Code as Manuscript: Analytical Practices for Data Efficiency and Reproducibility

Course Syllabus

Instructors
Luke Johnston, MSc, PhD (C)
luke.johnston@mail.utoronto.ca

Daiva Nielsen, PhD
daiva.nielsen@mail.utoronto.ca

Class day, time and location: *2 hours per class is requested*

Course description:
The purpose of this course is to introduce students to the concept of research reproducibility and present methods of conducting statistical analyses in a more efficient, less error prone way. The ultimate goal of the course is to demonstrate how to reduce the number of steps required to produce a manuscript from the data analysis stage. Students will be expected to bring their own fully-charged laptop to each class with the appropriate software installed (Git, a text editor and a statistical analysis software). Emphasis will be placed on hands-on practice activities integrated into each weekly class.

Course objectives:
At the end of the course, students will be able to:

- Track their research files with Git to control multiple versions of files.
- Operate GitHub to maintain a back-up of files and have an outlet for efficient collaboration.
- Create macros to condense statistical code writing.
- Use ODS to output results into a more useable form.

Course Schedule

Week 1	Introduction to version control with Git and GitHub for collaboration
Week 2	Statistical assumptions
Week 3	Macro-coding
Week 4	Output delivery system
Week 5	Data visualization
Week 6	Review and practice

Course Resource List

PLOS article (best practices in scientific computing)
Other articles on reproducibility
GitHub URLs and cheat sheets