MARKET
SEGMENTATION

SEGMENTATION

- Aarti Anil Zikre
- Andrews Truman
- Premkumar Janakbhai Patel
- Vitthlesh Sheth

ABOUT OUR PROJECT

Welcome to our marketing analytics project in DAB303. This project is designed to provide a deep understanding of a retail dataset, enabling us to perform various tasks related to exploratory data analysis, data visualization, and data modeling. Under the guidance of Professor Dr. Andreas S. Maniatis, we aim to gain valuable insights that can inform marketing strategies.

DATA

Transnational data set: UK online retail, 01/12/2010 - 09/12/2011, specializes in unique gifts, many wholesale customers.

MISSION

- Conduct EDA for dataset insights, use advanced ML like segmentation to find customer segments, and extract audience insights for improved marketing strategies.
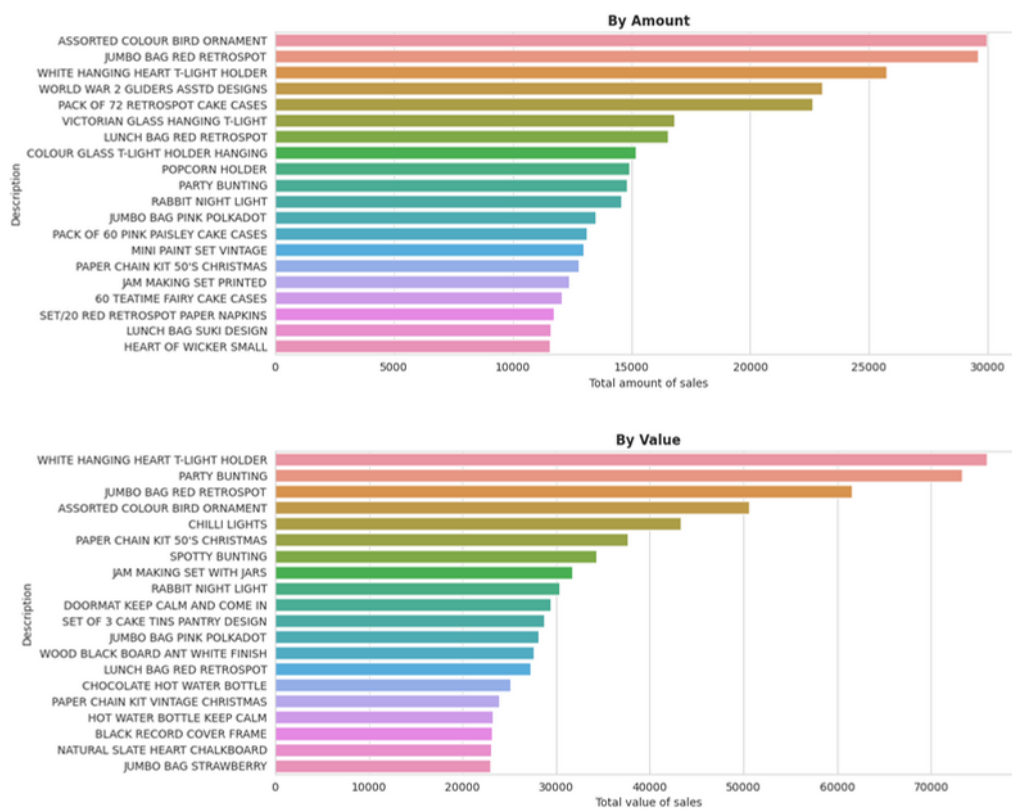
EXPLORATORY DATA ANALYSIS

Best Selling Products by Amount and Value

Best Selling Products by Amount and Value

we present a bar chart displaying the top 20 products that customers have purchased in the highest quantities. This visualization reveals the products with the greatest sales volume, giving us a clear understanding of which items are most popular among our customers.

EXPLORATORY DATA ANALYSIS

Finding the most returned items and the customers with the corresponding country


Most Returned Items

Customers with Most Returns
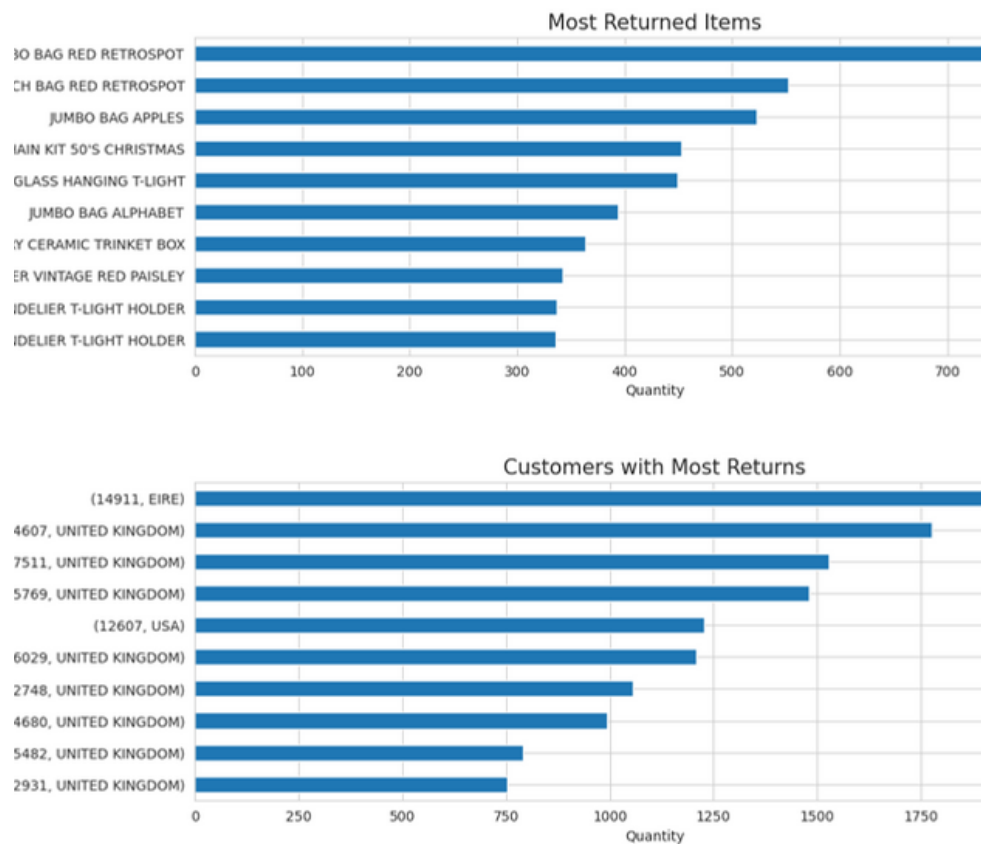
we present a chart highlighting the products that are returned most frequently by our customers. This visualization provides crucial insights into which items may have issues or are less satisfactory to our customer base. Understanding the reasons for returns is pivotal in improving product quality and customer satisfaction.

EXPLORATORY DATA ANALYSIS

## UnitPrice and Quantity Comparison
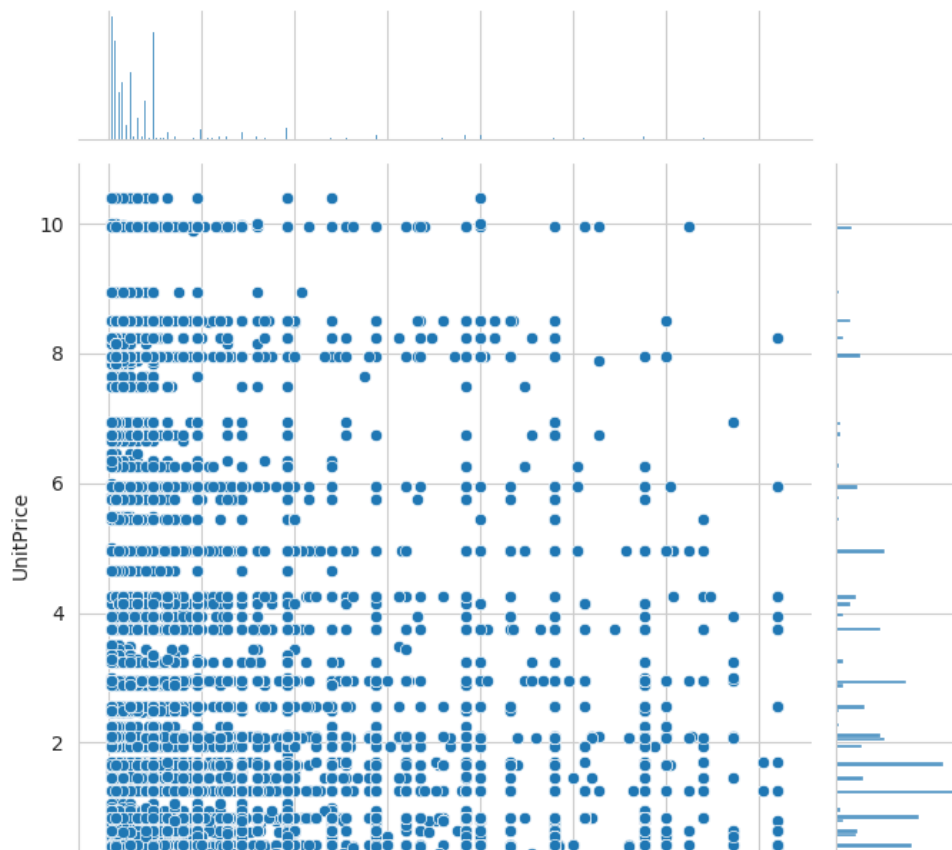
A jointplot depicts the connection between 'UnitPrice' and 'Quantity' of products purchased. It confirms that as prices rise, quantity sold decreases, and vice versa. This illustrates the crucial link between pricing and sales volume, guiding pricing and product strategies.

EXPLORATORY DATA ANALYSIS

Weekly Sales and Returns


Weekly Sales Quantity

Weekly Returns Quantity
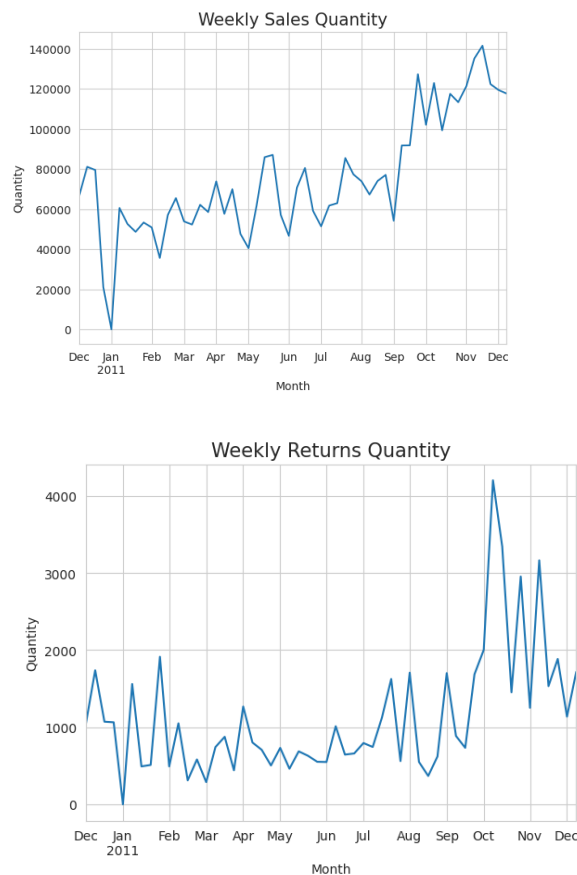
The weekly returns made by customers over the year. Returns appear relatively stable throughout the year, with minor fluctuations, except for a spike in the second week of October. This insight into return patterns helps us address customer satisfaction and product quality concerns more effectively.

EXPLORATORY DATA ANALYSIS
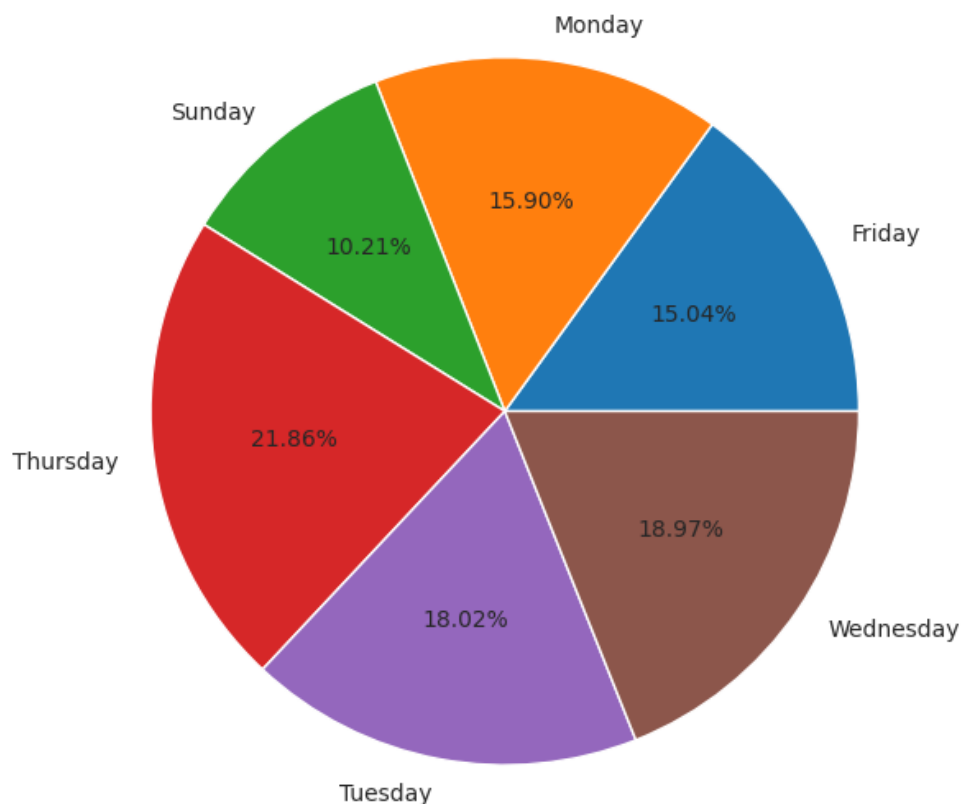
Sales in Foreign Countries
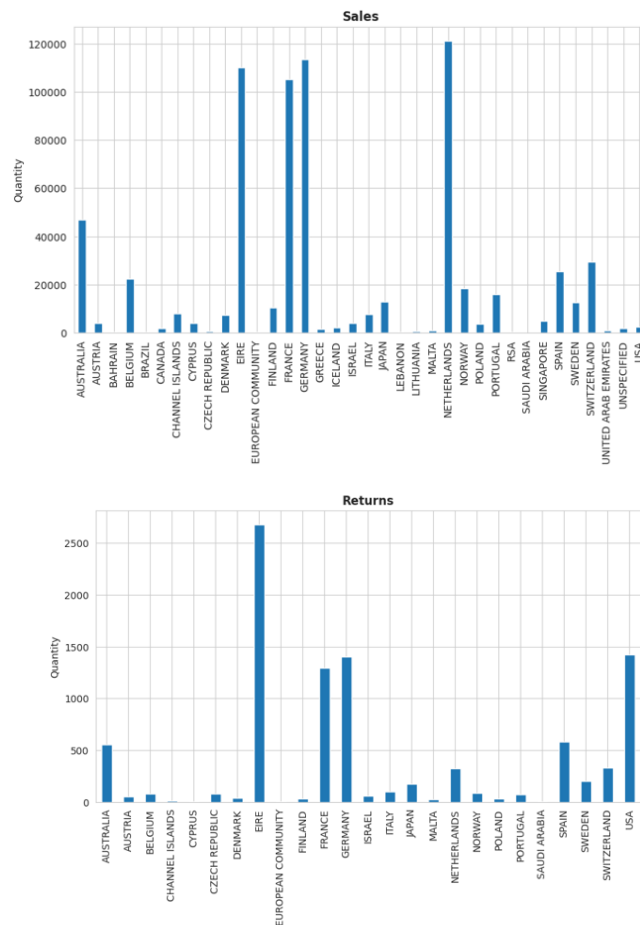
By focusing on countries other than the United Kingdom, we gain a clearer understanding of our international sales and return patterns, allowing us to make informed decisions to optimize our operations.
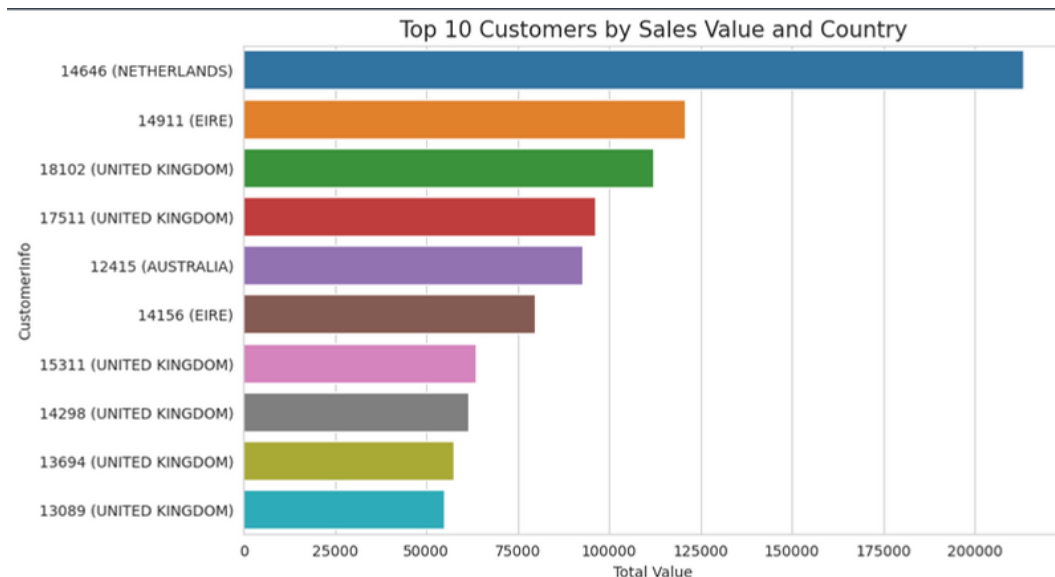
Percantages of Sales Value by Day of Week

The data reveals a notable pattern: Thursday consistently shows the highest sales value, while Sunday records the lowest sales value. This information sheds light on the variations in sales across different days of the week, which can be invaluable for optimizing sales and marketing strategies.

Top 10 Customers by Sales Value and Country

Top 10 Customers by Sales Value and Country

We present a chart that illustrates the sales value by each day of the week. The data reveals a notable pattern: Thursday consistently shows the highest sales value, while Sunday records the lowest sales value. This information sheds light on the variations in sales across different days of the week, which can be invaluable for optimizing sales and marketing strategies.

Understanding the sales patterns throughout the week allows us to align our marketing efforts and promotions with the days that exhibit the highest potential for sales, ultimately enhancing our business performance.

EXPLORATORY DATA ANALYSIS

Top Repeat Customers
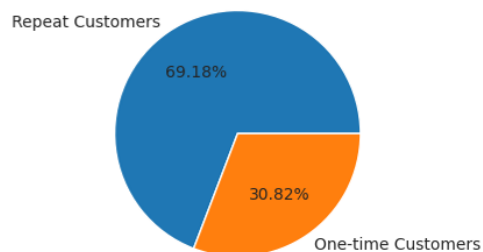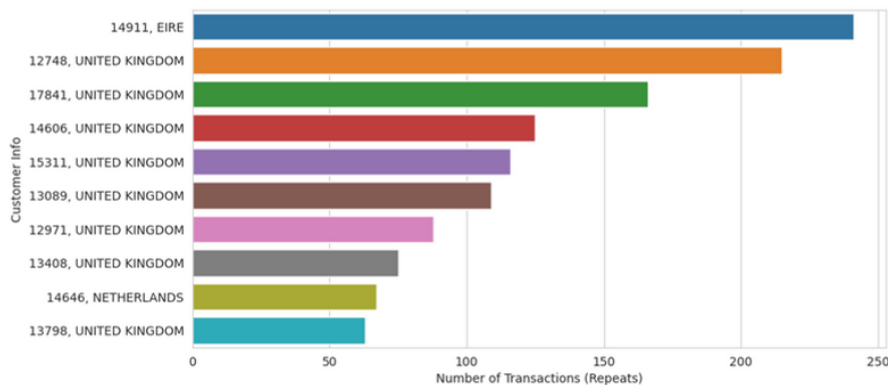
We present a chart that highlights our top customers based on the value they've contributed to the company. Additionally, we visualize the countries from which these valuable customers originate. This analysis allows us to appreciate and nurture our most significant customer relationships while gaining insights into the geographic distribution of these valued clients.

# EXPLORATORY DATA ANALYSIS

UnitPrice and Quantity Distribution
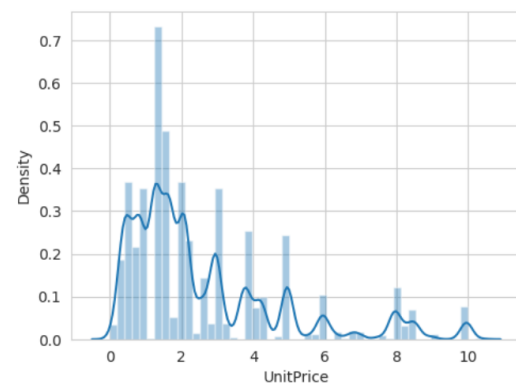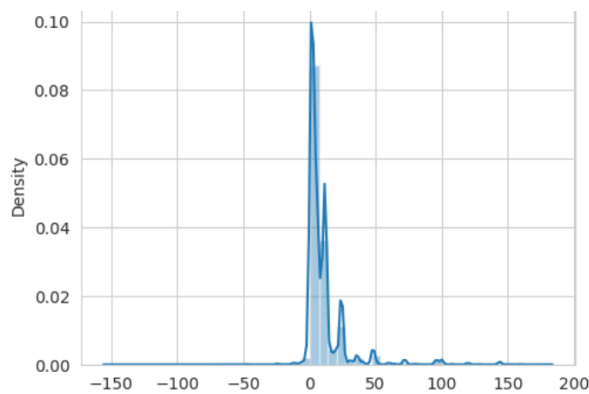
It represents the distribution of the 'UnitPrice' attribute. This plot provides insights into the distribution of unit prices for our products. Understanding the distribution of unit prices is crucial for pricing strategies and product positioning.

EXPLORATORY DATA ANALYSIS

Sales Value per Month and Day of Week

Sales Value per Month and Day of Week
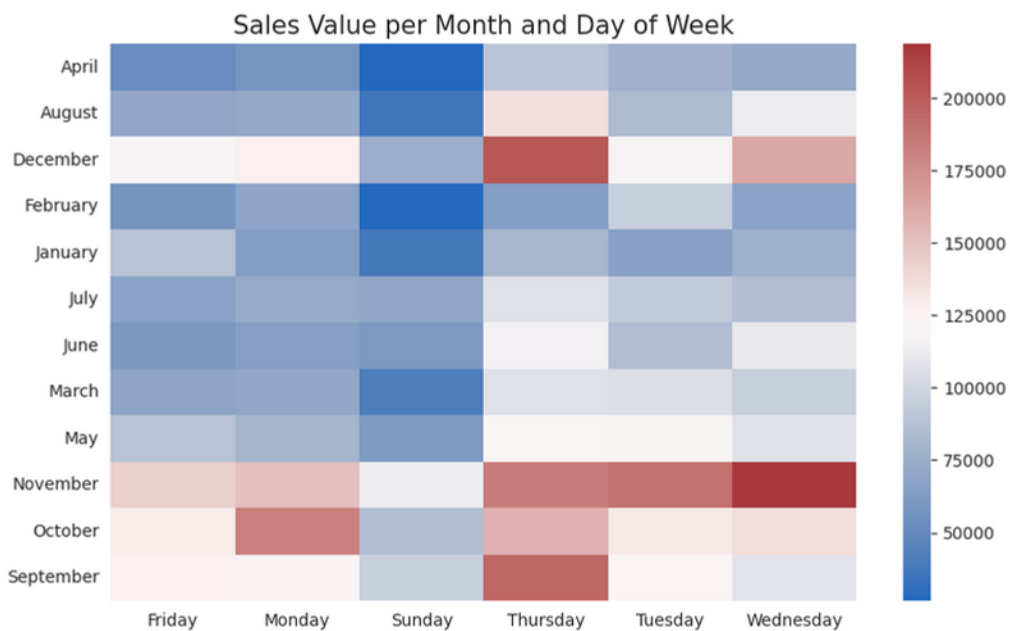
We present a heatmap chart that illustrates how sales are distributed among different months and days of the week. The color scale represents the total value of sales, helping us identify the patterns in sales activity over the course of the year.

ASSOCIATION RULE MINING

Rule 1

Customers who buy a green alarm clock are highly likely to purchase a red alarm clock, with a confidence of 67.1% and a lift of 13.834.

Rule 2

There is a strong association between the "Roses Regency Teacup and Saucer" and the "Green Regency Teacup and Saucer." Customers who purchase one are very likely to buy the other, with a confidence of 69.1% and a lift of 18.017.

Rule 3

Customers often purchase both "Jumbo Bag Pink Polka Dot" and "Jumbo Bag Red RetroSpot," with a confidence of 61.5% and a lift of 7.162.

Rule 4

Different lunch bag designs, like "Lunch Bag Red RetroSpot" and "Lunch Bag Suki Design," show a connection, as customers who choose one design are likely to choose another, with a confidence of 40.1% and a lift of 6.926.

CUSTOMER SEGMENTATION USING CLUSTERING

Customer Segmentation

Customers are categorized into "UK" and "non-UK" groups to consider location-based behavior differences.

Data Aggregation

Individual customer data is consolidated, and statistics like average quantity, price, repeat purchases, and product variety are calculated.

Scaling Features

Numerical features are scaled to ensure balanced influence in clustering analysis.

Clustering Algorithm

In this particular case, the silhouette score is highest when using 3 clusters, which suggests that segmenting customers into three distinct groups may be the most suitable choice for this dataset.

Clustering Metrics
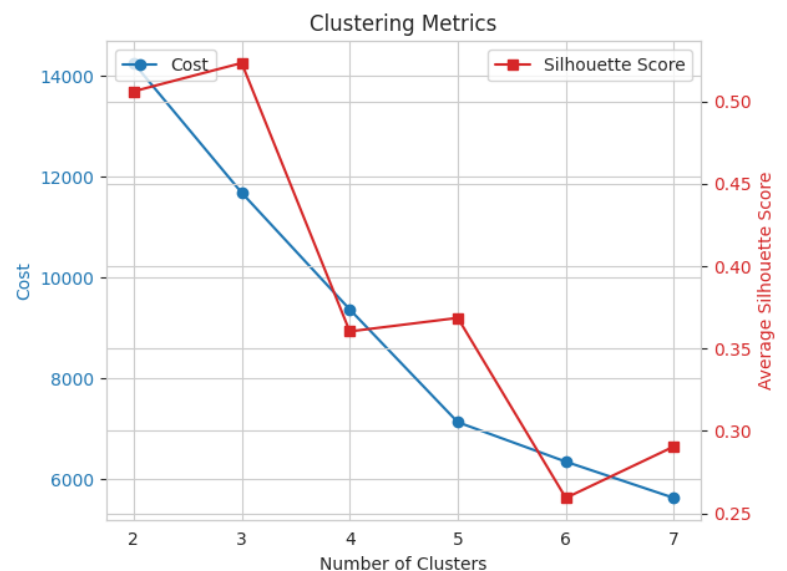
THANK YOU

Group 10

DAB - 303 Marketing Analytics