Introduction to GitHub & integration with R for project management

Kayla I Perry, PhD

Assistant Professor

The Ohio State University

Workshop Outline & Goals

Introduction to GitHub

Integrate R & Git

Link with a repository on GitHub

Create your own repository on GitHub

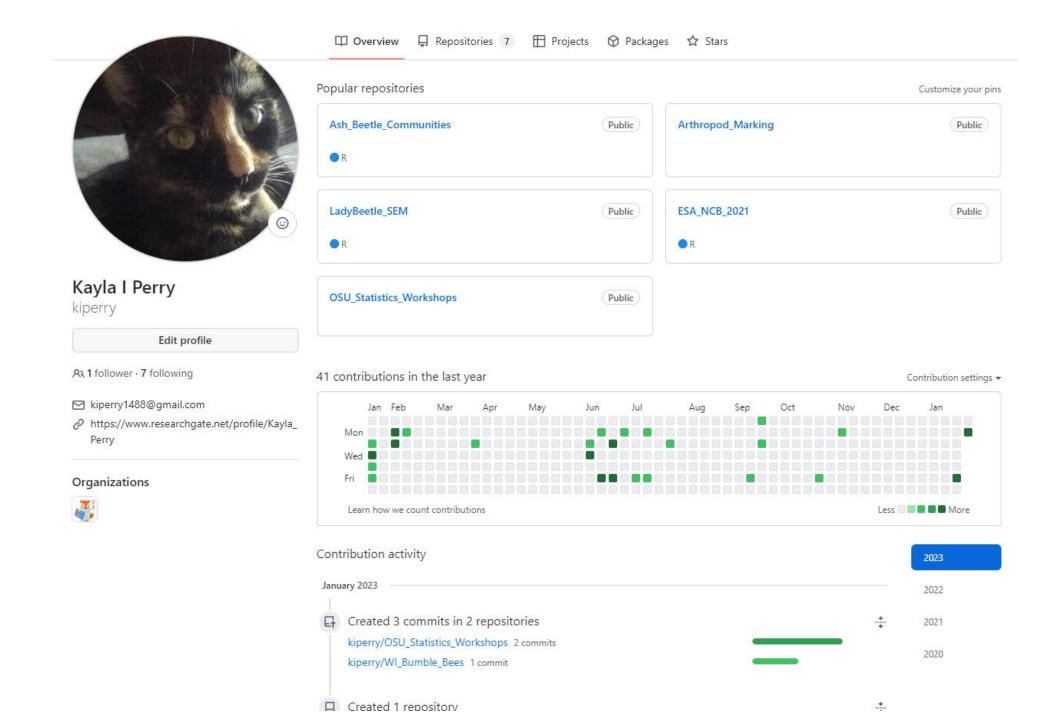
Git and GitHub – What are they?

Git: Software that handles version control on your repository

Working in the background when using GitHub

GitHub: Web interface that hosts your repository online

- Allows for collaboration on projects
- Interfaces with R/RStudio & Git



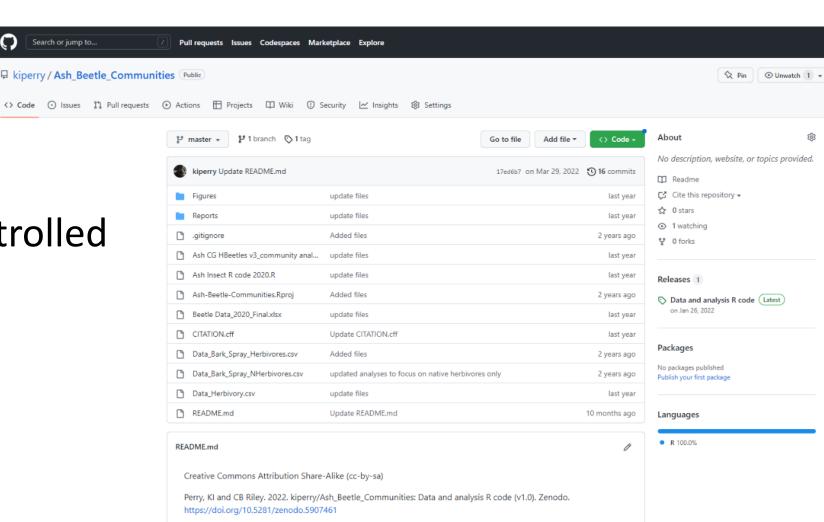
What is a repository (or repo)?

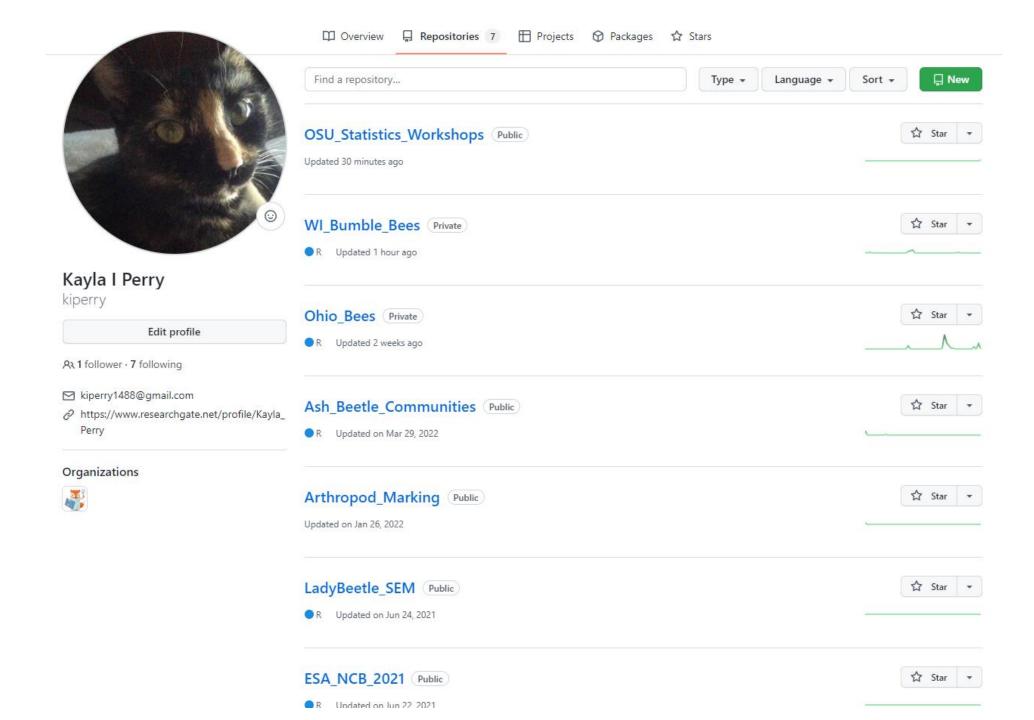
Place for all files associated with a project

With GitHub, your repo lives on your computer and online

Each file is version controlled with documented development history

Public or private





Make changes or updates to repo with commit

Save a version of a file, and provide notes on what you changed

When you commit a file in Git/GitHub, you are saving a new version, but also keeping a record of the changes you made 🕿 Commit changes

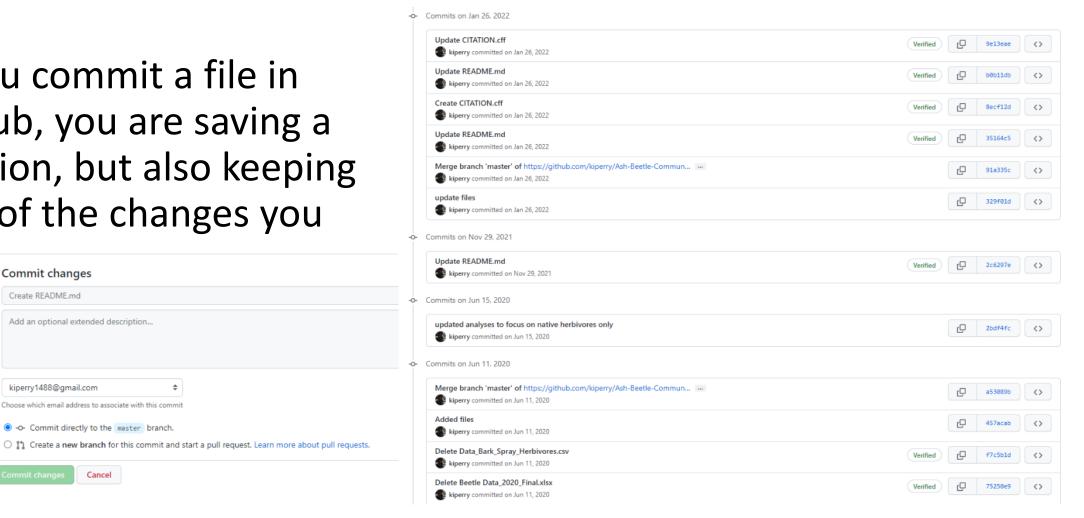
Create README.md

kiperry1488@gmail.com

Add an optional extended description...

Choose which email address to associate with this commit

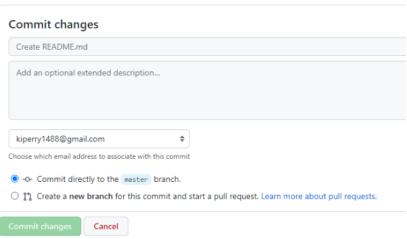
O- Commit directly to the master branch.

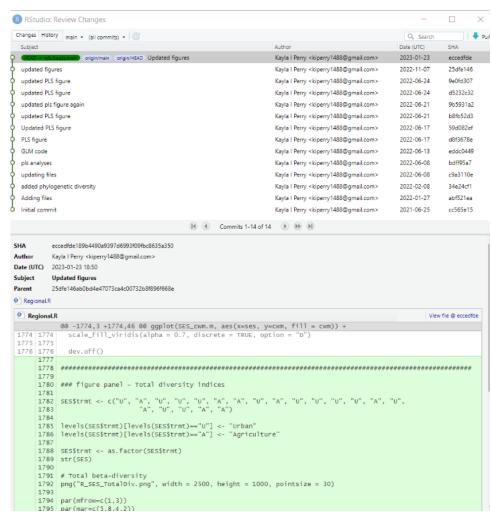


Make changes or updates to repo with commit

Save a version of a file, and provide notes on what you changed

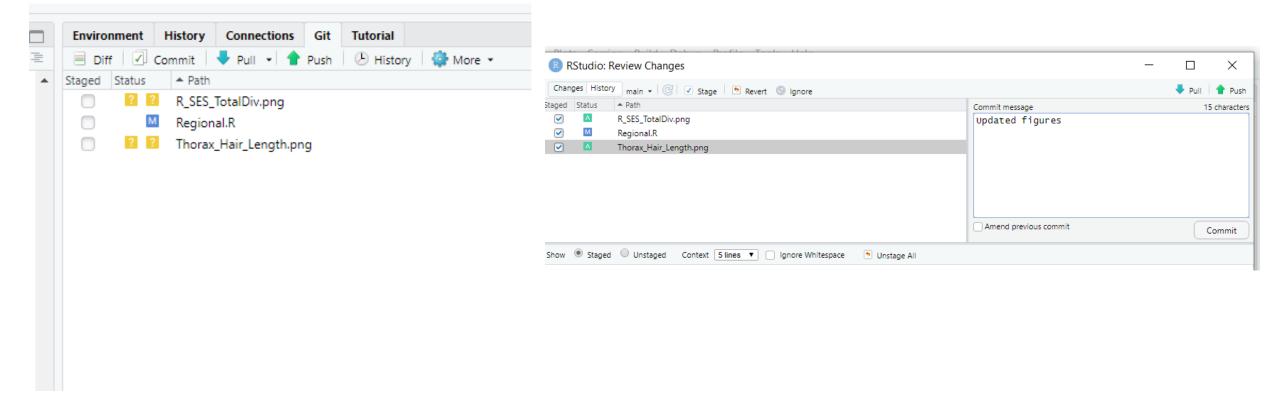
When you commit a file in Git/GitHub, you are saving a new version, but also keeping a record of the changes you made





Pull, Commit, then Push

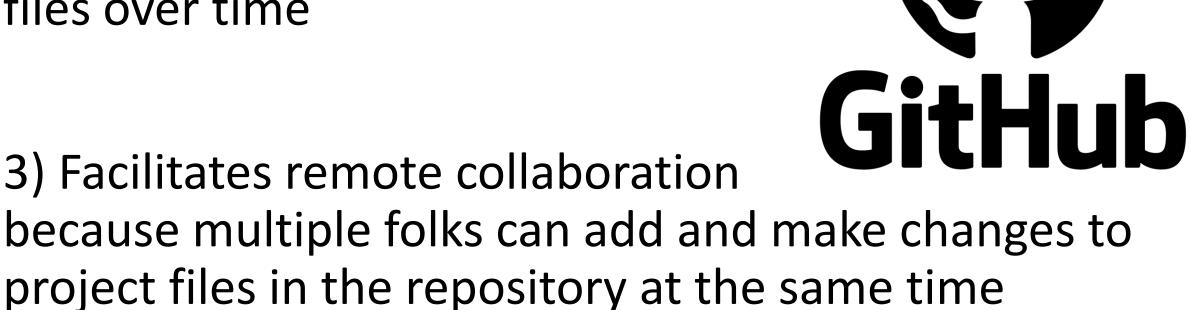
- 1. Pull from the online repository to update your local files
- 2. Commit to save a new version of a file(s)
- 3. **Push** those changes online to the repository



1) Sync project files locally on your computer and online

2) Make commits to record changes to files over time

3) Facilitates remote collaboration

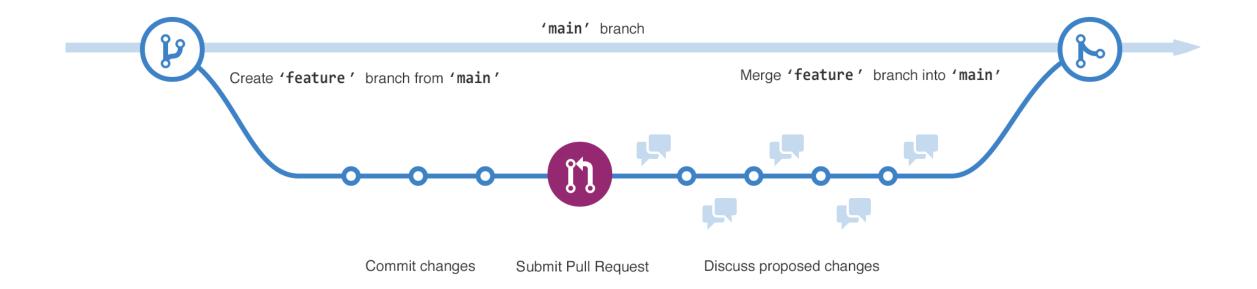


Additional features: Branch / Merge

Parallel version of your repository at a particular point in time

- Allows for experimentation/revision without affecting the original documents of the project
- Can merge the branch back into the project later

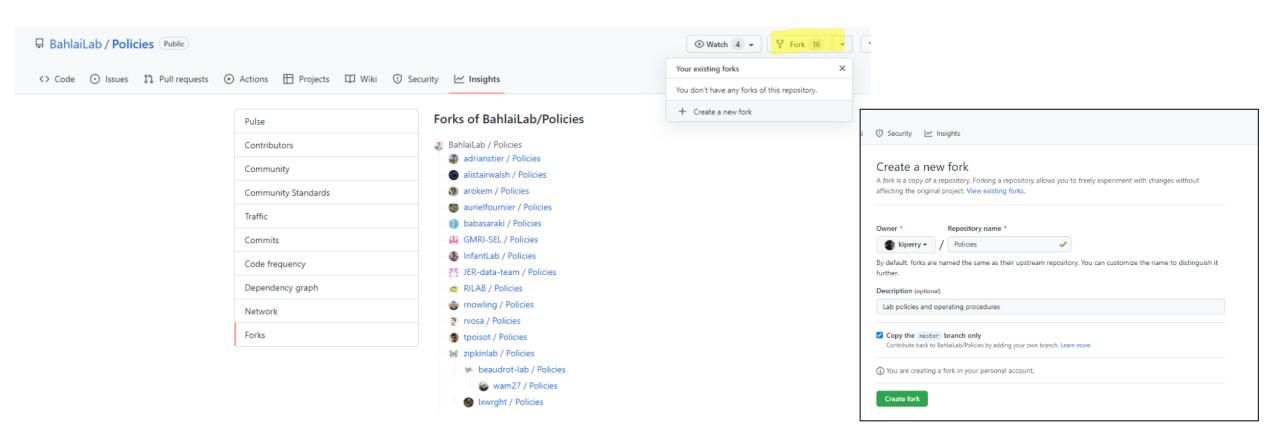
Contained within the repository, but does not impact the main branch



Additional features: Fork

Make a personal copy of a repository

Intention to contribute or improve, and then provide changes at a later point in time



Additional features: Fork

Make a personal copy of a repository

Intention to contribute or improve, and then provide changes at a later point in time

If you are not an official collaborator on the repository, need to submit a pull request

■ If the repository is public, folks have access to your files but cannot change them without permission

Starting point for your own project

Additional features: Clone

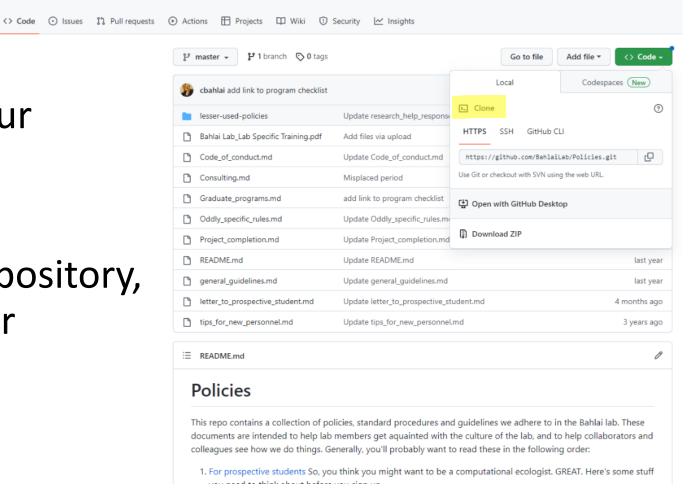
Make a copy of a repository on your computer

☐ BahlaiLab / Policies (Public

Access files locally



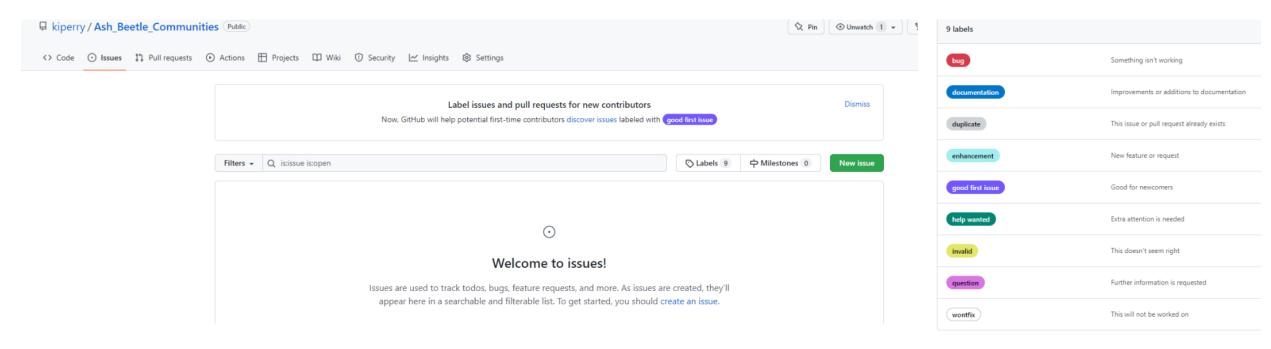
Added as a collaborator to a repository, clone the files to your computer



Additional features: <u>Issues</u>

List tasks associated with the project that need to be accomplished

Label issues, assign them to yourself or collaborators, and track completion



Review – What can we do with Git/GitHub?

1) Experiment on projects without breaking them – Branch

2) Make, assign, and keep track of tasks – Issues

3) Access existing projects made by others – Fork or Clone

4) Build on existing projects with collaborators – **Pull, Commit, Push**

Why use Git/GitHub with R?

Facilitates research transparency and reproducibility

Share data, code, and analyses with collaborators and scientific community

Track development history over time

Aligns with open science journal requirements





Before the workshop, you should have...

Installed R and RStudio

Installed Git

Created an account on GitHub



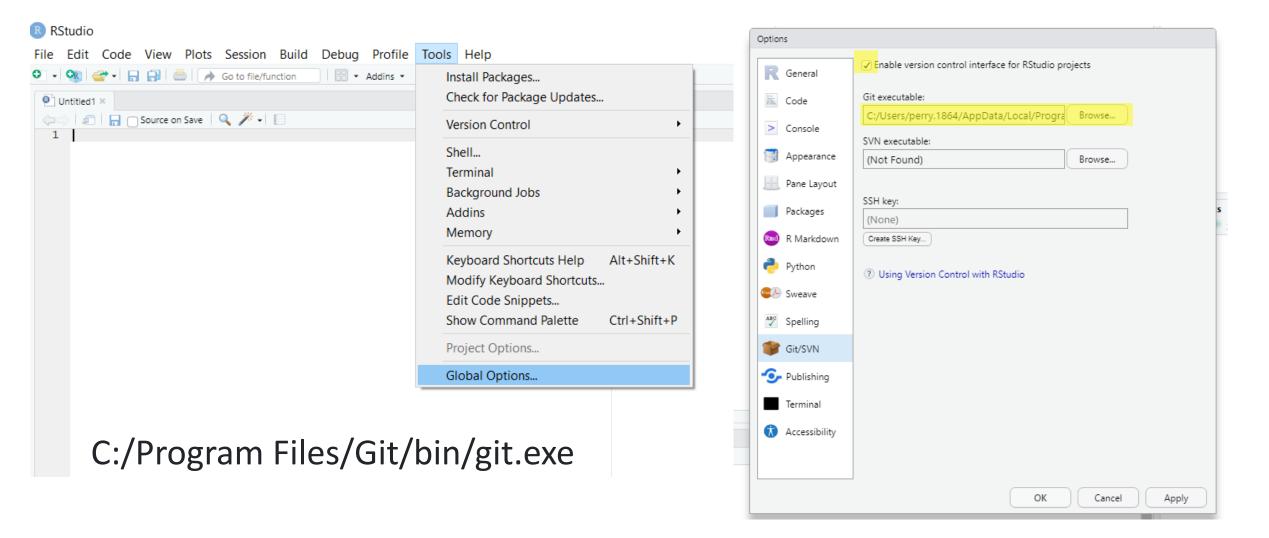


*May also be interested in GitDesktop (https://desktop.github.com/)

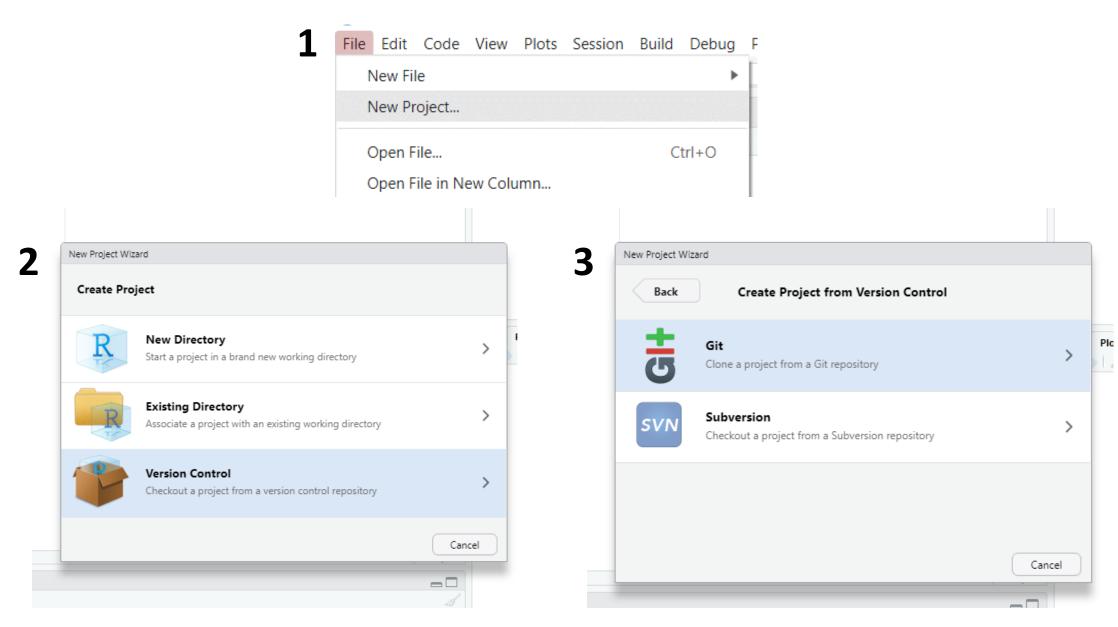
Integrate Git in RStudio

1) Check enable version control

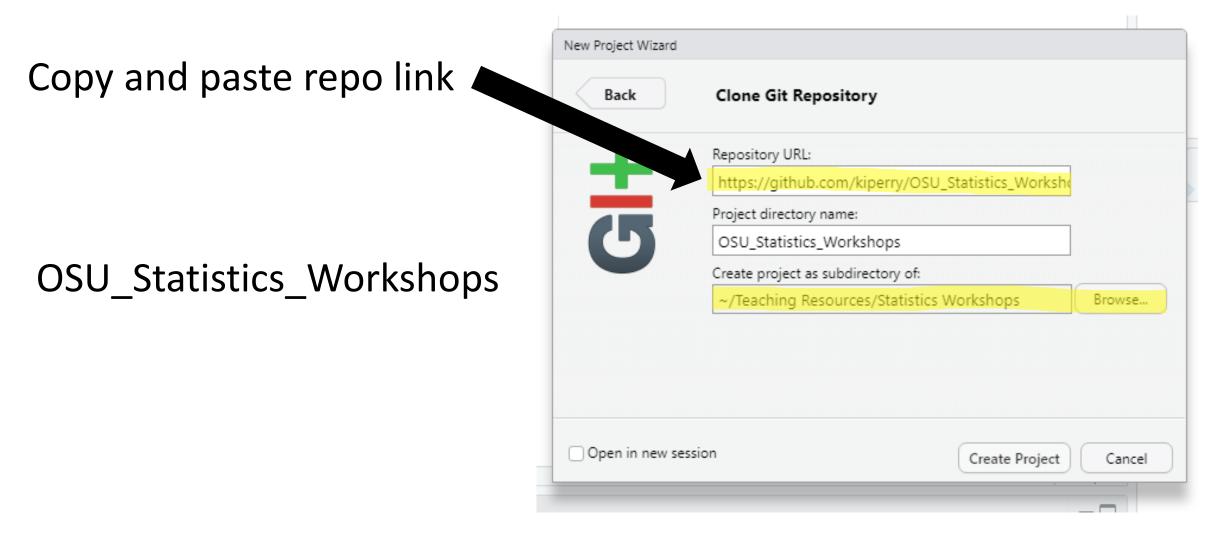
2) Set path to git.exe file



Check to make sure Git and RStudio are linked

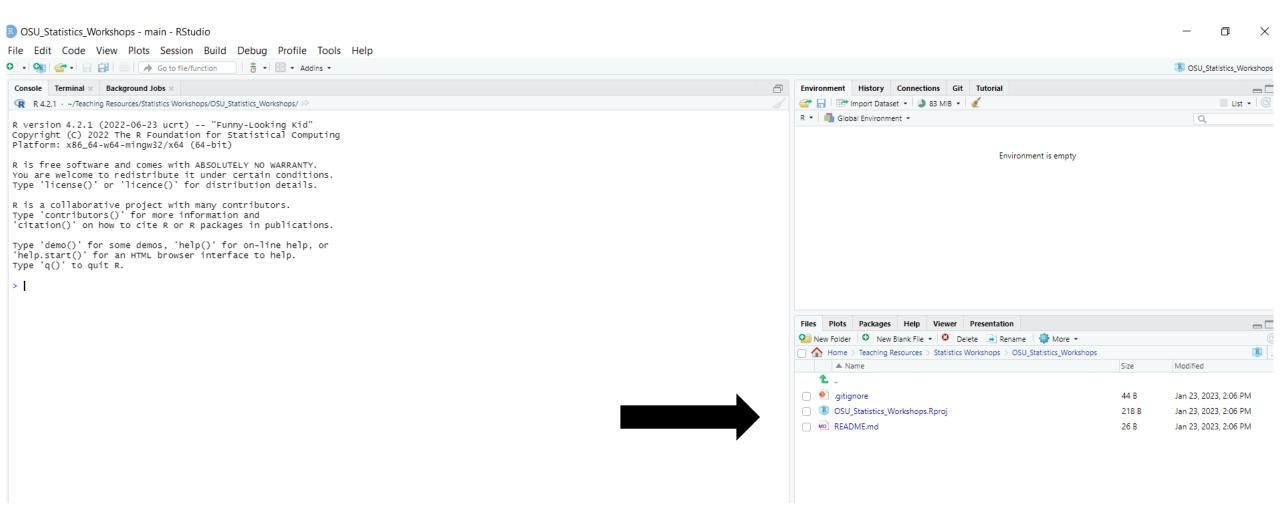


Clone a repository on GitHub to your computer

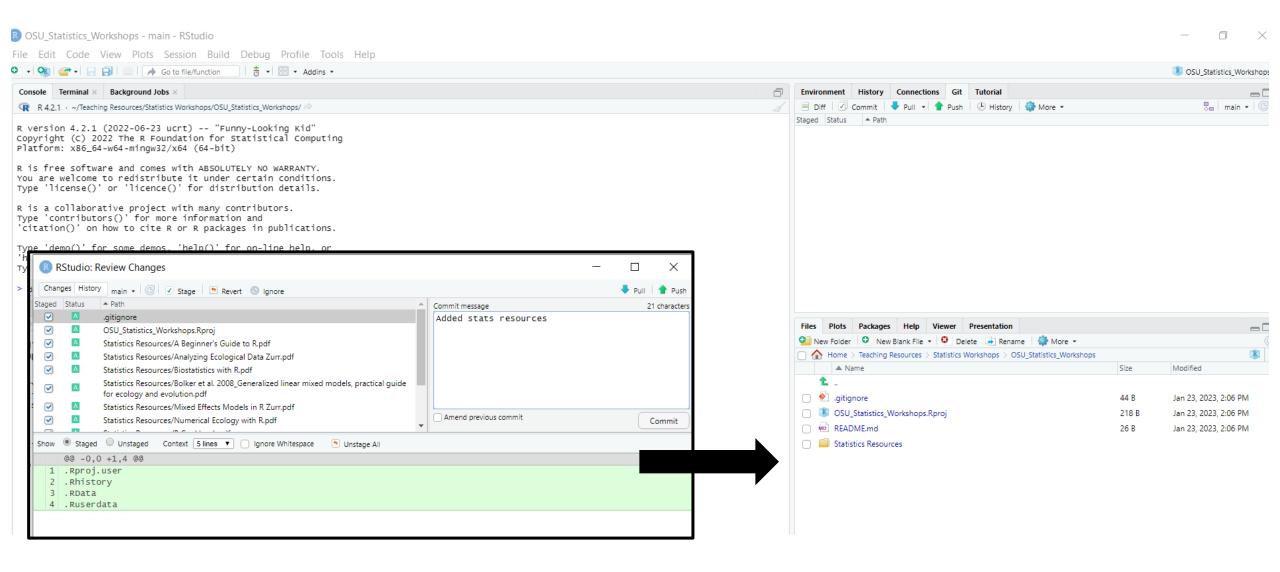


https://github.com/kiperry/OSU_Statistics_Workshops

Clone a repository on GitHub to your computer

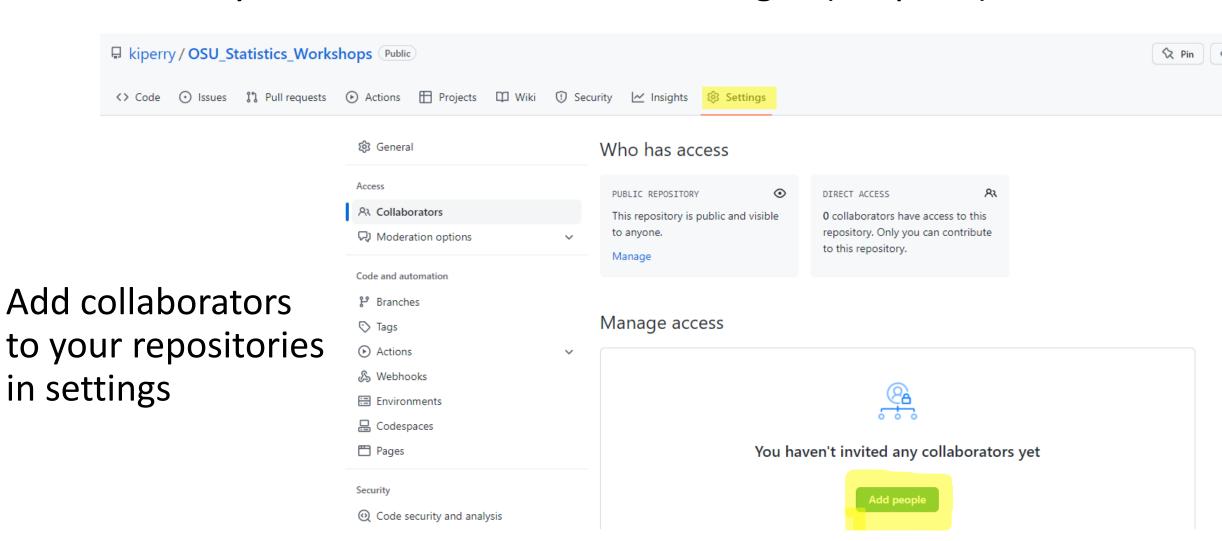


Clone a repository on GitHub to your computer



You can clone any public repositories on GitHub

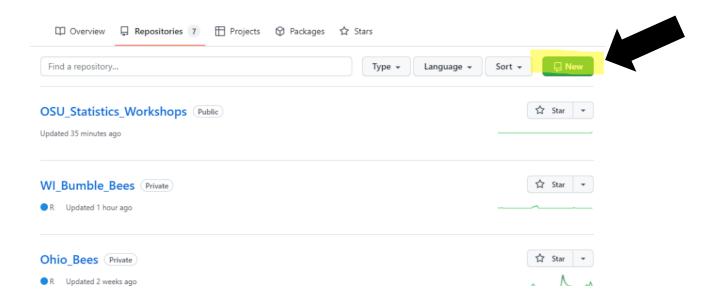
However, only collaborators can make changes (via push)



Create your own repository!

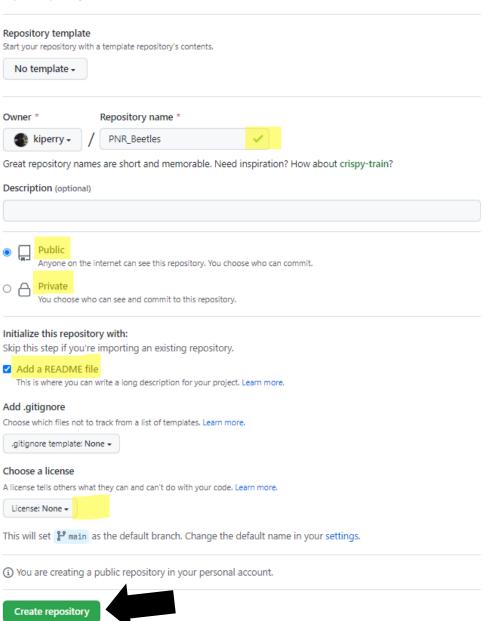
Create your own repository!

1. Log into your GitHub account, go to the Repositories tab, click New

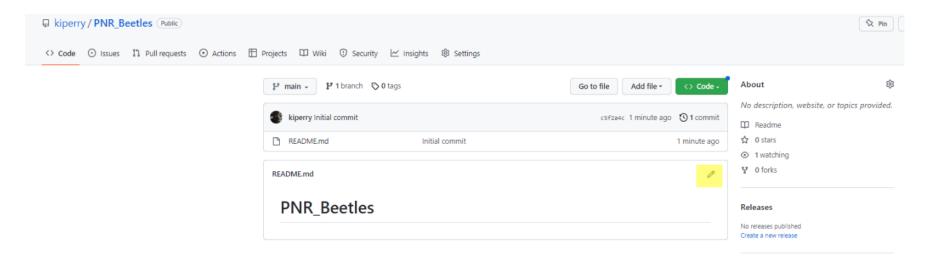


Create a new repository

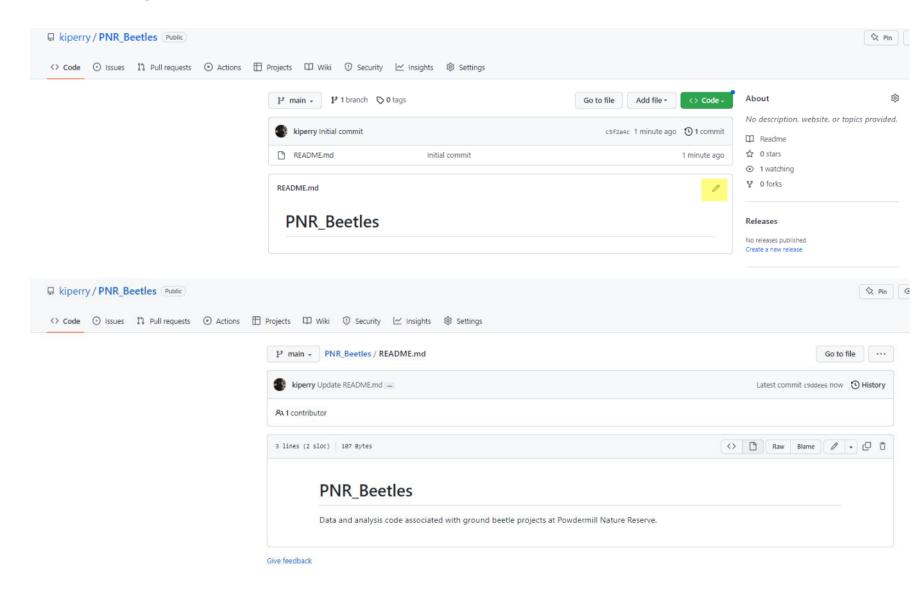
A repository contains all project files, including the revision history. Already have a project repository elsewhere? Import a repository.



Edit your README file

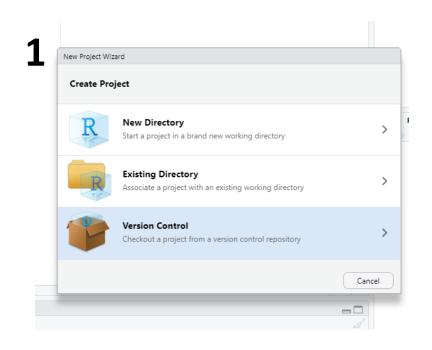


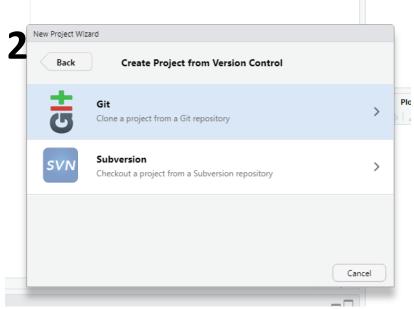
Edit your README file

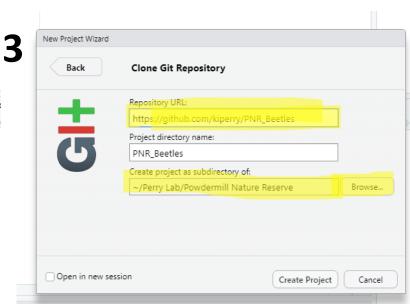


Update readme file with a commit

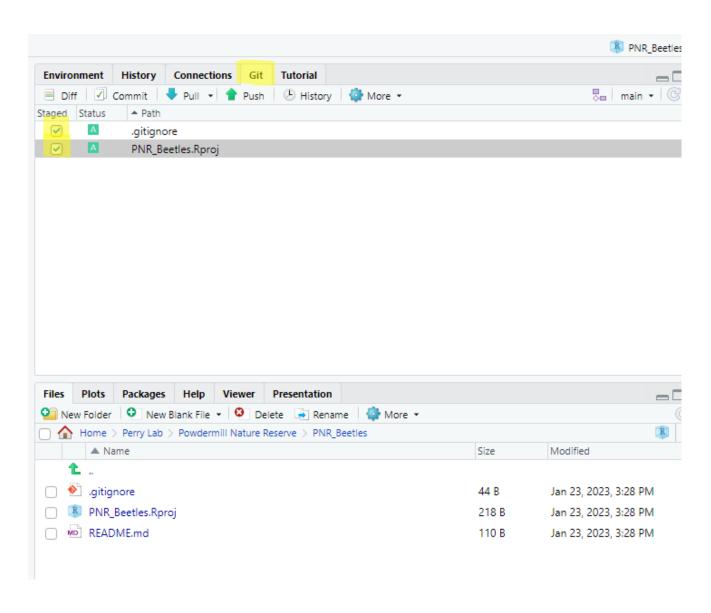
Clone your newly created GitHub repository using R



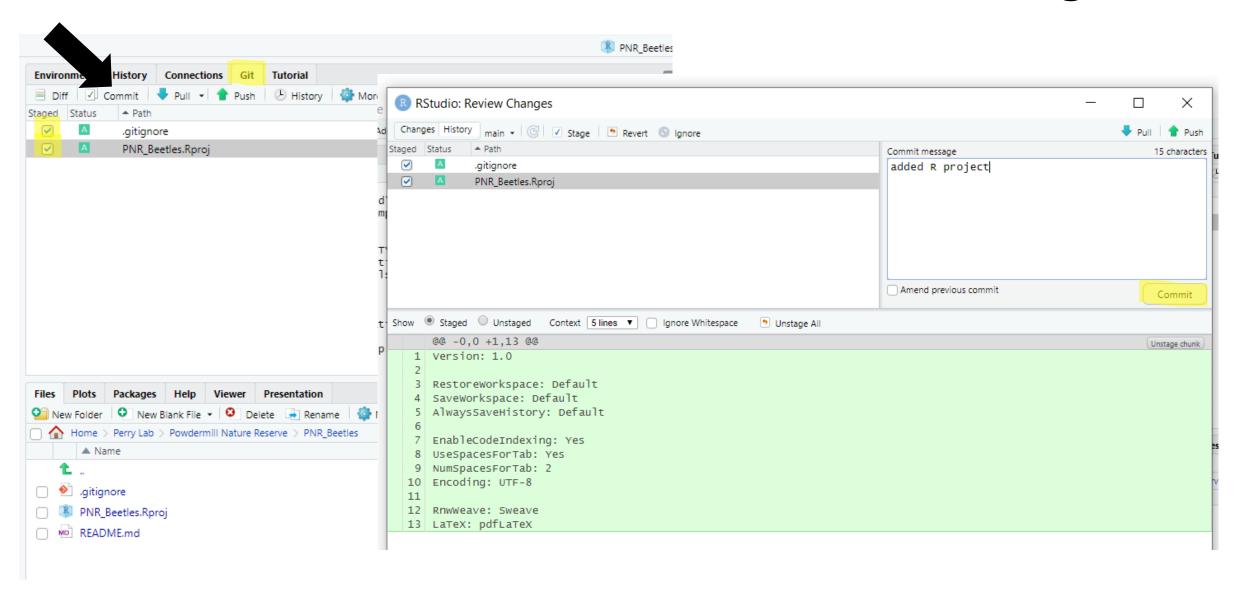




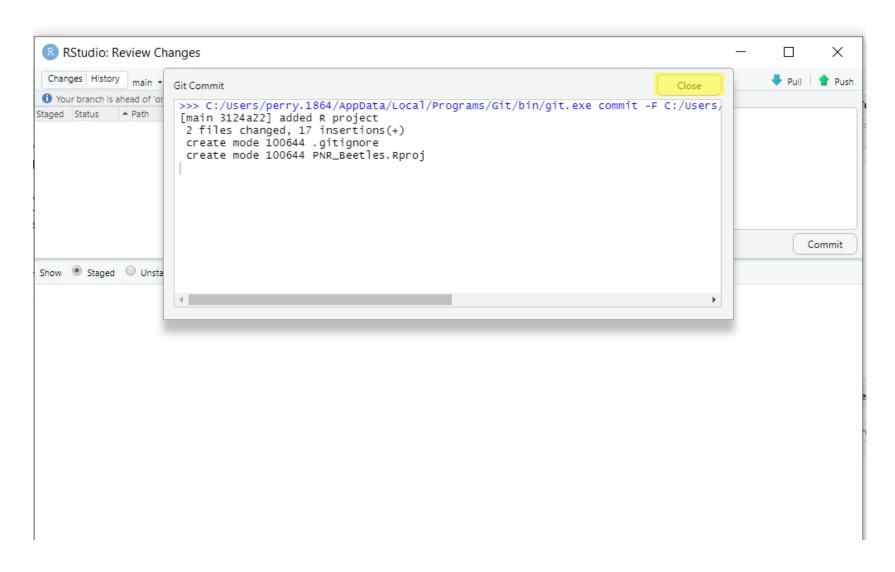
Select Git tab and make a commit to save changes



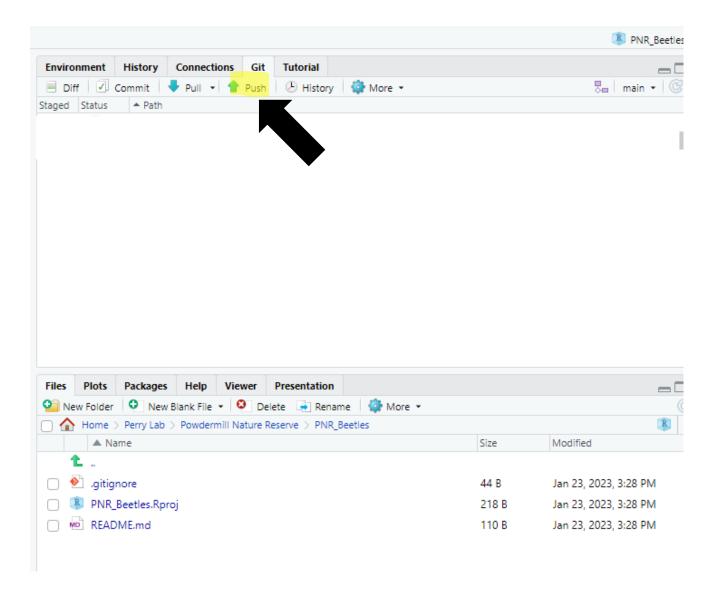
Select Git tab and make a commit to save changes



Select Git tab and make a commit to save changes



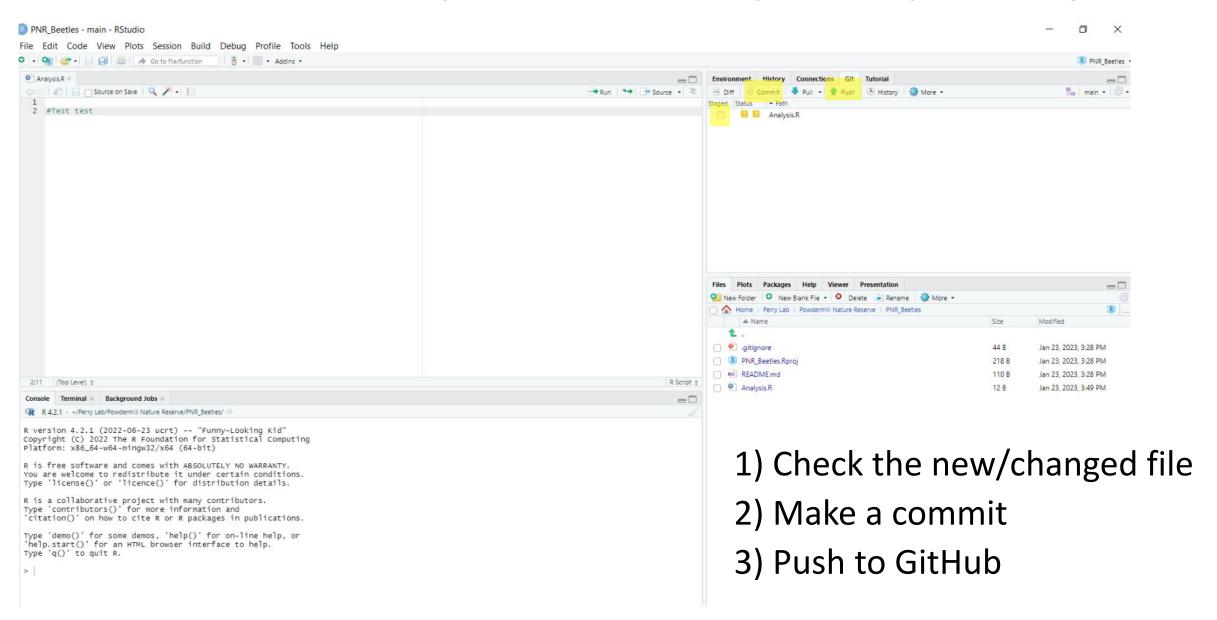
Push changes to your repository on GitHub



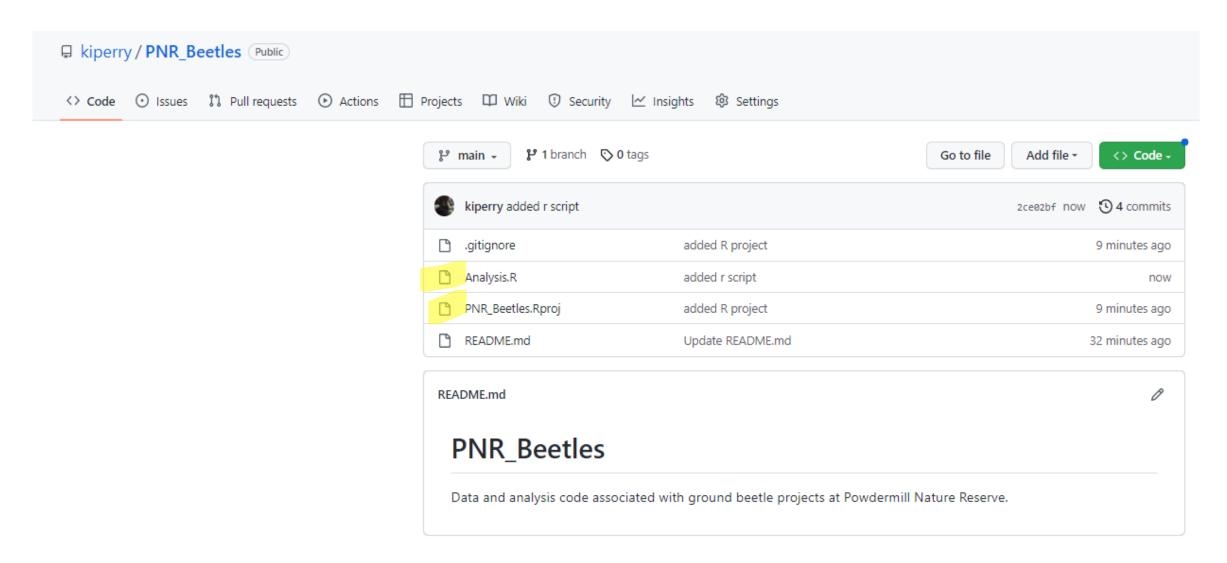
Link your newly created GitHub repository with R



Create a new r script and add to your repository

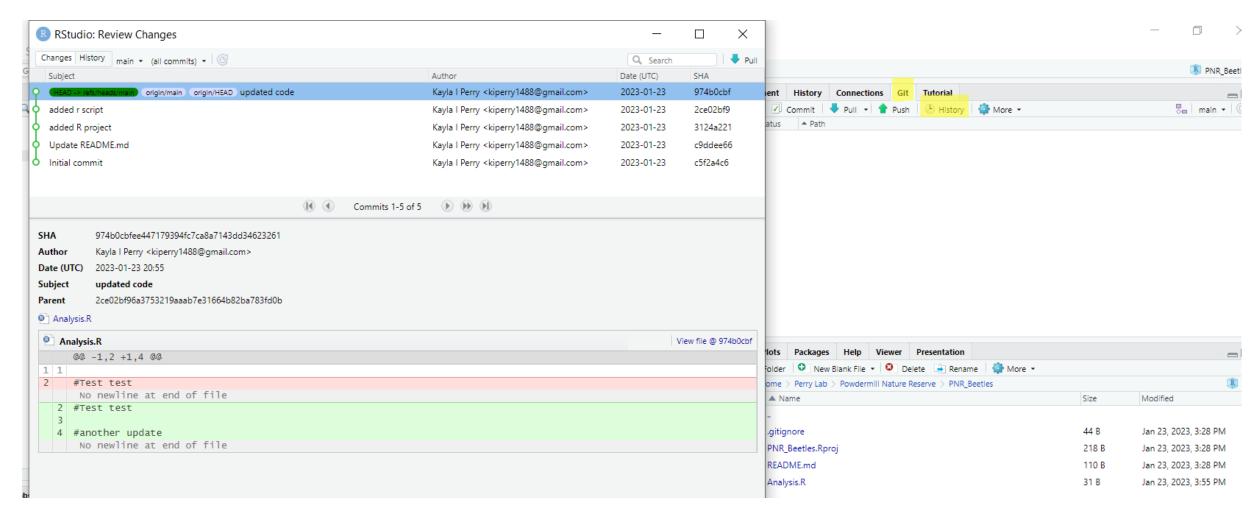


New R files now in GitHub repository



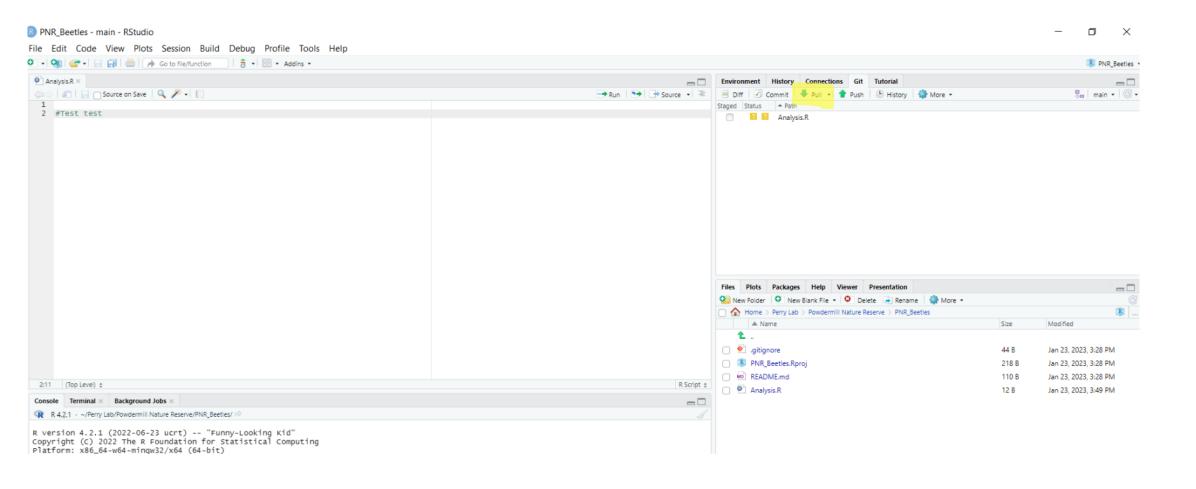
Review changes in your files over time

Version control that documents development history over time



If a collaborator is added to your repository....

Start by pulling changes from the online repository



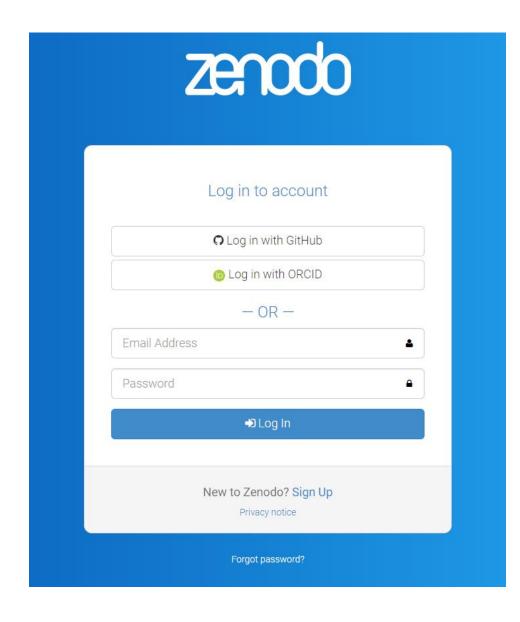
Submitting your manuscript for review?

Include your GitHub repository link in your open research/data availability statement

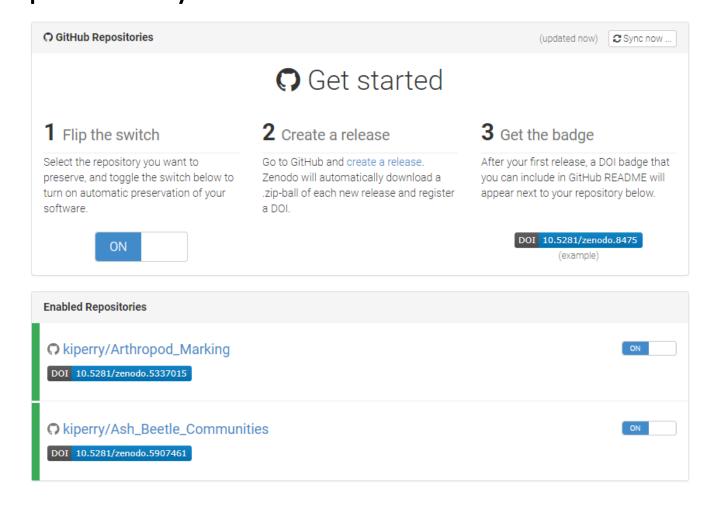
Many journals now require submission of data and code for peer-review

- Open Research Statement: Data are already published and publicly available, with those items properly
- 24 cited in this submission. This submission uses novel code, which is provided in an external repository to
- be evaluated during the peer review process and are available at
- 26 https://github.com/BahlaiLab/Ohio ladybeetles. If this paper is accepted for publication, data and code
- will be permanently archived in Zenodo.

Link GitHub repository with Zenodo for DOI



Developed under European OpenAIRE Program
Operated by CERN



Software and Data Products category on your CV!

January 26, 2022 Oper

Study data and analysis code

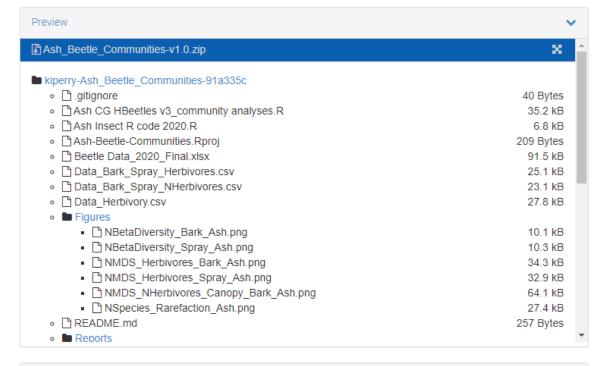
(b) Kayla I Perry; (b) Christopher B Riley

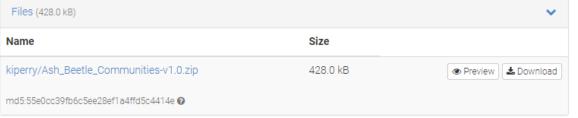
Study data and R code, first release v1.0

This dataset supports the following study:

Perry, KI, CB Riley, F Fan, J Radl, DA Herms, and MM Gardiner. The value of hybrid and nonnative ash for the conservation of ash specialists is limited following late stages of emerald ash borer invasion, Agricultural and Forest Entomology, doi.org/10.1111/afe.12499

Creative Commons Attribution Share-Alike (cc-by-sa)





New version

Edit

44

views

♣ downloads

See more details...



Publication date:

January 26, 2022

DOI:

DOI 10.5281/zenodo.5907461

Published in:

Agricultural and Forest Entomology; doi.org/10.1111/afe.12499:.

Related identifiers:

Supplement to

https://github.com/kiperry/Ash_Beetle_Communities/tree/v1.0

License (for files):

☐ Creative Commons Attribution Share Alike 4.0 International

Update readme file on GitHub with DOIs

