Customer Segmentation using K-Means Clustering
Internship Task 8

1. Objective

The objective of this project is to perform **unsupervised learning using K-Means clustering** on the **Mall Customer Segmentation dataset** to identify meaningful customer groups and derive **actionable business insights**.

2. Tools and Technologies Used

- Python
- Pandas
- NumPy
- Matplotlib
- Seaborn
- Scikit-learn

3. Methodology

Step 1: Data Loading & Exploration

The dataset was loaded using Pandas. Descriptive statistics and data visualization were performed to understand data distribution and detect anomalies.

Step 2: Feature Selection

Two main features were selected:

- Annual Income (k$)
- Spending Score (1–100)

These features best represent customer purchasing behavior.

Step 3: Feature Scaling

Standard Scaler was applied to normalize data because K-Means uses **distance-based calculations**.

Step 4: Elbow Method

The Elbow Method was used to determine the optimal number of clusters by plotting **WCSS vs number of clusters (K)**.
The optimal number of clusters was found to be **K = 5**.

Step 5: K-Means Model Training

A K-Means clustering model with **K = 5** was trained, and customers were assigned to clusters.

Step 6: Cluster Visualization

Clusters were visualized using scatter plots with color coding and centroid highlighting.

Step 7: PCA Visualization

Principal Component Analysis (PCA) was applied to reduce data into 2D space for **better visualization of cluster separation**.

Step 8: Evaluation using Silhouette Score

Silhouette Score was computed to measure clustering quality. A high silhouette score indicated **strong separation between clusters**.

5. Results and Observations

Five meaningful customer segments were identified:

1. **High Income – High Spending:** Premium Customers
2. **High Income – Low Spending:** Careful Customers
3. **Low Income – High Spending:** Target Customers
4. **Low Income – Low Spending:** Budget Customers
5. **Medium Income – Medium Spending:** Regular Customers

6. Business Insights

Customer segmentation enables businesses to:

- Design **targeted marketing campaigns**
- Improve **customer engagement**
- Optimize **inventory management**
- Increase **sales and customer retention**
- Enhance **decision-making strategies**

7. Conclusion

This project successfully demonstrates the application of **K-Means clustering for customer segmentation**. The results provide valuable insights that can significantly help businesses improve marketing strategies and customer experience.