AMERICAN UNIVERSITY OF BEIRUT

**Optimizing Growth: Sales and Market Expansion for GIFCO**

by

Mona Ghorayeb

Advisor(s)

Dr. Sirine Taleb

Dr. Ahmad El Hajj

A Capstone Project

submitted in partial fulfillment of the requirements

for the degree of Master’s in Business Analytics

to the Suliman S. Olayan School of Business

at the American University of Beirut

Beirut, Lebanon

August 2021

An Abstract of the Capstone Project of

Mona Ghorayeb for Master’s in Business Analytics (MSBA)

**Title: Optimizing Growth: Sales and Market Expansion for GIFCO**

This project addresses a strategic opportunity for GIFCO, a logistics company operating without a dedicated business analytics department. The motivation was to harness data analytics to enhance operational visibility, forecast performance, and inform future market expansion all for investment gathering. The absence of analytical tools limited the company’s ability to understand historical trends and forecast growth leading to better general performance in the market.

The project aimed to provide GIFCO with a digital decision-support solution by transforming raw operational data from 2019 to 2024 into actionable insights. After acquiring and cleaning both sales and shipment datasets, extensive feature engineering and exploratory data analysis (EDA) were conducted using Python to identify performance patterns, profitability trends, and freight characteristics. A Tableau dashboard was then developed to visualize these insights, offering GIFCO a comprehensive overview of sales distribution, client activity, freight efficiency, and seasonal shipment trends.

To project future performance and enable strategic planning, multiple forecasting models were implemented, including ARIMA, SARIMA, Holt-Winters, Prophet, and Theta. Model performance was evaluated using RMSE, with the Theta model emerging as the most accurate (RMSE around 190,000). It was further optimized using grid search to refine hyperparameters, achieving the lowest forecast error.

The final deployment involved a Streamlit web application integrating interactive dashboards and forecasts, providing GIFCO with a real-time analytics platform to share with investors. Major findings include significant growth potential in certain product families and shipment routes, as well as seasonality effects that could inform logistics optimization. The six-year forecast equips GIFCO with data-driven projections to guide investment leading to potential geographic expansion.

This capstone demonstrates how data-driven transformation can unlock new growth strategies, investment gathering and support international branching decisions for companies new to analytics.

Table of Contents

1. Introduction…………………………………………………………………………………..…. 4
2. Background & Related Work…………………………………………………………………… 7
3. Methodology…………………………………………………………………………………… 10
4. Data Overview……………………………………………………………………………... 10
5. Data Preprocessing: Cleaning……………………………………………………………... 11
6. Data Preprocessing: Feature Engineering…………………………………………………. 12
7. Exploratory Data Analysis (EDA)………………………………………………………….13
8. Modeling Framework……………………………………………………………………… 15
9. Modeling Evaluation and Selection……………………………………………………….. 16
10. Streamlit Web Application………………………………………………………………… 17
11. Tools and Technologies……………………………………………………………………. 17
12. Results…………………………………………………………………………………………. 18
13. Tableau Dashboard………………………………………………………………………… 18
14. Forecasting Models Results…………………………………………………………………25
15. Streamlit Web Application………………………………………………………………… 29
16. Discussions…………………………………………………………………………………….. 32
17. Conclusion……………………………………………………………………………………... 36
18. References……………………………………………………………………………………... 38

Introduction

In the current era of data and digital transformation, the capacity to harness data for informed decision-making has become not only a strategic advantage but a fundamental necessity. According to a 2022 report by McKinsey & Company, companies that extensively use customer and operational analytics are 23 times more likely to outperform their competitors in customer acquisition, 19 times more likely to achieve above-average profitability, and 7 times more likely to retain customers (McKinsey & Company, 2022). Similarly, the International Data Corporation estimates that global spending on big data and analytics reached $274 billion in 2022, highlighting a growing global shift toward evidence-based, predictive decision-making across sectors, including logistics (IDC, 2022). The logistics and freight forwarding industry, in particular, is undergoing a profound shift fueled by technological innovation, global interconnectedness, and the rising expectations of efficiency, transparency, and scalability. Businesses that once relied on manual tracking, intuitive planning, and static reports are now required to adapt rapidly, integrating dynamic data analytics to remain competitive. As the complexity of supply chains increases and customer demands grow more nuanced, companies that fail to leverage their operational data risk falling behind.

While data analytics and forecasting have been widely adopted in global markets, there remains a significant gap in their deployment across small and medium-sized enterprises (SMEs) in developing regions. The pace of adoption remains uneven, especially in these developing regions. In Lebanon, digital transformation is still in its early stages. A 2023 World Bank report noted that while over 80% of Lebanese firms collect operational data, fewer than 15% actually analyze that data systematically to inform business strategy (World Bank, 2023). Despite Lebanon’s strategic location, port access, and active trade connections with Africa and the Gulf, its logistics firms have only recently begun recognizing the strategic value of data. Amid economic volatility and regional instability, Lebanese companies that embrace data analytics are more likely to survive, adapt, and grow. One such firm at a pivotal moment in its operational trajectory is GIFCO.

GIFCO is a Lebanese freight forwarding company offering end-to-end logistics solutions by land, sea, and air. With a track record of growth and resilience, GIFCO manages a wide portfolio of clients and maintains operations in Lebanon and Juba since 1971 with over 684,000 shipments delivered and 129 milestones attained (GIFCO, 2025). Despite its regional footprint and substantial data accumulated over the years, GIFCO operates without a formal business analytics department. Decisions regarding pricing strategies, route optimization, profit assessment, and expansion planning have historically been made based on managerial intuition and experience, and fragmented Excel reports while still adopting the traditional way of doing things. While this has served the company for many years, it is increasingly insufficient in a landscape where agility and data-driven foresight are vital. In today’s competitive logistics and supply chain landscape, leveraging data to inform strategic decisions has become a necessity rather than a luxury. Businesses are increasingly relying on data-driven solutions to optimize operations, forecast future performance, and identify growth opportunities. Recognizing this, GIFCO management expressed the need for a robust yet accessible analytics system that could provide a clear overview of historical performance, generate forecasts for the upcoming years, and support strategic decisions such as investments gathering and market expansion. This capstone project was initiated to address this exact gap for GIFCO. The core problem was clear: how can GIFCO leverage its existing operational data to better understand its historical performance, uncover key business drivers, and build a forecasting and analytics framework that is scalable and aligned with the company’s strategic goals of gathering investment for expansion? In addressing this question, the project not only sought to equip GIFCO with the technical tools it lacked, but to demonstrate the broader potential of applied data analytics in the Lebanese logistics sector.

The methodology adopted for this project followed a structured, end-to-end data analytics pipeline, transforming GIFCO’s raw operational data into actionable business intelligence. The approach combined best practices in data engineering, exploratory analytics, and forecasting, all tailored to the company’s specific needs and contextual limitations. This project will provide GIFCO’s management and potential investors an interface to visualize historical performance, explore 6-year forecasts and data-driven insights, embedding analytics into their strategic planning process in a sustainable manner.

What distinguishes this project is not the novelty of the models themselves or the tools, many of which are established in the academic and industrial literature, but rather the contextualized application of these tools to a real-world Lebanese company operating in a traditionally non-digital sector. The fusion of data science with domain-specific knowledge, operational goals, and localized constraints creates a framework that is both innovative and replicable. Similar firms across Lebanon and the Middle East could benefit from comparable solutions, thereby fostering a culture of data-driven growth.

This project contributes to the evolving discourse on digital adoption in under-digitized economies. It highlights how even medium-sized firms in regions like Lebanon and Juba can harness modern analytics to navigate uncertainty, evaluate strategic options, and unlock new growth opportunities. Moreover, it reinforces the importance of accessibility in digital transformation by employing open-source tools and scalable methods, the project ensures that cost is not a barrier to intelligence.

In conclusion, this capstone represents a practical intervention into the operational challenges faced by logistics companies in emerging markets. By building an analytics foundation for GIFCO, the project lays the groundwork for smarter planning for targeted market expansion, strategy shifting and investments gathering. It also demonstrates that with the right tools and methodology, data can become the cornerstone of transformation even in the most resource-constrained environments.

Background and Related Work

In an era dominated by digital economies and real-time decision-making, data analytics has emerged as a cornerstone of operational efficiency and strategic foresight. Across industries, data-driven methods are being leveraged to streamline processes, identify hidden patterns, and forecast future trends. The logistics and freight forwarding industry, in particular, has witnessed transformative improvements due to the integration of analytics. From route optimization and demand forecasting to supply chain resilience and customer satisfaction, data analytics play a critical role in enabling companies to respond swiftly to market dynamics. According to Accenture (2021), 98% of supply chain executives reported that data analytics is essential to delivering a customer-centric supply chain, while 81% of leading firms use predictive analytics to improve supply chain performance.

The use of time series forecasting models in logistics has become especially significant in the wake of globalization and complex supply chain networks. Classical models such as ARIMA (Auto-Regressive Integrated Moving Average), SARIMA (Seasonal ARIMA), and Holt-Winters exponential smoothing are widely adopted due to their interpretability and effectiveness in capturing linear trends and seasonal components. More recently, machine learning-based approaches such as Facebook Prophet and PyCaret’s automated time series modeling have gained traction for their flexibility, scalability, and robustness to missing or noisy data (Taylor & Letham, 2018). These models have been deployed in various logistics-related contexts including inventory control, transportation scheduling, and revenue forecasting.

The power of data analytics in logistics is evident in several real-world applications. DHL, one of the global leaders in logistics, uses predictive analytics and machine learning to optimize parcel delivery times and fleet management. Through its Resilience360 platform, DHL monitors supply chain risks in real-time and uses big data to reroute shipments during disruptions. Similarly, FedEx has implemented AI-driven predictive maintenance on its vehicles to minimize downtime and operational costs. Amazon, renowned for its logistics infrastructure, uses sophisticated forecasting algorithms and real-time data streams to anticipate customer demand and dynamically allocate inventory across fulfillment centers. These examples showcase how analytics has evolved from simple reporting to a strategic function that enhances speed, reliability, and customer satisfaction.

In the academic realm, numerous studies have validated the effectiveness of data analytics in logistics and supply chain operations. Choi et al. (2020) demonstrated how big data and predictive models can reduce lead times, optimize inventory levels, and improve transportation efficiency. A study by Ghobakhloo (2018) and Waller and Fawcett (2013) highlighted that data analytics adoption is positively correlated with organizational agility, particularly in volatile markets fostering overall resilience through enhancing collaboration and demand sensing. A compelling example of data-driven market expansion comes from A.P. Moller on Maersk, which leveraged advanced analytics to guide its shift from ocean freight to integrated inland logistics. By analyzing container flow data, customer shipment behaviors, and trade route inefficiencies, Maersk identified high-demand corridors in underserved regions such as India and Sub-Saharan Africa. This insight led to the strategic opening of over 75 new inland logistics hubs, including dry ports and warehouses, between 2020 and 2023. Data models such as trade flow simulations, demand forecasting algorithms (like ARIMA and Prophet), and network optimization models enabled the company to simulate trade flow scenarios and demand forecasts before committing to infrastructure investments, ensuring targeted, efficient expansion (Maersk, 2023).

A notable illustration of how data analytics can reshape strategic planning comes from C.H. Robinson, one of the largest third-party logistics providers globally. The company implemented a centralized data analytics platform called Navisphere, which consolidated shipment tracking, route performance, and customer service metrics across all global operations. By analyzing this data in real time, C.H. Robinson uncovered inefficiencies in delivery routes, seasonal demand fluctuations, and inconsistencies in warehouse throughput. These insights led to significant changes in their logistics strategy, including the redesign of regional delivery networks, dynamic rerouting based on traffic and weather patterns, and customized service offerings for high-volume clients. As a result, the company reported enhanced on-time delivery rates, reduced costs, and greater customer satisfaction. This case demonstrates how logistics firms can use analytics not only to monitor business health but to actively adapt and optimize strategic operations (C.H. Robinson, 2022).

In this project, the PyCaret Theta model gave best RMSE and presented pretty accurate results. After research, an example of the PyCaret Theta model's application in logistics can be found in the work of logistics optimization research teams. In a study conducted by researchers at the Technical University of Munich (2022), the Theta model was used to forecast weekly inbound shipment volumes at a major European distribution hub. Compared to traditional ARIMA and Holt-Winters models, Theta achieved lower RMSE and better responsiveness to seasonal demand spikes, making it particularly suitable for logistics operations with fluctuating patterns. The insights derived from Theta’s forecasts enabled the company to optimize staffing schedules, allocate temporary storage units more effectively, and adjust vendor contracts in advance of high-volume periods (Schmid et al., 2022).

Previous studies have emphasized the importance of aligning data science with domain-specific challenges. For instance, Choi et al. (2020) emphasized the integration of analytics into the operational logic of logistics companies to create tangible value. Similarly, Christopher & Peck (2012) argue that the real impact of analytics is realized only when data models are embedded within business decision-making frameworks. This project aligns with that vision, ensuring that the analytical work is not isolated but deeply integrated into GIFCO’s strategic planning process.

Another practical study of sales forecasting in the logistics industry is seen in the case of DB Schenker, a global leader in supply chain management. Facing challenges with demand volatility and inconsistent forecasting accuracy, DB Schenker implemented a data-driven sales forecasting system that integrated time series models such as SARIMA, Prophet, and XGBoost. These models utilized historical sales data, macroeconomic indicators, customer profiles, and trade lane activity to predict monthly revenue across different regions and services. The transition from manual forecasting methods to this machine learning-based approach improved forecast accuracy by 25%, enabling more precise alignment of fleet deployment, staffing, and investment planning. By leveraging these forecasts, the company enhanced its operational efficiency and made more informed decisions regarding regional expansion and resource allocation (DB Schenker, 2021).

By tailoring forecasting and analytics to the specific constraints and goals of a logistics firm in Lebanon and Juba, this project contributes a rare, replicable case study to the broader literature on digital transformation in emerging markets. It demonstrates that when advanced analytics are made accessible and aligned with local realities, even firms in under-digitized economies can leverage them to drive operational intelligence, growth, and competitiveness.

Methodology

This methodology section describes the systematic approach followed to address the main goals of this project: to equip GIFCO with data-driven tools that enable forecasting and improve investment readiness to support market expansion decisions. The methodology combines data preprocessing including data cleaning and data engineering, exploratory data analysis (EDA), time series forecasting, and a Tableau dashboard comprising an overview of the business from 2019 till 2024. The final product would be a Streamlit app as an informative interface for decision-making.

Data

The datasets used in this project were provided directly by GIFCO and reflect operational records from 2019 to 2024. These datasets were exported from the company’s internal systems (Wizard Cloud), in excel format. The data provided were divided into 2 datasets:

* 1. Sales Data

The sales dataset spans from 2019 to 2024 and includes 13,755 observations and 36 columns. It captures details of every recorded service transaction across GIFCO’s operations in Lebanon. The features included: Service Attributes:

* **Service Attributes:** Family (sea freight, land freight, air freight), Sub-family (Export, Import, Storage, Clearance…).
* **Client Information:** Client Name, Branch, City, Client Category.
* **Transaction Metrics:** Invoice Number, Date, Product Code, Unit, Quantity.
* **Financial Fields:** Gross Amount (LBP & USD), Net Amount (LBP & USD), Amount in USD, Profit Margin.
  1. Shipment Data

The shipment dataset contains 12,536 observations from the year 2023-2024. This dataset was used to complement the sales view with information on freight movement and client activity across supply chains. The key features included:

* **Freight Details:** Freight Type (e.g., Land, Sea), Family, Commodity.
* **Client and Route:** Client Name, From (origin), To (destination), Shipper, Vessel.
* **Operational Data:** Shipment Number, Date, Packages, Gross Weight.

While the size of GIFCO’s datasets was moderate, their structure and format presented key challenges that required careful handling. The lack of consistency in formatting, naming conventions, and data completeness introduced complexity during preprocessing like for instance some From-To locations appeared as both “LBN - Lebanon” and “LBN - Lebanese Republic”, or sub-families were listed as “EXPORT”, “Export,” or left blank entirely, and key fields such as client names or shipment origins were missing in several entries.

Using Python’s pandas library, column names were standardized across both datasets to ensure compatibility throughout the cleaning and modeling pipeline. Service categories such as “Family”, “Sub-family”, “From”, “To” and others were cleaned to correct inconsistencies in capitalization, spelling, and spacing. Additionally, date fields were parsed into a unified datetime format, which served as the anchor for monthly and seasonal aggregations.

The datasets reflected operational diversity across multiple service lines (e.g., clearance, export, storage), routes, and freight types. Some records had missing or incomplete entries, particularly in client information or shipment routes. Although the datasets were not extensive, their variability in quality and structure required a focused effort to prepare them for forecasting. As a result, multiple preprocessing and enrichment operations were applied to transform the data into a reliable foundation for exploratory analysis and predictive modeling.

* 1. Data Preprocessing: Cleaning

Most of the efforts were dedicated to the cleaning data preprocessing step in order to make them suitable for analysis and time series forecasting. The following steps were done:

* **Column Reduction:** the redundant or irrelevant columns like Barcode, Address, Client Code, and Product Description were dropped to streamline the dataset. In addition, columns that had more than 50% of their data missing were dropped like Client Category. Moreover, for the sales dataset, the columns with the currency LBP were dropped as we are interested only in the USD exchanges (noting that the USD are the converted LBP amounts).
* **Handling Missing and Inconsistent Values:** key categorical fields like Family and Sub-family were missing in several records. These were filled using the most frequent value in the column. Fields like City, Client Category, and From/To in the shipment data were cleaned and imputed based on historical frequency or dominant entries (mode). In addition, rows that contained more than 50% missing values were also dropped.
* **Data Type Conversion:** financial and quantity fields (Net Amount USD, Gross Amount USD, Quantity) were originally read as object types due to inconsistent formatting. These were converted into float format using safe type coercion to allow arithmetic operations and modeling for the next steps in the project.
* **Outlier Detection and Removal:** records with clearly invalid entries, such as negative amounts, zero or null quantities, or incomplete invoices were flagged and removed. Duplicated records and test entries were also excluded to ensure accuracy in trend modeling.
* **Categorical Label Normalization:** variations in service and route names (like “Sea Freight” vs. “SEA FREIGHT”) were unified through string normalization techniques to avoid artificial fragmentation during grouping or aggregation. Also, “LBN - Lebanon” and “LBN - Lebanese Republic” were standardized to Lebanon in addition to the other countries through mapping. Finally, the number of categories in the “Family” column in the shipping data were reduced from 32 to 15 unified categories through mapping given from the company and inconsistencies in capitalization and missing words like “Prepared ” 🡪 “Prepared Foodstuff”.
* **Datetime Parsing and Temporal Alignment:** the date fields in both datasets were converted into Python datetime objects. Only the year was kept, since the goal is to analyze yearly trends, facilitating temporal resampling and time series modeling.
* **Client Mapping:** for privacy purposes, the client names were mapped into IDs were each client has an ID like Client\_1.
* **Duplicated Observations:** duplicated rows were checked in both datasets and none were detected, so no additional rows were dropped.

The resulting datasets were thoroughly cleaned and arranged, offering a strong basis for time series forecasting as well as exploratory research. These preprocessing efforts were critical in transforming GIFCO’s disorganized and incomplete data into a cohesive and reliable dataset, enabling accurate forecasting and informed strategic decision-making.

* 1. Data Preprocessing: Feature Engineering

After data cleaning, several new features were engineered from the cleaned sales and shipment datasets to enhance the predictive capability of the forecasting models and extract more actionable business insights. These derived variables allowed for deeper analysis of operational patterns and financial performance across services, clients, and time.

* **Profit Margin:** it was created by subtracting Gross Amount USD from Net Amount USD and dividing the result by Gross Amount USD. This ratio provided a normalized view of profitability per transaction and highlighted differences across service types.
* **Average Price per Unit:** it was calculated by dividing Net Amount USD by Quantity, offering insight into the pricing structure of various logistics services and enabling comparisons between product lines.

In the shipment dataset, the following features were created:

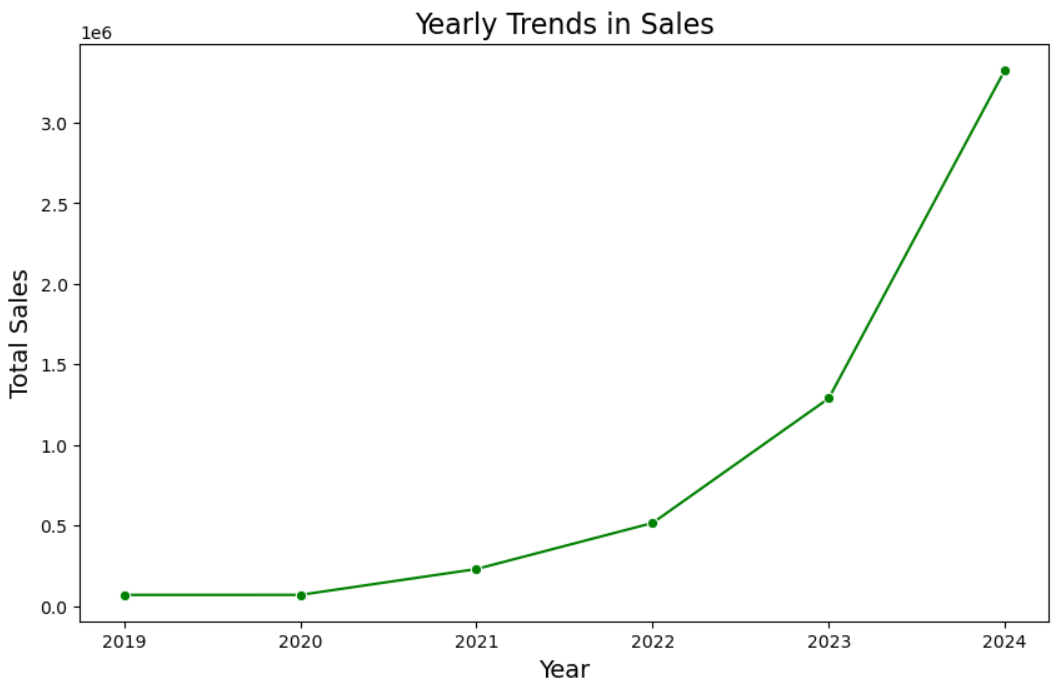
* **Route Field:** it was created by concatenating the From and To fields. This allowed for grouped analysis of high-traffic freight corridors and geographic performance patterns.
* **Weight per Package:** it was derived by dividing Gross Weight by Packages, this feature captured the logistical density of each shipment and was useful for understanding load efficiency across different clients and shipment types.
* **Season:** based on the extracted Month variable from the Date field, each record was assigned a seasonal label: Winter, Spring, Summer, or Fall. This classification enabled modeling of seasonal fluctuations in both sales and shipment activity.

These engineered features were incorporated into the forecasting and visualization pipeline and contributed significantly to improving trend detection, model performance, and interpretability of results. They also allowed for greater business alignment by contextualizing forecasts within operational and financial realities specific to GIFCO’s logistics services.

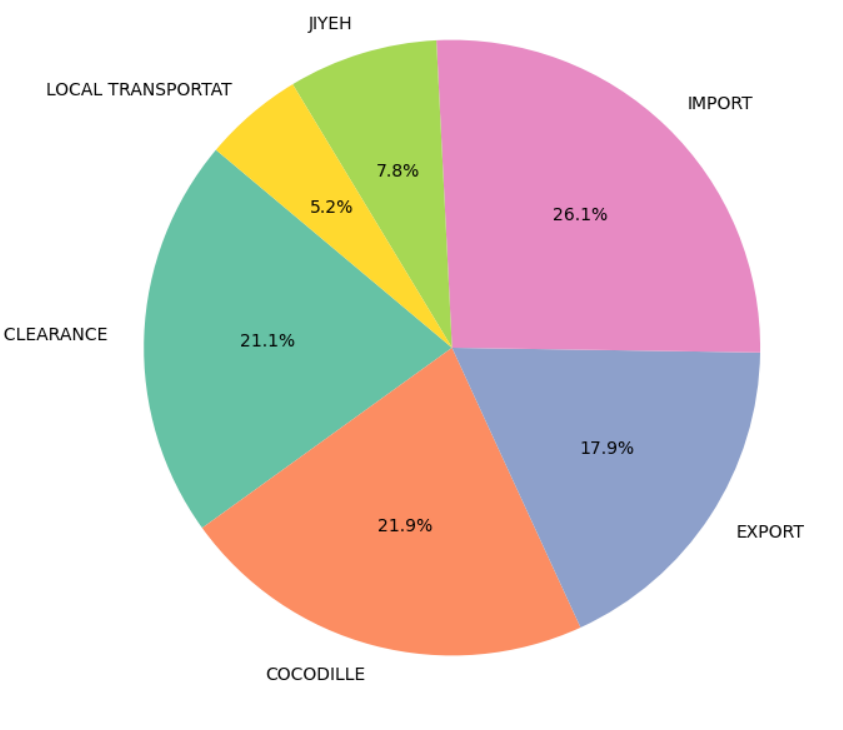
* 1. Exploratory Data Analysis (EDA)

Before moving into the modeling framework, an in-depth exploratory data analysis (EDA) phase was conducted to extract trends, detect anomalies, and generate insights that could guide feature engineering and model selection, using the matplotlib library in Python. The EDA process focused on uncovering sales dynamics, service-level behaviors, client segmentation, and seasonal patterns.

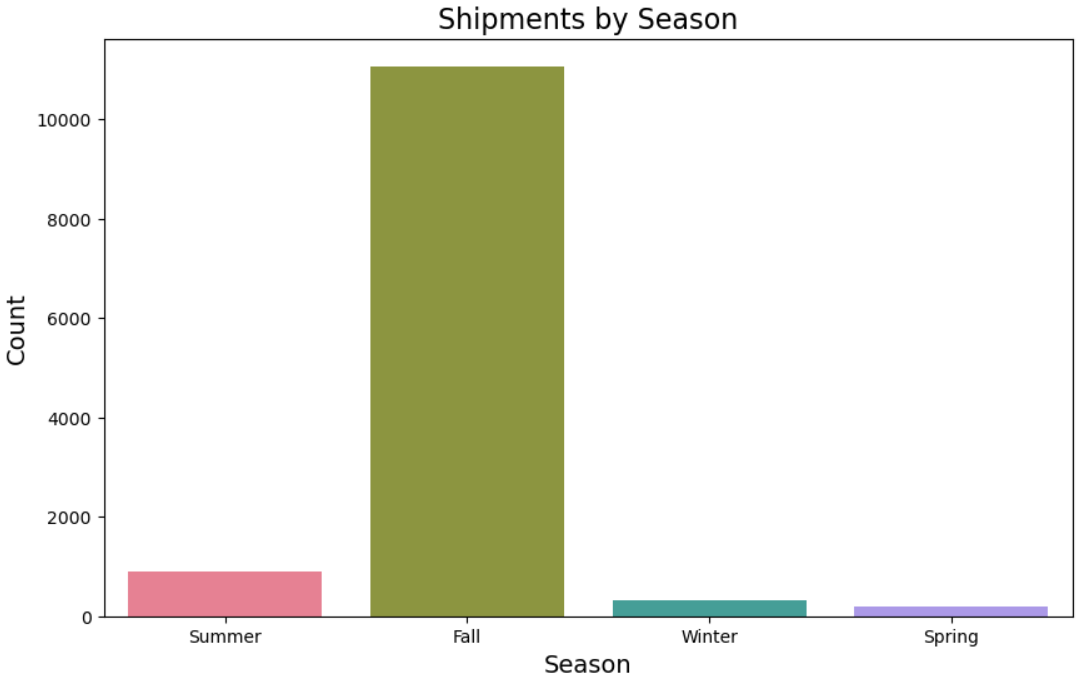
* **Sales Trends Over Time:** yearly aggregated revenue was plotted to observe the evolution of sales from 2019 to 2024. This revealed clear year-over-year growth, punctuated by seasonal peaks in late summer and fall periods.



* Service Family Performance: sales were grouped by service family (export, import, clearance…), exposing the services that they offer and is demanded from them the most over the years.



* Seasonality and Volume Patterns: distributions across seasons showed a repeatable pattern, with Fall being the most active in terms of shipment count. These patterns supported the assumption of strong seasonality in the time series.



* Freight and Weight Distribution: Using the shipment data, the distribution of “Gross Weight”, “Packages”, and “Weight per Package” was analyzed to understand logistics intensity and packaging efficiency.

A graph of different colored bars

AI-generated content may be incorrect.

Insights from EDA were used to inform the later dashboard creation, support the choice of seasonal models, and guide strategic interpretation of forecast results.

* 1. Tableau Dashboard

After EDA, a Tableau dashboard was created to provide GIFCO with an interactive overview of key business insights. The dashboard comprises nine distinct visualizations that offer a comprehensive view of the company’s sales performance, service-level trends, client contributions, and shipment patterns. These visuals allow for easy exploration of historical data, providing GIFCO’s management with actionable insights into profit margins, service performance, and geographic trends. This dashboard will later be integrated into the Streamlit web application, ensuring that GIFCO’s team can seamlessly interact with the data and forecasts in a user-friendly environment for real-time decision-making and strategic planning.

* 1. Modeling Framework

To address the dual goals of forecasting GIFCO’s future performance and supporting strategic planning, a comparative modeling framework was employed. Models were selected based on their ability to capture seasonality, linear and non-linear trends, and short-term fluctuations, while remaining interpretable for business users. Time series forecasting was performed using yearly aggregated “Net Amount USD”, covering the period from 2019 to 2024. A range of statistical and machine learning-based models were employed which allowed for automated model comparison, hyperparameter tuning, and metric-based evaluation:

* **ARIMA:** this model is a classical linear model that captures autocorrelation, differencing to remove trend, and moving average components. It was selected as a baseline model to assess the time series' stationarity and to model simpler linear patterns. While effective for non-seasonal trends, ARIMA was limited by its inability to account for recurring seasonal structures, which led to its use primarily as a benchmarking tool.
* **SARIMA:** this model extends ARIMA by incorporating seasonal differencing and seasonal autoregressive/moving average components. It was chosen due to the presence of clear seasonal patterns in GIFCO’s yearly and seasonal sales data. The model was useful in capturing both short-term dependencies and seasonal cycles, providing a strong statistical basis for comparison against other methods. However, when evaluating its RMSE, it showed that the model failed to capture the changing trends.
* **Holt-Winters Exponential Smoothing:** the Holt-Winters method was applied with additive seasonality to account for level, trend, and seasonal components of the time series. It was selected for its simplicity, interpretability, and effectiveness in modeling gradually evolving patterns. This model performed well for service lines with stable growth and regular seasonality but struggled with sharp fluctuations or irregular patterns.
* **Prophet:** Prophet which was developed by Meta, is designed to handle messy time series with missing data, outliers, and abrupt changepoints. It uses a decomposable model with trend, seasonality, and holiday effects. Prophet was selected because it is highly interpretable and accommodates real-world business scenarios such as sudden changes in activity, making it valuable for analyzing peak seasons and disruptions in sales and it showed good results.
* **Theta Model:** this model is known for its high accuracy in forecasting. It separates the time series into linear trend and curved components, then combines them to generate a robust and smoothed forecast. It was implemented using PyCaret and provided the best forecast performance. Its strength lies in smoothing noisy time series and capturing long-term trends, making it the most reliable option for GIFCO’s 6-year projection needs.

