This is a document describing a proposed workflow for survival 522 mini-projects. Describes roles, procedure, and technology we’ll use for collaborating.

# Responsibilities and Roles

There are four major groups (teams of two ideally, but can move around)

1. Coders
   * Conduct the actual analysis requested.
   * Includes all data cleaning/wrangling, regression analyses, and table / figure generation.
   * Good standardization of code is preferable by project. Enhances readability and prevents overwriting of shared function names. Coders should agree on which packages to use for an analysis. For example, try not to use two different survival analysis packages on one project, or two different data cleaning packages (data.table vs. Dplyr).
2. Writers
   * Write the actual paragraphs and text of the report.
   * Analyze and interpret results of the analyses.
3. Math / Researchers
   * Research the methods that will be used, and provide a short summary for the rest of the group to digest easily. For example, if we are supposed to use Cox PH regression for our analysis, a short summary on the assumptions, model formulation, estimation procedure, and helpful R packages. This summary will be very helpful for the coders and writers.
   * Write any math appendices and equations that are requested, methods sections of the actual report.
   * Act as support for both coders and writers with regard to the methods.
4. Formatter / Editors:
   * Transfer written paragraphs from word document to the RMarkdown.
   * Proofread and edit all final passages
   * Format tables and figures to look nice and fit within page limits.
   * Get approval from all group members before final rendering and submission

I suspect the least workload will fall to roles 3 and 4, so we can ask for help if work starts becoming imbalanced. For example, coders can ask for help making tables and figures and focus on the actual analysis.

# Procedure

1. **RESEARCH:** Researchers compile a methods summary. For the first project, this would be a summary of Weibull regression and its assumptions, model formulation
2. **CODING**: Coders (don’t need to wait for researchers) conduct analyses in R. Collaborate on coding using GitHub and try to push changes everyday you work on it so that rest of the group can see the progress of the code.
3. Consistently “knit” or render the RMarkdown into a pdf, which should automatically upload to the shared OneDrive folder. Once the code is done, let group know in Teams.
4. **WRITING**: Writers in the meantime will work on analyses-independent sections, like intro and conclusion. Once coders are done, writers can reference the rendered RMarkdown pdf with the tables, figures, and results to write the passages on analysis and interpretation. Use Word Document to collaborate and turn Track Changes on.
5. **FORMATTING**: Once writing is done, coding is done, and any appendices are done, the formatters can take the writing from the finalized word document and transfer it to the RMarkdown .rmd file. Make it look nice (no whitespace, appropriately centered figures and sizes, no typos). Once ready, send the pdf to the whole group for approval before submitting.

# Technologies

* **OneDrive Folder**: This onedrive folder has been shared with everyone and is a Git repository, which means its been tracked by Git. So if you can open any code in here within RStudio, any changes to code should be tracked by Git. Then when you’re ready to push, just follow standard procedure and it should show up in the public Git remote. Here’s a good page for getting started <https://guides.github.com/introduction/git-handbook/> . Link to the remote repository we’ll be using is here. <https://github.com/tXiao95/survival-522>
* **GitHub**: Everyone who works on code should make a GitHub account so that collaboration is easy. For those who are not experienced with Git, we can have a meeting with the coders and anyone else that wants to go through how to use it and procedures.
* **RMarkdown:** uses markdown and latex like syntax to easily integrate code and writing. Will be final format of our submission report.