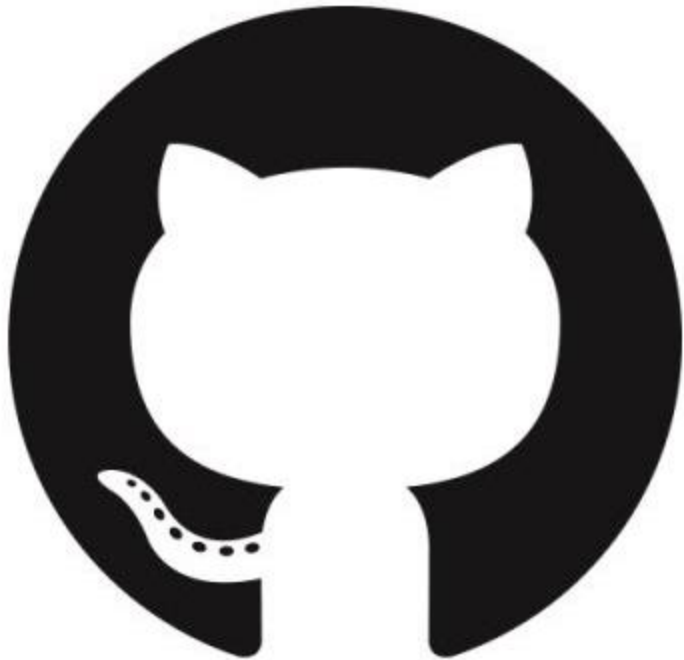


Git and Github

Code/slides: Joel Swift



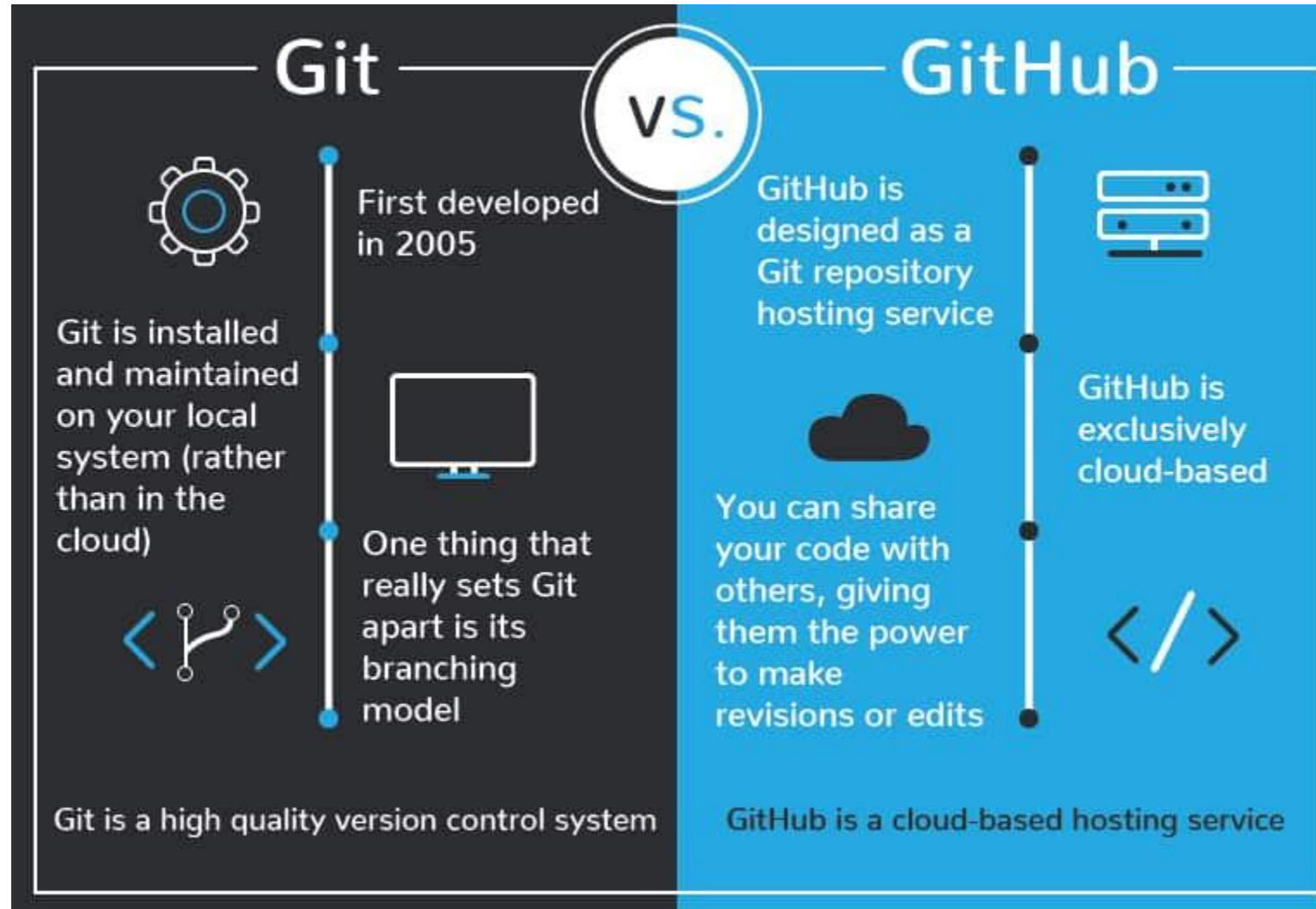
git



GitHub

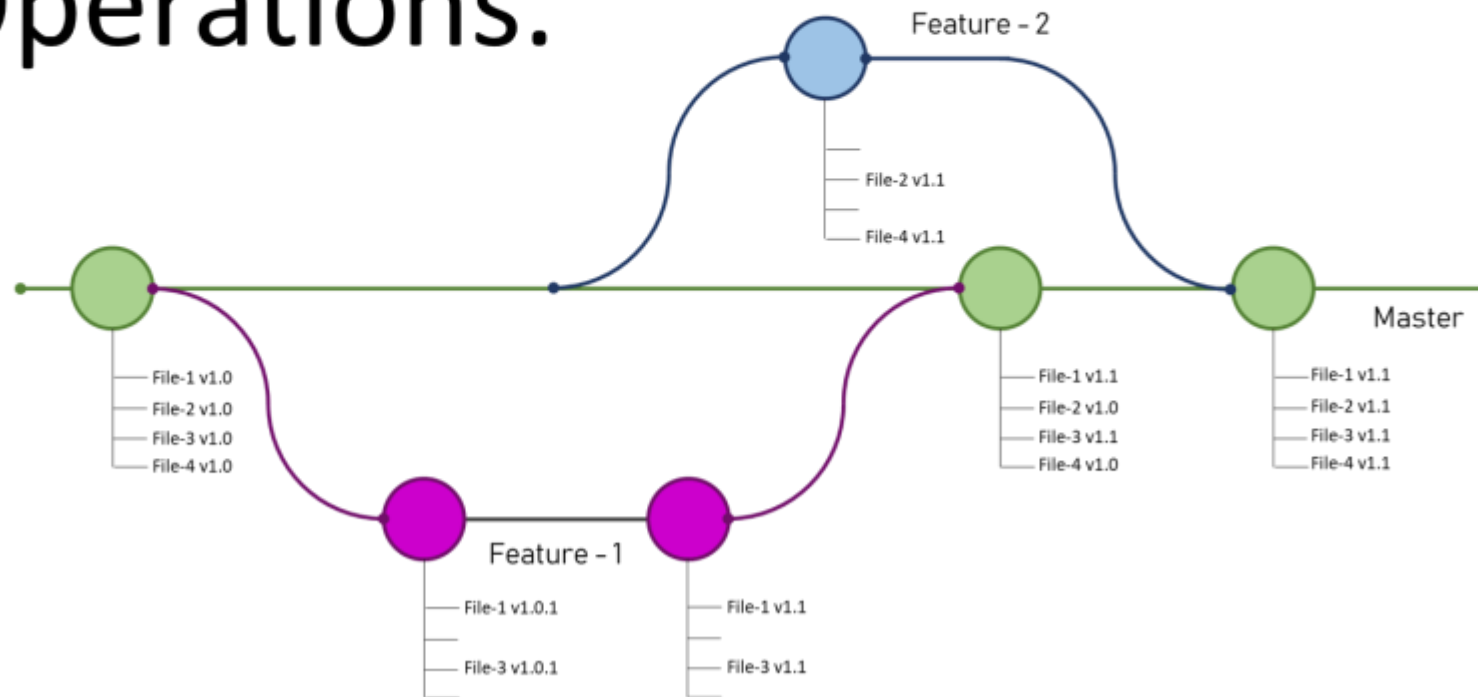
What the hell are git and github?

