Prepared By: Dibya Pandey Adhikari Date: 06/08/2023

Presented To:

Devin Fensterheim

University Of Maryland Baltimore County

Business Problem

Customer segmentation involves dividing customers into distinct groups based on a range of factors, including gender, age, location, and purchasing behavior. By categorizing customers into segments, businesses can gain a comprehensive understanding of the characteristics, preferences, and behaviors exhibited within each group.This understanding enables companies to provide personalized marketing services that cater to the specific needs and preferences of each customer segment. By delivering relevant and customized product information, companies can enhance their marketing effectiveness, engage customers more effectively, and increase the likelihood of generating positive responses.Customer segmentation enables businesses to move away from a one-size-fits-all approach and instead deliver targeted and personalized marketing efforts. By understanding customers' needs and preferences on a granular level, businesses can foster stronger customer relationships, drive growth, and gain a competitive advantage in the marketplace.

The objective of this study is to utilize an unsupervised learning model to categorize customers based on various characteristics such as gender, age, geography, and buying habits. This approach can be a valuable tool for organizations to enhance customer satisfaction, increase long-term retention, and drive revenue growth. By employing clustering algorithms, we aim to identify distinct customer segments that exhibit similar interests and behaviors.

An example of how customer segmentation can benefit organizations is illustrated by a winter jacket retailer. Through the process of customer segmentation, the retailer may identify a specific group of customers who place a high value on sustainability. This segment of environmentally conscious customers may prefer products that align with their values.

With this valuable insight, the retailer can respond by developing a new line of winter jackets made from recycled materials. By catering to the preferences and interests of this customer segment, the retailer can effectively target and attract their attention. This approach not only meets the specific needs of environmentally conscious customers but also differentiates the retailer from competitors in the market.

Followed approach and results:

Introduction

This report summarizes the outcomes of the customer segmentation analysis, which aimed to categorize customers based on their transactional behavior for targeted marketing strategies. The analysis involved data cleaning, exploratory data analysis, feature engineering, unsupervised machine learning techniques and model evaluation..

Data Cleaning and Preparation

The first stage of the analysis focused on cleaning and preparing the dataset. This involved addressing missing values, eliminating duplicates, standardizing data formats, and adding relevant columns for subsequent processes. These data cleansing steps ensured the dataset was suitable for further analysis and maintained data consistency.

Exploratory Data Analysis (EDA)

We performed Exploratory Data Analysis (EDA) to gain a deeper understanding of the dataset and the customer transactions it contained. This involved using descriptive statistics, data visualization techniques, and statistical tests to uncover valuable insights. Through EDA, we were able to explore customer demographics, transaction patterns, purchase frequencies, and monetary values, thereby obtaining a comprehensive understanding of the dataset's characteristics.

Feature Engineering

Feature engineering was conducted to derive meaningful features from the existing dataset. This process involved transforming or combining variables to extract additional insights. For instance, new features like "Recency," "Frequency," and "Monetary" were created using the transactional data. These engineered features were specifically designed to facilitate RFM (Recency, Frequency, Monetary) analysis, which is a widely used approach for customer segmentation and understanding customer behavior.

Unsupervised Machine Learning - Customer Segmentation

We utilized unsupervised machine learning techniques to segment customers based on their transactional behavior. Specifically, the K-means clustering algorithm was applied. This algorithm groups similar customers together based on their RFM scores, which were calculated using the engineered features. From K-means clustering we are able to group customers with similar transactional patterns and behaviors into distinct segments, allowing for a better understanding of their characteristics and enabling targeted marketing strategies.
We used silhouette score or the elbow method to determine the optimal number of clusters.In this analysis, we determined the cluster numbers as five, indicating that the dataset can be effectively segmented into five distinct customer groups with similar transactional characteristics. Each of these customer segments represents a unique cluster that can be targeted with tailored marketing strategies and personalized approaches.

Model Evaluation

We used the silhouette score for the model evaluation which gave the result 0.8710.After the evaluation we did the boxplot and histogram as boxplot help us understand the distribution of customers on each clusters according to their purchase characteristics and histogram helps us identify distribution of the overall data.We got the result of data as distributed to right. So, we

used median to identify the central tendency of each cluster.We got the range of each cluster as:

	cluster	median	min	max
0	0	132.0	111	155
1	1	344.0	312	355
2	2	234.0	212	255
3	3	445.0	412	455
4	4	555.0	533	555

Customer Segment Descriptions

Based on the model result, we segmented the customers based on recency score to focus on the losing customers and frequency score for the rewarding customers.

Recency score measures the time since the customer's last purchase. Customers who have a low Recency score (i.e., they made a purchase recently) are considered more engaged and active, while customers with a high Recency score (i.e., it has been a long time since their last purchase) are considered at risk of churn or disengagement).

Rewarding customers are those who have Repeat purchase behavior. From the frequency of their buying we have segmented them to rewarding and non rewarding segments in that above segmented data. We now have the rewarding and non rewarding customers from different segments defined.So that we can reward them by giving discounts, exclusive offers so they can invest more in our business.

Implications and Recommendations

The customer segmentation analysis provides actionable insights for targeted marketing strategies. By understanding the distinct characteristics and behaviors of each segment, we can tailor its marketing efforts and services to better meet the needs and preferences of different customer groups. The analysis can guide the following actions:

-Many businesses have loyalty programs to incentivize repeat purchases. Frequency of transactions is often a core metric in these programs, as it directly measures customer engagement and participation. High-frequency customers can earn more rewards, benefits, or exclusive offers, which can further reinforce their loyalty and encourage them to continue making frequent purchases.

-Implementing strategies such as personalized offers, reactivation campaigns, or targeted communication, can increase the chances of retaining and re-engaging losing customers under "Low Value" and "At Risk" segments.

Conclusion

The customer segmentation analysis, which involved data cleaning and preparation, EDA, feature engineering, unsupervised machine learning and model evaluation, has yielded valuable insights by identifying distinct customer segments. This segmentation enabled us to gain a comprehensive understanding of customer behaviors, preferences, and needs. By leveraging these insights, we can optimize our marketing strategies, improve customer satisfaction, and drive revenue growth.

However, it is crucial to acknowledge that customer behaviors and preferences are not static and may change over time. Therefore, it is recommended to periodically reassess and refine the segmentation model to ensure its continued effectiveness. By regularly updating the segmentation approach, we can adapt to evolving customer trends and maintain the relevance of our marketing strategies. This ongoing assessment and refinement process will allow us to stay responsive to the changing needs and preferences of our customers.