Unlocking Customer Insights: Data-Driven Strategies for Growth

BY SHIVANSH TRIPATHI

OBJECTIVES:

- Segment customers using RFM analysis and clustering techniques.
- Forecast sales at a country level to optimize inventory management.
- Identify product associations for cross-selling and up-selling strategies.
- Predict customer churn risk to improve retention initiatives.
- Derive actionable insights from transaction data to drive growth strategies.
- Demonstrate problem-solving skills and analytical thinking capabilities.

DATASET LINK

Dataset Overview:

- **Shape:** (525461, 8)
- **Columns:**
  - InvoiceNo: Invoice number (nominal, 6-digit)
  - StockCode: Product code (nominal, 5-digit)
  - Description: Product name (nominal)
  - Quantity: Quantity of products per transaction (numeric)
  - InvoiceDate: Invoice date and time (numeric)
  - UnitPrice: Product price per unit in GBP (numeric)
  - CustomerID: Customer number (nominal, 5-digit)
  - Country: Customer's country (nominal)

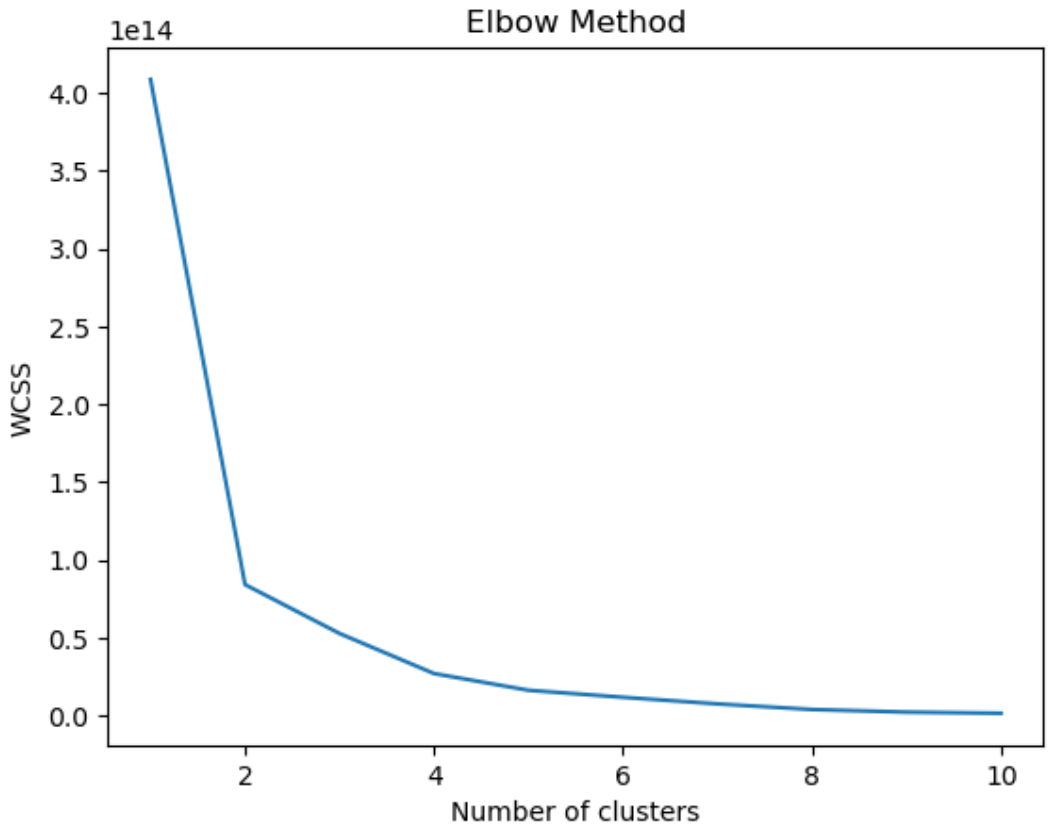
```
(525461, 8)
Index(['Invoice', 'StockCode', 'Description', 'Quantity', 'InvoiceDate',
       'Price', 'Customer ID', 'Country'],
      dtype='object')
Invoice             0
StockCode           0
Description      2928
Quantity            0
InvoiceDate         0
Price               0
Customer ID    107927
Country             0
dtype: int64
```

Number of missing values and their data types of each column

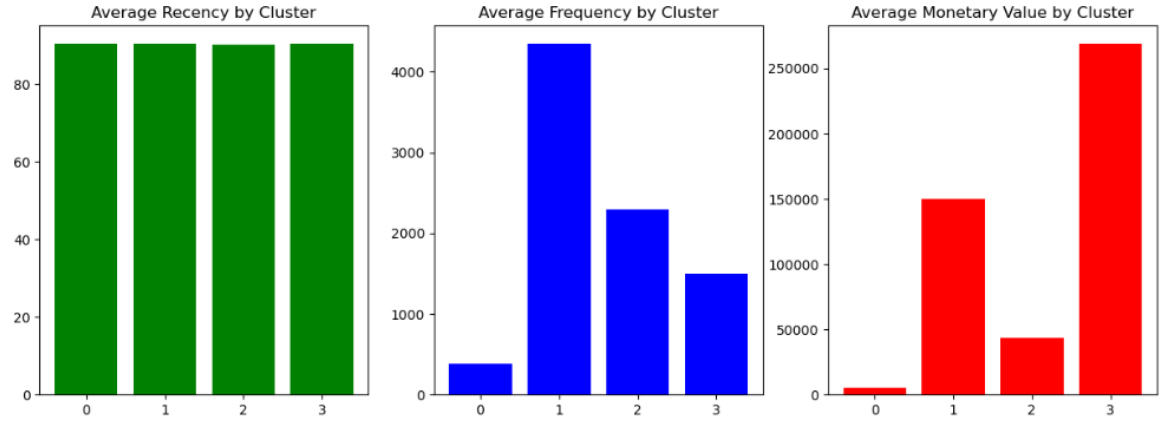
Customer Segmentation and RFM Analysis

- Clustered customers into 4 segments using K-Means clustering on RFM (Recency, Frequency, Monetary) features.
- Justified model selection using the elbow method for optimal number of clusters.
- Visualized cluster distribution on a RFM grid.
- Derived marketing strategies for each cluster based on RFM characteristics (e.g., targeted promotions, loyalty programs).

Elbow Method for optimal number of clusters



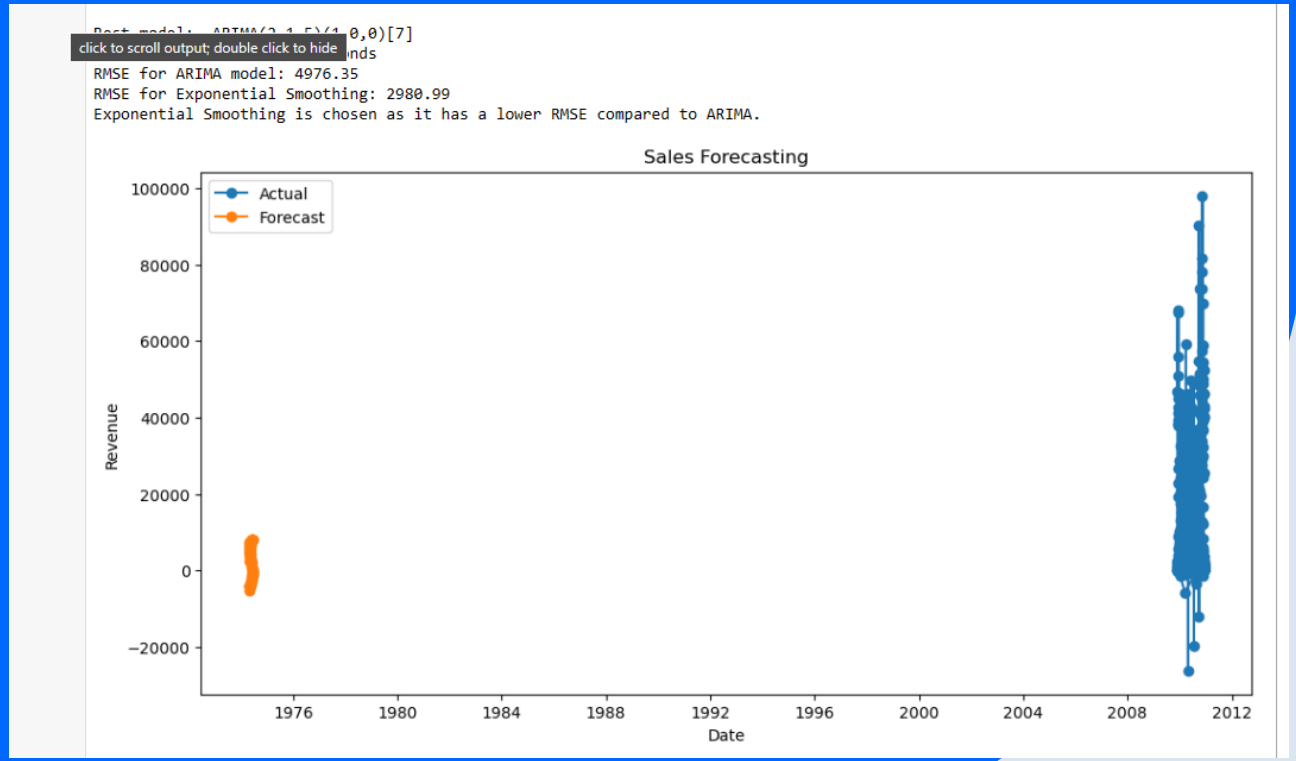
Cluster Distribution on a RFM grid




```
Cluster 0:
- This cluster contains inactive customers. Launch re-engagement campaigns.
- This cluster contains frequent buyers. Offer personalized recommendations.
- This cluster contains high-value customers. Offer premium services and exclusive deals.
Cluster 1:
- This cluster contains inactive customers. Launch re-engagement campaigns.
- This cluster contains frequent buyers. Offer personalized recommendations.
- This cluster contains high-value customers. Offer premium services and exclusive deals.
Cluster 2:
- This cluster contains inactive customers. Launch re-engagement campaigns.
- This cluster contains frequent buyers. Offer personalized recommendations.
- This cluster contains high-value customers. Offer premium services and exclusive deals.
Cluster 3:
- This cluster contains inactive customers. Launch re-engagement campaigns.
- This cluster contains frequent buyers. Offer personalized recommendations.
- This cluster contains high-value customers. Offer premium services and exclusive deals.
```

Sales Forecasting for Inventory Management

- Developed ARIMA model to forecast daily sales at the country level for next 2 months
- Justified model selection by comparing RMSE with Exponential Smoothing models
- Visualized forecasted sales with confidence intervals
- Insights on inventory levels required to meet forecasted demand in each country



Product Recommendations for Cross-Selling and Up-Selling

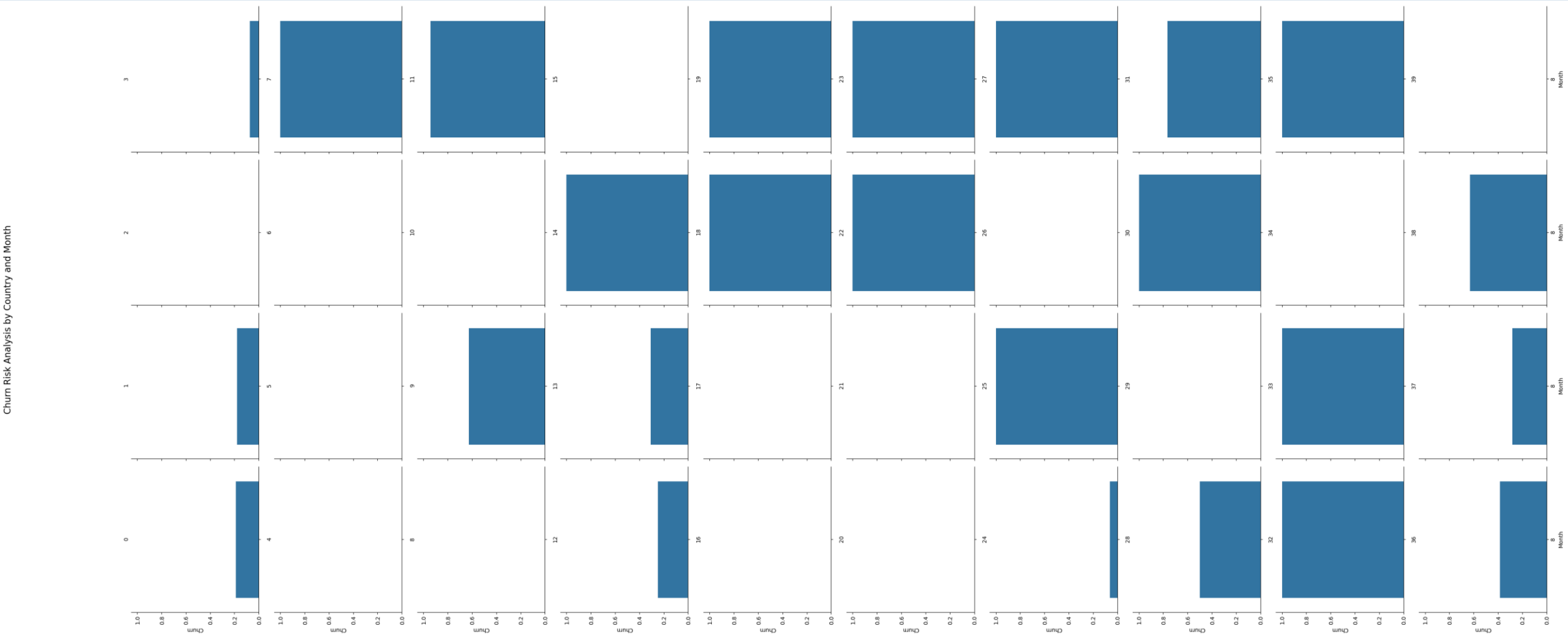
- Used association rules mining (Apriori algorithm) to identify product relationships
- Visualized product recommendations for top 10 selling items based on association rules
- Derived strategies for shelf arrangement to promote cross-selling and up-selling based on top association rules

Top 10 selling items

Customer Churn Prediction and Retention Strategies

- Developed Logistic Regression/Random Forest model to predict customer churn using RFM features
- Justified model selection using AUC scores and confusion matrices
- Visualized churn risk by country and month
- Derived customer retention strategies based on churn risk patterns (e.g., targeted offers, improved service)

Churn Risk by country and month



Thank you