

Visualization for Data Science

DS-4630 / CS-5630 / CS-6630

Managing Projects with Git



Git is a distributed **version-control** system

- Terminology: In git-speak, a “version” is called a “commit.”
- Git keeps track of the history of your commits, so you can go back and look at earlier versions or just give up on the current version and go back to some earlier version.
- Can be used to implement a variety of software configuration management models and workflows

Git is a **distributed** version-control system

- You keep your files in a *repository* on your local machine.
- You synchronize your repository with a remote repository on a server (in our case, GitHub).
 - You protect your code from system crashes by synchronizing with the server.
 - If you move from one machine to another, you can pick up the changes by synchronizing with the server.
 - If you work on a team, other people's uploads can be synchronized using the server.

Git Tools

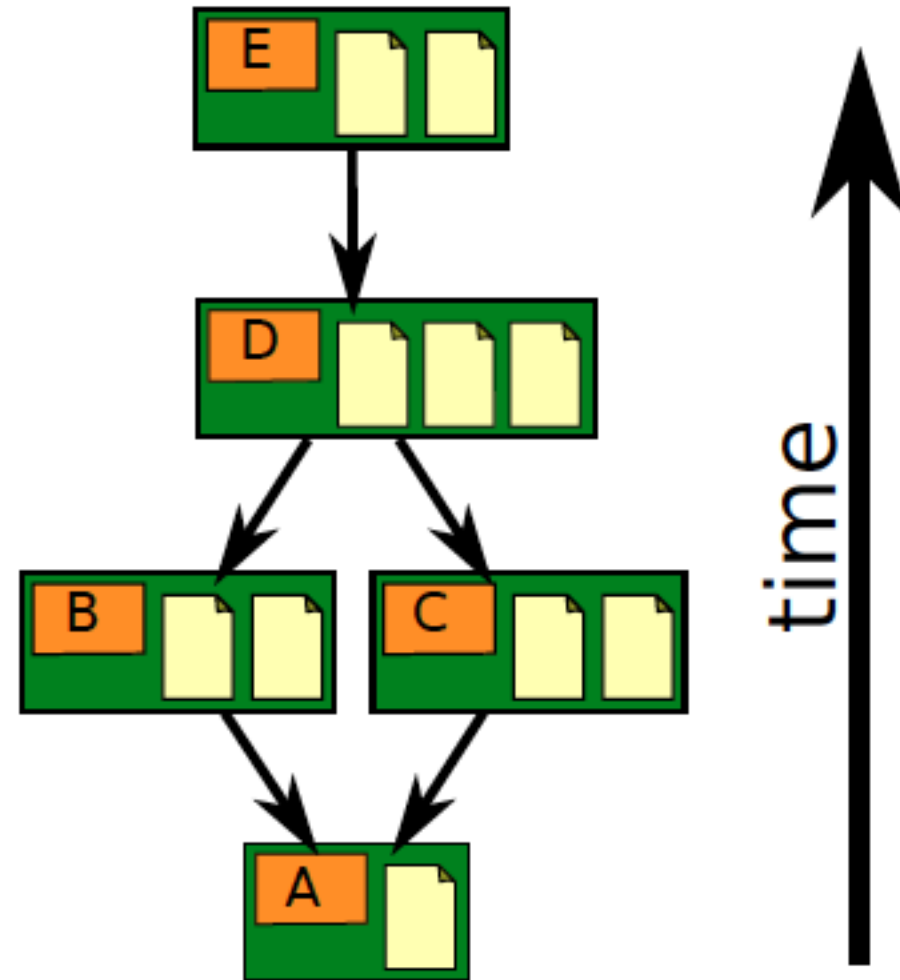
- A collection of many tools
 - Very flexible
- You can do anything the model permits
 - Including shooting yourself in the foot
- Need to understand the underlying model

Groups of **Git** commands

- Setup and branch management
 - **init**, **checkout**, **branch**, **clone**
- Modify
 - **add**, **delete**, **rename**, **commit**
- Get information
 - **status**, **diff**, **log**
- Create reference points
 - **tag**, **branch**
- Synchronization with remote
 - **push**, **pull**, **fetch**, **sync**

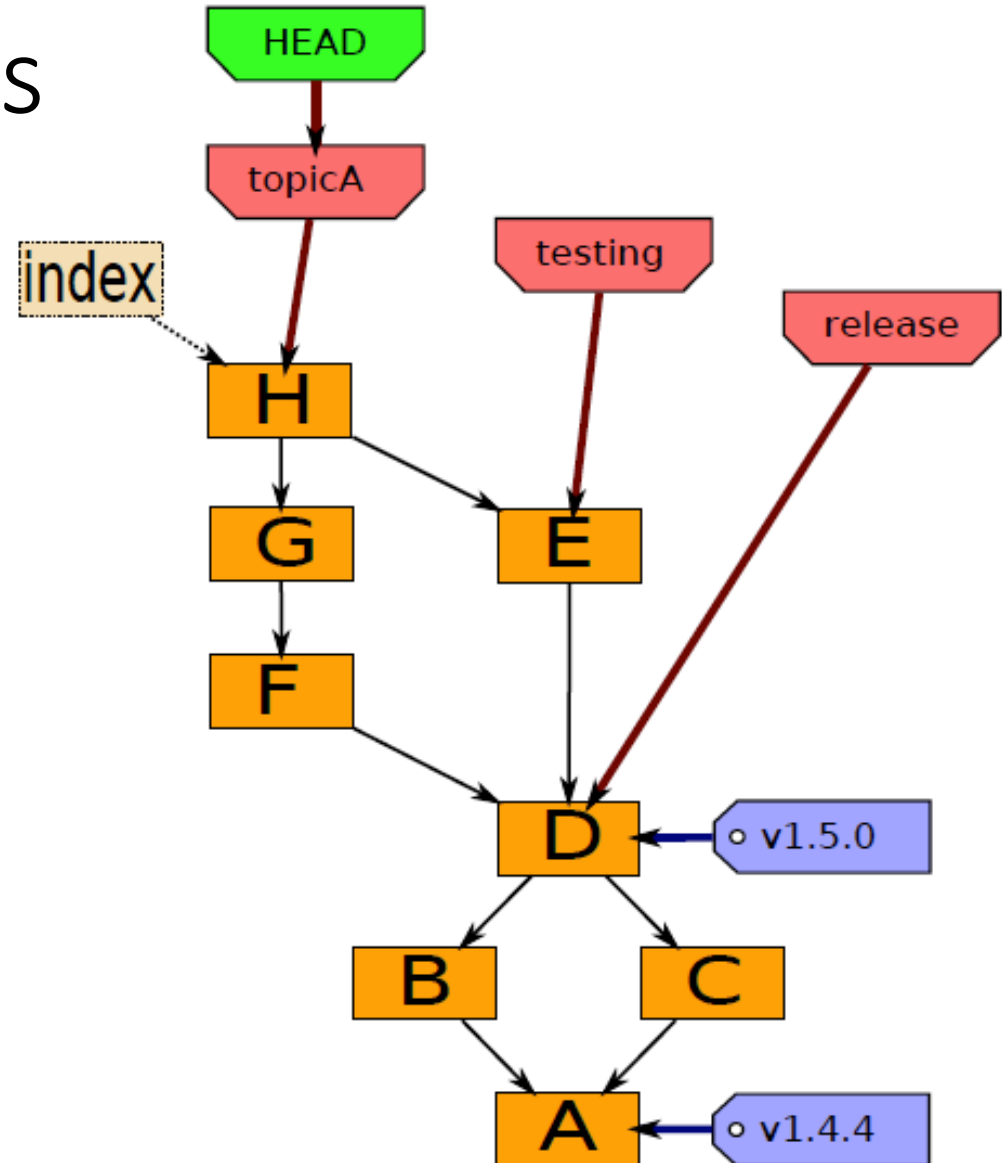
Repository Contains

- files & directories
- commits
- ancestry relationships

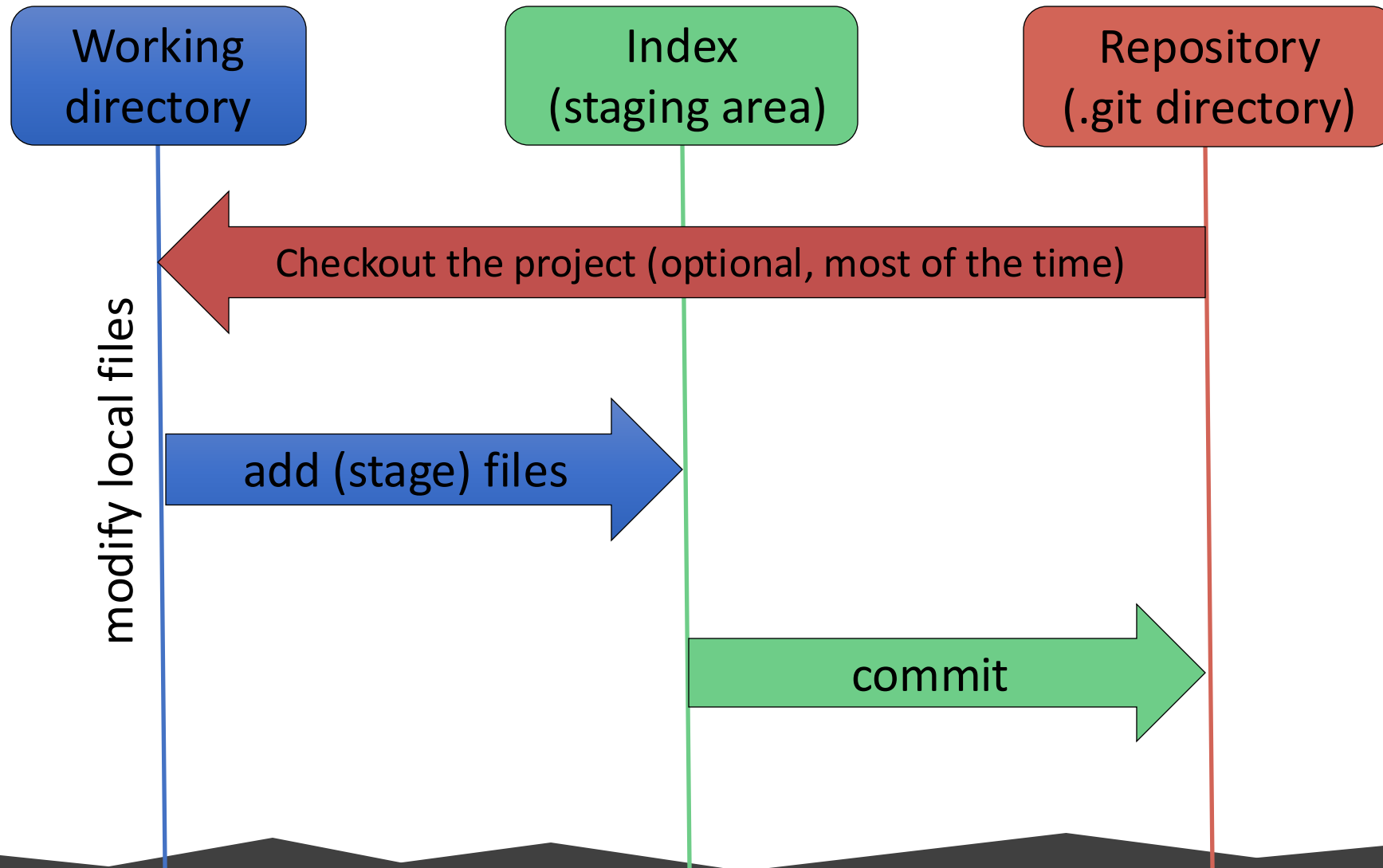


Ancestry graph features

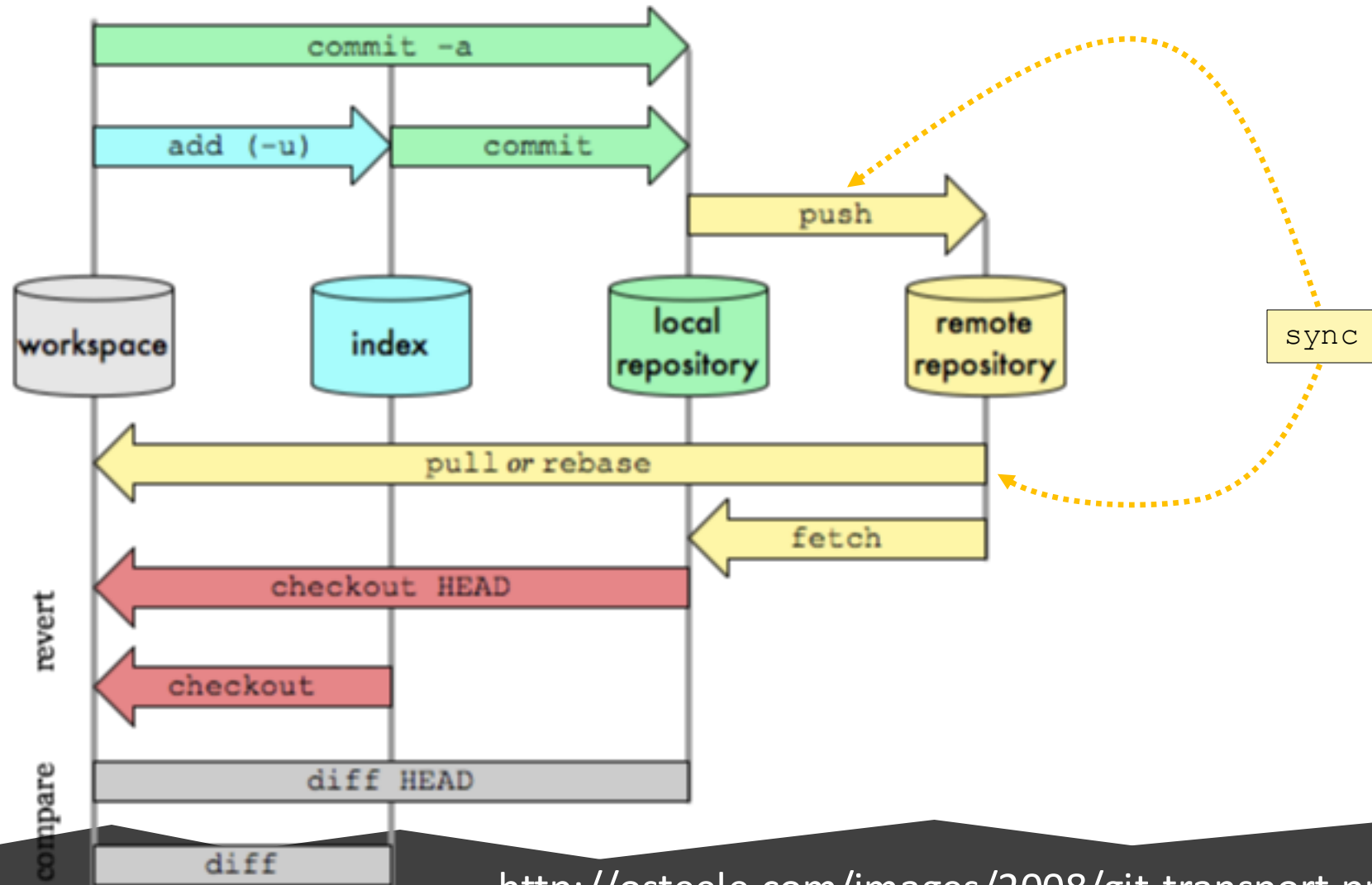
- form a directed acyclic graph (DAG)
- **Commits**
 - Snapshots of file status
- **Tags**
 - identify versions of interest
 - including “releases”
- **Branches**
 - divergent path for source code modification
- **HEAD**
 - is current checkout
 - usually points to a branch
- **Index**
 - “staging area”
 - what is to be committed



Local Operations



Git transport commands



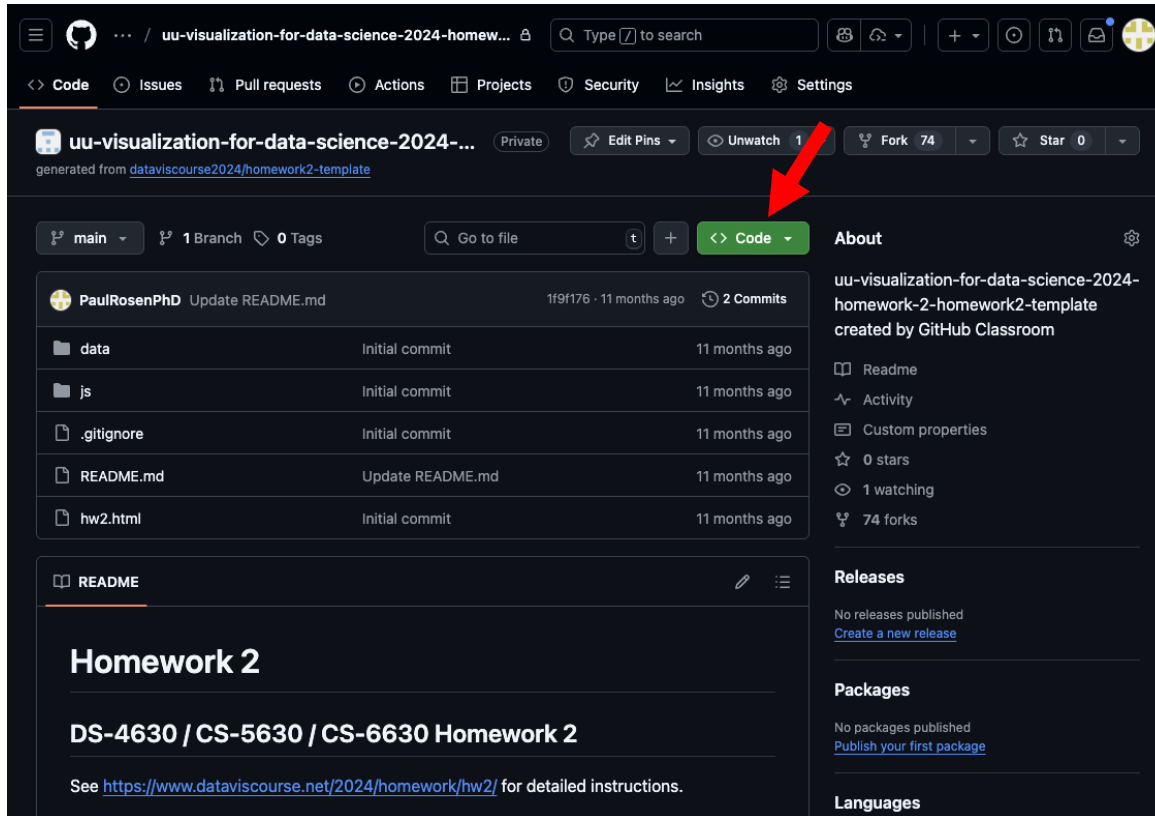
Git Software

- Windows
 - Git command line tools – <https://git-scm.com/download/win>
 - Git GUI – <https://tortoisegit.org/> (also requires download of command line tools)
- MAC
 - Install xcode and the command-line tools
 - <https://developer.apple.com/xcode/>
 - <http://railsapps.github.io/xcode-command-line-tools.html>
- Linux
 - git should already be installed. If not, use the appropriate package manager (e.g. apt or yum) to install it.

Getting Started

- Create a GitHub account, if you don't already have one (<https://github.com/>)
 - GitHub Education account is optional (https://education.github.com/discount_requests/application)
- Checkout the assignments for link to setup your repositories
- Once the repository is created (this can take a few minutes) determine the remote path and pick a local directory for code.

Finding Remote Path



This screenshot shows the GitHub repository page for 'uu-visualization-for-data-science-2024-homework2-template'. A red arrow points to the 'Code' button in the top right corner of the repository header.

Repository: uu-visualization-for-data-science-2024-homework2-template
generated from [dataviscourse2024/homework2-template](#)

Buttons: Edit Pins, Unwatch 1, Fork 74, Star 0

Branches: main (1 Branch), Tags: 0 Tags

Files:

- data (Initial commit, 11 months ago)
- js (Initial commit, 11 months ago)
- .gitignore (Initial commit, 11 months ago)
- README.md (Update README.md, 11 months ago)
- hw2.html (Initial commit, 11 months ago)

About: uu-visualization-for-data-science-2024-homework2-homework2-template created by GitHub Classroom

Releases: No releases published. [Create a new release](#)

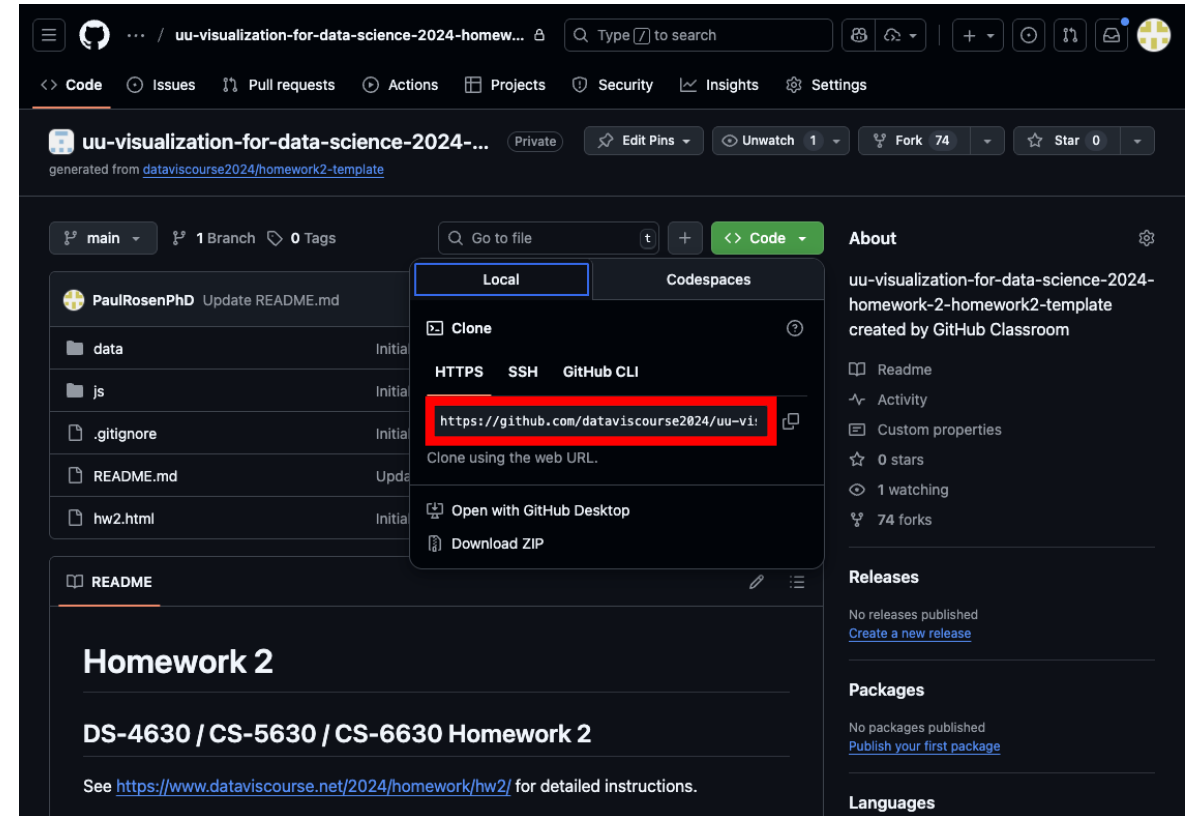
Packages: No packages published. [Publish your first package](#)

Languages

Homework 2

DS-4630 / CS-5630 / CS-6630 Homework 2

See <https://www.dataviscourse.net/2024/homework/hw2/> for detailed instructions.



This screenshot shows the same GitHub repository page, but with the 'Code' dropdown menu open. The 'Clone' option is selected, and the HTTPS URL is highlighted with a red box.

Repository: uu-visualization-for-data-science-2024-homework2-template
generated from [dataviscourse2024/homework2-template](#)

Buttons: Edit Pins, Unwatch 1, Fork 74, Star 0

Branches: main (1 Branch), Tags: 0 Tags

Files:

- data (Initial commit, 11 months ago)
- js (Initial commit, 11 months ago)
- .gitignore (Initial commit, 11 months ago)
- README.md (Update README.md, 11 months ago)
- hw2.html (Initial commit, 11 months ago)

About: uu-visualization-for-data-science-2024-homework2-homework2-template created by GitHub Classroom

Releases: No releases published. [Create a new release](#)

Packages: No packages published. [Publish your first package](#)

Languages

Homework 2

DS-4630 / CS-5630 / CS-6630 Homework 2

See <https://www.dataviscourse.net/2024/homework/hw2/> for detailed instructions.

Clone options:

- Local
- Codespaces
- Clone (selected)
- HTTPS (selected)
- SSH
- GitHub CLI

HTTPS URL: <https://github.com/dataviscourse2024/uu-vi>

Clone using the web URL.

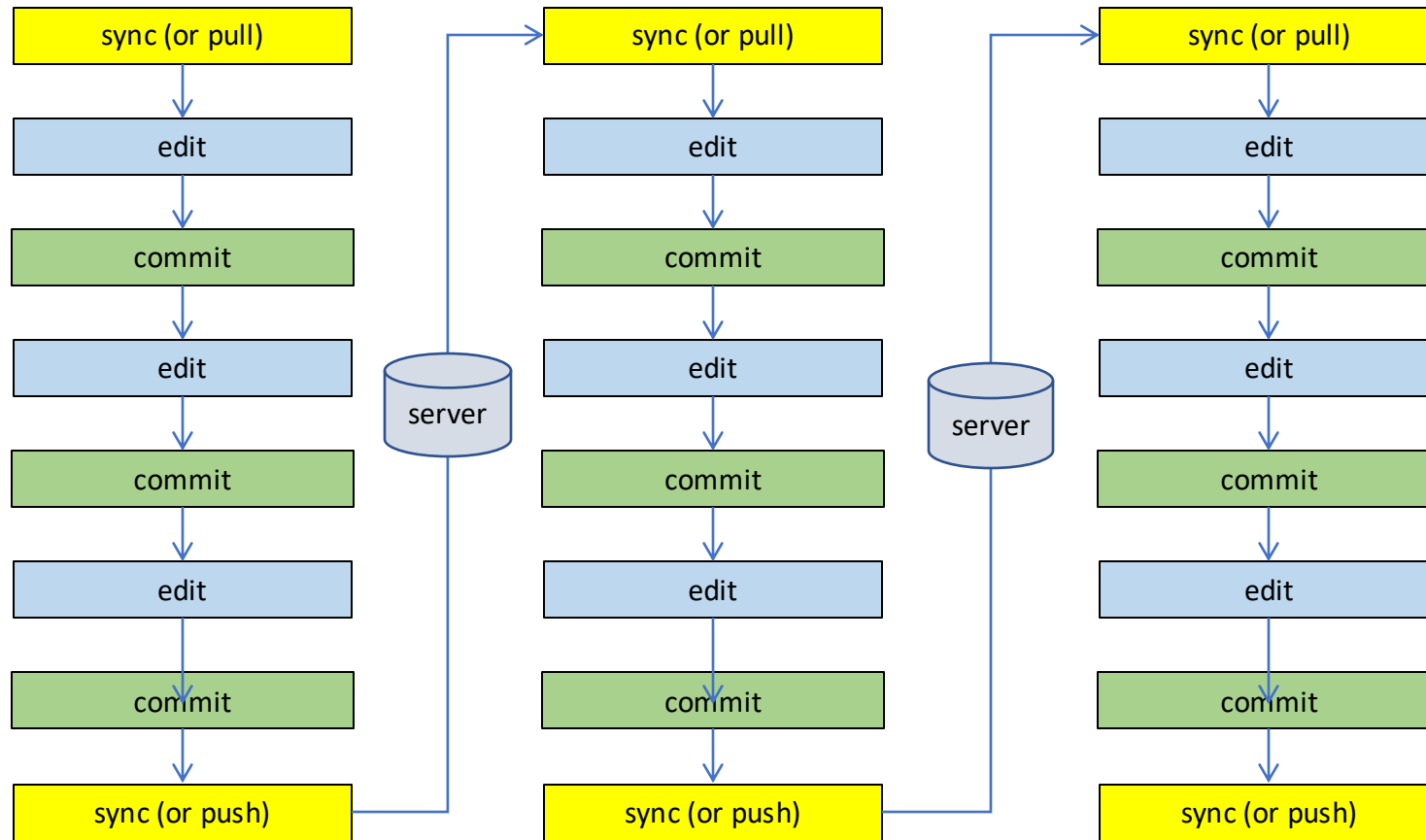
Open with GitHub Desktop

Download ZIP

Sample session commands

```
> git clone <remote_path> <local_directory>  
> cd <local_directory>  
> git pull  
> touch newfile.txt  
> git add newfile.txt  
> git commit -m "added a new file"  
> git push
```

Suggested workflow



**This is what
we grade
from!**

stop working /
start working

stop working/
change
computers

References

- <https://git-scm.com/book/en/v2>
- <http://www-cs-students.stanford.edu/~blynn/gitmagic/>
- Many YouTube videos
 - e.g., FreeCodeCamp <https://www.youtube.com/watch?v=RG0j5yH7evk>

