Fittlyf Sales Analysis

Introduction

- **Objective**

The primary objective of this analysis is to uncover valuable insights from one year of customer transaction data. By examining this data, we aim to derive data-backed strategies that address key questions posed by various departments within the organization, including Marketing, Operations, Retail Products, and Customer Relationship Management. The ultimate goal is to leverage these insights to improve customer segmentation, forecast sales, enhance product placement, and reduce customer churn.

(525461, 8)

Data Overview

The dataset comprises a comprehensive record of transactions over a ONE-year period, encapsulating 525,461 rows and 8 columns. Each row represents a unique transaction, with the following key features:

- Invoice: A nominal, 6-digit integral number uniquely assigned to each transaction. If prefixed with 'C', it indicates a cancellation.
- StockCode: A nominal, 5-digit integral code uniquely assigned to each product.
- Description: A nominal feature describing the product name.
- Quantity: A numeric feature indicating the quantity of each product per transaction.
- InvoiceDate: A numeric feature capturing the date and time of the transaction.
- Price: A numeric feature denoting the price per unit in sterling (£).
- Customer ID: A nominal, 5-digit integral number uniquely assigned to each customer.
- Country: A nominal feature specifying the country where the customer resides.

Marketing: Customer Clustering and RFM Analysis

- **Model Explanation and Justification**

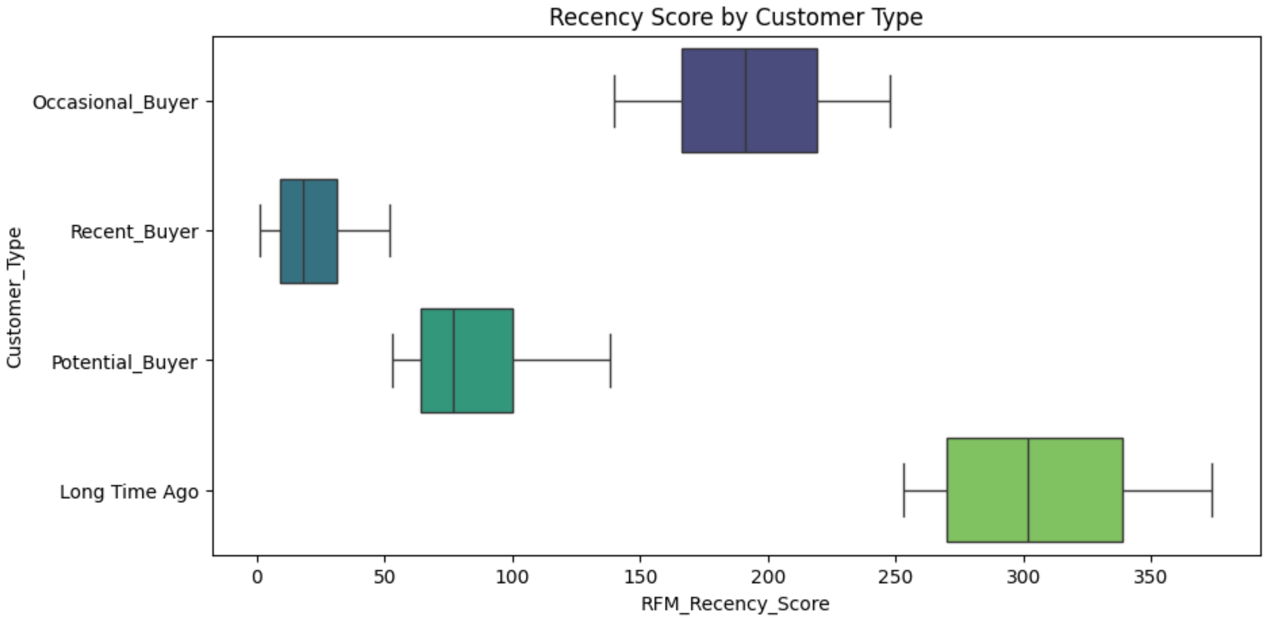
For Customer Segmentation, we Used RFM (Recency, Frequency, Monetary) techniques to segment our Customers. For further Segmentation we used the Clustering Model such as K Means Clustering ml model.

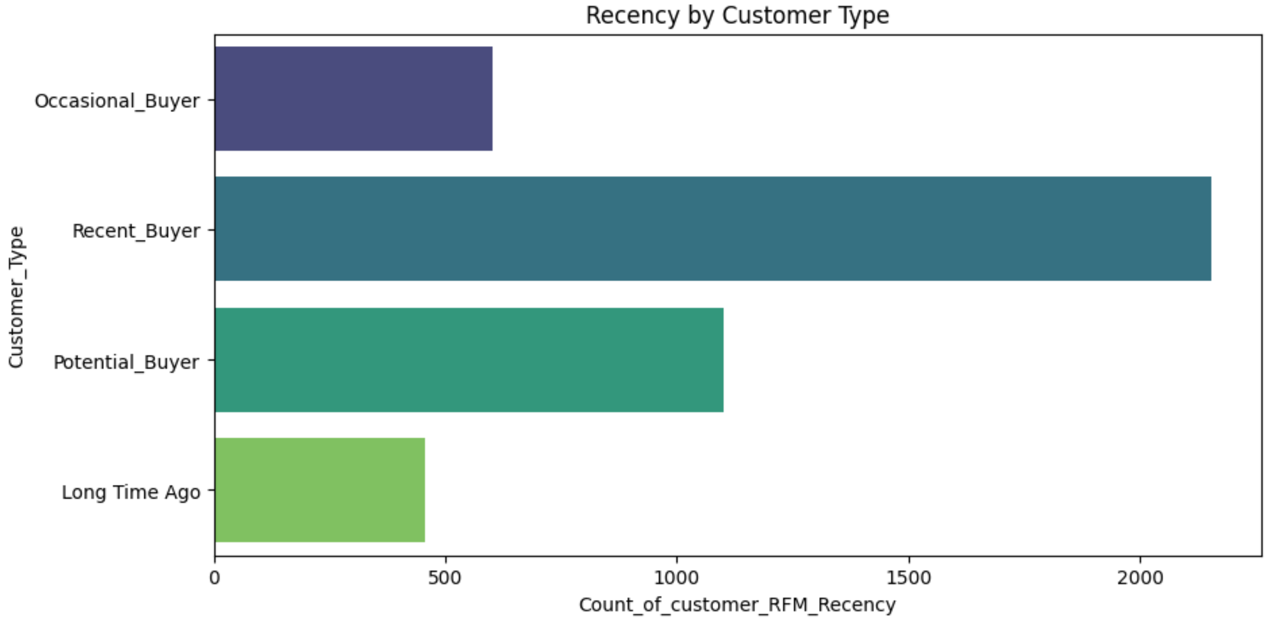
- Recency: Number of days since the last purchase.
- Frequency: Total number of purchases.
- Monetary: Total monetary value spent by the customer.

It is very good at clustering and we used the elbow method for detecting an Optimal number of clustering. This clustering is helping segmentation our customers.

Here Right Top Side ,This box plot reveals variations in customer recency scores across categories.

Right Bottom side, This bar chart visualizes the number of customers across different categories according to their recency score.





