1. Movie Recommendation systems are an example of:
Ans. d) 2 and 3

2. Sentiment Analysis is an example of:
Ans. d) 1, 2 and 4

3. Can decision trees be used for performing clustering?
Ans. a) True

4. Which of the following is the most appropriate strategy for data cleaning before performing clustering analysis, given less than desirable number of data points:
Ans. a) 1 only

5. What is the minimum no. of variables/ features required to perform clustering?
Ans. b) 1

6. For two runs of K-Mean clustering is it expected to get same clustering results?
Ans. b) No

7. Is it possible that Assignment of observations to clusters does not change between successive iterations in K-Means?
Ans. a) Yes

8. Which of the following can act as possible termination conditions in K-Means?
Ans. d) All of the above

9. Which of the following algorithms is most sensitive to outliers?
Ans. a) K-means clustering algorithm

10. How can Clustering (Unsupervised Learning) be used to improve the accuracy of Linear Regression model (Supervised Learning):
Ans. d) All of the above

11. What could be the possible reason(s) for producing two different dendrograms using agglomerative clustering algorithms for the same dataset?
Ans. d) All of the above

13. Why is K means better?
Ans. K-means is like the Exchange Sort algorithm. Easy to understand, helps one get into the topic, but should never be used for anything real, ever. In the case of Exchange Sort, even Bubble Sort is better because it can stop early if the array is partially sorted. In the case of K-means, the EM algorithm is the same algorithm but assumes Gaussian distributions for clusters instead of the uniform distribution assumption of K-means. K-means is an edge case of E-M when all clusters have diagonal covariance matrices. The Gaussian structure means that the clusters shrink-wrap themselves to the data in a very nice way. This gets around the serious objections you correctly raise in the question. And E-M is not much more expensive than K-means, really. (I can implement both in an Excel spreadsheet.) But for serious clustering applications, one should really look at the hierarchical spectrum from single-link to complete-link clustering.

14. Is K means a deterministic algorithm?
Ans. 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.