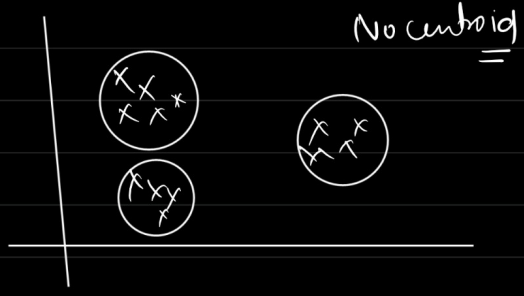
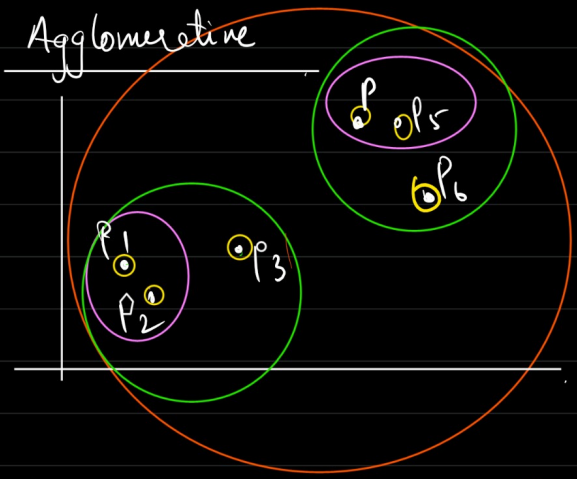
* Hierarchical Clustering



* Hierarchical Clustering.

① Agglomerative Clustering (combining)
② Divisive Clustering (dividing)

* Agglomerative



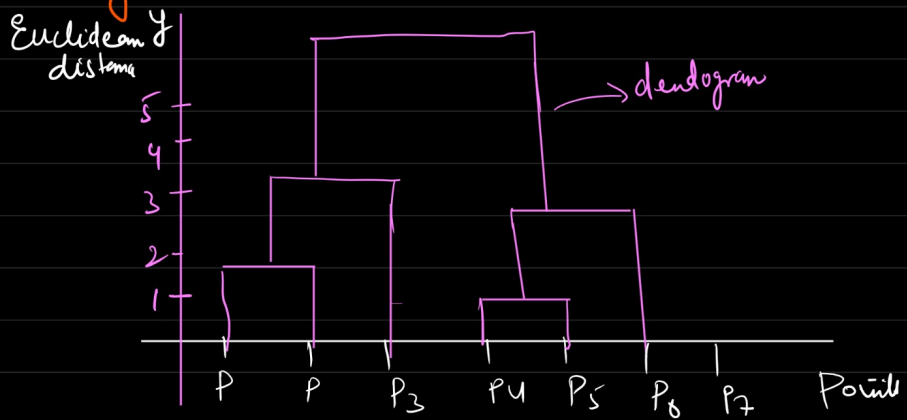
Steps (Agglomerative)

① Each point is a cluster in its own
② Find the nearest point and create a new cluster.
③ keep on doing step 2 untill we get a single cluster.

→ Euclidean dist
→ Manhattan distance
→ Cosine similarity → To calculate distance between two vector (Categorical data | strip variable)

Still how to select K ?

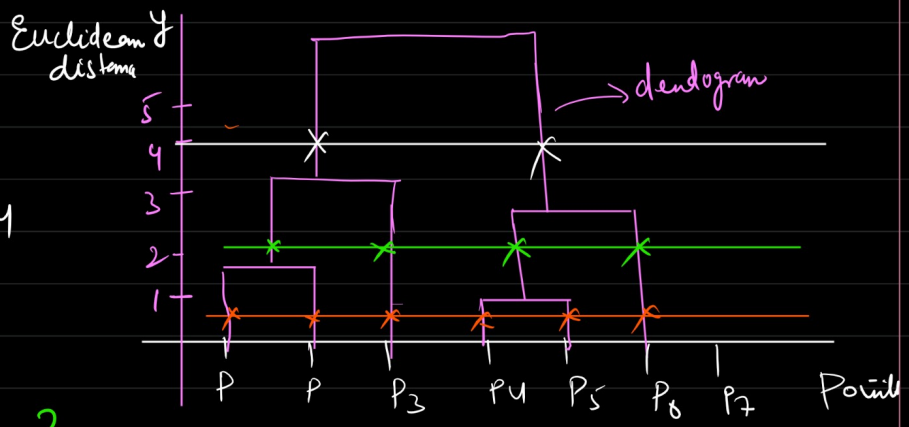
Using dendogram



To determine K

→ threshold on Euclidean distance. (Can be tricky job ⇒ business team)
→ trick → we can take longest vertical line of dendogram where none of the horizontal line of dendogram is passing.

① Using threshold on Euclidean distance



→ Say threshold = 4

⇓

K = 2

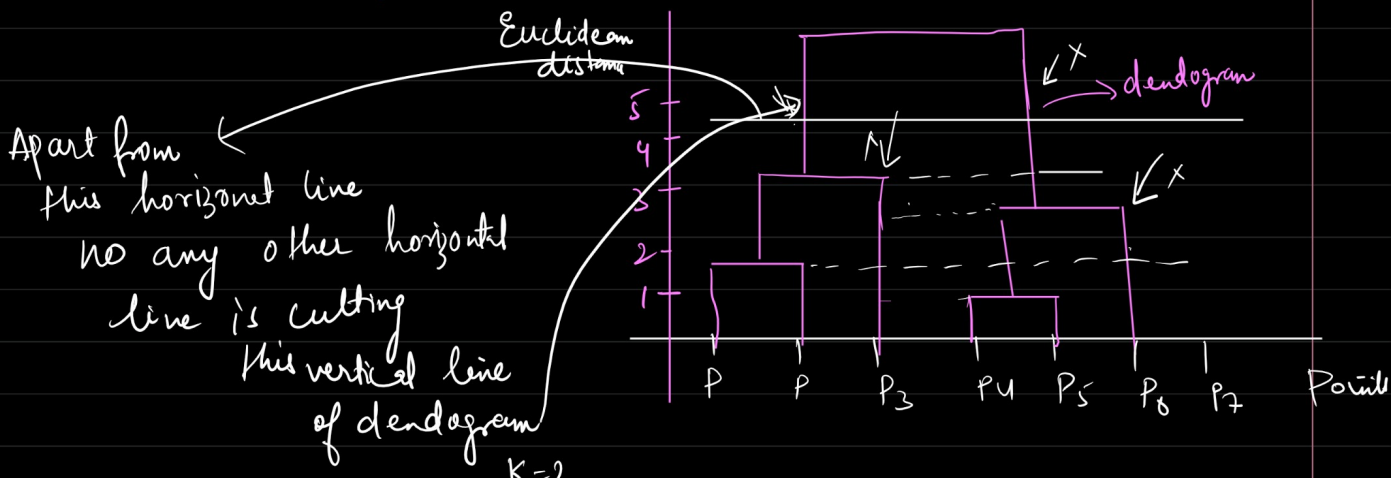
→ Say threshold is 2

K = 2

→ say threshold is 1

K = 6 (all the dP's is cluster in its own)

②

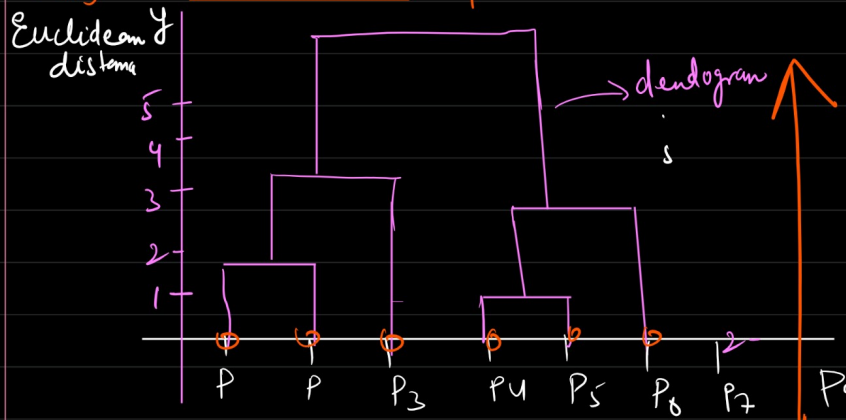


Apart from this horizontal line no any other horizontal line is cutting this vertical line of dendogram

K = 2

Conclusion → Select the longest vertical line such that no horizontal line passed through it.

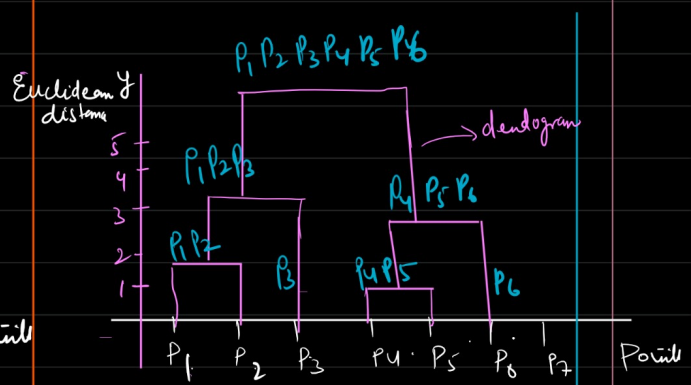
Agglomerative Clustering



→ Here every dP's are getting combined

Combining

Divisive clustering



Keep diving until every individual dp becomes Cluster in its own.

dividing

* K Means Vs Hierarchical (scalability & flexibility)

① Size of data → Huge → K Means
→ Small → Hierarchical

② K means — Numerical (only) data (Euclidean | Manhattan distance)

Hierarchical Clustering — Variety of data (Euclidean, Manhattan, cosine Similarity)

③ K means → Centroid (k) → Elbow method

Hierarchical — No centroid.