Machine Learning Assignment-2

1. A
2. D
3. A
4. A
5. B
6. B
7. A
8. D
9. A
10. D
11. D
12. k-means can be quite sensitive to outliers in your data set. The reason is simply that k-means tries to optimize the sum of squares. And thus a large deviation (such as of an outlier) gets a lot of weight.
13. K means is better because of the following reasons:
    i. Relatively simple to implement.
    ii. Scales to large data sets.
    iii. Guarantees convergence.
    iv. Can warn-start the positions of centroids.
    v. Easily adapts to new examples.
    vi. Generalizes clusters of different shapes and sizes, such as elliptical clusters.
14. 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.