Problem set 6
Hierarchical clustering

Excercise 1

Given the following set of 2-dimnesional points $X_1 = (2, 10), X_2 = (2, 5), X_3 = (8, 4), X_4 = (5, 8), X_5 = (7, 5), X_6 = (6, 4), X_7 = (1, 2), X_8 = (4, 9)$ apply single-linkage agglomerative clustering using Euclidean distance. Draw a dendrogram representing the resulting hierarchy of clusters.

Solution The euclidean distance between any pair of points is given below.

$$\begin{pmatrix} 0 & 5 & 8.49 & 3.61 & 7.07 & 7.21 & 8.06 & 2.24 \\ 5 & 0 & 6.08 & 4.24 & 5 & 4.12 & 3.16 & 4.47 \\ 8.49 & 6.08 & 0 & 5 & 1.41 & 2 & 7.28 & 6.4 \\ 7.21 & 4.12 & 2 & 4.12 & 1.41 & 0 & 5.39 & 5.39 \\ 8.06 & 3.16 & 7.28 & 7.21 & 6.71 & 5.39 & 0 & 7.62 \\ 2.24 & 4.47 & 6.4 & 1.41 & 5 & 5.39 & 7.62 & 0 \end{pmatrix}.$$

First Step of the clustering $\{X_1\}, \{X_2\}, \{X_3, X_5\}, \{X_4\}, \{X_6\}, \{X_7\}, \{X_8\}$
Second Step of the clustering $\{X_1\}, \{X_2\}, \{X_3, X_5\}, \{X_6\}, \{X_7\}, \{X_4, X_8\}$
Third Step of the clustering $\{X_1\}, \{X_2\}, \{X_3, X_5, X_6\}, \{X_7\}, \{X_4, X_8\}$
Fourth Step of clustering $\{X_1, X_4, X_8\}, \{X_2\}, \{X_3, X_5, X_6\}, \{X_7\}$
Fifth Step of clustering $\{X_1, X_4, X_8\}, \{X_2, X_7\}, \{X_3, X_5, X_6\}$
Sixth Step of clustering $\{X_1, X_4, X_8, X_3, X_5, X_6\}, \{X_2, X_7\}$
Seventh Step of clustering $\{X_1, X_4, X_8, X_3, X_5, X_6, X_2, X_7\}$

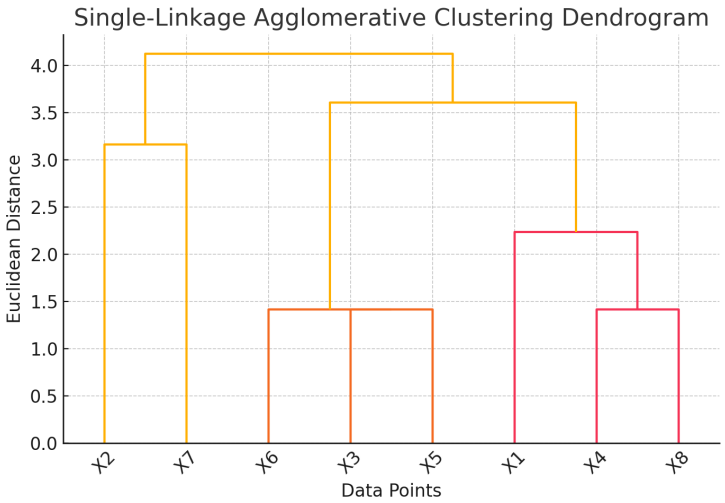

Figure 1:

Excercise 2

The dataset consists of four objects with pairwise distances given by the following distance matrix

$$\begin{pmatrix} & X_1 & X_2 & X_3 & X_4 \\ X1 & 0 & 2 & 6 & 2.5 \\ X_2 & 2 & 0 & 3 & 8 \\ X_3 & 6 & 3 & 0 & 4 \\ X_4 & 2.5 & 8 & 4 & 0 \end{pmatrix}.$$

1. Cluster dataset using single-linkage agglomerative clustering. Draw the corresponding dendrogram.

2. Cluster dataset using complete-linkage agglomerative clustering. Draw the corresponding dendrogram.

Solution First Step of Single Linkage Clustering $\{X_1, X_2\}, \{X_3\}, \{X_4\}$
Second Step of Single Linkage Clustering $\{X_1, X_2, X_4\}, \{X_3\}$
Third Step of Single Linkage Clustering $\{X_1, X_2, X_4, X_3\}$

First Step of Complete Linkage Clustering $\{X_1, X_2\}, \{X_3\}, \{X_4\}$
Second Step of Complete Linkage Clustering $\{X_1, X_2\}, \{X_3, X_4\}$
Third Step of Complete Linkage Clustering $\{X_1, X_2, X_4, X_3\}$

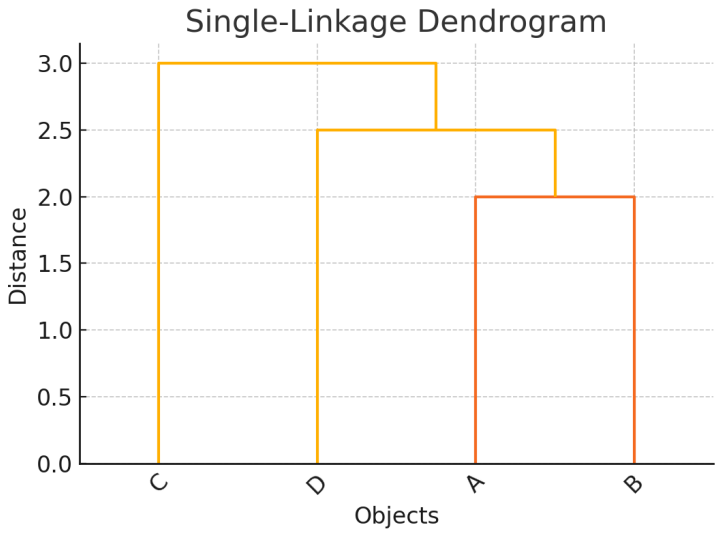

Figure 2:

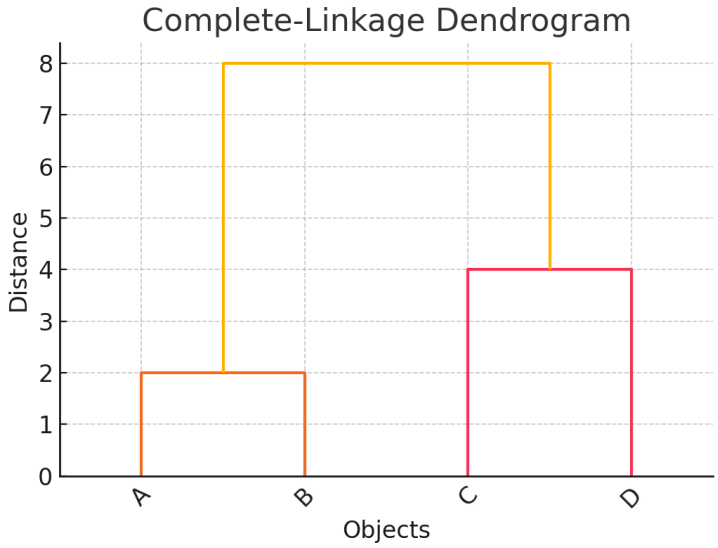

Figure 3: