Experiment Number: 8					
Date of Performance:					
Date of Submission:					
Program Execution/ formation/ correction/ ethical practices (07)	**Documentation (02)**	**Timely Submission (03)**	**Viva Answer to sample questions (03)**	**Experiment Total (15)**	**Sign**

Experiment 8

Aim: Implementation of Single Link Agglomerative Hierarchical Clustering method

Lab outcome: CSL 503.3: Implement clustering algorithms on a given set of data sample.

Problem Statement: To implement Single Link Agglomerative Hierarchical Clustering.

Theory:
Hierarchical clustering is another unsupervised machine learning algorithm, which is used to group the unlabelled datasets into a cluster and also known as **hierarchical cluster analysis** or HCA.
In this algorithm, we develop the hierarchy of clusters in the form of a tree, and this tree-shaped structure is known as the **dendrogram**.

Agglomerative: Agglomerative is a **bottom-up** approach, in which the algorithm starts with taking all data points as single clusters and merging them until one cluster is left.

Single Linkage: It is the Shortest Distance between the closest points of the clusters. Consider the below image:
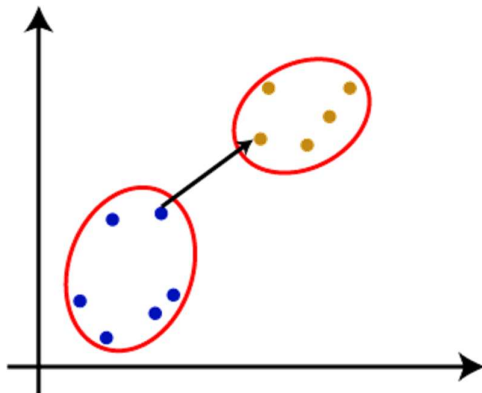
Program Listing and Output:

Code:

```
#importing
import numpy as np
import pandas as pd
import matplotlib.pyplot as plt

#matplotlib inline
import scipy.cluster.hierarchy as shc
from scipy.spatial.distance import squareform, pdist

#X-axis of point
x = [0.49,0.32,0.38,0.24,0.18,0.45]
```

```python
#Y-axis of point
y = [0.35,0.32,0.31,0.29,0.46,0.54]
point = ['P1','P2','P3','P4','P5','P6']
data = pd.DataFrame({'Point':point, 'x':np.round(x,2), 'y':np.round(y,2)})
data = data.set_index('Point')
#Printing all points

print(data)
print('\n\n')

#Plotting the points
plt.figure(figsize=(8,5))
plt.scatter(data['x'], data['y'], c='r', marker='*')
plt.xlabel('X-axis',fontsize=14,color='darkred')
plt.ylabel('Y-axis',fontsize=14,color='darkred')
plt.title('Plotting of Points',fontsize=16,color='purple')

for j in data.itertuples():
    plt.annotate(j.Index, (j.x, j.y), fontsize=15)
dist = pd.DataFrame(squareform(pdist(data[['x','y']]), 'euclidean'),
columns=data.index.values,
index=data.index.values)

#Displaying the dendrogram
plt.figure(figsize=(12,5))
plt.title("Dendrogram with Single Linkage",fontsize=18,color='purple')
dend = shc.dendrogram(shc.linkage(data[['x', 'y']], method='single'),
labels=data.index)
plt.show()
```
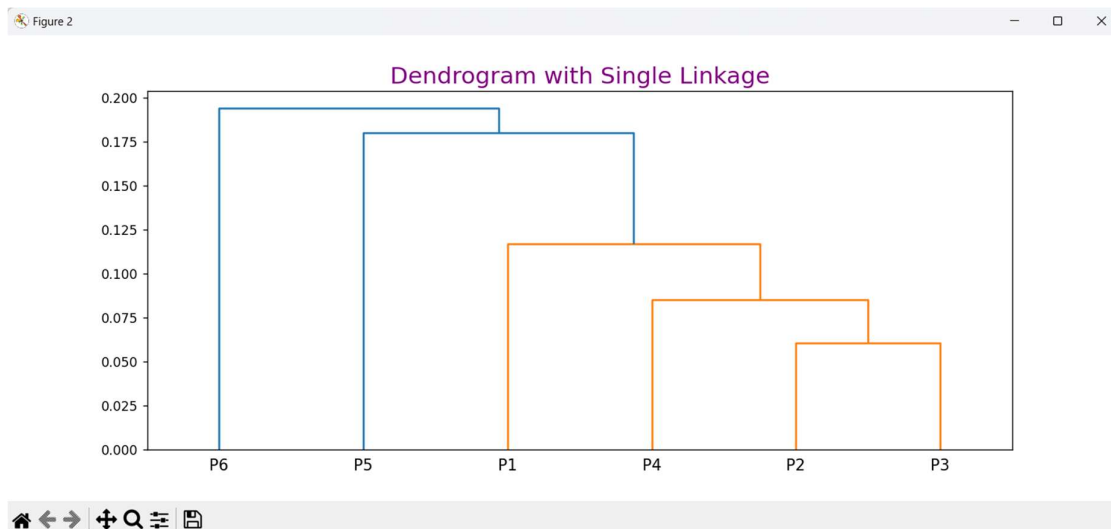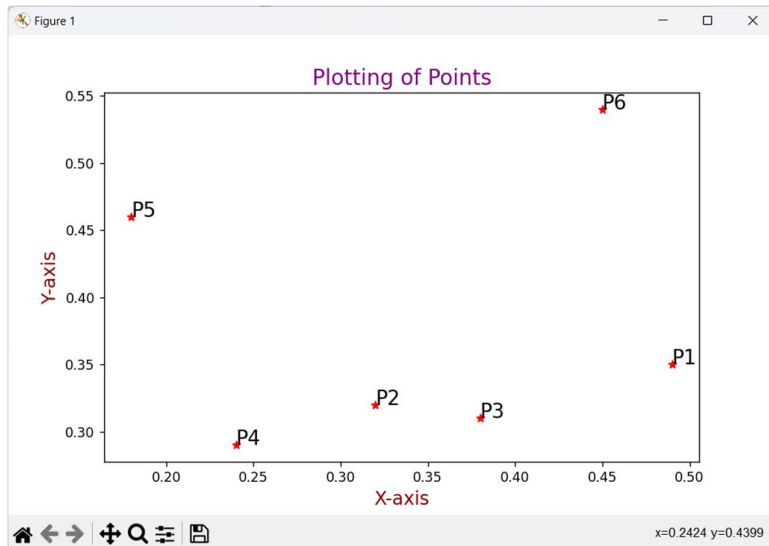
```
        x     y
Point
P1    0.49  0.35
P2    0.32  0.32
P3    0.38  0.31
P4    0.24  0.29
P5    0.18  0.46
P6    0.45  0.54
```

Figure 1 — Plotting of Points

Figure 2 — Dendrogram with Single Linkage

Conclusion: Here we implemented Single Link Agglomerative Hierarchical Clustering