```python
In [4]:   import numpy as np
          import pandas as pd
          import matplotlib.pyplot as plt
          from sklearn.cluster import KMeans
```

```python
In [5]:   df = pd.read_csv("/content/customer_data.csv")
```

```python
In [6]:   X = df[['Annual Income (k$)', 'Spending Score (1-100)']]
```

```python
In [7]:   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_)
```

```python
In [8]:   plt.plot(range(1, 11), wcss, marker='o')
          plt.title('Elbow Method')
          plt.xlabel('Number of clusters')
          plt.ylabel('WCSS')
          plt.show()
```
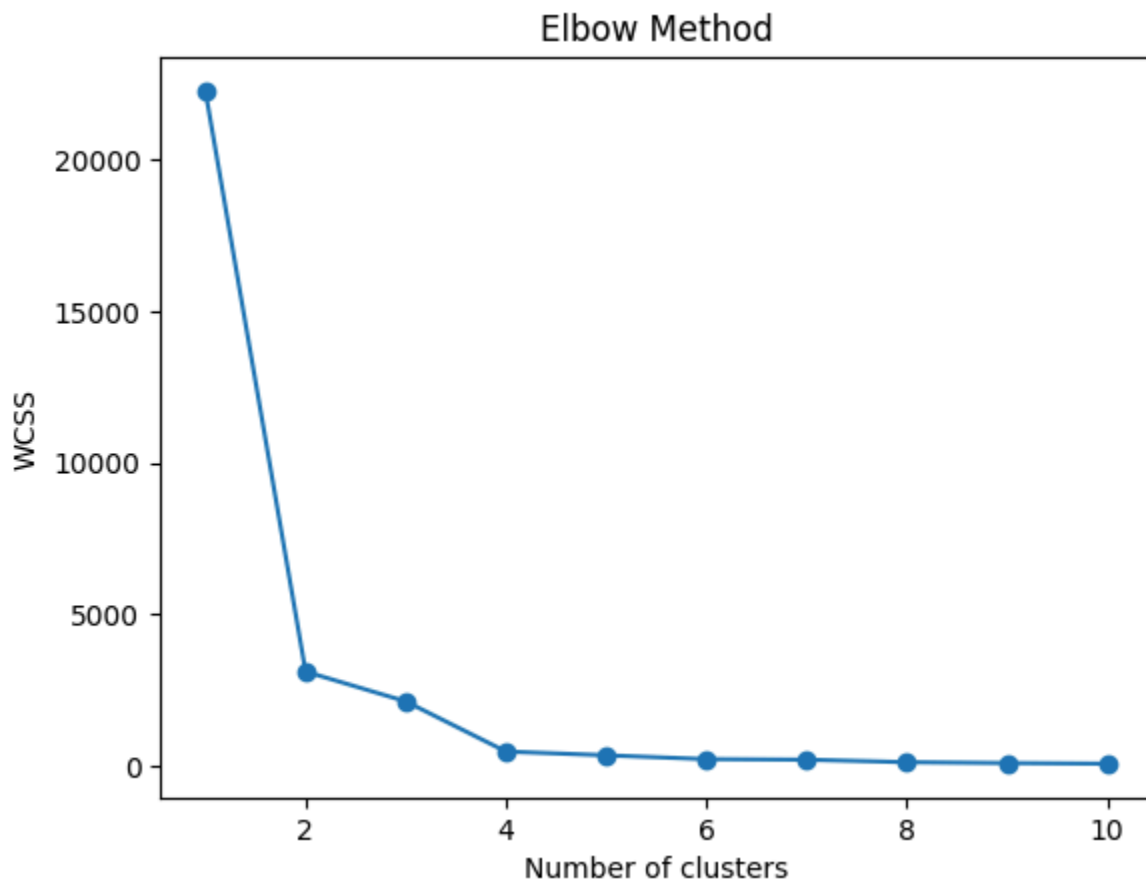


```python
In [9]:   kmeans = KMeans(n_clusters=3, init='k-means++',random_state=42)
          df['Cluster'] = kmeans.fit_predict(X)
```

```
In [10]: plt.figure(figsize=(8,6))
         plt.scatter(X['Annual Income (k$)'],X['Spending Score (1-100)'],
