

## Question 4

Yufei Yin

### Question 4

- List three benefits of kmeans and three drawbacks.

#### benefits

- Scales to large data sets.
- Easily adapts to new examples.
- Guarantees convergence

#### drawbacks

- K-means chooses the initial centroid point randomly, and since the clustering accuracy depends on the initial choice of centroids, the accuracy can be low if the chosen centroids aren't proper.
- Considering only the distance between centroids of the cluster may not be efficient for categorical data.
- As the cluster size grows, it becomes a challenge to distinguish between two clusters based on a few attributes.

## Reference

<https://www.section.io/engineering-education/k-means-from-scratch-r/>