

Customer Segmentation: Market Data Analysis

Billy L. Andrews

Bellevue University

### Abstract

Making commercial decisions requires careful consideration of customer segmentation. In the intense world of commerce, it is crucial to discover potential customers and meet their demands. These tasks must be completed at the appropriate time and in the appropriate way. In this white paper, a customer segmentation approach is offered. The most effective method of sorting them is clustering, and when evaluating clustering methods, K-Means algorithm is the most effective and user-friendly.

### Business Problem

*Customer segmentation* is the market divided into various unique consumer groups with related characteristics. Market segmentation is a valuable tool for identifying and satisfying customer needs. Market Basket Analysis was conducted using K-Means Clustering Algorithm and Unsupervised Machine Learning Techniques. To identify the target customers that can be quickly converged, a market basket analysis is done. To enable the marketing staff to prepare a plan for promoting new products to target customers who share their interests. Many businesses must segment their customer bases to generate future projections, create various marketing plans, or implement customer-specific marketing initiatives.

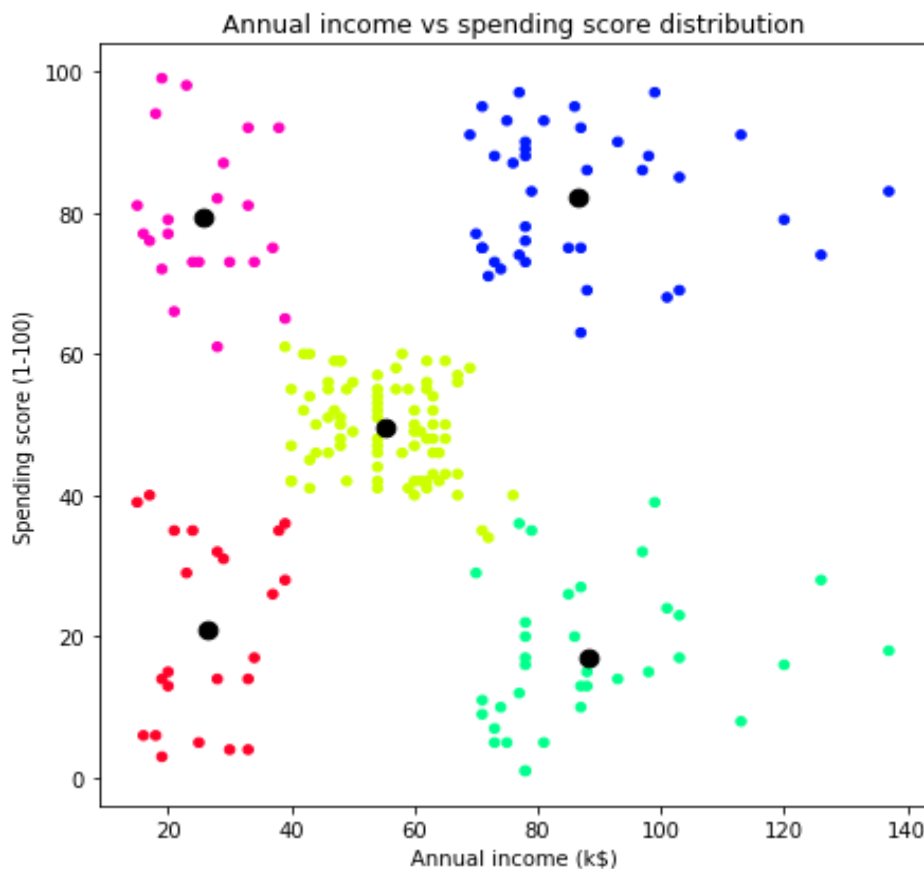
### Data Explanation

The data set is of 200 mall customers. It includes demographic variables such as Customer ID, Age, Gender, Annual Income, and Spending Score. The variables Customer ID and Gender were removed as they were determined not to be helpful in customer segmentation. The mean of the variables is Age 38.85, Annual Income of \$60.56K, and Spending Score of 50.20. The minimum age of the customer in the data is 18 years, and the maximum age is 70. The minimum annual income of a customer is 15k\$, and the maximum is 137k\$. A Spending

Score is assigned to the customer based on defined parameters like customer behavior and purchasing data. Here the minimum spending score assigned is one, and the maximum ranges to 99. Between 40 and 60 has the highest population of Spending Scores. The Ages between 20 and 40 have the highest Spending Scores. The Ages between 30 and 40 have the largest population. The amount between \$60K and \$80k has the highest population of annual income. The Ages between 30 and 50 have the highest annual income. None of the data types had to be changed, and there were no missing values to populate.

### **Methods**

K-Means Clustering algorithm used to cluster data. To implement K-Means clustering, we need to look at the Elbow Method. The Elbow method is used within cluster analysis to determine the appropriate number of clusters in a dataset. It is clear from the figure that we should take the number of clusters equal to 5, as the slope of the curve is not steep enough after it. The data(clusters) are plotted on a spending score Vs. Annual income curve.



### Analysis and Conclusion

This Clustering Analysis gives us a clear insight into the different segments of customers in the mall. There are five segments of Customers when using Annual Income and Spending Scores. It determined that they were the best variables to use to establish customer segments in a mall. Once we have the results, we can use different marketing strategies to capitalize on the spending scores of the customer in the mall. The marketing department no longer has to use a one size fits all campaign for all of the customers. The breakdown of the clusters into various customer personas is as follows:

**Cluster 1:** Customers with a lower annual income and lower spending score. These customers can be identified as “The Budget-Minded Customer.”

**Cluster 2:** Customers with a higher annual income and lower spending score. They are “More Marketing Focus” customers.

**Cluster 3:** Customers with a middle annual income and middle spending score are the “Average Joes.”

**Cluster 4:** Customers with a lower income and higher spending score - “Carefree shoppers.”

**Cluster 5:** Customers with a high income and high spending score - “Frequent Shoppers.”

### **Assumptions**

- One size fits all marketing campaign is not as effective and customer targeted campaigns.
- Customers cannot be identified by “gut feel” or experience. Clean and solid data through clustering will synch up the customer targets.

### **Challenges and Issues**

Data quality is one of the main problems with customer segmentation. Usually, inaccurate data in the source systems will lead to poor grouping. Properties of the segments will be erroneous if not preserved, making the material less helpful. Users will only use the segments if they are comfortable with the data quality. Issues with data quality also result from a lack of upkeep and routine cleaning to guarantee accuracy. The fact that business users need to be more accurate and understand segmentation definitions is another prevalent issue with consumer segmentation. The data within the segments represent the categorizations and when to use each segmentation for the appropriate analysis scenarios.

**Future Uses/Additional Applications**

- Understanding the mall's customers will allow the shoppers to increase brand awareness. For example, this could lead to higher email open rates on target products for customers. Customer segmentation allows for the correct information to be delivered to audiences. As a result, there is increasing awareness, engagement, and sales.
- Segmentation has been proven to reduce customer churn because businesses understand the wants of their customers. For example, if businesses send personalized email campaigns that emphasize consumers' favorite products, they are more likely to be active. Segmentation allows for higher ROI and customer lifetime value.

**Recommendations**

- Once subgroups are identified, businesses can continue to do so for subgroups further along the chain. Every sub-level presents unique marketing options that can increase revenue.
- They must also be able to keep track of and assess progress. By assessing how well they are currently addressing the demands of segments, segment evaluation assists in prioritizing and planning developments.
- The marketing message to a specific group of customers and customer segmentation helps any organization increase its profitability. Building segments implies that we get a more nuanced and detailed grasp of the results immediately. Utilize these client segmentation best practices to provide the data needed to enhance the performance values aiming.

**Implementation Plan**

- Concentrate on the most significant clients. It is crucial to concentrate on the consumer demographics that are most beneficial to the company.
- Create focused marketing campaigns that speak to these clients' needs and desires because they are the most likely to purchase.
- Test and try new things. Try various techniques and tactics to implement marketing techniques. Monitor outcomes. Finally, it is critical to monitor progress and change as necessary.

**Ethical Assessment**

Customer data needs to be applied for commercial purposes only. Manipulating customer data to achieve corporate objectives. Using customer data without transparency: Failed transparency is worse than no transparency. It exposes customers to abuse from various hostile actors. Data mining: Without permission or knowledge, data mining is obtaining and analyzing customer activity from various sources for personal or financial gain. Invasion of private rights: Federal law protects consumer privacy in the United States, and several state-level privacy regulations exist.

**References**

S. Koul and T. M. Philip, "Customer Segmentation Techniques on E-Commerce," 2021 International Conference on Advance Computing and Innovative Technologies in Engineering (ICACITE), 2021, pp. 135-138, doi: 10.1109/ICACITE51222.2021.9404659.

E. Altıparmakoglu, "Big Problems of Companies: Customer Segmentation," 2021 [Big Problems of Companies: Customer Segmentation | by Ece Altıparmakoglu | Medium](#)

Ishantha, Asith. (2021). MALL CUSTOMER SEGMENTATION USING CLUSTERING ALGORITHM.

Sagar, Abhinav. (2019) Customer Segmentation Using K Means Clustering. [Customer Segmentation Using K Means Clustering | by Abhinav Sagar | Towards Data Science](#)

Burke, Stephanie.(2022). The ethics of customer behavioral tracking. [Ethical Concerns Related to Digital Behavior Tracking | Lytics](#)

Wong, Sammi. (2022) The Importance of Customer Segmentation. [Importance of Customer Segmentation — 3 tier logic \(3tl.com\)](#)

Navaneetha. (2017). 1) Best Practices of Customer Segmentation. [10 Best Practices on Customer Segmentation \(aeroleads.com\)](#)