Git and Github: a primer

Alfredo Gimenez Zapiola

POLITECNICO
MILANO 1863

Nonparametric Statistics

AY 2024-2025

What is Git?

- Git is a Version Control System (VCS) designed to make it easier to have multiple versions of a code base, sometimes across multiple developers or teams
- It allows you to see changes you make to your code and easily revert them
- It is not github!

What is Github then?

- Github.com is a website that hosts git repositories on a remote server
- Hosting repositories on Github facilitates the sharing of codebases among teams by providing a GUI to easily fork or clone repos to a local machine
- By pushing your repositories to Github, you will pretty much automatically create your own developer portfolio as well!

Confirm that you have git

- Open your terminal and run "git"
- If you see a "command not recognized" error, you probably have not installed git yet
- Download it, according to your OS (Linux, Windows, MacOS), here.
- Good reference (mostly for R users!) can be found here
- Plenty of other free resources online

What do we need to know for the course?

Essentially nothing: meaning that you can simply treat the repository we will be using for the labs as a normal folder, and nothing about git will be asked during the exam. Yet:

- Using github will potentially help you during the NPS projects
- Most importantly, it will become fundamental if you will end up working in a data science/ programming related industry!

In practice

- Repository associated with the lab notebooks:

- Run the following commands in your terminal (in a dedicated path)
  - *git clone* *https://github.com/alfredo-g-zapiola/Nonparametric_Statistics_24_25.git* to clone it in your local folder
  - *git pull* to keep it updated.
- That's it!
- More to know if you want to be a proactive user (e.g., *git commit*, *git push* etc..), check this guide