After model creation, the best performing model out of those was selected for optimization using the GridSearch method.

* 1. Model Evaluation and Selection

All models were trained and tested using a rolling-origin evaluation (walk-forward validation). The key metric used for comparison was the Root Mean Squared Error (RMSE). Visual diagnostics, including forecast vs. actual plots and residual analysis, were used to validate temporal consistency.

Three different levels of interpretation were used to the forecasts:

* **Service-Level Trends:** the forecasts were disaggregated by service family (e.g., sea exports, storage, clearance…) to reveal growth trends and identify underperforming categories.
* **Expansion Potential:** based on projected growth and geographic shipment insights, new market opportunities were identified for regional expansion and valuation support.
* **Seasonality and Sales Insights:** the forecasts helped uncover recurring seasonal peaks and off-peak periods across key services, allowing GIFCO to align resource allocation, staffing, and promotional efforts with high-demand cycles, leading to better sales management and risk assessment.

This multi-model approach ensured not only forecasting accuracy but also flexibility, allowing GIFCO to explore different future scenarios, compare service performance, and strengthen its strategic planning capabilities.

* 1. Streamlit Web Application

To ensure the insights and forecasts generated by the project were accessible and actionable for GIFCO's management team and their potential future investors, a user-friendly web application was developed using Streamlit which is an open-source Python framework for building data apps. The app served as a centralized interface that allowed stakeholders to browse findings and get an overview of the company and its future in the coming 6 years. The application is structured into three pages:

* **Home Dashboard**: displays overview of GIFCO.
* **Tableau Integration**: embeds the Tableau dashboard for deeper sales and shipment analysis.
* **Forecasting Interface**: allows users to view 6-year sales projections generated.
  1. Tools and Technologies

This modeling framework was developed entirely in Python using the Jupyter Notebook environment. The key libraries and tools that were used included:

* **pandas, numpy:** for data cleaning and time series structuring.
* **PyCaret:** for forecasting model selection, training, and evaluation.
* **statsmodels, Prophet:** model-specific enhancements and diagnostics.
* **matplotlib, seaborn:** time series plotting and residual visualization.
* **Streamlit:** for building and deploying interactive web application, using Python.

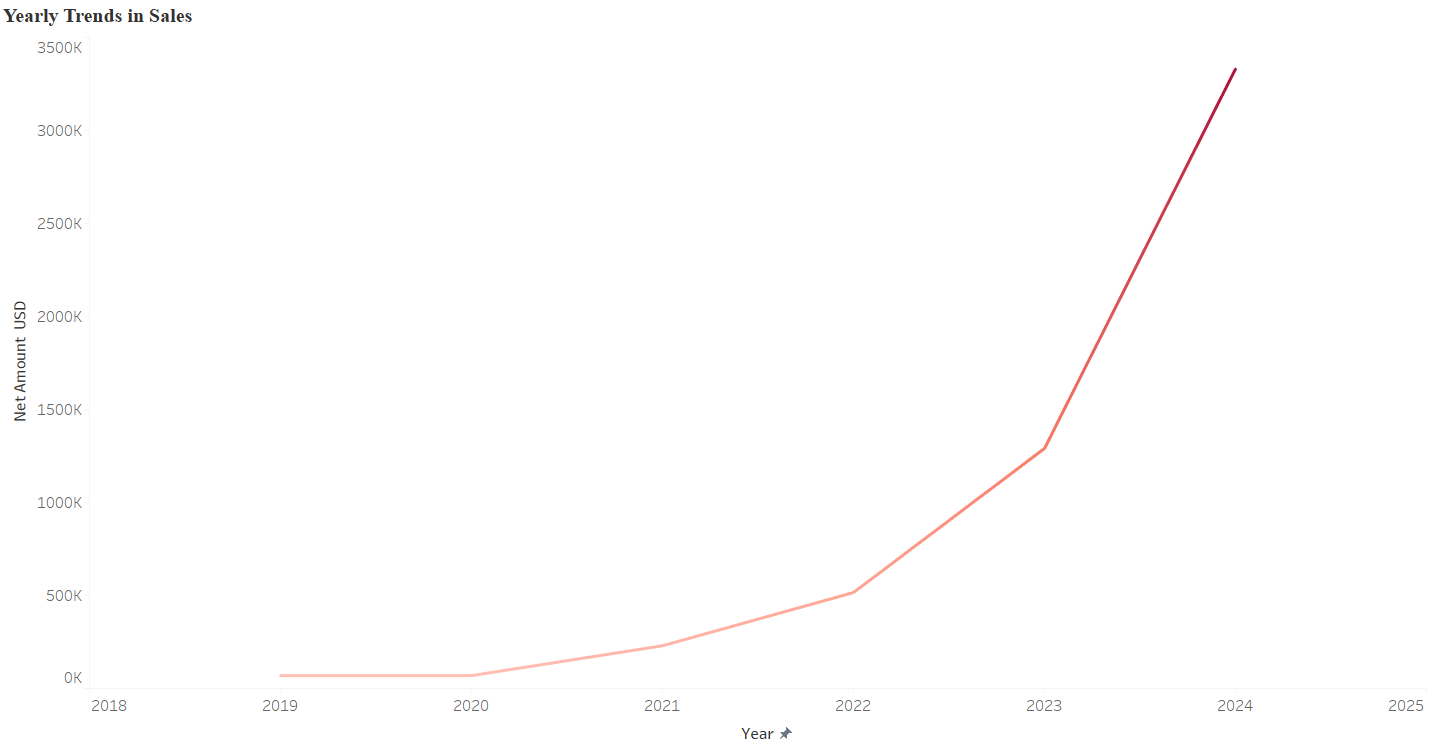
Results

In this section, we will explore the results derived from the three main levels of the methodology: the Tableau Dashboard, the Forecasting Models, and the Streamlit Web Application. These results provide actionable insights that support GIFCO’s decision-making process, offering a comprehensive view of their historical performance, future projections, and strategic opportunities.

* 1. Tableau Dashboard

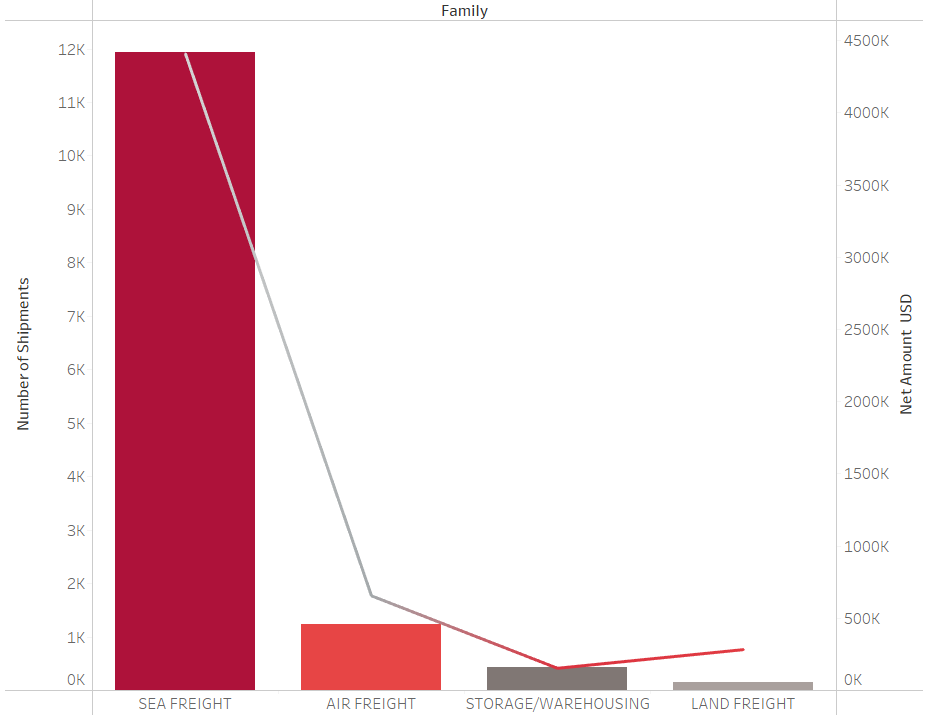
A Tableau dashboard was created with 9 visuals that give a clear overview of GIFCO’s past and present operations and business strategies (link [here](https://public.tableau.com/shared/KB6RTM2TD?:display_count=n&:origin=viz_share_link)). The visuals showed the following:

1. **Yearly Trends in Sales**

****

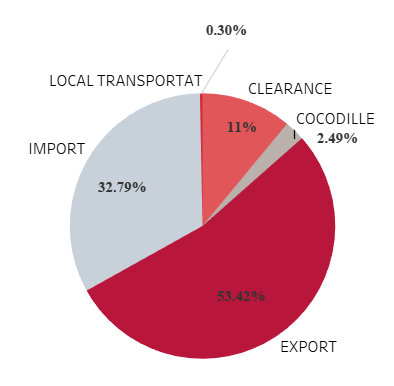
This visualization shows the yearly trends in sales, with the Net Amount in USD plotted across the years from 2019 to 2024. The graph highlights a significant increase in sales starting from 2021, with a sharp upward curve in 2023 and 2024. Prior to 2021, sales remained relatively flat, indicating that there were outside factors affecting the sales like the 2019 revolution, COVID-19 or other external factors. For GIFCO, this trend signifies the successful implementation of business changes or external factors that have driven revenue growth. The notable increase in 2023 and 2024 could indicate a recovery or expansion phase, potentially related to market penetration, improved services, or new client acquisitions. This trend will be key for future projections and can guide investment planning and resource allocation to sustain this growth trajectory.

1. **Sales and Number of Shipments by Freight Type**



This visualization illustrates the distribution of shipments across different service families which are sea freight, land freight, air freight and storage/warehousing, and their associated net revenue in USD. Sea freight emerged as the dominant category, with a significantly higher number of shipments compared to other service types (around 12,000 shipments). This is accompanied with a high net revenue of around $4.5 million over 5 years. We can also notice that air freight, despite having fewer shipments, generates a high net revenue per shipment, indicating its higher value per transaction. Storage/Warehousing and land freight show both lower shipment volumes and revenue, suggesting either underutilizing or less profitable compared to sea and air freight.

1. **Profit Margin Distribution by Sub-Family**

****

This profit margin distribution by sub-family pie chart provides a breakdown of profit margins across different sub-families within GIFCO’s service offerings. Export contributes the largest share, accounting for **53.42%** of the total profit margin, indicating that this service family is the most profitable. The second-largest contributor is Import, which makes up **32.79%**, followed by Local Transport at 11%. This chart emphasizes the importance of the Export service, both in terms of volume and profitability.

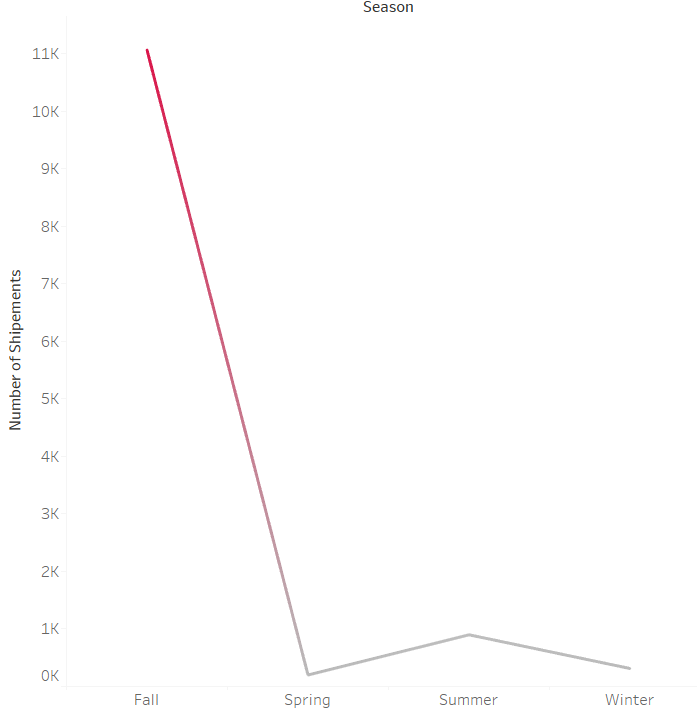
1. **Top 10 Clients by Shipment Volumes**

A screenshot of a graph

AI-generated content may be incorrect.

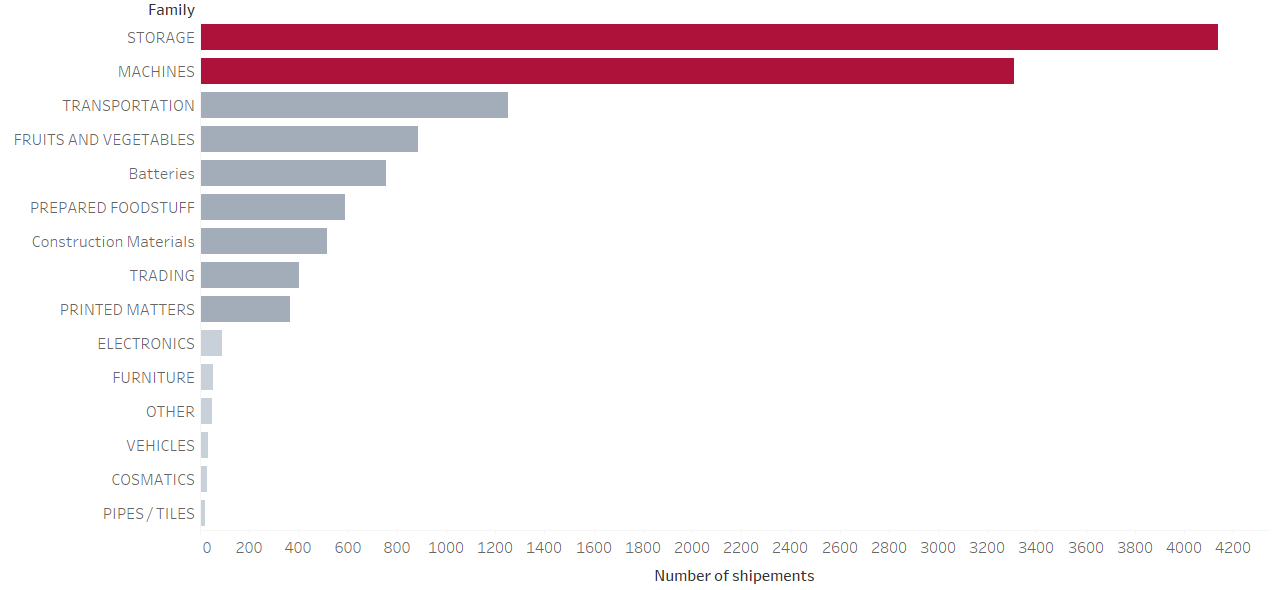
This bar chart of top clients by shipment volume highlights the disparity in shipment activity among GIFCO’s top clients. **Client\_2** stands out significantly, accounting for over 6,300 shipments, far surpassing the other clients. In contrast, the remaining clients in the top 10 contribute far fewer shipments, with Client\_7, Client\_6, and Client\_5 having volumes around 1,000 shipments or less. For GIFCO, this suggests that Client\_2 is a key contributor to their overall shipment volume, making them a high-value client for the business. The large volume of shipments from this client could mean the potential for further strategic development or special services. On the other hand, the other clients with much lower volumes might offer opportunities for growth or require closer attention to identify ways to increase shipment frequency and overall engagement. The data also implies a heavy reliance on Client\_2, so understanding their needs and strengthening this relationship could be critical for GIFCO's future strategy.

1. **Shipment Volumes by Season**



This seasonal shipment volume chart illustrates a pronounced peak in shipments during the Fall, with over 10,000 shipments processed during this period. However, the rest of the seasons show significantly lower shipment volumes, with the Winter period seeing the lowest. For GIFCO, this seasonal pattern suggests a highly concentrated demand in Fall, which could be linked to various factors like end-of-year logistics, increased sales during holiday periods, or seasonal business surges. The other seasons may present an opportunity for GIFCO to optimize operations or drive more business by addressing potential inefficiencies, offering promotions, or exploring ways to increase shipment volumes during typically slower periods.

1. **Shipments by Types of Goods**



This bar chart displaying shipment volumes by service family reveals the dominant categories within GIFCO’s operations. Storage (which refers to shipments that are sent to GIFCO’s warehouses) leads with the highest number of shipments, outpacing other service families. Following Storage, Machines and Transportation are the next largest categories, though they show much lower shipment volumes.

For GIFCO, this distribution highlights that Storage is a critical component of the business, with considerable operational activity. GIFCO may want to evaluate whether these service families can be further optimized or marketed to increase volume and revenue. The low shipment counts in categories like Cosmetics, Vehicles, and Pipes/Tiles suggest opportunities to either explore new markets or refine their offerings in these segments.

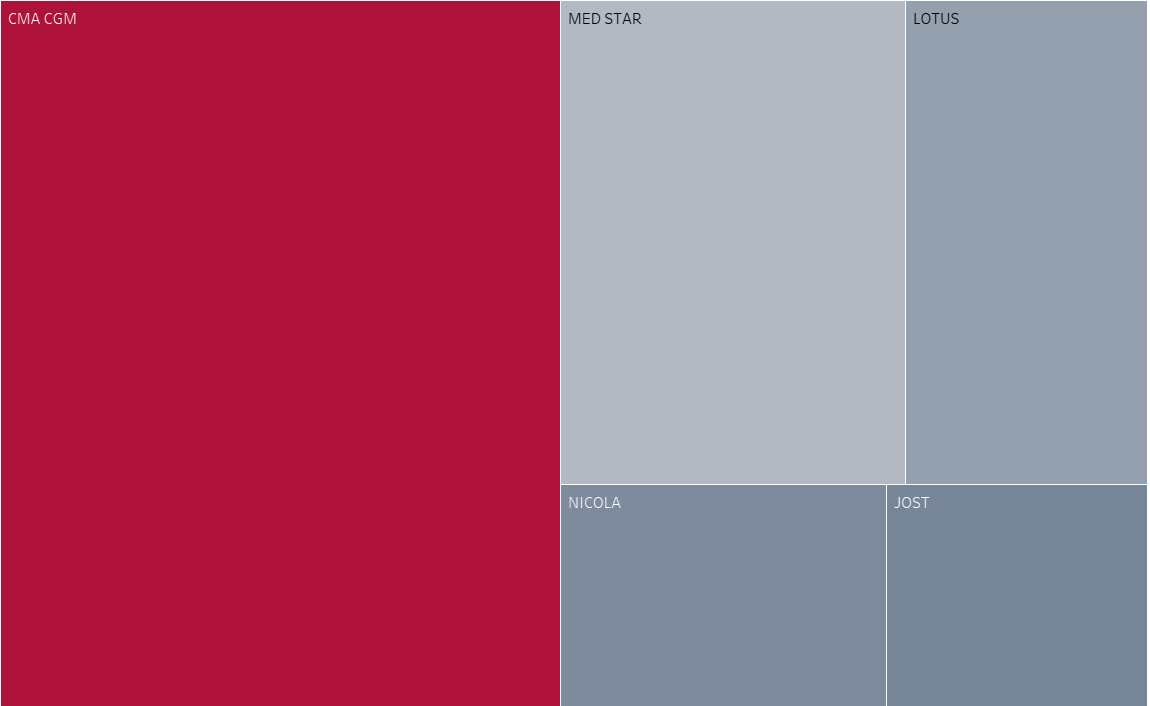
1. **Relationship between Quantity Sold and Net Amount USD Over the Years**

A graph showing the growth of a graph

AI-generated content may be incorrect.

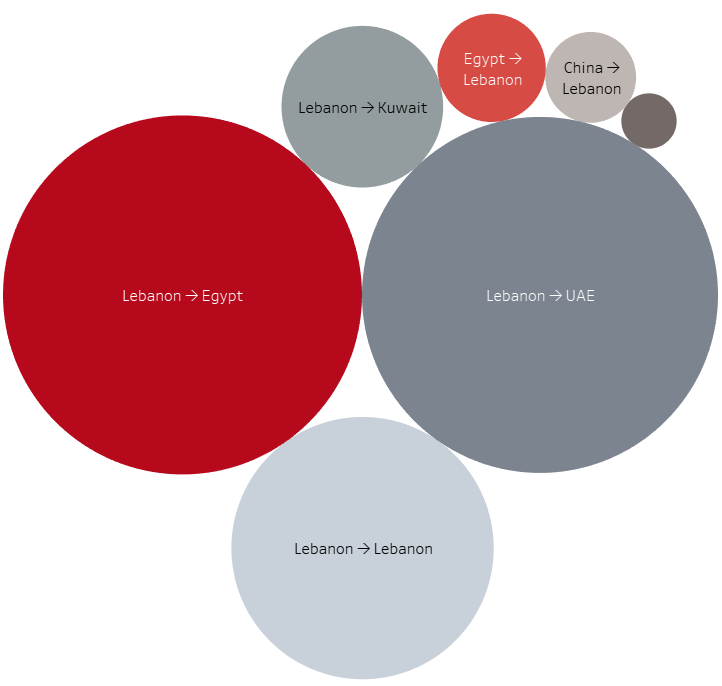
This yearly sales and quantity trend chart shows two distinct patterns in GIFCO’s sales performance. From 2018 to 2022, both quantity and net amount remained relatively flat, suggesting a period of stability and low growth. However, a significant shift occurred in 2023, where quantity increased modestly, but net amount rose sharply, peaking in 2024. This pattern suggests that in 2023, GIFCO implemented a premium pricing strategy and started offering more valuable services, leading to higher revenue despite relatively fewer shipments. This could be a strategic move to target high-value clients or markets that are less price-sensitive, offering higher-margin services or goods.

1. **Top Vessels Employed**



This vessel distribution chart shows the relative shipment volume handled by GIFCO’s top vessels. CMA CGM is the dominant vessel in terms of shipment volume, with a significantly larger proportion compared to the other vessels, such as Med Star, Lotus, Nicola, and Jost, which have much smaller shares. For GIFCO, this suggests that CMA CGM plays a central role in their operations, likely handling the majority of high-volume shipments. This reliance on a single vessel could represent both an advantage (in terms of specialization and efficiency) and a risk (in case of operational disruptions). GIFCO may want to explore optimizing their operations around CMA CGM to maintain or increase efficiency while also mitigating risks associated with dependency on one major asset.

1. **Top Routes of Shipment**

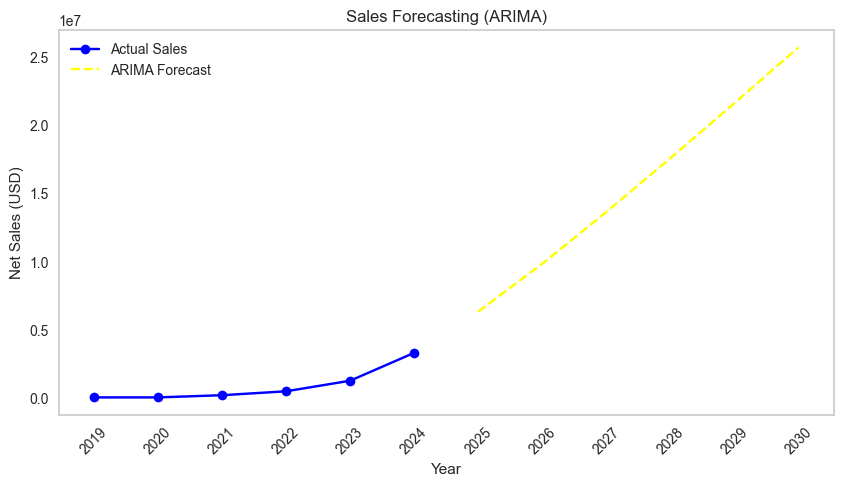


This bubble chart illustrates the volume of shipments between various countries and Lebanon. The largest bubble represents shipments from Lebanon to Egypt, indicating that this route is the most significant in terms of shipment volume. Notably, the Egypt to Lebanon route also stands out, suggesting a strong trade flow and presenting an expansion opportunity for GIFCO in Egypt. Expanding operations or establishing a local office in Egypt could further strengthen GIFCO's presence in the region. Other prominent routes include Lebanon to UAE and Lebanon to Kuwait, with the Lebanon to Lebanon internal shipments accounting for a substantial portion. The Lebanon to UAE route is particularly important, representing a significant market for GIFCO. Establishing operations in the UAE could also help leverage this high-volume route and tap into the growing demand within the Gulf region. Overall, the chart highlights the strategic importance of these routes. GIFCO can consider expanding into both Egypt and the UAE, which are key markets for potential growth and operational efficiency.

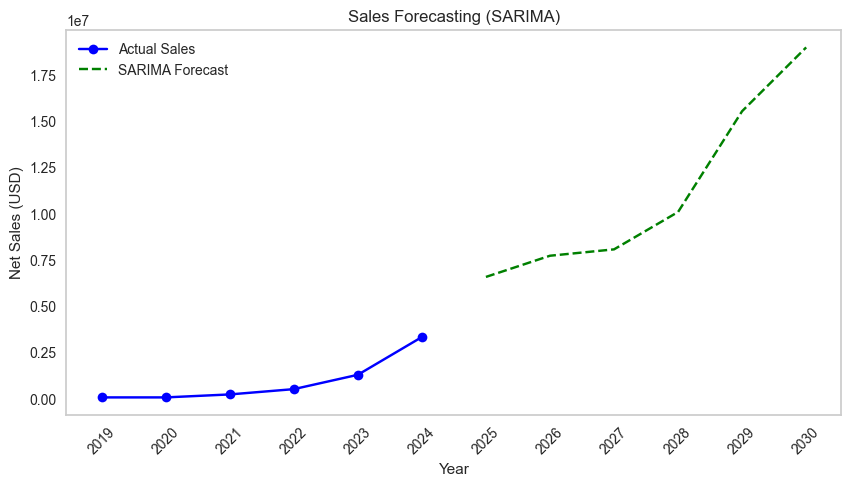
* 1. Forecasting Models Results

Five models were created for the 6 year forecasting of GIFCO’s sales. These are the results they have yielded with their RMSEs:

* **ARIMA Forecast:** predicted steady growth in sales over the next six years, with the forecast for 2025 at $6,340,378, rising to $25,682,490 by 2030. The model shows a consistent increase in forecasted sales, indicating an optimistic outlook. However, ARIMA struggled to capture seasonal variations effectively, leading to higher RMSE = $2,650,975.42. This made it less reliable for capturing cyclical patterns, which were important for accurate forecasting.



* **SARIMA Forecast:** forecasted similar growth trends to ARIMA, starting with $6,575,188 in 2025, reaching $18,970,860 by 2030. While it did account for seasonal variations, it was still less flexible in modeling irregular trends but still performed well with an RMSE = $445,764.96, which is the second best performing model.



* **Prophet Forecast:** predicted more moderate growth, starting at $3,643,872 in 2025, growing to $6,765,361 by 2030. While Prophet is great at handling trend changes, it is less effective in dealing with highly volatile data and outliers. The RMSE for this model was $773,059.15, indicating that it wasn’t as accurate in capturing sharp growth patterns or inflection points in sales as other models.

A graph with a line and dotted line

AI-generated content may be incorrect.

* **Holt-Winters Exponential Smoothing Forecast:** forecasted a sharp upward trend, starting at $5,351,892 in 2025 and peaking at $15,490,088 by 2030. Although it captured both seasonal effects and long-term trends well, it struggled with extreme fluctuations, often overshooting sales projections in periods of sudden market changes. The RMSE = $1,639,931.79 suggests it is still not ideal for highly fluctuating data.

**A graph with a dotted line and a dotted line

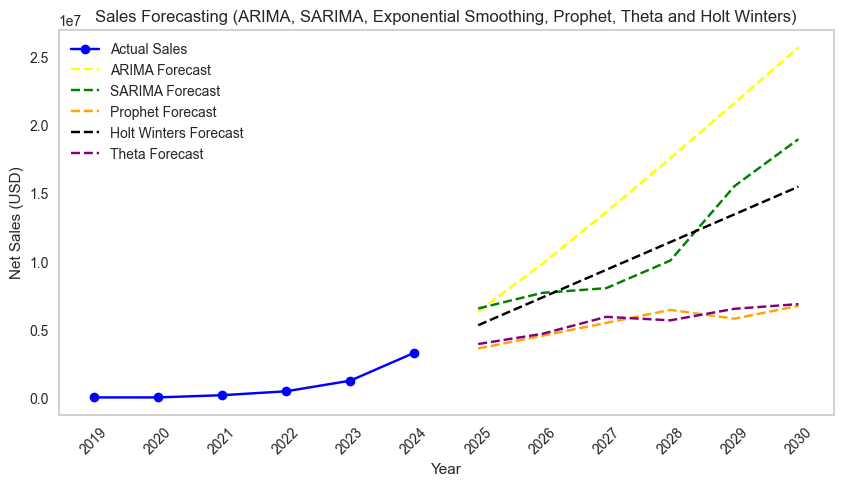
AI-generated content may be incorrect.**

* **Theta Forecast:** the best-performing model, predicted conservative but accurate sales projections, with $3,975,188 in 2025, reaching $6,897,086 by 2030. The Theta model effectively decomposes the time series into linear and curved components, smoothing out fluctuations. It outperformed other models with a RMSE = $189,453, showing its robustness in dealing with complex data and long-horizon forecasting.

**A graph with a line and a dotted line

AI-generated content may be incorrect.**

**All Model Comparison**

****

**Theta Model Optimization**

In the optimization step, the grid search technique was used to fine-tune the hyperparameters of the Theta model to achieve the best forecasting accuracy. Specifically, this approach aimed to identify the optimal combination of three key hyperparameters: seasonality, trend, and seasonal periods. The grid search was conducted using the following parameters:

* **Seasonality:** it defines the type of seasonal effect to be applied to the model. Two options were considered: “add” which assumes a linear relationship in the seasonal effects, and “mul” which assumes multiplicative seasonal effects.
* **Trend:** it controls the type of trend applied to the data. Two types were tested: “add” which represents a linear trend, adding a constant amount over time, and “mul” which represents a multiplicative trend, meaning the trend grows or shrinks exponentially.
* **Seasonal Periods:** it defines the length of the seasonal cycle, tested on three values: 4, 6, 12. These seasonal periods represent different cycles of data (monthly, quarterly, or annual).

For each combination of these parameters, the Theta model was trained and forecasts were generated. The model's performance was then evaluated using RMSE. The RMSE was calculated using the forecasted values (yhat) and the actual values from the dataset, which allowed for a direct comparison of the different configurations.

The optimization process involved running the grid search over all possible combinations of the parameters in the grid. For each combination, the model was fitted to the data, forecasts were generated, and the RMSE was computed to evaluate how well the model performed. The best parameters were selected based on the lowest RMSE, ensuring that the model could provide the most accurate long-term forecasts for GIFCO.

Ultimately, the grid search determined that the best hyperparameters for the Theta model were: Seasonality: “add”, Trend: “add”, and Seasonal Periods: 12, indicating the default Theta model. This combination resulted in the lowest RMSE of 189,453.56, which indicates the best-fit model for GIFCO’s data.

A graph of sales forecasting

AI-generated content may be incorrect.

* 1. Streamlit Web Application

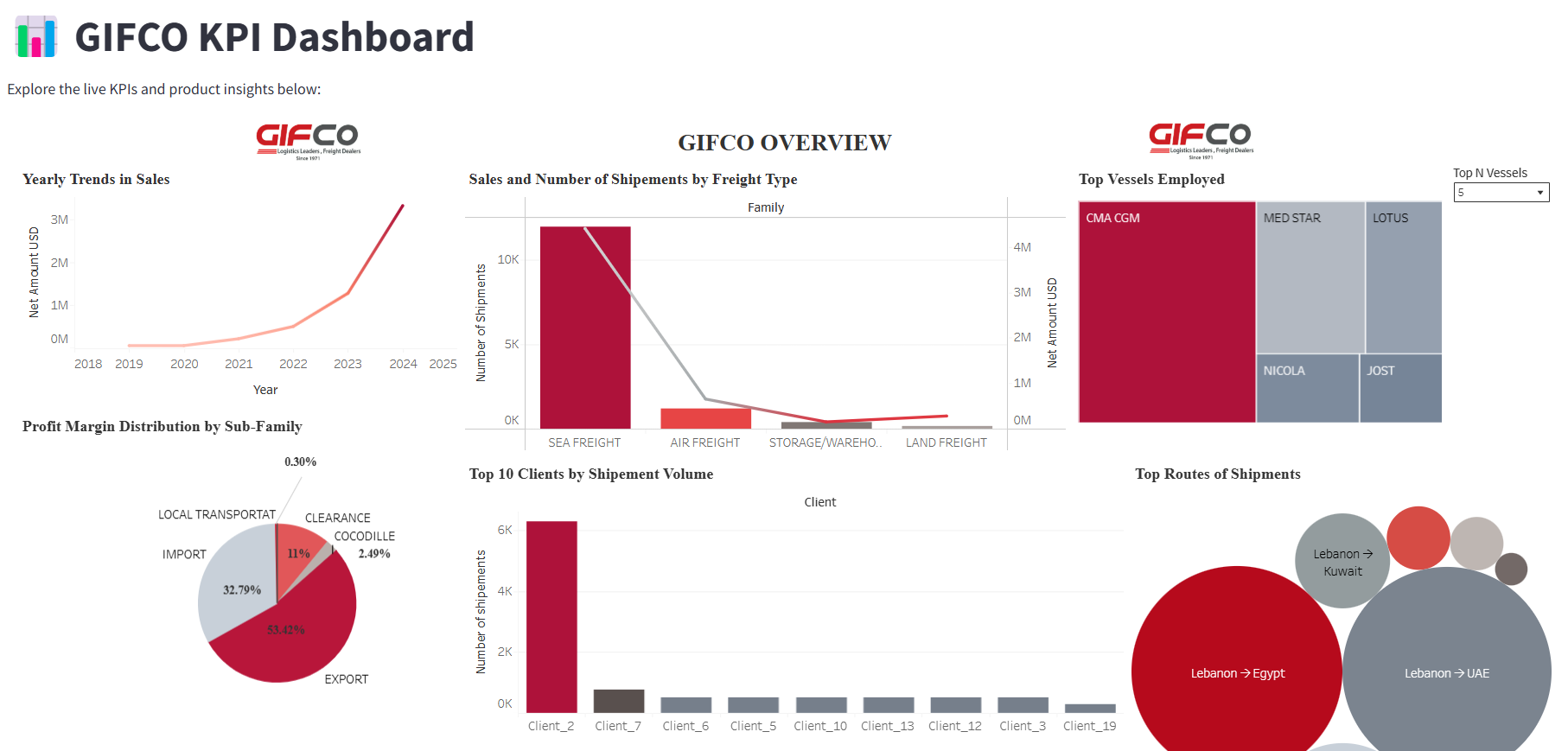
The final app deliverable was structured into three main pages that can be selected in the navigation bar to browse, each serving a distinct purpose:

* **Page 1 Home Page:** in this page, the overview of the company and the web app was added.

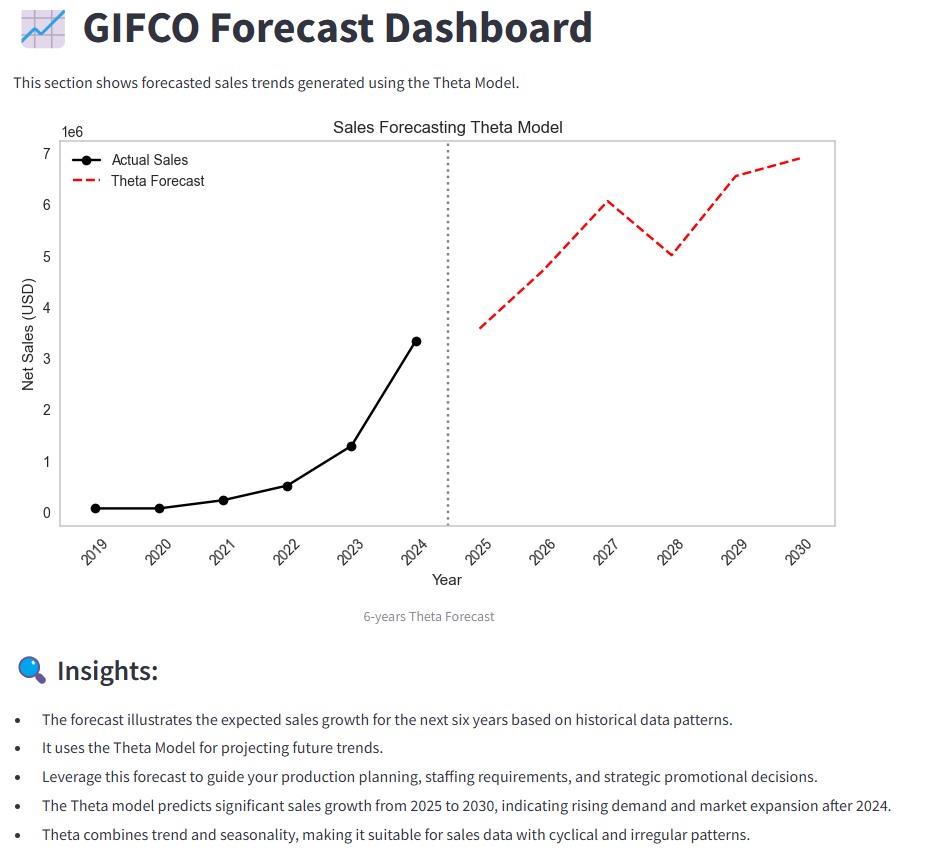
A red truck with white wheels

AI-generated content may be incorrect.

* **Page 2 Tableau Dashboard Integration:** in this page, the Tableau dashboard was embedded for visualizing shipment data, route analysis, and freight breakdowns which was interactive using filters for each visual.



* **Page 3 Forecasting Results and Insights:** in this page, the 6-year sales forecasts generated by the Theta model were displayed, in addition to insights about the forecasting and strategy.



This app was designed for browser access without requiring logins or local software installation, ensuring that GIFCO’s leadership could engage with the insights anytime and from any location. The Streamlit app served as the final delivery mechanism of the project, translating technical outputs into strategic tools for decision-making. It allowed GIFCO to explore the past, understand the present, and plan for the future in an intuitive and data-driven environment.

Discussion

**Optimization of Sales: Suggested Informed Business Strategy**

Based on the data analysis, GIFCO has demonstrated an impressive ability to grow revenue, but the growth patterns suggest that GIFCO is not solely focused on a premium pricing strategy. Instead, the company has effectively utilized a combination of strategic pricing and service diversification to cater to different segments of the market. Based on these insights, a new business strategy can be offered that encompasses the following ideas:

1. **Balanced Strategy: Growth through Quantity and Premium Services**

The yearly trends in sales show changing strategies in GIFCO’s sales. In 2019-2020, because of external factors like local instability and COVID-19, shipping saw a stable trend that is not increasing. In 2023, a significant growth was noted where the quantity of shipments dropped compared to the previous years, while sales increased significantly. This suggests that GIFCO adopted in 2023 a premium strategy where, after a 2-year period of stable and no-growth, they focused on increasing sales through targeting high-value clients or premium logistics services such as air freight or expedited services. Then, the year that followed they were able to get back on track with their normal shipping volumes and expanded market reach through competitive pricing strategies and getting more clients, especially with sea freight and exports, while also targeting premium services in parallel.

The strategy proposed to GIFCO would include continuing to balance between premium pricing for high-value services and competitive pricing for higher volume services like sea freight. Also, they should focus on their most prominent services of import and export since Lebanon is a country that relies mostly on exporting and importing goods. Therefore, GIFCO can maintain affordable offerings for certain services (like sea freight) while leveraging premium pricing for other specialized services, particularly in air freight and storage.

1. **Exploiting Export Opportunities**

As mentioned, the profit margin distribution shows that the export sub-family contributes to a substantial portion of GIFCO's profit. This aligns with the company's core business of shipping and logistics.

The strategy GIFCO should adopt is to continue to prioritize exports, particularly to high-demand markets such as Egypt, the UAE, and Kuwait. The data analysis suggests these countries are key trading partners, and further strengthening operations in these areas could lead to long-term profitability. GIFCO could also expand its export services by offering value-added services like customs clearance or special handling for certain high-value goods.

1. **Strengthening the Existing Client Base**

The data shows that a significant portion of GIFCO’s profit comes from existing clients and sometimes rely on specific customers most. The chart also highlights the top clients by shipment volume, which shows that some key clients account for the majority of shipments.

