

# GUI Git

Fernando Hoces de la Guardia  
BITSS

-

Slides at  
[https://github.com/BITSS/RT2\\_LOSANGELES\\_2018](https://github.com/BITSS/RT2_LOSANGELES_2018)

Los Angeles, September 2018

## Workflow & General Coding Suggestions

# The Claerbout Principle

*An article about computational science in a scientific publication is not the scholarship itself, it is merely advertising of the scholarship. The actual scholarship is the complete software development environment and the complete set of instructions which generated the figures.*

Buckheit & Donoho, 1995

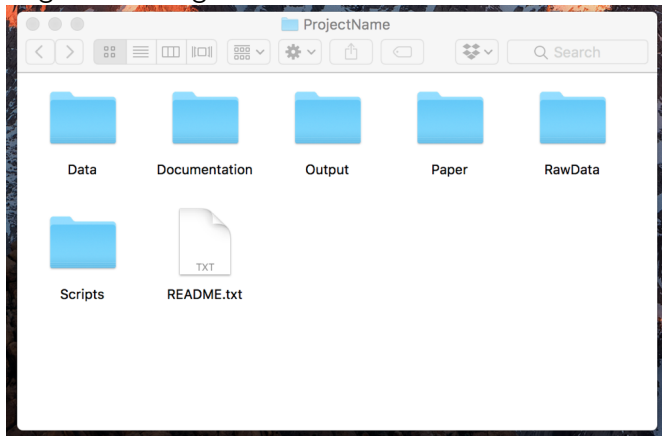
# Organizing Principles

Christensen, Miguel & Freese (2018)

- 1 - Use code (scripts), don't work by hand (Excel/spreadsheet, GUIs).
- 2 - Consider not saving statistical output, and just saving the code and raw data that generates it.
- 3 - Reproducibility—on your own machine across multiple runs, across machines, across researchers.

# File Management & Coding Suggestions

Begin with a logical file structure



## Version Control

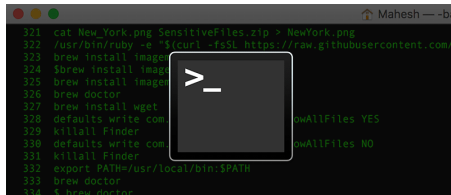
# What is a GUI and why the bad reputaion

## Graphical **U**ser **I**nterface

- ▶ For most of us (non-programmers): *GUI = Software*
- ▶ GUIs are behind the popularization of personal computers.
- ▶ Unfortunatley GUIs are pretty bad at keeping a record of actions taken (bad for reproducibility)

# What is not a GUI?

- ▶ Any software that is run in the command line (aka terminal, shell, bash, etc)

A screenshot of a macOS terminal window. The window title bar shows three colored window control buttons (red, yellow, green) on the left and the text 'Mahesh — -ba' on the right. The terminal content shows a series of commands and their outputs, including file operations, Homebrew package management, and system configuration. A white cursor icon, consisting of a greater-than sign followed by an underscore ('>\_'), is overlaid on the terminal text, specifically pointing to the 'brew doctor' command on line 334.

```
321 cat New_York.png SensitiveFiles.zip > NewYork.png
322 /usr/bin/ruby -e "$(curl -fsSL https://raw.githubusercontent.com/
323 brew install image
324 $brew install image
325 brew install image
326 brew doctor
327 brew install wget
328 defaults write com. owAllFiles YES
329 killall Finder
330 defaults write com. owAllFiles NO
331 killall Finder
332 export PATH=/usr/local/bin:$PATH
333 brew doctor
334 $ brew doctor
```

- ▶ Git was designed to run in the command line



# What is Git 1/2

- ▶ Git is a software designed to track the **entire** history of the code of a project.
- ▶ Designed originally for software development, it has gained important traction in the research community.
- ▶ Main appeal: facilitates full reproducibility and collaboration.
- ▶ Git is mainly meant to work as a non-GUI (in the command line) software. **However:** most of the key features can be used through a GUI.

## What is Git 2/2

- ▶ By code git understands any type of plain text file (`myfile.R`, `myfile.do`, `.tex/.md/.txt/.csv/.etc`).
- ▶ This types of files can be understood as “human readable” as machine and human see the same fie.
- ▶ Files that are “non-human readable” are called binary files (`myfile.docx`, `myfile.xlsx`, `.pdf/.exe/.dta/.etc`).
- ▶ Git can detect changes in binary files, but it cannot show those changes.

# What is Github

- ▶ Github is a company that provides two services (that we care of):
  - ▶ A web hosting service for all our files track with git (public free/private \$ or free if academic).
  - ▶ A GUI software (Desktop App) that provides user friendly access to git.
- ▶ Others hosting ss include: Bitbucket, GitLab, Gitkraken, etc.
- ▶ Other GUIs include: SourceTree, Gitkraken, Atom, RStudio.

# The Primary Goal of Version Control (for us)

**The Goal:** keep track of any potentially meaningful modification to your code.

# The Primary Goal of Version Control (for us)

**The Goal:** keep track of any potentially meaningful modification to your code.

**Secondary Goal:** Learn how to collaborate with others using Github.

# The Primary Goal of Version Control (for us)

**The Goal:** keep track of any potentially meaningful modification to your code.

**Secondary Goal:** Learn how to collaborate with others using Github.

**Bonus track:** get you excited about using open source statistical software (R, Python, Julia, etc)

## Strategy 1:

- 1 - Agree on a naming convention with you co-authors (eg: YYYYMMDDfilename\_INITIALS).
- 2 - Begin working from the last saved version (eg: 20180325demo\_FH.do).
- 3 - At the end of the day, save on a new version (eg: 20180327demo\_FH.do).

**Pros:** Easy adoption.

**Cons:** Error prone, hard to document, lots of files for each document.

## Strategy 2:

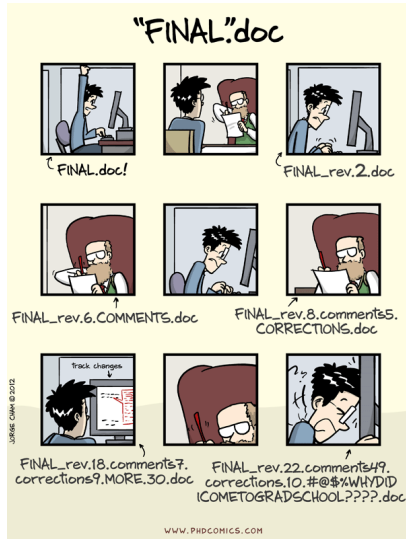
- 1 - Name your file `filename` (ideally `01_filename`)
- 2 - Take a snapshot of your work every time you complete relevant change (day, hour or minutes).
- 3 - Update your entire working folder to the cloud.

**Pros:** Error proof, seamless documentation, one file per document, track differences across all versions, meant to work with the cloud.

**Cons:** Harder adoption.



We want to avoid this situation:



## Other reasons to use git

- ▶ To access a whole new world of knowledge!
- ▶ Great tool for collaboration.
- ▶ Easier to test all sorts of ideas/models.

# Managing expectations



## Three Demos:

- 1 - **Simple but instructive.**
- 2 - Repeat but with a slightly more fun example. Collaborate.
- 3 - Repeat with a real-life example.

## Demo #1: We Start in the Cloud

- 1- Create github.com account and sign in.
- 2- Let's look at some **repos**.
- 3- First way to access content: download.
- 4- What if you want to have your own copy of the repo? **Fork** it!.
- 5- Now create your own repo. Initiate readme and make some edits.

## Demo #1: We move to our local computer

6- Clone the it. Explore the files and location.

7- Create new files, edit. And commit. Edit again, and commit again.

8- Push. Edit on github.com, and pull.

9- Simulate conflict (between local and remote) and start from a fresh copy!

## Three Demos:

1 - Simple but instructive.

**Review:** def repo, github.com, download, clone, destination folder, fork, create repo, commit, push, pull, delete & restart, search repo, download old version.

2 - **Repeat but with a slightly more fun example. Collaborate.**

3 - Repeat with a real-life example.

## Demo #2: Branches and collaboration

- 1- Create a branch from previous repo.
- 2- Add new content, commit a few times and merge.
- 3- Go back to main branch (master), observe file, merge.
- 4- Repeat 1-3 but now replace instead of adding content.



## Demo #2: Branches and collaboration

- 5- Fork repo test2, and clone it into your machine.
- 6- Edit fields of name, and birth date.
- 7- Save, commit and push.
- 8- Create your first pull request.
- 9- Let's see if I can manage all those PR very quickly.
- 10- All in the same project!

## Three Demos:

1 - Simple but instructive.

Review: def repo, github.com, download, clone, destination folder, fork, create repo, commit, push, pull, delete & restart, search repo, download old version.

2 - Repeat but with a slightly more fun example. Collaborate.

**Review:** All of the above, plus: branch, merge, resolve conflicts, collaborate: same proj, fork model+PR .

3 - **Repeat with a real-life example.**

Review: All of the above, plus: how does a real-life example looks like.

Demo #3: Look inside a half-way project (and collaborate!)

## Three Demos:

1 - Simple but instructive.

Review: def repo, github.com, download, clone, destination folder, fork, create repo, commit, push, pull, delete & restart, search repo, download old version.

2 - Repeat but with a slightly more fun example. Collaborate.

Review: All of the above, plus: branch, merge, resolve conflicts, collaborate: same proj, fork model+PR .

3 - **Repeat with a real-life example.**

**Review:** All of the above, plus: how does a real-life example looks like.

# Now go and explore!

Some good habits:

- Commit often ( $<1\text{hr}$ )
- Always pull before you start a new session of work. Also good to pull before pushing.
- Think of your remote as the most important set of files. Get used to deleting things in your local machine.

## Want to learn more:

- ▶ Great 20 min intro to Git by Alice Bartlett
- ▶ Great 2hr tutorial to Github by Jenny Bryan (git ninja)
- ▶ Jenny Bryan's Happy Git; Documentation from Matthew Gentzkow Jesse Shapiro; Karthik Ram's paper on Git for Research