Customer Segmentation Model Using K-Means Clustering

This report outlines a customer segmentation model utilizing K-Means clustering to categorize customers based on their annual income and spending scores.

Methodology

To determine the optimal number of clusters, I employed the Within-Cluster Sum of Squares (WCSS) method, which measures the distance between data points and the centroid of each cluster.

Aim: To effectively group customers who shop during the month based on their annual income and spending scores.

Insights:

200 customers

Female112 (56%)

Male 88 (44%)

Customers with Annual Income of $10k - $40k:

This segment generally exhibits lower monthly expenditures ranging from (1 to 40) in Spending score.

However, within this income bracket, there is a notable subset of customers who demonstrate high spending scores (60-100), indicating that while their overall income is low, they may prioritize certain purchases that reflect their preferences or needs.

Customers with Annual Income of $40k - $70k:

Customers in this income range show a spending score primarily between 40 and 60. This suggests a moderate level of spending, indicating that they are more willing to invest in products or services compared to the lower-income bracket, but still exercise some level of budgetary restraint.

Customers with Annual Income of $70k and Above:

This high-income segment presents varied spending behaviour.

Some customers in this group exhibit low spending scores (0-40), which may indicate a cautious approach to spending or a preference for saving.

Conversely, customers within this income range who have high spending scores (60 and above). This suggests that a portion of some high-income individuals are willing to spend significantly on discretionary items, luxury goods, or services.

Clustering Applications

Retail and Sales:

Recommendations for products can be tailored based on customers' shopping habits. Businesses can reach out to specific segments with new product alerts, offers, or discounts.

Media and Entertainment:

Similar clustering methods can be applied to recommend movies based on viewing preferences (e.g., comedy, horror), alerting customers when new releases align with their interests.

Banking:

Banks can analyze transaction patterns to detect fraud based on spending behaviour. They can also identify customers who might be eligible for upgrades based on their spending patterns.

Telecommunications:

Telecom companies can assess customer spending on data subscriptions and calls to tailor specific offers that match their usage patterns.

Market Research:

Cluster analysis is commonly used to segment customers based on buying behaviours, demographics, or other characteristics, enabling targeted marketing campaigns and product development.

Social Network Analysis:

In social network analysis, clustering helps identify subgroups within larger networks based on shared connections and characteristics.

Biology and Medicine:

This technique is also used to identify genetic markers associated with diseases or group patients with similar clinical characteristics, aiding in diagnosis and treatment.

Conclusion

Overall, implementing K-Means clustering can significantly enhance revenue and customer satisfaction by providing insights into customer behaviours and preferences.

Cluster Analysis Results

The results of the cluster analysis reveal distinct groupings among customers. The elbow plot indicates that the optimal number of clusters is 5, as there is no significant drop in WCSS values beyond this point.

Yellow Cluster: Customers with low income and high spending scores.

Grey Cluster: Customers with low income and low spending scores.

Green Cluster: Customers with average spending scores and salary ranges.

Blue Cluster: Customers with high salaries and high spending scores.

Additional Green Cluster: Customers with high salaries but low spending scores.

These insights can inform strategies for targeted marketing and product offerings tailored to each customer segment.

Recommendations

Tailored Promotions:

For High Spending, High-Income Customers (Blue Cluster): Offer exclusive promotions on luxury items, VIP shopping experiences, or early access to new collections to encourage continued spending.

For Low Income, High Spending Score Customers (grey and red Cluster): Provide discounts on essential items or loyalty rewards for frequent purchases to enhance their shopping experience while being budget-conscious.

Product Recommendations:

Utilize the clustering insights to recommend products based on spending patterns. For instance, if customers in a specific cluster frequently buy beauty products, highlight related products or complementary items during their next visit.

Personalized Marketing Campaigns:

Develop targeted marketing campaigns based on customer segments. For example, send tailored emails or notifications about new arrivals and promotions that align with their previous purchasing behaviour before salary is paid.

Customer Engagement Events:

Organize exclusive events for different customer segments to foster a community feeling.

Feedback Mechanisms:

Implement feedback channels to understand the needs and preferences of different customer segments better. This can help refine offerings and improve customer satisfaction.

Conclusion:

By leveraging the insights from K-Means clustering, the mall can enhance customer satisfaction, increase revenue through targeted marketing strategies, and foster long-term loyalty among its diverse customer base. Continued analysis and adaptation to changing customer behaviours will be crucial for maintaining relevance and appeal in a competitive retail environment.