

# Introduction to git and GitHub

Data Science option of BIO00058M Data Analysis.

Emma Rand

University of York, UK

# Aims & Learning Outcomes

The aim of these slides is to introduce you to git and GitHub

At the end of this presentation the successful student will be able to:

1. explain the purpose of using git and GitHub and understand the relationship between them.
2. recognise common terminology in git and GitHub use
3. reassure themselves that git is not assessed 😊

# What is git?



# git

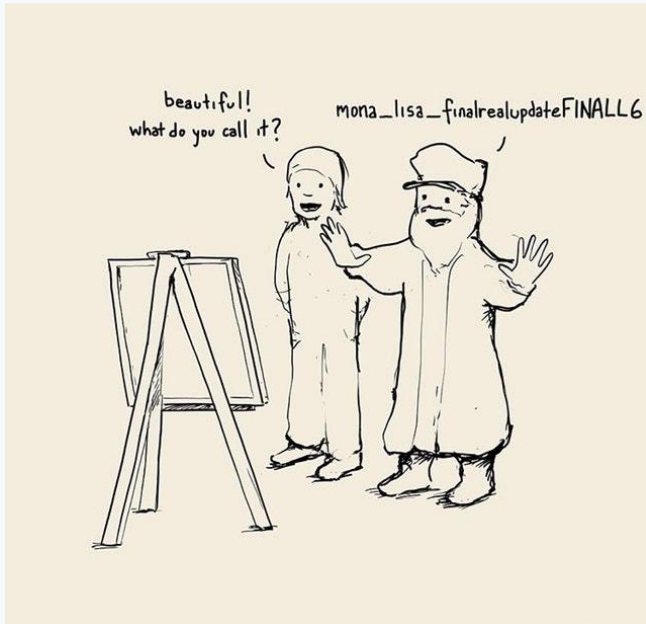
Git is a Version Control System (VCS).

A VCS monitors and records the changes made to a set of files. It allows you to:

- revert selected edits, files or entire projects back to a previous state
- collaborate on coding projects with any number of people
- compare versions, integrate and merge different versions
- resolve conflicts between versions
- contribute to open software/analysis projects in small (and big!) ways
- allows communication and collaboration to be embedded in the workflow

# You already use version control!

Or try to...



Think of it as like MS Word's Track changes or Googledoc's editing history but better!

# What is GitHub?



GitHub is a website that hosts git projects. It can also be used as a webserver for your project. Most of my git projects include a website.

Think of it as DropBox or googledrive but better.

Bitbucket and GitLab are alternative git hosting services.

# Git terminology

A set of files managed by git is called a **repository** or commonly **repo**

A repo would usually be all the files associated with a particular project.

You can look at my repos on GitHub here:

<https://github.com/3mmaRand?tab=repositories>

To **clone** a repo is to make a local copy of those files and their history.

# Git terminology

Instead of just saving files as you go about your work, you make a **commit** every so often. A commit is a snapshot of the entire project. That snapshot of the repo is stored locally.

Every so often you **push** your commits to Github.

If you are collaborating you would probably first **pull** from Github to incorporate any changes made to the GitHub repo in to your local repo.

Git will automatically **merge** versions and alert you if there are any **conflicts** to **resolve**.

# Git and GitHub terminology

To fork a repo is also to make a copy but the link between the original and the forked copy remains.

You can experiment with your forked copy without affecting the original.

If you want your changes to be incorporated in to the original you use a **Pull Request**



# In this module

The use of git and GitHub in this module is a "bare bones" introduction that will give you a workflow for using git and GitHub through RStudio.

Becoming fully proficient in git is a long, and often painful, process and its use in this module is designed to help you make the first steps and aid collaboration. It is not assessed.

The workflow taught is limited and confined but works. You are not expected to use git/GitHub for the assessment - please do not become stressed about git.

Think of it as a chance to take some risk free steps to learning an extremely valuable skill.

# Reading

A more thorough coverage of git and GitHub through RStudio is [Happy git and GitHub for the useR](#) (Bryan and Hester, a). I recommend reading Chapter 1 Why Git? Why GitHub.

## RStudio's Version Control with git and SVN

You don't have to use RStudio to use git and GitHub. [Git Handbook](#) from GitHub Guides

# References

Bryan, J. and J. Hester *Happy Git and GitHub for the useR.*

. Accessed: 2019-11-22.

Slides made with with xaringan (Xie, 2019) and xaringanExtra (Aden-Buie, 2020)

Emma Rand

[emma.rand@york.ac.uk](mailto:emma.rand@york.ac.uk)

Twitter: [@er13\\_r](https://twitter.com/er13_r)

GitHub: [3mmaRand](https://github.com/3mmaRand)

blog: <https://buzzrbeeline.blog/>



Data Science strand of BIO00058M by Emma Rand is licensed under a [Creative Commons Attribution-NonCommercial-ShareAlike 4.0 International License](https://creativecommons.org/licenses/by-nc-sa/4.0/).