

# Unsupervised Learning - WEEK1 - Clustering - QUIZ 1 - Clustering

Link: [Unsupervised Learning - WEEK1 - Clustering - QUIZ 1 - Clustering](#)

**Your grade: 100%**

Your latest: **100%** • Your highest: **100%** • To pass you need at least 80%. We keep your highest score.

Next item →

1. Which of these best describes unsupervised learning?

1 / 1 point

- ☐ A form of machine learning that finds patterns in data using only labels (y) but without any inputs (x) .
- ☐ A form of machine learning that finds patterns using labeled data (x, y)
- ☐ A form of machine learning that finds patterns without using a cost function.
- ☒ A form of machine learning that finds patterns using unlabeled data (x).

✓ Correct

Unsupervised learning uses unlabeled data. The training examples do not have targets or labels "y". Recall the T-shirt example. The data was height and weight but no target size.

2. Which of these statements are true about K-means? Check all that apply.

1 / 1 point

- ☒ If each example  $x$  is a vector of 5 numbers, then each cluster centroid  $\mu_k$  is also going to be a vector of 5 numbers.

✓ Correct

The dimension of  $\mu_k$  matches the dimension of the examples.

- ☒ If you are running K-means with  $K = 3$  clusters, then each  $c^{(i)}$  should be 1, 2, or 3.

✓ Correct

$c^{(i)}$  describes which centroid example( $i$ ) is assigned to. If  $K = 3$ , then  $c^{(i)}$  would be one of 1, 2 or 3 assuming counting starts at 1.

- ☒ The number of cluster assignment variables  $c^{(i)}$  is equal to the number of training examples.

✓ Correct

$c^{(i)}$  describes which centroid example( $i$ ) is assigned to.

3. You run K-means 100 times with different initializations. How should you pick from the 100 resulting solutions?

- ☐ Pick the last one (i.e., the 100th random initialization) because K-means always improves over time
- ☒ Pick the one with the lowest cost  $J$
- ☐ Pick randomly -- that was the point of random initialization.
- ☐ Average all 100 solutions together.

✓ Correct

K-means can arrive at different solutions depending on initialization. After running repeated trials, choose the solution with the lowest cost.

4. You run K-means and compute the value of the cost function  $J(c^{(1)}, \dots, c^{(m)}, \mu_1, \dots, \mu_K)$  after each iteration. Which of these statements should be true?

1 / 1 point

- ☒ The cost will either decrease or stay the same after each iteration. .
- ☐ There is no cost function for the K-means algorithm.
- ☐ The cost can be greater or smaller than the cost in the previous iteration, but it decreases in the long run.
- ☐ Because K-means tries to maximize cost, the cost is always greater than or equal to the cost in the previous iteration.

✓ Correct

The cost never increases. K-means always converges.

5. In K-means, the elbow method is a method to

1 / 1 point

- ☐ Choose the best number of samples in the dataset
- ☐ Choose the best random initialization
- ☐ Choose the maximum number of examples for each cluster
- ☒ Choose the number of clusters K

✓ Correct

The elbow method plots a graph between the number of clusters K and the cost function. The 'bend' in the cost curve can suggest a natural value for K. Note that this feature may not exist or be significant in some data sets.