

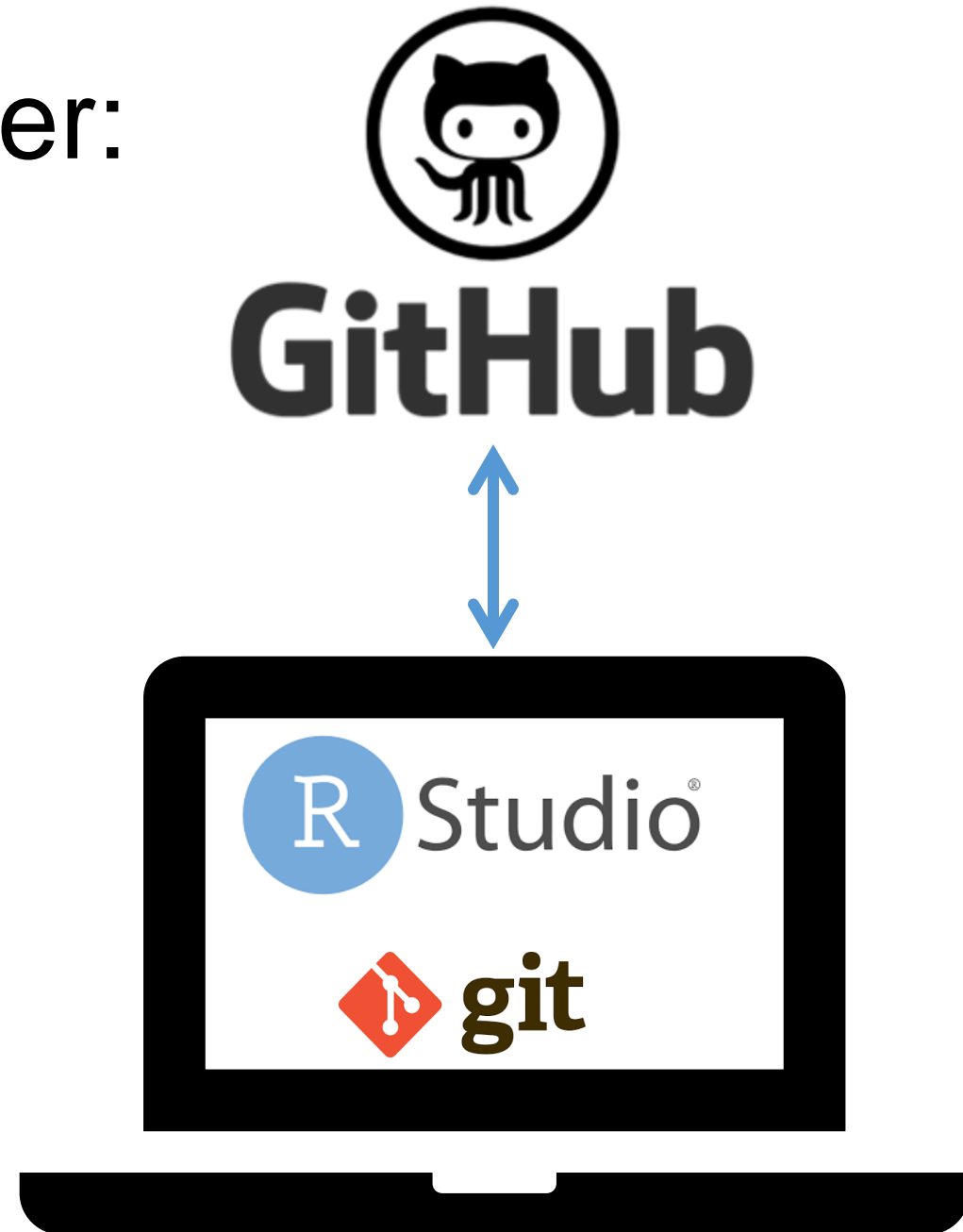
Version control using git in RStudio

Tobin Magle, PhD

Lead data management specialist
Research Computing and Data Services

Today we're going to cover:

- What is **version control**?
- What is **git**?
- Version control using **RStudio**
- Backing up your data to **GitHub**



Manual Version Control

“Renaming files”

- Creates many files
- Relies on your ability to name files consistently
- Not designed for collaboration

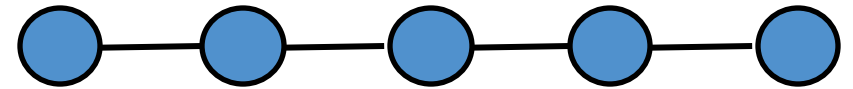


Version Control Systems

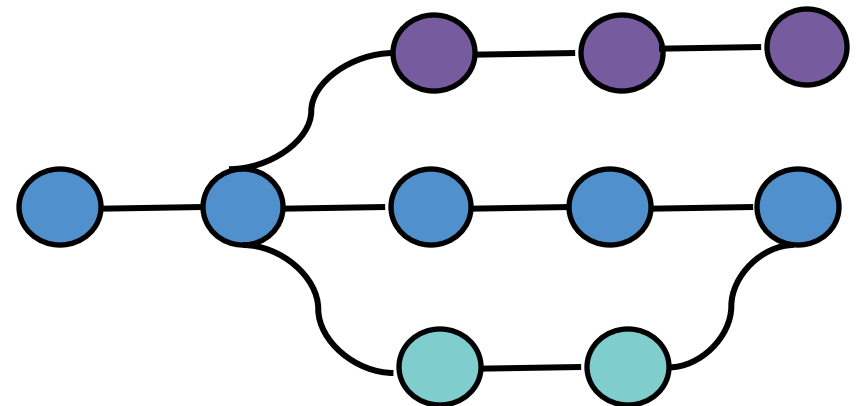
“Unlimited undo”

- One file, many versions
- Records who made what change when
- Good for **collaboration**: identifies and helps resolve conflicting changes

Solo



Collaborative



Today we're using git

- Built for large software development projects
- Can get complicated, but the basics are simple
- Free, open source
- Many online resources/tutorials



Many ways to use Git

Command line

```
C:\Windows\System32\cmd.exe

D:\ESC>git status
On branch master

Initial commit

Changes to be committed:
  (use "git rm --cached <file>..." to unstage)

    new file:   Basic Code/C++ Project.cbp
    new file:   Basic Code/README.md
    new file:   Basic Code/bin/Debug/C++ Project.exe
    new file:   Basic Code/data.in
    new file:   Basic Code/data.out
    new file:   Basic Code/main.cpp
    new file:   Basic Code/obj/Debug/main.o
    new file:   DS/stackPar.cpp
    new file:   DS/stackPar.exe
    new file:   DS/stackPar.o
    new file:   DS/stackPar.cpp
```

Desktop Clients

Current branch **esc-pr** #3972 ✓
 Fetch origin
Last fetched 2 minutes ago

Add event handler to dropdown component

iAmWillShepherd and Markus Olsson committed c79e71c 1 changed file

Co-Authored-By: Markus Olsson <niik@users.noreply.github.com>

app\src\ui\toolbar\dropdown.tsx			
	145	145	@@ -145,6 +145,10 @@ export class ToolbarDropdown extends React.Component<
	146	146	this.state = { clientRect: null }
	147	147	}
		148	+ private get isOpen() {
		149	+ return this.props.dropdownState === 'open'

GitHub Website

Edited the readme file [Browse files](#)

Changed the title to something better. Added a bulleted list

main

maglet committed 38 minutes ago Verified 1 parent 20225f0 commit e1b19cd

Showing 1 changed file with 5 additions and 2 deletions. Split Unified

7 README.md

...	@@ -1,3 +1,6 @@
1 - # github-gui-demo	1 + # Demo repository for using the GitHub GUI
2	2

RStudio IDE

Environment History Connections Git Tutorial

Diff Commit Pull Push History More main

Staged	Status	Path
<input type="checkbox"/>		script.R
<input type="checkbox"/>		data/2023-08-25-Globus_Usage_Transfer_Detail.csv
<input type="checkbox"/>		data/Globus_Usage_Transfer_Detail.csv
<input type="checkbox"/>		data/RDSS_users.csv

Many ways to use Git

Command line

```
C:\Windows\System32\cmd.exe

D:\ESC>git status
On branch master
Initial commit

Changes to be committed:
  (use "git rm --cached <file>..." to unstage)

    new file:   Basic Code/C++ Project.cbp
    new file:   Basic Code/README.md
    new file:   Basic Code/bin/Debug/C++ Project.exe
    new file:   Basic Code/data.in
    new file:   Basic Code/data.out
    new file:   Basic Code/main.cpp
    new file:   Basic Code/obj/Debug/main.o
    new file:   DS/stackPar.cpp
    new file:   DS/stackPar.exe
    new file:   DS/stackPar.o
    new file:   DS/stackPar.cpp
```

Desktop Clients

Current branch **esc-pr** #3972 ✓

Fetch origin
Last fetched 2 minutes ago

Add event handler to dropdown component

iAmWillShepherd and Markus Olsson committed c79e71c 1 changed file

Co-Authored-By: Markus Olsson <niik@users.noreply.github.com>

GitHub Website

Edited the readme file

Changed the title to something better. Added a bulleted list

main

maglet committed 38 minutes ago Verified 1 parent 20225f0 commit e1b19cd

Showing 1 changed file with 5 additions and 2 deletions.

7

README.md

...	@@ -1,3 +1,6 @@
1 - # github-gui-demo	1 + # Demo repository for using the GitHub GUI
2	2

RStudio IDE

Environment

History

Connections

Git

Tutorial

Diff Commit Pull Push History More main

Staged

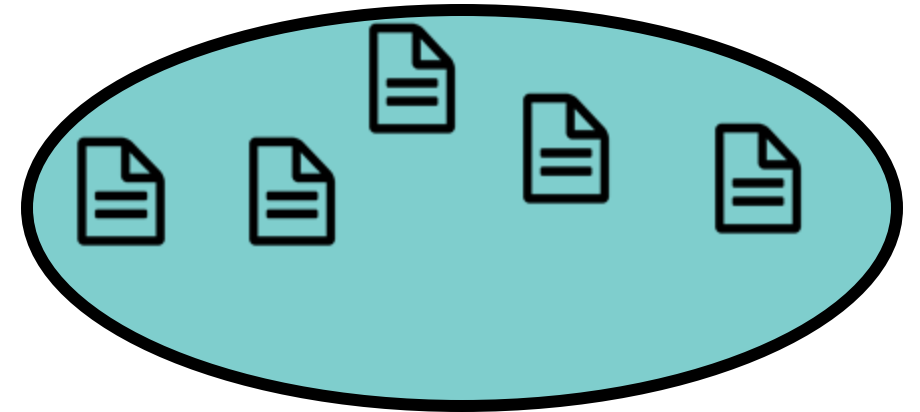
Status

Path

<input type="checkbox"/>		script.R
<input type="checkbox"/>		data/2023-08-25-Globus_Usage_Transfer_Detail.csv
<input type="checkbox"/>		data/Globus_Usage_Transfer_Detail.csv
<input type="checkbox"/>		data/RDSS_users.csv

Using git creates a repository

- **Repository** – A folder that is tracked by git



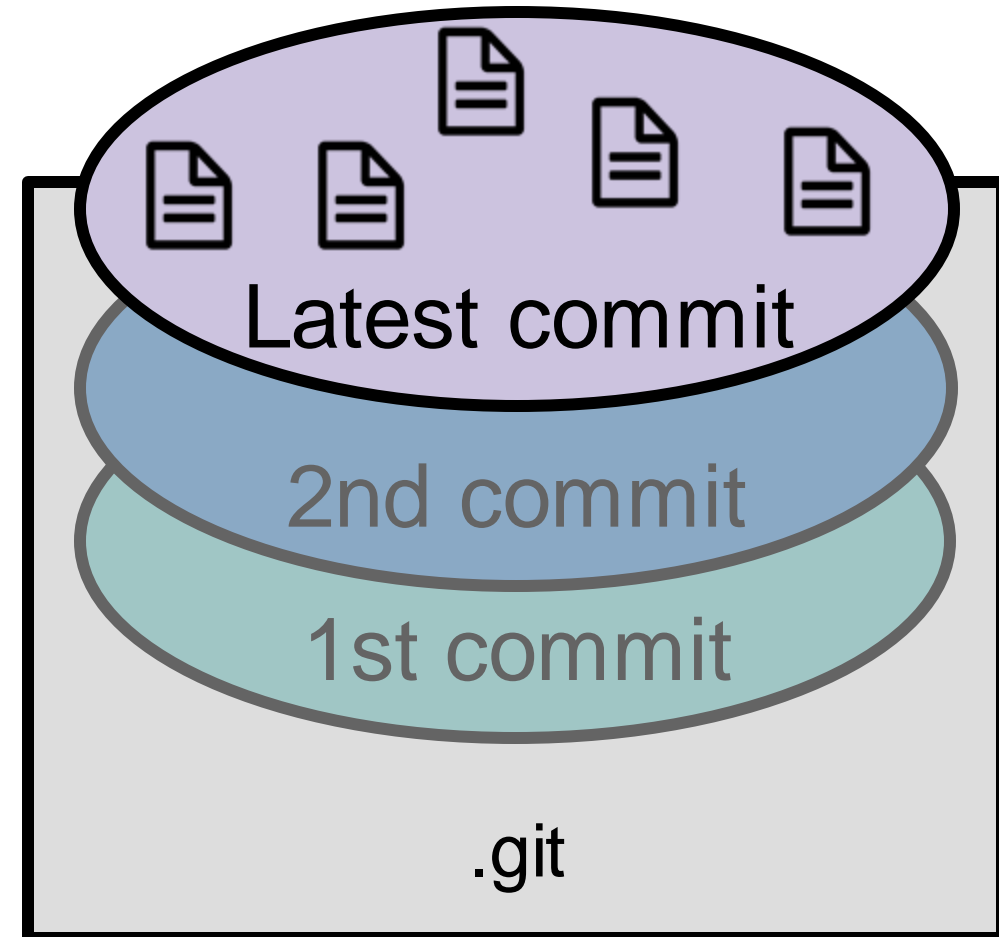
What is a repository?

