# Project 2

## Agreement

This Project is meant to be an assessment of your ability to creatively execute several learning objectives at once by improving the code from Project 1 or creating visualizations using a new data set.

In this project, you have the option to work individually or with another student in a group. If you are working with another student, both students will receive the same score for the project. Additionally, if you work with another student, you will be required to use Github and version control software (git) to complete the project.

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, you must include a section within the main documentation of the project detailing how and why you used generative AI including a link to the specific conversation(s) used to generate project content. **Failure to include this section will constitute unauthorized AI use and is subject to sanction as an academic integrity violation.**

## Instructions

- **Step 1:** Decide on how you would like to move forward with your project. Your options are:

  1. Improve your current project by yourself.
  2. Join together with another student and move one of your two projects forward.
  3. Decide to start on a different data set by yourself.
  4. Decide to start a new data set with a partner.

  You should submit your decision on Canvas by **Oct. 31, 2025 at 5 pm**. Any option other than #1 must be approved by the instructor *before* the deadline. If you are choosing option #2 or #4, submit the URL of your public github repository that you'll share with your partner when submitting your decision on Canvas.

- **Step 2:** Project 2 will be at least one data visualization (no more than two). The visualization(s) can be the same as project 1 or different. However, the code should be improved. The code producing your visualization should include at least **three elements from at least two categories** the following list:

  – a function created by you that is used at least twice;
  – a loop; and/or
  – a conditional statement.

  These elements need to exist in non-trivial way that adds to the complexity of the project. Your Project 2 code must conform to the Best Practices document in the Project 2 folder in the course repository. Refactor your project so that it conforms to these standards. Submit a draft of your project to Canvas by **Nov. 10, 2025 at 5 pm**.

- **Step 3:** Present a short summary of the changes you made to your project 1 to the class. It should be about 2-3 minutes, and list the major changes, including the required elements in Step 2. Be ready to

present this summary on Friday Nov. 14 during class. Submit a file with just the visualization(s) as a single Powerpoint slide or PDF to Canvas by **Wednesday Nov. 12 at 5 pm**.

- **Step 4:** Turn in your final version by **Friday Nov. 21 by 5 pm**. If you are working alone, you can submit either a zip file containing a folder with all project elements to successfully knit your project or a Github repository link. If you are working with a project, you will need to submit a Github repository link ONLY.

## Additional Information

Project 2 assesses your skills at creating visualizations based on real-world data sets with an improvement in the complexity of the code. Like Project 1, it also assesses your ability to create an R Project that contains analysis and visualizations that are replicable, in other words, the figures and analysis can be replicated on other machines (most importantly, the instructor's!).

In addition the the requirements listed in "Additional Information" in Project 1, Project 2 will be required to adhere to **standard project organization** and coding **best practices** in the R Language. More information for each of these can be found in the "Best Practices" document in this folder.

If the requirements of Project 2 conflict with those of Project 1, the instructions of Project 2 should supersede those of Project 1.

## How Project 2 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:

- Presence and completeness of each of the project items outlined above. Each file and section of the RMD file should be present to receive a complete.
- Only projects that contain the elements of Step 2 above (required code elements, adherence to standard organization, and adherence to best practices) will receive a complete.
- Meaningful improvement of the visualizations in response to feedback. There needs to be evidence that at least one item of feedback from the instructor and two items of feedback from peers were incorporated into the project. Please spell this out, the instructor can't keep track of everyone's changes!
- 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.