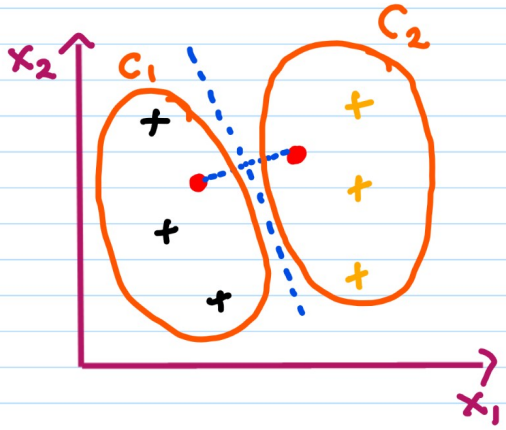
Machine Learning

- Supervised Machine Learning (Labels)
  - Regression
    - Linear Regression
  - Classification
    - Logistic Regression
- Unsupervised Machine Learning (No labels)
  - K-Means
    - e.g. Clustering.

K-Means

→ It is UnSupervised Machine Learning Algorithm, which basically clusters or groups the data together.

How K-Means Clustering Work?

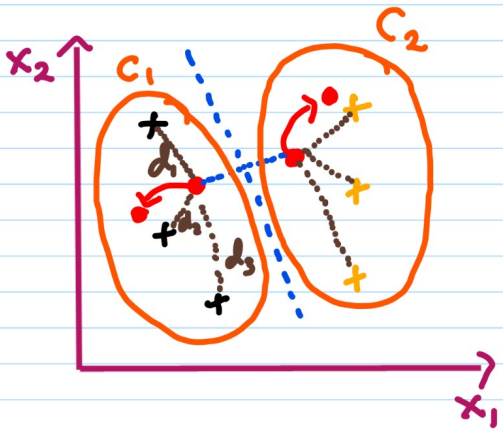


● → Centroid

K-Means Clustering
↳ Number of desired Clusters

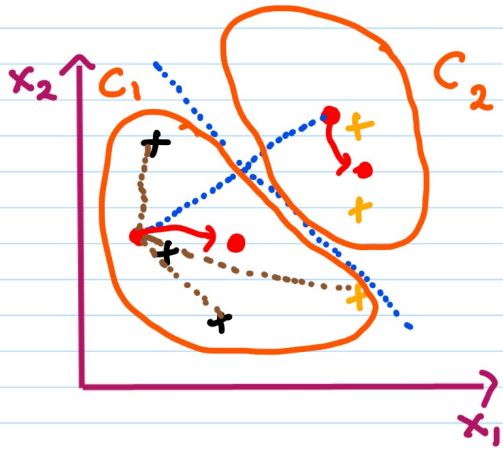
$K = 2$

1) Randomly choosing the centroid and assigning each datapoint to it's nearest centroid to form a Cluster.



$$\text{Avg } C_1 = \frac{d_1 + d_2 + d_3}{3}$$

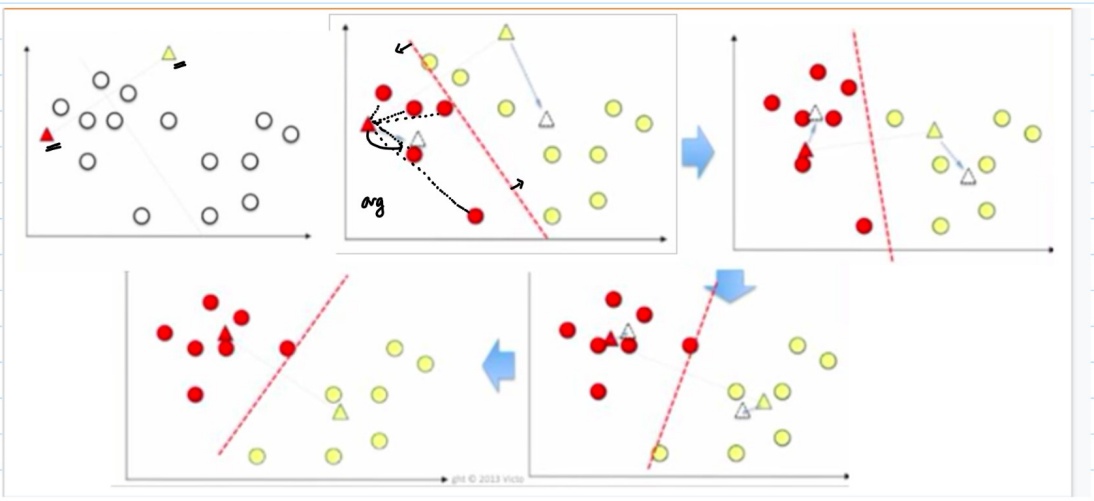
$= (\bullet)$



$$\frac{d_1 + d_2 + d_3 + d_4}{4}$$

Avg $= (\ )$

This entire process will continue, until and unless the movement of the centroid stops.



Data Scientist

↓

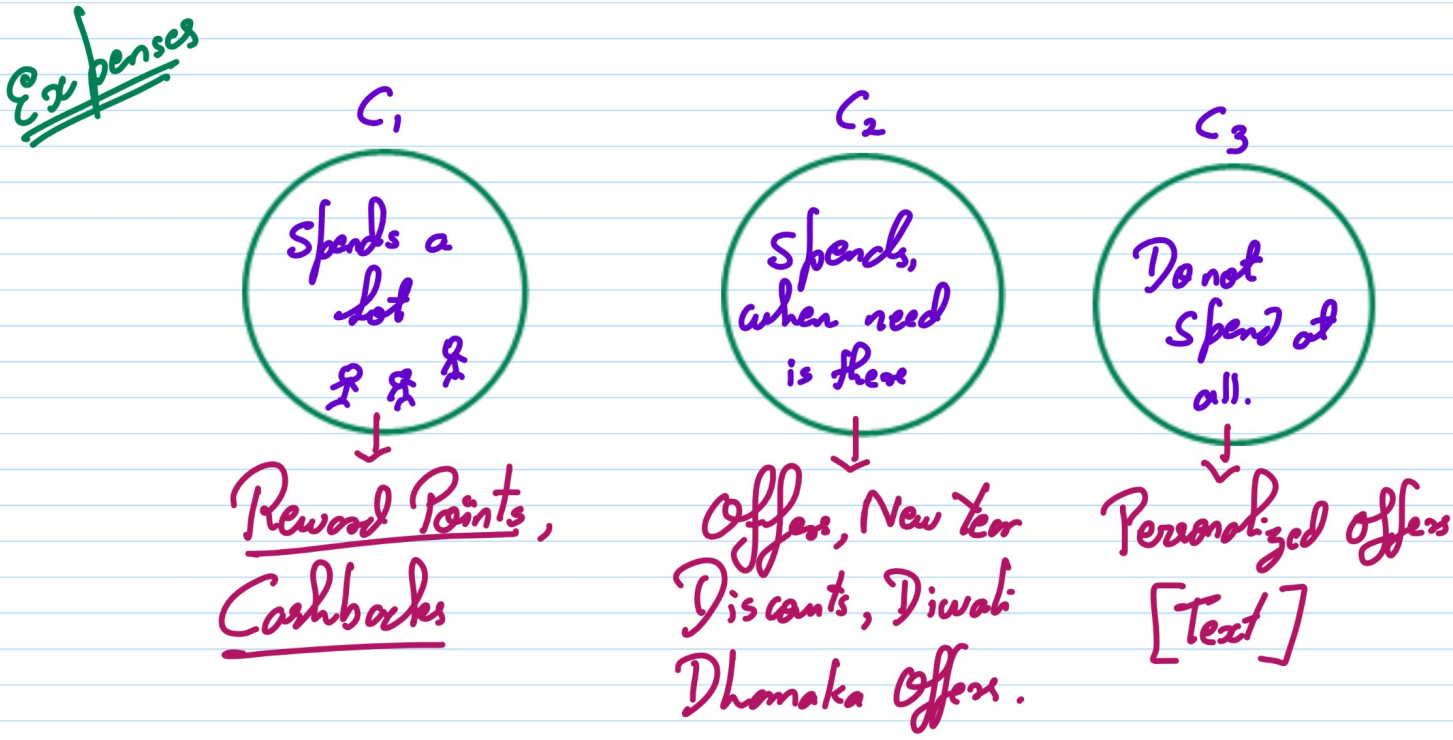
Amazon → [Increase the Sales].

↓

10 lakh

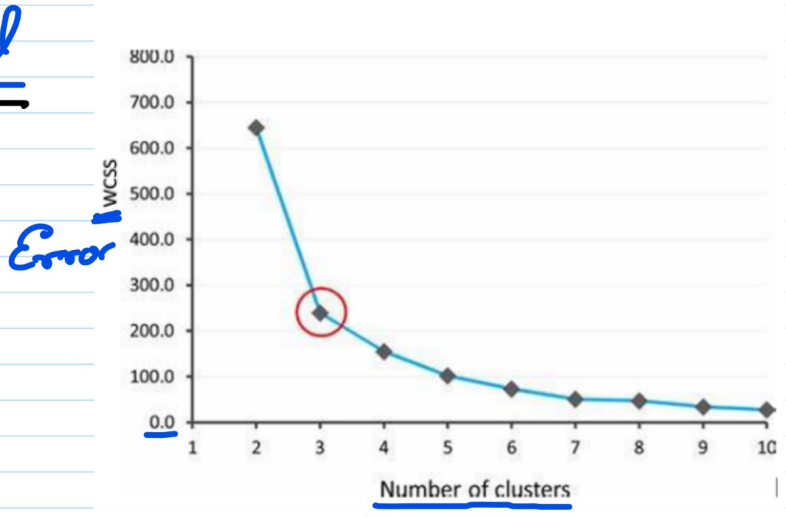
↳ Purchase History, What?, etc.....

Expenses



Q What is the Optimal number of Cluster?

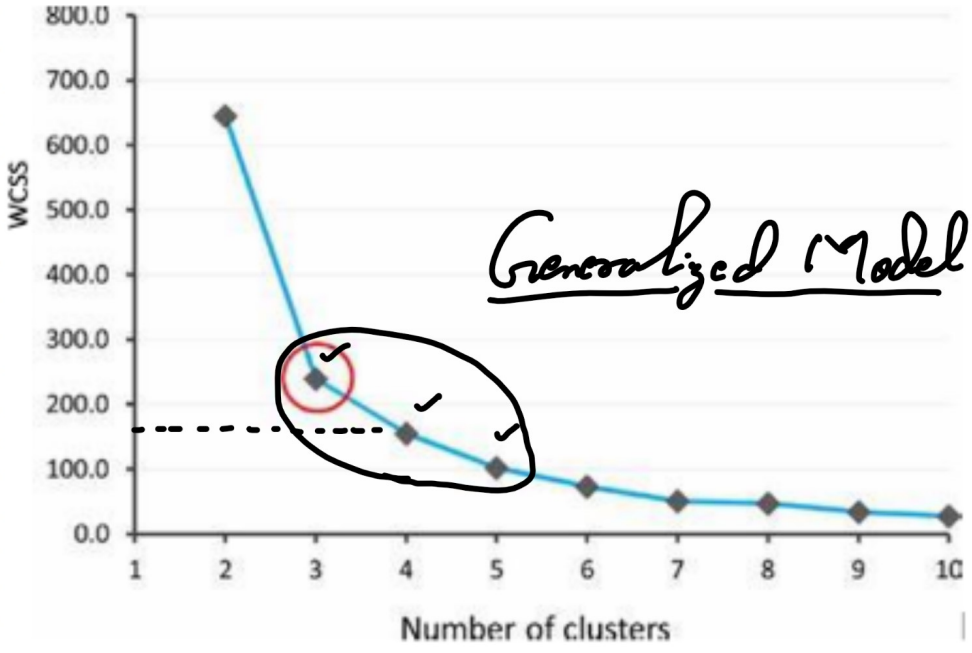
Elbow Method



WCSS / Inertia / Error.

↳ Within Cluster Sum of Squares

K → Clusters



Error

Generalized Model