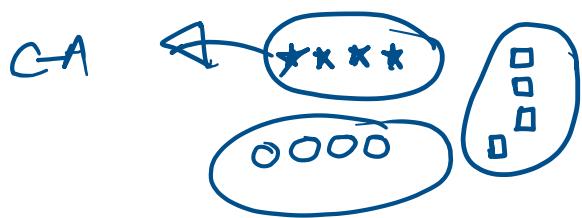
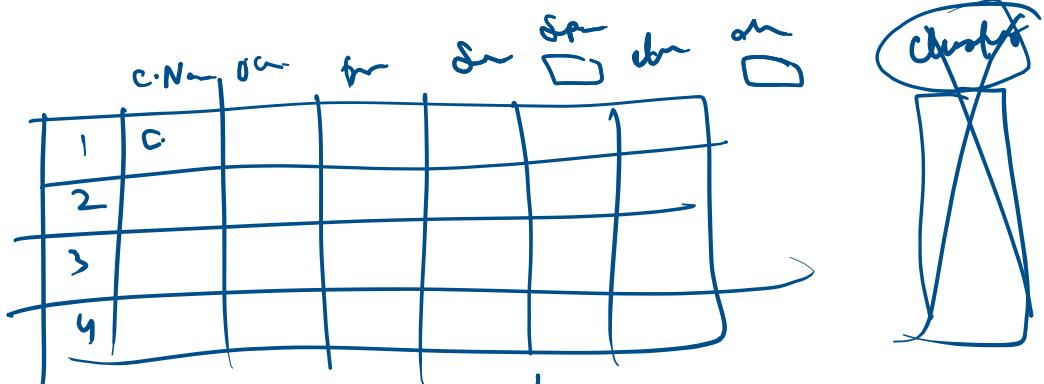
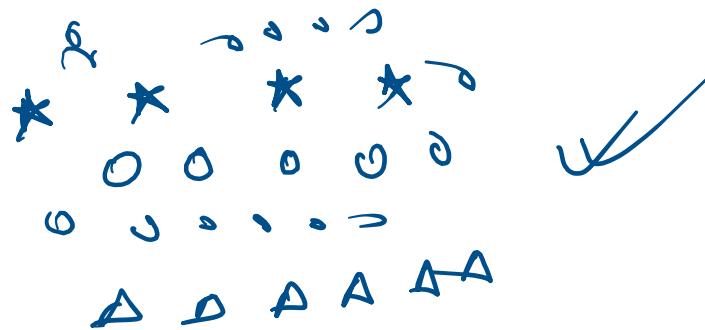


clustering :-

- Grouping
- unsupervised
- Find similar groups.



→ collection of object which are "Perceived" to be similar b/w them & dissimilar to the object belonging to diff. clusters.



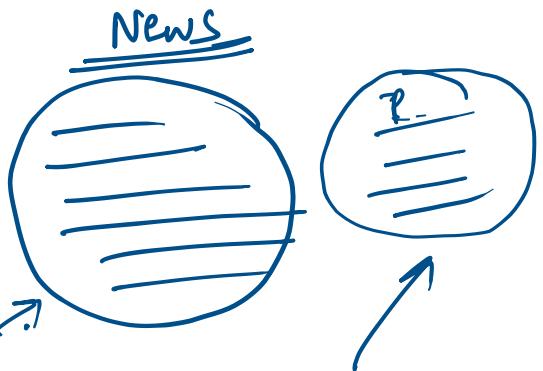
clustering

Hierarchical
(Dendrogram
based Cluster)

Non-Hierarchical
→ K-Means

Use Cases :

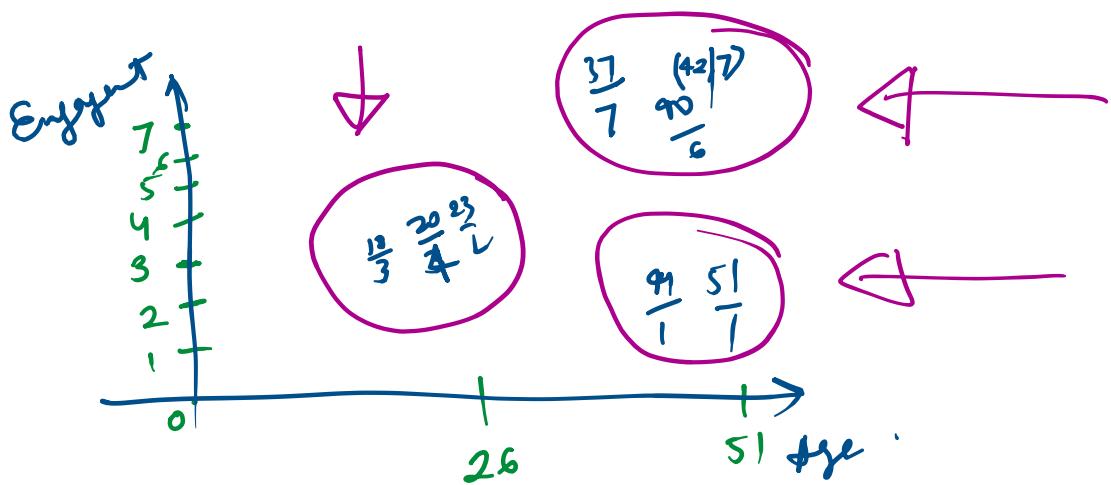
- ① Google News.
- ② Movie Recommendation
- ③ Launch a Product.



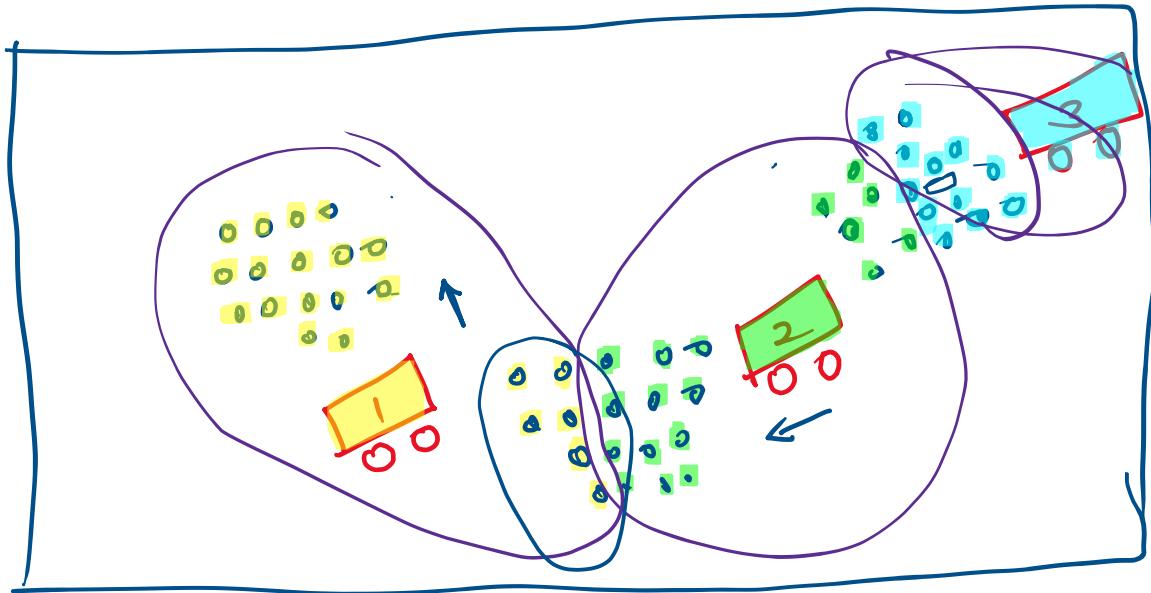
3 - Marketing

date → Age, Engaged in days/weeks.

| Age 92 | Age 18 | 23 | 49 | 37 | 51 | 90/20 |
|--------|--------|----|----|----|----|-------|
| cy 7 | 3 | 2 | 1 | 7 | 1 | 6 + |



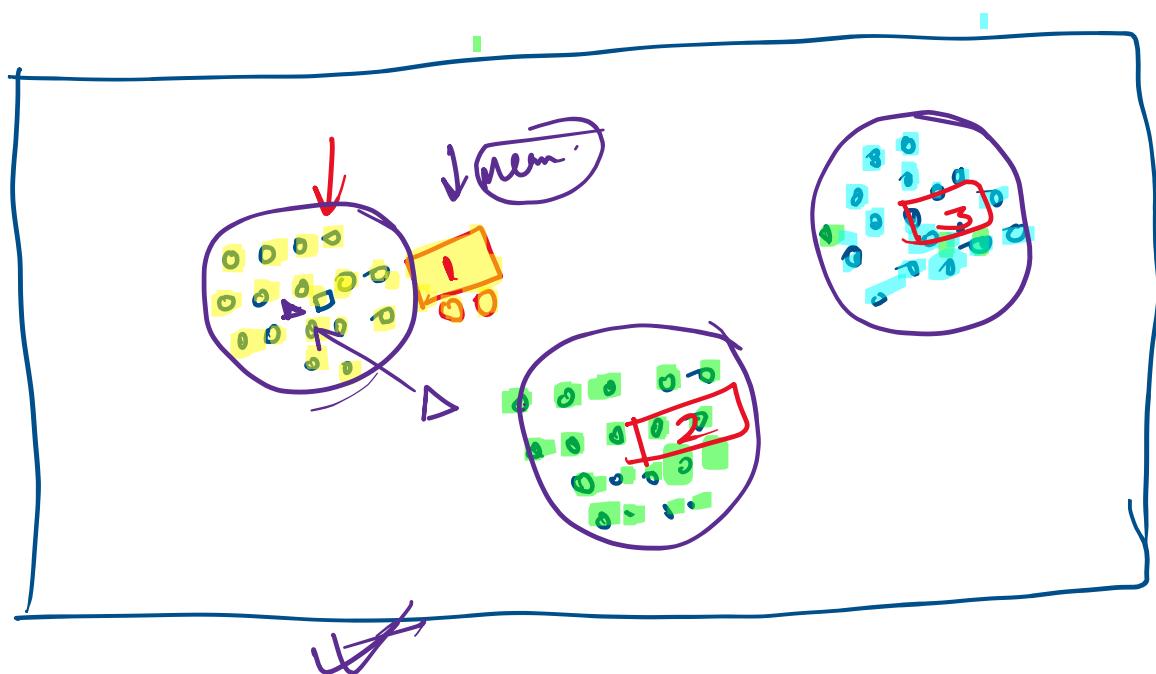
How does the clustering happen?



Step 1 : - locate 3 Random Food truck around the city.

Step 2 : there are some logic which when applied will make the decision.

- (a) everyone is going to their nearest F.T.
- (b) The best location to serve is always centre point of location, hence our truck will move to the centre point.
- (c) Repeat Step (a) & (b) - - - -

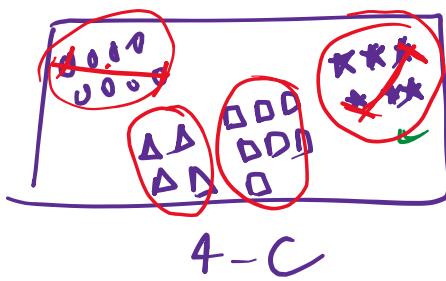
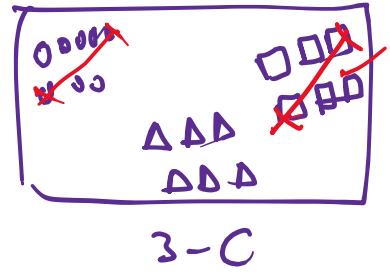
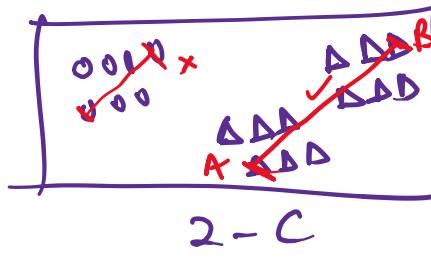
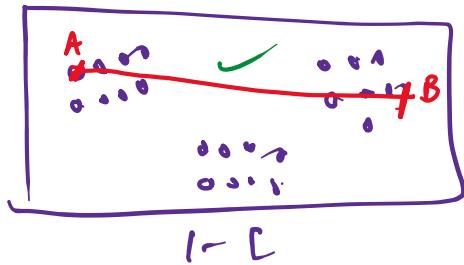


Trucks = Mean

I took the decision \Rightarrow No. of Trucks.

Q. How do we know how many clusters?

- A: Using elbow method (WSS).

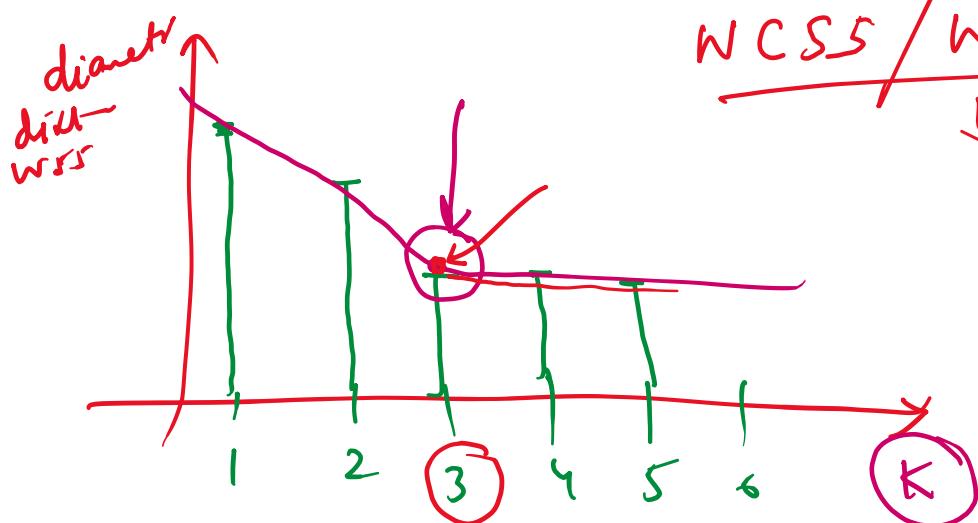


--

1 D →

1 D →

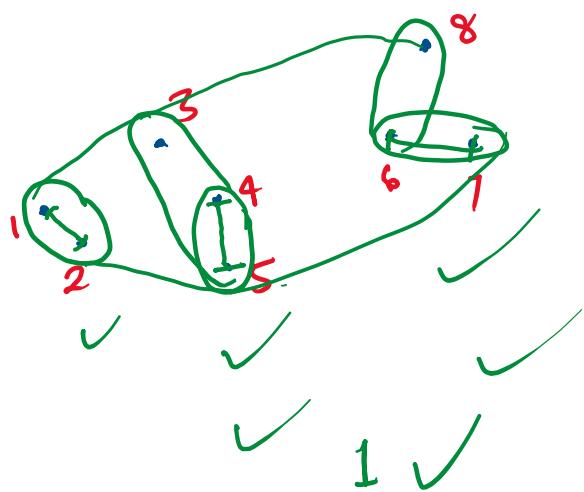
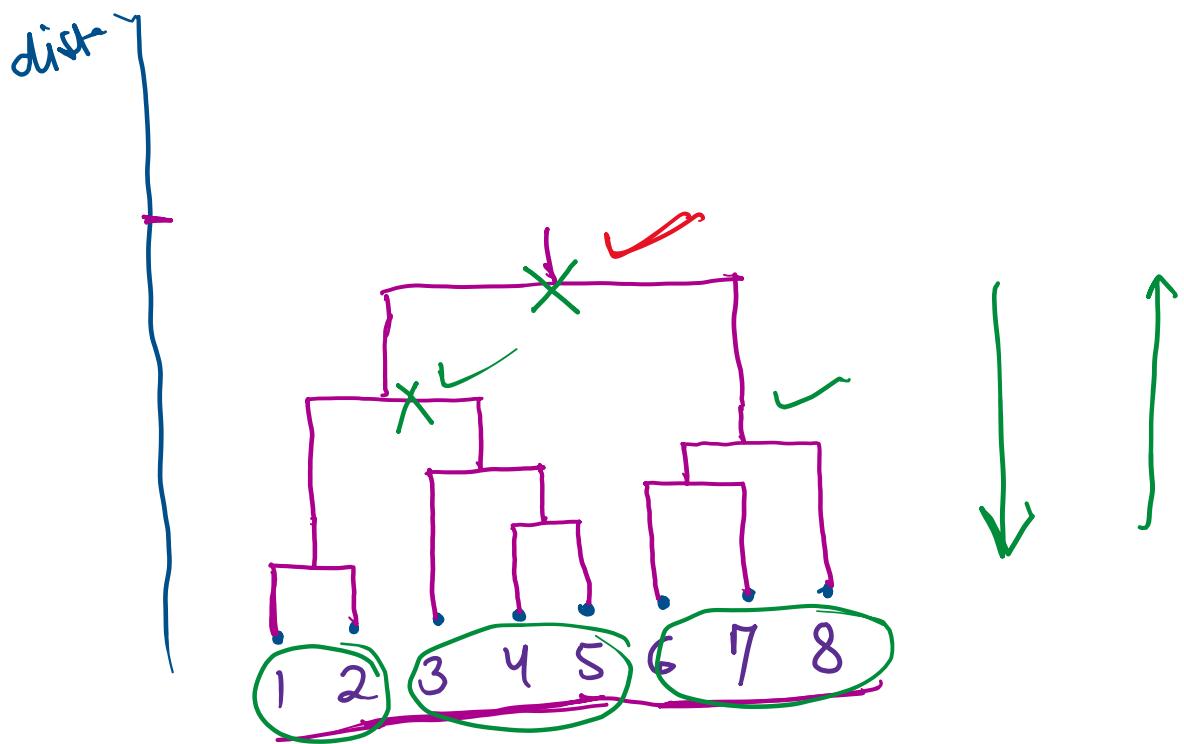
WCSS / WSS / elbow



$K = 3$

optimum.

Hierarchical clustering



- ① Bottom up — Agglomerative
- ② Top - down — divisive

Agglo

- * Bottom up approach
- * each observation starts its own cluster.
- pairs of clusters are merged when we move up in the hierarchy.
- subcomponent is visited first & then move to parent.

divisive

- * Top-down approach.
- * all observation starts in one cluster.
- splits are performed recursively as we move down the hierarchy.
- parent is visited first.

diff.

- ① Hierarchical clustering cannot handle large data well but K-Means can.

→ Time complexity of K-Means $O(n)$
 → " " " " Hier. $O(n^2)$

- ② Results are not always reproducible in K-Means (since we start randomly) whereas in Hierarchical it can be.

Linkage Method

- ① Single linkage
 - ② Complete linkage
 - ③ Ward's "
 - ④ Average "
 - ⑤ Centroid "

$$z = \sqrt{x^2 + y^2}$$