**Assignment 5**

**Name: Ashutosh Shinde**

**Roll no: 282006**

**Batch: B1**

**Problem Statement**

Apply clustering algorithms to segment customers of a cosmetics shop based on their spending behavior and demographics to predict responses to special offers.

**Objective**

1. To preprocess and analyze customer data.
2. To apply K-Means and Agglomerative Clustering for segmentation.
3. To evaluate and visualize customer clusters.

**Resources Used**

* **Software:** Google Colab
* **Libraries:**
  + Pandas, NumPy (Data Manipulation)
  + Matplotlib, Seaborn (Visualization)
  + Scikit-learn (Clustering Algorithms)

**Methodology**

**1. Data Loading & Exploration**

* **Dataset:** Mall\_Customers.csv (simulated cosmetics shop data).
* **Features Used:**
  + Annual Income (k$)
  + Spending Score (1-100)

**2. Data Preprocessing**

* Encoded categorical feature (Genre) using Label Encoding.
* Scaled numerical features using StandardScaler.

**3. Clustering Algorithms**

1. **K-Means Clustering:**
   * Used the Elbow Method to determine optimal clusters (K=5).
2. **Agglomerative Clustering:**
   * Applied hierarchical clustering with 5 clusters.

**4. Model Evaluation**

* **Silhouette Scores:**
  + K-Means: **0.55**
  + Agglomerative: **0.55**
* **Visualization:** Scatter plots of clusters.

**Conclusion**

1. **Segmentation Success:** Both algorithms identified 5 distinct customer groups with similar performance (Silhouette ≈ 0.55).
2. **Business Impact:** Clusters enable targeted marketing:
   * **High Income + High Spending:** Offer luxury items.
   * **Low Income + High Spending:** Promote affordable bundles.
3. **Future Work:**
   * Incorporate more features (e.g., purchase frequency).
   * Test DBSCAN for outlier detection.