In [6]:
```
import pandas as pd
data=pd.read_csv('Shopping data.csv')
data
```

Out[6]:

	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
...	...	...	...	...	...
195	196	Female	35	120	79
196	197	Female	45	126	28
197	198	Male	32	126	74
198	199	Male	32	137	18
199	200	Male	30	137	83

In [7]:
```
data.head(10)
```

Out[7]:

	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
5	6	Female	22	17	76
6	7	Female	35	18	6
7	8	Female	23	18	94
8	9	Male	64	19	3
9	10	Female	30	19	72

In [8]:
```
data.shape
```

Out[8]:
```
(200, 5)
```

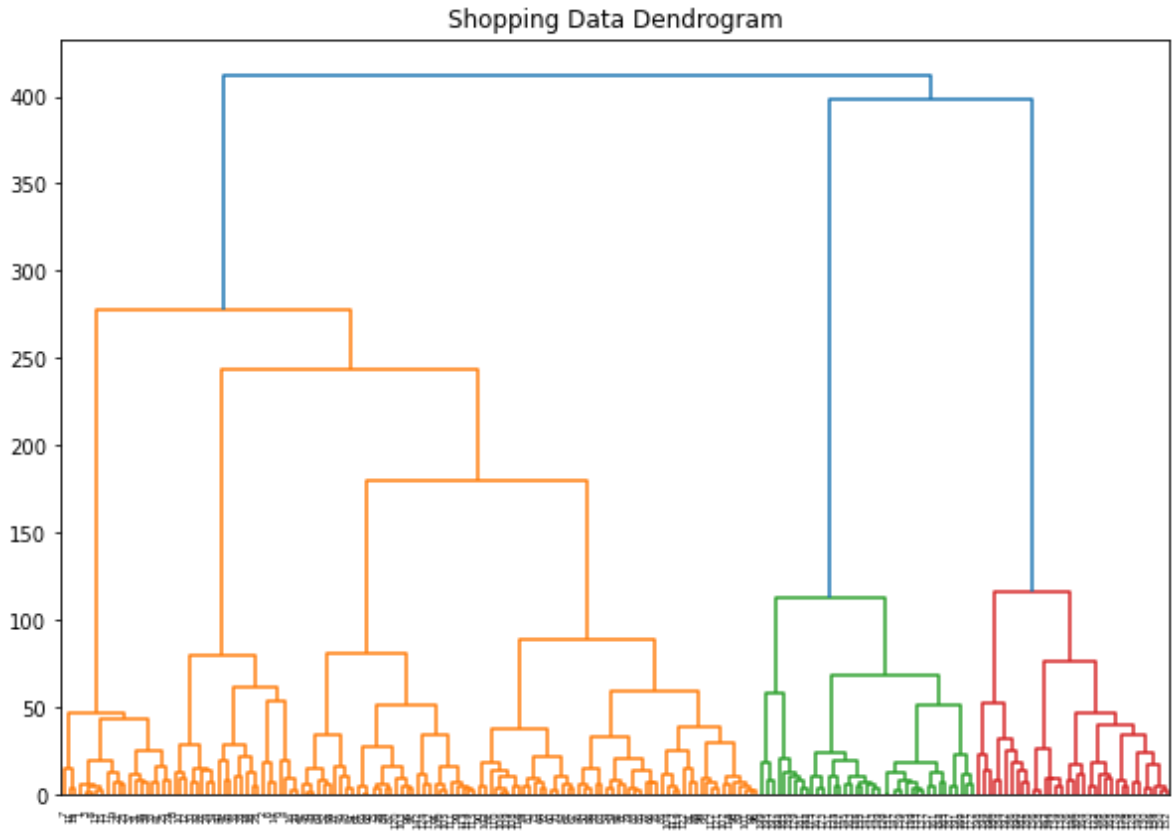
In [9]:
```
data=data.iloc[:,2:5]
```

In [10]:
```
data.head(10)
```

Out[10]:

	Age	Annual Income (k$)	Spending Score (1-100)
0	19	15	39
1	21	15	81
2	20	16	6
3	23	16	77
4	31	17	40
5	22	17	76
6	35	18	6
7	23	18	94
8	64	19	3
9	30	19	72

In [11]:
```
import matplotlib.pyplot as plt
import scipy.cluster.hierarchy as shc
plt.figure(figsize=(10,7))
plt.title("Shopping Data Dendrogram")
dend=shc.dendrogram(shc.linkage(data,
                method='ward'))
```



