Online Retail Data Analysis: Executive Summary

1. Objective

- Analyze online retail transactions to uncover sales trends, customer segments, and actionable business insights.

2. Methodology

Data Cleaning

- Removed administrative/non-sale transactions (e.g., StockCode: 'AMAZONFEE', 'BANK CHARGES', and others).

- Excluded rows with missing or blank CustomerID.

- Filtered out transactions with UnitPrice $= 0$.

Feature Engineering

- Split InvoiceDate into Date and Time columns.

- Created TotalCost $=$ UnitPrice $\times$ Quantity for each transaction.

Descriptive Analysis & Visualization

- Identified top products by quantity and revenue.

- Analyzed sales trends by month and by country.

- Visualized purchase patterns (e.g., UnitPrice vs Quantity scatter plot).

Customer Segmentation (RFM Analysis)

- Calculated Recency, Frequency, and Monetary values for each customer.

- Standardized RFM features and applied KMeans clustering.

- Labeled clusters: High-value, Loyal, At-risk, Low-value.

- Visualized clusters using pair plots and PCA.

3. Key Findings

- **Top Products**: Most sales volume and revenue are concentrated in a few products.

- **Sales Trends**: Clear seasonality and monthly fluctuations in sales.

- **Geographic Insights**: The UK dominates sales, but other countries contribute significant revenue.

- **Customer Segments**:

  - **High-value**: Recent, frequent, high-spending customers.

  - **Loyal**: Frequent buyers, moderate spend.

  - **At-risk**: Previously valuable, now inactive.

  - **Low-value**: Infrequent, low-spending customers.

4. Recommendations

- **High-value**: Retain with loyalty programs, exclusive offers.

- **Loyal:** Encourage referrals, maintain engagement.

- **At-risk**: Target with re-engagement campaigns and special offers.

- **Low-value**: Use introductory offers and education to increase engagement.

- **Product Strategy**: Focus on best-sellers for promotions and inventory.

- **Geographic Expansion**: Explore growth opportunities in high-revenue non-UK countries.

5. Visual Highlights

- **Bar charts**: Top products and countries.

- **Line plot**: Monthly sales trend.

- **Scatter plot**: UnitPrice vs Quantity.

- **RFM cluster visualizations**: Pair plot and PCA.