MACHINE LEARNING Worksheet 2

1. a
2. d
3. a
4. a
5. b
6. b
7. a
8. d
9. a
10. d
11. d
12. b
13. **The K-means clustering algorithm is sensitive to outliers**, because a mean is easily influenced by extreme values.
14. **K-means is better in number of ways:**
    - 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

15. One of the significant drawbacks of K-Means is its **non-deterministic nature**. K-Means starts with a random set of data points as initial centroids. This random selection influences the quality of the resulting clusters