

# Github: A Collaborative Code-Sharing Tool

Nelsha Athauda

2022-10-18

# What is Github?

"The complete developer platform to build, scale, and deliver secure software."

- Data Analysis
- Version History
  - See history of changes made to a document
  - Revert to older versions of a document
  - Try new things out without breaking things that work
- Collaboration

tl;dr: Stop saving fifty copies of the same file with different names

# Terminology

- Repository
  - A directory of (ideally) plain-text files (.html, .txt, .R, .Rmd, etc.) included in a project
- Commit
  - A unique flag for a particular state of your project
  - Every push to the server is an updated version = a new commit
- History
  - All of the commits for a project
- Branch
  - An independent line of commits of your repository -Fork
  - A copy of someone else's GitHub repository that you can edit

# Example repository

The screenshot shows a GitHub repository page for 'NelshaAth / Assignment4'. The repository is public and has 1 branch (main) and 0 tags. The commit history shows 8 commits by user 91cc739 on Dec 20, 2021. The file list includes folders Data, R, Rmd, figs, and renv, and files .Rprofile, Dockerfile, Makefile, README.md, and renv.lock. The README.md file is open, showing the title 'INFO-550 Data Science Toolkit Coursework' and a description of the final assignment. The right sidebar shows repository statistics: 0 stars, 1 watching, 1 fork, and no releases or packages published. The language distribution bar shows HTML at 98.0% and R at 2.0%.

NelshaAth / Assignment4 Public

Code Issues 2 Pull requests Actions Projects Wiki Security Insights Settings

main 1 branch 0 tags

Go to file Add file Code

About

INFO 550 Assignment 4

Readme 0 stars 1 watching 1 fork

Releases

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

Packages

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

Languages

HTML 98.0% R 2.0%

NelshaAth Removed docker build command 91cc739 on Dec 20, 2021 8 commits

Data	This is Assignment 4!	11 months ago
R	This is Assignment 4!	11 months ago
Rmd	This is Assignment 4!	11 months ago
figs	This is Assignment 4!	11 months ago
renv	This is Assignment 4!	11 months ago
.Rprofile	This is Assignment 4!	11 months ago
Dockerfile	Added Dockerfile	10 months ago
Makefile	This is Assignment 4!	11 months ago
README.md	Removed docker build command	10 months ago
renv.lock	This is Assignment 4!	11 months ago

README.md

## INFO-550 Data Science Toolkit Coursework

Hi! For the final assignment, I analyzed a data file from an particulate matter sensor (units 91) to see if the aerosol optical depth wavelength channels are aligned after the unit was serviced.

This is a repository of the code that I wrote for my Data Science Toolkit Final Exam. It compares data across two collocated air quality samplers. The repository is the directory of files. The README.md file describes the project and how to run it.

# Example commit

Showing 1 changed file with 3 additions and 7 deletions.

10 README.md	
@@ -14,22 +14,18 @@	A makefile is in the main project folder which is able to generate the report wi
14 14	make
15 15	```
16 16	
17	- ## Final Assignment :D
17	+ ## Final Assignment
18 18	Link to Docker Hub Image: <a href="https://hub.docker.com/repository/docker/nelshaath/info_550">https://hub.docker.com/repository/docker/nelshaath/info_550</a>
19 19	
20 20	To download the Docker image of this project, pull the image in your terminal:
21 21	```
22 22	docker image pull nelshaath/info_550:latest
23 23	```
24	- and then build the image
25	- ```
26	- docker build -t info_550 .
27	- ```
28 24	To see the output you will need to mount the directory to a local folder on your device. You must change the path to a folder on your device!!!
29 25	```
30 26	docker run -v /path/to/project/R:/info_550/Rmd -it info_550
31 27	```
32 28	
33	- The output is the Report.html file <3
29	+ The output is the Report.html file
34 30	
35	- Please email me at Nelsha.Athauda@emory.edu if you're having any problems! Happy end of term~
31	+ Please email me at Nelsha.Athauda@emory.edu if you're having any problems!

If you click on a commit on GitHub, you can see what lines were modified. The old version appears in red and the new version in green.

