



ΠΑΝΕΠΙΣΤΗΜΙΟ ΠΕΙΡΑΙΩΣ – ΤΜΗΜΑ ΠΛΗΡΟΦΟΡΙΚΗΣ

**Πρόγραμμα Μεταπτυχιακών Σπουδών
«Κυβερνοασφάλεια & Επιστήμη Δεδομένων»**

Μεταπτυχιακή Διατριβή

Τίτλος Διατριβής	Πρόβλεψη Πωλήσεων σε Αλυσίδα Καταστημάτων Τυχερών Παιχνιδιών με την χρήση προηγμένων μεθόδων Αναλυτικής & Οπτικοποίησης Revenue Forecasting using Advanced Analytical & Visualization Methods in Retail Betting Company
Όνοματεπώνυμο Φοιτητή	Κυριακοπούλου Μαρία
Πατρώνυμο	Γεώργιος
Αριθμός Μητρώου	ΜΠΚΕΔ21025
Επιβλέπων	Δημήτριος Αποστόλου, Καθηγητής

Ημερομηνία Παράδοσης **Μάιος 2024**

Στην ολοκλήρωση της παρούσας μεταπτυχιακής διατριβής, ιδιαίτερα σημαντική ήταν η συμβολή του Διδάσκοντα του ΠΜΣ κ. **Ανδρέα Ζάρα**, που προσέφερε επιστημονική και συμβουλευτική καθοδήγηση σε όλα τα στάδια εκπόνησής της.

Τριμελής Εξεταστική Επιτροπή

Δημήτριος Αποστόλου
Καθηγητής

Άγγελος Πικράκης
Επίκουρος Καθηγητής

Δρ. Γρηγόριος
Κωρονάκος
Διδάσκων ΠΜΣ

Ευχαριστίες

Θα ήθελα να εκφράσω την ειλικρινή μου εκτίμηση στον κ. Ανδρέα Ζάρα για την καθοδήγηση, τα σχόλια και την εξειδίκευσή του σε όλη τη διαδικασία έρευνας και συγγραφής. Η αφοσίωσή του στην προώθηση της κριτικής σκέψης, της επεξηγηματικής προσέγγισης και της ακαδημαϊκής αριστείας έπαιξε καθοριστικό ρόλο στη διαμόρφωση της ποιότητας αυτής της εργασίας.

Επίσης, θέλω να εκφράσω τις ευχαριστίες μου στην οικογένεια μου, στους συναδέλφους και τους φίλους μου για την ενθάρρυνση, την αμέριστη υποστήριξη και την κατανόησή σε αυτή την ακαδημαϊκή διαδρομή.

Περίληψη

Η παρούσα διπλωματική διατριβή διερευνά τη χρήση προηγμένων μεθόδων ανάλυσης, οπτικοποίησης και πρόβλεψης για τη βελτίωση της πρόβλεψης εσόδων σε αλυσίδα καταστημάτων τυχερών παιχνιδιών. Η μελέτη στοχεύει στην ανάπτυξη ενός αξιόπιστου και ακριβούς μοντέλου πρόβλεψης εσόδων για την υποστήριξη της λήψης αποφάσεων και του στρατηγικού σχεδιασμού στην εταιρεία. Η μεθοδολογία έρευνας περιλαμβάνει τη συλλογή ιστορικών δεδομένων πωλήσεων από τα καταστήματα και την εφαρμογή διαφόρων στατιστικών τεχνικών, συμπεριλαμβανομένων της ανάλυσης χρονοσειρών, της ανάλυσης παλινδρόμησης κ.α. Πέρα από τα δεδομένα πωλήσεων των καταστημάτων, λαμβάνονται υπόψιν διαφορετικά σύνολα δεδομένων όπως δημογραφικά στοιχεία ή σημεία ενδιαφέροντος τα οποία στηρίζουν την παρούσα ανάλυση και επαληθεύουν τα παραγόμενα αποτελέσματα. Η μελέτη εξετάζει επίσης την επίδραση εξωτερικών παραγόντων, όπως ο ανταγωνισμός, οι οικονομικές τάσεις και οι εκστρατείες μάρκετινγκ, στην ακρίβεια της πρόβλεψης εσόδων. Τα αποτελέσματα δείχνουν ότι το προτεινόμενο μοντέλο πρόβλεψης ξεπερνά τις παραδοσιακές μεθόδους όσον αφορά την ακρίβεια και την αξιοπιστία. Η μελέτη καταλήγει στο συμπέρασμα ότι η εφαρμογή προηγμένων μεθόδων ανάλυσης και πρόβλεψης μπορεί να βελτιώσει σημαντικά την πρόβλεψη εσόδων στα καταστήματα τυχερών παιχνιδιών και να βοηθήσει τις εταιρείες να λαμβάνουν τεκμηριωμένες αποφάσεις σε μια ανταγωνιστική αγορά.

Abstract

This thesis explores the use of advanced analysis, visualization and forecasting methods to improve revenue forecasting in a betting company. The study aims to develop a reliable and accurate revenue forecasting model to support decision making and strategic planning in the company. The research methodology includes the collection of historical sales data from the stores and the application of various statistical techniques, including time series analysis, regression analysis, etc. In addition to store sales data, different data sets such as demographics or points of interest are considered to support this analysis and verify the generated results. The study also examines the impact of external factors such as competition, economic trends and marketing campaigns on revenue forecast accuracy. The results show that the proposed prediction model outperforms traditional methods in terms of accuracy and reliability. The study concludes that the application of advanced analysis and forecasting methods can significantly improve revenue forecasting in gaming stores and help companies make informed decisions in a competitive market.

Key words: Big Data, Analytics, Demand Driven Forecasting, Consensus Forecasting, Sales & Operational planning, Budgeting, Promotional modelling, Price optimization, Quantitative Forecasting, Qualitative Forecasting, Data Visualization

Πίνακας περιεχομένων

Εισαγωγή.....	6
1. Big Data & Analytics	7
1.2 Η χρησιμότητα των Big Data & των Analytics στην έννοια της Ζήτησης	8
2. Πρόβλεψη πωλήσεων.....	10
2.1 Επιχειρησιακά Οφέλη Πρόβλεψης Πωλήσεων	10
3.2 Οι χρήσεις της Πρόβλεψης στα μοντέλα του promotional modelling & του price optimization	12
3. Χρονοσειρές.....	13
4.1 Εισαγωγή στην Ανάλυση Χρονοσειρών	13
4.2 Ιστορική αναδρομή Ανάλυσης Χρονοσειρών	14
4.3 Βασικές έννοιες στην ανάλυση Χρονοσειρών	15
4.4 Δοκιμές στην Ανάλυση Χρονοσειρών.....	16
4. Μέθοδοι Πρόβλεψης.....	17
5.1 Ποσοτική Πρόβλεψη.....	17
5.2 Ποσοτικές Μέθοδοι Πρόβλεψης	18
5.3 Ποιοτική Πρόβλεψη	25
5.4 Ποιοτικές Μέθοδοι Πρόβλεψης.....	26
5. Η εταιρεία ΟΠΑΠ	26
6.1 Γνωριμία με τον όμιλο ΟΠΑΠ	27
6.2 Εταιρική Υπευθυνότητα & Κοινωνική Ευθύνη	27
6.3 Δομή Ομίλου ΟΠΑΠ	28
6. Εφαρμογή Διαδικασίας Πρόβλεψης στον Όμιλο ΟΠΑΠ	28
6.1 Συλλογή Δεδομένων.....	29
6.2 Οπτικοποίηση Δεδομένων	30
7.2.1 Γεωχωρική Οπτικοποίηση Δεδομένων.....	30
7.2.2 Εξερεύνηση των Δεδομένων μέσω Οπτικοποίησης	33
6.3 Εφαρμογή Πρόβλεψης	37
7. Συμπεράσματα	64
8. Βιβλιογραφία.....	66

Εισαγωγή

Η πρόβλεψη εσόδων είναι μια κρίσιμη διαδικασία για τις εταιρείες που δραστηριοποιούνται σε ιδιαίτερα ανταγωνιστικές και δυναμικές αγορές. Η ακριβής πρόβλεψη εσόδων επιτρέπει στις εταιρείες να λαμβάνουν τεκμηριωμένες αποφάσεις και να εφαρμόζουν αποτελεσματικές στρατηγικές που μπορούν να μεγιστοποιήσουν την κερδοφορία και την ανάπτυξη. Η βιομηχανία τυχερών παιχνιδιών είναι μια άκρως ανταγωνιστική αγορά που έχει γνωρίσει τεράστια ανάπτυξη τα τελευταία χρόνια.

Οι εταιρείες τυχερών παιχνιδιών είναι επιχειρήσεις που δραστηριοποιούνται στον κλάδο των τυχερών παιχνιδιών, ο οποίος περιλαμβάνει την προσφορά διαφόρων μορφών παιχνιδιών και ευκαιριών στοιχηματισμού σε πελάτες με αντάλλαγμα χρήματα. Η βιομηχανία τυχερών παιχνιδιών είναι μια βιομηχανία πολλών δισεκατομμυρίων δολαρίων, με εταιρείες που δραστηριοποιούνται σε διαφορετικούς τομείς όπως τα παιχνίδια καζίνο, τα αθλητικά στοιχήματα, τα διαδικτυακά παιχνίδια, τα λαχεία κ.α. Οι εταιρείες τυχερών παιχνιδιών μπορούν να χωριστούν σε δύο κύριες κατηγορίες: επίγειες και διαδικτυακές. Οι επίγειες εταιρείες διαχειρίζονται φυσικά καζίνο, καταστήματα στοιχημάτων και άλλους χώρους τυχερών παιχνιδιών, ενώ οι διαδικτυακές εταιρείες παρέχουν υπηρεσίες τυχερών παιχνιδιών στο Διαδίκτυο μέσω ιστοτόπων και εφαρμογών για smart phones. Στη παρούσα διπλωματική διατριβή θα εστιάσουμε στο επίγειο παίγνιο και πιο συγκεκριμένα στη πρόβλεψη πωλήσεων που μπορούν να πετύχουν τα επίγεια καταστήματα τυχερών παιχνιδιών.

Ο κλάδος των τυχερών παιχνιδιών ελέγχεται κυβερνητικά στις περισσότερες χώρες για να διασφαλίζονται δίκαιες και υπεύθυνες πρακτικές τζόγου. Οι ρυθμιστικοί φορείς θέτουν πρότυπα και κατευθυντήριες γραμμές που πρέπει να ακολουθούν οι εταιρείες τυχερών παιχνιδιών για την προστασία των καταναλωτών και την πρόληψη των προβληματικών τυχερών παιχνιδιών. Επιπλέον, οι εταιρείες τυχερών παιχνιδιών υπόκεινται σε φόρους και τέλη αδειοδότησης, τα οποία ποικίλλουν ανάλογα με τη χώρα και το είδος της δραστηριότητας του στοιχηματισμού. Οι εταιρείες τυχερών παιχνιδιών αντιμετωπίζουν πολλές προκλήσεις, συμπεριλαμβανομένου του αυξημένου ανταγωνισμού, της αλλαγής των προτιμήσεων των καταναλωτών και των ρυθμιστικών περιορισμών. Ως αποτέλεσμα, οι εταιρείες στον κλάδο των τυχερών παιχνιδιών πρέπει να είναι καινοτόμες και προσαρμόσιμες για να επιβιώσουν και να ευδοκιμήσουν σε αυτήν την εξαιρετικά ανταγωνιστική αγορά.

Με το μέγεθος της παγκόσμιας αγοράς να φτάνει τα 180 δισεκατομμύρια δολάρια το 2022 και να αναμένεται να συνεχίσει να αυξάνεται, οι εταιρείες στη βιομηχανία τυχερών παιχνιδιών πρέπει να διαθέτουν αξιόπιστα και ακριβή μοντέλα πρόβλεψης εσόδων για να υποστηρίξουν τη λήψη αποφάσεων και τον στρατηγικό σχεδιασμό. Αυτή η διατριβή στοχεύει να διερευνήσει τη χρήση προηγμένων αναλυτικών και προγνωστικών μεθόδων για τη βελτίωση της πρόβλεψης εσόδων σε μια αλυσίδα καταστημάτων τυχερών παιχνιδιών. Η βιομηχανία τυχερών παιχνιδιών είναι μοναδική λόγω των ταχέως μεταβαλλόμενων προτιμήσεων των καταναλωτών και του συνεχώς εξελισσόμενου τεχνολογικού τοπίου. Ως εκ τούτου, οι εταιρείες σε αυτόν τον κλάδο πρέπει να είναι σε θέση να προσαρμοστούν γρήγορα στις αλλαγές της αγοράς και να λαμβάνουν τεκμηριωμένες αποφάσεις για να παραμείνουν μπροστά από τον ανταγωνισμό.

Στο πρώτο κεφάλαιο της παρούσας διπλωματικής διατριβής εξετάζεται η έννοια και τα χαρακτηριστικά των μεγάλων δεδομένων παρουσιάζοντας τις ιδιότητες και την χρησιμότητα τους στη πρόβλεψη πωλήσεων. Με λίγα λόγια ο όρος Μεγάλα Δεδομένα ή αλλιώς «Big data» χρησιμοποιείται για να περιγράψει σύνολα δεδομένων τόσο μεγάλα ή σύνθετα που ξεφεύγουν από τις δυνατότητες καταγραφής, αποθήκευσης και ανάλυσης των παραδοσιακών τεχνικών επεξεργασίας δεδομένων.

Στο δεύτερο κεφάλαιο μελετάται το πως τα Μεγάλα Δεδομένα αξιοποιούνται μέσω των analytics για την απόκτηση πληροφοριών και τη λήψη καλύτερων επιχειρηματικών αποφάσεων. Ουσιαστικά, ο όρος analytics σημαίνει οποιαδήποτε απόφαση λαμβάνεται με βάση τα δεδομένα.

Τα analytics μπορούν να εφαρμοστούν στο επιχειρηματικό κόσμο και να φανούν ωφέλιμα σε διάφορα τμήματα του οργανισμού όπως το marketing, το finance, το HR, οι πωλήσεις και τα operations. Τα Μεγάλα Δεδομένα είναι επίσης απαραίτητα σε πολλές μεθόδους πρόβλεψης πωλήσεων όπως το Demand Driven Forecasting και το Consensus forecasting.

Το επόμενο κεφάλαιο εστιάζει στη πρόβλεψη πωλήσεων. Η πρόβλεψη είναι το βασικότερο στοιχείο που βοηθάει στον αποτελεσματικό σχεδιασμό και στο σωστό προγραμματισμό με αποτέλεσμα την έγκαιρη λήψη αποφάσεων. Η πρόβλεψη των πωλήσεων έχει αποδειχθεί πως είναι ιδιαίτερα σημαντική στη διαχείριση αποθεμάτων, στο επίπεδο εξυπηρέτησης πελατών, στο operational planning, στο sales targeting και στο budgeting. Πέρα από τα παραπάνω η πρόβλεψη των πωλήσεων είναι στενά συνδεδεμένη με τις τεχνικές του promotional modelling και του price optimization τα οποία αναλύονται σε αυτή την μελέτη.

Στη συνέχεια της παρούσας διατριβής παρουσιάζονται οι μέθοδοι που θα χρησιμοποιηθούν για τη δημιουργία του κατάλληλου μοντέλου πρόβλεψης στην αλυσίδα καταστημάτων τυχερών παιχνιδιών. Η ποσοτική πρόβλεψη (quantitative forecasting) η οποία θα χρησιμοποιηθεί στη μελέτη περιλαμβάνει τη χρήση μαθηματικών και στατιστικών μοντέλων για την πραγματοποίηση προβλέψεων με βάση ιστορικά δεδομένα. Αυτά τα μοντέλα χρησιμοποιούν διάφορες τεχνικές, όπως ανάλυση παλινδρόμησης, ανάλυση χρονοσειρών, εκθετική εξομάλυνση κ.α. Η ποιοτική πρόβλεψη (qualitative forecasting) είναι ένας τύπος μεθόδου πρόβλεψης που βασίζεται σε γνώμες ειδικών, προσωπική κρίση και υποκειμενικές ερμηνείες δεδομένων για την πρόβλεψη μελλοντικών τάσεων και αποτελεσμάτων.

Το επόμενο κεφάλαιο είναι αφιερωμένο στη γνωριμία με τον όμιλο ΟΠΑΠ. Πιο συγκεκριμένα αναφέρεται το ποια είναι η εταιρεία, πότε δημιουργήθηκε, σε ποιες αγορές δραστηριοποιείται, ποια προϊόντα παρέχει στους πελάτες του, ποια είναι η δομή της εταιρείας κ.α. Επίσης, εξετάζεται ο τρόπος με το οποίο ο όμιλος πραγματοποιεί πρόβλεψη πωλήσεων και πώς θα μπορούσε να βελτιώσει τον τρόπο προβλέψεων.

Στη συνέχεια, προχωράμε με τη διαδικασία γνωριμίας και εξερεύνησης των δεδομένων μέσα από μια σειρά οπτικοποιήσεων τόσο με το γεωχωρικό λογισμικό ArcGIS όσο και με το λογισμικό PowerBI. Τέλος, αφού έχουμε κατανοήσει τα δεδομένα και τα διάφορα χαρακτηριστικά και μοτίβα που παρουσιάζουν προχωράμε με τη διαδικασία πρόβλεψης με τη χρήση της Python.

1. Big Data & Analytics

1.1 Αξιοποίηση Big Data μέσω Analytics

Σήμερα, η πληροφορία είναι ζωτικής σημασίας, επηρεάζοντας τη σκέψη, την εξέλιξη και την πρόοδο. Με την ανάπτυξη της τεχνολογίας και του διαδικτύου, παράγουμε δεδομένα με πρωτοφανείς ρυθμούς, δημιουργώντας μια κοινωνία βασισμένη σε αυτά. Τα Μεγάλα Δεδομένα ή αλλιώς "Big Data", είναι ένας συνδυασμός μη δομημένων, ημιδομημένων ή δομημένων δεδομένων που δεν μπορούν να επεξεργαστούν με παραδοσιακές μεθόδους λόγω του μεγάλου όγκου τους. Τα χαρακτηριστικά των Big Data περιλαμβάνουν τον όγκο, που αναφέρεται στην ποσότητα των δεδομένων, την ταχύτητα, που αφορά την παραγωγή, την μετακίνηση και την αποθήκευση των δεδομένων σε πραγματικό χρόνο για τη λήψη αποφάσεων, την ποικιλία, που αναφέρεται στην ποικιλομορφία των τύπων δεδομένων που συλλέγονται από διαφορετικές πηγές, την εγκυρότητα και την ακρίβεια, που είναι κρίσιμες για την εξαγωγή αξιόπιστων συμπερασμάτων από τα δεδομένα και την αξία, που εξαρτάται από το πώς χρησιμοποιούνται τα δεδομένα για να προσφέρουν χρήσιμες πληροφορίες στους οργανισμούς.

Οι επιχειρήσεις από όλους τους τομείς ενδιαφέρονται ολοένα και περισσότερο για τη διαχείριση και ανάλυση των Big Data, καθώς πρέπει να αντιδρούν γρήγορα στις αλλαγές και στις μελλοντικές προκλήσεις και να παίρνουν στρατηγικές αποφάσεις. Τα Μεγάλα Δεδομένα είναι ένα πολύ σημαντικό συστατικό για τους υπεύθυνους λήψης αποφάσεων, καθώς τους βοηθούν να επιτυγχάνουν τα καλύτερα δυνατά αποτελέσματα και να παραμένουν ανταγωνιστικοί στις απαιτήσεις της αγοράς. Όπως είπε ο Edwards Deming και ο Peter Drucker, "Δεν μπορείς να διαχειριστείς αυτό που δεν μετράς". Για να εξαχθεί λοιπόν γνώση και αξία από τα δεδομένα, είναι

απαραίτητη η ανάλυση τους, γνωστή ως analytics. Πιο συγκεκριμένα, η ανάλυση δεδομένων περιλαμβάνει τη λήψη αποφάσεων βάσει δεδομένων και διακρίνεται σε τέσσερις κατηγορίες: Descriptive Analytics (πληροφορίες για το παρελθόν), Diagnostic Analytics (κατανόηση του τι συμβαίνει και γιατί), Predictive Analytics (πρόβλεψη του μέλλοντος) και Prescriptive Analytics (συμβουλές για πιθανά αποτελέσματα). Ως αποτέλεσμα, η ανάλυση δεδομένων βοηθά στη λήψη καλύτερων αποφάσεων, στην εύρεση νέων επιχειρηματικών ευκαιριών, στον ανταγωνισμό με άλλες επιχειρήσεις, στη βελτίωση της απόδοσης και της αποτελεσματικότητας και στη μείωση του κόστους.

Για να γίνει λίγο πιο κατανοητό αυτό που αναφέρεται παραπάνω, μπορούμε να σκεφτούμε για παράδειγμα πως η πρόβλεψη πωλήσεων βελτιώνει την απόδοση και την αποτελεσματικότητα ολόκληρων επιχειρησιακών τμημάτων, όπως είναι τα Operations. Ένας επιχειρησιακός διευθυντής που εστιάζει σε μια γραμμή παραγωγής θέλει να διασφαλίσει την έγκαιρη παραγωγή προϊόντων, ελαχιστοποιώντας τα ελαττώματα και διατηρώντας επαρκές απόθεμα. Θέλει να γνωρίζει πόσες μονάδες υποβάλλονται σε επεξεργασία, τον χρόνο για τη διαδικασία, τον ρυθμό απόδοσης και τον αριθμό μονάδων που αποτυγχάνουν στη δοκιμή ποιότητας. Η πρόβλεψη πωλήσεων αξιοποιεί δεδομένα για να βοηθήσει τις επιχειρήσεις να προβλέψουν μελλοντικές πωλήσεις και να εκτιμήσουν τον αριθμό των πωλήσεων σε συγκεκριμένο χρονικό διάστημα.

1.2 Η χρησιμότητα των Big Data & των Analytics στην έννοια της Ζήτησης

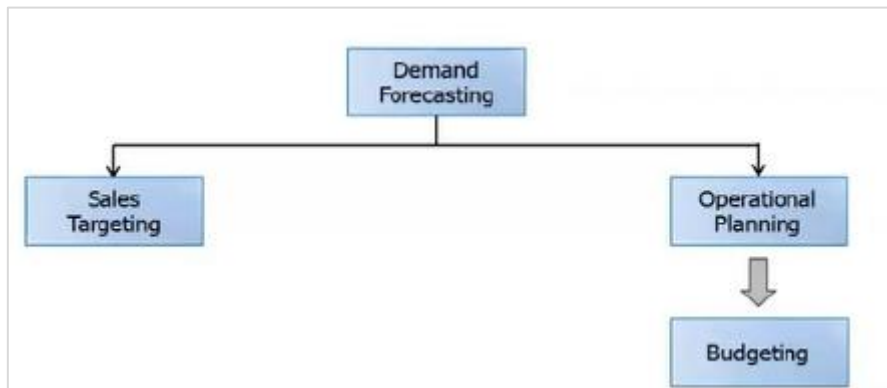
Στη παράγραφο αυτή θα εστιάσουμε στο κομμάτι του πως τα Big Data και κατ' επέκταση τα Analytics συμβάλουν καθοριστικά στην πρόβλεψη των πωλήσεων & της ζήτησης.

Το πρώτο θέμα το οποίο θα εξετάσουμε είναι το Demand Driven Forecasting. Το Demand Driven Forecasting είναι μια σημαντική προσέγγιση για την πρόβλεψη πωλήσεων, επειδή εστιάζει στην κατανόηση της ζήτησης των πελατών και στην ευθυγράμμιση της παραγωγής και της διαχείρισης αποθεμάτων με τις πραγματικές ανάγκες των πελατών. Αυτό έρχεται σε αντίθεση με τις παραδοσιακές μεθόδους πρόβλεψης που βασίζονται κυρίως σε ιστορικά δεδομένα πωλήσεων για την πρόβλεψη της μελλοντικής ζήτησης (Kilger & Wagner, 2008).

Το Demand Driven Forecasting χαρακτηρίζεται από ένα σύνολο σταδίων και βημάτων. Πιο αναλυτικά, το πρώτο στάδιο είναι η συλλογή των δεδομένων. Σε αυτό το πρώτο στάδιο, οι επιχειρήσεις συλλέγουν δεδομένα σχετικά με τις παραγγελίες των πελατών, τα επίπεδα αποθέματος, την παραγωγή, τις πωλήσεις και άλλους σχετικούς παράγοντες που επηρεάζουν τη ζήτηση. Τα δεδομένα αυτά τις περισσότερες φορές χαρακτηρίζονται από όλες τις ιδιότητες που παρουσιάζουν τα Big Data και κατά συνέπεια, οι επιχειρήσεις πρέπει να τα διαχειρίζονται και να τα επεξεργάζονται αναλόγως. Είναι πολύ σημαντικό να μην ξεχνάμε την σπουδαιότητα της ακρίβειας και της εγκυρότητας των δεδομένων μας όσο δύσκολο και αν είναι να επιτευχθεί κάτι τέτοιο σε τόσο μεγάλο όγκο. Ωστόσο ας μην ξεχνάμε και το γεγονός ότι μια ανάλυση η οποία βασίζεται σε ένα μεγάλο όγκο από ακριβή δεδομένα είναι πολύ πιο αντιπροσωπευτική από μια ανάλυση με μικρότερο όγκο δεδομένων.

Το δεύτερο στάδιο είναι το Demand Forecasting. Το Demand Forecasting είναι μια διαδικασία εκτίμησης της μελλοντικής ζήτησης για ένα προϊόν ή μια υπηρεσία. Περιλαμβάνει την ανάλυση ιστορικών δεδομένων πωλήσεων, τάσεων της αγοράς, συμπεριφοράς καταναλωτών και άλλων σχετικών παραγόντων για την πρόβλεψη της μελλοντικής ζήτησης για ένα προϊόν ή μια υπηρεσία. Ουσιαστικά, ζήτηση είναι η ακριβής ποσότητα ενός προϊόντος που επιθυμούν οι πελάτες, ακριβώς την στιγμή που το θέλουν. Ο πρωταρχικός στόχος της πρόβλεψης ζήτησης είναι να βοηθήσει τις επιχειρήσεις να λαμβάνουν τεκμηριωμένες αποφάσεις σχετικά με την παραγωγή, τη διαχείριση αποθεμάτων και την κατανομή των πόρων. Το Demand Forecasting μπορεί να χωριστεί σε δύο κατηγορίες: την ποιοτική (qualitative) και την ποσοτική (quantitative). Οι τεχνικές ποιοτικής πρόβλεψης βασίζονται σε γνώμες εμπειρογνομώνων και στην έρευνα αγοράς για την πρόβλεψη της μελλοντικής ζήτησης. Οι τεχνικές ποσοτικής πρόβλεψης, από την άλλη πλευρά, χρησιμοποιούν στατιστικά μοντέλα και μαθηματικούς τύπους για την πρόβλεψη της μελλοντικής ζήτησης με βάση ιστορικά δεδομένα. Μερικές από τις κοινώς χρησιμοποιούμενες μεθόδους ποσοτικής πρόβλεψης περιλαμβάνουν την ανάλυση χρονοσειρών, την ανάλυση παλινδρόμησης και την εκθετική εξομάλυνση. Αυτές οι μέθοδοι χρησιμοποιούν δεδομένα προηγούμενων

πωλήσεων για να προσδιορίσουν πρότυπα και τάσεις στη ζήτηση και χρησιμοποιούν αυτές τις πληροφορίες για να προβλέψουν τη μελλοντική ζήτηση. Η πρόβλεψη της ζήτησης θα πρέπει να χρησιμοποιείται πάντα ως εισροή στη στόχευση πωλήσεων, στον επιχειρησιακό σχεδιασμό και στον προϋπολογισμό.



Εικόνα 1 "Η Δομή του Demand Forecasting"

Στη συνέχεια έρχεται το Inventory Analysis ή αλλιώς Ανάλυση αποθέματος. Στο τρίτο στάδιο λοιπόν, οι επιχειρήσεις αναλύουν τα επίπεδα αποθέματος και αναδιατάσσουν τα σημεία για να διασφαλίσουν ότι η σωστή ποσότητα αποθέματος είναι διαθέσιμη για να καλύψει τη ζήτηση των πελατών. Αυτό μπορεί να περιλαμβάνει τη χρήση τεχνικών διαχείρισης αποθεμάτων just-in-time (JIT) για την ελαχιστοποίηση του πλεονάζοντος αποθέματος και τη μείωση του κόστους μεταφοράς.

Τελευταίο έρχεται το Sales & Operations Planning στο οποίο οι επιχειρήσεις ενσωματώνουν την πρόβλεψη ζήτησης με τον προγραμματισμό παραγωγής και αποθεμάτων για να διασφαλίσουν ότι οι πόροι κατανέμονται αποτελεσματικά και ότι η παραγωγή ευθυγραμμίζεται με την πραγματική ζήτηση των πελατών. Αυτό μπορεί να περιλαμβάνει τη χρήση εργαλείων συνεργατικού σχεδιασμού και πρόβλεψης που θα βοηθήσουν στην ευθυγράμμιση του οργανισμού γύρω από ένα κοινό σύνολο υποθέσεων και στόχων.

Το επόμενο στο οποίο αξίζει να αναφερθούμε είναι το Consensus Forecasting. Το Consensus Forecasting είναι μια διαδικασία ανάπτυξης μιας ενιαίας πρόβλεψης που βασίζεται σε εισροές και συνεργασία από πολλούς ενδιαφερόμενους φορείς σε έναν οργανισμό. Περιλαμβάνει τον συνδυασμό των προβλέψεων που δημιουργούνται από διαφορετικές επιχειρηματικές μονάδες, τμήματα ή άτομα για να καταλήξουμε σε μια ενιαία πρόβλεψη που αντιπροσωπεύει τη συναινετική άποψη του οργανισμού. Το consensus forecasting είναι μια σημαντική προσέγγιση επειδή μπορεί να βοηθήσει τους οργανισμούς να δημιουργήσουν πιο ακριβείς προβλέψεις αξιοποιώντας τη γνώση και την τεχνογνωσία διαφορετικών ενδιαφερομένων. Με τη συμμετοχή ενός διαφορετικού συνόλου προοπτικών, οι οργανισμοί μπορούν να αποκτήσουν μια πιο ολοκληρωμένη κατανόηση των παραγόντων που επηρεάζουν τη ζήτηση, την προσφορά και άλλες βασικές μεταβλητές που επηρεάζουν τις πωλήσεις και την παραγωγή. Αυτό μπορεί να οδηγήσει σε καλύτερη λήψη αποφάσεων, πιο αποτελεσματική κατανομή πόρων και βελτιωμένη λειτουργική αποτελεσματικότητα.

Όλα τα στάδια του Demand Driven Forecasting που αναφέρθηκαν παραπάνω συμβάλουν σημαντικά στη δημιουργία του Consensus Forecasting. Για να γίνω πιο σαφής, το στάδιο συλλογής δεδομένων είναι κρίσιμο γιατί παρέχει τα ακατέργαστα δεδομένα που απαιτούνται για τη δημιουργία ακριβών προβλέψεων ζήτησης. Επιπλέον, το στάδιο της πρόβλεψης (Demand Forecasting) της ζήτησης είναι σημαντικό επειδή βοηθά στον εντοπισμό τάσεων και προτύπων στα δεδομένα που μπορούν να χρησιμοποιηθούν για τη δημιουργία ακριβέστερων προβλέψεων. Το στάδιο της ανάλυσης αποθέματος είναι εξίσου ζωτικό για την δημιουργία του Consensus Forecasting γιατί διασφαλίζει ότι η σωστή ποσότητα αποθέματος είναι διαθέσιμη για να καλύψει τη ζήτηση των πελατών, χωρίς να φέρει υπερβολικό απόθεμα. Τέλος, το στάδιο του Sales & Operations planning είναι σημαντικό γιατί βοηθά στην ευθυγράμμιση της παραγωγής και της διαχείρισης αποθεμάτων με την πραγματική ζήτηση των πελατών και διασφαλίζει ότι οι πόροι κατανέμονται αποτελεσματικά.

Συμπερασματικά, το Demand Driven Forecasting είναι μια σημαντική προσέγγιση για την πρόβλεψη πωλήσεων, επειδή βοηθά τις επιχειρήσεις να κατανοήσουν καλύτερα τη ζήτηση των πελατών και να ευθυγραμμίσουν τη διαχείριση της παραγωγής και του αποθέματος με τις πραγματικές ανάγκες των πελατών. Τα στάδια του Demand Driven Forecasting παίζουν καθοριστικό ρόλο στη δημιουργία ακριβών προβλέψεων ζήτησης και στη διασφάλιση ότι οι πόροι κατανέμονται αποτελεσματικά για την κάλυψη των αναγκών των πελατών.

2. Πρόβλεψη πωλήσεων

2.1 Επιχειρησιακά Οφέλη Πρόβλεψης Πωλήσεων

Η πρόβλεψη πωλήσεων αποτελεί βασική διαδικασία στον επιχειρησιακό και στρατηγικό σχεδιασμό των επιχειρήσεων, προσφέροντας σημαντικά οφέλη σε διάφορους τομείς, όπως η διαχείριση αποθέματος, η βελτιστοποίηση του επιπέδου εξυπηρέτησης πελατών, στο operational planning, στο sales targeting καθώς επίσης και στο Budgeting. Στο δυναμικό τοπίο των σύγχρονων επιχειρήσεων, η ικανότητα πρόβλεψης μελλοντικών πωλήσεων με έναν βαθμό ακρίβειας είναι ανεκτίμητη. Η πρόβλεψη πωλήσεων αξιοποιεί ιστορικά δεδομένα και αναλυτικά μοντέλα για την προβολή μελλοντικών πωλήσεων, παρέχοντας ένα θεμελιώδες στοιχείο για τη λήψη στρατηγικών αποφάσεων σε πολλές πτυχές των επιχειρηματικών λειτουργιών.

Η αποτελεσματική διαχείριση αποθεμάτων εξαρτάται από τη λεπτή ισορροπία μεταξύ της ικανοποίησης της ζήτησης των πελατών και της ελαχιστοποίησης του κόστους διακράτησης. Η πρόβλεψη πωλήσεων διαδραματίζει κρίσιμο ρόλο στην επίτευξη αυτής της ισορροπίας, επιτρέποντας στις επιχειρήσεις να βελτιστοποιήσουν τα επίπεδα αποθεμάτων. Προβλέποντας τις διακυμάνσεις της ζήτησης, οι εταιρείες μπορούν να προσαρμόσουν ανάλογα τα χρονοδιαγράμματα προμήθειας και παραγωγής αποθεμάτων, μειώνοντας τον κίνδυνο εξάντλησης αποθεμάτων ή πλεονάζοντος αποθέματος. Αυτή η ευθυγράμμιση όχι μόνο περιορίζει το περιττό κόστος αποθήκευσης και απαξίωσης, αλλά διασφαλίζει επίσης τη διαθεσιμότητα των προϊόντων, συμβάλλοντας στην ικανοποίηση και την αφοσίωση των πελατών. Όπως αναφέρει και ο Michael Gilliland (*Fundamental Issues in Business Forecasting, 2003*) το πιο θεμελιώδες ερώτημα που πρέπει να κάνουμε πριν ξεκινήσουμε τη διαδικασία πρόβλεψης είναι το «*Τι προσπαθούμε να προβλέψουμε;*». Η Η συνήθης απάντηση λοιπόν, είναι ότι προσπαθούμε να προβλέψουμε τη «ζήτηση» των πελατών, με τη ζήτηση να ορίζεται ως «*τι θέλουν οι πελάτες και πότε το θέλουν*». Συνεπώς, η κατανόηση της ζήτησης περιλαμβάνει την αναγνώριση των προτύπων και των παραγόντων που επηρεάζουν την ποσότητα των προϊόντων ή των υπηρεσιών που απαιτούνται από την αγορά. Είναι ζωτικής σημασίας για την ακριβή πρόβλεψη, καθώς επιτρέπει στις επιχειρήσεις να προβλέπουν αλλαγές στη συμπεριφορά των καταναλωτών και να προσαρμόζουν ανάλογα τις στρατηγικές τους. Ένα επιπλέον σημαντικό στοιχείο που αναφέρει ο Michael Gilliland είναι πως μέσα σε έναν οργανισμό πρέπει να γίνεται η μέτρηση της απόδοσης πρόβλεψης που πραγματοποιήθηκε, καθώς είναι απαραίτητη για την αξιολόγηση της ακρίβειας των προβλέψεων και της αποτελεσματικότητας των μεθόδων πρόβλεψης που χρησιμοποιούνται. Αυτή η διαδικασία περιλαμβάνει τη σύγκριση των προβλεπόμενων αποτελεσμάτων με τα πραγματικά αποτελέσματα για τον εντοπισμό περιοχών για βελτίωση και τη βελτίωση των μοντέλων πρόβλεψης.

Η πρόβλεψη πωλήσεων είναι μια βασική διαδικασία που επηρεάζει σημαντικά τη βελτιστοποίηση του επιπέδου εξυπηρέτησης πελατών. Με την ακριβή πρόβλεψη των μελλοντικών πωλήσεων, οι οργανισμοί μπορούν να ευθυγραμμίσουν τις επιχειρησιακές τους δυνατότητες με τις απαιτήσεις της αγοράς, ενισχύοντας έτσι την ικανοποίηση και την αφοσίωση των πελατών. Η ικανότητα ενός οργανισμού να ανταποκρίνεται ή να υπερβαίνει τις προσδοκίες των πελατών με συνέπεια είναι ένας κρίσιμος καθοριστικός παράγοντας του ανταγωνιστικού πλεονεκτήματος και της επιτυχίας του στην αγορά. Οι Bowerman, O'Connell, Koehler (*Forecasting, Time Series, and Regression, 2010*) τονίζουν ότι η πρόβλεψη πωλήσεων διευκολύνει την καλύτερη κατανομή των πόρων, διασφαλίζοντας ότι οι απαιτήσεις των πελατών αντιμετωπίζονται έγκαιρα και αποτελεσματικά. Αυτή η στρατηγική ευθυγράμμιση μεταξύ των προβλεπόμενων πωλήσεων και της λειτουργικής εκτέλεσης διασφαλίζει ότι τα επίπεδα

εξυπηρέτησης πελατών όχι μόνο διατηρούνται αλλά βελτιστοποιούνται για να ανταποκρίνονται στις μεταβαλλόμενες δυναμικές της αγοράς.

Η πρόβλεψη πωλήσεων αποτελεί ένα βασικό στοιχείο για την αποτελεσματική υλοποίηση του operational planning, προσφέροντας γνώσεις που καθοδηγούν τον προγραμματισμό χωρητικότητας, την κατανομή του εργατικού δυναμικού και τα υλικοτεχνικά ζητήματα. Οι ακριβείς προβλέψεις επιτρέπουν στις επιχειρήσεις να κλιμακώνουν δυναμικά τις λειτουργίες τους, διασφαλίζοντας ότι οι πόροι χρησιμοποιούνται αποτελεσματικά. Αυτή η προσαρμοστικότητα είναι μεγάλης σημασίας για τη διατήρηση της επιχειρησιακής ευελιξίας και ανταγωνιστικότητας, ιδιαίτερα σε κλάδους που χαρακτηρίζονται από γρήγορες αλλαγές στην αγορά ή υψηλή μεταβλητότητα στη ζήτηση. Η διαδικασία πρόβλεψης πωλήσεων συνδυάζει ουσιαστικά τη στρατηγική της επιχείρησης και τις διαδικασίες που υλοποιούνται για την επίτευξη της. Προσφέροντας μια μακροπρόθεσμη προοπτική για τη ζήτηση, οι προβλέψεις πωλήσεων επιτρέπουν στους οργανισμούς να καταρτίζουν λεπτομερή επιχειρησιακά σχέδια που περιλαμβάνουν την παραγωγή, τη στελέχωση, τη διαχείριση αποθεμάτων και την εφοδιαστική (*Chase, Demand-Driven Forecasting: A Structured Approach to Forecasting, 2013*). Η ενσωμάτωση της πρόβλεψης πωλήσεων στο operational planning αποφέρει αρκετά απτά οφέλη. Ένα από αυτά είναι πως οι ακριβείς προβλέψεις πωλήσεων διασφαλίζουν ότι οι πόροι, συμπεριλαμβανομένης της εργασίας και των υλικών, κατανέμονται αποτελεσματικά, αποτρέποντας την υπερπαραγωγή και τη σπατάλη πόρων (*Bowerman 2010*). Αυτή η βελτιστοποίηση συμβάλλει στη μείωση του κόστους και τις προσπάθειες βιωσιμότητας. Με την πρόβλεψη των διακυμάνσεων της ζήτησης, η πρόβλεψη πωλήσεων βοηθά επίσης, στην ανάπτυξη ευέλικτων χρονοδιαγραμμάτων παραγωγής που μπορούν να προσαρμοστούν στις αλλαγές της αγοράς, βελτιώνοντας έτσι την αποδοτικότητα της παραγωγής και μειώνοντας τους χρόνους παράδοσης. Τέλος, οι προβλέψεις πωλήσεων επιτρέπουν στους οργανισμούς να ευθυγραμμίσουν τις επιχειρησιακές τους δραστηριότητες με στρατηγικούς επιχειρηματικούς στόχους, διασφαλίζοντας ότι τα επιχειρησιακά σχέδια υποστηρίζουν τους μακροπρόθεσμους στόχους τους.

Στο πλαίσιο της έννοιας του Sales Targeting, η πρόβλεψη βοηθά στον καθορισμό ρεαλιστικών στόχων πωλήσεων και στη βελτίωση των στρατηγικών μάρκετινγκ. Κατανοώντας τα μελλοντικά μοτίβα πωλήσεων, οι επιχειρήσεις μπορούν να προσαρμόσουν τις προσπάθειές τους για να κεφαλαιοποιήσουν τις προβλεπόμενες αλλαγές ζήτησης, βελτιστοποιώντας την απόδοση της επένδυσης για καμπάνιες μάρκετινγκ. Επιπλέον, οι προβλέψεις πωλήσεων βοηθούν στον εντοπισμό πιθανών ευκαιριών ανάπτυξης, καθοδηγώντας στρατηγικές αποφάσεις σχετικά με την επέκταση της αγοράς, την ανάπτυξη προϊόντων ή τη διαφοροποίηση. Σχετικά με τον ρεαλιστικό στόχο πωλήσεων που αναφέραμε παραπάνω, Michael Gilliland (*Fundamental Issues in Business Forecasting, 2003*) επισημαίνει πως στους περισσότερους οργανισμούς η πρόβλεψη είναι πιο συχνά μια έκφραση των στόχων ή των επιθυμιών του οργανισμού και όχι μια πρόβλεψη για το τι πρόκειται πραγματικά να συμβεί. Το Evangelical forecasting είναι μια προσέγγιση όπου η πρόβλεψη δίνεται «από τα πάνω» δηλαδή από τα ανώτερα ιεραρχικά επίπεδα. Σε αυτό το είδος περιβάλλοντος ένα στέλεχος καθορίζει την πρόβλεψη, η οποία μπορεί να μην είναι τίποτα περισσότερο από μια έκφραση του στόχου εσόδων. Αυτή η πρόβλεψη καθορίζεται σε χρηματικά ποσά (και όχι σε μονάδες), σε κάποιο υψηλό επίπεδο συγκέντρωσης όπως για παράδειγμα οι ανά πωλήσεις ανά επικράτεια και για ένα ευρύ χρονικό πλαίσιο (όπως έτος ή τρίμηνο). Σε αυτή την περίπτωση, η διαδικασία πρόβλεψης, αφορά απλώς τη προσαρμογή του όγκου του προϊόντος για να επιτευχθεί ο στόχος.

Ο αποτελεσματικός προϋπολογισμός (budgeting) και ο οικονομικός σχεδιασμός (financial planning) είναι δύο θεμελιώδεις έννοιες για την επιτυχία του οργανισμού, απαιτώντας ακριβή ευθυγράμμιση μεταξύ των οικονομικών πόρων και των στρατηγικών στόχων. Η πρόβλεψη πωλήσεων αναδεικνύεται ως μια κομβική διαδικασία σε αυτό το πλαίσιο, προσφέροντας μια προγνωστική προοπτική που ενημερώνει τη λήψη οικονομικών αποφάσεων. Σύμφωνα με τον Vic Richard (*Achieving and Maintaining Optimal Supply-and-Demand Alignment*) το financial planning είναι η διαδικασία που καθορίζει πώς η επιχείρηση θα έχει την οικονομική δυνατότητα να επιτύχει τους στρατηγικούς της στόχους. Συνήθως, μια εταιρεία δημιουργεί ένα οικονομικό σχέδιο (financial plan) αμέσως μετά τον ορισμό των στόχων. Το οικονομικό σχέδιο περιγράφει καθεμία από τις δραστηριότητες, τους πόρους, τον εξοπλισμό και τα υλικά που απαιτούνται για την επίτευξη αυτών των στόχων, καθώς και τα εμπλεκόμενα χρονικά πλαίσια. Η διαδικασία του

financial planning περιλαμβάνει διάφορες εργασίες όπως η αξιολόγηση του επιχειρηματικού περιβάλλοντος, η επιβεβαίωση του επιχειρηματικού οράματος, ο προσδιορισμός των πόρων που απαιτούνται για την επίτευξη αυτών των στόχων, η ποσοτικοποίηση της ποσότητας κάθε πόρου (εργασία, εξοπλισμός και υλικά), ο υπολογισμός του συνολικού κόστους κάθε τύπου πόρου, ο συντονισμός των παραπάνω για τη δημιουργία προϋπολογισμού και ο προσδιορισμός τυχόν κινδύνων και ζητημάτων με το σύνολο του προϋπολογισμού. Με την πρόβλεψη μελλοντικών πωλήσεων, οι οργανισμοί μπορούν να καταναείμουν τους οικονομικούς και λειτουργικούς πόρους πιο αποτελεσματικά, δίνοντας προτεραιότητα στις επενδύσεις σε τομείς με τις υψηλότερες αναμενόμενες αποδόσεις (Fildes & Goodwin, 2007). Αυτή η στρατηγική κατανομή υποστηρίζει την αποτελεσματική χρήση του κεφαλαίου και ενισχύει τη συνολική χρηματοοικονομική απόδοση (financial performance).

2.2 Οι χρήσεις της Πρόβλεψης στα μοντέλα του promotional modelling & του price optimization

Σε αυτή την ενότητα της παρούσας μελέτης θα εστιάσουμε στο promotional modelling και στο price optimization καθώς και για την σχέση τους με την πρόβλεψη πωλήσεων.

Το Promotional Modelling είναι ένας τύπος μοντελοποίησης που εστιάζει στην προώθηση προϊόντων ή υπηρεσιών σε εκδηλώσεις, εμπορικές εκθέσεις και άλλες δραστηριότητες μάρκετινγκ. Περιλαμβάνει τη χρήση ambassadors για να προσελκύσουν την προσοχή στα προϊόντα ή τις υπηρεσίες μιας εταιρείας και να δημιουργήσουν ενδιαφέρον μεταξύ των πιθανών πελατών. Το Promotional Modelling συνήθως χρησιμοποιείται για να αντιπροσωπεύσει το εμπορικό σήμα μιας εταιρείας και να προωθήσει τα προϊόντα ή τις υπηρεσίες της επιχείρησης σε εκδηλώσεις. Το μοντέλο εκπαιδεύεται να αλληλεπιδρά με πελάτες, να επιδεικνύει προϊόντα και να διανέμει διαφημιστικό υλικό. Μπορεί επίσης να είναι υπεύθυνο για τη δημιουργία δυναμικών πελατών, τη συλλογή δεδομένων πελατών και τη διεξαγωγή ερευνών για τη συλλογή σχολίων από τους πελάτες. Το promotional modelling είναι ένας αποτελεσματικός τρόπος για τις εταιρείες να δημιουργήσουν ντόρο γύρω από τα προϊόντα ή τις υπηρεσίες τους και να δημιουργήσουν ενδιαφέρον μεταξύ των πιθανών πελατών. Στο promotional modelling, η πρόβλεψη είναι ιδιαίτερα σημαντική γιατί μπορεί να βοηθήσει τις επιχειρήσεις να εντοπίσουν το βέλτιστο μείγμα προώθησης για τα προϊόντα ή τις υπηρεσίες τους. Η ακριβής πρόβλεψη δίνει τη δυνατότητα στις επιχειρήσεις να προβλέψουν τον αντίκτυπο των διαφημιστικών δραστηριοτήτων στις πωλήσεις και να καταναείμουν τους προϋπολογισμούς μάρκετινγκ πιο αποτελεσματικά. Είναι σημαντικό να αναφέρουμε ότι αναλύοντας ιστορικά δεδομένα και τη συμπεριφορά των πελατών, οι επιχειρήσεις μπορούν να προβλέψουν την αποτελεσματικότητα διαφορετικών διαφημιστικών δραστηριοτήτων και να προσαρμόσουν τα μηνύματα μάρκετινγκ σε συγκεκριμένα τμήματα πελατών. Επιπλέον, η πρόβλεψη μπορεί να βοηθήσει τις επιχειρήσεις να προσδιορίσουν τις τάσεις της αγοράς και να λάβουν αποφάσεις βάσει δεδομένων σχετικά με τον τρόπο βελτιστοποίησης των διαφημιστικών τους καμπανιών.

Το price optimization επωφελείται επίσης από την ακριβή πρόβλεψη. Το price optimization είναι η διαδικασία καθορισμού της σωστής τιμής για ένα προϊόν ή μια υπηρεσία για τη μεγιστοποίηση των εσόδων ή της κερδοφορίας. Περιλαμβάνει την ανάλυση δεδομένων και τάσεων της αγοράς για τον προσδιορισμό του βέλτιστου σημείου τιμής που εξισορροπεί τις ανάγκες της επιχείρησης με τις ανάγκες των πελατών. Το price optimization μπορεί να γίνει μέσω ποικίλων μεθόδων, όπως η έρευνα αγοράς, η ανταγωνιστική ανάλυση και τα πειράματα τιμολόγησης. Σε κάθε περίπτωση, ο στόχος είναι να καθοριστεί η σωστή τιμή που θα προσελκύσει πελάτες και θα μεγιστοποιήσει τα κέρδη για την επιχείρηση. Μια κοινή προσέγγιση για τη βελτιστοποίηση των τιμών είναι η δυναμική τιμολόγηση, η οποία περιλαμβάνει την προσαρμογή των τιμών σε πραγματικό χρόνο με βάση την προσφορά και τη ζήτηση. Για παράδειγμα, μια αεροπορική εταιρεία μπορεί να αυξήσει την τιμή μιας πτήσης καθώς πλησιάζει την ημερομηνία αναχώρησης εάν η ζήτηση είναι υψηλή ή να μειώσει την τιμή εάν υπάρχουν πολλές απούλητες θέσεις. Αυτό επιτρέπει στην αεροπορική εταιρεία να μεγιστοποιεί τα έσοδά της, διασφαλίζοντας παράλληλα την κάλυψη των θέσεων. Μια άλλη προσέγγιση για τη βελτιστοποίηση τιμών είναι η τμηματοποίηση, η οποία περιλαμβάνει τον καθορισμό διαφορετικών τιμών για διαφορετικά τμήματα πελατών. Για παράδειγμα, ένα ξενοδοχείο μπορεί να προσφέρει χαμηλότερες τιμές σε ταξιδιώτες αναψυχής

που κάνουν κράτηση εκ των προτέρων, ενώ χρεώνει υψηλότερες τιμές σε επαγγελματίες ταξιδιώτες που κάνουν κράτηση την τελευταία στιγμή. Με την πρόβλεψη λοιπόν της ζήτησης και την αξιολόγηση των ανταγωνιστικών πιέσεων, οι επιχειρήσεις μπορούν να βελτιστοποιήσουν τις στρατηγικές τιμολόγησης τους για να μεγιστοποιήσουν την κερδοφορία. Η ακριβής πρόβλεψη μπορεί να βοηθήσει τις επιχειρήσεις να καθορίσουν τα βέλτιστα σημεία τιμών για τα προϊόντα ή τις υπηρεσίες τους, να αξιολογήσουν τον αντίκτυπο των αλλαγών στην τιμολόγηση και να προβλέψουν τα αποτελέσματα της κυκλοφορίας νέων προϊόντων. Είναι φυσικό πως μελετώντας και αναλύοντας τη συμπεριφορά των ανταγωνιστών και τις ανάγκες των πελατών, οι επιχειρήσεις μπορούν να ορίσουν τιμές ανταγωνιστικές και μέγιστα κερδοφόρες. Επιπλέον, η πρόβλεψη μπορεί να βοηθήσει τις επιχειρήσεις να εντοπίσουν τους παράγοντες που επηρεάζουν τη ζήτηση και να λάβουν τεκμηριωμένες αποφάσεις σχετικά με τον τρόπο προσαρμογής των στρατηγικών τιμολόγησης τους ανάλογα.

Συνολικά, η ακριβής πρόβλεψη είναι απαραίτητη για τις επιχειρήσεις που επιδιώκουν να βελτιστοποιήσουν τις στρατηγικές μοντελοποίησης προώθησης και βελτιστοποίησης τιμών. Παρέχοντας πληροφορίες για τη συμπεριφορά των πελατών και προβλέποντας τη μελλοντική ζήτηση, οι επιχειρήσεις μπορούν να προσαρμόσουν τις στρατηγικές προώθησης και τιμολόγησης τους για να μεγιστοποιήσουν την κερδοφορία. Η πρόβλεψη μπορεί να προσφέρει τη βάση για τη λήψη στρατηγικών αποφάσεων και να επιτρέψει στις επιχειρήσεις να λαμβάνουν τεκμηριωμένες αποφάσεις που οδηγούν σε βελτιωμένη απόδοση και ανάπτυξη.

3. Χρονοσειρές

3.1 Εισαγωγή στην Ανάλυση Χρονοσειρών

Στον τομέα της στατιστικής καθώς και της ανάλυσης δεδομένων, η ανάλυση χρονοσειρών αποτελεί μια βασική τεχνική, που αντιμετωπίζει τις πολυπλοκότητες που είναι εγγενείς στα δεδομένα που εξαρτώνται και σχετίζονται με το χρόνο. Πιο συγκεκριμένα, στον πυρήνα της, η ανάλυση χρονοσειρών περιλαμβάνει τον έλεγχο των σημείων δεδομένων που συλλέγονται ή καταγράφονται σε διαδοχικά χρονικά διαστήματα. Αυτά τα διαστήματα μπορεί να κυμαίνονται από χιλιοστά του δευτερολέπτου στην περίπτωση δεδομένων συναλλαγών υψηλής συχνότητας έως χρόνια κατά την ανάλυση των δημογραφικών τάσεων. Η ουσία αυτής της ανάλυσης έγκειται στην κατανόηση, την ερμηνεία και την πρόβλεψη της χρονικής δυναμικής που είναι ενσωματωμένη στα δεδομένα. Οι μεθοδολογίες που χρησιμοποιούνται στην ανάλυση χρονοσειρών περιλαμβάνουν αλλά δεν περιορίζονται σε περιγραφική ανάλυση, συμπερασματικές στατιστικές μεθόδους και προηγμένα μοντέλα πρόβλεψης. Αυτή η ποικιλομορφία στις προσεγγίσεις, στις οποίες θα αναφερθούμε και στη συνέχεια αυτού του κεφαλαίου, επιτρέπει την πλήρη κατανόηση της χρονικής εξέλιξης των δεδομένων και των υποκείμενων προτύπων.

Είναι σημαντικό να αναφέρουμε πως το εύρος της ανάλυσης χρονοσειρών είναι τεράστιο, ενσωματώνοντας διάφορες στατιστικές τεχνικές προσαρμοσμένες για την αντιμετώπιση των μοναδικών προκλήσεων που παρουσιάζονται από τα χρονικά δεδομένα. Θεμελιώδης σε αυτήν την ανάλυση είναι ο εντοπισμός στοιχείων όπως οι τάσεις, οι εποχιακές διακυμάνσεις, τα κυκλικά μοτίβα και οι ακανόνιστες διακυμάνσεις. Τεχνικές όπως τα μοντέλα Autoregressive Integrated Moving Average (ARIMA), Seasonal Decomposition of Time Series (SDTS) τα οποία θα αναλυθούν περαιτέρω και στη συνέχεια, καθώς και προσεγγίσεις μηχανικής μάθησης, έχουν αναπτυχθεί για την αποτελεσματική πρόβλεψη αυτών των στοιχείων.

Η εφαρμογή αυτών των τεχνικών επιτρέπει τον μετασχηματισμό των ακατέργαστων χρονικών δεδομένων σε διορατικές, πρακτικές πληροφορίες οι οποίες μπορούν να χρησιμοποιηθούν από τους οργανισμούς και να επιφέρουν ωφέλιμα αποτελέσματα. Πιο συγκεκριμένα, στον σημερινό εταιρικό κόσμο που βασίζεται σε δεδομένα, η ανάλυση χρονοσειρών έχει αναδειχθεί ως βασικό αναλυτικό εργαλείο, προσφέροντας πολύτιμες πληροφορίες για τις τάσεις της αγοράς, τη συμπεριφορά των καταναλωτών και την οικονομική απόδοση. Αυτή η στατιστική μέθοδος, η οποία περιλαμβάνει την ανάλυση χρονικά ταξινομημένων σημείων δεδομένων, είναι ζωτικής σημασίας για τη λήψη τεκμηριωμένων αποφάσεων σε διάφορους επιχειρηματικούς τομείς. Μία από τις κύριες χρήσεις της ανάλυσης χρονοσειρών στις εταιρείες είναι η πρόβλεψη των τάσεων

της αγοράς. Εξετάζοντας δεδομένα προηγούμενων πωλήσεων, οι εταιρείες μπορούν να προβλέψουν τη μελλοντική ζήτηση, να προσαρμόσουν τα προγράμματα παραγωγής τους και να βελτιστοποιήσουν τα επίπεδα αποθεμάτων. Αυτή η πρόβλεψη είναι ιδιαίτερα σημαντική σε κλάδους με εποχιακές διακυμάνσεις, όπου η πρόβλεψη αλλαγών στη ζήτηση των καταναλωτών μπορεί να οδηγήσει σε πιο αποτελεσματικές λειτουργίες και υψηλότερα περιθώρια κέρδους. Επιπλέον, η ανάλυση χρονοσειρών επιτρέπει στις εταιρείες να κατανοήσουν και να προβλέψουν τη συμπεριφορά των καταναλωτών. Αναλύοντας τις τάσεις με την πάροδο του χρόνου, οι εταιρείες μπορούν να εντοπίσουν μοτίβα στις αγορές και τις προτιμήσεις των πελατών, οδηγώντας σε πιο στοχευμένες στρατηγικές μάρκετινγκ και ανάπτυξη προϊόντων. Στον τομέα των οικονομικών, η ανάλυση χρονοσειρών είναι επίσης απαραίτητη για τη λήψη εταιρικών αποφάσεων. Χρησιμοποιείται εκτενώς για την ανάλυση διάφορων τιμών και δεικτών παρέχοντας έτσι πληροφορίες για τη δυναμική της αγοράς και βοηθώντας τις εταιρείες να λαμβάνουν τεκμηριωμένες επενδυτικές αποφάσεις. Ως απόρροια, αυτή η ανάλυση βοηθά στον εντοπισμό πιθανών κινδύνων και ευκαιριών στην αγορά, επιτρέποντας στις εταιρείες να σχεδιάσουν αναλόγως τη στρατηγική τους. Η διαχείριση κινδύνου επωφελείται επίσης πολύ από την ανάλυση χρονοσειρών. Αναλύοντας ιστορικά δεδομένα, οι εταιρείες μπορούν να προβλέψουν πιθανούς χρηματοοικονομικούς κινδύνους και να εφαρμόσουν μέτρα για τον μετριασμό τους. Αυτή η προγνωστική ικανότητα είναι απαραίτητη για τη διατήρηση της χρηματοπιστωτικής σταθερότητας και τη διασφάλιση μακροπρόθεσμης ανάπτυξης. Η ανάλυση χρονοσειρών παίζει επίσης κρίσιμο ρόλο στην ενίσχυση της λειτουργικής αποτελεσματικότητας (operations efficiency). Εξετάζοντας ιστορικά δεδομένα σχετικά με τις διαδικασίες παραγωγής, οι εταιρείες μπορούν να εντοπίσουν αναποτελεσματικότητα και τομείς προς βελτίωση. Αυτό οδηγεί σε καλύτερη κατανομή πόρων, μείωση κόστους και συνολική λειτουργική βελτιστοποίηση. Ο στρατηγικός σχεδιασμός στις εταιρείες είναι ένας άλλος τομέας όπου η ανάλυση χρονοσειρών αποδεικνύεται ανεκτίμητη. Παρέχοντας σαφή κατανόηση των τάσεων του παρελθόντος και των μελλοντικών προβολών, βοηθά τα στελέχη να θέσουν ρεαλιστικούς στόχους και να διαμορφώσουν αποτελεσματικές στρατηγικές. Αυτή η προνοητικότητα είναι ζωτικής σημασίας για την πλοήγηση στο συνεχώς μεταβαλλόμενο επιχειρηματικό τοπίο και τη διατήρηση ενός ανταγωνιστικού πλεονεκτήματος.

3.2 Ιστορική αναδρομή Ανάλυσης Χρονοσειρών

Η εξέλιξη της ανάλυσης χρονοσειρών χαρακτηρίζεται από σημαντικές εξελίξεις, από στοιχειώδεις μεθόδους στην πρώιμη αστρονομία έως εξελιγμένους αλγόριθμους στη σημερινή ψηφιακή εποχή. Οι ρίζες της ανάλυσης χρονοσειρών μπορούν να εντοπιστούν στην αρχαία αστρονομία, όπου οι πρώτοι μελετητές παρατήρησαν ουράνια σώματα με την πάροδο του χρόνου. Αυτές οι παρατηρήσεις ήταν ουσιαστικά υποτυπώδεις μορφές δεδομένων χρονοσειρών. Ωστόσο, η επίσημη ανάπτυξη της ανάλυσης χρονοσειρών ξεκίνησε στα τέλη του 17ου αιώνα με την εμφάνιση της συστηματικής τήρησης αρχείων. Αστρονόμοι όπως ο Edmund Halley χρησιμοποίησαν αυτές τις μεθόδους για να προβλέψουν ουράνια γεγονότα, θέτοντας τις βάσεις για πιο προηγμένη στατιστική ανάλυση.

Ο 19ος αιώνας γνώρισε σημαντικές προόδους στην ανάλυση χρονοσειρών, κυρίως λόγω των αναγκών στη μετεωρολογία και την οικονομία. Οι μετεωρολόγοι χρησιμοποίησαν δεδομένα χρονοσειρών για να κατανοήσουν τα καιρικά μοτίβα και να κάνουν πρώιμες προβλέψεις. Ταυτόχρονα, οι οικονομολόγοι άρχισαν να χρησιμοποιούν αυτές τις τεχνικές για την ανάλυση χρηματοοικονομικών και οικονομικών δεδομένων. Αυτή η περίοδος σηματοδότησε την αρχή της χρήσης γραφικών μεθόδων για την αναπαράσταση δεδομένων χρονοσειρών, βοηθώντας στην αποτελεσματικότερη ερμηνεία των δεδομένων.

Οι αρχές του 20ου αιώνα σηματοδότησαν μια κομβική στιγμή στην ιστορία της ανάλυσης χρονοσειρών με το έργο πρωτοπόρων στατιστικολόγων όπως ο Francis Galton και ο Karl Pearson. Το έργο τους έθεσε τα θεμέλια για ανάλυση συσχέτισης και παλινδρόμησης, που αποτελούν αναπόσπαστο μέρος της σύγχρονης ανάλυσης χρονοσειρών. Η περίοδος είδε επίσης την ανάπτυξη της μεθόδου των ελαχίστων τετραγώνων, μια θεμελιώδη προσέγγιση για την προσαρμογή δεδομένων και την πρόβλεψη σε χρονοσειρές.

Η περίοδος μετά τον Β' Παγκόσμιο Πόλεμο σηματοδότησε μια σημαντική καμπή στον τομέα της ανάλυσης χρονοσειρών, που χαρακτηρίζεται από την εισαγωγή και την ευρεία υιοθέτηση της

μεθοδολογίας Box-Jenkins και των μοντέλων Autoregressive Integrated Moving Average (ARIMA). Αυτή η εποχή, που εκτείνεται σε μεγάλο βαθμό τις δεκαετίες του 1960 και του 1970, υπήρξε μάρτυρας πρωτοποριακών εξελίξεων που αναδιαμόρφωσαν το τοπίο των στατιστικών προβλέψεων και αναλύσεων. Ο George Box και ο Gwilym Jenkins, και οι δύο σημαντικοί στατιστικολόγοι, συνέβαλαν καθοριστικά στην ανάπτυξη μιας συστηματικής προσέγγισης για την κατασκευή, τη δοκιμή και την εφαρμογή μοντέλων αυτοπαλίνδρομου κινητού μέσου όρου (ARMA) και ARIMA για δεδομένα χρονοσειρών. Το έργο τους κορυφώθηκε με τη δημοσίευση του θεμελιώδους βιβλίου "Ανάλυση Χρονοσειρών: Πρόβλεψη και Έλεγχος" το 1970, το οποίο περιγράφει λεπτομερώς τη μεθοδολογία τους και τις εφαρμογές της. Η μεθοδολογία Box-Jenkins, όπως έγινε γνωστή, παρείχε ένα ολοκληρωμένο πλαίσιο για τον εντοπισμό, την εκτίμηση και τον έλεγχο μοντέλων για ανάλυση χρονοσειρών.

Πηγαίνοντας πιο κοντά στο σήμερα, η έλευση της ψηφιακής εποχής και η έκρηξη της υπολογιστικής ισχύος έφεραν επανάσταση στην ανάλυση χρονοσειρών. Η διαθεσιμότητα μεγάλων συνόλων δεδομένων και η ανάπτυξη εξελιγμένων αλγορίθμων έχουν επιτρέψει πιο σύνθετες τεχνικές μοντελοποίησης, όπως η μηχανική μάθηση και τα νευρωνικά δίκτυα. Αυτές οι εξελίξεις έχουν βελτιώσει σημαντικά την ακρίβεια των προβλέψεων και έχουν διευρύνει την εφαρμογή της ανάλυσης χρονοσειρών σε διάφορους τομείς, συμπεριλαμβανομένων των οικονομικών, του μάρκετινγκ και της βιολογίας.

3.3 Βασικές έννοιες στην ανάλυση Χρονοσειρών

Όπως αναφέρθηκε και παραπάνω, η ανάλυση χρονοσειρών αποτελεί μια βασική στατιστική τεχνική για την αποκρυπτογράφηση προτύπων και την πρόβλεψη μελλοντικών τάσεων στα διαδοχικά δεδομένα της οποίας η σημασία εκτείνεται σε πολλούς τομείς. Στη βάση της ανάλυσης χρονοσειρών βρίσκονται τέσσερις θεμελιώδεις έννοιες οι οποίες είναι η σταθερότητα, η αυτοσυσχέτιση, οι τάσεις και η εποχικότητα. Κάθε μια από αυτές παίζει καθοριστικό ρόλο στην κατασκευή και στην ερμηνεία των μοντέλων χρονοσειρών. Πιο αναλυτικά:

Η σταθερότητα (Stationarity) είναι μια κρίσιμη ιδιότητα των δεδομένων χρονοσειρών, η οποία σημαίνει ότι τα στατιστικά χαρακτηριστικά της σειράς δεν αλλάζουν με την πάροδο του χρόνου. Αυτή η συνέπεια στη μέση τιμή, τη διακύμανση και την αυτοσυσχέτιση είναι απαραίτητη για την εφαρμογή πολλών μοντέλων χρονοσειρών. Η Αυστηρή σταθερότητα (Strict Stationarity) είναι ένας τύπος σταθερότητας ο οποίος απαιτεί οι στατιστικές ιδιότητες της χρονοσειράς να είναι αμετάβλητες στις χρονικές μετατοπίσεις, υπονοώντας ότι η κοινή κατανομή οποιουδήποτε υποσυνόλου της σειράς δεν αλλάζει με την πάροδο του χρόνου. Η ασθενής σταθερότητα (Weak Stationarity) που εφαρμόζεται πιο συχνά απαιτεί σταθερό μέσο όρο και διακύμανση σε όλη τη σειρά, με την αυτοσυνδιακύμανση να εξαρτάται μόνο από την υστέρηση μεταξύ των παρατηρήσεων και όχι από τα πραγματικά χρονικά σημεία. Στις περιπτώσεις της Δοκιμής (Testing for Stationarity) για σταθερότητα χρησιμοποιούνται διάφορες στατιστικές δοκιμές για να εξακριβωθεί η σταθερότητα, συμπεριλαμβανομένης της δοκιμής Augmented Dickey-Fuller (ADF), της δοκιμής Phillips-Perron (PP) και της δοκιμής Kwiatkowski-Phillips-Schmidt-Shin (KPSS). Αυτές οι δοκιμές συνήθως επικεντρώνονται στην ανίχνευση μοναδιαίων ριζών, οι οποίες υποδεικνύουν μη σταθερότητα στη σειρά.

Η αυτοσυσχέτιση (Autocorrelation), ή η σειριακή συσχέτιση, αντιπροσωπεύει τον βαθμό ομοιότητας μεταξύ μιας χρονικής σειράς και μιας καθυστερημένης εκδοχής του εαυτού της σε διαδοχικά χρονικά διαστήματα. Είναι δηλαδή ένα μέτρο των εσωτερικών εξαρτήσεων μέσα στη σειρά. Η συνάρτηση αυτοσυσχέτισης (ACF) αξιολογεί τη συσχέτιση μεταξύ της σειράς και των τιμών της με καθυστέρηση, προσφέροντας πληροφορίες για επαναλαμβανόμενα μοτίβα στα δεδομένα. Η συνάρτηση μερικής αυτοσυσχέτισης (PACF) αντίθετα, απομονώνει τη συσχέτιση της σειράς με τις καθυστερήσεις της, αφαιρώντας την επιρροή των ενδιάμεσων καθυστερήσεων.

Οι τάσεις και η εποχικότητα είναι εγγενή χαρακτηριστικά πολλών συνόλων δεδομένων χρονοσειρών, που αντιπροσωπεύουν συστηματικά μοτίβα και κυκλικές διακυμάνσεις. Οι τάσεις αντιπροσωπεύουν μακροπρόθεσμες κινήσεις στα δεδομένα, είτε προς τα πάνω είτε προς τα κάτω. Ο εντοπισμός και ο ποσοτικός προσδιορισμός των τάσεων είναι ζωτικής σημασίας για την κατανόηση της συνολικής κατεύθυνσης της σειράς. Τεχνικές όπως γραμμικά και μη γραμμικά μοντέλα παλινδρόμησης, κινούμενοι μέσοι όροι και μέθοδοι εξομάλυνσης (π.χ. φίλτρο Hodrick-

Prescott) χρησιμοποιούνται συχνά για τον εντοπισμό και τη μοντελοποίηση αυτών των τάσεων. Η εποχικότητα περιλαμβάνει μοτίβα που επαναλαμβάνονται σε τακτά χρονικά διαστήματα, που συνήθως επηρεάζονται από εποχιακούς παράγοντες όπως η εποχή του έτους, ο μήνας ή η εβδομάδα. Η αναγνώριση αυτών των προτύπων είναι το κλειδί για τις προβλέψεις σε περιοχές με σαφείς εποχιακές επιρροές όπως είναι για παράδειγμα οι λιανικές πωλήσεις. Η εποχιακή αποσύνθεση χρονοσειρών (αποσύνθεση STL) και τα εποχιακά μοντέλα ARIMA (SARIMA) είναι δημοφιλείς μέθοδοι μοντελοποίησης και προσαρμογής για την εποχικότητα.

3.4 Δοκιμές στην Ανάλυση Χρονοσειρών

Στην ανάλυση χρονοσειρών, η χρήση διαφόρων στατιστικών δοκιμών είναι υψίστης σημασίας για την κατανόηση των υποκείμενων χαρακτηριστικών των δεδομένων με χρονική σειρά. Αυτές οι δοκιμές είναι απαραίτητα εργαλεία για τον εντοπισμό της σταθερότητας, της εποχικότητας, του λευκού θορύβου και άλλων κρίσιμων χαρακτηριστικών που επηρεάζουν την επιλογή και την αποτελεσματικότητα των τεχνικών μοντελοποίησης. Παρακάτω αναφέρονται τα πιο κοινά τεστ στην ανάλυση χρονοσειρών, διευκρινίζοντας τους σκοπούς, τις μεθοδολογίες και τη σημασία τους:

Unit Root Tests: Οι δοκιμές μοναδιαίας ρίζας είναι θεμελιώδεις για τον προσδιορισμό της σταθερότητας μιας χρονοσειράς. Μια μοναδιαία ρίζα υποδηλώνει ότι μια σειρά είναι μη ακίνητη και οι κραδασμοί της έχουν μόνιμο αποτέλεσμα. Οι πιο ευρέως χρησιμοποιούμενες δοκιμές μοναδιαίας ρίζας περιλαμβάνουν τη δοκιμή Dickey-Fuller, τη δοκιμή Augmented Dickey-Fuller (ADF) και τη δοκιμή Phillips-Perron. Η μέθοδος Dickey-Fuller και η επαυξημένη έκδοσή του, η ADF, είναι οι ακρογωνιαίοι λίθοι για τη σταθερότητα. Δοκιμάζουν τη μηδενική υπόθεση ότι μια μοναδιαία ρίζα υπάρχει σε μια χρονοσειρά. Η δοκιμή ADF επεκτείνει τη δοκιμή Dickey-Fuller συμπεριλαμβάνοντας διαφορές καθυστέρησης για να ληφθούν υπόψη οι αυτοπαλινδρομικές διαδικασίες υψηλότερης τάξης. Το τεστ Phillips-Perron βελτιώνει τα τεστ Dickey-Fuller επιτρέποντας πιο γενικές μορφές ετεροσκεδαστικότητας στον όρο σφάλματος. Είναι ιδιαίτερα χρήσιμο στην αντιμετώπιση χρονοσειρών που παρουσιάζουν αλλαγές στη διακύμανση με την πάροδο του χρόνου.

Seasonal Unit Root Tests: Οι εποχικές δοκιμές ρίζας μονάδας είναι ζωτικής σημασίας για την ανίχνευση της μη σταθερότητας στα εποχιακά δεδομένα. Βοηθούν στον προσδιορισμό του αν τα εποχιακά μοτίβα είναι σταθερά με την πάροδο του χρόνου ή εξελίσσονται. Το τεστ Hylleberg-Engle-Granger-Yoo (HEGY) είναι ένα χαρακτηριστικό παράδειγμα. Η δοκιμή HEGY έχει σχεδιαστεί για να ελέγχει τις εποχιακές μοναδιαίες ρίζες σε διαφορετικές εποχιακές συχνότητες. Αυτό το τεστ είναι καθοριστικό για την ανάλυση τριμηνιαίων και μηνιαίων δεδομένων όπου τα εποχιακά μοτίβα αποτελούν βασικό στοιχείο.

White Noise Tests: Οι δοκιμές λευκού θορύβου χρησιμοποιούνται για να ελέγξουν εάν μια σειρά είναι καθαρά τυχαία, κάτι που είναι μια σημαντική υπόθεση σε πολλά μοντέλα χρονοσειρών. Η δοκιμή Ljung-Box και η δοκιμή Box-Pierce είναι οι κύριες μέθοδοι που χρησιμοποιούνται για το σκοπό αυτό. Η δοκιμή Ljung-Box ελέγχει την απουσία αυτοσυσχέτισης σε πολλαπλά μήκη υστέρησης. Χρησιμοποιείται ευρέως στον διαγνωστικό έλεγχο μοντέλων για τον έλεγχο της τυχαιότητας των υπολειμμάτων από ένα μοντέλο προσαρμοσμένης χρονοσειράς. Η δοκιμή Box-Pierce είναι μια παλαιότερη έκδοση της δοκιμής Ljung-Box και εξυπηρετεί παρόμοιο σκοπό. Είναι λιγότερο ανθεκτική σε σύγκριση με τη δοκιμή Ljung-Box, αλλά εξακολουθεί να χρησιμοποιείται για την απλότητά του.

Stationarity Tests: Εκτός από τις δοκιμές μοναδιαίας ρίζας, χρησιμοποιούνται ειδικές δοκιμές όπως η δοκιμή KPSS (Kwiatkowski-Phillips-Schmidt-Shin) για την περαιτέρω διερεύνηση της σταθερότητας. Τα τεστ KPSS ακολουθούν μια διαφορετική προσέγγιση δοκιμάζοντας τη μηδενική υπόθεση ότι μια χρονοσειρά είναι ακίνητη γύρω από μια ντετερμινιστική τάση. Συμπληρώνει τις δοκιμές μοναδιαίας ρίζας κάνοντας διάκριση μεταξύ μιας σταθερής σειράς γύρω από μια τάση και μιας στάσιμης σειράς διαφοράς.

Τα διάφορα τεστ στην ανάλυση χρονοσειρών είναι αναπόσπαστα για την κατανόηση και την προετοιμασία δεδομένων για ακριβή μοντελοποίηση και πρόβλεψη. Αυτές οι δοκιμές καθοδηγούν τους αναλυτές στον εντοπισμό των κατάλληλων μετασχηματισμών και μοντέλων για μια δεδομένη χρονική σειρά. Καθώς το πεδίο της ανάλυσης χρονοσειρών συνεχίζει να αυξάνεται, ειδικά με τις

νέες μεθόδους συλλογής δεδομένων, ο ρόλος αυτών των δοκιμών γίνεται όλο και πιο κρίσιμος στην εξαγωγή ουσιαστικών και χρήσιμων πληροφοριών από χρονικά δεδομένα. Η σωστή κατανόηση και εφαρμογή αυτών των δοκιμών είναι ιδιαίτερα σημαντική αν θέλουμε να αξιοποιήσουμε πλήρως τις δυνατότητες της ανάλυσης χρονοσειρών σε πρακτικές εφαρμογές.

4. Μέθοδοι Πρόβλεψης

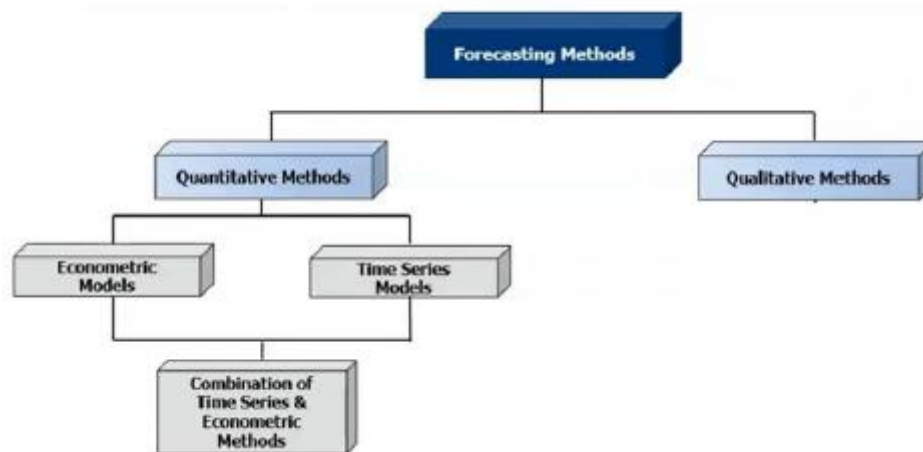
4.1 Ποσοτική Πρόβλεψη

Για τις ανάγκες της συγκεκριμένης μελέτης, θα χρησιμοποιηθεί quantitative forecasting δηλαδή ποσοτικές μέθοδοι πρόβλεψης. Συνεπώς είναι αρκετά σημαντικό να περιγράψουμε τις συγκεκριμένες μεθόδους και να μελετήσουμε τα χαρακτηριστικά τους αντίστοιχα.

Το quantitative forecasting ή αλλιώς ποσοτική πρόβλεψη, είναι μια στατιστική μέθοδος που χρησιμοποιείται για να γίνουν προβλέψεις σχετικά με μελλοντικά αριθμητικά δεδομένα, όπως στοιχεία πωλήσεων, επίπεδα αποθέματος ή οικονομικά δεδομένα, με βάση ιστορικά δεδομένα. Αυτός ο τύπος πρόβλεψης βασίζεται σε μαθηματικά μοντέλα και στατιστικές τεχνικές για τον εντοπισμό προτύπων, τάσεων και σχέσεων σε δεδομένα του παρελθόντος και τη χρήση τους για να κάνει προβλέψεις για το μέλλον. Το quantitative forecasting είναι χρήσιμο για επιχειρήσεις και οργανισμούς που θέλουν να προγραμματίσουν το μέλλον και να λάβουν τεκμηριωμένες αποφάσεις με βάση τις αναμενόμενες μελλοντικές τάσεις. Χρησιμοποιείται ευρέως σε διάφορους κλάδους, συμπεριλαμβανομένων των πωλήσεων, του finance, του marketing κ.α.

Υπάρχουν πολλές κοινές μέθοδοι που χρησιμοποιούνται για το quantitative forecasting, όπως η ανάλυση χρονοσειρών, η ανάλυση παλινδρόμησης και οι τεχνικές μηχανικής μάθησης. Η ανάλυση χρονοσειρών περιλαμβάνει τη μοντελοποίηση των ιστορικών δεδομένων ως χρονοσειρές και στη συνέχεια τη χρήση αυτού του μοντέλου για να γίνουν προβλέψεις σχετικά με τις μελλοντικές τιμές. Η ανάλυση παλινδρόμησης περιλαμβάνει τον προσδιορισμό της σχέσης μεταξύ της μεταβλητής που προβλέπεται και άλλων μεταβλητών που μπορεί να την επηρεάσουν, όπως η τιμή ή οι διαφημιστικές δαπάνες. Οι τεχνικές μηχανικής μάθησης χρησιμοποιούν αλγόριθμους για τον εντοπισμό μοτίβων σε ιστορικά δεδομένα και χρησιμοποιούν αυτά τα μοτίβα για να κάνουν προβλέψεις σχετικά με τις μελλοντικές τιμές.

Κατά την εκτέλεση του quantitative forecasting, είναι σημαντικό να χρησιμοποιούμε κατάλληλα δεδομένα και μοντέλα. Τα δεδομένα που χρησιμοποιούνται πρέπει να είναι σχετικά με την προβλεπόμενη μεταβλητή, ακριβή και αντιπροσωπευτικά της χρονικής περιόδου που προβλέπεται. Τα μοντέλα που χρησιμοποιούνται θα πρέπει να είναι κατάλληλα για τα δεδομένα που αναλύονται και τον τύπο της πρόβλεψης που γίνεται. Για παράδειγμα, ένα μοντέλο χρονοσειρών μπορεί να είναι κατάλληλο για την πρόβλεψη των τάσεων των πωλήσεων με την πάροδο του χρόνου, ενώ ένα μοντέλο παλινδρόμησης μπορεί να είναι κατάλληλο για την πρόβλεψη του αντίκτυπου μιας συγκεκριμένης καμπάνιας μάρκετινγκ στις πωλήσεις. Η ποσοτική πρόβλεψη μπορεί να αξιολογηθεί μετρώντας την ακρίβεια των προβλέψεων που δημιουργούνται από τα μοντέλα. Πιο συγκεκριμένα, οι συνήθεις μετρήσεις που χρησιμοποιούνται για την αξιολόγηση περιλαμβάνουν το μέσο απόλυτο σφάλμα (MAE), το μέσο τετραγωνικό σφάλμα (MSE) και το μέσο απόλυτο ποσοστό σφάλματος (MAPE). Αυτά τα μέτρα, τα οποία θα αναλυθούν και περαιτέρω στη συνέχεια της συγκεκριμένης μελέτης, βοηθούν στον προσδιορισμό του πόσο στενά ταιριάζουν οι προβλεπόμενες τιμές με τις πραγματικές τιμές.



Εικόνα 2 "Μέθοδοι Πρόβλεψης"

4.2 Ποσοτικές Μέθοδοι Πρόβλεψης

Σε αυτήν την ενότητα, θα εμβαθύνουμε στις μεθόδους της ποσοτικής πρόβλεψης, έναν στατιστικό τομέα που παρέχει μελλοντικές προβλέψεις με βάση ιστορικά δεδομένα. Θα εξετάσουμε σχολαστικά μια σειρά εξελιγμένων μεθόδων πρόβλεψης, ξεκινώντας με τις κλασικές τεχνικές Moving Averages και Exponential Smoothing, οι οποίες εξομαλύνουν τις ανωμαλίες για να αποκαλύψουν τις υποκείμενες τάσεις. Στη συνέχεια παρουσιάζονται πιο περίπλοκα μοντέλα όπως ARIMA, ARIMAX και SARIMA, τα οποία βελτιώνουν τις προβλέψεις υπολογίζοντας αυτοπαλινδρομικά στοιχεία, εξωτερικές μεταβλητές και εποχιακές διακυμάνσεις. Η μέθοδος intervention analysis θα διερευνηθεί ως μέσο για την αξιολόγηση του αντίκτυπου των εξωτερικών κραδασμών σε μια χρονοσειρά. Θα αναλυθεί επίσης η μέθοδος Unobserved Components Models (UCM), η οποία αναλύει τις χρονοσειρές σε στοιχεία όπως οι τάσεις που πιθανώς να εμφανίζονται σε μια χρονοσειρά.

Moving Averages: Η μέθοδος Moving Averages όροι είναι ένα βασικό στοιχείο της ανάλυσης και της πρόβλεψης χρονοσειρών, που βασίζεται στην αρχή της εξομάλυνσης των βραχυπρόθεσμων διακυμάνσεων για την αποκάλυψη μακροπρόθεσμων τάσεων. Οι κινούμενοι μέσοι όροι (MA) λειτουργούν με τη λήψη μέσου όρου ενός συγκεκριμένου αριθμού προηγούμενων σημείων δεδομένων για την παραγωγή μιας μοναδικής εξόδου ως προβλεπόμενη τιμή. Αυτή η μέθοδος είναι ιδιαίτερα αποτελεσματική για τον εντοπισμό τάσεων στα δεδομένα χρονοσειρών χωρίς να επηρεάζεται αδικαιολόγητα από προσωρινές αιχμές. Το Simple Moving Average (SMA) είναι ο πιο απλός τύπος της μεθόδου, όπου η προβλεπόμενη τιμή υπολογίζεται ως ο μέσος όρος ενός σταθερού αριθμού από τις πιο πρόσφατες παρατηρήσεις. Στον τύπο Weighted Moving Average (SMA) της μεθόδου τα πιο πρόσφατα σημεία δεδομένων έχουν μεγαλύτερη βαρύτητα στον μέσο όρο, αντικατοπτρίζοντας την πεποίθηση ότι οι νεότερες πληροφορίες είναι πιο σχετικές με την πρόβλεψη. Ο τύπος Exponential Moving Average (EMA) εφαρμόζει παράγοντες στάθμισης που μειώνονται εκθετικά για παλαιότερα σημεία δεδομένων, καθιστώντας τον πιο ευαίσθητο στις πρόσφατες αλλαγές στα δεδομένα. Ο τρόπος υπολογισμού των μεθόδων SMA, SMA και EMA εμφανίζονται παρακάτω:

SMA: Πρόσθεση των τιμών των τελευταίων n περιόδων και διαίρεση τους με το n .

$$SMA = (P_1 + P_2 + \dots + P_n) / n$$

Όπου P_1 έως P_n είναι οι παρατηρούμενες τιμές στις περιόδους 1 έως n .

WMA: Πολλαπλασιασμός κάθε σημείου δεδομένων με διαφορετικό βάρος και άθροισμα τους και στη συνέχεια διαίρεση με το άθροισμα των βαρών.

$$WMA = (w_1 \cdot P_1 + w_2 \cdot P_2 + \dots + w_n \cdot P_n) / (w_1 + w_2 + \dots + w_n)$$

EMA: Υπολογίζεται για κάθε σημείο χρησιμοποιώντας το προηγούμενο EMA και την τιμή της τρέχουσας περιόδου, συνυπολογίζοντας μια σταθερά εξομάλυνσης.

$$EMA_{today} = (Value_{today} \cdot (smoothing / (1 + days)) + EMA_{yesterday} \cdot (1 - (smoothing / (1 + days))))$$

Η εφαρμογή της μεθόδου περιλαμβάνει την αναγνώριση τάσης τις αποφάσεις συναλλαγών καθώς και τις προσαρμογές για εποχικότητα. Η μέθοδος Moving Averages προσφέρει ένα προσβάσιμο σημείο εισόδου στην πρόβλεψη, εξισορροπώντας την απλότητα με την αναλυτική χρησιμότητα. Παρέχει μια σαφή εικόνα της κατεύθυνσης της αγοράς για μια συγκεκριμένη περίοδο, αν και πρέπει να χρησιμοποιούνται με σύνεση, με κατανόηση των περιορισμών της και σε συνδυασμό με άλλες μεθόδους ανάλυσης για την επικύρωση των ευρημάτων. Είτε στα οικονομικά, είτε στα χρηματοοικονομικά είτε σε οποιονδήποτε τομέα όπου κυριαρχούν τα δεδομένα χρονοσειρών, η συγκεκριμένη μέθοδος αποτελεί ουσιαστικό εργαλείο για την εξομάλυνση των δεδομένων ώστε να διακρίνονται οι τάσεις και να καθοδηγείται η λήψη αποφάσεων.

Exponential Smoothing - Εκθετική Εξομάλυνση: Η εκθετική εξομάλυνση είναι μια διαχρονική μέθοδος πρόβλεψης που χρησιμοποιείται για την ανάλυση δεδομένων χρονοσειρών. Είναι ιδιαίτερα ευνοημένη για την απλότητα και την αποτελεσματικότητά της στην παραγωγή προβλέψεων που μπορούν να προσαρμοστούν γρήγορα στις αλλαγές των δεδομένων. Σε αντίθεση με την μέθοδο Moving Averages που αναλύθηκε παραπάνω, η οποία αντιμετωπίζει όλες τις παρατηρήσεις εξίσου, η εκθετική εξομάλυνση εκχωρεί εκθετικά μειούμενα βάρη σε προηγούμενες παρατηρήσεις καθώς υποχωρούν στο παρελθόν. Αυτή η μέθοδος εκτιμάται για την ικανότητά της να ενσωματώνει όλα τα δεδομένα του παρελθόντος, ενώ εξακολουθεί να δίνει έμφαση σε πιο πρόσφατες παρατηρήσεις, οι οποίες συχνά είναι πιο ενδεικτικές των μελλοντικών τάσεων. Η βασική αρχή της εκθετικής εξομάλυνσης είναι ότι η πρόβλεψη για την επόμενη περίοδο είναι ένας σταθμισμένος μέσος όρος όλων των προηγούμενων παρατηρήσεων, με τα βάρη να μειώνονται εκθετικά για παλαιότερα σημεία δεδομένων. Ο τύπος ενσωματώνει μια παράμετρο εξομάλυνσης, που συχνά συμβολίζεται ως α (άλφα), η οποία καθορίζει τον ρυθμό με τον οποίο μειώνονται τα βάρη. Η τιμή του α κυμαίνεται μεταξύ 0 και 1 και η επιλογή της τιμής του είναι ένα κρίσιμο μέρος της διαδικασίας μοντελοποίησης, καθώς εξισορροπεί τη βαρύτητα που δίνεται στα πρόσφατα σε σχέση με παλαιότερα δεδομένα. Η απλή εκθετική εξομάλυνση (SES) είναι κατάλληλη για δεδομένα χρονοσειρών χωρίς τάση ή εποχιακό μοτίβο. Η προβλεπόμενη τιμή F_{t+1} για την επόμενη περίοδο είναι ένας σταθμισμένος μέσος όρος της πραγματικής αξίας της τρέχουσας περιόδου A_t και της πρόβλεψη της προηγούμενης περιόδου:

$$F_{t+1} = \alpha A_t + (1 - \alpha) F_t$$

Αυτός ο αναδρομικός τύπος σημαίνει ότι η πρόβλεψη για την επόμενη περίοδο προσαρμόζεται προς την κατεύθυνση του σφάλματος που έγινε στην τρέχουσα περίοδο, με το α να καθορίζει την ταχύτητα προσαρμογής. Εάν το α είναι πιο κοντά στο 1, η πρόβλεψη προσαρμόζεται πιο γρήγορα στις αλλαγές στο επίπεδο της σειράς. Εάν το α είναι πιο κοντά στο 0, η πρόβλεψη ανταποκρίνεται πιο αργά. Όταν τα δεδομένα παρουσιάζουν μια τάση, το μοντέλο SES επεκτείνεται για να το εξηγήσει μέσω της μεθόδου Γραμμικής Τάσης του Holt. Αυτή η προσέγγιση εισάγει μια δεύτερη σταθερά εξομάλυνσης β (βήτα) για να καταγράψει την τάση. Η μέθοδος περιλαμβάνει δύο εξισώσεις: μία για το επίπεδο (παρόμοιο με το SES) και μία άλλη για την τάση. Για δεδομένα με εποχικότητα, εφαρμόζεται η εποχιακή μέθοδος Holt-Winters. Προσθέτει μια τρίτη εξίσωση για να συλλάβει την εποχικότητα και εισάγει μια νέα σταθερά εξομάλυνσης γ (γάμμα). Αυτή η μέθοδος απαιτεί τρεις ξεχωριστές σταθερές εξομάλυνσης: μία για το επίπεδο, μία για την τάση και μία για την εποχικότητα, και παρέχει μια ολοκληρωμένη προσέγγιση για τα δεδομένα πρόβλεψης τόσο

με στοιχεία τάσης όσο και με εποχιακά στοιχεία. Είναι σημαντικό να αναφέρουμε πως στην πράξη, οι σταθερές εξομάλυνσης α , β και γ συχνά βελτιστοποιούνται ελαχιστοποιώντας το μέσο τετραγωνικό σφάλμα των προβλέψεων σε ιστορικά δεδομένα. Αυτή η βελτιστοποίηση είναι ζωτικής σημασίας επειδή η ποιότητα της πρόβλεψης εκθετικής εξομάλυνσης εξαρτάται σε μεγάλο βαθμό από αυτές τις παραμέτρους. Δεν είναι ενιαία, αλλά πρέπει να είναι προσαρμοσμένα στα χαρακτηριστικά της συγκεκριμένης χρονοσειράς.

Παρότι η εκθετική εξομάλυνση είναι μια ισχυρή μέθοδος πρόβλεψης, έχει κάποιους περιορισμούς. Υποθέτει ότι τα μοτίβα στα ιστορικά δεδομένα θα συνεχιστούν και στο μέλλον, κάτι που μπορεί να μην συμβαίνει πάντα. Επιπλέον, η μέθοδος μπορεί να αντιμετωπίσει πολύ ασταθή δεδομένα και μπορεί να είναι ευαίσθητη στην επιλογή της αρχικής πρόβλεψης και των παραμέτρων εξομάλυνσης. Η εκθετική εξομάλυνση γενικά εξισορροπεί την ανάγκη για απλότητα και ανταπόκριση στην ανάλυση χρονοσειρών. Παρέχοντας ένα πλαίσιο που μπορεί να προσαρμοστεί σε διάφορα πρότυπα δεδομένων, από επίπεδο σε τάση και εποχικότητα, προσφέρει στους μετεωρολόγους ένα ευέλικτο εργαλείο. Η προσεκτική ρύθμιση των παραμέτρων του επιτρέπει να παρέχει διαφοροποιημένες προβλέψεις που μπορεί να είναι εξαιρετικά πολύτιμες για τη λήψη αποφάσεων στον έλεγχο των αποθεμάτων, στην πρόβλεψη ζήτησης και σε κάθε τομέα όπου η κατανόηση της μελλοντικής πορείας μιας χρονοσειράς είναι κρίσιμη.

ARIMA: Η ARIMA, που σημαίνει Autoregressive Integrated Moving Average, είναι μια σημαντική στατιστική μέθοδος που χρησιμοποιείται για την πρόβλεψη δεδομένων χρονοσειρών. Είναι ένα μείγμα μοντέλων αυτοπαλίνδρομου και κινούμενου μέσου όρου, ενσωματωμένο με την έννοια της διαφοράς για να διασφαλιστεί η σταθερότητα των δεδομένων. Το μοντέλο ARIMA είναι ιδιαίτερα ισχυρό ως προς την ευελιξία και την ικανότητά του να μοντελοποιεί ένα ευρύ φάσμα δεδομένων χρονοσειρών. Το μοντέλο ARIMA χαρακτηρίζεται από τρεις βασικές παραμέτρους: p , d και q . Το p αποτελεί το τμήμα AR του ARIMA και υποδηλώνει ότι η μεταβλητή ενδιαφέροντος υποχωρεί στις δικές της τιμές με καθυστέρηση (δηλαδή, προηγούμενες). Η παράμετρος p υποδηλώνει τον αριθμό των όρων καθυστέρησης που πρέπει να συμπεριληφθούν στο μοντέλο. Για παράδειγμα, εάν το p είναι 2, η τρέχουσα τιμή της σειράς σχετίζεται με τις δύο τελευταίες τιμές της. Το $d - I$ του μοντέλου - αντιπροσωπεύει τον αριθμό των πράξεων διαφοροποίησης που απαιτούνται για να είναι σταθερά τα δεδομένα χρονοσειράς. Η σταθερότητα είναι μια κρίσιμη πτυχή της ανάλυσης χρονοσειρών και υπονοεί ότι οι στατιστικές ιδιότητες της σειράς (όπως ο μέσος όρος και η διακύμανση) δεν αλλάζουν με την πάροδο του χρόνου. Η διαφοροποίηση βοηθά στην αφαίρεση των τάσεων και των εποχιακών δομών από τη χρονοσειρά, καθιστώντας την πιο σταθερή με την πάροδο του χρόνου. Η παράμετρος d δείχνει ουσιαστικά πόσες φορές πρέπει να διαφοροποιηθούν τα δεδομένα. Η παράμετρος q αποτελεί τη συνιστώσα MA του ARIMA καθώς μοντελοποιεί το σφάλμα του μοντέλου ως γραμμικό συνδυασμό όρων σφάλματος που συνέβησαν ταυτόχρονα και σε διάφορες χρονικές στιγμές στο παρελθόν. Με άλλα λόγια, η παράμετρος q καθορίζει τον αριθμό των καθυστερημένων σφαλμάτων πρόβλεψης στην εξίσωση πρόβλεψης.

Το μοντέλο ARIMA εκφράζεται συχνά ως $ARIMA(p, d, q)$, όπου:

- p είναι η σειρά (αριθμός χρονικών υστερήσεων) του αυτοπαλινδρομικού μοντέλου.
- d είναι ο βαθμός διαφοροποίησης (ο αριθμός των φορών που τα δεδομένα είχαν αφαιρέσει προηγούμενες τιμές).
- q είναι η σειρά του μοντέλου κινούμενου μέσου όρου.

Πιο συγκεκριμένα, το μοντέλο $AR(p)$ αναπαριστά την εξάρτηση μιας παρατήρησης από τις προηγούμενες τιμές της σειράς και έχει τη παρακάτω μορφή:

$$y_t = \phi_1 y_{t-1} + \phi_2 y_{t-2} + \dots + \phi_p y_{t-p} + \epsilon_t$$

όπου:

- y_t είναι η τιμή της χρονοσειράς τη χρονική στιγμή t
- ϕ_i είναι οι παράμετροι του μοντέλου
- ϵ_t είναι ο λευκός θόρυβος (white noise)

Η ολοκλήρωση (differencing) χρησιμοποιείται για να κάνει τη χρονοσειρά στατική (stationary). Αν μια χρονοσειρά δεν είναι στατική, μπορούμε να πάρουμε τις διαφορές των διαδοχικών τιμών μέχρι η σειρά να γίνει στατική. Για παράδειγμα, η πρώτη διαφορά είναι:

$$y_t' = y_t - y_{t-1}$$

και γενικά, η d διαφορά είναι:

$$y_t(d) = (1-B)^d y_t$$

όπου B είναι ο τελεστής καθυστέρησης (lag operator).

Το μοντέλο $MA(q)$ αναπαριστά την εξάρτηση μιας παρατήρησης από τα προηγούμενα σφάλματα (errors). Συγκεκριμένα, ένα μοντέλο $MA(q)$ έχει τη μορφή:

$$y_t = \epsilon_t + \theta_1 \epsilon_{t-1} + \theta_2 \epsilon_{t-2} + \dots + \theta_q \epsilon_{t-q}$$

όπου:

- θ_i είναι οι παράμετροι του μοντέλου
- ϵ_t είναι το λευκός θόρυβος.

Γενικά, το μοντέλο $ARIMA(p, d, q)$ συνδυάζει τα παραπάνω στοιχεία και αναπαριστάται από την εξίσωση:

$$(1 - \phi_1 B - \phi_2 B^2 - \dots - \phi_p B^p)(1 - B)^d y_t = (1 + \theta_1 B + \theta_2 B^2 + \dots + \theta_q B^q) \epsilon_t$$

Σε απλούστερη μορφή:

$$\Phi(B)(1-B)^d y_t = \Theta(B)\epsilon_t$$

Όπου:

- $\Phi(B) = 1 - \phi_1 B - \phi_2 B^2 - \dots - \phi_p B^p$ είναι το πολυώνυμο της αυτοπαλινδρόμησης
- $\Theta(B) = 1 + \theta_1 B + \theta_2 B^2 + \dots + \theta_q B^q$ είναι το πολυώνυμο του κινούμενου μέσου όρου
- B είναι ο τελεστής καθυστέρησης
- d είναι ο βαθμός της ολοκλήρωσης (δηλαδή πόσες φορές έχουμε πάρει διαφορές για να καταστήσουμε τη σειρά στατική).

Ενώ το μοντέλο $ARIMA$ είναι ισχυρό, παρουσιάζει και αυτό κάποιους περιορισμούς. Αρχικά, υποθέτει γραμμικές σχέσεις και μπορεί να μην λειτουργεί καλά με μη γραμμικά δεδομένα. Επίσης, δεν είναι κατάλληλο για χρονοσειρές με υψηλή μεταβλητότητα ή δομικές αλλαγές. Παράλληλα, η επιλογή του σωστού μοντέλου $ARIMA$ μπορεί να είναι περίπλοκη και απαιτεί καλή κατανόηση των υποκείμενων δεδομένων.

ARIMAX: Το $ARIMAX$ (Autoregressive Integrated Moving Average with exogenous inputs) είναι μια επέκταση του μοντέλου $ARIMA$ (Autoregressive Integrated Moving Average), ενσωματώνοντας εξωτερικές μεταβλητές στην εξίσωση πρόβλεψης. Αυτό το προηγμένο μοντέλο είναι ιδιαίτερα χρήσιμο σε σενάρια όπου η χρονοσειρά επηρεάζεται όχι μόνο από τις δικές της προηγούμενες τιμές αλλά και από άλλους εξωτερικούς παράγοντες. Το μοντέλο $ARIMAX$ ενσωματώνει τα ακόλουθα στοιχεία:

AR (Autoregressive): Αυτό το στοιχείο μοντελοποιεί την τρέχουσα τιμή της σειράς ως γραμμικό συνδυασμό των προηγούμενων τιμών της. Η σειρά p καθορίζει τον αριθμό των τιμών με καθυστέρηση που περιλαμβάνονται.

I (Integrated): Η διαφοροποίηση χρησιμοποιείται για να καταστήσει τη χρονοσειρά ακίνητη, δηλαδή για να σταθεροποιήσει τη μέση τιμή της σειράς. Η παράμετρος d υποδεικνύει τον αριθμό των απαιτούμενων διαφοροποιημένων λειτουργιών.

MA (Moving Average): Αυτό το τμήμα αντιπροσωπεύει το σφάλμα του μοντέλου ως συνδυασμό προηγούμενων όρων σφάλματος. Η παράμετρος q καθορίζει τον αριθμό των καθυστερημένων σφαλμάτων πρόβλεψης στην εξίσωση πρόβλεψης.

X (*exogenous inputs*): Πρόκειται για εξωτερικές μεταβλητές που μπορεί να επηρεάσουν τις χρονοσειρές αλλά δεν επηρεάζονται από αυτήν. Τα παραδείγματα περιλαμβάνουν οικονομικούς δείκτες, καιρικές συνθήκες ή εκστρατείες μάρκετινγκ.

Στη μέθοδο ARIMAX, το μοντέλο συνήθως συμβολίζεται ως $ARIMA(p, d, q)X$, όπου το X αντιπροσωπεύει τις εξωτερικές μεταβλητές που περιλαμβάνονται. Η συμπερίληψη του X προσθέτει σημαντική ευελιξία και δυνατότητα εφαρμογής στον πραγματικό κόσμο στο μοντέλο, επιτρέποντάς του να εξετάζει τις επιπτώσεις εκτός της ίδιας της χρονοσειράς. Η διαδικασία τοποθέτησης ενός μοντέλου ARIMAX περιλαμβάνει διάφορα βήματα όπως για παράδειγμα να προσδιοριστεί η κατάλληλη σειρά των συνιστωσών AR, MA και διαφοροποίησης (p, d, q) χρησιμοποιώντας στατιστικά εργαλεία όπως η αυτοσυσχέτιση και τα διαγράμματα μερικής αυτοσυσχέτισης. Επιπλέον, οι σημαντικές εξωτερικές μεταβλητές (X) πρέπει να προσδιορίζονται μέσω διερευνητικής ανάλυσης δεδομένων ή γνώσης τομέα. Επιπλέον, οι συντελεστές του μοντέλου ARIMA και οι συντελεστές για τις εξωτερικές μεταβλητές θα πρέπει να εκτιμώνται χρησιμοποιώντας στατιστικές μεθόδους, κοινώς Εκτίμηση Μέγιστης Πιθανότητας (MLE). Τα υπολείμματα από το προσαρμοσμένο μοντέλο αναλύονται για να ελεγχθεί εάν μοιάζουν με λευκό θόρυβο. Αυτό είναι κρίσιμο για να διασφαλιστεί ότι το μοντέλο έχει συλλάβει επαρκώς τις πληροφορίες στα δεδομένα. Εάν τα διαγνωστικά του μοντέλου είναι ικανοποιητικά, το μοντέλο μπορεί να χρησιμοποιηθεί για την πρόβλεψη μελλοντικών τιμών. Οι προβλέψεις θα ενσωματώσουν την επιρροή τόσο των ιστορικών τιμών των χρονοσειρών όσο και των εξωτερικών μεταβλητών.

Όσον αφορά τους περιορισμούς του μοντέλου αξίζει να αναφέρουμε πως η συμπερίληψη εξωτερικών μεταβλητών ενισχύει την πολυπλοκότητα του μοντέλου. Η επιλογή των σωστών εξωτερικών μεταβλητών και ο σωστός προσδιορισμός του μοντέλου μπορεί σαφώς να είναι δύσκολη.

Το ARIMAX απαιτεί δεδομένα όχι μόνο για την ίδια τη χρονοσειρά αλλά και για τις εξωτερικές μεταβλητές. Παράλληλα, υπάρχει κίνδυνος υπερβολικής προσαρμογής του μοντέλου, ειδικά εάν περιλαμβάνονται πάρα πολλές εξωτερικές μεταβλητές χωρίς επαρκή αιτιολόγηση.

SARIMA: Το SARIMA, το οποίο σημαίνει Εποχιακός Αυτοπαλινδρομικός Ολοκληρωμένος Κινητός Μέσος, είναι μια εξελιγμένη στατιστική μέθοδος που χρησιμοποιείται για την πρόβλεψη δεδομένων εποχιακών χρονοσειρών. Αυτή η μέθοδος επεκτείνει το μοντέλο ARIMA (Autoregressive Integrated Moving Average) προσθέτοντας εποχιακά στοιχεία για να ληφθούν υπόψη τα μοτίβα που επαναλαμβάνονται σε τακτά χρονικά διαστήματα. Το μοντέλο SARIMA συμβολίζεται ως $SARIMA(p, d, q)(P, D, Q)s$. Αυτό αντιπροσωπεύει τόσο τα μη εποχιακά όσο και τα εποχιακά στοιχεία του μοντέλου. Πιο συγκεκριμένα τα p, d, q συμβολίζουν μη εποχιακά στοιχεία και τα P, D, Q τα εποχιακά στοιχεία όπως περιγράφεται παρακάτω:

Μη εποχιακά στοιχεία:

p (*Autoregressive Part*): Αυτό καθορίζει τον αριθμό των παρατηρήσεων καθυστέρησης στο μοντέλο.

d (*Integrated Part*): Αυτό αντιπροσωπεύει τον βαθμό διαφοροποίησης που απαιτείται για να γίνει η σειρά ακίνητη.

q (*Moving Average Part*): Αυτό υποδεικνύει το μέγεθος του κινητού μέσου όρου.

Εποχιακά στοιχεία:

P (*Seasonal Autoregressive Part*): Ο αριθμός των εποχιακών αυτοπαλινδρομικών όρων.

D (*Seasonal Integrated Part*): Ο αριθμός των εποχιακών διαφορών που εφαρμόζονται στη σειρά.

Q (*Seasonal Moving Average Part*): Ο αριθμός των εποχιακών όρων κινητού μέσου όρου.

s - *Seasonality*: Αυτή η παράμετρος αντιπροσωπεύει τη διάρκεια του εποχιακού κύκλου.

Η μέθοδος SARIMA ουσιαστικά μοντελοποιεί τα δεδομένα χρονοσειρών σε δύο επίπεδα: τα μη εποχιακά και τα εποχιακά στοιχεία. Το εποχιακό μέρος του μοντέλου είναι ουσιαστικά ένα μοντέλο ARIMA που έχει προσαρμοστεί στις εποχιακές διαφορές της σειράς. Το πρώτο βήμα της μεθόδου

που είναι η εποχιακή διαφορά (Seasonal Differencing D & s) συχνά περιλαμβάνει τη σταθεροποίηση της σειράς μέσω της εποχιακής διαφοροποίησης, όπου οι παρατηρήσεις αφαιρούνται από προηγούμενες παρατηρήσεις που είναι ένας αριθμός χρονικών βημάτων ίσο με τη διάρκεια της σεζόν σε απόσταση. Στη συνέχεια, το μοντέλο περιλαμβάνει όρους για να ληφθούν υπόψη οι αυτοσυσχετίσεις σε εποχιακές καθυστερήσεις (P) και οι κινητοί μέσοι όροι για τις εποχιακές περιόδους (Q).

Η διαδικασία με την οποία χρησιμοποιείται ένα μοντέλο SARIMA συνήθως περιλαμβάνει διάφορα βήματα. Αρχικά, γίνεται η αναγνώριση του μοντέλου (Model Identification) χρησιμοποιώντας διαγράμματα που έχουν δημιουργηθεί από τα δεδομένα μας (όπως τα διαγράμματα αυτοσυσχέτισης και μερικής αυτοσυσχέτισης) για να επιλέξουμε αρχικές τιμές για τα p , d , q , P , D και Q . Στη συνέχεια ακολουθεί η εκτίμηση των παραμέτρων του μοντέλου χρησιμοποιώντας τεχνικές όπως η εκτίμηση μέγιστης πιθανότητας καθώς επίσης και ο έλεγχος του μοντέλου με τη χρήση διαγνωστικών ελέγχων για να διασφαλιστεί ότι το επιλεγμένο μοντέλο καταγράφει επαρκώς τις πληροφορίες στα δεδομένα. Εάν τα διαγνωστικά του μοντέλου είναι ικανοποιητικά, το μοντέλο μπορεί να χρησιμοποιηθεί για την πραγματοποίηση προβλέψεων.

Συμπερασματικά, το μοντέλο SARIMA είναι ιδιαίτερα χρήσιμο για σύνολα δεδομένων όπου το μοτίβο επαναλαμβάνεται σε τακτά χρονικά διαστήματα, όπως είναι για παράδειγμα τα δεδομένα μηνιαίων πωλήσεων. Ωστόσο αξίζει να αναφέρουμε ότι η συγκεκριμένη μέθοδος παρουσιάζει και κάποιους περιορισμούς. Αρχικά, ο προσδιορισμός της σωστής σειράς του μοντέλου μπορεί να είναι περίπλοκος και μπορεί να απαιτεί ουσιαστική δοκιμή και σφάλμα. Επίσης η μέθοδος SARIMA υποθέτει ότι το εποχιακό μοτίβο θα παραμείνει σταθερό με την πάροδο του χρόνου, κάτι που μπορεί να μην ισχύει για όλα τα σύνολα δεδομένων. Τέλος, για την αποτελεσματική αποτύπωση της εποχικότητας, τα μοντέλα SARIMA απαιτούν συνήθως αρκετούς κύκλους εποχιακών δεδομένων.

Intervention analysis: Η μέθοδος Intervention analysis, γνωστή και ως εκτίμηση επιπτώσεων, είναι μια στατιστική μέθοδος που χρησιμοποιείται για την αξιολόγηση των επιπτώσεων μιας εξωτερικής παρέμβασης σε ένα σύνολο δεδομένων χρονοσειράς. Αυτή η προσέγγιση είναι ιδιαίτερα χρήσιμη σε σενάρια όπου ένα συμβάν ή μια ενέργεια αναμένεται να έχει προκαλέσει σημαντική αλλαγή στο πρότυπο δεδομένων. Οι κοινές εφαρμογές περιλαμβάνουν την ανάλυση του αντίκτυπου των αλλαγών πολιτικής, των μεγάλων οικονομικών γεγονότων, των εκστρατειών μάρκετινγκ ή των φυσικών καταστροφών. Η θεμελιώδης έννοια της συγκεκριμένης μεθόδου είναι να προσδιορίσει εάν ένα εξωτερικό γεγονός έχει αλλάξει σημαντικά το επίπεδο, την τάση ή τη μεταβλητότητα μιας χρονοσειράς και να ποσοτικοποιήσει το μέγεθος και τη διάρκεια αυτής της αλλαγής. Η μέθοδος Intervention analysis έχει τα παρακάτω βασικά χαρακτηριστικά:

- *Δεδομένα χρονοσειρών (Time Series Data):* Το θεμελιώδες στοιχείο της ανάλυσης παρέμβασης είναι μια ακολουθία σημείων δεδομένων που καταγράφονται σε διαδοχικά χρονικά σημεία. Αυτά τα σημεία δεδομένων συχνά εμφανίζουν τάσεις, εποχικότητα και μοτίβα που αναλύονται με την πάροδο του χρόνου.

- *Συμβάν παρέμβασης (Intervention Event):* Πρόκειται για ένα συγκεκριμένο συμβάν ή ενέργεια που πιστεύεται ότι επηρέασε τη χρονοσειρά. Η παρέμβαση θα μπορούσε να είναι μια εφάπαξ εκδήλωση (όπως μια αλλαγή πολιτικής) ή μια πιο παρατεταμένη ενέργεια (όπως μια εκστρατεία μάρκετινγκ).

- *Μοντέλο βάσης (Baseline Model):* Πριν αξιολογήσετε τον αντίκτυπο μιας παρέμβασης, είναι σημαντικό να δημιουργήσετε ένα βασικό μοντέλο. Αυτό το μοντέλο αντιπροσωπεύει πώς θα συμπεριφερόταν η χρονοσειρά αν δεν είχε συμβεί η παρέμβαση. Τα κοινά βασικά μοντέλα περιλαμβάνουν τη μέθοδο ARIMA, η οποία λαμβάνει υπόψη διάφορα μοτίβα, όπως τάσεις και εποχικότητα στα δεδομένα.

- *Μοντέλο παρέμβασης (Intervention Model):* Αυτό το μοντέλο ενσωματώνει το βασικό μοντέλο με το αποτέλεσμα παρέμβασης. Στοχεύει στην ποσοτικοποίηση της αλλαγής στη χρονοσειρά λόγω της παρέμβασης.

- *Μεταβλητή παρέμβασης (Intervention Variable)*: Μια εικονική μεταβλητή που αντιπροσωπεύει την εμφάνιση της παρέμβασης. Συνήθως κωδικοποιείται ως 0 πριν από την παρέμβαση και 1 μετά την εμφάνισή της.

Το μοντέλο έχει την παρακάτω μορφή:

$$Y_t = ARIMA(p, d, q) + \delta D_t + \epsilon_t$$

Ο συντελεστής δ αποτυπώνει το μέγεθος και την κατεύθυνση της επίδρασης της παρέμβασης.

Το πρώτο βήμα είναι η δημιουργία ενός μοντέλου που αντικατοπτρίζει με ακρίβεια τα δεδομένα χρονοσειρών πριν από την παρέμβαση. Αυτό περιλαμβάνει τον εντοπισμό προτύπων, τάσεων και εποχικότητας στα δεδομένα και την τοποθέτηση ενός κατάλληλου μοντέλου, όπως το ARIMA, για να περιγράψει αυτά τα χαρακτηριστικά. Μόλις καθιερωθεί το βασικό μοντέλο, η παρέμβαση μοντελοποιείται. Αυτό γίνεται συχνά προσθέτοντας μια εικονική μεταβλητή στο βασικό μοντέλο. Ο συντελεστής αυτής της μεταβλητής θα υποδεικνύει το μέγεθος και την κατεύθυνση της επίδρασης της παρέμβασης. Πιο αναλυτικά, το προσθετικό αποτέλεσμα είναι ένας τύπος επιπτώσεων παρέμβασης το οποίο υποθέτει ότι η παρέμβαση προκαλεί μια ξαφνική και συνεχή μετατόπιση στο επίπεδο της σειράς. Από την άλλη μεριά, το πολλαπλασιαστικό αποτέλεσμα υποθέτει ότι η παρέμβαση έχει αναλογική επίδραση στη χρονοσειρά, που χρησιμοποιείται συχνά για παρεμβάσεις που επηρεάζουν τον ρυθμό ανάπτυξης ή την τάση. Σε συνέχεια των παραπάνω, ακολουθεί η εκτίμηση του μοντέλου χρησιμοποιώντας στατιστικές τεχνικές (όπως η Εκτίμηση Μέγιστης Πιθανότητας), εκτιμώνται οι παράμετροι του συνδυασμένου μοντέλου. Αυτή η διαδικασία περιλαμβάνει την προσαρμογή του μοντέλου στα πραγματικά δεδομένα και την εκτίμηση του αποτελέσματος της παρέμβασης. Μετά την εκτίμηση, το μοντέλο χρησιμοποιείται για την αξιολόγηση του αντίκτυπου της παρέμβασης. Αυτό περιλαμβάνει τη σύγκριση των δεδομένων μετά την παρέμβαση με τις προβλέψεις που δημιουργούνται από το βασικό μοντέλο. Η διαφορά αποκαλύπτει το αποτέλεσμα της παρέμβασης. Για να εξακριβωθεί η σημασία του αποτελέσματος παρέμβασης, συχνά διεξάγεται έλεγχος υποθέσεων. Αυτός ο έλεγχος καθορίζει εάν οι παρατηρούμενες αλλαγές στις χρονοσειρές μετά την παρέμβαση είναι στατιστικά σημαντικές και όχι μόνο σε τυχαίες διακυμάνσεις.

UCM: Τα μοντέλα Unobserved Components Models (UCM), γνωστά και ως μοντέλα δομικών χρονοσειρών, είναι μια κατηγορία μοντέλων που χρησιμοποιούνται στην ανάλυση και την πρόβλεψη χρονοσειρών. Αυτά τα μοντέλα είναι ιδιαίτερα χρήσιμα για την αποσύνθεση μιας χρονοσειράς στα υποκείμενα συστατικά της, τα οποία δεν παρατηρούνται άμεσα αλλά συνάγονται από τα παρατηρούμενα δεδομένα. Τα κύρια στοιχεία περιλαμβάνουν συνήθως στοιχεία τάσης, εποχιακά, κυκλικά και ακανόνιστα. Οι βασικές έννοιες που πρέπει να γνωρίζουμε για τα μοντέλα UCM είναι τα εξής:

Στοιχείο τάσης (Trend Component): Αντιπροσωπεύει τη μακροπρόθεσμη εξέλιξη της σειράς, καταγράφοντας το υποκείμενο επίπεδο και κατεύθυνση. Συχνά μοντελοποιείται ως ένα ομαλό, αργά εξελισσόμενο στοιχείο, το οποίο μπορεί να είναι γραμμικό ή μη γραμμικό.

Εποχιακό στοιχείο (Seasonal Component): Καταγράφει τακτικές, περιοδικές διακυμάνσεις στη σειρά. Αυτό το στοιχείο μοντελοποιεί μοτίβα που επαναλαμβάνονται σε γνωστά, σταθερά διαστήματα, όπως ημερήσια, μηνιαία ή τριμηνιαία εποχικότητα.

Κυκλική συνιστώσα (Cyclical Component): Λογίζει διακυμάνσεις στα δεδομένα χρονοσειρών που δεν είναι σταθερής συχνότητας. Αυτοί οι κύκλοι είναι συχνά ακανόνιστοι και μπορεί να ποικίλλουν σε διάρκεια, καταγράφοντας συμπεριφορά παρόμοια με τον επιχειρηματικό κύκλο.

Irregular Component: Αντιπροσωπεύει τον όρο τυχαίου θορύβου ή σφάλματος στη σειρά. Ουσιαστικά είναι το μέρος της σειράς που δεν μπορεί να αποδοθεί στην τάση, σε εποχιακά ή κυκλικά στοιχεία.

Το UCM λειτουργεί καθορίζοντας ένα ξεχωριστό μοντέλο για καθένα από αυτά τα στοιχεία και στη συνέχεια συνδυάζοντάς τα για να μοντελοποιήσει και να προβλέψει τις παρατηρούμενες χρονοσειρές. Αρχικά, γίνεται καθορισμός ξεχωριστών μοντέλων για στοιχεία τάσης, εποχιακά, κυκλικά και ακανόνιστα. Αυτά τα μοντέλα μπορεί να ποικίλλουν σε πολυπλοκότητα, από απλές

γραμμικές μορφές έως πιο σύνθετες δομές, όπως τυχαίους περιπάτους ή ομαλές τάσεις με ποικίλες κλίσεις. Τα μοντέλα UCM συχνά διαμορφώνονται σε ένα πλαίσιο κατάστασης χώρου. Σε αυτήν τη ρύθμιση, τα μη παρατηρούμενα στοιχεία αντιμετωπίζονται ως «καταστάσεις» που εξελίσσονται με την πάροδο του χρόνου σύμφωνα με καθορισμένες στοχαστικές διαδικασίες. Η παρατηρούμενη χρονοσειρά είναι τότε ένας συνδυασμός αυτών των καταστάσεων συν έναν όρο θορύβου. Οι παράμετροι του μοντέλου, συμπεριλαμβανομένων εκείνων που καθορίζουν την εξέλιξη των στοιχείων, εκτιμώνται χρησιμοποιώντας τεχνικές όπως το φιλτράρισμα Kalman και η εκτίμηση μέγιστης πιθανότητας. Τα UCM μπορούν να χρησιμοποιηθούν για την πρόβλεψη προβάλλοντας τα στοιχεία προς τα εμπρός στο χρόνο χρησιμοποιώντας τα αντίστοιχα μοντέλα τους. Στη συνέχεια, οι προβλέψεις από κάθε στοιχείο συνδυάζονται για να παραχθεί η τελική πρόβλεψη για τη χρονοσειρά.

Ένα τυπικό UCM για μια χρονολογική σειρά Y_t μπορεί να αναπαρασταθεί ως συνδυασμός των παρακάτω στοιχείων:

$$Y_t = T_t + S_t + C_t + I_t$$

Κάθε στοιχείο μοντελοποιείται χρησιμοποιώντας την αντίστοιχη διαδικασία του, όπως περιγράφεται παραπάνω.

Τα UCM χρησιμοποιούνται για την πρόβλεψη προωθώντας κάθε μοντέλο συστατικού προς τα εμπρός στο χρόνο. Για παράδειγμα, η πρόβλεψη για το επίπεδο της επόμενης περιόδου μπορεί να δοθεί από:

Προβλεπόμενη τάση: $T_{t+1} = T_t + S_t$ για ένα τοπικό γραμμικό μοντέλο τάσης.

Στη συνέχεια, οι προβλέψεις από κάθε στοιχείο συνδυάζονται για να παραχθεί η συνολική πρόβλεψη

$$Y_{t+1} = T_{t+1} + S_{t+1} + C_{t+1}$$

Τα UCM προσφέρουν ένα ισχυρό και ευέλικτο πλαίσιο για ανάλυση χρονοσειρών, ικανό να συλλαμβάνει πολύπλοκα μοτίβα μέσω της αποσύνθεσης της σειράς σε διακριτά στοιχεία. Με τη ρητή μοντελοποίηση αυτών των στοιχείων, τα UCM μπορούν να παρέχουν μια βαθύτερη κατανόηση της δυναμικής των χρονοσειρών και να δημιουργήσουν προβλέψεις που λαμβάνουν υπόψη τις τάσεις, την εποχικότητα, τους κύκλους και τις ακανόνιστες κινήσεις. Ωστόσο, η αποτελεσματική χρήση των UCM απαιτεί προσεκτική εξέταση των προδιαγραφών του μοντέλου και της εκτίμησης παραμέτρων, απαιτώντας συχνά υψηλότερο επίπεδο στατιστικής εμπειρογνομosύνης.

4.3 Ποιοτική Πρόβλεψη

Ένας άλλος όρος στον οποίο αξίζει να αναφερθούμε σε αυτή την μελέτη είναι το Qualitative forecasting. Το Qualitative forecasting είναι ένας τύπος μεθόδου πρόβλεψης που βασίζεται σε γνώμες ειδικών, κρίση και υποκειμενικές ερμηνείες δεδομένων για την πρόβλεψη μελλοντικών τάσεων και αποτελεσμάτων. Αυτή η μέθοδος χρησιμοποιείται συχνά σε καταστάσεις όπου υπάρχουν περιορισμένα ιστορικά δεδομένα ή όπου παράγοντες όπως κοινωνικά, πολιτικά ή οικονομικά γεγονότα είναι δύσκολο να προβλεφθούν. Οι ποιοτικές μέθοδοι πρόβλεψης περιλαμβάνουν τεχνικές όπως έρευνες, ομάδες εστίασης (target groups) και δημοσκοπήσεις από εμπειρογνώμονες. Αυτές οι μέθοδοι χρησιμοποιούνται για τη συλλογή πληροφοριών σχετικά με τις τάσεις της αγοράς, τη συμπεριφορά των καταναλωτών και άλλους παράγοντες που μπορεί να επηρεάσουν τα μελλοντικά αποτελέσματα.

Ένα πλεονέκτημα της ποιοτικής πρόβλεψης είναι ότι μπορεί να συλλάβει πληροφορίες που δεν είναι εύκολα ποσοτικοποιήσιμες, όπως οι προτιμήσεις των καταναλωτών ή οι αλλαγές στα κοινωνικά πρότυπα. Αυτό μπορεί να προσφέρει πολύτιμες πληροφορίες για επιχειρήσεις και οργανισμούς που επιθυμούν να προσαρμοστούν στις μεταβαλλόμενες συνθήκες της αγοράς. Ωστόσο, η ποιοτική πρόβλεψη υπόκειται επίσης σε μεροληψία και μπορεί να επηρεαστεί από τις απόψεις και τις προκαταλήψεις των ειδικών ή των ατόμων που παρέχουν τις πληροφορίες. Ως αποτέλεσμα, είναι σημαντικό να χρησιμοποιούνται πολλαπλές πηγές πληροφοριών και να

αξιολογείται προσεκτικά η αξιοπιστία των δεδομένων κατά τη χρήση ποιοτικών μεθόδων πρόβλεψης. Συνολικά, η ποιοτική πρόβλεψη μπορεί να είναι ένα χρήσιμο εργαλείο για την πρόβλεψη μελλοντικών αποτελεσμάτων όταν οι παραδοσιακές ποσοτικές μέθοδοι μπορεί να μην είναι εφαρμόσιμες ή αξιόπιστες. Με τη συλλογή πληροφοριών από ειδικούς και ενδιαφερόμενους φορείς, οι οργανισμοί μπορούν να προβλέπουν καλύτερα τις αλλαγές και να λαμβάνουν τεκμηριωμένες αποφάσεις σχετικά με τις στρατηγικές και τις λειτουργίες τους.

4.4 Ποιοτικές Μέθοδοι Πρόβλεψης

Μια προσέγγιση για το qualitative forecasting είναι η μέθοδος Delphi. Η μέθοδος Delphi είναι μια διαδικασία που περιλαμβάνει την αναζήτηση πληροφοριών από μια ομάδα ειδικών. Οι ειδικοί είναι συνήθως ανώνυμοι και μπορεί να προέρχονται από διαφορετικούς τομείς ή υπόβαθρα. Η διαδικασία περιλαμβάνει πολλούς γύρους ερωτηματολογίων, στους οποίους ζητείται από αυτούς τους ειδικούς να παράσχουν τις απόψεις ή τις εκτιμήσεις τους για ένα συγκεκριμένο θέμα. Μετά από κάθε γύρο, οι απαντήσεις συγκεντρώνονται και κοινοποιούνται στην επιτροπή και δίνεται η ευκαιρία στους ειδικούς να αναθεωρήσουν τις εκτιμήσεις τους με βάση τα σχόλια. Η διαδικασία συνεχίζεται μέχρι να επιτευχθεί συναίνεση ή έως ότου οι απαντήσεις συγκλίνουν σε ένα στενό εύρος.

Η μέθοδος Delphi έχει πολλά πλεονεκτήματα. Είναι πιο αντικειμενική και αυστηρή από άλλες ποιοτικές μεθόδους πρόβλεψης. Επιτρέπει την ενσωμάτωση διαφορετικών προοπτικών, καθώς η ομάδα των ειδικών μπορεί να προέρχεται από διαφορετικούς τομείς ή υπόβαθρα. Είναι επίσης χρήσιμη σε καταστάσεις όπου υπάρχει υψηλός βαθμός αβεβαιότητας, καθώς οι πολλαπλοί γύροι ερωτηματολογίων και ανατροφοδότησης επιτρέπουν στους ειδικούς να βελτιώσουν τις εκτιμήσεις τους. Ωστόσο, η μέθοδος Delphi μπορεί να είναι χρονοβόρα και δαπανηρή. Ένα πρόβλημα που προκύπτει αν σκεφτούμε τα χαρακτηριστικά αυτού του μοντέλου είναι ότι υπάρχει πάντα η πιθανότητα να υπόκειται σε προκαταλήψεις, όπως η ομαδική σκέψη ή η επιρροή κυρίαρχων προσωπικοτήτων. Οι εμπειρογνώμονες μπορεί επίσης να μην έχουν τις απαραίτητες γνώσεις ή την κατάλληλη εμπειρία για να κάνουν ακριβείς προβλέψεις.

Μια άλλη προσέγγιση για το qualitative forecasting είναι η κριτική επιτροπή της εκτελεστικής μεθόδου (jury of executives). Η κριτική επιτροπή της εκτελεστικής μεθόδου περιλαμβάνει την αναζήτηση πληροφοριών από μια ομάδα στελεχών ή διευθυντών μέσα σε έναν οργανισμό. Η διαδικασία περιλαμβάνει τη σύγκληση μιας ομάδας στελεχών που έχουν διαφορετικά υπόβαθρα και προοπτικές. Το πάνελ συνήθως διευθύνεται από έναν συντονιστή που καθοδηγεί τη συζήτηση και θέτει ερωτήσεις. Ο στόχος της συζήτησης είναι να καταλήξουμε σε συναίνεση για ένα συγκεκριμένο θέμα, όπως μια τάση της αγοράς ή μια ιδέα νέου προϊόντος.

Η κριτική επιτροπή της εκτελεστικής μεθόδου έχει πολλά πλεονεκτήματα. Είναι ταχύτερη και πιο οικονομική από άλλες ποιοτικές μεθόδους πρόβλεψης. Μπορεί επίσης να είναι πιο πρακτική, καθώς τα στελέχη μπορεί να έχουν καλύτερη κατανόηση των λειτουργιών και του περιβάλλοντος του οργανισμού. Ακόμη, είναι πιθανό η συγκεκριμένη μέθοδος να είναι χρήσιμη σε καταστάσεις όπου δεν απαιτείται υψηλός βαθμός ακρίβειας, όπως για βραχυπρόθεσμες προβλέψεις ή επιχειρησιακές αποφάσεις. Ωστόσο, η κριτική επιτροπή της εκτελεστικής μεθόδου μπορεί επίσης να υπόκειται σε προκαταλήψεις, όπως η ομαδική σκέψη ή η επιρροή κυρίαρχων προσωπικοτήτων. Τα στελέχη μπορεί επίσης να μην έχουν την απαραίτητη τεχνογνωσία ή εμπειρία για να κάνουν ακριβείς προβλέψεις όπως και στην περίπτωση της μεθόδου Delphi.

Καταλήγοντας, η επιλογή της ποιοτικής μεθόδου πρόβλεψης εξαρτάται από το συγκεκριμένο πρόβλημα πρόβλεψης και τους διαθέσιμους πόρους. Τόσο η μέθοδος Delphi όσο και η μέθοδος της κριτικής επιτροπής μπορούν να είναι χρήσιμα εργαλεία για την πραγματοποίηση ποιοτικών προβλέψεων. Ωστόσο, θα πρέπει να χρησιμοποιούνται σε συνδυασμό με άλλες μεθόδους πρόβλεψης, όπως το quantitative forecasting ή το scenario planning, για να διασφαλίζονται οι πιο ακριβείς και ολοκληρωμένες προβλέψεις.

5. Η εταιρεία ΟΠΑΠ

5.1 Γνωριμία με τον όμιλο ΟΠΑΠ

Ο Όμιλος ΟΠΑΠ (Οργανισμός Προγνωστικών Αγώνων Ποδοσφαίρου) είναι μια από τις μεγαλύτερες εταιρείες στον χώρο των τυχερών παιχνιδιών στην Ελλάδα. Η εταιρεία ιδρύθηκε το 1958 και έχει εξελιχθεί σε έναν από τους πιο αξιόπιστους και δυναμικούς οργανισμούς της χώρας. Σήμερα, ο ΟΠΑΠ είναι εισηγμένος στο Χρηματιστήριο Αθηνών και είναι μια από τις μεγαλύτερες εταιρείες στην Ελλάδα βάση κεφαλαιοποίησης. Έκτοτε, η εταιρεία επέκτεινε τις δραστηριότητές της σε άλλες χώρες, συμπεριλαμβανομένης της Κύπρου, της Ρουμανίας και της Ιταλίας

Η εταιρεία δραστηριοποιείται στους τομείς των αθλητικών στοιχημάτων, των καζίνο, των λαχειών και των VLTs. Τα προϊόντα της διατίθενται σε χιλιάδες σημεία πώλησης σε όλη τη χώρα αλλά και διαδικτυακά μέσω της ιστοσελίδας του ΟΠΑΠ. Η πλατφόρμα αθλητικών στοιχημάτων της εταιρείας είναι από τις πιο δημοφιλείς στην Ελλάδα και προσφέρει μια μεγάλη γκάμα επιλογών στους πελάτες, συμπεριλαμβανομένων των στοιχημάτων πριν από τον αγώνα και αλλά και κατά την διάρκεια αυτού, σε μια μεγάλη ποικιλία αθλημάτων. Τα κύρια προϊόντα της εταιρείας ωστόσο είναι τα παιχνίδια λοταρίας της, τα οποία είναι δημοφιλή στους πελάτες στην Ελλάδα και σε άλλες χώρες. Μερικά από τα πιο γνωστά τυχερά παιχνίδια του ΟΠΑΠ περιλαμβάνουν το KINO και το Τζόκερ. Το KINO είναι ένα παιχνίδι κλήρωσης αριθμών στο οποίο οι πελάτες επιλέγουν έναν συγκεκριμένο αριθμό αριθμών και εάν αυτοί οι αριθμοί ταιριάζουν με τους αριθμούς που κληρώθηκαν, ο πελάτης κερδίζει ένα χρηματικό έπαθλο. Το Joker είναι ένα παρόμοιο παιχνίδι με έναν επιπλέον αριθμό μπόνους που μπορούν να επιλέξουν οι πελάτες για μια επιπλέον ευκαιρία να κερδίσουν. Ένα ακόμα πολύ σημαντικό παιχνίδι του ομίλου και πολύ κερδοφόρο είναι τα τερματικά βιντεολαχειών (VLTs). Τα VLT είναι ηλεκτρονικά μηχανήματα τυχερών παιγνίων που προσφέρουν μια ποικιλία παιχνιδιών, συμπεριλαμβανομένων των κουλοχέρηδων και του βίντεο πόκερ.

5.2 Εταιρική Υπευθυνότητα & Κοινωνική Ευθύνη

Είναι ζωτικής σημασίας να αναφέρουμε σε αυτό το σημείο ότι ένα από τα βασικά χαρακτηριστικά της λειτουργίας του ΟΠΑΠ είναι η δέσμευσή του στο υπεύθυνο παιχνίδι. Η εταιρεία έχει εφαρμόσει ένα ολοκληρωμένο πρόγραμμα υπεύθυνου τυχερού παιχνιδιού που περιλαμβάνει μέτρα για την πρόληψη του τζόγου ανηλίκων, του προβληματικού τζόγου και του ξεπλύματος χρήματος. Το προσωπικό της εταιρείας είναι εκπαιδευμένο για να εντοπίζει και να βοηθά πελάτες που μπορεί να κινδυνεύουν να αναπτύξουν προβλήματα που σχετίζονται με τον τζόγο.

Ακόμη, εκτός από τις βασικές δραστηριότητες των τυχερών παιχνιδιών και του responsible gaming, ο ΟΠΑΠ συμμετέχει επίσης σε μια σειρά πρωτοβουλιών εταιρικής κοινωνικής ευθύνης. Η εταιρεία υποστηρίζει διάφορα εκπαιδευτικά και πολιτιστικά έργα, καθώς και αθλητικά προγράμματα και περιβαλλοντικές πρωτοβουλίες. Ο Όμιλος ΟΠΑΠ δεσμεύεται στην κοινωνική ευθύνη και έχει υλοποιήσει αρκετές πρωτοβουλίες για την υποστήριξη των τοπικών κοινωνιών στην Ελλάδα και σε άλλες χώρες όπου δραστηριοποιείται. Μία από τις κύριες κοινωνικές δράσεις είναι το πρόγραμμα «OPAP Forword» που ξεκίνησε το 2019, με στόχο την προώθηση της κοινωνικής και οικονομικής ανάπτυξης, με ιδιαίτερη έμφαση στη στήριξη των νέων και των μειονεκτούντων ομάδων. Το πρόγραμμα περιλαμβάνει κοινοτικές πρωτοβουλίες όπως το «OPAP Village», το οποίο χρηματοδοτεί κοινοτικά έργα σε απομακρυσμένα και απομονωμένα χωριά και το «OPAP Playground», που στοχεύει στην προώθηση της σωματικής δραστηριότητας και του υγιεινού τρόπου ζωής στα παιδιά. Το πρόγραμμα «OPAP Skills» παρέχει εκπαίδευση και υποστήριξη σε νέους που αναζητούν εργασία. Επιπρόσθετα, ο ΟΠΑΠ υποστηρίζει άλλες κοινωνικές πρωτοβουλίες όπως η «OPAP Racing Team», η οποία χρηματοδοτεί Έλληνες αθλητές που αγωνίζονται σε διεθνείς διοργανώσεις τριάθλου, και το «OPAP Running», μια σειρά εκδηλώσεων τρεξίματος που προωθούν τη σωματική δραστηριότητα και την ενασχόληση της κοινότητας. Το πρόγραμμα «OPAP Support» παρέχει στήριξη σε ευάλωτες ομάδες της ελληνικής κοινωνίας, συμπεριλαμβανομένων των προσφύγων και των ατόμων με αναπηρία. Η δέσμευση του ΟΠΑΠ για την προώθηση της κοινωνικής και οικονομικής ανάπτυξης, του υγιεινού τρόπου ζωής και της εταιρικής κοινωνικής ευθύνης αποδεικνύει την αφοσίωσή του στο να προσφέρει ανταπόδοση στις τοπικές κοινωνίες.

Ο ΟΠΑΠ έχει επίσης αποδείξει τη δέσμευσή του στην καινοτομία και την ψηφιοποίηση τα τελευταία χρόνια. Η εταιρεία έχει επενδύσει πολλά στις ψηφιακές της πλατφόρμες,

Σελίδα 27 από 67

συμπεριλαμβανομένων του ιστότοπού της και των εφαρμογών για smartphones, προκειμένου να προσφέρει στους πελάτες μια απρόσκοπτη και βολική εμπειρία παιχνιδιού. Η εταιρεία έχει επίσης διερευνήσει τις δυνατότητες των αναδυόμενων τεχνολογιών όπως το blockchain και η εικονική πραγματικότητα στη βιομηχανία τυχερών παιχνιδιών.

5.3 Δομή Ομίλου ΟΠΑΠ

Ο Όμιλος ΟΠΑΠ διαθέτει πάνω από 3.500 σημεία πώλησης σε όλη την Ελλάδα, τα οποία περιλαμβάνουν πρακτορεία, σημεία μικρής λιανικής όπως περίπτερα και μίνι μάρκετ, σουπερ μάρκετ, κ.α. Τα σημεία αυτά προσφέρουν μια ευρεία γκάμα από προϊόντα και υπηρεσίες του ΟΠΑΠ, συμπεριλαμβανομένων των λαχείων, των στοιχημάτων σε αθλητικά γεγονότα, των εικονικών αθλητικών αγώνων, των βιντεοτερματικών λαχείων, των VLTs, καθώς και των διαδικτυακών παιχνιδιών. Τα σημεία πώλησης του ΟΠΑΠ παρέχουν επίσης υπηρεσίες προς τους πελάτες, όπως την πληρωμή κερδών, την ανανέωση στοιχηματικών εισιτηρίων, και την παροχή πληροφοριών σχετικά με τους κανόνες και τις προϋποθέσεις των παιχνιδιών του ΟΠΑΠ.

Η εταιρεία έχει πάνω από 2.500 εργαζόμενους, με την μεγάλη πλειοψηφία αυτών να εργάζεται στην Ελλάδα. Είναι σημαντικό να αναφερθεί πως ο όμιλος ΟΠΑΠ είναι εργοδότης ίσων ευκαιριών και διαθέτει ποικίλο εργατικό δυναμικό που αντικατοπτρίζει τις κοινότητες που εξυπηρετεί. Επιπλέον, η εταιρεία προσφέρει συνεχείς ευκαιρίες εκπαίδευσης και εξέλιξης στους υπαλλήλους της, με έμφαση στην ανάπτυξη των δεξιοτήτων και των ταλέντων του εργατικού δυναμικού της.

Η οργανωτική δομή της εταιρείας είναι ιεραρχική, με διοικητικό συμβούλιο και εκτελεστική ομάδα διαχείρισης υπεύθυνη για τη συνολική στρατηγική και τη λήψη αποφάσεων. Η εκτελεστική ομάδα διαχείρισης αποτελείται από πολλές επιχειρηματικές μονάδες, συμπεριλαμβανομένων των τμημάτων του operations, του product, του technology, του finance και άλλων. Κάθε τμήμα διευθύνεται από ένα ανώτερο στέλεχος που αναφέρεται στον Διευθύνοντα Σύμβουλο.

6. Εφαρμογή Διαδικασίας Πρόβλεψης στον Όμιλο ΟΠΑΠ

Ο Όμιλος ΟΠΑΠ χρησιμοποιεί ποικίλες μεθόδους για την πρόβλεψη των πωλήσεων, όπως έρευνα αγοράς, ανάλυση συμπεριφοράς και τάσεων πελατών και χρηματοοικονομική μοντελοποίηση. Η εταιρεία παρακολουθεί επίσης τα δεδομένα πωλήσεών της σε πραγματικό χρόνο για να προσαρμόσει τις στρατηγικές της και να βελτιστοποιήσει τα έσοδά της. Για παράδειγμα, η εταιρεία μπορεί να αναλύσει δεδομένα πωλήσεων από συγκεκριμένες περιοχές ή δημογραφικές ομάδες για να κατανοήσει καλύτερα τις προτιμήσεις των πελατών και να προσαρμόσει τα προϊόντα και τις υπηρεσίες της σε αυτούς τους πελάτες. Μεγάλο μέρος της πρόβλεψης πωλήσεων που πραγματοποιεί ο οργανισμός βασίζεται σε δεδομένα σημείων ενδιαφέροντος (δημόσιοι χώροι, νοσοκομεία, εστιατόρια κοκ) καθώς και σε πληθυσμιακά δεδομένα. Με την χρήση κατάλληλων μοντέλων και λαμβάνοντας υπόψιν τα παραπάνω δεδομένα καθώς και την κατανομή των υφιστάμενων καταστημάτων υπολογίζεται η δυναμική της περιοχής και κατ' επέκταση ο μέγιστος αριθμός πωλήσεων που μπορεί να επιτύχει το κάθε σημείο.

Για να βελτιώσει περαιτέρω τις δυνατότητες πρόβλεψης και τη συνολική του απόδοση, ο Όμιλος ΟΠΑΠ θα μπορούσε να επενδύσει σε προηγμένα εργαλεία ανάλυσης, όπως η μηχανική μάθηση και η τεχνητή νοημοσύνη. Αυτά τα εργαλεία θα μπορούσαν να βοηθήσουν την εταιρεία να κατανοήσει καλύτερα τη συμπεριφορά και τις προτιμήσεις των πελατών και να βελτιώσει την ακρίβεια πρόβλεψής της. Το πιο δύσκολο κομμάτι στο οποίο θα ήταν ωφέλιμο να επενδύσει ο οργανισμός είναι η αυτόματη πρόβλεψη πωλήσεων ενός μελλοντικού καταστήματος. Στόχος είναι τα πρακτορεία του ομίλου να τοποθετούνται σε σωστά σημεία ώστε να μπορέσουν να αξιοποιήσουν την δυναμική της περιοχής και να πετυχαίνουν μέγιστες πωλήσεις. Για να επιτευχθεί αυτό λοιπόν, υπάρχει η ανάγκη να αναπτυχθεί ένα μοντέλο το οποίο να υπολογίζει τη δυναμική της περιοχής και κατ' επέκταση της δυναμικές πωλήσεις των υφιστάμενων αλλά και με μελλοντικών καταστημάτων.

Για να αντιληφθούμε το μέγεθος της ανάγκης για πρόβλεψη που προκύπτει, αρκεί να αναλογιστούμε ότι τα καθαρά έσοδα προ εισφορών (GGR) αυξήθηκαν κατά 26,0% σε ετήσια βάση, φτάνοντας τα €1.939 εκατομμύρια, σε σύγκριση με τα €1.538,8 εκατομμύρια του 2021. Αυτή η αύξηση προέκυψε από την αυξημένη ζήτηση των παιχνιδιών και τη διευρυμένη πελατειακή

Σελίδα 28 από 67

βάση στον τομέα του διαδικτυακού τυχερού παιχνιδιού. Επιπλέον, το μικτό κέρδος από παιχνίδια αυξήθηκε κατά 31,8% σε ετήσια βάση, φτάνοντας τα €798,8 εκατομμύρια το 2022, σε σύγκριση με τα €605,8 εκατομμύρια του 2021. Αυτή η αύξηση προήλθε από τα αυξημένα έσοδα σε συνδυασμό με υψηλότερα περιθώρια κέρδους. Τα καθαρά κέρδη χρήσης για το 2022 ανήλθαν σε €592,3 εκατομμύρια, σημειώνοντας αύξηση κατά 128,3% σε ετήσια βάση. Επιπλέον, τα καθαρά κέρδη του δ' τριμήνου του 2022 ανήλθαν σε €306,5 εκατομμύρια, αυξημένα κατά 277,6% σε σχέση με το αντίστοιχο διάστημα του 2021, ενσωματώνοντας κέρδος €181,3 εκατομμυρίων από την πώληση της Betano. Αξίζει να σημειωθεί ότι αυτά τα αποτελέσματα επιτεύχθηκαν ενώ διατηρήθηκε ένα χαμηλό επίπεδο δανεισμού, καθιστώντας την οικονομική θέση της εταιρείας ακόμα πιο ενισχυμένη.

6.1 Συλλογή Δεδομένων

Σε αυτή την ενότητα, επικεντρωνόμαστε στη συλλογή και οργάνωση των απαραίτητων δεδομένων για την πρόβλεψη των πωλήσεων του Ομίλου ΟΠΑΠ. Για να εκτελέσουμε μια ακριβή και αξιόπιστη πρόβλεψη, απαιτούνται δεδομένα που καλύπτουν ένα ευρύ χρονικό διάστημα και περιλαμβάνουν πληροφορίες σχετικές με τις μηνιαίες πωλήσεις, τη ζήτηση και την εποχικότητα.

Αρχικά, συλλέξαμε τα δεδομένα των μηνιαίων πωλήσεων για την περίοδο από το 2019 έως το 2021. Αυτά τα δεδομένα αποτελούν τη βάση της ανάλυσης μας και παρέχουν μια ιστορική εικόνα των πωλήσεων της εταιρείας ανά μήνα. Στη συνέχεια αναλύοντας αυτά τα δεδομένα με διάφορες προηγμένες αναλυτικές μεθόδους, εντοπίσαμε τις τάσεις και τα πρότυπα που επηρεάζουν τις αντίστοιχες πωλήσεις.

Τα δεδομένα ληφθήκαν σε μορφή Excel και αποτελούνται από στήλες όπως ο μοναδικός κωδικός του καταστήματος, οι πληροφορίες τοποθεσίας του (πόλη, δήμος, νομός καθώς και οι αντίστοιχες συντεταγμένες τους), μέγεθος καταστήματος, αριθμός προϊόντων vlt, αριθμός προϊόντων ssbt, καθώς και πωλήσεις του καταστήματος. Τα δεδομένα τροποποιήθηκαν κατάλληλα ώστε κάθε εγγραφή του συνόλου δεδομένων να αναφέρεται στις πωλήσεις του καταστήματος κάποιον συγκεκριμένο μήνα ώστε να αυξηθεί σημαντικά ο αριθμός των παρατηρήσεων που θα χρησιμοποιήσουμε στην ανάλυση μας. Επιπλέον, κάθε κατάσταση χαρακτηρίστηκε με βάση τα τετραγωνικά του σε μικρό, μεσαίο ή μεγάλο. Ο συνολικός αριθμός των εγγραφών οι οποίες θα ληφθούν υπόψιν για συγκεκριμένη διαδικασία πρόβλεψης ανέρχονται σε 5.328.

Code	City	Capodestrian	Municipality	Prefecture	Sqm	Size	VLTs	SSBTs	Date	Sales
100154	ΑΓΙΑ ΒΑΡΒΑΡΑ	ΑΓΙΑ ΒΑΡΒΑΡΑ ΑΤΤΙΚΗΣ	ΑΤΤΙΚΗΣ	ΑΤΤΙΚΗΣ	120	Large	11	4	1/1/2019	48.893 €
100154	ΑΓΙΑ ΒΑΡΒΑΡΑ	ΑΓΙΑ ΒΑΡΒΑΡΑ ΑΤΤΙΚΗΣ	ΑΤΤΙΚΗΣ	ΑΤΤΙΚΗΣ	120	Large	11	4	1/2/2019	45.542 €
100154	ΑΓΙΑ ΒΑΡΒΑΡΑ	ΑΓΙΑ ΒΑΡΒΑΡΑ ΑΤΤΙΚΗΣ	ΑΤΤΙΚΗΣ	ΑΤΤΙΚΗΣ	120	Large	11	4	1/3/2019	57.171 €
100154	ΑΓΙΑ ΒΑΡΒΑΡΑ	ΑΓΙΑ ΒΑΡΒΑΡΑ ΑΤΤΙΚΗΣ	ΑΤΤΙΚΗΣ	ΑΤΤΙΚΗΣ	120	Large	11	4	1/4/2019	60.461 €
100154	ΑΓΙΑ ΒΑΡΒΑΡΑ	ΑΓΙΑ ΒΑΡΒΑΡΑ ΑΤΤΙΚΗΣ	ΑΤΤΙΚΗΣ	ΑΤΤΙΚΗΣ	120	Large	11	4	1/5/2019	44.172 €
100154	ΑΓΙΑ ΒΑΡΒΑΡΑ	ΑΓΙΑ ΒΑΡΒΑΡΑ ΑΤΤΙΚΗΣ	ΑΤΤΙΚΗΣ	ΑΤΤΙΚΗΣ	120	Large	11	4	1/6/2019	33.954 €
100154	ΑΓΙΑ ΒΑΡΒΑΡΑ	ΑΓΙΑ ΒΑΡΒΑΡΑ ΑΤΤΙΚΗΣ	ΑΤΤΙΚΗΣ	ΑΤΤΙΚΗΣ	120	Large	11	4	1/7/2019	47.164 €
100154	ΑΓΙΑ ΒΑΡΒΑΡΑ	ΑΓΙΑ ΒΑΡΒΑΡΑ ΑΤΤΙΚΗΣ	ΑΤΤΙΚΗΣ	ΑΤΤΙΚΗΣ	120	Large	11	4	1/8/2019	41.820 €
100154	ΑΓΙΑ ΒΑΡΒΑΡΑ	ΑΓΙΑ ΒΑΡΒΑΡΑ ΑΤΤΙΚΗΣ	ΑΤΤΙΚΗΣ	ΑΤΤΙΚΗΣ	120	Large	11	4	1/9/2019	33.376 €
100154	ΑΓΙΑ ΒΑΡΒΑΡΑ	ΑΓΙΑ ΒΑΡΒΑΡΑ ΑΤΤΙΚΗΣ	ΑΤΤΙΚΗΣ	ΑΤΤΙΚΗΣ	120	Large	11	4	1/10/2019	35.303 €
100154	ΑΓΙΑ ΒΑΡΒΑΡΑ	ΑΓΙΑ ΒΑΡΒΑΡΑ ΑΤΤΙΚΗΣ	ΑΤΤΙΚΗΣ	ΑΤΤΙΚΗΣ	120	Large	11	4	1/11/2019	50.377 €
100154	ΑΓΙΑ ΒΑΡΒΑΡΑ	ΑΓΙΑ ΒΑΡΒΑΡΑ ΑΤΤΙΚΗΣ	ΑΤΤΙΚΗΣ	ΑΤΤΙΚΗΣ	120	Large	11	4	1/12/2019	37.275 €
100154	ΑΓΙΑ ΒΑΡΒΑΡΑ	ΑΓΙΑ ΒΑΡΒΑΡΑ ΑΤΤΙΚΗΣ	ΑΤΤΙΚΗΣ	ΑΤΤΙΚΗΣ	120	Large	11	4	1/1/2020	31.460 €
100154	ΑΓΙΑ ΒΑΡΒΑΡΑ	ΑΓΙΑ ΒΑΡΒΑΡΑ ΑΤΤΙΚΗΣ	ΑΤΤΙΚΗΣ	ΑΤΤΙΚΗΣ	120	Large	11	4	1/2/2020	39.216 €
100154	ΑΓΙΑ ΒΑΡΒΑΡΑ	ΑΓΙΑ ΒΑΡΒΑΡΑ ΑΤΤΙΚΗΣ	ΑΤΤΙΚΗΣ	ΑΤΤΙΚΗΣ	120	Large	11	4	1/3/2020	27.496 €
100154	ΑΓΙΑ ΒΑΡΒΑΡΑ	ΑΓΙΑ ΒΑΡΒΑΡΑ ΑΤΤΙΚΗΣ	ΑΤΤΙΚΗΣ	ΑΤΤΙΚΗΣ	120	Large	11	4	1/4/2020	11 €
100154	ΑΓΙΑ ΒΑΡΒΑΡΑ	ΑΓΙΑ ΒΑΡΒΑΡΑ ΑΤΤΙΚΗΣ	ΑΤΤΙΚΗΣ	ΑΤΤΙΚΗΣ	120	Large	11	4	1/5/2020	22.447 €
100154	ΑΓΙΑ ΒΑΡΒΑΡΑ	ΑΓΙΑ ΒΑΡΒΑΡΑ ΑΤΤΙΚΗΣ	ΑΤΤΙΚΗΣ	ΑΤΤΙΚΗΣ	120	Large	11	4	1/6/2020	49.977 €
100154	ΑΓΙΑ ΒΑΡΒΑΡΑ	ΑΓΙΑ ΒΑΡΒΑΡΑ ΑΤΤΙΚΗΣ	ΑΤΤΙΚΗΣ	ΑΤΤΙΚΗΣ	120	Large	11	4	1/7/2020	42.037 €
100154	ΑΓΙΑ ΒΑΡΒΑΡΑ	ΑΓΙΑ ΒΑΡΒΑΡΑ ΑΤΤΙΚΗΣ	ΑΤΤΙΚΗΣ	ΑΤΤΙΚΗΣ	120	Large	11	4	1/8/2020	27.543 €
100154	ΑΓΙΑ ΒΑΡΒΑΡΑ	ΑΓΙΑ ΒΑΡΒΑΡΑ ΑΤΤΙΚΗΣ	ΑΤΤΙΚΗΣ	ΑΤΤΙΚΗΣ	120	Large	11	4	1/9/2020	29.713 €
100154	ΑΓΙΑ ΒΑΡΒΑΡΑ	ΑΓΙΑ ΒΑΡΒΑΡΑ ΑΤΤΙΚΗΣ	ΑΤΤΙΚΗΣ	ΑΤΤΙΚΗΣ	120	Large	11	4	1/10/2020	34.066 €
100154	ΑΓΙΑ ΒΑΡΒΑΡΑ	ΑΓΙΑ ΒΑΡΒΑΡΑ ΑΤΤΙΚΗΣ	ΑΤΤΙΚΗΣ	ΑΤΤΙΚΗΣ	120	Large	11	4	1/11/2020	6.683 €
100154	ΑΓΙΑ ΒΑΡΒΑΡΑ	ΑΓΙΑ ΒΑΡΒΑΡΑ ΑΤΤΙΚΗΣ	ΑΤΤΙΚΗΣ	ΑΤΤΙΚΗΣ	120	Large	11	4	1/12/2020	- €
100154	ΑΓΙΑ ΒΑΡΒΑΡΑ	ΑΓΙΑ ΒΑΡΒΑΡΑ ΑΤΤΙΚΗΣ	ΑΤΤΙΚΗΣ	ΑΤΤΙΚΗΣ	120	Large	11	4	1/1/2021	- €
100154	ΑΓΙΑ ΒΑΡΒΑΡΑ	ΑΓΙΑ ΒΑΡΒΑΡΑ ΑΤΤΙΚΗΣ	ΑΤΤΙΚΗΣ	ΑΤΤΙΚΗΣ	120	Large	11	4	1/2/2021	39 €
100154	ΑΓΙΑ ΒΑΡΒΑΡΑ	ΑΓΙΑ ΒΑΡΒΑΡΑ ΑΤΤΙΚΗΣ	ΑΤΤΙΚΗΣ	ΑΤΤΙΚΗΣ	120	Large	11	4	1/3/2021	- €
100154	ΑΓΙΑ ΒΑΡΒΑΡΑ	ΑΓΙΑ ΒΑΡΒΑΡΑ ΑΤΤΙΚΗΣ	ΑΤΤΙΚΗΣ	ΑΤΤΙΚΗΣ	120	Large	11	4	1/4/2021	30.402 €

Εικόνα 3 "Δεδομένα οπτικοποίησης και πρόβλεψης"

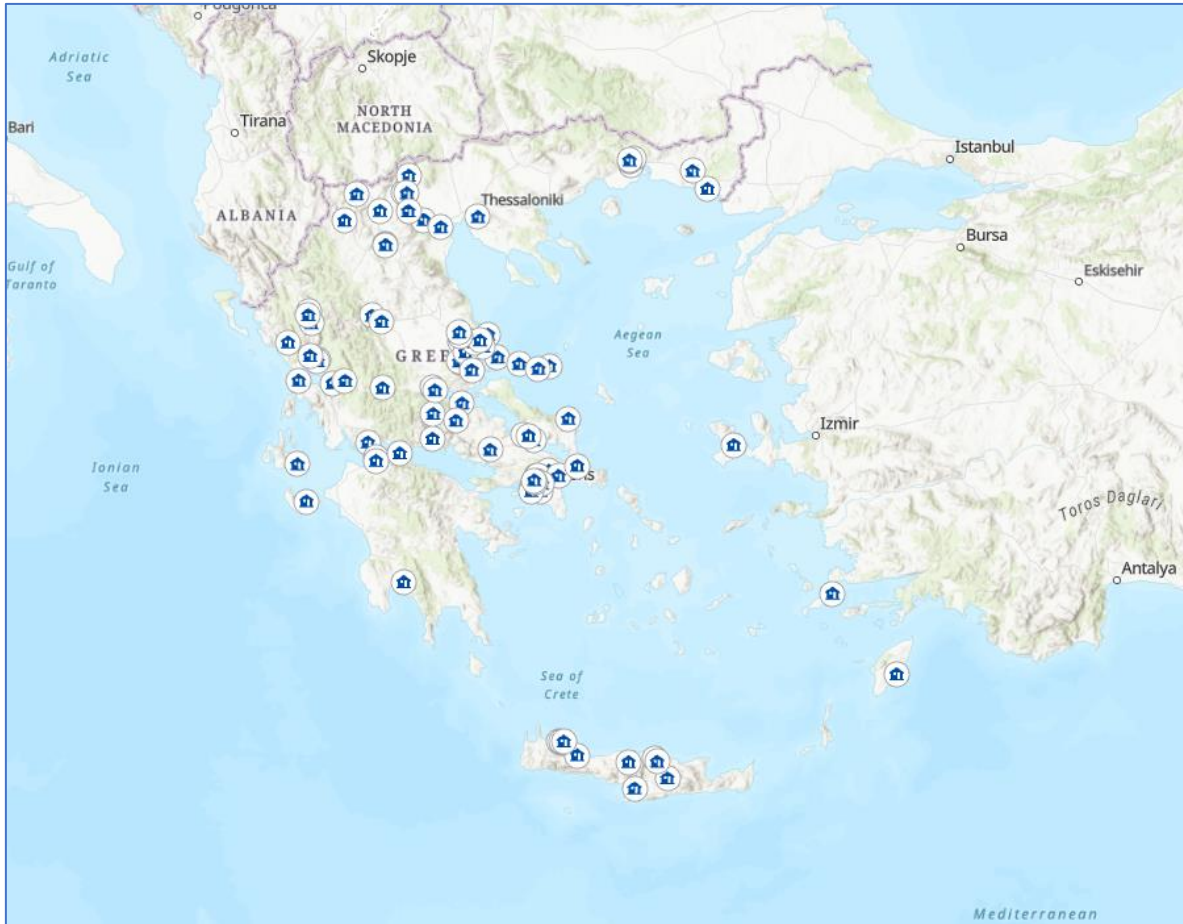
6.2 Οπτικοποίηση Δεδομένων

6.2.1 Γεωχωρική Οπτικοποίηση Δεδομένων

Παράλληλα με τα παραπάνω, και για την εξαγωγή επιπλέον συμπερασμάτων αλλά και για την επαλήθευση των αποτελεσμάτων χρησιμοποιήθηκαν και επιπλέον δεδομένα όπως σημεία ενδιαφέροντος (τράπεζες, εστιατόρια, δημόσιοι χώροι κ.α.) και πληθυσμιακά δεδομένα. Για να γίνει μια πρώτη αποτύπωση και εξοικείωση με τα δεδομένα, οπτικοποιήθηκαν στο γεωχωρικό λογισμικό ArcGIS. Κατά αυτό τον τρόπο, φαίνεται με γρήγορη ματιά αν τα καταστήματα είναι στημένα σε εμπορικά σημεία, καθώς και πόσο κόσμο δυνητικά θα μπορούσαν να εξυπηρετήσουν.

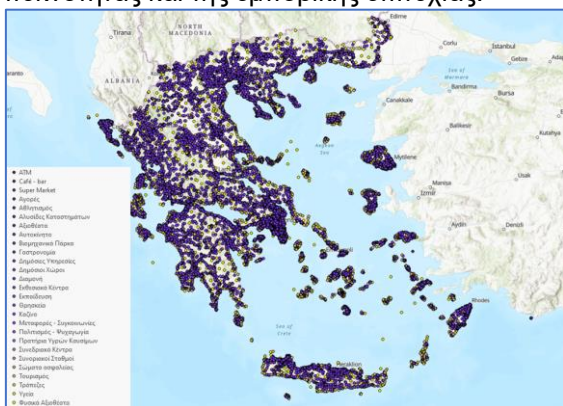
Οι διάφορες οπτικοποιήσεις που έγιναν παρουσιάζονται παρακάτω:

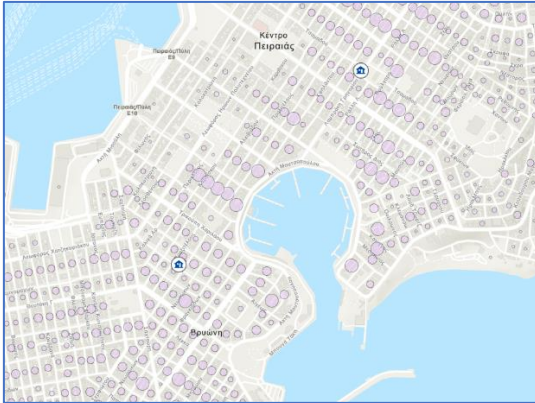
Ο παρακάτω χάρτης χρησιμεύει ως οπτική αναπαράσταση του κεντρικού σημείου της έρευνάς μας, δηλαδή, των διαφόρων καταστημάτων που θα εξεταστούν και θα αναλυθούν. Κάθε σημείο στο χάρτη αντιστοιχεί σε ένα συγκεκριμένο κατάστημα που έχει ελεγχθεί για διάφορες παραμέτρους, που κυμαίνονται από τις τάσεις της αγοράς και τη συμπεριφορά των καταναλωτών έως τη λειτουργική αποτελεσματικότητα. Η γεωγραφική διασπορά αυτών των καταστημάτων προσφέρει μια ολοκληρωμένη επισκόπηση της χωρικής κατανομής τους και βοηθά στην αποσαφήνιση μοτίβων και συσχετίσεων που μπορεί να είναι κρίσιμα για την κατανόηση του ευρύτερου τοπίου λιανικής.



Εικόνα 4 "Υπό εξέταση καταστήματα"

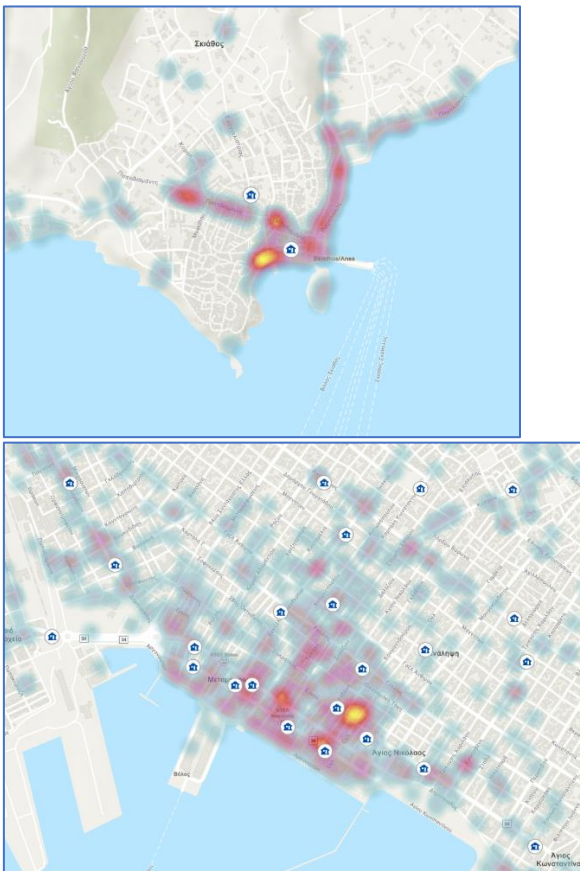
Τα σημεία ενδιαφέροντος, στα οποία αναφερθήκαμε και παραπάνω παρουσιάζουν που συγκεντρώνεται η εμπορικότητα σε μια περιοχή άρα και κατ' επέκταση που είναι πιθανό τα καταστήματα μας να έχουν μεγαλύτερες πωλήσεις. Ανάλογα με την κατηγορία στην οποία ανήκει κάθε σημείο ενδιαφέροντος παρουσιάζεται και με διαφορετικό χρώμα όπως φαίνεται στο υπόμνημα του χάρτη (1) παρακάτω. Επιπλέον, ο 2ος χάρτης που παρουσιάζεται παρακάτω ενσωματώνει μια κεντρική πτυχή της έρευνάς μας, απεικονίζοντας την κατανομή του πληθυσμού κοντά σε κάθε υπό εξέταση κατάστημα. Τα διαφορετικά μεγέθη σημείων που εμφανίζονται στον χάρτη μας δείχνουν ποιος είναι ο πληθυσμός κάθε οικοδομικού τετραγώνου. Ως θεμελιώδης αρχή, όσο μεγαλύτερος είναι ο κοντινός πληθυσμός, τόσο υψηλότερες είναι οι αναμενόμενες πωλήσεις για κάθε κατάστημα, υπογραμμίζοντας την κρίσιμη σχέση μεταξύ της δημογραφικής πυκνότητας και της εμπορικής επιτυχίας.





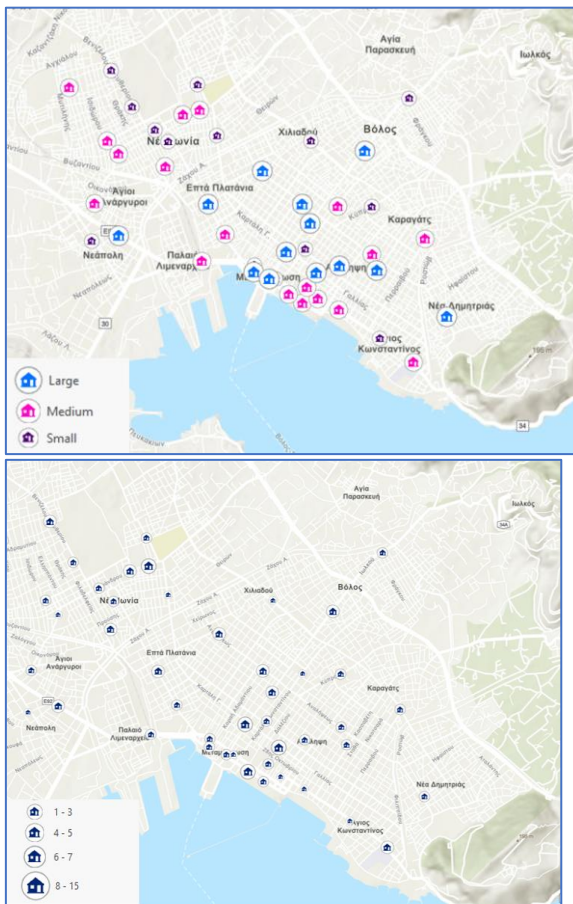
Εικόνα 5 "Συγκέντρωση εμπορικότητας"

Οι παρακάτω χάρτες χρησιμοποιούν έναν δυναμικό και οπτικά ενημερωτικό χάρτη θερμότητας (heat map) για να απεικονίσουν την εμπορικότητα των περιοχών που ερευνήθηκαν, λαμβάνοντας υπόψη τόσο την πυκνότητα του πληθυσμού όσο και την κατανομή των Σημείων Ενδιαφέροντος (POI). Οι χρωματικές διαβαθμίσεις στους χάρτες επικοινωνούν αποτελεσματικά την ένταση της εμπορικής δραστηριότητας, με θερμότερες αποχρώσεις που υποδηλώνουν υψηλότερη εμπορική πυκνότητα. Αυτή η οπτικοποίηση επιτρέπει την κατανόηση των περιοχών όπου η συγκέντρωση πληθυσμού συγκλίνει με μια ποικιλία σημείων ενδιαφέροντος, δηλώνοντας έτσι πιο εμπορικά ενεργά σημεία. Οι προβλέψεις δείχνουν μια θετική συσχέτιση, υποδηλώνοντας ότι τοποθεσίες που χαρακτηρίζονται από μια ζωντανή αλληλεπίδραση και των δύο αυτών παραγόντων είναι πιθανό να παρουσιάσουν υψηλότερες πωλήσεις για τα εξεταζόμενα καταστήματα.



Εικόνα 6 "Συγκέντρωση εμπορικότητας με χρήση heat map"

Στους παρακάτω χάρτες, τα υπό εξέταση καταστήματα παρουσιάζονται με διάφορα μεγέθη και χρώματα, καθένα από τα οποία είναι ενδεικτικό των διακριτών χαρακτηριστικών που σχετίζονται με την μελέτη μας. Η σκόπιμη χρήση διαφορετικών οπτικών στοιχείων επιτρέπει μια διαφοροποιημένη σύγκριση μεταξύ των καταστημάτων, ενισχύοντας τη βαθύτερη κατανόηση του τρόπου με τον οποίο συγκεκριμένα χαρακτηριστικά επηρεάζουν τη συνολική απόδοση του καταστήματος. Τα διαφορετικά μεγέθη των δεικτών στον χάρτη χρησιμεύουν ως απτή αντανάκλαση των ποσοτικών διαφορών μεταξύ των καταστημάτων (μέγεθος καταστήματος, αριθμός μηχανών κ.α.), ενώ τα ποικίλα χρώματα τονίζουν ξεχωριστά ποιοτικά χαρακτηριστικά. Μέσω αυτής της οπτικής αναπαράστασης, στοχεύουμε να διακρίνουμε μοτίβα και συσχετίσεις, εξάγοντας ουσιαστικά συμπεράσματα σχετικά με το πώς τα συγκεκριμένα χαρακτηριστικά, είτε πρόκειται για μέγεθος, τοποθεσία ή άλλους σχετικούς παράγοντες, συμβάλλουν ή επηρεάζουν τη συνολική απόδοση κάθε καταστήματος.



Εικόνα 7 "Οπτικοποίηση καταστημάτων με βάση τα χαρακτηριστικά τους"

6.2.2 Εξερεύνηση των Δεδομένων μέσω Οπτικοποίησης

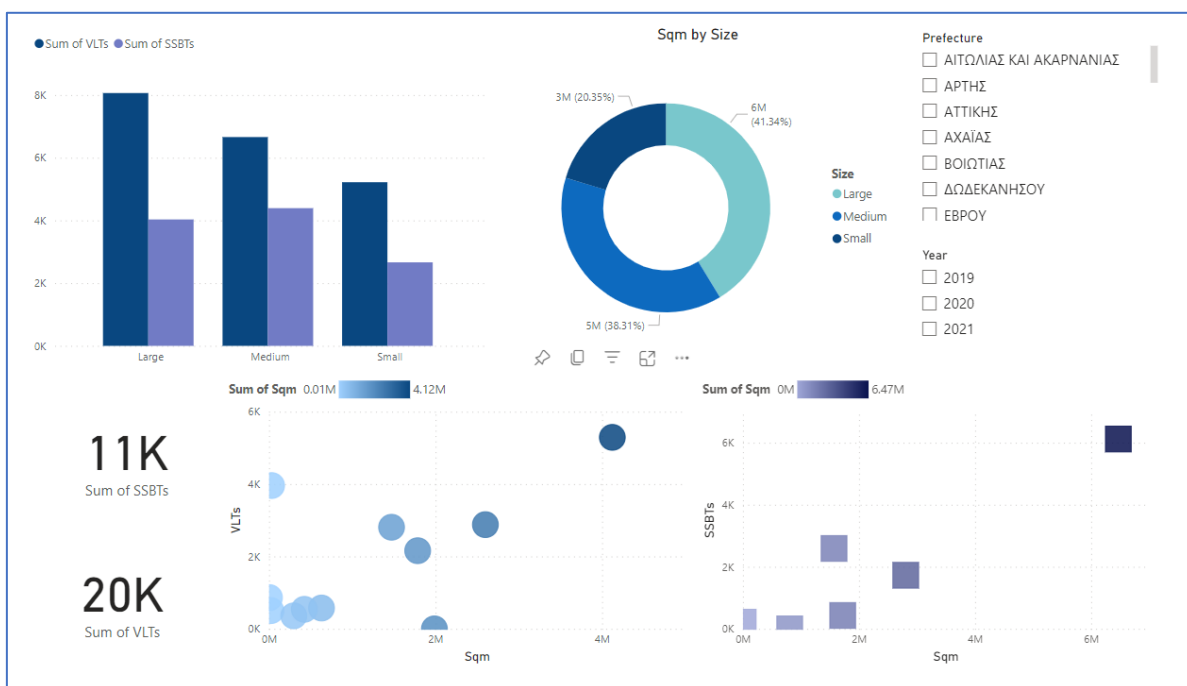
Αυτή η ενότητα της ανάλυσης, είναι αφιερωμένη στην εξέταση και τον προσδιορισμό των συσχετίσεων εντός του συνόλου δεδομένων μέσω μιας σειράς περιεκτικών απεικονίσεων. Για να το πετύχουμε αυτό, χρησιμοποιήσαμε το λογισμικό Microsoft Power BI, ένα ισχυρό και ευέλικτο εργαλείο επιχειρηματικής ανάλυσης. Το Power BI έχει σχεδιαστεί για να επιτρέπει στους χρήστες να συνδέονται με ένα ευρύ φάσμα πηγών δεδομένων, μετατρέποντας τα ανεπεξέργαστα δεδομένα σε ουσιαστικές πληροφορίες μέσω διαδραστικών αναφορών και πινάκων εργαλείων. Χρησιμοποιώντας το Power BI, αξιοποιούνται οι δυνατότητες μοντελοποίησης δεδομένων για τον εντοπισμό και την απεικόνιση συσχετίσεων στα δεδομένα. Αυτό το εργαλείο υπερέρχει στην ικανότητά του να δημιουργεί μια σειρά οπτικοποιήσεων—από βασικά γραφήματα ράβδων και γραμμών έως πολύπλοκα διαγράμματα διασποράς και χάρτες—που μας δίνει τη δυνατότητα να

Σελίδα 33 από 67

εξερευνούμε και να παρουσιάζουμε δεδομένα με τρόπο προσβάσιμο και κατατοπιστικό. Τα χαρακτηριστικά δυναμικής ανάλυσης δεδομένων μας επιτρέπουν να εμβαθύνουμε στις περιπλοκές του συνόλου δεδομένων μας, αποκαλύπτοντας μοτίβα και σχέσεις που διαφορετικά θα μπορούσαν να παραμείνουν κρυφές.

Κατά τη διάρκεια της ανάλυσης των δεδομένων μας, αναπτύχθηκαν 2 dashboards με τη χρήση του Power BI για τη λεπτομερή εξέταση του συνόλου δεδομένων. Αυτοί οι πίνακες έχουν σχεδιαστεί σχολαστικά ώστε όχι μόνο να παρουσιάζουν τα δεδομένα αλλά και να φωτίζουν τα υποκείμενα μοτίβα και τους συσχετισμούς που ενδέχεται να συμβάλλουν στη διαδικασία λήψης αποφάσεων.

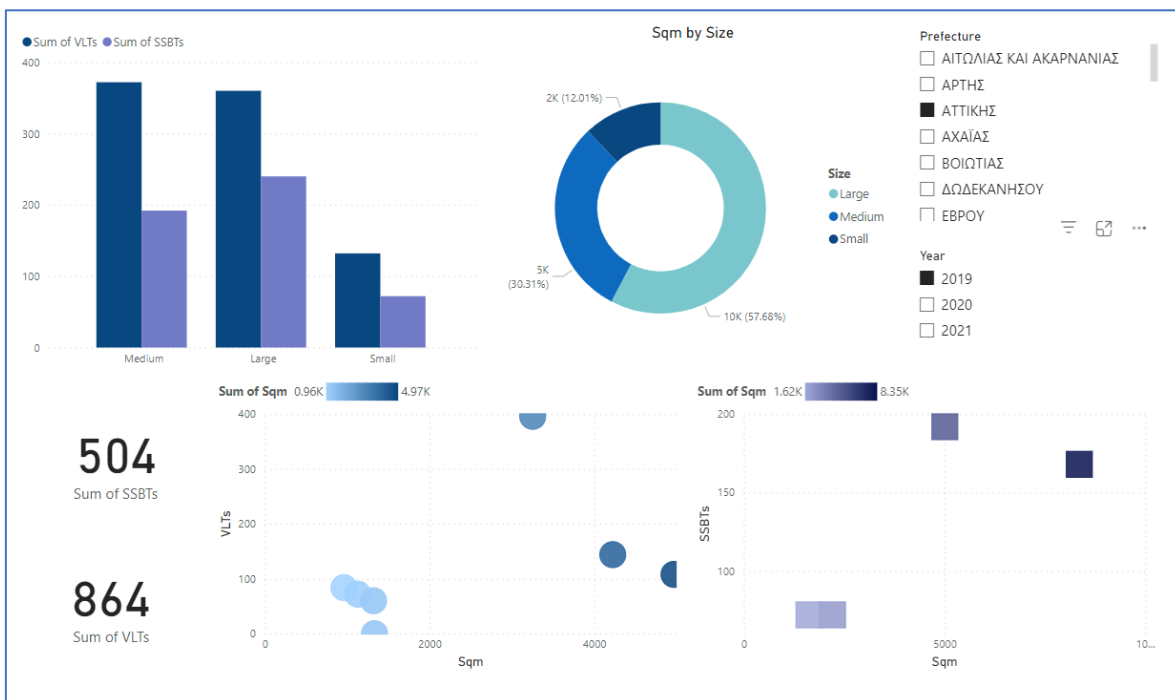
Ο πρώτος πίνακας εστιάζει στην παροχή μιας επισκόπησης υψηλού επιπέδου του συνόλου δεδομένων. Είναι δομημένος έτσι ώστε να παρουσιάζει βασικές μετρήσεις με μια ματιά, επιτρέποντας μια γρήγορη αλλά ολοκληρωμένη κατανόηση του συνολικού τοπίου δεδομένων. Όπως αναφέρθηκε και παραπάνω, για τη δημιουργία του μοντέλου πρόβλεψης τα δεδομένα μας έχουν χωριστεί σε 3 κατηγορίες ανάλογα με την εγγραφή τους στο πεδίο «τετραγωνικά». Ο συγκεκριμένος διαχωρισμός των δεδομένων έχει χρησιμοποιηθεί σε αυτόν τον πρώτο πίνακα ως βάση για την απεικόνιση και των υπόλοιπων δεδομένων. Πιο συγκεκριμένα, σε αυτόν τον πίνακα παρουσιάζονται το σύνολο των μηχανών vlts & ssbts ανά τύπο μαγαζιού (μικρό, μεσαίο, μεγάλο) καθώς και το σύνολο των τετραγωνικών που αντιστοιχούν σε κάθε κατηγορία. Επιπλέον, αναφέρεται ο συνολικός αριθμός μηχανών vlts & ssbts καθώς και η συσχέτιση που υπάρχει μεταξύ των πεδίων τετραγωνικά – μηχανές vlts & τετραγωνικά – μηχανές ssbts αντίστοιχα. Αυτό που αξίζει να αναφέρουμε είναι πως η διαδραστικότητα που προσφέρει το συγκεκριμένο λογισμικό παίζει κρίσιμο ρόλο στην ανάλυσή μας. Μας επιτρέπει να εμβαθύνουμε σε συγκεκριμένες πτυχές των δεδομένων μας, εξερευνώντας με ευκολία τις διαφορετικές διαστάσεις που εμφανίζουν. Ο πρώτος πίνακας παρουσιάζεται παρακάτω:



Εικόνα 8 "Dashboard 1"

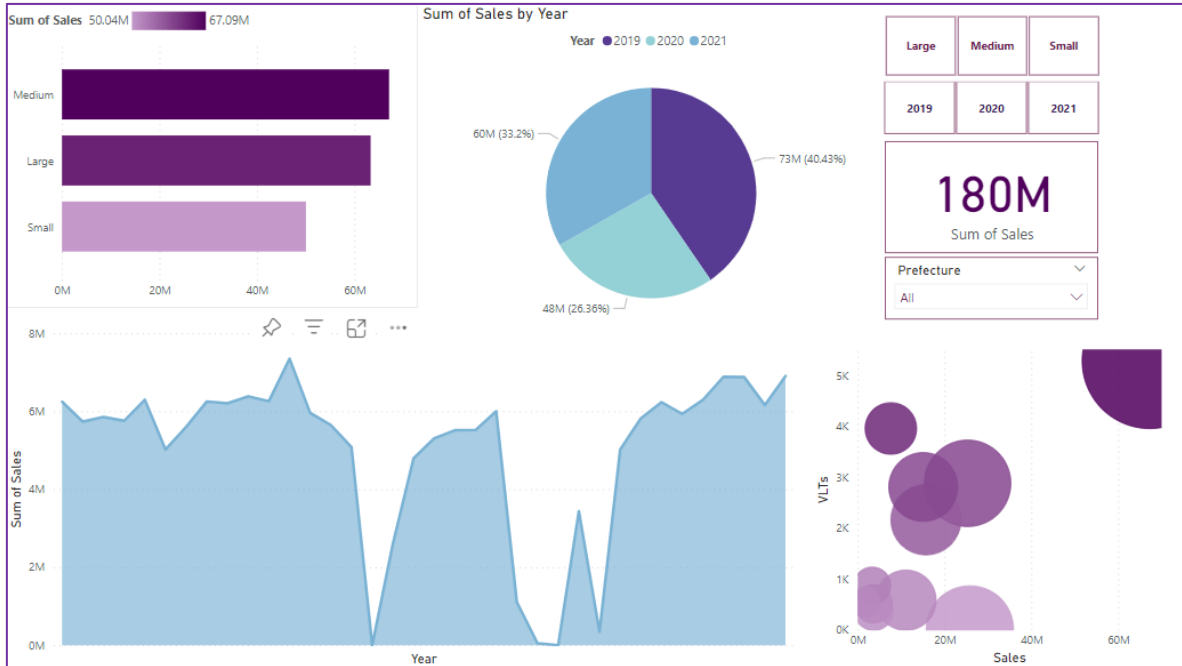
Παρατηρώντας τον παραπάνω πίνακα μπορούμε εύκολα να διακρίνουμε ότι καταστήματα που ανήκουν στην κατηγορία "Large" διαθέτουν μεγαλύτερο συνολικό αριθμό μηχανών vlts καθώς - και όπως είναι λογικό- μεγαλύτερο συνολικό αριθμό τετραγωνικών. Η θετική συσχέτιση που υπάρχει μεταξύ των τετραγωνικών και τον μηχανών vlts & ssbts αντίστοιχα εμφανίζεται στα γραφήματα στο κάτω μέρος του dashboard. Ένα σημαντικό χαρακτηριστικό του πίνακα είναι η δυνατότητα προβολής αποτελεσμάτων φιλτραρισμένα ανά συγκεκριμένες γεωγραφικές περιοχές

ή συγκεκριμένα έτη. Αυτή η λειτουργία μας επιτρέπει να προσαρμόσουμε την οπτικοποίηση δεδομένων στα συγκεκριμένα ενδιαφέροντα ή τις εκάστοτε ανάγκες μας, παρέχοντας μια πιο στοχευμένη και σχετική εικόνα. Για παράδειγμα, επιλέγοντας μια συγκεκριμένη γεωγραφία, μπορούμε να διακρίνουμε τάσεις και μοτίβα δεδομένων μοναδικά σε αυτήν την περιοχή, επιτρέποντας μια τοπική ανάλυση που μπορεί να αποκαλύψει περιφερειακές ανισότητες ή μοναδικά χαρακτηριστικά. Ομοίως, το φιλτράρισμα των δεδομένων κατά ένα συγκεκριμένο έτος επιτρέπει μια χρονική ανάλυση, προσφέροντας πληροφορίες για το πώς έχουν εξελιχθεί οι τάσεις με την πάροδο του χρόνου. Αυτό το επίπεδο ευκρίνειας στα δεδομένα όχι μόνο ενισχύει την ακρίβεια της ανάλυσης μας, αλλά παρέχει επίσης μια πιο ολοκληρωμένη κατανόηση της δυναμικής που παίζει, διασφαλίζοντας ότι τα συμπεράσματά μας είναι τόσο σχετικά με τα συμφραζόμενα όσο και χρονικά ακριβή.



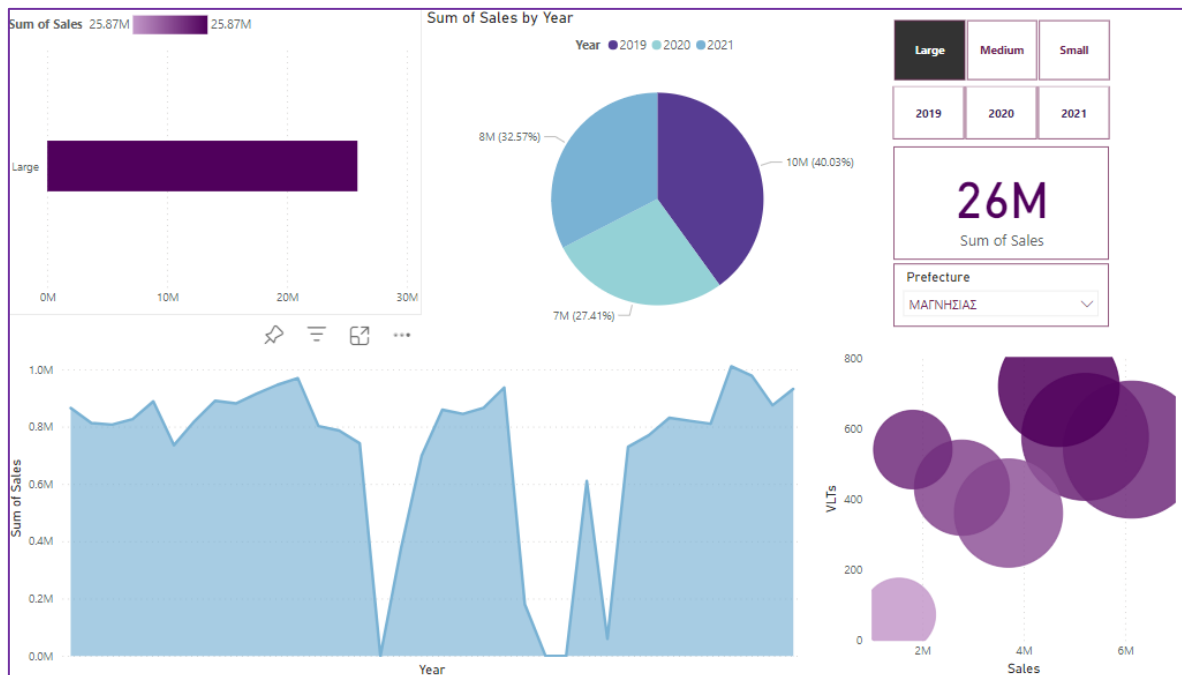
Εικόνα 9 "Αποτελέσματα πίνακα για την Αττική το 2019"

Το δεύτερο dashboard αντιπροσωπεύει μια ολοκληρωμένη αναλυτική προβολή των δεδομένων πωλήσεων, υπογραμμίζοντας τις συσχετίσεις μεταξύ των πωλήσεων και άλλων σχετικών μετρήσεων. Αυτός ο πίνακας είναι μια ποπτικοποίηση δεδομένων που παρέχει μια πολυδιάστατη ανάλυση των αριθμών πωλήσεων.



Εικόνα 10 "Dashboard 2"

Επάνω αριστερά, βλέπουμε ένα γράφημα ράβδων που αναλύει το άθροισμα των πωλήσεων ανά κατηγορία μεγέθους—μικρό, μεσαίο και μεγάλο—παρέχοντας μια σαφή οπτική αναπαράσταση του ποιο είδος καταστήματος συνεισφέρει περισσότερο στον συνολικό όγκο πωλήσεων. Αυτή η κατηγοριοποίηση μας επιτρέπει να διακρίνουμε αμέσως την κατανομή των πωλήσεων σε διαφορετικές κατηγορίες μεγεθών, παρέχοντας μια γρήγορη αξιολόγηση της απόδοσης της αγοράς τμηματοποιημένη ανά τύπο καταστήματος. Δίπλα στο γράφημα ράβδων υπάρχει ένα γράφημα πίτας που οριοθετεί το άθροισμα των πωλήσεων ανά έτος, με κάθε κομμάτι να αντιπροσωπεύει ένα έτος από το 2019 έως το 2021. Αυτή η χρονική ανάλυση διευκολύνει την κατανόηση της απόδοσης των πωλήσεων από έτος σε έτος, επιτρέποντας την παρακολούθηση της ανάπτυξης. Η κεντρική περιοχή του πίνακα διαθέτει ένα γραμμικό γράφημα, που δείχνει τις διακυμάνσεις στις πωλήσεις σε μια συνεχή περίοδο. Αυτή η οπτικοποίηση είναι το κλειδί για τον εντοπισμό προτύπων όπως για παράδειγμα οι εποχιακές τάσεις. Ο δεύτερος πίνακας, όπως και ο πρώτος, είναι διαδραστικός και επιτρέπει μια δυναμική εξερεύνηση των δεδομένων πωλήσεων σε διάφορες διαστάσεις, συμπεριλαμβανομένου του έτους, της γεωγραφίας και του μεγέθους του καταστήματος. Συνολικά, αυτός ο δεύτερος πίνακας χρησιμεύει ως στρατηγικό εργαλείο ανάλυσης, παρέχοντας μια πλούσια, διαδραστική και λεπτομερή εξερεύνηση των δεδομένων πωλήσεων. Μας επιτρέπει να αποκαλύψουμε συσχετισμούς μεταξύ πωλήσεων και διάφορων παραγόντων, διευκολύνοντας τη λήψη αποφάσεων βάσει δεδομένων και τη διαμόρφωση στρατηγικής.

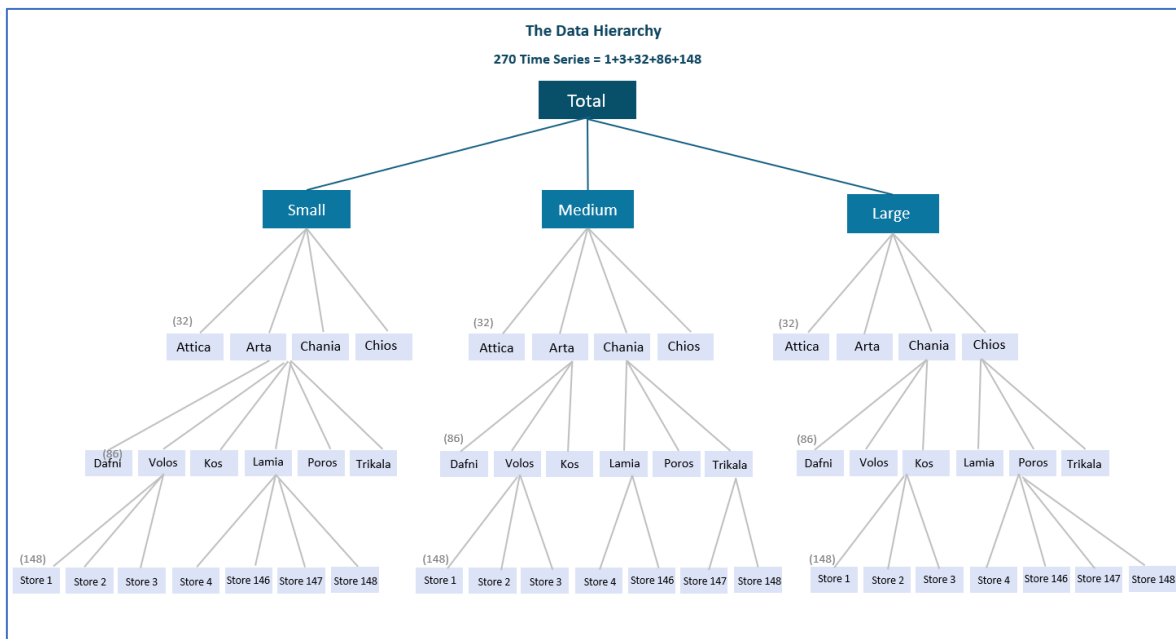


Εικόνα 11 "Αποτελέσματα πωλήσεων για τα Large καταστήματα της Μαγνησίας"

6.3 Εφαρμογή Πρόβλεψης

Σε αυτή την ενότητα, θα εμβαθύνουμε στη διαδικασία της πρόβλεψης χρησιμοποιώντας τη μέθοδο του Large Scale Hierarchical Forecasting. Αυτή η προηγμένη προσέγγιση είναι ιδιαίτερα κατάλληλη για σενάρια που περιλαμβάνουν πολύπλοκα σύνολα δεδομένων που εμφανίζουν ιεραρχικές δομές. Το Large Scale Hierarchical Forecasting είναι μια καινοτόμος μέθοδος που χειρίζεται και προβλέπει αποτελεσματικά δεδομένα σε διάφορα επίπεδα συγκέντρωσης εντός μιας ιεραρχικής δομής.

Το Large Scale Hierarchical Forecasting έχει σχεδιαστεί για να αντιμετωπίσει τις προκλήσεις που τίθενται από μεγάλα σύνολα δεδομένων που έχουν μια φυσική ιεραρχική ή ομαδοποιημένη δομή. Σε πολλά σενάρια πραγματικού κόσμου, τα δεδομένα οργανώνονται σε πολλαπλά επίπεδα – για παράδειγμα, μια εταιρεία λιανικής μπορεί να έχει δεδομένα πωλήσεων σε επίπεδο προϊόντος, καταστήματος, περιφερειακού και εθνικού επιπέδου. Κάθε ένα από αυτά τα επίπεδα αντιπροσωπεύει μια διαφορετική ευαισθησία των δεδομένων, από το πιο λεπτομερές έως το πιο συγκεντρωτικό. Αυτή η προσέγγιση εκμεταλλεύεται τις ιεραρχικές σχέσεις εντός των δεδομένων. Η ιεραρχία μπορεί να οπτικοποιηθεί ως δέντρο, όπου ο επάνω κόμβος αντιπροσωπεύει το σύνολο (όπως οι συνολικές εθνικές πωλήσεις) και οι χαμηλότεροι κόμβοι αντιπροσωπεύουν χωριστά δεδομένα (όπως οι τοπικές πωλήσεις και οι πωλήσεις καταστήματος). Οι χαμηλότεροι κόμβοι ενδέχεται να αντιπροσωπεύουν τα πιο αναλυτικά δεδομένα. Το Large Scale Hierarchical Forecasting περιλαμβάνει τη δημιουργία προβλέψεων σε κάθε επίπεδο της ιεραρχίας. Αυτή η διαδικασία δεν αφορά απλώς την εκτέλεση ανεξάρτητων προβλέψεων σε διαφορετικά επίπεδα. Αντίθετα, ενσωματώνει αυτές τις προβλέψεις για να εξασφαλίσει συνέπεια σε ολόκληρη την ιεραρχία. Οι προβλέψεις σε χαμηλότερα επίπεδα εμφανίζονται για να ταιριάζουν με τις προβλέψεις σε υψηλότερα επίπεδα και το αντίστροφο. Αυτή η ενοποίηση διασφαλίζει ότι οι προβλέψεις είναι συνεκτικές σε όλη την ιεραρχία, κάτι που είναι μια κρίσιμη πτυχή αυτής της μεθόδου. Τα υπό εξέταση δεδομένα της παρούσας διπλωματικής εργασίας αποτυπώνονται με τη μορφή ιεραρχικού δέντρου όπως απεικονίζονται παρακάτω:



Εικόνα 12 "Ιεραρχική ταξινόμηση δεδομένων"

Σε αυτή τη διατριβή, υιοθετήσαμε τη μέθοδο του Large Scale Hierarchical Forecasting, εφαρμόζοντάς την για την ανάλυση και την πρόβλεψη δεδομένων σε διαφορετικές κατηγορίες καταστημάτων, σε διαφορετικούς νομούς και πόλεις αλλά και σε επίπεδο καταστήματος. Αναγνωρίζοντας την εγγενή ποικιλομορφία στα μεγέθη των καταστημάτων και τη μοναδική λειτουργική τους δυναμική, ταξινόμησαμε τα καταστήματα σε τρεις διακριτές κατηγορίες: μικρά, μεσαία και μεγάλα. Αυτή η ταξινόμηση αποτελεί τη βάση της ιεραρχικής μας προσέγγισης, επιτρέποντάς μας να εμβαθύνουμε στις διαφοροποιημένες ανάγκες πρόβλεψης κάθε κατηγορίας καταστημάτων, διατηρώντας παράλληλα μια συνεκτική συνολική ανάλυση.

Η πρόβλεψη θα διεξαχθεί με τρόπο τέτοιο ώστε διερευνήσουμε την αποτελεσματικότητα των διαφόρων μεθόδων πρόβλεψης χρονοσειρών στην πρόβλεψη μελλοντικών τάσεων με βάση ιστορικά δεδομένα. Η κύρια εστίαση είναι στην αξιολόγηση και τη σύγκριση της απόδοσης πολλών ευρέως αναγνωρισμένων μοντέλων πρόβλεψης, το καθένα με τη μοναδική του προσέγγιση και τις παραδοχές του. Τα μοντέλα που επιλέχθηκαν για αυτήν την ανάλυση είναι και οι μέθοδοι που έχουμε αναλύσει σε προηγούμενες παραγράφους. Η υλοποίηση αυτών των μοντέλων θα πραγματοποιηθεί με χρήση Python, αξιοποιώντας βιβλιοθήκες όπως statsmodels και pandas, που προσφέρουν ολοκληρωμένα εργαλεία και λειτουργίες για ανάλυση χρονοσειρών. Τα δεδομένα θα υποβληθούν σε επεξεργασία και θα τροφοδοτηθούν σε καθένα από αυτά τα μοντέλα για τη δημιουργία προβλέψεων.

Στον τομέα της επιστήμης δεδομένων και της στατιστικής ανάλυσης, η ακεραιότητα και η πληρότητα των δεδομένων είναι πρωταρχικής σημασίας για ακριβή αποτελέσματα και γνώσεις. Ωστόσο, τα σύνολα δεδομένων που συχνά χρησιμοποιούνται συνοδεύονται από διάφορες ανωμαλίες όπως είναι για παράδειγμα οι ελλειπείς τιμές. Τα δεδομένα που λείπουν μπορεί να προκύψουν για διάφορους λόγους, συμπεριλαμβανομένων σφαλμάτων στη συλλογή δεδομένων, απωλειών μετάδοσης ή απλώς επειδή οι πληροφορίες δεν καταγράφηκαν ποτέ. Η διαδικασία αντιμετώπισης αυτών των κενών, γνωστή ως imputation, είναι κρίσιμη για την προετοιμασία του συνόλου δεδομένων για ανάλυση. Το **imputation** χρησιμεύει ως τεχνική στην προεπεξεργασία δεδομένων, διασφαλίζοντας ότι το υπό ανάλυση σύνολο δεδομένων είναι πλήρες, κάτι που είναι κρίσιμο για την αξιοπιστία των στατιστικών δοκιμών και των προβλέψεων. Περιλαμβάνει την αντικατάσταση των ελλιπών δεδομένων με υποκατάστατες τιμές, καθιστώντας το σύνολο δεδομένων χρησιμοποιήσιμο για περαιτέρω ανάλυση. Μεταξύ των διαφόρων τεχνικών imputation, το interpolation ξεχωρίζει για την ικανότητά της να εκτιμά τις τιμές που λείπουν με βάση τα υπάρχοντα σημεία δεδομένων. Αυτή η μέθοδος είναι ιδιαίτερα χρήσιμη σε δεδομένα χρονοσειρών ή σύνολα δεδομένων με εγγενή σειρά, όπου η σχέση μεταξύ διαδοχικών σημείων

δεδομένων μπορεί να καθοδηγήσει στη διαδικασία ακριβούς προσδιορισμού των ελλειπών τιμών. Στο πλαίσιο του παρεχόμενου συνόλου δεδομένων το οποίο περιλαμβάνει καταστήματα, πωλήσεις και ημερομηνίες κ.α, χρησιμοποιήθηκε η τεχνική του interpolation για τον υπολογισμό τιμών που λείπουν στη στήλη πωλήσεων. Συγκεκριμένα, η μέθοδος του linear interpolation επιλέχθηκε για την απλότητα και την αποτελεσματικότητά της. Η μέθοδος αυτή, υποθέτει ότι η αλλαγή μεταξύ δύο σημείων είναι γραμμική και χρησιμοποιεί αυτή την υπόθεση για να εκτιμήσει τις τιμές που λείπουν. Υπολογίζει την τιμή ενός σημείου δεδομένων που λείπει καθώς βρίσκεται στην ευθεία γραμμή μεταξύ των γειτονικών σημείων του στο σύνολο δεδομένων. Αυτή η μέθοδος είναι ιδιαίτερα κατάλληλη για σύνολα δεδομένων όπου οι αλλαγές μεταξύ διαδοχικών σημείων δεδομένων αναμένεται να είναι σταδιακές και όχι απότομες. Ο κώδικας που χρησιμοποιήθηκε για το imputation των δεδομένων εμφανίζεται παρακάτω:

```
import pandas as pd

data = pd.read_excel('Data_for_my_thesis.xlsx')

# Display the first few rows of the dataframe and summary of null values
data_info = data.isnull().sum()
print(data.head())
print(data_info)

# Fill missing values in 'Sales' column with interpolation
data['Sales'] = data['Sales'].interpolate(method='linear')

# Check again for null values to confirm imputation
data_imputed_info = data.isnull().sum()
print(data.head())
print(data_imputed_info)
```

Εικόνα 13 "Κώδικας Python - Interpolation"

Μέσω του interpolation δώθηκαν τιμές σε 585 καταχωρήσεις οι οποίες δεν διέθεταν προσυμπληρωμένες τιμές. Μερικά ενδεικτικά παραδείγματα εμφανίζονται παρακάτω:

Code	Sqm	Size	VLTs	SSBTs	Date	Sales after interpolation	Original Data
231779	96,25	Medium	3	1	2019-01-01 00:00:00	105.680 €	- €
231779	96,25	Medium	3	1	2019-02-01 00:00:00	94.635 €	- €
231779	96,25	Medium	3	1	2019-03-01 00:00:00	83.590 €	- €
231779	96,25	Medium	3	1	2019-04-01 00:00:00	72.545 €	- €
231771	170	Large	3	2	2020-04-01 00:00:00	67.907 €	- €
231779	96,25	Medium	3	1	2019-05-01 00:00:00	61.500 €	- €
232668	88,75	Medium	3	2	2020-04-01 00:00:00	60.418 €	- €
124437	69	Small	3	4	2020-04-01 00:00:00	57.010 €	- €
124437	69	Small	3	4	2021-01-01 00:00:00	56.246 €	- €
122695	75	Medium	4	2	2020-04-01 00:00:00	53.667 €	- €
232700	122,93	Large	10	2	2020-04-01 00:00:00	53.272 €	- €
231780	73,21	Medium	3	4	2020-04-01 00:00:00	51.564 €	- €
100433	76,6	Medium	7	4	2020-04-01 00:00:00	51.306 €	- €
232708	116,58	Large	4	4	2020-04-01 00:00:00	50.856 €	- €
231779	96,25	Medium	3	1	2019-06-01 00:00:00	50.455 €	- €
100071	99	Medium	3	4	2020-04-01 00:00:00	49.309 €	- €
100062	47	Small	0	2	2019-01-01 00:00:00	47.923 €	- €
100062	47	Small	0	2	2019-02-01 00:00:00	45.620 €	- €
100014	90	Medium	3	2	2020-04-01 00:00:00	45.291 €	- €
232765	255,98	Large	6	4	2020-04-01 00:00:00	44.923 €	- €
100148	45	Small	11	2	2019-01-01 00:00:00	44.505 €	- €
233260	75	Medium	5	4	2020-04-01 00:00:00	43.870 €	- €
124444	127	Large	3	4	2020-04-01 00:00:00	43.566 €	- €

Εικόνα 14 "Αποτελέσματα Interpolation"

Το επόμενο βήμα που απαιτείται για να προχωρήσουμε στη διαδικασία της πρόβλεψης, περιλαμβάνει τον διαχωρισμό του συνόλου δεδομένων σε δύο υπό σύνολα: ένα σετ εκπαίδευσης (training) και ένα σετ δοκιμής (testing). Το training test χρησιμοποιείται για την εκπαίδευση ή την προσαρμογή του μοντέλου. Είναι μέσω της έκθεσης σε αυτά τα δεδομένα που το μοντέλο μαθαίνει τα υποκείμενα μοτίβα, τις σχέσεις ή τις δομές που υπάρχουν μέσα στα δεδομένα. Το μέγεθος του συνόλου εκπαίδευσης μπορεί να επηρεάσει σημαντικά την ικανότητα γενίκευσης του μοντέλου. Συνήθως, ένα μεγαλύτερο μέρος του συνόλου δεδομένων κατανέμεται στην εκπαίδευση για να διασφαλιστεί ότι το μοντέλο έχει άφθονα δεδομένα για να μάθει. Οι συνήθεις διαχωρισμοί περιλαμβάνουν το 70% έως το 80% του συνόλου των δεδομένων που προορίζονται για εκπαίδευση. Το testing set από την άλλη πλευρά, είναι ένα ξεχωριστό μέρος των δεδομένων που δεν βλέπει το μοντέλο κατά τη φάση της εκπαίδευσης. Αυτό το σύνολο χρησιμοποιείται για την αξιολόγηση της απόδοσης του μοντέλου, δοκιμάζοντας την ικανότητά του να εφαρμόζει μαθημένα μοτίβα σε νέα, δεδομένα. Το training set παρέχει μια εκτίμηση για το πώς αναμένεται να αποδώσει το μοντέλο σε εφαρμογές πραγματικού κόσμου. Γενικά, το 20% έως το 30% του συνόλου δεδομένων προορίζεται για δοκιμή. Για του σκοπούς της πρόβλεψης, στη συγκεκριμένη εργασία χρησιμοποιήθηκε το 75% του συνόλου δεδομένων ως training set & το 25% ως testing set. Ο κώδικας που χρησιμοποιήθηκε εμφανίζεται παρακάτω:

```
# sort data by date
data['Date'] = pd.to_datetime(data['Date'])
data.sort_values('Date', inplace=True)

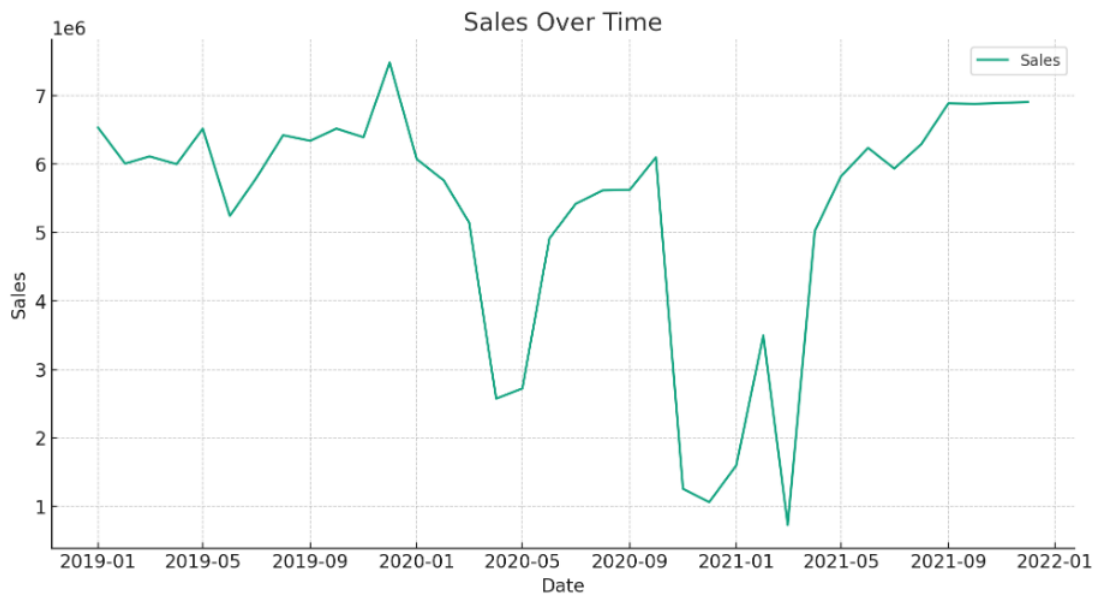
# Determine the split point for an 75-15 train-test split
split_point = int(len(data) * 0.8)

# Split the data into training and test sets
train_data = data[:split_point]
test_data = data[split_point:]
```

Εικόνα 15 "Διαχωρισμός Δεδομένων Training - Test"

Πρόβλεψη με τη μέθοδο Exponential Smoothing:

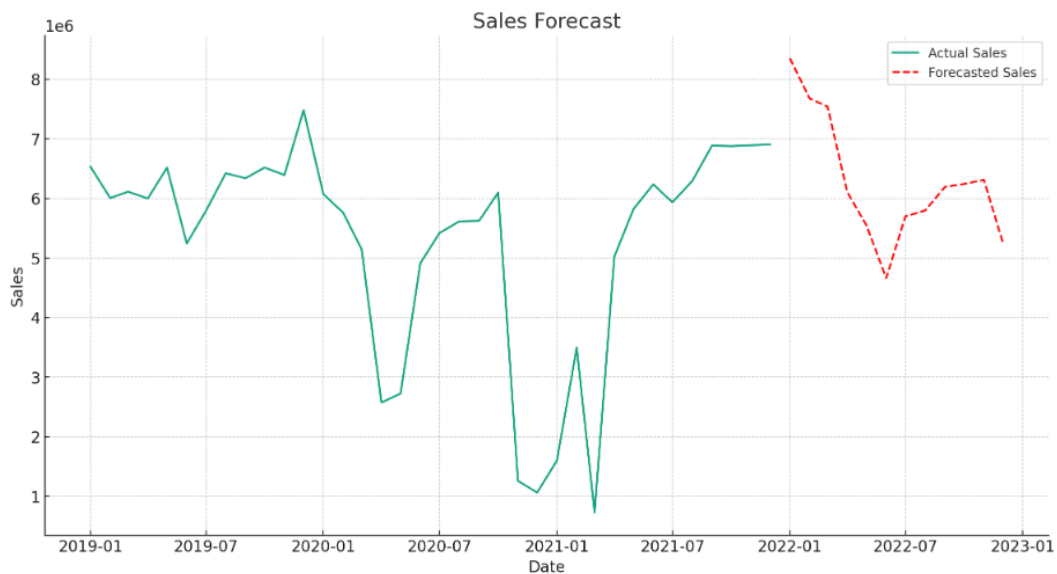
Η εκθετική εξομάλυνση είναι κατάλληλη για δεδομένα χρονοσειρών, ειδικά όταν τα δεδομένα δείχνουν μια τάση ή εποχιακό μοτίβο. Ωστόσο, από το δεδομένο σύνολο δεδομένων, δεν είναι αμέσως σαφές εάν τα δεδομένα είναι εποχιακά ή εάν υπάρχουν τάσεις που πρέπει να ληφθούν υπόψη. Θα ξεκινήσουμε λοιπόν αναλύοντας τις χρονοσειρές για κάθε κατηγορία μεγέθους για να κατανοήσουμε τα χαρακτηριστικά της και στη συνέχεια θα προχωρήσουμε με τις προβλέψεις της εκθετικής εξομάλυνσης.



Εικόνα 16 "Πωλήσεις διαχρονικά"

Το παραπάνω γράφημα δείχνει τις πωλήσεις με την πάροδο του χρόνου. Φαίνεται ότι υπάρχουν κάποιες διακυμάνσεις στις πωλήσεις, αλλά δεν είναι αμέσως σαφές εάν υπάρχει ένα ισχυρό εποχιακό μοτίβο ή μια σταθερή τάση.

Έχοντας αυτό υπόψη, θα προχωρήσουμε στην εφαρμογή της εκθετικής εξομάλυνσης για την πρόβλεψη πωλήσεων. Δεδομένης της φύσης των δεδομένων μας, θα χρησιμοποιήσουμε τη μέθοδο Holt-Winters Exponential Smoothing, η οποία μπορεί να εξηγήσει τις τάσεις και την εποχικότητα, παρόλο που δεν είναι έντονα εμφανείς. Αυτή η μέθοδος είναι ευέλικτη και μπορεί να προσαρμοστεί σε διάφορα χαρακτηριστικά χρονοσειρών. Δεδομένου ότι η συχνότητα των δεδομένων μας φαίνεται να είναι μηνιαία, η πρόβλεψη θα καλύπτει 12 μήνες στο μέλλον.



Εικόνα 17 "Πρόβλεψη με τη μέθοδο Exponential Smoothing"

Το γράφημα μας δείχνει τόσο τις πραγματικές πωλήσεις όσο και τις προβλεπόμενες πωλήσεις για το σύνολο των καταστημάτων. Η συμπαγής γραμμή αντιπροσωπεύει τα πραγματικά ιστορικά δεδομένα πωλήσεων ενώ διακεκομμένη γραμμή αντιπροσωπεύει τις προβλεπόμενες πωλήσεις για το επόμενο έτος. Ο κώδικας που χρησιμοποιήθηκε για την παραπάνω πρόβλεψη καθώς επίσης και τα αποτελέσματα αναλυτικά για τους επόμενους μήνες ανά κατηγορία καταστήματος εμφανίζεται παρακάτω:

```

train_data.set_index('Date ', inplace=True)

# Fit the Holt-Winters Exponential Smoothing model
# Assuming a yearly seasonality with a 12-month cycle
model = ExponentialSmoothing(train_data['Sales '], trend='add', seasonal='add', seasonal_periods=12).fit()

# Forecast the next 12 months
forecast = model.forecast(12)

# Creating a new DataFrame for the forecasted values to facilitate plotting
forecast_dates = pd.date_range(start=train_data.index[-1] + pd.DateOffset(months=1), periods=12, freq='MS')
forecast_df = pd.DataFrame(forecast.values, index=forecast_dates, columns=['Forecasted Sales'])

# Plotting
plt.figure(figsize=(14, 7))
plt.plot(train_data.index, train_data['Sales '], label='Actual Sales')
plt.plot(forecast_df.index, forecast_df['Forecasted Sales'], linestyle='--', color='red')
plt.title('Sales Forecast')
plt.xlabel('Date')
plt.ylabel('Sales')
plt.legend()
plt.grid(True)
plt.show()

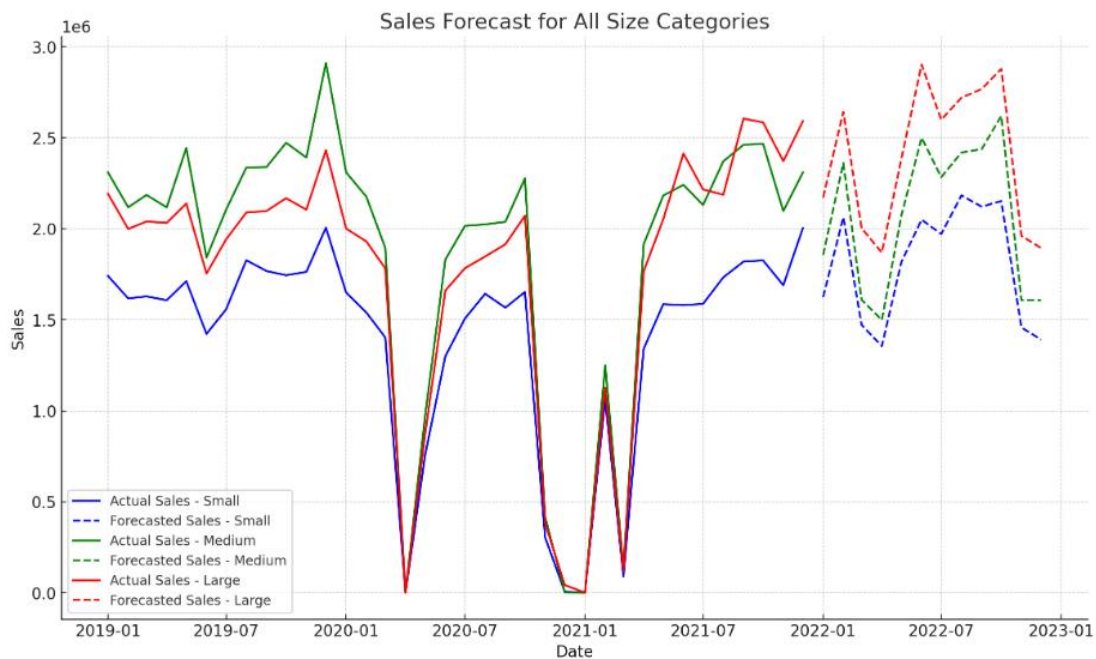
```

Εικόνα 18 "Κώδικας Python - Exponential Smoothing"

Date	Sales Forecast
31/1/2022	8350000
28/2/2022	7680000
31/3/2022	7540000
30/4/2022	6100000
31/5/2022	5550000

30/6/2022	4660000
31/7/2022	5700000
31/8/2022	5800000
30/9/2022	6190000
31/10/2022	6240000
30/11/2022	6310000
31/12/2022	5250000

Παρακάτω η ανάλυση για κάθε κατηγορία μεγέθους:



Εικόνα 19 "Πρόβλεψη με τη μέθοδο Exponential Smoothing για κάθε κατηγορία μεγέθους "

```
# Segment data by size category and forecast
sizes = data_size['Size'].unique()
forecast_dfs = {} # To store forecast DataFrames for each size

for size in sizes:
    size_data = data_size[data_size['Size'] == size].set_index('Date')
    model = ExponentialSmoothing(size_data['Sales'], trend='add', seasonal='add', seasonal_periods=12).fit()
    forecast = model.forecast(12)
    forecast_dates = pd.date_range(start=size_data['Date'].max() + pd.DateOffset(months=1), periods=12, freq='MS')
    forecast_dfs[size] = pd.DataFrame(forecast, index=forecast_dates, columns=['Forecasted Sales'])

# Plotting the historical and forecasted sales for all size categories on a single plot
plt.figure(figsize=(14, 8))

# Colors for each size category
colors = {'Small': 'blue', 'Medium': 'green', 'Large': 'red'}

for size in sizes:
    actual_data = data_size[data_size['Size'] == size]
    plt.plot(actual_data['Date'], actual_data['Sales'], label=f'Actual Sales - {size}', color=colors[size])
    plt.plot(forecast_dates, forecast_dfs[size], label=f'Forecasted Sales - {size}', linestyle='--', color=colors[size])

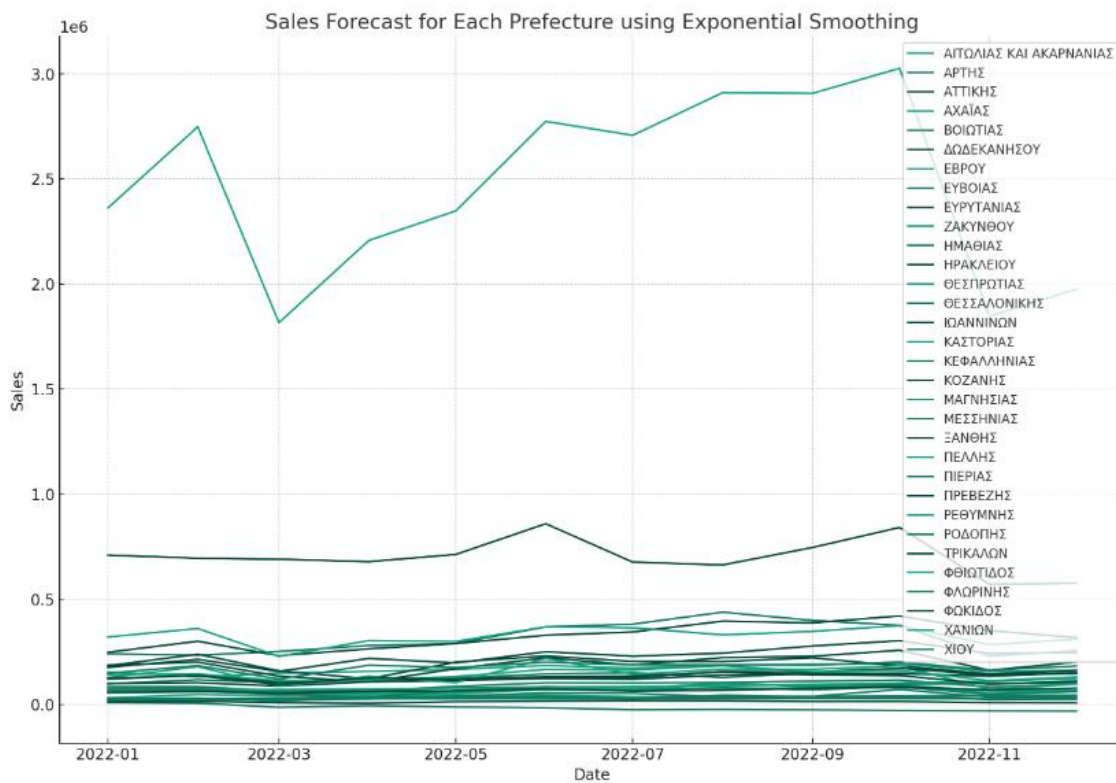
plt.title('Sales Forecast for All Size Categories')
plt.xlabel('Date')
plt.ylabel('Sales')
plt.legend()
plt.grid(True)
plt.show()
```

Εικόνα 20 "Κώδικας Python - Exponential Smoothing για κάθε κατηγορία μεγέθους "

Τα αναλυτικά αποτελέσματα της πρόβλεψης ανά μήνα και ανά κατηγορία μεγέθους καταστήματος εμφανίζονται παρακάτω:

Date	Small Size Forecast	Medium Size Forecast	Large Size Forecast
1/1/2022	1,624,267	1,855,696	2,171,137
1/2/2022	2,063,577	2,361,355	2,642,821
1/3/2022	1,472,082	1,609,989	2,003,476
1/4/2022	1,355,197	1,498,854	1,868,849
1/5/2022	1,817,816	2,073,645	2,385,566
1/6/2022	2,050,129	2,500,214	2,902,439
1/7/2022	1,970,679	2,283,901	2,598,844
1/8/2022	2,183,524	2,419,012	2,720,147
1/9/2022	2,121,405	2,437,465	2,767,913
1/10/2022	2,152,461	2,619,807	2,878,471
1/11/2022	1,457,496	1,607,606	1,961,173
1/12/2022	1,391,391	1,606,785	1,894,431

Παρακάτω η ανάλυση σε πλαίσιο νομού:



Εικόνα 21 " Πρόβλεψη με τη μέθοδο Exponential Smoothing για κάθε Νομό "

Το παραπάνω διάγραμμα, δείχνει τις προβλέψεις πωλήσεων για κάθε νομό με τη μέθοδο της εκθετικής εξομάλυνσης. Κάθε γραμμή αντιπροσωπεύει τις προβλεπόμενες πωλήσεις για έναν συγκεκριμένο νομό τους επόμενους 12 μήνες. Λόγω του αριθμού των νομών και της δυνατότητας οπτικής επικάλυψης, το αποτέλεσμα μπορεί να φαίνεται πυκνό, αλλά απεικονίζει τις προβλεπόμενες τάσεις σε διαφορετικές γεωγραφικές περιοχές.

```
# Extracting unique prefectures from the dataset
prefectures = data_prefecture['Prefecture'].unique()

# Forecasting sales for each prefecture and storing the forecast for plotting
forecast_prefecture_dfs = {}
for prefecture in prefectures:
    # Filter data for the current prefecture
    prefecture_data = data_prefecture[data_prefecture['Prefecture'] == prefecture].set_index('Date')
    # Fit the Exponential Smoothing model
    model = ExponentialSmoothing(prefecture_data['Sum'], trend='add', seasonal='add', seasonal_periods=12).fit()
    # Forecast the next 12 months
    forecast = model.forecast(12)
    # Store forecasts
    forecast_prefecture_dfs[prefecture] = forecast

# Plotting the forecasts for each prefecture
plt.figure(figsize=(15, 10))
for prefecture, forecast in forecast_prefecture_dfs.items():
    forecast_dates = pd.date_range(start=data_prefecture['Date'].max() + pd.DateOffset(months=1), periods=12, freq='MS')
    plt.plot(forecast_dates, forecast, label=prefecture)

plt.title('Sales Forecast for Each Prefecture using Exponential Smoothing')
plt.xlabel('Date')
plt.ylabel('Sales')
plt.legend(loc='upper left', bbox_to_anchor=(1, 1))
plt.grid(True)
plt.show()
```

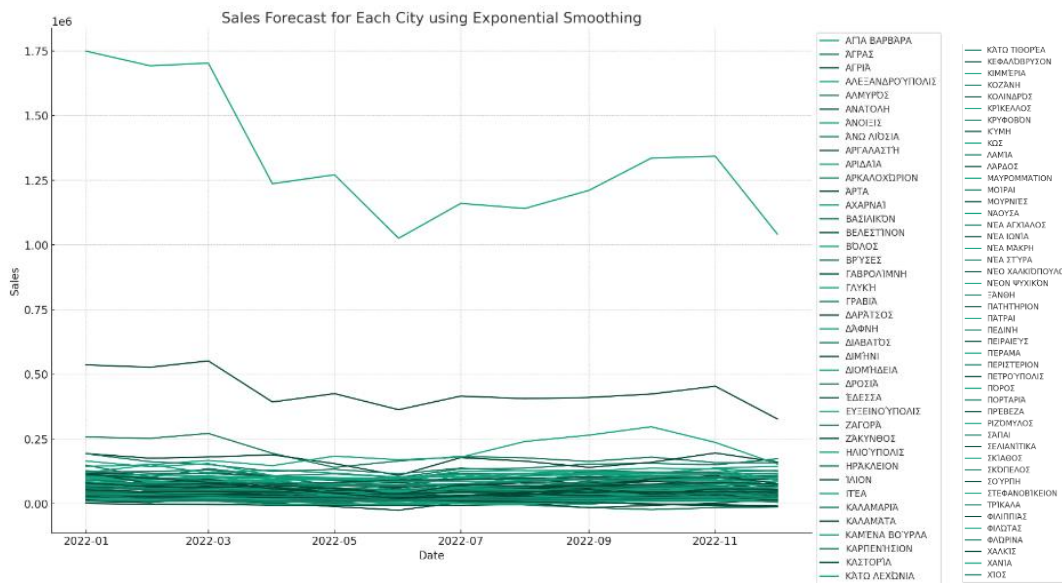
Εικόνα 22 "Κώδικας Python - Exponential Smoothing για κάθε Νομό "

Τα αναλυτικά αποτελέσματα της πρόβλεψης ανά μήνα και ανά Νομό εμφανίζονται παρακάτω:

Column1	2022-01-01 00:00:00	2022-02-01 00:00:00	2022-03-01 00:00:00	2022-04-01 00:00:00	2022-05-01 00:00:00	2022-06-01 00:00:00	2022-07-01 00:00:00	2022-08-01 00:00:00	2022-09-01 00:00:00	2022-10-01 00:00:00	2022-11-01 00:00:00	2022-12-01 00:00:00
ΑΙΤΩΛΙΑΣ ΚΑΙ ΑΚΑΡΝΑΝΙΑΣ	64.717 €	72.631 €	44.875 €	45.915 €	52.057 €	70.356 €	77.279 €	101.036 €	83.198 €	83.953 €	51.992 €	61.465 €
ΑΡΤΗΣ	98.086 €	116.893 €	103.632 €	114.529 €	125.431 €	144.336 €	150.422 €	127.194 €	150.219 €	144.906 €	136.686 €	149.424 €
ΑΤΤΙΚΗΣ	709.907 €	695.247 €	690.273 €	678.815 €	713.236 €	859.325 €	676.697 €	663.259 €	746.308 €	841.679 €	571.249 €	576.306 €
ΑΧΑΪΑΣ	154.229 €	130.256 €	148.778 €	156.646 €	165.927 €	184.707 €	172.197 €	187.333 €	190.872 €	193.260 €	156.060 €	162.329 €
ΒΟΙΩΤΙΑΣ	8.320 €	5.972 €	14.179 €	6.364 €	12.678 €	16.326 €	25.659 €	24.164 €	26.378 €	29.194 €	31.077 €	31.930 €
ΔΩΔΕΚΑΝΗΣΟΥ	175.757 €	214.088 €	156.663 €	218.390 €	197.106 €	250.245 €	229.195 €	243.506 €	277.528 €	302.936 €	242.309 €	246.746 €
ΕΒΡΟΥ	73.158 €	78.311 €	54.894 €	50.133 €	57.020 €	84.626 €	68.651 €	98.822 €	86.490 €	90.975 €	65.013 €	66.479 €
ΕΥΒΟΙΑΣ	240.018 €	233.836 €	251.487 €	279.825 €	289.513 €	368.431 €	381.469 €	438.678 €	399.246 €	376.148 €	229.252 €	255.573 €
ΕΥΡΥΤΑΝΙΑΣ	11.802 €	14.943 €	7.936 €	5.522 €	12.368 €	15.290 €	17.129 €	16.642 €	13.349 €	13.655 €	8.194 €	7.457 €
ΖΑΚΥΝΘΟΥ	24.175 €	26.174 €	28.563 €	41.753 €	42.946 €	44.248 €	36.730 €	37.570 €	38.114 €	40.567 €	42.364 €	44.035 €
ΗΜΑΘΙΑΣ	83.697 €	83.987 €	64.520 €	64.547 €	85.486 €	102.694 €	102.089 €	103.740 €	110.401 €	115.641 €	82.766 €	95.845 €
ΗΡΑΚΛΕΙΟΥ	247.521 €	300.304 €	231.577 €	263.952 €	289.476 €	328.999 €	344.110 €	395.950 €	387.695 €	420.738 €	350.571 €	317.827 €
ΘΕΣΣΑΛΟΝΙΚΗΣ	23.134 €	27.882 €	18.083 €	26.119 €	25.747 €	25.566 €	27.755 €	32.652 €	33.723 €	35.123 €	25.009 €	24.454 €
ΘΕΣΣΑΛΟΝΙΚΗΣ	20.630 €	21.470 €	23.158 €	24.726 €	39.388 €	54.812 €	52.521 €	35.656 €	37.647 €	70.978 €	34.106 €	32.691 €
ΙΟΑΝΝΙΝΩΝ	98.839 €	103.099 €	86.904 €	106.957 €	113.640 €	122.756 €	119.323 €	139.144 €	142.919 €	147.762 €	133.974 €	149.304 €
ΚΑΣΤΟΡΙΑΣ	30.913 €	46.936 €	41.281 €	54.869 €	63.359 €	71.094 €	76.580 €	75.597 €	77.877 €	78.689 €	63.336 €	60.100 €
ΚΕΦΑΛΛΗΝΙΑΣ	29.036 €	26.018 €	20.144 €	31.998 €	26.658 €	32.873 €	25.343 €	43.096 €	36.412 €	39.214 €	22.143 €	24.732 €
ΚΟΖΑΝΗΣ	183.568 €	238.355 €	159.211 €	121.669 €	200.861 €	230.069 €	204.313 €	203.626 €	221.408 €	182.689 €	144.390 €	160.313 €
ΜΑΓΝΗΣΙΑΣ	2.361.346 €	2.747.858 €	1.815.867 €	2.207.197 €	2.347.636 €	2.773.447 €	2.707.243 €	2.910.005 €	2.906.835 €	3.025.928 €	1.845.716 €	1.976.002 €
ΜΕΣΣΗΝΙΑΣ	57.965 €	60.060 €	61.806 €	63.150 €	79.194 €	81.912 €	84.778 €	88.683 €	92.587 €	96.491 €	99.782 €	102.408 €
ΞΑΝΘΗΣ	186.027 €	200.838 €	134.157 €	107.054 €	169.680 €	221.511 €	187.266 €	222.301 €	228.705 €	257.177 €	161.974 €	201.278 €
ΠΕΛΛΗΣ	127.720 €	179.054 €	86.315 €	133.782 €	103.014 €	218.039 €	142.541 €	167.964 €	166.957 €	183.192 €	70.700 €	96.313 €
ΠΕΡΙΑΣ	26.547 €	25.528 €	30.191 €	34.906 €	34.331 €	36.482 €	27.421 €	27.793 €	27.857 €	27.192 €	26.031 €	24.827 €
ΠΙΡΕΑΪΚΗΣ	120.225 €	138.264 €	106.259 €	120.335 €	117.473 €	123.538 €	127.335 €	156.697 €	154.321 €	171.713 €	139.981 €	146.086 €
ΡΕΘΥΜΝΗΣ	63.285 €	66.717 €	70.533 €	74.132 €	76.582 €	79.549 €	84.660 €	88.905 €	95.129 €	102.556 €	110.115 €	122.092 €
ΡΟΔΟΥ	69.849 €	71.559 €	47.774 €	33.084 €	66.541 €	74.429 €	65.408 €	79.795 €	69.225 €	83.241 €	47.560 €	61.141 €
ΤΡΙΚΑΛΩΝ	55.353 €	61.993 €	52.335 €	61.800 €	64.653 €	72.694 €	62.757 €	66.672 €	76.873 €	83.914 €	69.043 €	76.853 €
ΦΘΙΩΤΙΔΟΣ	141.419 €	145.417 €	112.519 €	115.745 €	132.241 €	166.027 €	163.421 €	183.500 €	155.985 €	161.363 €	112.836 €	130.351 €
ΦΛΩΡΙΝΗΣ	148.167 €	184.343 €	120.193 €	186.311 €	175.263 €	200.400 €	188.358 €	187.231 €	186.145 €	201.364 €	162.235 €	182.494 €
ΦΩΚΙΑΔΟΣ	124.848 €	136.195 €	93.849 €	126.907 €	118.959 €	131.800 €	133.878 €	153.956 €	140.579 €	137.185 €	96.340 €	110.137 €
ΧΑΝΙΩΝ	320.475 €	360.801 €	228.093 €	303.287 €	299.312 €	369.031 €	364.258 €	330.574 €	347.079 €	374.621 €	284.203 €	309.944 €
ΧΙΟΥ	27.894 €	28.771 €	16.279 €	34.429 €	27.841 €	38.196 €	27.777 €	29.225 €	34.880 €	36.342 €	38.790 €	41.388 €

Εικόνα 23 "Αποτελέσματα μεθόδου Exponential Smoothing ανά μήνα & ανά Νομό "

Παρακάτω η ανάλυση σε πλαίσιο πώλης:



Εικόνα 24 " Πρόβλεψη με τη μέθοδο Exponential Smoothing για κάθε Πόλη "

```
# Forecasting and plotting for each city
for city in cities:
    # Filter data for the current city
    city_data = data_city[data_city['City'] == city].set_index('Date')
    # Fit the Exponential Smoothing model
    model = ExponentialSmoothing(city_data['Sales'], trend='add', seasonal='add', seasonal_periods=12).fit()
    # Forecast the next 12 months
    forecast = model.forecast(12)

    # Generate dates for plotting the forecast
    forecast_dates = pd.date_range(start=data_city['Date'].max() + pd.DateOffset(months=1), periods=12, freq='MS')
    # Plot the forecast
    plt.plot(forecast_dates, forecast, label=city)

plt.title('Sales Forecast for Each City using Exponential Smoothing')
plt.xlabel('Date')
plt.ylabel('Sales')
plt.legend(loc='upper left', bbox_to_anchor=(1, 1)) # Adjust the legend position to display outside the plot
plt.grid(True)
plt.show()
```

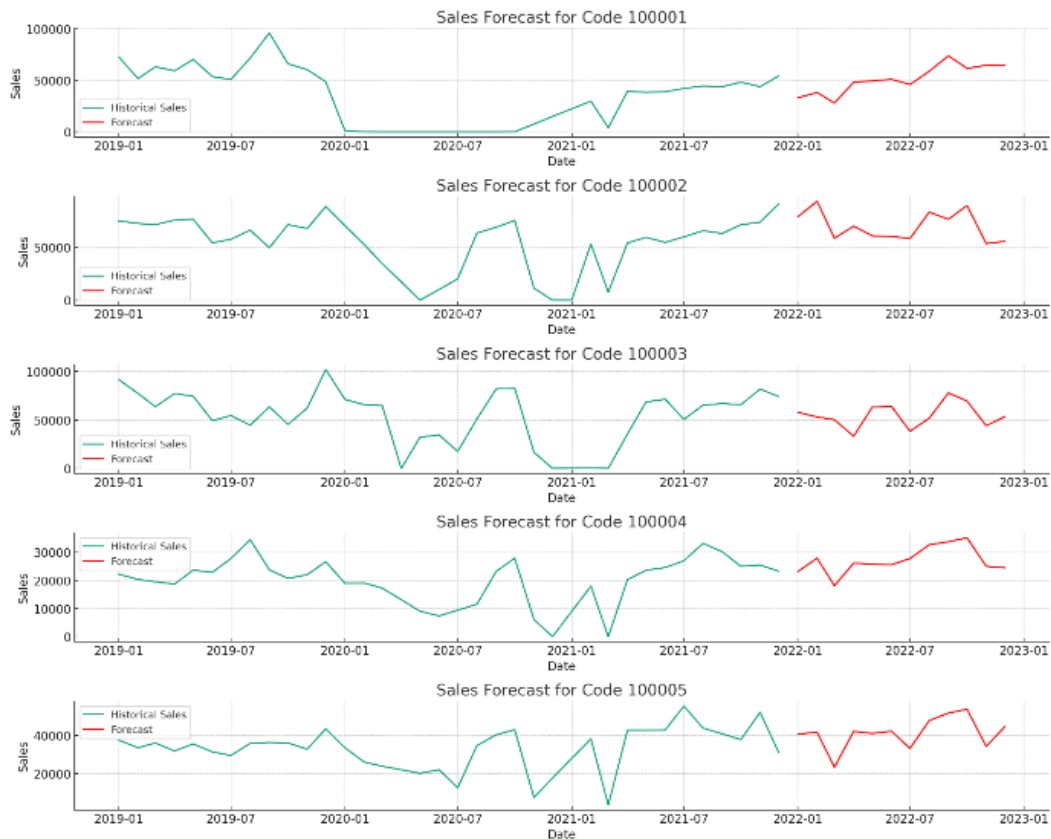
Εικόνα 25 "Κώδικας Python - Exponential Smoothing για κάθε Πόλη "

Παρακάτω εμφανίζεται ένα δείγμα των αποτελεσμάτων που παράχθηκαν από την πρόβλεψη των πωλήσεων για κάθε πόλη με τη μέθοδο Exponential Smoothing:

	1/1/2022	1/2/2022	1/3/2022	1/4/2022	1/5/2022	1/6/2022	1/7/2022	1/8/2022	1/9/2022	1/10/2022	1/11/2022	1/12/2022
ΑΠΑ ΒΑΡΒΑΡΑ	104.902 €	97.853 €	103.968 €	94.795 €	90.690 €	88.356 €	116.089 €	100.344 €	92.956 €	87.845 €	89.737 €	104.092 €
ΆΓΡΑΣ	39.676 €	31.193 €	17.875 €	8.528 €	2.554 €	4.898 €	5.157 €	4.825 €	15.078 €	22.963 €	15.418 €	12.421 €
ΑΓΡΙΑ	63.797 €	57.787 €	73.119 €	55.504 €	49.196 €	58.009 €	49.919 €	56.962 €	57.015 €	57.100 €	58.256 €	49.250 €
ΑΛΕΞΑΝΔΡΟΥΠΟΛ	53.897 €	68.972 €	74.310 €	54.791 €	46.166 €	44.359 €	69.801 €	63.207 €	80.697 €	66.834 €	59.582 €	34.807 €
ΑΛΜΥΡΟΣ	123.230 €	124.575 €	130.322 €	104.361 €	96.280 €	100.537 €	118.426 €	113.871 €	115.369 €	118.305 €	118.992 €	94.494 €
ΑΝΑΤΟΛΗ	43.528 €	47.635 €	51.362 €	55.860 €	59.319 €	63.274 €	69.876 €	80.033 €	81.616 €	87.809 €	89.021 €	96.428 €
ΑΝΟΙΞΙΣ	101.275 €	86.314 €	81.313 €	75.719 €	62.177 €	48.275 €	76.858 €	58.655 €	66.811 €	72.104 €	76.340 €	74.580 €
ΑΝΩ ΛΙΟΣΙΑ	78.271 €	62.971 €	69.791 €	61.888 €	21.892 €	22.820 €	22.084 €	33.057 €	22.403 €	21.551 €	29.532 €	34.515 €
ΑΡΓΑΛΑΣΤΗ	8.733 €	4.813 €	17.929 €	6.443 €	18.696 €	24.583 €	27.437 €	32.010 €	44.243 €	32.051 €	25.793 €	6.890 €
ΑΡΙΔΑΙΑ	108.539 €	91.686 €	109.199 €	75.340 €	80.752 €	52.001 €	114.575 €	75.851 €	83.802 €	93.601 €	93.158 €	44.015 €
ΑΡΚΑΛΟΧΩΡΙΟΝ	72.457 €	69.704 €	62.646 €	51.881 €	49.054 €	46.070 €	54.739 €	51.561 €	58.465 €	58.904 €	68.768 €	56.498 €
ΆΡΤΑ	120.213 €	111.015 €	119.608 €	104.633 €	89.032 €	81.364 €	94.231 €	97.504 €	74.209 €	85.236 €	77.570 €	67.301 €
ΑΧΑΡΝΑΙ	81.099 €	76.759 €	68.116 €	57.478 €	37.172 €	41.696 €	60.546 €	56.885 €	60.052 €	65.372 €	51.068 €	47.387 €
ΒΑΣΙΛΙΚΟΝ	126.820 €	118.298 €	117.016 €	103.008 €	72.049 €	73.178 €	95.593 €	107.233 €	123.056 €	105.300 €	103.245 €	69.347 €
ΒΕΛΕΣΤΙΝΟΝ	78.068 €	80.953 €	88.129 €	56.359 €	54.105 €	47.600 €	48.635 €	59.110 €	65.008 €	67.359 €	78.148 €	53.619 €
ΒΟΛΟΣ	1.749.227 €	1.692.449 €	1.703.224 €	1.236.785 €	1.271.349 €	1.025.966 €	1.160.674 €	1.140.960 €	1.211.134 €	1.335.514 €	1.343.417 €	1.042.351 €
ΒΥΡΥΣΣ	74.783 €	73.709 €	62.214 €	64.511 €	69.996 €	72.849 €	76.921 €	85.450 €	94.137 €	97.148 €	98.457 €	104.997 €
ΓΑΒΡΟΛΙΜΝΗ	847 €	1.941 €	2.660 €	5.846 €	5.679 €	5.789 €	6.476 €	2.863 €	5.807 €	2.451 €	7.547 €	11.906 €
ΓΛΥΚΗ	24.101 €	25.358 €	28.012 €	19.960 €	26.454 €	23.557 €	20.702 €	24.283 €	30.569 €	28.360 €	24.757 €	16.890 €
ΓΡΑΒΙΑ	36.168 €	33.052 €	32.128 €	24.552 €	27.358 €	21.812 €	27.462 €	28.686 €	36.531 €	31.148 €	32.393 €	23.428 €
ΔΑΡΑΤΣΟΣ	37.238 €	14.897 €	10.378 €	3.591 €	11.208 €	25.147 €	5.400 €	2.773 €	15.760 €	6.882 €	1.901 €	17.561 €
ΔΑΦΝΗ	22.910 €	18.356 €	18.055 €	14.603 €	10.711 €	9.729 €	1.890 €	2.511 €	2.175 €	14.695 €	17.824 €	5.165 €
ΔΙΑΒΑΤΟΣ	42.721 €	47.131 €	46.787 €	46.545 €	52.766 €	54.570 €	61.625 €	62.204 €	64.249 €	63.111 €	60.678 €	57.554 €
ΔΙΜΗΝΗ	40.164 €	43.488 €	37.639 €	30.831 €	34.959 €	28.383 €	31.172 €	26.444 €	24.558 €	18.916 €	31.352 €	35.532 €
ΔΙΟΜΗΔΕΙΑ	83.199 €	69.660 €	71.504 €	62.312 €	57.329 €	48.292 €	64.391 €	55.437 €	67.817 €	65.285 €	75.513 €	71.558 €
ΔΡΟΣΙΑ	91.551 €	78.018 €	76.320 €	82.013 €	81.042 €	72.939 €	72.830 €	91.343 €	93.081 €	92.908 €	68.215 €	49.670 €
ΕΔΕΣΣΑ	66.818 €	57.115 €	67.432 €	40.233 €	50.809 €	35.952 €	54.890 €	61.234 €	68.175 €	64.791 €	55.965 €	40.807 €
ΕΥΣΕΙΝΟΥΠΟΛΙΣ	63.060 €	33.323 €	24.413 €	20.557 €	13.633 €	310 €	9.872 €	10.590 €	24.304 €	24.896 €	36.890 €	11.382 €

Εικόνα 26 "Αποτελέσματα μεθόδου Exponential Smoothing ανά μήνα & ανά Πόλη "

Παρακάτω η ανάλυση σε πλαίσιο καταστήματος:



Εικόνα 27 "Αποτελέσματα μεθόδου Exponential Smoothing σε επίπεδο καταστήματος "

```
# Container for storing forecast results
forecasts = {}

# Get all unique codes
codes = train_data['Code'].unique()

for code in codes:
    # Select data for the current code
    code_sales = train_data[train_data['Code'] == code]['Sales']

    # Fit Exponential Smoothing model (you can adjust the parameters as needed)
    model = ExponentialSmoothing(code_sales, trend='add', seasonal='add', seasonal_periods=12)
    model_fit = model.fit()

    # Forecast the next 12 months
    forecast = model_fit.forecast(steps=12)
    forecasts[code] = forecast

    # Uncomment the following lines to plot the forecast for each code
    plt.figure(figsize=(10, 6))
    plt.plot(code_sales.index, code_sales, label='Historical Sales')
    plt.plot(pd.date_range(code_sales.index[-1], periods=13, closed='right')[1:], forecast, label='Forecast', color='red')
    plt.title(f'Sales Forecast for Code {code}')
    plt.xlabel('Date')
    plt.ylabel('Sales')
    plt.legend()
    plt.show()
```

Εικόνα 28 "Κώδικας Python - Exponential Smoothing για κάθε κατάσταση "

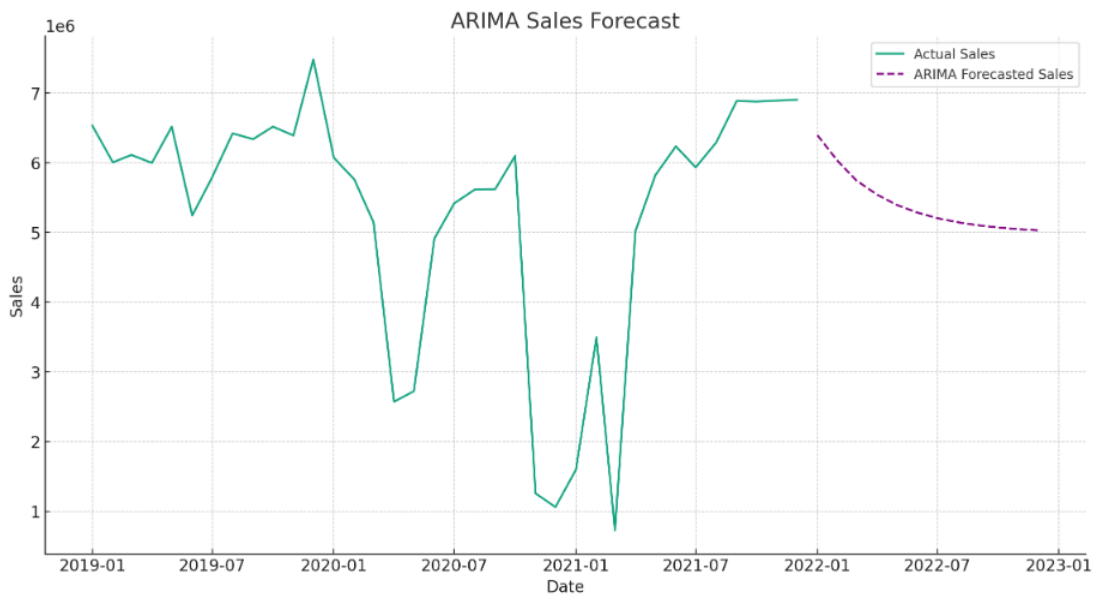
Παρακάτω εμφανίζεται ένα δείγμα των αποτελεσμάτων που παράχθηκαν από την πρόβλεψη των πωλήσεων για κάθε κατάσταση με τη μέθοδο Exponential Smoothing:

	1/1/2022	1/2/2022	1/3/2022	1/4/2022	1/5/2022	1/6/2022	1/7/2022	1/8/2022	1/9/2022	1/10/2022	1/11/2022	1/12/2022
100001	33.219 €	38.289 €	28.114 €	48.303 €	49.637 €	51.220 €	46.041 €	58.949 €	73.844 €	61.623 €	64.618 €	64.425 €
100002	84.058 €	97.881 €	73.909 €	86.113 €	84.828 €	84.035 €	92.115 €	117.111 €	116.915 €	130.732 €	103.908 €	110.027 €
100003	58.783 €	54.494 €	52.267 €	35.884 €	66.370 €	67.934 €	42.915 €	56.995 €	83.732 €	76.023 €	51.283 €	61.370 €
100004	23.270 €	27.774 €	18.432 €	26.104 €	25.998 €	25.778 €	28.108 €	33.066 €	33.977 €	35.081 €	25.447 €	24.808 €
100005	39.961 €	40.753 €	22.173 €	40.783 €	39.524 €	40.468 €	31.266 €	45.675 €	49.405 €	51.198 €	31.481 €	41.652 €
100006	25.863 €	18.157 €	18.459 €	15.022 €	24.870 €	29.619 €	18.985 €	29.554 €	28.626 €	45.154 €	18.104 €	22.804 €
100008	23.772 €	25.232 €	27.458 €	29.557 €	44.598 €	60.388 €	58.667 €	42.353 €	44.768 €	78.242 €	42.416 €	41.489 €
100009	62.835 €	66.065 €	69.638 €	72.952 €	75.077 €	77.677 €	82.380 €	86.179 €	91.915 €	98.816 €	105.809 €	117.183 €
100014	66.646 €	61.602 €	68.916 €	71.937 €	77.938 €	94.336 €	70.086 €	79.813 €	91.588 €	96.126 €	67.388 €	75.160 €
100015	29.327 €	26.461 €	20.743 €	32.755 €	27.577 €	33.957 €	26.595 €	44.519 €	38.010 €	40.989 €	24.098 €	26.870 €
100016	28.331 €	40.022 €	26.442 €	22.168 €	29.408 €	34.780 €	36.861 €	19.800 €	29.083 €	30.270 €	15.766 €	20.100 €
100022	75.961 €	71.879 €	66.351 €	79.693 €	80.299 €	90.906 €	89.570 €	94.476 €	99.311 €	102.486 €	109.021 €	116.183 €
100023	16.006 €	15.589 €	27.445 €	15.227 €	35.661 €	52.365 €	23.174 €	32.403 €	51.927 €	50.125 €	25.409 €	18.292 €
100031	55.935 €	58.776 €	54.522 €	56.789 €	57.064 €	64.360 €	53.121 €	51.027 €	58.611 €	76.563 €	63.596 €	67.740 €
100032	54.372 €	47.496 €	58.680 €	63.153 €	56.994 €	65.778 €	50.511 €	35.970 €	60.933 €	73.825 €	52.071 €	49.883 €
100037	25.258 €	24.022 €	21.035 €	22.724 €	26.089 €	31.742 €	38.967 €	39.788 €	33.098 €	30.731 €	20.162 €	19.430 €
100039	47.756 €	51.226 €	23.086 €	36.460 €	31.884 €	44.573 €	46.041 €	11.254 €	16.863 €	21.696 €	30.250 €	25.046 €
100046	35.244 €	32.454 €	29.263 €	25.853 €	25.946 €	26.406 €	26.247 €	24.591 €	22.948 €	21.325 €	19.573 €	17.539 €
100048	26.623 €	25.679 €	30.419 €	35.212 €	34.716 €	36.949 €	27.971 €	28.427 €	28.577 €	27.999 €	26.926 €	25.812 €
100049	28.148 €	32.603 €	24.864 €	32.563 €	38.456 €	34.890 €	32.234 €	33.372 €	34.408 €	35.606 €	30.041 €	27.993 €
100060	48.945 €	54.983 €	57.276 €	63.200 €	17.028 €	55.882 €	56.102 €	18.447 €	21.978 €	64.968 €	26.677 €	24.759 €
100062	39.234 €	41.546 €	48.662 €	54.440 €	56.838 €	60.293 €	55.430 €	56.753 €	57.756 €	66.397 €	58.071 €	56.549 €
100063	23.589 €	25.012 €	26.813 €	39.403 €	39.983 €	40.660 €	32.507 €	32.698 €	32.583 €	34.364 €	35.479 €	36.455 €
100068	13.787 €	15.176 €	15.379 €	17.457 €	15.505 €	20.279 €	16.245 €	17.249 €	16.593 €	11.246 €	12.910 €	13.410 €
100071	78.155 €	109.294 €	69.609 €	79.138 €	70.614 €	79.757 €	88.240 €	92.608 €	117.997 €	142.527 €	82.828 €	84.736 €

Εικόνα 29 "Αποτελέσματα μεθόδου Exponential Smoothing ανά μήνα & ανά κατάσταση "

Πρόβλεψη με τη μέθοδο ARIMA:

Για να χρησιμοποιήσουμε ένα μοντέλο ARIMA για πρόβλεψη, πρέπει πρώτα να καθορίσουμε τις κατάλληλες παραμέτρους του μοντέλου για κάθε κατηγορία μεγέθους. Αυτές οι παράμετροι είναι το p (αυτοπαλινδρομικό μέρος), το d (βαθμός διαφοροποίησης) και το q (κινούμενο μέσο μέρος). Η επιλογή αυτών των παραμέτρων τυπικά περιλαμβάνει την ανάλυση των διαγραμμάτων αυτοσυσχέτισης και μερικής αυτοσυσχέτισης και την εκτέλεση διαγνωστικών μοντέλων. Αρχικά, θα προχωρήσουμε με ένα βασικό μοντέλο ARIMA χρησιμοποιώντας ορισμένες κοινές προεπιλεγμένες παραμέτρους (όπως p=1, d=1, q=1). Εδώ φυσικά είναι σημαντικό να λάβουμε υπόψιν ότι αυτή η προσέγγιση ενδέχεται να μην αποφέρει το πιο ακριβές μοντέλο για τα δεδομένα μας.



Εικόνα 30 "Πρόβλεψη με τη μέθοδο ARIMA"

Όπως και στο παραπάνω διάγραμμα η συμπαγής γραμμή αντιπροσωπεύει τα πραγματικά ιστορικά δεδομένα πωλήσεων και η διακεκομμένη γραμμή αντιπροσωπεύει τις πωλήσεις που προβλέπονται από το βασικό μοντέλο ARIMA για τους επόμενους 12 μήνες. Αυτό που αξίζει να αναφέρουμε στη συγκεκριμένη περίπτωση είναι πως η ανάλυση παρέχει μια εκτίμηση των μελλοντικών τάσεων πωλήσεων για με βάση το μοντέλο ARIMA. Επομένως, η ακρίβεια αυτών των προβλέψεων εξαρτάται από την καταλληλότητα των παραμέτρων του μοντέλου και τις υποκείμενες υποθέσεις του μοντέλου. Η λεπτομερής ρύθμιση και η επικύρωση αυτών των μοντέλων με πρόσθετα δεδομένα ή στοιχεία είναι συχνά απαραίτητη για ακριβέστερη πρόβλεψη. Παρακάτω εμφανίζεται ο κώδικας που χρησιμοποιήθηκε καθώς και τα αναλυτικά αποτελέσματα:

```
# Define the ARIMA model parameters
p, d, q = 1, 1, 1 # Basic ARIMA(1,1,1) model configuration

# Fit the ARIMA model
arima_model = ARIMA(train_data['Sales '], order=(p, d, q)).fit()

# Forecast the next 12 months
arima_forecast = arima_model.forecast(steps=12)

# Creating forecast dates for plotting
forecast_dates = pd.date_range(start=train_data.index[-1] + pd.DateOffset(months=1), periods=12, freq='MS')

# Creating a DataFrame for the ARIMA forecasted values
arima_forecast_df = pd.DataFrame(arima_forecast.values, index=forecast_dates, columns=['ARIMA Forecasted Sales'])

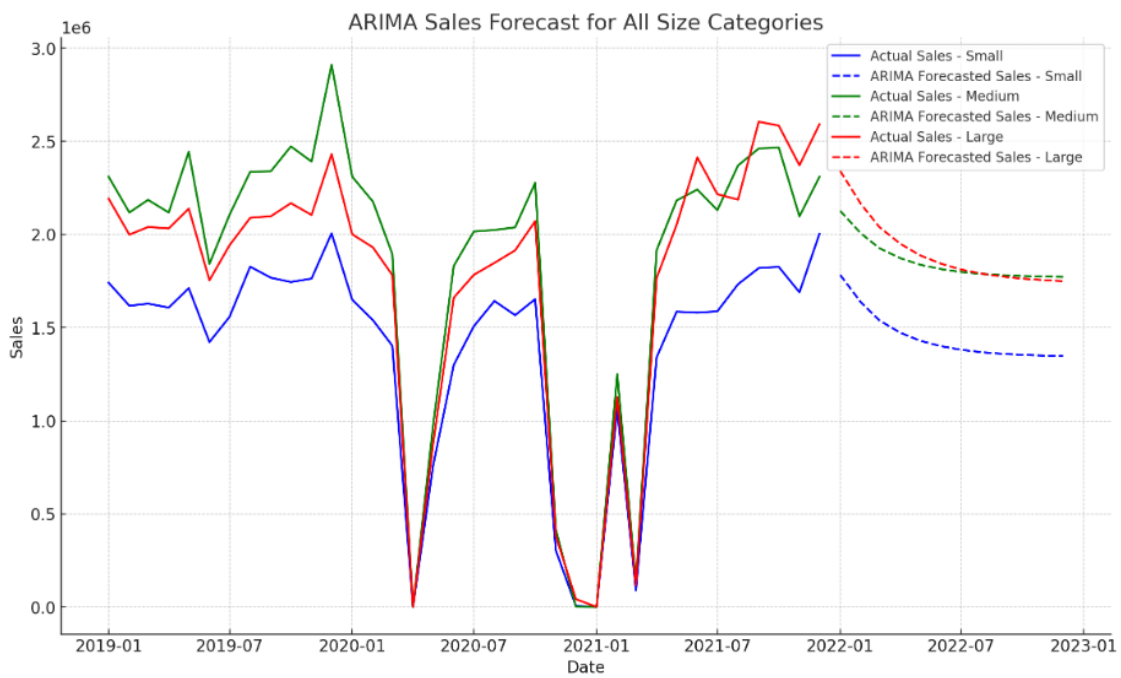
# Plotting the historical sales and forecasted values
plt.figure(figsize=(14, 7))
plt.plot(train_data.index, train_data['Sales '], label='Actual Sales')
plt.plot(arima_forecast_df.index, arima_forecast_df['ARIMA Forecasted Sales'], label='ARIMA Forecasted Sales', linestyle='--', color='purple')
plt.title('ARIMA Sales Forecast')
plt.xlabel('Date')
plt.ylabel('Sales')
plt.legend()
plt.grid(True)
plt.show()
```

Εικόνα 31 "Κώδικας Python - ARIMA"

Date	Sales Forecast
31/1/2022	6400000
28/2/2022	6020000

31/3/2022	5750000
30/4/2022	5540000
31/5/2022	5390000
30/6/2022	5290000
31/7/2022	5210000
31/8/2022	5150000
30/9/2022	5100000
31/10/2022	5070000
30/11/2022	5050000
31/12/2022	5030000

Παρακάτω η ανάλυση για κάθε κατηγορία μεγέθους:



Εικόνα 32 "Αποτελέσματα Πρόβλεψης με τη μέθοδο ARIMA για κάθε κατηγορία μεγέθους"

```
# Segment data by size category
sizes = data_size['Size'].unique()

# Colors for plotting each size category distinctly
colors = {'Small': 'blue', 'Medium': 'green', 'Large': 'red'}

plt.figure(figsize=(14, 8))

for size in sizes:
    # Filter the dataset for the current size category
    size_data = data_size[data_size['Size'] == size].copy()
    size_data.set_index('Date', inplace=True)

    # Fit an ARIMA model
    # Note: The choice of (1,1,1) for (p,d,q) is arbitrary and should be tailored to your specific dataset
    model = ARIMA(size_data['Sales'], order=(1,1,1))
    model_fit = model.fit()

    # Forecast the next 12 months
    forecast = model_fit.forecast(steps=12)

    # Plot the actual sales data
    plt.plot(size_data.index, size_data['Sales'], label=f'Actual Sales - {size}', color=colors[size])

    # Generate future dates for the forecasted data
    future_dates = pd.date_range(start=size_data.index[-1], periods=13, closed='right', freq='MS')

    # Plot the forecasted sales data
    plt.plot(future_dates, forecast, label=f'ARIMA Forecast - {size}', linestyle='--', color=colors[size])

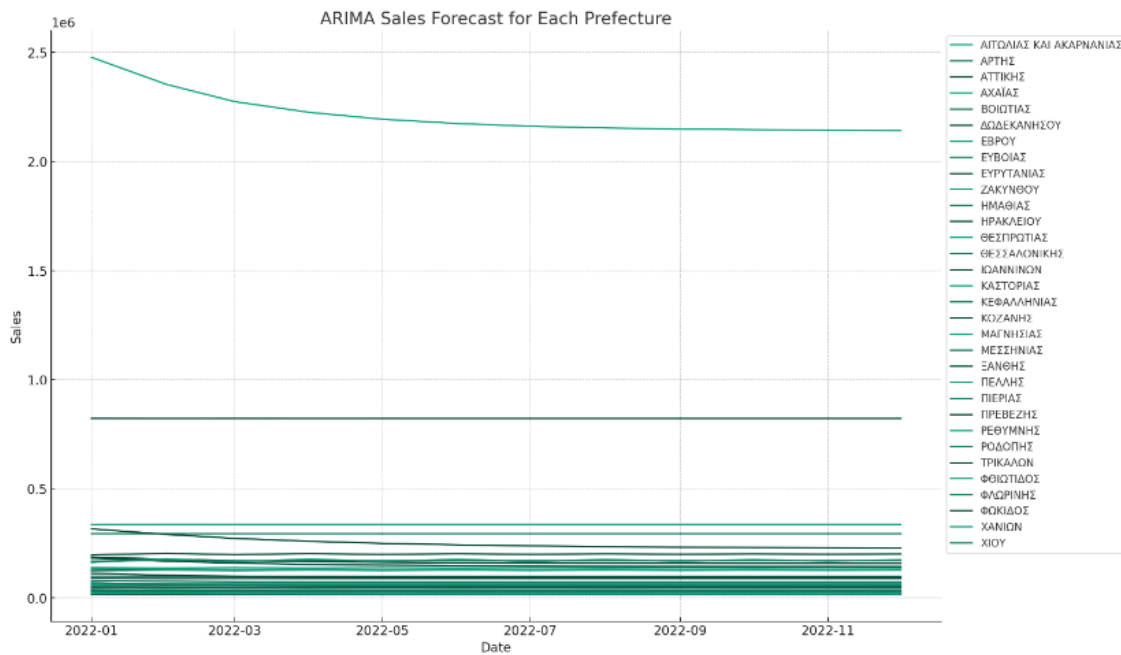
plt.title('Sales Forecast by Size Category using ARIMA')
plt.xlabel('Date')
plt.ylabel('Sales')
plt.legend()
plt.grid(True)
plt.show()
```

Εικόνα 33 "Κώδικας Python – μέθοδος ARIMA για κάθε κατηγορία μεγέθους"

Τα αναλυτικά αποτελέσματα της πρόβλεψης ανά μήνα και ανά κατηγορία μεγέθους καταστήματος εμφανίζονται παρακάτω:

Date	Small Size ARIMA Forecast	Medium Size ARIMA Forecast	Large Size ARIMA Forecast
1/1/2022	1,783,279	2,127,027	2,341,355
1/2/2022	1,636,296	2,006,073	2,164,200
1/3/2022	1,538,291	1,925,908	2,038,909
1/4/2022	1,472,944	1,872,778	1,950,296
1/5/2022	1,429,371	1,837,565	1,887,626
1/6/2022	1,400,318	1,814,226	1,843,302
1/7/2022	1,380,946	1,798,758	1,811,954
1/8/2022	1,368,030	1,788,507	1,789,784
1/9/2022	1,359,417	1,781,712	1,774,104
1/10/2022	1,353,674	1,777,209	1,763,014
1/11/2022	1,349,845	1,774,225	1,755,171
1/12/2022	1,347,292	1,772,247	1,749,624

Παρακάτω η ανάλυση σε πλαίσιο νομού:



Εικόνα 34 "Αποτελέσματα πρόβλεψης ARIMA ανά Νομό"

Το παραπάνω διάγραμμα δείχνει τις προβλέψεις πωλήσεων με το μοντέλο ARIMA για κάθε νομό τους επόμενους 12 μήνες. Κάθε γραμμή αντιπροσωπεύει έναν διαφορετικό νομό, απεικονίζοντας την ποικιλομορφία στις προβλεπόμενες τάσεις πωλήσεων μεταξύ των περιοχών.

```
# Assuming 'data_prefecture' is your DataFrame Loaded with prefecture sales data
prefectures = data_prefecture['Prefecture'].unique()

# Initialize a figure for plotting
plt.figure(figsize=(15, 10))

# Loop through each prefecture to fit an ARIMA model and forecast sales
for prefecture in prefectures:
    prefecture_data = data_prefecture[data_prefecture['Prefecture'] == prefecture].set_index('Date')['Sum']
    # Fit an ARIMA model
    arima_model = ARIMA(prefecture_data, order=(1, 1, 1)).fit()
    # Forecast the next 12 months
    arima_forecast = arima_model.forecast(steps=12)

    # Generate dates for plotting the forecast
    forecast_dates = pd.date_range(start=data_prefecture['Date'].max() + pd.DateOffset(months=1, periods=12, freq='MS')
    # Plot the forecast
    plt.plot(forecast_dates, arima_forecast, label=prefecture)

plt.title('ARIMA Sales Forecast for Each Prefecture')
plt.xlabel('Date')
plt.ylabel('Sales')
plt.legend(loc='upper left', bbox_to_anchor=(1, 1)) # Adjust the Legend position
plt.grid(True)
plt.show()
```

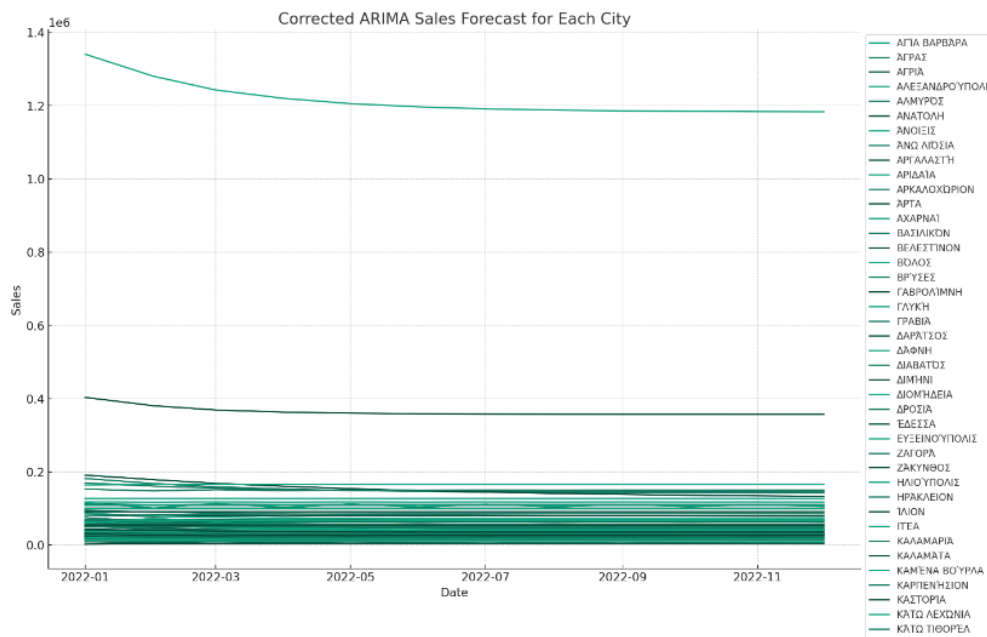
Εικόνα 35 "Κώδικας Python - Μοντέλο ARIMA ανά Νομό"

Τα αναλυτικά αποτελέσματα της πρόβλεψης ανά μήνα και ανά Νομό εμφανίζονται παρακάτω:

	1/1/2022	1/2/2022	1/3/2022	1/4/2022	1/5/2022	1/6/2022	1/7/2022	1/8/2022	1/9/2022	1/10/2022	1/11/2022	1/12/2022
ΑΙΤΩΛΙΑΣ ΚΑΙ ΑΚΑΡΝΑΝΙΑ	66.696 €	65.268 €	64.410 €	63.895 €	63.585 €	63.399 €	63.287 €	63.219 €	63.179 €	63.155 €	63.140 €	63.131 €
ΑΡΤΗΣ	99.268 €	98.893 €	98.858 €	98.855 €	98.855 €	98.855 €	98.855 €	98.855 €	98.855 €	98.855 €	98.855 €	98.855 €
ΑΤΤΙΚΗΣ	823.330 €	822.121 €	822.607 €	822.411 €	822.490 €	822.459 €	822.471 €	822.466 €	822.468 €	822.467 €	822.468 €	822.468 €
ΑΧΑΪΑΣ	138.915 €	136.775 €	138.608 €	137.038 €	138.383 €	137.231 €	138.218 €	137.372 €	138.097 €	137.476 €	138.008 €	137.552 €
ΒΟΙΩΤΙΑΣ	19.240 €	19.563 €	19.421 €	19.483 €	19.456 €	19.468 €	19.463 €	19.465 €	19.464 €	19.465 €	19.464 €	19.464 €
ΔΩΔΕΚΑΝΗΣΟΥ	196.866 €	204.224 €	197.978 €	203.280 €	198.779 €	202.600 €	199.356 €	202.109 €	199.773 €	201.756 €	200.072 €	201.502 €
ΕΒΡΟΥ	64.264 €	65.640 €	65.996 €	66.088 €	66.112 €	66.118 €	66.119 €	66.120 €	66.120 €	66.120 €	66.120 €	66.120 €
ΕΥΒΟΙΑΣ	293.782 €	293.608 €	293.616 €	293.616 €	293.616 €	293.616 €	293.616 €	293.616 €	293.616 €	293.616 €	293.616 €	293.616 €
ΕΥΡΥΤΑΝΙΑΣ	15.967 €	15.848 €	15.773 €	15.724 €	15.693 €	15.674 €	15.661 €	15.653 €	15.648 €	15.644 €	15.642 €	15.641 €
ΖΑΚΥΝΘΟΥ	20.190 €	21.722 €	20.395 €	21.544 €	20.549 €	21.410 €	20.665 €	21.310 €	20.751 €	21.235 €	20.816 €	21.179 €
ΗΜΑΘΙΑΣ	78.355 €	75.576 €	74.066 €	73.245 €	72.798 €	72.556 €	72.424 €	72.352 €	72.313 €	72.292 €	72.280 €	72.274 €
ΗΡΑΚΛΕΙΟΥ	316.315 €	290.864 €	272.567 €	259.413 €	249.956 €	243.157 €	238.269 €	234.755 €	232.228 €	230.412 €	229.106 €	228.167 €
ΘΕΣΠΡΟΤΙΑΣ	21.563 €	20.447 €	19.729 €	19.266 €	18.968 €	18.776 €	18.653 €	18.573 €	18.522 €	18.489 €	18.468 €	18.454 €
ΘΕΣΣΑΛΟΝΙΚΗΣ	35.539 €	33.957 €	33.721 €	33.685 €	33.680 €	33.679 €	33.679 €	33.679 €	33.679 €	33.679 €	33.679 €	33.679 €
ΙΩΑΝΝΙΝΩΝ	90.449 €	88.669 €	89.177 €	89.032 €	89.073 €	89.062 €	89.065 €	89.064 €	89.064 €	89.064 €	89.064 €	89.064 €
ΚΑΣΤΟΡΙΑΣ	35.537 €	37.969 €	36.012 €	37.586 €	36.320 €	37.339 €	36.519 €	37.179 €	36.648 €	37.075 €	36.731 €	37.008 €
ΚΕΦΑΛΛΗΝΙΑΣ	32.749 €	32.777 €	32.774 €	32.774 €	32.774 €	32.774 €	32.774 €	32.774 €	32.774 €	32.774 €	32.774 €	32.774 €
ΚΟΖΑΝΗΣ	182.078 €	168.412 €	159.499 €	153.686 €	149.895 €	147.423 €	145.811 €	144.759 €	144.073 €	143.626 €	143.334 €	143.144 €
ΜΑΓΝΗΣΙΑΣ	2.477.660 €	2.353.408 €	2.274.967 €	2.225.447 €	2.194.185 €	2.174.449 €	2.161.989 €	2.154.124 €	2.149.158 €	2.146.023 €	2.144.044 €	2.142.795 €
ΜΕΣΣΗΝΙΑΣ	56.979 €	57.154 €	57.279 €	57.368 €	57.432 €	57.478 €	57.510 €	57.533 €	57.550 €	57.561 €	57.570 €	57.576 €
ΞΑΝΘΗΣ	187.516 €	177.250 €	170.495 €	166.051 €	163.126 €	161.202 €	159.936 €	159.103 €	158.555 €	158.194 €	157.957 €	157.801 €
ΠΕΛΛΗΣ	162.557 €	179.841 €	164.495 €	178.121 €	166.022 €	176.765 €	167.226 €	175.696 €	168.175 €	174.853 €	168.924 €	174.188 €
ΠΙΕΡΙΑΣ	28.452 €	27.941 €	28.228 €	28.067 €	28.157 €	28.107 €	28.135 €	28.119 €	28.128 €	28.123 €	28.126 €	28.124 €
ΠΡΕΒΕΖΗΣ	124.621 €	128.227 €	125.003 €	127.886 €	125.308 €	127.613 €	125.552 €	127.394 €	125.747 €	127.220 €	125.903 €	127.080 €
ΡΕΘΥΜΝΗΣ	59.445 €	64.572 €	62.079 €	63.291 €	62.702 €	62.988 €	62.849 €	62.917 €	62.884 €	62.900 €	62.892 €	62.896 €
ΡΟΔΩΠΗΣ	69.229 €	62.097 €	57.931 €	55.498 €	54.077 €	53.247 €	52.763 €	52.479 €	52.314 €	52.218 €	52.161 €	52.128 €
ΤΡΙΚΑΛΩΝ	49.244 €	48.198 €	48.009 €	47.975 €	47.969 €	47.969 €	47.968 €	47.968 €	47.968 €	47.968 €	47.968 €	47.968 €
ΦΘΙΩΤΙΑΔΟΣ	133.165 €	130.657 €	129.117 €	128.173 €	127.593 €	127.237 €	127.019 €	126.885 €	126.802 €	126.752 €	126.721 €	126.702 €
ΦΛΩΡΙΝΗΣ	168.249 €	172.656 €	169.744 €	171.668 €	170.397 €	171.237 €	170.682 €	171.048 €	170.806 €	170.966 €	170.860 €	170.930 €
ΦΟΚΙΔΟΣ	112.443 €	103.403 €	98.909 €	96.674 €	95.563 €	95.011 €	94.737 €	94.600 €	94.532 €	94.498 €	94.482 €	94.473 €
ΧΑΝΙΩΝ	335.902 €	337.450 €	337.178 €	337.226 €	337.217 €	337.219 €	337.219 €	337.219 €	337.219 €	337.219 €	337.219 €	337.219 €
ΧΙΟΥ	38.047 €	35.622 €	36.942 €	36.224 €	36.615 €	36.402 €	36.518 €	36.455 €	36.489 €	36.470 €	36.481 €	36.475 €

Εικόνα 36 "Αποτελέσματα μεθόδου ARIMA ανά μήνα & ανά Νομό "

Παρακάτω η ανάλυση σε πλαίσιο πόλης:



- ΚΑΙΟ ΤΙΘΟΡΕΑ
- ΚΕΦΑΛΟΒΡΥΣΙΟΝ
- ΚΗΜΕΡΙΑ
- ΚΟΖΑΝΗ
- ΚΟΛΙΝΔΡΟΣ
- ΚΡΙΚΕΛΛΟΣ
- ΚΡΥΦΟΒΟΝ
- ΚΥΜΗ
- ΚΙΣ
- ΛΑΜΙΑ
- ΛΑΡΒΟΣ
- ΜΑΥΡΟΜΜΑΤΙΟΝ
- ΜΟΙΡΑΙ
- ΜΟΙΡΑΙΕΣ
- ΜΑΟΥΣΑ
- ΝΕΑ ΑΓΧΙΑΛΟΣ
- ΝΕΑ ΙΩΝΙΑ
- ΝΕΑ ΜΑΚΡΗ
- ΝΕΑ ΣΤΥΡΑ
- ΝΕΟ ΧΑΛΚΙΔΟΥΣΑΟ
- ΝΕΟΝ ΨΥΧΙΚΩΝ
- ΞΑΝΘΗ
- ΠΑΤΗΤΗΡΙΟΝ
- ΠΑΤΡΑΙ
- ΠΕΔΙΝΗ
- ΠΕΡΑΕΥΣ
- ΠΕΡΑΜΑ
- ΠΕΡΙΣΤΕΡΙΟΝ
- ΠΕΤΡΟΥΠΟΛΙΣ
- ΠΙΡΟΣ
- ΠΟΡΤΑΡΙΑ
- ΠΡΕΒΕΖΑ
- ΡΙΣΟΜΥΛΟΣ
- ΣΑΡΑΙ
- ΣΕΛΑΝΤΙΚΑ
- ΣΚΙΑΘΟΣ
- ΣΚΟΠΕΛΟΣ
- ΣΟΥΡΠΗ
- ΣΤΕΦΑΝΟΒΙΚΕΙΟΝ
- ΤΡΙΚΑΛΑ
- ΦΑΙΣΤΙΔΑΣ
- ΦΑΙΣΤΑΣ
- ΦΑΙΣΤΑΣ
- ΦΑΙΣΤΑΣ
- ΦΑΙΣΤΑΣ
- ΧΑΛΚΙΣ
- ΧΑΝΙΑ
- ΧΙΟΣ

Εικόνα 37 "Αποτελέσματα πρόβλεψης ARIMA ανά Πόλη"

```
# Forecasting and plotting for each city
for city in cities:
    city_data = data_city[data_city['City'] == city].set_index('Date')[correct_column_name]
    # Fit the ARIMA model
    arima_model = ARIMA(city_data, order=(1, 1, 1)).fit()
    # Forecast the next 12 months
    forecast = arima_model.forecast(steps=12)

    # Generate dates for plotting the forecast
    forecast_dates = pd.date_range(start=data_city['Date'].max() + pd.DateOffset(months=1), periods=12, freq='MS')
    # Plot the forecast
    plt.plot(forecast_dates, forecast, label=city)

plt.title('Corrected ARIMA Sales Forecast for Each City')
plt.xlabel('Date')
plt.ylabel('Sales')
plt.legend(loc='upper left', bbox_to_anchor=(1, 1)) # Adjust the Legend position
plt.grid(True)
plt.show()
```

Εικόνα 30 "Κώδικας Python - μοντέλο ARIMA ανά Πόλη"

Παρακάτω εμφανίζεται ένα δείγμα των αποτελεσμάτων που παράχθηκαν από την πρόβλεψη των πωλήσεων για κάθε πόλη με τη μέθοδο ARIMA:

	1/1/2022	1/2/2022	1/3/2022	1/4/2022	1/5/2022	1/6/2022	1/7/2022	1/8/2022	1/9/2022	1/10/2022	1/11/2022	1/12/2022
ΑΠΑ ΒΑΡΒΑΡΑ	117.966 €	116.924 €	117.799 €	117.064 €	117.682 €	117.163 €	117.599 €	117.233 €	117.540 €	117.282 €	117.499 €	117.317 €
ΑΓΡΑΣ	22.379 €	23.244 €	23.955 €	24.541 €	25.023 €	25.420 €	25.746 €	26.015 €	26.237 €	26.419 €	26.569 €	26.692 €
ΑΓΡΙΑ	50.839 €	45.621 €	42.990 €	41.662 €	40.993 €	40.655 €	40.485 €	40.399 €	40.356 €	40.334 €	40.323 €	40.318 €
ΑΛΕΞΑΝΔΡΟΥ	56.346 €	54.828 €	53.939 €	53.419 €	53.115 €	52.937 €	52.833 €	52.772 €	52.737 €	52.716 €	52.703 €	52.696 €
ΑΛΜΥΡΟΣ	112.570 €	110.445 €	111.365 €	110.967 €	111.139 €	111.065 €	111.097 €	111.083 €	111.089 €	111.086 €	111.087 €	111.087 €
ΑΝΑΤΟΛΗ	31.889 €	33.233 €	32.185 €	33.002 €	32.366 €	32.862 €	32.475 €	32.776 €	32.542 €	32.725 €	32.582 €	32.693 €
ΑΝΟΙΣΣ	69.017 €	63.536 €	59.678 €	56.962 €	55.051 €	53.706 €	52.760 €	52.093 €	51.624 €	51.294 €	51.062 €	50.899 €
ΑΝΩ ΛΙΟΣΙΑ	66.180 €	66.267 €	66.240 €	66.248 €	66.245 €	66.246 €	66.246 €	66.246 €	66.246 €	66.246 €	66.246 €	66.246 €
ΑΡΓΑΛΑΣΤΗ	9.450 €	4.894 €	9.163 €	5.163 €	8.911 €	5.399 €	8.689 €	5.607 €	8.495 €	5.789 €	8.324 €	5.949 €
ΑΡΙΔΑΙΑ	114.866 €	101.201 €	113.411 €	102.501 €	112.249 €	103.539 €	111.322 €	104.367 €	110.582 €	105.029 €	109.991 €	105.557 €
ΑΡΚΑΛΟΧΩΡΙΟ	66.952 €	64.950 €	64.418 €	64.277 €	64.240 €	64.230 €	64.227 €	64.227 €	64.226 €	64.226 €	64.226 €	64.226 €
ΑΡΤΑ	101.354 €	101.022 €	100.967 €	100.958 €	100.956 €	100.956 €	100.956 €	100.956 €	100.956 €	100.956 €	100.956 €	100.956 €
ΑΧΑΡΝΑΙ	56.359 €	51.150 €	47.529 €	45.012 €	43.263 €	42.047 €	41.201 €	40.613 €	40.205 €	39.921 €	39.724 €	39.586 €
ΒΑΣΙΛΙΚΟΝ	94.640 €	90.429 €	87.294 €	84.960 €	83.222 €	81.929 €	80.966 €	80.249 €	79.715 €	79.318 €	79.023 €	78.802 €
ΒΕΛΕΣΤΙΝΟΝ	81.157 €	78.542 €	80.285 €	79.123 €	79.897 €	79.381 €	79.725 €	79.496 €	79.649 €	79.547 €	79.615 €	79.569 €
ΒΟΛΟΣ	1.340.563 €	1.280.201 €	1.242.901 €	1.219.853 €	1.205.611 €	1.196.811 €	1.191.373 €	1.188.013 €	1.185.936 €	1.184.653 €	1.183.860 €	1.183.370 €
ΒΡΥΞΣ	69.857 €	69.807 €	69.847 €	69.814 €	69.841 €	69.819 €	69.837 €	69.823 €	69.834 €	69.825 €	69.833 €	69.826 €
ΓΑΒΡΟΛΙΜΝΗ	2.609 €	3.611 €	4.025 €	4.196 €	4.267 €	4.296 €	4.308 €	4.313 €	4.316 €	4.316 €	4.317 €	4.317 €
ΓΛΥΚΗ	21.609 €	20.528 €	19.834 €	19.389 €	19.104 €	18.921 €	18.804 €	18.729 €	18.680 €	18.649 €	18.629 €	18.617 €
ΓΡΑΒΙΑ	24.892 €	24.124 €	23.651 €	23.361 €	23.182 €	23.072 €	23.005 €	22.964 €	22.938 €	22.922 €	22.913 €	22.907 €
ΔΑΡΑΤΣΟΣ	55.913 €	56.078 €	56.090 €	56.091 €	56.091 €	56.091 €	56.091 €	56.091 €	56.091 €	56.091 €	56.091 €	56.091 €
ΔΑΦΝΗ	22.250 €	21.318 €	20.650 €	20.172 €	19.829 €	19.583 €	19.406 €	19.280 €	19.190 €	19.125 €	19.078 €	19.045 €
ΔΙΑΒΑΤΟΣ	37.718 €	36.918 €	36.340 €	35.920 €	35.616 €	35.396 €	35.236 €	35.120 €	35.037 €	34.976 €	34.932 €	34.900 €
ΔΙΜΗΝΙ	23.263 €	20.825 €	20.083 €	19.857 €	19.788 €	19.767 €	19.761 €	19.759 €	19.758 €	19.758 €	19.758 €	19.758 €
ΔΙΟΜΗΘΕΙΑ	53.351 €	50.146 €	48.296 €	47.229 €	46.613 €	46.258 €	46.053 €	45.934 €	45.866 €	45.827 €	45.804 €	45.791 €
ΔΡΟΣΙΑ	65.729 €	67.980 €	66.352 €	67.529 €	66.678 €	67.293 €	66.848 €	67.170 €	66.937 €	67.106 €	66.984 €	67.072 €

Εικόνα 38 "Αποτελέσματα μεθόδου ARIMA ανά μήνα & ανά Πόλη "

Παρακάτω η ανάλυση σε πλαίσιο καταστήματος:



Εικόνα 39 "Αποτελέσματα μεθόδου ARIMA σε επίπεδο καταστήματος "

```
# Container for storing forecast results
all_forecasts = {}

# Iterate through each unique code in the dataset
for code in train_data['Code'].unique():
    sales_data = train_data[train_data['Code'] == code]['Sales'].asfreq('MS')

    if not sales_data.empty:
        # Fit the ARIMA model
        model = ARIMA(sales_data, order=(1,1,1)) # Adjust ARIMA parameters as needed
        model_fit = model.fit()

        # Forecast the next 12 months
        forecast = model_fit.forecast(steps=12)

        # Store the forecast results
        all_forecasts[code] = forecast

# Convert the forecast results to a DataFrame
forecast_df = pd.DataFrame(all_forecasts)
```

Εικόνα 40 "Κώδικας Python μεθόδου ARIMA σε επίπεδο καταστήματος "

Παρακάτω εμφανίζεται ένα δείγμα των αποτελεσμάτων που παράχθηκαν από την πρόβλεψη των πωλήσεων για κάθε κατάσταση με τη μέθοδο ARIMA:

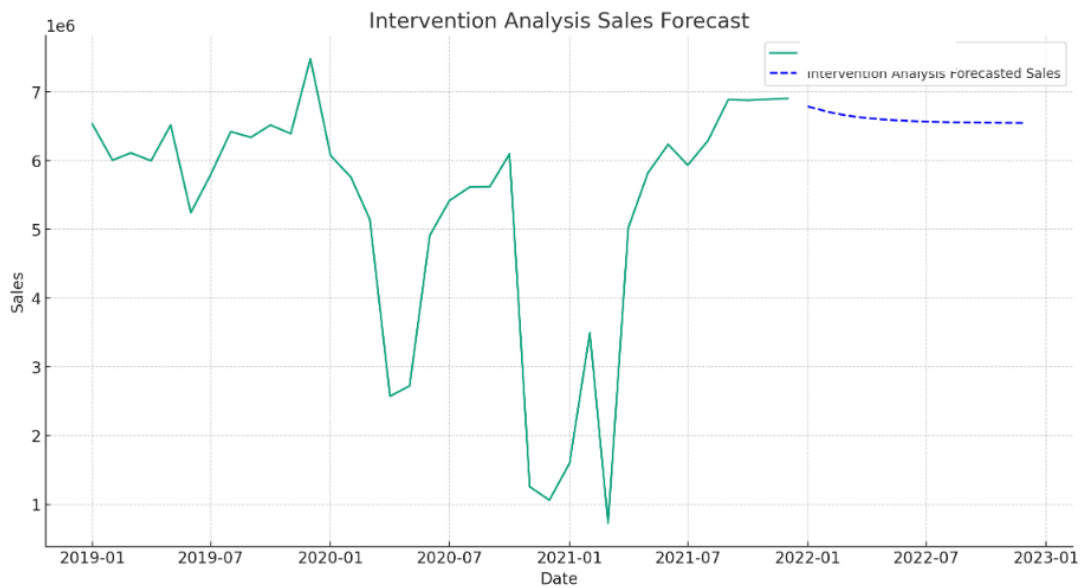
	1/1/2022	1/2/2022	1/3/2022	1/4/2022	1/5/2022	1/6/2022	1/7/2022	1/8/2022	1/9/2022	1/10/2022	1/11/2022	1/12/2022
100001	52.148 €	52.836 €	52.608 €	52.684 €	52.659 €	52.667 €	52.664 €	52.665 €	52.665 €	52.665 €	52.665 €	52.665 €
100002	86.872 €	90.474 €	87.277 €	90.114 €	87.596 €	89.831 €	87.847 €	89.608 €	88.045 €	89.432 €	88.201 €	89.294 €
100003	65.725 €	60.360 €	56.928 €	54.733 €	53.328 €	52.429 €	51.854 €	51.486 €	51.250 €	51.000 €	51.003 €	50.942 €
100004	21.563 €	20.447 €	19.729 €	19.266 €	18.968 €	18.776 €	18.653 €	18.573 €	18.522 €	18.489 €	18.468 €	18.454 €
100005	35.999 €	37.197 €	37.492 €	37.565 €	37.583 €	37.587 €	37.588 €	37.588 €	37.588 €	37.588 €	37.588 €	37.588 €
100006	30.095 €	28.889 €	29.454 €	29.189 €	29.313 €	29.255 €	29.282 €	29.270 €	29.276 €	29.279 €	29.274 €	29.274 €
100008	35.539 €	33.957 €	33.721 €	33.685 €	33.680 €	33.679 €	33.679 €	33.679 €	33.679 €	33.679 €	33.679 €	33.679 €
100009	59.445 €	64.572 €	62.079 €	63.291 €	62.702 €	62.988 €	62.849 €	62.917 €	62.884 €	62.900 €	62.892 €	62.896 €
100014	69.016 €	63.580 €	59.788 €	57.142 €	55.295 €	54.007 €	53.108 €	52.481 €	52.043 €	51.738 €	51.525 €	51.376 €
100015	32.749 €	32.777 €	32.774 €	32.774 €	32.774 €	32.774 €	32.774 €	32.774 €	32.774 €	32.774 €	32.774 €	32.774 €
100016	34.334 €	33.661 €	33.493 €	33.451 €	33.441 €	33.438 €	33.437 €	33.437 €	33.437 €	33.437 €	33.437 €	33.437 €
100022	73.498 €	70.848 €	72.744 €	71.387 €	72.358 €	71.663 €	72.161 €	71.805 €	72.060 €	71.877 €	72.008 €	71.914 €
100023	42.874 €	43.694 €	44.267 €	44.666 €	44.945 €	45.139 €	45.274 €	45.369 €	45.435 €	45.481 €	45.513 €	45.536 €
100031	65.613 €	65.819 €	65.769 €	65.781 €	65.778 €	65.779 €	65.778 €	65.779 €	65.779 €	65.779 €	65.779 €	65.779 €
100032	52.700 €	47.696 €	44.364 €	42.145 €	40.667 €	39.683 €	39.028 €	38.591 €	38.301 €	38.107 €	37.978 €	37.892 €
100037	26.909 €	26.854 €	26.891 €	26.866 €	26.883 €	26.872 €	26.879 €	26.874 €	26.878 €	26.875 €	26.877 €	26.876 €
100039	53.226 €	53.206 €	53.205 €	53.205 €	53.205 €	53.205 €	53.205 €	53.205 €	53.205 €	53.205 €	53.205 €	53.205 €
100046	36.300 €	35.671 €	35.394 €	35.271 €	35.217 €	35.199 €	35.182 €	35.178 €	35.176 €	35.175 €	35.174 €	35.174 €
100048	28.452 €	27.941 €	28.228 €	28.067 €	28.157 €	28.107 €	28.135 €	28.119 €	28.128 €	28.123 €	28.126 €	28.124 €
100049	28.326 €	27.274 €	27.138 €	27.120 €	27.118 €	27.118 €	27.118 €	27.118 €	27.118 €	27.118 €	27.118 €	27.118 €
100060	76.405 €	65.991 €	59.936 €	56.415 €	54.368 €	53.178 €	52.486 €	52.083 €	51.849 €	51.713 €	51.634 €	51.588 €
100062	38.313 €	38.313 €	38.313 €	38.313 €	38.313 €	38.313 €	38.313 €	38.313 €	38.313 €	38.313 €	38.313 €	38.313 €
100063	20.190 €	21.722 €	20.395 €	21.544 €	20.549 €	21.410 €	20.665 €	21.310 €	20.751 €	21.235 €	20.816 €	21.179 €
100068	14.866 €	14.848 €	14.846 €	14.845 €	14.845 €	14.845 €	14.845 €	14.845 €	14.845 €	14.845 €	14.845 €	14.845 €
100071	101.509 €	106.972 €	102.212 €	106.360 €	102.745 €	105.895 €	103.151 €	105.542 €	103.458 €	105.274 €	103.692 €	105.070 €
100089	14.870 €	13.482 €	12.385 €	11.519 €	10.834 €	10.294 €	9.867 €	9.530 €	9.264 €	9.053 €	8.887 €	8.756 €
100091	32.651 €	33.245 €	32.878 €	33.105 €	32.965 €	33.051 €	32.998 €	33.031 €	33.010 €	33.023 €	33.015 €	33.020 €
100092	22.566 €	22.744 €	22.613 €	22.709 €	22.639 €	22.691 €	22.653 €	22.680 €	22.660 €	22.675 €	22.664 €	22.672 €
100096	17.788 €	17.975 €	17.931 €	17.941 €	17.939 €	17.940 €	17.939 €	17.939 €	17.939 €	17.939 €	17.939 €	17.939 €

Εικόνα 41 "Αποτελέσματα μεθόδου ARIMA ανά μήνα & ανά Κατάστημα "

Πρόβλεψη με τη μέθοδο Intervention Analysis:

Η μέθοδος intervention analysis, που συχνά διεξάγεται με τη χρήση μοντέλων όπως το ARIMA με συνιστώσες παρέμβασης, χρησιμοποιείται για την αξιολόγηση του αντίκτυπου ενός εξωτερικού συμβάντος ή μιας παρέμβασης σε μια χρονοσειρά. Για να πραγματοποιήσουμε την ανάλυση παρέμβασης χρειαζόμαστε ένα σύνολο δεδομένων χρονοσειρών και πληροφορίες σχετικά με το χρόνο και τη φύση των παρεμβάσεων. Στη δίκη μας περίπτωση θα λάβουμε ως συνιστώσες παρέμβασης τον Covid-19.

Η ανάλυση περιλαμβάνει την προσαρμογή ενός μοντέλου χρονοσειρών θεωρώντας την παρέμβαση ως εξωγενή μεταβλητή. Αυτό θα μας βοηθήσει να αξιολογήσουμε τον αντίκτυπο της παρέμβασης στις πωλήσεις.



Εικόνα 42 "Πρόβλεψη με τη μέθοδο Intervention Analysis"

Στο παραπάνω διάγραμμα η συμπαγής γραμμή αντιπροσωπεύει τα πραγματικά ιστορικά δεδομένα πωλήσεων, ενώ η διακεκομμένη γραμμή αντιπροσωπεύει τις προβλεπόμενες πωλήσεις. Παρακάτω εμφανίζεται ο κώδικας που χρησιμοποιήθηκε καθώς και τα αναλυτικά αποτελέσματα:

```
# Define the ARIMA model parameters
p, d, q = 1, 1, 1 # Basic ARIMA(1,1,1) model configuration

# Fit the ARIMA model with intervention as an exogenous variable
arima_with_intervention_model = ARIMA(train_data['Sales'], exog=train_data[['intervention']], order=(p, d, q)).fit()

# Prepare the exogenous variable for forecasting (assuming no intervention in the forecast period)
future_intervention = pd.DataFrame({'intervention': [0] * 12}, index=forecast_dates)

# Forecast the next 12 months with the intervention model
arima_with_intervention_forecast = arima_with_intervention_model.get_forecast(steps=12, exog=future_intervention)
forecasted_values_with_intervention = arima_with_intervention_forecast.predicted_mean

# Creating a DataFrame for the forecasted values with intervention analysis
intervention_forecast_df = pd.DataFrame(forecasted_values_with_intervention.values, index=forecast_dates, columns=['Intervention Analysis Forecasted Sales'])

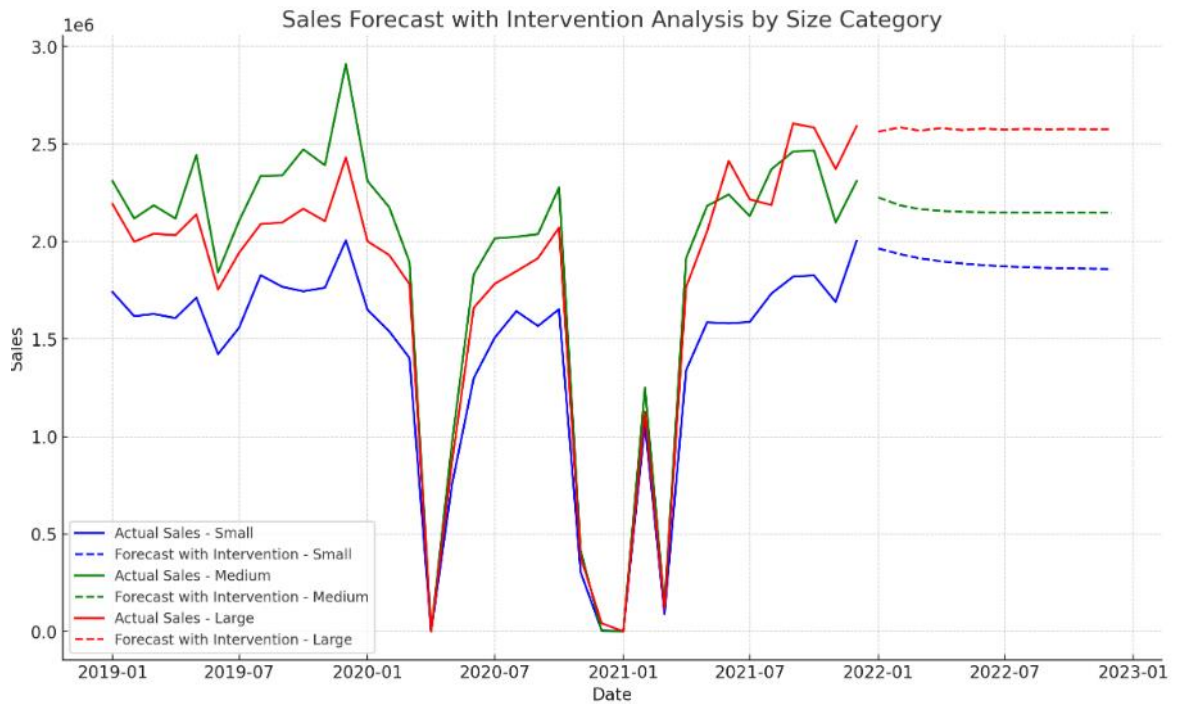
# Plotting the historical sales and forecasted values with intervention analysis
plt.figure(figsize=(14, 7))
plt.plot(train_data.index, train_data['Sales'], label='Actual Sales')
plt.plot(intervention_forecast_df.index, intervention_forecast_df['Intervention Analysis Forecasted Sales'], label='Intervention Analysis Forecasted Sales', linestyle='--', color='blue')
plt.title('Intervention Analysis Sales Forecast')
plt.xlabel('Date')
plt.ylabel('Sales')
plt.legend()
plt.grid(True)
plt.show()
```

Εικόνα 43 "Κώδικας Python - Intervention Analysis"

Date	Sales Forecast
31/1/2022	6395338
28/2/2022	6020779
31/3/2022	5745554
30/4/2022	5543319
31/5/2022	5394717
30/6/2022	5285525
31/7/2022	5205290

31/8/2022	5146334
30/9/2022	5103013
31/10/2022	5071181
30/11/2022	5047791
31/12/2022	5030604

Παρακάτω η ανάλυση για κάθε κατηγορία μεγέθους:



Εικόνα 44 "Αποτελέσματα πρόβλεψης με την μέθοδο Intervention Analysis για κάθε κατηγορία μεγέθους"

```
# Colors for plotting
colors = {'Small': 'blue', 'Medium': 'green', 'Large': 'red'}

plt.figure(figsize=(14, 8))

# Conducting ARIMA intervention analysis and plotting for each size category
for size in ['Small', 'Medium', 'Large']:
    size_data = data_size[data_size['Size'] == size].set_index('Date')
    exog = size_data[['Intervention']]
    model = ARIMA(endog=size_data['Sales'], exog=exog, order=(1,1,1))
    fit = model.fit()
    future_intervention = pd.DataFrame({'Intervention': [0]*12}, index=forecast_dates)
    forecast = fit.get_forecast(steps=12, exog=future_intervention)
    forecasted_sales = forecast.predicted_mean

# Plot actual sales
plt.plot(size_data.index, size_data['Sales'], label=f'Actual Sales - {size}', color=colors[size])

# Plot forecasted sales with intervention
plt.plot(forecast_dates, forecasted_sales, linestyle='--', label=f'Forecast with Intervention - {size}', color=colors[size])

plt.title('Sales Forecast with Intervention Analysis by Size Category')
plt.xlabel('Date')
plt.ylabel('Sales')
plt.legend()
plt.grid(True)
plt.show()
```

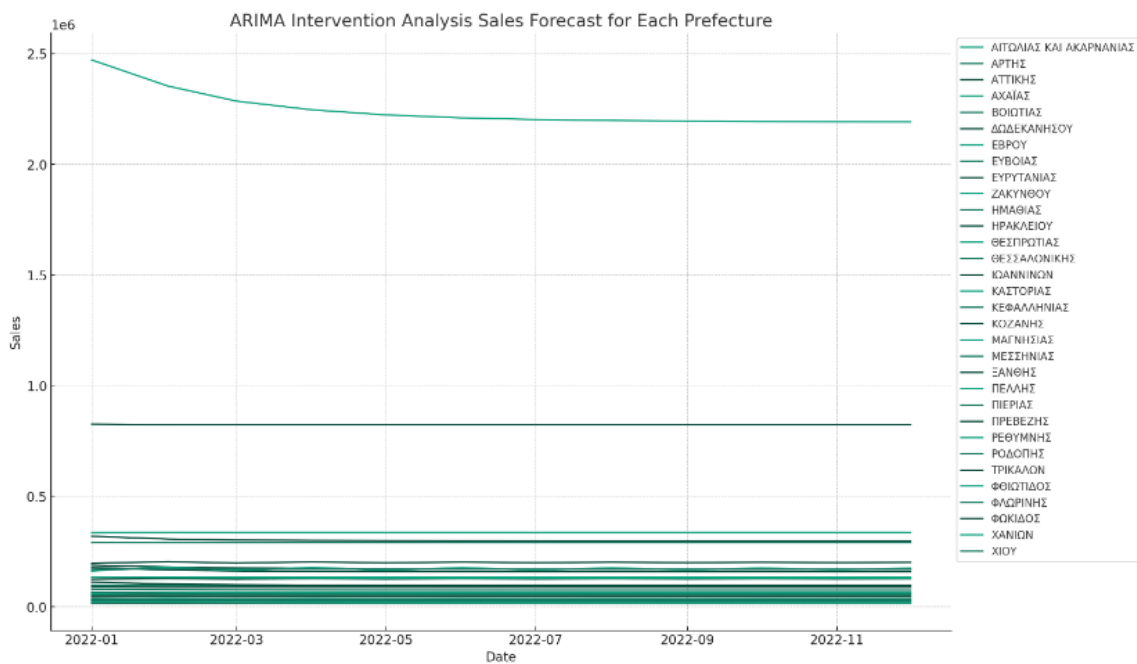
Εικόνα 45 "Αποτελέσματα πρόβλεψης με την μέθοδο Intervention Analysis για κάθε κατηγορία μεγέθους"

Τα αναλυτικά αποτελέσματα της πρόβλεψης ανά μήνα και ανά κατηγορία μεγέθους

καταστήματος εμφανίζονται παρακάτω:

Date	Small Size Intervention Forecast	Medium Size Intervention Forecast	Large Size Intervention Forecast
1/1/2022	1,964,312	2,226,115	2,563,673
1/2/2022	1,935,392	2,185,852	2,585,287
1/3/2022	1,914,168	2,166,416	2,568,701
1/4/2022	1,898,592	2,157,034	2,581,428
1/5/2022	1,887,160	2,152,505	2,571,662
1/6/2022	1,878,771	2,150,319	2,579,156
1/7/2022	1,872,613	2,149,264	2,573,405
1/8/2022	1,868,095	2,148,755	2,577,818
1/9/2022	1,864,778	2,148,509	2,574,432
1/10/2022	1,862,344	2,148,390	2,577,030
1/11/2022	1,860,558	2,148,333	2,575,036
1/12/2022	1,859,247	2,148,305	2,576,566

Παρακάτω η ανάλυση σε πλαίσιο νομού:



Εικόνα 46 "Αποτελέσματα πρόβλεψης με την μέθοδο Intervention Analysis για κάθε Νομό"

```
# Initialize a figure for plotting
plt.figure(figsize=(15, 10))

# Loop through each prefecture to conduct intervention analysis with ARIMA
for prefecture in prefectures:
    prefecture_data = data_prefecture[data_prefecture['Prefecture'] == prefecture].set_index('Date')
    exog = prefecture_data[['Intervention']]
    # Fit ARIMA model with 'Intervention' as exogenous variable
    model = ARIMA(endog=prefecture_data['Sum'], exog=exog, order=(1, 1, 1))
    fit = model.fit()
    # Assume no future interventions for the forecast period
    future_intervention = pd.DataFrame({'Intervention': [0] * 12}, index=forecast_dates)
    forecast = fit.get_forecast(steps=12, exog=future_intervention)
    forecasted_sales = forecast.predicted_mean

# Plot the forecast
plt.plot(forecast_dates, forecasted_sales, label=prefecture)

plt.title('ARIMA Intervention Analysis Sales Forecast for Each Prefecture')
plt.xlabel('Date')
plt.ylabel('Sales')
plt.legend(loc='upper left', bbox_to_anchor=(1, 1)) # Adjust the Legend position
plt.grid(True)
plt.show()
```

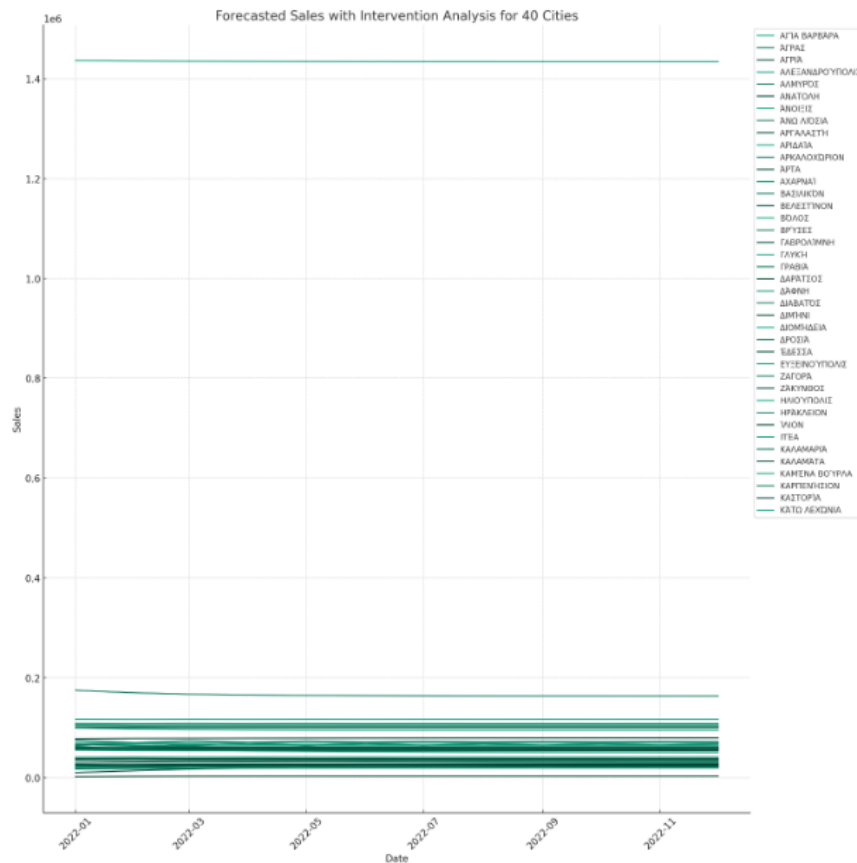
Εικόνα 47 "Κώδικας Python - Intervention Analysis για κάθε Νομό"

Τα αναλυτικά αποτελέσματα της πρόβλεψης ανά μήνα και ανά Νομό εμφανίζονται παρακάτω:

	1/1/2022	1/2/2022	1/3/2022	1/4/2022	1/5/2022	1/6/2022	1/7/2022	1/8/2022	1/9/2022	1/10/2022	1/11/2022	1/12/2022
ΑΙΤΩΛΙΑΣ ΚΑΙ ΑΚΑΡΝ	67.233 €	66.215 €	65.651 €	65.339 €	65.167 €	65.071 €	65.018 €	64.989 €	64.973 €	64.964 €	64.959 €	64.956 €
ΑΡΤΗΣ	98.151 €	97.789 €	97.764 €	97.762 €	97.762 €	97.762 €	97.762 €	97.762 €	97.762 €	97.762 €	97.762 €	97.762 €
ΑΤΤΙΚΗΣ	825.743 €	823.790 €	824.494 €	824.240 €	824.332 €	824.299 €	824.310 €	824.306 €	824.308 €	824.307 €	824.307 €	824.307 €
ΑΧΑΪΑΣ	133.354 €	133.776 €	133.718 €	133.726 €	133.725 €	133.725 €	133.725 €	133.725 €	133.725 €	133.725 €	133.725 €	133.725 €
ΒΟΙΩΤΙΑΣ	19.290 €	19.582 €	19.458 €	19.511 €	19.488 €	19.498 €	19.494 €	19.495 €	19.495 €	19.495 €	19.495 €	19.495 €
ΔΩΔΕΚΑΝΗΣΟΥ	196.856 €	204.225 €	197.968 €	203.280 €	198.769 €	202.600 €	199.347 €	202.109 €	199.764 €	201.755 €	200.064 €	201.500 €
ΕΒΡΟΥ	65.365 €	66.415 €	66.586 €	66.614 €	66.619 €	66.619 €	66.620 €	66.620 €	66.620 €	66.620 €	66.620 €	66.620 €
ΕΥΒΟΙΑΣ	291.080 €	290.700 €	290.816 €	290.781 €	290.791 €	290.788 €	290.789 €	290.789 €	290.789 €	290.789 €	290.789 €	290.789 €
ΕΥΡΥΤΑΝΙΑΣ	16.124 €	16.106 €	16.096 €	16.090 €	16.086 €	16.084 €	16.082 €	16.081 €	16.081 €	16.081 €	16.081 €	16.080 €
ΖΑΚΥΝΘΟΥ	20.220 €	21.724 €	20.423 €	21.548 €	20.575 €	21.416 €	20.689 €	21.318 €	20.774 €	21.244 €	20.838 €	21.189 €
ΗΡΑΚΛΕΙΑΣ	79.066 €	77.100 €	76.220 €	75.828 €	75.652 €	75.573 €	75.538 €	75.523 €	75.516 €	75.512 €	75.511 €	75.510 €
ΗΡΑΚΛΕΙΟΥ	318.893 €	306.426 €	301.691 €	299.892 €	299.209 €	298.949 €	298.850 €	298.813 €	298.799 €	298.793 €	298.791 €	298.791 €
ΘΕΣΣΑΛΟΝΙΚΗΣ	21.566 €	20.488 €	19.814 €	19.393 €	19.131 €	18.967 €	18.864 €	18.800 €	18.760 €	18.736 €	18.720 €	18.710 €
ΘΕΣΣΑΛΟΝΙΚΗΣ	35.561 €	34.224 €	34.054 €	34.033 €	34.030 €	34.030 €	34.030 €	34.030 €	34.030 €	34.030 €	34.030 €	34.030 €
ΙΩΑΝΝΙΝΩΝ	90.495 €	88.753 €	89.236 €	89.102 €	89.139 €	89.129 €	89.131 €	89.131 €	89.131 €	89.131 €	89.131 €	89.131 €
ΚΑΣΤΟΡΙΑΣ	35.493 €	37.968 €	35.970 €	37.583 €	36.281 €	37.332 €	36.484 €	37.168 €	36.616 €	37.062 €	36.702 €	36.992 €
ΚΕΦΑΛΛΗΝΙΑΣ	30.554 €	29.054 €	28.138 €	27.577 €	27.235 €	27.025 €	26.897 €	26.819 €	26.771 €	26.742 €	26.724 €	26.713 €
ΚΟΖΑΝΗΣ	177.408 €	166.712 €	162.248 €	160.384 €	159.606 €	159.281 €	159.146 €	159.089 €	159.066 €	159.056 €	159.052 €	159.050 €
ΜΑΓΝΗΣΙΑΣ	2.471.509 €	2.353.911 €	2.285.778 €	2.246.303 €	2.223.432 €	2.210.181 €	2.202.503 €	2.198.055 €	2.195.478 €	2.193.985 €	2.193.120 €	2.192.618 €
ΜΕΣΣΗΝΙΑΣ	55.896 €	56.222 €	56.095 €	56.145 €	56.125 €	56.133 €	56.130 €	56.131 €	56.131 €	56.131 €	56.131 €	56.131 €
ΞΑΝΘΗΣ	188.089 €	179.956 €	175.555 €	173.174 €	171.885 €	171.188 €	170.811 €	170.607 €	170.497 €	170.437 €	170.404 €	170.387 €
ΠΕΛΛΗΣ	162.118 €	179.925 €	163.995 €	178.246 €	165.497 €	176.902 €	166.699 €	175.827 €	167.661 €	174.966 €	168.431 €	174.278 €
ΠΕΡΙΑΣ	28.472 €	27.949 €	28.243 €	28.078 €	28.171 €	28.119 €	28.148 €	28.131 €	28.141 €	28.135 €	28.138 €	28.137 €
ΠΡΕΒΕΖΗΣ	124.401 €	128.162 €	124.837 €	127.777 €	125.178 €	127.475 €	125.444 €	127.240 €	125.652 €	127.056 €	125.815 €	126.912 €
ΡΕΘΥΜΝΗΣ	59.443 €	64.588 €	62.078 €	63.302 €	62.705 €	62.996 €	62.854 €	62.924 €	62.890 €	62.906 €	62.898 €	62.902 €
ΡΟΔΟΠΗΣ	66.824 €	61.006 €	58.690 €	57.769 €	57.402 €	57.256 €	57.198 €	57.175 €	57.165 €	57.162 €	57.160 €	57.160 €
ΤΡΙΚΑΛΩΝ	48.745 €	48.156 €	48.100 €	48.095 €	48.095 €	48.095 €	48.095 €	48.095 €	48.095 €	48.095 €	48.095 €	48.095 €
ΦΘΙΩΤΙΔΟΣ	134.098 €	132.326 €	131.331 €	130.772 €	130.458 €	130.281 €	130.182 €	130.127 €	130.095 €	130.078 €	130.068 €	130.062 €
ΦΛΩΡΙΝΗΣ	169.306 €	171.993 €	170.707 €	171.323 €	171.028 €	171.169 €	171.102 €	171.134 €	171.118 €	171.126 €	171.122 €	171.124 €
ΦΩΚΙΑΣ	111.329 €	103.230 €	99.831 €	98.404 €	97.805 €	97.554 €	97.449 €	97.404 €	97.386 €	97.378 €	97.375 €	97.373 €
ΧΑΝΙΩΝ	335.170 €	336.917 €	336.598 €	336.656 €	336.645 €	336.647 €	336.647 €	336.647 €	336.647 €	336.647 €	336.647 €	336.647 €
ΧΙΟΥ	38.225 €	35.602 €	37.088 €	36.247 €	36.723 €	36.453 €	36.606 €	36.519 €	36.568 €	36.541 €	36.556 €	36.547 €

Εικόνα 48 "Αποτελέσματα Intervention Analysis ανά μήνα & ανά Νομό "

Παρακάτω η ανάλυση σε πλαίσιο πόλης:



Εικόνα 49 "Αποτελέσματα μεθόδου Intervention Analysis ανά Πόλη "

```
# Loop through each city in the dataset
all_cities = data['City'].unique()
for city in all_cities:
    # Filter data for the current city
    city_data = data[data['City'] == city]
    sales_data = city_data['Sales'].asfreq('MS') # Assuming monthly data
    intervention_data = city_data['Intervention'].asfreq('MS').fillna(0)

    # Ensure there's enough data to fit the model
    if sales_data.dropna().shape[0] < 2:
        continue

    # Shape intervention data for forecast
    intervention_data_forecast = intervention_data.iloc[-1:].values.repeat(12).reshape(-1, 1)

    # Fit the SARIMAX model
    model = SARIMAX(sales_data, exog=intervention_data, order=(1, 1, 1), seasonal_order=(0, 0, 0, 12))
    model_fit = model.fit(dispatch=False)

    # Forecast the next 12 months
    forecast = model_fit.get_forecast(steps=12, exog=intervention_data_forecast)
    forecast_results_all_cities[city] = forecast.predicted_mean

# Compile all forecasts into a DataFrame
forecast_summary_all_cities = pd.DataFrame(forecast_results_all_cities)
```

Εικόνα 50 "Κώδικας Python – intervention analysis ανά Πόλη "

Παρακάτω εμφανίζεται ένα δείγμα των αποτελεσμάτων που παράχθηκαν από την πρόβλεψη των πωλήσεων για κάθε πόλη με τη μέθοδο Intervention Analysis:

	1/1/2022	1/2/2022	1/3/2022	1/4/2022	1/5/2022	1/6/2022	1/7/2022	1/8/2022	1/9/2022	1/10/2022	1/11/2022	1/12/2022
ΑΓΙΑ ΒΑΡΒΑΡΑ	117.153 €	116.780 €	117.105 €	116.821 €	117.069 €	116.853 €	117.041 €	116.877 €	117.020 €	116.895 €	117.004 €	116.909 €
ΆΓΡΑΣ	22.937 €	24.231 €	25.270 €	26.106 €	26.778 €	27.319 €	27.753 €	28.103 €	28.384 €	28.609 €	28.791 €	28.937 €
ΑΓΡΙΑ	60.743 €	60.755 €	60.755 €	60.755 €	60.755 €	60.755 €	60.755 €	60.755 €	60.755 €	60.755 €	60.755 €	60.755 €
ΑΛΕΞΑΝΔΡΟΥ	58.503 €	58.301 €	58.208 €	58.165 €	58.146 €	58.137 €	58.132 €	58.131 €	58.130 €	58.129 €	58.129 €	58.129 €
ΑΛΜΥΡΟΣ	108.274 €	107.780 €	108.177 €	107.858 €	108.115 €	107.909 €	108.074 €	107.941 €	108.048 €	107.962 €	108.031 €	107.976 €
ΑΝΑΤΟΛΗ	34.827 €	33.884 €	34.617 €	34.047 €	34.490 €	34.146 €	34.413 €	34.205 €	34.367 €	34.241 €	34.339 €	34.263 €
ΑΝΘΙΣΣΙΣ	69.690 €	65.236 €	62.447 €	60.702 €	59.610 €	58.926 €	58.498 €	58.230 €	58.062 €	57.957 €	57.891 €	57.850 €
ΑΝΘ ΛΙΟΣΙΑ	61.941 €	59.719 €	58.626 €	58.089 €	57.824 €	57.694 €	57.630 €	57.598 €	57.583 €	57.575 €	57.572 €	57.570 €
ΑΡΓΑΛΑΣΤΗ	10.082 €	14.152 €	17.166 €	19.398 €	21.051 €	22.275 €	23.182 €	23.853 €	24.350 €	24.718 €	24.991 €	25.193 €
ΑΡΙΔΑΙΑ	105.709 €	104.763 €	104.908 €	104.886 €	104.889 €	104.889 €	104.889 €	104.889 €	104.889 €	104.889 €	104.889 €	104.889 €
ΑΡΚΑΛΟΧΩΡΙΣ	66.478 €	63.720 €	62.772 €	62.445 €	62.333 €	62.294 €	62.281 €	62.277 €	62.275 €	62.274 €	62.274 €	62.274 €
ΑΡΤΑ	102.189 €	101.915 €	101.851 €	101.835 €	101.832 €	101.831 €	101.831 €	101.831 €	101.831 €	101.831 €	101.831 €	101.831 €
ΑΧΑΡΝΑΙ	59.840 €	56.997 €	54.982 €	53.555 €	52.543 €	51.826 €	51.318 €	50.958 €	50.703 €	50.522 €	50.394 €	50.303 €
ΒΑΣΙΛΙΚΟΝ	100.144 €	100.111 €	100.104 €	100.103 €	100.102 €	100.102 €	100.102 €	100.102 €	100.102 €	100.102 €	100.102 €	100.102 €
ΒΕΛΕΣΤΙΝΟΝ	77.887 €	78.347 €	78.669 €	78.895 €	79.054 €	79.166 €	79.244 €	79.299 €	79.337 €	79.364 €	79.383 €	79.397 €
ΒΟΛΟΣ	1.436.738 €	1.435.814 €	1.435.248 €	1.434.903 €	1.434.691 €	1.434.562 €	1.434.482 €	1.434.434 €	1.434.404 €	1.434.386 €	1.434.375 €	1.434.368 €
ΒΡΥΞΕΣ	64.097 €	69.373 €	64.489 €	69.010 €	64.824 €	68.700 €	65.112 €	68.433 €	65.358 €	68.205 €	65.569 €	68.010 €
ΓΑΒΡΟΛΙΜΝΗ	1.740 €	2.322 €	2.539 €	2.621 €	2.652 €	2.663 €	2.667 €	2.669 €	2.670 €	2.670 €	2.670 €	2.670 €
ΓΛΥΚΗ	23.524 €	23.620 €	23.660 €	23.676 €	23.683 €	23.686 €	23.687 €	23.688 €	23.688 €	23.688 €	23.688 €	23.688 €
ΓΡΑΒΙΑ	26.605 €	26.880 €	27.045 €	27.144 €	27.202 €	27.237 €	27.258 €	27.271 €	27.278 €	27.283 €	27.286 €	27.287 €
ΔΑΡΑΤΣΟΣ	55.480 €	56.249 €	56.555 €	56.677 €	56.725 €	56.744 €	56.752 €	56.755 €	56.756 €	56.757 €	56.757 €	56.757 €
ΔΑΦΝΗ	21.306 €	20.398 €	20.030 €	19.882 €	19.821 €	19.797 €	19.787 €	19.783 €	19.782 €	19.781 €	19.781 €	19.781 €
ΔΙΑΒΑΤΟΣ	38.198 €	37.709 €	37.324 €	37.022 €	36.784 €	36.597 €	36.449 €	36.333 €	36.242 €	36.171 €	36.114 €	36.070 €
ΔΙΜΗΝΗ	28.011 €	27.028 €	26.732 €	26.642 €	26.615 €	26.607 €	26.605 €	26.604 €	26.604 €	26.604 €	26.604 €	26.604 €
ΔΙΟΜΗΣΕΙΑ	55.186 €	54.159 €	53.876 €	53.798 €	53.777 €	53.771 €	53.769 €	53.769 €	53.769 €	53.769 €	53.769 €	53.769 €
ΔΡΟΣΙΑ	75.196 €	70.255 €	74.097 €	71.109 €	73.432 €	71.626 €	73.031 €	71.938 €	72.788 €	72.127 €	72.641 €	72.241 €

Εικόνα 51 "Αποτελέσματα Intervention Analysis ανά μήνα & ανά Πόλη "

Παρακάτω η ανάλυση σε πλαίσιο καταστήματος:



Εικόνα 52 "Αποτελέσματα μεθόδου Intervention Analysis σε επίπεδο καταστήματος "

```

# Container for storing forecast results
all_forecasts_intervention = {}

# Iterate through each unique code in the dataset
for code in train_data['Code'].unique():
    sales_data = train_data[train_data['Code'] == code]['Sales']
    intervention_data = train_data[train_data['Code'] == code]['intervention']

    # Ensure there's enough data to proceed with forecasting
    if not sales_data.empty:
        # Fit the SARIMAX model with the intervention as an exogenous variable
        model = SARIMAX(sales_data, exog=intervention_data, order=(1,1,1), seasonal_order=(0, 0, 0, 12))
        model_fit = model.fit()

        # Forecast the next 12 months
        exog_forecast = intervention_data.iloc[-1:].values.repeat(12).reshape(-1, 1) # Example approach
        forecast = model_fit.get_forecast(steps=12, exog=exog_forecast)

        # Store the forecast results
        all_forecasts_intervention[code] = forecast.predicted_mean

# Convert the forecast results to a DataFrame
forecast_df_intervention = pd.DataFrame(all_forecasts_intervention)

```

Εικόνα 53 "Κώδικας Python - Intervention Analysis σε επίπεδο καταστήματος "

Παρακάτω εμφανίζεται ένα δείγμα των αποτελεσμάτων που παράχθηκαν από την πρόβλεψη των πωλήσεων για κάθε πόλη με τη μέθοδο Intervention Analysis:

	1/1/2022	1/2/2022	1/3/2022	1/4/2022	1/5/2022	1/6/2022	1/7/2022	1/8/2022	1/9/2022	1/10/2022	1/11/2022	1/12/2022
100001	52.152 €	52.843 €	52.613 €	52.689 €	52.664 €	52.672 €	52.669 €	52.670 €	52.670 €	52.670 €	52.670 €	52.670 €
100002	86.692 €	90.416 €	87.143 €	90.019 €	87.492 €	89.713 €	87.761 €	89.476 €	87.969 €	89.293 €	88.130 €	89.152 €
100003	65.053 €	59.653 €	56.433 €	54.513 €	53.368 €	52.685 €	52.278 €	52.036 €	51.891 €	51.805 €	51.753 €	51.722 €
100004	21.566 €	20.488 €	19.814 €	19.393 €	19.131 €	18.967 €	18.864 €	18.800 €	18.760 €	18.736 €	18.720 €	18.710 €
100005	32.702 €	33.315 €	33.554 €	33.646 €	33.683 €	33.697 €	33.702 €	33.704 €	33.705 €	33.705 €	33.705 €	33.705 €
100006	29.670 €	28.583 €	29.133 €	28.855 €	28.996 €	28.925 €	28.961 €	28.942 €	28.952 €	28.947 €	28.949 €	28.948 €
100008	35.561 €	34.224 €	34.054 €	34.033 €	34.030 €	34.030 €	34.030 €	34.030 €	34.030 €	34.030 €	34.030 €	34.030 €
100009	59.443 €	64.588 €	62.078 €	63.302 €	62.705 €	62.996 €	62.854 €	62.924 €	62.890 €	62.906 €	62.898 €	62.902 €
100014	68.668 €	63.204 €	59.535 €	57.071 €	55.417 €	54.307 €	53.561 €	53.060 €	52.724 €	52.498 €	52.347 €	52.245 €
100015	30.554 €	29.054 €	28.138 €	27.577 €	27.235 €	27.025 €	26.897 €	26.819 €	26.771 €	26.742 €	26.724 €	26.713 €
100016	34.308 €	33.644 €	33.482 €	33.443 €	33.433 €	33.431 €	33.430 €	33.430 €	33.430 €	33.430 €	33.430 €	33.430 €
100022	73.207 €	70.762 €	72.514 €	71.258 €	72.158 €	71.513 €	71.975 €	71.644 €	71.882 €	71.711 €	71.833 €	71.746 €
100023	43.034 €	42.058 €	42.770 €	42.250 €	42.630 €	42.353 €	42.555 €	42.407 €	42.515 €	42.436 €	42.494 €	42.452 €
100031	66.031 €	66.168 €	66.124 €	66.138 €	66.133 €	66.135 €	66.134 €	66.134 €	66.134 €	66.134 €	66.134 €	66.134 €
100032	53.618 €	48.393 €	44.254 €	40.974 €	38.375 €	36.317 €	34.685 €	33.393 €	32.369 €	31.557 €	30.915 €	30.405 €
100037	25.516 €	24.679 €	24.145 €	23.805 €	23.588 €	23.449 €	23.361 €	23.305 €	23.269 €	23.246 €	23.231 €	23.222 €
100039	53.215 €	53.199 €	53.198 €	53.198 €	53.198 €	53.198 €	53.198 €	53.198 €	53.198 €	53.198 €	53.198 €	53.198 €
100046	36.895 €	36.509 €	36.330 €	36.247 €	36.208 €	36.190 €	36.181 €	36.178 €	36.176 €	36.175 €	36.174 €	36.174 €
100048	28.472 €	27.949 €	28.243 €	28.078 €	28.171 €	28.119 €	28.148 €	28.131 €	28.141 €	28.135 €	28.138 €	28.137 €
100049	29.142 €	24.125 €	20.673 €	18.299 €	16.665 €	15.542 €	14.769 €	14.237 €	13.871 €	13.620 €	13.447 €	13.327 €
100060	76.731 €	69.866 €	67.186 €	66.139 €	65.731 €	65.571 €	65.509 €	65.485 €	65.475 €	65.471 €	65.470 €	65.469 €
100062	38.302 €	38.309 €	38.305 €	38.307 €	38.306 €	38.307 €	38.306 €	38.306 €	38.306 €	38.306 €	38.306 €	38.306 €
100063	20.220 €	21.724 €	20.423 €	21.548 €	20.575 €	21.416 €	20.689 €	21.318 €	20.774 €	21.244 €	20.838 €	21.189 €
100068	14.843 €	14.833 €	14.832 €	14.832 €	14.832 €	14.832 €	14.832 €	14.832 €	14.832 €	14.832 €	14.832 €	14.832 €
100071	101.225 €	106.958 €	101.943 €	106.330 €	102.493 €	105.850 €	102.913 €	105.482 €	103.235 €	105.201 €	103.481 €	104.985 €
100089	14.857 €	13.449 €	12.329 €	11.439 €	10.732 €	10.169 €	9.722 €	9.366 €	9.083 €	8.858 €	8.680 €	8.537 €
100091	32.715 €	33.283 €	32.924 €	33.151 €	33.008 €	33.098 €	33.041 €	33.077 €	33.054 €	33.069 €	33.060 €	33.065 €

Εικόνα 54 "Αποτελέσματα Intervention Analysis σε επίπεδο καταστήματος "

Στο πλαίσιο των προβλέψεων που εκτελέστηκαν προηγουμένως - χρησιμοποιώντας Εκθετική εξομάλυνση, ARIMA και Intervention Analysis στο σύνολο των δεδομένων μας - ο προσδιορισμός της βέλτιστης μεθόδου εξαρτάται από διάφορους παράγοντες, συμπεριλαμβανομένων των στόχων της πρόβλεψης, της φύσης των δεδομένων και τις συγκεκριμένες απαιτήσεις της ανάλυσής.

Ο πρωταρχικός στόχος της παρούσας ανάλυσης είναι η πρόβλεψη των πωλήσεων για τις διαφορετικές ιεραρχίες που παρουσιάζονται στα δεδομένα μας, λαμβάνοντας υπόψη τυχόν παρεμβάσεις που θα μπορούσαν να έχουν συμβεί. Το σύνολο δεδομένων εμφάνισε μεταβλητότητα στα στοιχεία των πωλήσεων με την πάροδο του χρόνου, με μια στήλη παρέμβασης που υποδεικνύει συγκεκριμένες περιόδους ενδιαφέροντος. Αυτό το σενάριο απαιτούσε εγγενώς μια διαφοροποιημένη προσέγγιση που θα μπορούσε να λαμβάνει υπόψη τόσο τις ιστορικές τάσεις των πωλήσεων όσο και τον πιθανό αντίκτυπο των εξωτερικών παρεμβάσεων.

Το μοντέλο Exponential Smoothing ήταν έμπειρο στην καταγραφή των τάσεων και της εποχικότητας στα δεδομένα, παρέχοντας ένα απλό και υπολογιστικά αποδοτικό μοντέλο

πρόβλεψης. Ωστόσο, δεν είχε την ικανότητα να ενσωματώσει άμεσα το αποτέλεσμα των παρεμβάσεων.

Το μοντέλο ARIMA πρόσφερε ένα ευέλικτο πλαίσιο μοντελοποίησης ικανό να συλλάβει ένα ευρύ φάσμα συμπεριφορών χρονοσειρών μέσω της ενσωμάτωσης στοιχείων αυτοπαλίνδρομου και κινούμενου μέσου όρου. Παρά την ευελιξία του στο χειρισμό πολύπλοκων μοτίβων, τα τυπικά μοντέλα ARIMA δεν λαμβάνουν εγγενώς εξωτερικές παρεμβάσεις, απαιτώντας πρόσθετες τροποποιήσεις ή επεκτάσεις για να γίνει αυτό.

Το μοντέλο Intervention Analysis, ειδικά όταν ενσωματώθηκε στο πλαίσιο ARIMA (ως SARIMAX), παρείχε μια ισχυρή μέθοδο για την κατανόηση του αντίκτυπου των παρεμβάσεων στα δεδομένα πωλήσεων. Αυτή η προσέγγιση επέτρεψε τη ρητή μοντελοποίηση των παρεμβάσεων ως εξωγενών μεταβλητών, προσφέροντας πληροφορίες για το πώς αυτά τα γεγονότα θα μπορούσαν να έχουν επηρεάσει τις τάσεις των πωλήσεων στις διαφορετικές κατηγορίες μεγεθών.

Δεδομένων των χαρακτηριστικών του συνόλου δεδομένων και των στόχων της ανάλυσης, το μοντέλο Intervention analysis σε συνδυασμό με την ARIMA (SARIMAX) αναδεικνύεται ως η πιο ολοκληρωμένη και ενημερωτική μέθοδος μεταξύ αυτών που δοκιμάστηκαν. Αυτή η προσέγγιση όχι μόνο καταγράφει τις υποκείμενες τάσεις πωλήσεων και τα εποχιακά μοτίβα, αλλά ποσοτικοποιεί και τον αντίκτυπο των εξωτερικών παρεμβάσεων στα μεγέθη των πωλήσεων. Παρέχει μια ολιστική άποψη των δεδομένων, λαμβάνοντας υπόψη τόσο την εσωτερική δυναμική όσο και τις εξωτερικές επιρροές, κάτι που είναι ζωτικής σημασίας για τη λήψη τεκμηριωμένων επιχειρηματικών αποφάσεων. Σε σενάρια όπου η κατανόηση της επίδρασης συγκεκριμένων παρεμβάσεων ή εξωτερικών γεγονότων στην προβλεπόμενη μεταβλητή είναι πρωταρχικής σημασίας, η ικανότητα ενσωμάτωσης αυτών των παραγόντων απευθείας στο μοντέλο πρόβλεψης είναι ανεκτίμητη. Επομένως, για τη συγκεκριμένη ανάλυση -όπου ο στόχος εκτείνεται πέρα από την απλή πρόβλεψη για να συμπεριλάβει την αξιολόγηση των επιπτώσεων της παρέμβασης- το μοντέλο intervention analysis αντιπροσωπεύει την καλύτερη επιλογή μεταξύ των μεθόδων που αξιολογήθηκαν.

Αυτό το συμπέρασμα βασίζεται στα μεθοδολογικά πλεονεκτήματα του μοντέλου στον χειρισμό τόσο των εγγενών ιδιοτήτων της χρονοσειράς όσο και των εξωγενών παραγόντων που την επηρεάζουν. Ωστόσο, είναι σημαντικό να σημειωθεί ότι η "καλύτερη" μέθοδος για κάθε δεδομένη κατάσταση θα εξαρτάται πάντα από τους συγκεκριμένους στόχους της ανάλυσης, τη φύση των δεδομένων και τις πρακτικές εκτιμήσεις της διαδικασίας μοντελοποίησης.

7. Συμπεράσματα

Στη παρούσα διπλωματική εργασία, η οποία εμβαθύνει στις περιπλοκές της πρόβλεψης πωλήσεων γίνεται αρχικά αναφορά σε ένα ευρύ φάσμα ορολογιών και μεθοδολογιών, όπως τα Big Data, τα Analytics, το Demand Driven Forecasting, το Consensus Forecasting, το Sales & Operational Planning, το Budgeting, το Promotional Modelling, το Price Optimization, το Quantitative & Qualitative Forecasting, καθώς και το Data Visualization.

Στη συνέχεια, γίνεται η ανάλυση και η οπτικοποίηση των δεδομένων με την χρήση κατάλληλων τεχνικών και λογισμικών. Πιο συγκεκριμένα, υλοποιήθηκε οπτικοποίηση των δεδομένων με τη μορφή χαρτών – και τη χρήση του λογισμικού ArcGIS - μέσα από τους οποίους εμφανίζεται η θέση των υπό εξέταση καταστημάτων αλλά και η συγκέντρωση της εμπορικότητας η οποία επηρεάζει τις αναμενόμενες πωλήσεις του κάθε καταστήματος. Πέρα από αυτό, οπτικοποιήσεις πραγματοποιήθηκαν και για την καλύτερη κατανόηση των δεδομένων, όπως η συγκέντρωση μηχανών vlts & ssbts, το μέγεθος καταστήματος καθώς και για το πως μεταβάλλονται οι πωλήσεις με βάση την εποχικότητα με τη χρήση του λογισμικού Power BI. Για την πιο ακριβή εξαγωγή συμπερασμάτων, εξετάστηκαν και οι σχέσεις καθώς και τα μοτίβα που παρουσιάζονται στα υπό εξέταση δεδομένα.

Επιπλέον, χρησιμοποιήθηκαν διαφορετικά μοντέλα πρόβλεψης όπως η μέθοδος της εκθετικής εξομάλυνσης, η μέθοδος ARIMA καθώς και το μοντέλο Intervention Analysis. Το ARIMA μοντέλο είναι κατάλληλο όταν υπάρχουν πολύπλοκες χρονοσειρές με ποικιλία, εποχικότητα και τάση. Αυτό το μοντέλο είναι ευέλικτο και μπορεί να προσαρμοστεί σε πιο περίπλοκες δομές δεδομένων. Ωστόσο, αυξάνεται η πολυπλοκότητα της ερμηνείας των αποτελεσμάτων του. Αντίθετα, η

μέθοδος της εκθετικής εξομάλυνσης είναι πιο απλή και εύκολη στην ερμηνεία. Είναι κατάλληλη για περιπτώσεις όπου οι χρονοσειρές είναι πιο σταθερές και δεν υπάρχει έντονη εποχικότητα ή τάση. Αν και είναι λιγότερο πολύπλοκο, μπορεί να παρουσιάζει χειρότερη ακρίβεια σε περίπλοκες χρονοσειρές. Η μέθοδος Intervention Analysis, αποτελεί μια ισχυρή μέθοδο για την κατανόηση του αντίκτυπου των παρεμβάσεων στα δεδομένα πωλήσεων.

Δυστυχώς, ο βασικός παράγοντας της ποιότητας των δεδομένων έχει αναδειχθεί ως σημαντικός περιορισμός, εμποδίζοντας την ικανότητα εξαγωγής ισχυρών και αξιόπιστων συμπερασμάτων από τα προβλεπόμενα μοντέλα. Γίνεται προφανές ότι η αξιοπιστία των προβλέψεων συνδέεται περίπλοκα με την επάρκεια και τη συνέπεια των υποκείμενων δεδομένων. Ασυνήθιστες διακυμάνσεις ή ελλείψεις στην ποιότητα των δεδομένων μπορούν να θέσουν σε κίνδυνο σημαντικά την ακρίβεια των προβλέψεων, καθιστώντας ακόμη και τα πιο προηγμένα μοντέλα λιγότερο αποτελεσματικά.

Το συμπέρασμα που συνάγεται από αυτή τη μελέτη υπογραμμίζει ότι η επιτυχημένη πρόβλεψη πωλήσεων υπερβαίνει την απλή εφαρμογή σύγχρονων μεθόδων πρόβλεψης. Απαιτεί μια συνεχή δέσμευση για συνεχή εκπαίδευση, τακτικές ενημερώσεις των εισροών δεδομένων και μια προσαρμοστική προσέγγιση που ανταποκρίνεται στις περιβαλλοντικές αλλαγές και στις δυναμικές εξελίξεις της αγοράς. Σε ένα τοπίο όπου απρόβλεπτα γεγονότα, όπως η τρέχουσα παγκόσμια κατάσταση ή άλλες διαταραχές της αγοράς, μπορούν να ασκήσουν βαθιές επιπτώσεις, η ανθεκτικότητα και η προσαρμοστικότητα των μοντέλων πρόβλεψης αποκτούν πρωταρχική σημασία.

Κοιτάζοντας το μέλλον, η πορεία προς τη βελτίωση και την αύξηση της ακρίβειας των προβλέψεων περιλαμβάνει μια διπλή δέσμευση τόσο για την ποσότητα όσο και για την ποιότητα των δεδομένων.

Μελλοντικά, η βελτίωση των προβλέψεων μπορεί να επιτευχθεί με τη συλλογή περισσότερων και καλύτερων δεδομένων, καθώς και με την εξέταση πιο προηγμένων μεθόδων πρόβλεψης όπως νευρωνικά δίκτυα ή μέθοδοι βαθιάς μάθησης. Επίσης, η συνεχής παρακολούθηση των εξελίξεων στον χώρο της πρόβλεψης και η προσαρμογή των μοντέλων ανάλογα με τις απαιτήσεις της αγοράς είναι ουσιώδεις για την επιτυχία των προβλέψεων.

Ακόμη, πρέπει να λάβουμε υπόψη μας την εξέλιξη της τεχνολογίας και τη διαθεσιμότητα νέων εργαλείων και λογισμικών που μπορούν να βελτιώσουν την ακρίβεια των προβλέψεών μας. Η χρήση προηγμένων αλγορίθμων μηχανικής μάθησης, όπως οι νευρωνικοί δίκτυα, μπορεί να αντιμετωπίσει ακόμη πιο πολύπλοκες συσχετίσεις στα δεδομένα και να παρέχει προβλέψεις υψηλής ακρίβειας.

Επίσης, η συνεργασία με ειδικούς στον τομέα των προβλέψεων και της ανάλυσης δεδομένων μπορεί να φανεί αποφασιστική. Η συνδυασμένη γνώση των εμπειρογνομόνων με τη δύναμη των αλγορίθμων μπορεί να οδηγήσει σε βαθύτερη κατανόηση των προτύπων στα δεδομένα μας και να βοηθήσει στην ανάπτυξη πιο ακριβών μοντέλων πρόβλεψης.

Τέλος, πρέπει να είμαστε ευέλικτοι στην προσέγγισή μας. Οι συνθήκες στην αγορά και οι παράγοντες που επηρεάζουν τις πωλήσεις μπορεί να αλλάζουν συχνά. Επομένως, πρέπει να είμαστε έτοιμοι να προσαρμόζουμε τα μοντέλα και τις προσεγγίσεις μας ανάλογα με τις νέες εξελίξεις και τις απαιτήσεις της αγοράς.

Συνοψίζοντας, η σωστή πρόβλεψη των πωλήσεων απαιτεί όχι μόνο τη χρήση σύγχρονων μεθόδων και τεχνικών αλλά και τη συνεχή εκπαίδευση, την ανανέωση των δεδομένων και την προσαρμογή στο περιβάλλον και τις εξελίξεις της αγοράς. Με αυτήν την προσέγγιση, μπορούμε να είμαστε πιο βέβαιοι ότι οι προβλέψεις μας θα είναι ακριβείς και χρήσιμες για τη λήψη αποφάσεων στο μέλλον.

8. Βιβλιογραφία

SAS Institute Inc. (<https://www.sas.com/>)

Chiang, W., & Zheng, X. (2017). Big Data Analytics in Supply Chain Management: A State-of-the-Art Literature Review. *Expert Systems with Applications*, 89, 112-122.

Makridakis, S., Spiliotis, E., & Assimakopoulos, V. (2018). Statistical and Machine Learning forecasting methods: Concerns and ways forward. *PloS one*, 13(3), e0194889.

Tang, O., & Musa, S. N. (2011). Identifying risk issues and research advancements in supply chain risk management. *International Journal of Production Economics*, 133(1), 25-34.

Wilson, J. R. (2018). Sales & Operations Planning – Best Practices: Lessons Learned. *Business Horizons*, 61(5), 673-682.

Myers, M. B., & Cheung, M. S. (2008). Sharing global supply chain knowledge. *MIT Sloan Management Review*, 49(4), 67-73.

Chase, C. W., Jacobs, F. R., & Aquilano, N. J. (2006). *Operations Management for Competitive Advantage*. McGraw-Hill.

Makridakis, S., & Hibon, M. (2000). The M3-Competition: Results, conclusions and implications. *International Journal of Forecasting*, 16(4), 451-476.

Fildes, R., & Goodwin, P. (2007). Good and bad judgment in forecasting: Lessons from four companies. *International Journal of Forecasting*, 23(2), 321-331.

Wang, G., Gunasekaran, A., Ngai, E. W., & Papadopoulos, T. (2016). Big data analytics in logistics and supply chain management: Certain investigations for research and applications. *International Journal of Production Economics*, 176, 98-110.

Mentzer, J. T., Moon, M. A., & Mentzer, J. T. (2004). *Sales forecasting management: A demand management approach*. Sage Publications.

Ramanathan, R., & Bentley, Y. (2001). An activity-based cost analysis of demand-driven supply chain strategy. *International Journal of Physical Distribution & Logistics Management*, 31(3), 199-215.

Sanders, N. R. (2018). Sales & operations planning: A research review. *International Journal of Production Economics*, 200, 1-13.

Lei, D., & Srinivasan, R. (2012). Transforming strategic objectives into operational reality: A case study in a sales and operations planning process. *Journal of Operations Management*, 30(1-2), 143-158.

Wildeman, R. E., Hofman, E., & Meade, L. (2016). Best practices in demand-driven manufacturing. *Journal of Manufacturing Technology Management*, 27(7), 989-1016.

Hill, A. V. (2014). Supply chain forecasting: Theory, practice, their gap, and the future. *Operations Research*, 62(3), 536-549.

