Project Insights Report

The dataset Cleaned\_Mall\_Customers.csv consists of 5 columns and includes the following information about customers at a mall:

1. **CustomerID**: A unique identifier for each customer.
2. **Gender**: Encoded as 0 for Male and 1 for Female.
3. **Age**: The age of the customer.
4. **Annual Income (k$)**: The annual income of the customer in thousands of dollars.
5. **Spending Score (1-100)**: A score assigned by the mall based on customer behavior and spending nature.

The scatter plot (Annual Income vs. Spending Score, Colored by Gender) provides several key insights into customer behavior:

A graph with blue and orange dots

Description automatically generated

**Insights from the Scatter Plot:**

1. Spending Score Distribution:

- The Spending Score ranges from 0 to 100.

- There is a significant concentration of data points around the Spending Score of 50-60, particularly for Annual Incomes in the range of 20-40.

2. Annual Income Distribution:

- Annual Income varies from around 20 to 140.

- There is a noticeable clustering of data points around Annual Incomes of 60 and 80.

3. Gender Distribution:

- Both genders (0 and 1) are distributed across the entire range of Spending Scores and Annual Incomes.

- There is no clear segregation or pattern based on gender in the distribution of Spending Scores and Annual Incomes.

4. Clusters of Spending Patterns:

- There are distinct clusters observed:

- A cluster with high Spending Scores (above 80) for both genders, across a wide range of Annual Incomes.

- A cluster with medium Spending Scores (around 50-60) and moderate Annual Incomes (around 40-60).

- Another cluster with low Spending Scores (below 20) spread across a wide range of Annual Incomes.

5. High Annual Income Observations:

- For Annual Incomes above 100, the Spending Scores are dispersed, indicating no clear spending pattern based on high income.

6. Middle Income Group:

- Middle-income group (Annual Income 40-60) shows a wide range of Spending Scores from low to high.

* These observations indicate various spending behaviors across different income levels and gender, with some clustering that might suggest different market segments.

A graph of age distribution

Description automatically generatedA graph of income distribution

Description automatically generatedA graph of a number of columns

Description automatically generated with medium confidence

**K-Means Clustering Segmentation:**

A chart of colored dots

Description automatically generated

**Interpretation of the Plot:**

1. **Clustering**: The plot suggests that the customers have been divided into five distinct clusters. Each cluster represents a group of customers with similar characteristics based on their annual income and spending score.
   * **Cluster 0 (Purple)**: These customers have low to moderate annual income and a moderate spending score.
   * **Cluster 1 (Blue)**: These customers have a high annual income and a high spending score.
   * **Cluster 2 (Cyan)**: These customers have low annual income but a high spending score.
   * **Cluster 3 (Green)**: These customers have low annual income and a low spending score.
   * **Cluster 4 (Yellow)**: These customers have high annual income but a low spending score.
2. **Spending Behavior**:
   * Customers in Cluster 1 (blue) are likely high-income individuals who also spend a lot.
   * Customers in Cluster 2 (cyan) have lower incomes but spend significantly more,

indicating they might be younger or have different spending habits.

* + Customers in Cluster 3 (green) are more conservative with both their earnings and

spending.

* + Customers in Cluster 4 (yellow) have high incomes but prefer to save rather than spend.

**Marketing Strategies:**

Based on the customer segments shown in the scatter plot, here are some potential marketing strategies for each cluster:

**Cluster 0 (Purple)**

- Characteristics: Moderate annual income, moderate spending score

- Strategy: Introduce loyalty programs and personalized offers to encourage more spending. Highlight value and cost-effectiveness.

**Cluster 1 (Blue)**

- Characteristics: High annual income, high spending score

- Strategy: Focus on premium products and services. Offer exclusive deals and premium memberships to enhance their experience.

**Cluster 2 (Teal)**

- Characteristics\*\*: Low annual income, high spending score

- \*\*Strategy\*\*: Promote budget-friendly products with high value. Emphasize discounts and deals. Introduce layaway plans or installment payments.

**Cluster 3 (Green)**

- Characteristics: Low annual income, low spending score

- Strategy: Focus on basic and essential products. Offer significant discounts and promotions. Implement educational campaigns to highlight how certain products can improve their quality of life.

**Cluster 4 (Yellow)**

- Characteristics: High annual income, low spending score

- Strategy: Highlight high-quality and premium products. Create personalized marketing messages to convert them into higher spenders. Offer incentives for increased spending, such as cashback or high-value gift cards.