**Overview**

This project analyzes customer demographics and spending habits using python. The goal is to identify high value customer groups and suggest marketing strategies based on the insights found in the data.

**Tools Used**

* Python (Pandas, Matplotlib, Seaborn)
* Jupyter Notebook
* Dataset: Mall Customer Segmentation (Kaggle)

**Import Libraries and Load the Dataset**

We began by importing essential Python libraries and loading the dataset into a Pandas DataFrame. This dataset includes customer demographics, annual income, and spending scores. Below is a preview of the first five rows of the dataset to confirm successful loading.

A screenshot of a computer

AI-generated content may be incorrect.

**Check Data Structure and Missing Values**

Before proceeding, we checked the dataset’s structure to understand the data types and confirm there were no missing values. This ensures the dataset is clean and ready for analysis.

A screenshot of a computer

AI-generated content may be incorrect.

**Summary Statistics**

We reviewed the summary statistics of numerical features to understand the distribution of customer ages, incomes, and spending scores. This provides insights into central tendencies and variability.

A screenshot of a computer

AI-generated content may be incorrect.

**Gender Distribution**

We visualized the gender distribution to see if there was a dominant gender in the customer base. The data shows a fairly even split, which suggests balanced marketing potential.

A screenshot of a graph

AI-generated content may be incorrect.

**Age Distribution**

Analyzing the age distribution showed that most customers are between 20 and 40 years old, making this a prime target group for promotions and marketing campaigns.

A screenshot of a graph

AI-generated content may be incorrect.

**Annual Income vs. Spending Scores**

We plotted annual income against spending score to understand how income affects spending. The chart reveals that mid-income earners ($40k-$70k) have the highest spending scores, rather than high-income customers.

A screen shot of a graph

AI-generated content may be incorrect.

**Correlation Heatmap**

The heatmap identifies correlations between variables. In this dataset, no strong relationships appear between income and spending score, indicating other factors likely influence spending behavior.

A screenshot of a computer

AI-generated content may be incorrect.

**Key Findings**

Based on the findings, we suggest:

* Targeting marketing efforts towards the 20-40 age group.
* Offering loyalty programs to mid-income, high-spending customers.
* Customizing promotions by analyzing gender-based preferences.
* Avoid focusing only on high-income individuals; prioritize behavioral insights.

**Conclusion**

This analysis demonstrates how customer demographic data can identify profitable target groups and inform business strategies. By focusing on spending habits and demographic trends, businesses can optimize marketing efforts and improve customer engagement.