Machine Learning Worksheet 3

1. D)
2. D)
3. C)
4. B)
5. D)
6. C)
7. D)
8. A)
9. A)
10. B)
11. A)
12. B)
13. Clustering is useful for exploring data. If there are many cases and no obvious groupings, clustering algorithms can be used to find natural groupings. **Clustering** can also serve as a useful data-preprocessing step to identify homogeneous groups on which to build supervised models.
14. K-means clustering algorithm can be significantly improved by using a better initialization technique, and by repeating (re-starting) the algorithm. When the data has overlapping clusters, k-means can improve the results of the initialization technique. When the data has well separated clusters, the performance of k-means depends completely on the goodness of the initialization.