Git & Github Basics

Justin Post

Workflow

Important considerations for data analysis workflow:

- **Reproducibility**

  - be able to do everything from start to finish, get same results

- **Version control**

  - tracking changes to files so you don't lose anything

- **Collaboration**

  - sharing files amongst ppl, dropbox, google drive

What are git and github?

- **Git** is a version control software
- **Github** is an online hosting service for Git-based projects

What are git and github?

- **Git** is a version control software
- **Github** is an online hosting service for Git-based projects

Git Tracking

- You associate git with a folder (repo)
- Git keeps track of all files in the folder (repo)
- If you want to keep changes you've made, you **commit** and **push** the changes to the folder (repo)

Github

- Github allows you to have a remote file repository (folder) tracked by git
  - Let's create a repository on github.com
  - **Add** some files and **commit** to the changes
  - Modify some files on github
  - Investigate the version control!

Local vs Remote Work

Mostly you'll want to work on your local computer. Install `git` on your computer!

Local vs Remote Work

Mostly you'll want to work on your local computer. Install `git` on your computer!

Workflow

1. (Initially) **clone** the repo locally. (Later) **pull** to get most recent versions of files
2. Work and make changes
3. **add** and **commit** to changes you like
4. **push** changes to remote repo (on github)

Let's clone our repo and work on it locally!

Git & RStudio

Git and RStudio work great together!

- Works through **R Projects**
  - associates a folder + its files with a .Rproj file
- Start a new project from git repo
  - Update with command line or git menu!
    - working dir.
    - environment
    - RHistory