

# Agglomerative Clustering

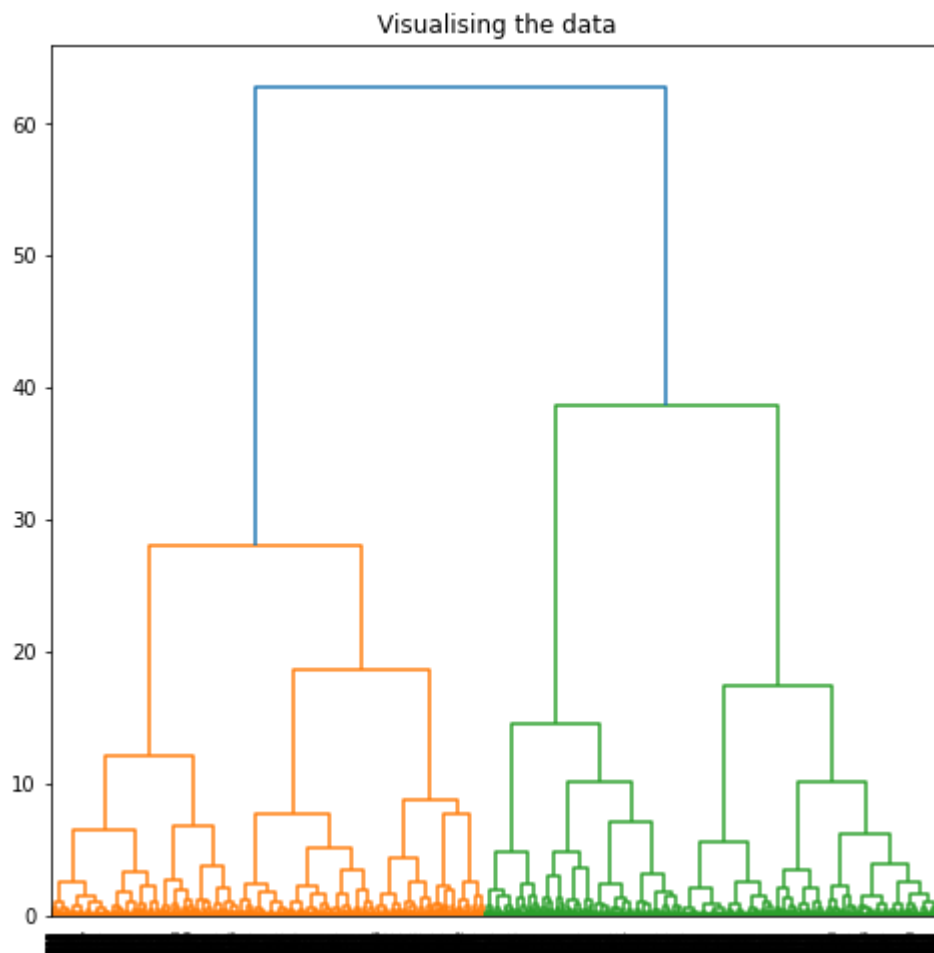
```
In [1]: import pandas as pd
import numpy as np
import matplotlib.pyplot as plt
from sklearn.decomposition import PCA
from sklearn.cluster import AgglomerativeClustering
from sklearn.preprocessing import StandardScaler, normalize
from sklearn.metrics import silhouette_score
import scipy.cluster.hierarchy as shc
```

```
In [2]: X = pd.read_csv('CC_GENERAL.csv')
X = X.drop('CUST_ID', axis = 1)
X.fillna(method = 'ffill', inplace = True)
```

