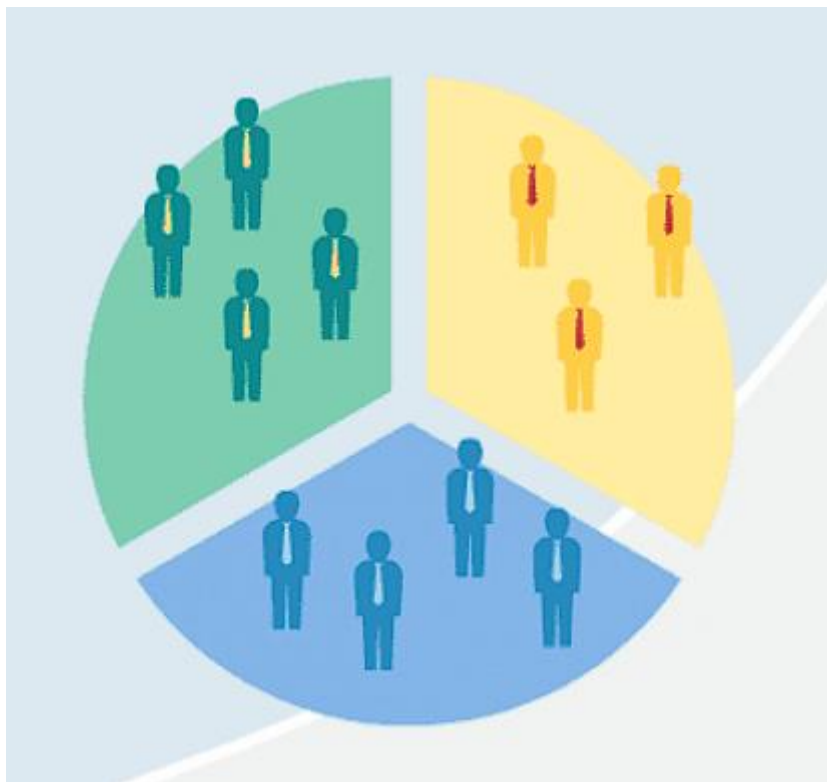


Data science Assignment-Zeotap

Customer segmentation & clustering



Customer Segmentation Report

Project Overview

The purpose of this report is to provide insights into customer behavior through segmentation. By analyzing customer and transaction data, we have grouped customers into distinct segments based on their characteristics and purchasing patterns. This segmentation will aid in targeted marketing strategies, personalized offers, and improved customer engagement.

1. Data Overview

The analysis is based on the following datasets:

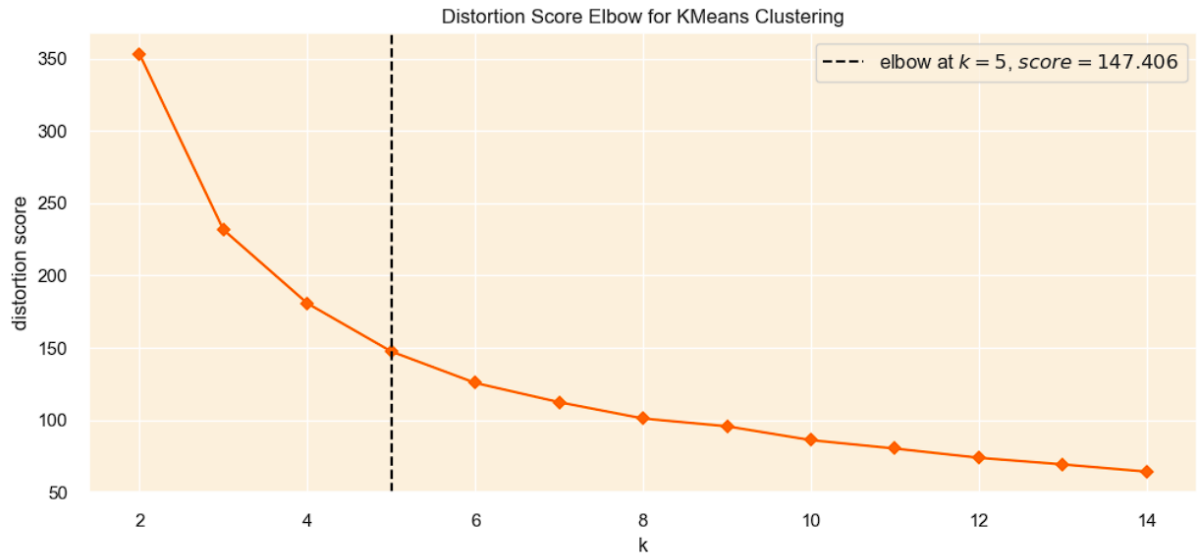
- **Customers.csv:** Contains demographic information such as age, gender, and location of customers.
- **Transactions.csv:** Includes transaction history, such as product purchases, transaction dates, and amounts.
- **Products.csv:** Details of the products purchased, including category and price.

2. Data Cleaning and Preprocessing

- **Handling Missing Values:** Missing or incomplete records were addressed by either imputing values or excluding records where necessary.
- **Feature Engineering:** Key features were created to enhance the segmentation process, such as:
 - **Total spending per customer**
 - **Frequency of purchases**
 - **Recency of purchases**
- **Data Transformation:** The data was scaled and normalized to ensure uniformity for clustering.

3. Clustering Algorithm Selection

- **Optimal Clusters:** Based on the DB Index, we determined that the ideal number of clusters is 5.



Clustering Results

The clustering algorithm has divided the customers into five distinct clusters, each characterized by their Recency, Frequency, and Monetary (RFM) values.

