

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

```
In [ ]: dataset = pd.read_csv("datasets/OnlineRetail.csv",encoding = 'latin1')
dataset.head()
```

```
Out[ ]: InvoiceNo StockCode Description Quantity InvoiceDate UnitPrice CustomerID Cou
```

0	536365	85123A	WHITE HANGING HEART T- LIGHT HOLDER	6	01-12-2010 08:26	2.55	17850.0	Un King
1	536365	71053	WHITE METAL LANTERN	6	01-12-2010 08:26	3.39	17850.0	Un King
2	536365	84406B	CREAM CUPID HEARTS COAT HANGER	8	01-12-2010 08:26	2.75	17850.0	Un King
3	536365	84029G	KNITTED UNION FLAG HOT WATER BOTTLE	6	01-12-2010 08:26	3.39	17850.0	Un King
4	536365	84029E	RED WOOLLY HOTTIE WHITE HEART.	6	01-12-2010 08:26	3.39	17850.0	Un King



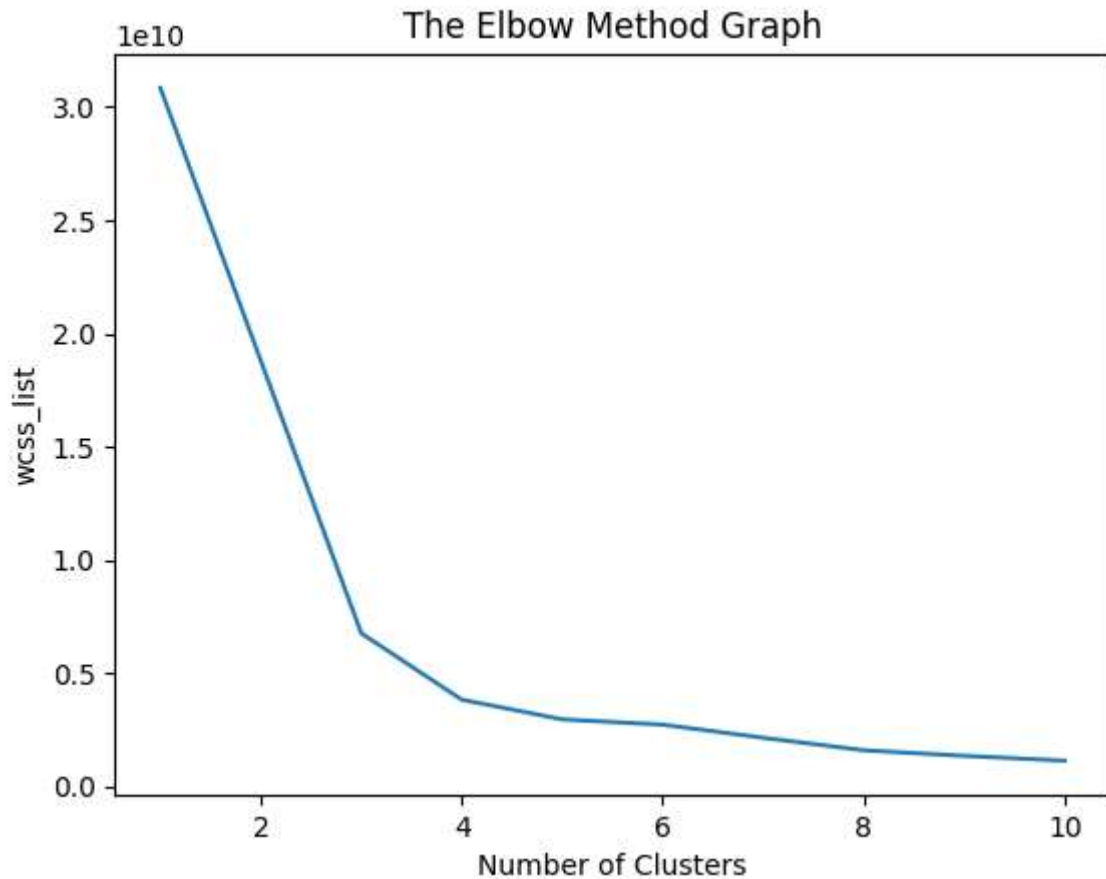
```
In [ ]: X= dataset.iloc[:,[3,5]].values
```

```
In [ ]: wcss_list = []

for i in range(1,11):
    kmeans = KMeans(n_clusters=i,init='k-means++',random_state = 42)
    kmeans.fit(X)
    wcss_list.append(kmeans.inertia_)
```

```
In [ ]: mtp.plot(range(1,11), wcss_list)
mtp.title("The Elbow Method Graph")
mtp.xlabel("Number of Clusters")
```

```
mtp.ylabel("wcss_list")
mtp.show()
```



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

```
In [ ]: mtp.scatter(X[y_predict == 0,0],X[y_predict == 0,1], s = 100,c = 'red',label='Clust
mtp.scatter(X[y_predict == 1,0],X[y_predict == 1,1], s = 100,c = 'green',label='Clu
mtp.scatter(X[y_predict == 2,0],X[y_predict == 2,1], s = 100,c = 'blue',label='Clus

mtp.title("Cluster of Online Retail")
mtp.xlabel("Quantity")
mtp.ylabel("Unit Price")
mtp.legend()
mtp.show()
```

