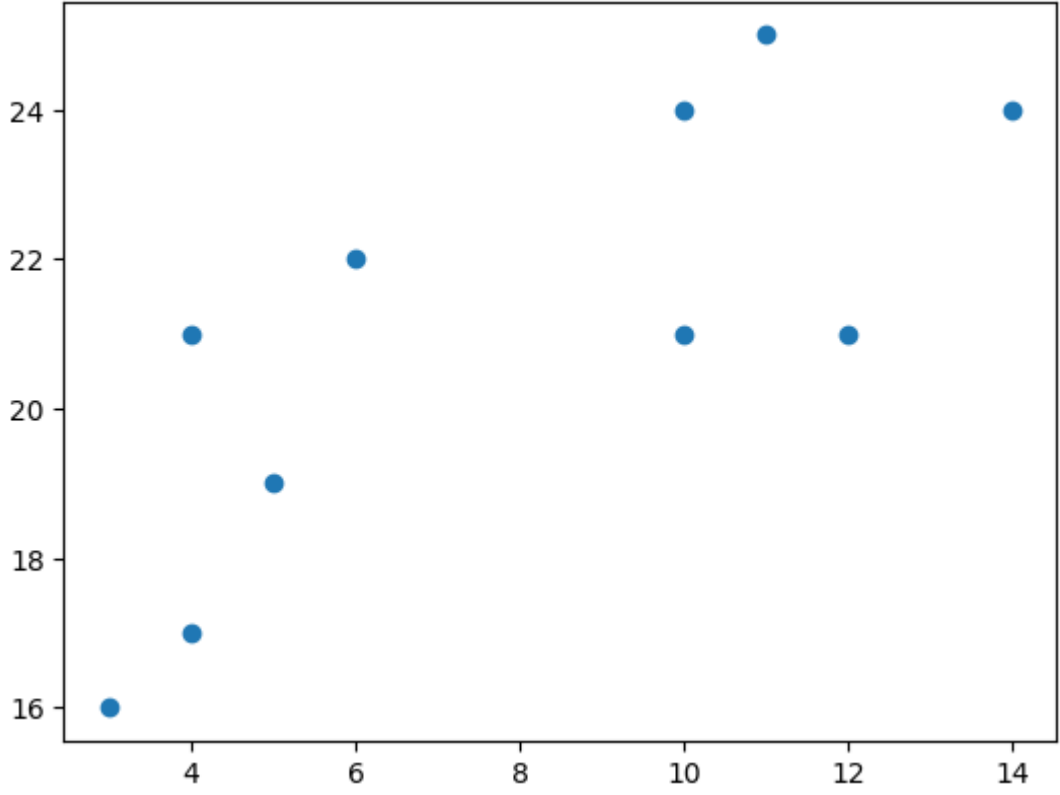


# Hierarchical Clustering

```
In [2]: import numpy as np
import matplotlib.pyplot as plt
```

```
In [3]: x = [4, 5, 10, 4, 3, 11, 14, 6, 10, 12]
y = [21, 19, 24, 17, 16, 25, 24, 22, 21, 21]

plt.scatter(x, y)
plt.show()
```



```
In [4]: from scipy.cluster.hierarchy import dendrogram, linkage
```

```
In [5]: x
```

```
Out[5]: [4, 5, 10, 4, 3, 11, 14, 6, 10, 12]
```

```
In [6]: y
```

```
Out[6]: [21, 19, 24, 17, 16, 25, 24, 22, 21, 21]
```

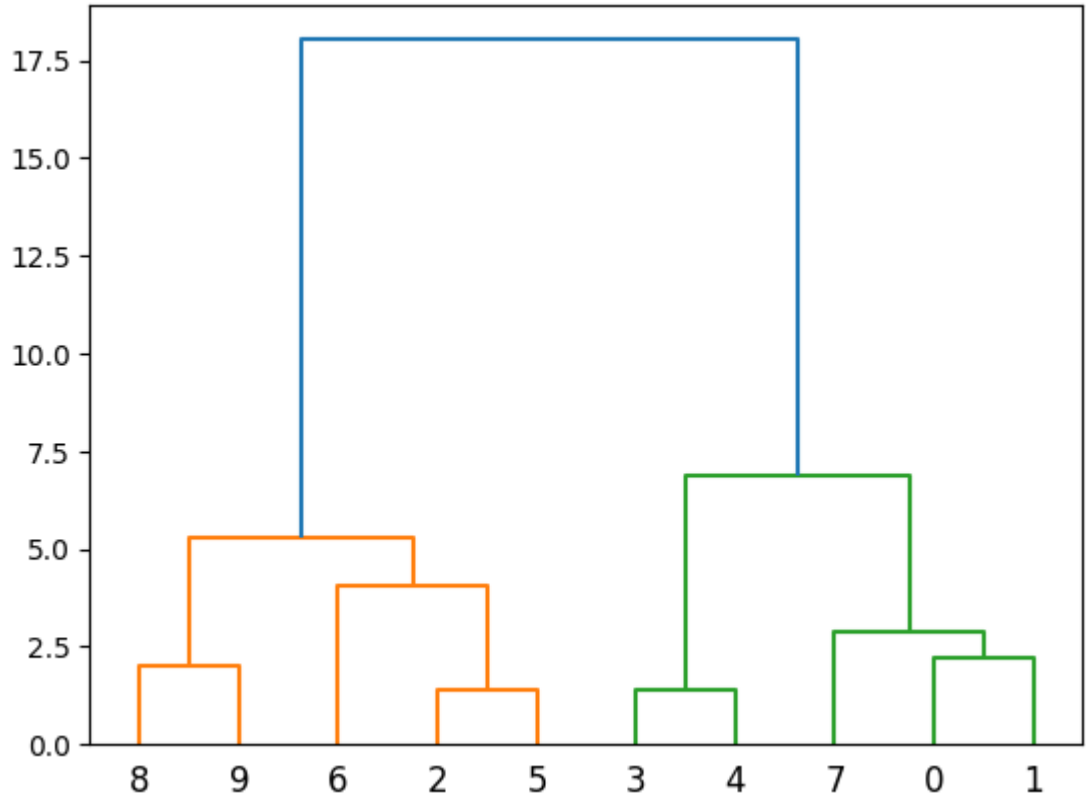
```
In [7]: data = list(zip(x, y))
data
```

```
Out[7]: [(4, 21),
(5, 19),
(10, 24),
(4, 17),
(3, 16),
(11, 25),
(14, 24),
(6, 22),
(10, 21),
(12, 21)]
```

```
In [8]: linkage_data = linkage(data, method="ward", metric="euclidean")

dendrogram(linkage_data)
plt.show
```

```
Out[8]: <function matplotlib.pyplot.show(close=None, block=None)>
```



Se aprecian dos grupos diferenciados

## Aglomerativas

```
In [10]: from sklearn.cluster import AgglomerativeClustering
```

```
In [11]: x
```

```
Out[11]: [4, 5, 10, 4, 3, 11, 14, 6, 10, 12]
```

```
In [12]: y
```

```
Out[12]: [21, 19, 24, 17, 16, 25, 24, 22, 21, 21]
```

```
In [13]: data
```

```
Out[13]: [(4, 21),
(5, 19),
(10, 24),
(4, 17),
(3, 16),
(11, 25),
(14, 24),
(6, 22),
(10, 21),
(12, 21)]
```

```
In [14]: hierarchical = AgglomerativeClustering(n_clusters=2, affinity="euclidean", linkage="ward")
labels= hierarchical.fit_predict(data)
labels
```

```
Out[14]: array([0, 0, 1, 0, 0, 1, 1, 0, 1, 1])
```

```
In [15]: plt.scatter(x, y, c=labels)
plt.show()
```

