```python
[1]: import numpy as np
     import pandas as pd
     import matplotlib.pyplot as plt
     import seaborn as sns

     %matplotlib inline

     df = pd.read_csv(r"C:\Users\nivet\OneDrive\Documents\Mall_Customers.csv")

     df.info()
```

```
<class 'pandas.core.frame.DataFrame'>
RangeIndex: 200 entries, 0 to 199
Data columns (total 5 columns):
 #   Column                  Non-Null Count  Dtype
---  ------                  --------------  -----
 0   CustomerID              200 non-null    int64
 1   Gender                  200 non-null    object
 2   Age                     200 non-null    int64
 3   Annual Income (k$)      200 non-null    int64
 4   Spending Score (1-100)  200 non-null    int64
dtypes: int64(4), object(1)
memory usage: 7.9+ KB
```
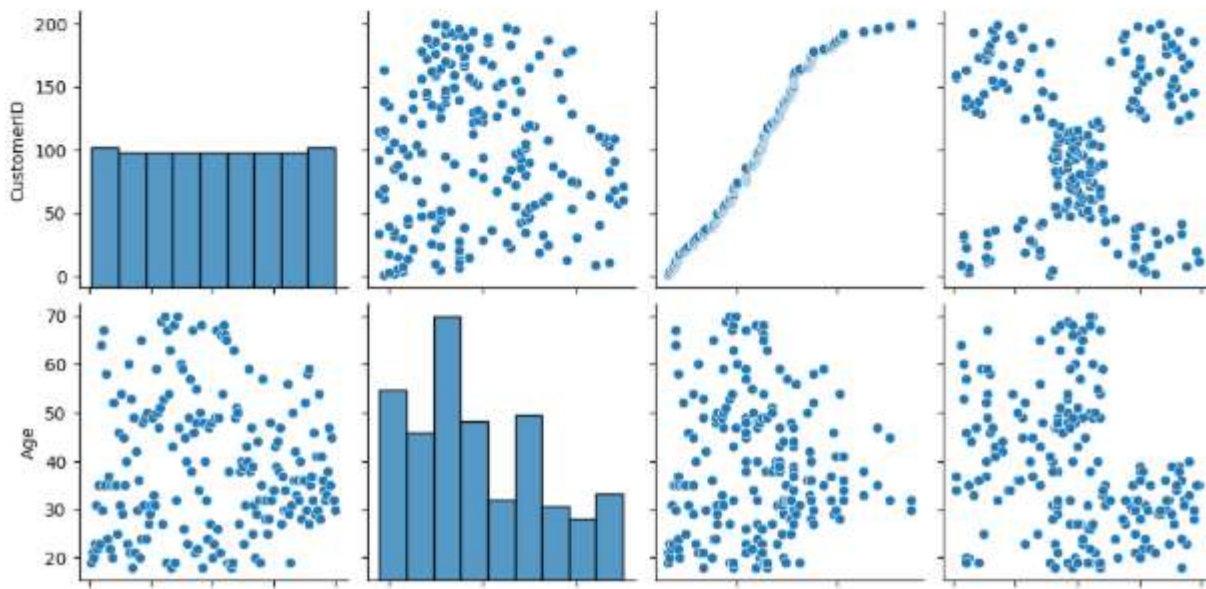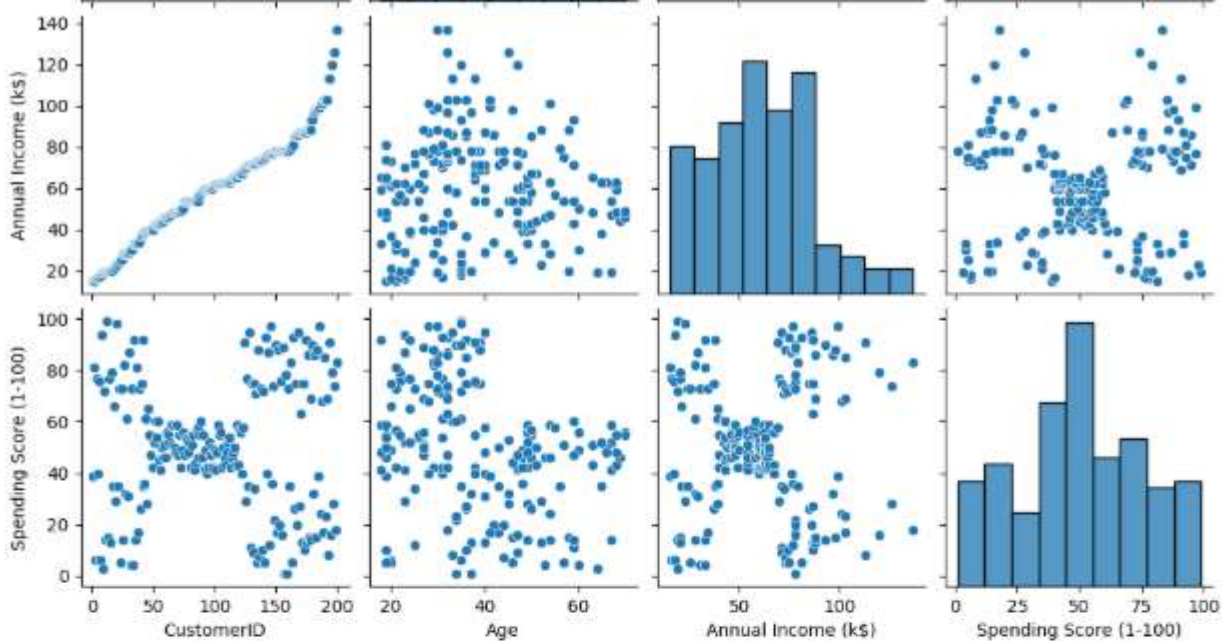
```python
[2]: df.head()
```

