SUMMER TRAINING PROJECT REPORT

# TITLE OF THE PROJECT – MALL CUSTOMERS SEGMENTATION

Submitted by :

DIVESH THAKUR

Registration No. : 12214492

DOMAIN : DATA SCIENCE

COURSE : Python , Data Science & Machine Learning Integrated - Hybrid

Under the Guidance of

Mr. ABHISHEK RAJ

Lovely School of  Computer Science Enginnering

Lovely Professional University, Phagwara

# INDEX

# INTRODUCTION

**Objective and Use Case**

This report presents insights and recommendations derived from the analysis of customer data collected from a mall. The objective was to segment customers based on their demographics and shopping behaviors to formulate targeted marketing strategies.

**Overview of the Dataset**

| | CustomerID | Genre | Age | Annual Income (k$) | Spending Score (1-100) |
|---|---|---|---|---|---|
| 0 | 1 | Male | 19 | 15 | 39 |
| 1 | 2 | Male | 21 | 15 | 81 |
| 2 | 3 | Female | 20 | 16 | 6 |
| 3 | 4 | Female | 23 | 16 | 77 |
| 4 | 5 | Female | 31 | 17 | 40 |
| 5 | 6 | Female | 22 | 17 | 76 |
| 6 | 7 | Female | 35 | 18 | 6 |
| 7 | 8 | Female | 23 | 18 | 94 |
| 8 | 9 | Male | 64 | 19 | 3 |
| 9 | 10 | Female | 30 | 19 | 72 |

**SOURCE** : Mall_Customers.csv  dataset is taken from [www.kaggle.com](www.kaggle.com)

The dataset consists of information about 200 customers, including:

- **CustomerID**: A unique identifier for each customer.
- **Gender**: The gender of the customer (Male/Female).
- **Age**: The age of the customer.
- **Annual Income**: The annual income of the customer in thousands.
- **Spending Score**: A score assigned by the mall based on customer behavior and spending nature (0-100).

# METHODOLOGY

**Data Collection and Cleaning**

- **The dataset was initially loaded and cleaned to handle missing values and encode categorical variables.**
- **Categorical variable 'Genre' was renamed to 'Gender' and encoded as 0 for Male and 1 for Female.**

**Exploratory Data Analysis (EDA)**

- Descriptive statistics were calculated to understand the distribution of age, annual income, and spending score.
- Visualizations including histograms and scatter plots were created to explore relationships and distributions.

**Customer Segmentation**

- Features were standardized before applying K-Means clustering to identify distinct customer segments.
- Five clusters were chosen based on silhouette scores and business context.

# FINDINGS

**Customer Segments Identified**

**Cluster 0: High Spenders**

- **Characteristics**: High annual income and high spending score.
- **Recommendations**: Focus on premium products and exclusive offers to maintain loyalty.

**Cluster 1: Moderate Spenders**

- **Characteristics**: Moderate income and moderate spending score.
- **Recommendations**: Develop personalized marketing campaigns to increase spending frequency.

### Cluster 2: Low Spenders

- **Characteristics**: Low income and low spending score.
- **Recommendations**: Introduce cost-effective products and discounts to stimulate spending.

### Cluster 3: Young Shoppers

- **Characteristics**: Younger age group with varying income levels and moderate spending.
- **Recommendations**: Engage with social media and digital marketing to reach younger demographics effectively.

### Cluster 4: Older Shoppers

- **Characteristics**: Older age group with stable income and conservative spending habits.
- **Recommendations**: Offer loyalty programs and personalized discounts to encourage repeat visits.

# RECOMMENDATIONS

**Marketing Strategies by Segment**

### Cluster 0: High Spenders

- Leverage data analytics to personalize shopping experiences.
- Introduce VIP services and exclusive previews.

### Cluster 1: Moderate Spenders

- Implement loyalty programs with rewards for frequent shoppers.
- Use targeted email campaigns based on purchase history.

**Cluster 2: Low Spenders**

- Create value bundles and limited-time offers to increase transaction sizes.
- Utilize referral programs to attract new customers.

**Cluster 3: Young Shoppers**

- Enhance online presence through influencer partnerships and social media engagement.
- Offer interactive experiences and gaming events in-store.

**Cluster 4: Older Shoppers**

- Provide personalized customer service and tailored product recommendations.
- Highlight product durability and reliability in marketing materials.

# CONCLUSION

In conclusion, the customer segmentation analysis has identified distinct groups with unique characteristics and behaviors. By implementing targeted marketing strategies tailored to each segment, the mall can enhance customer satisfaction, increase loyalty, and drive revenue growth effectively.

# BIBLIOGRAPHY

- [www.kaggle.com](www.kaggle.com)
- www.chatgpt.com
- [www.cipherschools.com](www.cipherschools.com)
- Lecture PPTs