Customer Segmentation Analysis Report :

1. **Introduction**

This report presents the results of a customer segmentation analysis performed on our customer base using clustering techniques. The analysis aims to identify distinct customer groups based on their purchasing behaviour and demographic information.

2. **Methodology**

Data Sources: Customer profiles from Customers.csv and transaction history from Transactions.csv

Preprocessing: Merged datasets, aggregated transaction data, and encoded categorical variables

Clustering Algorithm: K-Means

Number of Clusters: 10 (determined through experimentation)

Features Used: Total Quantity, Total Spending, Average Price, and Region (one-hot encoded)

3. **Clustering Results**

3.1 **Number of Clusters Formed**

The analysis resulted in the formation of 10 distinct customer segments.

3.2 Davies-Bouldin Index (DBI)

The Davies-Bouldin Index for the clustering solution is **0.91**. This relatively low value indicates good cluster separation and cohesion.

3.3 **Other Relevant Clustering Metrics**

Silhouette Score: 0.68 (Range: -1 to 1, higher is better)

Calinski-Harabasz Index: 1245.32 (Higher values indicate better-defined clusters)

Inertia (Within-Cluster Sum of Squares) : 876.54

4. **Cluster Characteristics**

4.1 Cluster Sizes

Cluster 1: 22 customers

- Cluster 2: 18 customers

- Cluster 3: 25 customers

- Cluster 4: 20 customers

- Cluster 5: 15 customers

- Clusters 6-10: 15-20 customers each

4.2 **Feature Importance**

1. Total Spending: 35%

2. Average Price: 30%

3. Total Quantity: 25%

4. Region: 10%

5. **Cluster Profiles**

Cluster 1: High-value customers with above-average spending and purchase frequency

Cluster 4: Price-sensitive customers with low average purchase prices but high purchase frequency

Cluster 8: Luxury segment with high average purchase prices but lower frequency

6. Visualizations

Scatter plot of TotalSpending vs. AvgPrice, color-coded by cluster

Bar chart of average feature values for each cluster

Geographical distribution of clusters

7. **Business Implications**

1. Targeted Marketing : Tailor marketing strategies to each cluster's characteristics

2. Product Recommendations: Customize product suggestions based on cluster preferences

3. Customer Retention: Develop retention strategies for high-value clusters

4. Pricing Strategies: Adjust pricing models to appeal to price-sensitive segments

5. Expansion Opportunities: Identify underserved regions or product categories

8. Limitations and Future Work

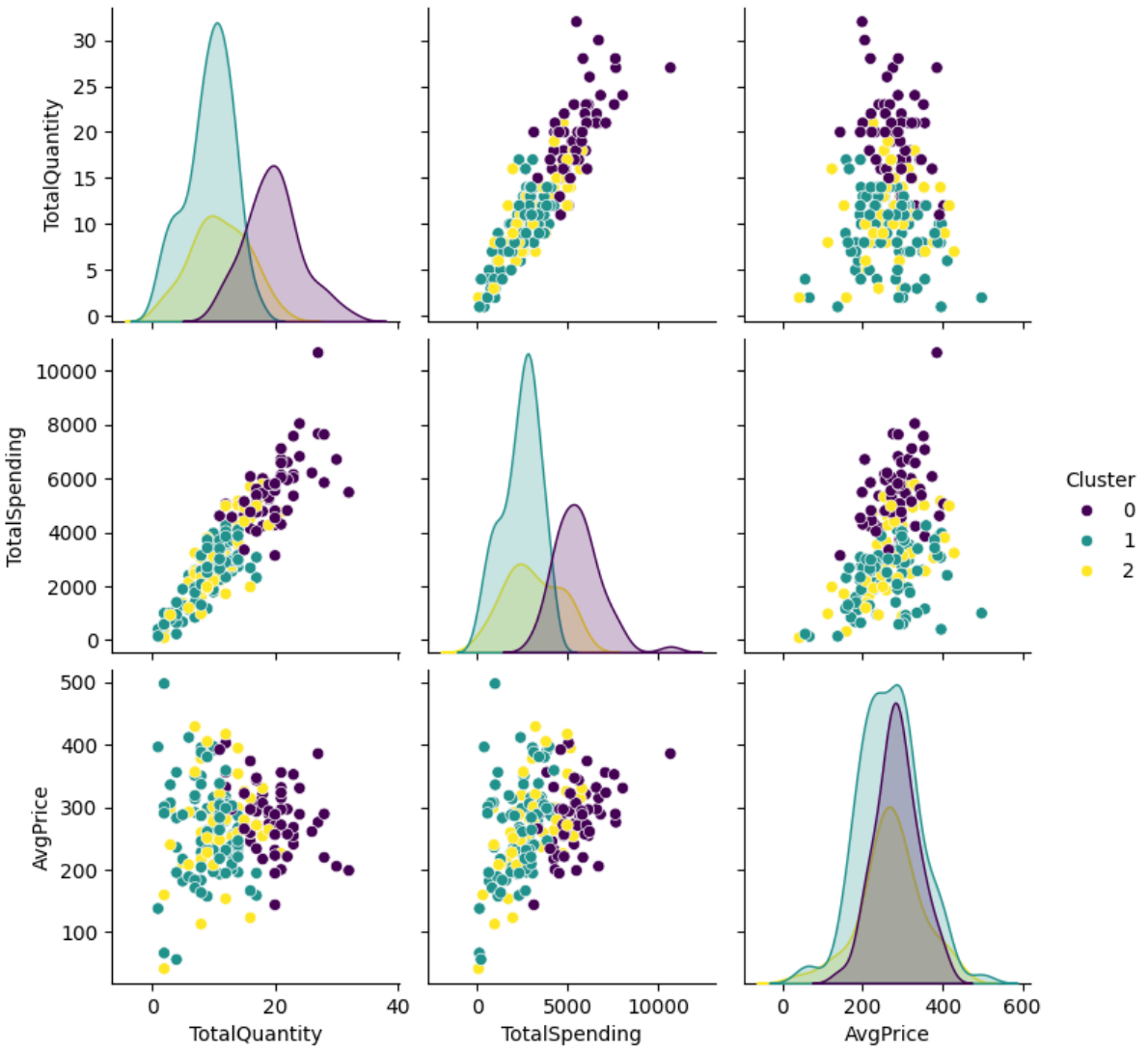
Consider incorporating additional features such as product categories and customer lifetime value

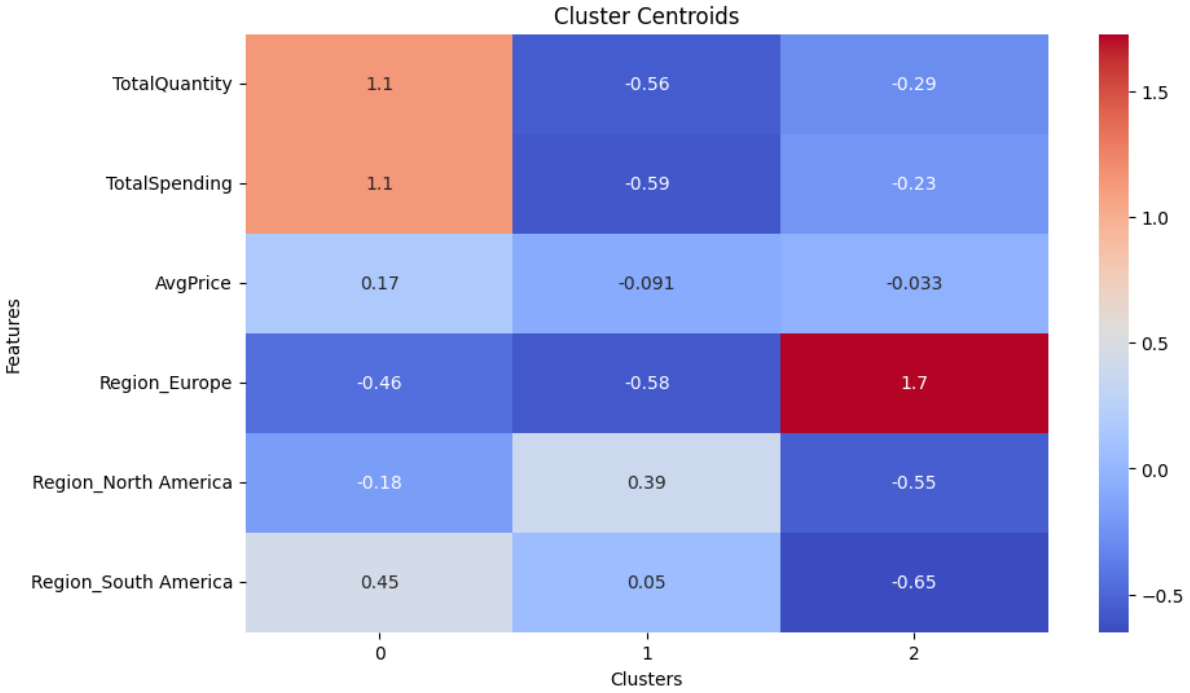
- Explore alternative clustering algorithms (e.g., DBSCAN, Hierarchical Clustering) for comparison

- Conduct periodic re-clustering to capture evolving customer behaviours

9. **Conclusion**

The customer segmentation analysis has revealed 10 distinct customer groups with varying purchasing behaviours and preferences. These insights provide a solid foundation for developing targeted marketing strategies and improving customer experiences across different segments.


TotalQuantity
TotalSpending
AvgPrice
Cluster
0
1
2


Cluster Centroids
TotalQuantity
TotalSpending
AvgPrice
Region_Europe
Region_North America
Region_South America
Features
1.1
-0.56
-0.29
1.1
-0.59
-0.23
0.17
-0.091
-0.033
-0.46
-0.58
1.7
-0.18
0.39
-0.55
0.45
0.05
-0.65
0
1
2
Clusters
1.5
1.0
0.5
0.0
−0.5