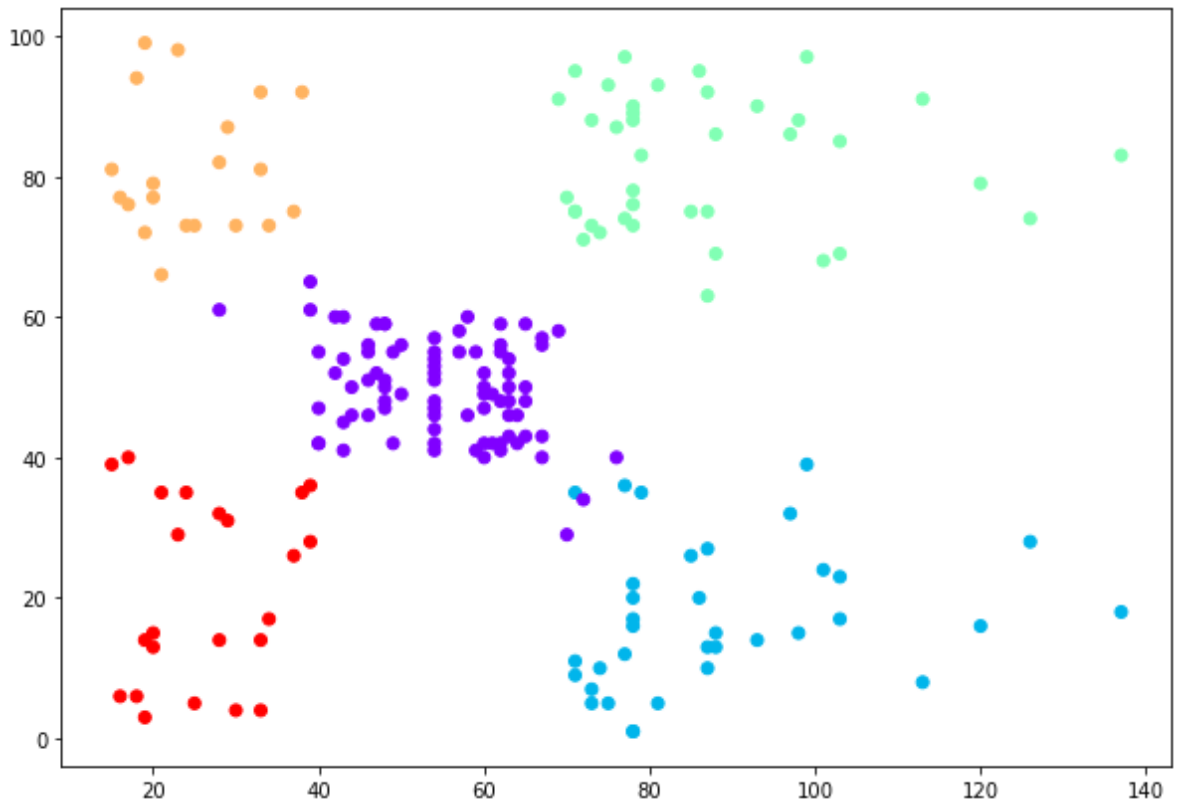
In [12]:
```
from sklearn.cluster import AgglomerativeClustering
cluster_model=AgglomerativeClustering(n_clusters=5,affinity='euclidean',
    linkage='ward')
```

In [13]:
```
cluster_model.fit(data)
print(cluster_model.labels_)
```

```
[4 3 4 3 4 3 4 3 4 3 4 3 4 3 4 3 4 3 4 3 4 3 4 3 4 0 4 3 4 3 4 3 4 3 4
 3 4 3 4 3 4 0 4 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0
 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0
 0 0 0 0 0 0 0 0 0 0 0 0 2 0 2 1 2 1 2 1 2 0 2 1 2 1 2 1 2 1 2 0 2 1 2 1 2
 1 2 1 2 1 2 1 2 1 2 1 2 1 2 1 2 1 2 1 2 1 2 1 2 1 2 1 2 1 2 1 2 1 2 1
 2 1 2 1 2 1 2 1 2 1 2 1 2 1 2]
```

