Project Title: Understanding Customer Segments and Improving Retail Strategy

Problem Statement:

The retail industry is highly competitive, and understanding customer behavior is crucial for businesses to thrive. In this project, the goal is to analyze customer data from a mall and identify distinct customer segments. By gaining insights into these segments, the aim is to help the mall optimize its marketing strategies, improve customer experience, and boost sales.

Design Thinking Process:

1. Empathize: Understand the needs of the mall and its customers.
2. Define: Clearly define the problem and set project objectives.
3. Ideate: Generate ideas and hypotheses about customer behavior.
4. Prototype: Create a model to segment customers based on available data.
5. Test: Validate the model and refine strategies based on insights.

Phases of Development:

Data Collection:

- ❖ Utilize the provided dataset from Kaggle: Mall Customers Dataset.

Data Preprocessing:

- ❖ Handle missing or irrelevant data.
- ❖ Encode categorical variables.
- ❖ Normalize/standardize numerical features if necessary.
- ❖ Explore data to understand distributions and correlations.

Analysis Techniques:

- Use unsupervised learning techniques (such as K-means clustering) to segment customers into distinct groups.
- Employ visualization tools (like matplotlib, seaborn) for data exploration and result presentation.
- Apply statistical methods to validate findings and test hypotheses.

Model Development:

- Apply K-means clustering algorithm to identify customer segments.
- Determine the optimal number of clusters using techniques like the elbow method or silhouette score.

Evaluation and Interpretation:

- Evaluate the model's performance and cluster quality.
- Interpret the clusters and identify key characteristics of each segment.
- Derive actionable insights from the clusters.

Recommendations:

- Develop targeted marketing strategies for each customer segment.
- Optimize product placement and pricing based on customer preferences.
- Enhance customer service and engagement tailored to different segments.
- Measure the impact of the implemented strategies and iterate if necessary.

Dataset Description:

The dataset contains the following columns:

- CustomerID: Unique identifier for each customer.
- Gender: Gender of the customer (Male/Female).
- Age: Age of the customer.
- Annual Income (k): Annual income of the customer in thousands of dollars.
- Spending Score (1-100): Spending score assigned to the customer based on their purchasing behavior and spending nature.

Data Preprocessing Steps:

Handling Missing Data:

- Check for missing values and decide on appropriate strategies (imputation or removal).

Encoding Categorical Variables:

- Convert categorical variable Gender into numerical format using label encoding (0 for Male, 1 for Female).

Normalization/Standardization:

- Normalize features like Age, Annual Income, and Spending Score to bring them to a similar scale.

Analysis Techniques Applied:

K-means Clustering:

- Use K-means algorithm to cluster customers into distinct groups based on their features.
- Experiment with different values of K (number of clusters) to find the optimal segmentation.

Data Visualization:

- Utilize various charts and graphs (scatter plots, bar charts) to visualize the clusters and their characteristics.
- Use visualization to explain the findings to stakeholders effectively.

Key Findings, Insights, and Recommendations:

Key Findings:

- Identified distinct customer segments (e.g., high spenders, moderate spenders, low spenders) based on their age, income, and spending score.
- Discovered that a specific gender group dominates a particular spending segment.

Insights:

- Younger customers tend to spend more on average compared to older customers.
- Customers with higher incomes do not necessarily have higher spending scores.
- Certain gender groups exhibit specific spending behaviors.

Recommendations:

- Target high-income, high-spending segments with premium products and personalized marketing.
- Improve store layout and promotions to attract and retain younger customers.
- Offer loyalty programs or discounts to enhance customer engagement and loyalty among various segments.
- Conduct periodic customer surveys to gather qualitative data and refine strategies further.