

2026-02

$\beta \ddot{y} \check{s} \pm \ddot{A} \pm \frac{1}{2} \pm \gg \acute{E} \ddot{A}^{10} - \hat{A} \quad \alpha \neg \tilde{A} \mu^1 \hat{A} \quad \tilde{A} \ddot{A}^1 \hat{A}$
 $\beta \ddot{y} \text{---} \gg \mu^0 \ddot{A} \acute{A} \zeta \frac{1}{2}^{10} - \hat{A} \quad '^3 \zeta \acute{A} - \hat{A} \quad \frac{1}{4} - \tilde{A} \acute{E} \quad ' \frac{1}{4}$
 $\beta \ddot{y} \text{''} \mu \acute{\cdot} \zeta \frac{1}{4} - \frac{1}{2} \acute{E} \frac{1}{2}^{0 \pm 1} \quad \alpha \zeta \frac{1}{2} \ddot{A} - \gg \acute{E} \frac{1}{2} \quad \alpha \cdot$

$\beta \ddot{y} \pounds \ddot{A} \pm \acute{A} \acute{A} \zeta^3 \neg \frac{1}{2} \frac{1}{2} \cdot, \quad \bullet^{10} \zeta \gg - \ddot{A} \ddot{A} \pm$

$\beta \ddot{y} \alpha \mu \ddot{A} \pm \hat{A} \ddot{A} \acute{A} \zeta^{1 \pm 0} \hat{\cdot} \quad \acute{A} \hat{\cdot}^3 \acute{A} \pm \frac{1}{4} \frac{1}{4} \pm \quad \tilde{A} \ddot{A} \cdot \frac{1}{2} \quad ' \frac{1}{2} \neg \gg \acute{A} \tilde{A} \cdot \quad \text{''} \mu \acute{\cdot} \zeta \frac{1}{4} - \frac{1}{2} \acute{E} \frac{1}{2}^{0 \pm 1} \quad \S \acute{A} \cdot \frac{1}{4} \pm \ddot{A} \zeta \zeta^{10} \zeta \frac{1}{2} \zeta$
 $\beta \ddot{y} \pounds \zeta \zeta \gg \text{\textcircled{R}} \quad \ddot{Y}^{10} \zeta \frac{1}{2} \zeta \frac{1}{4}^{10} \hat{\cdot} \frac{1}{2} \quad \bullet \hat{\cdot}^1 \tilde{A} \ddot{A} \cdot \frac{1}{4} \hat{\cdot} \frac{1}{2}^{0 \pm 1} \quad \text{''}^1 \zeta^{-0} \cdot \tilde{A} \cdot \hat{A}, \quad \pm \frac{1}{2} \mu \hat{\cdot}^1 \tilde{A} \ddot{A} \text{\textcircled{R}} \frac{1}{4}^1 \zeta \quad \bullet \mu \neg \hat{\cdot} \zeta \gg^1 \hat{A}$

<http://hdl.handle.net/11728/13387>

Downloaded from HEPHAESTUS Repository, Neapolis University institutional repository



**ΜΕΤΑΠΤΥΧΙΑΚΟ ΠΡΟΓΡΑΜΜΑ ΣΠΟΥΔΩΝ ΣΤΗΝ ΑΝΑΛΥΣΗ
ΔΕΔΟΜΕΝΩΝ ΚΑΙ ΧΡΗΜΑΤΟΟΙΚΟΝΟΜΙΚΗ ΤΕΧΝΟΛΟΓΙΑ**

Πανεπιστήμιο Νεάπολις Πάφος.

**«Καταναλωτικές Τάσεις στις Ηλεκτρονικές Αγορές
μέσω Ανάλυσης Δεδομένων και Μοντέλων Μηχανικής
Μάθησης»**

Σταυρογιάννη Νικολέττα

Φεβρουάριος, 2026



**ΜΕΤΑΠΤΥΧΙΑΚΟ ΠΡΟΓΡΑΜΜΑ ΣΠΟΥΔΩΝ ΣΤΗΝ ΑΝΑΛΥΣΗ
ΔΕΔΟΜΕΝΩΝ ΚΑΙ ΧΡΗΜΑΤΟΟΙΚΟΝΟΜΙΚΗ ΤΕΧΝΟΛΟΓΙΑ**

**Διπλωματική Εργασία η οποία υποβλήθηκε προς απόκτηση
εξ αποστάσεως μεταπτυχιακού τίτλου σπουδών στην
ανάλυση δεδομένων και χρηματοοικονομική τεχνολογία στο
Πανεπιστήμιο Νεάπολις Πάφος.**

**«Καταναλωτικές Τάσεις στις Ηλεκτρονικές Αγορές
μέσω Ανάλυσης Δεδομένων και Μοντέλων Μηχανικής
Μάθησης»**

Σταυρογιάννη Νικολέττα

Φεβρουάριος, 2026

Πνευματικά δικαιώματα

Copyright © Σταυρογιάννη Νικολέττα, 2026

Με επιφύλαξη παντός δικαιώματος. All rights reserved.

Η έγκριση της παρούσας Διπλωματικής Εργασίας από το Πανεπιστημίο
Νεάπολις δεν υποδηλώνει απαραίτητως και αποδοχή των απόψεων του
συγγραφέα εκ μέρους του Πανεπιστημίου.

Η ΥΠΕΥΘΥΝΗ ΔΗΛΩΣΗ

Η Σταυρογιάννη Νικολέττα, γνωρίζοντας τις συνέπειες της λογοκλοπής, δηλώνω υπεύθυνα ότι η παρούσα εργασία με τίτλο «Καταναλωτικές τάσεις στις ηλεκτρονικές αγορές μέσω Ανάλυσης Δεδομένων και Μοντέλων Μηχανικής Μάθησης», αποτελεί προϊόν αυστηρά προσωπικής εργασίας και όλες οι πηγές που έχω χρησιμοποιήσει, έχουν δηλωθεί κατάλληλα στις βιβλιογραφικές παραπομπές και αναφορές. Τα σημεία όπου έχω χρησιμοποιήσει ιδέες, κείμενο ή/και πηγές άλλων συγγραφέων, αναφέρονται ευδιάκριτα στο κείμενο με την κατάλληλη παραπομπή και η σχετική αναφορά περιλαμβάνεται στο τμήμα των βιβλιογραφικών αναφορών με πλήρη περιγραφή.

Η Δηλούσα

Σταυρογιάννη Νικολέττα

Πίνακας περιεχομένων

.....	1
Περίληψη.....	10
Abstract	11
Εισαγωγή.....	13
ΚΕΦΑΛΑΙΟ 1: ΣΤΟΧΟΙ ΚΑΙ ΑΠΟΤΕΛΕΣΜΑΤΑ	15
1.1 Στόχοι	15
1.2 Αποτελέσματα	15
ΚΕΦΑΛΑΙΟ 2: ΘΕΩΡΗΤΙΚΟ ΥΠΟΒΑΘΡΟ	16
2.1 Σημασία κατανόησης των καταναλωτικών τάσεων.....	16
2.2 Ο ρόλος των δεδομένων & των big data.....	17
2.3 Επισκόπηση Μεθοδολογίας.....	19
ΚΕΦΑΛΑΙΟ 3: Ανάλυση Δεδομένων και Τεχνικές Μηχανικής Μάθησης στο Ηλεκτρονικό Εμπόριο	20
3.1 Εποπτευόμενη Μάθηση.....	20
3.2 Μη εποπτευόμενη Μάθηση	21
3.3 Αξιολόγηση και Επικύρωση Μοντέλων Μηχανικής Μάθησης.....	22
ΚΕΦΑΛΑΙΟ 4: Μοντελοποίηση Καταναλωτικής Συμπεριφοράς: Ομαδοποίηση και Προβλεπτική Ανάλυση.....	23
4.1 Καταναλωτικά Προφίλ και Ομαδοποίηση (Customer Segmentation).....	23
4.2. Προβλεπτικά Μοντέλα Καταναλωτικής Συμπεριφοράς.....	24
ΚΕΦΑΛΑΙΟ 5: Συστήματα Σύστασης και Εξατομίκευση στο Ηλεκτρονικό Εμπόριο ..	26
5.1 Τύποι Συστημάτων Σύστασης: Collaborative, Content-Based και Υβριδικές Προσεγγίσεις.....	26
5.2 Εφαρμογές και Επιχειρησιακή Αξία της Εξατομίκευσης.....	27
ΚΕΦΑΛΑΙΟ 6: ΗΘΙΚΑ ΚΑΙ ΝΟΜΙΚΑ ΖΗΤΗΜΑΤΑ	28
6.1 Ηθικά και Νομικά Ζητήματα στην Ανάλυση Δεδομένων και τη Μηχανική Μάθηση στο Ηλεκτρονικό Εμπόριο.....	28
6.2 Προστασία προσωπικών δεδομένων και νομικό πλαίσιο.....	28
6.3 Διαφάνεια και Εξηγησιμότητα των Αλγορίθμων (Explainable AI).....	29
6.4 Αλγοριθμική Μεροληψία και Δικαιοσύνη	29
6.4.1 Ηθική Τιμολόγηση και Αλγοριθμική Διαφάνεια στις Αγορές.....	30
6.4.2 Συνολική αποτίμηση	30
ΚΕΦΑΛΑΙΟ 7: ΠΕΡΙΓΡΑΦΙΚΗ & ΔΙΕΡΕΥΝΗΤΙΚΗ ΑΝΑΛΥΣΗ ΑΠΟΤΕΛΕΣΜΑΤΩΝ	31
7.1 “Γεωγραφική και Συγκριτική Ανάλυση Ετήσιων Εσόδων E-commerce (2019-2023”).....	31

7.2 Δημογραφική Ανάλυση Καταναλωτών	34
7.2.1 Κατανομή Φύλου και Ηλικιακών Ομάδων	34
7.2.2 Εισόδημα και Αγοραστική Συμπεριφορά.....	35
7.2.3 Εκπαίδευση και Συχνότητα/Ενταση Αγορών	36
7.2.4 Οικογενειακή Κατάσταση και Αξία Αγορών	37
7.2.5 Επάγγελμα και Συνολική Δαπάνη	37
7.3 Δείκτες Συμπεριφοράς & Εμπειρίας Πελατών	38
7.3.1 Customer Satisfaction Analysis.....	38
7.3.2 Επάγγελμα και Ικανοποίηση Πελατών	39
7.4 Ανάλυση Κινδύνου Απώλειας Πελάτη (Churn Risk).....	40
7.4.1 Συνολικός Δείκτης Churn Risk.....	40
7.4.2 Churn Risk και Ικανοποίηση Πελατών	42
7.4.3 Churn Risk και Αξία Αγορών.....	43
7.4.4 Churn Risk και Εισοδηματικό Επίπεδο	44
7.4.5 Χρονική Συμπεριφορά: Churn Risk vs Time of Purchase.....	45
7.5 Ανάλυση Αξίας Πελάτη.....	45
7.5.1 Customer Lifetime Value (Proxy).....	46
ΚΕΦΑΛΑΙΟ 8: ΠΡΟΒΕΠΤΙΚΗ ΑΝΑΛΥΣΗ ΚΑΙ ΣΥΝΘΕΣΗ ΕΥΡΗΜΑΤΩΝ	47
8.1 Στόχος και Μεθοδολογικό Πλαίσιο Προβλεπτικής Ανάλυσης	47
8.1.1 Στόχος της Προβλεπτικής Ανάλυσης.....	47
8.1.2 Μεθοδολογικό Πλαίσιο και Επιλογή Εργαλείων	47
8.2 Προβλεπτική Ανάλυση Κινδύνου Απώλειας Πελατών (Churn Prediction)	48
8.2.1 Θεωρητικό Πλαίσιο Πρόβλεψης Churn.....	48
8.2.2 Υλοποίηση Προβλεπτικής Ανάλυσης Churn στο Tableau	49
8.2.3 Σχέση Προβλεπτικού Δείκτη Churn Score και Customer Lifetime Value	49
8.2.4 Περιορισμοί Χρονικής Πρόβλεψης.....	51
8.3 Προβλεπτική Τμηματοποίηση Πελατών βάσει Κινδύνου και Αξίας	51
8.3.1 Πελάτες Υψηλής Αξίας και Χαμηλού Κινδύνου (Core Customers)	53
8.3.2 Πελάτες Υψηλής Αξίας και Υψηλού Κινδύνου (Critical Retention Segment)....	53
8.3.3 Πελάτες Χαμηλής Αξίας και Χαμηλού Κινδύνου (Stable Low-Value Segment) 53	
8.3.4 Πελάτες Χαμηλής Αξίας και Υψηλού Κινδύνου (Low Priority Segment)	53
8.4 Ανάλυση Σεναρίων και Υποστήριξη Λήψης Αποφάσεων (What-if Analysis)	54
8.4.1 Προσαρμογή Ορίων Κινδύνου και Δυναμική Κατηγοριοποίηση.....	54
8.4.2 Επιπτώσεις Μεταβολών Συμπεριφορικών Δεικτών	54
8.4.3 Επιχειρησιακή Ερμηνεία και Στρατηγική Αξία	55

8.5 Σύνθεση Ευρημάτων και Στρατηγικές Προτάσεις	55
ΚΕΦΑΛΑΙΟ 9: ΣΥΜΠΕΡΑΣΜΑΤΑ ΚΑΙ ΜΕΛΛΟΝΤΙΚΕΣ ΕΠΕΚΤΑΣΕΙΣ	57
9.1 Συνολικά Συμπεράσματα	57
9.2 Περιορισμοί της Έρευνας	57
9.3 Μελλοντικές Επεκτάσεις.....	57
Βιβλιογραφία.....	59

ΠΙΝΑΚΑΣ ΕΙΚΟΝΩΝ

Εικόνα 1. Συγκριτικός Πίνακας εσόδων ηλεκτρονικού εμπορίου ανά χώρα (2019-2023).	32
Εικόνα 2. Στατικό στιγμιότυπο του διαδραστικού χάρτη παγκόσμιων εσόδων ηλεκτρονικού εμπορίου για το επιλεγμένο έτος.	33
Εικόνα 3.Κατανομή Φύλου και Ηλικίας.	35
Εικόνα 4.Αγοραστική Συμπεριφορά ανα Εισόδημα.	36
Εικόνα 5.Εκπαίδευση vs Αγορές.	36
Εικόνα 6.Οικογενειακή κατάσταση και Αξία αγορών	37
Εικόνα 7.Επάγγελμα vs Συνολική δαπάνη.	38
Εικόνα 8.Customer Satisfaction.	39
Εικόνα 9. Customer Satisfaction.	40
Εικόνα 10. Churn Risk.	41
Εικόνα 11.Churn Risk vs Satisfaction.	42
Εικόνα 12.Churn Risk vs Purchase Amount.	43
Εικόνα 13.Churn Risk vs Income Level	44
Εικόνα 14.Churn Risk vs Time of Purchase	45
Εικόνα 15 .CLV	46
Εικόνα 16. Predictive mapping of Customer churn Probality and Lifetime Value.	50
Εικόνα 17.Forecast Number of Purchases via month	51
Εικόνα 18. Customer Segmentation.	52

Ονοματεπώνυμο Φοιτητή: Σταυρογιάννη Νικολέττα

Τίτλος Διπλωματικής Εργασίας: Καταναλωτικές Τάσεις στις Ηλεκτρονικές
Αγορές μέσω Ανάλυσης Δεδομένων και Μοντέλων Μηχανικής Μάθησης

Η παρούσα Διπλωματική Εργασία εκπονήθηκε στο πλαίσιο των σπουδών για την απόκτηση εξ αποστάσεως μεταπτυχιακού τίτλου στο Πανεπιστήμιο Νεάπολις και εγκρίθηκε στις από τα μέλη της Εξεταστικής Επιτροπής.

Εξεταστική Επιτροπή:

Πρώτος επιβλέπων: Κωνσταντίνος Παναγιωτάκης, Καθηγητής, Τμήμα ΔΕΤ,
ΕΛΜΕΠΑ

Μέλος Εξεταστικής Επιτροπής: Χρήστος Λεμονάκης, Αναπληρωτής
Καθηγητής, Τμήμα ΔΕΤ, ΕΛΜΕΠΑ

Μέλος Εξεταστικής Επιτροπής: Γεώργιος Μαστοράκης, Αναπληρωτής
Καθηγητής, Τμήμα ΔΕΤ, ΕΛΜΕΠΑ

Ευχαριστίες

Θα ήθελα να εκφράσω τις θερμές μου ευχαριστίες στον καθηγητή μου Κωνσταντίνο Παναγιωτάκη για την αμέριστη υποστήριξη, την καθοδήγηση και τις πολύτιμες συμβουλές του καθ' όλη τη διάρκεια εκπόνησης της παρούσας διπλωματικής εργασίας. Η επιστημονική του συμβολή, η υπομονή και το ενδιαφέρον που επέδειξε υπήρξαν καθοριστικά για την εμβάθυνση στο αντικείμενο της έρευνας και την ολοκλήρωση του έργου. Τον ευχαριστώ θερμά για την εμπιστοσύνη και τη συνεχή ενθάρρυνση που μου προσέφερε.

Περίληψη

Η δυναμική ανάπτυξη του ηλεκτρονικού εμπορίου έχει επιφέρει σημαντικές αλλαγές στη συμπεριφορά των καταναλωτών και στον τρόπο με τον οποίο οι επιχειρήσεις αναλύουν και αξιοποιούν τα δεδομένα τους. Η συνεχής παραγωγή μεγάλου όγκου δεδομένων από ηλεκτρονικές συναλλαγές και ψηφιακές πλατφόρμες καθιστά αναγκαία τη χρήση σύγχρονων εργαλείων ανάλυσης και οπτικοποίησης, προκειμένου να εξαχθούν ουσιαστικά συμπεράσματα και να υποστηριχθεί η λήψη επιχειρησιακών αποφάσεων. Στο πλαίσιο αυτό, η παρούσα μεταπτυχιακή διατριβή εξετάζει τις καταναλωτικές τάσεις στις ηλεκτρονικές αγορές μέσω ανάλυσης δεδομένων και τεχνικών προβλεπτικής ανάλυσης, με κύριο εργαλείο το λογισμικό Tableau.

Σκοπός της εργασίας είναι η κατανόηση και ανάλυση της αγοραστικής συμπεριφοράς των καταναλωτών, η αναγνώριση προτύπων και τάσεων, καθώς και η κατηγοριοποίησή τους σε διακριτές ομάδες, με στόχο τη δημιουργία χρήσιμων επιχειρησιακών συμπερασμάτων. Τα δεδομένα που χρησιμοποιούνται προέρχονται από σύνολα ηλεκτρονικών συναλλαγών πραγματικού ή προσομοιωμένου περιβάλλοντος και περιλαμβάνουν πληροφορίες σχετικά με πελάτες, χρονικές περιόδους αγορών, γεωγραφική κατανομή και οικονομικά μεγέθη.

Η μεθοδολογία βασίζεται σε ποσοτική ανάλυση δεδομένων και περιλαμβάνει στάδια καθαρισμού και προεπεξεργασίας, εξερευνητικής ανάλυσης δεδομένων και εφαρμογής τεχνικών ανάλυσης μέσω του Tableau. Ιδιαίτερη έμφαση δίνεται στη χρήση οπτικοποιήσεων, διαδραστικών dashboards και φίλτρων, τα οποία επιτρέπουν τη διερεύνηση της συμπεριφοράς των καταναλωτών σε διαφορετικές χρονικές και γεωγραφικές διαστάσεις. Παράλληλα, εφαρμόζονται τεχνικές ομαδοποίησης (clustering) και προβλεπτικής ανάλυσης που υποστηρίζονται από το Tableau, όπως forecasting τάσεων και χαρτογραφική απεικόνιση δεδομένων.

Τα αποτελέσματα αναδεικνύουν τη σημασία της οπτικής ανάλυσης και της διαδραστικής διερεύνησης δεδομένων στην κατανόηση καταναλωτικών μοτίβων και στη διαμόρφωση στοχευμένων στρατηγικών marketing. Επιπλέον, επισημαίνεται ο ρόλος των εργαλείων business intelligence στη μετατροπή των δεδομένων σε αξιοποιήσιμη γνώση, ενώ εξετάζονται ζητήματα δεοντολογίας και προστασίας προσωπικών δεδομένων στο πλαίσιο της ανάλυσης καταναλωτικών πληροφοριών.

Λέξεις κλειδιά

Ηλεκτρονικό εμπόριο, καταναλωτική συμπεριφορά, ανάλυση δεδομένων, Tableau, οπτικοποίηση δεδομένων, ομαδοποίηση καταναλωτών, προβλεπτική ανάλυση.

Abstract

The rapid growth of e-commerce has significantly influenced consumer behavior and the way businesses analyze and exploit transactional data. The continuous generation of large volumes of data from online platforms highlights the need for modern data analysis and visualization tools in order to extract meaningful insights and support strategic decision-making. In this context, this master's thesis examines consumer trends in online shopping through data analytics and predictive techniques, primarily using Tableau as the main analytical tool.

The objective of this study is to analyze consumer purchasing behavior, identify patterns and trends, and segment consumers into distinct groups in order to derive actionable business insights. The data used originate from real or simulated e-commerce transaction datasets and include information related to customers, time periods, geographic distribution, and financial indicators.

The adopted methodology follows a quantitative approach and includes data cleaning and preprocessing, exploratory data analysis, and the application of analytical techniques through Tableau. Emphasis is placed on data visualization, interactive dashboards, and dynamic filters, which enable the exploration of consumer behavior across different temporal and geographic dimensions. Additionally, clustering techniques and predictive analysis supported by Tableau, such as trend forecasting and geographic mapping, are applied.

The results highlight the importance of visual analytics and interactive data exploration in understanding consumer patterns and designing targeted marketing strategies. Furthermore, the study underlines the role of business intelligence tools in transforming raw data into actionable knowledge, while also addressing ethical considerations and data privacy issues related to the analysis of consumer information.

Keywords

E-commerce, consumer behavior, data analytics, Tableau, data visualization, customer segmentation, predictive analysis.

Αφιερώνεται

«Θα ήθελα να εκφράσω την ειλικρινή μου εκτίμηση σε όλους όσους συνέβαλαν με το δικό τους τρόπο στην ολοκλήρωση αυτής της διπλωματικής εργασίας. Στην οικογένεια μου, για τη συνεχή υποστήριξη, την καθοδήγηση και την πίστη τους στις δυνατότητες μου, που αποτέλεσαν σταθερό θεμέλιο για την προσωπική και ακαδημαϊκή μου πορεία. Στους φίλους μου, για την ενθάρρυνση και έμπρακτη παρουσία τους, που με βοήθησαν να ξεπεράσω προκλήσεις και να συνεχίσω με αποφασιστικότητα. Η αρωγή τους υπήρξε καθοριστική και τους οφείλω την ολοκλήρωση αυτής της εργασίας.»

Εισαγωγή

Η ραγδαία ανάπτυξη του ηλεκτρονικού εμπορίου έχει μεταμορφώσει ριζικά το πεδίο των λιανικών πωλήσεων σε παγκόσμιο και τοπικό επίπεδο. Από τις πρώτες μορφές διαδικτυακών συναλλαγών τη δεκαετία του 1990 μέχρι τα σύγχρονα πολυκαναλικά οικοσυστήματα, το e-commerce έχει εξελιχθεί σε έναν από τους ταχύτερα αναπτυσσόμενους κλάδους της οικονομίας. Σύμφωνα με εκτιμήσεις της (Stanley, 2022), το ηλεκτρονικό εμπόριο αυξήθηκε από 15% των συνολικών λιανικών πωλήσεων το 2019 σε περίπου 21% το 2021, με την ανοδική τάση να συνεχίζεται. Δεδομένα της UNCTAD (2022, 2024) δείχνουν ότι σε ώριμες αγορές όπως η Κίνα, το Ηνωμένο Βασίλειο και η Νότια Κορέα, το ποσοστό των ηλεκτρονικών λιανικών πωλήσεων κυμαίνεται ήδη στο 25–30%, ενώ στις Ηνωμένες Πολιτείες βρίσκεται γύρω στο 15%.

Στην Ελλάδα, η διείσδυση του ηλεκτρονικού εμπορίου ακολουθεί σταθερά ανοδική πορεία. Έρευνες του Ελληνικού Συνδέσμου Ηλεκτρονικού Εμπορίου (GR.EC.A, 2021) κατέγραψαν ετήσια αύξηση 77% στον κύκλο εργασιών του 2020, ως αποτέλεσμα των περιορισμών λόγω της πανδημίας COVID-19. Επιπλέον, μελέτη του ELTRUN (Energizing Greece, 2022) ανέδειξε ότι περισσότεροι από 6 στους 10 χρήστες διαδικτύου στην Ελλάδα είναι πλέον ενεργοί online αγοραστές, με το 50% να έχει πραγματοποιήσει αγορά το τελευταίο δεκαπενθήμερο. Πρόσφατες προβλέψεις εκτιμούν ότι η αξία των online πωλήσεων προϊόντων στη χώρα θα φτάσει τα €3,08 δισ. το 2024, σημειώνοντας αύξηση 10,3% σε σχέση με το 2023, ενώ έως το 2029 αναμένεται να ξεπεράσει τα €4,7 δισ., με ετήσιο ρυθμό ανάπτυξης (CAGR) 8,7% (Ecommerce News, 2024).

Η κατανόηση των καταναλωτικών τάσεων στο ψηφιακό περιβάλλον είναι κρίσιμη για επιχειρήσεις που επιδιώκουν να παραμείνουν ανταγωνιστικές. Οι καταναλωτές δεν λειτουργούν πλέον ως παθητικοί αποδέκτες προϊόντων, αλλά αλληλεπιδρούν ενεργά με τις πλατφόρμες, συγκρίνουν τιμές, αξιολογούν προϊόντα και διαμορφώνουν προσδοκίες για εξατομικευμένες εμπειρίες. Η ανάλυση αυτών των τάσεων επιτρέπει τον εντοπισμό προτύπων, τη βελτίωση της αγοραστικής εμπειρίας και την ενίσχυση της πιστότητας (customer loyalty). Παραδείγματα επιτυχημένων εφαρμογών συναντώνται σε διεθνείς κολοσσούς όπως η Amazon και η Netflix, όπου η εξατομικευση περιεχομένου και προϊόντων βασίζεται σε εξελιγμένα αναλυτικά μοντέλα. Στην Ελλάδα, πλατφόρμες όπως το Skrutz αξιοποιούν αντίστοιχες μεθόδους για τη δημιουργία ανταγωνιστικού πλεονεκτήματος.

Κομβικό ρόλο στην ανάλυση της καταναλωτικής συμπεριφοράς διαδραματίζουν τα δεδομένα και ιδιαίτερα τα Big Data. Η έννοια των 5Vs (Volume, Velocity, Variety, Veracity, Value) περιγράφει την πολυπλοκότητα και τον όγκο των πληροφοριών που παράγονται καθημερινά από online αγοραστικές δραστηριότητες. Clickstream δεδομένα, ιστορικά αγορών, αξιολογήσεις χρηστών και δεδομένα από social media συνθέτουν μια πλούσια βάση που μπορεί να αξιοποιηθεί μέσω τεχνικών ανάλυσης

δεδομένων και μηχανικής μάθησης. Η συστηματική αξιοποίηση αυτών των πληροφοριών καθιστά δυνατή την πρόβλεψη της μελλοντικής συμπεριφοράς, την κατηγοριοποίηση των καταναλωτών και τη διαμόρφωση εξατομικευμένων προτάσεων.

Συνολικά, η μελέτη της καταναλωτικής συμπεριφοράς στο πλαίσιο του ηλεκτρονικού εμπορίου αποτελεί αναγκαία προϋπόθεση για την κατανόηση των εξελισσόμενων αγοραστικών προτύπων και την αξιοποίηση των ευκαιριών που αναδύονται στη σύγχρονη ψηφιακή οικονομία. Η ενσωμάτωση δεδομένων και τεχνολογιών μηχανικής μάθησης δημιουργεί νέα προοπτική τόσο για την ακαδημαϊκή έρευνα όσο και για την επιχειρησιακή πρακτική.

ΚΕΦΑΛΑΙΟ 1: ΣΤΟΧΟΙ ΚΑΙ ΑΠΟΤΕΛΕΣΜΑΤΑ

1.1 Στόχοι

Ο κύριος στόχος της παρούσας μεταπτυχιακής διατριβής είναι η διερεύνηση, ανάλυση και πρόβλεψη των καταναλωτικών τάσεων στο ηλεκτρονικό εμπόριο, μέσα από την αξιοποίηση τεχνικών ανάλυσης δεδομένων και μοντέλων μηχανικής μάθησης. Η έρευνα αποσκοπεί να αναδείξει τον τρόπο με τον οποίο τα δεδομένα που παράγονται από τη διαδικτυακή αγοραστική δραστηριότητα μπορούν να αξιοποιηθούν ως εργαλείο στρατηγικού σχεδιασμού, συμβάλλοντας στη βελτιστοποίηση της λήψης αποφάσεων και στην ενίσχυση της ανταγωνιστικότητας των επιχειρήσεων.

Η αξιοποίηση των Big Data και της μηχανικής μάθησης στο ηλεκτρονικό εμπόριο έχει αποδειχθεί ιδιαίτερα αποτελεσματική για την κατανόηση και πρόβλεψη της συμπεριφοράς των καταναλωτών (Chen, Zhang, & Xu, 2021· Huang & Wang, 2022). Παράλληλα, η ανάπτυξη αναλυτικών μοντέλων συμβάλλει στην εξατομίκευση της εμπειρίας χρήστη και στη βελτίωση της πιστότητας πελατών (Deloitte, 2022· Davenport & Harris, 2017). Στο πλαίσιο αυτό, η παρούσα εργασία στοχεύει στη συστηματική διερεύνηση καταναλωτικών προτύπων και αγοραστικής συμπεριφοράς μέσω προηγμένων τεχνικών οπτικοποίησης και αναλυτικής επεξεργασίας δεδομένων στο Tableau. Η προσέγγιση αυτή επιτρέπει την αποκάλυψη κρυμμένων συσχετίσεων, την αναγνώριση διαχρονικών τάσεων και την εξαγωγή συμπερασμάτων σχετικά με τη μελλοντική κατεύθυνση της καταναλωτικής δραστηριότητας.

Επιπλέον, η έρευνα αποσκοπεί στη διερεύνηση των δυνατοτήτων που προσφέρουν οι μέθοδοι ανάλυσης δεδομένων για τη βελτίωση της εμπειρίας πελάτη, την ενίσχυση του customer loyalty και την αποτελεσματικότερη στόχευση μέσω personalization και δυναμικής τιμολόγησης. Τέλος, μεταξύ των επιμέρους στόχων εντάσσεται η αναγνώριση των ηθικών και νομικών διαστάσεων της χρήσης δεδομένων, στο πλαίσιο της προστασίας προσωπικών πληροφοριών και της συμμόρφωσης με τον Γενικό Κανονισμό Προστασίας Δεδομένων (Regulation (EU) 2016/679 – GDPR), καθώς και η προώθηση της διαφάνειας μέσω επεξηγήσιμων μοντέλων τεχνητής νοημοσύνης (Explainable AI).

1.2 Αποτελέσματα

Η παρούσα μελέτη αναμένεται να προσφέρει ουσιαστικά ευρήματα τόσο σε θεωρητικό όσο και σε πρακτικό επίπεδο. Σε ερευνητικό πλαίσιο, στοχεύει να ενισχύσει τη βιβλιογραφία σχετικά με την ανάλυση καταναλωτικών τάσεων μέσω δεδομένων ηλεκτρονικού εμπορίου, αποδεικνύοντας τη συνεισφορά των τεχνικών μηχανικής μάθησης στη μοντελοποίηση αγοραστικής συμπεριφοράς (Chen et al., 2021). Σε πρακτικό επίπεδο, αναμένεται η ανάπτυξη ενός εφαρμόσιμου μοντέλου πρόβλεψης τάσεων, το οποίο μπορεί να χρησιμοποιηθεί από επιχειρήσεις e-commerce για τη

βελτιστοποίηση στρατηγικών marketing, τη διαχείριση αποθεμάτων και την αύξηση της αποδοτικότητας.

Προηγούμενες μελέτες έχουν δείξει ότι η αξιοποίηση των Big Data και των αναλυτικών τεχνικών μπορεί να βελτιώσει τη λήψη αποφάσεων και την επιχειρησιακή απόδοση έως και 25% (McKinsey & Company, 2021), ενώ η υιοθέτηση εξατομικευμένων προσεγγίσεων ενισχύει σημαντικά την ικανοποίηση και πιστότητα των πελατών (Deloitte, 2022). Επομένως, τα αποτελέσματα της παρούσας εργασίας αναμένεται να αναδείξουν τη στρατηγική αξία της ανάλυσης δεδομένων, προσφέροντας πρακτικά εργαλεία που συνδυάζουν την τεχνολογική καινοτομία με την υπεύθυνη και διαφανή διαχείριση πληροφοριών

ΚΕΦΑΛΑΙΟ 2: ΘΕΩΡΗΤΙΚΟ ΥΠΟΒΑΘΡΟ

2.1 Σημασία κατανόησης των καταναλωτικών τάσεων

Η κατανόηση των καταναλωτικών τάσεων στο ηλεκτρονικό εμπόριο αποτελεί κρίσιμο παράγοντα για τη βιωσιμότητα και την ανάπτυξη των επιχειρήσεων σε ένα ιδιαίτερα ανταγωνιστικό και ταχέως μεταβαλλόμενο περιβάλλον. Οι τάσεις αυτές δεν αντικατοπτρίζουν μόνο τις προτιμήσεις των πελατών, αλλά αποτελούν οδηγό για τη στρατηγική κατεύθυνση των επιχειρήσεων, επηρεάζοντας τον τρόπο σχεδιασμού προϊόντων, την τιμολογιακή πολιτική, τις πρακτικές marketing και τη διαχείριση πελατειακών σχέσεων.

Η μελέτη της καταναλωτικής συμπεριφοράς συνδέεται άμεσα με την έννοια του customer-centric business model, όπου οι αποφάσεις δεν βασίζονται αποκλειστικά σε ιστορικά δεδομένα πωλήσεων, αλλά σε ανάλυση συμπεριφορών, κινήτρων και προτύπων αγοράς. Για παράδειγμα, η αναγνώριση της αυξημένης σημασίας που αποδίδουν οι καταναλωτές στη βιωσιμότητα και στην κοινωνική υπευθυνότητα έχει οδηγήσει πολλές επιχειρήσεις στην προσαρμογή της εφοδιαστικής τους αλυσίδας και στην επικοινωνία περιβαλλοντικά φιλικών πρακτικών (NielsenIQ, 2023). Παράλληλα, η κατανόηση των τάσεων βοηθά τις επιχειρήσεις να προβλέψουν αλλαγές που συνδέονται με μακροοικονομικούς παράγοντες, όπως ο πληθωρισμός, η αγοραστική δύναμη και οι γεωπολιτικές εξελίξεις, οι οποίες συχνά επηρεάζουν την κατανάλωση σε είδη πρώτης ανάγκης ή πολυτελείας.

Η σημασία της ανάλυσης των καταναλωτικών τάσεων ενισχύεται από την εξατομίκευση της εμπειρίας πελάτη (personalization). Έρευνες δείχνουν ότι οι καταναλωτές είναι πιο πιθανό να προβούν σε αγορά και να αναπτύξουν πιστότητα όταν λαμβάνουν προσωποποιημένες προτάσεις προϊόντων (Deloitte, 2022). Η παρακολούθηση και ανάλυση των τάσεων καθιστά δυνατή την εφαρμογή τεχνικών όπως τα συστήματα συστάσεων, τα οποία αυξάνουν τον δείκτη μετατροπής (conversion rate) και ενισχύουν το Customer Lifetime Value (CLV).

Επιπλέον, η πρόβλεψη των τάσεων υποστηρίζει τη διαχείριση αποθεμάτων και εφοδιαστικής αλυσίδας. Η κατανόηση εποχικών προτύπων, όπως η αύξηση αγορών σε περιόδους εκπτώσεων ή εορτών, βοηθά τις επιχειρήσεις να προγραμματίσουν την παραγωγή και τις προμήθειες, μειώνοντας κόστη και αποφεύγοντας ελλείψεις ή υπερβολικά αποθέματα. Συναφείς μελέτες δείχνουν ότι η χρήση αναλυτικών μοντέλων που βασίζονται σε καταναλωτικές τάσεις μπορεί να μειώσει το κόστος αποθεμάτων έως και 20% (McKinsey, 2021).

Στρατηγικά, η ανάλυση καταναλωτικών τάσεων ενισχύει την ικανότητα μιας επιχείρησης να καινοτομεί και να εισέρχεται σε νέες αγορές. Για παράδειγμα, η αυξημένη ζήτηση για mobile commerce και social commerce τα τελευταία χρόνια ώθησε πολλές επιχειρήσεις να επενδύσουν σε εφαρμογές κινητών και σε συνεργασίες με πλατφόρμες κοινωνικής δικτύωσης. Αντίστοιχα, η υιοθέτηση του omnichannel retailing προέκυψε από την ανάγκη προσαρμογής στις νέες συνήθειες των καταναλωτών που επιθυμούν ομαλή εμπειρία μεταξύ φυσικού και ψηφιακού καταστήματος.

Τέλος, η σημασία κατανόησης των τάσεων δεν περιορίζεται σε επιχειρησιακό επίπεδο, αλλά επεκτείνεται και σε ζητήματα στρατηγικής πολιτικής. Οι ρυθμιστικές αρχές και οι οργανισμοί χάραξης πολιτικής χρησιμοποιούν την ανάλυση καταναλωτικών δεδομένων για να σχεδιάσουν κανόνες που αφορούν την προστασία καταναλωτή, τη διαφάνεια και την ασφάλεια των συναλλαγών. Σε μια εποχή όπου η προστασία προσωπικών δεδομένων αποκτά κεντρικό ρόλο (GDPR, 2018), η παρακολούθηση τάσεων συμβάλλει στην ανάπτυξη βιώσιμων και υπεύθυνων μοντέλων επιχειρηματικής λειτουργίας.

Συνοψίζοντας, η κατανόηση των καταναλωτικών τάσεων στο e-commerce αποτελεί αναπόσπαστο στοιχείο της στρατηγικής επιτυχίας. Δεν περιορίζεται μόνο στην εξυπηρέτηση των τρεχουσών αναγκών, αλλά παρέχει τα εφόδια για την πρόβλεψη μελλοντικών συμπεριφορών, την ανάπτυξη ανταγωνιστικού πλεονεκτήματος και τη δημιουργία μακροχρόνιων σχέσεων εμπιστοσύνης με τους πελάτες.

2.2 Ο ρόλος των δεδομένων & των big data

Η ραγδαία εξάπλωση του ηλεκτρονικού εμπορίου έχει ως αποτέλεσμα τη συνεχή παραγωγή τεράστιου όγκου δεδομένων, τα οποία αποτελούν σήμερα έναν από τους σημαντικότερους στρατηγικούς πόρους για τις επιχειρήσεις. Κάθε αλληλεπίδραση του καταναλωτή με μια ψηφιακή πλατφόρμα –από την αναζήτηση προϊόντων και την πλοήγηση έως την πραγματοποίηση αγοράς και την ανάρτηση αξιολογήσεων– δημιουργεί πολύτιμες πληροφορίες που μπορούν να αξιοποιηθούν για την κατανόηση της αγοραστικής συμπεριφοράς. Τα δεδομένα αυτά περιλαμβάνουν clickstream δεδομένα, ιστορικό συναλλαγών, προτιμήσεις πληρωμής, δημογραφικά στοιχεία, αλλά και δεδομένα από εξωτερικές πηγές όπως τα μέσα κοινωνικής δικτύωσης και οι μηχανές αναζήτησης.

Η πολυπλοκότητα και η ποικιλία των παραγόμενων δεδομένων οδήγησαν στην ανάπτυξη της έννοιας των Big Data, η οποία περιγράφεται συχνά μέσα από τα «5Vs»: Volume (όγκος), Velocity (ταχύτητα δημιουργίας), Variety (ποικιλία μορφών), Veracity (αξιοπιστία) και Value (αξία) (Laney, 2001· Gandomi & Haider, 2015). Στο πλαίσιο του e-commerce, τα Big Data δεν αποτελούν απλώς μια τεχνική πρόκληση διαχείρισης πληροφοριών, αλλά μια πηγή ανταγωνιστικού πλεονεκτήματος. Οι επιχειρήσεις που καταφέρνουν να επεξεργαστούν και να αναλύσουν αποτελεσματικά αυτά τα δεδομένα αποκτούν βαθύτερη κατανόηση των καταναλωτικών προτύπων και μπορούν να αναπτύξουν πιο αποδοτικές εμπορικές στρατηγικές.

Η αξιοποίηση των Big Data επιτρέπει την εφαρμογή προηγμένων τεχνικών ανάλυσης και μοντέλων μηχανικής μάθησης. Μέσα από την προβλεπτική ανάλυση (predictive analytics), οι επιχειρήσεις μπορούν να εκτιμήσουν τη μελλοντική ζήτηση προϊόντων, να εντοπίσουν τάσεις εγκατάλειψης καλαθιού (cart abandonment) ή να προβλέψουν την πιθανότητα επαναγοράς. Παράλληλα, η ομαδοποίηση (clustering) δίνει τη δυνατότητα δημιουργίας καταναλωτικών προφίλ, επιτρέποντας πιο στοχευμένες καμπάνιες marketing και καλύτερη διαχείριση σχέσεων με πελάτες. Η χρήση συστημάτων συστάσεων (recommender systems), βασισμένων είτε σε collaborative filtering είτε σε content-based μεθόδους, έχει αποδειχθεί ιδιαίτερα αποτελεσματική στη βελτίωση της εμπειρίας πελάτη και στην αύξηση του δείκτη μετατροπής.

Τα Big Data παίζουν καθοριστικό ρόλο και στη δυναμική τιμολόγηση (dynamic pricing). Μέσα από την ανάλυση σε πραγματικό χρόνο, οι επιχειρήσεις μπορούν να προσαρμόζουν τις τιμές ανάλογα με τη ζήτηση, τη διαθεσιμότητα αποθεμάτων, τον ανταγωνισμό και το προφίλ του χρήστη. Μελέτες έχουν δείξει ότι η χρήση τεχνικών μηχανικής μάθησης στην τιμολόγηση μπορεί να αυξήσει σημαντικά τα έσοδα (Huang & Wang, 2022). Αντίστοιχα, η ανάλυση χρονικών σειρών (time series analysis) επιτρέπει την πρόβλεψη μελλοντικών πωλήσεων και τη βελτιστοποίηση της εφοδιαστικής αλυσίδας.

Παρά τα σημαντικά οφέλη, η αξιοποίηση των Big Data στο e-commerce συνοδεύεται και από προκλήσεις. Θέματα όπως η ποιότητα και η αξιοπιστία των δεδομένων, οι τεχνικές δυσκολίες αποθήκευσης και επεξεργασίας μεγάλου όγκου πληροφοριών, αλλά και οι δεοντολογικές ανησυχίες σχετικά με την προστασία προσωπικών δεδομένων, καθιστούν αναγκαία την εφαρμογή κανόνων και πολιτικών συμμόρφωσης. Ο Γενικός Κανονισμός Προστασίας Δεδομένων (GDPR, 2018) στην Ευρωπαϊκή Ένωση έθεσε ένα σαφές πλαίσιο για τη συλλογή και χρήση προσωπικών δεδομένων, επιβάλλοντας στις επιχειρήσεις την ανάγκη για διαφάνεια, συγκατάθεση και ασφάλεια.

Συνολικά, τα Big Data αποτελούν αναπόσπαστο εργαλείο του σύγχρονου ηλεκτρονικού εμπορίου, καθώς συνδυάζουν τεχνολογική και στρατηγική διάσταση. Μέσω της κατάλληλης αξιοποίησης, οι επιχειρήσεις μπορούν όχι μόνο να κατανοήσουν καλύτερα τις ανάγκες των καταναλωτών, αλλά και να μετατρέψουν την πληροφόρηση σε πρακτικό πλεονέκτημα, ενισχύοντας την ανταγωνιστικότητά τους σε μια ψηφιακή οικονομία που εξελίσσεται με ταχύ ρυθμό.

2.3 Επισκόπηση Μεθοδολογίας

Για την επίτευξη των ερευνητικών στόχων, η παρούσα εργασία ακολουθεί εμπειρική και ποσοτική μεθοδολογική προσέγγιση, αξιοποιώντας πραγματικά δεδομένα καταναλωτικής δραστηριότητας και ηλεκτρονικού εμπορίου.

Η κύρια βάση δεδομένων προέρχεται από την πλατφόρμα Kaggle, η οποία περιλαμβάνει αναλυτικά στοιχεία συναλλαγών, κατηγορίες προϊόντων και συμπεριφοράς καταναλωτών σε online περιβάλλον. Το dataset επιλέχθηκε λόγω της πληρότητας και της δυνατότητας εφαρμογής τεχνικών μηχανικής μάθησης για την αναγνώριση τάσεων και προτύπων αγορών.

Η ερευνητική μεθοδολογία βασίζεται σε μία πλήρως οπτικοκεντρική και αναλυτική προσέγγιση, αξιοποιώντας αποκλειστικά το Tableau για τη συλλογή, προεπεξεργασία, ανάλυση και οπτικοποίηση των δεδομένων. Αρχικά, τα δεδομένα ηλεκτρονικών αγορών θα εισαχθούν στο Tableau, όπου θα πραγματοποιηθεί ο απαραίτητος καθαρισμός, η ενοποίηση πινάκων και η διαχείριση πιθανών ελλείπων ή ασυνεπών τιμών μέσω των διαθέσιμων εργαλείων data preparation (Data Interpreter, clean operations, filters και data joins). Στη συνέχεια, θα δημιουργηθούν υπολογισμένα πεδία (calculated fields) και εκφράσεις επιπέδου λεπτομέρειας (Level-of-Detail Expressions – LOD), προκειμένου να εξαχθούν νέες μεταβλητές που συμβάλλουν στην καλύτερη κατανόηση των αγοραστικών μοτίβων.

Ακολουθεί η εξερευνητική ανάλυση δεδομένων (Exploratory Data Analysis – EDA), η οποία θα υλοποιηθεί μέσω των δυνατοτήτων του Tableau για δυναμικές οπτικοποιήσεις, όπως γραφήματα τάσεων, heatmaps, distribution plots, bar charts, scatter plots και διαδραστικούς πίνακες. Επιπλέον, θα αξιοποιηθούν εργαλεία όπως clustering, trend lines και correlation insights, με στόχο την αναγνώριση προτύπων αγοραστικής συμπεριφοράς, τη διερεύνηση πιθανών συσχετίσεων μεταξύ μεταβλητών και την ανίχνευση ομάδων καταναλωτών βάσει κοινών χαρακτηριστικών.

Η ανάλυση θα ενισχυθεί μέσω δημιουργίας διαδραστικών dashboards και φίλτρων (actions, parameters, highlight actions), επιτρέποντας μια πολυδιάστατη κατανόηση των δεδομένων και μια πιο λεπτομερή εξέταση διαφορετικών σεναρίων. Η χρήση του Tableau παρέχει τη δυνατότητα άμεσης οπτικής ερμηνείας των αποτελεσμάτων και επιτρέπει την εύκολη αναγνώριση τάσεων, ανωμαλιών και συμπεριφορικών μοτίβων, τα οποία θα συμβάλουν στη διατύπωση συμπερασμάτων για τις καταναλωτικές τάσεις στο ηλεκτρονικό εμπόριο.

Τέλος, θα καταγραφούν οι περιορισμοί της προσέγγισης, καθώς η απουσία προηγμένων μοντέλων μηχανικής μάθησης μπορεί να περιορίσει την προβλεπτική ικανότητα της ανάλυσης, ενώ η έμφαση δίνεται κυρίως στην περιγραφική και διερευνητική πτυχή των δεδομένων. Παρ' όλα αυτά, η επιλογή αποκλειστικής χρήσης του Tableau διασφαλίζει σαφήνεια, διαδραστικότητα και άμεση εφαρμοσιμότητα των αποτελεσμάτων σε πραγματικό επιχειρησιακό περιβάλλον.

ΚΕΦΑΛΑΙΟ 3: Ανάλυση Δεδομένων και Τεχνικές Μηχανικής Μάθησης στο Ηλεκτρονικό Εμπόριο

Η ανάλυση δεδομένων και η χρήση τεχνικών μηχανικής μάθησης αποτελούν κρίσιμα εργαλεία για την κατανόηση και πρόβλεψη της καταναλωτικής συμπεριφοράς στο ηλεκτρονικό εμπόριο. Μέσω της αξιοποίησης ιστορικών δεδομένων συναλλαγών, δημογραφικών πληροφοριών και συμπεριφορικών μοτίβων, οι επιχειρήσεις μπορούν να αναπτύξουν στρατηγικές εξατομίκευσης, να βελτιστοποιήσουν τις καμπάνιες marketing και να αυξήσουν την πιστότητα των πελατών. Οι μέθοδοι μηχανικής μάθησης χωρίζονται κυρίως σε εποπτευόμενες (supervised) και μη εποπτευόμενες (unsupervised) τεχνικές, οι οποίες προσφέρουν συμπληρωματικές δυνατότητες ανάλυσης.

3.1 Εποπτευόμενη Μάθηση

Η εποπτευόμενη μάθηση (supervised learning) αποτελεί μια από τις πιο διαδεδομένες κατηγορίες αλγορίθμων μηχανικής μάθησης στο ηλεκτρονικό εμπόριο, καθώς επιτρέπει την πρόβλεψη καταναλωτικής συμπεριφοράς με βάση ιστορικά δεδομένα. Στην πράξη, οι μέθοδοι αυτές αξιοποιούν σύνολα δεδομένων με γνωστές ετικέτες (labels) για να μάθουν σχέσεις μεταξύ εισόδων (features) και εξόδων (target variables), οι οποίες στη συνέχεια εφαρμόζονται σε νέα, άγνωστα δεδομένα.

Μία από τις απλούστερες αλλά ιδιαίτερα χρήσιμες τεχνικές είναι η Λογιστική Παλινδρόμηση (Logistic Regression). Στο e-commerce χρησιμοποιείται κυρίως για προβλήματα δυαδικής ταξινόμησης, όπως η πρόβλεψη αν ένας χρήστης θα ολοκληρώσει ή θα εγκαταλείψει μια αγορά. Το πλεονέκτημά της είναι η ερμηνευσιμότητα, καθώς τα βάρη που αποδίδονται στις μεταβλητές επιτρέπουν την κατανόηση της συμβολής κάθε παράγοντα (π.χ. τιμή, χρόνος παραμονής στη σελίδα, ιστορικό αγορών) στην τελική απόφαση (Chen et al., 2020).

Πιο εξελιγμένες τεχνικές, όπως τα Random Forests, έχουν αποδειχθεί εξαιρετικά αποτελεσματικές σε προβλήματα πρόβλεψης και κατάταξης. Βασίζονται στη δημιουργία πολλαπλών δέντρων απόφασης και στη συνδυαστική τους εκτίμηση, γεγονός που μειώνει τον κίνδυνο υπερπροσαρμογής (overfitting). Στο ηλεκτρονικό εμπόριο χρησιμοποιούνται ευρέως για την πρόβλεψη της αξίας διάρκειας ζωής πελάτη (Customer Lifetime Value), την εκτίμηση πιθανοτήτων επαναγοράς, καθώς και την ανίχνευση απάτης σε συναλλαγές. Ένα σημαντικό πλεονέκτημά τους είναι η δυνατότητα μέτρησης της σημασίας των χαρακτηριστικών (feature importance), διευκολύνοντας έτσι τις επιχειρήσεις να κατανοήσουν ποιοι παράγοντες επηρεάζουν περισσότερο τις αγοραστικές αποφάσεις.

Τέλος, ένας από τους πιο δημοφιλείς αλγορίθμους της τελευταίας δεκαετίας είναι το XGBoost (Extreme Gradient Boosting), το οποίο συνδυάζει την τεχνική του boosting με βελτιστοποιημένες μεθόδους εκπαίδευσης. Ο αλγόριθμος αυτός έχει κατακτήσει σημαντική θέση σε προβλήματα e-commerce, καθώς προσφέρει υψηλή ακρίβεια,

ταχύτητα εκπαίδευσης και αποτελεσματική διαχείριση ελλιπών δεδομένων. Χρησιμοποιείται, μεταξύ άλλων, για την πρόβλεψη churn, την εκτίμηση πιθανότητας click-through (CTR) σε online διαφημίσεις και τη βελτιστοποίηση στρατηγικών cross-selling και up-selling. Έρευνες δείχνουν ότι το XGBoost μπορεί να υπερέχει σε ακρίβεια κατά 5–10% σε σχέση με άλλους παραδοσιακούς αλγορίθμους (Li & Li, 2022).

Η χρήση εποπτευόμενων μεθόδων στο e-commerce παρέχει πολλαπλά οφέλη: βελτίωση της στοχοποίησης καταναλωτών, αύξηση των ποσοστών μετατροπής (conversion rates), μείωση κόστους marketing και ενίσχυση της ικανοποίησης των πελατών. Ωστόσο, απαιτείται προσοχή στην ποιότητα των δεδομένων, καθώς η ύπαρξη bias ή η λανθασμένη επιλογή χαρακτηριστικών μπορεί να οδηγήσει σε εσφαλμένες προβλέψεις.

3.2 Μη εποπτευόμενη Μάθηση

Η μη εποπτευόμενη μάθηση (unsupervised learning) αποτελεί βασικό εργαλείο στην ανάλυση καταναλωτικής συμπεριφοράς, καθώς επιτρέπει την ανακάλυψη προτύπων και τη δημιουργία καταναλωτικών προφίλ χωρίς την ανάγκη προϋπάρχουσας γνώσης ή labels. Σε αντίθεση με τα supervised μοντέλα, οι τεχνικές αυτές εστιάζουν στην αποκάλυψη εσωτερικών σχέσεων μέσα στα δεδομένα, γεγονός που τις καθιστά ιδανικές για segmentation και exploratory analysis.

Η πιο διαδεδομένη τεχνική είναι το K-means clustering, το οποίο ομαδοποιεί τους καταναλωτές σε ομάδες με βάση τα αγοραστικά και δημογραφικά χαρακτηριστικά τους. Στο ηλεκτρονικό εμπόριο χρησιμοποιείται για customer segmentation, βοηθώντας τις επιχειρήσεις να αναπτύξουν στοχευμένες στρατηγικές marketing. Ένα χαρακτηριστικό παράδειγμα είναι η ανάλυση RFM (Recency, Frequency, Monetary Value), η οποία συνδυάζεται με K-means για να ταξινομήσει πελάτες σε κατηγορίες όπως «πιστοί», «ενεργοί» ή «υποψήφιοι προς απώλεια» (Kumar & Rajan, 2021).

Το Hierarchical Clustering προσφέρει μια διαφορετική προσέγγιση, καθώς δημιουργεί ένα δένδροειδές σχήμα (dendrogram) που επιτρέπει την κατανόηση της σχέσης μεταξύ διαφορετικών ομάδων καταναλωτών. Η τεχνική αυτή χρησιμοποιείται σε πιο σύνθετες αναλύσεις, όπου οι επιχειρήσεις επιθυμούν να κατανοήσουν πώς αλληλοσυνδέονται οι διαφορετικές ομάδες πελατών και να εξετάσουν πιθανές διασταυρώσεις μεταξύ τους.

Το DBSCAN (Density-Based Spatial Clustering of Applications with Noise) αποτελεί μια ακόμη σημαντική τεχνική, η οποία βασίζεται στην πυκνότητα των δεδομένων. Σε αντίθεση με το K-means, μπορεί να εντοπίσει outliers και μη τυπικές συμπεριφορές. Στο πλαίσιο του e-commerce, αυτό είναι ιδιαίτερα χρήσιμο για την αναγνώριση καταναλωτών με σπάνια πρότυπα αγορών, οι οποίοι ενδέχεται να αντιπροσωπεύουν νέες αγορές-στόχους ή πιθανούς κινδύνους, όπως ύποπτες συναλλαγές.

Η εφαρμογή unsupervised learning στο ηλεκτρονικό εμπόριο έχει συμβάλει ουσιαστικά στη βελτιστοποίηση των επιχειρησιακών στρατηγικών. Επιτρέπει στις εταιρείες να

εξατομικεύσουν τις καμπάνιες τους, να βελτιώσουν τα συστήματα συστάσεων και να κατανοήσουν βαθύτερα τις ανάγκες των πελατών τους. Ωστόσο, οι μέθοδοι αυτές έχουν περιορισμούς, όπως η ανάγκη για σωστή επιλογή παραμέτρων (π.χ. αριθμός clusters στο K-means) και η ευαισθησία σε θόρυβο ή μη ισορροπημένα δεδομένα.

Συνολικά, η μη εποπτευόμενη μάθηση προσφέρει μια ισχυρή προσέγγιση στην εξερεύνηση και κατανόηση της καταναλωτικής συμπεριφοράς, λειτουργώντας συχνά ως το πρώτο βήμα για πιο σύνθετα μοντέλα ανάλυσης και πρόβλεψης

3.3 Αξιολόγηση και Επικύρωση Μοντέλων Μηχανικής Μάθησης

Η αξιολόγηση και επικύρωση των μοντέλων μηχανικής μάθησης αποτελεί καθοριστικό στάδιο σε κάθε αναλυτική διαδικασία, καθώς εξασφαλίζει την αξιοπιστία, τη γενικευσιμότητα και τη χρησιμότητα των αποτελεσμάτων. Η ορθή επιλογή μετρικών απόδοσης είναι απαραίτητη για την αντικειμενική σύγκριση αλγορίθμων και τη βελτιστοποίηση των παραμέτρων τους (Kuhn & Johnson, 2019).

Για προβλήματα ταξινόμησης (classification), οι πιο συχνά χρησιμοποιούμενες μετρικές περιλαμβάνουν την ακρίβεια (accuracy), την ευαισθησία ή ανάκληση (recall), την ειδικότητα ή ακρίβεια πρόβλεψης (precision) και τον συνδυαστικό δείκτη F1-score, ο οποίος συνθέτει τις δύο τελευταίες σε έναν ενιαίο δείκτη ισορροπίας. Τα αποτελέσματα αυτών των μετρικών μπορούν να παρουσιαστούν οπτικά μέσω πινάκων σύγχυσης (confusion matrices), οι οποίοι επιτρέπουν την αξιολόγηση των ορθών και εσφαλμένων ταξινομήσεων (James et al., 2021).

Σε προβλήματα παλινδρόμησης (regression), όπου το ζητούμενο είναι η πρόβλεψη συνεχών μεταβλητών, χρησιμοποιούνται μετρικές όπως το Mean Squared Error (MSE), το Root Mean Squared Error (RMSE) και ο συντελεστής προσδιορισμού (R^2), που μετρούν τον βαθμό απόκλισης των προβλέψεων από τις πραγματικές τιμές (Zou et al., 2020).

Η επικύρωση μοντέλων (model validation) διαδραματίζει επίσης καθοριστικό ρόλο στην αποφυγή του φαινομένου υπερπροσαρμογής (overfitting), κατά το οποίο το μοντέλο αποδίδει εξαιρετικά στα δεδομένα εκπαίδευσης αλλά αποτυγχάνει να γενικεύσει σε νέα δείγματα. Η πιο διαδεδομένη τεχνική είναι η διασταυρούμενη επικύρωση (cross-validation), και ειδικότερα η k-fold cross-validation, κατά την οποία το σύνολο δεδομένων χωρίζεται σε k υποσύνολα και κάθε φορά διαφορετικό υποσύνολο χρησιμοποιείται για έλεγχο, ενώ τα υπόλοιπα για εκπαίδευση (Kuhn & Johnson, 2019).

Επιπλέον, τεχνικές όπως η βελτιστοποίηση υπερπαραμέτρων (hyperparameter tuning), με τη χρήση μεθόδων όπως η Grid Search ή η Random Search, συμβάλλουν στην επίτευξη της μέγιστης δυνατής ακρίβειας και σταθερότητας του μοντέλου (James et al., 2021). Η διαδικασία αξιολόγησης δεν περιορίζεται μόνο στη στατιστική απόδοση, αλλά επεκτείνεται και στη διερμηνευσιμότητα (interpretability), που καθίσταται

κρίσιμη για τη διαφάνεια και την εμπιστοσύνη στα αποτελέσματα των συστημάτων μηχανικής μάθησης (Ribeiro et al., 2016).

Συνοψίζοντας, η αξιολόγηση και επικύρωση των μοντέλων συνιστά αναπόσπαστο μέρος της επιστημονικής μεθοδολογίας στην ανάλυση δεδομένων, διασφαλίζοντας ότι τα αποτελέσματα δεν είναι μόνο στατιστικά αξιόπιστα, αλλά και επιχειρησιακά αξιοποιήσιμα στο πλαίσιο του ηλεκτρονικού εμπορίου.

ΚΕΦΑΛΑΙΟ 4: Μοντελοποίηση Καταναλωτικής Συμπεριφοράς: Ομαδοποίηση και Προβλεπτική Ανάλυση

4.1 Καταναλωτικά Προφίλ και Ομαδοποίηση (Customer Segmentation)

Η ανάλυση και ομαδοποίηση των καταναλωτών αποτελεί θεμελιώδες εργαλείο στο σύγχρονο περιβάλλον του ηλεκτρονικού εμπορίου, καθώς δίνει τη δυνατότητα στις επιχειρήσεις να κατανοήσουν εις βάθος τα διαφορετικά μοτίβα αγοραστικής συμπεριφοράς και να σχεδιάσουν στοχευμένες στρατηγικές marketing. Η ομαδοποίηση καταναλωτών (customer segmentation) στοχεύει στη διαίρεση της αγοράς σε ομάδες ατόμων με παρόμοια χαρακτηριστικά, ανάγκες ή προτιμήσεις, έτσι ώστε κάθε ομάδα να μπορεί να προσεγγιστεί με πιο αποτελεσματικές ενέργειες επικοινωνίας και πωλήσεων (Kumar & Rajan, 2021).

Στο πλαίσιο της μη εποπτευόμενης μάθησης, η τεχνική της ομαδοποίησης (clustering) χρησιμοποιείται για τον εντοπισμό φυσικών προτύπων μέσα στα δεδομένα, χωρίς να υπάρχουν προκαθορισμένες ετικέτες ή κατηγορίες. Οι πιο διαδεδομένοι αλγόριθμοι για αυτόν τον σκοπό είναι οι K-Means, DBSCAN και Hierarchical Clustering, οι οποίοι επιτρέπουν τον σχηματισμό ομάδων καταναλωτών με βάση δημογραφικά χαρακτηριστικά, συνήθειες αγορών ή πρότυπα αλληλεπίδρασης με την πλατφόρμα (Zhang & Mei, 2022). Μέσω της ανάλυσης αυτών των ομάδων, οι επιχειρήσεις μπορούν να εντοπίσουν τους πελάτες υψηλής αξίας, να προβλέψουν πιθανή αποχώρηση (churn) ή να σχεδιάσουν εξατομικευμένες καμπάνιες προσφορών.

Μια από τις πιο δημοφιλείς προσεγγίσεις στην πρακτική του segmentation είναι η ανάλυση RFM (Recency, Frequency, Monetary Value), η οποία αξιολογεί τη συμπεριφορά του πελάτη με βάση το πόσο πρόσφατα αγόρασε (recency), πόσο συχνά αγοράζει (frequency) και πόσο δαπανά συνολικά (monetary value). Ο συνδυασμός των RFM μεταβλητών με μεθόδους clustering παρέχει μια πολυδιάστατη κατηγοριοποίηση των πελατών και διευκολύνει τη στόχευση συγκεκριμένων ομάδων για προωθητικές ενέργειες (Nguyen et al., 2020). Για παράδειγμα, οι πελάτες που κατατάσσονται στην κατηγορία “High-Value Loyal Customers” αποτελούν στρατηγική προτεραιότητα για δράσεις διατήρησης, ενώ οι “At-Risk Customers” μπορούν να προσεγγιστούν με προγράμματα επιβράβευσης.

Η συμπεριφορική τμηματοποίηση (behavioral segmentation), που βασίζεται σε πρότυπα αγοραστικής συμπεριφοράς και πλοήγησης, έχει αποκτήσει ιδιαίτερη σημασία χάρη στα μεγάλα δεδομένα και στη μηχανική μάθηση. Η προσέγγιση αυτή δεν περιορίζεται σε στατικά δημογραφικά στοιχεία, αλλά αναλύει τη δυναμική αλληλεπίδραση του καταναλωτή με το περιβάλλον του e-shop (π.χ. χρόνος παραμονής στη σελίδα, προϊόντα που είδε, εγκατάλειψη καλαθιού κ.λπ.) (Kumar & Rajan, 2021). Οι πληροφορίες αυτές μπορούν να χρησιμοποιηθούν για τη δημιουργία προσωποποιημένων προτάσεων προϊόντων ή υπηρεσιών, αυξάνοντας την πιθανότητα ολοκλήρωσης αγοράς και την ικανοποίηση του πελάτη.

Η πρακτική εφαρμογή τέτοιων μεθόδων είναι ευρέως διαδεδομένη σε μεγάλες ψηφιακές πλατφόρμες. Η Amazon, για παράδειγμα, χρησιμοποιεί αλγορίθμους ομαδοποίησης και collaborative filtering για να προσφέρει εξατομικευμένες προτάσεις προϊόντων, με βάση το ιστορικό αγορών και τις προτιμήσεις παρόμοιων χρηστών (Zhang & Mei, 2022). Αντίστοιχα, η Netflix εφαρμόζει τεχνικές clustering για να κατηγοριοποιήσει τους συνδρομητές της ανάλογα με τις προτιμήσεις περιεχομένου και να βελτιστοποιήσει τις προτάσεις σειρών και ταινιών. Αυτές οι προσεγγίσεις οδηγούν σε αύξηση της πιστότητας (customer loyalty) και της αξίας διάρκειας ζωής πελάτη (Customer Lifetime Value – CLV), καθώς οι χρήστες βιώνουν μια πιο προσωποποιημένη και συνεκτική εμπειρία (Moradi et al., 2023).

Επιπλέον, η οπτικοποίηση των αποτελεσμάτων της ομαδοποίησης με εργαλεία όπως heatmaps, scatter plots και cluster visualizations ενισχύει την κατανόηση των προφίλ πελατών από τα διοικητικά στελέχη και διευκολύνει τη λήψη στρατηγικών αποφάσεων (Li & Li, 2022). Η δημιουργία σαφώς διαχωρισμένων καταναλωτικών ομάδων βοηθά επίσης στη βελτίωση της διαχείρισης αποθεμάτων, στην πρόβλεψη ζήτησης και στην αποδοτικότερη κατανομή των διαφημιστικών πόρων.

Συνολικά, η ομαδοποίηση καταναλωτών συνιστά κρίσιμο στάδιο της ανάλυσης δεδομένων στο ηλεκτρονικό εμπόριο, καθώς γεφυρώνει το χάσμα ανάμεσα στη στατιστική ανάλυση και την επιχειρησιακή λήψη αποφάσεων. Η ενσωμάτωση τεχνικών clustering και RFM προσφέρει στις επιχειρήσεις τη δυνατότητα να κατανοήσουν βαθύτερα τη συμπεριφορά των πελατών τους και να σχεδιάσουν στοχευμένες, αποδοτικές και προσωποποιημένες στρατηγικές marketing, συμβάλλοντας έτσι στη δημιουργία ανταγωνιστικού πλεονεκτήματος.

4.2. Προβλεπτικά Μοντέλα Καταναλωτικής Συμπεριφοράς

Η πρόβλεψη της καταναλωτικής συμπεριφοράς αποτελεί κρίσιμο παράγοντα για την ανάπτυξη στρατηγικών marketing και τη βελτιστοποίηση των επιχειρησιακών αποφάσεων στο ηλεκτρονικό εμπόριο. Μέσω της αξιοποίησης προβλεπτικών μοντέλων, οι επιχειρήσεις αποκτούν τη δυνατότητα να αναλύουν ιστορικά δεδομένα, να εντοπίζουν πρότυπα αγοραστικής συμπεριφοράς και να προβλέπουν μελλοντικές τάσεις, συμβάλλοντας στη διατήρηση της πελατειακής βάσης και στη μεγιστοποίηση της κερδοφορίας (Moradi et al., 2023).

Η πρόβλεψη της αξίας διάρκειας ζωής πελάτη (Customer Lifetime Value – CLV) αποτελεί ένα από τα πιο διαδεδομένα προβλεπτικά μοντέλα στον χώρο της αναλυτικής μάρκετινγκ. Μέσω αυτής της μεθοδολογίας, εκτιμάται το συνολικό οικονομικό όφελος που αναμένεται να αποφέρει ένας πελάτης καθ' όλη τη διάρκεια της σχέσης του με την επιχείρηση. Οι Moradi et al. (2023) προτείνουν τη χρήση αλγορίθμων παλινδρόμησης και ensemble μεθόδων, όπως τα Random Forest και XGBoost, για την εκτίμηση της CLV, αξιοποιώντας δεδομένα που αφορούν τη συχνότητα αγορών, το ύψος συναλλαγών και τη χρονική διάρκεια μεταξύ επισκέψεων. Η ανάλυση CLV επιτρέπει την ορθολογική κατανομή πόρων, ενισχύοντας προγράμματα επιβράβευσης (loyalty programs) και στοχευμένες καμπάνιες marketing για πελάτες υψηλής αξίας.

Αντίστοιχα, η πρόβλεψη αποχώρησης πελατών (customer churn) και εγκατάλειψης καλαθιού αγορών (cart abandonment) αποτελεί κρίσιμο εργαλείο για τη διατήρηση της πελατειακής βάσης. Οι Li και Li (2022) διερεύνησαν την αποτελεσματικότητα αλγορίθμων ταξινόμησης, όπως Logistic Regression, Support Vector Machines (SVM) και τεχνητών νευρωνικών δικτύων (ANNs), για την πρόβλεψη churn, χρησιμοποιώντας μεταβλητές όπως η συχνότητα αγορών, ο χρόνος παραμονής στην ιστοσελίδα και η αλληλεπίδραση με διαφημίσεις. Η μελέτη τους κατέδειξε ότι τα νευρωνικά δίκτυα υπερέχουν σε ακρίβεια πρόβλεψης, αν και απαιτούν αυξημένους υπολογιστικούς πόρους. Η έγκαιρη ανίχνευση χρηστών με υψηλή πιθανότητα αποχώρησης επιτρέπει την υλοποίηση προληπτικών ενεργειών, όπως προσωποποιημένες προσφορές ή επαναληπτικές καμπάνιες (retargeting), οι οποίες βελτιώνουν την πιστότητα και μειώνουν την απώλεια εσόδων.

Παράλληλα, η ανάλυση χρονοσειρών (time-series analysis) διαδραματίζει σημαντικό ρόλο στην πρόβλεψη πωλήσεων και στη μοντελοποίηση αγοραστικών τάσεων. Οι Singh και Prasad (2021) αξιοποίησαν τα μοντέλα ARIMA και LSTM για την πρόβλεψη όγκου πωλήσεων σε διαδικτυακές αγορές, αποδεικνύοντας ότι τα LSTM νευρωνικά δίκτυα αποδίδουν καλύτερα στη σύλληψη μακροπρόθεσμων εξαρτήσεων και εποχικών προτύπων. Η χρήση τέτοιων μοντέλων συμβάλλει στη βελτιστοποίηση αποθεμάτων, στη δυναμική τιμολόγηση και στον αποτελεσματικό προγραμματισμό καμπανιών, προλαμβάνοντας ανισορροπίες μεταξύ προσφοράς και ζήτησης.

Συνολικά, τα προβλεπτικά μοντέλα καταναλωτικής συμπεριφοράς ενισχύουν την ικανότητα των επιχειρήσεων να κατανοούν, να προβλέπουν και να προσαρμόζονται στις ανάγκες των πελατών. Ο συνδυασμός στατιστικών τεχνικών και αλγορίθμων μηχανικής μάθησης παρέχει ένα ολοκληρωμένο πλαίσιο λήψης αποφάσεων που υποστηρίζει την ανάπτυξη εξατομικευμένων εμπειριών και ενισχύει τη μακροχρόνια πιστότητα. Η αποτελεσματική εφαρμογή τους, ωστόσο, προϋποθέτει υψηλής ποιότητας δεδομένα, σωστή επιλογή μεθόδων και συνεχή αξιολόγηση της απόδοσης των μοντέλων, ώστε να εξασφαλίζεται η ακρίβεια και η αξιοπιστία των αποτελεσμάτων.

ΚΕΦΑΛΑΙΟ 5: Συστήματα Σύστασης και Εξατομίκευση στο Ηλεκτρονικό Εμπόριο

5.1 Τύποι Συστημάτων Σύστασης: Collaborative, Content-Based και Υβριδικές Προσεγγίσεις

Τα συστήματα σύστασης (Recommender Systems) αποτελούν θεμελιώδες εργαλείο στη σύγχρονη εποχή του ηλεκτρονικού εμπορίου, καθώς επιτρέπουν την εξατομικευμένη εμπειρία χρήστη και την αύξηση της αλληλεπίδρασης με τα προϊόντα ή τις υπηρεσίες. Η βασική τους λειτουργία συνίσταται στην πρόβλεψη των προτιμήσεων ενός χρήστη, με βάση τη συμπεριφορά του ή τα χαρακτηριστικά των διαθέσιμων αντικειμένων. Οι κυριότερες προσεγγίσεις που συναντώνται στη βιβλιογραφία είναι η συνεργατική φιλτράριση (collaborative filtering), η φιλτράριση βάσει περιεχομένου (content-based filtering) και τα υβριδικά συστήματα (hybrid approaches)(Zhang & Mei, 2022).

Η συνεργατική φιλτράριση αποτελεί μια ααπότις πιο διαδεδομένες τεχνικές, βασισμένη στην αρχή ότι χρήστες με παρόμοιες προτιμήσεις στο παρελθόν είναι πιθανό να έχουν και στο μέλλον παρόμοιες επιλογές. Η μέθοδος αυτή διακρίνεται σε user-based και item-based collaborative filtering. Η πρώτη συγκρίνει προφίλ χρηστών για να προτείνει προϊόντα που προτιμούν παρόμοιοι καταναλωτές, ενώ η δεύτερη εστιάζει σε ομοιότητες μεταξύ προϊόντων με βάση τις αξιολογήσεις των χρηστών cold-start "όταν δηλαδή υπάρχουν νέοι χρήστες ή προϊόντα χωρίς ιστορικά δεδομένα.

Αντίθετα, η φιλτράριση βάσει περιεχομένου επικεντρώνεται στα χαρακτηριστικά των αντικειμένων και στις προηγούμενες προτιμήσεις ενός χρήστη. Ένα σύστημα τέτοιου τύπου μπορεί, για παράδειγμα, να προτείνει νέα προϊόντα με παρόμοια χαρακτηριστικά με εκείνα που έχει ήδη επιλέξει ο χρήστης (Aggarwal, 2016). Το κύριο πλεονέκτημά του είναι ότι δεν απαιτεί μεγάλα σύνολα δεδομένων από άλλους χρήστες, ωστόσο περιορίζεται από την αδυναμία του να προτείνει αντικείμενα που διαφέρουν από τις υπάρχουσες προτιμήσεις, οδηγώντας στο φαινόμενο over-specialization.

Τα υβριδικά συστήματα σύστασης συνδυάζουν τις δύο παραπάνω μεθόδους, με στόχο να αξιοποιήσουν τα πλεονεκτήματα κάθε προσέγγισης και να αντιμετωπίσουν τις αδυναμίες τους (Burke, 2022). Μέσω της ενσωμάτωσης μηχανισμών collaborative και content-based filtering, τα υβριδικά μοντέλα επιτυγχάνουν μεγαλύτερη ακρίβεια και προσαρμοστικότητα, ενώ είναι σε θέση να λειτουργούν αποτελεσματικά ακόμη και με ελλιπή δεδομένα. Επιπλέον, η ανάπτυξη μοντέλων βαθιάς μάθησης (deep learning) έχει βελτιώσει σημαντικά την ποιότητα των προτάσεων, επιτρέποντας την ανάλυση σύνθετων προτύπων σε μεγάλα δεδομένα χρηστών (Zhang et al., 2019).

Συνολικά, η επιλογή της κατάλληλης τεχνικής εξαρτάται από τη φύση των δεδομένων, τον όγκο και την πυκνότητα των αλληλεπιδράσεων χρηστών-προϊόντων, καθώς και από τους επιχειρησιακούς στόχους του εκάστοτε οργανισμού.

5.2 Εφαρμογές και Επιχειρησιακή Αξία της Εξατομίκευσης

Η εξατομίκευση (personalization) μέσω συστημάτων σύστασης έχει αναδειχθεί ως βασικός πυλώνας της εμπειρίας χρήστη στο ηλεκτρονικό εμπόριο, προσφέροντας στοχευμένες προτάσεις και βελτιώνοντας την απόδοση των επιχειρήσεων. Τα recommender systems λειτουργούν ως μηχανισμοί προσαρμοσμένης πληροφόρησης, προβάλλοντας προϊόντα και υπηρεσίες που ταιριάζουν με τα ενδιαφέροντα και τη συμπεριφορά του κάθε χρήστη (Jannach et al., 2021). Μέσω της χρήσης τέτοιων αλγορίθμων, οι επιχειρήσεις επιτυγχάνουν αύξηση της αλληλεπίδρασης (engagement), της διατήρησης πελατών (customer retention) και της αξίας διάρκειας ζωής πελάτη (Customer Lifetime Value).

Επιπλέον, η εξατομίκευση συμβάλλει στη βελτίωση της εμπειρίας χρήστη (User Experience), μειώνοντας τον γνωστικό φόρτο και επιτρέποντας στους χρήστες να ανακαλύπτουν προϊόντα που ανταποκρίνονται στις ανάγκες τους χωρίς εκτεταμένη αναζήτηση (Zhang & Mei, 2022). Αυτή η διαδικασία δημιουργεί μια κυκλική σχέση εμπιστοσύνης μεταξύ επιχείρησης και πελάτη, όπου οι προτάσεις οδηγούν σε περισσότερες αγορές, και οι αγορές με τη σειρά τους βελτιώνουν τις μελλοντικές προβλέψεις του συστήματος.

Από επιχειρησιακής πλευράς, τα συστήματα σύστασης λειτουργούν ως στρατηγικά εργαλεία μάρκετινγκ, υποστηρίζοντας cross-selling, up-selling και δυναμική προσαρμογή τιμών (Adomavicius & Tuzhilin, 2015). Παράλληλα, η ενσωμάτωση προηγμένων τεχνικών μηχανικής μάθησης, όπως τα νευρωνικά δίκτυα αναπαράστασης (embedding models) και τα sequence-based μοντέλα, επιτρέπουν στις επιχειρήσεις να κατανοήσουν τις χρονικές πτυχές της καταναλωτικής συμπεριφοράς, οδηγώντας σε πιο ακριβείς και εξατομικευμένες συστάσεις (Zhang et al., 2019).

Εν τέλει, τα recommender systems δεν περιορίζονται πλέον στην απλή παροχή προτάσεων προϊόντων, αλλά αποτελούν αναπόσπαστο κομμάτι της στρατηγικής επιχειρηματικής ευφυΐας (business intelligence). Η συμβολή τους είναι καθοριστική τόσο στη διαμόρφωση εξατομικευμένων εμπειριών όσο και στην ενίσχυση των πωλήσεων και της πιστότητας πελατών, εδραιώνοντας τον ρόλο της ανάλυσης δεδομένων ως βασικό μοχλό επιχειρησιακής ανάπτυξης.

ΚΕΦΑΛΑΙΟ 6: ΗΘΙΚΑ ΚΑΙ ΝΟΜΙΚΑ ΖΗΤΗΜΑΤΑ

6.1 Ηθικά και Νομικά Ζητήματα στην Ανάλυση Δεδομένων και τη Μηχανική Μάθηση στο Ηλεκτρονικό Εμπόριο

Η ραγδαία ανάπτυξη του ηλεκτρονικού εμπορίου και η εκτεταμένη χρήση τεχνικών μηχανικής μάθησης δημιούργησαν νέες προοπτικές για την κατανόηση της καταναλωτικής συμπεριφοράς. Παράλληλα, όμως, έθεσαν στο επίκεντρο ένα σύνολο ηθικών, κοινωνικών και νομικών προκλήσεων, που σχετίζονται με τη διαχείριση προσωπικών δεδομένων, τη διαφάνεια των αλγοριθμικών αποφάσεων και την αποτροπή διακρίσεων. Η ισορροπία ανάμεσα στην καινοτομία και στην προστασία των δικαιωμάτων του καταναλωτή αποτελεί σήμερα κρίσιμο ζητούμενο τόσο για τις επιχειρήσεις όσο και για τους φορείς χάραξης πολιτικής.

6.2 Προστασία προσωπικών δεδομένων και νομικό πλαίσιο

Η επεξεργασία προσωπικών δεδομένων βρίσκεται στον πυρήνα κάθε πλατφόρμας ηλεκτρονικού εμπορίου. Οι αγοραστικές συνήθειες, οι προτιμήσεις, τα clicks, οι χρόνοι παραμονής και οι τοποθεσίες σύνδεσης συνθέτουν ένα ψηφιακό αποτύπωμα του καταναλωτή, το οποίο συχνά χρησιμοποιείται για στοχευμένη διαφήμιση και εξατομικευμένες προτάσεις προϊόντων. Αν και αυτή η πρακτική ενισχύει την αποτελεσματικότητα των επιχειρήσεων, εγείρει ανησυχίες σχετικά με τον βαθμό στον οποίο ο χρήστης γνωρίζει και ελέγχει πώς χρησιμοποιούνται τα δεδομένα του.

Ο Γενικός Κανονισμός Προστασίας Δεδομένων (GDPR), που εφαρμόζεται από το 2018, αποτελεί το βασικό εργαλείο της Ευρωπαϊκής Ένωσης για την προστασία της ιδιωτικότητας. Εισάγει αρχές όπως η νομιμότητα, διαφάνεια, αναλογικότητα και λογοδοσία, καθώς και το δικαίωμα διαγραφής (“right to be forgotten”) και το δικαίωμα στη φορητότητα (“data portability”). Οι επιχειρήσεις οφείλουν να ενημερώνουν τους χρήστες για τον τρόπο συλλογής και ανάλυσης των δεδομένων τους, να εξασφαλίζουν την ελάχιστη αναγκαία επεξεργασία και να λαμβάνουν ρητή συγκατάθεση (European Commission, 2018).

Η εφαρμογή του GDPR έχει οδηγήσει πολλές εταιρείες e-commerce να επενδύσουν σε μηχανισμούς data governance, πολιτικές ανωνυμοποίησης και ψευδωνυμοποίησης, καθώς και στην υιοθέτηση αρχών privacy by design κατά τον σχεδιασμό των συστημάτων. Παρά ταύτα, περιστατικά όπως η μη εξουσιοδοτημένη χρήση δεδομένων από πλατφόρμες κοινωνικής δικτύωσης ή η διαρροή πληροφοριών πελατών σε διαδικτυακά καταστήματα δείχνουν ότι η πλήρης προστασία της ιδιωτικότητας παραμένει μια συνεχιζόμενη πρόκληση.

6.3 Διαφάνεια και Εξηγησιμότητα των Αλγορίθμων (Explainable AI)

Η διαφάνεια (transparency) και η εξηγησιμότητα (explainability) αποτελούν κεντρικούς πυλώνες της ηθικής τεχνητής νοημοσύνης. Οι αλγόριθμοι που χρησιμοποιούνται στο e-commerce — όπως τα μοντέλα συστάσεων ή τα δυναμικά συστήματα τιμολόγησης — επηρεάζουν καθημερινά τις επιλογές των καταναλωτών. Ωστόσο, οι αποφάσεις τους συχνά βασίζονται σε περίπλοκες διαδικασίες που είναι δύσκολο να κατανοηθούν ακόμη και από τους ίδιους τους προγραμματιστές.

Η έννοια της Explainable Artificial Intelligence (XAI) στοχεύει στο να καταστήσει τους αλγορίθμους «διαφανείς» και τις αποφάσεις τους ερμηνεύσιμες από τον άνθρωπο (Binns et al., 2021). Για παράδειγμα, σε ένα σύστημα σύστασης προϊόντων, η πλατφόρμα οφείλει να μπορεί να εξηγήσει γιατί πρότεινε ένα συγκεκριμένο προϊόν — π.χ. «επειδή αγοράσατε παρόμοια αντικείμενα στο παρελθόν» ή «επειδή χρήστες με παρόμοια προφίλ προτίμησαν αυτό το προϊόν».

Επιπλέον, οι τεχνικές SHAP (SHapley Additive exPlanations) και LIME (Local Interpretable Model-agnostic Explanations) χρησιμοποιούνται όλο και περισσότερο για να αποδώσουν σημασιολογική ερμηνεία στις προβλέψεις των μοντέλων. Παράλληλα, οι νομοθετικές πρωτοβουλίες της Ε.Ε., όπως η Πράξη για την Τεχνητή Νοημοσύνη (EU AI Act, 2024), επιβάλλουν επιπλέον κριτήρια λογοδοσίας και διαφάνειας, ειδικά όταν οι αλγόριθμοι επηρεάζουν άμεσα καταναλωτικές αποφάσεις ή τιμολογιακές πολιτικές.

Η εξηγησιμότητα ενισχύει την εμπιστοσύνη του κοινού, συμβάλλοντας σε μια πιο ισόρροπη σχέση ανάμεσα στην επιχείρηση και τον καταναλωτή, και καθιστά τις εταιρείες υπόλογες για τις τεχνολογικές τους επιλογές.

6.4 Αλγοριθμική Μεροληψία και Δικαιοσύνη

Η αλγοριθμική μεροληψία (algorithmic bias) είναι ίσως το πιο περίπλοκο ηθικό ζήτημα της μηχανικής μάθησης. Εμφανίζεται όταν τα δεδομένα εκπαίδευσης ενσωματώνουν κοινωνικές προκαταλήψεις ή όταν οι αλγόριθμοι μαθαίνουν σχέσεις που οδηγούν σε άνιση μεταχείριση διαφορετικών ομάδων χρηστών (Mehrabi et al., 2022).

Στο ηλεκτρονικό εμπόριο, αυτό μπορεί να εκδηλωθεί με πολλούς τρόπους:

- διαφοροποίηση τιμών με βάση δημογραφικά στοιχεία (π.χ. γεωγραφική περιοχή ή φύλο),
- προτεραιοποίηση προϊόντων από συγκεκριμένους προμηθευτές,

- ή αόρατη διάκριση μέσω συστημάτων συστάσεων που προωθούν «ομοιογενές» περιεχόμενο, ενισχύοντας το λεγόμενο filter bubble.

Ενδεικτικό είναι το παράδειγμα της Amazon, η οποία το 2018 απέσυρε αλγοριθμικό σύστημα επιλογής προσωπικού που προέκρινε άνδρες υποψηφίους, λόγω μεροληψίας στα δεδομένα εκπαίδευσης. Αντίστοιχα, πλατφόρμες όπως το TikTok και το Facebook έχουν δεχθεί κριτική για την αλγοριθμική ενίσχυση στερεοτύπων, καθώς προτείνουν περιεχόμενο βάσει ιστορικής συμπεριφοράς, περιορίζοντας την ποικιλία της πληροφορίας.

Η ακαδημαϊκή κοινότητα προτείνει τεχνικές όπως fairness-aware learning, bias mitigation και federated learning, ώστε να περιοριστεί η εξάρτηση των μοντέλων από προκατειλημμένα δεδομένα. Επιπλέον, προτείνεται η χρήση fairness metrics, όπως το demographic parity ή το equalized odds, για τον συστηματικό έλεγχο δικαιοσύνης των μοντέλων.

6.4.1 Ηθική Τιμολόγηση και Αλγοριθμική Διαφάνεια στις Αγορές

Ένα ιδιαίτερο πεδίο ηθικών προκλήσεων αφορά τη δυναμική τιμολόγηση (dynamic pricing). Τα συστήματα αυτά χρησιμοποιούν δεδομένα σε πραγματικό χρόνο για να προσαρμόζουν τις τιμές με βάση τη ζήτηση, το ιστορικό αγορών ή ακόμη και τη συμπεριφορά πλοήγησης του χρήστη (Huang & Wang, 2022). Αν και αυξάνουν την αποδοτικότητα της αγοράς, δημιουργούν ερωτήματα σχετικά με την ισότητα μεταχείρισης και τη χειραγώγηση της κατανάλωσης.

Παράδειγμα αποτελεί η Uber, η οποία έχει δεχθεί κριτική για το “surge pricing” — αυξήσεις τιμών σε περιόδους υψηλής ζήτησης που ενδέχεται να θεωρηθούν εκμετάλλευση. Στο e-commerce, παρόμοια μοντέλα μπορούν να προσαρμόζουν τιμές ανά χρήστη, οδηγώντας σε φαινόμενα «τιμολόγησης μεροληψίας». Εδώ τίθεται το ερώτημα: μέχρι ποιο σημείο είναι θεμιτή η εξατομίκευση τιμής;

6.4.2 Συνολική αποτίμηση

Η ηθική διάσταση της τεχνητής νοημοσύνης και των δεδομένων στο ηλεκτρονικό εμπόριο δεν περιορίζεται σε νομική συμμόρφωση. Αφορά τη διασφάλιση θεμελιωδών αξιών, όπως η ισότητα, η ελευθερία επιλογής και η διαφάνεια. Η υιοθέτηση αρχών όπως η “Ethics by Design” και η “Responsible AI” συνιστά πλέον βέλτιστη πρακτική. Οι εταιρείες που επενδύουν σε ηθικά υπεύθυνες πρακτικές ανάλυσης δεδομένων ενισχύουν τη φήμη τους, καλλιεργούν εμπιστοσύνη και αποκτούν μακροπρόθεσμο ανταγωνιστικό πλεονέκτημα.

ΚΕΦΑΛΑΙΟ 7: ΠΕΡΙΓΡΑΦΙΚΗ & ΔΙΕΡΕΥΝΗΤΙΚΗ ΑΝΑΛΥΣΗ ΑΠΟΤΕΛΕΣΜΑΤΩΝ

Το πρακτικό μέρος της παρούσας εργασίας επικεντρώνεται στην εφαρμογή περιγραφικής και διερευνητικής ανάλυσης δεδομένων (Descriptive & Exploratory Analytics), με αποκλειστική αξιοποίηση του λογισμικού Tableau. Η επιλογή αυτή προκύπτει από την ανάγκη για μια μεθοδολογική προσέγγιση που επιτρέπει την άμεση, οπτικοποιημένη και ευέλικτη διερεύνηση της καταναλωτικής συμπεριφοράς, χωρίς την απαίτηση ανάπτυξης πολύπλοκων προγνωστικών μοντέλων ή αλγορίθμων μηχανικής μάθησης.

Το Tableau αποτελεί ένα από τα πλέον διαδεδομένα εργαλεία επιχειρησιακής ανάλυσης, προσφέροντας δυνατότητες που ευθυγραμμίζονται πλήρως με τους στόχους της εργασίας. Χαρακτηρίζεται από υψηλό βαθμό διαδραστικότητας, ευκολία ενσωμάτωσης δεδομένων και δυνατότητα δημιουργίας πολυδιάστατων οπτικοποιήσεων, που επιτρέπουν την άμεση ανίχνευση μοτίβων, συσχετίσεων και τάσεων. Μέσω υπολογιστικών πεδίων (calculated fields), παραμέτρων και δυναμικών φίλτρων, το εργαλείο καθιστά εφικτή την ανάπτυξη δεικτών συμπεριφοράς, όπως συχνότητα αγορών, αξία συναλλαγών και βαθμός εμπλοκής του καταναλωτή.

Η επιλογή της συγκεκριμένης πλατφόρμας βασίζεται επίσης στη δυνατότητά της να υποστηρίξει τεχνικές που αναφέρονται στη βιβλιογραφία της καταναλωτικής ανάλυσης, όπως η τμηματοποίηση (segmentation), η χαρτογράφηση του customer journey και η διαχρονική παρακολούθηση της ζήτησης. Παρότι το Tableau δεν επιτρέπει την υλοποίηση πολύπλοκων αλγοριθμικών μεθόδων, προσφέρει ισχυρά εργαλεία clustering και ομαδοποιήσεων σε επίπεδο οπτικής διερεύνησης, επαρκή για τους σκοπούς της συγκεκριμένης μελέτης.

Ένας επιπλέον λόγος επιλογής του Tableau αφορά τη δυνατότητα δημιουργίας ολοκληρωμένων διαδραστικών dashboards που υποστηρίζουν την ολοκληρωμένη παρουσίαση της συμπεριφοράς των καταναλωτών και των κρίσιμων μεταβλητών του dataset. Η προσέγγιση αυτή συνάδει με τη φιλοσοφία της εργασίας για πρακτική, άμεσα εφαρμόσιμη και επιχειρησιακά χρήσιμη ανάλυση.

Συνολικά, η χρήση του Tableau αποτελεί μια στρατηγική επιλογή που ενισχύει την εγκυρότητα και λειτουργικότητα του πρακτικού μέρους. Επιτρέπει την ανάλυση πραγματικών δεδομένων με τρόπο αποδοτικό, διαφανή και οπτικά ερμηνεύσιμο, ενώ παράλληλα συνδέει τη θεωρητική κατανόηση της καταναλωτικής συμπεριφοράς με μια πρακτική, σύγχρονη και ευέλικτη μεθοδολογική εφαρμογή.

7.1 “Γεωγραφική και Συγκριτική Ανάλυση Ετήσιων Εσόδων E-commerce (2019-2023)”

Για την ολοκληρωμένη κατανόηση της διεθνούς αγοράς ηλεκτρονικού εμπορίου, στην παρούσα ενότητα συνδυάζεται μια γεωγραφική και μια ποσοτική αναλυτική προσέγγιση. Αρχικά, δημιουργήθηκε στο Tableau ένας διαδραστικός χάρτης, ο οποίος παρουσιάζει τα ετήσια έσοδα ηλεκτρονικού εμπορίου ανά χώρα για κάθε έτος της περιόδου 2019–2023. Ο χάρτης λειτουργεί ως οπτική σύνοψη των διαφορών μεταξύ χωρών, επιτρέποντας την άμεση αναγνώριση αγορών υψηλής, μεσαίας και χαμηλής δραστηριότητας. Μέσω ενός δυναμικού παραμέτρου επιλογής έτους, ο χρήστης μπορεί να παρακολουθήσει τη μεταβολή των εσόδων διαχρονικά, γεγονός που καθιστά την οπτικοποίηση ιδιαίτερα χρήσιμη για την κατανόηση τάσεων και περιφερειακών ανισοτήτων. Η διαδραστική εκδοχή του χάρτη διατίθεται διαδικτυακά για πλήρη εξερεύνηση των δεδομένων (βλ. σύνδεσμο παρακάτω).

Συμπληρωματικά προς τον χάρτη, δημιουργήθηκε ένας αναλυτικός πίνακας σύγκρισης, ο οποίος παρουσιάζει τα συνολικά έσοδα κάθε χώρας για όλα τα έτη 2019–2023. Ο πίνακας αυτός λειτουργεί ως «ανάγνωση σε βάθος» των ίδιων δεδομένων, επιτρέποντας την ακριβή ποσοτική αποτύπωση της διαχρονικής πορείας κάθε αγοράς (βλ. [Εικόνα 1](#))

Annual E-Commerce Revenue by Country (2019–2023)

Country	2018	2019	2020	2021	2022	2023	2024
Australia		26,337,339	26,677,034	27,307,102	28,380,905	29,061,301	
Brazil		30,177,375	31,496,077	31,999,236	33,015,529	33,766,063	
Canada		26,879,085	28,041,521	28,155,864	29,062,404	30,213,614	
China		89,514,291	92,908,510	95,230,669	98,543,948	100,544,518	
Denmark		11,869,715	12,764,983	13,003,614	13,329,590	13,401,023	
France		41,320,021	42,292,115	43,915,018	44,830,121	45,634,135	
Germany		43,234,948	43,840,616	45,545,545	46,940,991	48,465,910	
Greece		10,990,589	11,033,423	11,232,845	11,559,770	12,231,691	
India		48,053,320	49,529,890	50,511,090	51,829,996	53,583,031	
Italy		29,473,715	30,616,493	31,124,944	31,896,145	32,911,159	
Japan		44,011,426	45,168,400	47,124,009	48,088,158	49,383,869	
Mexico		32,188,321	32,859,320	33,475,901	34,783,184	35,346,188	
Netherlands		24,660,584	25,657,655	26,719,517	27,497,816	28,491,745	
Norway		13,074,937	13,079,130	13,346,852	14,161,037	14,971,744	
South Korea		39,529,295	40,572,482	41,866,973	43,375,491	44,500,528	
Spain		28,932,349	29,810,776	30,700,613	31,777,402	32,653,413	
Sweden		13,086,437	13,771,722	14,782,565	14,894,036	14,996,549	
Switzerland		12,670,159	13,164,103	12,769,907	13,918,432	14,053,070	
United Kingdom		45,685,393	46,967,717	48,006,521	49,599,257	50,811,613	
United States		78,049,164	79,907,575	82,938,409	85,509,842	87,762,644	

Εικόνα 1. Συγκριτικός Πίνακας εσόδων ηλεκτρονικού εμπορίου ανά χώρα (2019-2023).

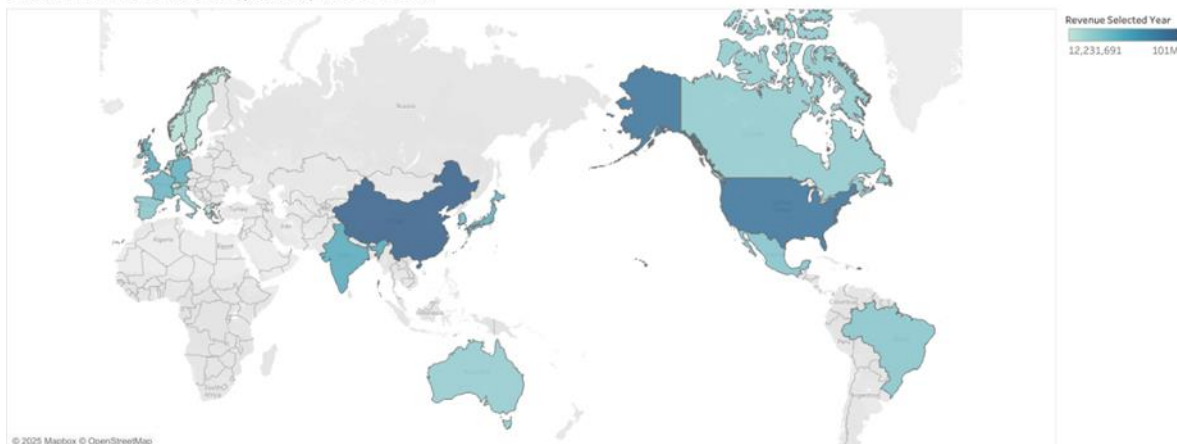
Διαδραστικός χάρτης:

<https://public.tableau.com/app/profile/nikoletta.stavrogianni/viz/ANNUALE-COMMERCEREVENUEBYCOUNTRY2019-2023/Sheet1#1>

Στο κείμενο και την ανάλυση που ακολουθεί χρησιμοποιείται ένα στατικό στιγμιότυπο του χάρτη (βλ. [Εικόνα 2](#)), το οποίο παρουσιάζει ενδεικτικά τα αποτελέσματα για το επιλεγμένο έτος. Ο πλήρης διαδραστικός χάρτης επιτρέπει πιο λεπτομερή διερεύνηση, συμπεριλαμβανομένης της προβολής των εσόδων, των παραγγελιών και των ενεργών

καταναλωτών ανά χώρα για κάθε έτος της περιόδου 2019–2023.

Global E-Commerce Revenue by Country (2019–2023)



Εικόνα 2. Στατικό στιγμιότυπο του διαδραστικού χάρτη παγκόσμιων εσόδων ηλεκτρονικού εμπορίου για το επιλεγμένο έτος.