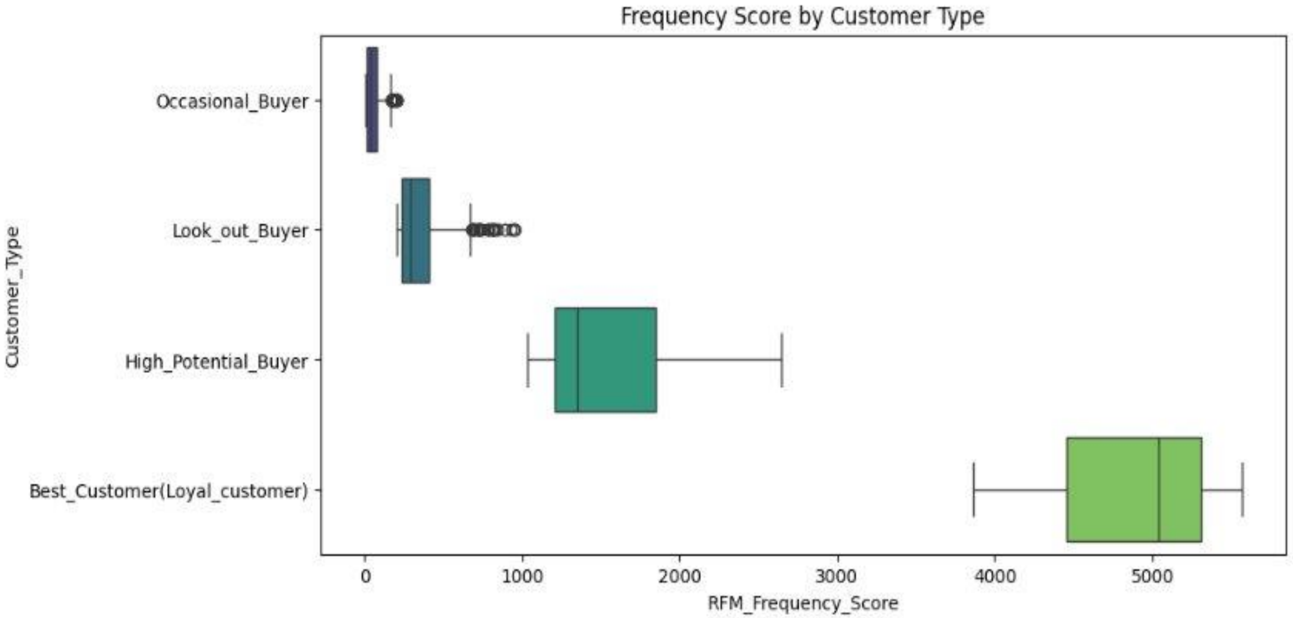
Result

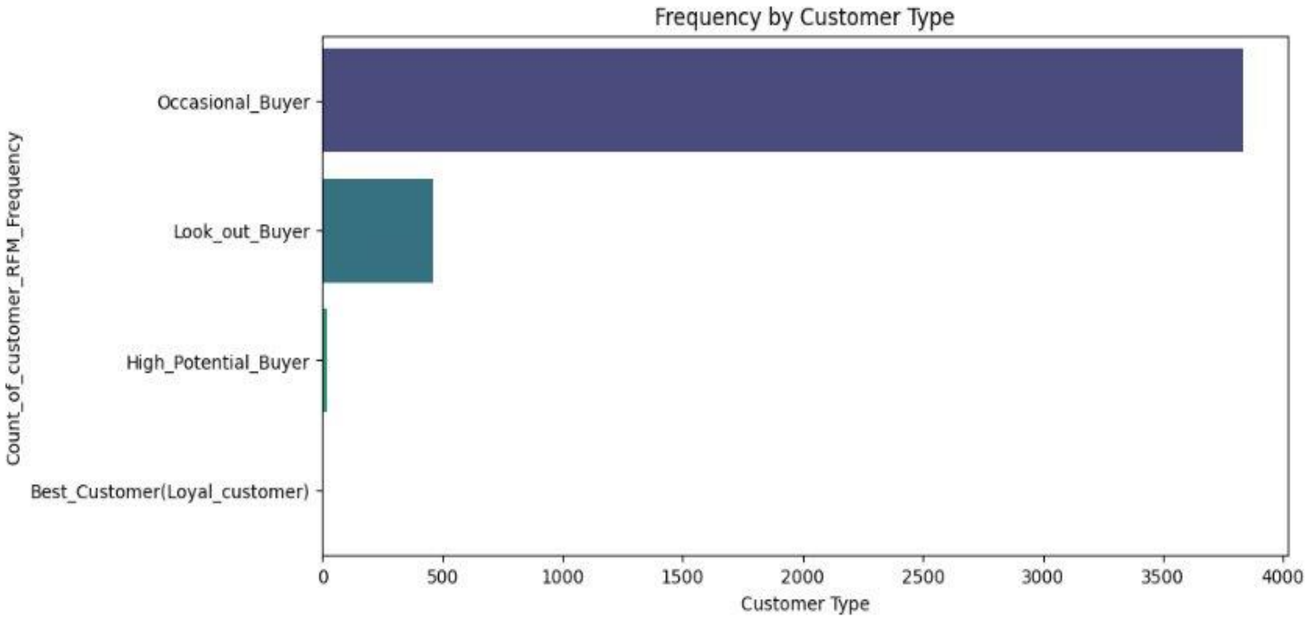
Our Four Clusters are

- Recent Buyer: These are our Customers whose Recency score lie between 0 to 50. It is the highest in Quanity.
- Potential Buyer: Recency Score range is 50 to 150 as per the Box plot and through proper marketing strategy we can make this category Loyal Buyer.
- Occasional Buyer : Recency Score Range 150 to 250.
- Long time Ago : Recency Score range is 250 to 350.And it lowest in Total customer.

Here Right Top Side ,This box plot reveals variations in customer frequency scores across categories.

Right Bottom side, This bar chart visualizes the number of customers across different categories according
to their frequency score.





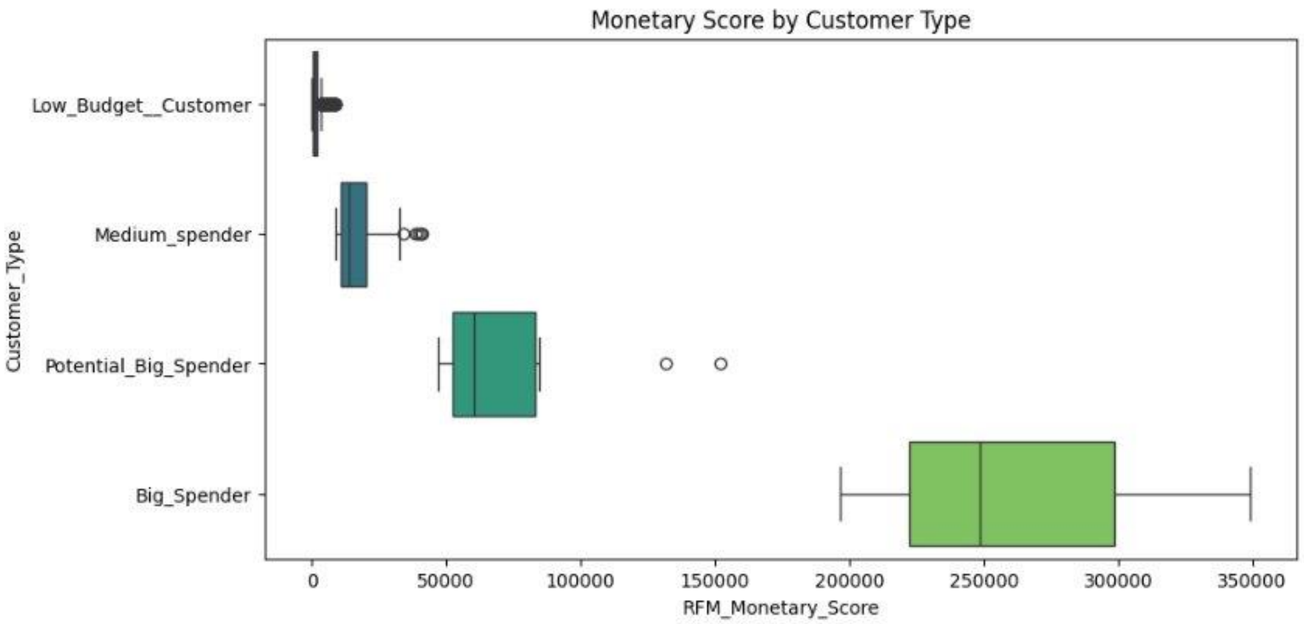
Result

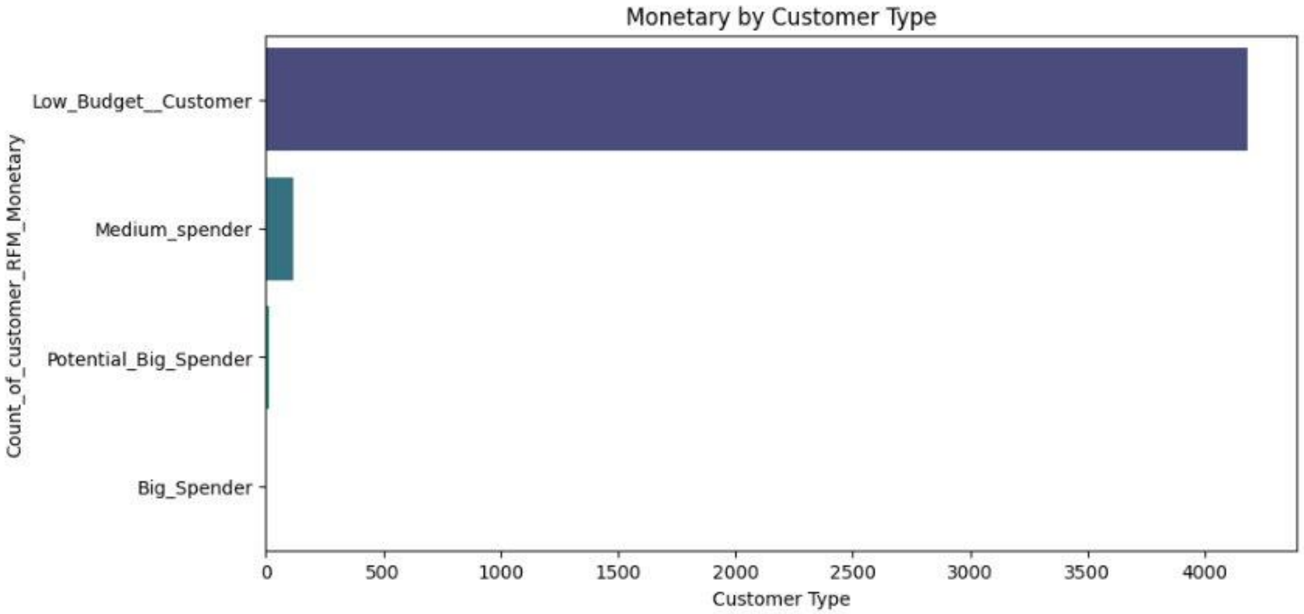
Our Four Clusters are

- Best Customer: These are our Customers whose Frequency Score is very high but they are less number.
- High Potential Buyer: Frequecy Score is in between 1000 to 3000 .
- LookOut Buyer : Frequency Score are Medium Range.
- Occasional Buyer : Frequency Score is between 500 But they are almost 95 percent of Customer So we can focus this Category for more Sales.

Here Right Top Side ,This box plot reveals variations in customer Monetary scores across categories.

Right Bottom side, This bar chart visualizes the number of customers across different categories according to their Monetary score.





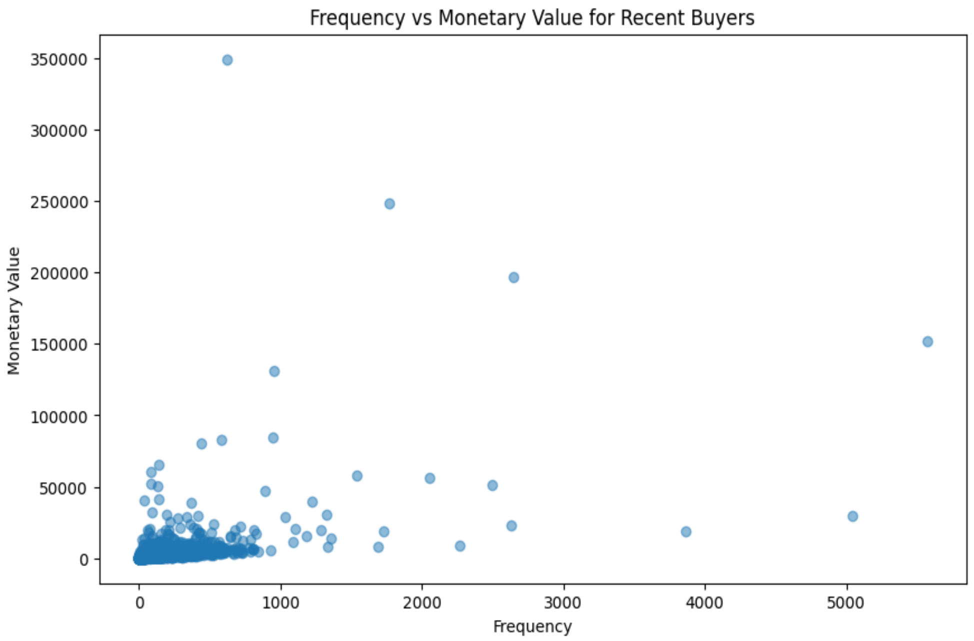
Result

Our Four Clusters are

- Big Spender : These are our Customers whose Monetary Score lie between 200000 to 350000. They have the capacity of High Spending.
- Potential Big Spender: Monetary Score Lie beteen 50000 to 100000.
- Medium Spender : They spend less than 50000.
- Low Budget Customer : They are 95 percent of Customer, we can focus on their purchasing Practices for improve Sales

Key Visualisation

ScatterPlot of frequency vs monetary value For Recent Buyers



Top 10 Recent Buyer who has max Frequency In purchasing Product

	Customer ID	Frequency	Monetary
0	14911.0	5570	152147.57
1	17841.0	5043	29892.50
2	14606.0	3866	18704.41
3	14156.0	2648	196566.74
4	12748.0	2634	22879.66
5	17850.0	2499	51208.87
6	16549.0	2269	9027.55
7	15311.0	2055	56003.26
8	14646.0	1774	248396.50
9	14527.0	1729	19283.48

Top 10 Recent Buyer who bought highest

	Customer ID	Frequency	Monetary
0	18102.0	627	349164.35
1	14646.0	1774	248396.50
2	14156.0	2648	196566.74
3	14911.0	5570	152147.57
4	13694.0	957	131443.19
5	17511.0	948	84541.17
6	15061.0	584	83284.38
7	16684.0	441	80489.21
8	16754.0	140	65500.07
9	17949.0	87	60117.60

Strategical Points For Marketting

1.***A1 Customers:***

It's great to focus on **recent customers** based on their monetary and frequency value. If we understand their purchasing behavior and needs, we can deliver more products tailored to their preferences.

2. ***A2 Customers:***

Our next target will be **our occasional buyers**, prioritized by frequency, as they represent the largest group in quantity. For these buyers, we can sell similar products or offer discounts to increase their participation and buying frequency.

3. ***A3 Customers:***

If we analyze **monetary value clustering**, we find that our low-budget customers are the most numerous. Therefore, by creating quality products within a lower budget, we can significantly boost our sales.

Operation Analysis: Sales Forecasting

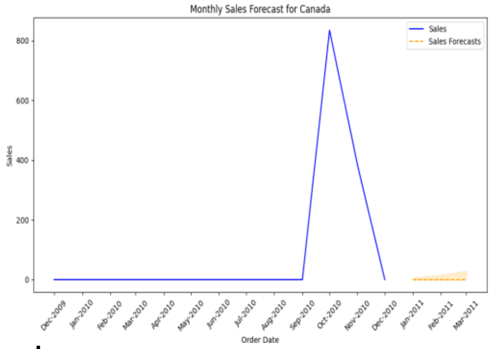
- **Model Explanation and Justification**

To forecast sales at the country level for the next quarter, we used the SARIMA (Seasonal Autoregressive Integrated Moving Average) model. The SARIMA model is well-suited for time series data with seasonal patterns, making it ideal for forecasting monthly sales data.

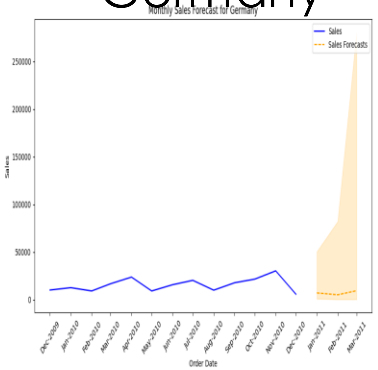
The SARIMA model was chosen due to its ability to handle complex seasonality in time series data. That's why it is goof fit.

G7 COUNTRIES

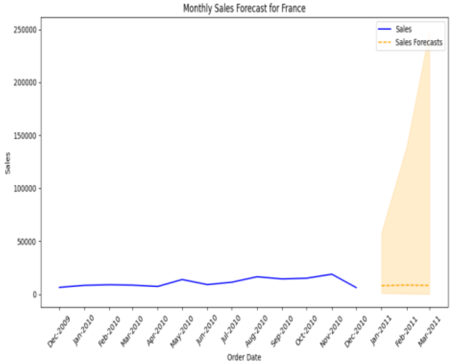
Canada



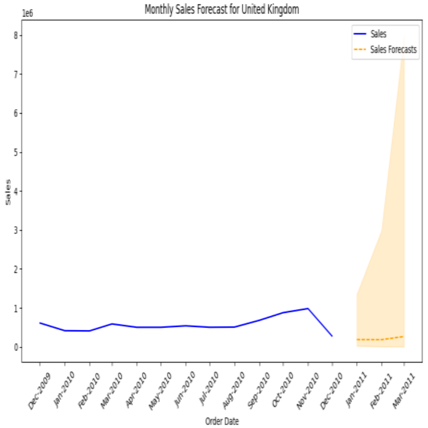
Germany



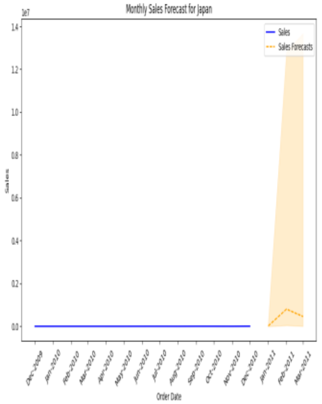
FRANCE



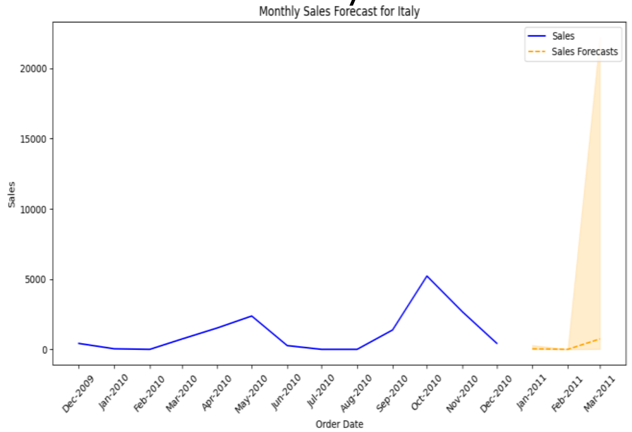
UK



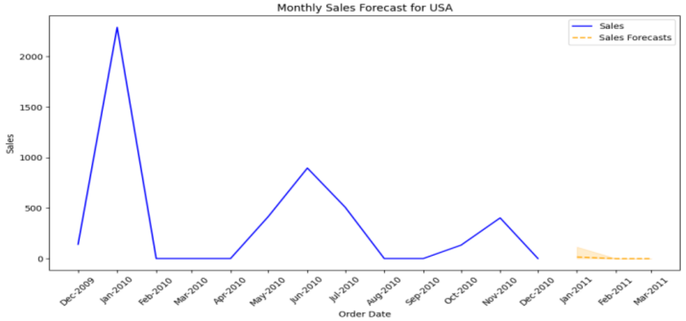
Japan



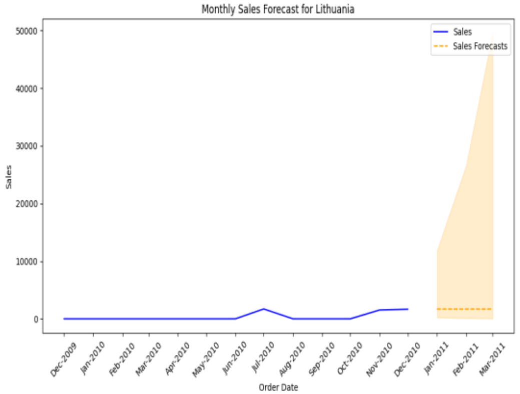
Italy



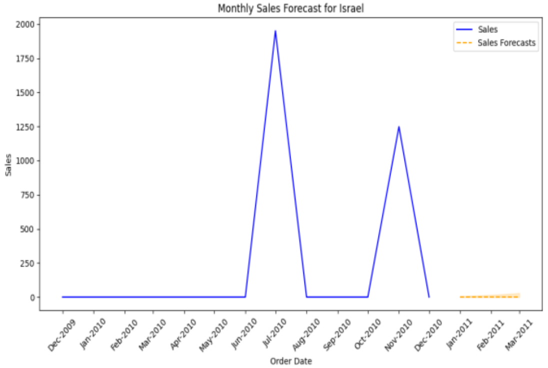
USA



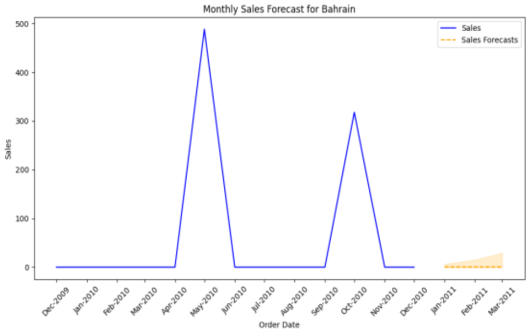
Lithuania



Israel

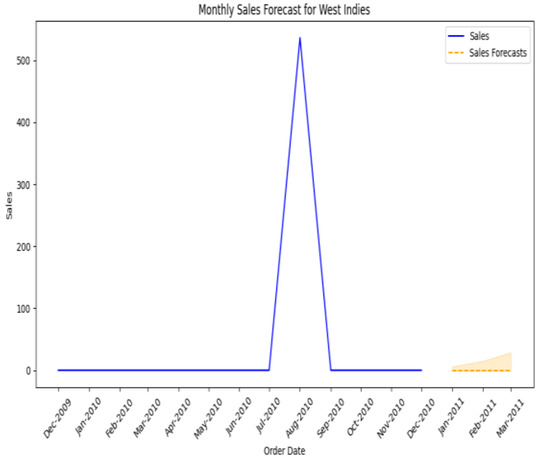


Bahrain

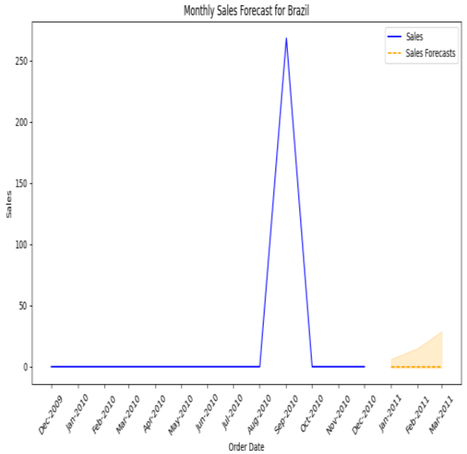


Thailand

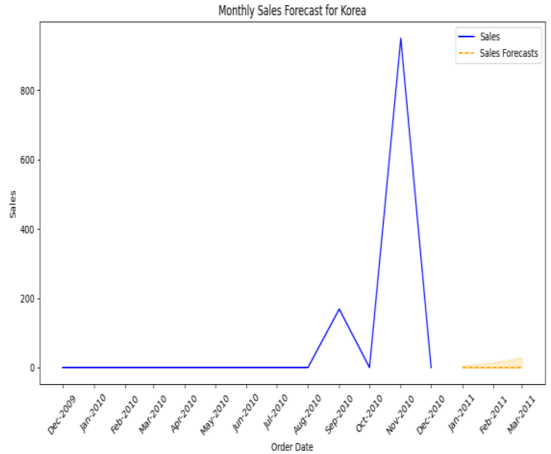
