

Detailed Insight Report for Mall Customer Dataset

Introduction:

Objective:

The primary objective of this project is to analyze customer data from a retail store and segment customers into distinct groups based on their purchasing behavior. By identifying these segments, the retail store can tailor its marketing strategies to better meet the needs of each customer group, ultimately enhancing customer satisfaction and boosting sales.

By leveraging customer segmentation, the retail store can implement more effective marketing strategies, improve operational efficiency, and ultimately achieve a competitive advantage in the market.

Use Case:

Customer segmentation is a crucial aspect of customer relationship management (CRM) and marketing strategies. In the context of a retail store, understanding different customer segments allows the store to:

- **Develop Targeted Marketing Campaigns:** Tailor promotions and advertisements to specific customer groups based on their purchasing habits and preferences.
- **Personalize Customer Experiences:** Offer personalized recommendations and services to improve customer satisfaction and loyalty.
- **Optimize Product Offerings:** Adjust inventory and product offerings to align with the preferences of different customer segments.
- **Increase Customer Retention:** Implement strategies to retain high-value customers and reduce churn rates.
- **Enhance Sales and Revenue:** Identify opportunities for cross-selling and up-selling to maximize sales and revenue.

Overview of the Dataset

Dataset Description: The dataset used in this project is the "Mall Customers" dataset, which provides information about customers from a mall. The dataset contains demographic and behavioral attributes of the customers, which can be used to perform segmentation. The dataset includes the following columns:

1. CustomerID: Unique identifier for each customer.
2. Gender: Gender of the customer (Male/Female).
3. Age: Age of the customer.
4. Annual Income (k\$): Annual income of the customer in thousands of dollars.
5. Spending Score (1-100): Spending score assigned by the mall based on customer behavior and spending nature (1 being lowest and 100 being highest).

Attributes:

- CustomerID: A numerical identifier unique to each customer.
- Gender: Categorical variable indicating the customer's gender.
- Age: Numerical variable indicating the customer's age.
- Annual Income (k\$): Numerical variable indicating the customer's annual income in thousands of dollars.
- Spending Score (1-100): Numerical variable indicating the spending score, a metric assigned by the mall based on customer spending behavior.

Purpose :

The dataset is used to perform customer segmentation analysis. By examining the demographic and behavioral attributes of the customers, we aim to identify distinct groups of customers who exhibit similar purchasing behaviors. These

insights will enable the retail store to develop targeted marketing strategies and enhance overall customer satisfaction.

Data Source: The dataset is publicly available and can be downloaded from Kaggle at the following link: [Mall Customers Dataset on Kaggle](#).

Data Analysis and Segmentation:

- Clean the data to handle missing values and ensure consistency.
- Perform Exploratory Data Analysis (EDA) to understand the distribution and relationships within the data.
- Apply K-Means Clustering to segment customers into distinct groups.
- Visualize the results using Matplotlib and Power BI to gain actionable insights.

Demographic Distribution

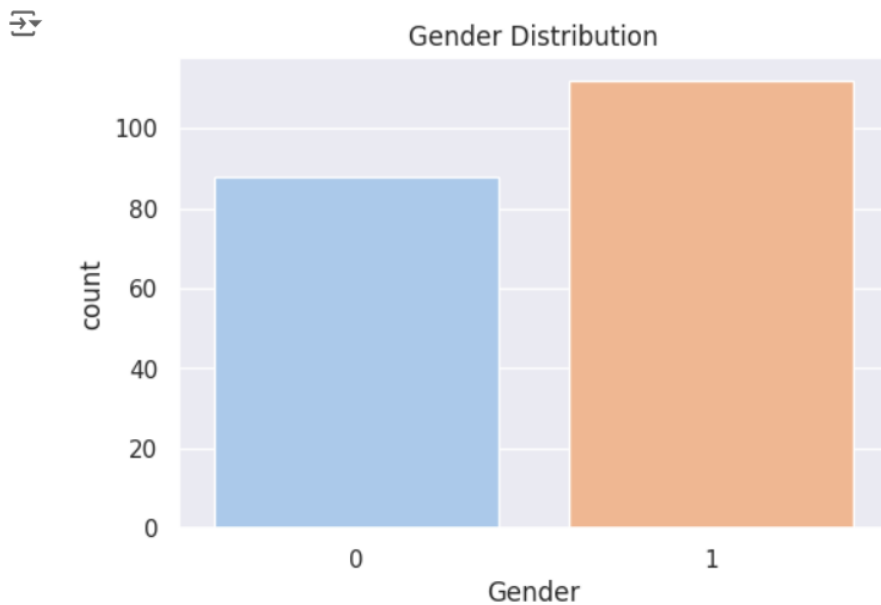
- **Gender Distribution :**

The dataset contains a total of 200 customers, with the following gender breakdown:

