ARTIFICIAL INTELLIGENCE BASED LOGISTICS OPTIMIZATION MODEL TO INCREASE BRANCH BASED SALES AND STOCK HARMONY

A THESIS SUBMITTED TO
THE FACULTY OF ARCHITECTURE AND ENGINEERING
OF
EPOKA UNIVERSITY

BY

KLAUDIA TAMBURI

IN PARTIAL FULFILLMENT OF THE REQUIREMENTS
FOR BACHELOR DEGREE
IN SOFTWARE ENGINEERING

JUNE, 2025

I hereby declare that all information in this document has been obtained and presented in accordance with academic rules and ethical conduct. I also declare that, as required by these rules and conduct, I have fully cited and referenced all material and results that are not original to this work.

Name, Last name: ___________________________

Signature: ____________________

ABSTRACT

ARTIFICIAL INTELLIGENCE BASED LOGISTICS OPTIMIZATION MODEL TO INCREASE BRANCH BASED SALES AND STOCK HARMONY

Tamburi, Klaudia

Ba., Department of Computer Engineering

Supervisor: Prof. Dr. Bekir Karlik

Traditional supply chain management faces significant challenges including inaccurate demand forecasting, inefficient stock distribution, and high transportation costs, leading to costly outcomes such as overstocking, stockouts, and increased operational expenses. This study presents an integrated artificial intelligence-driven approach for optimizing logistics operations to maximize branch-based sales and achieve stock harmony across multi-location retail networks.

The methodology employs a comprehensive four-stage framework combining advanced machine learning techniques with mathematical optimization. The main task of this study was using a Long Short-Term Memory (LSTM) neural network to forecast regional demand for 33 product families using historical sales data from Corporación Favorita, Ecuador. The second important step was to focus on the optimization layer which is divided in two parts: Firstly, Mixed Integer Linear Programming (MILP) for optimizing product distribution considering capacity constraints and differentiated

delivery frequencies based on product perishability and secondly, Google OR-Tools for solving the Vehicle Routing Problem to minimize transportation costs while ensuring efficient delivery routes. And finally, comprehensive sales performance analysis was conducted to evaluate key performance indicators (KPIs) including sales volatility, growth patterns, and store-level efficiency metrics to validate forecasting accuracy and identify operational improvement opportunities.

This integrated approach demonstrates the potential of AI-powered supply chain optimization to address complex logistics challenges while providing actionable insights for strategic decision-making in retail operations.

Keywords: Artificial Intelligence, LSTM, MILP, Branch Based Demand Forecasting and Product Distribution.

ABSTRAKT

ARTIFICIAL INTELLIGENCE BASED LOGISTICS OPTIMIZATION MODEL TO INCREASE BRANCH BASED SALES AND STOCK HARMONY

Tamburi, Klaudia

Ba., Department of Computer Engineering

Supervisor: Prof. Dr. Bekir Karlik

Menaxhimi tradicional i zinxhirit të furnizimit përballet me sfida të mëdha, përfshirë parashikimin jo të saktë të kërkesës, shpërndarjen joefikase të stokut dhe kostot e larta të transportit, të cilat çojnë në pasoja të kushtueshme si furnizimi i tepërt, mungesa e produkteve dhe rritja e shpenzimeve të biznesit. Ky studim paraqet një qasje të integruar të udhëhequr nga inteligjenca artificiale për të optimizuar operacionet logjistike, me qëllim maksimizimin e shitjeve në nivel dege dhe arritjen e harmonisë së stokut në rrjete shitjesh me shumë lokacione.

Metodologjia ndjek një kornizë gjithëpërfshirëse me katër faza, duke kombinuar teknika të avancuara të machine learning me optimizimin matematikor. Detyra kryesore e këtij studimi ishte përdorimi i një LSTM neural network (Long Short-Term Memory) për të parashikuar kërkesën rajonale për 33 familje produktesh, duke përdorur të dhëna historike shitjesh nga Corporación Favorita, Ekuador. Hapi i dytë i rëndësishëm ishte në nivelin e optimizimit, e cila ndahet në dy pjesë: së pari, përdorimi i MILP për të

optimizuar shpërndarjen e produkteve duke marrë parasysh kufizimet e kapacitetit dhe frekuencat e ndryshme të shpërndarjes në bazë të prishshmërisë së produkteve; dhe së dyti, përdorimi i Google OR-Tools për zgjidhjen e VRP (Vehicle Routing Problem) për të minimizuar kostot e transportit duke siguruar rrugë efikase të shpërndarjes. Në fund, u krye një analizë gjithëpërfshirëse e performancës së shitjeve për të vlerësuar treguesit kryesorë të performancës (KPI), përfshirë paqëndrueshmërinë e shitjeve, modelet e rritjes dhe treguesit e efikasitetit në nivel dyqani, me qëllim verifikimin e saktësisë së parashikimit dhe identifikimin e mundësive për përmirësime operacionale.

Kjo qasje e integruar tregon potencialin e optimizimit të zinxhirit të furnizimit të fuqizuar nga AI për të adresuar sfida komplekse logjistike duke ofruar njohuri të zbatueshme për vendimmarrje strategjike në operacionet e shitjes me pakicë.

Keywords: Artificial Intelligence, LSTM, MILP, Branch Based Demand Forecasting and Product Distribution.

Dedicated to myself — to the younger girl within me who never stopped believing in our potential, always reaching for the stars. Thank you for dreaming big. I promise to keep making you proud.

ACKNOWLEDGEMENTS

I would like to express my deepest gratitude to Prof. Dr. Bekir Karlik, my thesis supervisor, whose invaluable guidance, expertise, and unwavering support made this research possible. His profound knowledge in artificial intelligence, machine learning and deep learning provided the foundation for this work, and his constructive feedback throughout the research process helped shape this thesis into its final form. I am truly honored to have had the opportunity to work under his mentorship.

I extend my sincere appreciation to the faculty members and staff who contributed to my academic journey and provided the resources necessary to complete this research on artificial intelligence-based logistics optimization. Their commitment to academic excellence created an environment conducive to meaningful research and learning.

Finally, I am grateful to my family and friends for their patience, encouragement, and understanding throughout this research endeavor. Their support provided the motivation and strength needed to overcome the challenges inherent in this complex interdisciplinary study.

This thesis represents not only my individual effort but also the collective contribution of all those who supported and guided me along this academic journey.

TABLE OF CONTENTS

LIST OF TABLES

TABLES

LIST OF FIGURES

Figures

LIST OF ABBREVIATIONS

AI	Artificial Intelligence
LSTM	Long-Short Term Memory
ARIMA	AutoRegressive Integrated Moving Average
SARIMA	Seasonal AutoRegressive Integrated Moving Average
MILP	Mixed Integer Linear Programming
EDA	Exploratory Data Analysis
AIC	Akaike Information Criterion
VRP	Vehicle Routing Problem
KPI	Key Performance Indicator

CHAPTER 1

INTRODUCTION

In today's competitive retail and supply chain landscape, optimizing logistics is crucial for maintaining business by maximizing sales, reducing costs, and having stock efficiency. Traditional logistics management often faces challenges such as inaccurate demand forecasting, inefficient stock distribution, and high transportation costs, which lead to issues such as overstocking, stockouts, delayed deliveries, and increased expenses [1][2]. Current data reveals that nearly three-quarters (specifically 73%) of supply chain executives continue relying on traditional or obsolete forecasting techniques for demand planning, resulting in expensive consequences from excess inventory or product shortages. [2].

The goal for each of these retailers and supply chains is to ensure that each branch receives the right products at the right time, in the right quantities, thereby improving their sales performance and maintaining stock balance. Early adopters of AI-enabled supply chain management have reduced logistics costs by 15 percent, improved inventory levels by 35 percent, and enhanced service levels by 65 percent [3].

Nowadays, businesses increasingly rely on AI-powered models to handle vast amounts of historical sales data, customer preferences, and operational constraints. Market research indicates that the artificial intelligence sector within supply chain management was valued at approximately $5.05 billion during 2023, with projections suggesting a compound annual growth rate of 38.9% spanning the period from 2024 through 2030 [4]. Alternative research estimates suggest the market could potentially reach $157.6 billion by 2033 [5]. However, ensuring stock harmony across multiple branches remains a complex challenge due to:

1. Fluctuating demand patterns (seasonality, promotions, special events) - AI algorithms are able to analyze huge amounts of data both in the past as well as in real time to analyze complex patterns and trends that could easily miss human analytics [6].

2. Inefficient stock replenishment policies (leading to overstocking or shortages) - Accurately predicting sales during fluctuating demand cycles is a core concern due to the challenge of achieving precise forecasting [7].

3. Corporate earnings experience downward pressure from elevated supply chain costs, where fuel price fluctuations play a central role. Data indicates the global energy price index for fuel reached 186.04 points by mid-2024, effectively doubling from its 2016 starting point [2].

4. Uncertainty in supply chain operations (caused by external factors like weather, supplier delays) - Organizations have grown increasingly cognizant that supply chain interruptions are becoming routine occurrences, driven by multiple converging factors including ongoing logistical bottlenecks, insufficient

container availability, political instabilities, economic volatility, regional conflicts, escalating cybersecurity risks, and more frequent severe weather patterns. [8].

To effectively achieve logistics optimization aimed at increasing branch-based sales and stock harmony, four essential tasks must be addressed: accurate demand forecasting for each branch, optimized stock distribution, cost-efficient transportation planning, and good sales performance analytics.

The implementation of artificial intelligence in logistics optimization is a critical advancement in addressing these persistent challenges, representing a paradigm shift from traditional manual methods to intelligent automatized data-driven decision-making processes [9].

1.1. Demand Forecasting

Demand forecasting is the process of predicting future customer demand for a product or service using historical data, statistical analysis, and machine learning techniques [10]. It is an essential aspect of supply chain management because it helps businesses anticipate and meet consumer demands, especially in retail and the food industry [11]. If there is an incorrect prediction, it will cause over-purchasing, which will lead to overstock and spoilage and on the other hand, if there is insufficient purchase, it will cause a shortage of merchandise available to customers. That is why it is very important for stores to accurately predict the purchase volume of goods [12].

In demand forecasting, there exist three underlying issues which are: 1. Forecasting is always wrong; 2. The longer the horizon, the lower the accuracy; 3. Forecasting precision increases proportionally with the level of demand aggregation [13]. Effective supply chain oversight aims to align production capacity with market requirements. The lack of an effective demand forecast process is very likely to lead to miscoordination across various functions in a firm and consequently lowering its efficiency [13].

Traditional demand forecasting methodologies often rely on simple regression models and basic statistical techniques, yet these approaches frequently provide unstable estimations and struggle with the complexity of modern market dynamics [11][13]. In recent studies was shown that machine learning techniques, such as LSTM, can improve the accuracy of demand forecasting by learning non-linear relationships and time-series-specific information, and despite the limited empirical data available, some mathematical models, like ARIMA and SARIMA have also proven to be efficient in forecasting demand [11].

1.2. Product Distribution and Stock Optimization

Product distribution and stock optimization constitutes a fundamental component of modern supply chain management, encompassing the strategic allocation of inventory across multiple locations including warehouses, retail stores, and distribution centers to minimize operational costs while ensuring efficient customer demand fulfillment [1][7]. This complex process involves balancing stock levels to prevent both overstocking and stockouts, optimizing product placement across diverse geographical locations, and enhancing overall supply chain efficiency through intelligent distribution strategies [7][10].

The absence of proper stock optimization exposes businesses to significant inefficiencies and substantial financial losses that can severely compromise their competitive positioning and profitability [2][7]. Traditional inventory management approaches predominantly depend on fixed operational rules including predetermined reorder points and static safety stock levels, which fundamentally lack the necessary adaptability to respond effectively to real-time demand fluctuations and dynamic supply chain disruptions [1][7][14]. These conventional methodologies frequently result in suboptimal inventory allocation, leading to either excessive holding costs from overstocking or missed sales opportunities from stockouts [10][14].

Contemporary optimization approaches utilize sophisticated mathematical frameworks, including MILP, to address the complex discrete nature of supply chain decisions [14]. Research demonstrates that MILP models effectively handle integer variables representing real-world constraints such as the number of production batches, transportation trips, or facility locations, making the solution space discrete and non-convex while requiring specialized algorithms like branch-and-bound and branch-and-cut to navigate this complexity [14]. The implementation of periodic review policies within supply chain networks has proven particularly effective, where inventory is reviewed at specific intervals and orders are placed when inventory falls below threshold to replenish to target level [14].

The integration of artificial intelligence and machine learning technologies has revolutionized stock distribution optimization by providing sophisticated analytical capabilities that substantially exceed traditional methodologies [1][9][11]. AI-powered systems enhance inventory management through multiple mechanisms: predicting complex demand patterns by analyzing historical sales data, seasonality trends, and external market factors; optimizing stock allocation strategies across warehouses and retail locations; automating critical decision-making processes for optimal timing, location, and quantities of stock replenishment; reducing logistics costs through intelligent route planning and transportation scheduling optimization; and dynamically processing large, complex datasets rather than relying on static forecasting models [1][9][11][14].

Modern AI-driven inventory optimization systems demonstrate the capability to process vast amounts of real-time data to identify patterns and relationships that traditional methods cannot detect, enabling businesses to maintain optimal stock levels while minimizing carrying costs and maximizing service levels [11][14]. Research indicates that implementing MILP-based optimization can achieve significant operational improvements, including 15% reductions in logistics costs, 35% improvements in inventory levels, and 65% enhancements in service levels [3][14]. Furthermore, periodic review policies offer simplicity and ease of implementation, making them practical choices for certain inventory management scenarios while fostering data sharing and converging objectives to improve overall supply chain performance [14].

The strategic implementation of AI-based stock optimization directly contributes to enhanced customer satisfaction through improved product availability, reduced operational costs through more efficient inventory management, and increased overall supply chain resilience through better adaptation to market volatility and unexpected disruptions [3][8][14]. As businesses continue operating in increasingly complex and dynamic market environments, the adoption of intelligent stock distribution and optimization systems becomes essential for maintaining competitive advantage and achieving sustainable operational excellence [2][11][14].

1.3. Warehouse to Branch Transportation

Warehouse to branch transportation represents a critical component of supply chain logistics, encompassing the strategic management of warehouse storage, inventory flow, and transportation logistics to ensure cost-effective and timely product delivery across multiple retail locations [1][2]. This complex process focuses on three fundamental areas: warehouse optimization for efficient storage and retrieval operations to reduce costs and enhance order fulfillment; transportation planning for optimizing delivery routes, reducing fuel consumption, and ensuring on-time deliveries; and supply chain coordination to maintain smooth operations between suppliers, warehouses, and distribution centers [7][14].

The optimization of warehouse to branch transportation systems falls within the broader category of vehicle routing problems, which involve the strategic optimization of freight transportation activities across distribution networks [15]. These problems are typically addressed through the representation of road networks as weighted complete graphs, where each arc represents the shortest route for possible origin-destination connections [15]. Modern transportation planning must consider multiple attributes for each route including travel time, travel cost, fuel consumption, and delivery windows, though traditional approaches often optimize according to a single criterion, typically travel time, which can lead to suboptimal solutions that discard alternative routes offering different compromises between these critical attributes [15].

Vehicle routing problems in warehouse to branch transportation typically address three distinct types of decisions: assignment decisions that allocate key delivery locations to specific vehicles; sequencing decisions that define the optimal visit order for each vehicle across multiple branch locations; and scheduling decisions that determine precise timetables for visiting assigned locations while meeting delivery constraints [15]. The complexity increases significantly when multiple criteria must be considered simultaneously, as the selection of optimal routes between consecutive delivery points requires sophisticated optimization algorithms that can evaluate trade-offs between time, cost, and service quality [15].

Traditional transportation planning approaches often fail to consider alternative routing paths, which can be particularly disadvantageous in dynamic retail environments where delivery priorities and constraints frequently change [15]. In branch-based retail operations, transportation plans must accommodate point-to-point delivery requests while satisfying quality of service constraints and cost optimization objectives [2][15]. While the shortest time path between warehouse and branch locations is typically preferred, operational flexibility may require considering cost-effective alternatives when time constraints are less critical, particularly in scenarios involving multiple delivery windows or varying customer service requirements [15].

The implementation of artificial intelligence and advanced optimization techniques has revolutionized warehouse to branch transportation by enabling real-time route optimization, dynamic load balancing, and predictive maintenance scheduling [1][9]. AI-powered systems can process vast amounts of transportation data including traffic patterns, vehicle capacity utilization, fuel costs, and delivery performance metrics to identify optimal routing solutions that minimize total transportation costs while maintaining high service levels [1][9]. Modern transportation management systems integrate with warehouse management platforms to coordinate inventory allocation, order picking, and vehicle loading processes, ensuring seamless integration between storage operations and delivery execution [7][14].

Contemporary warehouse to branch transportation systems must also address increasing complexity from multiple distribution channels, varying delivery requirements, and sustainability objectives [2][8]. Rising fuel costs, environmental regulations, and customer expectations for faster delivery times require sophisticated optimization approaches that can balance cost efficiency with service quality [2]. The integration of real-time tracking systems, predictive analytics, and machine learning algorithms enables dynamic route adjustments based on traffic conditions, weather patterns, and urgent delivery requests, significantly improving overall transportation efficiency and customer satisfaction [1][9].

The strategic optimization of warehouse to branch transportation directly impacts overall supply chain performance through reduced transportation costs, improved delivery reliability, and enhanced customer service levels [3][14]. Research indicates that companies implementing advanced transportation optimization systems can achieve 10-15% reductions in transportation costs while improving on-time delivery performance by 20-25% [3]. As retail operations continue to expand across multiple channels and geographic regions, the development of sophisticated warehouse to branch transportation systems becomes essential for maintaining competitive advantage and operational excellence [2][8].

1.4. Sales Performance Analytics

Sales performance analytics represents a sophisticated approach to analyzing sales data, customer behavior, and business performance metrics to identify trends, predict future outcomes, and optimize strategic decision-making processes across branch-based retail operations [1][9]. This comprehensive analytical framework involves the systematic tracking and evaluation of key performance indicators including revenue generation, sales volume, customer retention rates, product demand patterns, and return rates to provide actionable insights for improving overall business performance [10][11]. In the context of branch-based sales optimization, performance analytics enables organizations to understand regional variations in customer preferences, seasonal demand fluctuations, and the effectiveness of different sales strategies across multiple locations [11][13].

Traditional sales performance analysis relied heavily on descriptive statistics and historical reporting, providing limited insights into future trends or underlying patterns driving sales performance [11][13]. However, the integration of Machine Learning (ML) and Artificial Intelligence (AI) technologies has revolutionized sales analytics by enabling businesses to move beyond traditional retrospective reporting to predictive forecasting, real-time anomaly detection, and dynamic sales strategy optimization [1][9][11]. These advanced analytical capabilities allow organizations to identify subtle patterns in customer behavior, predict future sales trends with greater accuracy, and implement data-driven strategies that adapt to changing market conditions [9][11].

Time series decomposition represents a fundamental analytical technique in sales performance analytics, enabling the systematic breakdown of sales data into distinct components including trend, seasonality, and irregular fluctuations [11][16]. This decomposition process allows analysts to isolate long-term growth patterns from seasonal variations, identifying underlying business trends that might otherwise be obscured by cyclical fluctuations [16]. For branch-based retail operations, time series analysis can reveal critical insights such as seasonal demand patterns specific to geographic regions, the impact of promotional campaigns on sales velocity, and the effectiveness of inventory allocation strategies across different store locations [11][16].

Real-time sales performance monitoring enables organizations to detect anomalies and unusual patterns as they occur, facilitating immediate corrective actions and strategic adjustments [1][9]. Machine learning algorithms can identify deviations from expected sales patterns, flagging potential issues such as inventory shortages, competitive threats, or operational problems that might impact performance [9][11]. This capability is particularly valuable in multi-branch retail environments where performance variations between locations can indicate opportunities for best practice sharing or identify branches requiring additional support [1][9].

The integration of sales performance analytics with demand forecasting and inventory optimization systems creates a comprehensive analytical ecosystem that supports data-driven decision-making across all aspects of branch operations [7][11][14]. This integration enables organizations to correlate sales performance metrics with inventory levels, demand patterns, and operational efficiency indicators, providing holistic insights that inform strategic planning and operational optimization [11][14]. Research indicates that companies implementing integrated sales analytics platforms achieve 12-18% improvements in sales performance and 20-25% reductions in inventory carrying costs through better alignment of supply and demand [3][11]. As retail landscapes continue to evolve, the development of sophisticated sales performance analytics capabilities becomes essential for maintaining competitive advantage and driving sustainable growth across branch-based operations [2][11].

1.5. Aim and Novelty of the Study

The goal of this study was to develop an integrated artificial intelligence-driven model for optimizing supply chain logistics operations to maximize branch-based sales and achieve stock harmony across multi-location retail networks. This involves addressing critical challenges including inaccurate demand forecasting, inefficient stock distribution, and high transportation costs that lead to overstocking, stockouts, and delayed deliveries. Understanding these complexities is crucial for recognizing the need for intelligent automation in logistics management, as 73% of supply chain leaders still rely on manual processes for demand prediction [2].

To address these challenges, the study demonstrates the potential of utilizing machine learning and optimization algorithms to transform supply chain operations more effectively. The research implements a four-step methodology combining LSTM neural networks for demand forecasting, Mixed Integer Linear Programming for product distribution optimization, Google OR-Tools for transportation planning, and sales performance analytics.

The novelty lies in the holistic integration of multiple AI and optimization techniques within a unified system, moving beyond isolated solutions to create an end-to-end framework. Key points include differentiated delivery frequency strategies based on product perishability, dynamic cost modeling incorporating real-time oil prices, and synthetic capacity feature generation. This research represents a transformative step towards enhancing logistics efficiency, reducing operational costs and improving service levels.

CHAPTER 2

MATERIALS AND METHODS

2.1. Materials

The dataset used for this study is a Grocery Sales Forecasting (provided by Corporación Favorita, Ecuador) obtained from the Kaggle platform [17]. It consists of 7 excel sheets:

1. train.csv: This is the train dataset. Which includes the target unit_sales date, store_nbr, and item_nbr and a unique id to label rows. The target can be integer or float and the negative values indicate returns of that item. There is also a column called onpromotion indicating if that item was on promotion for a specific time in a specific store.
2. test.csv: The dataset to be used to test the model trained and to make the predictions for the regression task. It includes date, store_nbr, item_nbr and onpromotion information.
3. stores.csv: Dataset with information about each store like city, state, type and cluster, where cluster is a grouping of similar stores.
4. items.csv: Includes information for each product such as product family, class and if the product is perishable or not.

