Overleaf + R Part 1: Practicing Open-Source, Cloud-Based Workflows for Science Writing

Recorded Peer Scholars Workshop

Kate Jones
PhD Candidate, Center for Geospatial Analytics

Zoom Chat Course Materials Link

Installing R/RStudio

See hyperlinks at top of page for installing for Mac/Windows/Linux.

https://techvidvan.com/tutorials/install-r/#install-r-windows

If you have R/RStudio, but the versions are older than a year or so, please consider updating.

If you're installing R, please type "installing" in the chat.

Chat Write In

Rank: 1 (first time) **– 5** (use daily)

1. **Familiarity with R**
2. **Familiarity with Overleaf**
3. **Mac/Windows/Linux?**

Disclaimer:

I am still learning. All questions & comments are welcome to improve this material and ensure that we all learn something new & accurate.

Please feel free to unmute and share your thoughts or ask questions.

o·pen-source

adjective COMPUTING

adjective: **open-source**

denoting software for which the original source code is made freely available and may be redistributed and modified.

cloud-based

adjective COMPUTING

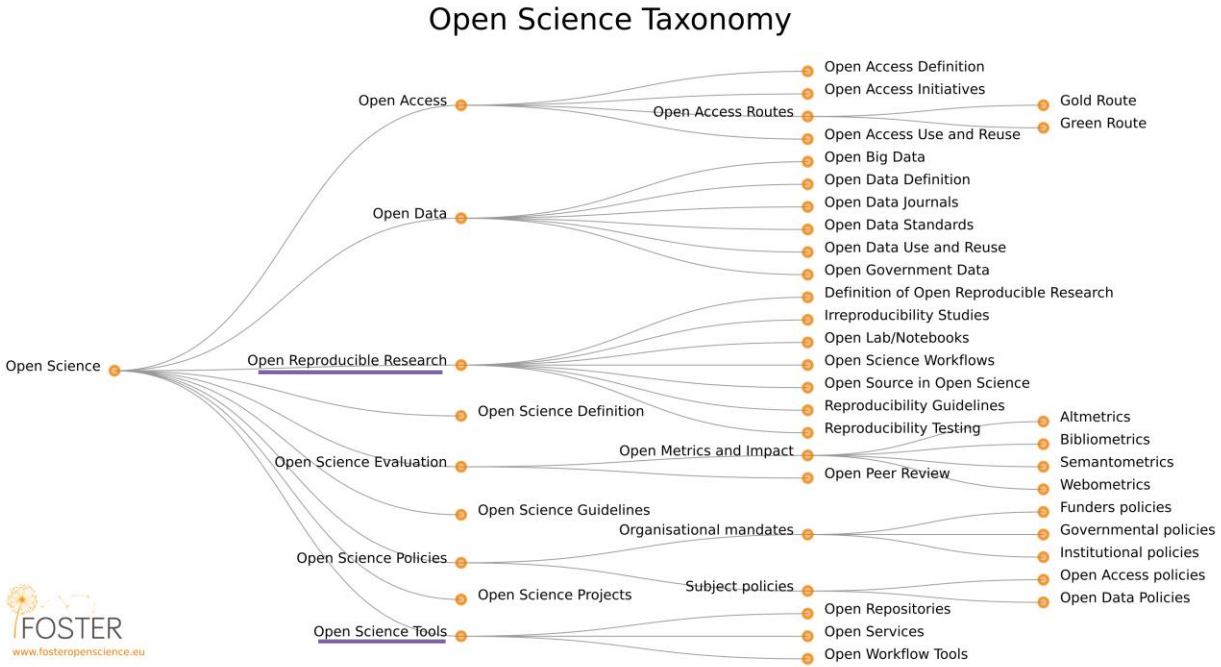
(of digital data) stored, managed, and processed on a network of remote servers hosted on the internet, rather than on local servers or personal computers.
"a commercial cloud-based platform"

Open Source begets Open Science

Open Source

- Free & Reproducible
- Open Source software evolves
- No reliance on single entity or license subscription
- It's common for developers to also be end-users

Open Science represents a whole ecosystem of ideas.



To learn more, check out NC State's:
Micah Vandegrift, @micahvandegrift

Why commit to learning new writing tools?

1. Open source -> Open Science

2. Work in the cloud = no/less stress about backups

3. EFFICIENT Formatting (journal + CV + other formats!)

Today's Goal

Each person leaves this workshop with a BASIC working example connecting **R/RStudio**, **Git/GitHub**, and **Overleaf**.

Part 1 is set up. After **Part 2**, you will be able to write, share figures, & back up code seamlessly across these platforms.

*If for some reason you are unable to follow along, this powerpoint (with steps to set up) and working code example are provided in the shared Google Drive.

Necessary Tools/Software

1. R/RStudio Desktop

R is a programming language and free software environment for statistical computing and graphics supported by the R Foundation for Statistical Computing. The R language is widely used among statisticians and data miners for developing statistical software and data analysis.

RStudio is an integrated development environment (IDE) for R. It includes a console, syntax-highlighting editor that supports direct code execution, as well as tools for plotting, history, debugging and workspace management.

2. Git/GitHub

Git is a version control system that lets you manage and keep track of your source code history.

GitHub is a cloud-based hosting service that lets you manage **Git** repositories.

3. Overleaf & LaTeX

Overleaf is an **online LaTeX** editor that's easy to use. No installation, real-time collaboration, version control, hundreds of **LaTeX** templates, and more.

LaTeX, which is pronounced «Lah-tech» or «Lay-tech» (to rhyme with «blech» or «Bertolt Brecht»), is a document preparation system for high-quality typesetting.

https://www.latex-project.org/about/

https://en.wikipedia.org/wiki/RStudio, blog.devmountain.com/git-vs-github-whats-the-difference#:~:text=Simply%20put%2C%20Git%20is%20a%20version%20control%20system,is%20designed%20to%20help%20you%20better%20manage%20them, https://www.overleaf.com/learn.

Why LaTex/Overleaf?

- What-You-See-Is-What-You-Get (WYSWYG) text editors
  - 1 line of code vs. 1,000 clicks
- Functions like google docs (simultaneous editing!)
  - Not tied to computer or subscription
- Organization/References
- Minimal cost – free through university

Why R/RStudio?

- Open source (https://github.com/rstudio)
- CRAN (Comprehensive R Archive Network)
- R vs. Excel
- Friendly IDE
- Kate is most familiar!
- Alternatives: R & Python vs. Matlab

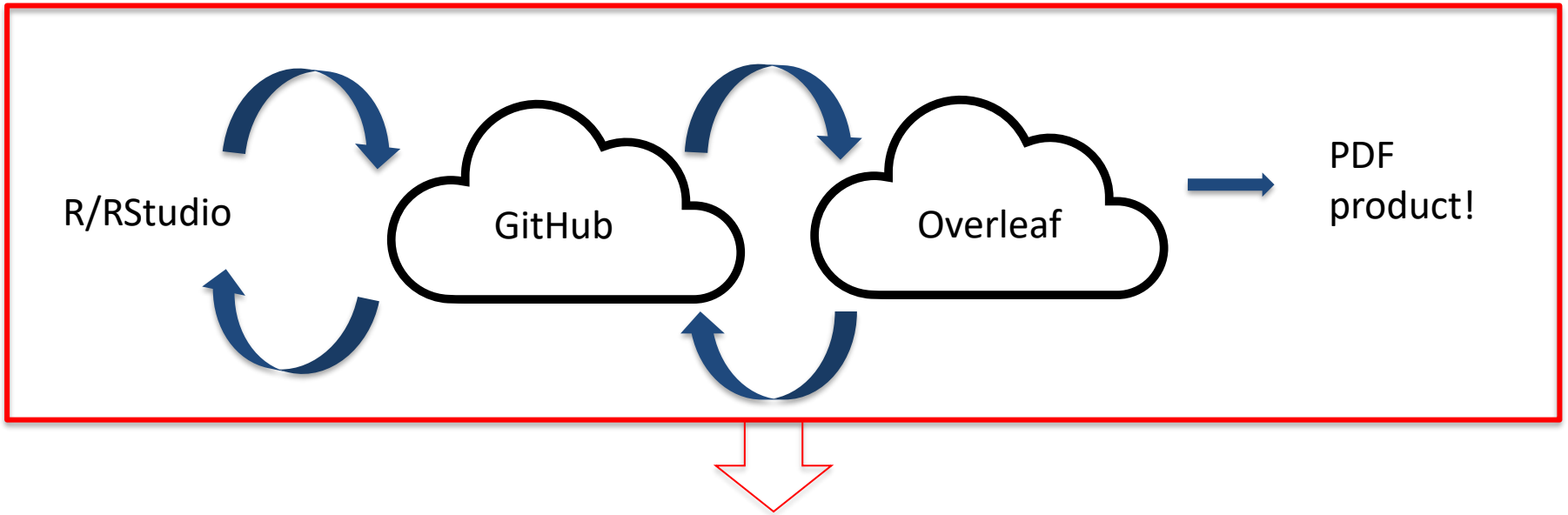
Why Overleaf + RStudio for writing?

- Reproducibility, reproducibility, reproducibility
  - Blending of open source tools
- Consistency – no manual data transfer/formatting
- Uses GitHub! – Back it up. Repeat.

Situations where this workflow is very useful:

- Recurring reports/assignments
- Manuscripts (corresponding authors)
- It’s been a loooonnnng time since you’ve looked at something (code/writing)
- Submitting to a different journal
- Others we can think of?

Our Desired Workflow



But first, we need to set up & connect these tools!

***Screenshots follow instructions**

Setup - Step 0 & 1

0. Everyone should have working installations of R & RStudio

1. Open RStudio to install necessary packages

- In the console type:

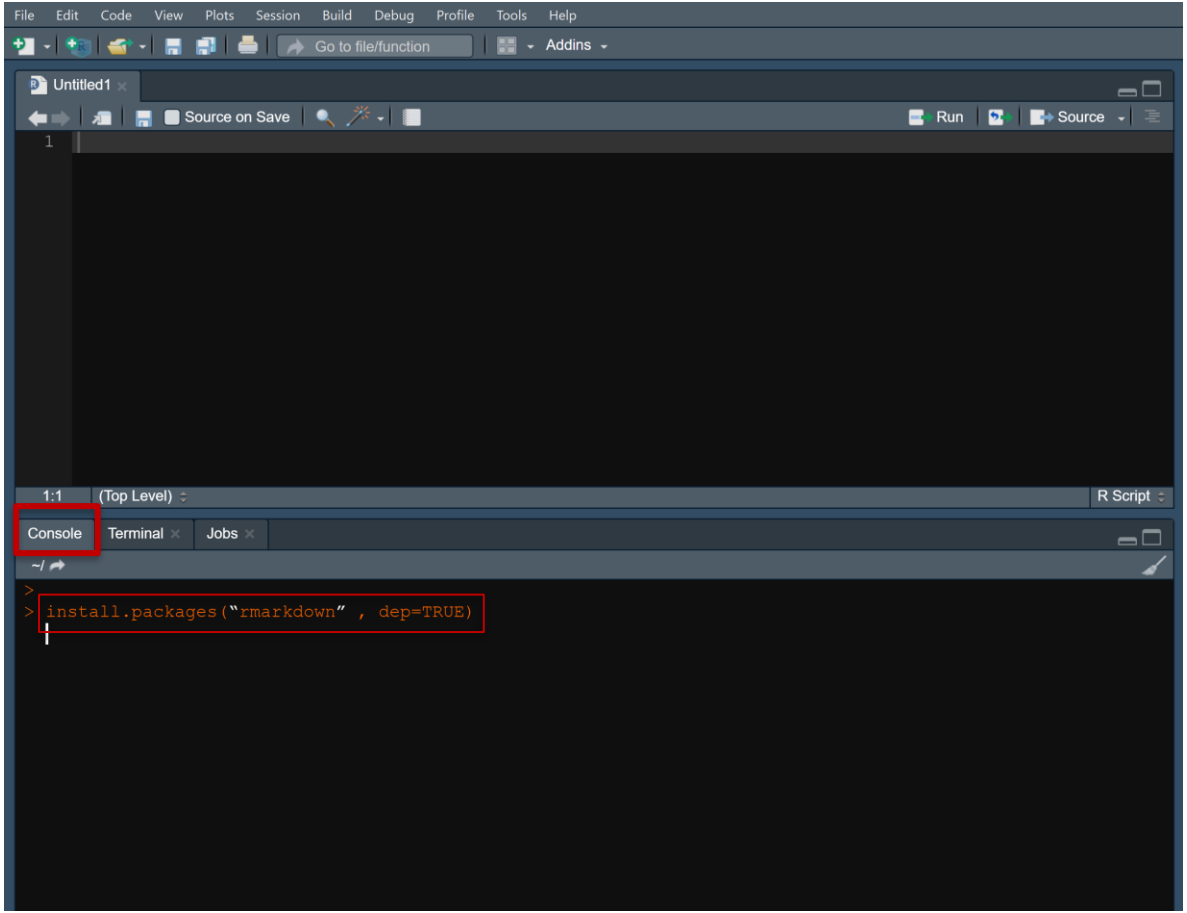
```
install.packages('rmarkdown' , dep=TRUE)
```

Setup - Step 1

1. Open RStudio to install necessary packages

- In the console type:

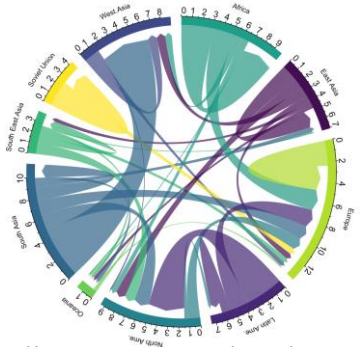
  install.packages(“rmarkdown” , dep=TRUE)



What is an R package?

Packages extend the base functions of R, such as:

- **Niftier plots**



https://www.data-to-viz.com/graph/chord.html

- **Ability to read in different file types:**
  .csv = Base R vs. .xlsx = readxl package

- **And so much more!**

What is RMarkdown?

- An R package ... that allows you to make readable reports directly from RStudio.

- Donalds Knuth's "literate programming"

Excellent DETAILED Intro to RMarkdown here:
https://rmarkdown.rstudio.com/articles_intro.html

More info on RMarkdown later

Setup - Step 2 & 3

2. Create GitHub account: https://github.com/

***This workflow will not work with Enterprise (NCSU) Github.**

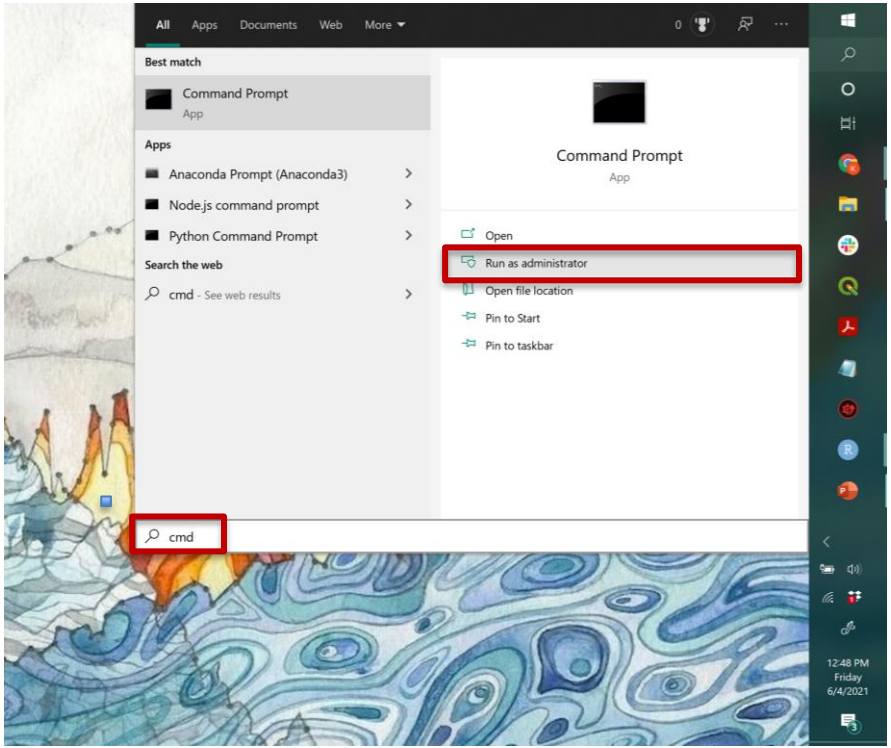
3. Check to see if Git is installed on your computer

1. Windows Application Search & type “cmd”
2. Right click black screen icon, select “Run as Administrator”
3. In console, type: “git --version” *note the double dash
4. If returns a version, type “where git”
5. If this returns a file path on your computer, you don’t need to install git! **COPY THIS PATH, paste in text file.

Setup - Step 2 & 3

1. Windows Application Search & type “cmd”

2. Right click black screen icon, select “Run as Administrator”



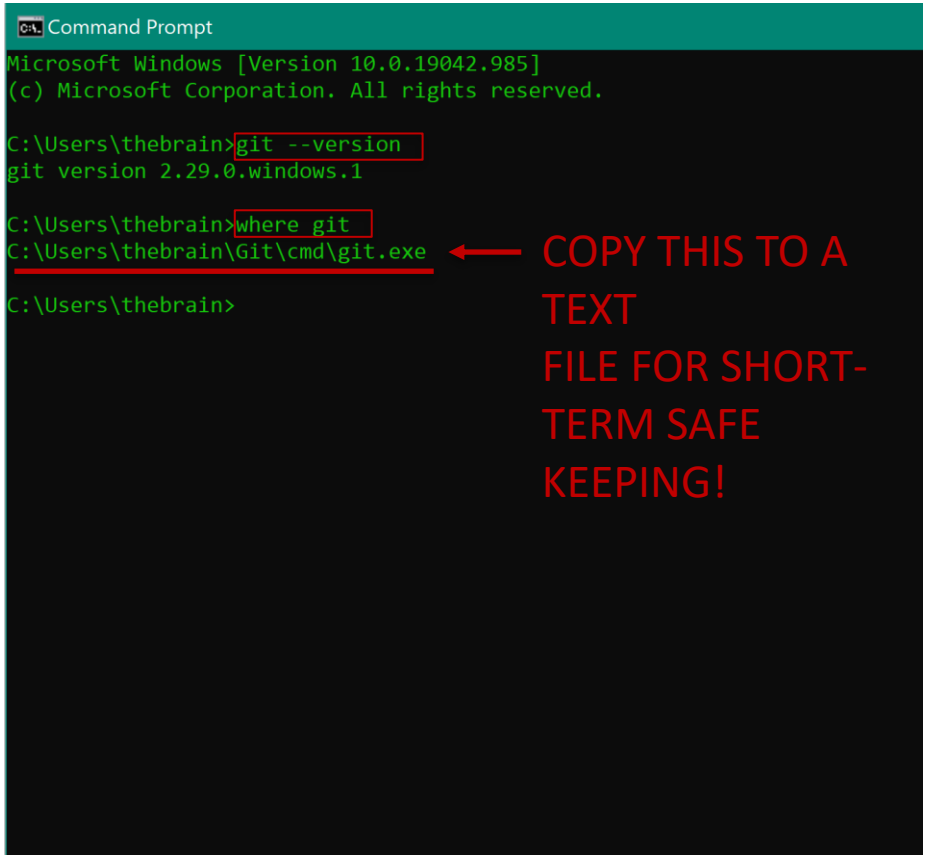
Setup - Step 2 & 3

3. In console, type: “git --version” *note the double dash

4. If returns a version, type “where git”

5. If this returns a file path on your computer, you don’t need to install git!

5b. If no git version appears, see next slide to download Git!



Setup - Steps 4 (ONLY IF NO GIT INSTALLATION)

4. If needed, install Git to your machine: https://git-scm.com/downloads

* Accept all defaults for Git installation
* Choose a directory to install w/ NO SPACES IN PATH!
* **Copy and paste installation directory path to text file ☺**

Setup – Step 5

5. Open RStudio > Tools > Global Options > Git/SVN > Git executable > Browse:

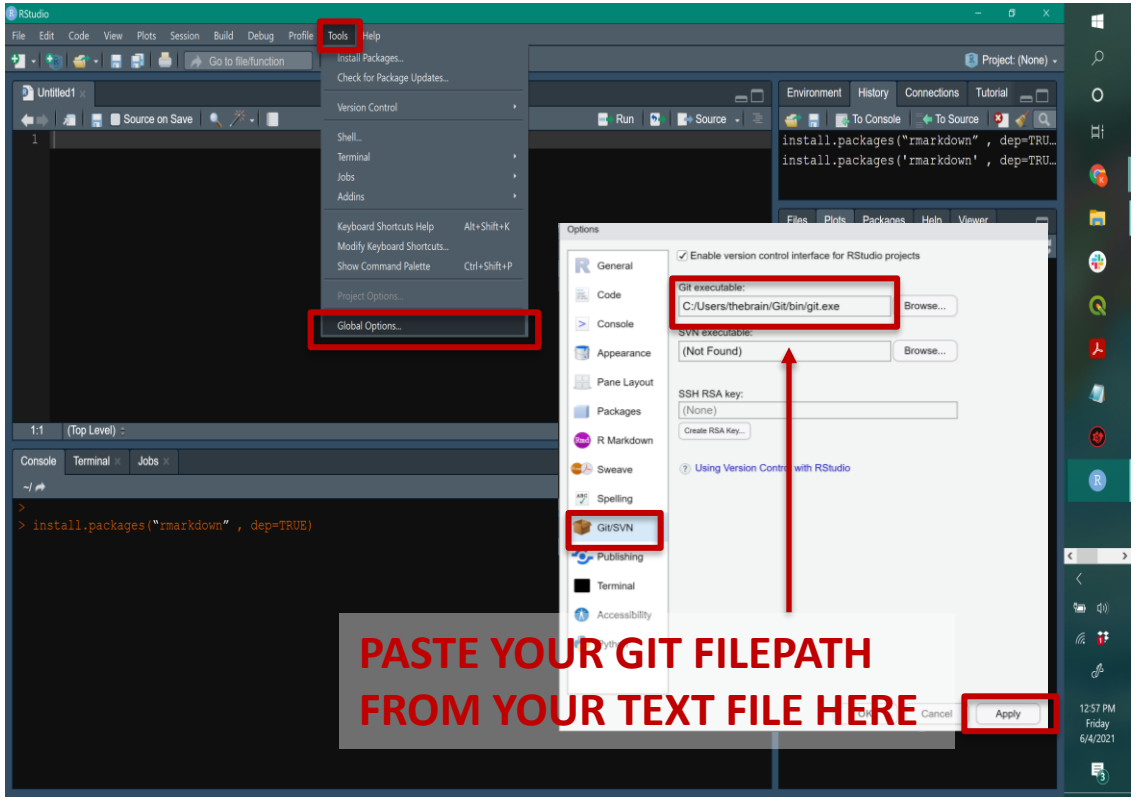
Path_pasted_in_textfile_above**/bin/git.exe**

(.exe will show as Git Application in directory files)

Setup - Steps 5

5. Open RStudio > Tools > Global Options > Git/SVN > Git executable > Browse:

Path_pasted_in_textfile_above/**bin/git.exe**



.exe will show as Git Application in directory files

Setup - Steps 6 & 7

6. Create Overleaf account (free with NCSU email): https://www.overleaf.com/edu/ncsu

***This workflow <u>should</u> be executed with NCSU Overleaf because of required functionality.**

7. New Project > Scroll to “Templates” & select “AMATH582 homework” (or use a template of choice)

-**Rename** Overleaf template/project to “PeerScholars_working”

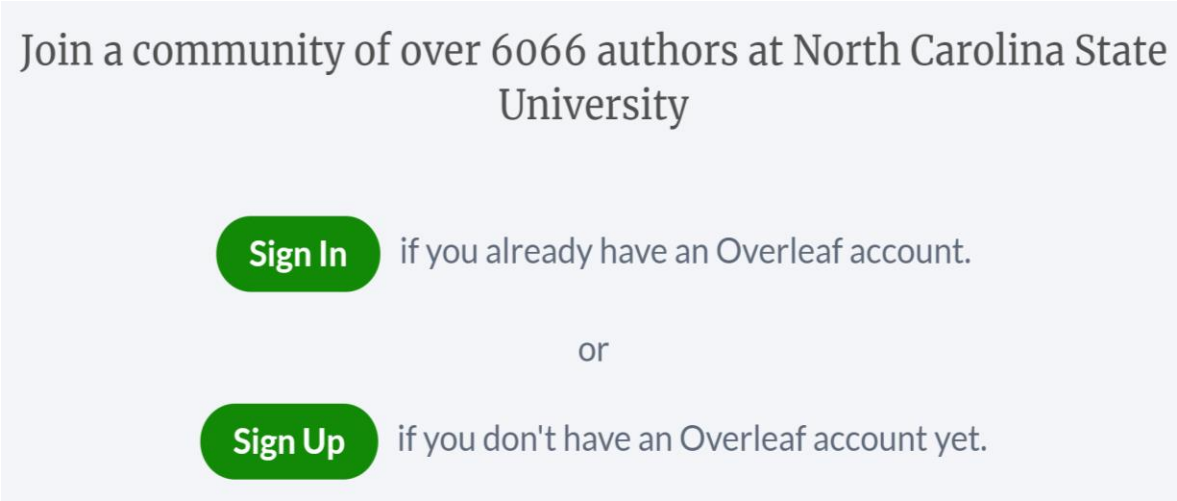
PAUSE: take 2-3 minutes to find an Overleaf template for a relevant journal

Setup - Steps 6 & 7

6. Create Overleaf account (free with NCSU email):

https://www.overleaf.com/edu/ncsu

Scroll to bottom of page:



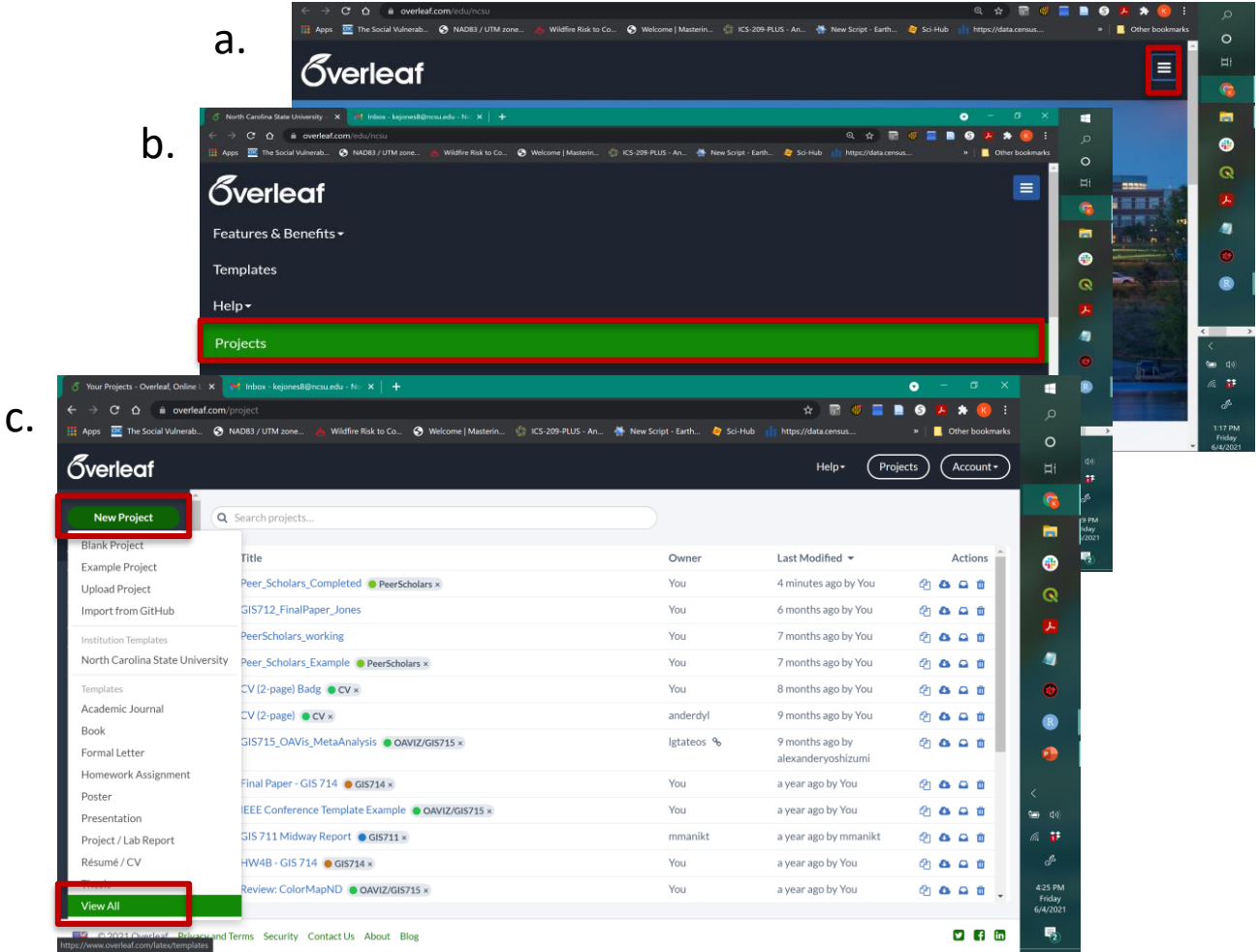
Setup - Steps 6 & 7

7. New Project > Scroll to "Templates" & select "AMATH582 homework" (or use a template of choice)

a.

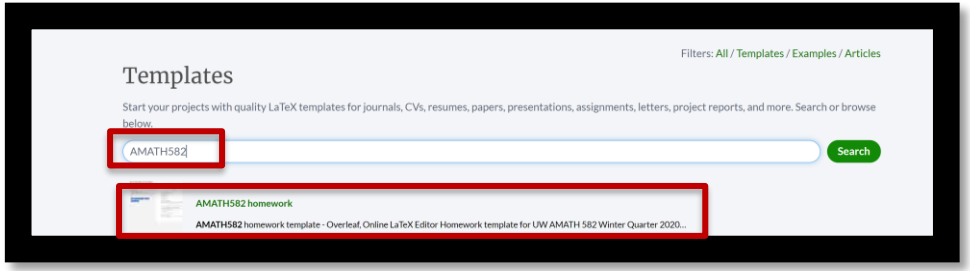
b.

c.

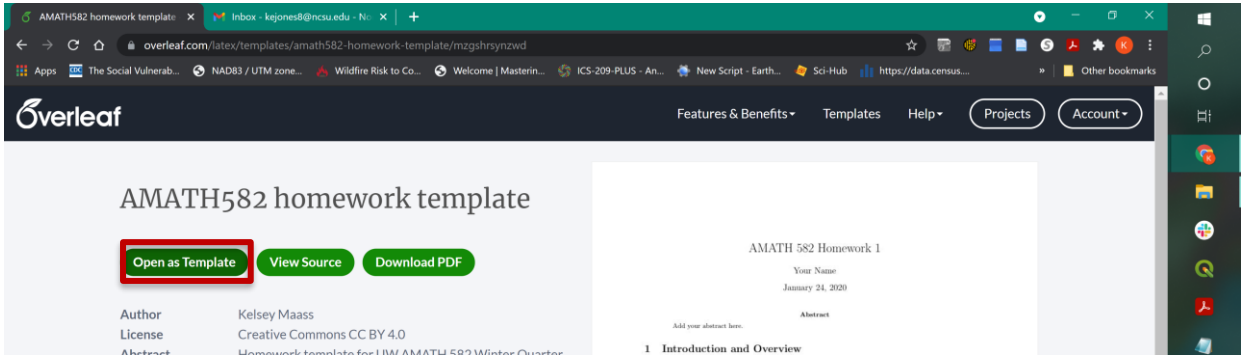


7. New Project > Scroll to “Templates” & select “AMATH582 homework” (or use a template of choice)

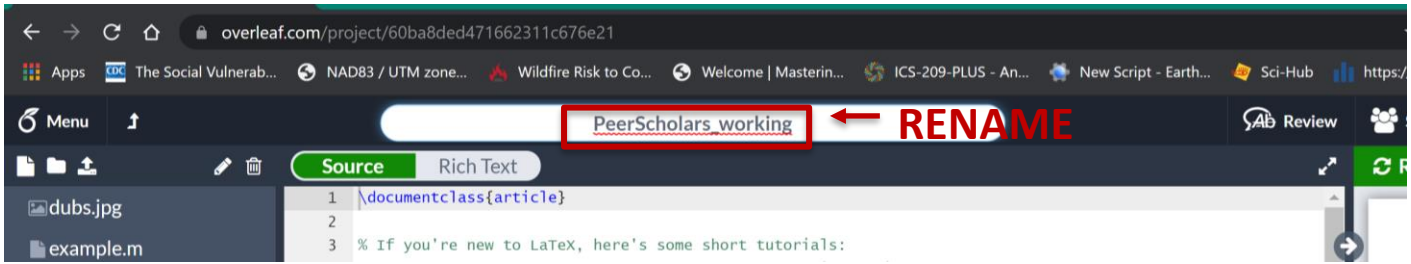
d.



e.



f.



Setup - Step 8

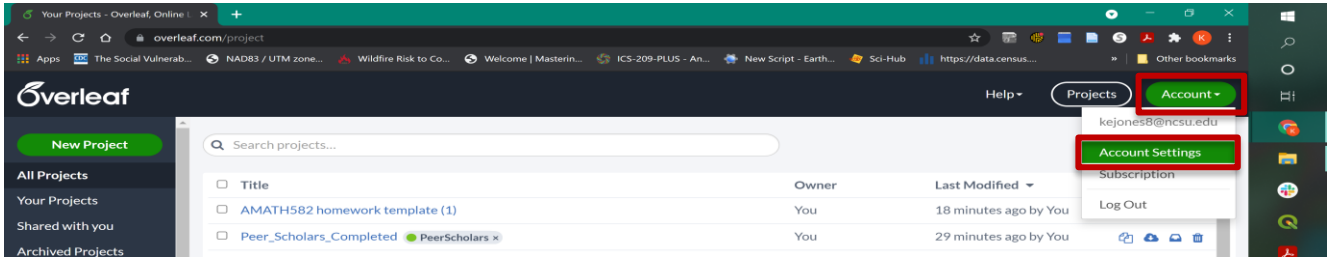
8. Create GitHub repository from Overleaf

1. Overleaf upper right corner of home page click  > Account > Account Settings
2. Scroll to “GitHub Integration” & Enter GitHub credentials
3. Back to your template “PeerScholars_working”
4. Upper left corner click Menu > GitHub > Create GitHub repository
5. Your GitHub ID should show up – can accept default project name
   *Make public repository 

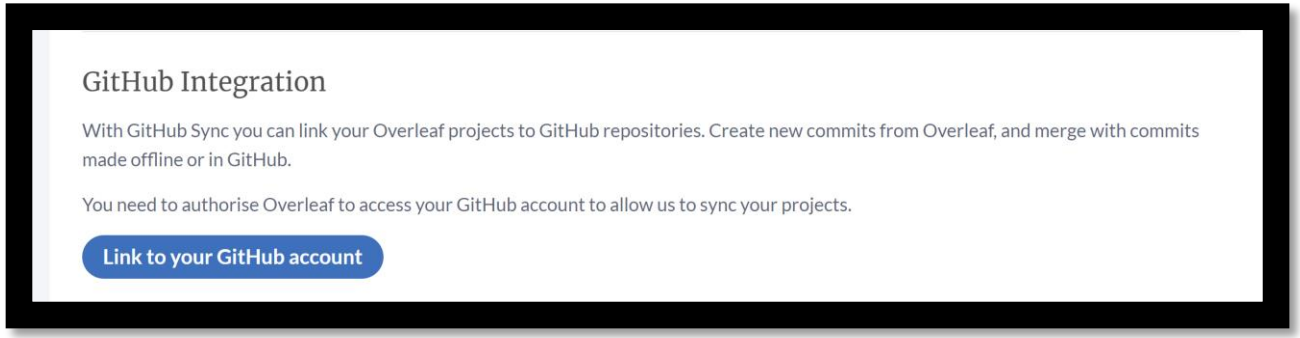
6. **It’ll tell you it’s linked! Click on** YourGitHubID\projectname
7. **It will open to your GitHub repository, leave this open in your browser**

8. Create GitHub repository from Overleaf

1. Overleaf upper right corner of home page click > Account > Account Settings



2. Scroll to “GitHub Integration” & Enter GitHub credentials



3. Select template “PeerScholars_working” (or whatever you may have named your selected Overleaf template)
4. Upper left corner click Menu > GitHub > Create GitHub repository
5. Your GitHub ID should show up – can accept default project name *Make public repository
6. **It’ll tell you it’s linked! In blue, click on** YourGitHubID\projectname
7. Click link, and it will open to your GitHub repository, **leave this open in your browser!**

Setup - Steps 9 & 10

Back in RStudio….

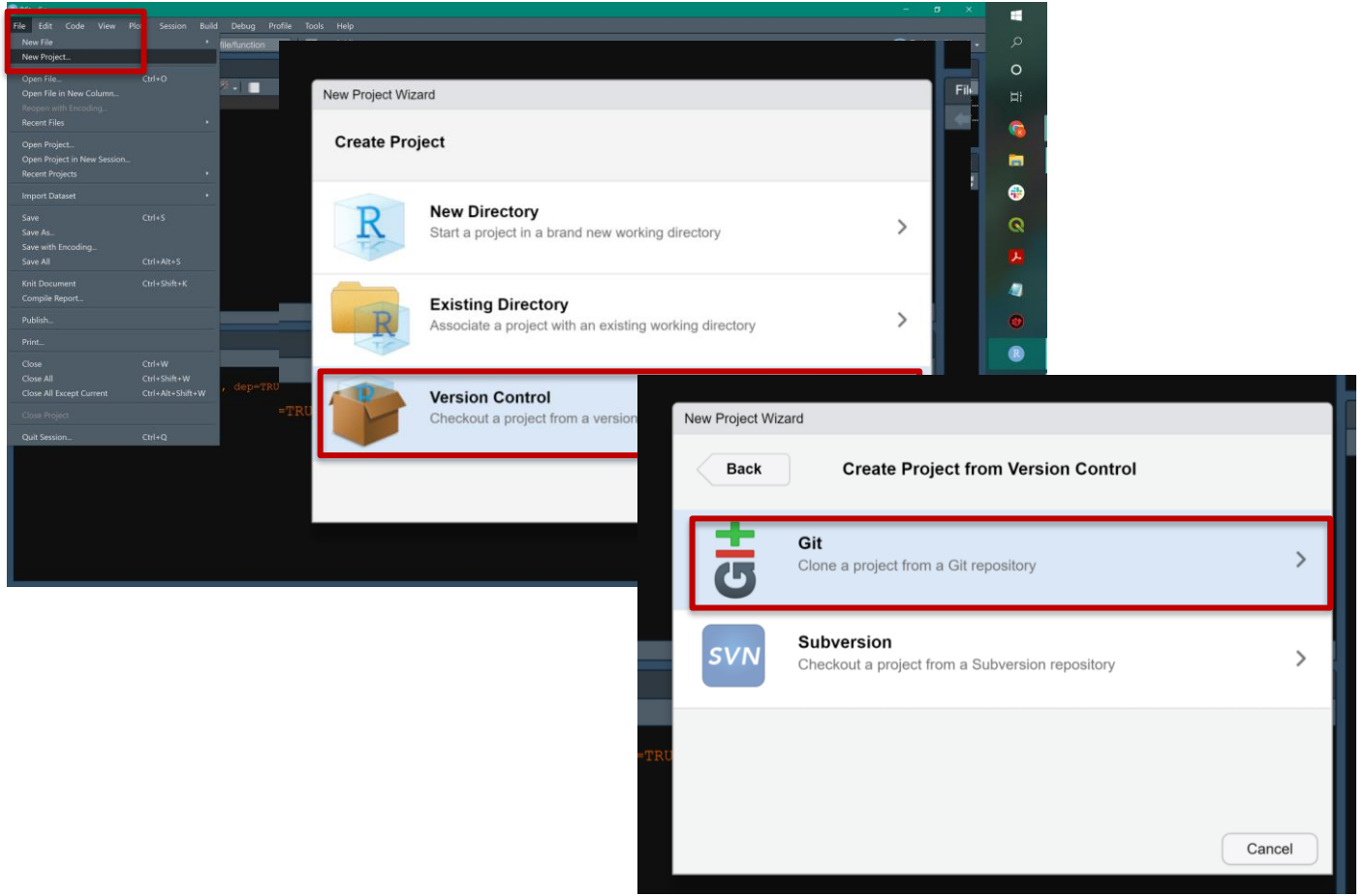
9. File > New Project : Select from Version Control > Git

10. Copy and Paste GitHub Repository URL (same repository created in Overleaf)

- Name directory (I recommend same as repository, as this option auto populates)
  - This directory/folder will be the repository linked to Github
- Where do you want the directory created on your computer?
  - Make directory a sub-folder somewhere in existing file structure
- Click radio button “Open in New Session”

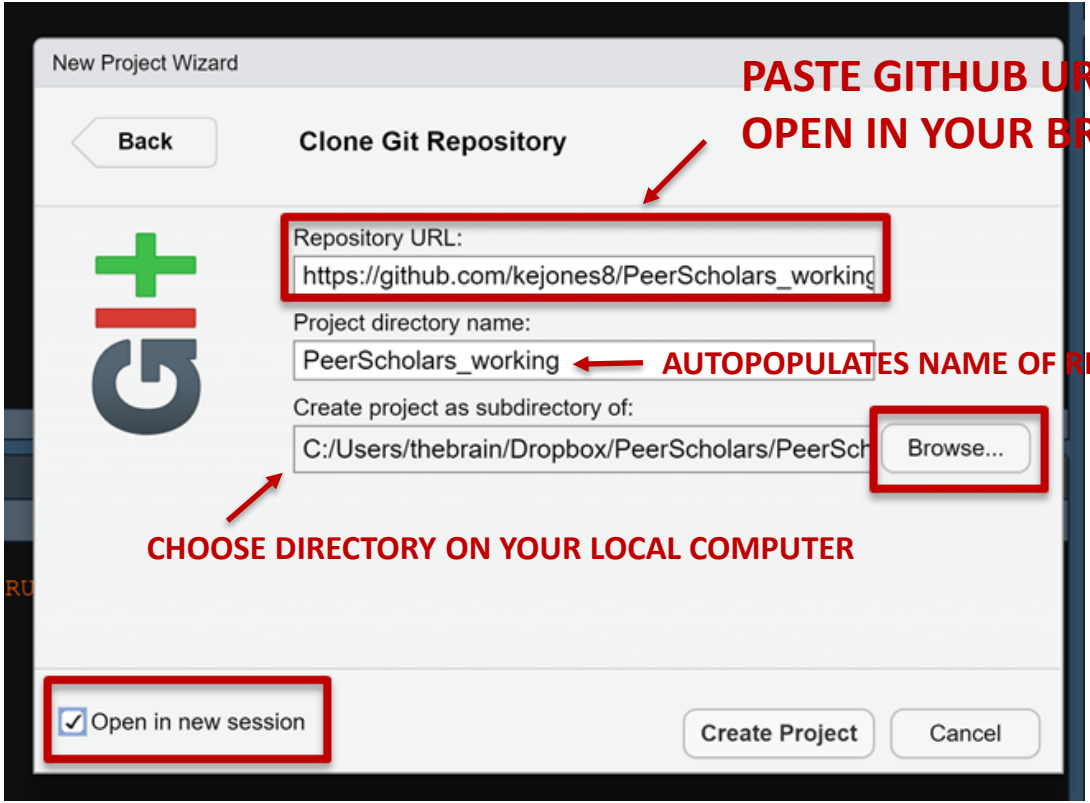
Setup - Steps 9 & 10

9. File > New Project : Select from Version Control > Git



10. Copy and Paste GitHub Repository URL (same repository created in Overleaf)

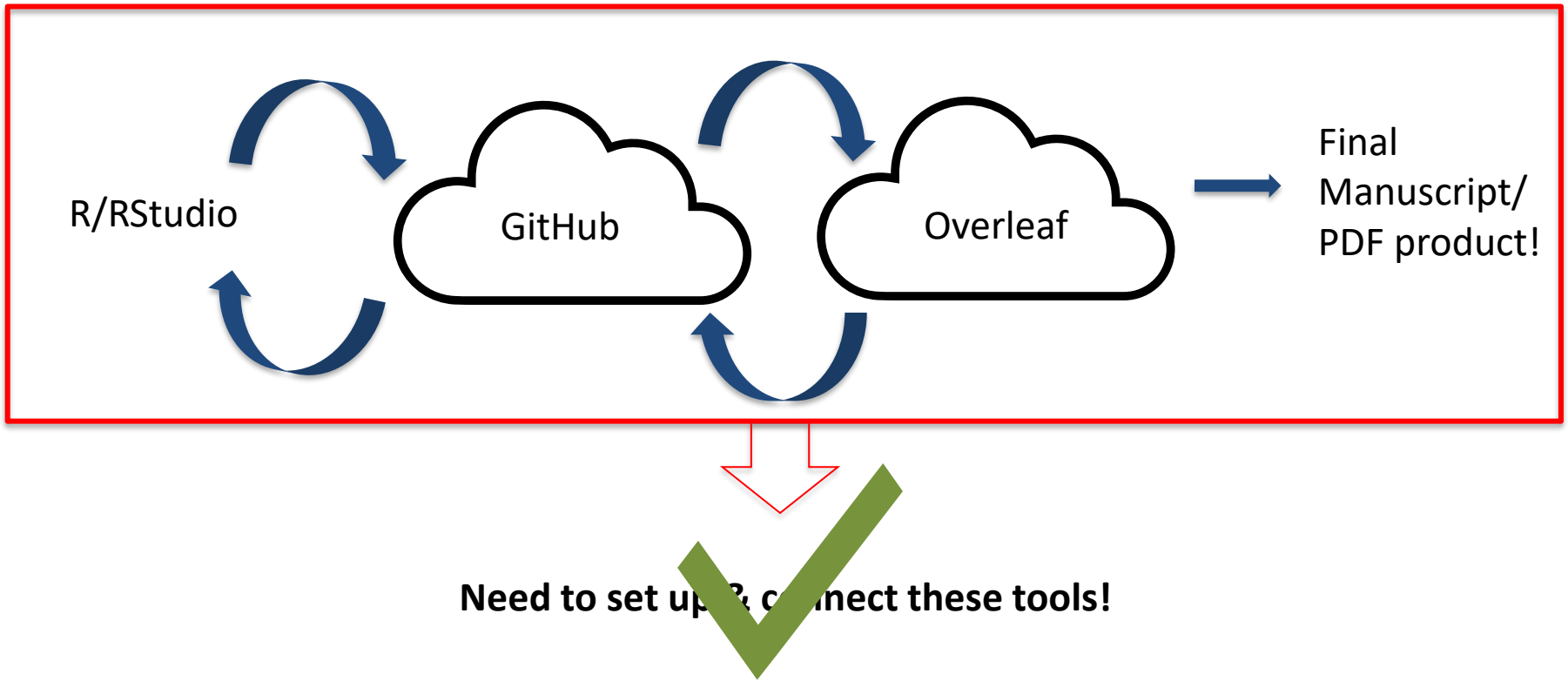
- Name directory (I recommend same as repository)
  - This directory/folder will be the repository linked to Github
- Where do you want the directory created on your computer?
  - Make directory a sub-folder somewhere in existing file structure
- Click radio button “Open in New Session”



Setup - Step 11

11. Move the Example_Rnotebook.Rmd shared google drive to your repository (aka working directory) where you just created your R Project, then from within the repository open (in RStudio File>Open) the Rmd.

Our Workflow



What steps will need to be repeated for future projects?

~~**0.** Working installation of R/RStudio~~

~~**1.** Open RStudio to install necessary packages~~

~~**2.** Create GitHub account: https://it.engr.ncsu.edu/services/github/~~

~~**3.** Install Git to your machine: https://git-scm.com/downloads~~

~~**4.** Open RStudio > Tools > Global Options > Git/SVN > Git executable > Browse: git.exe~~

~~**5.** Create Overleaf account (free with NCSU email): https://www.overleaf.com/edu/ncsu~~

6. In Overleaf: New Project > Scroll to “Templates” & select “AMATH582 homework” (or use a template of choice)

7. In Overleaf: Create GitHub repository from Overleaf

8. In R: File > New Project : Select from Version Control > Git

9. In R: Copy and Paste GitHub Repository URL (same repository created in Overleaf)

10. Open your own .Rmd, annotate, name chunks, plot to figures folders! AND:
PULL > KNIT > STAGE > COMMIT > PUSH > Overleaf Sync

Intro to R IDE (Integrated Development Environment)

- ✓ View all panes
- ✓ Files? Have we pulled?
- ✓ Console, Terminal, RMarkdown, Jobs
- ✓ Projects & Working Directories
- ✓ Libraries/Packages
- ✓ Built-in Datasets
- ✓ Global Options

install.packages("ggplot2")
install.packages("ggthemes")
install.packages("ggalt")

***before KNITTING RMD**

RNotebooks (using Rmarkdown)

- ✓ .Rmd vs. .R
- ✓ Code outputs in-line vs. Knitting
- ✓ YAML vs. HTML vs. R Code
- ✓ .Rmd header
- ✓ Naming Chunks

https://rmarkdown.rstudio.com/

Let's look at some possible RMD outputs!

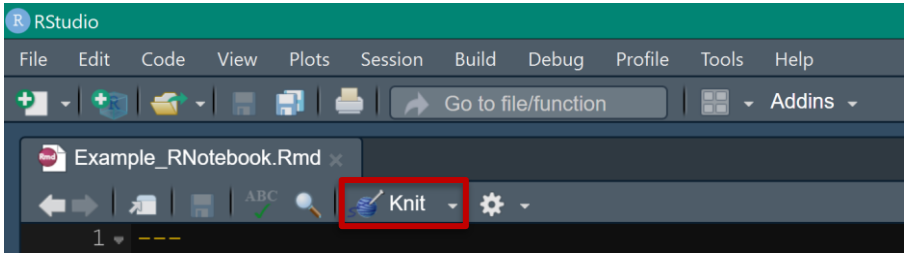
Knitting?

If your code has errors, it won't knit…

Knit = Render

OR

Turn Your Code into Pretty Text Format



HTML, PDF, etc.

If you have problems knitting to pdf, do this:

install.packages("tinytex")

tinytex::install_tinytex()

…your knit document will get overwritten every time you 'knit'!

Why not write EVERYTHING in RMarkdown???

Markdown vs. LaTex

https://www.slant.co/versus/1903/13783/~markdown_vs_latex

Markdown Cheatsheet: https://rstudio.com/.../02/rmarkdown-cheatsheet.pdf
LaTex Cheatsheet: https://fr.overleaf.com/latex/templates/a-quick-guide-to-latex-overleaf-version/bphpqrdgjyqy

Useful Resources:

Intro to R (opt. 1)
Intro to R (opt. 2)
Intro to R (opt. 3)
Intro to R (opt. 4)

Intro to RMarkdown
Intro to RNotebooks (opt2)

More on R & Github
More on R & Version Control (i.e. Github or Subversion)

In the second half of this workshop (Part 2) we will:

- **Work seamlessly between RStudio & Overleaf**

(R) PULL > (R) KNIT > (R) COMMIT > (R) PUSH > Sync (Overleaf)

- Overleaf tips & tricks
- Helpful file & document management strategies

Questions

Kate Jones, kejones8@ncsu.edu