

```
In [16]: df.head()
```

Out[16]:	CustomerID	Genre	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

```
In [17]: df.shape
```

```
Out[17]: (200, 5)
```

```
In [18]: df.columns
```

```
Out[18]: Index(['CustomerID', 'Genre', 'Age', 'Annual Income (k$)',
              'Spending Score (1-100)'],
              dtype='object')
```

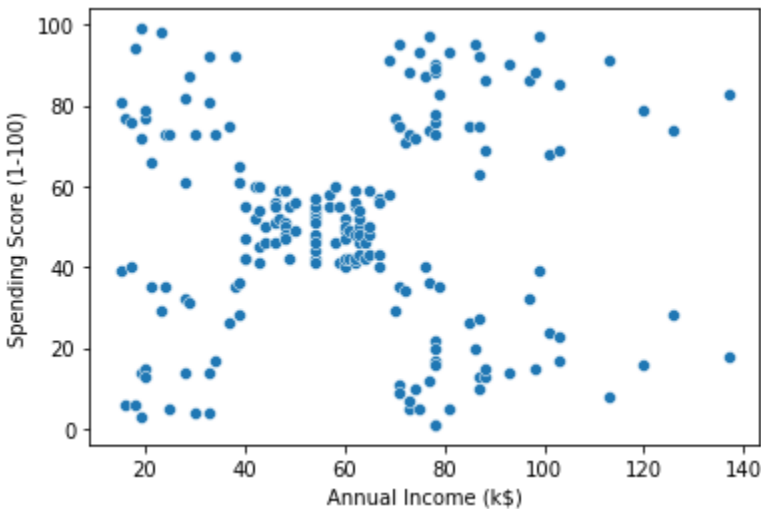
```
In [19]: df1 = df.drop(columns=['CustomerID', 'Genre', 'Age'])
```

```
In [20]: df1.head()
```

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

```
In [21]: sns.scatterplot(data = df1, x='Annual Income (k$)',y='Spending Score (1-100)')
```

```
Out[21]: <AxesSubplot:xlabel='Annual Income (k$)', ylabel='Spending Score (1-100)'>
```

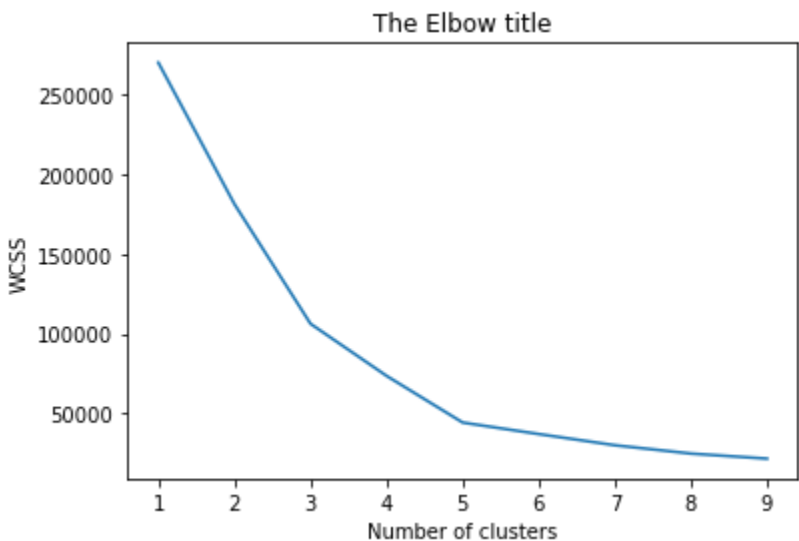


```
In [22]: from sklearn.cluster import KMeans
```

```
In [23]: wcss = []
          for i in range(1,10):
              kmeans = KMeans(i)
              kmeans.fit(df1)
              wcss.append(kmeans.inertia_)

          number_clusters = range(1,10)
          plt.plot(number_clusters, wcss)
          plt.title('The Elbow title')
          plt.xlabel('Number of clusters')
          plt.ylabel('WCSS')
```

```
Out[23]: Text(0, 0.5, 'WCSS')
```



```
In [28]: # selecting no. of clusters to be 3
Kmeans = KMeans(5)
Kmeans.fit(df1)
```

```
Out[28]: KMeans(n_clusters=5)
```

```
In [29]: Kmeans.labels_
```

[illegible]

In []: