WATER DATA ANALYTICS

ENVIRON 790.02

Spring 2022

WELCOME TO WATER DATA ANALYSIS

Kateri Salk, PhD (she/her)

Adjunct Assistant Professor of Water Resources

Aquatic Ecologist, Tetra Tech

kateri.salk@duke.edu

Background:

- Limnology, biogeochemistry
- Environmental informatics
- Empirical and process-based modeling

OFFICE HOURS

I will set office hours based on student schedules. Fill out the poll!

https://www.when2meet.com/?13981797-YJeci

- 1 hour of virtual office hours each week
- Before & after class in person

COURSE OBJECTIVES

1. Synthesize information on fundamental and applied topics in water resources using quantitative analysis
2. Apply the appropriate steps of the data analytics pipeline to answer questions about aquatic systems
3. Develop marketable skills in data management, analysis, and communication for the aquatic sciences field

COURSE SCHEDULE

Week 1: Intro and R boot camp

Weeks 2-4: Physical properties of lakes & rivers

Weeks 4-7: Water quality in lakes & rivers

Weeks 8-12: Time series analysis, spatial analysis, high frequency data

Week 13: Final project workshop

Finals week: Project presentations, project report

ASSIGNMENTS

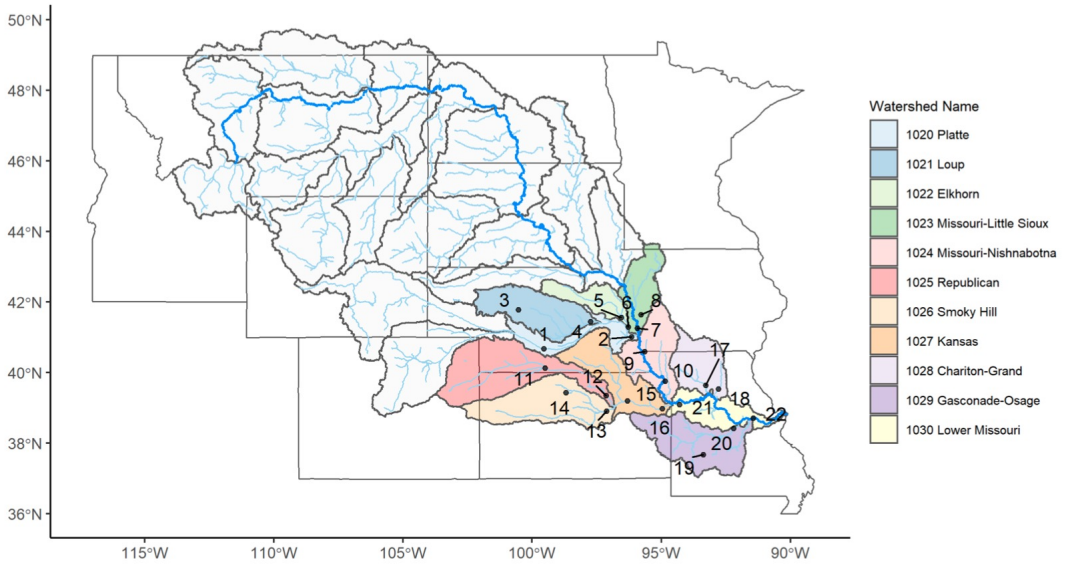
- Each week will have an assignment (homework) (80 % of grade)
- Competency based
- I will provide a key – try to complete the assignment without it
- Following completion, take the corresponding survey on Sakai
- Due dates in syllabus are targets, not deadlines

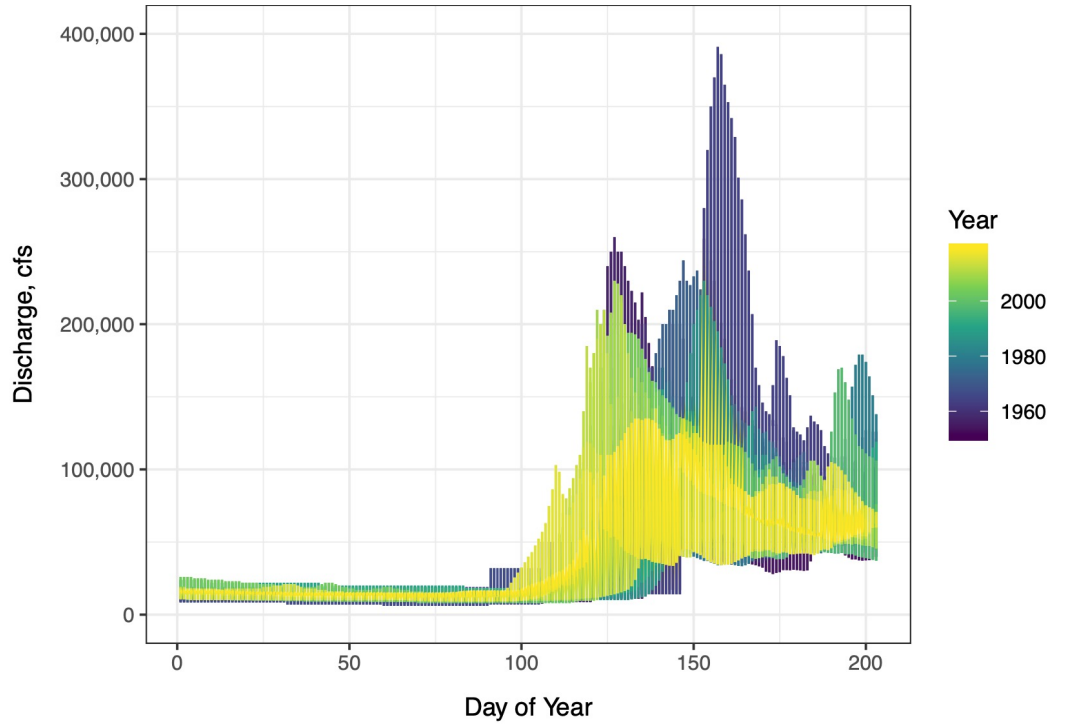
FINAL PROJECT

Choose a water-focused question and perform a quantitative analysis to answer this question

Project stages (% final grade)

- Initial idea (2%)
- Workshop session (5%)
- Final report (10%)
- Final presentation (3%)





OTHER SYLLABUS STUFF

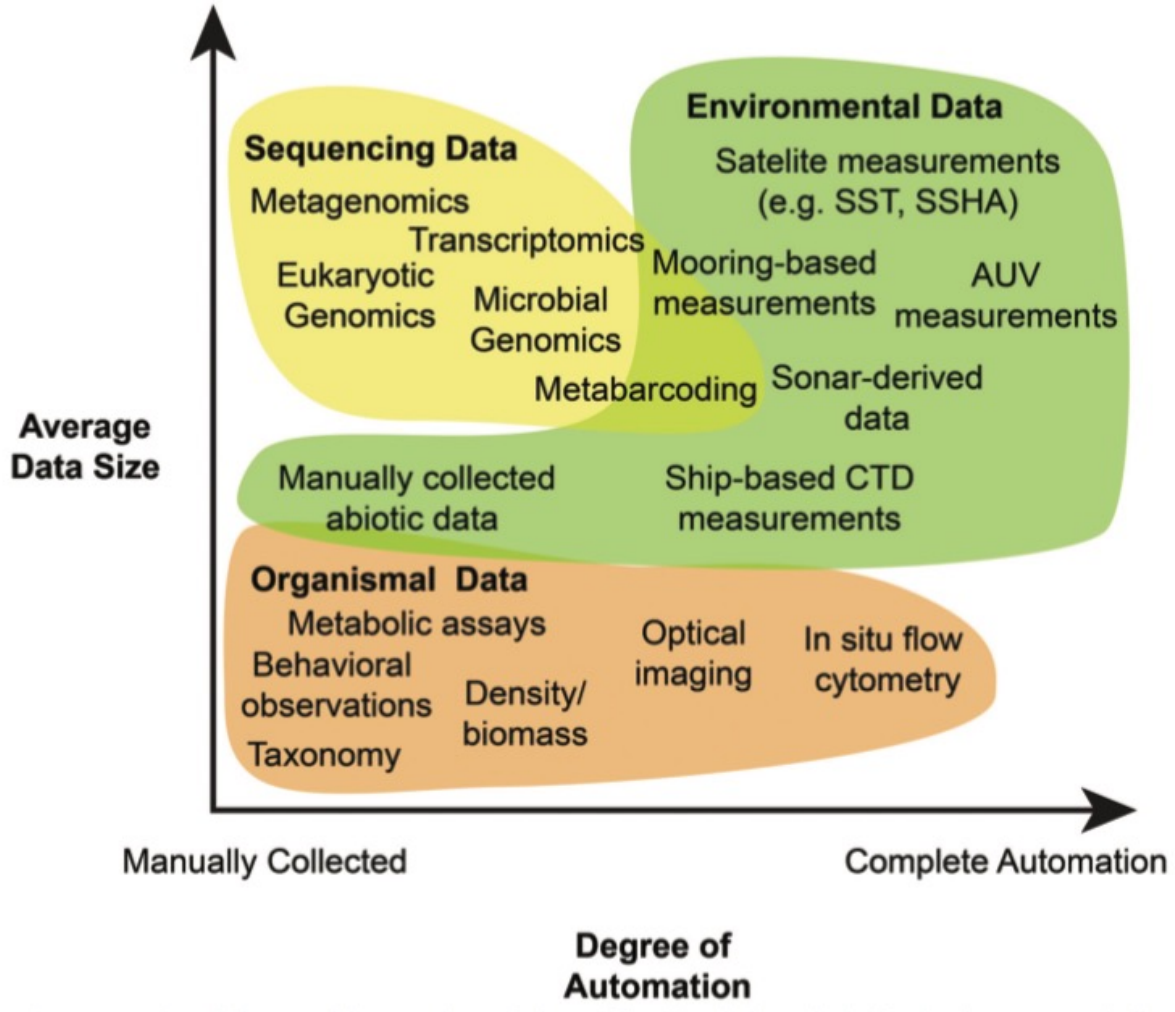
- Schedule specifics
- COVID procedures
- Accommodations

INTEGRATING BIG DATA INTO AQUATIC ECOLOGY

6 key challenges:

1. Recognizing big data
2. Data handling
3. Analytical techniques
4. Verification
5. Data sharing
6. Developing knowledge infrastructure

Durden et al. 2017



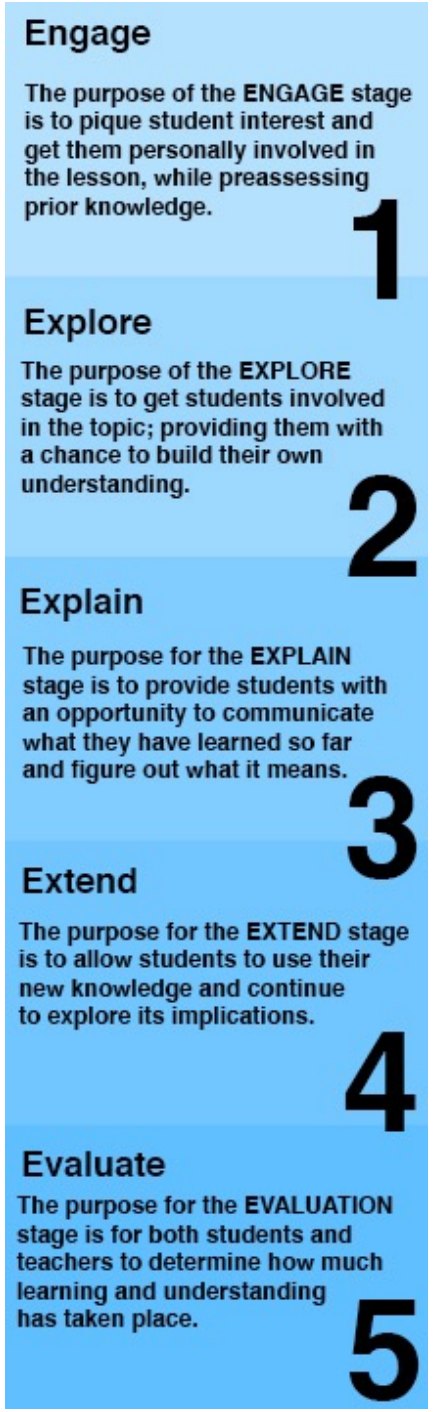
INQUIRY-BASED LEARNING

Construction of knowledge through scientific practices

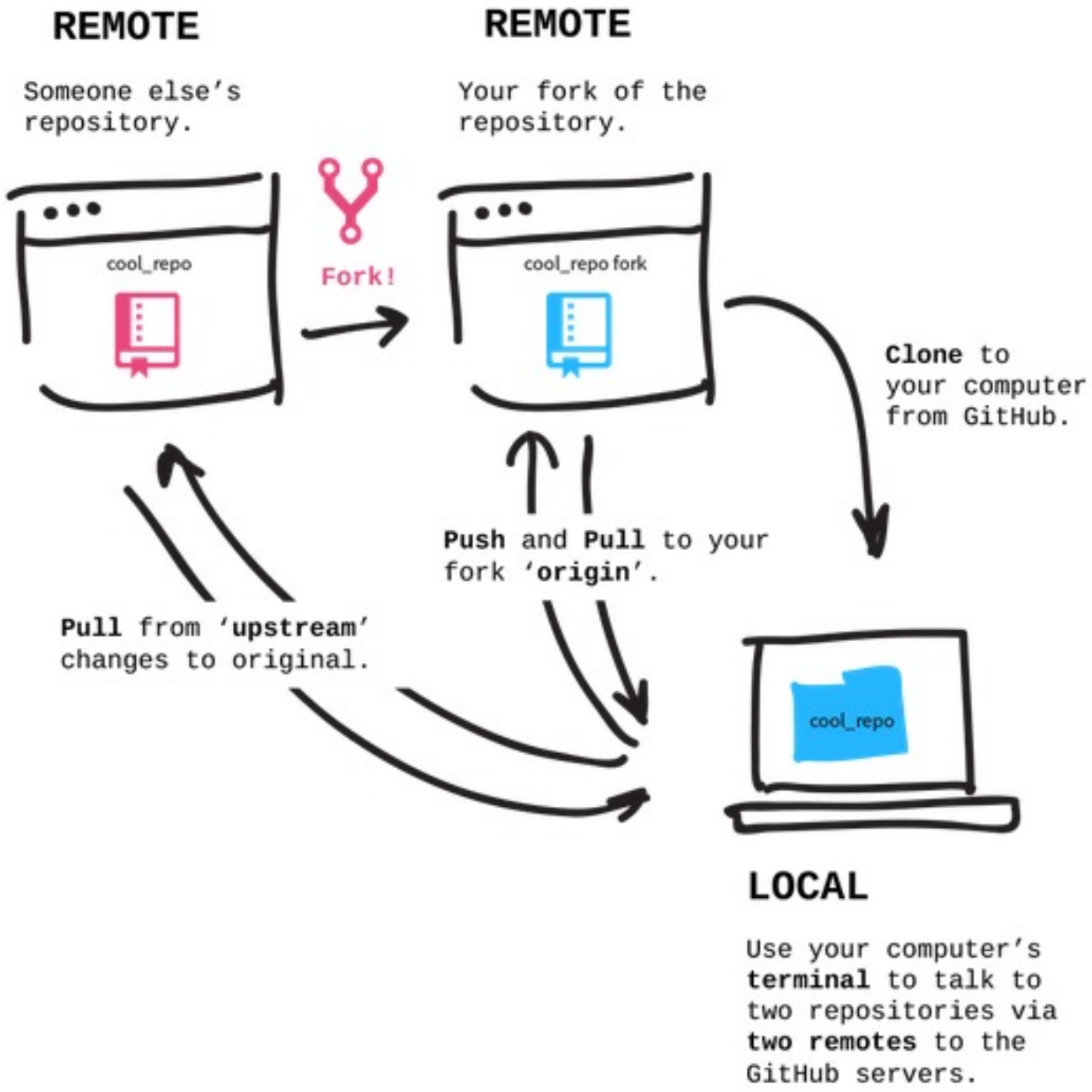
Involves:

- Problem solving skills
- Active participation
- Knowledge discovery by the learner
- Inductive and/or deductive approach

Outcomes: inquiry based learning > traditional instruction

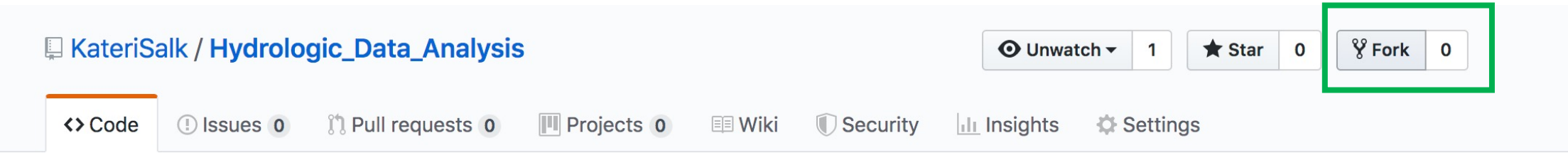


GITHUB SETUP



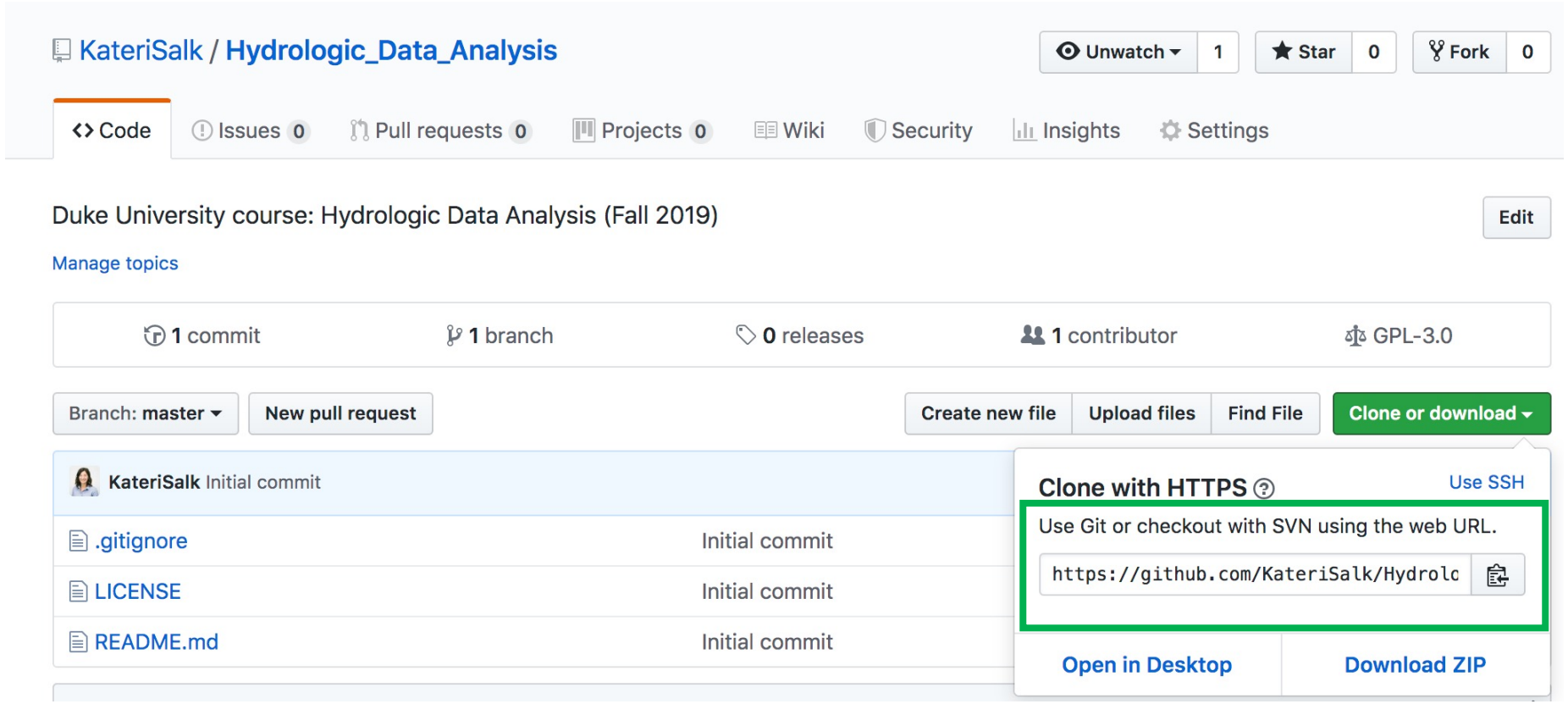
GITHUB SETUP: FORKING

1. Navigate to https://github.com/KateriSalk/Water_Data_Analytics_2022
2. Fork the repository to your GitHub account



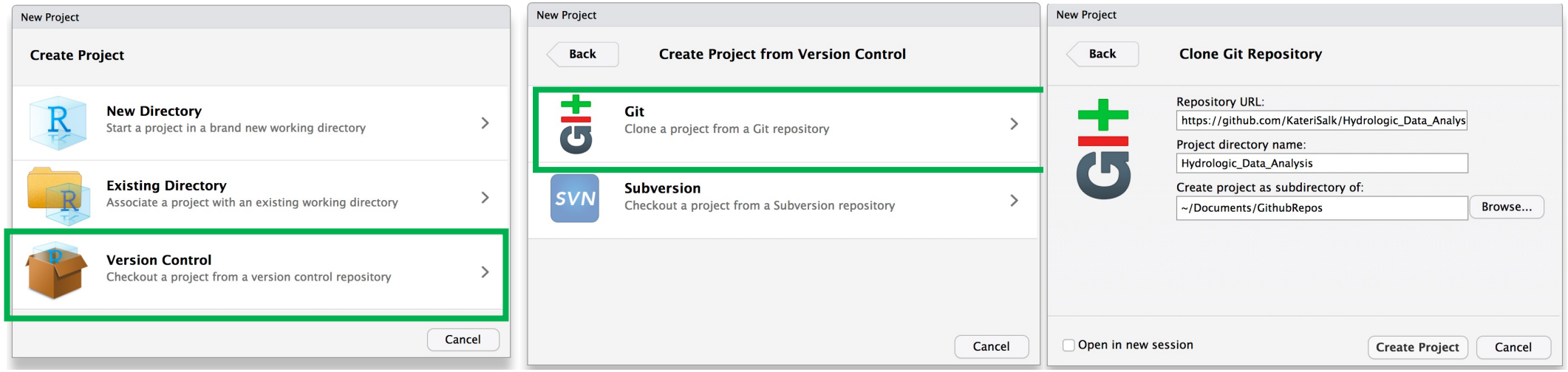
GITHUB SETUP: CLONING

3. Copy the link to your forked repository



GITHUB SETUP: CLONING

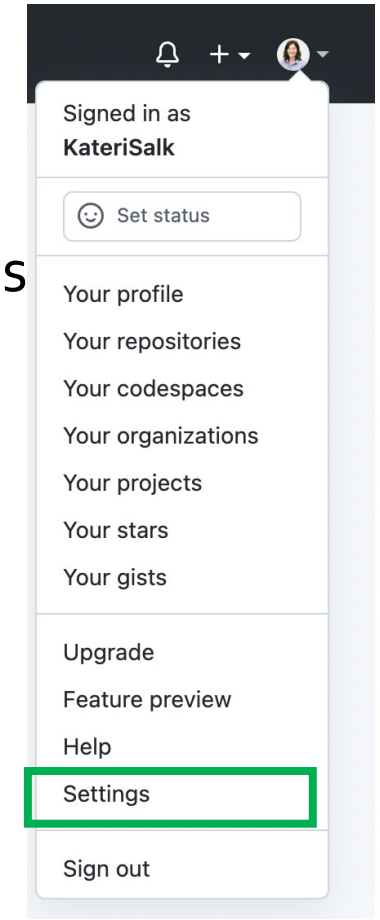
4. Open RStudio and go to File > New Project…
5. Select “Version Control”, then “Git”
6. Paste your forked repo URL and choose a folder where the local repo will be saved

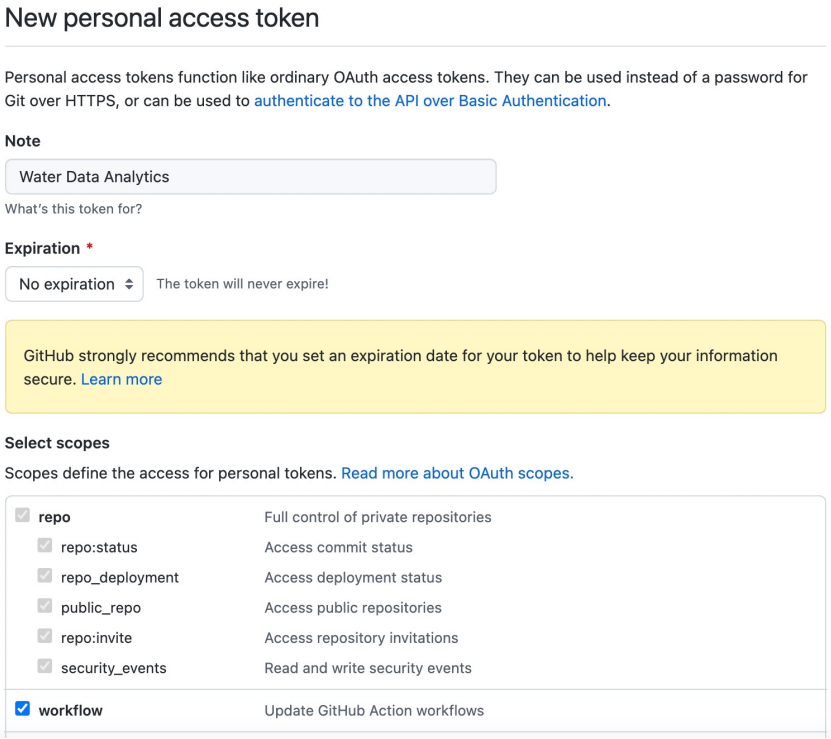


GITHUB SETUP: ERROR ON CLONING?

Set up a personal access token

1. From your profile menu, select Settings
2. Go to Developer Settings
3. Go to Personal access tokens
4. Click "Generate new token"
5. Add info
   - Note: suggest "Water Data Analytics"
   - Expiration
   - Repo,workflow, and user scope





GITHUB SETUP: PERSONAL ACCESS TOKEN

Copy access token from GitHub, then enter the following into R:

install.packages("gitcreds")
library(gitcreds)
gitcreds_set()

Then, paste the access token when prompted

From: https://happygitwithr.com/https-pat.html

GITHUB SETUP: COMMIT AND PUSH

- Open the Git_Help file and follow the instructions in the Editing, Committing, Pushing section.
- Familiarize yourself with how to keep the local, remote, and upstream remote repositories up to date with each other.