Cluster	Recency (Avg Days)	Frequency (Avg Orders)	Monetary (Avg Spend)	Customer Profile
0	33.94	8.28	₹6045.33	High-Value, Frequent Buyers
1	37.78	3.66	₹2055.94	Mid-Tier, Occasional Buyers
2	274.92	1.33	₹764.80	Inactive Customers
3	66.62	5.84	₹4216.83	Loyal, High-Spending Customers
4	134.52	3.90	₹3001.03	Moderate-Value Customers

Cluster Profiling & Insights

Cluster 0: High-Value, Frequent Buyers

- **Characteristics:** Customers in this group have a low recency of around 34 days, high purchase frequency of about 8.28 orders, and the highest monetary contribution of approximately ₹6045.
- **Strategy:** Retain them with VIP programs, exclusive deals, and priority customer service.

Cluster 1: Mid-Tier, Occasional Buyers

- **Characteristics:** These customers have a moderate recency of about 38 days, lower purchase frequency of 3.66 orders, and moderate spending at ₹2055.
- **Strategy:** Encourage repeat purchases through personalized offers, discounts, and reminder emails.

Cluster 2: Inactive Customers

- **Characteristics:** This segment has the highest recency of about 275 days, very low purchase frequency of 1.33 orders, and the lowest monetary contribution at ₹764.
- **Strategy:** Re-engage them through win-back campaigns, special discounts, and feedback collection to understand why they left.

Cluster 3: Loyal, High-Spending Customers

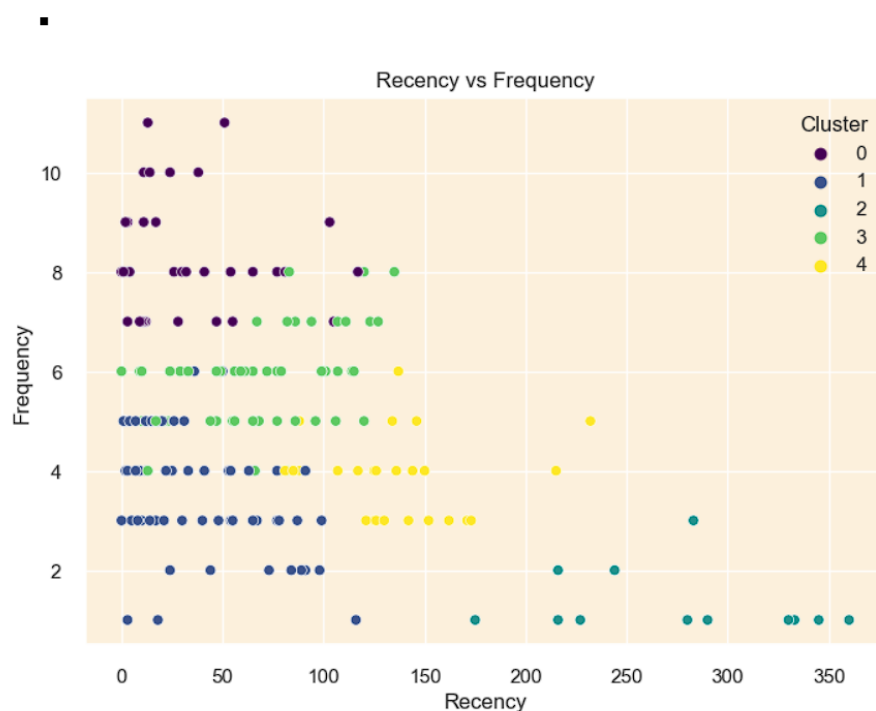
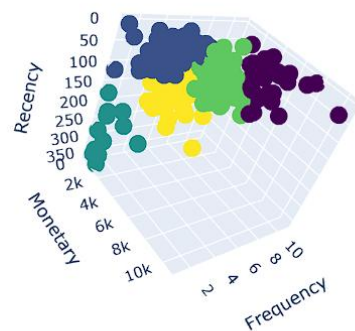
- **Characteristics:** Customers in this cluster have a recency of around 66 days, high purchase frequency of 5.84 orders, and a high monetary contribution of ₹4216.
- **Strategy:** Strengthen engagement through referral programs, exclusive product launches, and personalized recommendations.

Cluster 4: Moderate-Value Customers

- **Characteristics:** Customers in this group have a mid-range recency of approximately 134 days, moderate purchase frequency of 3.90 orders, and a reasonable spending level of ₹3001.
- **Strategy:** Improve retention by offering tailored promotions, bundle deals, and incentives for higher-order values.

Clustering Performance Evaluation

- **Davies-Bouldin Index (DBI): 0.6591**
 - A lower DBI indicates well-separated clusters. Since this value is relatively low, the clusters are reasonably well-defined.
- **Silhouette Score: 0.3125**
 - This score ranges from -1 to 1. A value of 0.3 suggests moderate clustering quality, indicating some overlap between clusters.



Final Recommendations

1. **Focus on Clusters 0 & 3** → These are high-value and repeat customers. Keep them engaged with VIP programs, rewards, and special deals.
2. **Convert Cluster 1 into Cluster 3** → This group has potential; provide personalized offers and engagement strategies to increase purchase frequency.
3. **Re-engage Cluster 2 (Lost Customers)** → Run win-back campaigns, send targeted discounts, and gather feedback to understand their inactivity.
4. **Monitor Cluster Movement** → Track how customers transition between clusters over time to refine marketing and retention strategies.
5. **Optimize Marketing Efforts** → Allocate more resources to clusters with the highest return on investment while minimizing acquisition costs for low-value segments.