**Overview of the Project – Customer Segmentation Analysis**

**1. Project Introduction:**

* **Title:** Customer Segmentation Using KMeans and Exploratory Data Analysis
* **Goal:** The primary goal of this project was to explore customer demographics and behavior to segment customers into different groups using KMeans clustering. This segmentation could help the mall understand customer patterns and target different groups with personalized marketing strategies.

**2. Overview of the Dataset:**

* The dataset consisted of 200 customers with the following features:
  + **Customer ID**
  + **Gender**
  + **Age**
  + **Annual Income (in $)**
  + **Spending Score** (a value from 1 to 100)

**3. Steps Undertaken:**

* **Data Loading and Cleaning:** Loaded the dataset, ensured it was clean (no missing values), and conducted initial exploration using descriptive statistics like mean, median, and variance.
* **Univariate and Bivariate Analysis:**
  + Visualized the distribution of **Age**, **Gender**, and **Annual Income** to get an initial understanding of the customers.
  + Analyzed relationships between **Age, Income, and Spending Score** to find patterns in customer spending.
* **KMeans Clustering:**
  + Used KMeans clustering to group customers based on their **Annual Income** and **Spending Score**.
  + The algorithm identified key customer segments that could be used for marketing purposes, such as high-income, high-spending customers or low-income, low-spending customers.

**4. Challenges and Learnings:**

* One of the challenges was selecting the optimal number of clusters. You likely used techniques like the **Elbow method** or **Silhouette score** to determine the right number of clusters, which helped refine the customer segmentation.

**5. Business Impact:**

* The insights gained from the segmentation can help businesses tailor marketing campaigns. For example, the high-spending segment could be targeted with loyalty programs, while the low-spending group could be engaged with promotional offers.

**6. Tools Used:**

* **Pandas** for data manipulation.
* **Matplotlib and Seaborn** for data visualization.
* **KMeans Clustering** from **scikit-learn** for segmentation.