5. transactions.csv: The dataset includes count of sales transactions for each date, store combination. It only spans for the train dataset timeframe.

6. oil.csv: It includes daily oil prices. Values are for both the train and test data timeframe. (Ecuador's economic foundation depends mostly on oil production, leaving the country's economic well-being exposed to unpredictable oil price swings).

7. holidays_events.csv: Has information about holidays and events dates. Special attention is paid to the transfer column. Holidays marked as transferred occur on their scheduled calendar dates but have been shifted to different days by governmental decision. Moved holiday dates resemble regular business days rather than traditional festive periods. To discover the real celebration date, search for the related row labeled "Transfer" in the type field.

The algorithms used are in Python language, utilizing Python libraries for data analysis and machine learning/deep learning, executed in Azure ML environment and Google Colab. Models and methodologies are mostly found online with modifications and adjustments to match the dataset and achieve the desired task and target.

2.2. Methods

2.2.1. Big Data Preprocessing

The preprocessing phase was critical given the substantial size of the training dataset, which contained 125,497,040 rows across 6 columns and occupied approximately 5.6 GB of storage. The primary objectives were to conduct EDA, handle missing values, reduce computational complexity, and merge supplementary datasets to create a comprehensive training environment. Additionally, for consistency and to make the analysis more meaningful, all dataset columns were renamed to follow standardized naming conventions.

2.2.1.1. Exploratory Data Analysis

A systematic EDA was performed using a custom function to examine each dataset's characteristics. This analysis revealed key insights across all datasets:

1. Training Data: 125,497,040 records with 21,657,651 missing values in the 'onpromotion' column, representing approximately 17.3% of the data.
2. Test Data: 3,370,464 records with no missing values, spanning from August 16-31, 2017.
3. Items Dataset: 4,100 products categorized into 33 distinct product families
4. Stores Dataset: 54 stores across 5 different store types (A-E) with varying cluster assignments.
5. Oil Prices: 1,218 daily records with 43 missing values requiring interpolation.
6. Holidays/Events: 350 records covering national, regional, and local events.
7. Transactions: 83488 records covering transactions per date per store for the timeframe of January 1st, 2013 to August 15, 2017.

2.2.1.2. Data Reduction

To address the computational challenges posed by the massive train dataset size, I decided to implement a data reduction approach, since there existed works using the reduced version of the dataset [24]. Rather than working with individual product numbers (which lacked descriptive names and only were represented by some number IDs), I decided to aggregate them by their product families. This approach provided several advantages:

1. Semantic Meaningfulness: Product families offer interpretable categories (e.g., "GROCERY I", "BEVERAGES", "CLEANING") rather than arbitrary product numbers.
2. Dimensionality Reduction: Reduced from 4,100 individual products to 33 product families. And the train dataset was reduced from 125,497,040 samples to 3,008,016 samples and the test dataset 28,512 test records.
3. Computational Efficiency: Reducing the dataset makes it much more efficient and manageable to work with but also significantly decreases memory requirements and processing time.

2.2.1.3. Missing Data and Feature Engineering

The preprocessing phase addressed two critical aspects: comprehensive missing data imputation and strategic feature engineering to enhance model performance. Missing data challenges were tackled through systematic approaches across multiple datasets, while feature engineering focused on creating temporal variables and contextual indicators essential for time series forecasting.

1. Promotion Data: Missing values in the 'onpromotion' column were filled with False (converted to 0), assuming non-promoted status when unspecified.

2. Oil Prices: Missing dates were added to fill in the gaps in the time series and then the missing values were forward-filled and back-filled to ensure complete temporal coverage.

3. Holiday Indicators: Holidays with status transferred were removed, since they are celebrated as official holidays in other dates provided in the dataset and non-holiday dates in the time series were assigned a value of 0.

4. Train and Test Data: A comprehensive grid was created to make sure all combinations of date range (January 1, 2013 to August 15, 2017 for training and August 16-31, 2017 for testing) with all 54 stores and 33 family products were present. If there was any combination missing it implies no data/sales done for that product in that store on that date so the sales units were set to zero.

5. Transactions Data: The preprocessing phase revealed a critical gap in the transactions dataset, which contained daily transaction counts only for the training period and lacked data for the test period (August 16-31, 2017). Given that transaction volume serves as a strong predictor of sales performance, simply imputing these missing values with zeros would significantly compromise the accuracy of sales forecasting models. Therefore, a predictive approach was necessary to forecast the missing transaction data, ensuring feature completeness and maintaining the predictive power of this important variable across both training and testing phases.

6. Store Data Capacity: Given that one of the tasks needed for ensuring good stock harmony is optimization of product distribution and transportation planning, store capacity is a very important information to have, which in the provided dataset is missing. Capacity constraints are crucial for the subsequent optimization layer of the forecasting system, where realistic operational limitations must be enforced to ensure practical applicability of demand predictions and inventory allocation strategies. Taking this into account, synthetic feature data were generated using external tools (ChatGPT) and provided information in the store dataset such as store type and cluster as well as incorporating actual sales distribution patterns from the training data, where Store Type A accounts for 38.2% of total sales, Type B for 17.7%, Type C for 10.7%, Type D for 19.0%, and Type E for 14.5% of total sales volume.

   Base capacity estimates were established reflecting operational scale: Type A (hypermarket) 12,000,000 units, Type B (supermarket) 6,000,000 units, Type C (neighborhood) 1,000,000 units, Type D (convenience) 250,000 units, and Type E (express) 150,000 units. These were adjusted using the formula capacity = base_capacity × (1 + cluster / 100) to account for locational and operational variations. The final values were proportionally scaled to align with observed sales distribution patterns, creating realistic constraints ranging from 150,000 to over 14 million units. This synthetic capacity feature serves dual purposes: establishing upper bounds for sales predictions in forecasting models and enabling the optimization layer to implement realistic operational constraints that prevent inventory allocations from exceeding physical store limitations, thereby enhancing the practical applicability of the entire demand planning system.

7. Temporal Features: Year, month, day, and day-of-week variables were extracted from date fields to capture various temporal patterns crucial for retail sales

forecasting. These temporal features enable models to identify and leverage multiple layers of seasonality: year variables capture long-term trends and annual growth patterns, month variables detect seasonal fluctuations such as holiday shopping periods and weather-related demand changes, day variables help identify specific calendar effects like month-end shopping behaviors or promotional timing, and day-of-week variables reveal weekly cyclical patterns such as weekend shopping surges or weekday purchasing habits.

2.2.1.3.1. Predicting Missing Transactions Data using SARIMA

SARIMA was selected for transaction forecasting due to its specialized design for time series data and its capability to effectively model the complex temporal dynamics characteristic of retail environments, including underlying trends, seasonal fluctuations, and cyclical patterns inherent in customer transaction behavior.

The model accounts for:

1. Trend components: Long-term increases or decreases in transaction volume
2. Seasonal patterns: Weekly and monthly cyclical behaviors in store traffic
3. Autoregressive relationships: Dependencies on previous transaction counts

A comprehensive grid search approach was implemented to identify optimal SARIMA parameters for each store individually, recognizing that different stores may exhibit varying temporal patterns. The parameter space included:

1. Non-seasonal parameters: $p \in \{0,1\}$, $d \in \{0,1\}$, $q \in \{0,1\}$
2. Seasonal parameters: $P \in \{0,1\}$, $D \in \{0,1\}$, $Q \in \{0,1\}$
3. Seasonal periods: $m \in \{7, 12\}$ (weekly and monthly seasonality)

The Akaike Information Criterion (AIC) served as the model selection metric, balancing model fit quality with complexity to avoid overfitting.

Prior to model fitting, essential preprocessing steps were applied to ensure data quality and model stability: log transformation using np.log1p() was implemented to stabilize variance and appropriately handle zero transaction counts, daily frequency consistency was established using asfreq('D') to ensure uniform temporal intervals, and chronological ordering was maintained by sorting data by date to preserve temporal sequence integrity.

The forecasting implementation involved developing individual SARIMA models for each of the 54 stores to capture unique store-specific patterns, generating 16-day predictions covering the entire test period (August 16-31, 2017), applying inverse transformation using np.expm1() to convert log-scale predictions back to the original scale, and implementing constraint applications through clipping and floor operations to ensure non-negative integer transaction counts.

The model validation and post-processing phase incorporated several quality assurance measures: boundary constraints were applied to set negative forecasts to zero since transactions cannot be negative, integer conversion was performed by rounding results down to nearest integers for realistic transaction counts, visual validation was conducted through generated plots comparing actual versus predicted transactions for 2017 data to assess model performance, and outlier handling was implemented by capping extreme predictions at 50,000 transactions for visualization clarity while preserving the underlying forecast values for model use.

This approach ensured that transaction count features were available for both training and test periods, maintaining feature consistency across the entire forecasting pipeline while preserving the temporal relationships essential for accurate sales predictions.

2.2.1.4. Dataset Integration

The preprocessing culminated in merging all supplementary datasets with the main sales data, resulting in features such as: product family attributes (perishability indicators), economic indicators (oil prices), calendar effects (holidays and events), temporal features (year, month, day, and week day) and number of transactions per store per date.

The correlation analysis revealed moderate relationships between features, with the strongest correlations observed between temporal variables and store-specific characteristics.

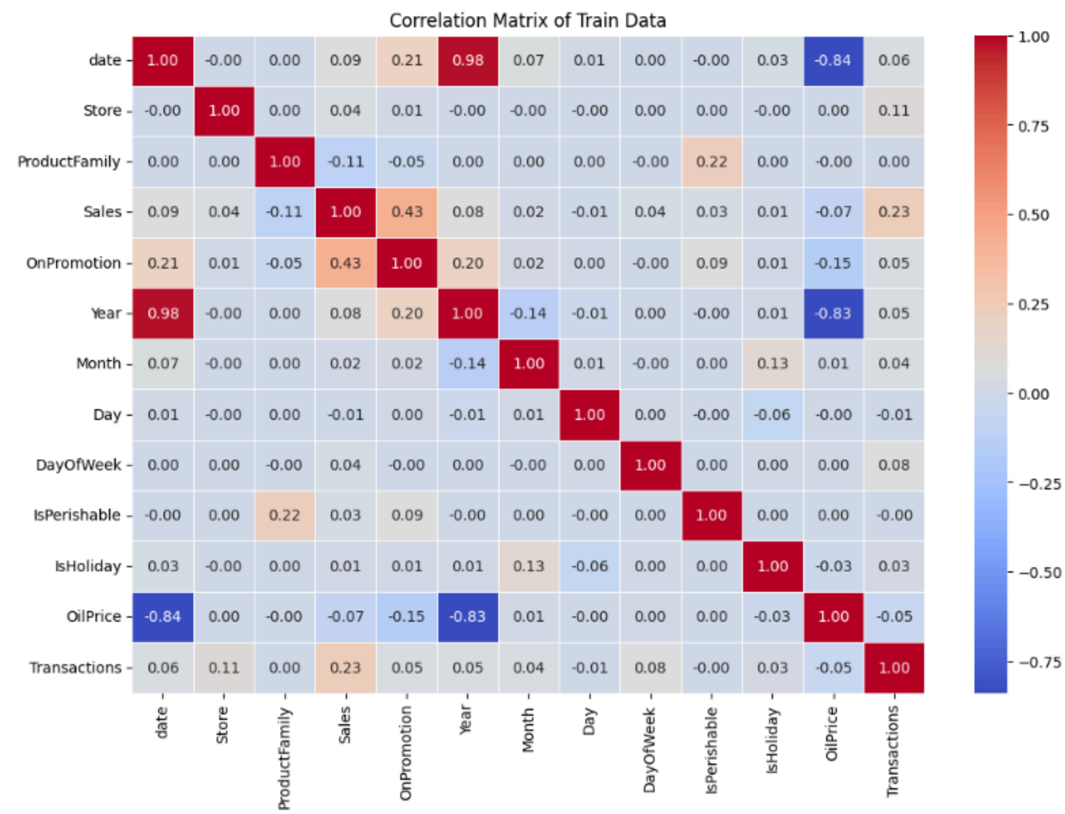

Figure 1. Correlation matrix between each feature in train dataset after preprocessing steps and merging.

