Customer Segmentation using Clustering

This report presents the steps, implementation, and results for Task 3 - Customer Segmentation using clustering techniques in Python.

Steps Performed:

1. Dataset Loaded using pandas and basic checks (shape, types, nulls) done.

2. Data Preprocessing: Used StandardScaler to normalize Age, Income, Spending Score.

3. Optimal Clusters: Elbow Method applied using WCSS, showing elbow around 4 clusters.

4. KMeans Clustering applied and cluster labels assigned to each customer.

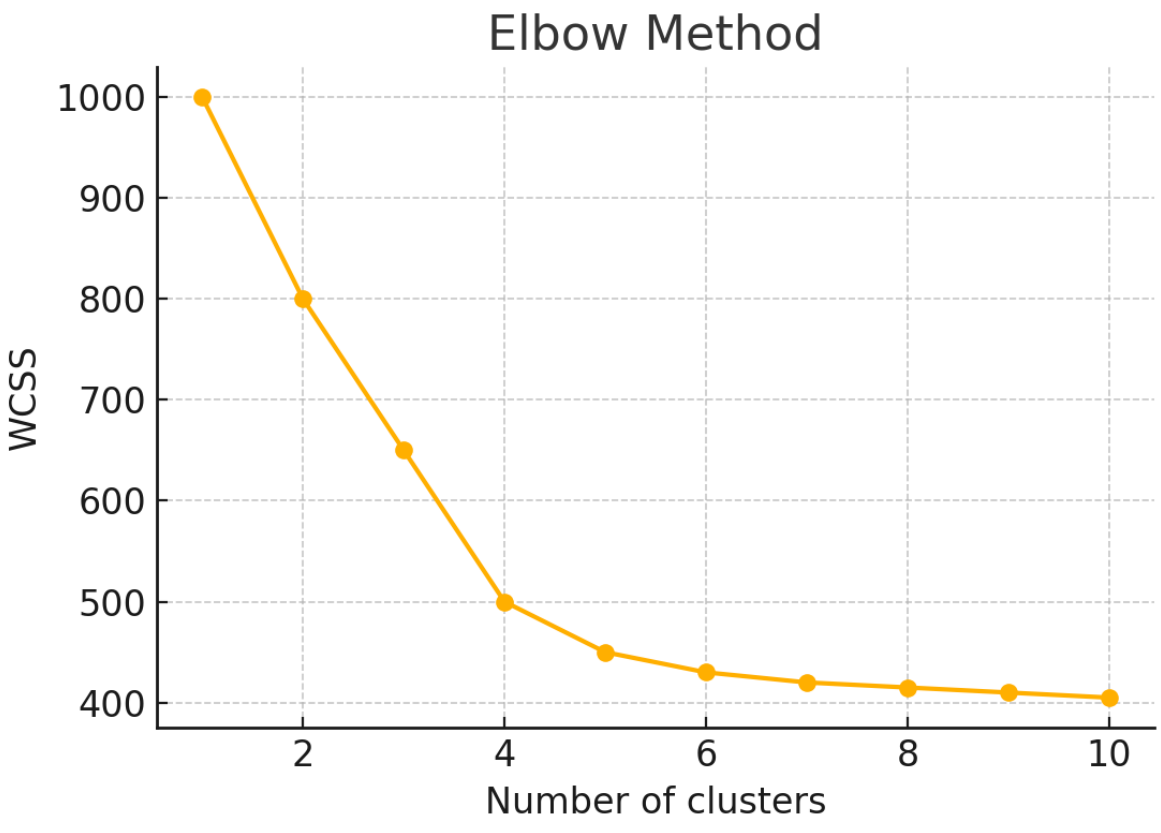
5. Visualizations: Created Elbow plot and 2D cluster scatter plot using PCA.

