STA 2300 and MATH398: Lecture 26 Using github

Getting started

This material was abstracted from the detailed and excellent tutorial by Jenny Bryan (an R Studio Black Belt). https://happygitwithr.com/

Set up a github account

Go to https://github.com/ and create your account. Remember your user name! It is different from the email you use.

Install git on your laptop

Now install github software on your laptop this allows your computer to communicate with the github repository. This install will depend on whether you have a PC, Mac or Linux. I found this link helpful for the simplest routes for each system.

https://git-scm.com/book/en/v2/Getting-Started-Installing-Git

(In the case of Macs when RStudio was first opened it gave a prompt to download the mac command line tools that includes git. Pretty smart.)

Here are more details (that I actually found a bit too much!): Chapter 6 in the git/R tutorial https://happygitwithr.com/install-git.html

Once github is installed, Rstudio configure your local github version automatically or give a specific username using the code below.

```
library(usethis)
use_git_config(user.name = "Dane Joe", user.email = "jane@example.org")
```

Here the email *must* be the same as the one you used to create your github account but the user name can be different.

Making an R project folder

It is easy to create ordinary R project folders.

file > New Project * and dont selecet Version Control*

What we want is a new folder that is also a github repository.

Create the new repositoryin github

Log onto github and create a new repository and choose to add a README.md file. When you are in this repository the URL link is just the web address in your browser.

E.g. My example looks like: https://github.com/dnychka/MowaterExample

This is publicly accessible so you should be able to use my link and see the repository, download its contents, etc.

Here are the steps in detail:

- Go to https://github.com and make sure you are logged in.
- Click the green "New repository" button. Or, if you are on your own profile page, click on "Repositories", then click the green "New" button.

How to fill this in:

- Repository name: Choose wisely!
- Description: At least a sentence what this is for.
- Public. YES Initialize this repository with a README. For everything else, just accept the default.
- Click the big green button "Create repository."
- Copy the HTTPS clone URL to your clipboard via the green "Clone or Download" button.

You now have a new github reposistory!

You can share the URL, add files to this, recover previous files that were overwritten or removed and delete contents. And like Fort Knox it will also still be there if your laptop meets a disaster.

Adding a new project connected to github in RStudio

The way we will use github is as a way to keep a secure copy of your work in R. Having created the new repository we will now work on the R project side using tools in R studio.

In RStudio, start a new Project:

- File > New Project > Version Control > Git. In "Repository URL", paste the URL of your new GitHub repository. It will be something like this https://github.com/dnychka/MowaterExample
- Accept the default project directory name, e.g. myrepo, which coincides with the GitHub repo name.

- Think carefully where you want the project folder to sit. Organize your folders so you can find them later.
- Check "Open in new session", as that's what you'll usually want in real life.
- Click "Create Project".

The sandbox and the vault

The use of git for keeping track of files can be very complex, adapting to large software development efforts by many people. Here the key is to manage different *versions* of the software and that is why git is referred to as a version control system. For this course and for much of your data science work, howeever, you will just have to use a few core git operations: **commit**, **push**, and **pull**. These actions move files back aand forth between the remote github repository (your vault) and your laptop (your sandbox)

Using git and R is simplified by the handy Git GUI that will appear as a tab in Rstudio when you open up a git-based project.

Your vault 

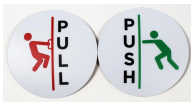
This is the copy of your Rstudio project folder on gihub and gives access to the current commited version as well as all previous files that have been committed. It is also permanent and secure being backed up through the github facility.

Your sandbox 

This is the copy of your Rstudio project on your laptop. You can add, delete, and modify files as you work on analysis and then periodically updtee the git repository with new work.This model is especially useful when several people work on the project. They develop some ideas separately in their sandboxes and when the code or files are ready then the central git repository is updated.

The process of updating a file or files to the git repository involves three steps "staging", **committing** and **pushing**

- Only files that been "staged", *committed* and *pushed* (**SCPed**) will be appear in the github the repository (the Vault). That means if your laptop crashes before updating you lose your sandbox and its contents.
- the R Studio GUI makes this convenient. Just click all the boxes for files to stage, click on **commit** and add a commit message, and then click on **push**.

- 

Clicking on **pull** will update the sandbox with any new files from the repository. Always **pull** before you **push**. If more than one person is updating the repository it is possible that your sandbox is out of date. E.g. Rachel updates fileA then Emily updates fileA. If Rachel now wants to work more on fileA it is important to have Emily's new copy with her excellent additions.

Once you quit R studio

To get back into the project folder in a new R studio session to continue your work open with

File> Open Project

Giving others write access to your respository

For others to save files to your repository you need to give them permission.

Sign on to github and go to the repository page. At the far right of the toolbar is **Settings**

Settings > Manage Access (in left menu) > Invite Collaborator

Type in the account name of the person to give access and it will send an email invite to them. Once they respond they will have access.

OK that is about it to get started!