

```
1 import pandas as pd
2 from sklearn.cluster import KMeans
3 import matplotlib.pyplot as plt
```

```
1 import pandas as pd
2 df = pd.read_csv ("/content/drive/MyDrive/Colab Notebooks/Dataset/Mall_Customers.csv")
3 df.head()
```

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

```
1 # Select relevant features for clustering
2 features = data[['Age', 'Annual Income (k$)', 'Spending Score (1-100)']]
```

```
1 # Choose the number of clusters (you may need to adjust this)
2 k = 3
```

```
1 # Create a KMeans instance
2 kmeans = KMeans(n_clusters=k, random_state=42)
```

```
1 # Fit the model to the data
2 kmeans.fit(features)
```

```
/usr/local/lib/python3.10/dist-packages/sklearn/cluster/_kmeans.py:870: FutureWarning:
warnings.warn(
```

```
▼ KMeans
KMeans(n_clusters=3, random_state=42)
```

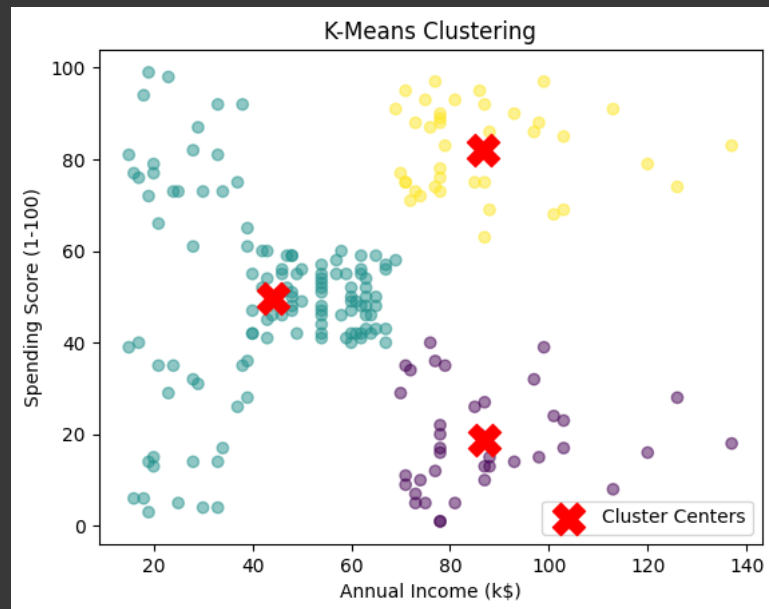
```
1 # Add cluster labels to the original dataset
2 data['Cluster'] = kmeans.labels_
```

```
1 # Print the cluster centers (optional)
2 print("Cluster Centers:")
3 print(kmeans.cluster_centers_)
```

```
Cluster Centers:
[[40.39473684  87.          18.63157895]
 [40.32520325  44.15447154  49.82926829]
 [32.69230769  86.53846154  82.12820513]]
```

1

```
1 # Visualize the clusters (2D plot, considering two features at a time)
2 plt.scatter(data['Annual Income (k$)'], data['Spending Score (1-100)'], c=data['Cluster'], cmap='viridis', alpha=0.5)
3 plt.scatter(kmeans.cluster_centers_[0, 1], kmeans.cluster_centers_[0, 2], s=300, c='red', marker='X', label='Cluster Centers')
4 plt.xlabel('Annual Income (k$)')
5 plt.ylabel('Spending Score (1-100)')
6 plt.title('K-Means Clustering')
7 plt.legend()
8 plt.show()
```



```
1 inertia = kmeans.inertia_
2 print("Inertia:", inertia)
```

Inertia: 143342.751571706

```
1 from sklearn.metrics import silhouette_score
2 # Assuming 'features' is your feature matrix
3 silhouette_avg = silhouette_score(features, kmeans.labels_)
4 print("Silhouette Score:", silhouette_avg)
```

Silhouette Score: 0.3839349967742105