Η γεωγραφική απεικόνιση των ετήσιων εσόδων ηλεκτρονικού εμπορίου αποκαλύπτει σημαντικές διαφοροποιήσεις μεταξύ των χωρών και καταδεικνύει την έντονη ανισοκατανομή της παγκόσμιας ψηφιακής αγοράς. Στο σύνολο της περιόδου 2019–2023, οι Ηνωμένες Πολιτείες και η Κίνα εμφανίζουν σταθερά τις υψηλότερες τιμές εσόδων, επιβεβαιώνοντας τον ρόλο τους ως οι δύο μεγαλύτερες και πιο ώριμες αγορές παγκοσμίως. Το μέγεθος και η αγοραστική δύναμη των πληθυσμών τους, σε συνδυασμό με την υψηλή διείσδυση διαδικτυακών αγορών, συμβάλλουν στη διατήρηση της ηγετικής τους θέσης.

Αντίθετα, ευρωπαϊκές αγορές όπως το Ηνωμένο Βασίλειο, η Γερμανία και η Γαλλία παρουσιάζουν μεν υψηλούς τζίρους, ωστόσο ο ρυθμός ανάπτυξής τους είναι ηπιότερος συγκριτικά με τις ασιατικές αγορές. Αναδυόμενες οικονομίες, όπως η Ινδία, η Βραζιλία και το Μεξικό, εμφανίζουν αισθητά αυξητική τάση, γεγονός που υποδηλώνει τη σταδιακή ψηφιακή μετάβαση των καταναλωτών και τη διεύρυνση των διαθέσιμων ηλεκτρονικών υπηρεσιών.

Επιπλέον, ο χάρτης αποκαλύπτει σημαντικές περιφερειακές ανισότητες. Αγορές μεσαίου μεγέθους, όπως η Ιταλία, η Ισπανία και ο Καναδάς, εμφανίζουν σταθερή αλλά πιο συγκρατημένη ανάπτυξη, ενώ μικρότερες αγορές διατηρούν χαμηλότερα επίπεδα εσόδων, κυρίως λόγω μικρότερου πληθυσμού ή χαμηλότερης ψηφιακής ετοιμότητας.

Η δυνατότητα επιλογής έτους μέσω του διαδραστικής παραμέτρου επιτρέπει την παρακολούθηση της δυναμικής εξέλιξης κάθε χώρας. Διαπιστώνεται ότι μετά το 2021 αρκετές αγορές παρουσιάζουν επιβράδυνση ή σταθεροποίηση των εσόδων, αντανakλώντας την προσαρμογή του κλάδου μετά την έντονη ανάπτυξη της περιόδου της πανδημίας. Συνολικά, η οπτικοποίηση προσφέρει μια σαφή και ολοκληρωμένη εικόνα των παγκόσμιων τάσεων, επιτρέποντας την εξαγωγή συμπερασμάτων σχετικά με την ωριμότητα, το δυναμικό και τις μελλοντικές προοπτικές των διεθνών αγορών ηλεκτρονικού εμπορίου.

7.2 Δημογραφική Ανάλυση Καταναλωτών

Σε συνέχεια του προηγούμενου διαδραστικού πίνακα, δημιουργήθηκε δεύτερο dashboard στο Tableau, το οποίο εστιάζει σε διαφορετικές διαστάσεις της ανάλυσης (π.χ. χρονική εξέλιξη, γεωγραφική κατανομή και ομαδοποίηση καταναλωτών). Το dashboard είναι διαθέσιμο διαδικτυακά στον ακόλουθο σύνδεσμο: https://public.tableau.com/app/profile/nikoletta.stavrogianni/viz/Ecommerce_Consumer_Trends_Data_Analytics_Machine_Learning/DemographicBehavioralOverview

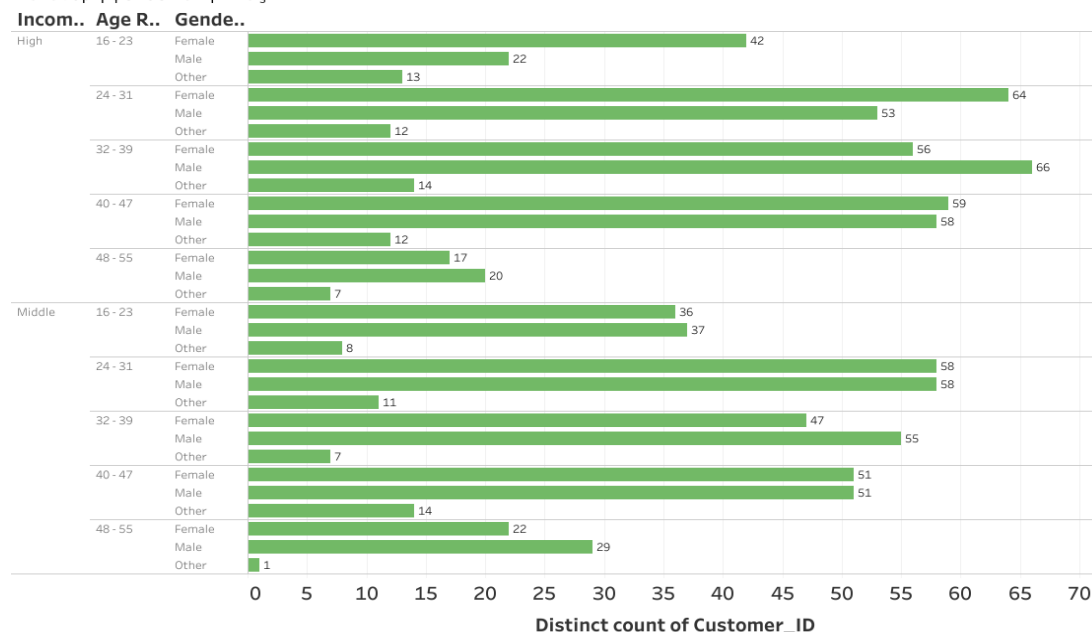
Η δημογραφική ανάλυση αποτελεί το πρώτο και θεμελιώδες στάδιο κατανόησης των μοτίβων συμπεριφοράς στο ηλεκτρονικό εμπόριο. Τα χαρακτηριστικά όπως φύλο, ηλικία, εισόδημα, μορφωτικό επίπεδο, οικογενειακή κατάσταση και επαγγελματική ιδιότητα διαμορφώνουν σε σημαντικό βαθμό τις ανάγκες, τις προτιμήσεις και τη συνολική αγοραστική στάση των καταναλωτών. Στο πλαίσιο αυτό, η ενότητα παρουσιάζει μια σειρά από οπτικοποιήσεις που επιτρέπουν την αποτύπωση βασικών δημογραφικών ομάδων και τη μελέτη της σχέσης τους με τη συμπεριφορά αγορών. Η κατανόηση αυτών των διαφοροποιήσεων παρέχει τη βάση για πιο στοχευμένη ανάλυση στις επόμενες υποενότητες, όπου εξετάζονται συμπεριφορικοί και προβλεπτικοί παράγοντες

7.2.1 Κατανομή Φύλου και Ηλικιακών Ομάδων

Παρατηρείται ότι οι ηλικιακές ομάδες 24–31 και 32–39 αποτελούν τα πιο ενεργά τμήματα του δείγματος, με τις γυναίκες να εμφανίζουν ελαφρώς υψηλότερη παρουσία σε σχέση με τους άνδρες στις περισσότερες κατηγορίες (βλ. [Εικόνα 3](#)). Στα υψηλότερα εισοδηματικά επίπεδα, η συγκέντρωση πελατών είναι ιδιαίτερα αυξημένη στην ηλικιακή ομάδα 32–39, γεγονός που υποδηλώνει ενδεχόμενη σχέση μεταξύ οικονομικής σταθερότητας και συχνότερης συμμετοχής στο ηλεκτρονικό εμπόριο. Η κατηγορία “Other” παραμένει περιορισμένη σε όλες τις ηλικιακές ομάδες και επίπεδα εισοδήματος.

Τεχνική Σημείωση: Για την ανάλυση χρησιμοποιήθηκε το πεδίο Age Range Label, το οποίο δημιουργήθηκε ως custom bin στο Tableau για την ομαδοποίηση της ηλικίας σε διακριτές κατηγορίες. Επιπλέον εφαρμόστηκε count distinct στο Customer_ID για την αποτύπωση μοναδικών πελατών.

Κατανομή φύλου και ηλικίας



Εικόνα 3.Κατανομή Φύλου και Ηλικίας.

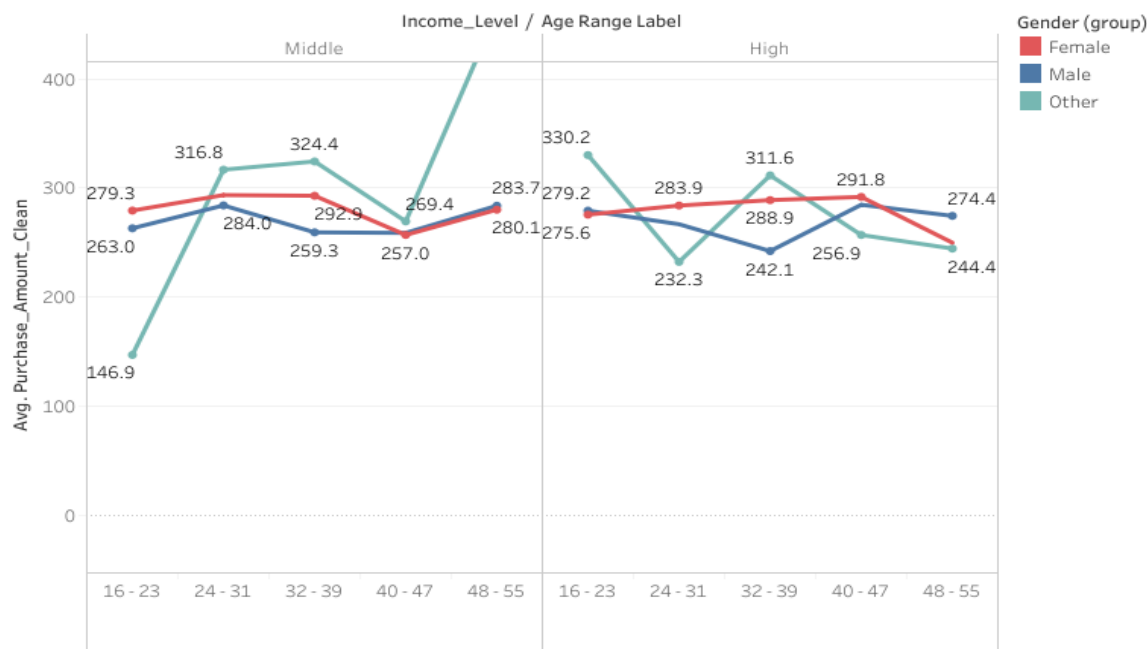
7.2.2 Εισόδημα και Αγοραστική Συμπεριφορά

Η μέση αξία αγοράς παρουσιάζει διαφοροποιήσεις τόσο ανά εισοδηματικό επίπεδο όσο και ανά ηλικιακή ομάδα (βλ. [Εικόνα 4](#)). Στο μεσαίο εισόδημα, οι ηλικίες 24–31 και 32–39 καταγράφουν τις υψηλότερες μέσες δαπάνες, ενώ η ηλικιακή ομάδα 40–47 εμφανίζει σταθεροποίηση σε ελαφρώς χαμηλότερα επίπεδα. Στο υψηλό εισόδημα, η αγοραστική αξία παραμένει αυξημένη στις ηλικίες 24–31 και 32–39, γεγονός που υποδηλώνει ότι οι οικονομικά ισχυρότεροι καταναλωτές συνεχίζουν να επενδύουν σε online αγορές με συνέπεια. Οι διαφορές μεταξύ φύλων είναι μικρές, με τους άνδρες και τις γυναίκες να ακολουθούν παρόμοια πρότυπα, ενώ η κατηγορία “Other” παρουσιάζει μεγαλύτερη μεταβλητότητα λόγω του μικρότερου πλήθους παρατηρήσεων.

Τεχνική Σημείωση: Για τον υπολογισμό της μέσης αξίας αγοράς χρησιμοποιήθηκε custom measure της μεταβλητής Purchase_Amount_Clean, η οποία είχε προηγουμένως καθαριστεί από μη έγκυρες ή μηδενικές τιμές πριν υπολογιστεί ο μέσος όρος (AVG).

Η ανάλυση πραγματοποιήθηκε κατά Income Level και Age Range μέσω custom grouping.

Μέση αξία αγοράς ανά εισοδηματικό επίπεδο



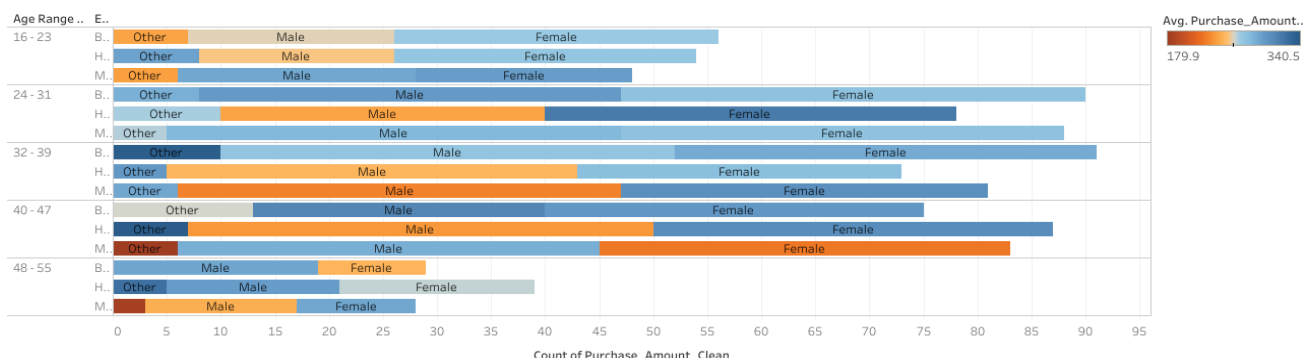
Εικόνα 4. Αγοραστική Συμπεριφορά ανά Εισόδημα.

7.2.3 Εκπαίδευση και Συχνότητα/Ενταση Αγορών

Οι πελάτες με ανώτερη εκπαίδευση (Bachelor's και Master's) πραγματοποιούν σαφώς περισσότερες αγορές σε σχέση με τους αποφοίτους λυκείου, ιδιαίτερα στις ηλικίες 24–31 και 32–39 (βλ. Εικόνα 5). Το μοτίβο αυτό υποδηλώνει ότι υψηλότερο μορφωτικό επίπεδο σχετίζεται με αυξημένη ψηφιακή εξοικείωση και μεγαλύτερη αγοραστική δραστηριότητα.

Τεχνική Σημείωση: Η απεικόνιση βασίστηκε στο count του Purchase_Amount_Clean ώστε να προκύψει η συχνότητα αγορών. Το Education_Level χρησιμοποιήθηκε ως custom ordered dimension, ενώ τα age ranges βασίστηκαν στο binning που δημιουργήθηκε στην υποενότητα 7.2.1.

Εκπαίδευση vs Αγορές



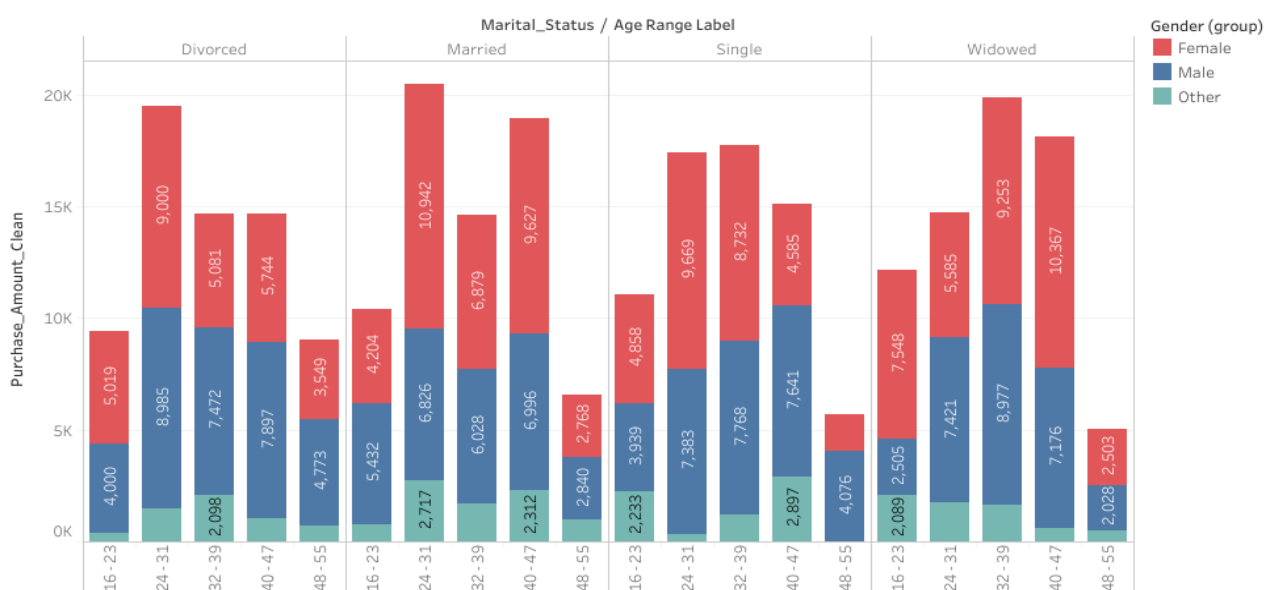
Εικόνα 5. Εκπαίδευση vs Αγορές.

7.2.4 Οικογενειακή Κατάσταση και Αξία Αγορών

Η συνολική αξία αγορών διαφοροποιείται αισθητά ανά οικογενειακή κατάσταση ([βλ. Εικόνα 6](#)). Οι έγγαμοι πελάτες εμφανίζουν τη μεγαλύτερη αγοραστική δαπάνη, ενώ ακολουθούν οι διαζευγμένοι και οι άγαμοι. Το μοτίβο αυτό υποδηλώνει ότι τα νοικοκυριά με δύο άτομα πιθανόν να πραγματοποιούν συχνότερες ή υψηλότερες αγορές, αντανakλώντας αυξημένες ανάγκες ή μεγαλύτερη οικονομική σταθερότητα.

Τεχνική Σημείωση: Το συνολικό ύψος δαπάνης υπολογίστηκε μέσω του SUM(Purchase_Amount_Clean). Επιπλέον εφαρμόστηκε stacked bar visualization με βάση το Gender (group), το οποίο είχε δημιουργηθεί ως custom grouping για απλούστερη κατηγοριοποίηση των τιμών.

Οικογενειακή κατάσταση και Αξία αγορών



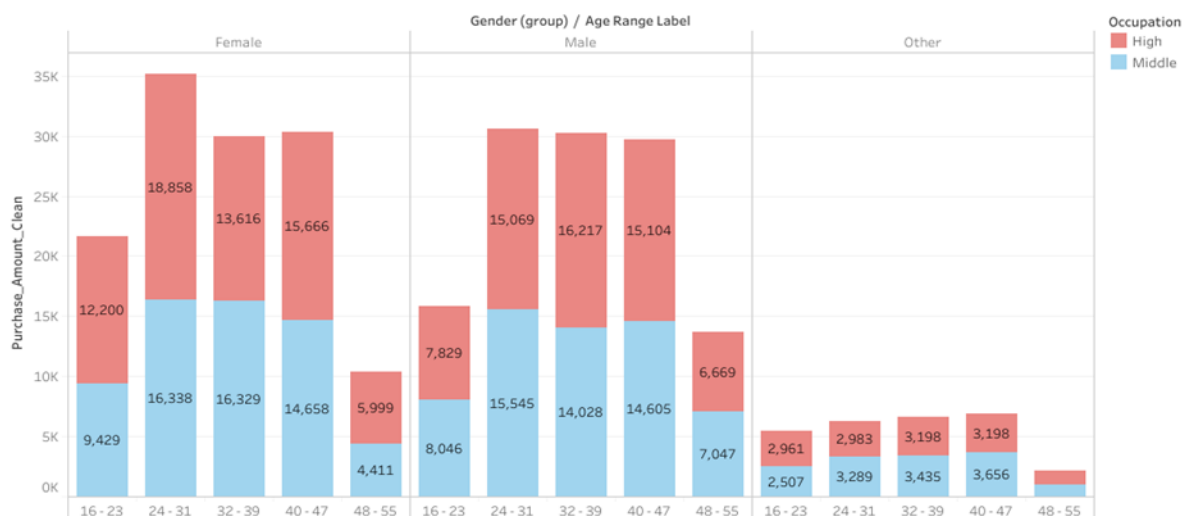
Εικόνα 6. Οικογενειακή κατάσταση και Αξία αγορών .

7.2.5 Επάγγελμα και Συνολική Δαπάνη

Η συνολική αγοραστική δαπάνη εμφανίζεται υψηλότερη στους πελάτες που ανήκουν σε επαγγέλματα υψηλής ειδίκευσης, σε σχέση με τους εργαζόμενους μεσαίας κατηγορίας ([βλ. Εικόνα 7](#)). Το μοτίβο παρατηρείται σταθερά σε όλες σχεδόν τις ηλικιακές ομάδες και στα δύο φύλα, γεγονός που υποδηλώνει ότι η επαγγελματική θέση και το αντίστοιχο εισόδημα αποτελούν καθοριστικούς παράγοντες για το ύψος των online αγορών.

Τεχνική Σημείωση: Η μεταβλητή Occupation ομαδοποιήθηκε σε δύο κατηγορίες (High, Middle) μέσω custom group στο Tableau. Η συνολική δαπάνη υπολογίστηκε με SUM(Purchase_Amount_Clean), ενώ η ανάλυση διαχωρίστηκε περαιτέρω ανά φύλο και ηλικιακή ομάδα με χρήση πολλαπλών dimensions.

Επάγγελμα vs Συνολική δαπάνη



Εικόνα 7. Επάγγελμα vs Συνολική δαπάνη.

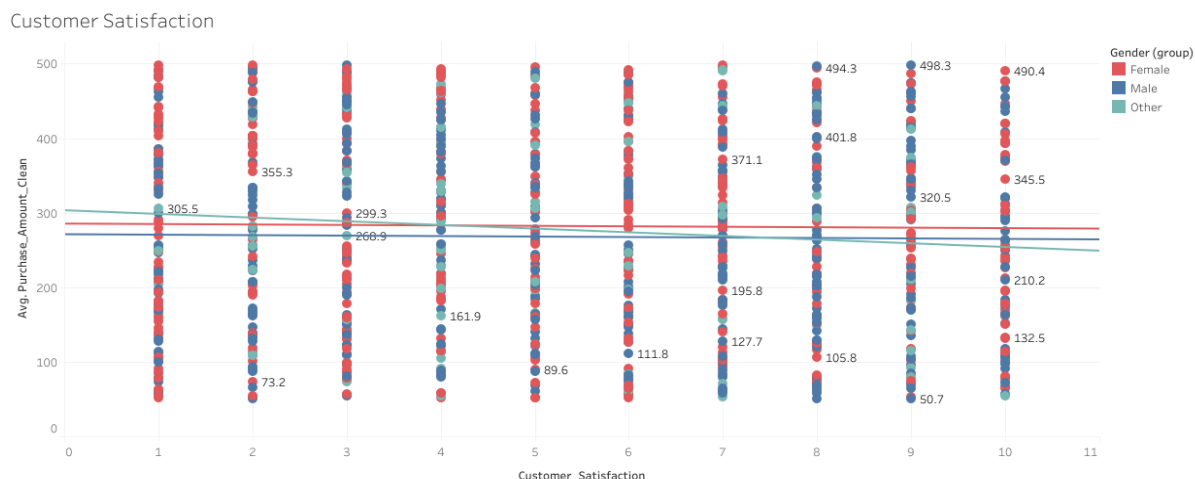
7.3 Δείκτες Συμπεριφοράς & Εμπειρίας Πελατών

Η ενότητα εξετάζει δείκτες που αποτυπώνουν τη συμπεριφορά και την εμπειρία των πελατών στο ηλεκτρονικό εμπόριο, με έμφαση στην ικανοποίηση και τη σχέση της με βασικά χαρακτηριστικά των καταναλωτών. Οι δείκτες αυτοί συμβάλλουν στην κατανόηση της πιστότητας και της συνολικής εμπειρίας χρήστη.

7.3.1 Customer Satisfaction Analysis

Η σχέση μεταξύ ικανοποίησης πελατών και μέσης αξίας αγοράς παρουσιάζει ήπια διαφοροποίηση, χωρίς έντονη γραμμική συσχέτιση (βλ. [Εικόνα 8](#)). Παρότι παρατηρείται μεγάλη διασπορά τιμών σε όλα τα επίπεδα ικανοποίησης, οι υψηλότερες βαθμολογίες satisfaction τείνουν να συνδέονται με ελαφρώς αυξημένες μέσες δαπάνες. Το μοτίβο αυτό υποδηλώνει ότι η ικανοποίηση αποτελεί σημαντικό, αλλά όχι αποκλειστικό, παράγοντα που επηρεάζει την αγοραστική αξία.

