

Teaching Version Control in Introductory and Advanced Courses With GitHub Classroom: Experiences and Resources

Jacob Fiksel, BA

Leah Jager, PhD

Margaret Taub, PhD

Johanna Hardin (Pomona College), PhD

May 22, 2018



JOHNS HOPKINS

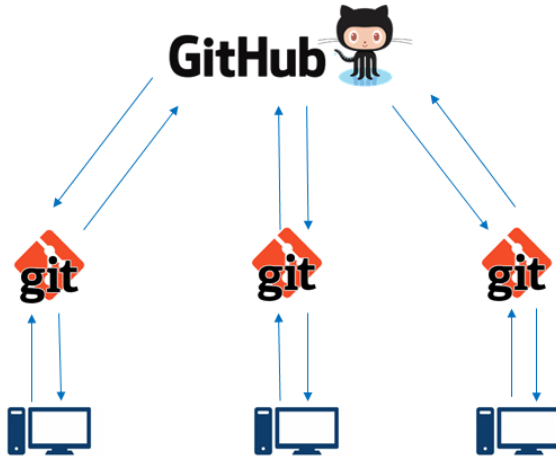
BLOOMBERG SCHOOL
of PUBLIC HEALTH

Data analysis is a team effort

- Businesses, governments, and researchers make data-driven decisions
- Data projects are collaborative—data, code, and results must constantly be updated and shared with the analysis team
- Need a method for managing large analysis projects
 - Git and GitHub!



Git and GitHub allow for easy management and sharing of data analytic content




GitHub allows you to track changes made during the course of a project

Improved GitHub pagination (#3)

[Browse files](#)

- * Can pull existing repos and add date to push commit
- * Can scroll through pages in GitHub API

🔗 master (#3)

 jfkisel authored and konzy committed on Sep 8, 2017 1 parent 1baf96f commit 3fdfe3dbeb2cc2c541c89a522f7177ce2d3b6628

Showing 1 changed file with 3 additions and 4 deletions. Unified Split

7 clone_all.sh View

	@@ -42,20 +42,19 @@ else
42	42 # Get the first page of repo results (100 entries)
43	43 rawJSON=\$(curl --user "\$githubUsername:\$githubPassword" "https://api.github.com/orgs/\$organization/repos?per_page=100")
44	44 # Get the line that tells if this is the last page
45	- linkTag=\$(echo "\$rawJSON" grep "< Link: <https:")
45	+ numRepos=\$(echo "\$rawJSON" grep -o "full_name" wc -l)
46	46 page=2
47	47
48	48 # While we have not seen the last page
49	- while [[\$linkTag == *">; rel=\"last\"*]]; do
49	+ while [["\$numRepos" -eq "100"]]; do



Why teach Git and GitHub in class?

1. Exposure to commonly used tool

- Kaggle industry-wide survey: 58.4% of the 16,000 respondents said that Git was the main tool used for sharing code in their workplace



Why teach Git and GitHub in class?

2. Group projects

- Students get to practice common industry work sharing pipelines
- Much easier than Google docs for updating and sharing code
- Teachers see contribution of each student



Why teach Git and GitHub in class?

3. Providing feedback during assignments

- Can give feedback directly through GitHub, rather than emailing the file back and forth
- Teachers can easily grab all pieces needed to run the student's code

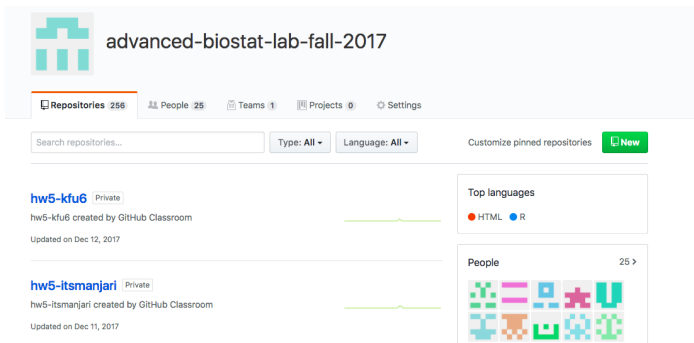


GitHub Classroom allows for integration of Git and GitHub into a course

- <https://classroom.github.com/>
- Can organize starter code for assignments
- Automated assignment (repository) creation for students
 - All students start off with same scripts, data, etc...



GitHub Classroom allows for integration of Git and GitHub into a course



The screenshot shows the GitHub Classroom interface for a repository named "advanced-biostat-lab-fall-2017". The repository is owned by a user with a teal logo. The interface includes a navigation bar with "Repositories 256", "People 25", "Teams 1", "Projects 0", and "Settings". Below the navigation bar is a search bar and filters for "Type: All" and "Language: All". The main content area displays two pinned repositories: "hw5-kfu6" and "hw5-itsmanjari", both created by GitHub Classroom and updated in December 2017. To the right, there are sections for "Top languages" (HTML and R) and "People" (25 >).

advanced-biostat-lab-fall-2017

Repositories 256 People 25 Teams 1 Projects 0 Settings

Search repositories... Type: All Language: All Customize pinned repositories **New**

hw5-kfu6 Private
hw5-kfu6 created by GitHub Classroom
Updated on Dec 12, 2017

hw5-itsmanjari Private
hw5-itsmanjari created by GitHub Classroom
Updated on Dec 11, 2017

Top languages
HTML R

People 25 >



GitHub Classroom can be introduced in both introductory and advanced courses

- In the Fall 2017 semester, we introduced GitHub Classroom in:
 - An introduction to statistics lab at JHU—students had little-to-no previous computing or statistics experience
 - Taught by: Leah Jager and Margaret Taub
 - Teaching assistant: Jacob Fiksel
 - ~ 20 students
 - An advanced computational statistics course at Pomona College—students had experience with statistical computing, but not Git or GitHub
 - Taught by: Johanna Hardin
 - ~ 70 students



Lessons from using GitHub Classroom in an introductory course

- Challenge: introduce Git and GitHub to students with no computational experience
 - Git and GitHub “bootcamp” before first class
 - High start-up cost
- Student feedback
 - “I did not enjoy setting [Git and GitHub] up and the instructions seemed really tedious”
 - “I did not like that [Git] was quite hard to install and that issues occur when you [update] your computer”



Lessons from using GitHub Classroom in an introductory course

- Students appreciated learning version control (eventually)
 - “I really enjoyed using GitHub because it’s applicable to the other things that I do in my lab”
 - “I liked that we can save our past work on GitHub and that it taught me how to take advantage of it for other projects”



Recommendation: tailor use of GitHub Classroom to student background

- For advanced courses: use GitHub Classroom to start and have students work on team projects
- For introductory courses: introduce Git and GitHub as its own unit after students have gotten more comfortable with computing



Resources for implementing GitHub Classroom

- We have created a detailed guide for teachers to start their own GitHub Classroom
 - <https://github.com/jfiksel/github-classroom-for-teachers>
 - Step-by-step instructions on everything from assignment creation to grading workflow
- We also have a guide and videos for beginner students who have never used Git or GitHub
 - <https://github.com/jfiksel/github-classroom-for-students>
 - Covers set-up on both Windows and Apple
- Questions, comments, and collaborations on these guides are more than welcome!