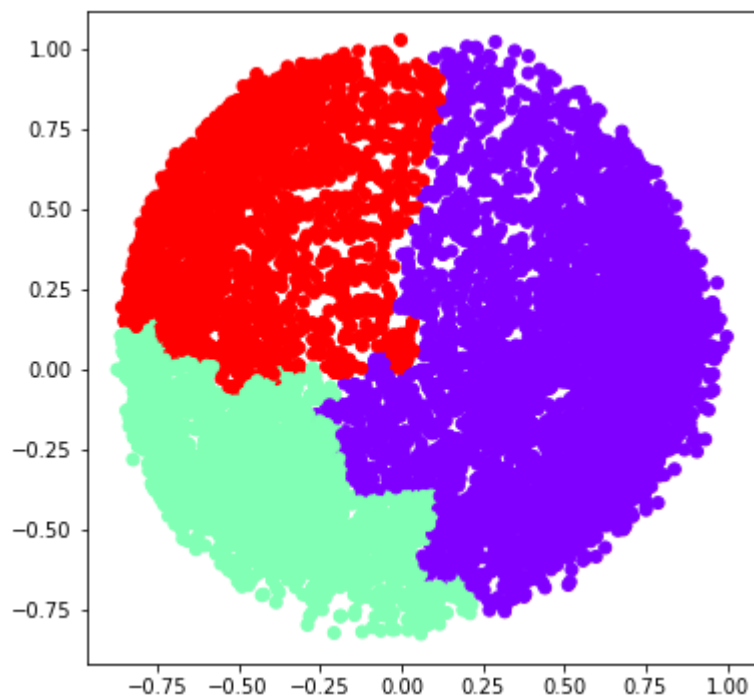
```
In [3]: scaler = StandardScaler()
X_scaled = scaler.fit_transform(X)
X_normalized = normalize(X_scaled)
X_normalized = pd.DataFrame(X_normalized)
```

```
In [4]: pca = PCA(n_components = 2)
X_principal = pca.fit_transform(X_normalized)
X_principal = pd.DataFrame(X_principal)
X_principal.columns = ['P1', 'P2']
```

```
In [5]: plt.figure(figsize =(8, 8))
plt.title('Visualising the data')
Dendrogram = shc.dendrogram((shc.linkage(X_principal, method ='ward')))
```

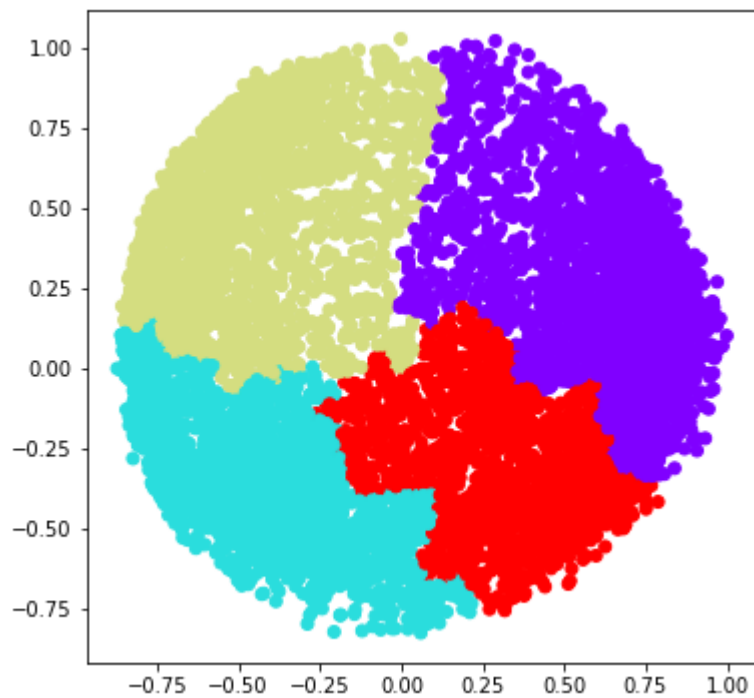


```
In [6]: ac = AgglomerativeClustering(n_clusters = 3)
plt.figure(figsize = (6, 6))
plt.scatter(X_principal['P1'], X_principal['P2'],
            c = ac.fit_predict(X_principal), cmap = 'rainbow')
plt.show()
```



```
In [7]: ac = AgglomerativeClustering(n_clusters = 4)

plt.figure(figsize =(6, 6))
plt.scatter(X_principal['P1'], X_principal['P2'],
            c = ac.fit_predict(X_principal), cmap ='rainbow')
plt.show()
```



```
In [ ]:
```