This homework is meant to help you get started with R.

R is a language and environment for statistical computing and graphics, available as a free open-source software. It is one of the prominent data science and statistics languages, which is why it is in your own great professional interest to take this assignment seriously and devote sufficient time and effort to it.

When it comes to learning and mastering R, practice is the key, and self-learning "by doing" is the way here. In fact, most R-users would argue it is the only way, which is why learning to code in R will be largely your own responsibility. In my own experience, the most effective way to learn R is by tinkering with it on one's own.

There is an immense amount of tutorials, books, videos, and free courses available online. Also, always remember that virtually any questions you will ever have about coding in R have already been asked by someone else and answered on numerous R-help forums. Just google it (our use ChatGPT)!

To receive full credit for this homework, your D2L submission should be two files:

1. Your solutions for Problem 2 in the file "Lastname_Firstname_HW1_P2.R"
2. Your solutions for Problem 3 in the file "Lastname_Firstname_HW1_P3.pdf"

This assignment is worth 100 points (10%) of your grade.

Problem 1 (20 Points)

Watch the following introductory video-course on YouTube made by datalab.cc: R Programming Tutorial - Learn the Basics of Statistical Computing (clickable link).

You can also access this video-course from the datalab.cc website (clickable link).

- The video-course will show you how to install R and R-Studio as well as provide an excellent overview of basics to get you acquainted with the software/language.
- This playlist contains 20 short videos and, in total, is 2 hours 11 minutes long.
- When watching, you can skip the "Hierarchical Clustering" and "Principal Components" segments.
- To maximize benefit, you are required to not just watch the video but to follow along in R-Studio. Download the course files here (clickable link). All R-scripts are enumerated according to the video segments for easy navigation.

Problem 2 (35 Points)

Watch a different introductory video-course on YouTube, this time made by Dynamic Data Script: R Programming for Beginners — Complete Tutorial — R & RStudio (clickable link).

Expect some overlap between this and the video-tutorial from Problem 1. That is okay: repetition is great for learning.

- The video-course provides an excellent overview of basics to get you acquainted with the software/language.
- The video is 49 minutes long and is conveniently sliced into 16 short segments.
- To maximize benefit, do not just watch the video, but follow along in R-Studio by replicating every command that the instructor is executing in the R console in the video.

  To motivate you to follow along with the video: **to receive credit for this problem, you are required to submit an R-script of all commands executed in this video**:

  (i). The code should be annotated using comments;

  (ii). The R-script should be saved as as .R file (i.e., do NOT convert to PDF) using the "Lastname_Firstname_HW1_P2.R" naming convention where Lastname corresponds to your last name and Firstname corresponds to your first name.

Problem 3 (45 Points)

Install the swirl package and complete the interactive R_Programming_E course.

- To install the package and the assigned course, run the following code in R:
  1. install.packages("swirl")
  2. library(swirl)
  3. install_course_github("swirldev", "R_Programming_E")
- After installing swirl, in order to use it, you just need to call the package every time you start a new R session via library(swirl) (as you would with any other package) followed by the swirl() command. That is, you would need to run the following code:
  1. library(swirl)
  2. swirl()

Running this code will initialize the interactive swirl environment in which you will be prompted to complete various R courses.

- After initializing swirl, just follow the instructions shown in the R console. When swirl prompts you to choose a course to complete, type "1" to select 1: R Programming E.

  You will then be asked to choose a lesson, of which there are 15 in total. **To receive full credit for this problem, you are required to complete the following lessons (all but lesson 8)**:

  1. Basic Building Blocks
  2. Workspace and Files
  3. Sequences and Numbers
  4. Vectors
  5. Missing Values
  6. Subsetting Vectors
  7. Matrices and Data Frames
  9. Functions
  10. lapply and sapply
  11. vapply and tapply
  12. Looking at Data
  13. Simulation
  14. Dates and Times
  15. Base Graphics

  After you complete each lesson, you will be asked:

  Would you like to inform someone about your successful completion of this lesson via email?

  Type "2" to select "2: No." BUT you need to record (by copy-pasting) the following:

  (i). The penultimate (second to last) swirl instruction of the lesson (usually, in red color). For instance, the penultimate swirl instruction for the lesson Basic Building Blocks is "You can type the first two letters of the variable name, then hit the Tab key (possibly more than once). Most programming environments will provide a list of variables that you've created that begin with "my". This is called auto-completion and can be quite handy

when you have many variables in your workspace. Give it a try. (If auto-completion doesn't work for you, just type my_div and press Enter.)"

(ii). The corresponding "progress percentage" that swirl showed on the screen in the R console before prompting the email question. For instance, the progress percentage that swirl showed on the screen in the R console before prompting the email question for the lesson Basic Building Blocks is 97%.

To receive credit for this problem, you are required to submit the record containing (i). and (ii). for each of the 15 lessons to D2L as the PDF "Lastname_Firstname_HW1_P3.pdf" where Lastname corresponds to your last name and Firstname corresponds to your first name.

- You can exit swirl and return to the R prompt (>) at any time by pressing the Esc key. If you are already at the prompt, type bye() to exit and save your progress. When you exit properly, you'll see a short message letting you know you've done so.
- When at the R prompt (>) in the swirl environment:
  - typing play() lets you experiment with R on your own; swirl will ignore what you do UNTIL you type nxt() which will regain swirl's attention
  - typing bye() causes swirl to exit; your progress will be saved
  - typing main() returns you to swirl's main menu
  - typing info() displays these basic command options