

**Project Title:** Customer Segmentation for a Retail Store

**Project Manager:** Ritu

**Registration no:** 12223863

**Program:** B. Tech. CSE

**Course:** Python, Data Science & Machine Learning  
Integrated – Hybrid

**College:** Lovely Professional University

**Submitting To:** Cipher School

**Start Date:** 10-07-2024

**End Date:** 25-07-2024

**Objective:** To segment customers into distinct groups based on their purchasing behaviour.

**Scope:** Data cleaning, EDA, customer segmentation using K-Means, visualization using Matplotlib and Power BI.

# Table of Contents

- 1) Problem Statement
- 2) Abstract
- 3) Introduction
- 4) Insights
- 5) Conclusion
- 6) References

# **PROBLEM STATEMENT**

To develop a machine learning model that can identify and group customers based on their shopping patterns and behavior. The model should use customer data such as age, gender, annual income, and spending score to segment customers into different groups. This segmentation will enable mall owners to target their marketing efforts more effectively, improve customer experience, and optimize their business operations.

## **ABSTRACT**

Mall customer segmentation refers to the process of grouping customers based on their shopping patterns and behavior in a mall. This is achieved by analyzing customer data, including demographic information such as age, gender, and income, and behavioral data such as purchase history and frequency. The purpose of mall customer segmentation is to enable mall owners to better understand their customers and provide targeted marketing efforts that cater to their needs and preferences. By segmenting customers, mall owners can identify trends, patterns, and preferences that can be used to create more personalized shopping experiences and improve customer satisfaction. Additionally, mall owners can optimize their business operations by targeting specific customer segments with promotions, discounts, and other incentives. Furthermore, mall customer segmentation can also help in identifying high-value customers who contribute significantly to the mall's revenue, allowing mall owners to focus on retaining and nurturing these customers. It can also assist in identifying customer groups that require more attention and support, such as first-time visitors or customers who have experienced issues during their shopping experience. Overall, mall customer segmentation is an essential tool in the retail industry that allows mall owners to better understand their customers, optimize their operations, and provide personalized experiences that cater to the needs and preferences of their customers.

# INTRODUCTION

Implemented a Mall Customer Segmentation based Machine Learning model which focuses on different clustering methods. The clustering methods implemented are:

- 1) K means clustering
- 2) Mean-Shift Clustering
- 3) Hierarchical Clustering

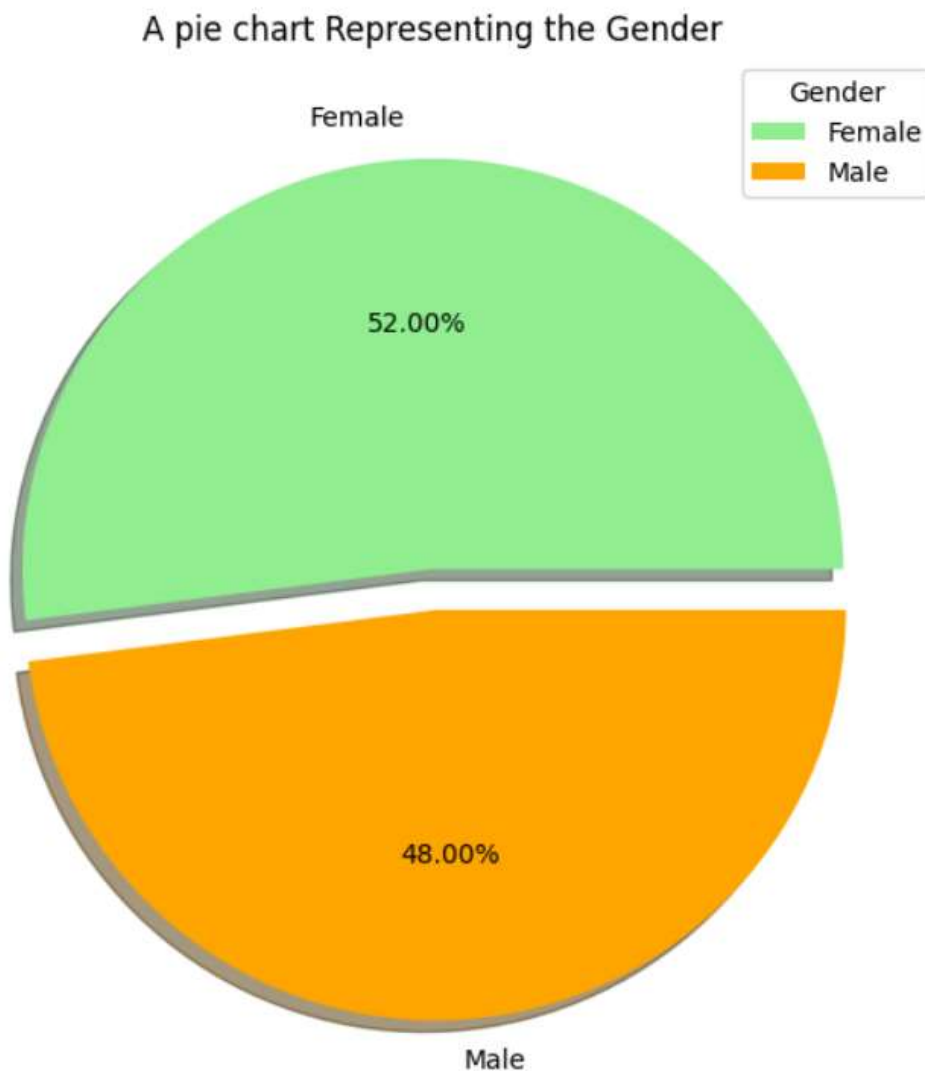
Various types of graphs are plotted for comparing the results, namely:

- 1) Normal Scatter Plot
- 2) Plot with Clusters (K-means)
- 3) Distplot
- 4) Pie Chart
- 5) Bar Graph
- 6) Pair plot
- 7) Heatmap
- 8) Scatter Matrix
- 9) Elbow Method Plot

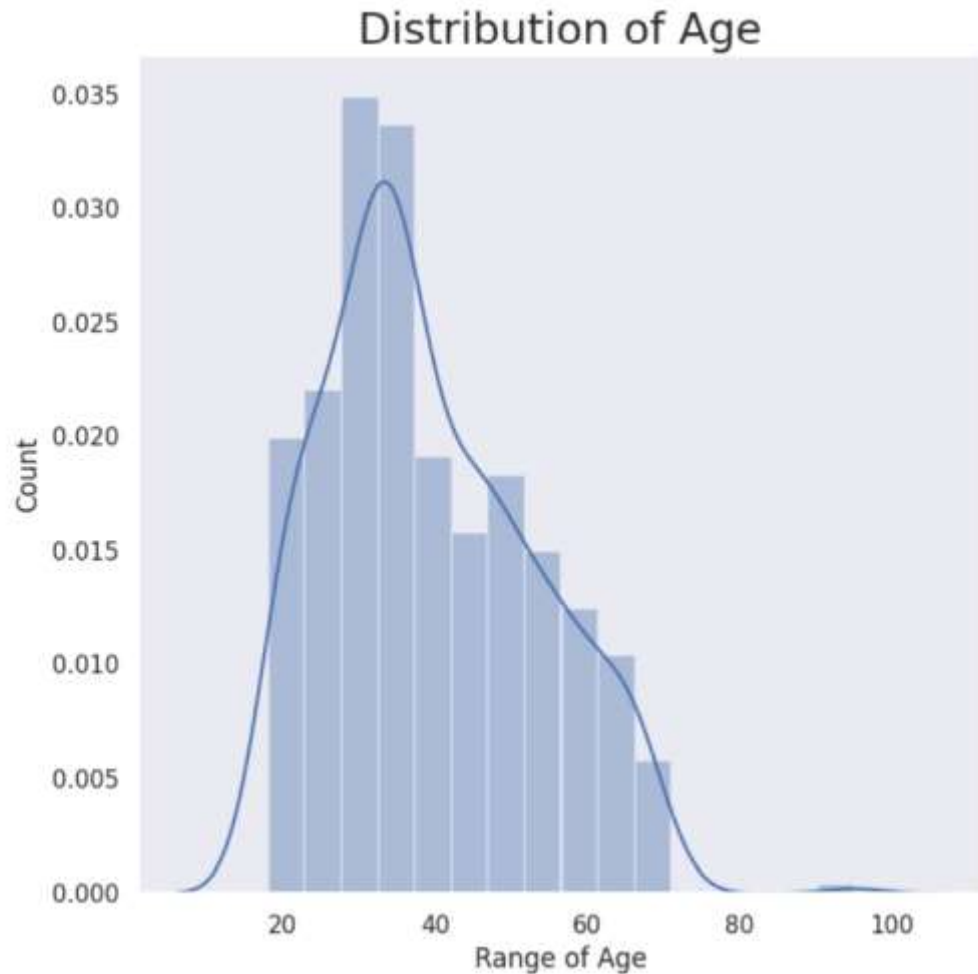
Based on the results obtained through these plots, we have compared the two clustering methods to find out which clustering method is more effective.

## INSIGHTS

- **There are more females customers than male customers. We should introduce more ways to make these customers permanent and also to engage new customers.**



- **Maximum customers are from age group 20-40 which could be used strategically to attract customers from other age groups also.**



## **CONCLUSION**

The analysis has identified different customer segments based on factors such as demographics, spending habits, purchasing behavior etc. The conclusion of the analysis is recommendations for the mall or retailers within the mall on how to better cater to the needs and preferences of different customer segments. For example, the analysis has found that one segment of customers is more price-sensitive and prefers discounts and promotions, while another segment values high-quality products and is willing to pay a premium price. Based on these findings, the mall or retailers can tailor their marketing and merchandising strategies to appeal to different customer segments and ultimately increase sales and customer satisfaction.

## **FUTURE SCOPE**

Possible area for exploration/ future scope:

1. Feature engineering: In this implementation, we have used only two features (Annual Income and Spending Score) to perform customer segmentation. However, there are several other features in the dataset that can be explored to identify different segments of customers.
2. Visualization techniques: In the current implementation, we have used scatterplots and pair plots to visualize the clusters. However, there are several other visualization techniques such as t-SNE, PCA, etc. that can be explored to gain a better understanding of the clusters.
3. Association rule mining: Association rule mining can be applied to the customer data to identify patterns in customer behavior, such as which products are frequently purchased together or which customers are likely to purchase a particular product.
4. Predictive analytics: Machine learning algorithms such as decision trees, random forests, etc. can be applied to the customer data to predict customer behavior, such as which customers are likely to make a purchase or which customers are likely to churn.
5. Customer feedback analysis: Customer feedback data can be analyzed to gain insights into customer preferences, satisfaction levels, etc. This can help businesses improve their products and services and provide a better customer experience.

Overall, there are several possibilities for future exploration in mall customer segmentation, and businesses can benefit from implementing these techniques to gain a better understanding of their customers and improve their operations.

## **REFERENCES**

### **Books**

- 1) Machine Learning (Authors: S. Sridhar, M. Vijayalakshmi)

2) Artificial Intelligence and Machine Learning (Author: Vinod Chandra)

Online Resources:

- 1) <https://www.kaggle.com/datasets>
- 2) <https://www.analyticsvidhya.com/blog/2021/11/understanding-k-means-clustering-in-machine-learning-with-examples/>