Congratulations! You passed!

Grade received 91.67% To pass 80% or higher

Go to next item

1. Fill in the blank: K-means is an unsupervised partitioning algorithm used to organize _____ data into clusters. 1 / 1 point

- (•) unlabeled
- ( ) subcategorized
- ( ) presorted
- ( ) hierarchical

Correct
K-means is an unsupervised partitioning algorithm used to organize unlabeled data into clusters. It does this by creating a logical scheme to make sense of the data.

2. In k-means, what term describes the point at which each cluster is defined? 1 / 1 point

- (•) Centroid
- ( ) Core
- ( ) Commonality
- ( ) Coordinate

Correct
In k-means, the centroid is the point at which each cluster is defined. Its position represents the center of the cluster, also known as the mathematical mean.

3. Which of the following situations in the underlying data can be problematic for a K-means model? Choose all that apply. 0.75 / 1 point

- [ ] Unlabeled data
- [x] Too many clusters

This should not be selected
K-means models can handle any number of clusters. If the data truly contains many clusters, a K-means model may be suitable.

- [x] Too many dimensions

Correct
K-means can suffer from the "curse of dimensionality," whereby having too many dimensions in your data increases the distance between observations to such a degree that they are all too far away from each other to assign them to meaningful clusters.

- [x] Long, narrow clusters

Correct
K-means works best when the clusters are approximately circular, because it uses Euclidean distance from its centroids to assign cluster membership.