```
import pandas as pd
```

```
dt = pd.read_csv('/content/Mall_Customers.csv')
```

Show command palette (Ctrl+Shift+P)

```
dt.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
```

```
x = dt.iloc[:,3:]
```

```
x.head()
```

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

```
from sklearn.cluster import KMeans
```

```
wcss = []
for i in range(1,11):
  kmeans = KMeans(n_clusters=i,init = "k-means++", random_state= 42)
  kmeans.fit(x)
  wcss.append(kmeans.inertia_)
```

```
import matplotlib.pyplot as plt
plt.plot(range(1,11),wcss)
plt.title("The Elbow Method")
plt.xlabel("Number of clusters")
plt.ylabel("WCSS")
plt.show()
```

The Elbow Method

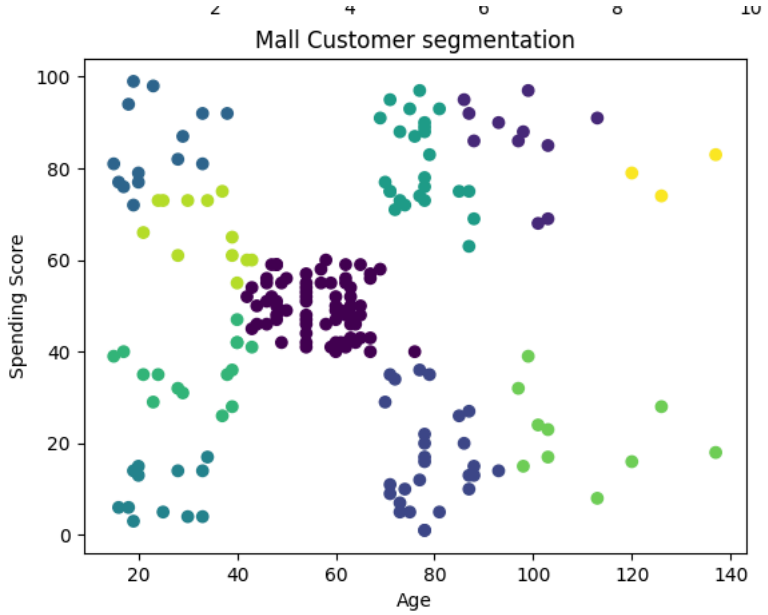
```
kmeans = KMeans(n_clusters=i,init ="k-means++", random_state=42 )
kmeans.fit(x)
```

Show command palette (Ctrl+Shift+P)

```
KMeans(n_clusters=10, random_state=42)
```

```
labels = kmeans.labels_
dt['cluster'] = labels
```

```
plt.scatter(dt['Annual Income (k$)'],dt['Spending Score (1-100)'],c=dt['cluster'])
plt.title('Mall Customer segmentation')
plt.xlabel('Age')
plt.ylabel('Spending Score')
plt.show()
```



``