

1. **Compare and contrast K-means clustering and hierarchical clustering.**
 - **K-means Clustering:**
 - It's a partitioning method that divides a dataset into k distinct, non-overlapping subsets (or clusters).
 - Requires the number of clusters to be specified beforehand.
 - It's generally faster and suitable for large datasets.
 - It may converge to a local optimum depending on the initialization of clusters.
 - **Hierarchical Clustering:**
 - It creates a tree of clusters. The hierarchy starts with each point as an individual cluster and merges them step by step.
 - Doesn't require the number of clusters to be pre-specified.
 - Generally slower and may not be suitable for large datasets.
 - Provides more detailed information, allowing one to choose the number of clusters by cutting the dendrogram at a level.
2. **Briefly explain the steps of the K-means clustering algorithm.**
 - Initialization: Start by selecting k initial centroids, where k is the number of clusters you want.
 - Assignment: Assign each data point to the nearest centroid, creating k clusters.
 - Recomputation: Recalculate the centroid of each cluster based on the members of that cluster.
 - Iteration: Repeat the assignment and recomputation steps until the centroids no longer change significantly, or a set number of iterations is reached.
3. **How is the value of 'k' chosen in K-means clustering?**
 - **Statistical Aspect:** Methods like the Elbow Method or Silhouette Analysis are used. The Elbow Method involves plotting the number of clusters against within-cluster sum of squares, looking for a "bend" or "elbow" which suggests the optimal number of clusters.
 - **Business Aspect:** Sometimes the choice of 'k' might be driven by business needs or domain-specific knowledge, where the organization has a predefined number of clusters in mind.
4. **Explain the necessity of scaling/standardisation before performing clustering.**
 - Clustering algorithms, like K-means, use distance measures to determine the similarity between data points. Features with higher scales (larger numeric ranges) can unduly influence the cluster assignment. By scaling or standardizing, every feature contributes equally to the distance computation, ensuring a balanced clustering.

5. **Explain the different linkages used in hierarchical clustering.**

- **Single Linkage:** The distance between two clusters is defined as the shortest distance between points in the two clusters.
- **Complete Linkage:** The distance between two clusters is defined as the longest distance between points in the two clusters.
- **Average Linkage:** It considers the average distance between points in two clusters.
- **Centroid Linkage:** It uses the distance between the centroids of two clusters.