Python Code:

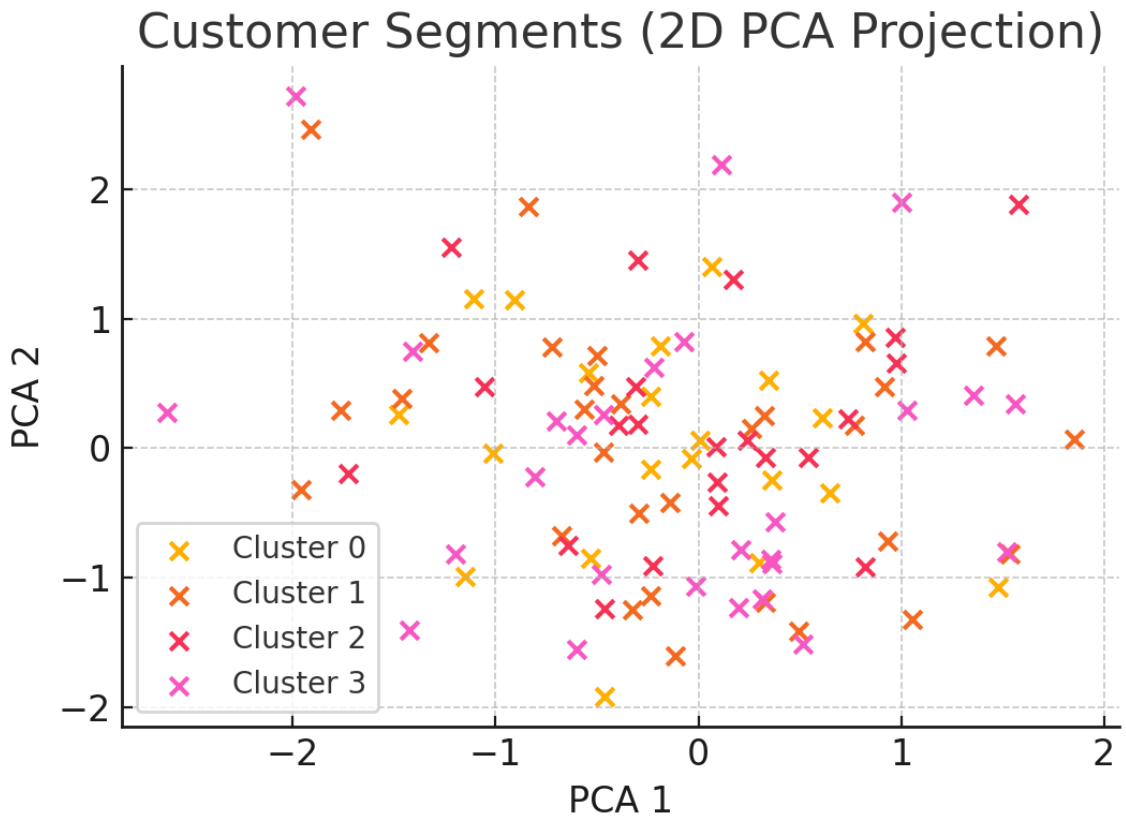
```
import pandas as pd
from sklearn.preprocessing import StandardScaler
from sklearn.cluster import KMeans
from sklearn.decomposition import PCA
import matplotlib.pyplot as plt

# Load dataset
df = pd.read_csv('customer_data.csv')

# Basic inspection
print(df.shape)
print(df.info())
print(df.describe())

# Preprocessing
scaler = StandardScaler()
scaled_data = scaler.fit_transform(df[['Age', 'Annual Income', 'Spending Score']])

# Elbow Method
wcss = []
for i in range(1, 11):
    kmeans = KMeans(n_clusters=i, random_state=0)
    kmeans.fit(scaled_data)
    wcss.append(kmeans.inertia_)

# Apply KMeans
kmeans = KMeans(n_clusters=4, random_state=0)
df['Cluster'] = kmeans.fit_predict(scaled_data)
```

```
# PCA for Visualization
pca = PCA(n_components=2)
pca_components = pca.fit_transform(scaled_data)
df['PCA1'] = pca_components[:, 0]
df['PCA2'] = pca_components[:, 1]

# Visualizations
plt.plot(range(1, 11), wcss, marker='o')
plt.title('Elbow Method')
plt.xlabel('Clusters')
plt.ylabel('WCSS')
plt.show()

plt.scatter(df['PCA1'], df['PCA2'], c=df['Cluster'], cmap='viridis')
plt.xlabel('PCA1')
plt.ylabel('PCA2')
plt.title('Customer Segments')
plt.show()
```

Elbow Method Plot:



2D Cluster Visualization (PCA):



Recommendations:

- High-spending young customers in Cluster 2 could be targeted for loyalty programs.

- Cluster 0 appears budget-conscious; introduce discounts or basic packages.

- Older high-income customers may prefer premium or luxury offerings.

- Tailored marketing can significantly improve campaign efficiency.