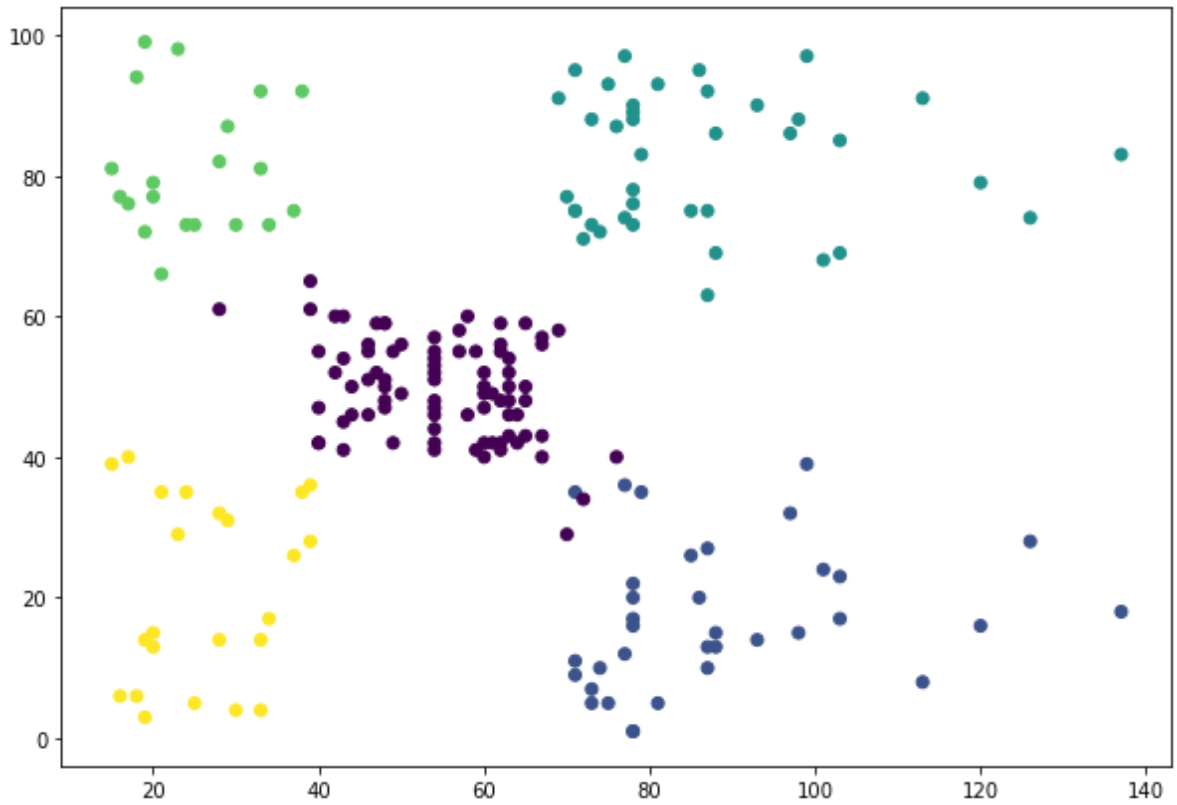
In [14]:
```
plt.figure(figsize=(10,7))
plt.scatter(data['Annual Income (k$)'],data['Spending Score (1-100)'],
            c=cluster_model.labels_,cmap='rainbow')
```

Out[14]: <matplotlib.collections.PathCollection at 0x1d3915c4070>



In [15]:
```
plt.figure(figsize=(10,7))
plt.scatter(data['Annual Income (k$)'],data['Spending Score (1-100)'],
            c=cluster_model.labels_)
```

Out[15]:
```
<matplotlib.collections.PathCollection at 0x1d3923846a0>
```



In [ ]:
```
from sklearn.cluster import AgglomerativeClustering
cluster_model=AgglomerativeClustering(n_clusters=4,affinity='euclidean',
linkage='ward')
```

In [16]:
```
cluster_model.fit(data)
print(cluster_model.labels_)
```

```
[4 3 4 3 4 3 4 3 4 3 4 3 4 3 4 3 4 3 4 3 4 3 4 3 4 0 4 3 4 3 4 3 4 3 4
 3 4 3 4 3 4 0 4 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0
 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0
 0 0 0 0 0 0 0 0 0 0 0 0 2 0 2 1 2 1 2 1 2 0 2 1 2 1 2 1 2 1 2 0 2 1 2 1 2
 1 2 1 2 1 2 1 2 1 2 1 2 1 2 1 2 1 2 1 2 1 2 1 2 1 2 1 2 1 2 1 2 1 2 1
 2 1 2 1 2 1 2 1 2 1 2 1 2 1 2]
```

In [ ]: