Lesson: Reproducibility in R

BiGG - Summer 2021

What is Reproducibility?

The ability for you (or someone else) to understand your code and replicate the results and output produced by it.

Basic Idea

The idea of reproducibility can be broken down into a couple components:

- Keeping track of your changes (version control)
- Making sure others can understand your code (readability)
- Ensuring others can replicate the results your code produces (reusability)

Version Control

We want to keep track of incremental changes to our code so that:

A. We don't lose our progress
B. We have previous iterations of our code in case we make a mistake
C. We have a history of previous tests/models we generated in case we need to revert back to a prior iteration

We will be using GitHub (https://github.com/) for version control over the Summer.

Readability

Our code needs to be easy to follow in order for us to make use of it again at a later date (or explain it to others). The main things that help with this are:

- Comments
- Good variable names
- Code style
- Program structure (if-statements, loops)

Reusability

We also want others to be able to replicate our results and borrow from our code if necessary (open source). This is best accomplished through:

- Modular programming (breaking code into smaller chunks)
- Functions
- Preventing duplicative code/hard-coding

Agenda

Go to
https://github.com/ababjac/BiGG-test/tree/main/reproducability_R_lesson
and download the materials for today.

- Configure version control for GitHub/R (using GitHub-R_tutorial.pdf)
- Work through readability notes (using Readability.Rmd)
- Work through reusability notes (using Reusability.Rmd)