- Git is a version control system.
 - Basically it allows for branched workflows, where multiple branches of a file or code can co-exist and be merged.
- Github
 - A cloud based hosting service for hosting git projects (repositories, repos).
 - GUI support.
 - Was purchased by Microsoft for 7.5 billion in !

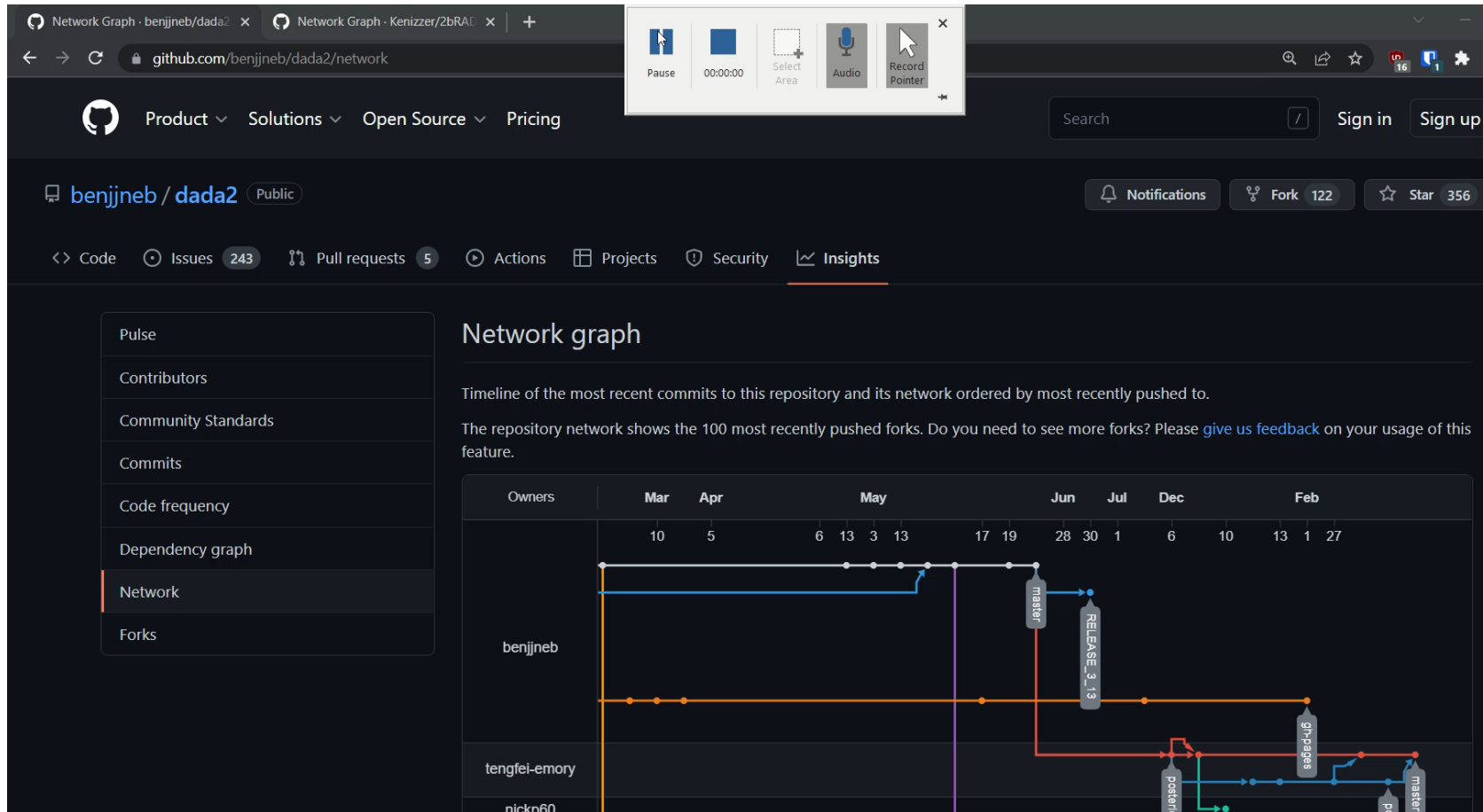


Branching model of Git

GIT Branch and its Operations.



Real world examples (dada2 and my repo)



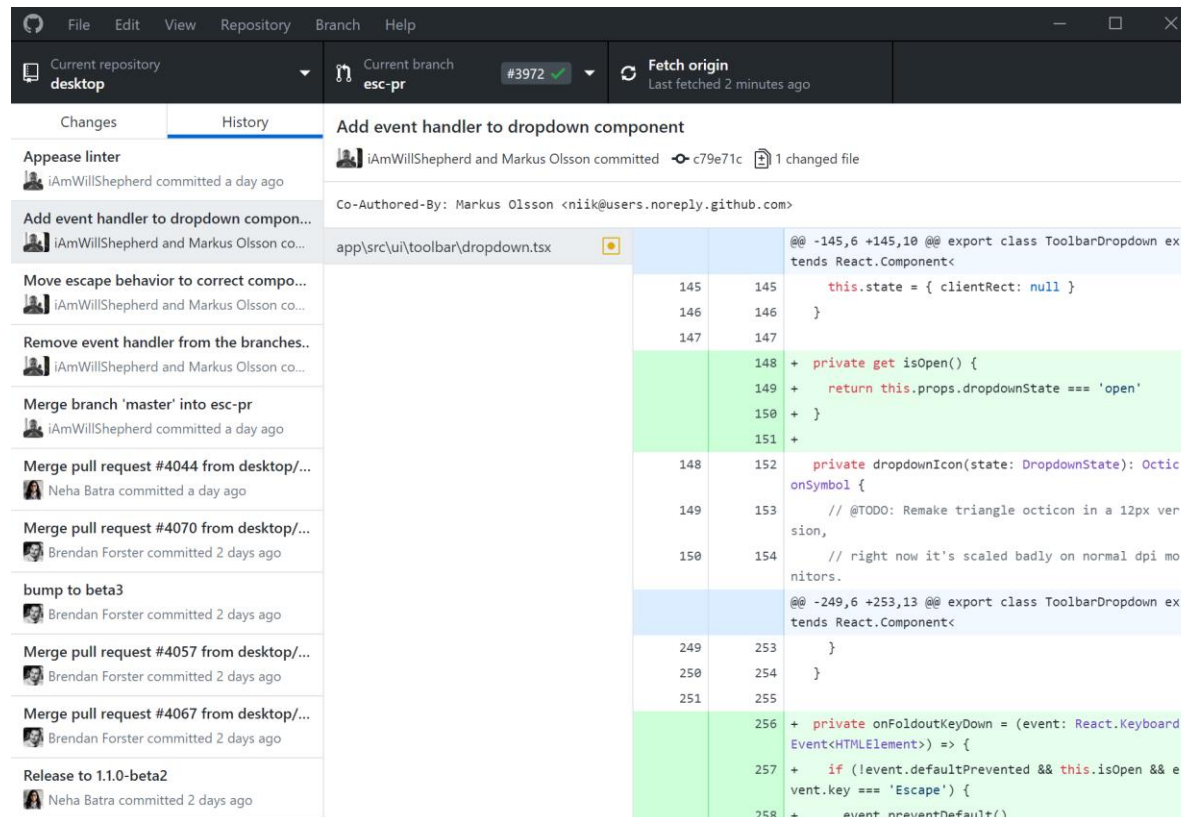
Github... make an account

- Make an account
- Setup 2FA (w/ mobile or hardware if you have it)

Github Desktop



- <https://desktop.github.com/>



Why use github?

- Collaborative coding is greatly simplified in a git/github environment.
- To share the code required to ensure your analysis is reproducible.
- Its often a journal requirement prior to publication.
- To better the environment for others that come after you, sometimes the solution to ones problems have been solved before...

```
### 30 year average (1990-2021) of precipitation for KS counties
county_precip_KS <- read.csv('KS_precipitation1895to2021.csv', sep=
which(colnames(county_precip_KS)=="X1990") # index 99
which(colnames(county_precip_KS)=="X2021") # index 130
# Average the rainfall annually for 1990-2021
county_precip_KS$AVG_30_years <- rowMeans(county_precip_KS[,99:130])
# Make a simpler data frame for plotting
county_fip_30yrAVG <- data.frame('county_fips' = county_precip_KS$county_fips,
# Fips # to character to play nice with plotting function
county_fip_30yrAVG$county_fips <- as.character(county_fip_30yrAVG$county_fips)
# Sanity check
county_fip_30yrAVG[1:5,]
summary(county_fip_30yrAVG)

# Color palette
# Taken from the colors in https://climate.k-state.edu/basics/
colfunc<-colorRampPalette(c("#fffe7a",
```

Fitzpatrick, Connor R., et al. "Assembly and ecological function of the root microbiome across angiosperm plant species." *PNAS* 115.6 (2018)

```
1591 ##### 4) Differential abundance testing #####
1592 # Thank you to Maggie Wagner for making her R code publically available! #
1593 # Many of the analytical operations were based on her code #
1594 # Also thank you to Michael Love for patient responses on the DESeq
1595 # Will use DESeq2 to estimate fold change differences across factors and test
1596 # for differential abundance.
```

Repositories/Repos on github

https://github.com/Kenizzer/California_Transect_Microbiome

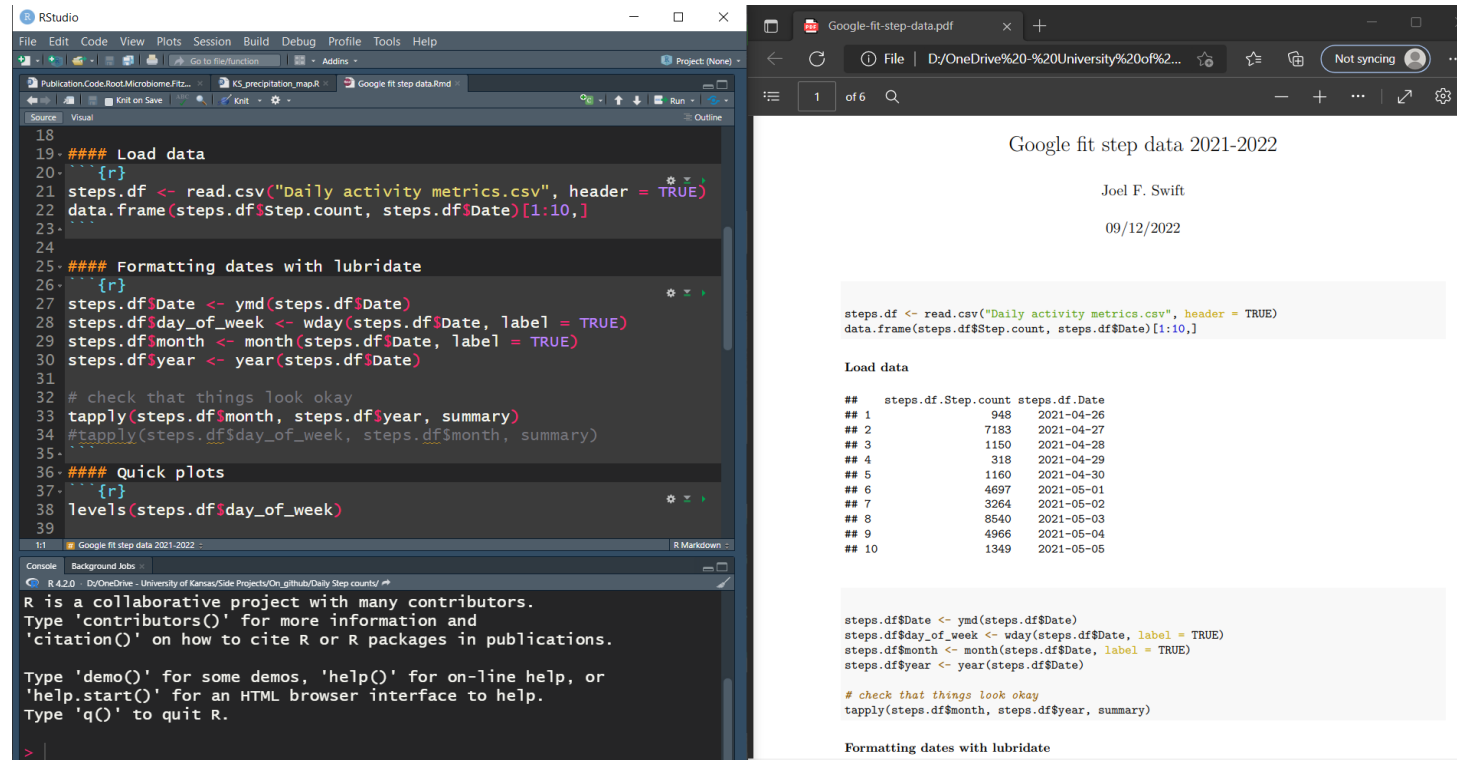
The screenshot shows the GitHub repository page for 'Kenizzer/California_Transect_Microbiome'. The repository is public and has 1 branch and 0 tags. The main branch is 'main'. The repository description is 'Repository of code for "Grafted grapevine microbiota varies across space, time, and scion/rootstock combinations"'. The repository has 0 stars, 1 watching, and 0 forks. The repository is licensed under GPL-3.0. The repository has no releases published. The repository has no packages published. The repository has a README.md file. The repository has a LICENSE file. The repository has a .gitattributes file. The repository has a Qiime_files folder. The repository has a Metadata folder. The repository has an Experimental_design_image folder. The repository has a Data_files folder. The repository has a Code folder. The repository has a commit history table.

| File/Folder | Commit Message | Commit Date |
|---------------------------|--------------------------------|--------------|
| Code | Abstract and a function change | 6 months ago |
| Data_files | 1st push | 7 months ago |
| Experimental_design_image | 1st push | 7 months ago |
| Metadata | 1st push | 7 months ago |
| Qiime_files | 1st push | 7 months ago |
| .gitattributes | Initial commit | 7 months ago |
| LICENSE | Initial commit | 7 months ago |
| README.md | Abstract and a function change | 6 months ago |

