

```
import pandas as pd
import matplotlib.pyplot as plt
import seaborn as sns
from sklearn.cluster import KMeans
from sklearn.preprocessing import StandardScaler
```

```
df = pd.read_csv("Ecommerce Customers.csv")
print(df.head(5))
print(df.describe())
```

```

Email \
0      mstephenson@fernandez.com
1      hduke@hotmail.com
2      pallen@yahoo.com
3      riverarebecca@gmail.com
4      mstephens@davidson-herman.com

Address      Avatar \
0      835 Frank Tunnel\nWrightmouth, MI 82180-9605      Violet
1      4547 Archer Common\nDiazchester, CA 06566-8576      DarkGreen
2      24645 Valerie Unions Suite 582\nCobbborough, D...      Bisque
3      1414 David Throughway\nPort Jason, OH 22070-1220      SaddleBrown
4      14023 Rodriguez Passage\nPort Jacobville, PR 3...      MediumAquaMarine

Avg. Session Length      Time on App      Time on Website      Length of Membership \
0      34.497268      12.655651      39.577668      4.082621
1      31.926272      11.109461      37.268959      2.664034
2      33.000915      11.330278      37.110597      4.104543
3      34.305557      13.717514      36.721283      3.120179
4      33.330673      12.795189      37.536653      4.446308
```

```

Yearly Amount Spent
0      587.951054
1      392.204933
2      487.547505
3      581.852344
4      599.406092
```

```

Avg. Session Length      Time on App      Time on Website \
count      500.000000      500.000000      500.000000
mean      33.053194      12.052488      37.060445
std      0.992563      0.994216      1.010489
min      29.532429      8.508152      33.913847
25%      32.341822      11.388153      36.349257
50%      33.082008      11.983231      37.069367
75%      33.711985      12.753850      37.716432
max      36.139662      15.126994      40.005182
```

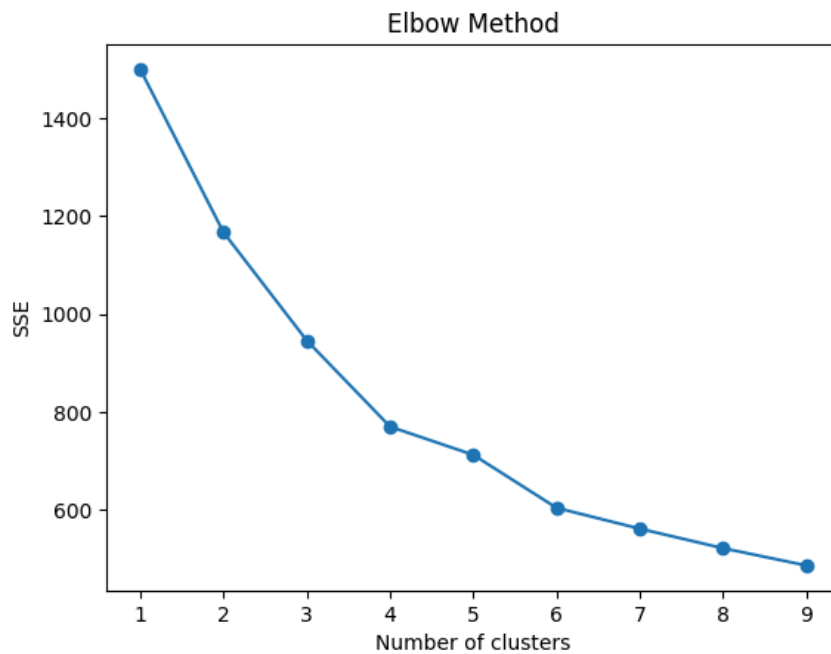
```

Length of Membership      Yearly Amount Spent
count      500.000000      500.000000
mean      3.533462      499.314038
std      0.999278      79.314782
min      0.269901      256.670582
25%      2.930450      445.038277
50%      3.533975      498.887875
75%      4.126502      549.313828
max      6.922689      765.518462
```

```
X = df[['Time on Website', 'Time on App', 'Length of Membership']]
scaler = StandardScaler()
X_scaled = scaler.fit_transform(X)
```

```
sse = []
for k in range(1, 10):
    km = KMeans(n_clusters=k, random_state=42)
    km.fit(X_scaled)
    sse.append(km.inertia_)
```

```
plt.plot(range(1, 10), sse, marker='o')
plt.xlabel('Number of clusters')
plt.ylabel('SSE')
plt.title('Elbow Method')
plt.show()
```



```
kmeans = KMeans(n_clusters=3, random_state=42)
df['Cluster'] = kmeans.fit_predict(X_scaled)

import pandas as pd
import matplotlib.pyplot as plt
from sklearn.cluster import KMeans

url = "https://raw.githubusercontent.com/Premalatha-success/Datasets/main/Mall_Customers.csv"
df = pd.read_csv(url)

X = df[['Annual Income (k$)', 'Spending Score (1-100)']]

kmeans = KMeans(n_clusters=5, random_state=42)
df['Cluster'] = kmeans.fit_predict(X)

plt.figure(figsize=(8, 6))
plt.scatter(X['Annual Income (k$)'], X['Spending Score (1-100)'], c=df['Cluster'], cmap='rainbow')
plt.xlabel('Annual Income (k$)')
plt.ylabel('Spending Score (1-100)')
plt.title('Customer Segments')
plt.grid(True)

plt.savefig('cluster_plot.png')
plt.show()
```

