CSC148 - Introduction to Assignment 1: Forming Student Groups

Below is a dataset containing students' answers to a survey:

Name	Year	College
Priya	3	Victoria
Alain	2	New
Zoe	3	Woodsworth
Francesco	3	Victoria
Mohammed	4	Woodsworth
Xiaoyuan	5	New
Rohit	2	New
Yimin	3	Trinity
Grace	5	Woodsworth
Claire	1	Woodsworth
Kai	1	Woodsworth

1. Suppose we are interested in using this information to put the students into groups of 4 for a course project. Discuss with one or more other students who you think should be put together. (Aside: What did you notice about group sizes?)

2. A simple way to put students into groups is to just count them off, so that Priya, Alain, Zoe and Francesco go in the first group, and so on. We'll call this **slicing** into groups. Mark the table above to show the groups.

3. For Assignment 1, you'll write a program to automate the grouping task using different algorithms. Your program will allow the professor to specify what they want the groups to be like. For instance, they may want this:

   (a) A group should be as homogeneous as possible in terms of what year the student is in.

   We call this a **criterion**. Which group best satisfies this criterion? Which group worst satisfies this criterion?

4. Your program is going to need to be able score a group on how well it meets a criterion. Suppose we score a group as follows:

   - Define the **scaled distance** between two student answers as the distance between their answers divided by the range of possible answers. (We'll say that the range is 5, supposing a student is in year 1 to 6, and $6 - 1 = 5$).
   - Define the **similarity** of two student answers as 1 minus their scaled distance.
   - Calculate the similarity for every pair of students in the group, and average these results.

   Compute the group score on criterion (a) for each group. Feel free to keep your answers as fractions.

5. What is the highest score a group could get on this criterion? The lowest?

Here is that dataset again:

Name	Year	College
Priya	3	Victoria
Alain	2	New
Zoe	3	Woodsworth
Francesco	3	Victoria
Mohammed	4	Woodsworth
Xiaoyuan	5	New
Rohit	2	New
Yimin	3	Trinity
Grace	5	Woodsworth
Claire	1	Woodsworth
Kai	1	Woodsworth

6. Your program will allow a professor to specify more than one criterion. Suppose they want to also use this one:

   (b) A group should have no member who is the only one from their college

   and that a group's score on this criterion is 1 if it satisfies the rule and 0 if it doesn't. (Yes, this is an all-or-nothing score.) Compute each group's score for criterion (b).

7. Let's say these two criteria are not of equal importance. Criterion (a) has relative weight 80 and criterion (b) has relative weight 20. What is the **total score** for each group based on **both** of these criteria? Feel free to give an approximate whole number answer.

8. What is the highest total score a group can earn, given these weightings? What is the lowest?

9. Compute the average score across all the groups. This is one way to measure how good the entire grouping is. Again, feel free to approximate.

10. We did simple "slicing" to form these groups. Can you come up with a better set of groups, producing a higher score for the entire grouping?