West Indies



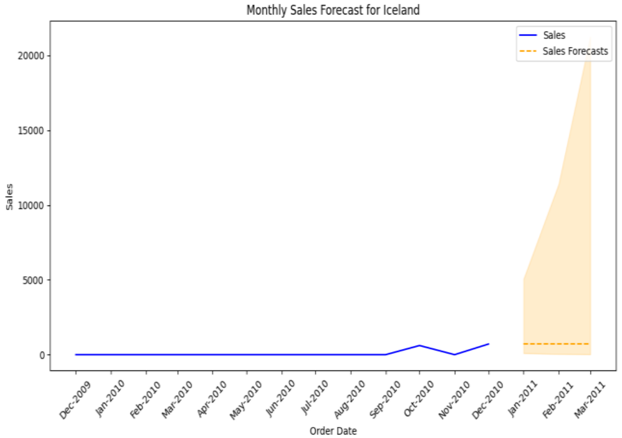
Brazil



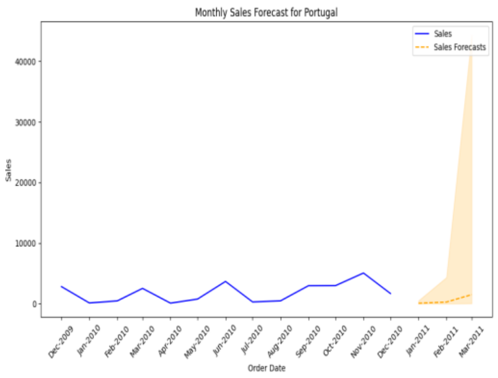
South Korea



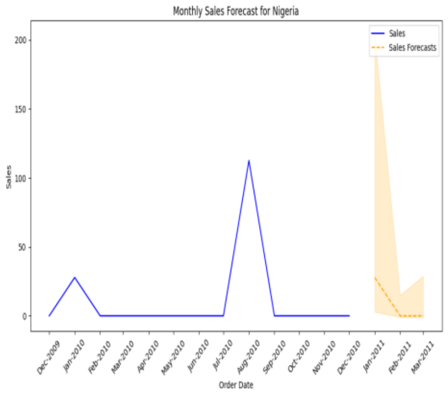
Iceland



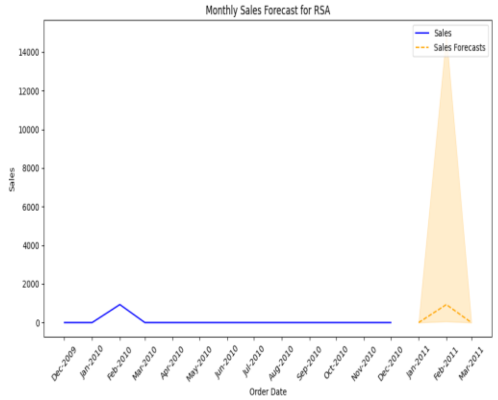
Portugal



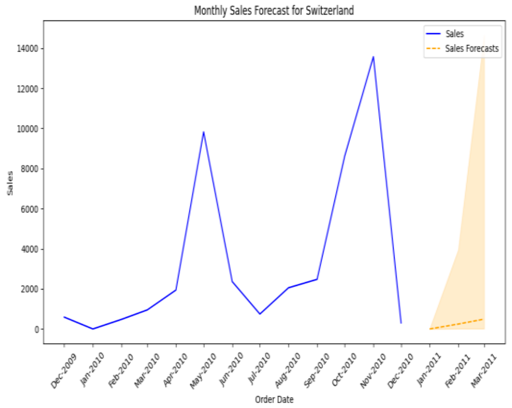
Nigeria



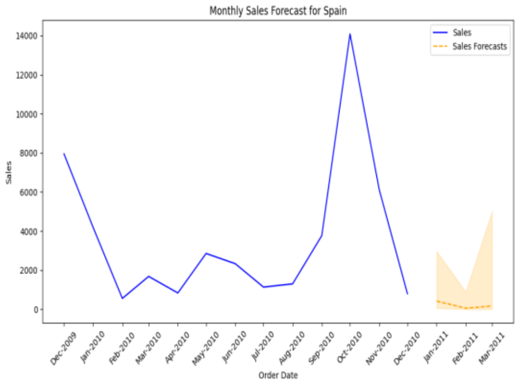
Russia



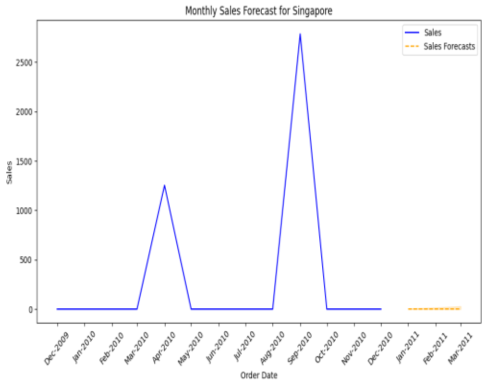
Switerland



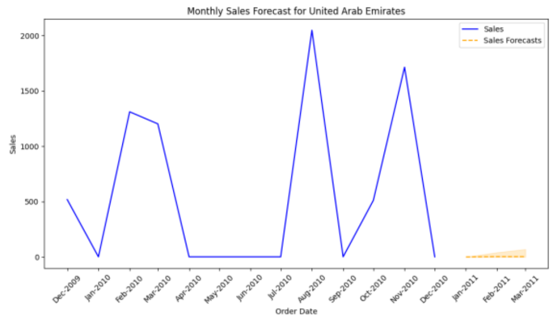
Spain



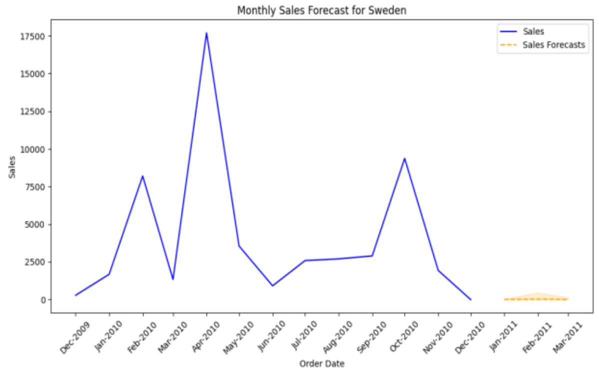
Singapore



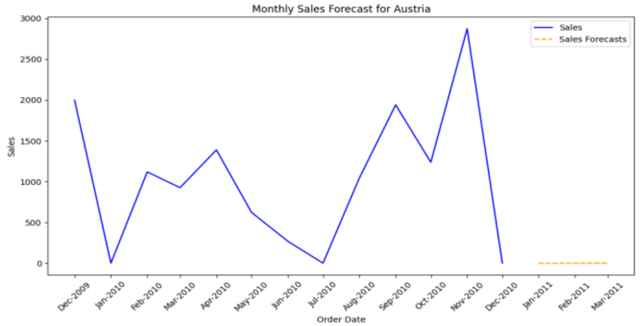
UAE



Sweden

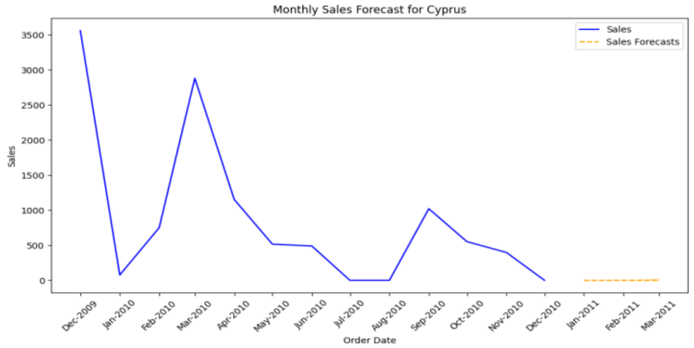


Australia

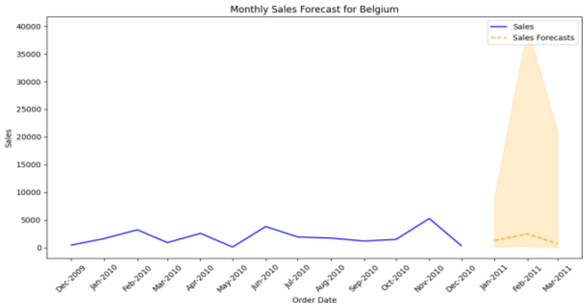


Greece

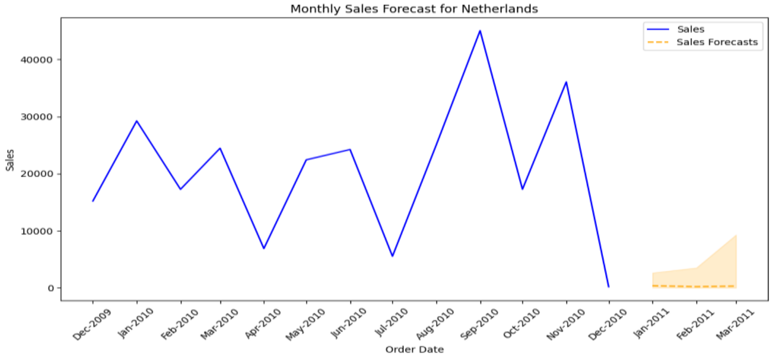
Cyprus