California transect microbiomes

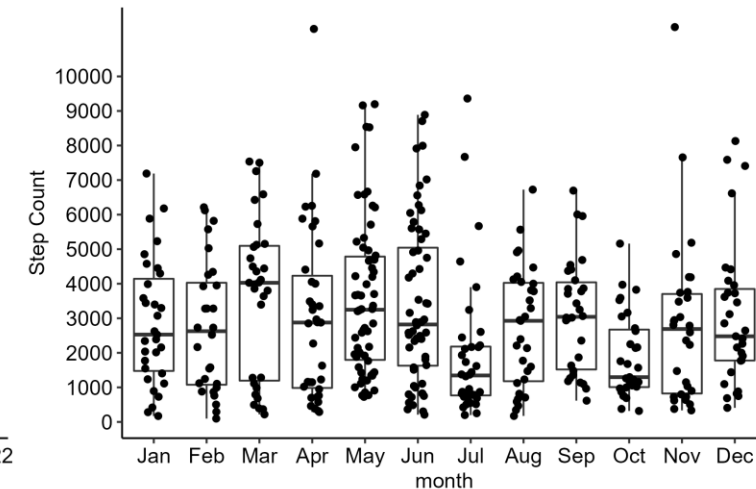
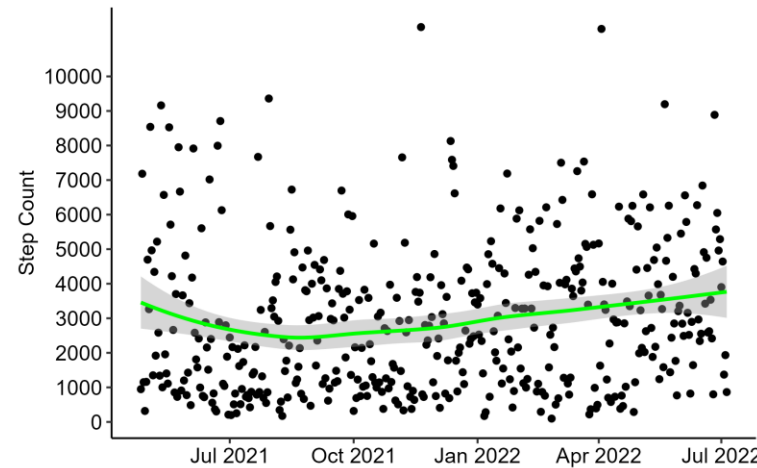
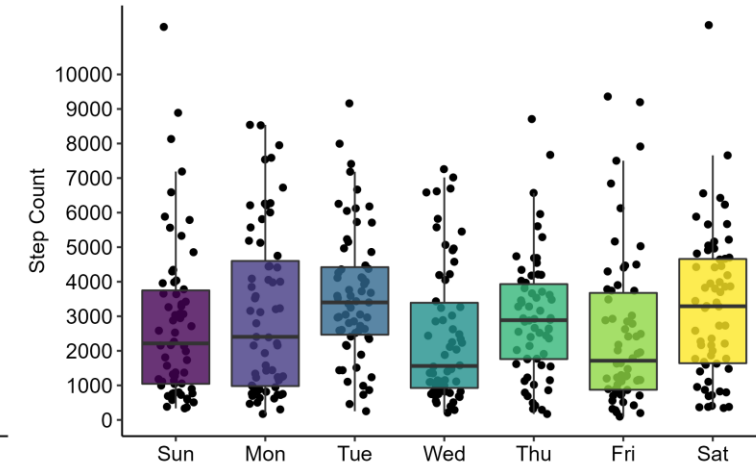
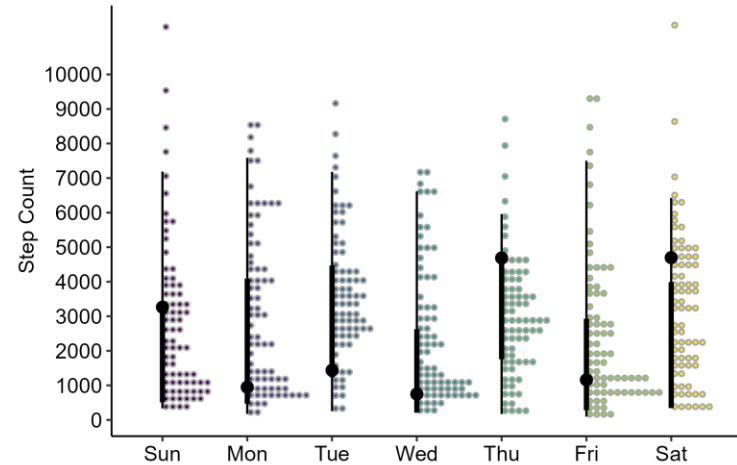
Repositories/Repos on github

- The basics are:
 - A readme file that will be kind of like the home page of the repo.
 - Something you are wanting to share (i.e., code/analyses)
- Readme files:
 - These are typically written in markdown (not covered today)



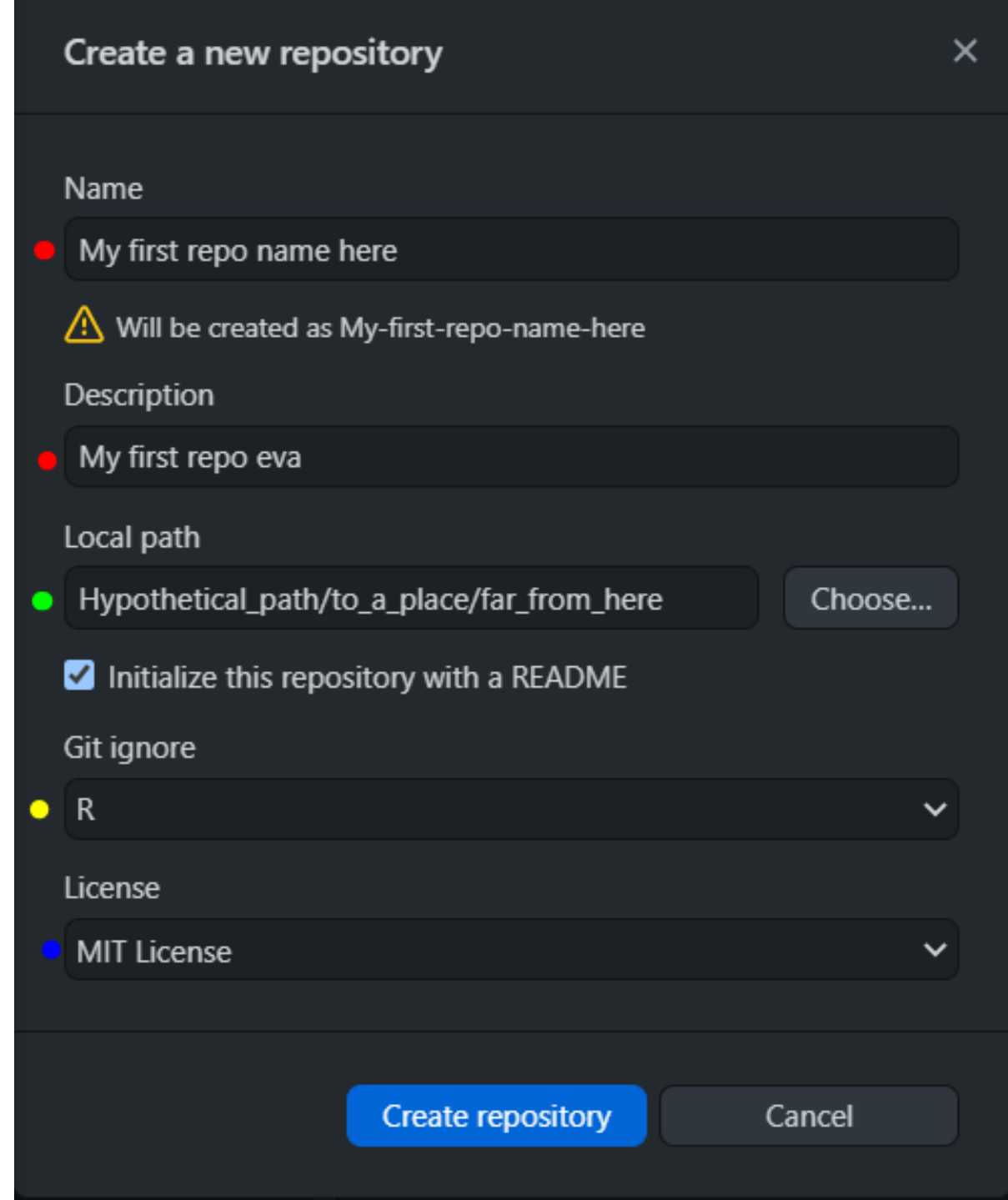
Example repo – Google fit step count data

- I wanted to look for trends in the number of steps I took per day.
- Data was collected by the phone application google fit.
- ~04-2021 to 07-2022.



Adding a Repo

- Name and description.
- Local path to the repo (where the files are on your computer).
- Git ignore template to use (in our case R will work).
- License to use for the repo MIT and GNU are likely fine (particularly important for patentable research)



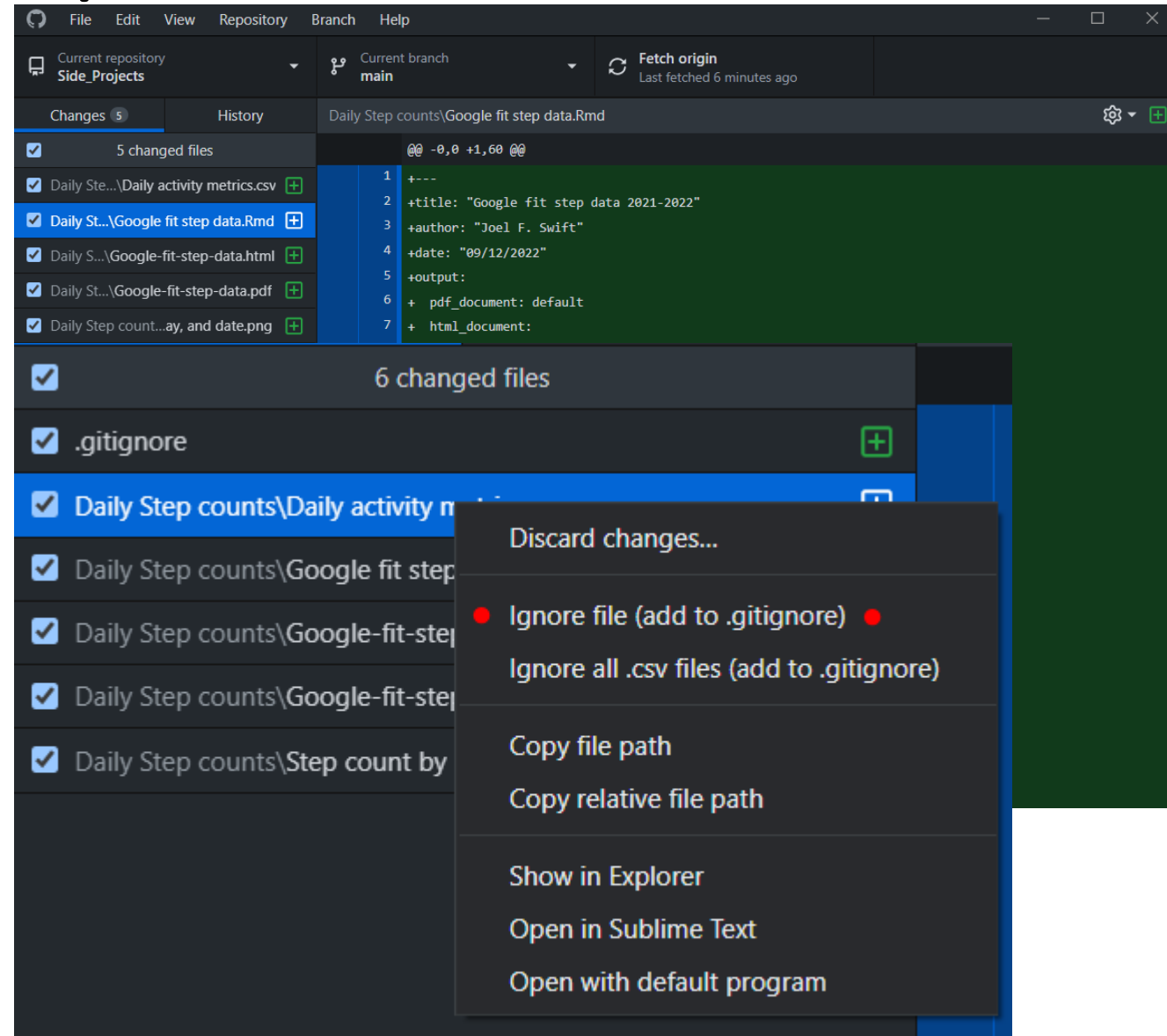
The screenshot shows the 'Create a new repository' dialog in GitHub. It has a dark theme. At the top, the title 'Create a new repository' is in white, with a close button (X) on the right. Below the title, there are several input fields and options:

- Name:** A text input field containing 'My first repo name here'. Below it, a warning icon (yellow triangle with an exclamation mark) and the text 'Will be created as My-first-repo-name-here'.
- Description:** A text input field containing 'My first repo eva'.
- Local path:** A text input field containing 'Hypothetical_path/to_a_place/far_from_here'. To its right is a button labeled 'Choose...'. Below this is a checked checkbox labeled 'Initialize this repository with a README'.
- Git ignore:** A dropdown menu showing 'R'.
- License:** A dropdown menu showing 'MIT License'.

At the bottom of the dialog, there are two buttons: 'Create repository' (in blue) and 'Cancel' (in gray).

Committing changes to a Repo

- Once you have a repo, adding any files into the repo folder will show up as a change in github desktop.
- We can choose to ignore a file (aka don't change or add to repo) if we like.
 - An example might be a dataset that contains GPS data you don't want to share



Committing changes to a Repo

- Commit messages can be important to remember what changes were made.



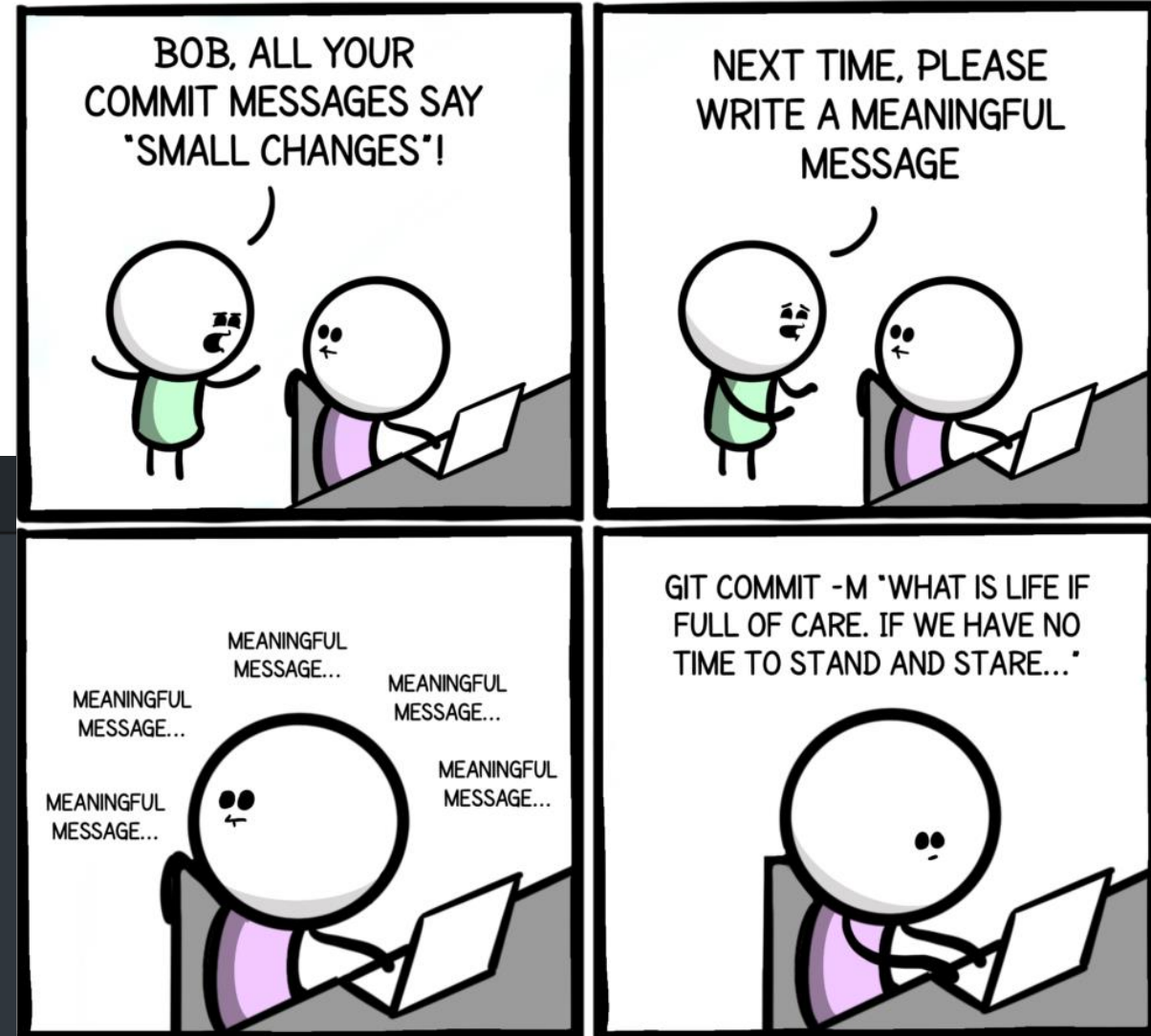
Added new project, google fit step count

Google fit data was downloaded for 2021 and 2022. I looked at trends in the data by day of the week, month, and date.

Data is included in .gitignore as it contains GPS data I'd rather not share :)



Commit to **main**



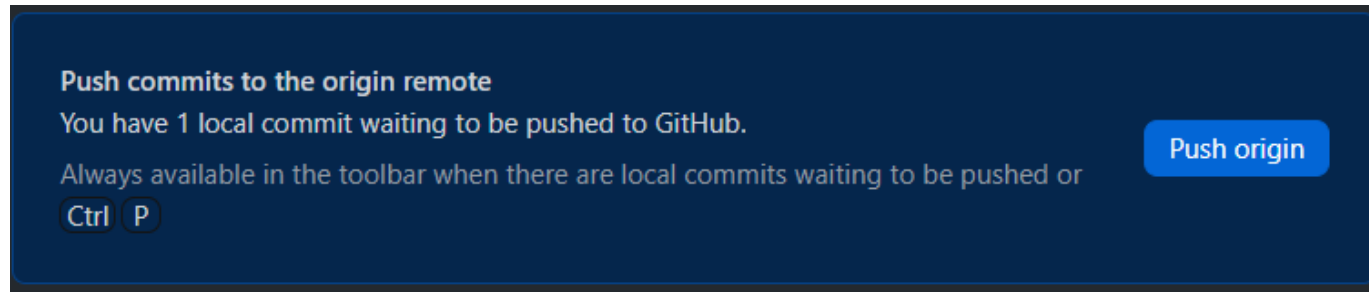
@_workchronicles

follow on *Twitter* & *Instagram* for more (and don't forget the *underscore*)

workchronicles.com

Committing changes to a Repo

- Commit the change.
- Choose the branch.
- Push the change to the branch.




Lets look at the repo that was changed

- https://github.com/Kenizzer/Side_Projects

er / **Side_Projects** Public Pin


Issues Pull requests Actions Projects Wiki Security Insights Settings

main 1 branch 0 tags Go to file Add file Code

 **Kenizzer** Added new project, google fit step count ...

d02ba81 1 hour ago 13 commits

| | | |
|---------------------------|--|---------------|
| 📁 Daily Step counts | Added new project, google fit step count | 1 hour ago |
| 📁 NFL margin of victory | Updated box scores and new goodies NFL stats | 3 days ago |
| 📁 NFL win-lose Prediction | Updated box scores and new goodies NFL stats | 3 days ago |
| 📁 NFL | Initial commit | 10 months ago |
| 📁 Punk rock lyrics | Initial commit | 10 months ago |
| 📄 .gitattributes | Initial commit | 10 months ago |
| 📄 .gitignore | Added new project, google fit step count | 1 hour ago |
| 📄 README.md | Updated box scores and new goodies NFL stats | 3 days ago |

 **Kenizzer** Added new project, google fit step count ...

1 hour ago History

..

| | | |
|------------------------------|--|------------|
| 📄 Google fit step data.Rmd | Added new project, google fit step count | 1 hour ago |
| 📄 Google-fit-step-data.html | Added new project, google fit step count | 1 hour ago |
| 📄 Google-fit-step-data.pdf | Added new project, google fit step count | 1 hour ago |
| 📄 Step count by month, da... | Added new project, google fit step count | 1 hour ago |

Working with others

Branching and Forking

Fork vs Branch

Comparison Chart

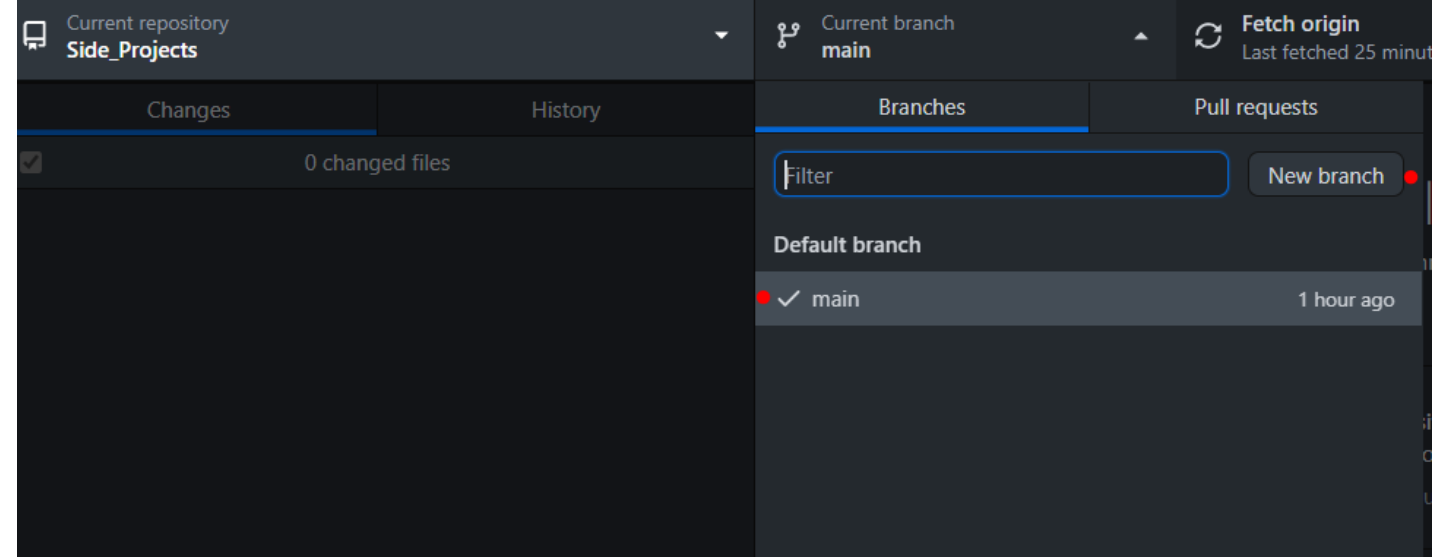
| Fork | Branch |
|---|---|
| A fork is nothing but a copy of the repository to your own scope without affecting the original project repo. | A branch is an isolated environment to add, modify or delete a portion of the code without messing with the main code base. |
| Forking is a cloning operation in Git that is executed on the entire repository level. | Branching is a cloning operation in Git executed on a single repository. |
| Forking creates a full copy of the original repository which sits in your account. | Branching creates a branch to implementing your changes without affecting other developers. |
| The purpose is to improve someone else's project by adding some new features to the existing repo. | The purpose is to divert from the original code base without affecting other developers' work. |

- Fork on.
- Create a new branch to work on.

laborator
re
the main



Branches



- Typically there will be a single main branch (often named main or master¹) for a repo.
- When you would like to work in a new environment, you make a new branch.
 - Why? Often its to try to fix or add things that have the potential to break other parts of the code. This way an important software can have a stable (main) branch and a dev branch for experimenting.

