Project 4

Agreement

This Project is meant to be an assessment of your ability to present the work from Project 3 in a written report format.

This project is an individual evaluation. Each student, even if working together for other projects, is to write and submit their own, unique report. Visualizations may be duplicated between students working on the same project but text should not.

You may use any resource, either online or physical, to complete the work. This includes:

- Any help forum or website (e.g. StackOverflow) questions that already exist.
- Any notes, code, slides, papers, or previous feedback from the instructor.
- Any books, online or physical.
- Scholarly works such as papers.
- Help from generative artificial intelligence such as ChatGPT.

It DOES NOT include:

- Help from homework websites such as Course Hero or Chegg.
- Help from students or persons outside of those currently enrolled in the current semester of CPSC 292.

If you use work outside normal course resources (textbooks, lecture notes, slides, code, or instructor feedback), you are expected to cite the work by providing a URL to the source near the place that the code was used. If generative AI is used, the use must conform to the authorized use discussed in the course syllabus and you must disclose use on the report. **Failure to disclose AI use or use that is inconsistent with the guidelines on the syllabus will constitute unauthorized AI use and is subject to sanction as an academic integrity violation.**

Instructions

- **Step 1:** You will work with your visualizations from Project 3. In your R Project, create a "Project4" RMD file. Create a report, similar to a lab report, with the sections and requirements outlined below. This should have the same structure as Project 3 in a written format with more room to expand on ideas.
- **Step 2:** Turn in the final version of your full R Project as a compressed zip folder by **Wednesday, December 10 by 5 pm**.

Additional Information

The report should have the following sections:

1. **Background**: This section is a narrative of the background and motivation behind the visualizations in Project 1 or 2. It gives the reader the appropriate background to understand the motivation and the motivation itself. The background should build and reference previous works. It should clearly lay out any information needed for the reader to understand the question and the visualization (but not anything else!). It should end with the question that motivates the visualization.
2. **Analysis**: This section is what you used and what you did. It should be detailed enough so that the reader could *replicate*, or repeat, the analysis if they wished. It should also be a narrative (not a list of materials or a set of numbered steps) with as much details about the analysis as possible. You should include any statistical tests you performed and what software you used. Any data loading, cleaning, or manipulation should be included in this section as code included in the report. Take this space to embed your reasoning and explanation with the code itself!
3. **Results & Discussion**: You report and interpret results in this hybrid section. You should reference figures (and tables, if present) in your reporting of results. Report the results of any statistical tests you performed (if present). Interpret the results for your readers to tell them what the visualizations

show or what unique insights the visualizations provide. You must include at least one visualization from Project 1 or 2, but you can add additional ones if you wish. These additional visualizations should only serve to highlight the main analysis, rather than be extra decorations.

4. **Conclusions**: The main takeaway points of your study. This can be a bullet list if you wish!
5. **References**: A list of full citations that correspond to the papers you referenced in the text of your report. In-text citations are required to indicate the position in the text where the info in your reference was used!

Other required elements of the report (RMD file):

- Code: All code should be included in the final report *but not visible*. Code chunks should include `include=FALSE` if the chunk contains no graphics and `echo=FALSE` if the code chunk contains graphics.
- At least one visualization in your Project 1 or 2.
- A full reference to your data set, including the location/URL of the set.
- **Include at least two references to peer-reviewed papers** using an in-text citations at their point of reference in the text and a full reference at the end of the document. Remember that peer-reviewed publications include journal articles, review papers, and academic books, but NOT lecture notes, talks, newspaper articles, blog posts, textbooks, etc. A reference to your data set is not part of the two peer-reviewed references (so you should have a minimum of three references total.)
- In-text citations for all references.
- **All data plots must be generated within the R Markdown file itself. The R Markdown file must knit to a HTML or DOC.**
- Captioning figures: You will need to caption figures in order to refer to them in your text. This is really easy, just add `fig.cap = "your text"` to the header along with echo and warning. Be sure to separate each command in the header with a comma.

For the final report, you will turn in a zip file that contains all the necessary elements to reproduce your report using RMarkdown and RStudio. This could include the following:

- An R Project file and folder that contains all work necessary to generate the report. All file should be in standard project format or modified standard project format.
- An RMD file which will knit to either HTML or DOC.
- Data files in a `data/` directory.
- Image files if there are pictures in `doc/` or `document/`.
- bib files for references, should be in `doc/` or `document/`.
- Other files, if necessary, in appropriate locations.

The best way to check to see if you are turning in a complete package is to zip everything up, move it to a different location of your computer, delete the html or doc knitted file, restart R and clear your global environment, then try to knit the file. If it works, you're ready to submit!

Project 4 files are expected to adhere to the same best practices as Project 2.

How Project 4 will be graded

As with all work in the course, Project 2 will be assessed for completeness, in other words, you will receive a complete (2 awarded course points) or incomplete (0 awarded course points). Completion will be scored on:

- Content: Presence and completeness of each of the project items and requirements outlined above. Each file and section of the RMD file should be present to receive a complete. See "Project4-Additional.Rmd" for the specific rubric used to assess content.
- Replication: Successful replication of the Project by the instructor. The replication process is as follows:
    1. The zip file is downloaded and uncompressed.
    2. The RStudio project is opened using the RProj file.
    3. The RMD is opened and "Knit" button is pressed.

    If the RMD fails to knit, or if there are errors produced in the code, the project will receive an incomplete.