1.a(only 2)

2.d(1,2and 4)

3.a(true)

4.1(only a)

5.b(1)

6.b(no)

7.a(yes)

8.d(All of the above)

9.a(K-means clustering algorithm)

10. d(All of the above)

11. d(All of the above)

12.(A) **Is K sensitive to outliers**

Answer-The *K*-means clustering algorithm is sensitive to outliers, because a mean is easily influenced by extreme values. *K*-medoids clustering is a variant of *K*-means that is more robust to noises and outliers. Instead of using the mean point as the center of a cluster, *K*-medoids uses an actual point in the cluster to represent it. Medoid is the most centrally located object of the cluster, with minimum sum of distances to other points. Figure [1] shows the difference between mean and medoid in a 2-D example. The group of points in the right form a cluster, while the rightmost

point is an outlier. Mean is greatly influenced by the outlier and thus cannot represent the correct cluster center, while medoid is robust to the outlier and correctly represents the cluster center.

(b) **Why is K means better**

Answer-Relatively simple to implement.

Scales to large data sets.

Guarantees convergence.

Can warm-start the positions of centroids.

Easily adapts to new examples.

Generalizes to clusters of different shapes and sizes, such as elliptical clusters

(C) **. Is K means a deterministic algorithm**

Answer- The basic k-means clustering is based on a non-deterministic algorithm.

This means that running the algorithm several times on the same data, could give different results..