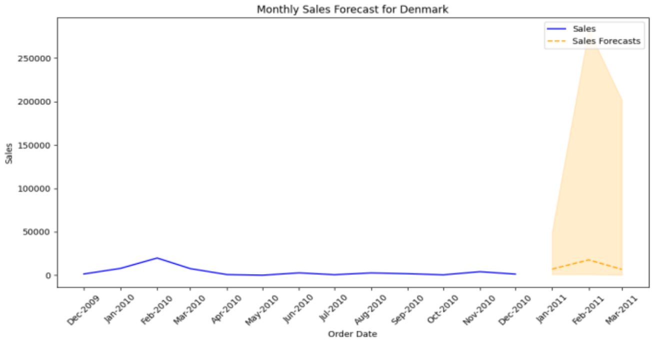
Belgium



Netherland



Denmark



Strategy For Operation Team

1. If We see G7 Country which are most Developed USA,UK,FRANCE,GERMANY,ITALY,CANADA,JAPAN, There is more Customers on there consuming Basis we have enourmous Opportunity for operating There.
2. Another Strategy is where is Most of our NRI stay This is easy for us operation there and Expand for market Base. These Countries are : USA,UK,UAE,AUSTRAILIA,Singapore,Bahrain.
3. Then we can see our existing base from our previous Database EXAMPLE: Netherland,South korea, Brazil,Switzerland,sweden,Spain etc.

Retail Product Analysis

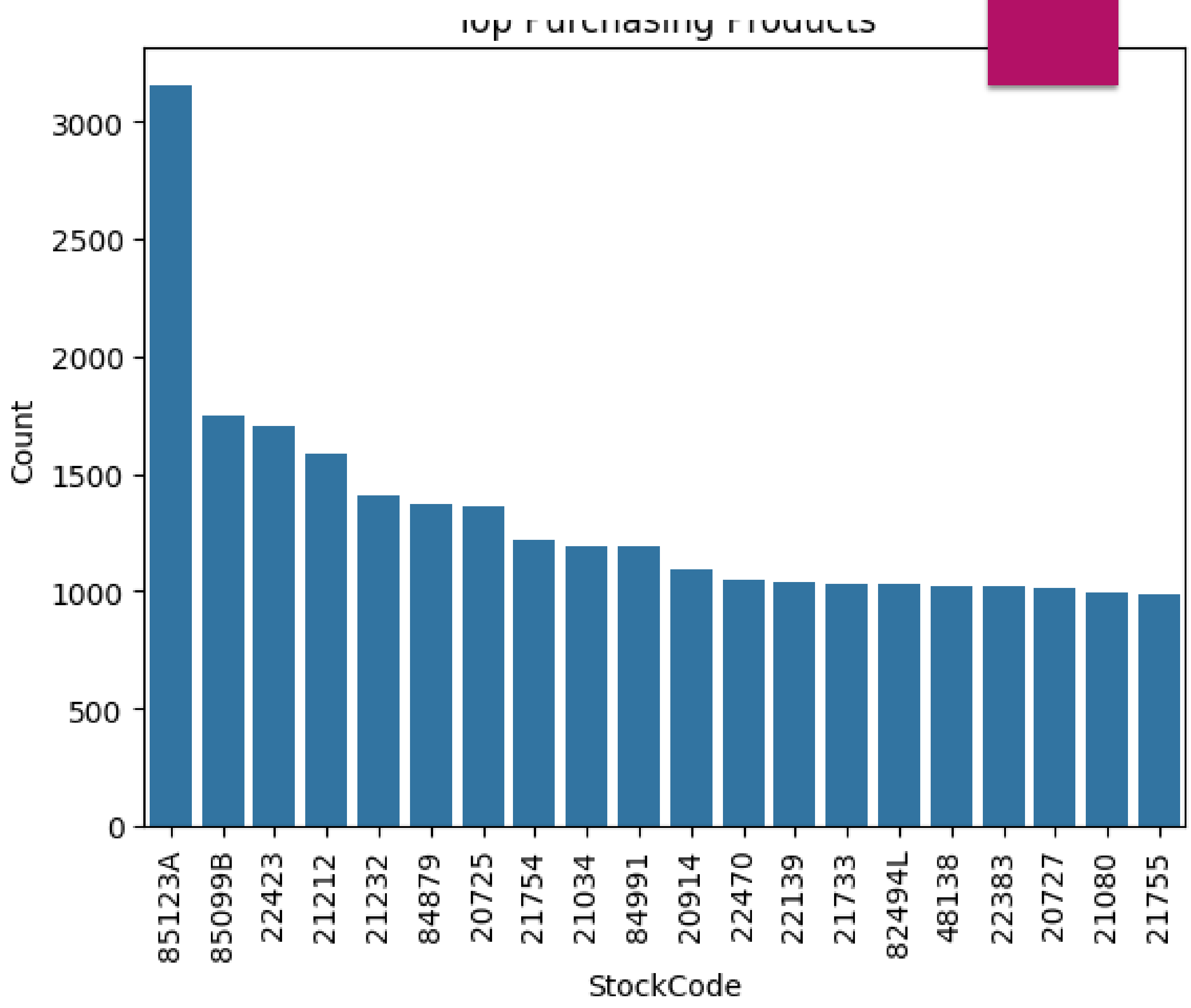
- **Model Explanation and Justification**

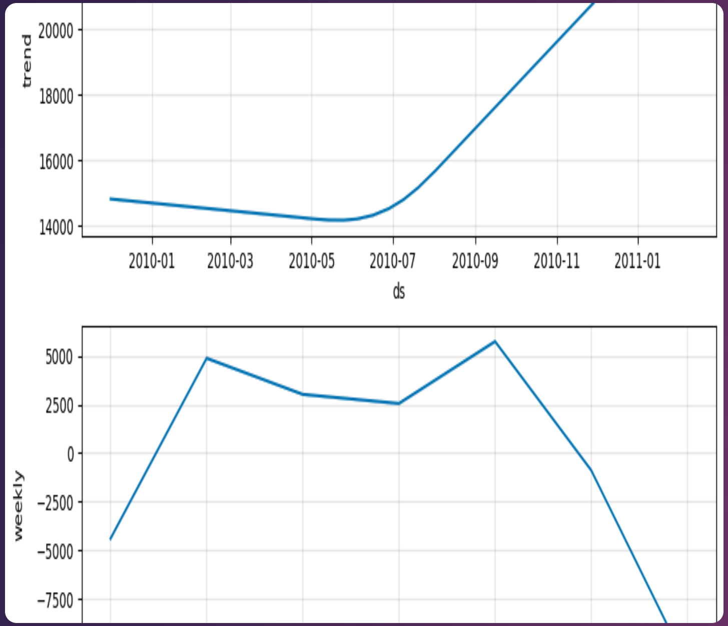
For Daily sails Forecasting we used Prophet library which is help us Time Series Forecasting,Handling Seasonality,Trend Detection etc.

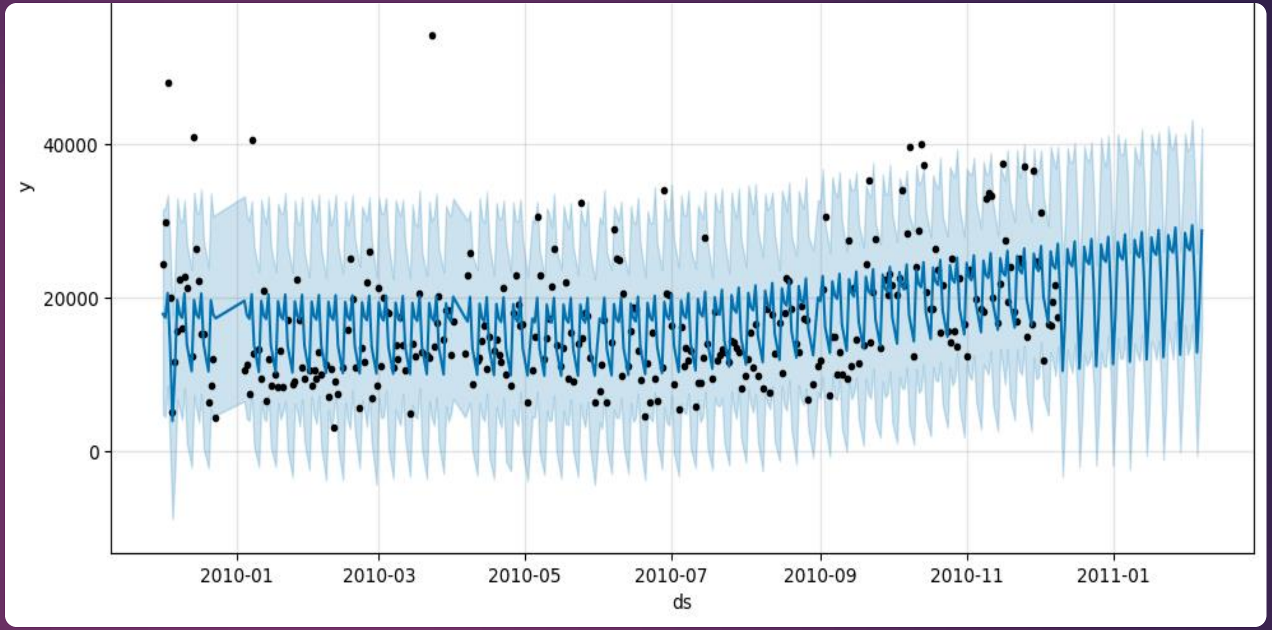
The Apriori algorithm is effective for finding frequent itemsets and association rules, which helps in understanding product co-purchase patterns and determining the next best product to place alongside top-selling items.

Top Purchasing Products

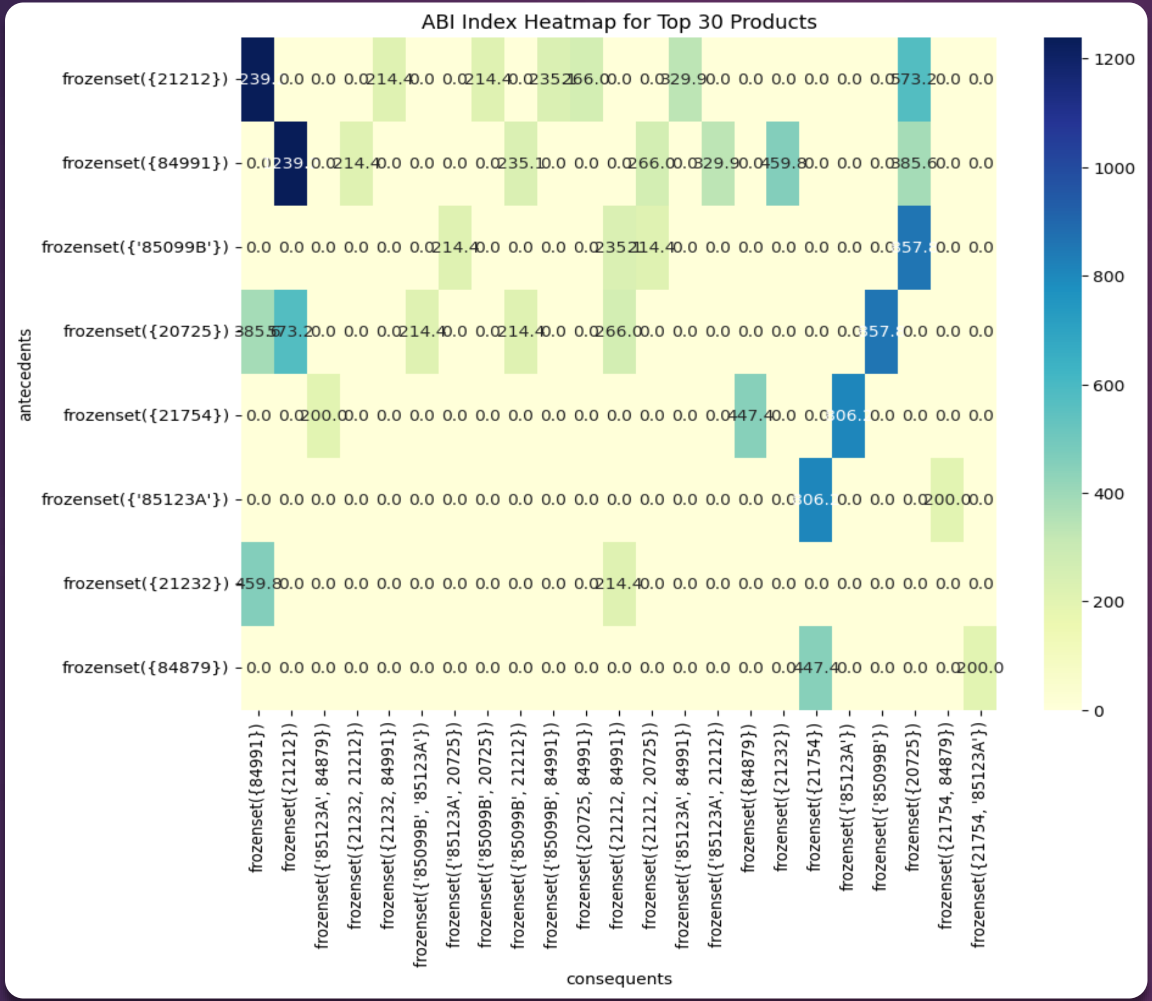
We can see Top 20 most demanding Product in the market.







