Homework 5

1. *Final grading with K-means.* So the semester is finally over. It was quite a large class with 56 hard-working students, and the instructor was not.

   As everyone probably remembers, there will not be a final exam for this course; students will be graded by the performances on their homeworks (HW#1 through HW#4). The professor has always been terribly lazy and this time he decided to run a K-means model on the students' homework scores, in order to categorize the students into three clusters (A, B, and C), so that the final grading is automatically assigned. Awesome...

   Download the anonymized data file, `homework_scores.csv`, from the course webpage, which contains 56×4 matrix where each row represents a student's four homework scores. Fit a K-means model with three clusters to the given data, by initializing the model with random centers. You may want to run the model several times to make it sure that your model achieve the lowest cost.

   (a) What are the representative homework scores for each group?

   (b) How many students will get A, B, and C?

   Don't worry by the way. Your professor will never do this.