**Report**

**Approach Used**

**Exploratory Data Analysis (EDA):**

* Checked for missing values and summary statistics.
* Dropped unnecessary columns (**CustomerID, Gender**) since they do not contribute to numerical clustering.

**Feature Scaling & Clustering:**

* Used **StandardScaler** to normalize numerical features.
* Applied the **Elbow Method** to determine the optimal number of clusters.
* Used **K-Means Clustering (k=5)** to segment customers.

**Visualization & Interpretation:**

* Created a **2D scatter plot** (Age vs Spending Score) to analyze clusters.
* Created a **3D visualization** (Age, Income, Spending Score) to better understand customer groups.

**Challenges Faced**

**Choosing the Right Number of Clusters:**

* Initially, we were unsure how many clusters to use. The **Elbow Method** helped determine that **k=5** was the optimal choice.

**Handling Scaling Issues:**

* K-Means is sensitive to different feature scales, so **StandardScaler** was used to normalize data.

**Model Performance & Improvements**

**Customer segmentation successfully grouped customers into 5 clusters** based on their demographics and spending habits.  
**The visualization clearly showed patterns** (e.g., high-income, low-spending customers vs. low-income, high-spending customers).

**Improvements for Future Work:**  
🔹 Using **PCA (Principal Component Analysis)** to reduce dimensionality and improve visualization.