¹ <https://www.theserverside.com/feature/Why-GitHub-renamed-its-master-branch-to-main>

Branches

Create a branch

Name

Test_branch

Your new branch will be based on your currently checked out branch (**main**). **main** is the **default branch** for your repository.

Create branch

Cancel

- Change committed!

File Edit View Repository Branch Help

Current repository
Side_Projects

Current branch
Test_branch

↑

Publish branch

Publish this branch to GitHub

Changes 1

History

1 changed file

README.md

@@ -1,5 +1,8 @@

1 1 # Side projects

2 2

3 +### Google fit step data

4 +***Current state: Get to stepping, needs more data***

5 +

3 6 ### Lyrics for US punk rock bands in a machine learning data object.

4 7 ***Current state: Machine learning data object built but needs cleaning up

5 8

.....

↓

Updated the readme for the google fit project

Small edit to readem

+

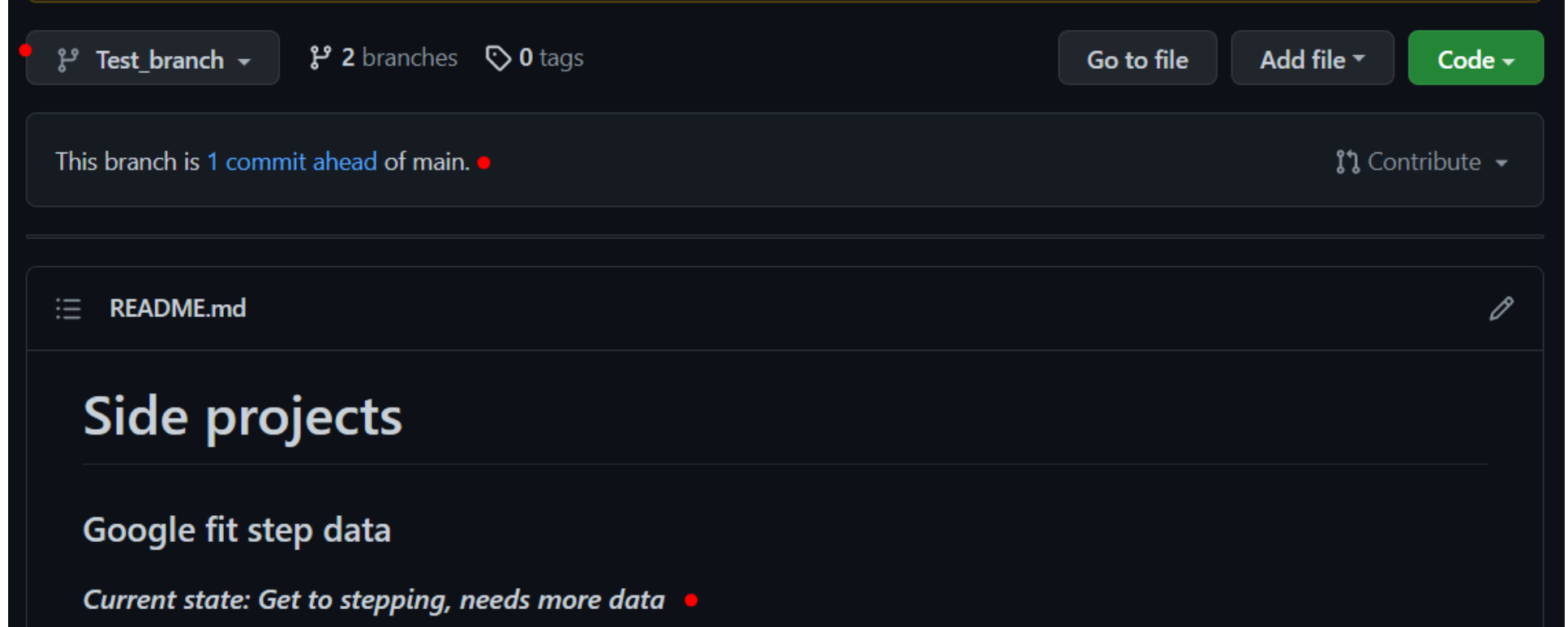
Commit to Test_branch

The dreaded Merge Conflict

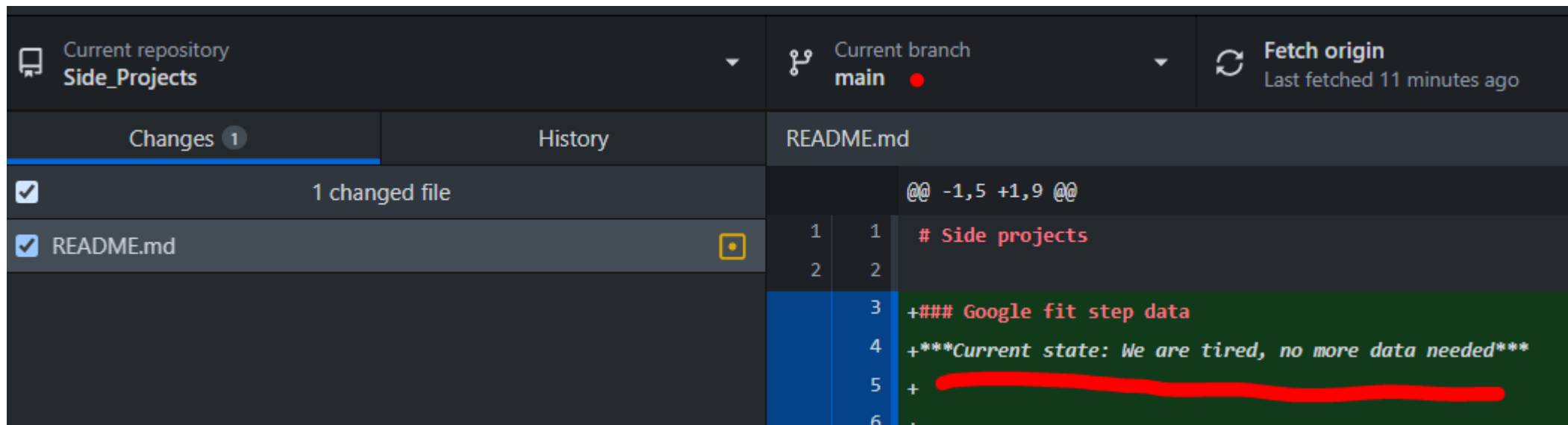


- What happens when we attempt to merge branches that have changes to the same file? Which version do we choose?
- Merge Conflicts!!

- Test ->




- Main ->




Pull request

Open a pull request

Create a new pull request by comparing changes across two branches. If you need to, you can also [compare across forks](#).

 base: main ▾ ← compare: Test_branch ▾ ✖ Can't automatically merge. Don't worry, you can still create the pull request.




Updated the readme for the google fit project

Write

Preview

H B I ≡ <> 🔗 ≡ ≡ ≡ @ ↗ ↶

Small edit to readme

Attach files by dragging & dropping, selecting or pasting them. 

