BSDS 100: Intro to Data Science with `R`
Assignment 5

by Abbie M Popa (University of San Francisco)

Directions: For all questions in this assignment, write complete sentences and fully answer any question that is asked, and use `R` to answer each question. Provide all `R` code and solutions by *knitting* your final RStudio file into a single file named `[your_name]_CA5.pdf`. Remember to use 'echo = TRUE' to print your code as well as your results. This assignment will be submitted on canvas by the due date before class (2:39 PM). The total assignment is worth 30 points, and late assignments will automatically have 10 points deducted.

1. (6 pts) Create the following vectors, populated with information about the courses for which you are enrolled this year in addition to one course (any course) that you are not enrolled.

   - `courseNum`: course number of each course

   - `coursename`: course name of each course

   - `courseProf`: name of the instructor for each course

   - `enrolled`: a logical vector indicating whether or not you are enrolled in the course

   - `anticipatedGrade`: your anticipated letter grade in each course, with an `NA` for any course for which you are **not** enrolled

   - `anticipatedHours`: your anticipated hours spent on each class per week based on on your experience so far, with an `NA` for any course for which you are **not** enrolled.

2. (5 pts) Create and print a data frame called `MyCourseDataFrame` by combining all of the above vectors. Assign the names of each column to be the names of the original vectors. Summarize the `type` of each column. Do the data frame variables retain their original classes? Formally test this using appropriate `R` code.

3. (5 pts) Combine the vectors from (1) into a list called `MyCourseDataList`, where each vector is an element of the list. Assign the names of each element to be the names of the original vectors. Do the elements of the list maintain their original classes? Formally test this using appropriate `R` code.

4. Write code that returns the following values. As always, use 'echo = TRUE' so that your code as well as your output is displayed after each calculation:

   - (2 pts) The values in `courseNum`, excluding the fourth value

   - (2 pts) The total number of hours you anticipate spending on coursework per week

   - (2 pts) A data frame with only the third row and first two columns of `MyCourseDataFrame`

   - (2 pts) The first value in the second element of `MyCourseDataList`

5. Convert the `anticipatedGrade` variable in `MyCourseDataFrame` into an ordered factor where your best anticipated grade is the maximum and your lowest anticipated grade is the minimum using the

function `factor()`. Note: to get an ordering of values, you'll have to use the argument `is.ordered` `= TRUE`. Look at the documentation of `factor()` to understand how to do this. Now write code to answer the following questions, and output the answers.

(a) (2 pts) What is the maximum letter grade you anticipate receiving this semester?

(b) (2 pts) What is the minimum number of hours you expect to work per week in a class this semester? Is this in this course?

(c) (2 pts) For (a) and (b), what is the name and course number of each class? Your code should provide the result as a single textual output with *both* course number and course name separated by a colon, e.g. 'BSDS100: Intro to Data Science with R'.