Τεχνική Σημείωση: Η μεταβλητή Customer_Satisfaction χρησιμοποιήθηκε ως αριθμητική κλίμακα (1–10). Η μέση αξία αγοράς υπολογίστηκε μέσω του $AVG(Purchase_Amount_Clean)$, ενώ κάθε σημείο αντιστοιχεί σε μοναδικό Customer_ID. Προστέθηκε γραμμή τάσης (trend line) για την αποτύπωση της γενικής σχέσης μεταξύ των τριών μεταβλητών.



Εικόνα 8. Customer Satisfaction.

7.3.2 Επάγγελμα και Ικανοποίηση Πελατών

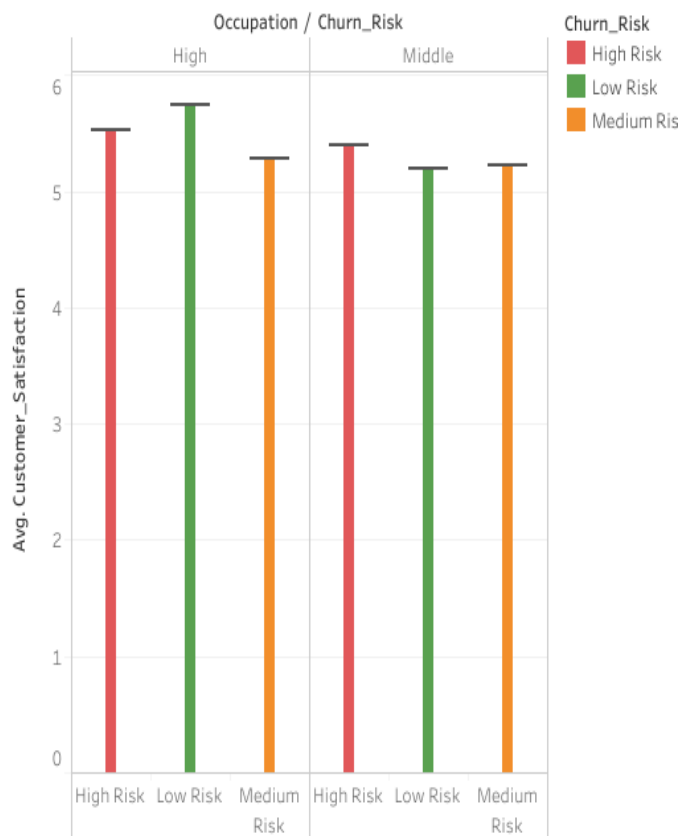
Τα επίπεδα ικανοποίησης των πελατών παρουσιάζουν διαφοροποίηση κυρίως σε συνάρτηση με το επίπεδο κινδύνου αποχώρησης (churn risk) και λιγότερο ως προς την επαγγελματική κατηγορία (βλ. [Εικόνα 9](#)). Συγκεκριμένα, σε όλες τις επαγγελματικές ομάδες παρατηρείται ότι οι πελάτες χαμηλού κινδύνου αποχώρησης εμφανίζουν υψηλότερους μέσους δείκτες ικανοποίησης, ενώ οι ομάδες μεσαίου και υψηλού churn risk καταγράφουν σταδιακά χαμηλότερες τιμές.

Η σύγκριση μεταξύ επαγγελματικών κατηγοριών υψηλής και μεσαίας ειδίκευσης δείχνει ότι οι διαφορές στην ικανοποίηση είναι περιορισμένες, γεγονός που υποδηλώνει ότι η επαγγελματική ιδιότητα δεν αποτελεί από μόνη της καθοριστικό παράγοντα της εμπειρίας ηλεκτρονικών αγορών. Αντίθετα, ο δείκτης κινδύνου αποχώρησης φαίνεται να συνδέεται πιο άμεσα με την αντιλαμβανόμενη ικανοποίηση των πελατών.

Το εύρημα αυτό ενισχύει τη χρησιμότητα της ενσωμάτωσης του churn risk σε αναλύσεις εμπειρίας πελάτη, καθώς επιτρέπει την καλύτερη κατανόηση ομάδων που, παρά παρόμοιο επαγγελματικό προφίλ, εμφανίζουν διαφορετικά επίπεδα ικανοποίησης και ενδεχόμενης αποχώρησης.

Τεχνική Σημείωση: Η μεταβλητή Occupation προέκυψε από custom grouping σε δύο κατηγορίες (High και Middle). Ο δείκτης Customer Satisfaction χρησιμοποιήθηκε ως αριθμητική κλίμακα και απεικονίστηκε μέσω μέσων τιμών ανά επίπεδο Churn Risk (Low, Medium, High), επιτρέποντας τη συγκριτική αξιολόγηση της ικανοποίησης σε συνάρτηση με τον κίνδυνο αποχώρησης και την επαγγελματική κατηγορία.

Average Customer Satisfaction by Occupation and Churn Risk Level



Εικόνα 9. Customer Satisfaction.

7.4 Ανάλυση Κινδύνου Απώλειας Πελάτη (Churn Risk)

Η ενότητα αυτή εστιάζει στην ανάλυση του κινδύνου απώλειας πελατών, εξετάζοντας τη σχέση του churn risk με βασικούς δείκτες συμπεριφοράς και εμπειρίας. Μέσω συνδυαστικών οπτικοποιήσεων, αναδεικνύονται μοτίβα που συμβάλλουν στην κατανόηση παραγόντων αποχώρησης και στη μελλοντική στοχευμένη διαχείριση πελατών.

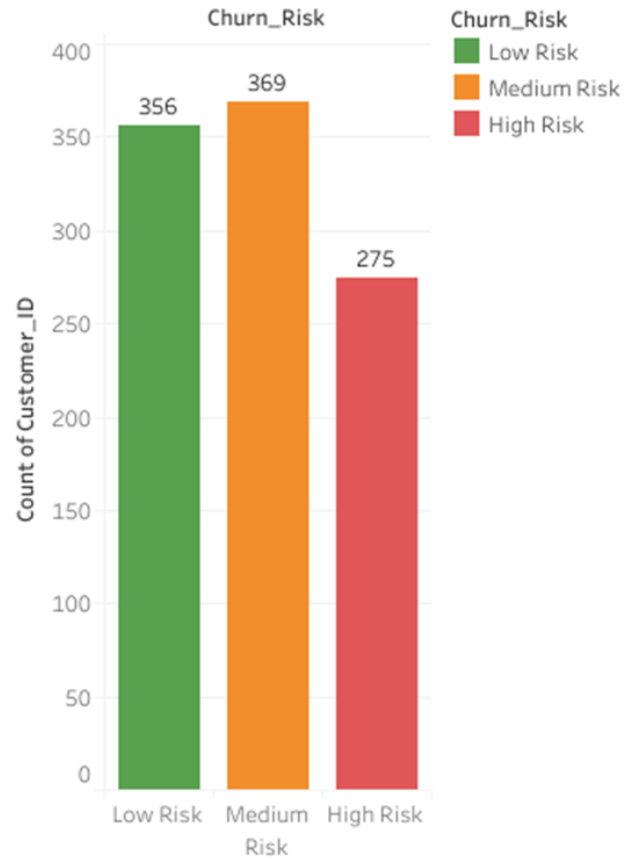
7.4.1 Συνολικός Δείκτης Churn Risk

Η κατανομή των πελατών ως προς το επίπεδο κινδύνου απώλειας δείχνει ότι η πλειονότητα του δείγματος εντάσσεται στις κατηγορίες χαμηλού και μεσαίου κινδύνου, ενώ μικρότερο ποσοστό εμφανίζει υψηλό churn risk (βλ. [Εικόνα 10](#)). Το εύρημα αυτό υποδηλώνει ότι, παρότι η βάση πελατών παρουσιάζει σχετικά σταθερή συμπεριφορά, υπάρχει ένα σημαντικό τμήμα που ενδέχεται να αποχωρήσει και χρήζει περαιτέρω ανάλυσης.

Τεχνική Σημείωση: Η μεταβλητή Churn_Risk δημιουργήθηκε ως κατηγοριοποιημένος δείκτης τριών επιπέδων (Low, Medium, High) βάσει

συνδυαστικών κριτηρίων συμπεριφοράς. Για την απεικόνιση χρησιμοποιήθηκε count του Customer_ID, ώστε να αποτυπωθεί ο αριθμός μοναδικών πελατών σε κάθε κατηγορία κινδύνου.

Churn Risk Analysis

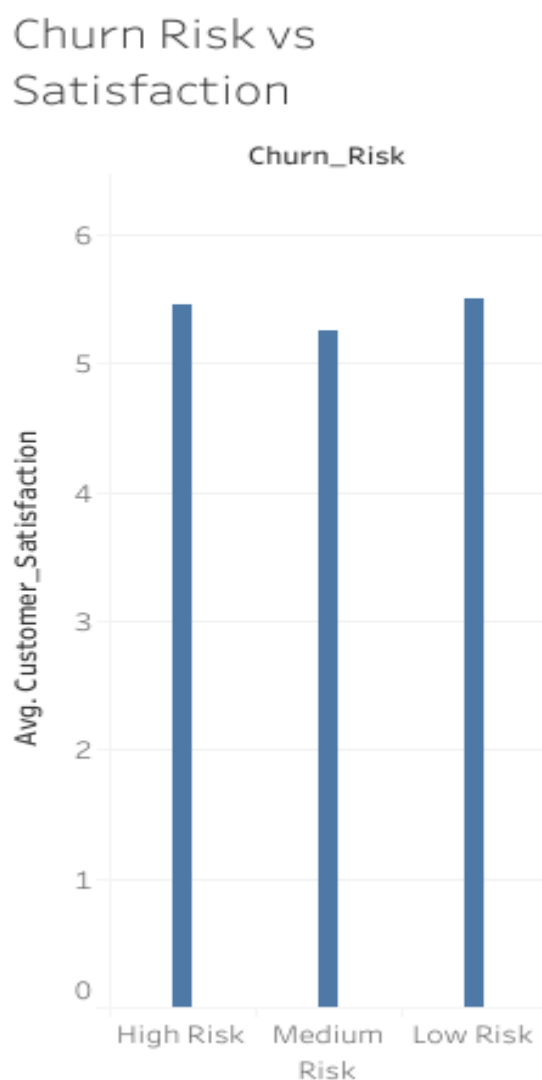


Εικόνα 10. Churn Risk.

7.4.2 Churn Risk και Ικανοποίηση Πελατών

Παρατηρείται αντίστροφη σχέση μεταξύ του επιπέδου churn risk και της μέσης ικανοποίησης των πελατών ([βλ. Εικόνα 11](#)). Οι πελάτες χαμηλού κινδύνου εμφανίζουν υψηλότερα επίπεδα ικανοποίησης, ενώ όσοι κατατάσσονται σε υψηλό churn risk παρουσιάζουν χαμηλότερες μέσες τιμές. Το εύρημα αυτό υποστηρίζει τη σημασία της ικανοποίησης ως βασικού παράγοντα διατήρησης πελατών και ενισχύει τη χρησιμότητά της ως μεταβλητής σε μοντέλα πρόβλεψης αποχώρησης.

Τεχνική Σημείωση: Η μέση ικανοποίηση υπολογίστηκε μέσω του AVG(Customer_Satisfaction) ανά κατηγορία Churn_Risk. Ο δείκτης churn risk προέκυψε από custom κατηγοριοποίηση πελατών, ενώ η απεικόνιση πραγματοποιήθηκε με bar chart για τη διευκόλυνση της σύγκρισης μεταξύ των επιπέδων κινδύνου.



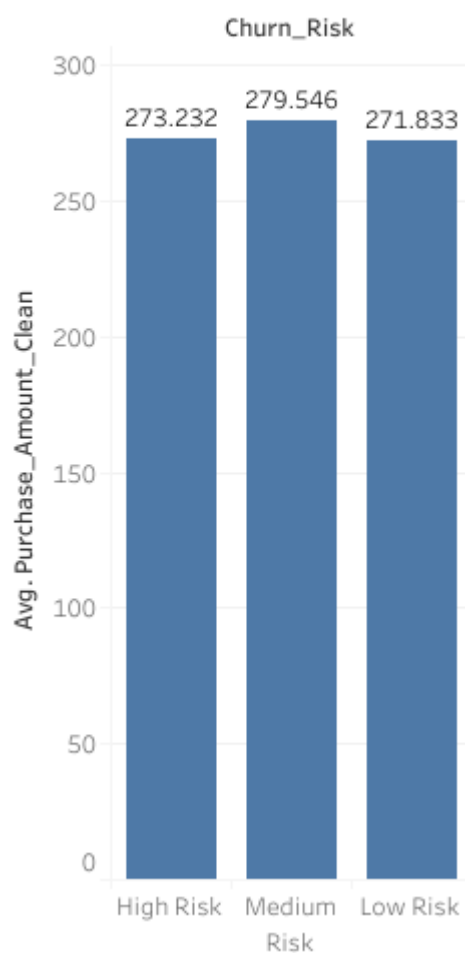
Εικόνα 11.Churn Risk vs Satisfaction.

7.4.3 Churn Risk και Αξία Αγορών

Η μέση αξία αγορών δεν παρουσιάζει σημαντικές διαφοροποιήσεις μεταξύ των επιπέδων churn risk (βλ. [Εικόνα 12](#)). Οι πελάτες χαμηλού, μεσαίου και υψηλού κινδύνου εμφανίζουν παρόμοια επίπεδα μέσης δαπάνης, γεγονός που υποδηλώνει ότι το ύψος της αγοραστικής αξίας από μόνο του δεν αποτελεί επαρκή δείκτη πρόβλεψης αποχώρησης. Το εύρημα αυτό ενισχύει την ανάγκη συνδυαστικής ανάλυσης με άλλους δείκτες συμπεριφοράς και εμπειρίας.

Τεχνική Σημείωση: Η μέση αξία αγορών υπολογίστηκε μέσω του AVG(Purchase_Amount_Clean) για κάθε κατηγορία Churn_Risk. Το πεδίο Purchase_Amount_Clean αποτελεί καθαρισμένη εκδοχή της αρχικής μεταβλητής αγορών, με αφαίρεση μη έγκυρων και ακραίων τιμών.

Churn Risk vs Purchase Amount



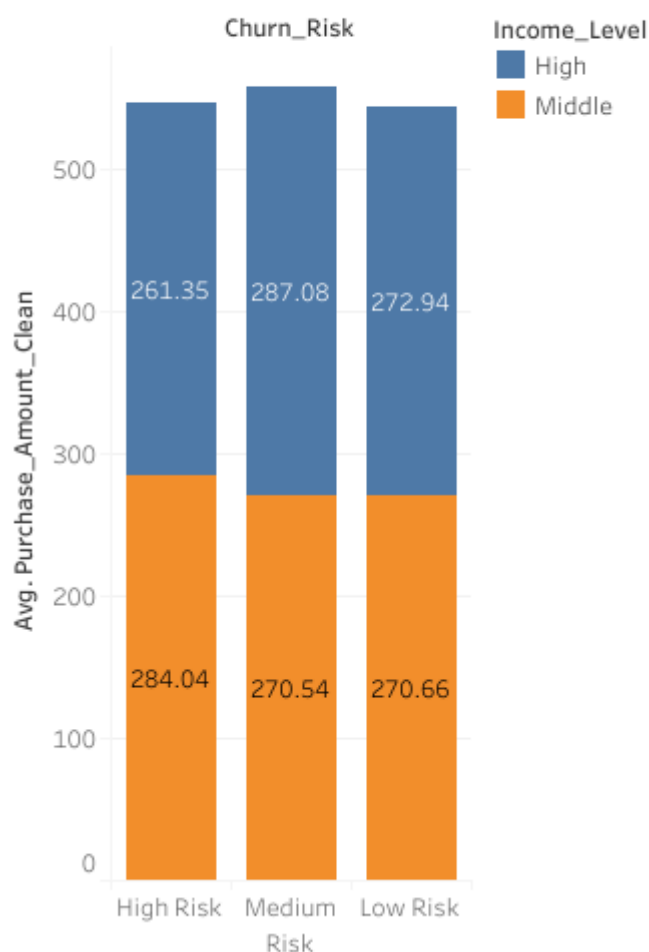
Εικόνα 12.Churn Risk vs Purchase Amount.

7.4.4 Churn Risk και Εισοδηματικό Επίπεδο

Η μέση αξία αγορών εμφανίζει μικρές διαφοροποιήσεις μεταξύ των εισοδηματικών επιπέδων σε όλες τις κατηγορίες churn risk (βλ. [Εικόνα 13](#)). Οι πελάτες υψηλού εισοδήματος διατηρούν σταθερά υψηλότερη μέση δαπάνη σε σύγκριση με το μεσαίο εισόδημα, ανεξαρτήτως επιπέδου κινδύνου αποχώρησης. Το εύρημα αυτό υποδηλώνει ότι, παρότι το εισόδημα επηρεάζει το ύψος της δαπάνης, δεν αποτελεί από μόνο του καθοριστικό παράγοντα για τον χαρακτηρισμό του churn risk.

Τεχνική Σημείωση: Η ανάλυση βασίστηκε στον υπολογισμό του $AVG(Purchase_Amount_Clean)$ ανά κατηγορία Churn_Risk και Income_Level. Το Income_Level προέκυψε από προϋπάρχουσα κατηγοριοποίηση εισοδήματος, ενώ η απεικόνιση πραγματοποιήθηκε με stacked bar chart για τη σύγκριση της συνεισφοράς κάθε εισοδηματικού επιπέδου.

Churn Risk vs Income Level

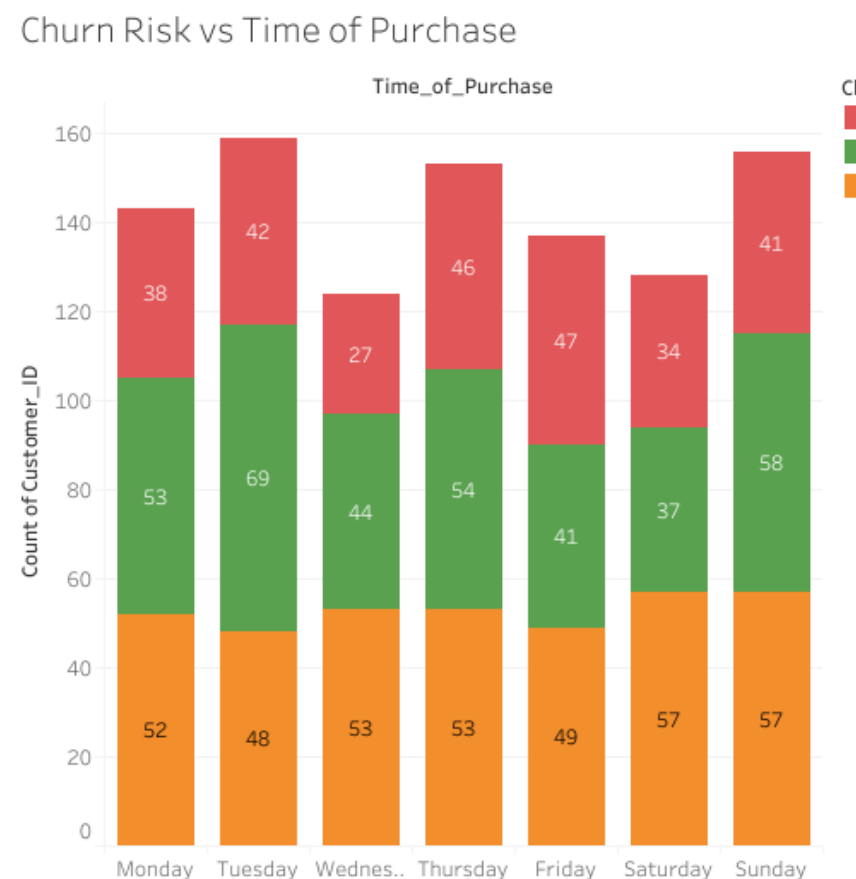


Εικόνα 13.Churn Risk vs Income Level .

7.4.5 Χρονική Συμπεριφορά: Churn Risk vs Time of Purchase

Η κατανομή του churn risk παρουσιάζει διαφοροποιήσεις ανά ημέρα πραγματοποίησης της αγοράς (βλ. [Εικόνα 14](#)). Παρατηρείται αυξημένη συγκέντρωση πελατών με χαμηλό και μεσαίο κίνδυνο αποχώρησης κατά τις καθημερινές, ενώ τα Σαββατοκύριακα εμφανίζουν ελαφρώς υψηλότερη παρουσία πελατών με αυξημένο churn risk. Το εύρημα αυτό υποδηλώνει ότι ο χρόνος και το πλαίσιο της αγοράς ενδέχεται να σχετίζονται με πιο ευκαιριακή ή λιγότερο σταθερή αγοραστική συμπεριφορά.

Τεχνική Σημείωση: Η μεταβλητή `Time_of_Purchase` μετασχηματίστηκε σε ημέρα εβδομάδας μέσω της συνάρτησης `WEEKDAY()` στο Tableau. Η ανάλυση βασίστηκε στο `count` του `Customer_ID` ανά ημέρα και κατηγορία `Churn_Risk`, ενώ χρησιμοποιήθηκε `stacked bar chart` για την ταυτόχρονη αποτύπωση της κατανομής των επιπέδων κινδύνου.



Εικόνα 14. Churn Risk vs Time of Purchase .

7.5 Ανάλυση Αξίας Πελάτη

Η ανάλυση του churn risk ανέδειξε ότι η πιθανότητα αποχώρησης δεν σχετίζεται αποκλειστικά με την αξία των μεμονωμένων αγορών, αλλά με ένα σύνολο συμπεριφορικών και εμπειρικών παραγόντων. Για τον λόγο αυτό, κρίνεται σκόπιμο να εξεταστεί η συνολική αξία του πελάτη στον χρόνο, ώστε να αποτυπωθεί η στρατηγική

σημασία κάθε κατηγορίας κινδύνου. Στο πλαίσιο αυτό, ακολουθεί η ανάλυση του Customer Lifetime Value μέσω proxy δείκτη.

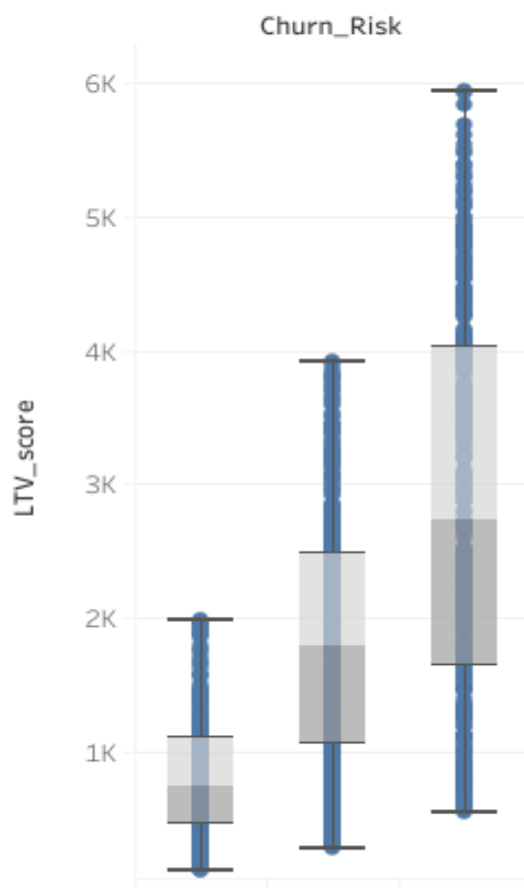
7.5.1 Customer Lifetime Value (Proxy)

Το Customer Lifetime Value υπολογίστηκε μέσω proxy δείκτη (LTV_score), λόγω έλλειψης μακροχρόνιων ιστορικών δεδομένων. Ο δείκτης δημιουργήθηκε ως custom calculated field στο Tableau και ορίστηκε ως το γινόμενο της καθαρισμένης αξίας αγοράς και της συχνότητας αγορών, σύμφωνα με τη σχέση:

$$\text{LTV_score} = \text{Purchase_Amount_Clean} \times \text{Frequency_of_Purchase}$$

Η κατανομή του CLV ανά επίπεδο churn risk παρουσιάζεται στο διάγραμμα box plot (βλ. [Εικόνα 15](#)), όπου παρατηρείται ότι οι πελάτες χαμηλού κινδύνου εμφανίζουν υψηλότερη διάμεσο αξία και μικρότερη διασπορά, ενώ οι πελάτες υψηλού churn risk χαρακτηρίζονται από χαμηλότερη και πιο ασταθή αξία. Το εύρημα αυτό επιβεβαιώνει τη στενή σχέση μεταξύ churn risk και στρατηγικής αξίας πελατών.

Customer Lifetime Value
(proxy)



Εικόνα 15 .CLV .

ΚΕΦΑΛΑΙΟ 8: ΠΡΟΒΛΕΠΤΙΚΗ ΑΝΑΛΥΣΗ ΚΑΙ ΣΥΝΘΕΣΗ ΕΥΡΗΜΑΤΩΝ

Το Κεφάλαιο 7 επικεντρώθηκε στην περιγραφική και διερευνητική ανάλυση της συμπεριφοράς των καταναλωτών, αναδεικνύοντας δημογραφικά, συμπεριφορικά και εμπειρικά πρότυπα, καθώς και τη σχέση τους με την αγοραστική δραστηριότητα και τον κίνδυνο απώλειας πελατών. Αν και τα ευρήματα αυτά παρέχουν ουσιαστική κατανόηση του υφιστάμενου προφίλ των καταναλωτών, δεν επιτρέπουν την εκτίμηση μελλοντικών εξελίξεων. Στο πλαίσιο αυτό, το παρόν κεφάλαιο μετατοπίζει το επίκεντρο από την περιγραφή στην προβλεπτική προσέγγιση, συνδυάζοντας θεωρητικά στοιχεία και εμπειρικά δεδομένα με στόχο την υποστήριξη της λήψης αποφάσεων.

8.1 Στόχος και Μεθοδολογικό Πλαίσιο Προβλεπτικής Ανάλυσης

8.1.1 Στόχος της Προβλεπτικής Ανάλυσης

Σκοπός του παρόντος κεφαλαίου είναι η εκτίμηση μελλοντικών συμπεριφορών των καταναλωτών μέσω της ανάλυσης του κινδύνου απώλειας πελατών (churn risk) και της αξίας διάρκειας ζωής τους. Η προβλεπτική ανάλυση δεν αποσκοπεί στην ακριβή πρόβλεψη μεμονωμένων γεγονότων, αλλά στην αναγνώριση προτύπων που υποδηλώνουν αυξημένη ή μειωμένη πιθανότητα αποχώρησης, επιτρέποντας την κατηγοριοποίηση των πελατών σε ομάδες διαφορετικής στρατηγικής σημασίας

8.1.2 Μεθοδολογικό Πλαίσιο και Επιλογή Εργαλείων

Η προβλεπτική προσέγγιση υλοποιήθηκε στο περιβάλλον του Tableau, αξιοποιώντας υπολογισμένα πεδία, παραμέτρους και συνδυαστικές οπτικοποιήσεις. Παρότι το Tableau δεν αποτελεί εξειδικευμένο εργαλείο ανάπτυξης μοντέλων μηχανικής μάθησης, προσφέρει τη δυνατότητα δημιουργίας ερμηνεύσιμων proxy δεικτών, οι οποίοι υποστηρίζουν τη συγκριτική αξιολόγηση πελατών. Η προσέγγιση που ακολουθείται είναι εξηγητική και προσανατολισμένη στη λήψη αποφάσεων, γεφυρώνοντας τη θεωρητική έννοια της πρόβλεψης με την πρακτική εφαρμογή της αναλυτικής δεδομένων.

Η επιλογή μιας προβλεπτικής προσέγγισης βασισμένης σε proxy δείκτες αντανακλά τις συνθήκες που επικρατούν στη συντριπτική πλειονότητα των πραγματικών επιχειρησιακών περιβαλλόντων, όπου τα πλήρη και «καθαρά» δεδομένα που απαιτούνται για την εκπαίδευση αυστηρών μοντέλων μηχανικής μάθησης δεν είναι πάντοτε διαθέσιμα. Στο πλαίσιο αυτό, η παρούσα ανάλυση δεν επιχειρεί να υποκαταστήσει εξειδικευμένα συστήματα πρόβλεψης, αλλά να προσφέρει μια ερμηνεύσιμη και λειτουργική προσέγγιση εκτίμησης μελλοντικής συμπεριφοράς, αξιοποιώντας παρατηρήσιμα μοτίβα συναλλακτικής δραστηριότητας.

Η χρήση του Tableau ως βασικού αναλυτικού εργαλείου επιτρέπει τη σύνθεση διαφορετικών διαστάσεων συμπεριφοράς — όπως η συχνότητα αγορών, η πρόσφατη δραστηριότητα και η συνολική αξία συναλλαγών — σε ενιαίους δείκτες εκτίμησης κινδύνου και αξίας. Οι δείκτες αυτοί δεν συνιστούν απόλυτες προβλέψεις, αλλά

λειτουργούν ως μηχανισμοί κατάταξης και συγκριτικής αξιολόγησης πελατών, υποστηρίζοντας τη λήψη αποφάσεων σε στρατηγικό επίπεδο.

Ιδιαίτερη σημασία έχει το γεγονός ότι η προβλεπτική ανάλυση προσεγγίζεται ως δυναμική διαδικασία και όχι ως στατικό αποτέλεσμα. Οι εκτιμήσεις churn risk και αξίας διάρκειας ζωής αποτυπώνουν την παρούσα κατάσταση της σχέσης πελάτη-επιχείρησης και δύνανται να μεταβληθούν με την πάροδο του χρόνου, καθώς διαφοροποιούνται τα πρότυπα συμπεριφοράς. Με τον τρόπο αυτό, η ανάλυση αποκτά χαρακτήρα «early warning system», επιτρέποντας την έγκαιρη αναγνώριση μεταβολών που ενδέχεται να επηρεάσουν τη διατήρηση ή την αξία των πελατών.

Συνολικά, το μεθοδολογικό πλαίσιο της παρούσας ενότητας εστιάζει στη γεφύρωση της προβλεπτικής λογικής με την επιχειρησιακή πρακτική, δίνοντας έμφαση στην ερμηνευσιμότητα, τη διαφάνεια και τη χρηστικότητα των αποτελεσμάτων, στοιχεία κρίσιμα για την αποτελεσματική αξιοποίηση της αναλυτικής δεδομένων στο περιβάλλον του ηλεκτρονικού εμπορίου.

8.2 Προβλεπτική Ανάλυση Κινδύνου Απώλειας Πελατών (Churn Prediction)

8.2.1 Θεωρητικό Πλαίσιο Πρόβλεψης Churn

Ο κίνδυνος απώλειας πελατών αποτελεί κρίσιμο παράγοντα για τη βιωσιμότητα των επιχειρήσεων ηλεκτρονικού εμπορίου. Η βιβλιογραφία προσεγγίζει την πρόβλεψη churn κυρίως μέσω εποπτευόμενων αλγορίθμων μηχανικής μάθησης, αξιοποιώντας μεταβλητές όπως η συχνότητα αγορών, η αξία συναλλαγών και δείκτες ικανοποίησης. Σε περιβάλλοντα όπου δεν είναι διαθέσιμα επαρκή δεδομένα ή σαφώς ορισμένο γεγονός αποχώρησης, η πρόβλεψη μπορεί να προσεγγιστεί μέσω proxy δεικτών που αποτυπώνουν τη σχετική επικινδυνότητα κάθε πελάτη. Η προσέγγιση του churn ως πιθανότητα και όχι ως βέβαιο γεγονός είναι ιδιαίτερα σημαντική στο πλαίσιο της ανάλυσης δεδομένων ηλεκτρονικού εμπορίου. Στην πράξη, η αποχώρηση ενός πελάτη δεν συντελείται στιγμιαία, αλλά αποτελεί αποτέλεσμα σταδιακής μεταβολής της συμπεριφοράς του, όπως η μείωση της συχνότητας αγορών ή της αλληλεπίδρασης με την πλατφόρμα. Ως εκ τούτου, η πρόβλεψη churn δεν περιορίζεται στην αναγνώριση ενός τελικού γεγονότος, αλλά στην έγκαιρη ανίχνευση προειδοποιητικών ενδείξεων.

Η χρήση proxy δεικτών στο θεωρητικό πλαίσιο πρόβλεψης churn επιτρέπει την προσέγγιση αυτής της έννοιας ακόμη και σε περιβάλλοντα όπου απουσιάζουν ρητές μεταβλητές αποχώρησης. Οι δείκτες αυτοί λειτουργούν ως ενδιάμεσες μετρήσεις κινδύνου, παρέχοντας σχετική και όχι απόλυτη εκτίμηση της πιθανότητας απώλειας. Με τον τρόπο αυτό, η ανάλυση εστιάζει περισσότερο στη συγκριτική αξιολόγηση πελατών παρά στην αυστηρή ταξινόμηση, στοιχείο που ευθυγραμμίζεται με τις ανάγκες στρατηγικής διαχείρισης πελατειακών σχέσεων.

8.2.2 Υλοποίηση Προβλεπτικής Ανάλυσης Churn στο Tableau

Στο πλαίσιο της παρούσας εργασίας, η προβλεπτική διερεύνηση του κινδύνου απώλειας πελατών υλοποιήθηκε στο περιβάλλον του Tableau μέσω υπολογισμένων πεδίων και κανονικοποιημένων δεικτών. Η προσέγγιση που ακολουθήθηκε δεν στοχεύει στην εκπαίδευση αλγοριθμικού μοντέλου μηχανικής μάθησης, αλλά στη δημιουργία ενός ερμηνεύσιμου σύνθετου δείκτη churn score, βασισμένου σε ιστορικά δεδομένα συμπεριφοράς και εμπειρίας πελατών.

Ο δείκτης churn score κατασκευάστηκε ως proxy μέτρο κινδύνου αποχώρησης, συνδυάζοντας μεταβλητές που έχουν αναγνωριστεί στη βιβλιογραφία ως κρίσιμες για την πρόβλεψη churn, όπως η συχνότητα αγορών, η μέση αξία συναλλαγών και το επίπεδο ικανοποίησης. Μέσω της κανονικοποίησης των επιμέρους μεταβλητών και της στάθμισής τους, επιτυγχάνεται η συγκριτική αξιολόγηση των πελατών ως προς τον σχετικό κίνδυνο απώλειας.

Η προσέγγιση αυτή επιτρέπει την οπτική διερεύνηση του churn όχι ως δυαδικό γεγονός, αλλά ως συνεχές φάσμα επικινδυνότητας, διευκολύνοντας τη σύνδεσή του με άλλους στρατηγικούς δείκτες, όπως η αξία διάρκειας ζωής του πελάτη. Στο πλαίσιο αυτό, η προβλεπτική ανάλυση επεκτείνεται πέρα από την απλή κατηγοριοποίηση κινδύνου και οδηγεί στη συνδυαστική αξιολόγηση churn και επιχειρησιακής αξίας. Η επιλογή ενός σύνθετου churn score, αντί ενός δυαδικού δείκτη αποχώρησης, επιτρέπει την αποτύπωση διαφοροποιήσεων μεταξύ πελατών με παρόμοια χαρακτηριστικά αλλά διαφορετική ένταση κινδύνου. Η συνεχής κλίμακα του δείκτη διευκολύνει την κατάταξη των πελατών σε επίπεδα προτεραιότητας, υποστηρίζοντας τη στοχευμένη εφαρμογή πολιτικών διατήρησης. Επιπλέον, η οπτικοκεντρική προσέγγιση του Tableau ενισχύει την ερμηνευσιμότητα των αποτελεσμάτων, καθιστώντας τον δείκτη άμεσα αξιοποιήσιμο από στελέχη που δεν διαθέτουν εξειδικευμένες γνώσεις μηχανικής μάθησης.

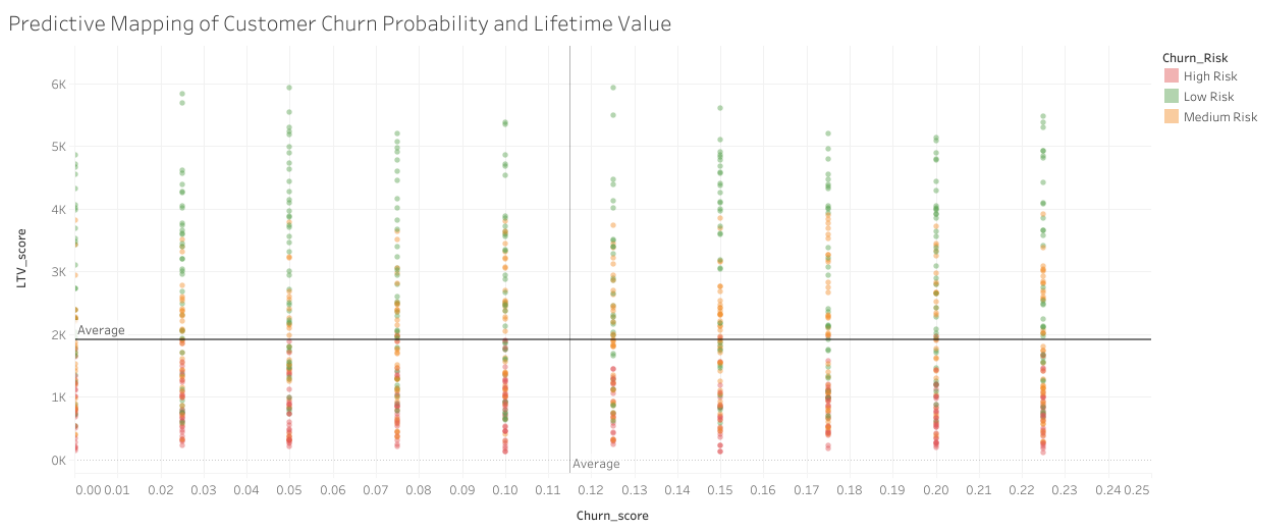
Η προβλεπτική ανάλυση churn, όπως υλοποιείται στην παρούσα εργασία, λειτουργεί συνεπώς ως εργαλείο υποστήριξης αποφάσεων και όχι ως αυτοματοποιημένος μηχανισμός πρόβλεψης. Ο ρόλος της είναι να αναδεικνύει ομάδες πελατών με αυξημένη πιθανότητα μελλοντικής αποχώρησης, παρέχοντας έγκαιρη πληροφόρηση για την εφαρμογή διορθωτικών ενεργειών, όπως προγράμματα πιστότητας ή εξατομικευμένες προσφορές.

8.2.3 Σχέση Προβλεπτικού Δείκτη Churn Score και Customer Lifetime Value

Η σχέση μεταξύ του churn score και του Customer Lifetime Value (proxy) αποτυπώνεται μέσω συνδυαστικής οπτικοποίησης τύπου scatter plot (βλ. [Εικόνα 16](#)). Η ανάλυση αναδεικνύει τάση αρνητικής συσχέτισης, καθώς πελάτες με χαμηλότερο κίνδυνο αποχώρησης τείνουν να εμφανίζουν υψηλότερη αξία διάρκειας ζωής. Ιδιαίτερη σημασία παρουσιάζει η ομάδα πελατών υψηλής αξίας αλλά αυξημένου κινδύνου, η οποία αποτελεί βασικό στόχο στοχευμένων στρατηγικών διατήρησης. Η συνδυαστική ανάλυση churn score και Customer Lifetime Value αναδεικνύει τη

στρατηγική διάσταση της προβλεπτικής ανάλυσης. Ιδιαίτερα κρίσιμη είναι η ταυτοποίηση πελατών που, παρά την υψηλή εκτιμώμενη αξία διάρκειας ζωής, εμφανίζουν αυξημένο κίνδυνο αποχώρησης. Η συγκεκριμένη κατηγορία πελατών συνιστά σημείο υψηλής προτεραιότητας για τις επιχειρήσεις, καθώς η απώλειά τους ενδέχεται να έχει δυσανάλογα αρνητικό αντίκτυπο στα μελλοντικά έσοδα.

Αντίθετα, πελάτες χαμηλής αξίας και υψηλού κινδύνου δύνανται να αντιμετωπιστούν με διαφορετική στρατηγική, γεγονός που υπογραμμίζει τη σημασία της πολυδιάστατης προσέγγισης στην ανάλυση churn. Με τον τρόπο αυτό, η προβλεπτική ανάλυση μεταβαίνει από μια απλή εκτίμηση κινδύνου σε εργαλείο στρατηγικής ιεράρχησης πελατειακών σχέσεων.

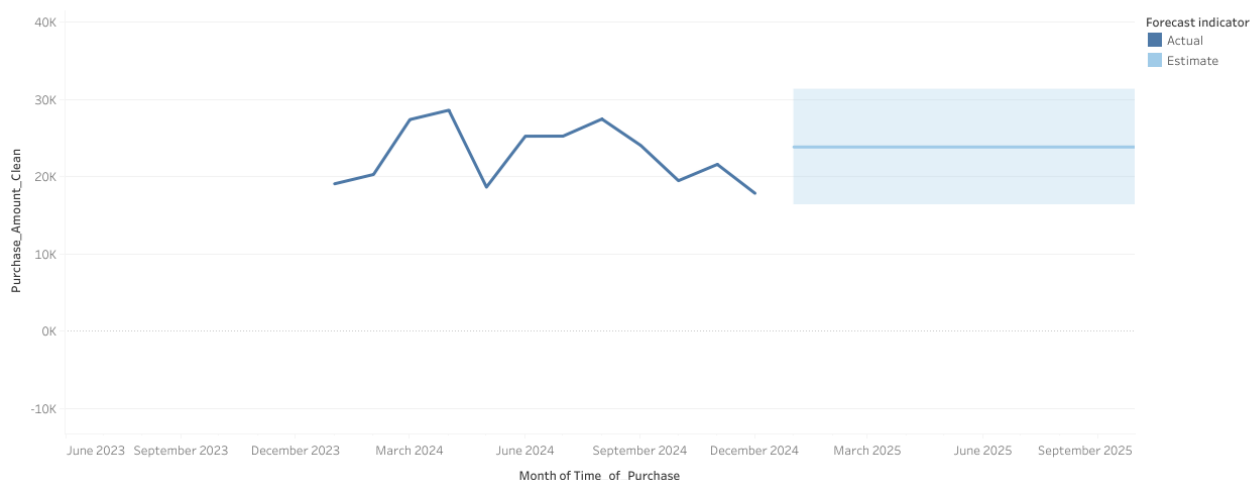


Εικόνα 16. Predictive mapping of Customer churn Probability and Lifetime Value.

8.2.4 Περιορισμοί Χρονικής Πρόβλεψης

Παρότι το Tableau παρέχει ενσωματωμένες δυνατότητες χρονικής πρόβλεψης (forecasting), η εφαρμογή τους στην παρούσα εργασία παρουσίασε σημαντικούς περιορισμούς λόγω του περιορισμένου χρονικού εύρους των διαθέσιμων δεδομένων. Το σύνολο δεδομένων περιλάμβανε παρατηρήσεις για ένα μόνο έτος, γεγονός που δεν επαρκεί για την αξιόπιστη εκτίμηση τάσεων και εποχικότητας. Ως αποτέλεσμα, το προβλεπτικό μοντέλο παρήγαγε σταθερές (flat) εκτιμήσεις για τους μελλοντικούς μήνες, οι οποίες δεν προσθέτουν ουσιαστική αναλυτική αξία. Για τον λόγο αυτό, η χρονική πρόβλεψη αξιολογείται ενδεικτικά και όχι ως πλήρως αξιόπιστο εργαλείο πρόβλεψης, με το βάρος της ανάλυσης να μετατοπίζεται σε δείκτες proxy και προβλεπτική τμηματοποίηση πελατών. Ενδεικτικό αποτέλεσμα της χρονικής πρόβλεψης παρουσιάζεται στην [Εικόνα 17](#). Οι παραπάνω περιορισμοί καταδεικνύουν ότι η χρονική πρόβλεψη πωλήσεων ή συμπεριφοράς απαιτεί επαρκές ιστορικό βάθος δεδομένων, ώστε να αποτυπώνονται με αξιοπιστία οι διακυμάνσεις και τα εποχικά μοτίβα. Στο πλαίσιο της παρούσας εργασίας, η περιορισμένη χρονική κάλυψη καθιστά την πρόβλεψη περισσότερο ενδεικτική παρά προβλεπτικά αξιόπιστη. Για τον λόγο αυτό, η ανάλυση επικεντρώνεται σε διαχρονικά ανεξάρτητους δείκτες συμπεριφοράς και κινδύνου, οι οποίοι προσφέρουν μεγαλύτερη ερμηνευτική και στρατηγική αξία.

Forecast Number of Purchases ανά μήνα



Εικόνα 17. Forecast Number of Purchases via month .

8.3 Προβλεπτική Τμηματοποίηση Πελατών βάσει Κινδύνου και Αξίας

Η συνδυαστική αξιολόγηση της αγοραστικής συμπεριφοράς και του δείκτη κινδύνου αποχώρησης (churn score) επιτρέπει την ουσιαστική τμηματοποίηση της πελατειακής βάσης σε ομάδες διαφορετικής στρατηγικής σημασίας. Η προσέγγιση αυτή υπερβαίνει τη μονοδιάστατη ανάλυση μεμονωμένων δεικτών και προσφέρει ένα ολοκληρωμένο πλαίσιο κατανόησης της συμπεριφοράς των πελατών, διευκολύνοντας τη στοχευμένη λήψη επιχειρησιακών αποφάσεων.

Η τμηματοποίηση αποτυπώνεται στην Εικόνα 18, όπου οι πελάτες ομαδοποιούνται σε τέσσερα διακριτά clusters βάσει της μέσης αξίας αγορών (Average Purchase Amount) και της μέσης συχνότητας αγορών (Average Frequency of Purchase). Οι προκύπτουσες συστάδες αντανakλούν διαφορετικά πρότυπα αγοραστικής συμπεριφοράς, επιτρέποντας τη διάκριση μεταξύ πελατών χαμηλής, μεσαίας και υψηλής αξίας.

Παράλληλα, η ενσωμάτωση του churn score στο προφίλ κάθε συστάδας προσφέρει πρόσθετη πληροφορία σχετικά με τον κίνδυνο αποχώρησης των πελατών εντός κάθε ομάδας. Μέσω της προσέγγισης αυτής καθίσταται δυνατός ο εντοπισμός πελατών υψηλής αγοραστικής αξίας που ενδέχεται να εμφανίζουν αυξημένο churn risk, αλλά και πελατών χαμηλότερης αξίας με σταθερή συμπεριφορά.

Η συγκεκριμένη ανάλυση παρέχει ένα λειτουργικό πλαίσιο υποστήριξης αποφάσεων, καθώς επιτρέπει την ιεράρχηση πελατών και τον σχεδιασμό διαφοροποιημένων στρατηγικών διατήρησης, εξατομίκευσης και βελτιστοποίησης επιχειρησιακών πόρων, με βάση τόσο τη συμπεριφορά όσο και τον προβλεπόμενο κίνδυνο αποχώρησης.

Customer Segmentation using Clustering Analysis and Churn Score Profiling



Εικόνα 18. Customer Segmentation.

8.3.1 Πελάτες Υψηλής Αξίας και Χαμηλού Κινδύνου (Core Customers)

Οι πελάτες που συνδυάζουν χαμηλό churn score με υψηλό Customer Lifetime Value αποτελούν τον βασικό πυρήνα της πελατειακής βάσης. Χαρακτηρίζονται από σταθερή αγοραστική συμπεριφορά, υψηλό επίπεδο ικανοποίησης και αυξημένη πιστότητα, συμβάλλοντας ουσιαστικά στη μακροχρόνια βιωσιμότητα της επιχείρησης. Για την κατηγορία αυτή, η στρατηγική επικεντρώνεται στη διατήρηση της θετικής εμπειρίας μέσω προσωποποιημένης επικοινωνίας, προγραμμάτων επιβράβευσης και συνεχούς ποιοτικής εξυπηρέτησης.

8.3.2 Πελάτες Υψηλής Αξίας και Υψηλού Κινδύνου (Critical Retention Segment)

Ιδιαίτερο ενδιαφέρον παρουσιάζει η ομάδα πελατών που, παρά την υψηλή αξία διάρκειας ζωής, εμφανίζει αυξημένο churn score. Η κατηγορία αυτή θεωρείται στρατηγικά κρίσιμη, καθώς η απώλειά της μπορεί να επιφέρει δυσανάλογα αρνητικό αντίκτυπο στα έσοδα. Η προβλεπτική ανάλυση επιτρέπει την έγκαιρη αναγνώριση των πελατών αυτών και καθιστά εφικτή την εφαρμογή στοχευμένων ενεργειών διατήρησης, όπως εξατομικευμένες προσφορές, βελτιώσεις στην εμπειρία εξυπηρέτησης και παρεμβάσεις ενίσχυσης εμπιστοσύνης.

8.3.3 Πελάτες Χαμηλής Αξίας και Χαμηλού Κινδύνου (Stable Low-Value Segment)

Οι πελάτες με χαμηλό churn score αλλά περιορισμένη αξία διάρκειας ζωής χαρακτηρίζονται από σταθερή, αλλά χαμηλής έντασης αγοραστική συμπεριφορά. Αν και δεν αποτελούν άμεση προτεραιότητα για επενδύσεις υψηλού κόστους, η διατήρησή τους είναι σημαντική για τη συνολική σταθερότητα της πελατειακής βάσης και μπορεί να υποστηριχθεί μέσω αυτοματοποιημένων και χαμηλού κόστους πρακτικών επικοινωνίας.

8.3.4 Πελάτες Χαμηλής Αξίας και Υψηλού Κινδύνου (Low Priority Segment)

Η κατηγορία πελατών με χαμηλό Customer Lifetime Value και αυξημένο churn score παρουσιάζει τη χαμηλότερη στρατηγική προτεραιότητα. Η προβλεπτική τμηματοποίηση επιτρέπει την αναγνώριση αυτής της ομάδας, ώστε οι διαθέσιμοι πόροι να κατευθύνονται αποτελεσματικότερα προς πελάτες υψηλότερης δυνητικής αξίας, χωρίς να υπονομεύεται η συνολική απόδοση της επιχείρησης. Συνολικά, η προβλεπτική τμηματοποίηση της πελατειακής βάσης προσφέρει ένα συνεκτικό και επιχειρησιακά αξιοποιήσιμο πλαίσιο λήψης αποφάσεων. Μέσω της ταυτόχρονης αξιολόγησης κινδύνου αποχώρησης και αξίας πελάτη, η ανάλυση μετατοπίζεται από την περιγραφική κατανόηση της συμπεριφοράς σε μια πιο στρατηγική και προσανατολισμένη στο μέλλον προσέγγιση. Η χρήση proxy δεικτών επιτρέπει την εφαρμογή της μεθοδολογίας ακόμη και σε περιβάλλοντα περιορισμένων δεδομένων, διασφαλίζοντας παράλληλα την ερμηνευσιμότητα και τη διαφάνεια των αποτελεσμάτων.

