Group Project: Part II

Data Programming in R
Business Analytics/Professional MBA Program
MSCI:6060:0EXB/C Fall 2018 (Cedar Rapids)

This is the second portion of the group project. Please use the same group as you have for Part I when completing Part II

Summary of Deliverables and Deadlines

What?	When?	Where?
A single R script with the solution	6:00PM on Sunday, October 28th	ICON upload

Download the dataset entitled `Salary_Information_for_State_Authorities_NY.csv` which can be found on ICON and proceed to write a script so that, when sourced, the script provides the first and last name of the individual with the highest base annualized salary during a fiscal year end date that occurs at some point in 2015. Absolutely zero adjustments should be made to the Excel file, i.e. do not 'Save As' in Excel after downloading the file. The script should begin with the following 2 lines of code:

1. `rm(list=ls())`
2. `df <−read.csv(‘‘Salary_Information_for_State_Authorities_NY.csv’’)`

Key things to know about this portion of the project.

1. You will be graded on both the accuracy and elegance of your solution. In particular, if you find the correct person by just going through the Excel spreadsheet line-by-line and then write that name in your script, you would likely receive a 0. Instead, the script should be able to produce the correct answer even if new rows were added/subtracted/altered in the data set.
2. I will be providing zero guidance as to how to complete this portion of the project. Successfull completion of this part will likely require you to utilizie functions that we have not directly covered in class. It is expected that you will consult other resources (such as Google, books on R, etc.) in order to find a way to complete this problem. This problem may take some groups 5 minutes, and it may take other groups considerably longer, i.e. hours and hours.
3. Please comment your code to indicate what you are doing, especially when utilizing functions that we have not specifically used in class. Regardless of what functions you use (even if they are all taken from class), you must outline your methodology with comments in your code. Do not expect to receive full credit without well documented code.
4. All other guidelines from a typical HW exercise apply to the script that you will be turning in.

5. Of the 25 points allocated for Part II, 10 points will be for the correct answer and 15 points will be for the methodology. This means that an incorrect answer will result in a maximimum possible score of 15 out of 25, even if there was just a single type-o causing the error in an otherwise perfectly thought out script. Additionally, a correct answer stemming from an unnacceptable methodology (such as copying another group, guessing, or just viewing the Excel file, for example) will receive no credit (and the methodology will also receive no credit).