**1. (True/False) k-means always converges to a local optimum.**

True

**False**

**True**

False

**2. (True/False) The clustering objective is non-increasing throughout a run of k-means.**

**True**

False

**3. (True/False) Running k-means with a larger value of k always enables a lower possible final objective value than running k-means with smaller k.**

**True**

False

**4. (True/False) Any initialization of the centroids in k-means is just as good as any other.**

True

**False**

**5. (True/False) Initializing centroids using k-means++ guarantees convergence to a global optimum.**

True

**False**

**6. (True/False) Initializing centroids using k-means++ costs more than random initialization in the beginning, but can pay off eventually by speeding up convergence.**

**True**

False

**7. (True/False) Using k-means++ can only influence the number of iterations to convergence, not the quality of the final assignments (i.e., objective value at convergence).**

True

**False**

**8. Consider the following dataset:**

|  |  |  |
| --- | --- | --- |
|  | X1 | X2 |
| Data point 1 | -1.88 | 2.05 |
| Data point 2 | -0.71 | 0.42 |
| Data point 3 | 2.41 | -0.67 |
| Data point 4 | 1.85 | -3.80 |
| Data point 5 | -3.69 | -1.33 |

**Perform k-means with k=2 until the cluster assignment does not change between successive iterations. Use the following initialization for the centroids:**

|  |  |  |
| --- | --- | --- |
|  | X1 | X2 |
| Cluster 1 | 2.00 | 2.00 |
| Cluster 2 | -2.00 | -2.00 |

**Which of the five data points changed its cluster assignment most often during the k-means run?**

Data point 1

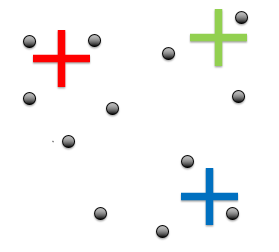
**Data point 2**

Data point 3

Data point 4

Data point 5

**9. Suppose we initialize k-means with the following centroids**



**Which of the following best describes the cluster assignment in the first iteration of k-means?**

