K-Means Clustering and Elbow Method (Simple Explanation)

1. What is K-Means?

K-Means is a method used to divide data into groups (called clusters). Each group has points that are close to each other. The center of each group is called a **centroid**.

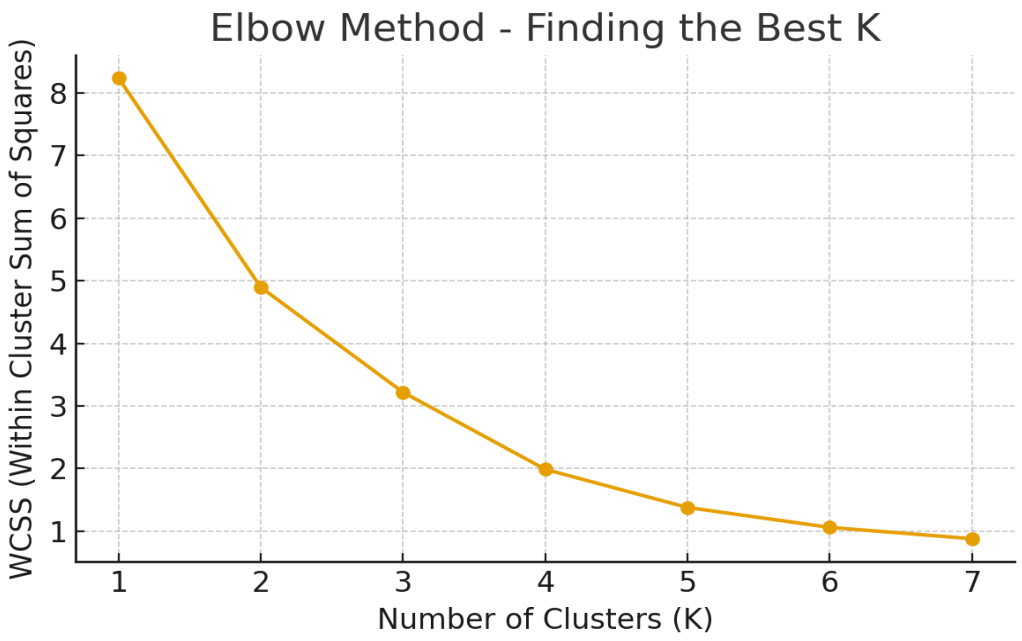
2. Steps of K-Means

1. Choose how many groups (K) you want.
2. Pick K random points as starting centers (centroids).
3. Assign each data point to the nearest center.
4. Move each center to the average of its group's points.
5. Repeat steps 3 and 4 until centers stop moving.

3. How to Find the Best K (Elbow Method)

The Elbow Method helps you find how many clusters (K) to choose.

Steps:
1. Run K-Means for K = 1, 2, 3, … up to 8 or 10.
2. For each K, calculate **WCSS** (Within Cluster Sum of Squares) — it tells how close points are to their centers.
3. Draw a graph of K (x-axis) vs WCSS (y-axis).
4. Look for the "elbow" — the point where the line starts to bend slowly. That is the best K value.



In the picture, the bend (elbow) shows the best number of clusters. Before the bend, WCSS drops fast; after it, the improvement is small. The elbow point is the right choice for K.

4. Summary

- K-Means divides data into K groups.
- Each group has a center called a centroid.
- The Elbow Method helps find the best K.
- The “elbow” point gives a balance between too few and too many clusters.

This makes K-Means simple, fast, and very useful for grouping data in real life.