Tentative Timeline and Expectations
## Summary 2022 Internship

This document will be updated regularly.

Updated: August 10, 2022

Overview

The visualization below is an overview of the summer 2022 internship timeline with events and landmarks. Each vertical line is one day and the time labels are shown in one week intervals. Note that the timeline is tentative and can change based on unique circumstances.
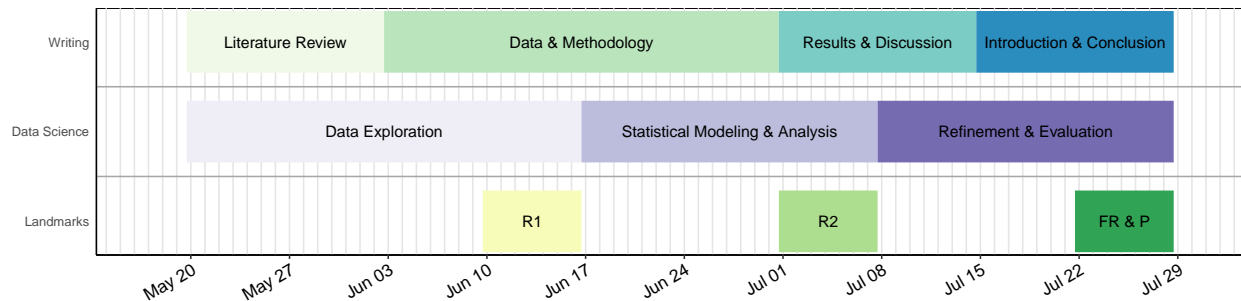
Expectations

- The student will write short weekly reports using pdf-knitted RMarkdown, which are expected to be done each Friday.

- The weekly report shall contain the following:
    - A summarized written understanding of the topic for the week.
    - A summary of the tasks completed and a progress report related to the main project.
    - A short reflection of the week explaining the time and energy it took to complete the tasks and a short list of item on what can the supervisor do to support you and to ameliorate some limitations.

- The student will write landmark reports (R1, R2, FR & P) related to the main project.

- The student will turn-in (upload/push) all relevant work into the project's Github repository https://github.com/stressosaurus/police-violence-study.

- The student is only expected to work at most 20 hours per week and they can freely choose how to allocate their time. If any problems or concerns arises, please let the supervisor now.

Books and Online Resources List

Textbooks

[**AUSCD**] Walker, K. (2022). Analyzing US Census Data: Methods, Maps, and Models in R. CRC Press.

[**PSD**] Darrin, S., & Bryan, C. (2021). Probability, Statistics, and Data: A Fresh Approach Using R.

[**TMR**] Kuhn, M., & Silge, J. (2020). Tidy Modeling with R.

[**RMC**] Xie, Y., Dervieux, C., & Riederer, E. (2020). R markdown cookbook. Chapman and Hall/CRC.

[**GR**] Lovelace, R., Nowosad, J., & Muenchow, J. (2019). Geocomputation with R. Chapman and Hall/CRC.

[**B**] Xie, Y. (2016). Bookdown: authoring books and technical documents with R markdown. Chapman and Hall/CRC.

[**TMR**] Silge, J., & Robinson, D. (2017). Text mining with R: A tidy approach. " O'Reilly Media, Inc.".

[**ISLR**] James, G., Witten, D., Hastie, T., & Tibshirani, R. (2013). An introduction to statistical learning (Vol. 112, p. 18). New York: springer.

[**HGR**] Happy Git with R

Timeline Details and Reading Materials

Note: The "Readings" subsection in the list below contains a number on which it refers to a numbered item in the Books and Online Resources List. For example "[PSD]" refers to the first item in the list, which is our main text book titled "Probability, Statistics, and Data: A Fresh Approach Using R".

Week 1 (5/20-5/27)

Readings and Exercises:

- [PSD] Chapter 1, 6, and 7 | Exercises: NA
- [HGR] Chapter 9-14, 16 & 18, and 20-23 | Exercises: NA
- [RMC] Chapter 2, 3 and 4.5 | Exercises: NA

Development Tasks:

- Test making a branch and push your branch to the Github repository.
- Merge your branch to the main branch, and push the updates to the Github repository.
- Add test reference and citation within an R Markdown file and include it into the weekly report.

