

Recitation 6: Git and Github

















Eugene Tan (TA)

10/16/2020

Outline for today

- ▶ Version Control
- ▶ Git commits as snapshots
- ▶ A sequence of commits as version history
- ▶ Repositories
- ▶ Collaborating on github
- ▶ Troubleshooting - commits at the same time
 - ▶ merging
 - ▶ rebasing

Version Control

 lac_redraft_ET4.docx		Jul 6, 2019 at 4:32 PM	18.3 MB
 lac_redraft_ET5.docx		Aug 8, 2019 at 5:56 PM	21.5 MB
 lac_writeup_AJ.docx		Jun 10, 2019 at 7:53 AM	13.1 MB
 lac_writeup_ET2.docx		Jun 17, 2019 at 5:38 PM	13.2 MB
 lac_writeup_ET3.docx		Jun 24, 2019 at 5:36 PM	13.1 MB
 lac_writeup.docx		Apr 19, 2019 at 5:19 PM	324 KB
 Unbreakable_ARG_zero_draft_ET.docx		May 16, 2019 at 7:10 PM	312 KB
 Unbreakable_ARG_zero_draft.docx		May 16, 2019 at 12:49 PM	147 KB

- ▶ Emailing your file back and forth
- ▶ Working over zoom
- ▶ Time difference

Version Control

Git solves this problem for code. Code is different from normal writing because your end depends on your beginning. Uses multiple files.

From datacamp:

- ▶ Nothing that is saved to Git is ever lost, so you can always go back to see which results were generated by which versions of your programs.
- ▶ Git automatically notifies you when your work conflicts with someone else's, so it's harder (but not impossible) to accidentally overwrite work.
- ▶ Git can synchronize work done by different people on different machines, so it scales as your team does.

Version Control - production and accountability

Nothing that is saved to Git is ever lost, so you can always go back to see which results were generated by which versions of your programs.

Version Control - preventing bugs in collaboration

Git automatically notifies you when your work conflicts with someone else's, so it's harder (but not impossible) to accidentally overwrite work.

Version Control - encourage collaboration on code

Git can synchronize work done by different people on different machines, so it scales as your team does.

Version Control - preventing bugs in collaboration

```
* 2df59fd - (5 months ago) Fixes a quite bad link (#849) - Garrett Grolemond (HEAD -> exercises, origin/master, origin/HEAD, master)
| * 01246a4 - (5 months ago) Fixes the other instance of a very bad link - Garrett Grolemond (origin/garr
ettgman-patch-3-1)
|/
* 9dda73b - (9 months ago) typo in 16.4.5 Exercises 5 (#732) - Zhuoer Dong
* 5f85b4c - (9 months ago) Remove use of mu in exercises (#748) - Jeffrey Arnold
* 5082dd8 - (9 months ago) Add equations to 19.2 Ex 4. (#749) - Jeffrey Arnold
* 6194b38 - (9 months ago) Rename "Practice" sections to "Exercises" (#750) - Jeffrey Arnold
* e25d3ac - (9 months ago) use `stat()` instead of `..` in "3.7 Statistical transformations" (#751) - Zhuoer Dong
* 6eb53ea - (9 months ago) update tidyverse loadings (#754) - Jacek Kolacz
* 1dd2934 - (9 months ago) Remove duplicate id - Hadley Wickham
* 50e3db4 - (9 months ago) Assuring minimal version of Rstudio (#758) - Edwin Thoen
* a5fbbb5 - (9 months ago) Revert typo (#760) - Jeffrey Arnold
* 92baa5d - (9 months ago) Update model-basics.Rmd (#765) - MattWittbrodt
* 37b3029 - (9 months ago) inserting forgotten of (#767) - Edwin Thoen
* 3b91773 - (9 months ago) typos (#773) - Stéphane Guillou
* 4708c0d - (9 months ago) Improve grammatical consistency (#774) - Benjamin Yeh
* 06fb008 - (9 months ago) Improve exercise wording (#781) - Kunal Marwaha
* be32c6c - (9 months ago) Update model-building.Rmd (#782) - BirgerNi
* 2712ba1 - (9 months ago) Fixed text to correspond with code output (#784) - Bianca Peterson
* 602baae - (9 months ago) Update import.Rmd (#787) - Pablo E
* 2182fff6 - (9 months ago) name hour_perc should be hour_prop because proportion not percentage (#795) - a-rosenberg
* ba3892e - (9 months ago) Update to say that there are over 40 geoms (#801) - Nicholas Tierney
* 4e2d2a1 - (9 months ago) good news for kakapos (#804) - Stéphane Guillou
* 8784e41 - (9 months ago) Update Chp 8 exercise on quote arg of read_csv(). Close #696. (#806) - John Blischak
* 0830626 - (9 months ago) Fix typo (#809) - Mine Cetinkaya-Rundel
* 80aec62 - (9 months ago) Correct typo or explanation (#818) - robertchu03
* e8e0218 - (9 months ago) Eliminate devtools dep - Hadley Wickham
* 09fd870 - (9 months ago) Fix quotes in caption - Hadley Wickham
* 6a65eb9 - (9 months ago) No longer need dev versions of packages - Hadley Wickham
* 2459702 - (9 months ago) Updates R4ds to use pivot_ functions (#820) - Garrett Grolemond
* ce824d0 - (1 year, 9 months ago) Add google analytics (#745) - Mara Averick
```


What my git repo looks like for this class

The *repository* is the folder in which everything is tracked A *commit* is a snapshot

Public repo Private repo

```
* ca079c6 - (5 minutes ago) recitation 6 init - Eugene Tan Perk Han (HEAD -> master)
| * 0411cbb - (23 hours ago) some progress hw2 - Eugene Tan Perk Han (hw2)
| * a186011 - (4 days ago) ignacia's edits to hw2 - Eugene Tan Perk Han (origin/hw2)
| * 1c30c15 - (4 days ago) final edits hw1 answer key - Eugene Tan Perk Han (origin/master)
| * 4592948 - (4 days ago) hw2 init - Eugene Tan Perk Han
|/
* daec0ea - (4 days ago) cleanup - Eugene Tan Perk Han (public/master)
| * 0ed0725 - (4 days ago) private datacamp progress - Eugene Tan Perk Han (origin/grading, grading)
| * 331004d - (4 days ago) private outline - Eugene Tan Perk Han
| * 828be15 - (4 days ago) sent out mds for feedback - Eugene Tan Perk Han
| * 618334d - (5 days ago) gabriel add file - Eugene Tan Perk Han
| * be6b6aa - (5 days ago) remove .mds - Eugene Tan Perk Han
| * f0a7279 - (5 days ago) add luna and liz - Eugene Tan Perk Han
| * 70667bf - (5 days ago) qingyi, amy, zora done. gabriel/chandni midway - Eugene Tan Perk Han
| * 89039a4 - (5 days ago) julio comments - Eugene Tan Perk Han
| * e178b5e - (5 days ago) add julio and beiwen - Eugene Tan Perk Han
| * 8899cfc - (6 days ago) WIP on grading: 12211e0 add mds - Eugene Tan Perk Han
| \
| | * 1faa450 - (6 days ago) index on grading: 12211e0 add mds - Eugene Tan Perk Han
| | /
| * 12211e0 - (6 days ago) add mds - Eugene Tan Perk Han
| * 6d1794c - (6 days ago) grade zhongling and yiwen - Eugene Tan Perk Han
| * 1fc7d53 - (6 days ago) filename and zip removal - Eugene Tan Perk Han
| * 23e5302 - (6 days ago) bastian and francisco diff and comments, muhammad and mehulika edits - Eugene Tan Perk Han
| * a117b0b - (6 days ago) bastian francisco init - Eugene Tan Perk Han
| * cc742e7 - (6 days ago) add junbin and yuchen comments - Eugene Tan Perk Han
| * ffd935f - (6 days ago) add naveen comments - Eugene Tan Perk Han
| * 0aaf13c - (6 days ago) naveen comments - Eugene Tan Perk Han
| * c6b4db3 - (6 days ago) naveen's v1 - Eugene Tan Perk Han
| * 6fc199b - (6 days ago) graded angelas' - Eugene Tan Perk Han
| * 12ec111 - (6 days ago) add initial homeworks - Eugene Tan Perk Han (tag: subhw1)
|
|
```

Github

Github is just a place to store your git repos online. Allows collaboration.

Baby command-line tools

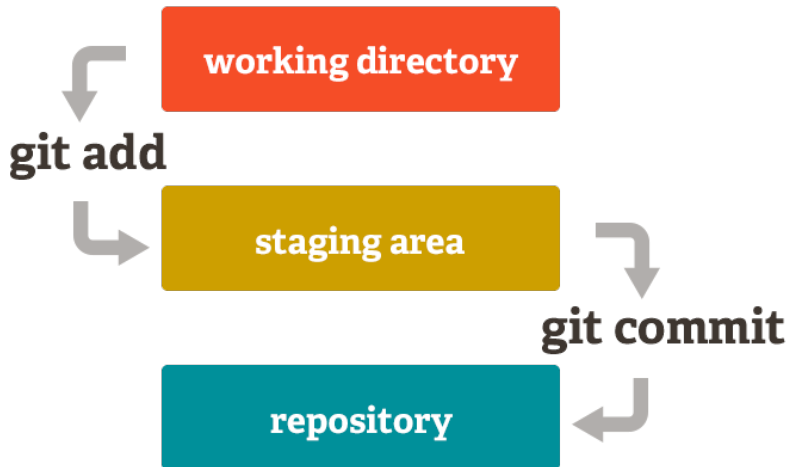
On git bash (windows) or on terminal (mac)

`computer_name:folder username$`

- ▶ `ls (-a, -l)`
- ▶ `pwd`
- ▶ `cd (., .., ~, -)`
- ▶ `mkdir/rmdir`
- ▶ `open (-a)`
- ▶ `tab-autocomplete`
- ▶ `man`
- ▶ `control-a control-e`

A commit is a snapshot.

`git add` to stage a commit. `git commit` to commit the staged items.



Demonstration.

Edit this file in this folder.

- ▶ `git status` checks the current status - relative to the last commit, what has changed, what has been staged, what is untracked?
- ▶ `git log` checks the version history through time (if you added the aliases for `lg1` and `lg2`, `git lg1` does the same thing)
- ▶ `git diff` checks the differences between my current working directory and the last commit
- ▶ `git diff hash_before hash_after` checks the differences between two commits
- ▶ `git add filename` stages an edited file
- ▶ `git commit -m "my commit message"` commits with the commit message 'my commit message'
- ▶ `git push` pushes to github.

Diff

What is the difference between the status now and the last commit?

Basic workflow

- ▶ Configure your name and email
- ▶ Configure security
- ▶ Create a git repository `git init` or `git clone`
- ▶ After you've made changes
 - ▶ Stage your changes `git add`
 - ▶ Commit a change `git commit`
 - ▶ Push your commits to GitHub `git push`
- ▶ After someone in your team has made changes
 - ▶ Get their commits from github `git pull`
 - ▶ Check what the differences they made are `git diff`
 - ▶ Work on it yourself

Configure your git identity

Set your username \$ `git config --global user.name "John Doe"`

Set your email \$ `git config --global user.email johndoe@example.com`

Configure others \$ `git config --global init.defaultBranch main`

Get a list of your current setup \$ `git config --list`

Some extra shortcuts to show the pretty logs I use above.

```
Make an alias for git lg1 (from here) $ git config --global
alias.lg1 "log --graph --abbrev-commit --decorate
--format=format:'%C(bold blue)%h%C(reset) - %C(bold
green)(%ar)%C(reset) %C(white)%s%C(reset) %C(dim
white)- %an%C(reset)%C(bold yellow)%d%C(reset)'
--all"
```

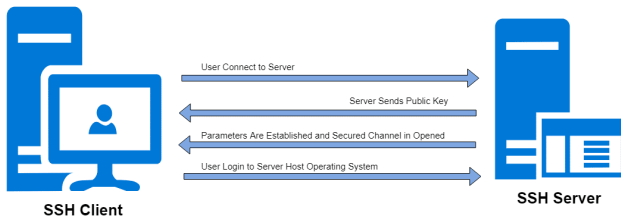
```
Make an alias for git lg2 $ git config --global alias.lg2
"log --graph --abbrev-commit --decorate
--format=format:'%C(bold blue)%h%C(reset) - %C(bold
cyan)%aD%C(reset) %C(bold green)(%ar)%C(reset)%C(bold
yellow)%d%C(reset)%n' '%C(white)%s%C(reset)
%C(dim white)- %an%C(reset)' --all"
```

Cryptography (Setup is advanced)

If you don't want to enter your github password every time, then set a key using this link

- ▶ One-way security
- ▶ You hold a private key that is linked to a public key

Secured Socket Layer (SSH) Process



Demonstration

- ▶ Volunteer will make a private repository on github called 'ee2'.
- ▶ Share it with me and her group member.
- ▶ Clone it to her folder for this class.
- ▶ Adds the first part of her homework ee2. Commit and push.
- ▶ I will clone it.
- ▶ Make an edit, add, commit, push.
- ▶ Liz will pull the changes, diff to see what I have done, and then continue working.

Now we'll cause an error

What happens if we make a change at the same time?

- ▶ She will make an edit to the file and commit
- ▶ She will push
- ▶ I will make an edit to the file and commit
- ▶ I will try to push

What is the problem here?

I need to pull before I edit the file. Editing the file at the same time causes conflicts.

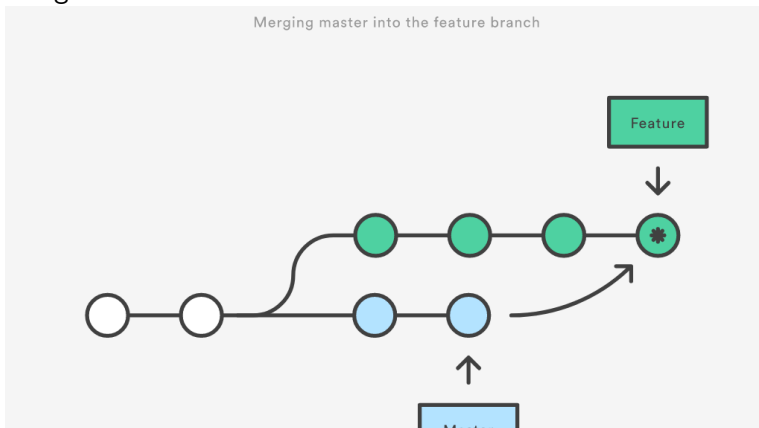
Dealing with conflicts

Easy way

- ▶ look at my file and her file and manually sort out the differences, commit that

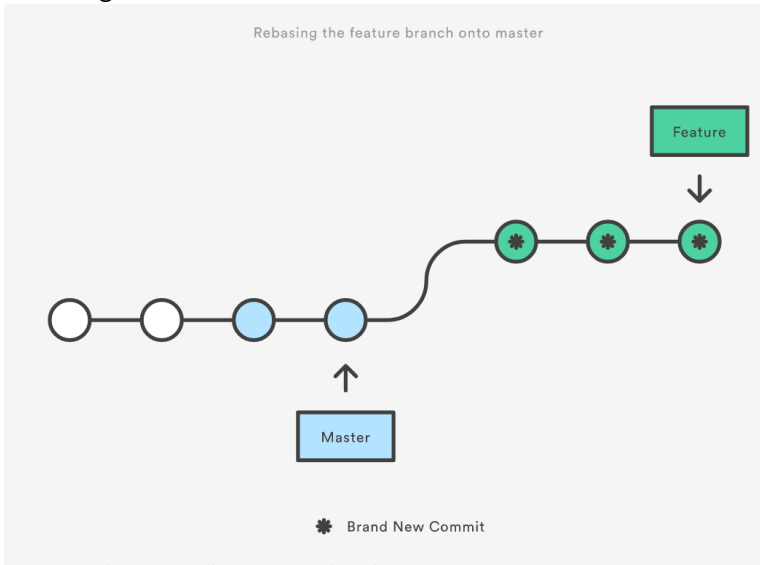
Intermediate way

- ▶ Merge branches



Advanced way

► Rebasing



Branches (Intermediate)



Tips

- ▶ Set up your repo on github.
- ▶ share your repo with your partner before you add anything
- ▶ commit whenever you paste something in,
- ▶ commit before and after you get each thing working
- ▶ commit often
- ▶ push often
- ▶ write informative commit messages