Q1. Hierarchical clustering is a method of clustering data points into a hierarchy of nested clusters. It differs from other clustering techniques in that it does not require the number of clusters to be specified in advance and produces a tree-like structure (dendrogram) that illustrates the relationships between data points and clusters at different levels of granularity.

Q2. The two main types of hierarchical clustering algorithms are:

- Agglomerative hierarchical clustering: This bottom-up approach starts with each data point as its own cluster and iteratively merges the closest pairs of clusters until only one cluster remains.
- Divisive hierarchical clustering: This top-down approach begins with all data points in a single cluster and recursively divides the dataset into smaller clusters until each data point is in its own cluster.

Q3. The distance between two clusters in hierarchical clustering is determined using a distance metric, which measures the dissimilarity between clusters. Common distance metrics include:

- Euclidean distance
- Manhattan distance
- Pearson correlation
- Cosine similarity

Q4. The optimal number of clusters in hierarchical clustering can be determined using methods such as:

- Observation of the dendrogram: Visual inspection of the dendrogram to identify the level at which merging or splitting clusters leads to meaningful partitions.
- Cutting the dendrogram: Applying a threshold to the dendrogram to cut it into a specific number of clusters.
- Measures of cluster cohesion and separation: Calculating metrics such as the silhouette score or within-cluster sum of squares to identify the number of clusters that optimize cluster compactness and separation.

Q5. Dendrograms in hierarchical clustering are tree-like diagrams that illustrate the arrangement of clusters at different levels of similarity or dissimilarity. They are useful in analyzing the results of hierarchical clustering by visually representing the hierarchical relationships between data points and clusters, allowing users to interpret the

structure of the data and make decisions about the number of clusters to retain.

Q6. Yes, hierarchical clustering can be used for both numerical and categorical data. The distance metrics used for numerical data (e.g., Euclidean distance) measure the dissimilarity between data points based on their numerical values. For categorical data, specialized distance metrics such as the Jaccard distance or Hamming distance are used, which account for differences in the categorical attributes.

Q7. Hierarchical clustering can be used to identify outliers or anomalies in data by examining the structure of the dendrogram. Outliers are often located in singleton clusters or clusters with very few members that are distant from the main body of clusters. By inspecting the dendrogram, outliers can be identified based on their position and isolation within the hierarchy of clusters. Additionally, techniques such as pruning the dendrogram or setting distance thresholds can be used to automatically detect and flag outliers in hierarchical clustering results.