

Customer Segmentation Analysis Report

Executive Summary

This report details a customer segmentation analysis using transaction data from an online retail store. The objective was to use unsupervised machine learning to group customers into distinct segments based on their purchasing behavior. By applying data preparation techniques, clustering algorithms, and visualizations, we successfully identified four key customer segments. The findings provide a clear foundation for developing targeted marketing campaigns and customer retention strategies.

Methodology and Key Findings

The analysis was performed using a structured approach to ensure reliable and interpretable results:

- Data Preparation:** The raw dataset was cleaned to address missing CustomerID values and outliers in total spending. Key features were then engineered for each customer, including their total number of orders (NumOrders), total spend (TotalSpend), total quantity of items purchased (TotalQuantity), and average unit price (AvgUnitPrice). These features were then scaled using StandardScaler to ensure they were all given equal importance during the clustering process.
- Clustering Models:** The K-Means clustering algorithm was applied to the prepared data. The **Elbow Method** and **Silhouette Score** were used to determine the optimal number of clusters, with **k=4** providing a good balance between model performance and the interpretability of the customer segments.
- Dimensionality Reduction: Principal Component Analysis (PCA)** was used to reduce the data from four dimensions to two. This allowed for the creation of a clear, two-dimensional scatter plot to visualize the customer clusters.

The analysis successfully segmented the customer base into four distinct groups, each with unique purchasing characteristics.

Cluster Profiles and Actionable Business Insights

The four clusters represent distinct customer segments, each requiring a tailored business strategy. The following table summarizes the characteristics of each cluster and provides actionable business insights.

Cluster	Profile Description	Key Metrics (Average)	Actionable Business Insights
0	Low-Value / Regular Customers	- NumOrders: 2.5 - TotalSpend: \$604.89	This is the largest segment. Focus on targeted upselling and cross-selling campaigns to increase their average order value.
1	High-Value Customers	- NumOrders: 23.5 - TotalSpend: \$8,746.78	This is the most profitable and loyal segment. Offer VIP loyalty programs , personalized recommendations, and exclusive access to new products to ensure retention.

Cluster	Profile Description	Key Metrics (Average)	Actionable Business Insights
2	Frequent / Medium-Value Customers	- NumOrders: 9.5 - TotalSpend: \$3,039.76	These customers are frequent buyers with a medium spend. Develop retention campaigns with special offers to prevent them from churning.
3	High-Unit-Price Customers	- NumOrders: 5.0 - AvgUnitPrice: \$8,055.78	This is a small, unique segment that makes high-value, high-ticket purchases. Promote new premium products and luxury items to this group.