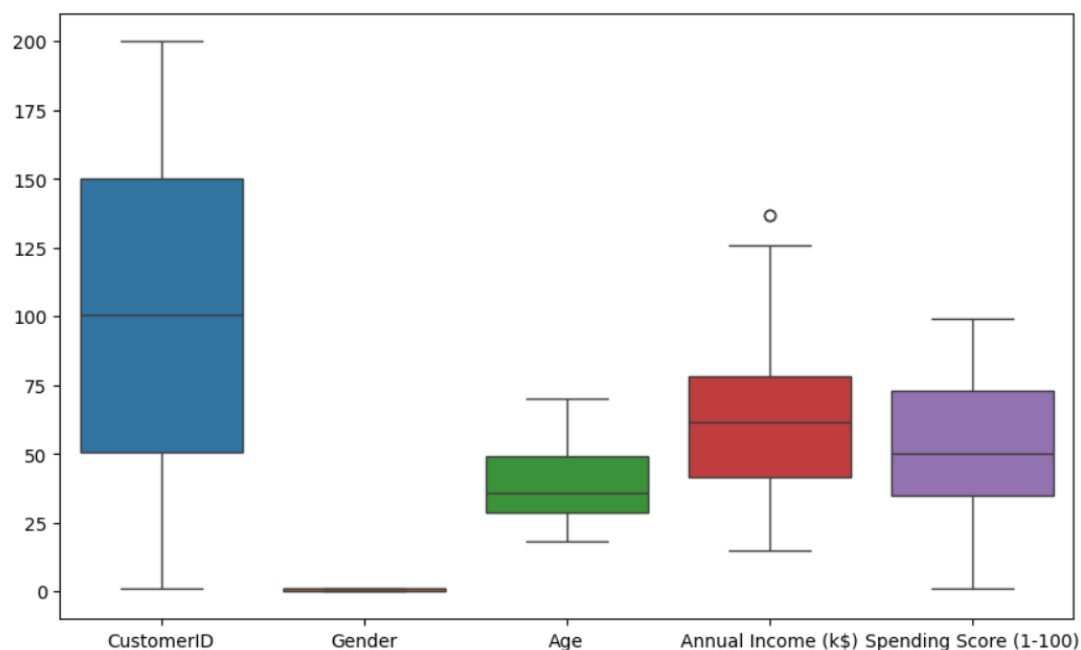
- **Female:** 112 customers (56%)
- **Male:** 88 customers (44%)

Distribution of the Spending Score:

```
plt.figure(figsize=(6, 4))
sns.countplot(x='Gender', data=df, palette='pastel')
plt.title('Gender Distribution')
plt.show()
```



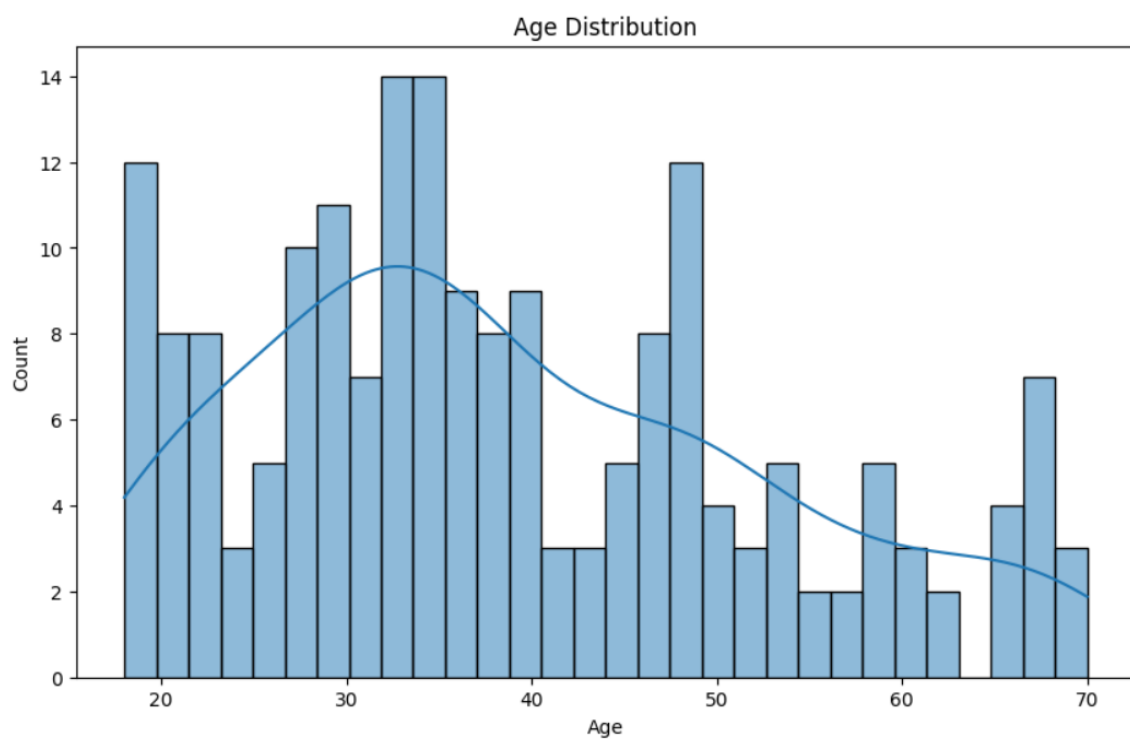
Analysis: The slight majority of customers are female. This can influence the types of products and services offered, as well as marketing strategies. Understanding the preferences and behaviors of female customers can help the mall tailor its offerings to meet their needs more effectively.



- **Age Distribution:**

- **Age Range:** 18 to 70 years

- **Common Age Groups:**
 - Young adults (18-25): A significant portion of the customer base consists of young adults, who may have different shopping habits compared to older customers.
 - Adults (26-35): This age group is typically entering their peak earning years and may have more disposable income.
 - Middle-aged adults (36-50): Often have established careers and families, potentially increasing their spending power.
 - Older adults (51-70): This group may be more selective in their spending, focusing on quality and value.



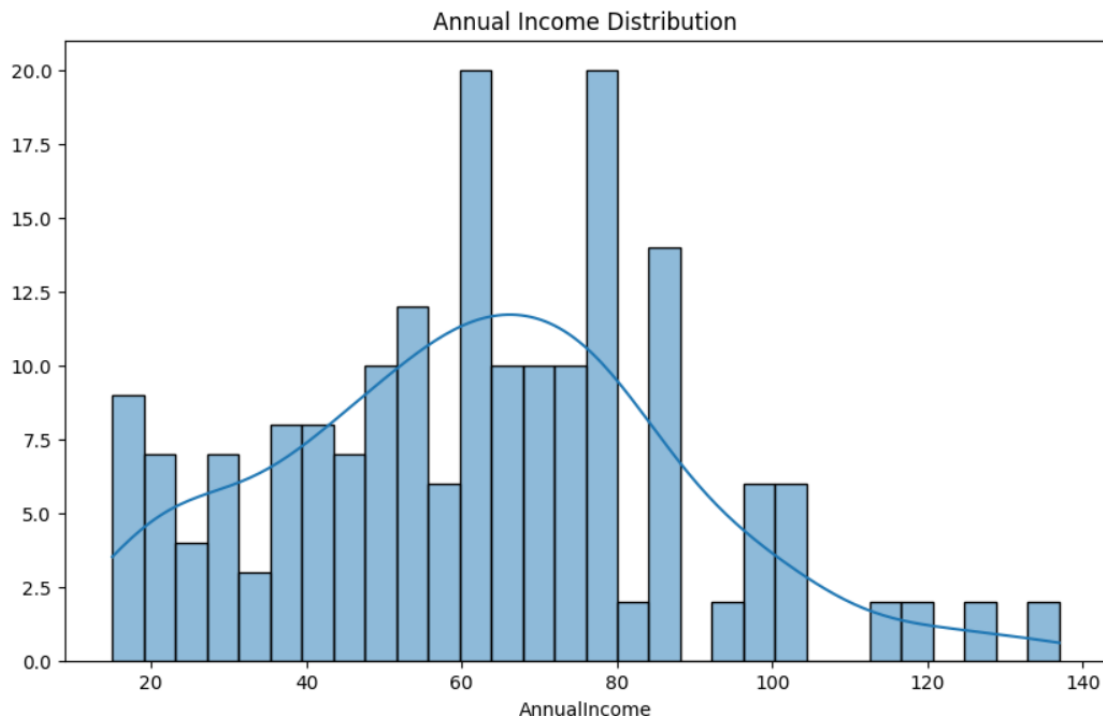
Age Distribution Insights:

- The most common age group is between 30 and 40 years.

- This age group is likely to have stable incomes and family responsibilities, influencing their purchasing decisions towards family-oriented and utility products.
- Younger customers (18-25) might be more inclined towards fashion, entertainment, and technology products.
- Older customers might prioritize health, wellness, and luxury items.

1. Income Analysis :

- **Income Distribution**
- **Annual Income Range:** \$15,000 to \$137,000
- **Income Brackets:**
 - Low income (under \$40,000): These customers may be budget-conscious and look for discounts and value deals.
 - Middle income (\$40,000 to \$80,000): The bulk of customers fall into this category. They are likely to balance quality with affordability and may be responsive to promotions and loyalty programs.
 - High income (over \$80,000): These customers might be less price-sensitive and more focused on premium products and exclusive services.



Income Distribution Insights:

- The most common income range is between \$50,000 and \$85,000, indicating a strong middle-income customer base.
- Peaks in the \$60,000 to \$75,000 range suggest a concentration of customers with moderate spending power, which can be leveraged by offering mid-range products and services.
- Marketing strategies should highlight value, quality, and potential savings to appeal to the largest income segment.

2. Spending Score Analysis

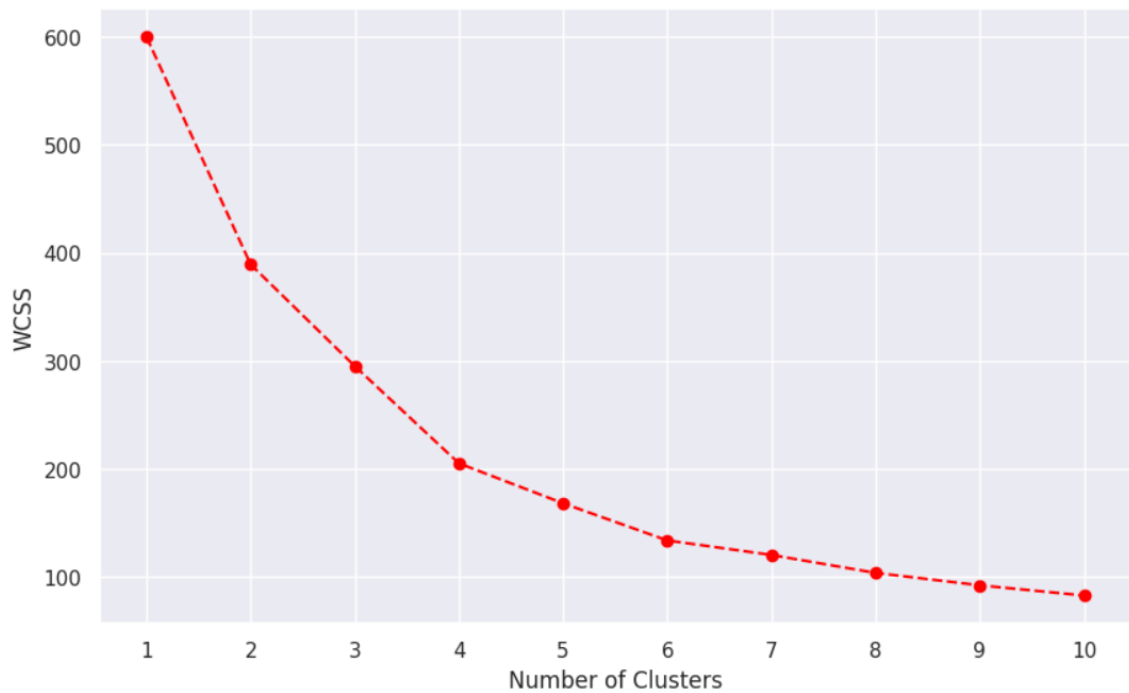
Spending Score Distribution

- **Spending Score Range:** 1 to 99
- **Spending Behavior:**

- Low spenders (1-40): These customers may be infrequent shoppers or those who spend cautiously. Offering targeted discounts and promotions could encourage higher spending.
- Moderate spenders (41-60): The bulk of the customer base falls into this category. They represent a stable revenue stream with potential for increased spending through loyalty programs and personalized marketing.
- High spenders (61-100): These customers likely contribute a significant portion of the mall's revenue. They might be interested in premium products, exclusive deals, and high-end services.

Spending Score Insights:

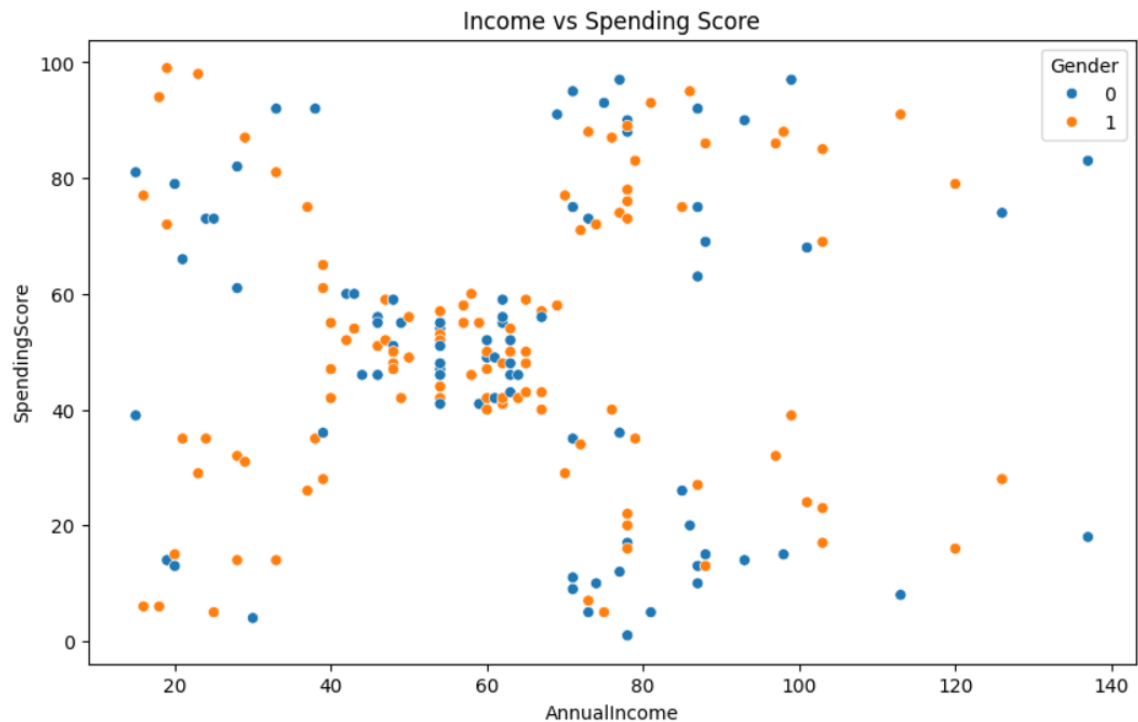
- The spending score distribution is fairly uniform, with a slight increase in scores between 40 and 60.
- Customers with spending scores between 50 and 75 are numerous, indicating moderate spending behavior.
- This balanced distribution suggests opportunities to increase spending across all customer segments through targeted marketing and personalized incentives.
- **Clusters :**



3. Relationships between Variables

Income vs. Spending Score

- **Correlation Analysis:**
 - No strong correlation between annual income and spending score.
 - Customers with low, middle, and high incomes exhibit a wide range of spending scores.
 - This indicates that income alone is not a predictor of spending behavior.



Income vs. Spending Score Insights:

- Marketing strategies should not solely focus on high-income customers for high spending but rather consider other factors such as shopping preferences, loyalty, and customer engagement.
- Understanding the underlying motivations and preferences of customers across different income levels can help tailor marketing efforts more effectively.

Age vs. Spending Score

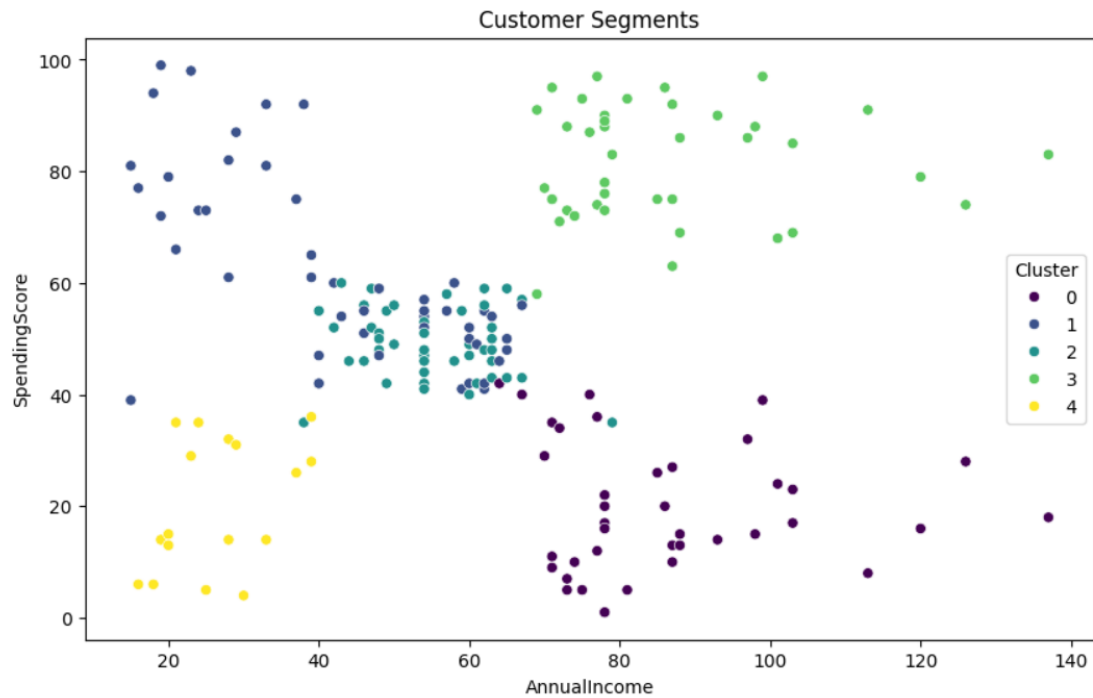
- **Correlation Analysis:**
 - No clear pattern or correlation between age and spending score.
 - Customers of all age groups exhibit varied spending scores, indicating diverse spending behaviors.

Age vs. Spending Score Insights:

- This suggests that spending behavior is influenced by factors beyond age, such as lifestyle, personal preferences, and individual financial situations.
- Marketing strategies should consider segmenting customers based on behaviors and preferences rather than age alone.
- Personalized marketing approaches that cater to individual interests and needs are likely to be more effective.



- **Customer Segments :**



Summary of Key Insights:

1. Demographic Insights:

- The customer base is predominantly female, with a slight majority.
- The age range is broad, with a significant number of customers in their 30s and 40s.

2. Income Insights:

- Most customers fall within the middle-income bracket, with annual incomes ranging from \$50,000 to \$85,000.
- Marketing strategies should highlight value and quality to appeal to this segment.

3. Spending Behavior:

- Spending scores are widely distributed, with many customers showing moderate spending behavior.
- There is no strong correlation between income and spending score or between age and spending score, suggesting that spending behavior is influenced by multiple factors.

4. Targeted Marketing Opportunities:

- Focus on middle-aged and middle-income customers, who form a significant portion of the customer base.
- Personalized marketing strategies based on individual preferences and behaviors can enhance customer engagement and spending.
- Promotions and loyalty programs can be tailored to different spending segments to encourage higher spending.