EDRM 718 Project A
Writing Simple Scripts and Using Available Scripts

Directions: This project will require you to fork and clone a GitHub repository and then write a script that uses the data from that repository. You will also modify one of the scripts in that repository so that the script can be used to obtain additional data that you will analyze. Your data analysis should be in an R notebook, with narrative, that you should push to your GitHub repository prior to the deadline.
(100 points possible)

1. The website FiveThirtyEight published an interactive graphic titled Gun Deaths in America that illustrated the nature of gun deaths for the years 2012 to 2014. The data and code for the project are in a GitHub repo. Fork and clone this repo.

2. Write a report to analyze the data from 2012. Specifically, we are interested in the proportions of different types of gun deaths, the proportions of each gender type, and the proportion of each race type. Your report should include graphs, descriptive statistics, and accompanying narrative. You can see Instructor Report 1 for an example and watch the accompanying videos for an example of completing this type of analysis. (Your report does not need a press release.)

   Some possible functions to consider: subset, table, prop.table, barplot

3. The file CDC_parser.R in the repo was used to download the data from the Center for Disease Control (CDC) and to prepare it for analysis. Save this script with a different name (e.g., CDC_parser_19.R) so that you can modify it and retain the original script. Modify the script to download the 2019 data, which is the most recent data available from the CDC. Listed below are the changes you need to make. Note that as you make changes the line numbers may change. The line numbers below are based on the original file, so you may want to have your two scripts opened side-by-side when you make the changes to your revised script.

   line 151: change the year to 2019
   line 152: change the 2013 in the URL to 2019
   line 166: change the number in the file name to 19
   lines 167 & 168: delete these lines
   line 170: replace the rbind function (including arguments) with guns_19
   line 207: change the file name from all_guns.RData to all_guns_19.RData
   line 208: change the file name from full_data.csv to full_data_19.csv

4. Install two packages: tidyverse and magrittr

5. Run your script. (You may want to go get a beer. This will take a while.)

6. Add to your notebook an analysis of the 2019 data by addressing the same questions you addressed for 2012.

7. For each of the variables of interest, compare proportions for the two years (2012 and 2019) with graphs and descriptive statistics.