Customer Segmentation Report

1. Introduction

In this project, we aimed to understand customer behaviors by grouping them into segments based on their transaction and profile data. By using **K-Means Clustering**, we were able to classify customers into distinct groups, each showing similar purchasing patterns. This segmentation is valuable because it helps businesses tailor marketing and sales strategies to different customer types.

After analyzing the data and considering different factors, we decided to create **5 clusters**, which seemed to provide the most meaningful separation between customer behaviors.

---

2. Clustering Metrics

To evaluate the quality of the clusters, we used a couple of key metrics:

- **Davies-Bouldin Index (DBI)**: This index helps assess the separation and compactness of the clusters. The lower the value, the better the clustering. Our model returned a **DBI score of 1.24**, which indicates the clusters are reasonably well-separated, though some slight overlap may exist.
- **Silhouette Score**: This score provides a measure of how similar a customer is to others within the same cluster compared to customers in different clusters. Our **Silhouette Score** came out to be **0.56**, meaning the clusters are fairly well-formed but could benefit from further refinement to be more distinct.

---

3. Visual Representation

(Insert Scatter Plot Image Here)
Here, you can see a scatter plot visualizing how customers have been segmented into clusters. Each customer is represented as a point, and different colors indicate different clusters. The markers show the centroids, or the "average" customer of each cluster.

---

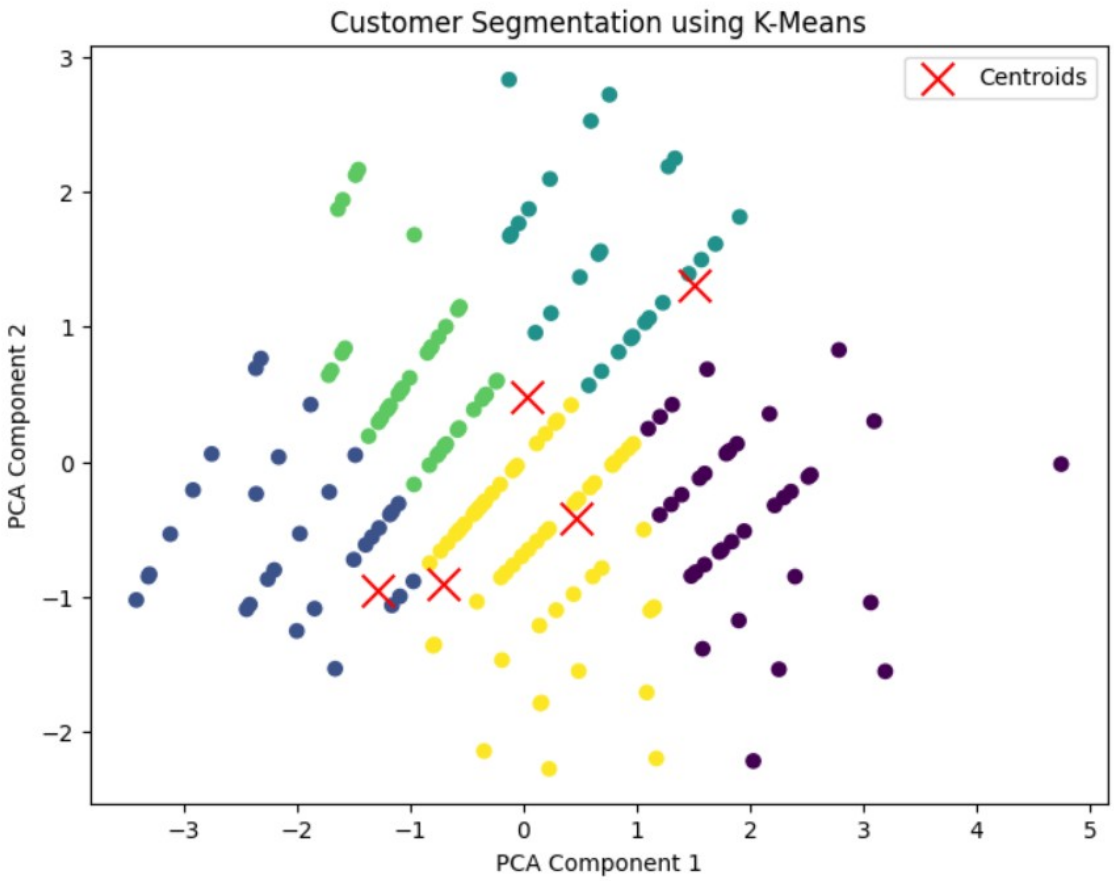
4. Cluster Insights

Each of the 5 clusters provides a unique profile of customer behavior, which can be used to inform business strategies. Below is a breakdown of the key characteristics of each cluster:

- **Cluster 1 (Low spend, frequent purchases)**:
    - Customers in this group tend to buy frequently but in smaller quantities.
    - **What this means**: These customers might be interested in frequent but low-value products. Offering loyalty rewards or incentives for repeat purchases could encourage them to buy even more.
- **Cluster 2 (High spend, low frequency)**:
    - These customers tend to make large purchases but don’t shop as often.

    - **What this means**: This group is valuable because of the high transaction amounts. To increase their frequency, personalized promotions or special offers on high-ticket items could encourage them to shop more often.
- **Cluster 3 (Moderate spend, moderate frequency)**:
    - Customers here exhibit a balanced buying behavior. They purchase with moderate frequency and spend a reasonable amount.
    - **What this means**: This is your average customer group. Cross-selling and up-selling products based on their past purchases can boost their overall spend.
- **Cluster 4 (Very low spend, very low frequency)**:
    - These customers purchase rarely and spend little when they do.
    - **What this means**: This segment could represent either dormant or new customers. Engaging these customers with special discounts, promotions, or newsletters could help convert them into more active buyers.
- **Cluster 5 (High spend, high frequency)**:
    - This group contains your most valuable customers, who make frequent and large purchases.
    - **What this means**: These customers are highly loyal and engaged. Offering VIP programs, early access to new products, or exclusive deals can further strengthen the relationship and maximize their lifetime value.

Appendix (Images)



Cluster Visualization

5. Conclusion

Accroding to my insights, The K-Means Clustering model has successfully identified 5 customer segments that exhibit distinct purchasing behaviors. While the **Davies-Bouldin Index** and **Silhouette Score** indicate that the clusters are relatively well-separated, there's always room for optimization in future iterations.

The insights from this segmentation will allow the business to implement targeted marketing strategies. Each group represents a different type of customer, and understanding these behaviors can help refine product offerings, improve customer retention, and increase revenue.