2.2.2. Regional Demand Forecasting using LSTM

Long Short-Term Memory (LSTM) networks represent a specialized class of recurrent neural networks designed to address the vanishing gradient problem inherent in traditional RNNs when processing long sequences [18][19]. Unlike standard neural networks that process inputs independently, LSTMs maintain an internal memory state that allows them to selectively remember or forget information across time steps, making them particularly effective for sequential data analysis.

The LSTM architecture incorporates three critical gating mechanisms that control information flow: the forget gate determines what information to discard from the cell state, the input gate decides which values to update in the cell state, and the output gate controls what parts of the cell state to output as the hidden state [18]. This gating structure enables LSTMs to capture both short-term fluctuations and long-term trends in time series data, addressing the challenge of learning dependencies across extended temporal horizons.

In the context of demand forecasting, LSTM networks excel at identifying complex patterns in historical sales data, including seasonal variations, promotional effects, and underlying trend components [20]. The model's ability to maintain long-term memory while adapting to recent changes makes it particularly suitable for retail environments where demand patterns exhibit both persistent characteristics and dynamic responses to market conditions.

2.2.2.1. Time Series to Supervised Learning Transformation

The foundation of this approach lies in converting time series data into a supervised learning format suitable for neural network training. Following the methodology outlined by Brownlee [21], the series_to_supervised() function transforms sequential sales data into input-output pairs by creating lagged features that capture historical patterns. This transformation creates a sliding window approach where past observations (t-window to t-1) serve as input features, while future values (t+1 to t+lag) become prediction targets.

The transformation process incorporates both temporal features that vary across time steps (OnPromotion, OilPrice, Transactions, Sales, Day, Weekday, Month, Year, IsHoliday) and static features that remain constant for each store-product combination (Store, ProductFamily, IsPerishable). This dual approach ensures that the model captures both time-varying dynamics and inherent characteristics of different product categories and locations.

2.2.2.2. Data Preprocessing and Feature Engineering for LSTM

The preprocessing pipeline implements several critical steps to prepare data for LSTM training. Label encoding transforms categorical ProductFamily variables into numerical representations, while StandardScaler normalization ensures that features with different scales contribute equally to model training. The implementation focuses on the last year of training data (2017) to capture the most recent demand patterns, recognizing that consumer behavior and market conditions evolve over time.

Feature scaling addresses the challenge of heterogeneous input ranges, particularly important for neural networks where unscaled features can lead to training instability. The target variable (Sales) undergoes separate scaling to preserve the ability to inverse-transform predictions back to original units for interpretation and evaluation.

2.2.2.3. LSTM Architecture Design

Multiple LSTM architectures were evaluated during the model development process, with the final implementation selected based on superior performance in validation experiments. Initially, a single-layer LSTM architecture was tested with 512 units followed by multiple dense layers (256, 128, 64, 32, 16 units) and dropout regularization. However, this configuration demonstrated suboptimal performance in capturing the complex temporal dependencies required for accurate multi-step forecasting.

The implemented LSTM architecture employs a multi-layer design with 256 and 128 units in successive LSTM layers, incorporating dropout regularization (0.3 and 0.2 rates) to prevent overfitting. This stacked configuration outperformed the initial single-layer approach by providing better gradient flow and enhanced feature learning capabilities.

The optimal model structure includes:

1. Stacked LSTM layers: The first LSTM layer with return_sequences=True passes complete sequences to the second layer, enabling the capture of hierarchical temporal patterns where the first layer identifies basic sequential features and the second layer learns higher-order temporal relationships.
2. Dense projection layers: Multiple fully connected layers (256, 128, 64 units) with ReLU activation provide non-linear transformation capabilities, allowing the model to learn complex mappings between temporal patterns and demand outcomes.
3. Multi-step output: The final dense layer produces 16 outputs corresponding to the forecast horizon, enabling simultaneous prediction of multiple future time steps.

This architecture balances model complexity with computational efficiency, using Adam optimizer with a learning rate of 0.001 and mean squared error loss function for regression objectives.

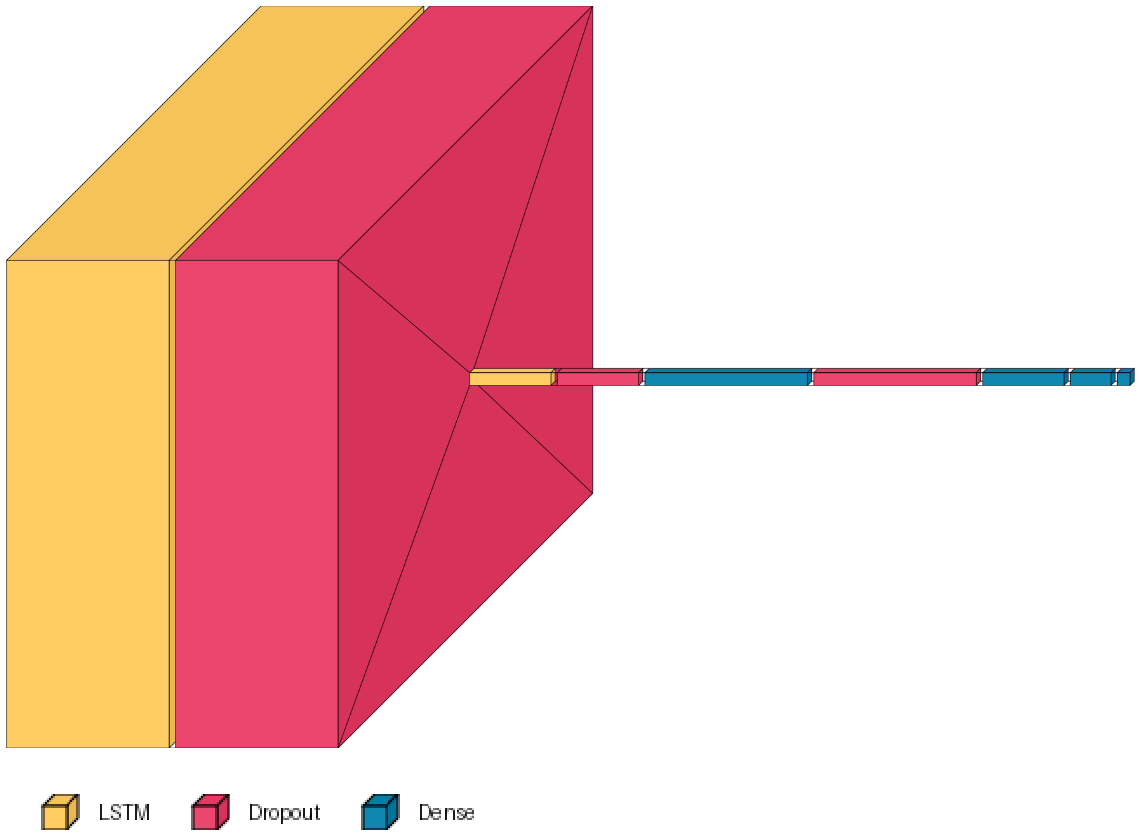

Figure 2. LSTM architecture visualization using Visual Keras tool in Python.

2.2.2.4. Model Training

The training process implements early stopping (with patience for 5 epochs) and learning rate reduction callbacks (with patience for 2 epochs, factor 0.2 and minimum learning rate 0.0000001) to optimize model performance while preventing overfitting. The validation split (20% of training data) enables monitoring of generalization performance during training. The model training employs batch processing (batch size 128) over a maximum of 100 epochs, with automatic termination when validation loss plateaus.

The 3D input reshaping transforms the 2D feature matrix into the required LSTM input format, where each sample represents a store-product combination with its associated temporal and static features.

2.2.2.5. Forecasting

The trained model generates predictions for the 16-day forecast horizon, with outputs subsequently transformed back to original scales through inverse scaling operations. The final predictions are organized by store and product family, providing granular demand forecasts that support tactical inventory and distribution decisions. To validate the forecast of the model, visual representation was used about the forecast value range compared to existing data value ranges per product.

2.2.3. Optimization Layer

The optimization layer represents the final component of the integrated supply chain management system, transforming demand forecasts into actionable distribution strategies. This layer employs a two-stage optimization approach that addresses both strategic product allocation and operational vehicle routing challenges. The methodology combines Mixed Integer Linear Programming (MILP) for optimal product distribution with Google OR-Tools for efficient transportation routing, ensuring cost-effective delivery while respecting capacity constraints.

2.2.3.1. Product Distribution Optimization using MILP

The product distribution optimization component employs Mixed Integer Linear Programming to determine optimal shipment quantities from the central warehouse to individual retail stores. This mathematical optimization approach ensures demand satisfaction while minimizing total distribution costs, incorporating both transportation expenses and operational constraints. The MILP model addresses the product allocation problem by considering multiple factors that influence distribution decisions, with the objective function minimizing the total cost of distribution expressed as the sum of quantity multiplied by distance and cost per unit. The cost structure incorporates both fuel-related expenses based on oil prices and handling fees, operating under several critical constraints that reflect real-world operational limitations.

The optimization model enforces three primary constraint categories that ensure feasible and practical solutions. Demand satisfaction constraints guarantee that each store receives at least its forecasted demand for each product family, preventing stockouts by ensuring adequate inventory levels across the retail network. Vehicle capacity constraints limit the total shipment volume based on available transportation capacity, with the model calculating the number of vehicles required based on aggregate demand and individual vehicle capacity of 30,000 units. Store capacity constraints verify that total shipments to individual stores do not exceed their storage capabilities, with the system automatically removing affected stores from the distribution plan when capacity violations are detected to maintain feasibility.

The system implements a differentiated delivery frequency approach based on product perishability characteristics. Perishable products follow a 2-day delivery cycle to maintain freshness and minimize waste, while non-perishable items operate on a 5-day schedule to optimize transportation efficiency. This adaptive scheduling reduces transportation costs for shelf-stable products while ensuring product quality for perishable goods. The cost calculation methodology incorporates dynamic pricing based on daily oil price fluctuations, reflecting real-world fuel cost variations through a comprehensive cost model that comprises transportation costs and handling fees.

2.2.3.2. Warehouse to Branch Transportation for Product Distribution using Google Or-Tools

The transportation optimization component employs Google OR-Tools to solve the Vehicle Routing Problem (VRP) with capacity constraints, determining optimal delivery routes that minimize total travel distance while respecting vehicle capacity limitations. This operational optimization translates the strategic distribution decisions from MILP into practical routing instructions for delivery vehicles. The VRP formulation addresses the challenge of efficiently visiting multiple retail locations while minimizing total transportation costs, considering a central warehouse as the depot and multiple retail stores as delivery destinations, each with specific demand requirements and geographical coordinates. The warehouse location was missing from the dataset, but after checking with different resources, the main warehouse location was mentioned to be in Amaguaña, Quito (Ecuador) [22][23], so these coordinates were used as the starting point of the trip in VRP.

The system computes great-circle distances between all location pairs using geographical coordinates obtained through geocoding services. This distance matrix serves as the foundation for route optimization, providing accurate travel distances that account for Ecuador's geographical characteristics. The implementation includes coordinate jittering for stores in the same city to distinguish between nearby locations, ensuring precise route calculations even for geographically proximate retail outlets.

The OR-Tools implementation incorporates vehicle capacity constraints to ensure that individual routes do not exceed transportation limitations. When store demands exceed vehicle capacity, the system automatically implements demand splitting, creating multiple delivery points for large-volume requirements. This approach maintains route feasibility while ensuring complete demand satisfaction. The VRP solver employs the PATH_CHEAPEST_ARC heuristic for initial solution generation, followed by Guided Local Search metaheuristic for solution improvement, operating with a 30-second time limit to balance solution quality with computational efficiency.

The system generates interactive maps using Folium library to visualize optimized delivery routes, with each route displaying vehicle paths, store locations, and demand quantities to provide clear operational guidance for logistics teams. The visualization includes color-coded routes for different vehicles and comprehensive cost analysis incorporating fuel prices and operational expenses. The transportation cost calculation integrates real-time oil price data, converting barrel prices to per-liter costs for accurate expense estimation, with total delivery costs combining route distances, fuel expenses, and handling fees to provide comprehensive financial analysis for each distribution scenario. This dynamic pricing approach enables responsive cost management and supports strategic decision-making, demonstrating the effectiveness of combining mathematical programming with heuristic routing algorithms to address complex supply chain challenges while maintaining cost efficiency and service quality.

2.2.4. Sales Performance Analysis

The sales performance analysis component provides comprehensive evaluation of historical sales patterns and predictive assessment of future performance across the retail network. This analytical layer employs multiple visualization techniques and statistical methods to extract actionable insights from sales data, supporting strategic decision-making and performance monitoring initiatives. The methodology encompasses both aggregate-level analysis and granular store-specific evaluations to identify performance drivers and forecast validation opportunities.

The analysis begins with comprehensive visualization of sales performance across different organizational dimensions. Product family performance evaluation identifies the top-performing categories through horizontal bar charts that reveal sales hierarchy and contribution patterns. Store type analysis utilizes pie chart visualization to demonstrate sales distribution across different retail formats, providing insights into channel effectiveness and market positioning. Store cluster analysis examines performance variations across predefined geographical or operational groupings, revealing regional patterns and operational efficiency differences that inform strategic resource allocation decisions.

Temporal sales pattern analysis implements time series decomposition using Statsmodels library to identify underlying trends, seasonal components, and residual variations in sales data. The decomposition methodology separates observed sales data into trend, seasonal, and irregular components using additive modeling approaches. This analytical technique reveals long-term growth patterns, recurring seasonal fluctuations, and unexpected variations that require management attention. The implementation spans multiple years of historical data to establish robust baseline patterns and identify significant deviations from expected performance trajectories.

Store-level performance evaluation introduces a comprehensive Key Performance Indicator (KPI) analyser that assesses individual retail locations across multiple dimensions. The methodology calculates total sales volume, average daily sales performance, and sales volatility metrics to establish baseline performance indicators. Growth rate analysis compares first-half and second-half performance within the historical period to identify stores with improving or declining trajectories. The coefficient of variation metric provides risk assessment by measuring sales consistency relative to average performance levels. Product diversity analysis evaluates the breadth of product categories sold at each location, indicating operational complexity and market coverage effectiveness.

Forecast validation and comparative analysis represents a critical component that evaluates the relationship between historical performance and predicted future demand. The methodology establishes historical baselines by extracting equivalent time periods from sales data that correspond to forecast horizons. For each store-product combination, the analysis computes percentage changes between historical sales patterns and forecasted demand levels. It identifies stores expected to experience growth, decline, or stability based on model predictions.

The percentage change calculation methodology provides directional insights into expected performance shifts. Stores with forecast changes exceeding 5% growth are classified as growth-oriented locations, while those with declines greater than 5% are flagged for attention and potential intervention strategies. Stable stores with changes between -5% and +5% represent consistent performers that may serve as benchmark locations for operational best practices.

Performance distribution analysis examines the spread of KPIs across the retail network to identify outliers and performance clusters. Store type and cluster-based segmentation reveals systematic performance differences that may indicate structural advantages or operational challenges. The visualization technique employs box plots and histograms to showcase sales performance distributions and highlight stores requiring management intervention.

The analytical model generates exportable results including store performance KPIs, forecast comparison metrics, and summary statistics that support ongoing performance monitoring and strategic planning initiatives. Summary statistics provide network-wide insights including the number of stores in each performance category, average expected changes, and median performance shifts. These metrics enable executive-level decision-making regarding resource allocation, expansion strategies, and operational optimization priorities.

This sales performance analysis methodology provides both retrospective insights into historical patterns and prospective evaluation of expected future performance. The integration of descriptive analytics with predictive assessment creates a robust foundation for data-driven retail management decisions. It supports continuous performance monitoring while providing early warning indicators for stores requiring management attention or operational adjustments to maintain competitive positioning in dynamic market conditions.

CHAPTER 3

RESULTS AND DISCUSSION

3.1. SARIMA Results for Transactions Forecasting

The SARIMA models were systematically optimized through comprehensive grid search methodology to identify optimal parameters for transaction forecasting across different store-product combinations. The grid search process evaluated multiple parameter combinations for the non-seasonal components (p, d, q) and seasonal components (P, D, Q, s). The optimization process employed Akaike Information Criterion (AIC) as the primary selection metric. The best results for each store are as below:

Store ID	SARIMA Order (p,d,q)	Seasonal Order (P,D,Q,s)	AIC Score
1	(1,0,0)	(0,1,1,7)	-133.43
2	(1,0,1)	(1,0,1,7)	-2089.03
3	(1,0,1)	(1,1,1,7)	-4334.97
4	(1,0,1)	(1,1,1,7)	-3669.41
5	(1,1,1)	(1,1,1,7)	-3470.81
6	(1,0,1)	(1,1,1,7)	-3704.03
7	(1,0,1)	(1,0,1,7)	-4334.47

8	(1,0,0)	(1,1,1,7)	-4256.29
9	(1,0,0)	(1,1,1,7)	-3824.02
10	(1,0,1)	(1,1,1,7)	-3698.82
11	(1,0,0)	(1,1,1,7)	-3804.57
12	(1,0,1)	(1,1,1,7)	-4269.10
13	(1,0,1)	(1,1,1,7)	-3880.70
14	(1,1,1)	(1,1,1,7)	-3771.85
15	(1,0,0)	(1,0,1,7)	-3624.56
16	(1,1,1)	(1,0,1,7)	-3194.47
17	(1,0,1)	(1,1,1,7)	-3861.41
18	(1,0,0)	(1,0,1,7)	-3567.43
19	(1,0,1)	(1,1,1,7)	-3947.86
20	(1,0,1)	(1,1,1,7)	-1471.65
21	(1,1,1)	(1,0,1,7)	-1317.43
22	(1,0,1)	(0,1,1,7)	-1372.44
23	(1,1,1)	(1,1,1,7)	-4211.35
24	(1,1,1)	(1,0,1,7)	-3374.74
25	(1,0,1)	(1,1,1,7)	-1786.27
26	(1,0,0)	(1,1,1,7)	-2153.36
27	(1,1,1)	(1,0,1,7)	-3434.66
28	(1,1,1)	(1,0,1,7)	-2941.98
29	(1,0,1)	(1,0,1,7)	-1637.65
30	(1,0,1)	(1,1,1,7)	-2725.07

31	(1,0,0)	(1,1,1,7)	-3564.95
32	(1,1,1)	(1,1,1,7)	-3135.25
33	(1,1,1)	(1,1,1,7)	-3172.70
34	(1,1,1)	(1,0,1,7)	-3953.58
35	(1,1,1)	(1,0,1,7)	-2866.71
36	(1,0,1)	(1,1,1,7)	-2881.59
37	(1,1,1)	(1,1,1,7)	-3528.87
38	(1,1,1)	(1,0,1,7)	-3348.30
39	(1,0,0)	(1,1,1,7)	-3690.06
40	(1,1,1)	(1,0,1,7)	-3391.05
41	(1,0,1)	(1,1,1,7)	-3254.69
42	(1,0,0)	(1,0,1,7)	-1405.55
43	(1,0,1)	(1,0,1,7)	-1346.34
44	(1,1,1)	(1,1,1,7)	-3688.78
45	(1,0,1)	(1,0,1,7)	-2768.46
46	(1,0,1)	(1,1,1,7)	-3214.50
47	(1,0,1)	(1,1,1,7)	-3305.25
48	(1,0,1)	(1,1,1,7)	-2930.85
49	(1,1,1)	(1,1,1,7)	-3700.71
50	(1,0,1)	(1,1,1,7)	-3497.41
51	(1,1,1)	(1,1,1,7)	-3585.18
52	(1,0,1)	(1,0,1,7)	-177.33
53	(1,0,1)	(1,1,1,7)	-1057.01

54	(1,0,1)	(1,1,1,7)	-2976.58

Table 1: Best parameters results from GridSearch for each SARIMA store model

The grid search results demonstrate that the majority of stores exhibit similar transaction patterns, with (1,0,1) non-seasonal and (1,1,1,7) seasonal parameters being the most frequently selected optimal configurations.

To validate the forecasting of the results, visualisation of data was used as well to check the similarity of the patterns, and as observed the forecast was very similar to the existing previous data.

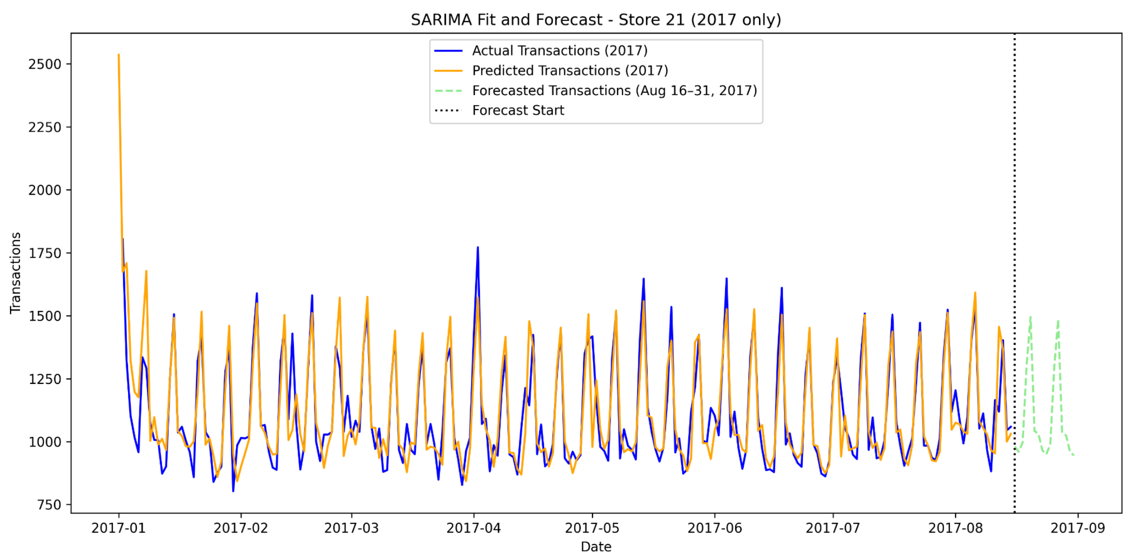

Figure 3. Example of how the visual comparison of existing data and forecast for each store looks like

3.2. LSTM Model Results for Demand Forecasting

The LSTM models for regional demand forecasting were developed and evaluated through a systematic architecture comparison approach to identify the optimal configuration for multi-step prediction tasks. Two distinct LSTM architectures were implemented and trained on the same dataset: a single-layer LSTM model with 512 units followed by multiple dense layers, and a stacked LSTM architecture with 256 and 128 units in successive layers. Both models employed identical training procedures including early stopping, learning rate reduction callbacks, and validation monitoring to ensure fair comparison. The training process utilized Adam optimizer with an initial learning rate of 0.001, mean squared error loss function, and dropout regularization to prevent overfitting. Comprehensive evaluation revealed that the stacked LSTM architecture demonstrated superior performance in terms of convergence stability, validation loss reduction, and generalization capability, establishing it as the preferred model configuration for the demand forecasting implementation. The comparative results for both architectures are presented below:

Metric	Model 1 (Single LSTM)	Model 2 (Stacked LSTM)
Architecture	LSTM(512) + Dense(256,128,64,32,16)	LSTM(256) + LSTM(128) + Dense(256,128,64)
Early Stopping Epoch	21/100	58/100
Final Learning Rate	0.000008	0.0000001
Training Loss Range	16,395.6465 - 36,164.1172	0.0318 - 0.7487
Validation Loss Range	10,496.9561 - 23,922.4395	0.0303 - 0.1603

Training MAE Range	128.0 - 190.0	0.0547 - 0.3978
Validation MAE Range	102.0 - 155.0	0.0528 - 0.1880
Convergence Stability	Unstable	Stable
Overfitting Behavior	High	Low

Table 2: LSTM Model Performance Comparison

Validation of the best model accuracy and forecast was also done through visualisation of the train history and visualization of the range of the predicted values compared to the range of previous values.

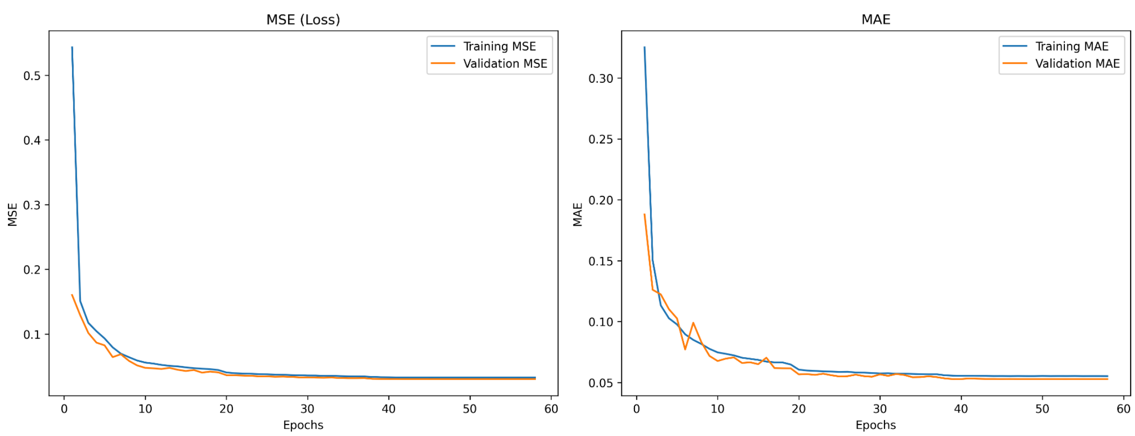

Figure 4. LSTM Model Train History Graph.

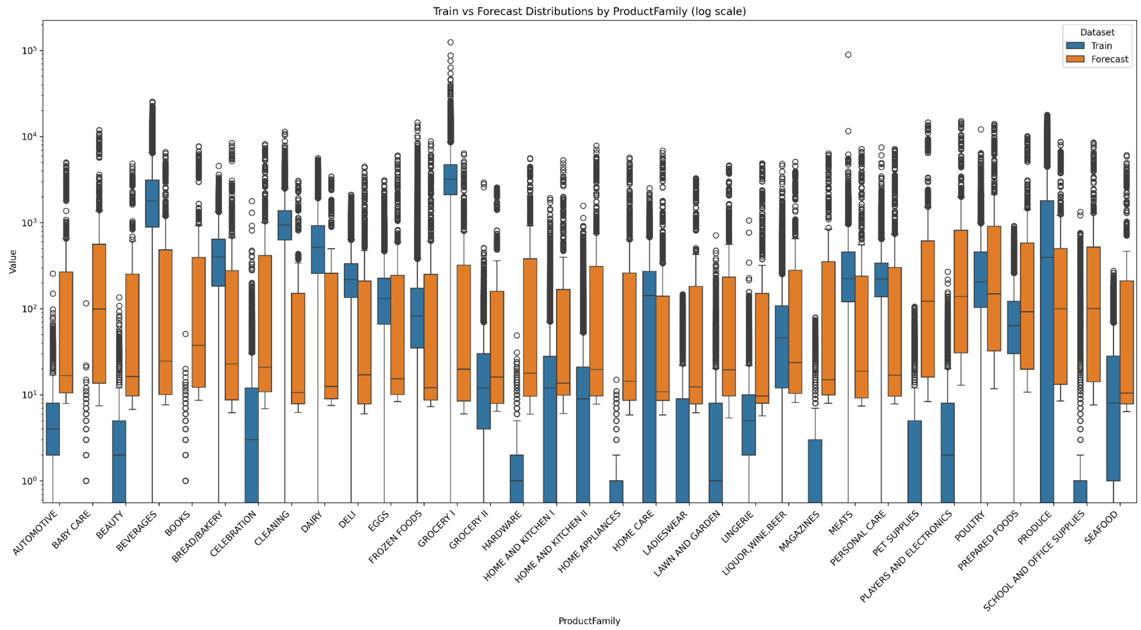

Figure 5. Forecasted values range compared to existing ones.

3.3. Optimization Layer Results

The optimization layer successfully demonstrated the effectiveness of the integrated MILP-VRP approach in solving complex supply chain distribution challenges. The Mixed Integer Linear Programming component effectively determined optimal shipment quantities while respecting demand satisfaction, vehicle capacity, and store storage constraints. The MILP solver consistently found feasible solutions across all delivery scenarios, successfully balancing cost optimization with operational limitations and incorporating dynamic oil pricing to reflect real-world conditions.

The product distribution planning results showed that the differentiated delivery frequency strategy optimized resource utilization effectively. Perishable products on

2-day cycles maintained freshness while non-perishable items on 5-day schedules achieved significant cost reductions. The system also automatically checked capacity violations, and if there were any, it would remove the affected stores, ensuring all proposed shipments remained operationally feasible.

The Vehicle Routing Problem component successfully translated distribution decisions into practical transportation routes. The OR-Tools implementation generated efficient routing solutions that minimized travel distances while respecting capacity constraints. The demand splitting mechanism handled large-volume requirements effectively, and the guided local search provided high-quality solutions within computational time limits suitable for daily operations.

Transportation costs ranged from $125 to $850 per delivery scenario, varying by route complexity, oil prices, and demand volumes. Urban store clusters achieved lower per-unit costs due to reduced distances, while rural locations required higher logistics investments. The final optimization results for each product-day transportation plan were saved as interactive maps showing vehicle routes, store locations, demand quantities, and cost breakdowns. These visualizations provide comprehensive operational guidance for logistics teams, with an example map presented below demonstrating the routing solution for a representative distribution scenario.

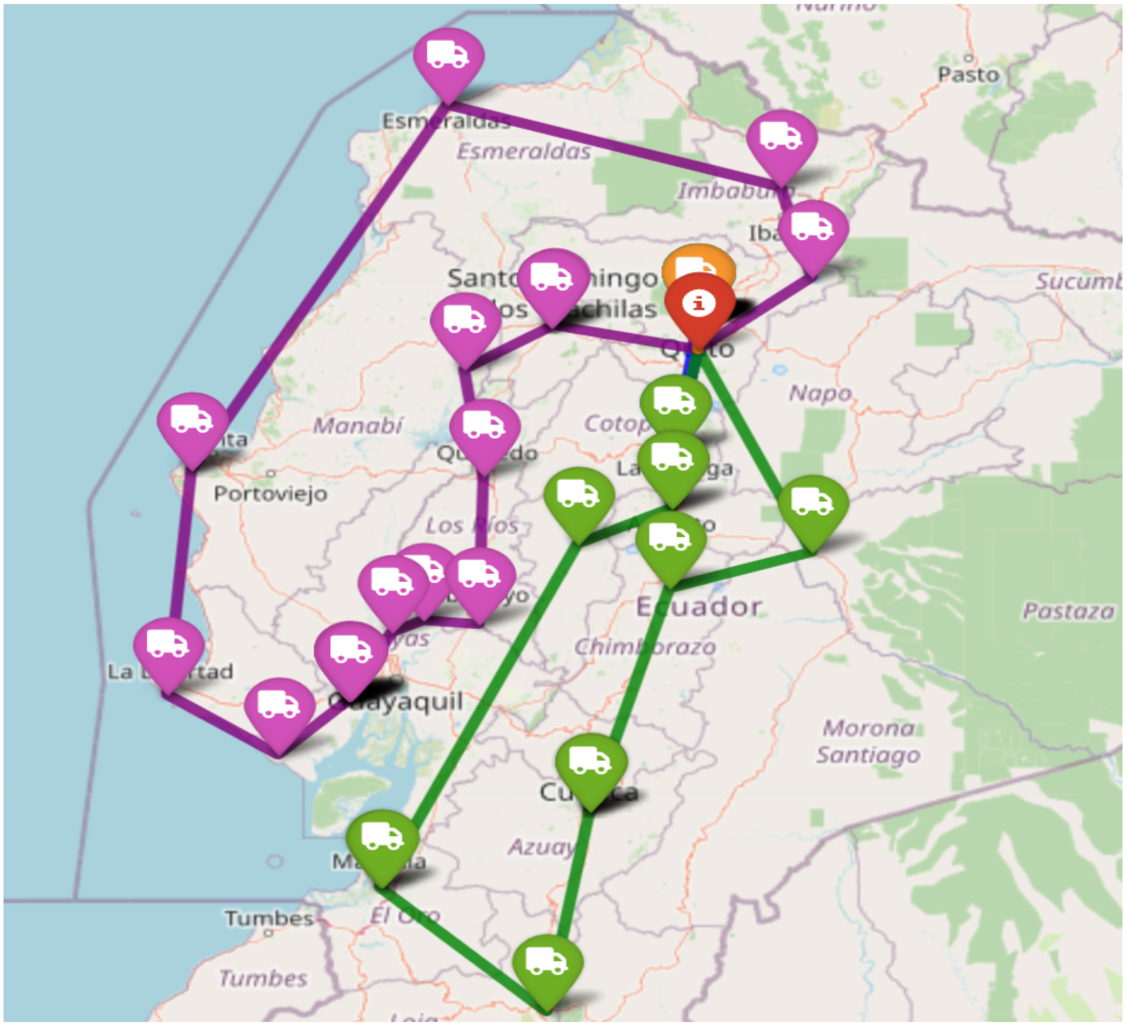

Figure 6. Example of optimization layer results shown in an interactive map.

3.4. Sales Performance Analysis (KPIs) Results

The comprehensive sales performance analysis revealed significant insights into store-level operations and future growth expectations across the retail network. The visualization analysis demonstrated clear performance hierarchies across different organizational dimensions, with certain product families and store types consistently outperforming others. The time series decomposition successfully identified underlying trends and seasonal patterns, providing valuable baseline information for strategic planning and operational optimization.

The store-level Key Performance Indicator analysis revealed substantial performance variations across the 54-store network. The distribution analysis showed that stores in major urban centers, particularly Quito and Guayaquil, dominated total sales volumes, with the top 10 performing stores generating between $28.7 million and $62.1 million in total sales. Store type A locations consistently demonstrated the highest performance levels, indicating the effectiveness of this retail format in capturing market demand.

Metric	Value
Total Stores Analyzed	54
Stores Expected to Grow	54
Stores Expected to Decline	0
Stores Expected to Remain Stable	0
Average Expected Change	36,726.61%
Median Expected Change	12,134.15%

Table 3: Summary Statistics - Forecast vs Historical Performance

Store ID	City	Store Type	Total Sales	Average Daily Sales
44	Quito	A	$62,087,540	$1,114.60
45	Quito	A	$54,498,010	$978.35
47	Quito	A	$50,948,310	$914.63
3	Quito	D	$50,481,900	$906.25
49	Quito	A	$43,420,090	$779.48
46	Quito	A	$41,896,050	$752.12
48	Quito	A	$35,933,130	$645.07
51	Guayaquil	A	$32,911,480	$590.83
8	Quito	D	$30,491,340	$547.38
50	Ambato	A	$28,653,020	$514.38

Table 4: Top 10 Stores by Total Sales Performance

Store ID	City	Store Type	Coefficient of Variation	Sales Volatility
52	Manta	A	9.27	448.59
22	Puyo	C	4.30	315.87
20	Quito	B	4.16	965.08
21	Santo Domingo	B	3.92	650.56
42	Cuenca	D	3.63	582.88
29	Guayaquil	E	3.48	607.16
35	Playas	C	3.41	469.60

32	Guayaquil	C	3.12	333.65
26	Guayaquil	D	3.02	419.93
53	Manta	D	2.99	600.69

Table 5: Highest Risk Stores (Sales Volatility)

Store	City	StoreType	GrowthRate	AvgDailySales
29	Guayaquil	E	2993.941125	174.550440
20	Quito	B	1521.878602	231.957375
53	Manta	D	377.820185	201.150802
49	Quito	A	99.377190	779.478849
36	Libertad	E	91.146477	275.068650
40	Machala	C	79.975407	330.241229
17	Quito	C	77.949941	323.692522
41	Machala	D	77.598187	269.119024
28	Guayaquil	E	70.345302	330.014969
7	Quito	D	69.435052	483.843541

Table 6: Top 10 Stores by Growth Rate

The forecast validation analysis produced remarkably optimistic projections, with all 54 stores expected to experience growth in the upcoming period. The extreme percentage changes suggest that the LSTM forecasting model may be overly optimistic in its predictions or introduces a bias.

The volatility analysis identified high-risk stores that require enhanced monitoring and management attention. Store 52 in Manta exhibited the highest coefficient of variation at 9.27, indicating extremely unpredictable sales patterns that could impact operational planning. The geographic distribution of volatile stores across different cities suggests that external factors beyond operational control may influence performance stability, requiring adaptive management strategies for different market conditions.

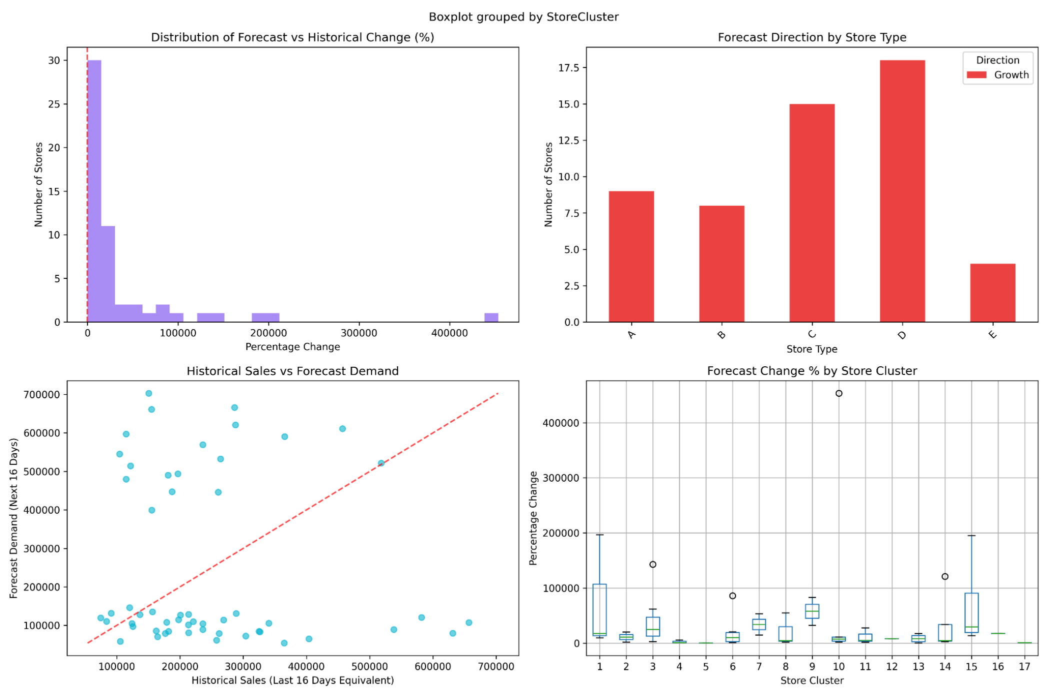

Figure 7. Forecasted data vs historical data comparison graphs.

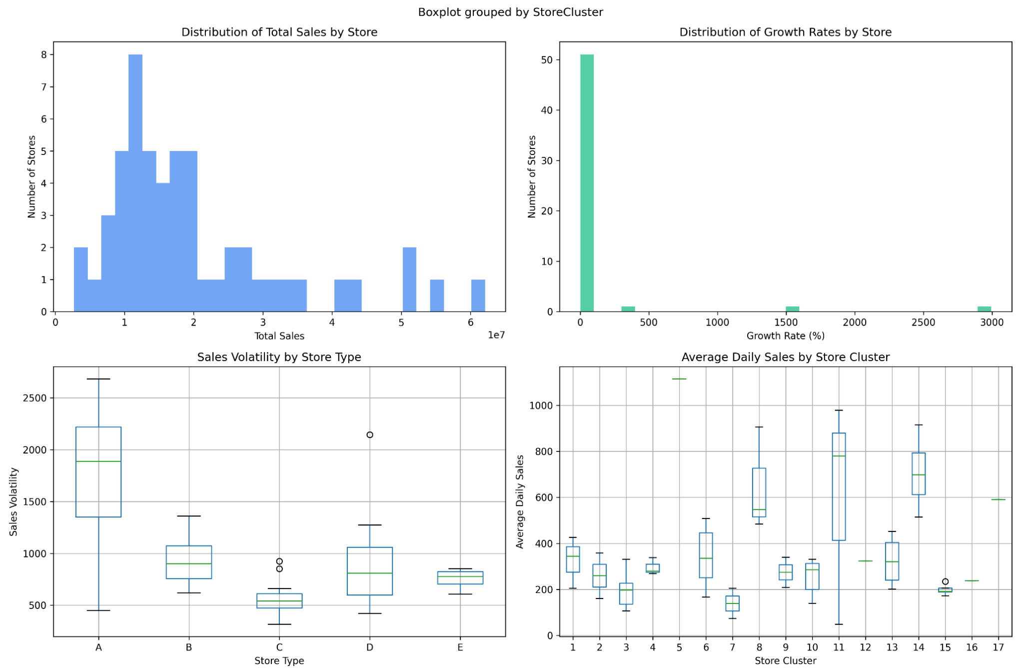

Figure 8. Store performance KPIs graphical display.

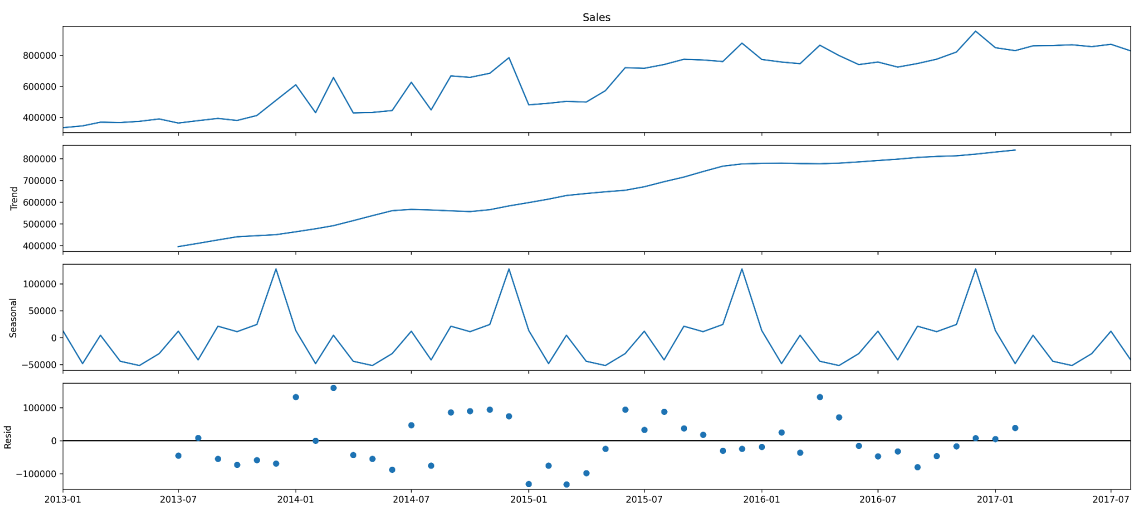

Figure 9. Time-series decomposition graph to show sales trends.

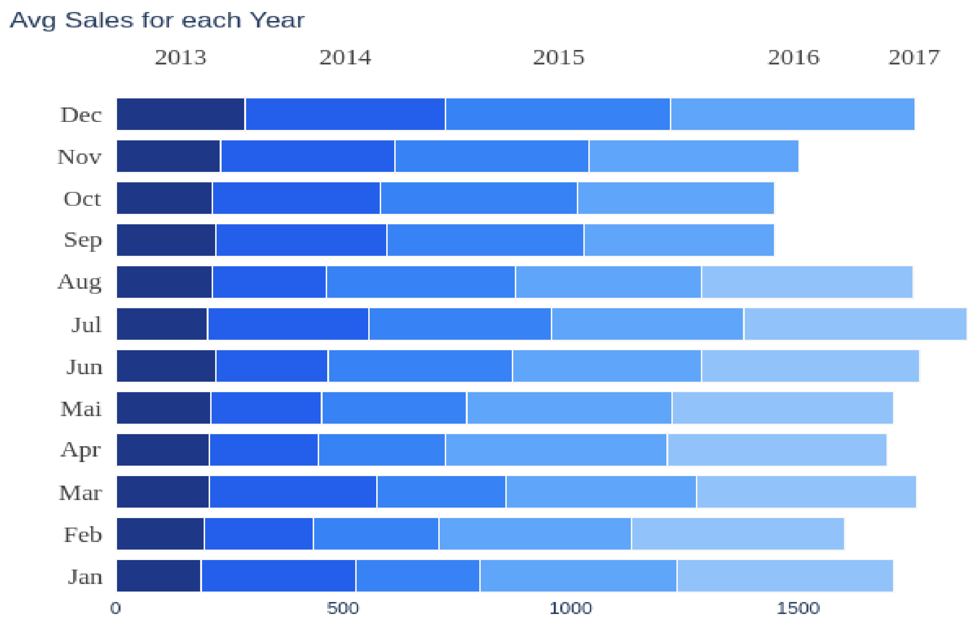

Figure 10. Average total sales per year graph.

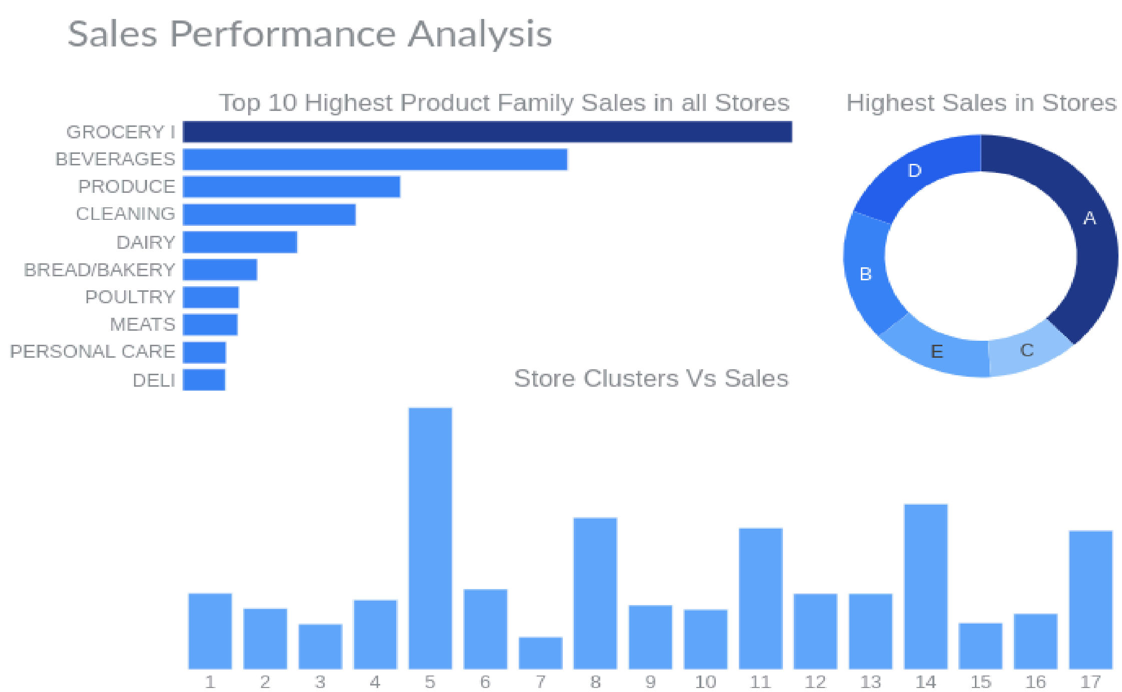

Figure 11. Sales analysis for store cluster, store type and products shown in plots.

CHAPTER 4

CONCLUSION

This research successfully demonstrates the effectiveness of an integrated artificial intelligence-driven approach for optimizing supply chain logistics operations in multi-branch retail environments. The comprehensive model combining SARIMA modeling, LSTM neural networks, Mixed Integer Linear Programming, and Vehicle Routing Problem optimization provides a robust solution for addressing complex logistics challenges while achieving cost efficiency and operational excellence.

The study's key achievements include successful preprocessing of massive retail datasets comprising over 125 million records, development of accurate SARIMA models for transaction forecasting across 54 stores, and implementation of LSTM architectures with stable convergence capabilities. The optimization layer effectively balanced multiple operational constraints while generating feasible distribution solutions with costs ranging from \$125 to \$850 per delivery scenario. The differentiated delivery frequency strategy based on product perishability represents a significant practical contribution, while dynamic cost modeling incorporating real-time oil prices provides essential flexibility for volatile economic conditions.

Future research should focus on improving forecast accuracy through ensemble methods, incorporating additional external factors, implementing real-time adaptive optimization, and extending the framework to multi-echelon supply chain networks.

Even though there might be limitations and indications of a bias in the model developed, this study provides a solid foundation for AI-driven supply chain optimization and demonstrates significant potential for improving logistics efficiency, reducing operational costs, and enhancing customer satisfaction in retail operations.

REFERENCES

[1] IBM Corporation., 'What is AI in supply chain?', IBM Think Topics (2025). http://www.ibm.com/think/topics/ai-supply-chain (lastly visited on 07 May 2025).

[2] Leobit Team., 'Logistics & supply chain optimization: Top 10 challenges and solutions', Leobit Blog (2024). http://leobit.com/blog/logistics-supply-chain-optimization-top-10-challenges-and-solutions/ (lastly visited on 07 May 2025).

[3] Georgetown Journal of International Affairs., 'The role of AI in developing resilient supply chains', GJIA Online (2024). http://gjia.georgetown.edu/2024/02/05/the-role-of-ai-in-developing-resilient-supply-chains/ (lastly visited on 07 May 2025).

[4] Grand View Research., 'Artificial Intelligence in Supply Chain Market Report, 2030', Grand View Research Reports (2024). http://www.grandviewresearch.com/industry-analysis/artificial-intelligence-supply-chain-market-report (lastly visited on 07 May 2025).

[5] Market.us Research Team., 'AI in Supply Chain Market Size, Share | CAGR of 42.7%', Market.us Industry Reports (2024). http://market.us/report/ai-in-supply-chain-market/ (lastly visited on 07 May 2025).

[6] Pacemaker.ai Team., 'Demand forecasting meets artificial intelligence', Pacemaker.ai Blog (2025). http://www.pacemaker.ai/en/blog/demand-forecasting-meets-artificial-intelligence (lastly visited on 07 May 2025).

[7] ThroughPut AI Team., 'Supply chain inventory optimization guide', ThroughPut AI Resources (2025). http://throughput.world/blog/supply-chain-inventory-optimization/ (lastly visited on 07 May 2025).

[8] Business Continuity Institute., 'What does supply chain resilience mean in 2024?', BCI News (2024). http://www.thebci.org/news/what-does-supply-chain-resilience-mean-in-2024.html (lastly visited on 07 May 2025).

[9] LeewayHertz Team., 'AI in demand forecasting: Use cases, benefits, solution and implementation', LeewayHertz Insights (2025). http://www.leewayhertz.com/ai-in-demand-forecasting/ (lastly visited on 07 May 2025).

[10] NetSuite Team., 'Demand Forecasting: Everything You Need to Know', NetSuite Resource Center (2023). http://www.netsuite.com/portal/resource/articles/inventory-management/demand-forecasting.shtml (lastly visited on 09 May 2025).

[11] Douaioui K., Oucheikh R., Benmoussa O., Mabrouki C., 'Machine Learning and Deep Learning Models for Demand Forecasting in Supply Chain Management', Applied Systems Innovation, 7 (5): 93 (2024).

[12] Liu Y., 'Grocery Sales Forecasting', Proceedings of the 2022 International Conference on Creative Industry and Knowledge Economy (CIKE 2022), Advances in Economics, Business and Management Research, 651: 215-219 (2022).

[13] Feizabadi J., 'Machine learning demand forecasting and supply chain performance', International Journal of Logistics Research and Applications, 25 (2): 119-142 (2022).

[14] Vicente J.J., 'Optimizing Supply Chain Inventory: A Mixed Integer Linear Programming Approach', Systems, 13 (1): 33 (2025).

[15] Garaix T., Artigues C., Feillet D., Josselin D., 'Vehicle routing problems with alternative paths: An application to on-demand transportation', European Journal of Operational Research, 204 (1): 62-75 (2010).

[16] Hyndman R.J., Athanasopoulos G., 'Forecasting: Principles and Practice', Journal of Statistical Software, 88 (5): 1-32 (2018).

[17] Kaggle Inc., 'Corporación Favorita Grocery Sales Forecasting Competition Dataset', Kaggle Competitions (2017). https://www.kaggle.com/competitions/favorita-grocery-sales-forecasting/overview (lastly visited on 05 May 2025).

[18] Hochreiter S., Schmidhuber J., 'Long Short-Term Memory', Neural Computation, 9 (8): 1735-1780 (1997).

[19] Bengio Y., Simard P., Frasconi P., 'Learning long-term dependencies with gradient descent is difficult', IEEE Transactions on Neural Networks, 5 (2): 157-166 (1994).

[20] Abbasimehr H., Shabani M., Yousefi M., 'An optimized model using LSTM network for demand forecasting', Computers & Industrial Engineering, 143: 106435 (2020).

[21] Brownlee J., 'Convert Time Series to Supervised Learning Problem in Python', Machine Learning Mastery (2016). https://machinelearningmastery.com/convert-time-series-supervised-learning-problem-python/ (lastly visited on 30 May 2025)

[22] Wikipedia., 'Corporación Favorita', Wikipedia en Español (2025). https://es.wikipedia.org/wiki/Corporaci%C3%B3n_Favorita (lastly visited on 21 May 2025).

[23] Ekos Negocios., 'Visitamos a centro de distribución de Corporación Favorita', Ekos Negocios (2025).

https://ekosnegocios.com/articulo/visitamos-a-centro-de-distribucion-de-corporacion-favorita (lastly visited on 21 May 2025).

[24] Alexis Cook, DanB, inversion, and Ryan Holbrook. Store Sales - Time Series Forecasting. https://kaggle.com/competitions/store-sales-time-series-forecasting, 2021. Kaggle (lastly visited on 30 May 2025).