                     c=df['Cluster'],cmap='viridis',s=100, alpha=0.7, edgecolors='k')
         plt.scatter(kmeans.cluster_centers_[:,0], kmeans.cluster_centers_[:,1],
                     s=300, c='red', marker='X', label='Centroids')
         plt.title('Customer Segments (K-Means Clustering)')
         plt.xlabel('Annual Income (k$)')
         plt.ylabel('Spending Score (1-100)')
         plt.legend()
         plt.show()
```
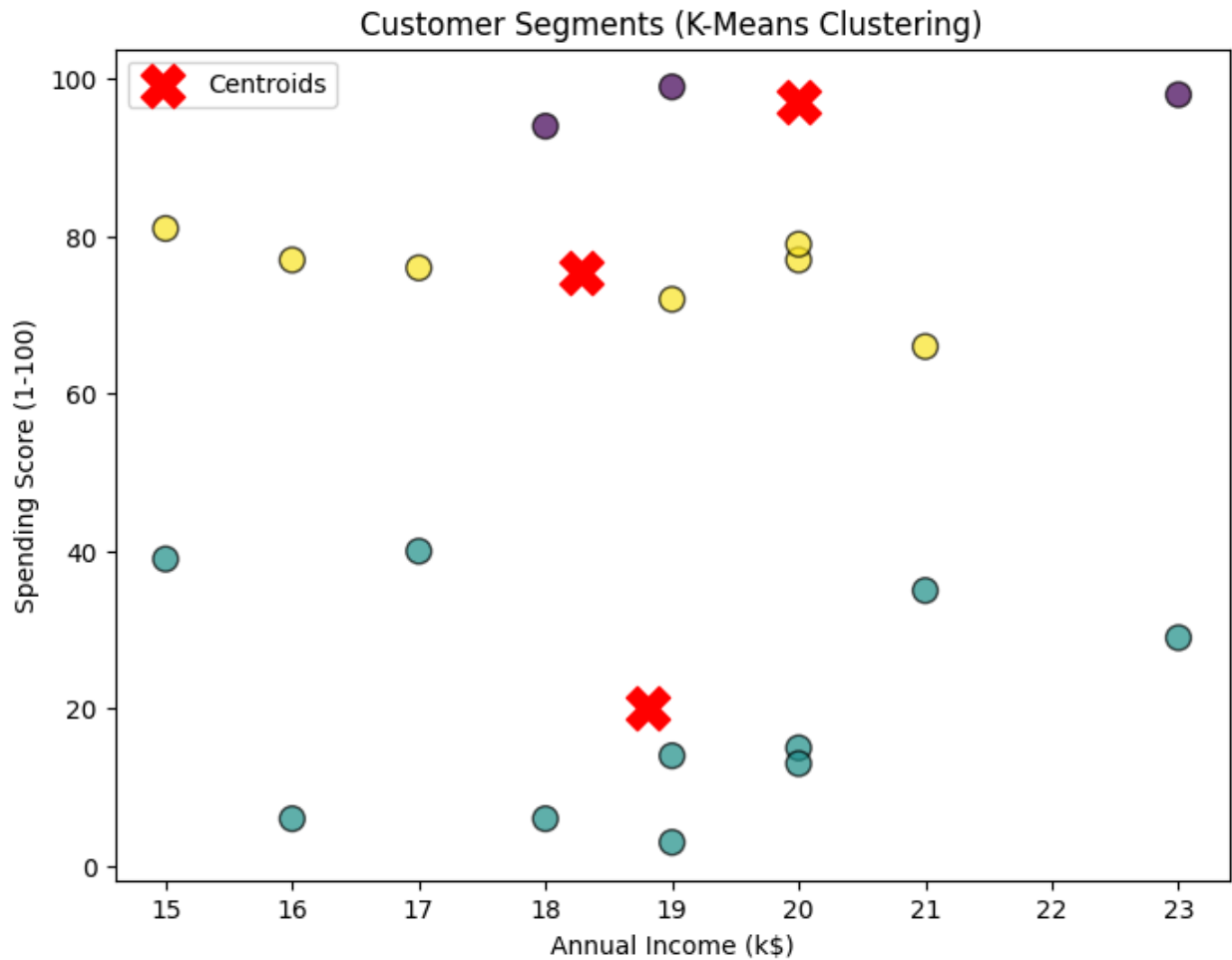


```
In [11]: print(df[['CustomerID','Annual Income (k$)','Spending Score (1-100)','Cluster'
```

	CustomerID	Annual Income (k$)	Spending Score (1-100)	Cluster
0	1	15	39	1
1	2	15	81	2
2	3	16	6	1
3	4	16	77	2
4	5	17	40	1
5	6	17	76	2
6	7	18	6	1
7	8	18	94	0
8	9	19	3	1
9	10	19	72	2
10	11	19	14	1
11	12	19	99	0
12	13	20	15	1
13	14	20	77	2
14	15	20	13	1
15	16	20	79	2
16	17	21	35	1
17	18	21	66	2
18	19	23	29	1
19	20	23	98	0

In [11]: