

Basics of Using git & GitHub

Justin Post

Workflow

Important considerations for data analysis workflow:

- **Reproducibility**

- **Version control**

- **Collaboration**

What are git and github?

- **Git** is a version control software
- **Github** is an online hosting service for Git-based projects

What are git and github?

- **Git** is a version control software
- **Github** is an online hosting service for Git-based projects

Git Tracking

- You associate git with a folder (repo)
- Git keeps track of all files in the folder (repo)
- If you want to keep changes you've made, you **commit** and **push** the changes to the folder (repo)

Github

- Github allows you to have a remote file repository (folder) tracked by git
 - Let's create a repository on github.com
 - **Add** some files and **commit** to the changes
 - Modify some files on github
 - Investigate the version control!

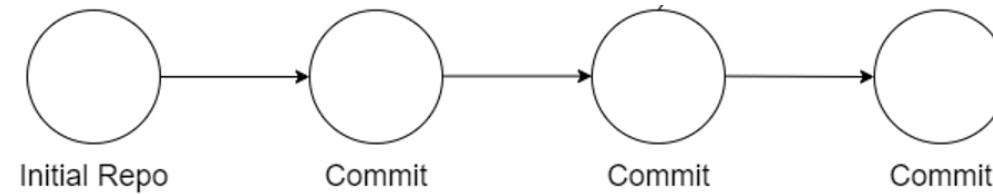
Local vs Remote Work

Mostly we'll be working on Google Colab for now. If we were working on our local machine, we'd need to install `git` on our computer!

- When we get to working on our local `JupyterHub`, we'll just use the web-based interface with `gitHub`.
- Next, let's talk about how we'd collaborate using `git` and `gitHub`

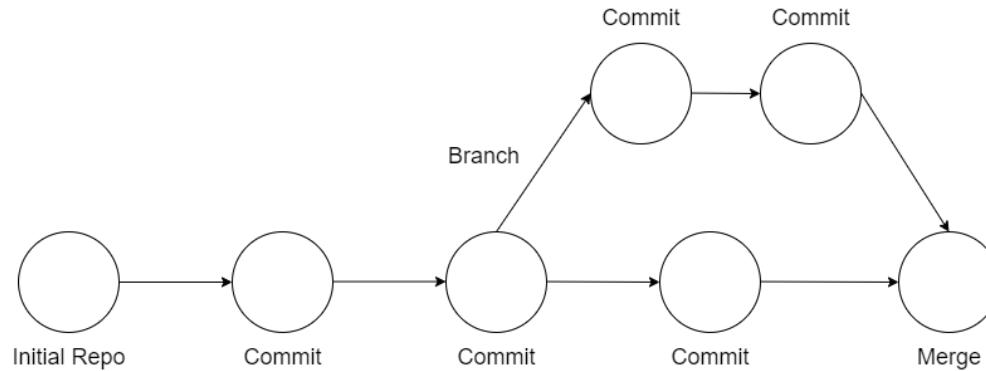
Collaboration Idea

- Everyone can work on the same **branch** and just take turns working



Working on Branches

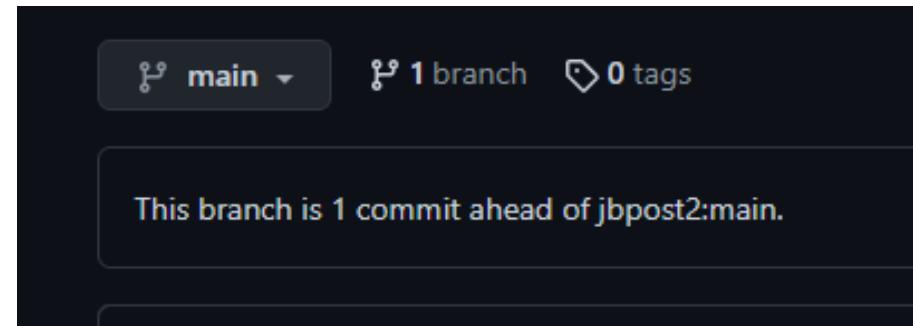
- Alternatively, you can have separate **branches** of the repo



- Work on a branch is similar to working on the main branch
- Can merge when happy!

Forking

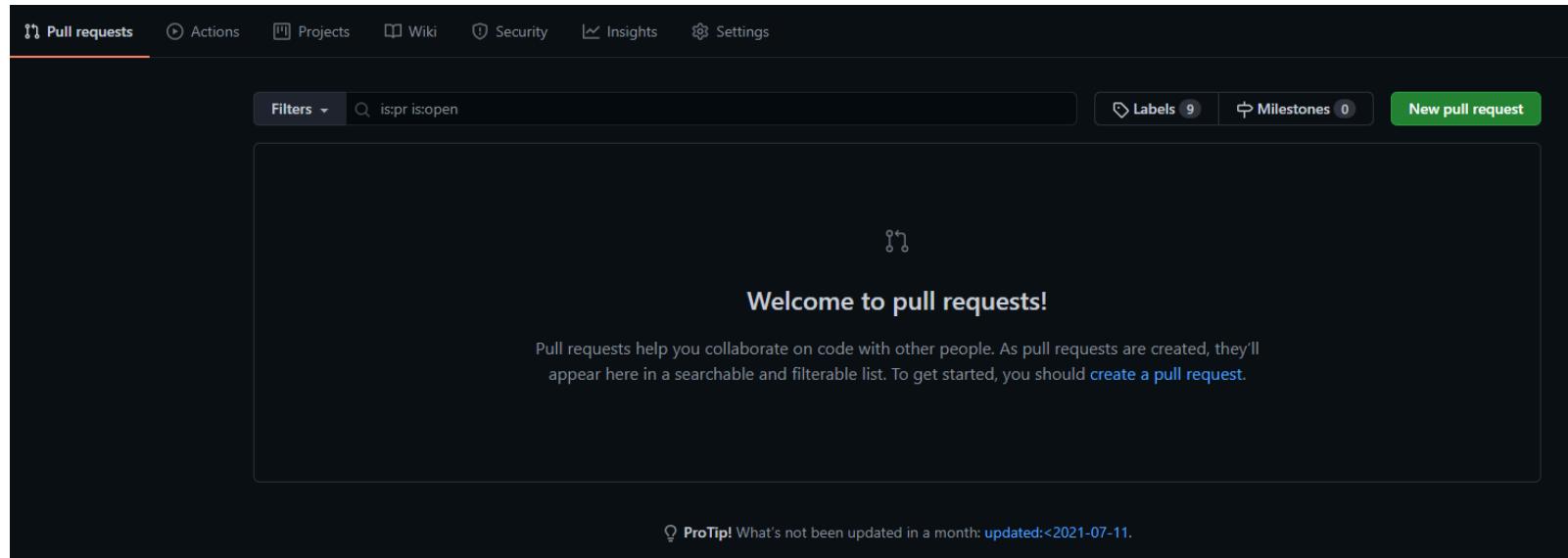
- People often **fork** the repo
- This creates a copy of the repo on your account
- You can then work as normal
- If you do a commit on your branch, you may notice something like this



Merging Branches

Suppose you like your commit and you think the original owner will too!

- You can do a **pull** request



Merging Branches

Suppose you like your commit and you think the original owner will too!

- You can do a **pull** request

Comparing changes

Choose two branches to see what's changed or to start a new pull request. If you need to, you can also [compare across forks](#).

base repository: [jbpot2/myBook](#) [base: main](#) [head repository: JustinBPost/myBook](#) [compare: main](#) ✓ **Able to merge.** These branches can be automatically merged.

Discuss and review the changes in this comparison with others. [Learn about pull requests](#) [Create pull request](#)

-o- 1 commit [1 file changed](#) 0 comments 1 contributor

Commits on Aug 12, 2021

Update 05-summary.Rmd [Verified](#) 2f7ad8f

Showing 1 changed file with 2 additions and 0 deletions. [Unified](#) [Split](#)

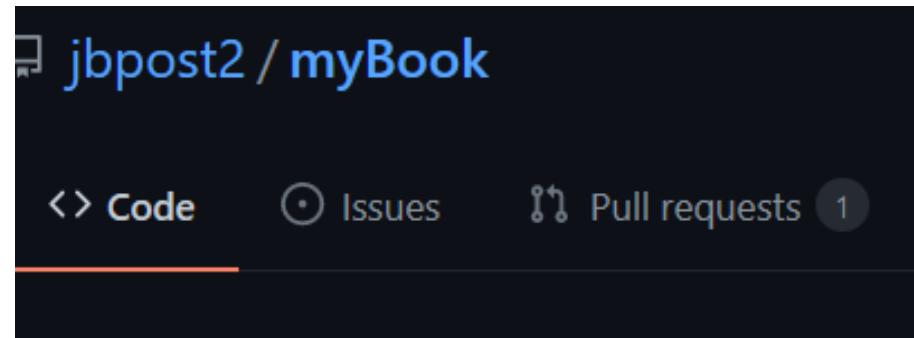
05-summary.Rmd

```
... ... @@ -1,3 +1,5 @@  
1 1 # Final Words  
2 2  
3 3 We have finished a nice book.  
4 +  
5 + Or did we?
```

Merging Branches

If you are lucky, there won't be any merge conflicts.

- Allows the owner of the original repo to accept the pull request without needing to modify things
- The owner will get a notification that a pull request has been made



Merging Branches

Owner can then investigate the request and choose whether or not to accept it or they can ask for more details

Update 05-summary.Rmd #1

[Open](#) JustinBPost wants to merge 1 commit into `jbpost2:main` from `JustinBPost:main` 

Conversation 0 Commits 1 Checks 0 Files changed 1

 JustinBPost commented 4 minutes ago First-time contributor  

No description provided.

 Update 05-summary.Rmd Verified 2f7ad8f

Add more commits by pushing to the `main` branch on [JustinBPost/myBook](#).

 **Continuous integration has not been set up**
GitHub Actions and several other apps can be used to automatically catch bugs and enforce style.

 **This branch has no conflicts with the base branch**
Merging can be performed automatically.

[Merge pull request](#) You can also open this in GitHub Desktop or view command line instructions.

Dealing with conflicts

- Sometimes changes requested conflict with changes already made

Comparing changes

Choose two branches to see what's changed or to start a new pull request. If you need to, you can also [compare across forks](#).

base repository: [jbpost2/myBook](#) [base: main](#) ← head repository: [JustinBPost/myBook](#) [compare: main](#)

X **Can't automatically merge.** Don't worry, you can still create the pull request.

Discuss and review the changes in this comparison with others. Learn about [pull requests](#) [Create pull request](#)

-o 1 commit [1 file changed](#) 0 comments 1 contributor

Commits on Aug 12, 2021

-o [Update 02-literature.Rmd](#) [Verified](#) 366353a

Showing 1 changed file with 1 addition and 0 deletions. [Unified](#) [Split](#)

1 02-literature.Rmd

```
... ... @@ -1,3 +1,4 @@
1 1 # Literature
2 2 Yes I like this line.
3 3 Here is a review of existing methods.
4 + This line is better than yours, approve this change.
```

Dealing with conflicts

Owner sees a notification about conflicts that must be resolved

Update 02-literature.Rmd #2

[Open](#) JustinBPost wants to merge 1 commit into `jbpost2:main` from `JustinBPost:main` 

 Conversation 0  1  0  1

 **JustinBPost** commented 26 seconds ago Contributor  

No description provided.

 Update 02-literature.Rmd Verified 366353a

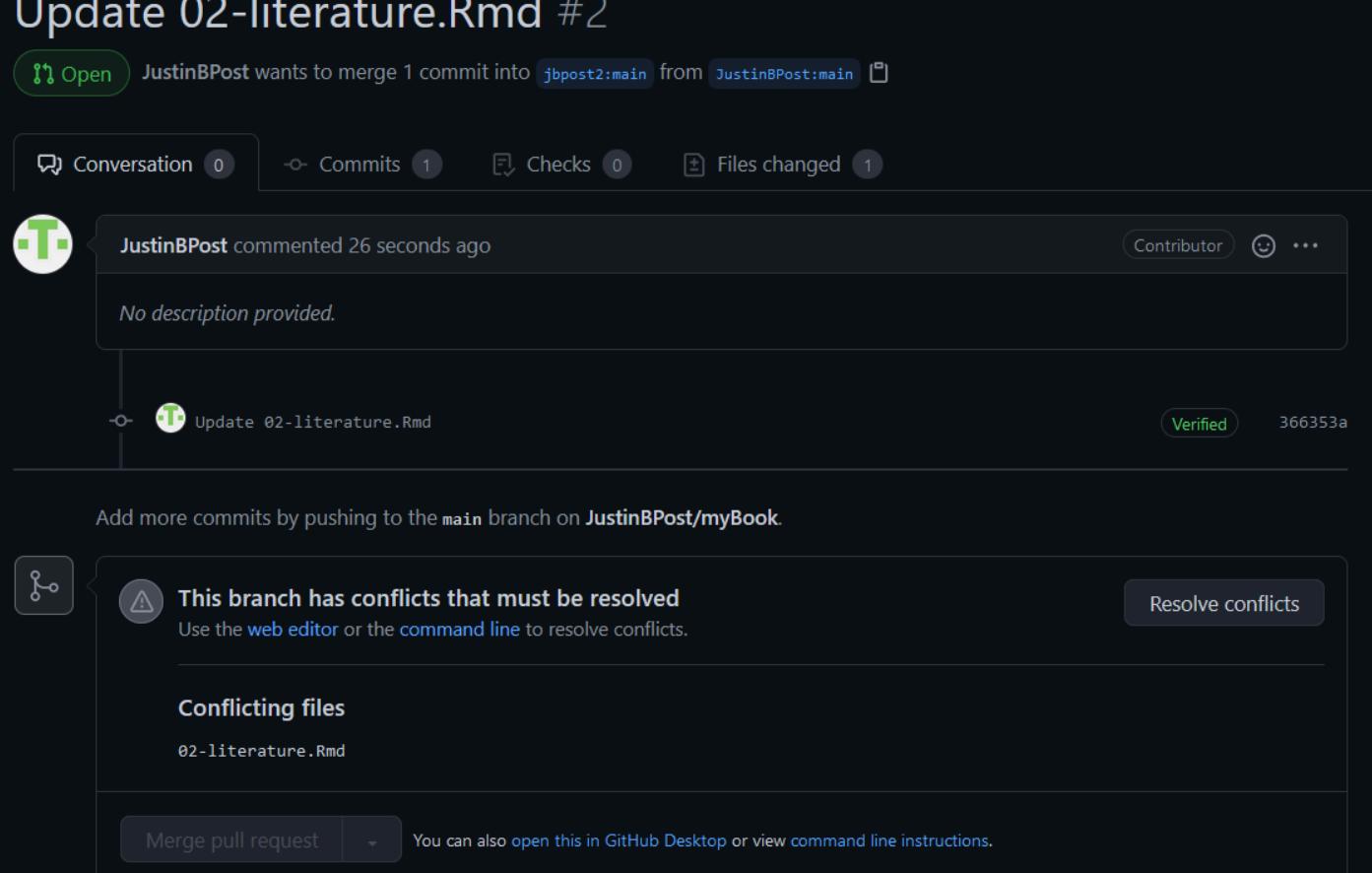
Add more commits by pushing to the `main` branch on [JustinBPost/myBook](#).

 **This branch has conflicts that must be resolved** Resolve conflicts

Use the [web editor](#) or the [command line](#) to resolve conflicts.

Conflicting files
`02-literature.Rmd`

[Merge pull request](#)  You can also open this in [GitHub Desktop](#) or view [command line instructions](#).



Dealing with conflicts

They can view the issues and pick which to include or to include both with a modification

<<<<< is a conflict marker

Update 02-literature.Rmd #2

Resolving conflicts between `JustinBPost:main` and `jbpost2:main` and committing changes → `JustinBPost:main`

1 conflicting file	02-literature.Rmd
<code>02-literature.Rmd</code> 02-literature.Rmd	<pre>1 # Literature 2 Yes I like this line. 3 Here is a review of existing methods. 4 <<<<< main 5 This line is better than yours, approve this change. 6 ===== 7 This line isn't so good. 8 >>>> main 9</pre>

- Figure out what to do and delete the <<< ==>>> lines

Recap

- **Git** is a version control software
 - Associated with a folder (repo)
 - Tracks changes to files
- **Github** is an online hosting service for Git-based projects
- Workflow:
 - Create a repo on `gitHub`
 - Open `.ipynb` files in Colab
 - Save a copy on `gitHub`!