

**EX NO:** HIERARCHICAL CLUSTERING: WRITE A PROGRAM TO IMPLEMENT AGGLOMERATIVE CLUSTERING  
**DATE:** TECHNIQUE, WRITE A PROGRAM TO IMPLEMENT DIVISIVE HIERARCHICAL CLUSTERING TECHNIQUE

**AIM:**

## **BACKGROUND THEORY:**

### **1) HIERARCHICAL CLUSTERING:**

Hierarchical clustering is a connectivity-based clustering model that groups the data points together that are close to each other based on the measure of similarity or distance. The assumption is that data points that are close to each other are more similar or related than data points that are farther apart.

A dendrogram, a tree-like figure produced by hierarchical clustering, depicts the hierarchical relationships between groups. Individual data points are located at the bottom of the dendrogram, while the largest clusters, which include all the data points, are located at the top. In order to generate different numbers of clusters, the dendrogram can be sliced at various heights.

### **2) DIVISIVE HIERARCHICAL CLUSTERING:**

It is also known as a top-down approach. This algorithm also does not require to prespecify the number of clusters. Top-down clustering requires a method for splitting a cluster that contains the whole data and proceeds by splitting clusters recursively until individual data have been split into singleton clusters.

## **PROCEDURE:**

### **1. Load Data:**

- o Drag the "File" widget to the canvas.
- o Load your dataset file (e.g., dataset.csv).

### **2. Agglomerative Clustering:**

- o Drag the "Hierarchical Clustering" widget to the canvas.
- o Connect the "File" widget to the "Hierarchical Clustering" widget.
- o Configure the widget to use agglomerative clustering.

### **3. Visualize Clustering:**

- o Drag the "Dendrogram" widget to the canvas.
- o Connect the "Hierarchical Clustering" widget to the "Dendrogram" widget.

## OUTPUT:

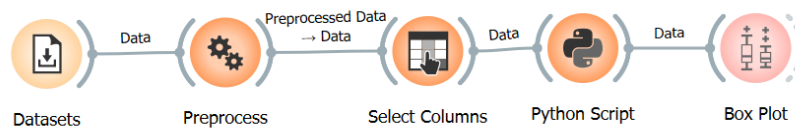


FIG 9.1: IMPLEMENTATION OF DIVISIVE HIERARCHICAL CLUSTERING TECHNIQUE

## PYTHON SCRIPT:

```
from sklearn.cluster import AgglomerativeClustering
import matplotlib.pyplot as plt
import scipy.cluster.hierarchy as sch
data = in_data.X
plt.figure(figsize=(10, 7))
dendrogram = sch.dendrogram(sch.linkage(data, method='ward'))
plt.show()
hc = AgglomerativeClustering(n_clusters=5, affinity='euclidean', linkage='ward')
cluster_labels = hc.fit_predict(data)
out_data = in_data.copy()
out_data.Y = cluster_labels
```

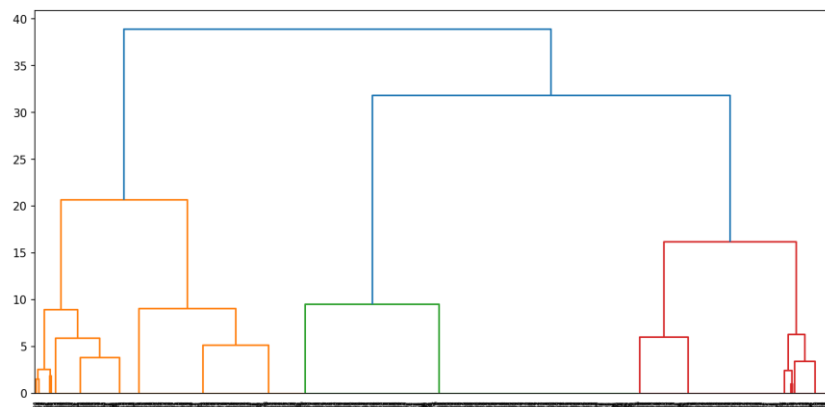


FIG 9.2: CLUSTERING USING PYTHON SCRIPT

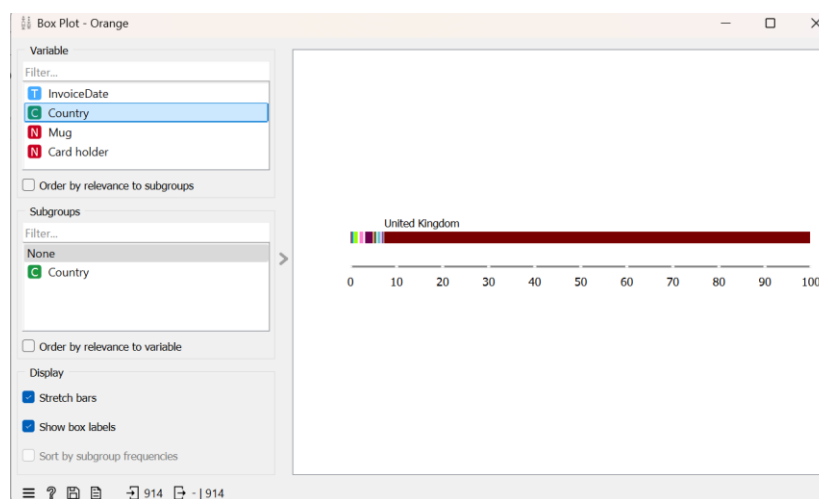


FIG: 9.3: DISPLAYING USING A BOX PLOT

## RESULT: