

A.

kNN uses the nearest k neighbors of an observation to put it into a group.

Decision trees split the data into smaller rectangular sections until the observations in a region are similar.

B.

kMeans clustering works by randomly assigning observations to a closest centroid, then recalculating the centroids to repeat until a desired convergence is reached.

Hierarchical clustering works by placing each observation into its own cluster, then combining the two closest clusters iteratively until all of them are connected.

Model-based clustering assumes the data was generated by a model and tries to find a matching model and clusters.

C.

PCA reduces the number of variables until it reaches a certain threshold for accuracy. It is useful for identifying less useful variables and can simplify the data.

LDA finds a combination of variables that maximize separation. It is particularly useful for classification. LDA is supervised while PCA is not.