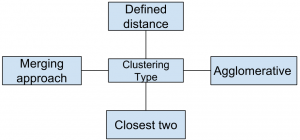
**Data Science Questions and Answers – Clustering**

**This set of Data Science Multiple Choice Questions & Answers (MCQs) focuses on “Clustering”.**

1. Which of the following clustering type has characteristic shown in the below figure ?  
[](https://www.sanfoundry.com/wp-content/uploads/2015/10/data-science-questions-answers-clustering-q1.png)  
a) Partitional  
b) Hierarchical  
c) Naive Bayes  
d) None of the Mentioned  
View Answer

Answer: b  
Explanation: Hierarchical clustering groups data over a variety of scales by creating a cluster tree or dendrogram.

2. Point out the correct statement:  
a) The choice of an appropriate metric will influence the shape of the clusters  
b) Hierarchical clustering is also called HCA  
c) In general, the merges and splits are determined in a greedy manner  
d) All of the Mentioned  
View Answer

Answer: d  
Explanation: Some elements may be close to one another according to one distance and farther away according to another.

3. Which of the following is finally produced by Hierarchical Clustering ?  
a) final estimate of cluster centroids  
b) tree showing how close things are to each other  
c) assignment of each point to clusters  
d) all of the Mentioned  
View Answer

Answer: b  
Explanation: Hierarchical clustering is an agglomerative approach.

4. Which of the following is required by K-means clustering ?  
a) defined distance metric  
b) number of clusters  
c) initial guess as to cluster centroids  
d) all of the Mentioned  
View Answer

Answer: d  
Explanation: K-means clustering follows partitioning approach.

5. Point out the wrong statement:  
a) k-means clustering is a method of vector quantization  
b) k-means clustering aims to partition n observations into k clusters  
c) k-nearest neighbor is same as k-means  
d) None of the Mentioned  
View Answer

Answer: c  
Explanation: k-nearest neighbor has nothing to do with k-means.

6. Which of the following combination is incorrect ?  
a) Continuous – euclidean distance  
b) Continuous – correlation similarity  
c) Binary – manhattan distance  
d) None of the Mentioned  
View Answer

Answer: d  
Explanation: You should choose a distance/similarity that makes sense for your problem.

7. Hierarchical clustering should be primarily used for exploration.  
a) True  
b) False  
View Answer

Answer: a  
Explanation: Hierarchical clustering is deterministic.

8. Which of the following function is used for k-means clustering ?  
a) k-means  
b) k-mean  
c) heatmap  
d) None of the Mentioned  
View Answer

Answer: a  
Explanation: K-means requires a number of clusters.

9. Which of the following clustering requires merging approach ?  
a) Partitional  
b) Hierarchical  
c) Naive Bayes  
d) None of the Mentioned  
View Answer

Answer: b  
Explanation: Hierarchical clustering requires a defined distance as well.

10. K-means is not deterministic and it also consist of number of iterations.  
a) True  
b) False  
View Answer

Answer: a  
Explanation: K-means clustering produces final estimate of cluster centroids.