Problem Statement

To Segregate Customers Based On The Data In Form Of Optimal Clusters

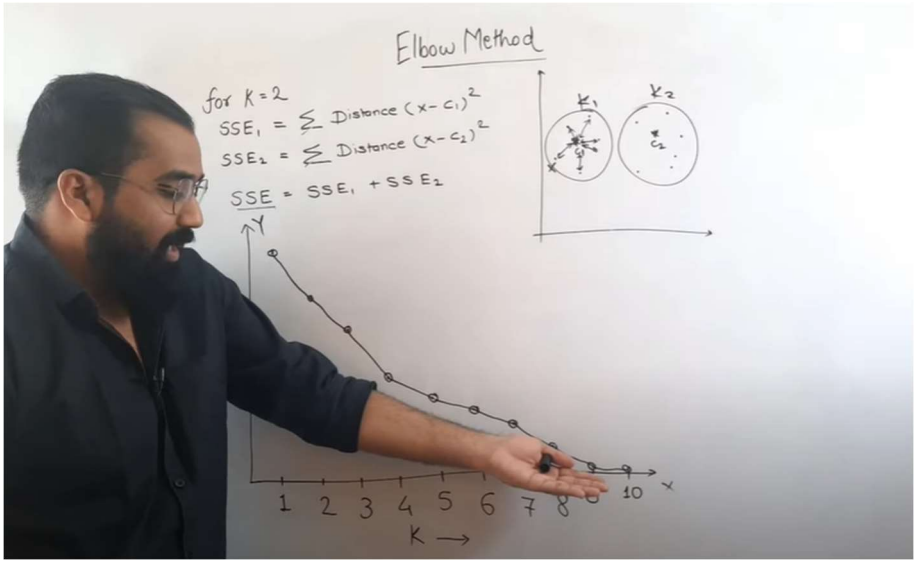
Steps:

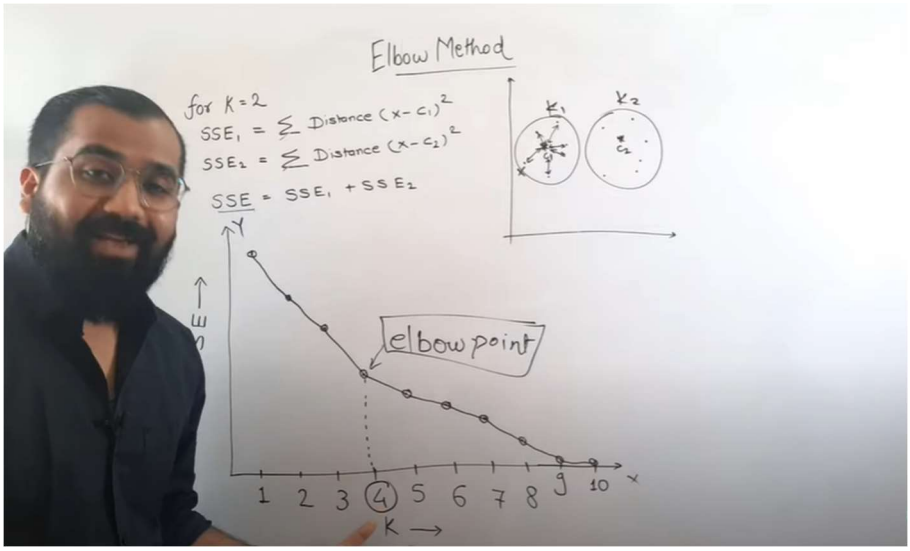
- Understanding The Data

    - Importing Useful Libraries Such As Pandas(For Data Manipulation). ,Numpy(Numeric Calculation , For Complex Mathametical Operations) ,Seaborn (Visualization) ,Matplotlib (Visualization),Etc.
    - Head() To Print First Rows.
    - Tail() To Print Bottom Rows.
    - Shape To Print Size Of Table (Eg (200,5)).
    - Info() To Check Null Value Count And Column Data Type.
    - Preparing Data For Clustering.
    - Iloc[] Method To Extract The Column Based On Index Number And Saving Them In New Data Frame “X”.

- Performed Elbow Method To Find Optimal No. Of Clusters

    - Using K-Means To Iterate From 2 To 10 Clusters And Plotting A Elbow Plot.
    - Deciding Optimal No Of Clusters To Be Used.
    - Here K Defines Pre Label Clusters
    - Importing Clusters,Kmeans From Sklearn
    - Wcss Stands For Within Cluster Sum Of Square. It Calculate Distance Between Each Point And Centroid In The Cluster.
    - Kmean++ Used For Smarter Initialization Of Cluster





- Training A Model Using Unsupervised Learning Algorithm (K-Means)
  - Initializing K-Means Model With Selected Optimal No. Of Clusters.
  - Plot Of Clusters And Gaining Intuitions Regarding Our Customers
  - Legend() Adding In Plot

- Plotting The Cluster.

  - ✓ Customer 1== Having Medium Salary And Spending Is
  - ✓ Customer 2== Having Less Salary And Spending Is High
  - ✓ Customer 3== Having Less Salary And Spending Is High
  - ✓ Customer 4== Having High Salary And Spending Is Low
  - ✓ Customer 5== Having High Salary And Spending Is Low

Reference Link:

Https://Www.Youtube.Com/Watch?V=7j3mz9lmjqg