Project Tasks:

- List all variables of the data set with their detailed descriptions and variable types (numerical, categorical, etc.).
- Under the "literature" directory: These documents will be regularly updated when new references are discovered.
  - "references.bib": Create a bibtex file which contains the bibtex references of the project.
  - "references-list.Rmd": Create a separate document that lists the research papers and references.
  - "annotated-bibliography.Rmd": Create a separate document where we can put annotated bibliographies.

Week 2 (5/27-6/3)

Readings and Exercises:

- [PSD] Chapter 2, 3, and 4 | Exercises: NA
- [PSD] Chapter 5 | Exercises: 5.8, 5.13, 5.14, 5.16, 5.19, 5.24, 5.30, 5.31, 5.32, 5.33
- [PSD] Chapter 8 | Exercises: 8.29, 8.36, 8.39, 8.46, 8.47, 8.49, 8.50, 8.51, 8.52, and 8.55
- [PSD] Chapter 9 | Exercises: 9.1, 9.2, 9.3, 9.5, 9.7, 9.11, 9.12, 9.20, 9.24, and 9.27
- [GR] Chapter 1 | Exercises (External): Go to MATH241 S22 Module 3 and do the exercises in Section I.

Project Tasks:

- Continue literature review.
- Continue exploring the data sets.

Week 3 (6/3-6/10)

Readings and Exercises:

- [PSD] Chapter 11 | Exercises: 11.3, 11.9, 11.11, 11.13, 11.15, 11.19, 11.25, 11.29, 11.32, and 11.33
- [PSD] Chapter 12 | Exercises: 12.2, 12.3, 12.4, 12.5, 12.7, 12.9, 12.12, 12.13, 12.21, and 12.22
- [PSD] Chapter 13 | Exercises: 13.1, 13.2, 13.3, 13.4, 13.5, 13.6, 13.7, 13.8, 13.9, and 13.10

Project Tasks:

- Continue literature review.
- Continue exploring the data sets.

Week 4 (6/10-6/17)

Project Tasks:

- **Landmark Report R1 Guidelines:** The list below outlines the expectations of the R1 document. All relevant R codes should exist throughout the report and a short introduction. Please follow standard bibtex referencing and citations using the APA format.
  - *Relevant literature review and background research:* The report should contain a summary of the related research papers and background research of the project topic. Give some context for the topic: why it's fascinating, who's interested, what's known about it, some references, and so on.
  - *Detailed descriptions of all data sets:* The report should contain descriptions of the data set variables - how are they obtained, when are they taken, what are their variable types.
  - *Data explorations:* The report should contain the explorations of any associations/relationships between the variables. This includes all visualizations of distributions and associations between the variables, and any descriptive statistics. Include any discussions about the results.

Week 5 (6/17-6/24)

Readings and Exercises:

- [TMR] Chapters 7 - 12 | Exercises: NA

Project Tasks:

- Work on the FBI crime data (see raw-data/fbi-cd/scraper.R file).
- Continue working on the producing results.

Week 6 (6/24-7/1)

Readings and Exercises:

- [TMR] Chapter 16 - 21 | Exercises: NA

Project Tasks:

- Continue working on the FBI Crime data.
- Continue working on the producing results.

Week 7 (7/1-7/8)

Project Tasks:

- **Landmark Report R2 Guidelines:** The list below outlines the expectations of the R2 document. All relevant R codes should exist throughout the report and a short introduction. Please follow standard bibtex referencing and citations using the APA format.
  - *Research questions:* Draft three or more research questions during data exploration and applying methods. We will discuss these questions later and choose which ones can we present - depending on the results.
  - *Results and Discussions:* The report should contain the relevant results which includes tables, figures, and statistical model summaries. Please include initial analyses and discussions. We will further discuss these results later.

Week 8 (7/8-7/15)

Project Tasks:

- Continue working on the R2 report.

Week 9 (7/15-7/22)

<<<<

Week 10 (7/22-7/29)

======= Project Tasks:

- **Final Report Guidelines:** The list below outlines the expectations of the Final Report document. All relevant R codes should exist throughout the report and a short introduction. Please follow standard bibtex referencing and citations using the APA format.
    - *The final report should contain the introduction, methods, results, conclusions, and future work:* Most of these sections are already written from the previous two reports. The task here is to summarize the data wrangling steps, visualizations (if any), any modeling methods used, and discussion of the results.

Week 10 (7/22-7/29)

Project Tasks:

- Work on the project poster that summarizes the data wrangling, visualizations, and results.