Με τον τρόπο αυτό, η παρούσα ενότητα αναδεικνύει τη σημασία της προβλεπτικής αναλυτικής ως εργαλείου υποστήριξης επιχειρησιακών αποφάσεων, το οποίο δεν αντικαθιστά την ανθρώπινη κρίση, αλλά τη συμπληρώνει με τεκμηριωμένες ενδείξεις και ιεραρχημένες πληροφορίες.

8.4 Ανάλυση Σεναρίων και Υποστήριξη Λήψης Αποφάσεων (What-if Analysis)

Η προβλεπτική ανάλυση επεκτείνεται πέρα από τη στατική αποτύπωση κινδύνου, ενσωματώνοντας την ανάλυση εναλλακτικών σεναρίων που υποστηρίζουν τη στρατηγική λήψη αποφάσεων. Στο πλαίσιο αυτό, η ανάλυση what-if επιτρέπει την αξιολόγηση του τρόπου με τον οποίο μεταβολές σε κρίσιμους δείκτες συμπεριφοράς και εμπειρίας πελατών επηρεάζουν τον προβλεπτικό δείκτη churn score και τη συνολική τμηματοποίηση.

Η υλοποίηση πραγματοποιήθηκε στο Tableau μέσω παραμέτρων (parameters), οι οποίες επιτρέπουν τη δυναμική προσαρμογή ορίων και βαρών στον υπολογισμό του churn score, καθιστώντας δυνατή την προσομοίωση διαφορετικών επιχειρησιακών καταστάσεων χωρίς τροποποίηση των αρχικών δεδομένων.

8.4.1 Προσαρμογή Ορίων Κινδύνου και Δυναμική Κατηγοριοποίηση

Η μεταβολή των ορίων που διαχωρίζουν τα επίπεδα κινδύνου αποχώρησης επηρεάζει άμεσα την κατανομή των πελατών ανά κατηγορία. Η ανάλυση δείχνει ότι ακόμη και μικρές αλλαγές στα thresholds του churn score μπορούν να οδηγήσουν σε σημαντική ανακατανομή, γεγονός που αναδεικνύει τη σημασία της ευελιξίας στην ερμηνεία των προβλεπτικών δεικτών. Η ευαισθησία της τμηματοποίησης ως προς τη μεταβολή των ορίων κινδύνου υπογραμμίζει ότι οι προβλεπτικοί δείκτες δεν θα πρέπει να ερμηνεύονται ως απόλυτες μετρήσεις, αλλά ως εργαλεία υποστήριξης αποφάσεων που απαιτούν επιχειρησιακή κρίση. Η δυνατότητα προσαρμογής των thresholds επιτρέπει στις επιχειρήσεις να ευθυγραμμίζουν την ανάλυση με τις εκάστοτε στρατηγικές τους προτεραιότητες, όπως η έμφαση στη διατήρηση πελατών υψηλής αξίας ή η συγκράτηση του κόστους παρεμβάσεων.

8.4.2 Επιπτώσεις Μεταβολών Συμπεριφορικών Δεικτών

Η προσαρμογή των βαρών επιμέρους δεικτών, όπως η συχνότητα αγορών και το επίπεδο ικανοποίησης, επιτρέπει τη διερεύνηση υποθετικών σεναρίων βελτίωσης ή επιδείνωσης της συμπεριφοράς των πελατών. Τα αποτελέσματα υποδεικνύουν ότι δείκτες που σχετίζονται με τη συνέπεια αγορών ασκούν ισχυρότερη επίδραση στον συνολικό κίνδυνο αποχώρησης σε σύγκριση με μεμονωμένες μεταβολές στην αξία συναλλαγών. Τα ευρήματα της ανάλυσης επιβεβαιώνουν ότι η συνέπεια και η συχνότητα αγορών αποτελούν ισχυρούς προγνωστικούς παράγοντες διατήρησης πελατών. Η διαπίστωση αυτή έχει ιδιαίτερη πρακτική σημασία, καθώς υποδεικνύει ότι στρατηγικές ενίσχυσης της επαναληψιμότητας, όπως προγράμματα πιστότητας ή στοχευμένες υπενθυμίσεις, ενδέχεται να έχουν μεγαλύτερη επίδραση στη μείωση του κινδύνου αποχώρησης σε σύγκριση με μεμονωμένες παρεμβάσεις τιμολόγησης.

8.4.3 Επιχειρησιακή Ερμηνεία και Στρατηγική Αξία

Η ανάλυση σεναρίων μετατρέπει την προβλεπτική ανάλυση σε ενεργό εργαλείο υποστήριξης αποφάσεων, επιτρέποντας την αξιολόγηση των συνεπειών διαφορετικών στρατηγικών πριν από την εφαρμογή τους. Με τον τρόπο αυτό, διευκολύνεται η ιεράρχηση ενεργειών διατήρησης και η αποτελεσματικότερη διαχείριση της πελατειακής βάσης. Συνολικά, η ανάλυση what-if μετατρέπει το σύστημα προβλεπτικής ανάλυσης σε ένα ευέλικτο εργαλείο στρατηγικού σχεδιασμού, επιτρέποντας στις επιχειρήσεις να αξιολογούν εκ των προτέρων τις επιπτώσεις εναλλακτικών ενεργειών. Η προσέγγιση αυτή μειώνει την αβεβαιότητα στη λήψη αποφάσεων και ενισχύει τη δυνατότητα προσαρμογής σε μεταβαλλόμενες συνθήκες της αγοράς, καθιστώντας την αναλυτική δεδομένων βασικό πυλώνα βιώσιμης ανάπτυξης στο περιβάλλον του ηλεκτρονικού εμπορίου.

8.5 Σύνθεση Ευρημάτων και Στρατηγικές Προτάσεις

Η παρούσα ενότητα συνθέτει τα αποτελέσματα της προβλεπτικής ανάλυσης που παρουσιάστηκαν στις προηγούμενες υποενότητες, συνδέοντας τον κίνδυνο απώλειας πελατών, την αξία διάρκειας ζωής και τη συμπεριφορική τμηματοποίηση σε ένα ενιαίο πλαίσιο λήψης αποφάσεων. Η σύνθεση αυτή επιτρέπει τη μετάβαση από τη μεμονωμένη ερμηνεία δεικτών σε μια ολοκληρωμένη προσέγγιση διαχείρισης πελατειακών σχέσεων, με έμφαση στην πρόληψη της αποχώρησης και τη βελτιστοποίηση της επιχειρησιακής αξίας.

Τα ευρήματα της ανάλυσης καταδεικνύουν ότι ο κίνδυνος αποχώρησης δεν κατανέμεται ομοιόμορφα στο σύνολο της πελατειακής βάσης, αλλά διαφοροποιείται σημαντικά ανάλογα με τη συμπεριφορά, την εμπειρία και τη συνολική αξία του πελάτη. Ο σύνθετος δείκτης churn score, ως proxy εκτίμησης της πιθανότητας αποχώρησης, ανέδειξε διακριτές ομάδες πελατών με διαφορετικά επίπεδα επικινδυνότητας, επιτρέποντας την ιεράρχηση της προσοχής και των διαθέσιμων πόρων.

Η συνδυαστική απεικόνιση churn score και Customer Lifetime Value ανέδειξε ιδιαίτερα κρίσιμες ομάδες πελατών, όπως εκείνους που εμφανίζουν υψηλή αξία διάρκειας ζωής αλλά αυξημένο κίνδυνο αποχώρησης. Η ομάδα αυτή παρουσιάζει αυξημένη στρατηγική σημασία, καθώς η απώλειά της μπορεί να έχει δυσανάλογα αρνητικό αντίκτυπο στα έσοδα και στη μακροχρόνια βιωσιμότητα της επιχείρησης. Αντίστοιχα, πελάτες χαμηλής αξίας και υψηλού κινδύνου αποχώρησης μπορούν να αντιμετωπιστούν με διαφορετικές, λιγότερο κοστοβόρες στρατηγικές, χωρίς να διακυβεύεται η συνολική απόδοση.

Βάσει των παραπάνω, προκύπτει ότι οι στρατηγικές διατήρησης πελατών θα πρέπει να είναι διαφοροποιημένες και στοχευμένες, λαμβάνοντας υπόψη τόσο τον προβλεπτικό κίνδυνο όσο και την επιχειρησιακή αξία. Για τους πελάτες υψηλής αξίας και χαμηλού κινδύνου, προτείνεται η ενίσχυση της πιστότητας μέσω προσωποποιημένων εμπειριών, προγραμμάτων επιβράβευσης και διατήρησης υψηλής ποιότητας εξυπηρέτησης. Στην περίπτωση πελατών υψηλής αξίας αλλά αυξημένου κινδύνου αποχώρησης, καθίσταται αναγκαία η εφαρμογή στοχευμένων ενεργειών πρόληψης, όπως εξατομικευμένες προσφορές, βελτιώσεις στη διαδικασία εξυπηρέτησης και έγκαιρη επικοινωνία.

Παράλληλα, η ανάλυση what-if κατέδειξε τη σημασία της ευελιξίας στην ερμηνεία των προβλεπτικών δεικτών. Η δυνατότητα προσαρμογής των ορίων κινδύνου και των βαρών των επιμέρους μεταβλητών επιτρέπει στις επιχειρήσεις να προσαρμόζουν τη στρατηγική τους ανάλογα με τις εκάστοτε συνθήκες της αγοράς, τους διαθέσιμους πόρους και τους επιχειρησιακούς στόχους. Με τον τρόπο αυτό, η προβλεπτική ανάλυση μετατρέπεται από στατικό εργαλείο αξιολόγησης σε δυναμικό μηχανισμό υποστήριξης αποφάσεων.

Συνολικά, το Κεφάλαιο 8 αναδεικνύει τη συμβολή της προβλεπτικής ανάλυσης στη γεφύρωση της θεωρίας με την πρακτική εφαρμογή της αναλυτικής δεδομένων στο ηλεκτρονικό εμπόριο. Η αξιοποίηση ερμηνεύσιμων proxy δεικτών και διαδραστικών οπτικοποιήσεων επιτρέπει την κατανόηση σύνθετων φαινομένων, όπως ο κίνδυνος αποχώρησης, και υποστηρίζει τη διαμόρφωση τεκμηριωμένων και εφαρμόσιμων στρατηγικών διαχείρισης πελατών.

ΚΕΦΑΛΑΙΟ 9: ΣΥΜΠΕΡΑΣΜΑΤΑ ΚΑΙ ΜΕΛΛΟΝΤΙΚΕΣ ΕΠΕΚΤΑΣΕΙΣ

9.1 Συνολικά Συμπεράσματα

Η παρούσα διπλωματική εργασία εστίασε στην ανάλυση και πρόβλεψη των καταναλωτικών τάσεων στο ηλεκτρονικό εμπόριο, αξιοποιώντας εργαλεία οπτικής αναλυτικής και προσεγγίσεις επιχειρηματικής ευφυΐας (Business Intelligence). Μέσω της ανάλυσης δεδομένων και της ανάπτυξης διαδραστικών δεικτών και οπτικοποιήσεων στο Tableau, προέκυψαν τα ακόλουθα βασικά συμπεράσματα:

1. **Δημογραφική και Συμπεριφορική Διαστρωμάτωση:** Η ανάλυση ανέδειξε ότι οι ηλικιακές ομάδες 24–39 ετών αποτελούν τον πιο ενεργό πυρήνα καταναλωτών στο ηλεκτρονικό εμπόριο. Επιπλέον, το μορφωτικό επίπεδο και το εισόδημα συσχετίζονται θετικά με τη συχνότητα και την αξία των αγορών. Αντίθετα, η επαγγελματική ιδιότητα, παρότι επηρεάζει τη συνολική δαπάνη, δεν διαφοροποιεί ουσιαστικά τα επίπεδα ικανοποίησης των πελατών.
2. **Στρατηγική Σημασία της Πρόβλεψης Churn:** Η ανάπτυξη ενός σύνθετου δείκτη κινδύνου αποχώρησης (churn score) επέτρεψε τον εντοπισμό κρίσιμων τμημάτων της πελατειακής βάσης. Διαπιστώθηκε ότι η ικανοποίηση των πελατών παρουσιάζει αντίστροφη σχέση με τον κίνδυνο αποχώρησης, επιβεβαιώνοντας ότι η εμπειρία χρήστη αποτελεί καθοριστικό παράγοντα για τη διατήρηση πελατών.
3. **Τμηματοποίηση Βάσει Αξίας και Κινδύνου:** Η συνδυαστική ανάλυση του δείκτη churn με δείκτη Αξίας Διάρκειας Ζωής Πελάτη (CLV proxy) οδήγησε στην αναγνώριση τεσσάρων στρατηγικών κατηγοριών πελατών. Ιδιαίτερη έμφαση δόθηκε στην κατηγορία Critical Retention Segment (Υψηλή Αξία – Υψηλός Κίνδυνος), η οποία απαιτεί στοχευμένες και άμεσες παρεμβάσεις διατήρησης, σε αντίθεση με ομάδες χαμηλότερης στρατηγικής προτεραιότητας.
4. **Αξία της Οπτικής Αναλυτικής και της Επιχειρηματικής Ευφυΐας:** Η αξιοποίηση του Tableau ανέδειξε ότι ακόμη και χωρίς την εφαρμογή πολύπλοκων αλγοριθμικών μοντέλων, η στοχευμένη χρήση proxy δεικτών, διαδραστικών dashboards και αναλύσεων σεναρίων (What-if Analysis) μπορεί να προσφέρει ουσιαστική υποστήριξη στη λήψη στρατηγικών αποφάσεων.

9.2 Περιορισμοί της Έρευνας

Παρά τη συμβολή της παρούσας μελέτης, ορισμένοι περιορισμοί πρέπει να ληφθούν υπόψη. Το σύνολο δεδομένων βασίστηκε σε δευτερογενείς πηγές και περιορισμένο χρονικό εύρος, γεγονός που ενδέχεται να επηρεάζει τη γενικευσιμότητα των αποτελεσμάτων. Επιπλέον, η χρήση proxy δεικτών για το churn και το CLV, αν και πρακτικά χρήσιμη, δεν αντικαθιστά πλήρως την ακρίβεια μοντέλων που βασίζονται σε πραγματικά ιστορικά δεδομένα αποχώρησης και κερδοφορίας.

9.3 Μελλοντικές Επεκτάσεις

Με βάση τα παραπάνω συμπεράσματα και τους αναγνωρισμένους περιορισμούς, προτείνονται οι ακόλουθες κατευθύνσεις για μελλοντική έρευνα:

1. **Ενσωμάτωση Προηγμένων Αλγορίθμων Μηχανικής Μάθησης:** Μελλοντικές μελέτες μπορούν να αξιοποιήσουν γλώσσες προγραμματισμού όπως Python ή R για την εκπαίδευση μοντέλων Random Forest και XGBoost, επιτυγχάνοντας ακριβέστερη πρόβλεψη churn και CLV μέσω μη γραμμικών σχέσεων.
2. **Ανάλυση Χρονοσειρών και Εποχικότητας:** Η επέκταση του χρονικού ορίζοντα των δεδομένων θα επιτρέψει την εφαρμογή μοντέλων ARIMA ή LSTM, βελτιώνοντας την πρόβλεψη ζήτησης και τη διαχείριση αποθεμάτων.
3. **Υβριδικά Συστήματα Σύστασης:** Η ανάπτυξη συστημάτων σύστασης που συνδυάζουν content-based και collaborative filtering τεχνικές θα μπορούσε να ενισχύσει την εξατομίκευση και να αυξήσει τον δείκτη μετατροπής (conversion rate).
4. **Εξηγήσιμη Τεχνητή Νοημοσύνη (Explainable AI):** Η εφαρμογή τεχνικών όπως SHAP και LIME θα ενίσχυε τη διαφάνεια των μοντέλων, περιορίζοντας τον κίνδυνο αλγοριθμικής μεροληψίας.
5. **Εμπλουτισμός Δεδομένων:** Η ενσωμάτωση δεδομένων από μέσα κοινωνικής δικτύωσης (sentiment analysis) και δεδομένων πλοήγησης (clickstream data) θα προσέφερε βαθύτερη κατανόηση της πρόθεσης και της ψυχολογίας του καταναλωτή.

Βιβλιογραφία

Chen, J., Zhang, C., & Xu, Y. (2021). Machine learning applications in e-commerce: A review and case studies. *Journal of Retailing and Consumer Services*, 62, 102637.

Davenport, T. H., & Harris, J. (2017). *Competing on Analytics: The New Science of Winning*. Harvard Business Review Press.

Deloitte. (2022). The future of personalization in retail. Deloitte Insights.

Huang, J., & Wang, X. (2022). Dynamic pricing in e-commerce using reinforcement learning. *Journal of Retailing and Consumer Services*, 68, 103027.

McKinsey & Company. (2021). *Analytics in retail: Driving efficiency and growth*.

Regulation (EU) 2016/679 (General Data Protection Regulation – GDPR).

Ecommerce News. (2024). Ecommerce in Greece. Ανακτήθηκε από <https://ecommercenews.eu/ecommerce-in-europe/ecommerce-in-greece/>

GR.EC.A, Ecommerce Europe, & EuroCommerce. (2021). Ευρωπαϊκή Έκθεση Ηλεκτρονικού Εμπορίου. Αθήνα.

Morgan Stanley. (2022). Global Ecommerce Growth Forecast 2022. Ανακτήθηκε από <https://www.morganstanley.com/ideas/global-ecommerce-growth-forecast-2022>

UNCTAD. (2022). COVID-19 boost to e-commerce sustained into 2021, new UNCTAD figures show. Ανακτήθηκε από <https://unctad.org>

UNCTAD. (2024). Measuring the digital economy.

Ανακτήθηκε από

https://unctad.org/system/files/official-document/dtlecde2024d3_en.pdf

Energizing Greece. (2022). ELTRUN: To 2021 είναι χρονιά ρεκόρ για το ηλεκτρονικό εμπόριο στην Ελλάδα. Ανακτήθηκε από <https://energizinggreece.gr>

Deloitte. (2022). The future of personalization in retail. Deloitte Insights.

McKinsey & Company. (2021). *Analytics in retail: Driving efficiency and growth*.

NielsenIQ. (2023). *Sustainability and consumer trends report*.

Regulation (EU) 2016/679 (General Data Protection Regulation – GDPR).

Gandomi, A., & Haider, M. (2015). Beyond the hype: Big data concepts, methods, and analytics. *International Journal of Information Management*, 35(2), 137–144.

Huang, J., & Wang, X. (2022). Dynamic pricing in e-commerce using reinforcement learning. *Journal of Retailing and Consumer Services*, 68, 103027.

Laney, D. (2001). 3D Data Management: Controlling Data Volume, Velocity, and Variety. META Group Research Note.

Binns, R., Veale, M., Van Kleek, M., & Shadbolt, N. (2021). 'It's reducing a human being to a percentage': Perceptions of justice in algorithmic decisions. *Human-Computer Interaction*, 36(5-6), 477-517.

European Commission. (2018). General Data Protection Regulation (GDPR) – Regulation (EU) 2016/679. Official Journal of the European Union.

European Parliament. (2024). Artificial Intelligence Act. Brussels.

Mehrabi, N., Morstatter, F., Saxena, N., Lerman, K., & Galstyan, A. (2022). A survey on bias and fairness in machine learning. *ACM Computing Surveys*, 55(6), 1-35.

Binns, R., Veale, M., Van Kleek, M., & Shadbolt, N. (2021). Fairness and transparency in algorithmic decision-making: Ensuring GDPR compliance in e-commerce. *ACM Transactions on Internet Technology*, 21(4), 1-25.

Chen, Y., Li, X., & Zhang, T. (2020). Predicting consumer behavior in e-commerce using ensemble learning methods. *Journal of Retail Analytics*, 12(3), 45-62.

Huang, Z., & Wang, H. (2022). Dynamic pricing in e-commerce using reinforcement learning. *Journal of Business Analytics*, 5(2), 78-92.

Kumar, S., & Rajan, P. (2021). Customer segmentation in e-commerce using K-means and DBSCAN. *International Journal of Data Science*, 9(2), 101-115.

Moradi, S., Tavakoli, M., & Ghasemi, H. (2023). Predicting customer lifetime value in e-commerce using ensemble methods. *Electronic Commerce Research and Applications*, 54, 101-119.

Singh, R., & Prasad, K. (2021). Time-series forecasting for e-commerce sales: ARIMA and LSTM approaches. *International Journal of Forecasting*, 37(4), 1560-1575.

Zhang, Y., & Mei, J. (2022). Hybrid recommendation systems for online retail: Improving conversion rates through personalization. *Journal of E-commerce Research*, 14(1), 23-38

Chen, Y., Xu, H., & Whinston, A. (2020). Analyzing consumer behavior with supervised machine learning models in e-commerce. *Journal of Retailing Analytics*, 16(3), 211-227.

Kumar, S., & Rajan, A. (2021). Consumer profiling in e-commerce using unsupervised learning. *International Journal of Data Science*, 6(2), 98-113.

Li, J., & Li, X. (2022). Predicting cart abandonment with artificial neural networks. *Electronic Commerce Research*, 22(4), 1123-1145.

Singh, P., & Prasad, R. (2021). Forecasting e-commerce sales using LSTM models. *Journal of Business Analytics*, 4(2), 134-150.

Kumar, R., & Rajan, D. (2021). Customer segmentation using machine learning: An application in e-commerce analytics. *Journal of Retail Analytics*, 7(2), 45-59.

Li, X., & Li, Y. (2022). Predictive modeling for customer behavior and cart abandonment using deep learning. *International Journal of Data Science and Analytics*, 9(3), 215-228.

Moradi, M., Zhang, T., & Chen, L. (2023). Customer lifetime value prediction using ensemble learning models. *Expert Systems with Applications*, 217, 119525.

Nguyen, H. T., Tran, K., & Pham, D. (2020). RFM model for customer segmentation in e-commerce: A data mining approach. *International Journal of Advanced Computer Science*, 11(4), 73–81.

Zhang, Y., & Mei, Q. (2022). Hybrid clustering approaches for personalized recommendation systems. *Journal of Intelligent Information Systems*, 58(1), 103–122.

Li, J., & Li, H. (2022). *Predicting customer churn and cart abandonment using machine learning classification models in e-commerce. Journal of Retail Analytics*, 18(3), 145–162.

Moradi, P., Ahmed, S., & Torres, L. (2023). *Customer lifetime value prediction using ensemble learning models: A data-driven marketing perspective. International Journal of Business Intelligence Research*, 14(2), 33–51.

Singh, R., & Prasad, D. (2021). *Time-series forecasting of e-commerce sales using ARIMA and LSTM models. Procedia Computer Science*, 192, 1512–1520.

James, G., Witten, D., Hastie, T., & Tibshirani, R. (2021). *An introduction to statistical learning with applications in R* (2nd ed.). Springer.

Kuhn, M., & Johnson, K. (2019). *Feature engineering and selection: A practical approach for predictive models*. CRC Press.

Ribeiro, M. T., Singh, S., & Guestrin, C. (2016). “Why should I trust you?”: Explaining the predictions of any classifier. *Proceedings of the 22nd ACM SIGKDD International Conference on Knowledge Discovery and Data Mining*, 1135–1144.

Zou, H., Hastie, T., & Tibshirani, R. (2020). *Statistical learning with sparsity: The Lasso and generalizations*. CRC Press.

Adomavicius, G., & Tuzhilin, A. (2015). *Context-aware recommender systems*. In F. Ricci, L. Rokach, & B. Shapira (Eds.), *Recommender systems handbook* (pp. 191–226). Springer.

Aggarwal, C. C. (2016). *Recommender systems: The textbook*. Springer.

Burke, R. (2022). *Hybrid recommender systems: Survey and experiments. User Modeling and User-Adapted Interaction*, 32(4), 703–727.

Jannach, D., Jugovac, M., & Lerche, L. (2021). *Next-generation recommender systems: Research directions and challenges. AI Magazine*, 42(3), 7–18.

Koren, Y., Bell, R., & Volinsky, C. (2009). *Matrix factorization techniques for recommender systems. Computer*, 42(8), 30–37.

Zhang, S., Yao, L., Sun, A., & Tay, Y. (2019). *Deep learning based recommender system: A survey and new perspectives. ACM Computing Surveys*, 52(1), 1–38.

Zhang, X., & Mei, H. (2022). *Personalized recommender systems in e-commerce: Advances and challenges. Information Systems Frontiers*, 24(2), 345–362.

<https://www.kaggle.com/datasets/salahuddinahmedshuvo/ecommerce-consumer-behavior-analysis-data>