Customer Segmentation Analysis Report

This report presents an in-depth analysis of customer data using the K-Means Clustering algorithm. The project focuses on segmenting customers based on their annual income and spending score. Such segmentation is crucial for businesses to better understand their customer base and optimize their marketing and operational strategies.

1. Data Collection and Preprocessing

The dataset comprises two key features:
- **Annual Income (k$):** Represents the yearly income of customers.
- **Spending Score (1-100):** Indicates the spending behavior of customers on a scale of 1 to 100.

These features were selected for their relevance in identifying distinct customer groups. The dataset was cleaned to ensure it was free from missing values, enabling accurate clustering.

2. Optimal Number of Clusters: Elbow Method

The Elbow Method was applied to determine the optimal number of clusters. Within-Cluster Sum of Squares (WCSS) was calculated for different numbers of clusters, ranging from 1 to 10. This metric helps evaluate the compactness of clusters, with lower WCSS indicating tighter clusters.

The Elbow Point, where the reduction in WCSS slows significantly, was identified as 5 clusters. This balance ensures meaningful segmentation without overfitting the data.
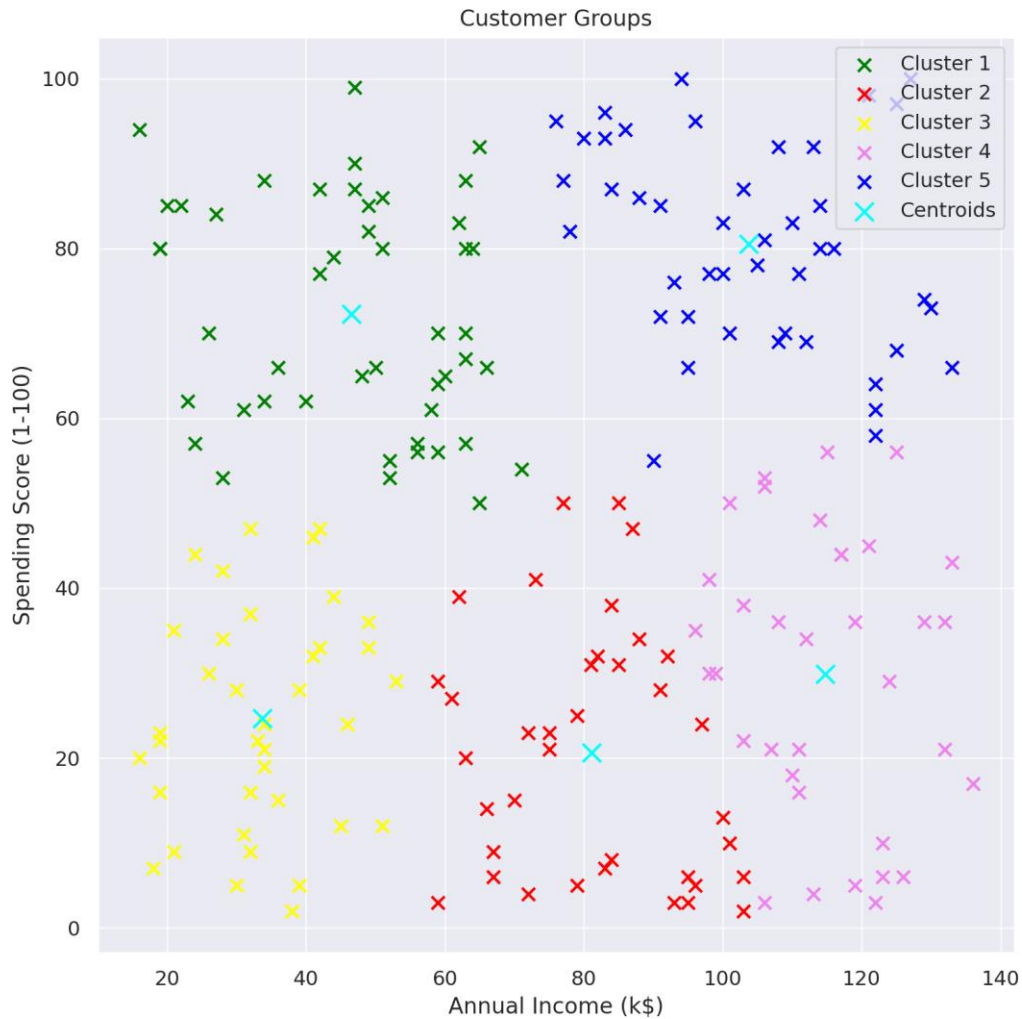
Customer Segmentation Analysis Report

The Elbow Point Graph
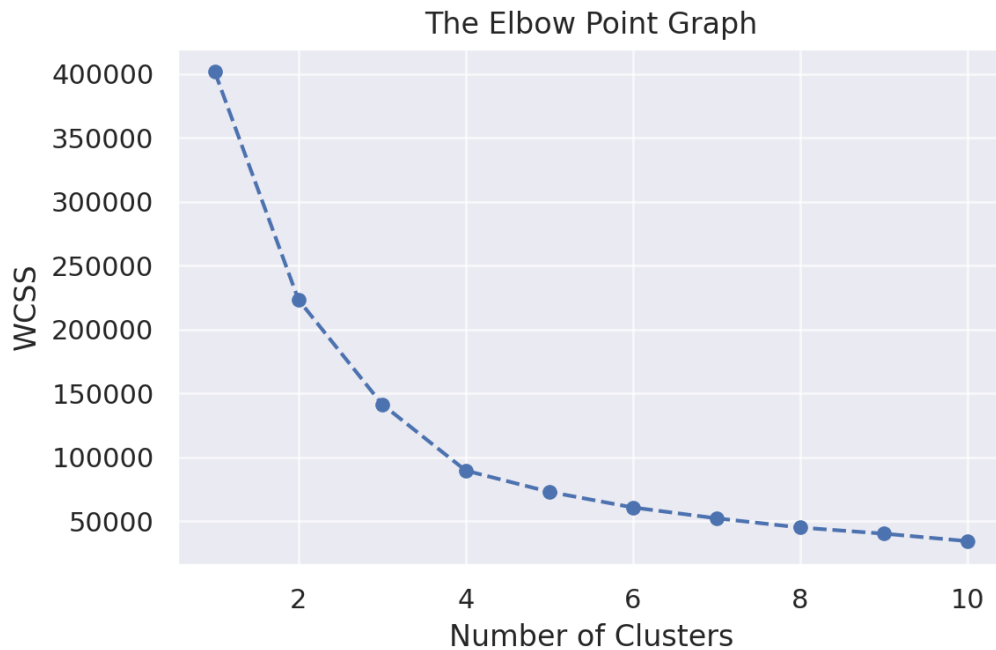
The Elbow Graph above highlights the optimal number of clusters as 5, which was used for further analysis.

3. K-Means Clustering and Visualization

The K-Means algorithm was applied to the dataset, segmenting customers into 5 distinct clusters. Each cluster represents a group of customers with similar characteristics in terms of income and spending behavior. The centroids, marked in cyan, denote the center of each cluster and help summarize the overall group behavior.

Customer Segmentation Analysis Report

Customer Groups

The above scatter plot visually depicts the clusters, where:
- Each color represents a unique customer group.
- The spatial distribution highlights the differences in income and spending behavior among the groups.

Customer Segmentation Analysis Report

4. Insights and Key Findings

The analysis reveals valuable insights into the customer base:

Cluster Descriptions:

- Cluster 1 (Green): Low-income customers with high spending scores. These customers may value premium services despite limited income, making them ideal for targeted discounts or loyalty programs.
- Cluster 2 (Red): High-income customers with high spending scores. They represent a lucrative segment with significant purchasing power. Personalized premium services and exclusive offers can strengthen relationships with this group.
- Cluster 3 (Yellow): Moderate-income customers with moderate spending scores. This group reflects average purchasing behavior and may benefit from general promotions.
- Cluster 4 (Violet): Low-income customers with low spending scores. These customers are price-sensitive and may respond to cost-effective solutions or budget-friendly options.
- Cluster 5 (Blue): High-income customers with low spending scores. This group may require engagement strategies to increase spending, such as showcasing high-value offerings or personalized product recommendations.

Practical Applications:

1. Targeted Marketing Campaigns: Segmentation enables businesses to design tailored marketing campaigns for each group, maximizing return on investment.
2. Resource Allocation: Companies can allocate resources efficiently by focusing on high-value customer segments, such as Cluster 2.
3. Customer Retention Strategies: Insights from clusters help develop strategies to retain customers with low engagement, such as Cluster 5.
4. Product Development: Understanding customer preferences across clusters aids in creating products that cater to diverse needs.