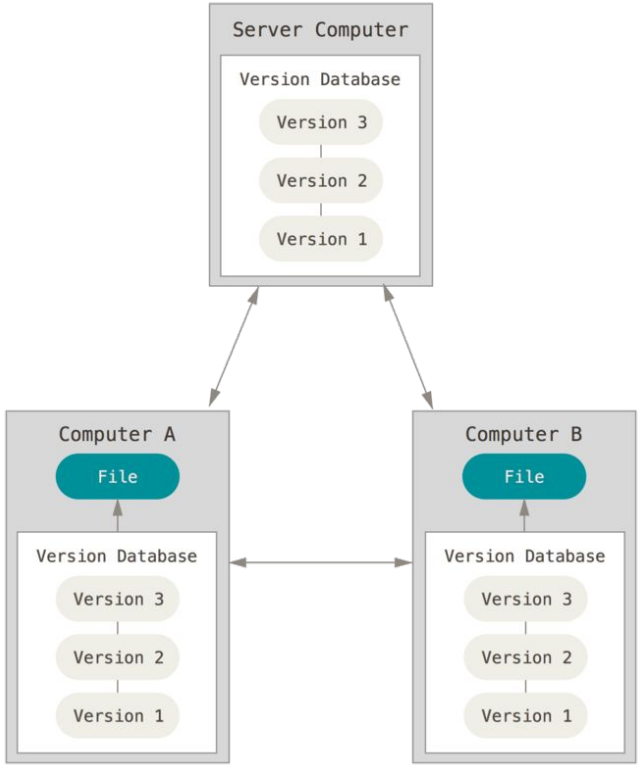
DATA SCIENCE
ACADEMY

Master in Data Science - Git

Version Control | **1**

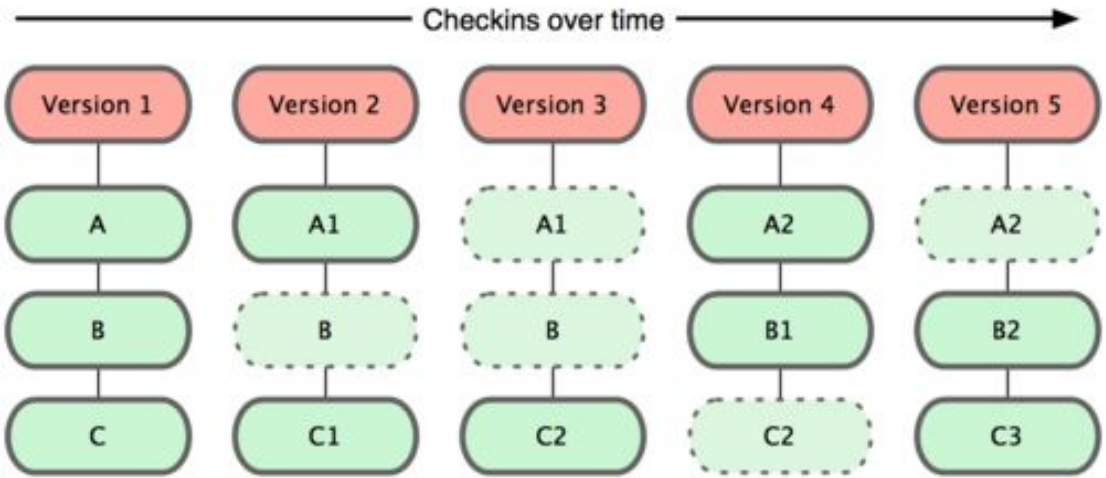
Basic Intro to Git

- Version Control
  a. A system controlling the different version of a file or a series of files.
  b. It can be *local* or *remote* (online or on a local network).
  c. A remote version control system can be *centralised* or *distributed*.
  d. Initial goals:
     i. Speed
     ii. Support for non-linear development (thousands of parallel branches)
     iii. Fully distributed
     iv. Able to handle large projects like Linux efficiently



Basic Intro to Git

- Git
  a. Git considers different versions as *snapshots*.
  b. Everytime we modify a file, the *git status* of the project changes.
  c. To be efficient, if a file does not change, git creates a link to the previous snapshot of the file.



How to use Git

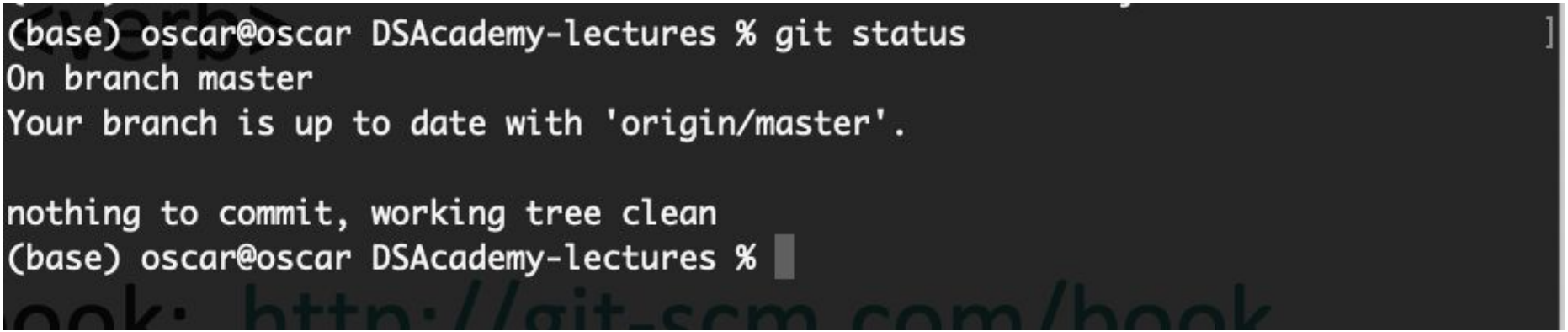
- Command line

  $ git help <verb>

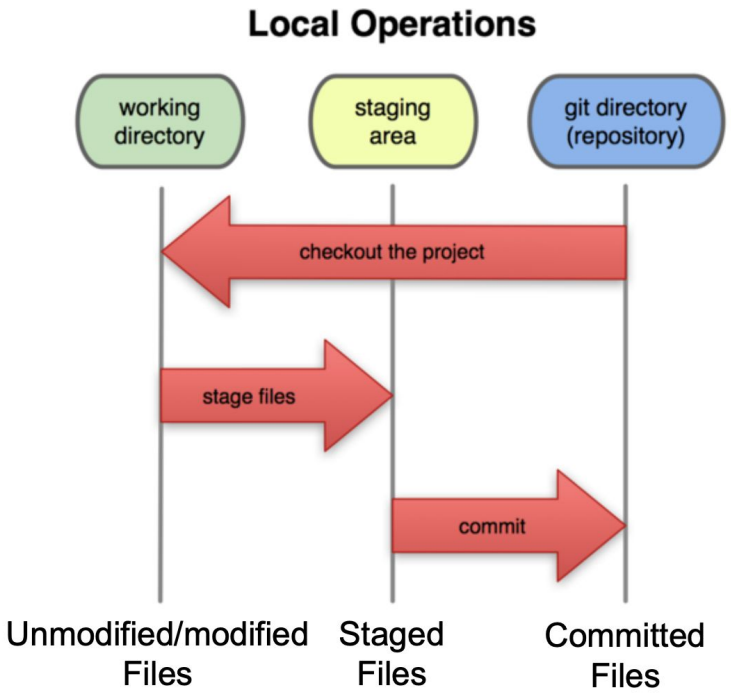
  $ git <verb> --help

  $ man git-<verb>

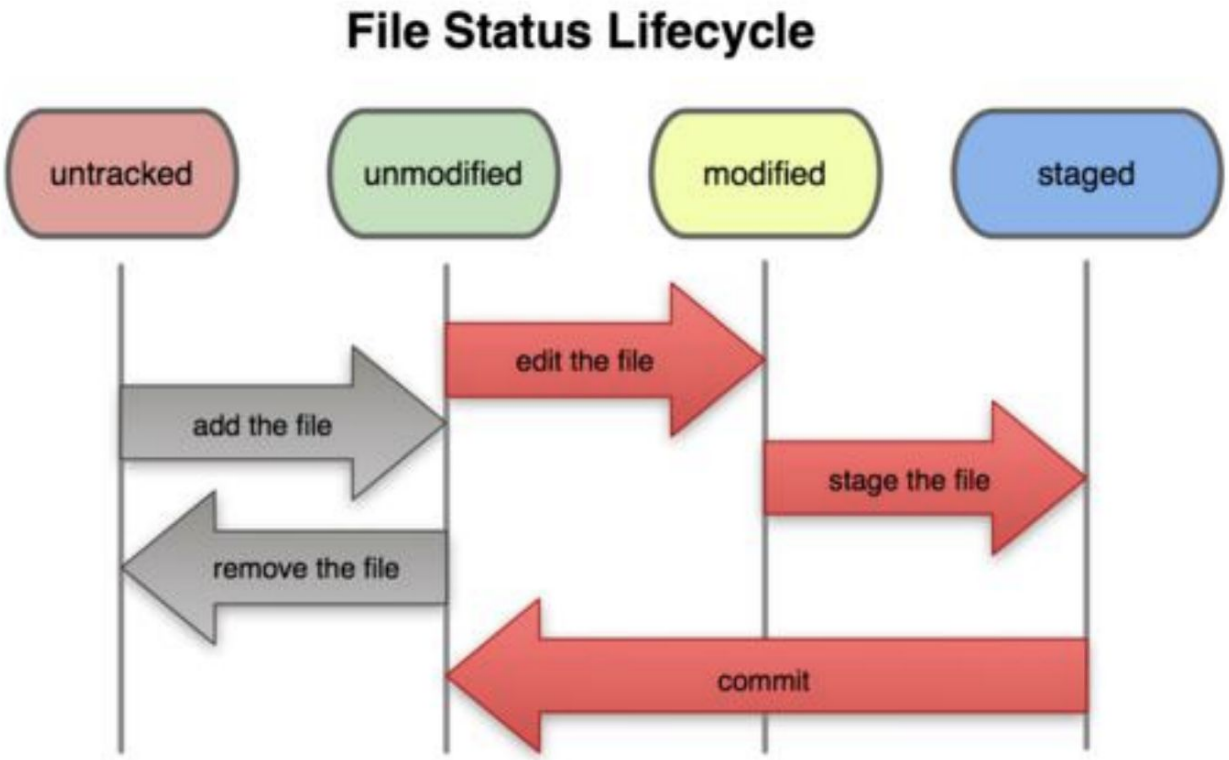
- <verb> = config, add, commit, etc.



A *local* git project has three areas



Git file lifecycle



Basic workflow

- Basic git workflow
  a. **Modify** files in your working directory
  b. **Stage** files, adding snapshots of them to your staging area.
  c. Do a **commit**, which takes the files as they are in the staging area and stores that snapshot permanently on your git directory.

Notes:

- If a particular version of a file is in the git directory, it's considered committed.
- If it's modified but has been added to the staging area, it is staged.
- If it was changed since it was checked out but has not been staged, it is modified.

GitHub | **2**

Hands-on example of working with git and github

So what is github?

- GitHub.com is a site for online storage of Git repositories.
- Many open source projects use it, such as the Linux kernel.
- You can get free space for open source projects or you can pay for private projects.

Question: Do I have to use github to use Git?

Answer: No!

- you can use Git completely locally for your own purposes, or
- you or someone else could set up a server to share files, or
- you could share a repo with users on the same file system.

git commands

command	description
`git clone` ***`url [dir]`***	copy a git repository so you can add to it
`git add` ***`files`***	adds file contents to the staging area
`git commit`	records a snapshot of the staging area
`git status`	view the status of your files in the working directory and staging area
`git diff`	shows diff of what is staged and what is modified but unstaged
`git help` ***`[command]`***	get help info about a particular command
`git pull`	fetch from a remote repo and try to merge into the current branch
`git push`	push your new branches and data to a remote repository
others: `init, reset, branch, checkout, merge, log, tag`	

Have Fun!

Contacts:
oscar.defelice@datascienceacademy.it

marica.acconcia@datascienceacademy.it
carlotta.reggioli@datascienceacademy.it