 Due by 2:30pm, 28^{th} Sept 2018

CMPSC 301
Data Analytics
Fall 2018

Lab 3: Introduction to R
21^{th} Sept 2018

Objectives

To enhance the understanding of the basic R functionality, including the use of R Studio and producing data visualizations.

Reading Assignment

Please read Chapters assigned for this week's lessons which you will find in the class slides, in addition to reviewing your notes. Please take some time to gain experience with using Markdown to write your work. See *Mastering Markdown* `https://guides.github.com/features/mastering-markdown/` for more details about Markdown.

GitHub Starter Link

https://classroom.github.com/a/bCaJqLR3

To use this link, please follow the steps below.

- Click on the link and accept the assignment.
- Once the importing task has completed, click on the created assignment link which will take you to your newly created GitHub repository for this lab.
- Clone this repository (bearing your name) and work on the lab locally.
- As you are working on your lab, you are to commit and push regularly. You can use the following commands to add a single file, you must be in the directory where the file is located (or add the path to the file in the command):
  - `git commit <nameOfFile> -m ''Your notes about commit here''`
  - `git push`

  Alternatively, you can use the following commands to add multiple files from your repository:
  - `git add -A`
  - `git commit -m ''Your notes about commit here''`
  - `git push`

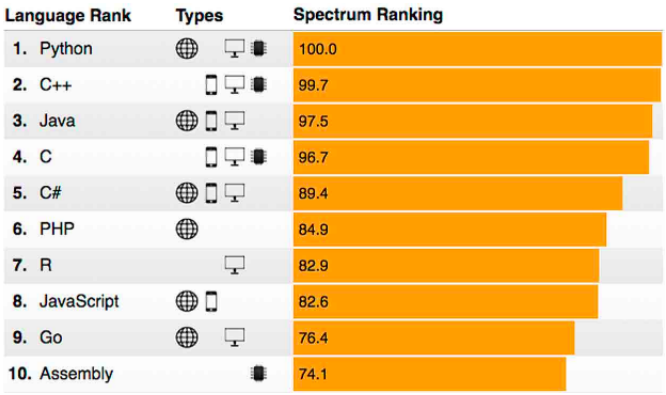

Figure 1: IEEE's list of 2018's most popular languages.

Exploring R Programming

As shown in Figure 1 (IEEE's listing of 2018's top programming languages, `https://spectrum.ieee.org/at-work/innovation/the-2018-top-programming-languages`), R has become increasingly popular as an extensive language for industrial and academic applications. Thanks to its community of developers who write its open source libraries, R continues to grow. Besides its popularity, R is the most recommended language to learn when getting started with data analytics as it is created with statistics and data in mind. In this lab you are invited to complete some questions out of the book where you are to implement code in your responses.

In this lab you are asked to read the assigned sections of the book and complete all exercises from Chapters 3, 4 and 6 of the "R for Data Science" textbook (online version). As you remember, the chapter numbering in the printed version of the book is not the same as the numbering in the online version of the book. This lab's assigned exercises correspond to the following exercises in the online version of the book, found in `http://r4ds.had.co.nz/`.

For each question, please be sure to complete all its parts. The questions to answer are the following and have been taken from the book's web site.

{3.2.4, 3.3.1, 3.5.1, 3.6.1, 3.7.1, 4.4, 6.3.1 (See Twitter application note below}

Note: For the question 6.3.1 (Twitter application), the link will take you to a listing of tips and tricks from the community which may be used in R to obtain results from Twitter data. Find an interesting tip, explain it what it does and then discuss how it is useful to you.

Submission Information

Your answers to the assigned questions are to be typed up using the Markdown file `writing/responses.md`. Please be sure to add and label each question and its parts.

Required Deliverables

1. Your **labeled** answers to the assigned exercises are to be placed in `writing/responses.md`

When you have finished, please ensure that the GitHub web site has your pushed work by visiting your repository at the site. Please see the instructor if you have any questions about assignment submission.