενη μέθοδος διαχωρισμού των αγορών με βάση χαρακτηριστικά όπως η ηλικία, το φύλο, το εισόδημα, η εκπαίδευση, το επάγγελμα και η οικογενειακή κατάσταση. Οι δημογραφικές μεταβλητές είναι εύκολο να συλλεχθούν και συχνά επηρεάζουν άμεσα τη συμπεριφορά των καταναλωτών, καθιστώντας αυτόν τον τύπο τμηματοποίησης μια πρακτική επιλογή για όσους ασχολούνται με το εμπόριο (Dolnicar, 2004).

Για παράδειγμα, οι Chandon κ.ά. (1997) διερεύνησαν πώς δημογραφικά στοιχεία όπως το εισόδημα και το επάγγελμα επηρεάζουν την αγοραστική συμπεριφορά, αποδεικνύοντας ότι τα άτομα με υψηλότερο εισόδημα είναι πιο πιθανό να αγοράσουν αγαθά πολυτελείας. Ομοίως, οι Wedel και Kamakura (2000) τόνισαν ότι η δημογραφική τμηματοποίηση είναι θεμελιώδης σε κλάδους όπως τα καταναλωτικά συσκευασμένα προϊόντα (consumer packaged goods - CPG) και το λιανικό εμπόριο, όπου η ηλικία, το μέγεθος της οικογένειας και το εισόδημα έχουν άμεση επίδραση στην επιλογή προϊόντων και στις καταναλωτικές συνήθειες.

* ***Γεωγραφική τμηματοποίηση***

Η γεωγραφική τμηματοποίηση διαχωρίζει τους πελάτες με βάση την τοποθεσία τους, όπως χώρα, περιοχή ή πόλη. Η προσέγγιση αυτή αναγνωρίζει ότι γεωγραφικοί παράγοντες όπως το κλίμα, ο πολιτισμός και οι τοπικές προτιμήσεις μπορούν να επηρεάσουν σε μεγάλο βαθμό την αγοραστική συμπεριφορά. Για παράδειγμα, οι Cavusgil et al. (2004) τόνισαν τη σημασία της γεωγραφίας στο διεθνές μάρκετινγκ, σημειώνοντας ότι οι πολυεθνικές εταιρείες συχνά προσαρμόζουν τις προσφορές τους ώστε να ανταποκρίνονται στις συγκεκριμένες απαιτήσεις των διαφόρων περιοχών ή χωρών.

Επιπλέον, οι Hofstede κ.ά. (2010) επέκτειναν αυτή την κατανόηση συζητώντας πώς οι γεωγραφικοί παράγοντες και οι κουλτούρες των εθνών διαμορφώνουν τις προτιμήσεις των καταναλωτών και τη συμπεριφορά της αγοράς, καθιστώντας έτσι τη γεωγραφική τμηματοποίηση ζωτικής σημασίας για τις παγκόσμιες μάρκες (brands). Το σύστημα PRIZM, που αναπτύχθηκε από την Claritas τη δεκαετία του 1970, αποτελεί ένα παράδειγμα χρήσης τόσο γεωγραφικών όσο και δημογραφικών δεδομένων για τη δημιουργία προφίλ καταναλωτών (Weiss, 1988).

* ***Ψυχογραφική τμηματοποίηση***

Η ψυχογραφική τμηματοποίηση περιλαμβάνει την κατηγοριοποίηση των πελατών με βάση τον τρόπο ζωής, τις αξίες, τις προσωπικότητες και τις στάσεις τους. Αυτός ο τύπος τμηματοποίησης υπερβαίνει τις πιο απτές δημογραφικές ή γεωγραφικές μεθόδους προσπαθώντας να κατανοήσει τα υποκείμενα κίνητρα και τις στάσεις που καθοδηγούν τη συμπεριφορά των καταναλωτών. Ο Demby (1974) εισήγαγε τα ψυχογραφικά στοιχεία στο μάρκετινγκ και ο Wells (1975) επέκτεινε αργότερα τη σημασία του τρόπου ζωής στην τμηματοποίηση των καταναλωτών, υποστηρίζοντας ότι τα χαρακτηριστικά της προσωπικότητας και οι αξίες του ατόμου προσφέρουν μια βαθύτερη κατανόηση των αναγκών των καταναλωτών. Για παράδειγμα, ο Plummer (1974) διερεύνησε τα ψυχογραφικά προφίλ, κατηγοριοποιώντας τους πελάτες με βάση τα χαρακτηριστικά της προσωπικότητας και τις συμπάθειές τους για συγκεκριμένα προϊόντα. Οι Kahle και Kennedy (1988) συζήτησαν επίσης τη σημασία της κατανόησης των αξιών των πελατών για την τμηματοποίηση, τονίζοντας ότι τα ψυχογραφικά στοιχεία μπορούν να αποκαλύψουν γιατί οι καταναλωτές λαμβάνουν ορισμένες αγοραστικές αποφάσεις.

Η ψυχογραφική τμηματοποίηση παραμένει ιδιαίτερα χρήσιμη σε κλάδους όπως τα είδη πολυτελείας, η υγεία και η ευεξία και η μόδα, όπου τα κίνητρα και οι αξίες των καταναλωτών συχνά καθορίζουν την αγοραστική συμπεριφορά περισσότερο από τα δημογραφικά χαρακτηριστικά (Schiffman & Kanuk, 2007).

* ***Τμηματοποίηση με βάση την συμπεριφορά των καταναλωτών***

Η συμπεριφορική (?) τμηματοποίηση επικεντρώνεται στις αλληλεπιδράσεις των πελατών με μια εταιρεία ή ένα προϊόν, ομαδοποιώντας τους με βάση παράγοντες όπως το ιστορικό αγορών, η χρήση του προϊόντος, η αφοσίωση και οι απαντήσεις σε εκστρατείες μάρκετινγκ. Σύμφωνα με τους Wedel και Kamakura (2000), αυτός ο τύπος τμηματοποίησης παρέχει πιο εφαρμόσιμες πληροφορίες για τις επιχειρήσεις, καθώς τους επιτρέπει να προσαρμόζουν τις προσπάθειες μάρκετινγκ σε διαφορετικές ομάδες πελατών με βάση την πραγματική συμπεριφορά τους και όχι τα συμπερασματικά χαρακτηριστικά τους.

Οι Blattberg κ.ά. (2008) τόνισαν την αξία της τμηματοποίησης συμπεριφοράς στη διαχείριση πελατειακών σχέσεων (CRM), σημειώνοντας ότι οι επιχειρήσεις μπορούν να στοχεύουν σε πελάτες υψηλής αξίας που παρουσιάζουν συχνή αγοραστική συμπεριφορά ή υψηλή χρηματική αξία. Η τμηματοποίηση συμπεριφοράς χρησιμοποιείται συνήθως στο ηλεκτρονικό εμπόριο και το ψηφιακό μάρκετινγκ, όπου τα δεδομένα των πελατών είναι άμεσα διαθέσιμα, επιτρέποντας τη στόχευση και την εξατομίκευση σε πραγματικό χρόνο (Kumar et al., 2006).

Για παράδειγμα, οι Bauer και Hammerschmidt (2005) χρησιμοποίησαν δεδομένα συμπεριφοράς για να τμηματοποιήσουν τους πελάτες ενός διαδικτυακού λιανοπωλητή, εντοπίζοντας συγκεκριμένα μοτίβα περιήγησης και αγοραστικής συμπεριφοράς που επέτρεψαν στην εταιρεία να βελτιστοποιήσει τις στρατηγικές μάρκετινγκ και προώθησης.

* ***Τμηματοποίηση της επιχείρησης***

Στις αγορές B2B (business-to-business), χρησιμοποιείται συνήθως η εταιριογραφική τμηματοποίηση. Η μέθοδος αυτή διαιρεί τις εταιρείες με βάση χαρακτηριστικά όπως ο κλάδος, το μέγεθος της εταιρείας, τα έσοδα και η τοποθεσία. Σύμφωνα με τους Gordon και Chatterjee (2003), η εταιριογραφική τμηματοποίηση είναι απαραίτητη για τις επιχειρήσεις που πωλούν προϊόντα ή υπηρεσίες σε άλλες επιχειρήσεις, διότι τους επιτρέπει να στοχεύουν σε επιχειρήσεις που είναι πιθανότερο να έχουν ανάγκη για τις προσφορές τους με βάση αυτά τα χαρακτηριστικά. Για παράδειγμα, οι Shaw και Adamson (2005) μελέτησαν την εφαρμογή των εταιριογραφικών δεδομένων στον κλάδο των υπηρεσιών πληροφορικής, δείχνοντας ότι οι επιχειρήσεις με μεγαλύτερα έσοδα και βάσεις εργαζομένων είχαν σημαντικά διαφορετική αγοραστική συμπεριφορά σε σύγκριση με τις μικρότερες επιχειρήσεις.

## Ανάλυση RFM

### 2.1 Προέλευση και υλοποίηση

Η ανάλυση RFM είναι μια τεχνική τμηματοποίησης που χρησιμοποιείται στο μάρκετινγκ και τη διαχείριση πελατειακών σχέσεων (Customer Relationship Management) για την κατηγοριοποίηση των πελατών με βάση την αγοραστική τους συμπεριφορά. Η μέθοδος αναπτύχθηκε κατά τις δεκαετίες του 1960 και 1970 αλλά εδραιώθηκε στον τομέα του μάρκετινγκ βάσεων δεδομένων με το έργο του Arthur Hughes (1994), ο οποίος την παρουσίασε ως ένα πρακτικό εργαλείο άμεσου μάρκετινγκ (direct marketing) για τη βελτίωση των ποσοστών ανταπόκρισης στις εκστρατείες μάρκετινγκ. Η ονομασία της προέρχεται από τα αρχικά των παραμέτρων Recency, Frequency και Monetary, οι οποίες αναφέρονται στα εξής δεδομένα:

* Recency (R): Η πιο πρόσφατη αγορά ενός πελάτη, και συγκεκριμένα το διάστημα που μεσολάβησε από την τελευταία αγορά που πραγματοποίησε μέχρι την χρονική στιγμή που ορίζεται στην ανάλυση.
* Συχνότητα (F): Η συχνότητα των αγορών που πραγματοποιεί ο καταναλωτής, και ειδικότερα ο αριθμός των συναλλαγών που πραγματοποίησε σε μια ορισμένη χρονική περίοδο.
* Νομισματική αξία (Μ): Tο ποσό των χρημάτων που ο πελάτης διέθεσε στην επιχείρηση κατά την ορισμένη αυτή χρονική περίοδο.

Σύμφωνα με τον Hughes (1994), το πρώτο βήμα της μεθόδου RFM είναι η ταξινόμηση των εγγραφών των πελατών ανάλογα με το πόσο πρόσφατα έχουν αγοράσει από την επιχείρηση (Recency). Στη συνέχεια, η βάση δεδομένων χωρίζεται σε ίσα πεμπτημόρια (quintiles) και σε αυτά τα πεμπτημόρια αποδίδονται οι βαθμολογίες 5 έως 1. Επομένως, το 20% των πελατών που αγόρασαν πιο πρόσφατα από την εταιρεία λαμβάνει τη βαθμολογία 5, το επόμενο 20% λαμβάνει τον αριθμό 4 και ούτω καθεξής. Το επόμενο βήμα περιλαμβάνει την ταξινόμηση των πελατών σε κάθε πεμπτημόριο Recency με βάση την συχνότητα των αγορών τους (Frequency). Τους αποδίδεται μία βαθμολογία από το 5 έως το 1 για τη συχνότητα. Κάθε μία από αυτές τις ομάδες (25 ομάδες) ταξινομείται ανάλογα με το πόσα χρήματα έχουν ξοδέψει οι πελάτες στην εταιρεία (Monetary). Οι κατηγορίες αυτές χωρίζονται σε πεμπτημόρια και τους αποδίδονται βαθμολογίες από 5 έως 1. Επομένως, η βάση δεδομένων χωρίζεται σε 125 ίσες ομάδες σύμφωνα με τις τρεις μετρικές. Ο Miglautsch (2000) σημειώνει το πλεονέκτημα της μεθόδου των πεμπτημορίων για την προβολή της συμπεριφοράς των πελατών εφόσον τα σχήματα τμηματοποίησης δημιουργούνται περιοδικά. Ωστόσο, υποστηρίζει ότι το κύριο μειονέκτημα της μεθόδου των πεμπτημορίων είναι η τάση της να «ομαδοποιεί πελάτες που έχουν πολύ διαφορετική αγοραστική συμπεριφορά (στην κορυφή) και να διαχωρίζει αυθαίρετα τους πελάτες που έχουν πανομοιότυπη συμπεριφορά (στη βάση)».

O Hughes (1994) τόνισε πώς η ανάλυση των δεδομένων αγοράς των πελατών μέσω των παραμέτρων RFM θα μπορούσε να βοηθήσει τις επιχειρήσεις να εντοπίσουν τους πιο πολύτιμους πελάτες τους και να σχεδιάσουν εξατομικευμένες προσφορές, οδηγώντας σε πιο ισχυρή δέσμευση και αύξηση των πωλήσεων. Η δημοτικότητα της ανάλυσης RFM πηγάζει από την απλότητα και την αποτελεσματικότητα της εφαρμογής της, καθώς είναι μία μέθοδος εύκολα κατανοητή από τις διοικήσεις και τα επιχειρησιακά στελέχη που λαμβάνουν τις αποφάσεις (Marcus, 1998). Επιπλέον, χρησιμοποιείται ευρέως επειδή συνδέει άμεσα τη συμπεριφορά των πελατών με την κερδοφορία, επιτρέποντας στις επιχειρήσεις να βελτιστοποιήσουν τις στρατηγικές μάρκετινγκ τους εστιάζοντας στους πελάτες που είναι πιο πιθανό να ανταποκριθούν σε μελλοντικές προωθητικές ενέργειες (Fader, 2009).

Οι Dursun και Caber (2016) αναφέρουν τα εξής ως κύρια πλεονεκτήματα της RFM ανάλυσης:

* Είναι ένα ισχυρό εργαλείο για την αξιολόγηση της αξίας διάρκειας ζωής των πελατών (customer life value – CLV), το οποίο είναι επίσης ικανό να συνδυαστεί με τεχνικές εξόρυξης συχνών προτύπων/μοτίβων (Hu & Yeh, 2014).
* Θεωρείται ως «βάση για μια συνεχή ροή τεχνικών για τη βελτίωση της τμηματοποίησης των πελατών» (Elsner, Krafft, & Huchzemeier, 2003).
* Είναι ένα αποτελεσματικό εργαλείο για την πρόβλεψη της ανταπόκρισης των πελατών σε εκστρατείες μάρκετινγκ και την ενίσχυση των κερδών της εταιρείας σε σύντομο χρονικό διάστημα (Baecke & Van den Poel, 2011).

Από την άλλη μεριά, οι συγγραφείς συγκέντρωσαν τα παρακάτω μειονεκτήματα της ανάλυσης RFM:

* Η μέθοδος, χρησιμοποιώντας μόνο τις τρεις μετρικές R, F και M, αγνοεί ορισμένα άλλα σημαντικά χαρακτηριστικά των πελατών όπως η ηλικία, το εισόδημα, ο τρόπος ζωής και οι παραλλαγές των προϊόντων, γεγονός που την καθιστά ανεπαρκή για τη δημιουργία επιτυχημένων προγραμμάτων μάρκετινγκ (Fitzpatrick, 2001).
* Οι τρεις μετρικές τείνουν να συσχετίζονται σε μεγάλο βαθμό, με τις υψηλότερες συσχετίσεις να παρατηρούνται μεταξύ των μετρικών Frequency και Monetary (Olson et al., 2009).
* Η ανάλυση δεν λαμβάνει υπ’όψιν της τους δυνητικούς και τους μη κερδοφόρους πελάτες μιας επιχείρησης όπως και την διαφορετική σημασία που έχουν οι μετρικές R, F και M από κλάδο σε κλάδο (Băcilă, Rădulescu, & Marar, 2012).

### 2.2 Παραλλαγές της ανάλυσης RFM

Η ανάλυση RFM έχει εξελιχθεί σημαντικά από την αρχική της διάδοση, με διάφορους ερευνητές και επαγγελματίες να εισάγουν τροποποιήσεις και επεκτάσεις ώστε να ενισχύσουν την αποτελεσματικότητά της σε διαφορετικά επιχειρηματικά πλαίσια. Κάποιες από τις εν λόγω παραλλαγές ενσωματώνουν πρόσθετες διαστάσεις/μεταβλητές συμπεριφοράς, ενώ κάποιες αποσκοπούν στη βελτίωση της ακρίβειας τμηματοποίησης ή στην προσαρμογή της ανάλυσης RFM σε συγκεκριμένους κλάδους. Ακολουθούν ορισμένες σημαντικές παραλλαγές της ανάλυσης RFM που έχουν δημοσιευτεί με την πάροδο των ετών.

Ένα από τα πιο δημοφιλή μοντέλα τμηματοποίησης είναι το **LRFM** μοντέλο. Οι Reinartz και Kumar (2000) ισχυρίστηκαν ότι το μοντέλο RFM δεν μπορεί να τμηματοποιήσει τους πελάτες που έχουν είτε μακροχρόνια είτε βραχυχρόνια σχέση με την εταιρεία. Υπό αυτές τις συνθήκες, οι Chang και Tsay (2004) εισήγαγαν το μοντέλο LRFM προσθέτοντας την μεταβλητή *Length* (L) στο μοντέλο RFM η οποία αναφέρεται στη διάρκεια της σχέσης του πελάτη με την επιχείρηση, αφού υπολογίζεται ως το χρονικό διάστημα, σε ημέρες, μεταξύ της πρώτης και της τελευταίας επίσκεψης του πελάτη. Όσο μεγαλύτερη είναι η τιμή της μεταβλητής L ενός πελάτη, τόσο μεγαλύτερη αφοσίωση δείχνει ο πελάτης στην επιχείρηση ή στο προϊόν όταν οι άλλες τρεις μετρικές παραμένουν σταθερές. Οι Kao et al. (2011) χρησιμοποίησαν το LRFM μοντέλο σε συνδυασμό με τον αλγόριθμο K-means και τμηματοποίησαν τους πελάτες μιας επιχείρησης ανδρικών ρούχων σε 12 ομάδες.

Οι Peker, Kocyigit και Eren (2017) εισήγαγαν μία νέα προσέγγιση τμηματοποίησης πελατών, το μοντέλο **LRFMP**, μια επέκταση του παραδοσιακού μοντέλου LRFM. Το μοντέλο LRFMP προσθέτει τη διάσταση *Periodicity* (P), ενισχύοντας τη δυνατότητα αποτελεσματικότερης ταξινόμησης των πελατών με βάση την αγοραστική τους συμπεριφορά. Η μετρική P αντικατοπτρίζει το αν οι πελάτες επισκέπτονται τακτικά τα καταστήματα και ορίζεται ως η τυπική απόκλιση των χρόνων μεταξύ των επισκέψεων των πελατών. Εάν ένας πελάτης έχει χαμηλή τιμή περιοδικότητας, σημαίνει ότι ο πελάτης αυτός επισκέπτεται ή πραγματοποιεί αγορές σε σχετικά σταθερά διαστήματα και μπορεί να χαρακτηριστεί ως τακτικός. Επιπλέον, η μετρική Recency τροποποιήθηκε για την συγκεκριμένη ανάλυση, καθώς για τον υπολογισμό της λήφθηκαν υπόψη οι τελευταίες N συναλλαγές του πελάτη αντί να λαμβάνεται υπόψη μόνο η πιο πρόσφατη συναλλαγή του. Συγκεκριμένα, η μεταβλητή R ορίστηκε ως η μέση τιμή του αριθμού των ημερών μεταξύ των ημερομηνιών των Ν πρόσφατων επισκέψεων του πελάτη και της τελευταίας ημερομηνίας της περιόδου παρατήρησης. Στη μεθοδολογία οι συγγραφείς χρησιμοποίησαν το μοντέλο LRFMP σε συνδυασμό με τον αλγόριθμο μηχανικής μάθησης K-means για την τμηματοποίηση των πελατών μιας αλυσίδας καταστημάτων που δραστηριοποιείται στον κλάδο λιανικής πώλησης ειδών παντοπωλείου στην Τουρκία. Οι ερευνητές τόνισαν την αναγκαιότητα για αποτελεσματική τμηματοποίηση των πελατών των επιχειρήσεων του συγκεκριμένου κλάδου, καθώς ο ανταγωνισμός είναι ιδιαίτερα έντονος, κάτι που παράγει επιπλέον πίεση στους επιχειρηματίες οι οποίοι πασχίζουν να διαχειριστούν αποδοτικά την πελατειακή τους βάση και να αποκτήσουν ανταγωνιστικό πλεονέκτημα μέσα σε αυτό το κορεσμένο περιβάλλον. Σε αυτή την κατεύθυνση το μοντέλο LRFMP μπορεί να τους παράσχει χρήσιμες πληροφορίες για τα διαφορετικά προφίλ πελατών, να βοηθήσει τους υπεύθυνους λήψης αποφάσεων να αναπτύξουν αποτελεσματικές σχέσεις με τους πελάτες και να κατανείμουν με αποδοτικό τρόπο τους διαθέσιμους πόρους των διαφόρων στρατηγικών μάρκετινγκ.

Στην έρευνά τους, οι Liu και Shih (2005) ανέπτυξαν μία νέα μεθοδολογία σύστασης προϊόντων (product recommendation system) δίνοντας ιδιαίτερη έμφαση στην αξία διάρκειας ζωής του πελάτη (CLV) η οποία προσδιορίζεται από τη σταθμισμένη RFM (**Weighted RFM**). Η προτεινόμενη μεθοδολογία χρησιμοποιεί κυρίως τις τεχνικές της Mεθόδου Αναλυτικής Ιεράρχησης (Analytic Hierarchy Process), συσταδοποίησης (clustering) και εξόρυξης κανόνων συσχέτισης (association rule mining techniques). Αρχικά οι τιμές R, F και M χρησιμοποιήθηκαν για την τμηματοποίηση των πελατών σε ομάδες με παρόμοιες τιμές. Η μέθοδος AHP χρησιμοποιήθηκε για την απόδοση βάρους (weight) σε κάθε μεταβλητή R, F, και M. Στη συνέχεια, με βάση τη σταθμισμένη RFM, χρησιμοποιήθηκε η ομαδοποίηση K-means για την ταξινόμηση των καταναλωτών με συγκρίσιμες αξίες διάρκειας ζωής ή επίπεδα αφοσίωσης. Έπειτα εφαρμόστηκε μια προσέγγιση εξόρυξης κανόνων συσχέτισης για την εξαγωγή κανόνων σύστασης, δηλαδή συχνών μοτίβων αγοράς από κάθε ομάδα πελατών. Τα εξαγόμενα συχνά μοτίβα αγοράς αντιπροσωπεύουν την κοινή αγοραστική συμπεριφορά των πελατών με παρόμοιες αγορές προϊόντων. Συνεπώς η συγκεκριμένη μεθοδολογία συνιστά προϊόντα στους πελάτες με βάση τα συχνά μοτίβα αγοράς πελατών με παρόμοιες αγορές προϊόντων.

Οι Lang et al. (2022) παρατηρούν ότι στην εποχή των «μεγάλων» δεδομένων, η εφαρμογή στατικών μοντέλων βαρών αποδεικνύεται λιγότερο αποτελεσματική, οδηγώντας σε μη βέλτιστη τμηματοποίηση και στόχευση πελατών. Οι συγγραφείς προτείνουν μια δυναμική προσέγγιση στάθμισης (**dynamic weighted RFM approach**) για κάθε μία από τις μετρικές R, F και M. Αρχικά υπολογίζεται το *υποκειμενικό* βάρος (subjective weight) με τη χρήση της Mεθόδου Αναλυτικής Ιεράρχησης (Analytic Hierarchy Process - AHP), ενώ στη συνέχεια εφαρμόζεται η μέθοδος Εντροπίας (Entropy) για τον υπολογισμό του *αντικειμενικού* βάρους (objective weight). Το τελικό βάρος της κάθε μετρικής προκύπτει από την Integrated Weighting Method. Οι συγγραφείς αρχικά εφάρμοσαν την εν λόγω μέθοδο χρησιμοποιώντας δεδομένα πωλήσεων CD ενός ηλεκτρονικού καταστήματος. Στη συνέχεια χρησιμοποίησαν ένα πολύ ευρύτερο σύνολο δεδομένων που αφορά το ηλεκτρονικό εμπόριο (e-commerce) της Βραζιλίας, θέλοντας να επικυρώσουν την επεκτασιμότητα και την αποδοτικότητα του προτεινόμενου μοντέλου τους σε ένα μοντέρνο και ποικιλόμορφο περιβάλλον ηλεκτρονικού εμπορίου. Οι δύο αναλύσεις έδειξαν ότι η δυναμική απόδοση βαρών στο μοντέλο RFM επηρεάζει θετικά την απόδοση τμηματοποίησης πελατών.

Η έρευνα των Yeh et al. (2009) επέκτεινε το μοντέλο RFM σε μοντέλο **RFMTC** με τη συμπερίληψη των εξής δύο παραμέτρων: της *Time since first purchase* (Τ) και της *Churn probability* (C). Η μετρική *T* αναφέρεται στον χρόνο που μεσολαβεί από την πρώτη αγορά που πραγματοποίησε ο πελάτης από την επιχείρηση, ενώ η μετρική *C* αναφέρεται στην πιθανότητα αποχώρησής του από το πελατολόγιο της επιχείρησης. Βασιζόμενοι στην ακολουθία Bernoulli της θεωρίας πιθανοτήτων, οι συγγραφείς εξήγαγαν τύπους (φόρμουλες) που μπορούν να προβλέψουν την πιθανότητα ότι ένας πελάτης θα αγοράσει την επόμενη φορά από την επιχείρηση, καθώς και τον αναμενόμενο συνολικό αριθμό των φορών που ο πελάτης θα αγοράσει στο μέλλον από την επιχείρηση. Υπογραμμίζεται ότι το μοντέλο RFMTC αποτελεί ένα ακριβές μοντέλο ποσοτικής πρόβλεψης της συμπεριφοράς των πελατών το οποίο, σε αντίθεση με το παραδοσιακό μοντέλο RFM, δεν χρησιμοποιεί πρόχειρες μεθόδους κατηγοριοποίησής τους, όπως ο χωρισμός τους σε πέντε ισόποσα τμήματα. Ένα ακόμη πλεονέκτημα της συγκεκριμένης μεθόδου είναι ότι δεν προσαρμόζει κάθε φορά τα βάρη των παραμέτρων R, F, M ανάλογα με τις διάφορες βιομηχανίες, αλλά μπορεί να οικοδομήσει αυτόματα το βέλτιστο προγνωστικό μοντέλο με βάση τα δεδομένα των βάσεων δεδομένων μάρκετινγκ διαφορετικών κλάδων. Αναφέρεται επίσης ότι το μοντέλο RFMTC δεν χρειάζεται να τμηματοποιήσει τους πελάτες σε διαφορετικές ομάδες ώστε να επιβεβαιώσει το ποσοστό ανταπόκρισης κάθε ομάδας στις ενέργειες μάρκετινγκ της επιχείρησης αλλά χρησιμοποιεί μία μόνο ομάδα πελατών, μειώνοντας σημαντικά την αναγκαία ποσότητα δοκιμών σε πελάτες.

Η έρευνα των Heldt et al. (2021) αποσκοπεί στη βελτίωση της πρόβλεψης της αξίας των πελατών εισάγοντας μία νέα επέκταση του παραδοσιακού μοντέλου RFM, το μοντέλο *RFM per Product* (**RFM/P**). Η μεταβλητή P (product) ενσωματώνει στην ανάλυση δεδομένα για τα συγκεκριμένα προϊόντα που αγοράζουν οι πελάτες, επιτρέποντας έτσι στις επιχειρήσεις να διαφοροποιούν την αξία των πελατών ανά προϊόν, αντί να συγκεντρώνουν όλες τις αγορές μαζί. Το προτεινόμενο μοντέλο παρέχει τον προσδιορισμό των προϊόντων που αφορούν τους πιο πολύτιμους πελάτες και των πελατών που αγοράζουν τα πιο κερδοφόρα προϊόντα. Η εστίαση αποκλειστικά στην κερδοφορία των προϊόντων μπορεί να οδηγήσει την εταιρεία στη διαδικασία που είναι γνωστή ως «σπείρα θανάτου». Από την άλλη πλευρά, η εστίαση μόνο στην κερδοφορία των πελατών μπορεί να οδηγήσει σε αυξημένο συνολικό κίνδυνο της επιχείρησης, ενθαρρύνοντας ενδεχομένως την υπερβολική συγκέντρωση των προσπαθειών μάρκετινγκ σε μια μικρή ομάδα πελατών. Το μοντέλο RFM/P συνδυάζει τις δύο προοπτικές – προϊόντων και πελατών – δίνοντας τη δυνατότητα στους διαχειριστές να εντοπίσουν ευκαιρίες για βελτιώσεις προϊόντων και υπηρεσιών ώστε να ταιριάζουν καλύτερα οι προσφορές της εταιρείας στους πελάτες-κλειδιά, να δρομολογήσουν επεκτάσεις εμπορικών σημάτων για πολύτιμες υφιστάμενες κατηγορίες προϊόντων για την απόκτηση νέων πελατών και να επιτρέψουν στρατηγικές μάρκετινγκ που έχουν θετικό αναμενόμενο αντίκτυπο στην αξία των πελατών (CLV).

### 2.3 Τμήματα RFM: Χαρακτηριστικά & προτεινόμενες στρατηγικές μάρκετινγκ

Οι υπεύθυνοι αποφάσεων και οι διαχειριστές είναι σε θέση να κατανοήσουν και να εφαρμόσουν εύκολα αυτήν την τεχνική εξόρυξης δεδομένων (Marcus, 1998). Από την τμηματοποίηση που προκύπτει από την εφαρμογή του μοντέλου, μπορούν να αναπτυχθούν προσαρμοσμένες στρατηγικές μάρκετινγκ για συγκεκριμένες ομάδες πελατών (Ernawati et al., 2021).

* ***Champions***

Πρόκειται για τους πελάτες που έχουν πολύ πρόσφατα πραγματοποιήσει αγορές από την επιχείρηση (υψηλό R), έχουν αγοράσει πολλές φορές (υψηλό F) καταναλώνοντας μεγάλα ποσά χρηματικών μονάδων (υψηλό M). Στόχος της επιχείρησης είναι να διατηρήσει την ικανοποίησή τους σε υψηλά επίπεδα καθώς οι πελάτες αυτοί είναι υπεύθυνοι για ένα μεγάλο μερίδιο των εσόδων της αλλά και για την διαφήμιση των προιόντων της μέσω θετικών σχολίων (positive word of mouth). Είναι επίσης οι πελάτες που είναι πιο πιθανό να αγοράσουν πρώτοι τα νέα προϊόντα που θα λανσάρει η επιχείρηση θέτοντας τη βάση για την καθιέρωσή τους στην αγορά. Η επιχείρηση θα πρέπει να εστιάσει σε ενέργειες μάρκετινγκ με στόχο την επιβράβευση αυτής της ομάδας πελατών με ενέργειες όπως η παροχή αποκλειστικών προσφορών, η προτεραιότητα στη διάθεση των προϊόντων και στην εξυπηρέτηση, και οι εξατομικευμένες προσπάθειες προσέγγισής τους. Τέλος, σε περιπτώσεις που υλοποιείται έρευνα για τον εντοπισμό πιθανών μειονεκτημάτων/αστοχιών των παραχόμενων προϊόντων/υπηρεσιών αλλά και νέων τάσεων στις προτιμήσεις των καταναλωτών, η ομάδα ‘champions’ θα παρέχουν την πιο αξιόπιστη ανατροφοδότηση.

* ***Loyal Customers***

Οι πελάτες αυτής της κατηγορίας αγόρασαν πολύ πρόσφατα από την επιχείρηση (υψηλό R), παρουσιάζουν συνήθως υψηλή συχνότητα αγορών (υψηλό F) για τις οποίες διαθέτουν σημαντικά χρηματικά ποσά (υψηλό Μ). Είναι πολύτιμοι και ιδιαίτερα αφοσιωμένοι πελάτες καθώς ανταποκρίνονται θετικά στις καμπάνιες μάρκετινγκ. Λαμβάνουν ενημερωτικά δελτία και ενημερώνονται για τα νέα της επιχείρησης. Η επιχείρηση θα πρέπει να τους κρατήσει χρησιμοποιώντας στρατηγικές upselling και cross-selling για να «κατασκευάσει» νέες ανάγκες που θα τους κεντρίσουν το ενδιαφέρον καταλήγοντας σε περισσότερες ή πιο δαπανηρές αγορές. Θα πρέπει να εφαρμόζονται στην πράξη εξατομικευμένες ενέργειες μάρκετινγκ, προσωποποιημένη εξυπηρέτηση καθώς και κάθε άλλη δραστηριότητα που μπορεί να ενισχύσει την αφοσίωσή τους στην επιχείρηση.

* ***Potential Loyalists***

Οι πρόσφατες αγορές (υψηλό R), οι συχνές αγορές (υψηλό F) και οι δαπάνες μέτριου ύψους (μέτριο M) αποτελούν τυπικά χαρακτηριστικά συμπεριφοράς αυτής της κατηγορίας καταναλωτών. Αν και οι πελάτες αυτοί έχουν πραγματοποιήσει στο παρελθόν αρκετές αγορές από την επιχείρηση, το μέγεθος του καλαθιού τους δεν ήταν πολύ μεγάλο. Οι ενέργειες μάρκετινγκ θα πρέπει να επικεντρωθούν στην ενθάρρυνση των συγκεκριμένων πελατών να αυξήσουν τις αγορές τους μέσω προτάσεων για συμπληρωματικά προϊόντα ή επιπλέον προϊόντα αλλά και μέσω ενεργειών που θα τους κάνουν να νιώσουν πολύτιμοι και να επομένως να αυξήσουν την αφοσίωσή τους στην επιχείρηση.

* ***Promising***

Οι πελάτες αυτής της κατηγορίας πραγματοποίησαν πρόσφατα κάποια αγορά (υψηλό R), αγοράζουν με μέτρια συχνότητα από την επιχείρηση (μέτριο F) και έχουν διαθέσει ένα αρκετά μεγάλο χρηματικό ποσό (υψηλό Μ). Παρόλο που αγόρασαν πολύ πρόσφατα από την επιχείρηση, δεν είναι οι πελάτες που πραγματοποιούν αγορές σε τακτά χρονικά διαστήματα. Για να τους προσελκύσει η επιχείρηση θα πρέπει να εστιάσει σε στρατηγικές μάρκετινγκ που θα τους δελεάσουν να αυξήσουν την συχνότητα των αγορών τους, όπως είναι η προσφορά προγραμμάτων συνδρομής, η παροχή προτάσεων για προϊόντα, η αποστολή μικρών δώρων με κάθε αγορά καθώς και η ενθάρρυνσή τους να κάνουν αξιολόγηση των προϊόντων/υπηρεσιών της επιχείρησης.

* ***New Customers***

Οι πελάτες αυτής της ομάδας έχουν πρόσφατα αγοράσει από την επιχείρηση (υψηλό R), δεν αγοράζουν όμως συχνά (χαμηλό F) και έχουν δαπανήσει ποσά μετρίου ή χαμηλού ύψους (μέτρια M). Για κάποιους από αυτούς ίσως είναι η πρώτη φορά που αγοράζουν από την επιχείρηση. Η στρατηγική μάκρετινγκ που συνήθως ακολουθείται στοχεύει κυρίως στη διερεύνηση των απαιτήσεων και των προτιμήσεών των συγκεκριμένων πελατών όπως και στην προσφορά κινήτρων για αγορά. Η επιχείρηση μπορεί να τους παρέχει ένα εκπτωτικό κουπόνι για μελλοντικές αγορές ή να τους προσφέρει ένα δώρο με την πρώτη αγορά. Επιπλέον μια έρευνα ικανοποίησης μέσω τηλεφωνικού μάρκετινγκ ή μιας ιστοσελίδας όπου θα αξιολογήσουν τα προϊόντα, μπορεί να τους κινητοποιήσει σημαντικά. Ο στόχος των προσπαθειών μάρκετινγκ είναι να ξεχωρίσει η επιχείρηση από τον ανταγωνισμό στα μάτια του καταναλωτή και συνεπώς να στραφεί ολοκληρωτικά προς αυτή.

* ***Need Attention***

Οι πελάτες αυτής της κατηγορίας αγόρασαν πρόσφατα ή σχετικά πρόσφατα από την επιχείρηση (μέτριο R), αγοράζουν με μέτρια συχνότητα (μέτριο F) και διαθέτουν υψηλά ή μέτρια χρηματικά ποσά (μέτριο M). Για να μετατρέψει αυτούς τους πελάτες σε τακτικούς πελάτες, η επιχείρηση είναι κρίσιμο να αναδείξει τις ιδιαίτερες ιδιότητες των αγαθών/υπηρεσιών που προσφέρει και να παρέχει κίνητρα για περισσότερες αγορές. Οι προσφορές περιορισμένης χρονικής διάρκειας, οι ενέργειες εξατομικευμένου μάρκετινγκ, το επιθετικό μάρκετινγκ, οι εξατομικευμένες επικοινωνίες και τα χαμηλού κόστους προγράμματα αυτοματοποιημένης προώθησης αποδεικνύονται συνήθως αποτελεσματικές ενέργειες για την προσέλκυση αυτής της ομάδας πελατών.

* ***About to Sleep***

Οι πελάτες αυτοί έχουν λιγότερες πρόσφατες αγορές (μέτριο R), μέτρια ή χαμηλή συχνότητα αγορών (μέτριο F) και χαμηλές ή μέτριες δαπάνες (χαμηλή M). Οι πελάτες της συγκεκριμένης κατηγορίας δεν έχουν αγοράσει από την επιχείρηση για μεγάλο χρονικό διάστημα, αλλά όχι σε βαθμό που να είναι απρόσιτοι. Μπορούν να γίνουν ξανά ενεργοί με την χρήση των κατάλληλων τακτικών όπως η συνεχής πληροφόρηση. Κάποια κίνητρα για να τους προσελκύσει ξανά θα μπορούσαν να περιλαμβάνουν εκπτώσεις, προσφορές δώρων 1+1, μάρκετινγκ μέσω ηλεκτρονικού ταχυδρομείου, μηνύματα υπενθύμισης για προϊόντα με έκπτωση ή υπενθυμίσεις ότι έχουν καιρό να αγοράσουν, προϊόντα με έκπτωση παρόμοια με τις προηγούμενες προτιμήσεις τους.

* ***Can’t Lose Them***

Η ομάδα αυτή περιλαμβάνει πελάτες που έχουν καιρό να πραγματοποιήσουν κάποια αγορά (χαμηλό R), έχουν μέτρια ή υψηλή συχνότητα αγορών (μέτριο F) αλλά έχουν επενδύσει σηματνικά χρηματικά ποσά στην επιχείρηση (υψηλό M). Παρά το γεγονός ότι δεν πραγματοποιούν συχνές ή πρόσφατες αγορές, αυτοί οι καταναλωτές είναι πολύτιμοι για την επιχείρηση, επειδή τείνουν να πραγματοποιούν μεγάλες αγορές. Οι εξατομικευμένες τηλεφωνικές κλήσεις, οι εκστρατείες επιστροφής κέρδους, τα προγράμματα επιβράβευσης και η απαλλαγή από πρόσθετες χρεώσεις (όπως τα έξοδα αποστολής ή οι δωρεάν επιστροφές προϊόντων) αποτελούν στρατηγικές μάρκετινγκ που συνιστώνται για να τους κινητοποιήσουν. Μια εκστρατεία με προσφορές (δώρα ή εκπτώσεις) σε αγαθά που έχουν ήδη αγοράσει ή αναζητήσει θα ήταν μια κατάλληλη τακτική εκ νέου προσέλκυσής τους. Θα ήταν επίσης χρήσιμο για την επιχείρηση να λάβει τη συμβολή τους προκειμένου να κατανοήσει καλύτερα τις απαιτήσεις και τις προτιμήσεις τους. Θα πρέπει να διατεθούν πόροι στο τμήμα αυτό, δεδομένου ότι είναι ζωτικής σημασίας για την κερδοφορία του οργανισμού.

* ***At risk***

Οι καταναλωτές αυτοί συνήθως δεν ξοδεύουν αρκετά χρήματα (μέτριο Μ), δεν  
έχουν κάνει αρκετές αγορές πρόσφατα (χαμηλό R) και έχουν μέτρια  
συχνότητα (μέτριο F). Παρόλο που ένα σημαντικό ποσοστό των συναλλαγών που  
πραγματοποιούνται από αυτούς τους καταναλωτές δεν ολοκληρώνεται, είναι πιθανό  
να απαντήσουν σε προσπάθειες προσαρμογής της αγοραστικής τους εμπειρίας ή σε  
μηνύματα ηλεκτρονικού ταχυδρομείου που τούς υπενθυμίζουν τα είδη που δεν  
κατάφεραν να προσθέσουν στο καλάθι τους. Ανεξάρτητα από αυτό, πρέπει να  
διερευνηθεί το σκεπτικό που κρύβεται πίσω από τις ανολοκλήρωτες αγορές τους.

* ***Hybernating***

Η συμπεριφορά αυτής της ομάδας πελατών χαρακτηρίζεται συνήθως από  
χαμηλή συχνότητα (χαμηλό F), χαμηλές δαπάνες (χαμηλό M) και μη πρόσφατες αγορές  
(μέτριο R). Δεδομένου ότι αυτή η ομάδα πελατών έχει ήδη επιδείξει απροθυμία να  
συμμετάσχει σε πρωτοβουλίες προσέγγισης, δεν συνιστάται η διάθεση ειδικών πόρων  
για αυτούς.

* ***Lost***

Πρόκειται για τους καταναλωτές με τις χαμηλότερες δαπάνες (χαμηλό M), τις  
παλαιότερες αγορές (χαμηλό R) και τη χαμηλότερη συχνότητα αγορών (χαμηλό F). Η  
προσέγγιση νέων πελατών είναι συνήθως πιο επικερδής για μια επιχείρηση από το  
να δαπανά χρόνο και πόρους προσπαθώντας να διατηρήσει τους υπάρχοντες πελάτες.  
Τεχνικές όπως η αποστολή εκπαιδευτικών μηνυμάτων ηλεκτρονικού ταχυδρομείου, η  
προώθηση κατά τη διάρκεια των διακοπών ή των εκπτώσεων και η προσφορά  
εκπτωτικών κουπονιών με μέτρια ποσά αγοράς θα μπορούσαν να χρησιμεύσουν ως  
κάποιο είδος κινήτρου για μια νέα αγορά.

## Μηχανική Μάθηση

Η μηχανική μάθηση είναι ένας κλάδος της τεχνητής νοημοσύνης που επικεντρώνεται στη δημιουργία στατιστικών μοντέλων και αλγορίθμων που επιτρέπουν στους υπολογιστές να μαθαίνουν από δεδομένα και να κάνουν κρίσεις ή προβλέψεις χωρίς ρητό προγραμματισμό (Ανδρουτσόπουλος, 2019). Το θέμα αυτό έχει προσελκύσει το ενδιαφέρον πολλών ερευνητών τα τελευταία χρόνια λόγω της ικανότητάς του να εξάγει σημαντικές γνώσεις από μεγάλα σύνολα δεδομένων, καθιστώντας το ιδιαίτερα αποτελεσματικό στην επιχειρησιακή ανάλυση και τη διαχείριση πελατειακών σχέσεων (Brynjolfsson & McAfee, 2017). Οι Alzubi et al. (2018) αναφέρουν ότι η μηχανική μάθηση εφαρμόζεται σε πληθώρα προβλημάτων του πραγματικού κόσμου τα οποία παρουσιάζουν υψηλή πολυπλοκότητα. Για παράδειγμα, μέσω αυτής σχεδιάζονται και προγραμματίζονται αλγόριθμοι υψηλής απόδοσης για το φιλτράρισμα ανεπιθύμητων μηνυμάτων ηλεκτρονικού ταχυδρομείου (spam), την ανίχνευση απάτης σε κοινωνικά δίκτυα (fraud detection), τις διαδικτυακές χρηματιστηριακές συναλλαγές, την ανίχνευση προσώπων και σχημάτων, την ιατρική διάγνωση, την πρόβλεψη της κυκλοφορίας στους δρόμους της πόλης, την αναγνώριση χαρακτήρων και τα συστήματα σύστασης προϊόντων (product recommendation). Τα αυτοκινούμενα αυτοκίνητα της Google, το Netflix που παρουσιάζει τις ταινίες και τις σειρές που μπορεί να αρέσουν σε ένα άτομο, οι μηχανές ηλεκτρονικών συστάσεων -όπως οι προτάσεις φίλων στο Facebook, τα «περισσότερα αντικείμενα προς εξέταση» και «πάρε κάτι για τον εαυτό σου» στο Amazon και η ανίχνευση απάτης με πιστωτικές κάρτες- είναι όλα παραδείγματα εφαρμογής της μηχανικής μάθησης στον πραγματικό κόσμο.

Η θεμελιώδης ιδέα της μηχανικής μάθησης είναι ότι, σε αντίθεση με τον ρητό προγραμματισμό γενικών εφαρμογών, τα συστήματα είναι ικανά να μαθαίνουν από την εμπειρία και να λαμβάνουν αποφάσεις. Αυτό επιτρέπει στους αλγορίθμους μηχανικής μάθησης να βελτιώνουν την απόδοσή τους με την πάροδο του χρόνου και να προσαρμόζονται σε νέα δεδομένα (Mitchell, 1997).

### 3.1 Κατηγορίες Μηχανικής Μάθησης

Οι Mohammed et al. (2016) αναφέρουν ότι οι τέσσερις βασικές κατηγορίες αλγορίθμων μηχανικής μάθησης είναι η επιβλεπόμενη μάθηση (supervised learning), η μη επιβλεπόμενη μάθηση (unsupervised learning), η ημιεπιβλεπόμενη μάθηση (semi-supervised learning) και η ενισχυτική μάθηση (reinforcement learning). Παρακάτω παρουσιάζονται συνοπτικά τα κύρια χαρακτηριστικά της κάθε κατηγορίας:

* *Επιβλεπόμενη Mάθηση (Supervised Learning)*

Η έρευνα των Sandhya και Charanjeet (2016) αναφέρει ότι στη μάθηση με επίβλεψη παρέχεται ένα σύνολο δεδομένων που αποτελείται τόσο από χαρακτηριστικά όσο και από ετικέτες. Το έργο της επιβλεπόμενης μάθησης είναι να κατασκευάσει έναν εκτιμητή που είναι σε θέση να προβλέψει την ετικέτα ενός αντικειμένου δεδομένου του συνόλου χαρακτηριστικών του. Ο αλγόριθμος μάθησης λαμβάνει ένα σύνολο χαρακτηριστικών ως είσοδο (inputs) μαζί με τις αντίστοιχες σωστές εξόδους (outputs) και «μαθαίνει» συγκρίνοντας την πραγματική του έξοδο με τις σωστές εξόδους για να βρει σφάλματα. Στη συνέχεια τροποποιεί το μοντέλο αναλόγως. Η επιβλεπόμενη μάθηση χρησιμοποιείται συνήθως σε εφαρμογές όπου τα ιστορικά δεδομένα προβλέπουν πιθανά μελλοντικά γεγονότα. Κατά τη διαδικασία της εκπαίδευσης, ο αλγόριθμος επιβλεπόμενης μάθησης κατασκευάζει το μοντέλο πρόβλεψης. Μετά την εκπαίδευση, το προσαρμοσμένο μοντέλο θα προσπαθήσει να προβλέψει τις πιο πιθανές ετικέτες για ένα νέο σύνολο δειγμάτων X στα δεδομένα δοκιμής. Οι αλγόριθμοι επιβλεπόμενης μάθησης χρησιμοποιούνται κυρίως για την επίλυση προβλημάτων ταξινόμησης (classification) και προβλημάτων παλινδρόμησης (regression) (Sarker, 2021). Παραδείγματα αλγορίθμων αυτής της κατηγορίας αποτελούν τα δέντρα αποφάσεων (decision trees), η λογιστική παλινδρόμηση (logistic regression), και η γραμμική παλινδρόμηση (linear regression).

* *Μη επιβλεπόμενη Μάθηση (Unsupervised Learning)*

Η μη επιβλεπόμενη μάθηση λειτουργεί με δεδομένα χωρίς ετικέτες (unlabelled data) και αναζητά κρυφές ομαδοποιήσεις ή μοτίβα (Phill, 2024). Σε αυτή την κατηγορία ανήκουν οι αλγόριθμοι συσταδοποίησης (clustering algorithms), όπως είναι ο K-means, ένας από τους πιο απλούς στην εφαρμογή τους αλγορίθμους που όμως χρησιμοποιείται ευρέως από τους ερευνητές για την επίλυση ποικίλων προβλημάτων του πραγματικού κόσμου. Δουλεύοντας με δεδομένα που δεν έχουν ετικέτες, οι αλγόριθμοι μη επιβλεπόμενης μάθησης επιδιώκουν να εντοπίσουν μια φυσική ομαδοποίηση μέσα στα δεδομένα. Εξαιτίας αυτού είναι για παράδειγμα ιδανικοί για τον εντοπισμό υποκείμενων τάσεων στη συμπεριφορά των καταναλωτών χωρίς να χρειάζεται εκ των προτέρων γνώση των επιδιωκόμενων αποτελεσμάτων.

* *Ημιεπιβλεπόμενη Μάθηση (Semi-supervised Learning)*

Η μηχανική μάθηση με ημιεπίβλεψη είναι ένας συνδυασμός επιβλεπόμενων και μη επιβλεπόμενων μεθόδων μηχανικής μάθησης. Ενδέχεται να υπάρχουν περιπτώσεις όπου ορισμένες παρατηρήσεις είναι εφοδιασμένες με ετικέτες, αλλά η πλειονότητα των παρατηρήσεων δεν είναι επισημασμένες λόγω του υψηλού κόστους της επισήμανσης και της έλλειψης εξειδικευμένης ανθρώπινης γνώσης. Σε τέτοιες περιπτώσεις, οι αλγόριθμοι με ημιεπίβλεψη είναι οι καταλληλότεροι για τη δημιουργία μοντέλων (Alzubi et al., 2018). Η μάθηση με ημιεπίβλεψη μπορεί να χρησιμοποιηθεί σε προβλήματα όπως η ταξινόμηση, η παλινδρόμηση και η πρόβλεψη (Sandhya & Charanjeet, 2016).

* *Ενισχυτική Μάθηση (Reinforcement Learning)*

Η εκπαίδευση πρακτόρων λογισμικού και μηχανών ώστε να συμπεριφέρονται στο περιβάλλον τους με τρόπο που μεγιστοποιεί μία ανταμοιβή είναι γνωστή ως ενισχυτική μάθηση (Sutton & Barto, 2018). Ένας αλγόριθμος μαθαίνει πώς να συμπεριφέρεται σε ένα δεδομένο περιβάλλον μέσω της ενισχυτικής μάθησης, σύμφωνα με την οποία οι ενέργειες ανταμείβονται ή τιμωρούνται. Οι Portugal et al. (2018) αναφέρουν σαν παράδειγμα έναν αλγόριθμο μηχανικής μάθησης που παίζει παιχνίδια στον υπολογιστή εναντίον ενός αντιπάλου. Οι κινήσεις που οδηγούν σε νίκες (θετική ανατροφοδότηση) στο παιχνίδι πρέπει να μαθαίνονται και να επαναλαμβάνονται, ενώ οι κινήσεις που οδηγούν σε ήττες (αρνητική ανατροφοδότηση) πρέπει να αποφεύγονται. Σύμφωνα με τον Sarker (2021) η ενισχυτική μάθηση είναι ένα ισχυρό εργαλείο για την εκπαίδευση μοντέλων τεχνητής νοημοσύνης που μπορούν να βοηθήσουν στην αύξηση της αυτοματοποίησης ή στη βελτιστοποίηση της λειτουργικής αποδοτικότητας εξελιγμένων συστημάτων, όπως η ρομποτική, οι εργασίες αυτόνομης οδήγησης, η μεταποίηση και η εφοδιαστική αλυσίδα, ωστόσο, δεν προτιμάται για την επίλυση βασικών ή απλών προβλημάτων.

### 3.2 Αλγόριθμοι Συσταδοποίησης

Μια βασική μέθοδος της μη επιβλεπόμενης μηχανικής μάθησης είναι η συσταδοποίηση (clustering), η οποία χωρίζει μια συλλογή σημείων σε ομάδες ανάλογα με τις ομοιότητες και τις διαφορές τους (Sinaga & Yang, 2020). Αυτή η προσέγγιση χρησιμοποιείται συχνά στην αναγνώριση προτύπων και στην ανάλυση δεδομένων, ιδίως όταν γίνεται ανάλυση RFM για την τμηματοποίηση των καταναλωτών. Με βάση τις τιμές των μετρικών Recency, Frequency και Monetary, η συσταδοποίηση μπορεί να βοηθήσει στην ανάλυση RFM στη διαίρεση των πελατών σε διακριτές ομάδες (Devarapalli et al., 2022). Επειδή οι αλγόριθμοι συσταδοποίησης περιλαμβάνουν πολλές παραμέτρους, λειτουργούν συχνά σε χώρους πολλών διαστάσεων και πρέπει να διαχειριστούν θορυβώδη, ελλιπή και δειγματοληπτικά δεδομένα, η απόδοσή τους μπορεί να διαφέρει σημαντικά για διαφορετικές εφαρμογές και τύπους δεδομένων (Rodriguez, 2019). Για τους λόγους αυτούς, στη βιβλιογραφία έχουν προταθεί αρκετές διαφορετικές προσεγγίσεις συσταδοποίησης, μερικές από τις οποίες παρουσιάζονται στην παρούσα ενότητα.

* *Ιεραρχική Συσταδοποίηση (Hierarchical Clustering)*

Η ιεραρχική συσταδοποίησηδημιουργεί μια ιεραρχία συστάδων διαχωρίζοντας ή συνδυάζοντας προϋπάρχουσες συστάδες (Ding & He, 2002). Αυτή η μέθοδος είναι ιδιαίτερα χρήσιμη όταν ο αριθμός των συστάδων είναι ακανόνιστος ή αβέβαιος, επειδή μπορεί να αποκαλύψει εμφωλευμένες δομές συστάδων. Η δομή ενός δέντρου που δημιουργείται από την ιεραρχική συσταδοποίηση απεικονίζει τις σχέσεις μεταξύ των συστάδων. Εξαιτίας αυτού, μπορεί να είναι χρήσιμη για την παρουσίαση των συνδέσεων μεταξύ των διαφόρων τμημάτων καταναλωτών που βρέθηκαν από την ανάλυση RFM (John et al., 2023).

* *Αλγόριθμος K-Means*

Ένας από τους πιο δημοφιλείς και συχνά εφαρμοζόμενους αλγορίθμους συσταδοποίησης είναι ο αλγόριθμος K-means. Στόχος του είναι να χωρίσει n παρατηρήσεις σε k συστάδες, στις οποίες κάθε παρατήρηση ανήκει στη συστάδα που έχει τον πλησιέστερο μέσο όρο, χρησιμεύοντας ως πρωτότυπο της συστάδας. Ο μέσος όρος των παρατηρήσεων σε μια συγκεκριμένη συστάδα ορίζει το κέντρο της συστάδας (Alzubi et al., 2018). Ο εν λόγω αλγόριθμος φημίζεται για την απλότητα και την αποτελεσματικότητά του και λειτουργεί ιδιαίτερα καλά με σφαιρικές συστάδες. Ο αλγόριθμος K-means αναθέτει τα σημεία δεδομένων στο πλησιέστερο κεντροειδές με επαναληπτική ενημέρωση των κεντροειδών (centroids). Για πολλές εφαρμογές, εξακολουθεί να είναι μια από τις πιο επιτυχημένες και αποδοτικές τεχνικές ομαδοποίησης, παρόλο που προϋποθέτει σφαιρικές συστάδες [12]. Στην επόμενη ενότητα του παρόντος κεφαλαίου παρουσιάζονται αναλυτικότερα ο τρόπος εφαρμογής, τα πλεονεκτήματα καθώς και οι περιορισμοί του αλγορίθμου K-means.

* *Συσταδοποίηση βασισμένη στην πυκνότητα*

Η εύρεση συστάδων διαφορετικών μεγεθών και μορφών μπορεί να επιτευχθεί με τη βοήθεια αλγορίθμων συσταδοποίησης με βάση την πυκνότητα, ο πιο διάσημος από τους οποίους είναι ο αλγόριθμος DBSCAN (Density-Based Spatial Clustering of Applications with Noise). Οι αλγόριθμοι αυτής της κατηγορίας είναι ιδιαίτερα χρήσιμοι σε περιπτώσεις όπου ο αριθμός των συστάδων είναι αβέβαιος ή όταν τα δεδομένα περιλαμβάνουν μεγάλα επίπεδα θορύβου (Ester et al., 1996). Ο αλγόριθμος DBSCAN ομαδοποιεί τα σημεία ανάλογα με το πόσο κοντά και πυκνά είναι το ένα στο άλλο. Αυτό αυξάνει την ανθεκτικότητά του στο θόρυβο και τις ακραίες τιμές, γεγονός που θα μπορούσε να οδηγήσει σε ακριβέστερη κατάτμηση στην ανάλυση RFM (Monalisa & Kurnia, 2019).

### Συσταδοποίηση K-means

Οι Jin και Han (2011) αναφέρουν ότι τα βασικά βήματα της δημιουργίας συστάδων με τη χρήση του αλγορίθμου K-means είναι τα εξής:

1. Αρχικοποίηση μέσω τυχαίας επιλογής K κεντροειδών (centroids)
2. Σε κάθε σημείο δεδομένων ανάθεση του πλησιέστερου κεντροειδούς με βάση την Ευκλείδεια απόσταση (Maheswari, 2019).
3. Καθορισμός των νέων κεντροειδών που προκύπτουν από τον υπολογισμό του μέσου όρου των σημείων που έχουν ανατεθεί στην κάθε συστάδα
4. Επανάληψη των βημάτων 2 και 3 έως ότου ικανοποιηθεί κάποιο κριτήριο σύγκλισης (διαφορά στην τιμή της συνάρτησης-παραμόρφωσης) ή διακοπής (περάτωση προκαθορισμένου αριθμού επαναλήψεων).

Δεδομένων των αρχικών συνθηκών διασποράς των κεντροδειών, ο k-means μπορεί να προσδιορίσει τις βέλτιστες αναθέσεις συστάδων μέσω της παραπάνω επαναληπτικής διαδικασίας. Ωστόσο το τελικό αποτέλεσμα της συσταδοποίησης μπορεί να επηρεαστεί σημαντικά από την επιλογή των αρχικών κεντροειδών.

Η συσταδοποίηση με τον αλγόριθμο K-means προσφέρει τα ακόλουθα πλεονεκτήματα όταν συνδυάζεται με την ανάλυση RFM:

* **Κλιμακωσιμότητα (Scalability):** Οι σύγχρονες επιχειρήσεις έχουν να αντιμετωπίσουν μεγάλα σύνολα δεδομένων τα οποία ο αλγορίθμος K-means μπορεί να διαχειριστεί αποτελεσματικά. Η ανάλυση της συμπεριφοράς των πελατών σε πραγματικό χρόνο με βάση τις μετρικές RFM καθίσταται δυνατή χάρη στην υπολογιστική αποδοτικότητα του αλγορίθμου, η οποία του επιτρέπει να χειρίζεται γρήγορα τεράστιους όγκους δεδομένων (Shindler et al., 2011).
* **Ερμηνευσιμότητα (Interpretability):** Οι προκύπτουσες συστάδες μπορούν να επισημανθούν σύμφωνα με τις ιδιότητές τους και να ερμηνευθούν εύκολα. Μια συστάδα με υψηλή βαθμολογία Recency, για παράδειγμα, θα μπορούσε να ονομαστεί «Πρόσφατοι αγοραστές». Προκειμένου οι επιχειρήσεις να κατανοήσουν τα υποκείμενα μοτίβα στη συμπεριφορά των καταναλωτών και να δημιουργήσουν προσαρμοσμένες στρατηγικές ως απάντηση, η ερμηνευσιμότητα είναι απαραίτητη στην ανάλυση RFM.
* **Ταχύτητα:** Η μέθοδος K-means είναι κατάλληλη για εφαρμογές πραγματικού χρόνου και για συχνή επανασυγκέντρωση σε ομάδες όταν διατίθενται νέα δεδομένα, καθώς είναι συνήθως ταχύτερη από άλλες τεχνικές ομαδοποίησης. Λόγω της ταχύτητας του k-means, οι επιχειρήσεις μπορούν να παρακολουθούν συνεχώς τις αλλαγές στη συμπεριφορά των πελατών και να προσαρμόζουν γρήγορα τη στρατηγική τους ως απάντηση στις τάσεις της αγοράς ή των προτιμήσεων των πελατών (Fadaei & Khasteh, 2019).

Η συσταδοποίηση K-means χρησιμοποιείται ευρέως, ωστόσο έχει ορισμένα μειονεκτήματα και περιορισμούς:

* **Καθορισμός του βέλτιστου αριθμού συστάδων:** Η εύρεση του βέλτιστου αριθμού συστάδων (k) στις εφαρμογές K-means είναι ένα από τα πιο δύσκολα ζητούμενα. Για το σκοπό αυτό απαιτείται συχνά δοκιμή-και-σφάλμα (trial-and-error) ή η εφαρμογή άλλων μέτρων, όπως η μέθοδος του «αγκώνα» (Elbow Method) (Syakur et al., 2018) ή η βαθμολογία «σιλουέτας» (Silhouette Method) (Shutaywi & Kachouie, 2021). Παρόλο που οι τεχνικές αυτές μπορούν να είναι χρήσιμες για τον προσδιορισμό του ιδανικού αριθμού συστάδων, πρέπει και οι ίδιες να ερμηνεύονται προσεκτικά.
* **Ευαισθησία στις αρχικές συνθήκες:** Το τελικό αποτέλεσμα της ομαδοποίησης επηρεάζεται έντονα από την αρχική τοποθέτηση των κεντροειδών. Ακόμη και για το ίδιο σύνολο δεδομένων, διαφορετικές αρχικοποιήσεις μπορεί να οδηγήσουν σε διαφορετικές αναθέσεις συστάδων. Εξαιτίας αυτής της ευαισθησίας στις αρχικές συνθήκες, συνιστάται η εκτέλεση του k-means πολλές φορές χρησιμοποιώντας εναλλακτικές αρχικοποιήσεις, προκειμένου να διασφαλιστεί η συνέπεια των αποτελεσμάτων σε όλες τις αναλύσεις [24].
* **Υπόθεση σφαιρικών συστάδων:** Η μέθοδος K-means προϋποθέτει ότι κάθε συστάδα έχει την ίδια διακύμανση, κάτι που μπορεί να μην ισχύει σε πραγματικές καταστάσεις όπου οι συστάδες έχουν διαφορετικά μεγέθη και σχήματα. Τα τμήματα πελατών μπορεί να εμφανίζουν διαφορετικά επίπεδα ομοιογένειας στην ανάλυση RFM, γεγονός που αντιβαίνει στην υπόθεση των σφαιρικών συστάδων. Λόγω αυτού του περιορισμού, απαιτούνται εναλλακτικές στρατηγικές συσταδοποίησης ή προσαρμογές στον συμβατικό αλγόριθμο k-means [25].

## Συνδυασμός RFM και K-means

Στην παρούσα ενότητα παρουσιάζονται έρευνες που συνδυάζουν τη μέθοδο RFM με τον αλγόριθμο μηχανικής μάθησης K-means για την τμηματοποίηση των πελατών σε επιχειρήσεις που ανήκουν σε διαφορετικούς κλάδους. Ο αυξανόμενος αριθμός των ερευνών που συνδυάζουν πολλαπλές μεθόδους για την ανάπτυξη νέων μοντέλων τμηματοποίησης υποδηλώνει την κρισιμότητα της αποτελεσματικής διαχείρισης των πελατών τόσο για την βελτίωση της κερδοφορίας των επιχειρήσεων όσο και για την διατήρηση της ικανοποίησης των πελατών σε υψηλά επίπεδα.

Οι Maryani και Riana (2017) προσπάθησαν να δώσουν λύση στην πρόκληση μιας επιχείρησης να εντοπίσει δυνητικούς πελάτες ώστε να αξιοποιήσει το σύστημα διαχείρισης πελατειακών σχέσεων (CRM) για να υλοποιήσει το βέλτιστο δυνατό σχέδιο μάρκετινγκ για την προσέγγισή τους και τελικά για την αποκόμιση κέρδους για την ίδια. Οι συγγραφείς χρησιμοποίησαν το μοντέλο RFM σε δεδομένα συναλλαγών μιας εταιρίας της βιομηχανίας εξάτμισης μοτοσικλετών και αυτοκινήτων για την κατηγοριοποίηση των πελατών της. Τα στάδια της έρευνας που διεξήχθη αποτελούνται από τέσσερις δραστηριότητες. Αρχικά πραγματοποιήθηκε η ομαδοποίηση των δεδομένων συναλλαγών με τη χρήση του αλγορίθμου K-means. Το δεύτερο στάδιο ήταν ο προσδιορισμός των χαρακτηριστικών κάθε συστάδας με τη μέθοδο του δέντρου αποφάσεων (Decision Tree Method). Κατά το τρίτο στάδιο πραγματοποιήθηκε αξιολόγηση του προφίλ του πελάτη με τη χρήση της οικονομικής θεωρίας του Grid Hill, η οποία διαμορφώνει τη σύσταση 4 χαρακτηριστικών πελατών που ταιριάζουν με τη συναλλαγή εξόρυξης δεδομένων. Στο τελικό στάδιο δημιουργήθηκαν εφαρμογές χαρτογράφησης πελατών με την παραγωγή συστάσεων σε κάθε προφίλ πελάτη ώστε να μεγιστοποιηθεί η ικανοποίησή τους και έτσι η επιχείρηση να μπορεί να τους διατηρεί στο πελατολόγιό της. Τα αποτελέσματα της παρούσας μελέτης μπορούν να χρησιμοποιηθούν ως σύστημα υποστήριξης αποφάσεων στον κλάδο των μέσων ενημέρωσης για τη χαρτογράφηση πελατών και τη γνώση δυνητικών πελατών.

Η έρευνα των Gustriansyah et al. (2020) βασίστηκε στην ιδέα ότι η διαδικασία διατήρησης βάσεων δεδομένων προϊόντων για τη διαχείριση αποθεμάτων γίνεται όλο και πιο δύσκολη καθώς αυξάνεται ο όγκος των συναλλαγών. Οι συγγραφείς ισχυρίζονται ότι μια πιο αποτελεσματική στρατηγική για την επίλυση αυτού του ζητήματος θα ήταν η κατάτμηση όλων των προϊόντων στον κατάλληλο αριθμό συστάδων με βάση κάποιες από τις ομοιότητές τους με την χρήση κάποιας μεθόδου εξόρυξης δεδομένων. Οι αξίες των διαφόρων ομάδων μπορούν στη συνέχεια να υπολογιστούν και να αξιολογηθούν ώστε η διοίκηση να λάβει τεκμηριωμένες αποφάσεις και να κατανείμει τους πόρους με τον πλέον ορθολογικό τρόπο. Για την τεκμηρίωση των προαναφερθέντων οι ερευνητές χρησιμοποίησαν τον αλγόριθμο K-Means για να ομαδοποιήσουν τα δεδομένα των προϊόντων ενός φαρμακείου στην Palembang της Ινδονησίας με βάση τις τιμές RFM. Ο προσδιορισμός του βέλτιστου αριθμού k συστάδων στη μέθοδο k-Means αξιολογήθηκε με τη χρήση των εξής οκτώ δεικτών εγκυρότητας: της μεθόδου Elbow, του δείκτη Silhouette, του δείκτη Calinski-Harabasz, του δείκτη Davies-Bouldin, του δείκτη Ratkowski, του δείκτη Hubert, του δείκτη Ball-Hall και του δείκτη Krzanowski -Lai. Ο λόγος που χρησιμοποιήθηκαν όλοι οι παραπάνω δείκτες ήταν για τη βελτίωση της ακρίβειας στη διαδικασία διαχείρισης των αποθεμάτων και την υλοποίηση μιας πιο αντικειμενικής ομαδοποίησης των προϊόντων. Για τον αριθμό k που προέκυψε από κάθε δείκτη εγκυρότητας πραγματοποιήθηκε έλεγχος για την ποιότητα των αποτελεσμάτων τμηματοποίησης. Ο εν λόγω έλεγχος είχε σαν μετρική την διακύμανση *R* η οποία αναφέρεται στην τιμή του λόγου μεταξύ της μέσης απόστασης των δεδομένων στην ίδια συστάδα (απόσταση εντός συστάδας) και της μέσης απόστασης των δεδομένων στις άλλες συστάδες (απόσταση μεταξύ συστάδων). Μια τιμή R κοντά στο 0 υποδηλώνει ότι τα δεδομένα στις ίδιες συστάδες παρουσιάζουν μεγάλη ομοιότητα μεταξύ τους. Τα αποτελέσματα της αξιολόγησης έδειξαν ότι ο βέλτιστος αριθμός συστάδων k ήταν οι τρεις συστάδες με τιμή διακύμανσης R=0,19113.

Οι Sarvari, Ustundag και Takci (2016) αξιολόγησαν την απόδοση διαφορετικών προσεγγίσεων τμηματοποίησης πελατών που ενσωματώνουν τόσο μετρήσεις RFM όσο και δημογραφικές πληροφορίες, όπως η ηλικία, το φύλο και η τοποθεσία, στην πρόβλεψη της συμπεριφοράς των πελατών. Οι συγγραφείς χρησιμοποίησαν διάφορες τεχνικές τμηματοποίησης, συμπεριλαμβανομένων της ομαδοποίησης K-means, των αυτό-οργανωτικών χαρτών (Self-Organizing Maps - SOM) και των δέντρων απόφασης (decision trees), οι οποίες έπειτα αξιολογήθηκαν με βάση την ικανότητά τους να ομαδοποιούν τους πελάτες σε τμήματα με ακρίβεια αλλά και με νόημα. Σύμφωνα με τα αποτελέσματα της ανάλυσης, διαπιστώθηκε ότι οι προσεγγίσεις τμηματοποίησης που ενσωματώνουν δημογραφικά δεδομένα με την ανάλυση RFM αποδίδουν σημαντικά καλύτερα από την χρήση του παραδοσιακού μοντέλου RFM από μόνο του. Η ενισχυμένη προσέγγιση προσφέρει στις επιχειρήσεις βαθύτερη κατανόηση των προφίλ των πελατών, οδηγώντας στη βελτίωση της διαχείρισης των πελατειακών σχέσεων, σε καλύτερη στόχευση και αποτελεσματικότερες στρατηγικές μάρκετινγκ.

Στην έρευνά τους οι Christy et al. (2018) τόνισαν τη σημασία της τμηματοποίησης των πελατών για τη βελτίωση της κερδοφορίας της επιχείρησης καθώς και για την διατήρηση του πελατολογίου της καθώς υπογραμμίζουν ότι αποτελεί πιο συμφέρουσα στρατηγική σε σχέση με την προσέγγιση νέων πελατών. Για την ανάλυσή τους σε δεδομένα συναλλαγών μιας επιχείρησης λιανεμπορίου εφάρμοσαν αρχικά το παραδοσιακό μοντέλο RFM ενώ στη συνέχεια επέκτειναν την έρευνά τους χρησιμοποιώντας τρεις αλγόριθμους συσταδοποίησης: τον αλγόριθμο K-means, τον αλγόριθμο Fuzzy C-means και τέλος, τον αλγόριθμο Repetitive Median K-means (RM K-means). Πριν προχωρήσουν στην υλοποίηση των εν λόγω αλγορίθμων και έπειτα στη σύγκριση των αποτελεσμάτων τους, χρειάστηκε να κανονικοποιήσουν τις τιμές R, F και M καθώς παρατηρήθηκε ιδιαίτερα μεγάλη λοξότητα στις κατανομές τους. Ο αλγόριθμος Fuzzy C-Means επιτρέπει σε ένα συγκεκριμένο data point να ανήκει σε περισσότερες από μία συστάδες. Οι συγγραφείς αναφέρουν ότι το πλεονέκτημα αυτής της μεθόδου έναντι του K-Means είναι ότι εφόσον ένας πελάτης μπορεί να ανήκει σε περισσότερες από μία ομάδες, η επιχείρηση αυξάνει την πιθανότητα διατήρησης των πελατών της καθώς τους προσεγγίζει με ποικίλους τρόπους. Τέλος υλοποιείται ο αλγόριθμος RM K-means, μία προτεινόμενη από τους συγγραφείς παραλλαγή του K-means, η οποία χρησιμοποιεί τις διαμέσους (medians) των μετρικών R, F και M για την αρχικοποίηση των κεντροειδών (centroids) των συστάδων. Ο αλγόριθμος αυτός αποδεικνύεται και ο πιο αποδοτικός συγκριτικά με τους άλλους δύο, καθώς ο χρόνος εκτέλεσής του είναι μικρότερος και οι επαναλήψεις που χρειάζεται να γίνουν λιγότερες.

Οι Wu et al. (2020) χρησιμοποίησαν επίσης τον συνδυασμό της ανάλυσης RFM και του αλγορίθμου K-means για την τμηματοποίηση των πελατών μιας online επιχείρησης στο Πεκίνο της Κίνας. Οι συγγραφείς αφού πραγματοποίησαν αρχικά την προεπεξεργασία και τον καθαρισμό των δεδομένων και υπολόγισαν τις τιμές R, F και M, προχώρησαν σε κανονικοποίηση των μετρικών με τη χρήση της μεθόδου min-max. Στη συνέχεια χρησιμοποιείται η Ανάλυση Κύριων Συνιστωσών PCA (Principal Component Analysis) για την ανάθεση βαρών στο μοντέλο RFM δεδομένου ότι το σύνολο δεδομένων που χρησιμοποιήθηκε για την συγκεκριμένη ανάλυση χαρακτηρίζεται από μεγάλο όγκο εγγραφών. Η ανάλυση PCA είναι μια μέθοδος στατιστικής ανάλυσης που καταφέρνει να μετατρέψει το αρχικό σύνολο δεδομένων που περιέχει μεγάλο αριθμό μεταβλητών σε ένα σύνολο με λιγότερες μεταβλητές μέσω της τεχνικής μείωσης των διαστάσεων του συνόλου δεδομένων. Οι συγγραφείς όρισαν ότι το βάρος κάθε μετρικής ισούται με το ποσοστό συνεισφοράς της διακύμανσης της κύριας συνιστώσας. Έπειτα πραγματοποιήθηκε ομαδοποίηση των πελατών με τη χρήση του αλγορίθμου K-means σε τέσσερις ομάδες και προτάθηκαν στρατηγικές διαχείρισης της κάθε ομάδας για την απόκτηση υψηλού επιπέδου ικανοποίησης των πελατών. Οι συγγραφείς ισχυρίζονται ότι η υιοθέτηση της προτεινόμενης μεθόδου έχει ως αποτέλεσμα την αύξηση του συνολικού όγκου αγορών και του συνολικού ποσού κατανάλωσης.

## Συστήματα επιχειρηματικής ευφυΐας και οπτικοποίηση δεδομένων

Τα συστήματα επιχειρηματικής ευφυΐας (Business Intelligence tools) είναι τύποι λογισμικού που χρησιμοποιούνται για τη συλλογή, την οργάνωση, την οπτικοποίηση και την ανάλυση δεδομένων που συγκεντρώνονται μέσω των επιχειρηματικών λειτουργιών για την ανάδειξη τάσεων και μοτίβων που επιτρέπουν τη λήψη αποφάσεων βάσει δεδομένων (Tripathi et al., 2020). Οι Saabith et al. (2022) κάνοντας μια καταγραφή των πλεονεκτημάτων της χρήσης των BI εργαλείων αναφέρουν μεταξύ άλλων ότι τα εργαλεία BI:

* Βελτιώνουν την ανάλυση μάρκετινγκ καθώς επιτρέπουν την αξιολόγηση των στρατηγικών ή των εκστρατειών μάρκετινγκ, συνεισφέρουν στην καλύτερη ανάλυση της συμπεριφοράς των πελατών και στον εντοπισμό νέων ευκαιριών
* Βοηθούν στην κατανόηση των αγοραστικών προτύπων των πελατών και στη βελτίωση της εμπειρίας τους, επιτρέποντας την υλοποίηση στοχευμένων ενεργειών μάρκετινγκ, όπως η παροχή εκπτώσεων ή προσφορών
* Επιτρέπουν στους μη τεχνικούς χρήστες της επιχείρησης να αποκομίσουν τις πληροφορίες που χρειάζονται λαμβάνοντας εξατομικευμένες αναφορές (tailor made reports)
* Βοηθούν στην ταχύτερη ανίχνευση προβλημάτων ή σφαλμάτων, καθώς μπορούν να παρέχουν ιστορικές, σε πραγματικό χρόνο και προγνωστικές αναφορές

Οι συγγραφείς συνέκριναν 10 διαφορετικά εργαλεία BI με βάση τις διαφορετικές λειτουργίες που μπορούν να υποστηρίξουν. Μία από τις θεματικές που εξετάζουν αφορά την οπτικοποίηση των δεδομένων και συγκεκριμένα το αν υποστηρίζουν τις παρακάτω λύσεις-εφαρμογές:

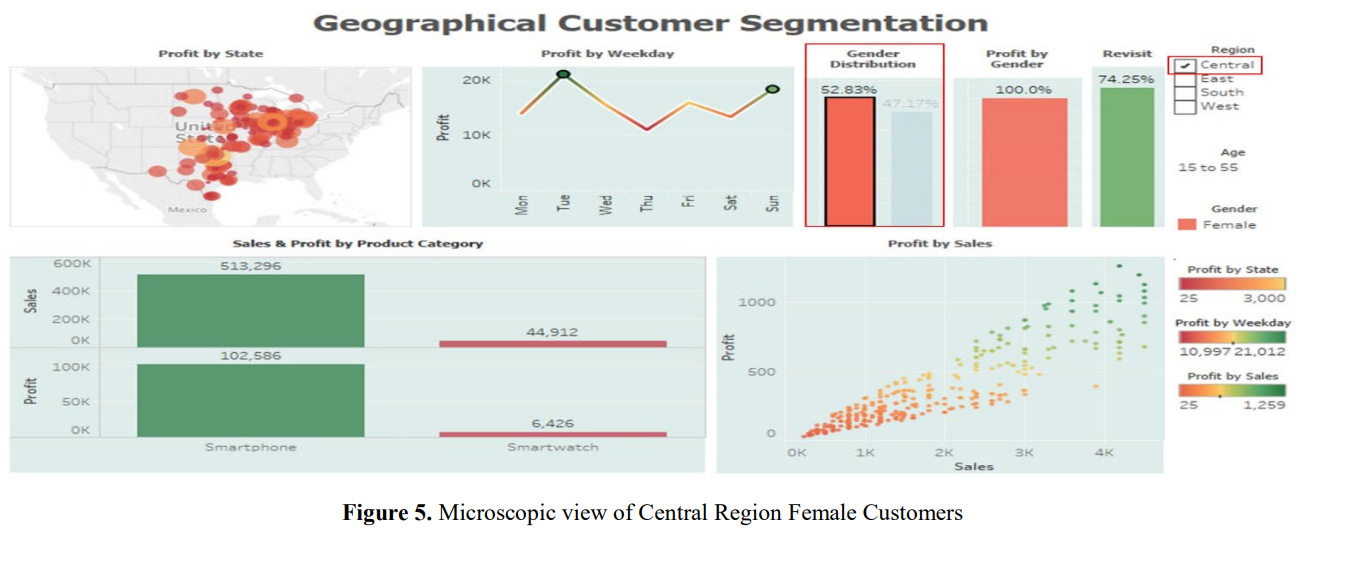
1. ***Advanced Visualizations using Python and R***: δημιουργία προηγμένων και εξελιγμένων οπτικοποιήσεων με τη χρήση βιβλιοθηκών και πακέτων προγραμματισμού Python και R
2. ***Animations***: παρουσίαση δεδομένων ως κινούμενη εικόνα, κυρίως για την παρουσίαση αλλαγών σε πολλαπλές ομάδες ή χρονικές περιόδους
3. ***Auto-charting***: καθοδήγηση των χρηστών προτείνοντας την καταλληλότερη οπτικοποίηση για τη γραφική αναπαράσταση των επιλεγμένων δεδομένων
4. ***Auto-refresh***: αυτόματη ανανέωση των διαγραμμάτων και των οπτικοποιήσεων σε ένα ταμπλό σε τακτά χρονικά διαστήματα
5. ***Dashboard Rebranding***: δυνατότητα αλλαγής των ρυθμίσεων μορφοποίησης (γραμματοσειρά, λογότυπο, χρώμα κλπ) ενός ταμπλό για την ευθυγράμμισή του με το εμπορικό σήμα του οργανισμού
6. ***Dashboards***: δημιουργία οπτικοποιήσεων που παρέχουν μια στιγμιαία προβολή σε μία οθόνη διαφόρων KPIs, μετρήσεων επιχειρηματικής ανάλυσης και κρίσιμων σημείων δεδομένων
7. ***Embed Dashboards and Visualizations in Webpages***: ενσωμάτωση ταμπλό και οπτικοποιήσεων σε άλλες ιστοσελίδες
8. ***Interactive Data Visualizations***: διαγράμματα, γραφήματα και οπτικοποιήσεις με αλληλεπιδράσεις όπως κλιμάκωση (scaling), συνδεσιμότητα (linking) και υπομνήσεις (tooltips)
9. ***Visualizations with Drill-down and Drill-up***: δυνατότητες drill-down και drill-up για την εξερεύνηση πολυδιάστατων και ιεραρχικών δεδομένων απευθείας από οπτικοποιήσεις

Οι δυνατότητες οπτικοποίησης και διαδραστικότητας που προσφέρουν σε όλο και ευρύτερο φάσμα τα εργαλεία BI (με κυρίαρχα το Microsoft Power BI και το Tableau) μπορούν να βοηθήσουν σημαντικά σε αναλύσεις μάρκετινγκ. Οι Bačić και Fadlalla (2016) υποστήριξαν ότι η τμηματοποίηση πελατών μέσω τεχνικών οπτικοποίησης αυξάνει την ερμηνευσιμότητα των τμημάτων, γεγονός που διευκολύνει την παρουσίαση των αποτελεσμάτων της τμηματοποίησης. Παρόλα αυτά, οι Sheng και Subramanian (2019) σημειώνουν ότι οι διαδραστικές τεχνικές οπτικοποίησης χρησιμοποιούνται ελάχιστα στην υπάρχουσα βιβλιογραφία για την εκτέλεση στοιχειώδους τμηματοποίησης πελατών. Οι συγγραφείς χρησιμοποιούν το Tableau ως ένα ισχυρό εργαλείο για την εφαρμογή της προτεινόμενης μεθόδου διαδραστικής οπτικοποίησης με βάση τη βαθμίδα κατάταξης (*rank-based stepwise interactive visualization method*). Η στοιχειώδης τμηματοποίηση των πελατών πραγματοποιείται με τη χρήση δημογραφικών στοιχείων για την ανακάλυψη των χαρακτηριστικών των πελατών που έλκονται από ένα συγκεκριμένο προϊόν. Ως εκ τούτου, κατασκευάζεται ένας διαδραστικός πίνακας (dashboard) με τη χρήση του Tableau που επιτρέπει την επισήμανση και το φιλτράρισμα με βάση τα χαρακτηριστικά που σχετίζονται με τα δημογραφικά στοιχεία των πελατών και παρουσιάζεται στο σχήμα **??.** Τα διαδραστικά ταμπλό των εργαλείων BI, τα προσαρμόσιμα φίλτρα και οι δυνατότητες κατάταξης (ranking) επιτρέπουν στον χρήστη να εξερευνεί δυναμικά τα πολυδιάστατα δεδομένα μέσω διαφόρων τύπων γραφημάτων για την εις βάθος κατανόηση των αγοραστικών μοτίβων των πελατών και της σχέσης τους με την επιχείρηση.

**A close-up of a graph

Description automatically generatedA close-up of a graph

Description automatically generated**

****

**References**

1. Bass, F. M. (1969). A New Product Growth for Model Consumer Durables. *Management Science*.
2. Bauer, H. H., & Hammerschmidt, M. (2005). Customer-Based Corporate Valuation. *Marketing Research and Modeling: Progress and Prospects*.
3. Blattberg, R. C., Kim, B. D., & Neslin, S. A. (2008). *Database Marketing: Analyzing and Managing Customers*. Springer. -------- ΔΕΝ βρίσκω ακριβώς αυτό. Στο scholar https://link.springer.com/chapter/10.1007/978-0-387-72579-6\_4
4. Cavusgil, S. T., Knight, G., & Riesenberger, J. (2004). *International Business: The New Realities*. Pearson.
5. Chandon, P., Morwitz, V. G., & Reinartz, W. J. (1997). Do intentions really predict behavior? *Journal of Marketing*.
6. Demby, E. H. (1974). Psychographics and from whence it came. *Journal of Consumer Research*.
7. Dolnicar, S. (2004). Beyond “Commonsense Segmentation”: A Systematics of Segmentation Approaches in Tourism. *Journal of Travel Research*.
8. Fader, P. S. (2009). *Customer Centricity: Focus on the Right Customers for Strategic Advantage*. Wharton Digital Press.
9. Hughes, A. M. (1994). *Strategic Database Marketing*. McGraw-Hill.
10. Jain, A. K., Murty, M. N., & Flynn, P. J. (1999). Data Clustering: A Review. *ACM Computing Surveys*.
11. Kotler, P. (1980). *Marketing Management: Analysis, Planning, and Control*. Prentice-Hall.
12. MacQueen, J. (1967). Some methods for classification and analysis of multivariate observations. *Proceedings of the Fifth Berkeley Symposium on Mathematical Statistics and Probability*.
13. Smith, W. R. (1956). Product Differentiation and Market Segmentation as Alternative Marketing Strategies. *Journal of Marketing*.
14. Wells, W. D. (1975). Psychographics: A critical review. *Journal of Marketing Research*.
15. Wedel, M., & Kamakura, W. A. (2000). *Market Segmentation: Conceptual and Methodological Foundations*. Springer.
16. Wind, Y., & Bell, D. R. (2007). Market Segmentation. *Wharton on Dynamic Competitive Strategy*.
17. Jiang T., T. A., 2009. Improving Personalization Solutions through Optimal Segmentation of Customer Bases. IEEE TRANSACTIONS ON KNOWLEDGE AND DATA ENGINEERING, 21(3), pp. 305-320.
18. Vijay Shinde, N. R. S. R. S. C. P. A. S. S., 2023. An Implementation on User Centered Website Using Customer Segmentation. International Journal for Research in Applied Science & Engineering Technology (IJRASET), May, 11(V), pp. 5212-5217.