Daily Sales Forecasting



The darker the color in the heatmap, the higher the ABI index value for that particular combination of antecedent and consequent products. This indicates a stronger association between those two products being purchased together in the same transaction.

ABI Index

Strategy For Retail Products

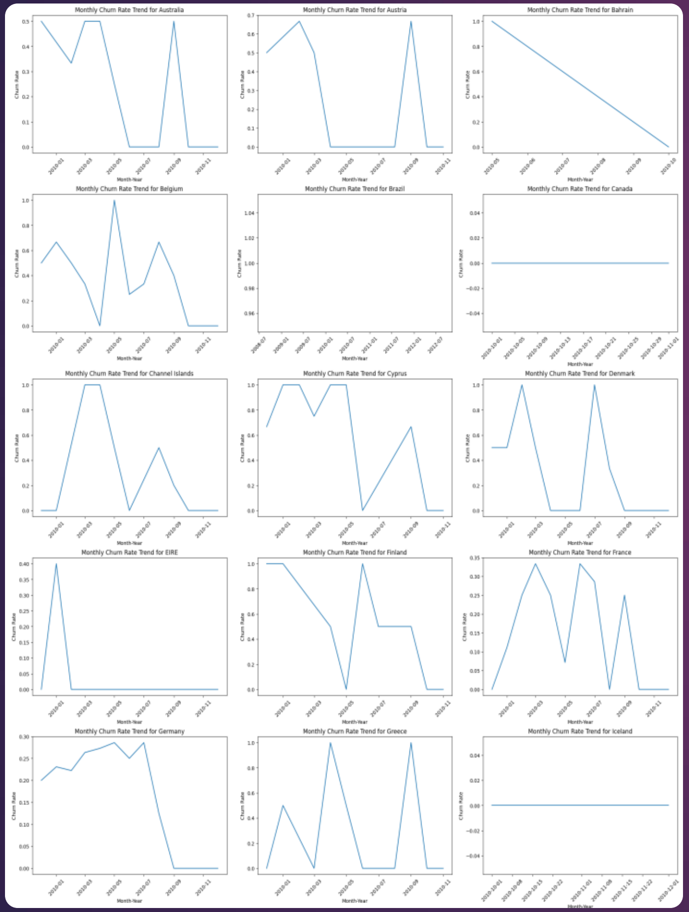
1. We have Top 20 purchasing products .So we can focus on Upselling these products as per the consumer demands . As these are Top Products we can arrange them infront in our physical Store.
2. From Daily Sailes Forecasting, we have seen that our sales are constanly increasing and our 3 months Forecasting is also great So we have good in Future from this Data.
3. ABI Index Heatmap tell us the Crossselling Example: These two products {21212, 84991} are High Potential of Cross Selling we can arrange side by side in the Store .

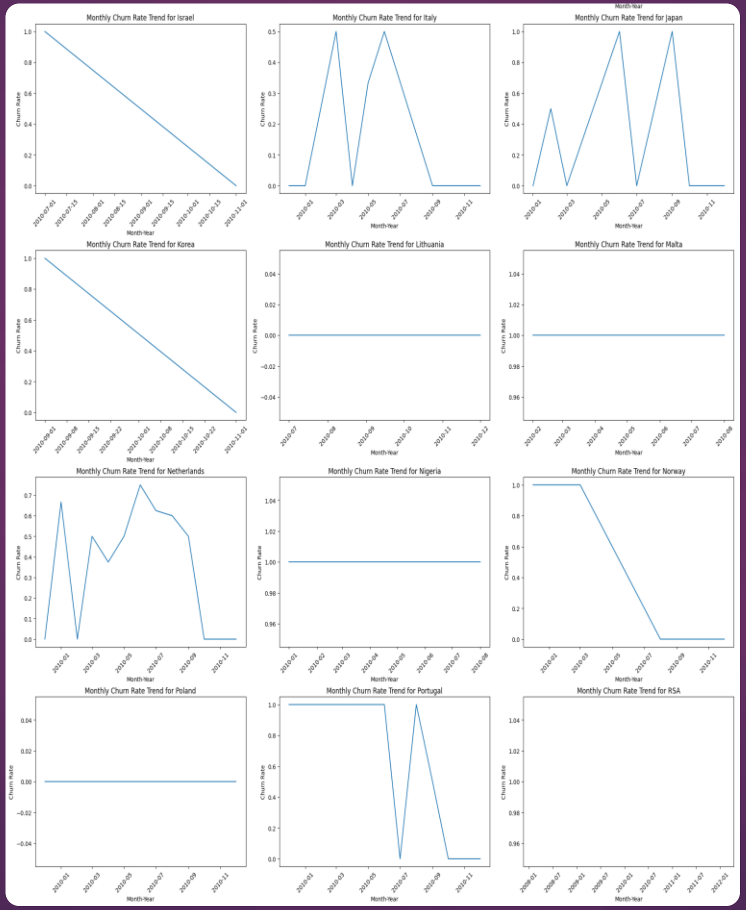
Customer Relationship Management:Churn Detection

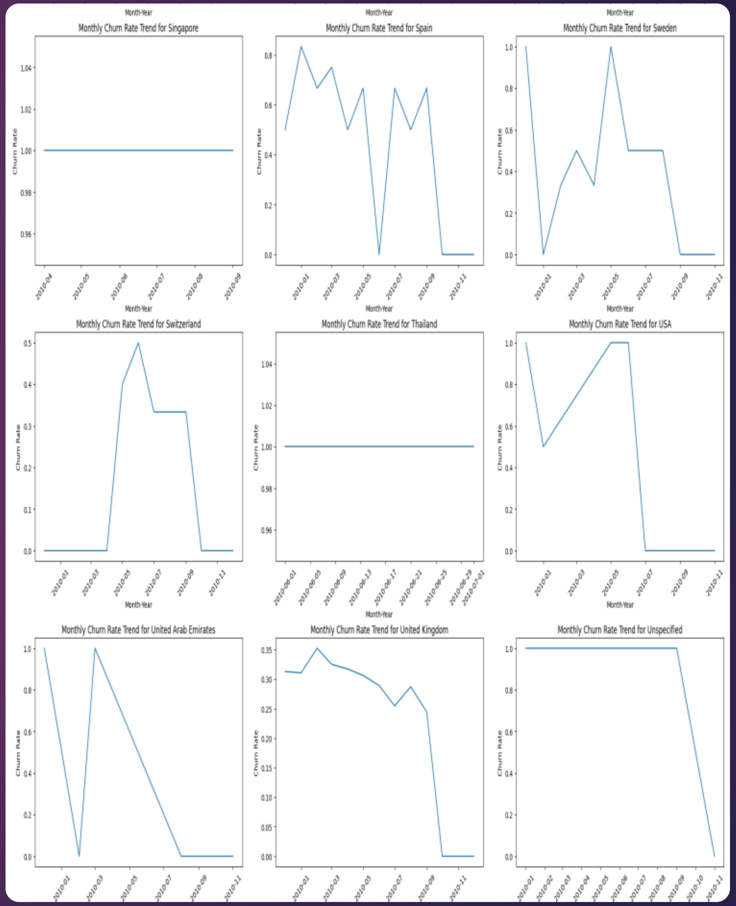
- **Model Explanation and Justification**

By using the Random Forest Classifier, the organization can leverage its ability to handle non-linear relationships, understand feature importance, and achieve robust performance in predicting customer churn at the country and month levels. The quantitative evaluation metrics further support the choice of this model by providing a measurable assessment of its predictive capabilities.

Customer Churn of Each Country on Monthly Basis







Strategy For reduction in Churn rate

According to the visuals we divided our Customers three Categories:

3 or more peak Countries like Spain, the Netherlands, France, and Greece show higher purchasing rates, indicating a high churn rate frequency among customers from these regions.

Less than 3 peak Like Brazil , Lithuania They are comparatively less churn frequency rate.

THANK YOU