# Commit History

main ▾

Commits on Dec 20, 2021

- Removed docker build command  
NelshaAth committed on Dec 20, 2021 Verified 91cc739 <>
- Added Dockerfile  
NelshaAth committed on Dec 20, 2021 Verified ce45aca <>

Commits on Dec 4, 2021

- I... finally got it to work  
NelshaAth committed on Dec 4, 2021 Verified 59e0ef4 <>

Commits on Dec 3, 2021

- Added link  
NelshaAth committed on Dec 3, 2021 Verified 828a3a1 <>
- first final update  
NelshaAth committed on Dec 3, 2021 Verified 8130567 <>

Commits on Nov 3, 2021

- Updated Readme to include make line  
NelshaAth committed on Nov 3, 2021 214baf2 <>
- Updated Readme to include Renv  
NelshaAth committed on Nov 3, 2021 624c887 <>
- This is Assignment 4!  
NelshaAth committed on Nov 3, 2021 f55220b <>

This is the full project history, each commit is a different version.

# Using a terminal

Unless you want to upload your files directly to the website (inconvenient), you will need to learn how to use your computer's terminal.

All devices have a command line (Command Prompt) but you can download a WSL (Windows Subsystem for Linux) like Ubuntu which can download R itself and be used as a one-stop console and terminal.

WSL or Command Prompt can be selected in your IDE (I use RStudio or Sublime Text Editor 4)

# Starting with Github

Go to <https://github.com/> and create a GitHub account. Also go into your settings and create a PAT (Personal Access Token) which will be required to access your Github account from a terminal.

Copy and paste your PAT somewhere easy to find or you'll have to make a new one!!



# Using Simple bash commands to create a local directory

```
cd ~/Desktop # Change Directory to where your project is
mkdir myproject # Make directory titled ___
cd myproject/ # Navigate to the project

git init # Initialize repository
```

Important: For any file addresses, use forward slashes only! A back slash in bash means to disregard the next character (used if you have spaces in the address)

Copied from File explorer: C:\Users\nrathauda\OneDrive - University of Alaska\R

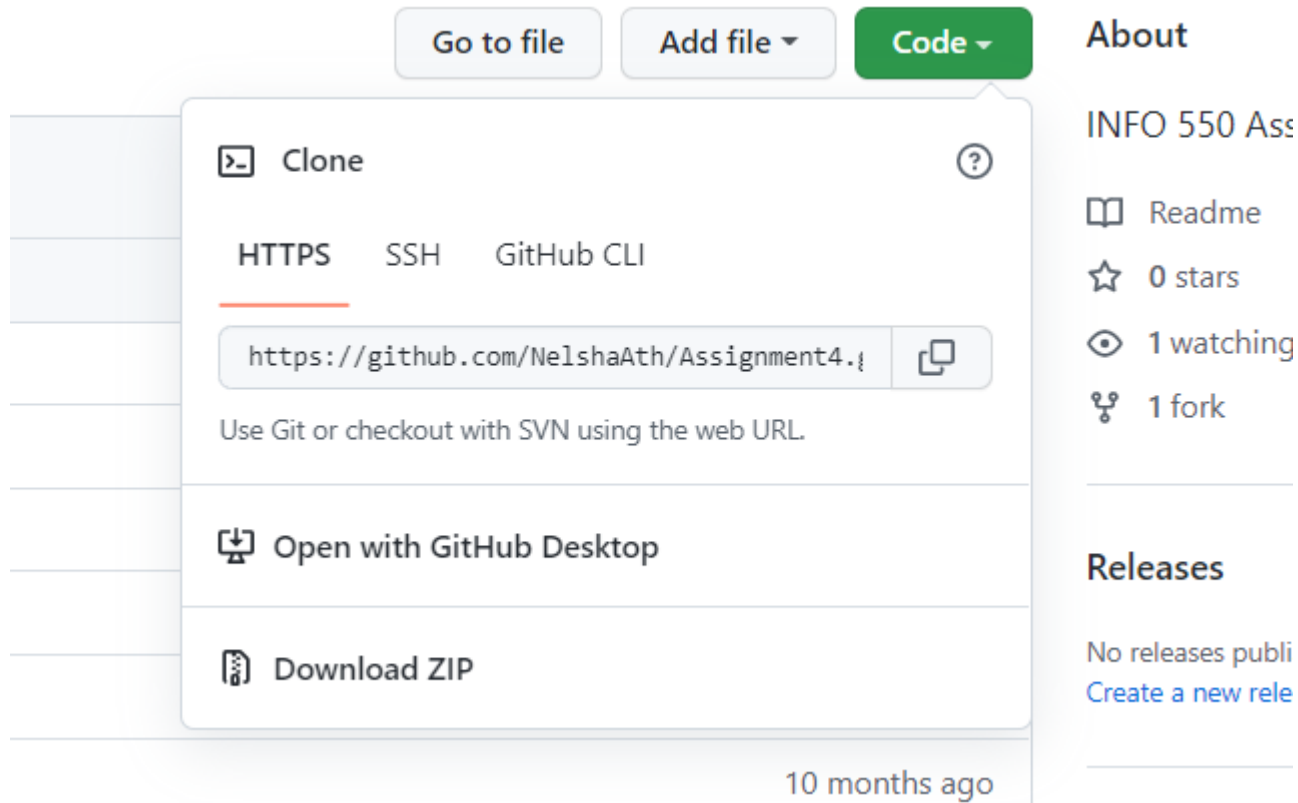
Adapted to bash: C:/Users/nrathauda/OneDrive\ -\ University\ of\ Alaska/R

# Maintain your local repository

```
git add # tell git which files to add to this commit  
git add . # The period will add all files in the folder  
git commit -m "I fixed so many things!" # make a commit with a message  
git status # shows what files have changed since your last commit
```

# Create a online repository (GitHub)

GitHub is the online platform. Create a repository on the website and copy the HTTPS share link to create a remote (bridge) between the Hub and your local repository



The screenshot shows a GitHub repository interface. At the top, there are three buttons: "Go to file", "Add file", and "Code". The "Code" button is highlighted in green. Below it, a dropdown menu is open, showing options to clone the repository. The "Clone" option is selected, and the "HTTPS" method is chosen. The URL displayed is `https://github.com/NelshaAth/Assignment4.1`. Below the URL, there is a text input field with the same URL and a copy icon. To the right of the dropdown, there is a section titled "About" with the following information: "INFO 550 Ass", "Readme", "0 stars", "1 watching", and "1 fork". Below this, there is a section titled "Releases" with the text "No releases publi" and a link "Create a new rele". At the bottom right, there is a timestamp "10 months ago".

Go to file Add file Code

Clone

HTTPS SSH GitHub CLI

`https://github.com/NelshaAth/Assignment4.1`

Use Git or checkout with SVN using the web URL.

Open with GitHub Desktop

Download ZIP

About

INFO 550 Ass

Readme

0 stars

1 watching

1 fork

Releases

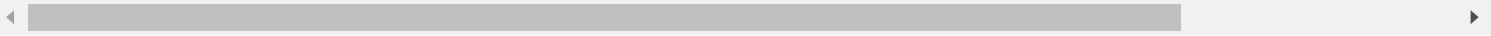
No releases publi

Create a new rele

10 months ago

# Pushing your local git repo onto your online GitHub repo

```
git remote add origin https://github.com/[your-username]/[repository-  
git push -u origin main  
# Pushing your origin (local repo) to the Hub (online repo) on a brai
```



After sending git push, your terminal will ask for your GitHub username and password/PAT

# Check your GitHub profile

If everything pushed successfully, there should be a recent commit!



# Using Branches

If you want to try something risky but don't want to mess up your current working code you can use branches Commit the working version on your normal branch (Main) Start working on your risky changes but commit the updates to a new branch (Ex. Branch1) If your new branch works out, you can merge the changes you made onto your original Main branch

```
# make a branch called Branch1  
git branch Branch1  
# checkout is used to switch between branches  
git checkout Branch1  
# see all branches  
git branch
```

# Using Branches (cont.)

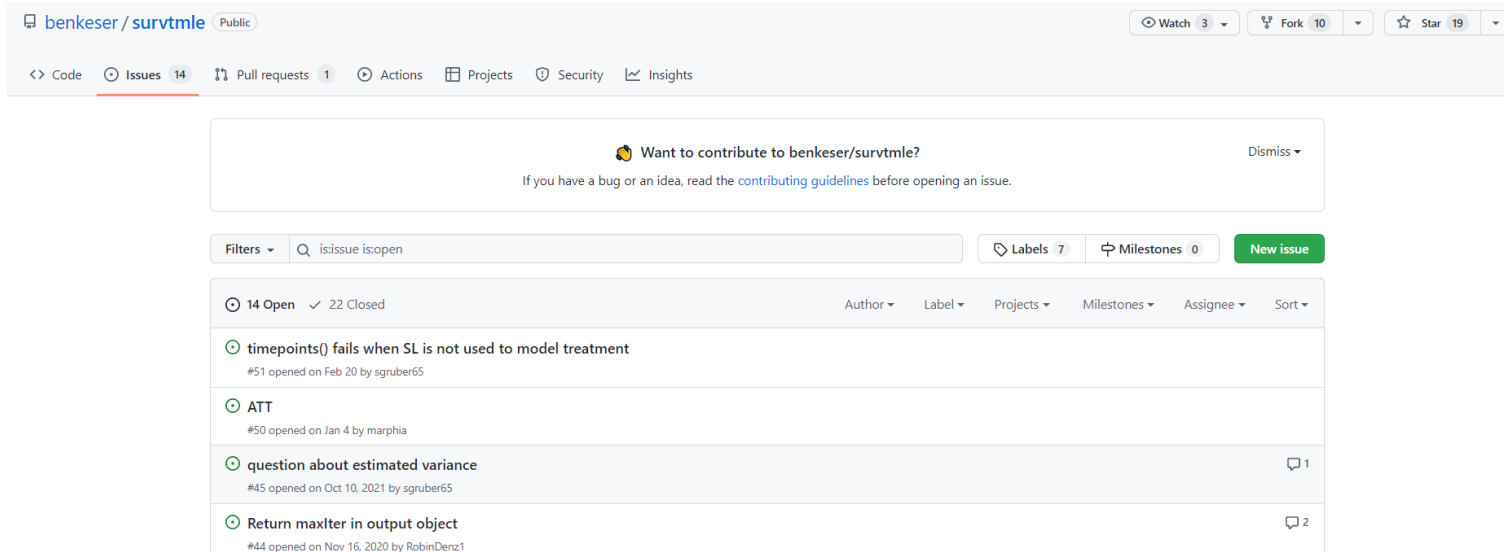
Both branches can be edited, merging will add/delete to the branch you want it to act on

```
# Switch back to the original branch  
git checkout Main  
# Merge the changes on Branch1 to Main  
git merge Branch1
```

Merge conflicts will come up if you edited the same lines in different branches and will have to be resolved manually (git will show you which lines)

# Other ways to collaborate

If you notice a bug in someone's code you can file an issue (Similar to a YouTube comment)



The screenshot shows the GitHub repository page for `benkeser/survtmle`. The repository is public and has 3 watchers, 10 forks, and 19 stars. The navigation bar includes links to Code, Issues (14), Pull requests (1), Actions, Projects, Security, and Insights. A banner at the top asks if the user wants to contribute to the repository, with a link to the contributing guidelines. Below the banner, there is a search bar with the filter `is:issue is:open` and buttons for Labels (7) and Milestones (0). A green button labeled "New issue" is also present. The list of issues shows 14 open and 22 closed issues. The first three issues are:

- `timepoints()` fails when SL is not used to model treatment (#51 opened on Feb 20 by sgruber65)
- ATT (#50 opened on Jan 4 by marphia)
- question about estimated variance (#45 opened on Oct 10, 2021 by sgruber65)

The last issue shown is "Return `maxlter` in output object" (#44 opened on Nov 16, 2020 by RobinDenz1).

Or you can create a pull request

- Fork the repo on GitHub
- git clone to download to local machine
- Modify code (read contributing guidelines first!)
- Commit changes
- git push back to GitHub



# Pull Request Workflow

```
# add friend's repo as a remote branch
git remote add wrongdoer https://github.com/wrongdoer/repo
# downloads friend's branch, but do not merge it yet
git fetch wrongdoer master
# view all local and remote branches
git branch -a
# checkout friends remote branch
git checkout remotes/wrongdoer/master
# make a local branch based on friend's repo
git checkout -b wrongdoer
# test out the branch; make sure it works as expected
[...]
# checkout local master, merge, push
git checkout master
git merge wrongdoer
git push origin master
```

Now the fixed code is in YOUR repository

To give the fixed code to the person who created the version with a bug, create a pull request on Github, in their repository

The screenshot shows the GitHub interface for the repository 'benkeser/survtmle'. At the top, there is a navigation bar with tabs for Code, Issues (14), Pull requests (1), Actions, Projects, Security, and Insights. The 'Pull requests' tab is selected and highlighted with a red underline.

Below the navigation bar, there is a message box titled 'First time contributing to benkeser/survtmle?' with a 'Dismiss' link. The message text reads: 'If you know how to fix an issue, consider opening a pull request for it. You can read this repository's contributing guidelines to learn how to open a good pull request.'

Below the message box, there is a search bar with the text 'is:pr is:open'. To the right of the search bar, there are buttons for 'Labels 7' and 'Milestones 0', and a green button labeled 'New pull request'.

Below the search bar, there is a table of pull requests. The table has columns for 'Author', 'Label', 'Projects', 'Milestones', 'Reviews', 'Assignee', and 'Sort'. The first row shows a pull request titled '[WIP] Known observation-level weights' with a red 'X' icon and a blue 'enhancement' label. The pull request is marked as '1 Open' and '14 Closed'. Below the title, it says '#43 opened on Jun 6, 2020 by nhejazi'.

At the bottom of the page, there is a 'ProTip!' message: 'Adding nolabel will show everything without a label.'

# On the receiving end of a pull request?

Merge it on GitHub itself To add it to your local copy, git fetch or git pull

git fetch

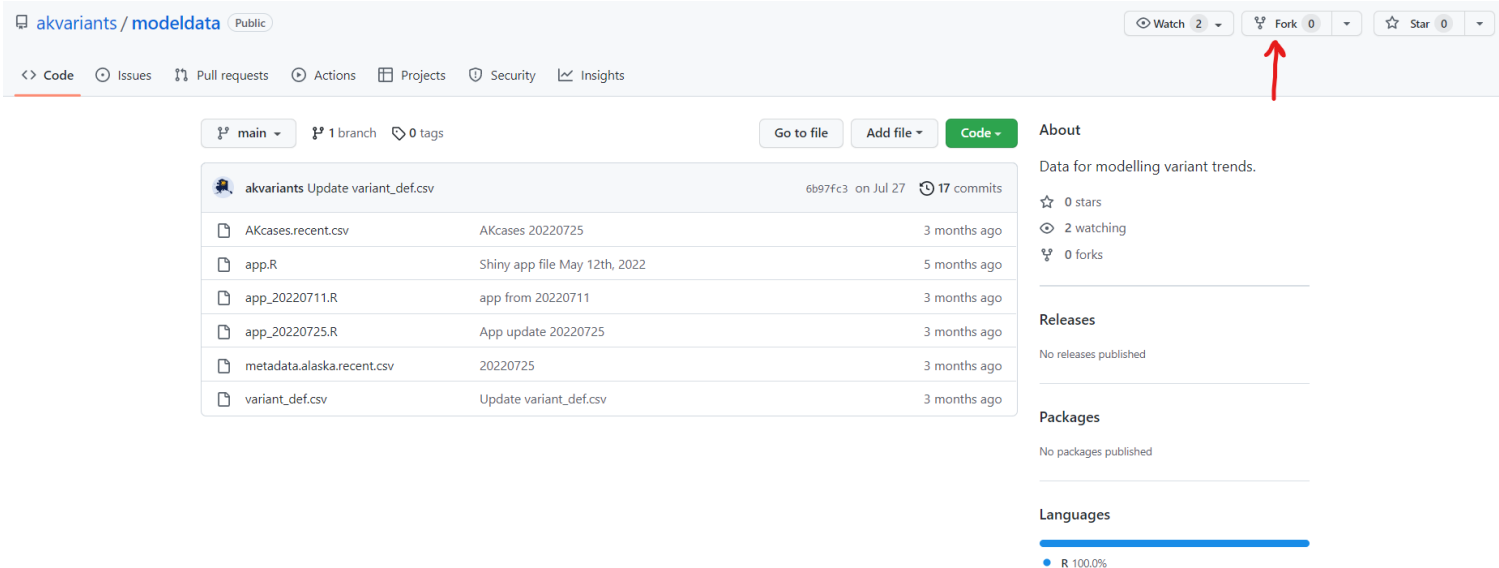
- Download code separately from your local repository.
- Explicitly merge into your local repository.

git pull

- fetch and merge at the same time
- Faster, but you can't verify that the code works before changing your local repository.

# Want to see/use/adapt someone else's code?

Find the repository on GitHub and fork it to create a copy that you can alter.



akvariants / modeldata Public

Watch 2 Fork 0 Star 0

Code Issues Pull requests Actions Projects Security Insights

main 1 branch 0 tags

Go to file Add file Code

akvariants Update variant\_def.csv 6b97fc3 on Jul 27 17 commits

AKcases.recent.csv	AKcases 20220725	3 months ago
app.R	Shiny app file May 12th, 2022	5 months ago
app_20220711.R	app from 20220711	3 months ago
app_20220725.R	App update 20220725	3 months ago
metadata.alaska.recent.csv	20220725	3 months ago
variant_def.csv	Update variant_def.csv	3 months ago

About

Data for modelling variant trends.

0 stars  
2 watching  
0 forks

Releases

No releases published

Packages

No packages published

Languages

R 100.0%

To bring the copy to your local device, use git clone

```
git clone https://github.com/Your-name/The-Forks-name
```

# Great Job!

Git has an incredibly steep learning curve but its version control and reproducibility features make it worth it.