- **Repository** – A folder that is tracked by git
- **Commit** – a version of what the folder looks like when it was created



Repository structure

- **Repository** – A folder that is tracked by git
- **Commit** – a version of what the folder looks like when it was created
- **Structure is hidden** (.git), you see the most recent version



RStudio workflow



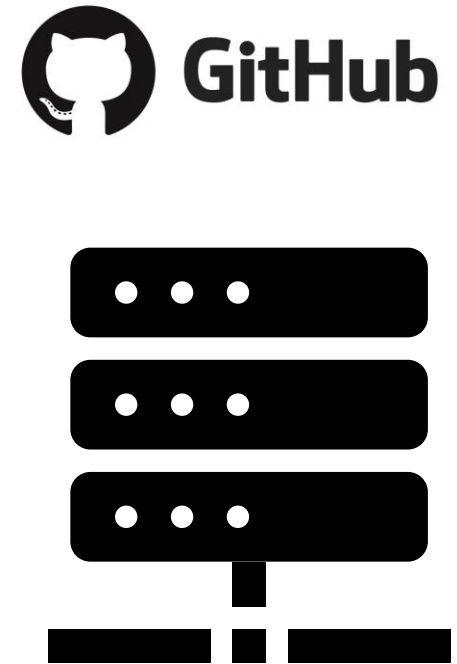
Coding happens on your computer

- Create and edit files in RStudio
- Use git version control in the RStudio interface
- Creates a local repository

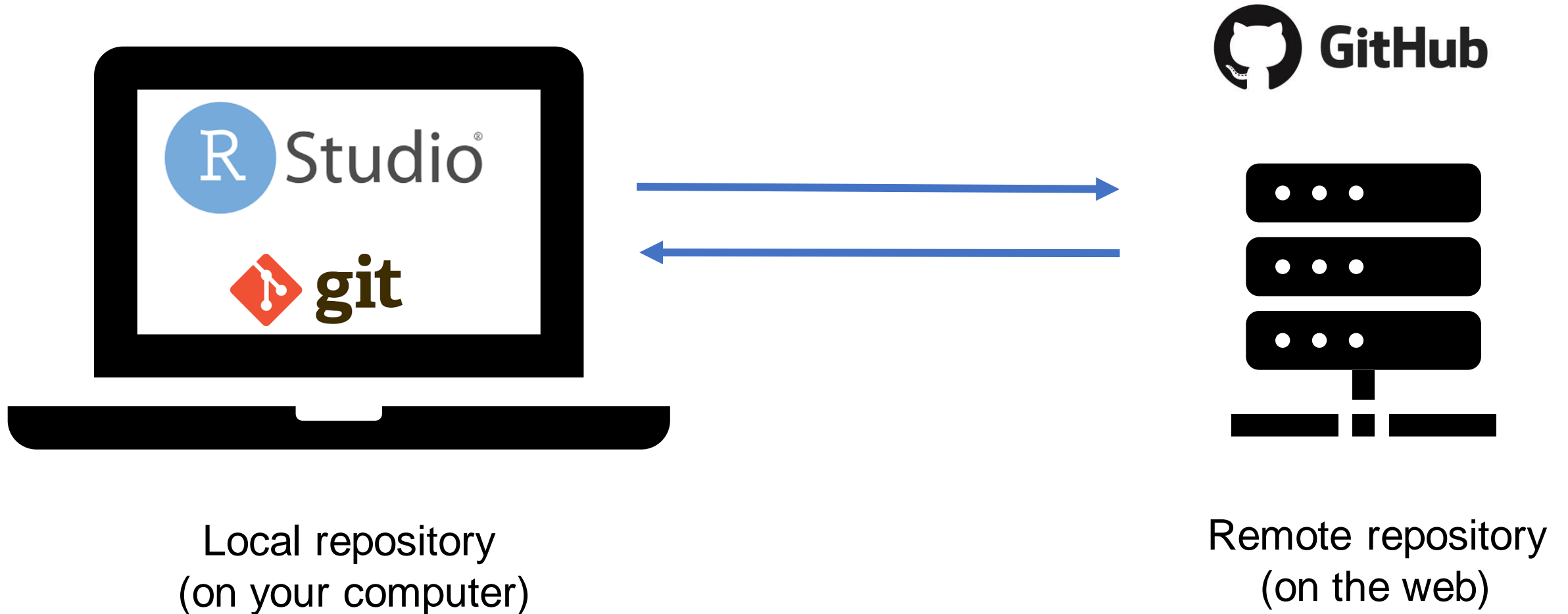
Git Workflow

You can save your work on GitHub

- Acts as a backup
- Share with colleagues
- Collaboration: Users with access can copy, modify and suggest changes to code

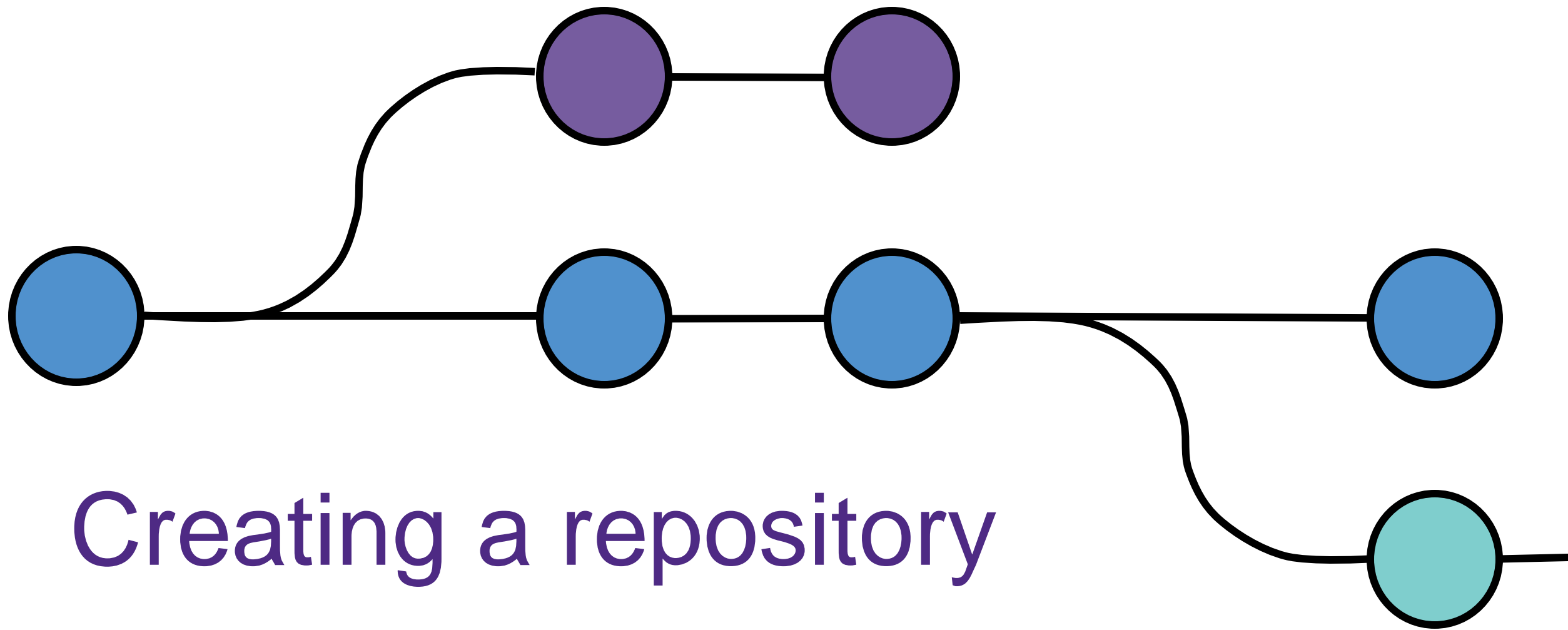


In practice

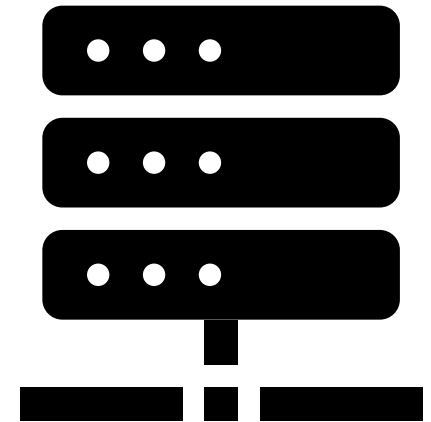


Pre-workshop Setup

- Install R installed
- Install RStudio
- Install git

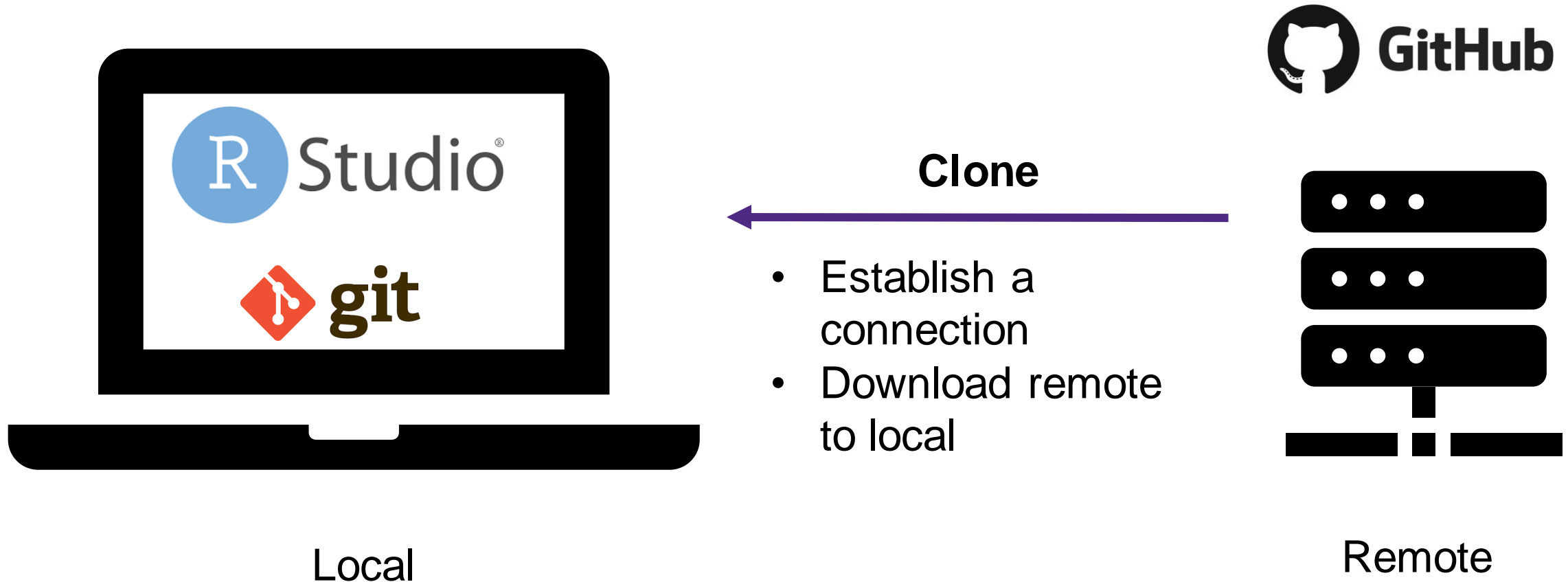


Workflow



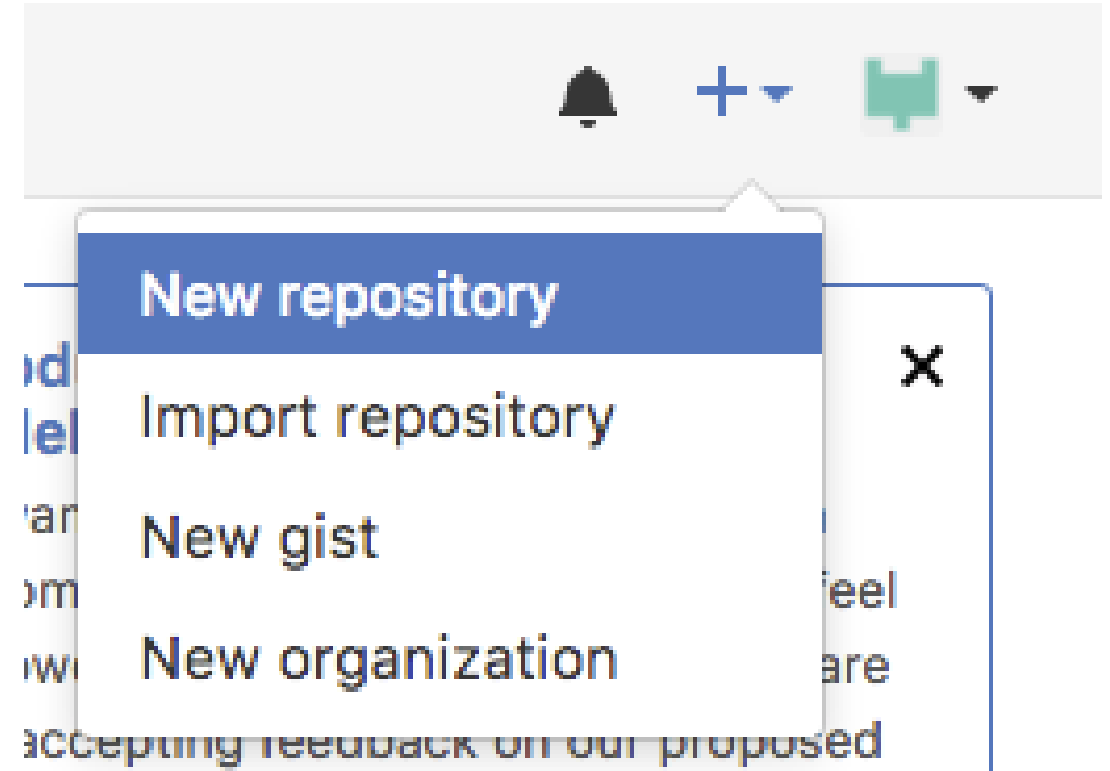
Remote repository
(on the web)