[2]:

	CustomerID	Gender	Age	Annual Income (k$)	Spending Score (1-100)
0	1	Male	19	15	39
1	2	Male	21	15	81
2	3	Female	20	16	6
3	4	Female	23	16	77
4	5	Female	31	17	40

```python
[3]: sns.pairplot(df)
```

[3]: <seaborn.axisgrid.PairGrid at 0x19eeef6a660>

```
[4]: features = df.iloc[:, [3, 4]].values

     from sklearn.cluster import KMeans
     model = KMeans(n_clusters=5, random_state=0)
     model.fit(features)
```

```
[4]: • KMeans  ○ ○

     ▶ Parameters
```

```
[5]: Final = df.iloc[:, [3, 4]].copy()
     Final['label'] = model.predict(features)
```
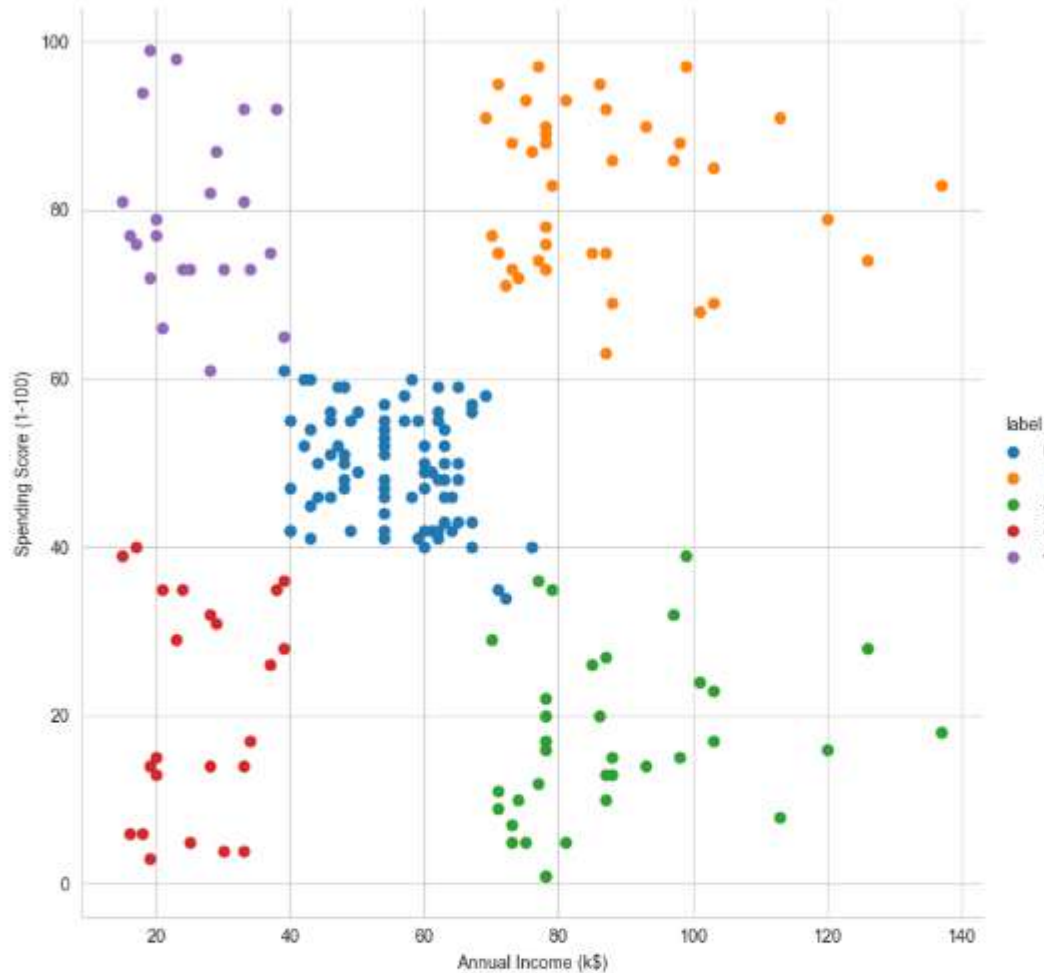
```
[6]: Final.head()
```

[6]:

	Annual Income (k$)	Spending Score (1-100)	label
0	15	39	3
1	15	81	4
2	16	6	3
3	16	77	4
4	17	40	3

```
[7]: sns.set_style("whitegrid")
     sns.FacetGrid(Final, hue="label", height=8) \
         .map(plt.scatter, "Annual Income (k$)", "Spending Score (1-100)") \
         .add_legend()

     plt.show()
```

```
[8]: features_el = df.iloc[:, [2, 3, 4]].values

     wcss = []
     for i in range(1, 10):
         model = KMeans(n_clusters=i, random_state=0)
         model.fit(features_el)
         wcss.append(model.inertia_)

     plt.plot(range(1, 10), wcss)
     plt.title("Elbow Method")
     plt.xlabel("Number of Clusters")
     plt.ylabel("WCSS")
     plt.show()
```

Elbow Method

WCSS vs Number of Clusters

[ ]: