Aravinth R - 19MIC0053

Lab Experiment - 6

In [ ]:
```python
import pandas as pd
import numpy as np
import matplotlib.pyplot as plt
%matplotlib inline
```

In [ ]:
```python
data = pd.read_csv('/Users/aravinth/Desktop/Data Warehousing/Assignment 5/(
data.head()
```

Out[ ]:

	Channel	Region	Fresh	Milk	Grocery	Frozen	Detergents_Paper	Delicassen
0	2	3	12669	9656	7561	214	2674	1338
1	2	3	7057	9810	9568	1762	3293	1776
2	2	3	6353	8808	7684	2405	3516	7844
3	1	3	13265	1196	4221	6404	507	1788
4	2	3	22615	5410	7198	3915	1777	5185

In [ ]:
```python
from sklearn.preprocessing import normalize
data_scaled = normalize(data)
data_scaled = pd.DataFrame(data_scaled, columns=data.columns)
data_scaled.head()
```

Out[ ]:

|	Channel	Region	Fresh	Milk	Grocery	Frozen	Detergents_Paper	Delicas
0	0.000112	0.000168	0.708333	0.539874	0.422741	0.011965	0.149505	0.074
1	0.000125	0.000188	0.442198	0.614704	0.599540	0.110409	0.206342	0.111
2	0.000125	0.000187	0.396552	0.549792	0.479632	0.150119	0.219467	0.489
3	0.000065	0.000194	0.856837	0.077254	0.272650	0.413659	0.032749	0.115
4	0.000079	0.000119	0.895416	0.214203	0.284997	0.155010	0.070358	0.205

In [ ]:
```python
import scipy.cluster.hierarchy as shc
plt.figure(figsize=(10, 7))
plt.title("Dendrograms")
dend = shc.dendrogram(shc.linkage(data_scaled, method='ward'))
```
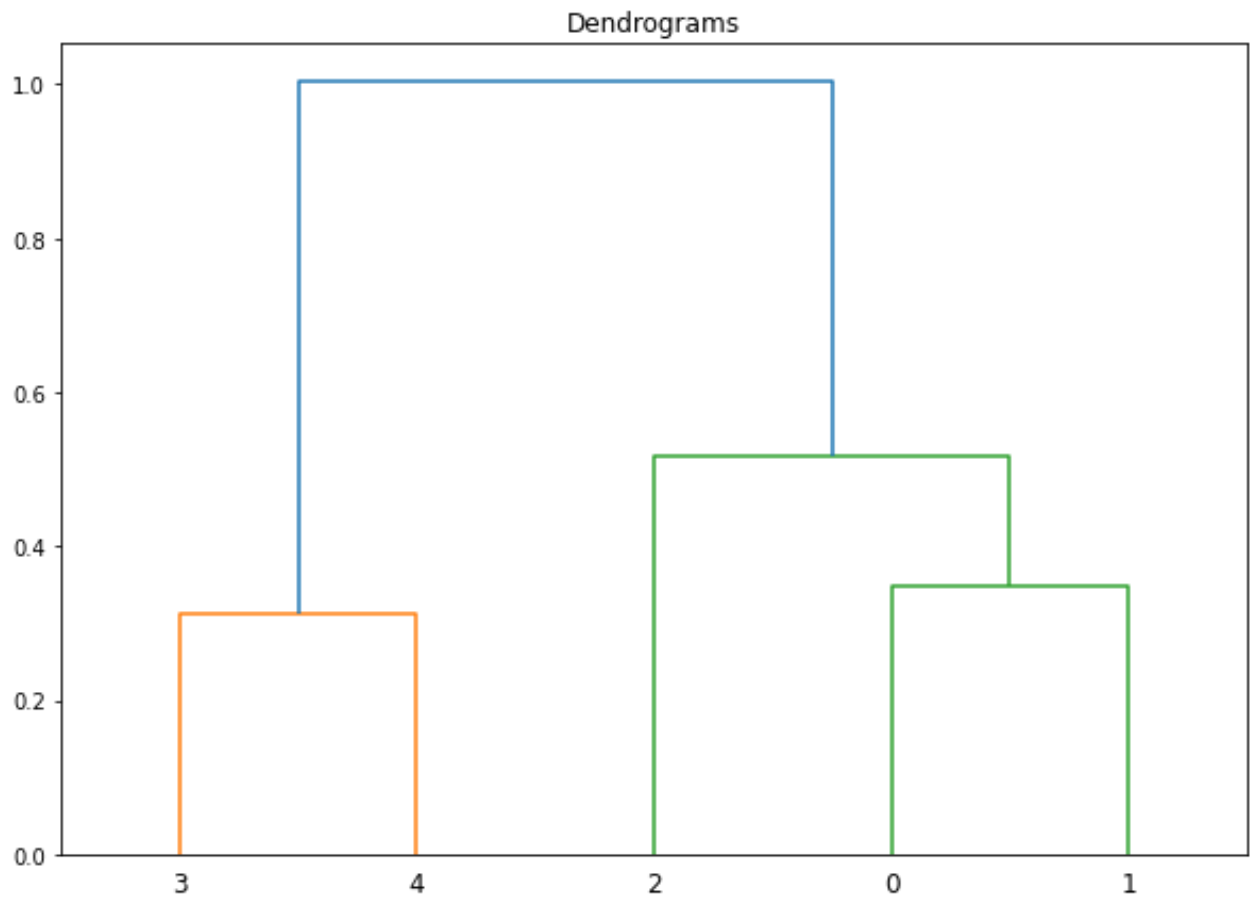
Dendrograms


```
In [ ]:    plt.figure(figsize=(10, 7))
           plt.title("Dendrograms")
           dend = shc.dendrogram(shc.linkage(data_scaled, method='ward'))
           plt.axhline(y=6, color='r', linestyle='--')
```

Out[ ]:   `<matplotlib.lines.Line2D at 0x7fda18f95400>`

Dendrograms
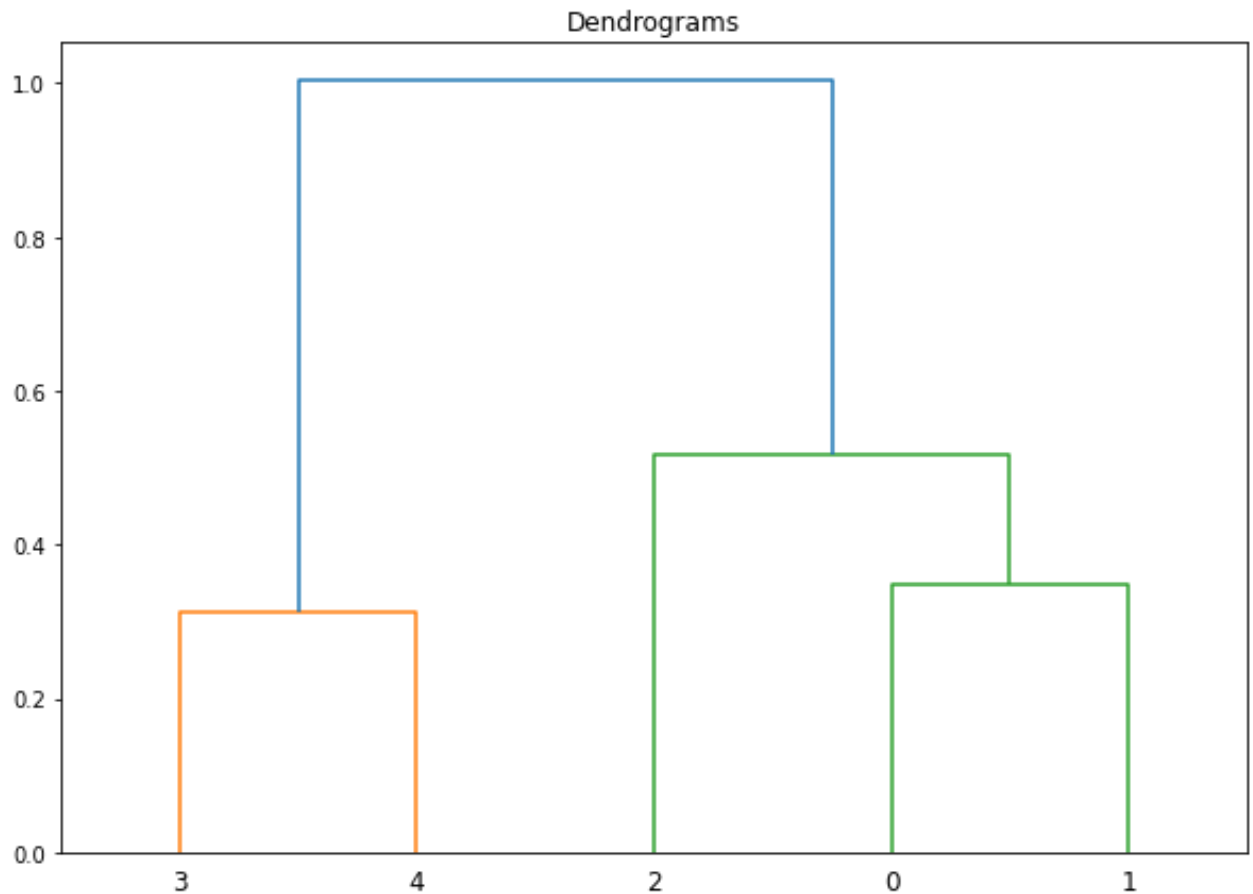
In [ ]:
```python
from sklearn.cluster import AgglomerativeClustering
cluster = AgglomerativeClustering(n_clusters=2, affinity='euclidean', linka
cluster.fit_predict(data_scaled)
```

Out[ ]:   `array([0, 0, 0, 1, 1])`

In [ ]:
```python
plt.figure(figsize=(10, 7))
plt.scatter(data_scaled['Milk'], data_scaled['Grocery'], c=cluster.labels_
```

Out[ ]:   `<matplotlib.collections.PathCollection at 0x7fd9c8f55790>`