PHC 6068 Biostatistical Computing
FALL 2019 Final Project

Overall description: The goal of the final project is to develop an R package that will be useful to other statisticians and R users. Students can either form teams (at most 3 people for each team) by themselves or work individually. Readable R documentations are necessary. Students will be encouraged to use higher level knowledges from the class to the R package. (E.g., Rcpp, GitHub). The final R package and a short report of how to use the R package is due on Mon Dec 9 (11:59 pm). One team only need to submit one copy of report and R package.

R package

Basic requirement

1. Topic: any functions or procedures you feel useful or interesting. The project topics in previous years can be found here https://caleb-huo.github.io/teaching/finalProject/finalProject_all.html.
2. Package documentation should be complete and easy to understand.
    a. Function descriptions
    b. Argument and return value
    c. Return values
    d. A toy example
3. Wrap up as a package and later demonstrate how to use it in the final report.

Higher-level feature

1. Utilize Rcpp
2. GitHub
3. Or other features you feel useful or advanced (e.g., R Shiny app).
4. Very heavy workload

Final Presentation

There is no final presentation.

Final Report and R package

Both final report and R package are due on Mon Dec 9 (11:59 pm).

R package

Should be tested such that the instructor can directly install the package.

Final Report

Suggested to use Rmd (like HW) to write your final report since Rmd can integrate your code and result. In the final report, you need to demonstrate how to use **your R package** to generate your results. There is no page requirement for your final report, below are suggested format.

- Introduction: background info, literature survey, problem definition and motivation, etc, if any.
- Methods
    - Details of your proposed method

    - Describe your functions in the package and your workflow (similar to your package documentation but a bit more detailed description)
- Results
    - Describe the datasets (or simulation) you used to test your package, if there is any.
    - Show some outputs of your package. Please be organized and make sure your output is readable.
- Conclusion: conclusion, discussion and future work, if any.

Grading

All members in one team will have the same grade for the final project. The grading breakdown for the final report and R package is as follows:

R package (50%): the instructors/TA must be able to install your package via tar.gz or GitHub

- 15% for documentation of your R package.
- 15% some higher-level features.
- 10% Vignettes: a nice and overall introduction of your package.
- 10% pass CRAN checks without errors or warnings.

Report (50%):

- 5% for your name, other team member names if you are working on a group.
- 5% for a title of your report.
- 10% for introduction or motivation so we understand what is the problem.
- 10% for quality of writing (clarity, organization, flow, etc.)
- 10% for correctness and completeness of your proposed method.
- 10% for correctness and completeness of results by using the R package.

Again, the final R package and a short report of how to use the R package is due on Mon Dec 9 (11:59 pm). One team only need to submit one copy of report (e.g., html file) and R package (e.g., tar.gz file). They should be submitted electronically via canvas https://elearning.ufl.edu.