Part 2 :

B: Hierarchical Clustering

Hierarchical Agglomerative Clustering Initializes with number of clusters equals to number of data points such that each data point represents cluster.

Clusters : {10}, {20}, {40}, {80}, {85}, {121}, {160}, {168}, {195}

*Then Starts Merging **Most Similar Pair of Clusters** one by one until reach remains only single cluster. Clusters similarity is defined by **Similarity/Distance Measure method** (Euclidean Distance in this case) & **Linkage Measures** (Single Linkage & Complete Linkage in this case).*

*Initialize **Distances Matrix** that contains **Euclidian Distance Measure** that represents distance between each pair of clusters.*

Cluster Set	{10}	{20}	{40}	{80}	{85}	{121}	{160}	{168}	{195}
{10}	0	-------	-------	-------	-------	-------	-------	-------	-------
{20}	10	0	-------	-------	-------	-------	-------	-------	-------
{40}	30	20	0	-------	-------	-------	-------	-------	-------
{80}	70	60	40	0	-------	-------	-------	-------	-------
{85}	75	65	45	5	0	-------	-------	-------	-------
{121}	111	101	81	41	36	0	-------	-------	-------
{160}	150	140	120	80	75	39	0	-------	-------
{168}	158	148	128	88	83	47	8	0	-------
{195}	185	175	155	115	110	74	35	27	0

1- Single Linkage :

Single Linkage Evaluate Clusters Similarity Based on Smallest distance between 2 nearest observations, one from each cluster.

Clustering Steps

Distances Matrix 1 :

Cluster Set	{10}	{20}	{40}	{80}	{85}	{121}	{160}	{168}	{195}
{10}	0	-------	-------	-------	-------	-------	-------	-------	-------
{20}	10	0	-------	-------	-------	-------	-------	-------	-------
{40}	30	20	0	-------	-------	-------	-------	-------	-------
{80}	70	60	40	0	-------	-------	-------	-------	-------
{85}	75	65	45	5	0	-------	-------	-------	-------
{121}	111	101	81	41	36	0	-------	-------	-------
{160}	150	140	120	80	75	39	0	-------	-------
{168}	158	148	128	88	83	47	8	0	-------
{195}	185	175	155	115	110	74	35	27	0

*Minimum Distance Among Clusters Pairwise Distance is **5** which is between **{80}** & **{85}** Clusters So, those two clusters will be **Merged** together Into single cluster Resulting new **Distance Matrix**.*

Distances Matrix 2 :

Cluster Set	{10}	{20}	{40}	{80, 85}	{121}	{160}	{168}	{195}
{10}	0	---------	---------	---------	---------	---------	---------	---------
{20}	10	0	---------	---------	---------	---------	---------	---------
{40}	30	20	0	---------	---------	---------	---------	---------
{80, 85}	70	60	40	0	---------	---------	---------	---------
{121}	111	101	81	36	0	---------	---------	---------
{160}	150	140	120	75	39	0	---------	---------
{168}	158	148	128	83	47	8	0	---------
{195}	185	175	155	110	74	35	27	0

Repeat Same Steps Until all points are in the single cluster

Distances Matrix 3 :

Cluster Set	{10}	{20}	{40}	{80, 85}	{121}	{160, 168}	{195}
{10}	0	---------	---------	---------	---------	---------	---------
{20}	10	0	---------	---------	---------	---------	---------
{40}	30	20	0	---------	---------	---------	---------
{80, 85}	70	60	40	0	---------	---------	---------
{121}	111	101	81	36	0	---------	---------
{160, 168}	150	140	120	75	39	0	---------
{195}	185	175	155	110	74	27	0

Distances Matrix 4 :

Cluster Set	{10, 20}	{40}	{80, 85}	{121}	{160, 168}	{195}
{10, 20}	0	---------	---------	---------	---------	---------
{40}	20	0	---------	---------	---------	---------
{80, 85}	60	40	0	---------	---------	---------
{121}	101	81	36	0	---------	---------
{160, 168}	140	120	75	39	0	---------
{195}	175	155	110	74	27	0

Distances Matrix 5 :

Cluster Set	{10, 20, 40}	{80, 85}	{121}	{160, 168}	{195}
{10, 20, 40}	0	---------	---------	---------	---------
{80, 85}	40	0	---------	---------	---------
{121}	81	36	0	---------	---------
{160, 168}	120	75	39	0	---------
{195}	155	110	74	27	0

Distances Matrix 6 :

Cluster Set	{10, 20, 40}	{80, 85}	{121}	{160, 168, 195}
{10, 20, 40}	0	---------	---------	---------
{80, 85}	40	0	---------	---------
{121}	81	36	0	---------
{160, 168, 195}	120	75	39	0

Distances Matrix 7 :

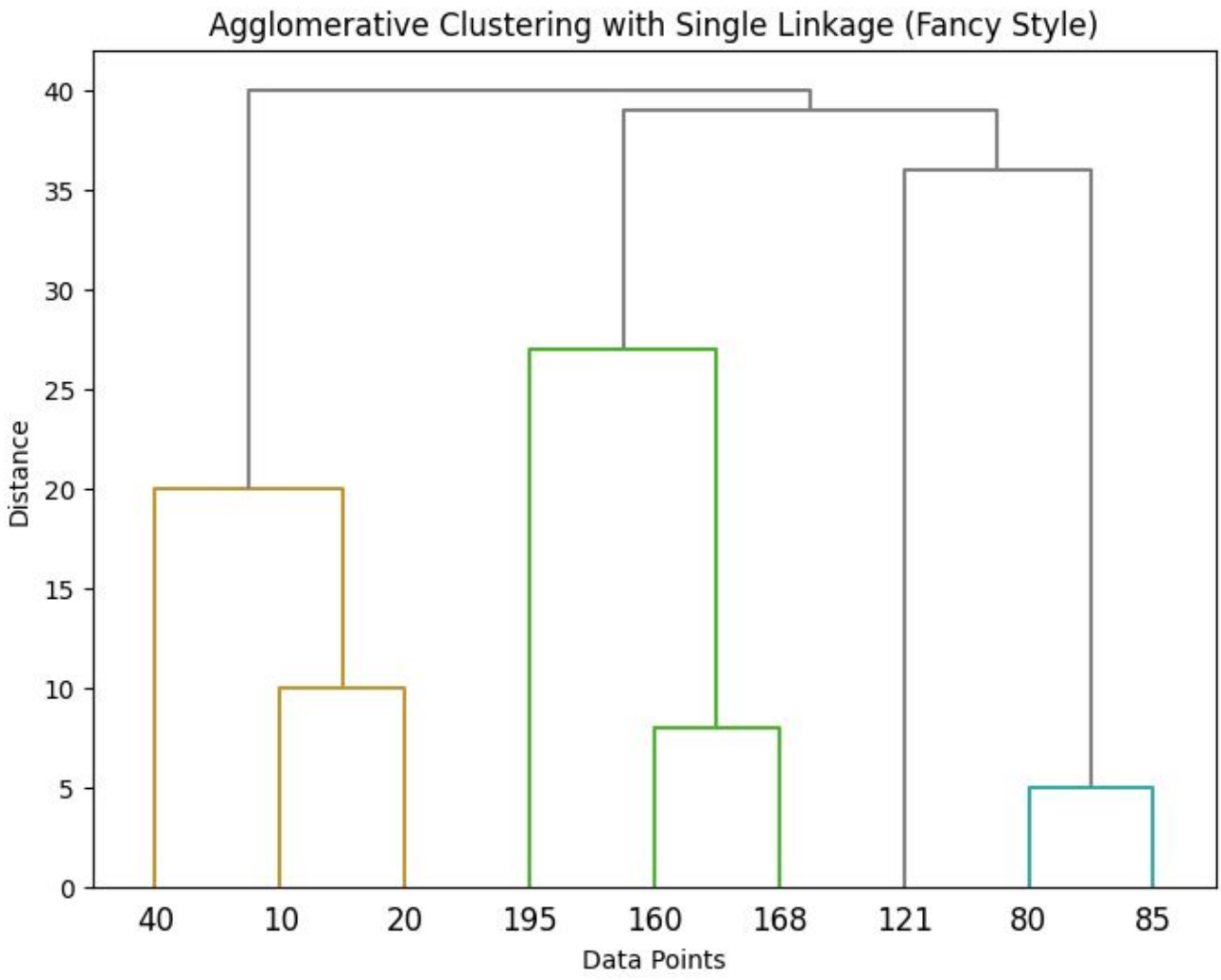
Cluster Set	{10, 20, 40}	{80, 85, 121}	{160, 168, 195}
{10, 20, 40}	0	---------	---------
{80, 85, 121}	40	0	---------
{160, 168, 195}	120	39	0

Distances Matrix 8 :

Cluster Set	{10, 20, 40}	{80, 85, 121, 160, 168, 195}
{10, 20, 40}	0	----------
{80, 85, 121, 160, 168, 195}	40	0

*The final clusters are **{80, 85, 121, 160, 168, 195}** & **{10, 20, 40}** at Distance **40**.*

Dendrogram Steps



2- Complete Linkage :

Complete Linkage Evaluate Clusters Similarity Based on Largest distance between an element in one cluster and an element in the other.

Clustering Steps

Distances Matrix 1 :

Cluster Set	{10}	{20}	{40}	{80}	{85}	{121}	{160}	{168}	{195}
{10}	0	-------	-------	-------	-------	-------	-------	-------	-------
{20}	10	0	-------	-------	-------	-------	-------	-------	-------
{40}	30	20	0	-------	-------	-------	-------	-------	-------
{80}	70	60	40	0	-------	-------	-------	-------	-------
{85}	75	65	45	5	0	-------	-------	-------	-------
{121}	111	101	81	41	36	0	-------	-------	-------
{160}	150	140	120	80	75	39	0	-------	-------
{168}	158	148	128	88	83	47	8	0	-------
{195}	185	175	155	115	110	74	35	27	0

*Minimum Distance Among Clusters Pairwise Distance is **5** which is between **{80}** & **{85}** Clusters So, those two clusters will be **Merged** together Into single cluster Resulting new **Distance Matrix**.*

Distances Matrix 2 :

Cluster Set	{10}	{20}	{40}	{80, 85}	{121}	{160}	{168}	{195}
{10}	0	---------	---------	---------	---------	---------	---------	---------
{20}	10	0	---------	---------	---------	---------	---------	---------
{40}	30	20	0	---------	---------	---------	---------	---------
{80, 85}	75	65	45	0	---------	---------	---------	---------
{121}	111	101	81	41	0	---------	---------	---------
{160}	150	140	120	80	39	0	---------	---------
{168}	158	148	128	88	47	8	0	---------
{195}	185	175	155	115	74	35	27	0

*The only difference here that we updated distance matrix values of the new cluster with **Maximum** distance of the two merged clusters instead of the **Minimum** to consider Largest distance between an element in one cluster & an element in the other.*

Distances Matrix 3 :

Cluster Set	{10}	{20}	{40}	{80, 85}	{121}	{160, 168}	{195}
{10}	0	---	---	---	---	---	---
{20}	10	0	---	---	---	---	---
{40}	30	20	0	---	---	---	---
{80, 85}	75	65	45	0	---	---	---
{121}	111	101	81	41	0	---	---
{160, 168}	158	148	128	88	47	0	---
{195}	185	175	155	115	74	35	0

Repeat Same Steps Until all points are in the single cluster

Distances Matrix 4 :

Cluster Set	{10, 20}	{40}	{80, 85}	{121}	{160, 168}	{195}
{10, 20}	0	---	---	---	---	---
{40}	30	0	---	---	---	---
{80, 85}	75	45	0	---	---	---
{121}	111	81	41	0	---	---
{160, 168}	158	128	88	47	0	---
{195}	185	155	115	74	35	0

Distances Matrix 5 :

Cluster Set	{10, 20, 40}	{80, 85}	{121}	{160, 168}	{195}
{10, 20, 40}	0	---	---	---	---
{80, 85}	75	0	---	---	---
{121}	111	41	0	---	---
{160, 168}	158	88	47	0	---
{195}	185	115	74	35	0

Distances Matrix 6 :

Cluster Set	{10, 20, 40}	{80, 85}	{121}	{160, 168, 195}
{10, 20, 40}	0	---	---	---
{80, 85}	75	0	---	---
{121}	111	41	0	---

Cluster Set	{10, 20, 40}	{80, 85}	{121}	{160, 168, 195}
{160, 168, 195}	185	115	74	0

Distances Matrix 7 :

Cluster Set	{10, 20, 40}	{80, 85, 121}	{160, 168, 195}
{10, 20, 40}	0	---------	---------
{80, 85, 121}	111	0	---------
{160, 168, 195}	185	115	0

Distances Matrix 8 :

Cluster Set	{10, 20, 40, 80, 85, 121}	{160, 168, 195}
{10, 20, 40, 80, 85, 121}	0	---------
{160, 168, 195}	185	0

Finally, we merge the remaining clusters at distance of 185 & Since we have only one cluster left, the clustering process is complete.

Dendrogram Steps