A recommendation for GIFCO would be to leverage its existing client base by offering more tailored services. GIFCO could engage in account management strategies, ensuring that the needs of high-value clients are met and even exceeded, fostering long-term relationships. Additionally, GIFCO can create loyalty programs or special offers for top clients, particularly in the storage/warehousing segment, to boost repeat business. In addition, they should also consider expanding slowly their customer base, avoiding the risk of overly relying on specific customers for shipments.

Therefore, by focusing on balancing pricing strategies, capitalizing on the existing client base, and expanding high-margin service offerings like storage/warehousing and exports, GIFCO can position itself as a leading logistics provider while ensuring long-term profitability and growth.

**Investment Gathering for Expansion and Strategic Market Entry**

As part of GIFCO’s long-term growth strategy, one of the key objectives is to expand its operations into new markets. This expansion will require significant capital investment to establish a strong local presence, optimize infrastructure, and improve service offerings. Based on the data analysis, several key insights have emerged that can guide the company’s expansion strategy and inform its investment gathering approach:

1. **Identifying High-Growth Markets for Expansion**

* The shipment data clearly shows that Egypt and the UAE are significant trade partners, with a high volume of shipments from Lebanon to these regions. The largest shipment volume to Egypt and UAE suggests that these regions offer considerable growth potential.
* GIFCO should consider expanding into Egypt and UAE, as these markets show **high demand** for logistics services and already have strong trading relationships with Lebanon. Expanding operations into these regions can help optimize costs, increase market share, and take advantage of the established trade flows.
* **Egypt:** as the Lebanon to Egypt route has shown a dominant volume, GIFCO should explore investment opportunities in Egypt, such as opening new offices or distribution hubs in key cities like Cairo or Alexandria. Egypt’s strategic location as a gateway to Africa makes it an attractive market for logistics firms, especially since they are established also in the African market with their office in Juba.
* **UAE:** The Lebanon to UAE route has shown consistent activity. Expanding into the UAE would help GIFCO tap into the Gulf Cooperation Council (GCC) market, where demand for premium logistics services is high. Establishing a presence in Dubai or Abu Dhabi could significantly strengthen GIFCO's position in the region.

1. **Expansion into Local Markets with High Domestic Demand**

* The Lebanon to Lebanon internal shipments represent a significant portion of GIFCO's operations. This suggests that while international expansion is critical, local market growth should not be overlooked.
* Local market efficiency can be improved by investing in technology upgrades, such as advanced warehouse management systems (WMS), and exploring domestic expansion opportunities in other regions of Lebanon. By improving operational efficiency domestically, GIFCO can increase profit margins while also supporting international operations.

1. **Attracting Investors Based on Forecasted Growth**

* The Theta model’s forecast also provides additional informative insights for investors as to how the company would perform in the future and if the company is worth taking a risk on. This forecast provides a reliable outlook that can help attract investors.
* So, GIFCO can present this forecasted growth to potential investors as evidence of its scalable operations and long-term viability. By offering a clear investment roadmap based on real data and forecasts, GIFCO can secure funds for market entry and expansion initiatives.

Therefore, GIFCO’s expansion strategy should focus on both international growth in markets like Egypt and UAE and domestic optimization. By securing investments in key expansion areas, GIFCO can solidify its position in the logistics industry and extend its market reach in the MENA region. Leveraging its existing relationships, tapping into underexploited markets, and expanding service offerings will enable GIFCO to drive future growth while maximizing the ROI for investors.

The insights provided by this project have several implications for GIFCO’s future operations. First, the company is now equipped with a data-driven framework and online web interface that can be shared with the stakeholders to assess business performance, forecast future trends, and make strategic decisions in a competitive landscape. This project has demonstrated how leveraging historical data, alongside forecasting models, can offer real-time insights that inform pricing decisions, investment gathering, and market entry.

The impact of the data model forecasts and exploratory analysis is profound, particularly in helping GIFCO navigate uncertainties like market volatility, economic instability, and demand shifts. The Theta model, as the most accurate forecasting tool, has provided a long-term outlook that will guide the company’s strategic decisions, particularly in the context of investment gathering for expansion into regions like Egypt and UAE. So, by implementing these data-driven insights, GIFCO can achieve greater operational efficiency, ensuring that its strategies are based on real data rather than just intuition, thus reducing risk and increasing profitability.

While the project successfully met the goal of providing a comprehensive forecasting framework for GIFCO, certain areas could be refined. One area of improvement is the granularity of the client segmentation analysis, which could have been more detailed. While high-value clients were identified, further disaggregation could have offered a more targeted approach to account management. Moreover, the economic conditions of the targeted expansion markets (i.e., Egypt and UAE) were not explored in-depth in this analysis, which could have provided a more nuanced understanding of external factors influencing market entry success. Despite these limitations, the overall project outcome provides a reliable forecast and a clear path forward for GIFCO’s market expansion and business strategy.

The primary objective of this project was to assist GIFCO in expanding its operations, both through strategic forecasting and by identifying potential market entry points. The project successfully achieved this by offering forecast models for sales projections and expansion recommendations based on shipment data and profit margins. Additionally, by identifying Egypt and UAE as key targets for expansion, the project provided strategic recommendations for investment gathering to support GIFCO’s future growth.

Conclusion

This capstone project was developed to address the growing need for data-driven decision-making at GIFCO in an increasingly competitive logistics and freight forwarding market. For years, GIFCO's reliance on traditional methods and fragmented data limited their ability to leverage the full potential of their operational insights. The company’s pain point stemmed from underutilizing its data, relying on intuition-based decisions that failed to capture the complexity of their operations and market dynamics. This project aimed to equip GIFCO with a predictive forecasting framework capable of supporting both sales growth and market expansion. To thrive in this rapidly changing environment, GIFCO needed to shift from relying on historical, fragmented data to adopting a digital mindset, embracing data as a strategic asset. The project involved the development and implementation of a comprehensive forecasting model built on time series analysis and advanced machine learning techniques. By integrating sales data, shipment volumes, and key performance metrics, the methodology provided valuable insights into service-level performance, profit margins, and client behaviors. Multiple models were tested, including ARIMA, SARIMA, Prophet, Holt-Winters, and Theta, with the Theta model emerging as the most accurate and reliable tool for predicting sales and guiding long-term strategic planning. This robust framework not only provided forecasts but also offered a deeper understanding of operational dynamics, revealing areas of strength and opportunities for improvement. Beyond purely technical forecasting, the project also emphasized strategic market expansion and investment gathering as key themes. Insights from the forecasts suggested that GIFCO should prioritize growth in high-demand regions like Egypt and the UAE, where both sales volume and potential market opportunities are substantial. The findings also indicated that storage and warehousing services, as well as export and import, should be a focal point for expansion, as these services are poised to capitalize on both domestic and international demand. Incorporating these insights, GIFCO can now move forward with a clearer roadmap for expansion into strategic markets and optimize its pricing strategy across services. The data-driven approach empowers GIFCO to make more informed decisions, aligning its operational goals with projected market trends and client needs.

Looking ahead, future developments could involve integrating real-time data streams from shipments, expanding geospatial analytics, and incorporating external factors like economic conditions or global trade fluctuations. These enhancements would enable GIFCO to further refine its forecasts and improve adaptability in an ever-evolving logistics landscape. Moreover, the continued adoption of data-driven decision-making will foster a culture of agility and resilience within the company, allowing GIFCO to respond proactively to market shifts. The long-term impact of this project will be the transition from reactive management to forward-thinking, strategic planning, positioning GIFCO as a leader in the logistics industry.

References

Accenture. (2021). *Supply Chain of the Future: Resilient, Agile, and Sustainable*. [https://www.accenture.com](https://www.accenture.com/)

C.H. Robinson. (2022). *Navisphere: Real-time data transforming global supply chains*, <https://www.chrobinson.com/en-us/technology/navisphere/>

Choi, T. M., Wallace, S. W., & Wang, Y. (2020), *Big Data Analytics in Operations Management*. Production and Operations Management, <https://onlinelibrary.wiley.com/doi/abs/10.1111/poms.12838?msockid=024a406db70d61a01c4e538db61b604b>

Christopher, M., & Peck, H. (2012). *Marketing Logistics*, Routledge, <https://www.taylorfrancis.com/books/mono/10.4324/9780080496429/marketing-logistics-martin-christopher-helen-peck>

Ghosh, S., & Shah, D. (2020). *Applied Time Series Analysis and Forecasting with Python*. Apress, <https://link.springer.com/book/10.1007/978-3-031-13584-2>

Ghobakhloo, M. (2018). The role of digitalization in business model innovation: A review of the literature, *Technological Forecasting and Social Change*, <https://www.researchgate.net/publication/369591819_The_role_of_digitalization_in_business_and_management_a_systematic_literature_review>

IDC Worldwide Big Data and Analytics Spending Guide (2022) –  
*Worldwide spending on big data and business analytics*, <https://www.idc.com/getdoc.jsp?containerId=prUS49451322>

Maersk. (2023). *Annual Report 2022: Integrated logistics and inland expansion strategy*, <https://www.maersk.com/news/articles/2023/01/24/annual-report-logistics-expansion-strategy>

McKinsey & Company. (2021). *The rise of analytics in logistics*, <https://www.mckinsey.com/industries/travel-logistics-and-infrastructure/our-insights/the-rise-of-analytics-in-logistics>

McKinsey & Company (2022) - *The data-driven enterprise of 2025*, <https://www.mckinsey.com/business-functions/mckinsey-digital/our-insights/the-data-driven-enterprise-of-2025>

OECD. (2022). *SME Policy Index: The Mediterranean Middle East and North Africa 2022*. <https://www.oecd.org/publications/sme-policy-index-mena-2022-0e92d39a-en.htm>

Prophet Forecasting Model. (2024). <https://facebook.github.io/prophet/>

PyCaret Documentation. (2024). <https://pycaret.gitbook.io/docs/>

Taylor, S. J., & Letham, B. (2018). *Forecasting at Scale*. The American Statistician, <https://doi.org/10.1080/00031305.2017.1380080>

Waller, M. A., & Fawcett, S. E. (2013). Data science, predictive analytics, and big data: a revolution that will transform supply chain design and management. *Journal of Business Logistics*, <https://papers.ssrn.com/sol3/papers.cfm?abstract_id=2279482>

World Bank. (2023). *Doing Business in Lebanon: Enhancing SME Competitiveness*. <https://www.worldbank.org/en/country/lebanon/publication/doing-business>