19. Marcus, C. (1998) ‘A practical yet meaningful approach to customer segmentation’, Journal of Consumer Marketing, 15(5), pp. 494–504.

20. Sarvari, P. A., Ustundag, A., & Takci, H. (2016). Performance evaluation of different customer segmentation approaches based on RFM and demographics analysis. Kybernetes, 45(7), 1129–1157. doi:10.1108/k-07-2015-0180.

21. Peker, S., Kocyigit, A., & Eren, P. E. (2017). LRFMP model for customer segmentation in the grocery retail industry: a case study. Marketing Intelligence & Planning, 35(4), 544–559. doi:10.1108/mip-11-2016-0210

22. H. H. Chang and S. F. Tsay, Integrating of SOM and K-mean in data mining clustering: an empirical study of CRM and profitability evaluation, Journal of Information Management, Vol. 11 (4) (2004), pp. 161–203.

23. Reinartz, W.J. and Kumar, V. (2000), “On the profitability of long-life customers in a noncontractual setting: An empirical investigation and implications for marketing”, *Journal of Marketing*, American Marketing Association, Vol. 64 No. 4, pp. 17–35.

24. Kao, Y.-T., Wu, H.-H., Chen, H.-K., & Chang, E.-C. (2011). *A case study of applying LRFM model and clustering techniques to evaluate customer values. Journal of Statistics and Management Systems, 14(2), 267–276.* doi:10.1080/09720510.2011.10701555

25. Dursun, A., & Caber, M. (2016). *Using data mining techniques for profiling profitable hotel customers: An application of RFM analysis. Tourism Management Perspectives, 18, 153 160.* doi:10.1016/j.tmp.2016.03.001

26. Hu, Y. -H., & Yeh, T. -W. (2014). Discovering valuable frequent patterns based on RFM

analysis without customer identification information. Knowledge-Based Systems, 61, 76–88.

27. Elsner, R., Krafft, M., & Huchzemeier, A. (2003). Optimizing Rhenania's mail-order business through dynamic multilevel modeling (DMLM). Interfaces, 33(1), 50–66.

28. Baecke, P., & Van den Poel, D. (2011). Data augmentation by predicting spending pleasure using commercially available external data. Journal of Intelligent Information Systems, 36(3), 367–383.

29. Fitzpatrick, M. (2001). Statistical analysis for direct marketers—In plain English. Data Management, 64(4), 54–56.

30. Olson, D. L., Cao, Q., Gu, C., & Lee, D. (2009). Comparison of customer response models. Service Business, 3, 117–130.

31. Băcilă, M. -F., Rădulescu, A., & Marar, I. L. (2012). RFM based segmentation: An analysis of a telecom company's customers, marketing from information to decision (marketing from information to decision)5. (pp. 52–62), 52–62.

32. Miglautsch, J. R. (2000). Thoughts on RFM scoring. Journal of Database Marketing, 8(1), 67–72.

33. Heldt, R., Silveira, C. S., & Luce, F. B. (2021). Predicting customer value per product: From RFM to RFM/P. Journal of Business Research, 127, 444–453. doi:10.1016/j.jbusres.2019.05.001

34. Lin Lang, Shuang Zhou, Minjuan Zhong, Guang Sun, Bin Pan, Peng Guo (2022) *A Big Data Based Dynamic Weight Approach for RFM Segmentation*, Computers, Materials and Continua, Volume 74, Issue 2, Pages 3503-3513, ISSN 1546-2218

35. Yeh, I.-C., Yang, K.-J., & Ting, T.-M. (2009). *Knowledge discovery on RFM model using Bernoulli sequence. Expert Systems with Applications, 36(3), 5866–5871.* doi:10.1016/j.eswa.2008.07.018

36. Liu, D.-R. and Shih, Y.-Y. (2005). Integrating AHP and datamining for product recommendation based on customer lifetime value, *Information & Management*, 42(3), pp. 387–400.

37. Gustriansyah, Rendra & Suhandi, Nazori & Antony, Fery. (2020). Clustering optimization in RFM analysis Based on k-Means. *Indonesian Journal of Electrical Engineering and Computer Science*. 18. 470. 10.11591/ijeecs.v18.i1.pp470-477.

38. Maryani, I., & Riana, D. (2017). *Clustering and profiling of customers using RFM for customer relationship management recommendations. 2017 5th International Conference on Cyber and IT Service Management (CITSM).* doi:10.1109/citsm.2017.8089258

39. Christy, A. J., Umamakeswari, A., Priyatharsini, L., & Neyaa, A. (2018). *RFM Ranking – An Effective Approach to Customer Segmentation. Journal of King Saud University - Computer and Information Sciences.* doi:10.1016/j.jksuci.2018.09.004

40. Wu, J., Shi, L., Lin, W.-P., Tsai, S.-B., Li, Y., Yang, L., & Xu, G. (2020). *An Empirical Study on Customer Segmentation by Purchase Behaviors Using a RFM Model and K-Means Algorithm. Mathematical Problems in Engineering, 2020, 1–7.* doi:10.1155/2020/8884227

41. Ανδρουτσόπουλος, Ι. (2019). Τεχνητή νοημοσύνη και μηχανική μάθηση. Εκδόσεις Καστανιώτη.

42. Brynjolfsson, E., & McAfee, A. (2017). The Second Machine Age: Work, Progress, and Prosperity in a Time of Brilliant Technologies. W.W. Norton & Company.

43. Mitchell, T. M. (1997). Machine Learning. McGraw-Hill Science/Engineering/Math.

44. Phill, Lindsey. (2024). Unsupervised Learning: A Comprehensive Exploration of Algorithms, Applications, and Challenges.

45. Sutton, R. S., & Barto, A. G. (2018). Reinforcement Learning: An Introduction. MIT Press. ISBN 9780262039246

46. Alzubi, J., Nayyar, A., & Kumar, A. (2018). *Machine Learning from Theory to Algorithms: An Overview. Journal of Physics: Conference Series, 1142*

47. Mohammed, M., Khan, M.B., & Bashier, E.B.M. (2016). Machine Learning: Algorithms and Applications (1st ed.). *CRC Press*.

48. Sarker, I. H. (2021). *Machine Learning: Algorithms, Real-World Applications and Research Directions. SN Computer Science, 2(3).*

49. Mahesh, B. (2020). Machine learning algorithms-a review. *International Journal of Science and Research (IJSR).[Internet]*, *9*(1), 381-386.

50. Sandhya, N., & Charanjeet, K. R. (2016). A review on machine learning techniques. *International Journal on Recent and Innovation Trends in Computing and Communication*, *4*(3), 451-458.

51. Sinaga, K.P., & Yang, M. (2020). Unsupervised K-Means Clustering Algorithm. *IEEE Access, 8*, 80716-80727.

52. Devarapalli, Dharmaiah & Virajitha, Ayinavilli & Geddam, Sai Veera Venkata Satya Sunanda & Sunanda, Satya & Sravya, Amudalapalli & Keerthi, Boddu & Devi, Allada. (2022). Analysis of RFM Customer Segmentation Using Clustering Algorithms. Journal of Mechanical Engineering. 7. 6375-6381.

53. Rodriguez, M. Z., Comin, C. H., Casanova, D., Bruno, O. M., Amancio, D. R., Costa, L. da F., & Rodrigues, F. A. (2019). *Clustering algorithms: A comparative approach. PLOS ONE, 14(1), e0210236.*

54. Ding, Chris & He, Xiaofeng. (2002). Cluster Merging and Splitting in Hierarchical Clustering Algorithms. *Proceedings - IEEE International Conference on Data Mining, ICDM*. 139- 146. 10.1109/ICDM.2002.1183896.

55. John, J.M., Shobayo, O., Ogunleye, B. (2023). An Exploration of Clustering Algorithms for Customer Segmentation in the UK Retail Market. *Analytics*. *2*. 809-823.

56. [12] What is k-means clustering? IBM <https://www.ibm.com/topics/k-means-clustering>

57. Portugal, I., Alencar, P., & Cowan, D. (2018). The use of machine learning algorithms in recommender systems: A systematic review. Expert Systems with Applications, 97, 205–227. doi:10.1016/j.eswa.2017.12.020

58. Ester, M., Kriegel, H. P., Sander, J., & Xu, X. (1996). A density-based algorithm for discovering clusters in large spatial databases with noise. In *Proceedings of the Second International Conference on Knowledge Discovery and Data Mining (KDD’96)*. AAAI Press, 226–231

|  |
| --- |
|  |

59. Monalisa, S., & Kurnia, F. (2019). Analysis of DBSCAN and K-means algorithm for evaluating outlier on RFM model of customer behaviour. *Telkomnika (Telecommunication Computing Electronics and Control)*, *17*(1), 110-117

60. Shindler, M., Wong, A., & Meyerson, A. (2011). Fast and accurate k-means for large datasets. *Advances in neural information processing systems*, *24*.

61. Jin, X., Han, J. (2011). *K*-Means Clustering. In: Sammut, C., Webb, G.I. (eds) Encyclopedia of Machine Learning. Springer, Boston, MA. <https://doi.org/10.1007/978-0-387-30164-8_425>

62. Maheswari, K. (2019). Finding best possible number of clusters using k-means algorithm. *International Journal of Engineering and Advanced Technology*, *9*(1S4), 533-538.

63. Fadaei, A., & Khasteh, S. H. (2019). Enhanced K-means re-clustering over dynamic networks. *Expert Systems with Applications*, *132*, 126-140.

64. Syakur, M. A., Khotimah, B. K., Rochman, E. M. S., & Satoto, B. D. (2018). Integration k-means clustering method and elbow method for identification of the best customer profile cluster. In *IOP conference series: materials science and engineering* (Vol. 336, p. 012017). IOP Publishing.

65. Shutaywi, M., & Kachouie, N. N. (2021). Silhouette analysis for performance evaluation in machine learning with applications to clustering. *Entropy*, *23*(6), 759.

66. [24] The Drawbacks of K-Means Algorithm | Baeldung on Computer Science <https://www.baeldung.com/cs/k-means-flaws-improvements>

67. [25] Demonstration of K-Means Assumptions - GeeksforGeeks <https://www.geeksforgeeks.org/demonstration-of-k-means-assumptions/>

68. Tripathi, Anuj & Bagga, Teena & Aggarwal, Rashmi. (2020). Strategic Impact of Business Intelligence : A Review of Literature. Prabandhan: Indian Journal of Management. 13. 35.

69. Saabith, A. S., Vinothraj, T., & Fareez, M. (2022). Business Intelligence Tools-Systematic Review. *Int. J. Res. Eng. Sci. ISSN*, *10*(10), 394-408.

70. Bačić, D., & Fadlalla, A. (2016). Business information visualization intellectual contributions: An integrative framework of visualization capabilities and dimensions of visual intelligence. *Decision Support Systems*, *89*, 77-86.

71. Sheng, T. K., & Subramanian, P. (2019, January). Proposition of rank-based stepwise interactive visualization for customer segmentation in e-commerce. In Proceedings of the 2nd International Conference on Software Engineering and Information Management (pp. 244-248).

72. Janvrin, D. J., Raschke, R. L., & Dilla, W. N. (2014). Making sense of complex data using interactive data visualization. Journal of Accounting Education, 32(4), 31-48.