Collaborating on Data Projects using GitHub

HBS workshop led by

Ista Zahn

Roadmap

Session is informal, example-based, hands-on

- I will introduce a topic or demonstrate a technique
- We will discuss or practice as a group

Topics may include:

- Discussion of factors that make collaboration easier
- Version control as a collaborative practice
- Using pull requests to facilitate project contributions
- Issue tracking and project planning
- Creating project documentation with GitHub pages
- Project organization and structure

Accounts and Software

- Sign-up / into https://github.com and https://code.harvard.edu
- Install GitHub Desktop from https://desktop.github.com/
- Install VSCode: https://code.visualstudio.com/download
- Navigate to the class repository at https://github.com/izahn/demo

Collaboration basics

- Use established conventions and workflows
- Set up project management – issue tracking, task assignment etc.
- Keep project communication with the project itself
- Establish a process for reviewing and accepting contributions
- Write both technical and non-technical documentation
- Set up and use a shared development environment

GitHub: The most popular software development platform in the world

- Collaborate via issues, pull requests, discussions
- Easy web presence for your projects
- Creates a record of the work you do
- Track and compare changes to project files over time
- Automate with continuous integration, bots
- Document your projects with simple web hosting

You can do a lot with GitHub

- **Create a repository**
- Add / Edit files and track changes
- Fork an existing repository
- Make a documentation website for your project
- Open / close / track issues
- Manage settings and collaborators
- Submit / approve / deny pull requests

And more! https://lab.github.com/ is a good place to learn all about it.

Create a new repository

- Sign in to https://github.com
- Click the + sign in the upper right and select "New repository"
- Name it whatever you like
- Select "Public"
- Select "Initialize with a README"
- Click the "Create project" button
- Click the "README" link and add something to the README file.

You can do a lot with GitHub

- Create a repository
- **Add / Edit files and track changes**
- Fork an existing repository
- Make a documentation website for your project
- Open / close / track issues
- Manage settings and collaborators
- Submit / approve / deny pull requests

And more! https://lab.github.com/ is a good place to learn all about it.

Editing files on github.com

You can add, edit and commit using the web interface. In your new repository:

- Click the README file then the edit icon
- Make some changes, write a commit message, commit

- Click "Add file" => Upload, message, commit
- You can see how your project has changed over time.

- Click the "clock" icon
- Select a commit to view the changes made

Try it out!

Take some time to explore the github.com website, trying out things we've discussed or other things you discover.

You can do a lot with the GitHub web interface

- Create a repository
- Add / Edit files and track changes
- **Fork an existing repository**
- Make a documentation website for your project
- Open / close / track issues
- Manage settings and collaborators
- Submit / approve / deny pull requests

And more! https://lab.github.com/ is a good place to learn all about it.

Forking and pull requests

GitHub has powerful collaboration features, including the famous "pull request", AKA a "PR".

- Navigate to https://github.com/izahn/demo and click the "Fork" icon in the upper right!
- Make some changes in your fork
- Click the "Pull requests" icon and select "new pull request" then "Create pull request"
- Add a comment and click "Create pull request"

You can do a lot with the GitHub web interface

- Create a repository
- Add / Edit files and track changes
- Fork an existing repository
- **Make a documentation website for your project**
- Open / close / track issues
- Manage settings and collaborators
- Submit / approve / deny pull requests

And more! https://lab.github.com/ is a good place to learn all about it.

Websites with GitHub page

You can easily turn your repository into a web page:

- Click "settings"
- Scroll down to "GitHub Pages" and select "master branch"
- Copy the URL and navigate to it

You can do a lot with Github

- Create a repository
- Add / Edit files and track changes
- Fork an existing repository
- Make a documentation website for your project
- **Open / close / track issues**
- Manage settings and collaborators
- Submit / approve / deny pull requests

And more! https://lab.github.com/ is a good place to learn all about it.

Create and manage issues

- Navigate to https://github.com/izahn/demo
- Click the "Issues" button and create a new issue
- Write a message requesting to be added as a collaborator
- Click the "Submit New Issue" button

You can do a lot with GitHub

- Create a repository
- Add / Edit files and track changes
- Fork an existing repository
- Make a documentation website for your project
- Open / close / track issues
- **Manage settings and collaborators**
- Submit / approve / deny pull requests

And more! https://lab.github.com/ is a good place to learn all about it.

Collaborators and settings

GitHub has powerful collaboration features, including the ability to add collaborators and assign issues.

- Navigate to your fork of https://github.com/izahn/demo and click the "Settings" icon in the upper right
- Click "Manage access" in the menu on the left, then click the green "Invite a collaborator" button
- Invite me ("izahn") as a collaborator on your project

You can do a lot with GitHub

- Create a repository
- Add / Edit files and track changes
- Fork an existing repository
- Make a documentation website for your project
- Open / close / track issues
- Manage settings and collaborators
- **Submit / approve / deny pull requests**

And more! https://lab.github.com/ is a good place to learn all about it.

Branch and create pull requests

Branches are used to work on adding capabilties or fix bugs without interfering with other work being done.

- Navigate to https://github.com/izahn/demo/branches and click the "New Branch" button in the upper right
- Name your branch "thank-<username>", replacing "<username>" with your GitHub username
- Edit the "thanks.txt" file and add your name to it
- Create a pull request for this change

Break / Free Play!

Suggestions:

Search GitHub for a repository you might be interested in and explore it. Examine the pull request history and issue tracker to develop an understanding of how the project is managed.

Explore other features we've covered or new things you've discovered.

Configuring and using GitHub Desktop

- Install GitHub Desktop from https://desktop.github.com/
- Start GitHub Desktop on your local computer
- Sign in to Github
- Clone your fork of https://github.com/izahn/demo
- Make changes locally (e.g., using your favorite editor or IDE) and commit using GitHub Desktop
- Push changes to GitHub using GitHub Desktop

Configuring and using GIT from your IDE
(RStudio as an example)

- Start RStudio on your local computer
- Create a new RStudio project with version control enabled
  - https://happygitwithr.com/rstudio-git-github.html#clone-the-new-github-repository-to-your-computer-via-rstudio
- Make changes locally and commit
  - https://happygitwithr.com/rstudio-git-github.html#make-local-changes-save-commit
- Push changes to GitHub (using RStudio or Git from a terminal)
  - https://happygitwithr.com/rstudio-git-github.html#make-local-changes-save-commit

Interesting public data projects

- https://github.com/ImperialCollegeLondon/covid19model
- https://github.com/gcampede/terrorism-metagraphs
- https://github.com/nychealth/coronavirus-data
- https://github.com/github/covid19-dashboard
- https://github.com/altsalt/rms-letter-comparison