Workflow



Make an empty GitHub repo

- Go to <http://github.com>
- Click the plus sign on the upper right hand side of the screen
- Select New repository



Make an empty repository

- Name your repository



- Provide a description



- Keep it public




- Initialize with a readme



Create a new repository

A repository contains all project files, including the revision history.

Owner

 maglet ▾

Repository name *

/

Great repository names are short and memorable. Need inspiration? How about [miniature-octo-broccoli](#)?

Description (optional)



Public

Anyone can see this repository. You choose who can commit.



Private

You choose who can see and commit to this repository.

☒ Initialize this repository with a README

This will let you immediately clone the repository to your computer. Skip this step if you're importing an existing repository.

Add .gitignore: None ▾

Add a license: None ▾

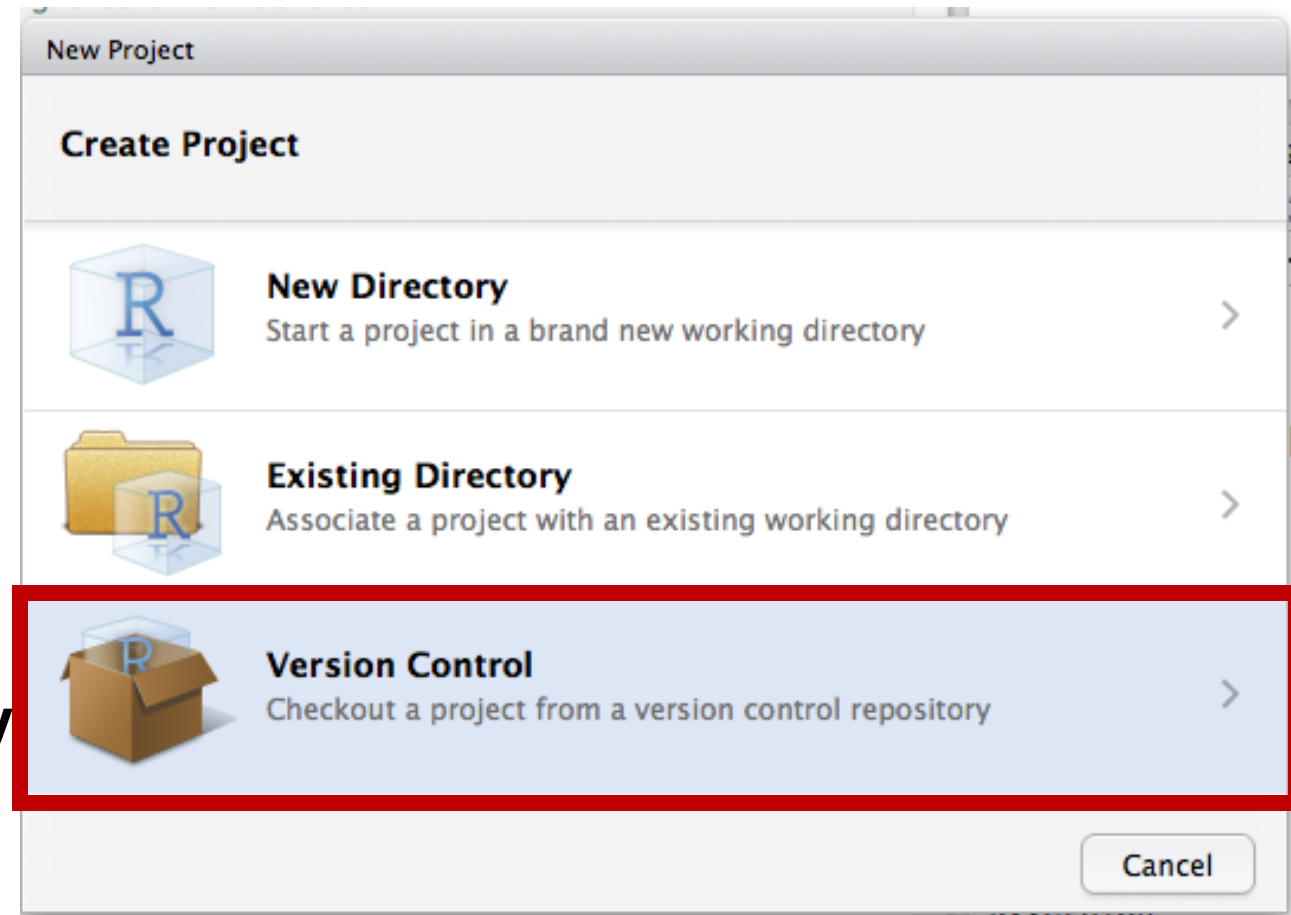


Create repository

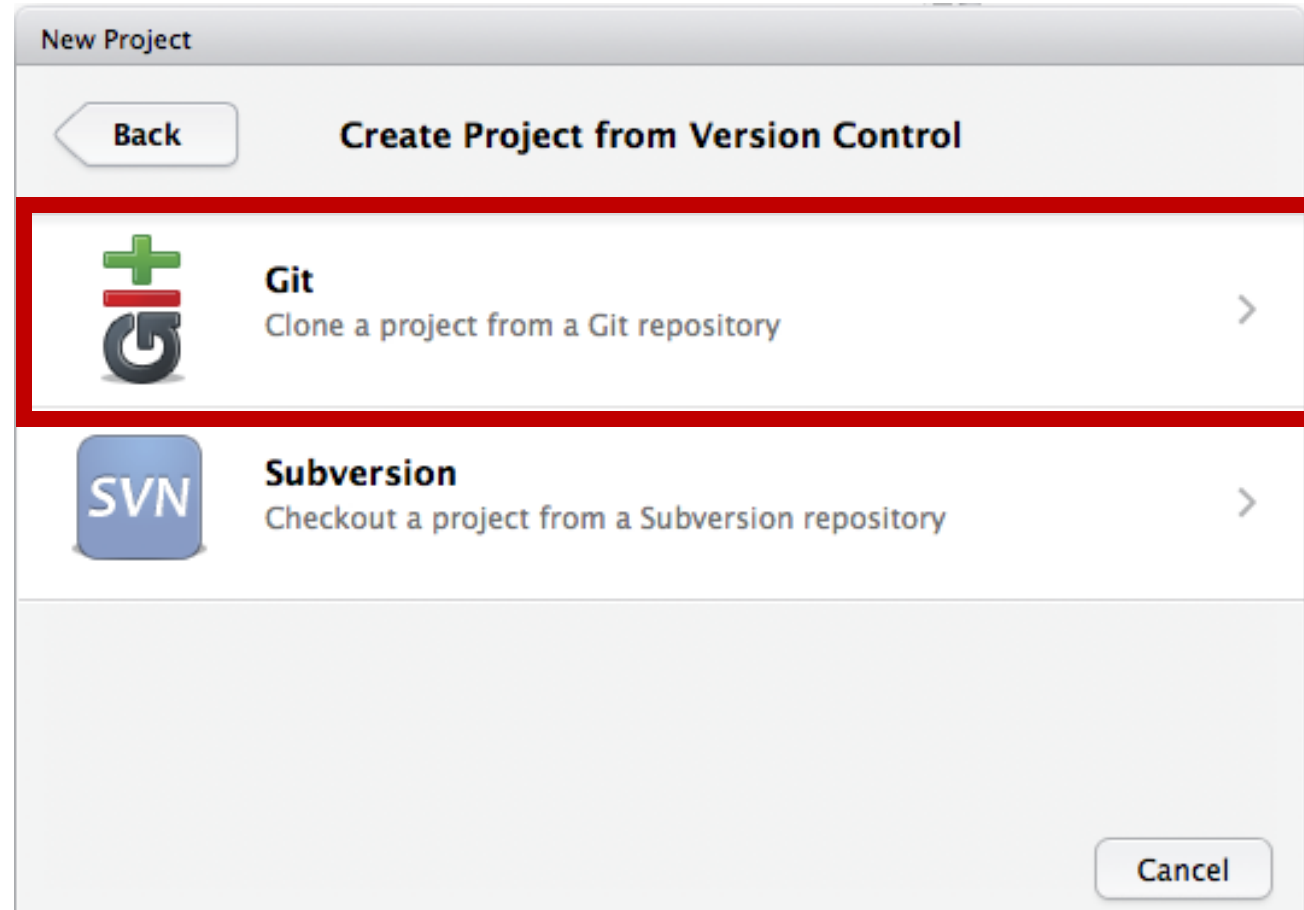
Select Version Control

File > New Project →

- no version control
- auto-detects existing version control
- **creates a version-controlled directory**




Select git



Link to GitHub

New Project

[Back](#) **Clone Git Repository**



Repository URL:

Project directory name:

Create project as subdirectory of:
 [Browse...](#)

☐ Open in new session

[Create Project](#) [Cancel](#)

- URL from github
- New dir name
- Parent directory

Get GitHub URL

The screenshot shows the GitHub repository page for 'maglet / fac-cred'. The repository has 1 commit, 1 branch, 0 releases, and 0 contributors. The 'Clone or download' button is highlighted with a red box. Below the button, the 'Clone with HTTPS' option is selected, and the URL 'https://github.com/maglet/fac-cred.git' is displayed in a text box, also highlighted with a red box. The repository contains files: 'data', '.gitignore', and 'fac_cred_analysis.R'. A prompt at the bottom encourages adding a README.

maglet / fac-cred

Unwatch 1 Star 0 Fork 0

Code Issues 0 Pull requests 0 Projects 0 Wiki Insights Settings

No description, website, or topics provided. Edit

Add topics

1 commit 1 branch 0 releases 0 contributors

Branch: master New pull request Create new file Upload files Find file Clone or download

Clone with HTTPS Use SSH

https://github.com/maglet/fac-cred.git

Open in Desktop Download ZIP

Add a README


• Click

• Copy

Paste GitHub URL into RStudio

New Project

Back Clone Git Repository



Repository URL:

Project directory name:

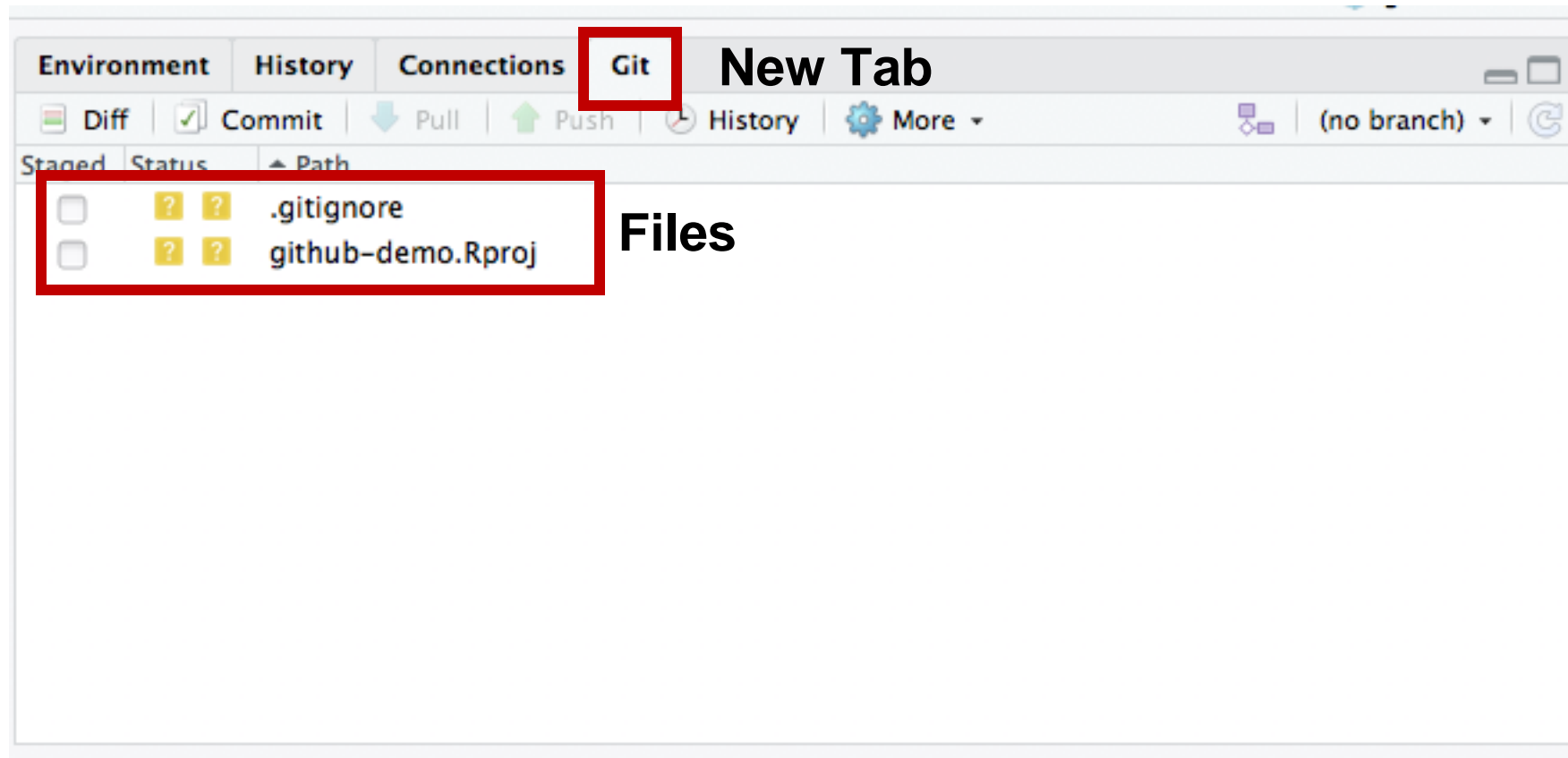
Create project as subdirectory of:
 Browse...

☐ Open in new session

Create Project Cancel

- Paste
- Match git repo name
- Pick

Upper right panel should look like this:



Untracked

Edited files recognized by git,
not acted upon

A yellow square containing a white question mark, located in the bottom right corner of the Untracked quadrant.

Staging Area

Edited files to be added to
your repository

A green square containing a white letter 'A', located in the bottom right corner of the Staging Area quadrant.

Ignored

Files you don't want to
go in your repository

Repository

Files added to your repository

Workspace

Edited files recognized by git,
not acted upon

ignore

Ignored

Files you don't want to
go in your repository

Add / stage

Staging Area

Edited files to be added to
your repository

commit

Repository

Files added to your repository



Workspace



.Rproj



.gitignore

(created by RStudio)

?

Staging Area

A

Ignored

Repository

README.md

(from GitHub)

Workspace



.Rproj .git

(created by RStudio)



Staging Area

Goal

Commit all changes to the repository
Empty workspace and staging area

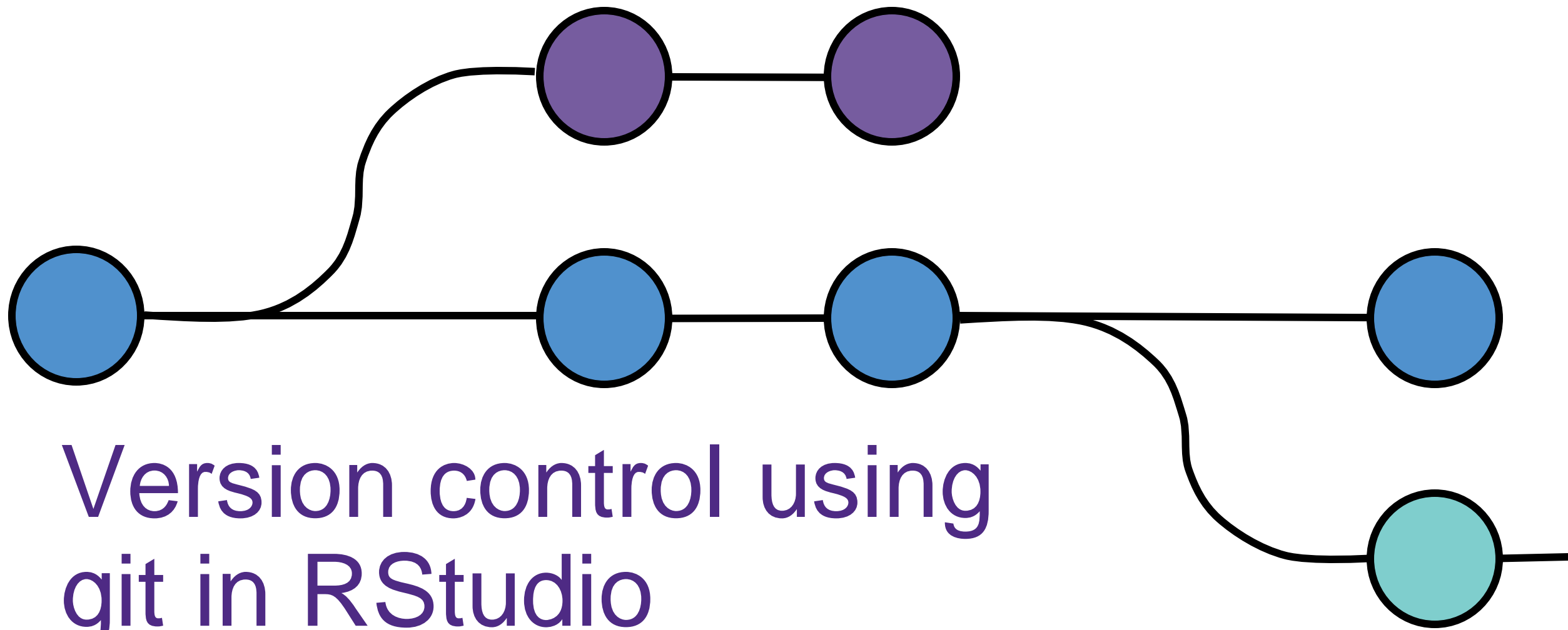
Ignore

Repository



README.md

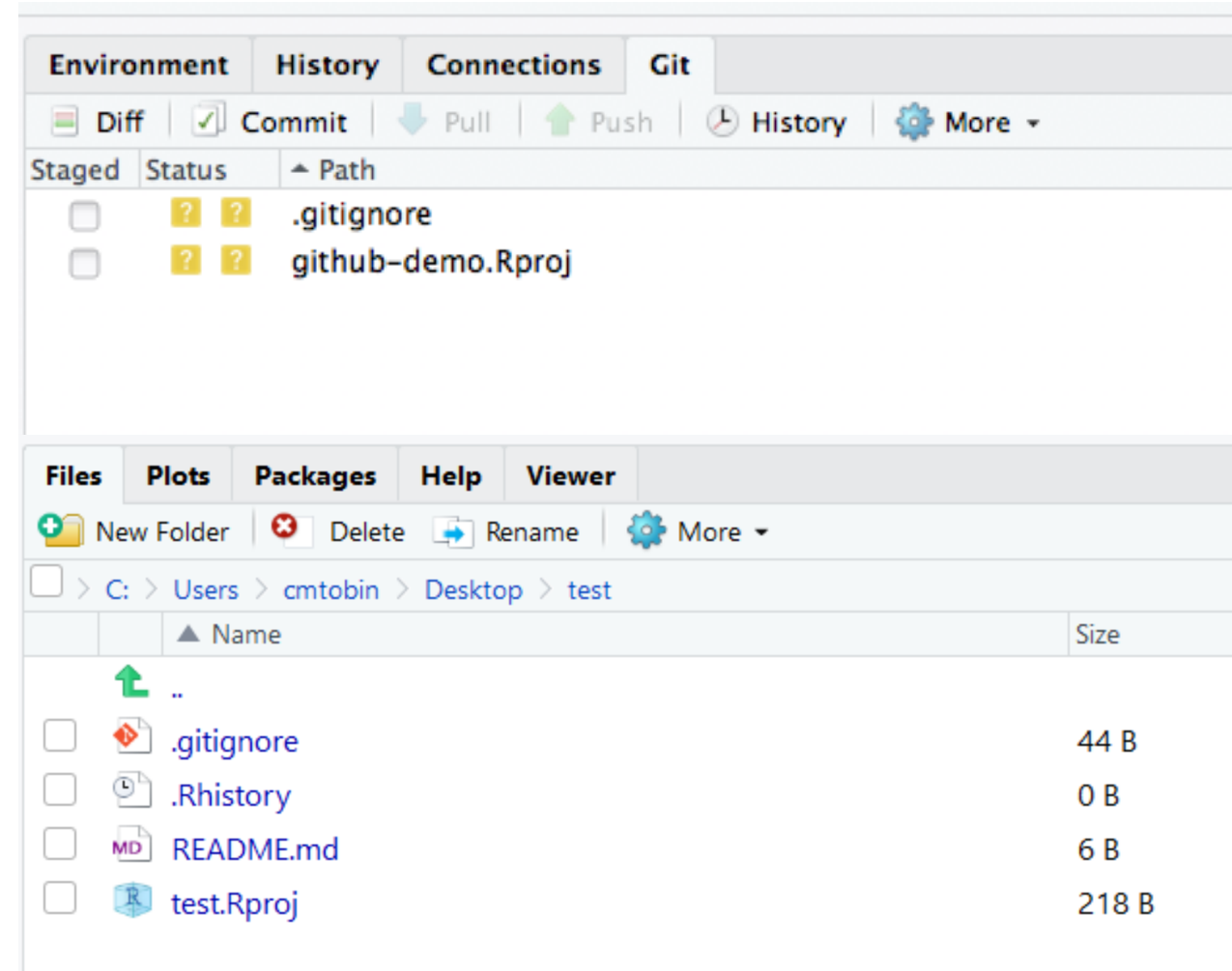
(from GitHub)



Version control using
git in RStudio

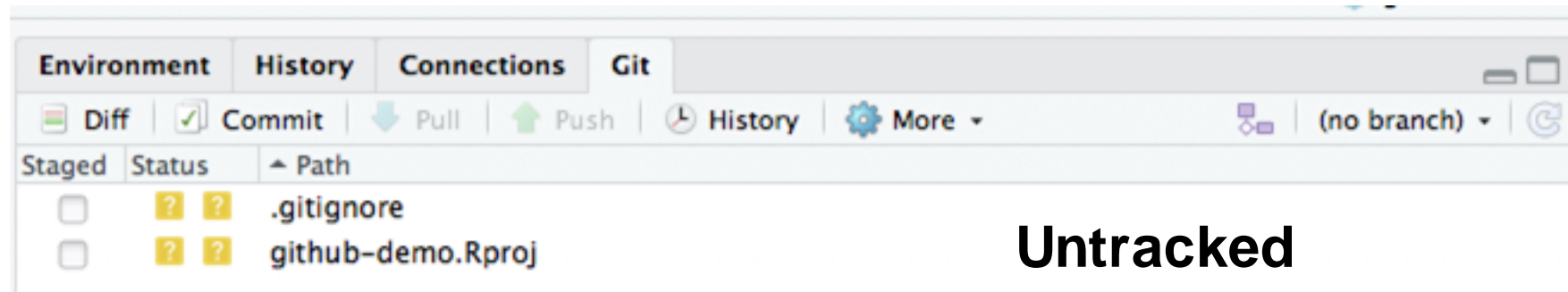
Setup results

- Downloaded files from Github
 - README.md
- New files added to your project folder
 - .gitignore, .Rproj, .Rhistory
- Ignore files
 - .Rhistory, .Rproj.user, .Rdata, .Ruserdata



Git tab

- Only files with uncommitted changes appear
- README.md was already in the repository (committed)
- .Rhistory is being ignored (will address this later)



Workspace



.Rproj



.gitignore

?

Staging Area

A

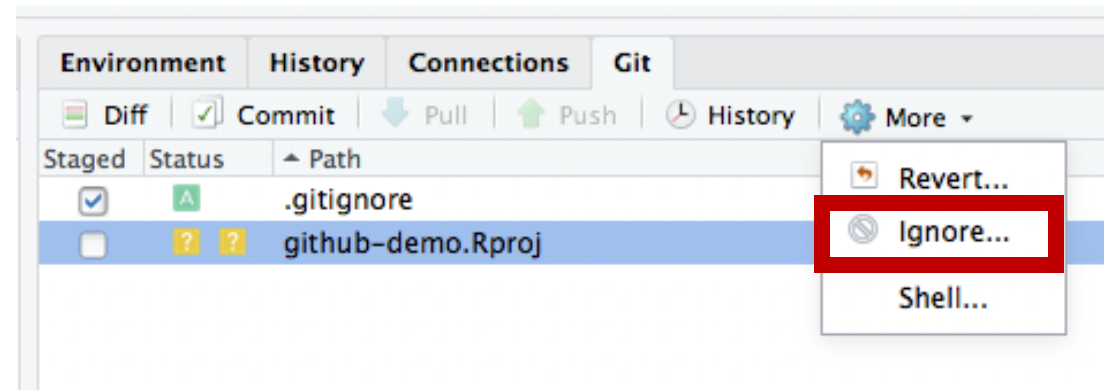
Ignored

Repository

README.md

Ignoring files

- Git by default “sees” all files in its directory unless you tell it not to
- Add the filename to the .gitignore file to stop tracking



Types of files git works best with

“Good” formats (text based)

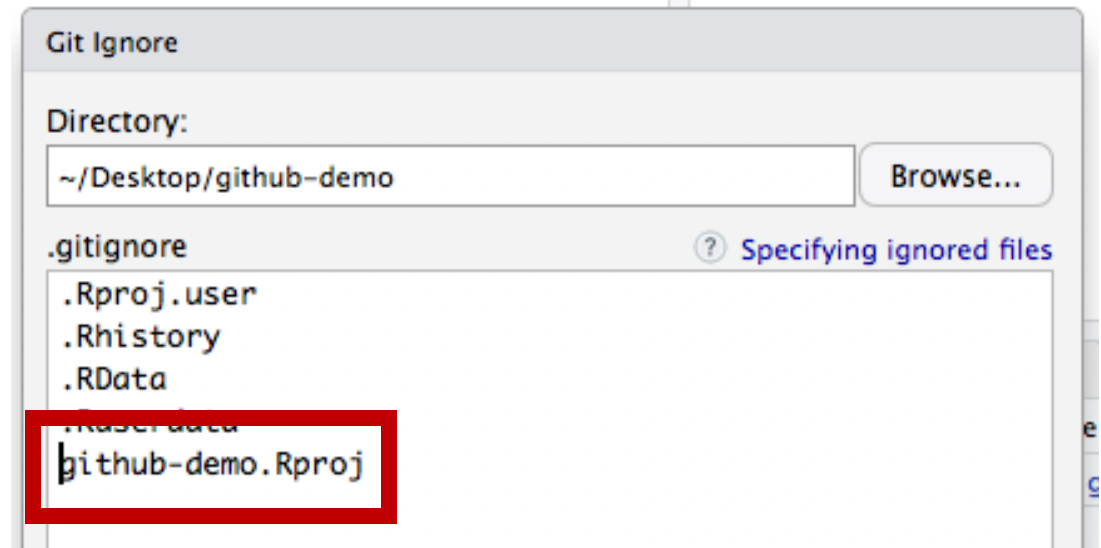
- Documents (.txt, .tex, .rtf, .Rmd)
- Tabular Data (.csv)
- Source code (.R, .py, .c, .sh)

“Bad” formats (binary files)

- Documents (.docx, .pdf, .ppt)
- Excel spreadsheets (.xlsx)
- Media (.jpg, .mp3, .mp4)
- Databases (.mdb, .sqlite)

Ignoring files

- Added the R.proj file to the ignore list



Workspace



.gitignore

ignore

?



Ignored



.Rhistory

Staging Area

A

Repository

README.md

Workspace



.gitignore

?

Staging Area

A

Ignored



.Rhistory



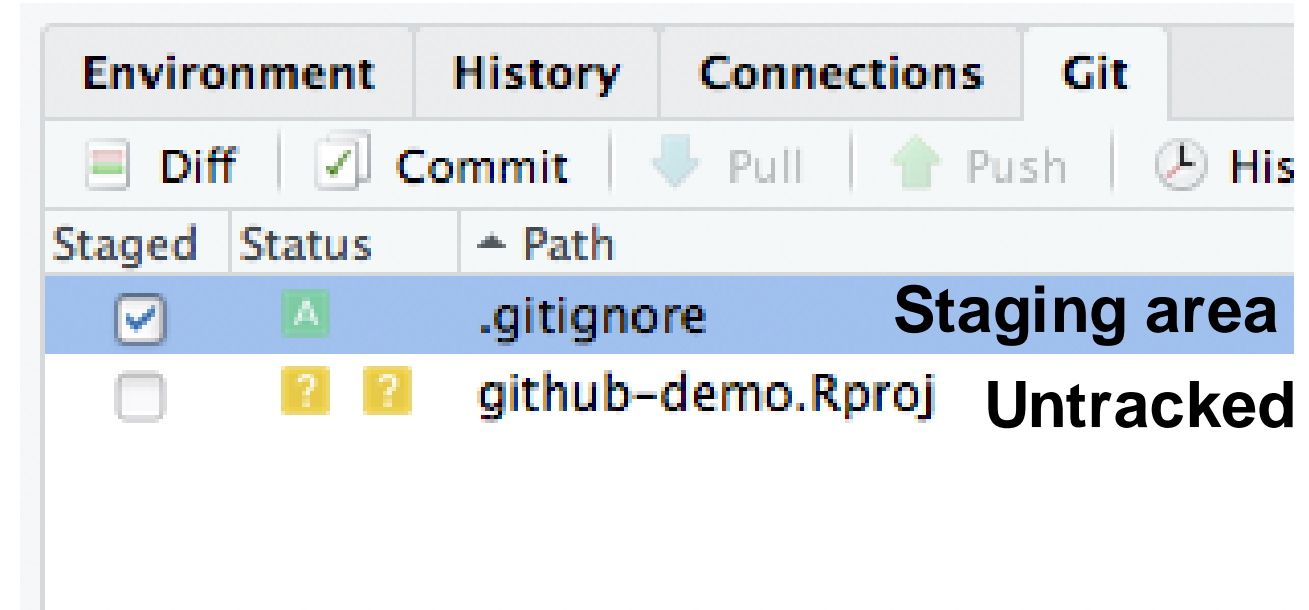
.Rproj

Repository

README.md

Stage (Add) files

- Check boxes to add files to the “staging area”
- **Status:** Staged
- Staging area compiles changes to be committed to the repository



Workspace



.gitignore

Add / stage



Staging Area

?

A

Ignored



.Rhistory



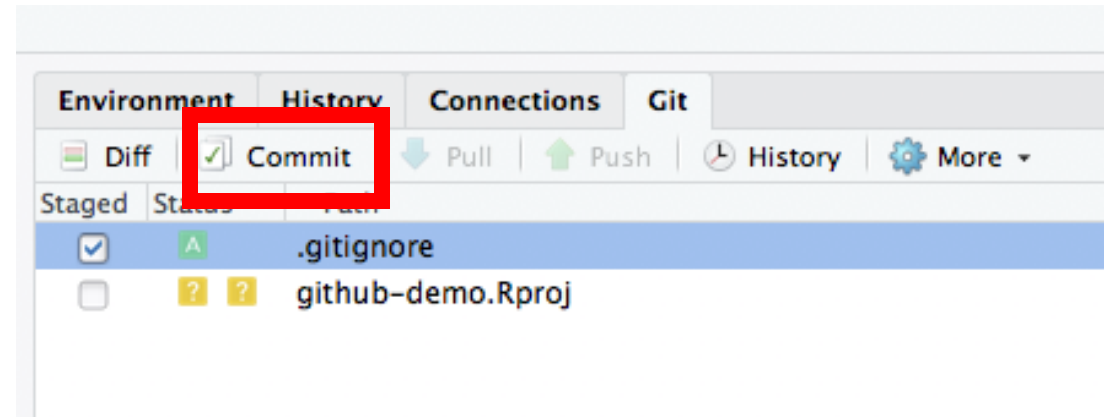
.Rproj

Repository

README.md

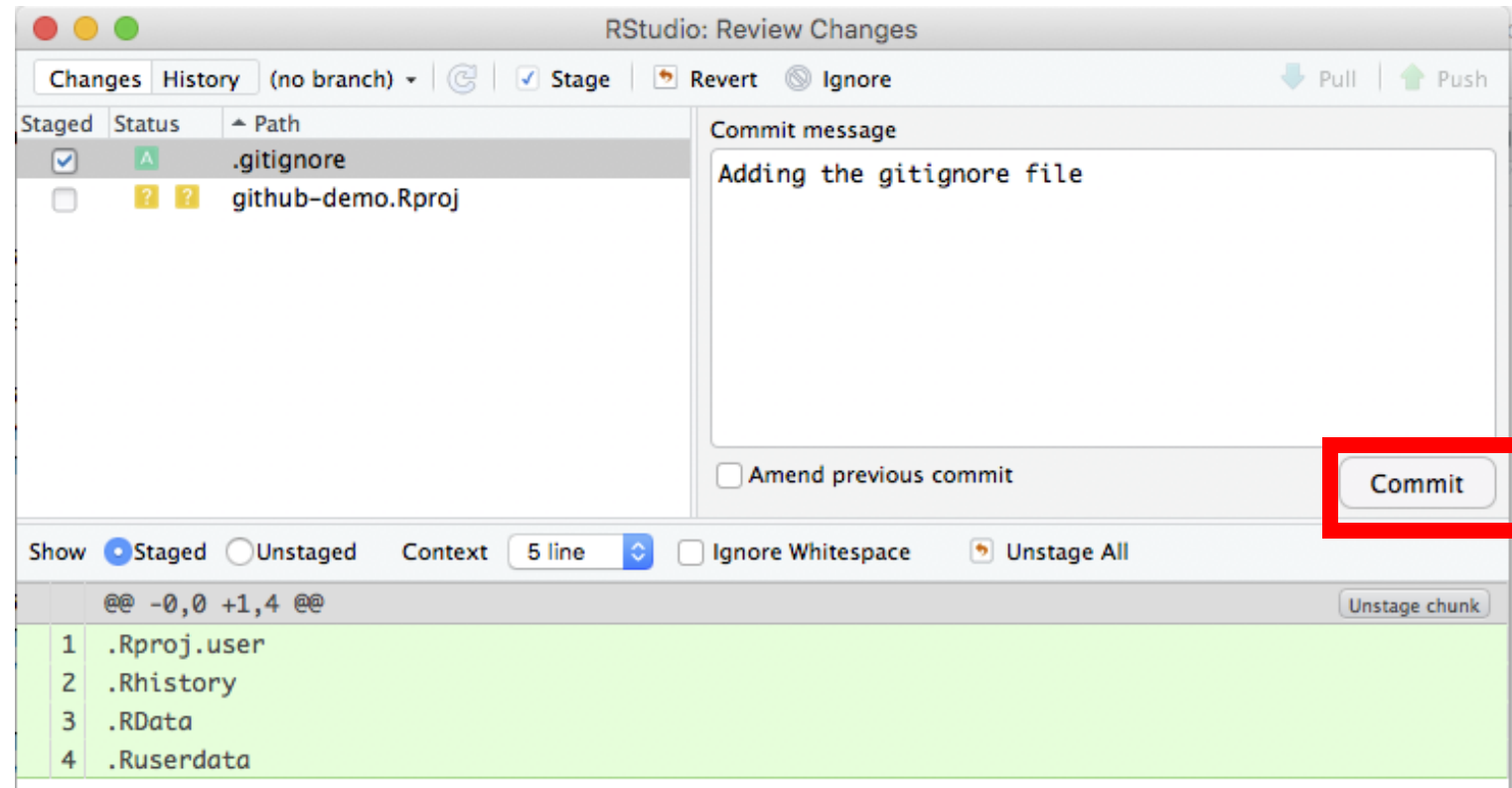
Commit files

- Clicking “commit” adds all files from the staging area to the repository
- Opens a new window



Commit message

- Add a message to say what changed in this commit
- Can see the contents of the file and what changed
 - Green = added
 - Red = removed



Workspace

Staging Area



.gitignore

?

commit

A



Ignored



.Rhistory



.Rproj

Repository

README.md

Workspace

Staging Area

?

A

Ignored



.Rhistory



.Rproj

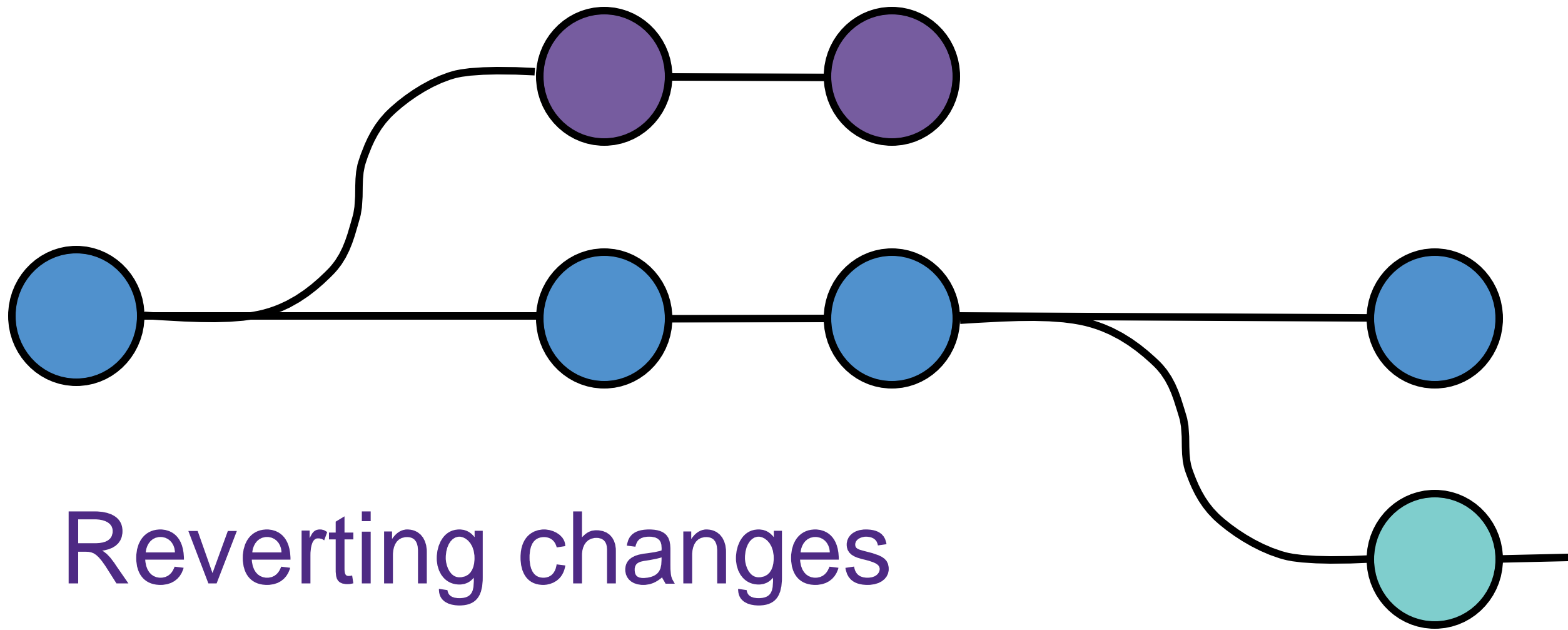
Repository

.gitignore

README.md

Exercise:

- Edit your README.md file
- **Stage changes: Add** the changes file to the **staging area**
- **Commit** your staging area to the **repository**



History

The screenshot shows the 'RStudio: Review Changes' window. At the top, there are tabs for 'Changes' and 'History', with 'History' selected. Below the tabs, there's a search bar and a 'Pull' button. The main area displays a list of commits with columns for Subject, Author, Date, and SHA. The first commit is highlighted in blue. Below the list, there are navigation controls and a summary of the selected commit. At the bottom, a diff view for the '.gitignore' file is shown, with line 5 highlighted in green.

Subject	Author	Date	SHA
HEAD -> refs/heads/master origin/master origin/HEAD ignored .Rproj	Tobin Magle <cmtobin@wisc.edu>	2019-03-25	c4b0d865
Adding the .gitignore file	Tobin Magle <cmtobin@wisc.edu>	2019-03-22	8c56ce95
Initial commit	maglet <ctobmag@gmail.com>	2019-03-22	ec9ad922

Commits 1-3 of 3

SHA c4b0d865
Author Tobin Magle <cmtobin@wisc.edu>
Date 2019-03-25 22:03
Subject ignored .Rproj
Parent 8c56ce95

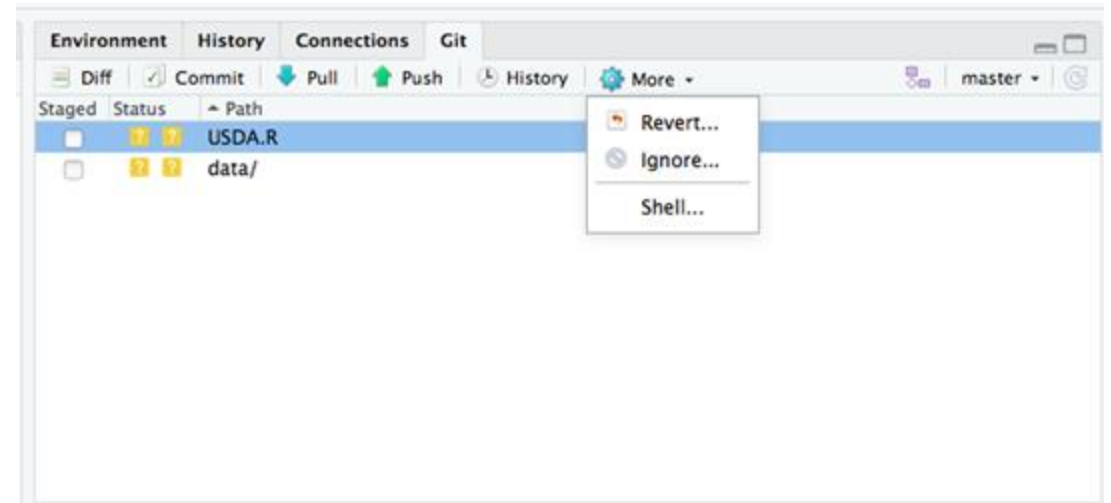
.gitignore

```
@@ -2,3 +2,4 @@
2 2 .Rhistory
3 3 .RData
4 4 .Ruserdata
5 5 test.Rproj
```

View file @ c4b0d865

Revert

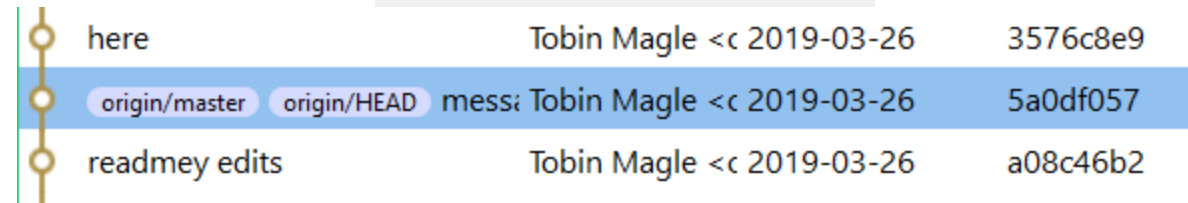
- Removes all uncommitted changes
- Includes staged changes
- Can't revert to committed changes from R Studio



Go back to previous commit

- Need to use the terminal
- `git checkout <commit_id> <filename>`
- File will show as modified, add and commit like normal

Find Commit ID

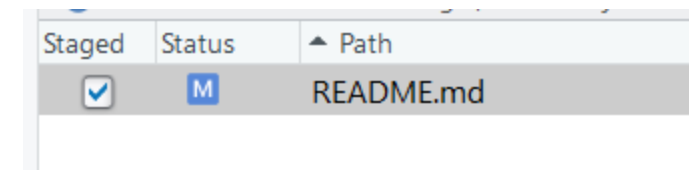


here	Tobin Magle <c 2019-03-26	3576c8e9
origin/master	origin/HEAD mess: Tobin Magle <c 2019-03-26	5a0df057
readmey edits	Tobin Magle <c 2019-03-26	a08c46b2

Run the checkout command in the terminal

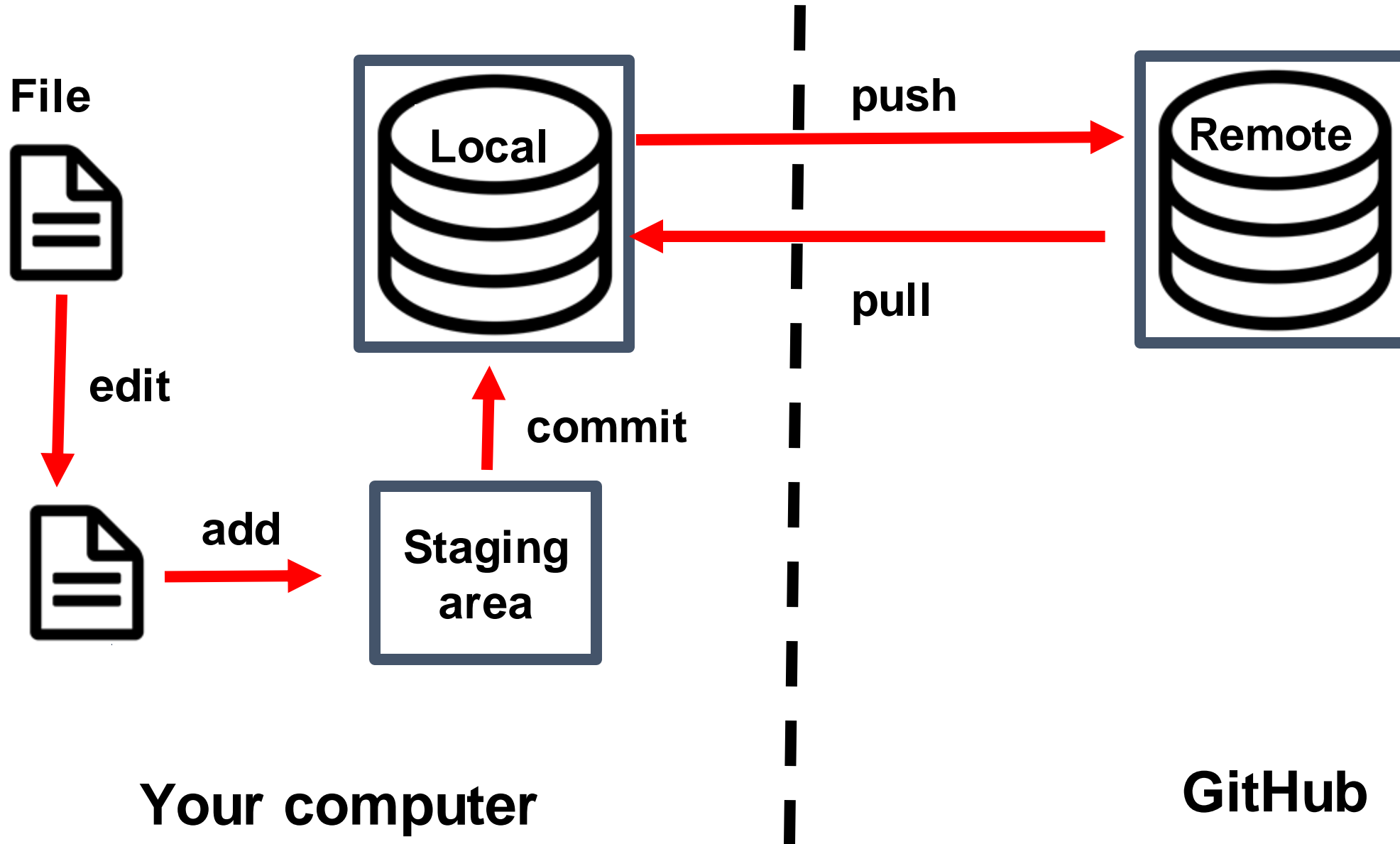
```
cmtobin@GLS-LAP8228 ~/Desktop/test2 (master)  
$ git checkout 5a0df057 README.md
```

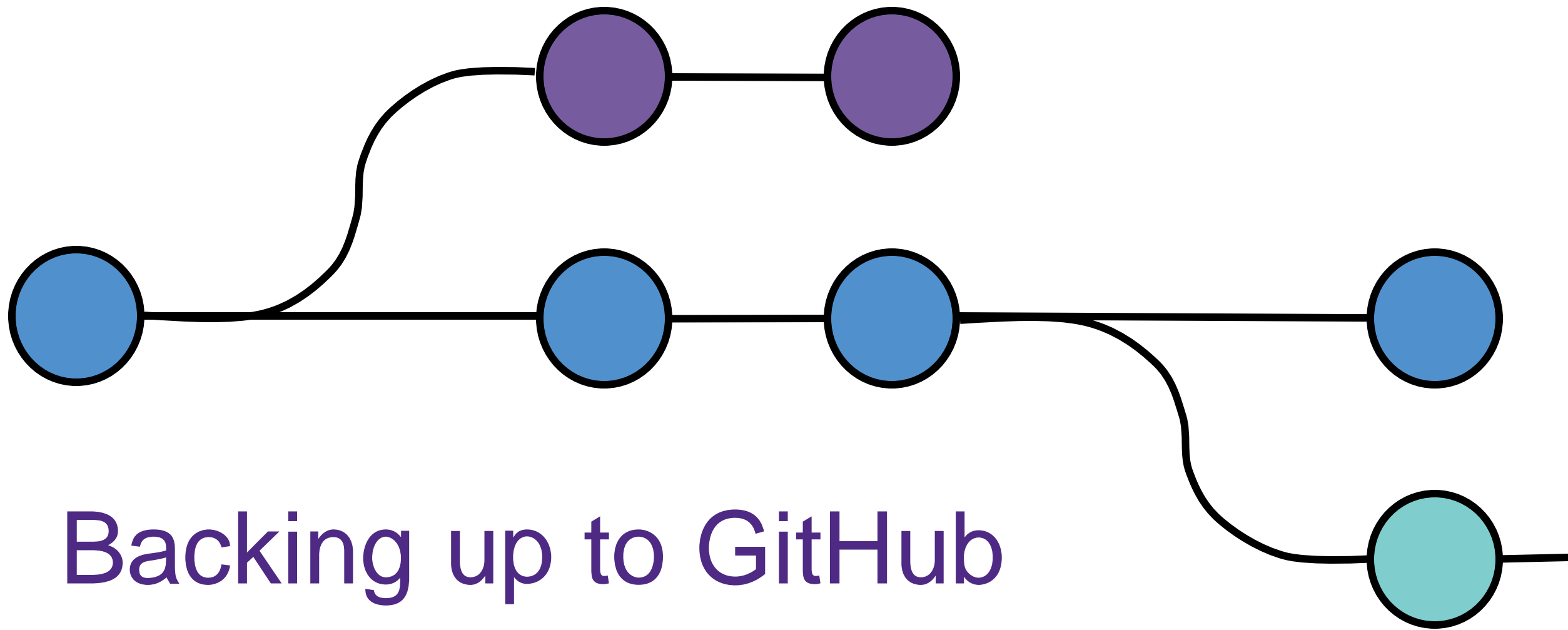
file has been modified



Staged	Status	Path
<input checked="" type="checkbox"/>	M	README.md

Git workflow

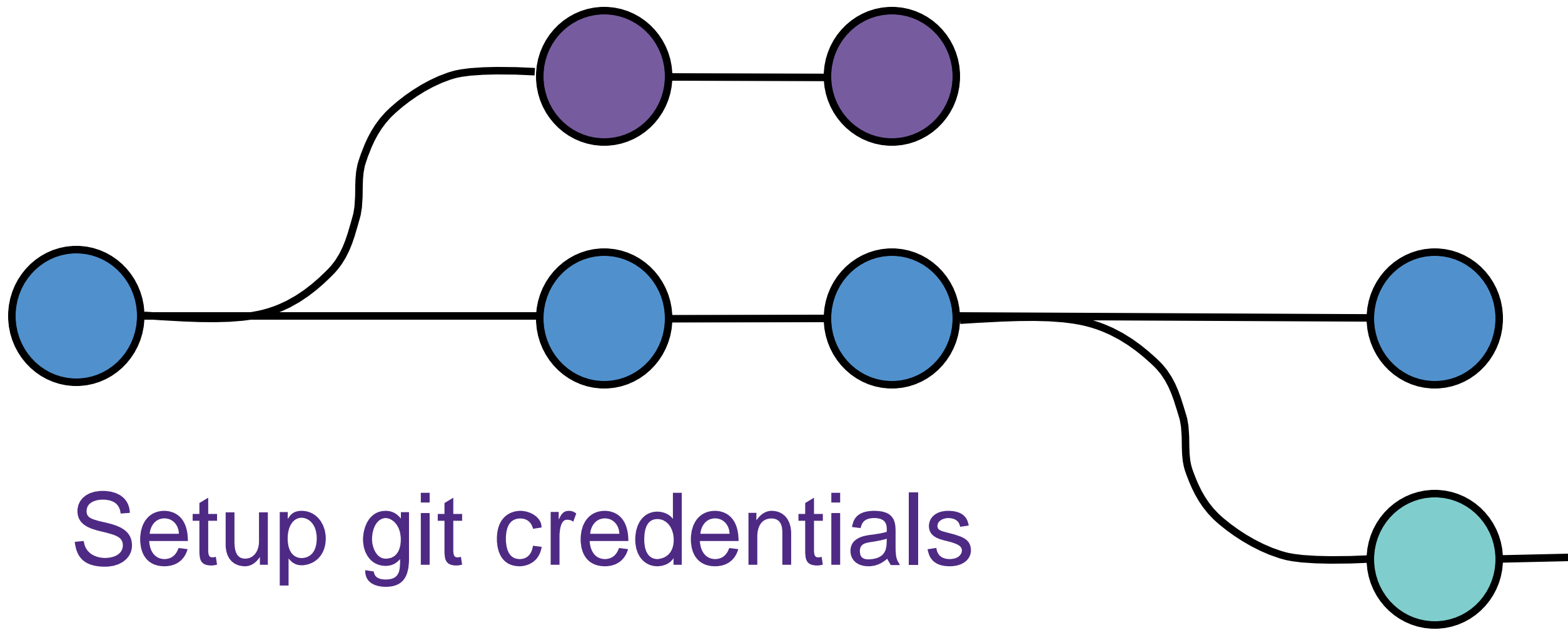




GitHub best practices

- Store **code and documentation**, not data
- Never store **protected or personal data** even in a private repo
- Store **credentials** (private keys or tokens) in a separate file outside of your repository
 - ie, not in your code
- Add a **personal email** to your GitHub account, so you can use it when you leave Northwestern





Setup

During the workshop

- GitHub: Create a personal access token
- RStudio:
 - Install usethis package
 - **Configure** your git user
 - Set up **git authentication**

Configure git on your computer

In RStudio console

1. Install and load `usethis`

```
install.packages("usethis")  
library(usethis)
```

2. Set git username and email

- Commit history shows username
- **Use your GitHub account email**

```
usethis::use_git_config(user.name="Jane Doe",  
user.email="jane@example.org")
```

3. Set GitHub authentication

```
usethis::create_github_token()
```

Create GitHub Token

This page will open

Enter a description

Note exp date

Keep default scope



Generate token

More about scopes: <https://docs.github.com/en/apps/oauth-apps/building-oauth-apps/scopes-for-oauth-apps>

GitHub Personal Access Token

Copy token (it won't be visible once you leave the page)

Personal access tokens

Generate new token

Revoke all

Tokens you have generated that can be used to access the [GitHub API](#).

Make sure to copy your personal access token now. You won't be able to see it again!

✓ ghp_PFg4Gazp0FicBKmxsuGk8ppYxBXd41LwiaM 

Delete

downloading individual files programatically from metrics db — repo, workflow

Last used within the last 5 months

Delete

⚠ This token has no expiration date.

Personal access tokens function like ordinary OAuth access tokens. They can be used instead of a password for Git over HTTPS, or can be used to [authenticate to the API over Basic Authentication](#).

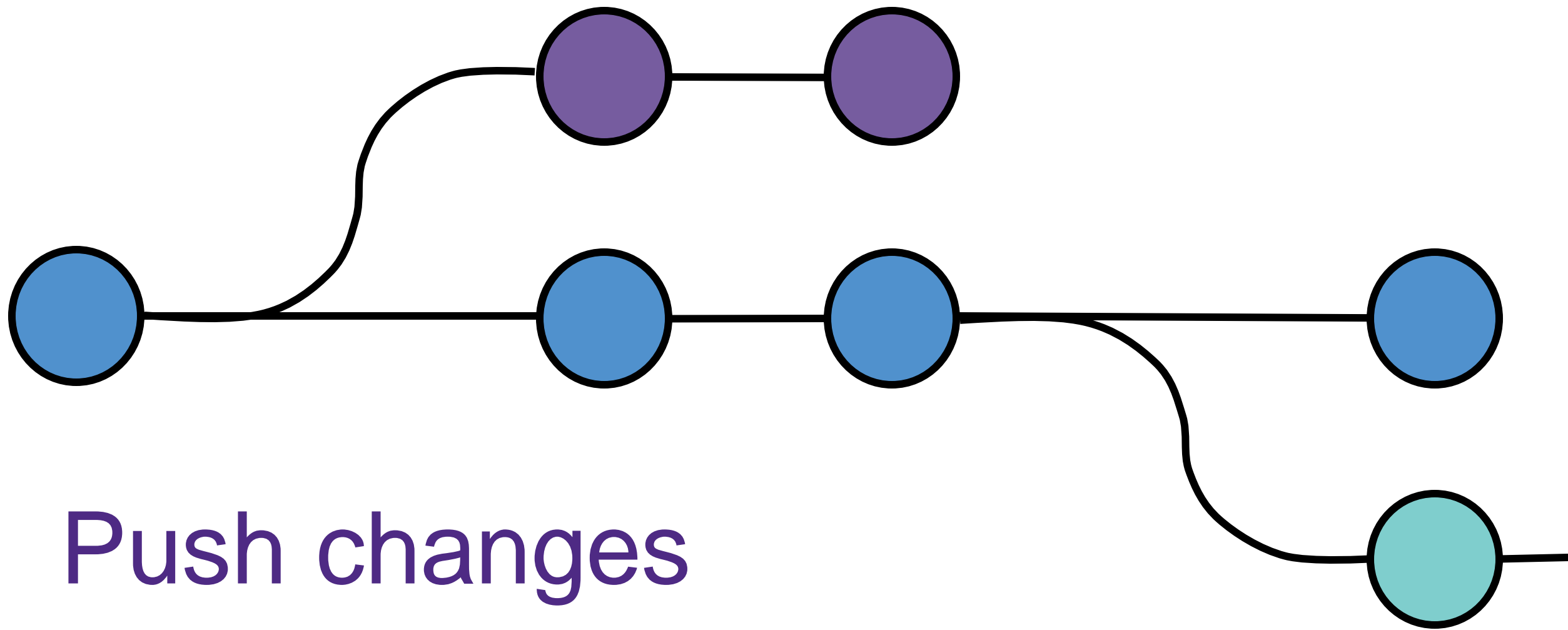
Set GitHub Credentials

In R Console:

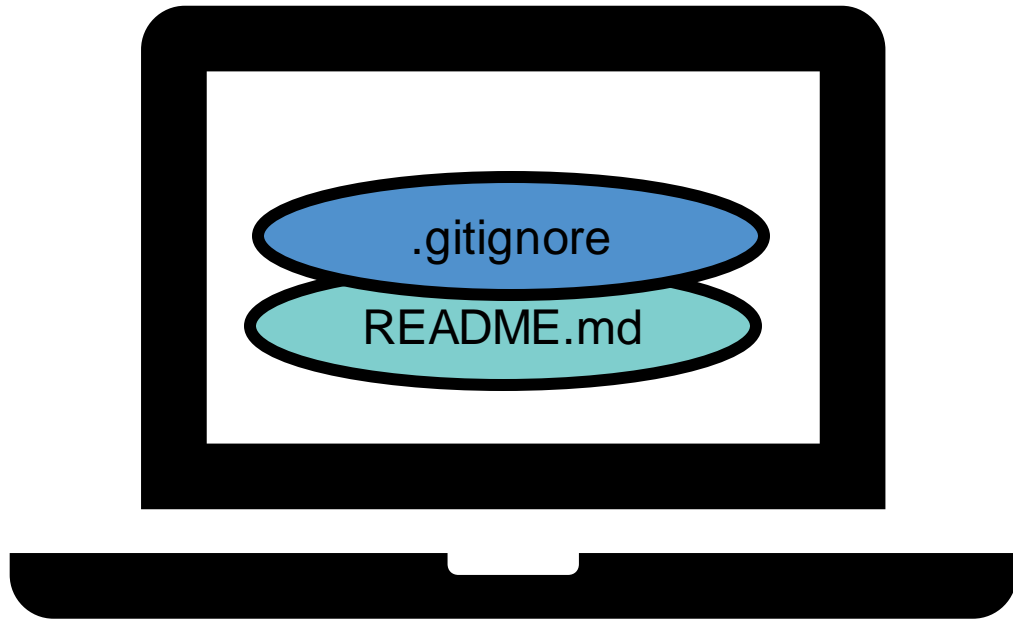
```
gitcreds::gitcreds_set()
```

Follow the prompts

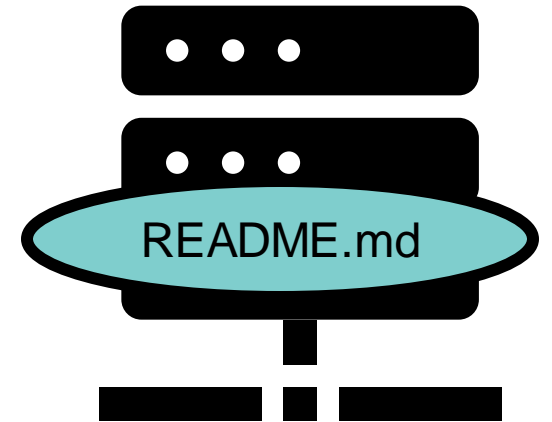
- Paste in the token when asked
- If you had existing credentials (including username/password which are no longer allowed), you may need to choose option 2: Replace these credentials
- git saves these credentials for you



Local changes not backed up

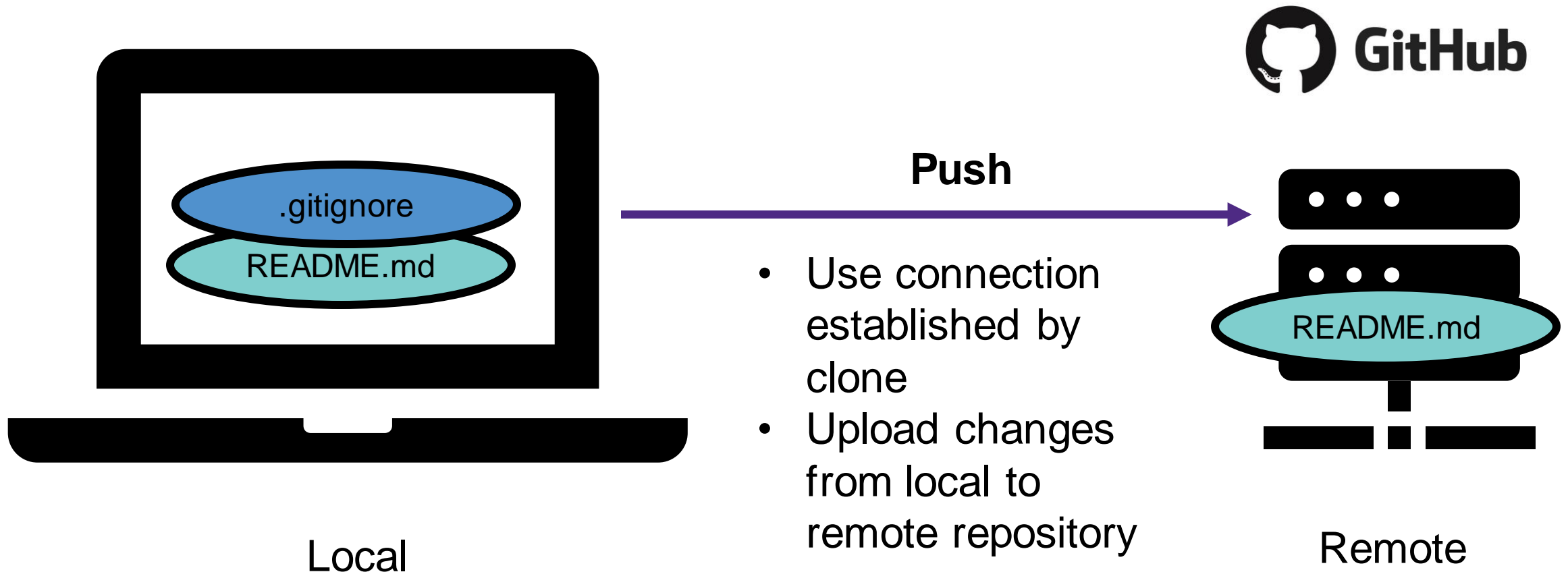


Local



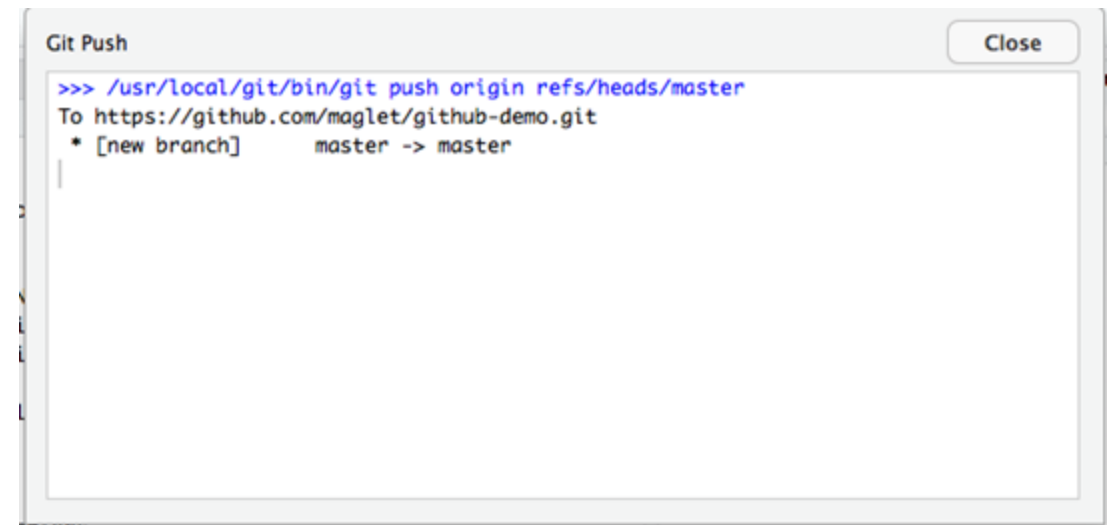
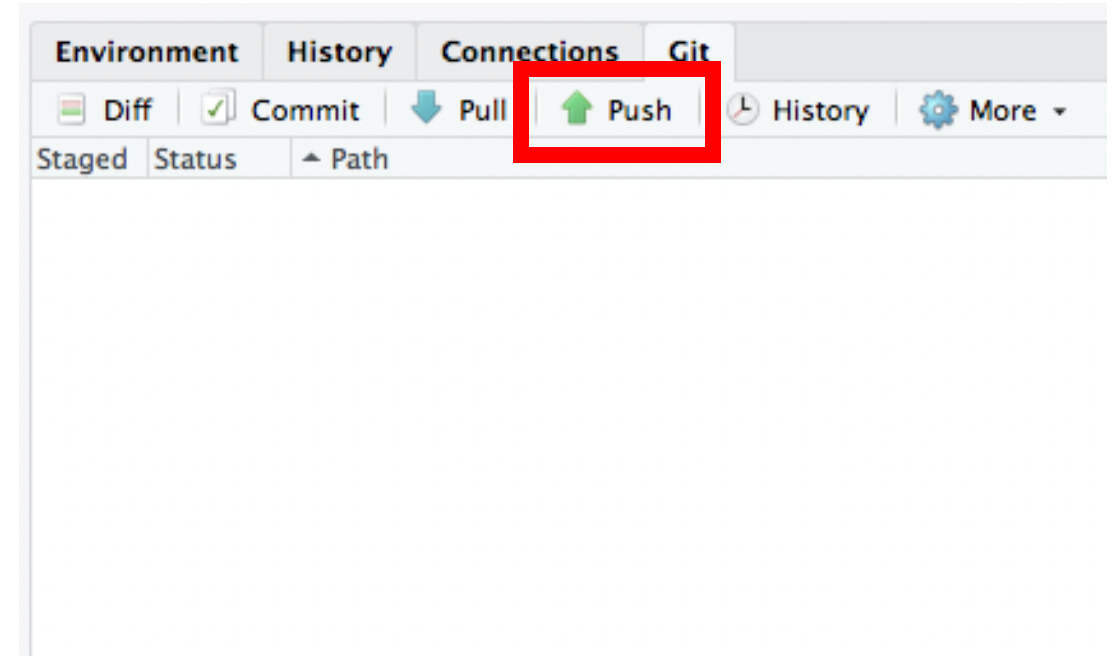
Remote

Push to remote

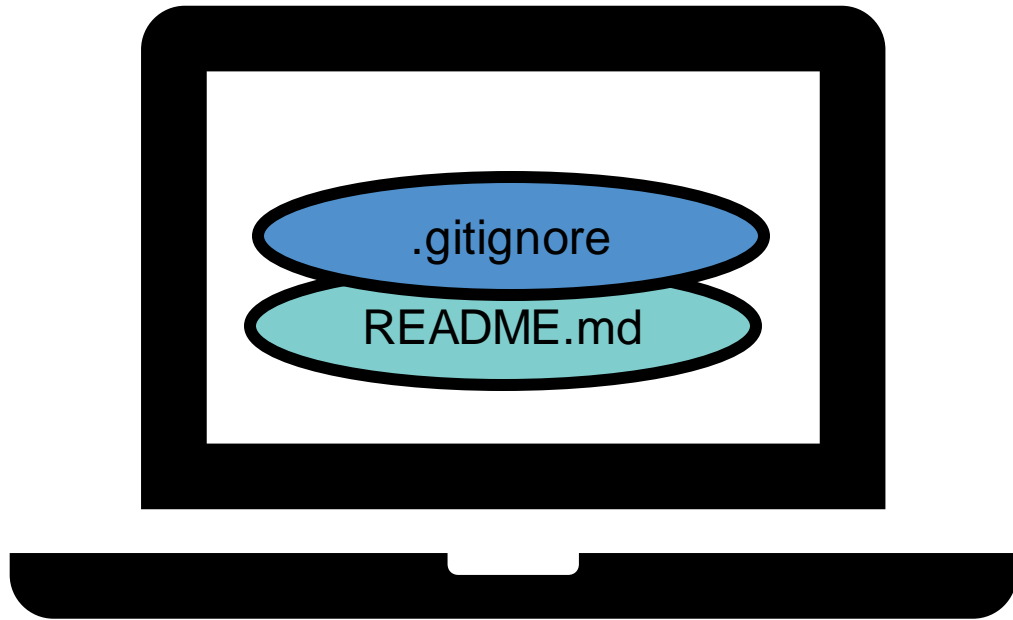


Push

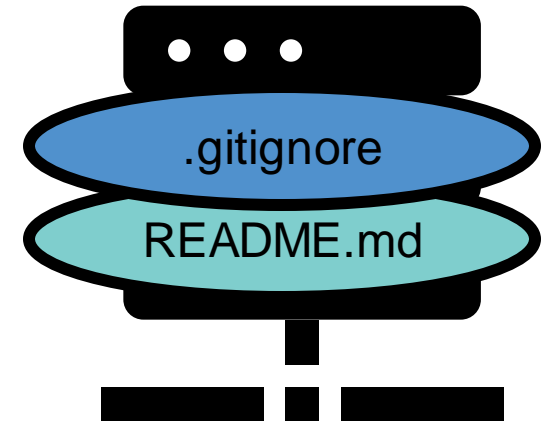
- Save local changes to remote repository
- Changes must be committed before pushing
- Requires that you have authentication set up on your machine



Changes pushed




Local



Remote

Files in remote

 **maglet / github-demo**

Unwatch 1

Star 0

Fork 0

<> Code

Issues 0

Pull requests 0

Projects 0

Wiki

Insights

Settings

No description, website, or topics provided.

[Add topics](#)

2 commits

1 branch

0 releases

0 contributors

Branch: master


New pull request

Create new file


Upload files

Find file

Clone or download

 Tobin Maple Added .Rproj to the .gitignore

Latest commit 4138947 10 minutes ago

 [.gitignore](#)

Added .Rproj to the .gitignore

10 minutes ago

Help people interested in this repository understand your project by adding a README.

Add a README

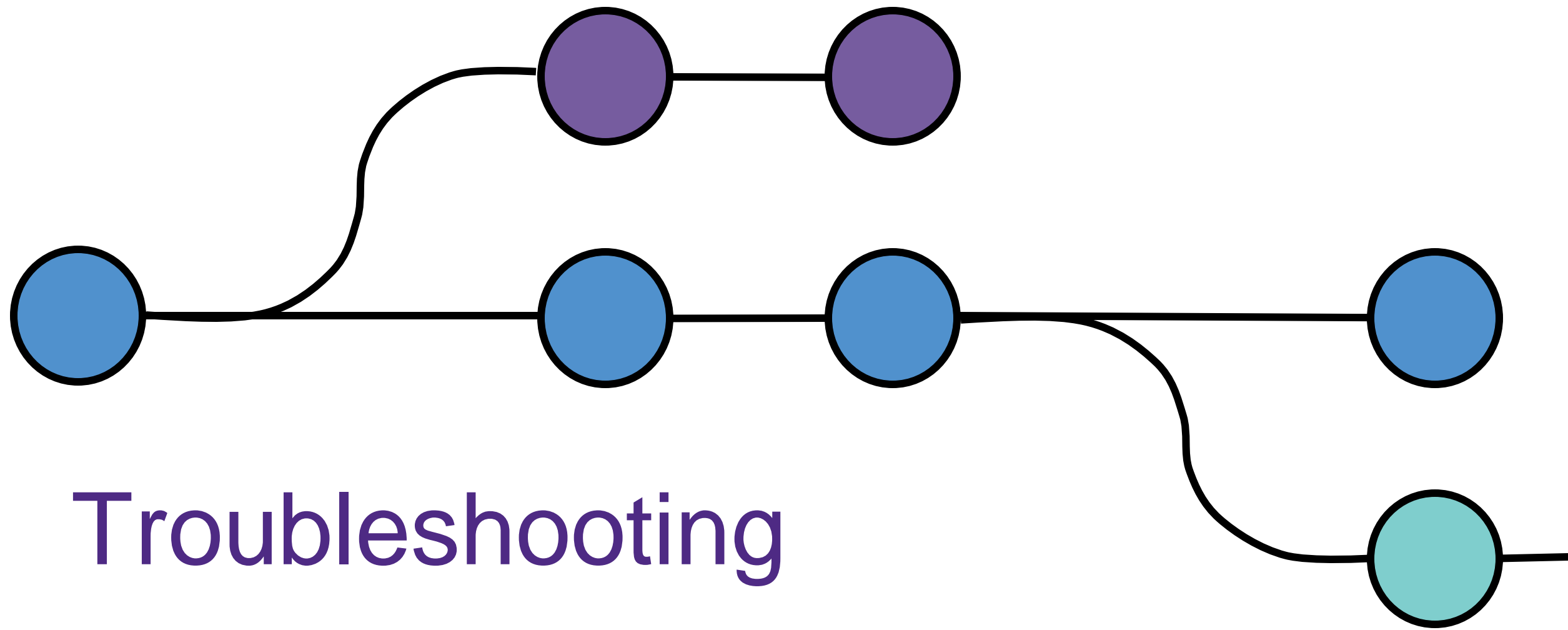
Practice:

- Make a new R script
- Save it to the R project directory
- Add some R code
- Do the steps necessary to get this file into your GitHub repo

Need help?

Workshop materials: <https://github.com/nuitrcs/git-RStudio>

- Request a consult: bit.ly/rcdsconsult
- Other Resources
 - [RCDS: Written Setup Instructions](#)
 - [RCDS: Other ways to set up Git repos in R](#)
 - What git is and how it works: [Git for Humans](#)
 - Why you should use git: [Academic benefits of using git](#)
 - How to use git with R and RStudio: [Happy git with R](#)



Troubleshooting

Q: I don't see the git tab

A: Check that RStudio sees git

Mac: RStudio > Preferences

Windows: Tools > Global Options

Mac: type

which git

in the Terminal to get the path