Create pull request ▾

Reviewers

No reviews

Assignees

No one—assign yourself

Labels

None yet

Projects

None yet

Milestone

No milestone

Add more commits by pushing to the **Test_branch** branch on **Kenizzer/Side_Projects**.



This branch has conflicts that must be resolved

Use the [web editor](#) or the [command line](#) to resolve conflicts.

Conflicting files

README.md

Resolve conflicts

Merge pull request



You can also [open this in GitHub Desktop](#) or view [command line instructions](#).

Updated the readme for the google fit project #1

Resolving conflicts between `Test_branch` and `main` and committing changes → `Test_branch`

1 conflicting file

README.md



README.md
README.md

```
1  # Side projects
2
3  ### Google fit step data
4  <<<<<<< Test_branch
5  ***Current state: Get to stepping, needs more data***
6  |=====
7  ***Current state: We are tired, no more data needed***
8
9  >>>>>>> main
10
```

- Resolved by deleting one of the lines about the google fit project.

README.md

1 conflict

Prev ^

Next v



Mark as resolved

```
1  # Side projects
2
3  ### Google fit step data
4  |***Current state: Get to stepping, needs more data***
5
6
7  ### Lyrics for US punk rock bands in a machine learning data object.
8  ***Current state: Machine learning data object built but needs cleaning up before analysis***
9
10
11  ### NFL margin of victory for all NFL teams 2017-Current.
12  ***Current state: Get scores for games farther back***
13
14
15  ### NFL boxscore stats scrapped from ESPN.com 2002-2019.
16  ***Current state: Making interactive vizualizations with ggplotly and exploring data***
```


Commit has been merged after resolving the conflict

Conversation 0

Commits 1

Checks 0

Files changed 1





Kenizzer commented 3 minutes ago

Owner


😊 ...


Small edit to readme




 Updated the readme for the google fit project ...


616d0a3




 Merge branch 'main' into Test_branch

Verified 345a79e



 Kenizzer merged commit **fec3692** into **main** now

Revert



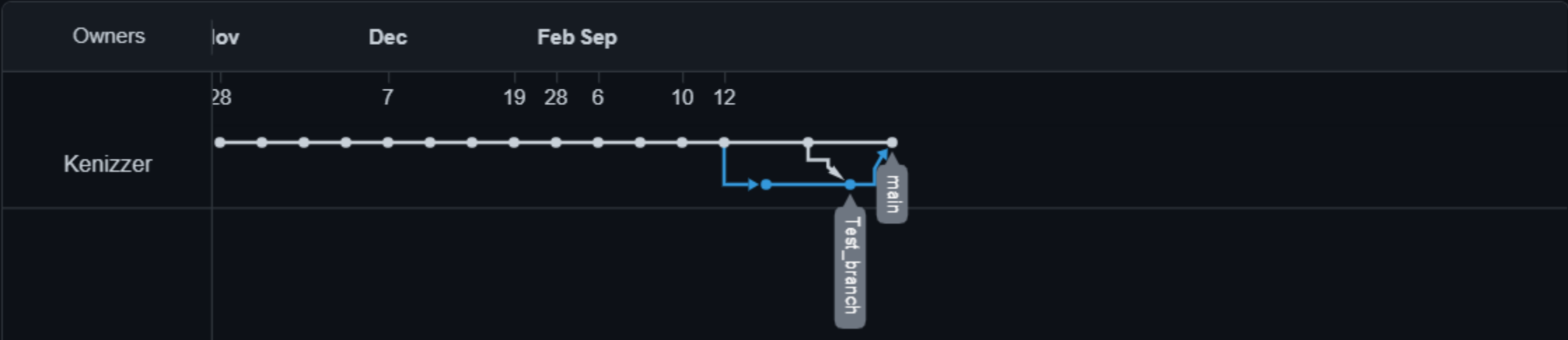
Pull request successfully merged and closed

You're all set—the **Test_branch** branch can be safely deleted.

Delete branch

Network graph

Timeline of the most recent commits to this repository and its network ordered by most recently pushed to.




Conversation 0

Commits 1

Checks 0



Files changed 1

 Kenizzer commented 7 minutes ago



Owner

...



Small edit to readme

  Updated the readme for the google fit project



616d0a3

  Merge branch 'main' into Test_branch

Verified 345a79e

  Kenizzer merged commit **fec3692** into **main** 4 minutes ago

Revert

  Kenizzer deleted the **Test_branch** branch now

Restore branch

Forks


- Create forks when you would like to contribute to a project you did not create.
- Example: The 2022 KU symposium had an repo on github, I wanted to make the 2023 repo.

Create a new fork

A *fork* is a copy of a repository. Forking a repository allows you to freely experiment with changes without affecting the original project.

Owner *

Repository name *

 Kenizzer ▾

 /


KUCG_SymposiumPlanning ✓

By default, forks are named the same as their parent repository. You can customize the name to distinguish it further.

Description (optional)

Forked to add 2023 folder for Symposium planning

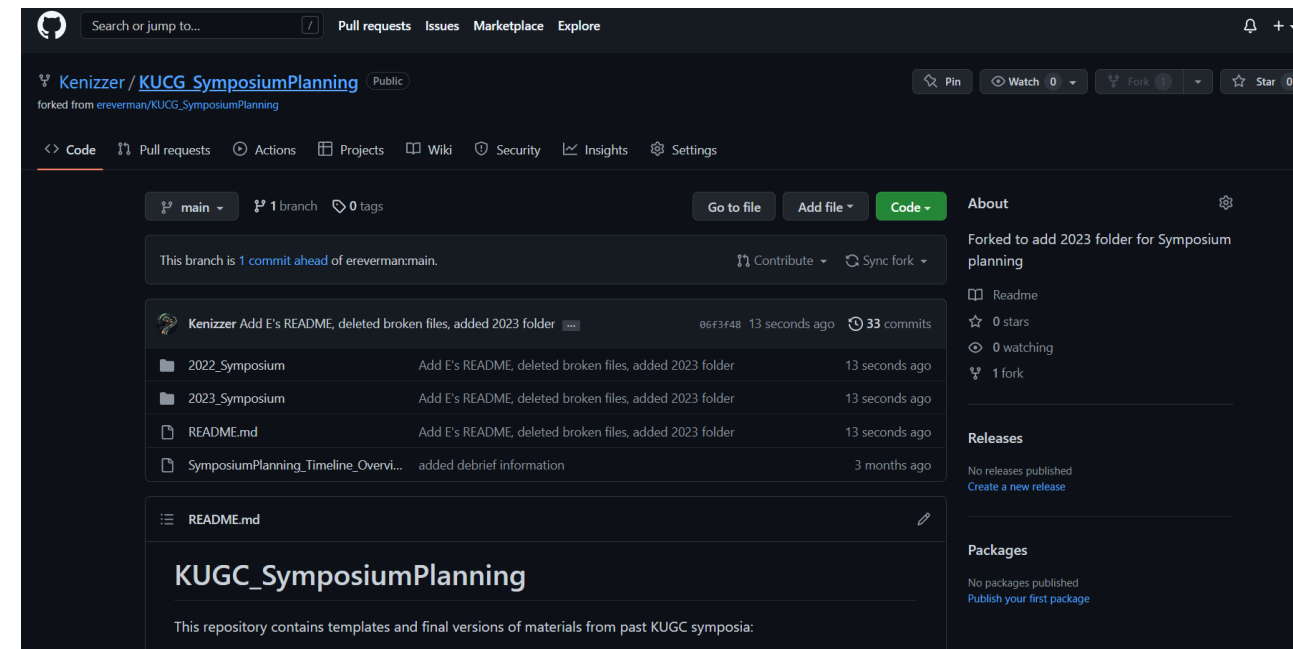
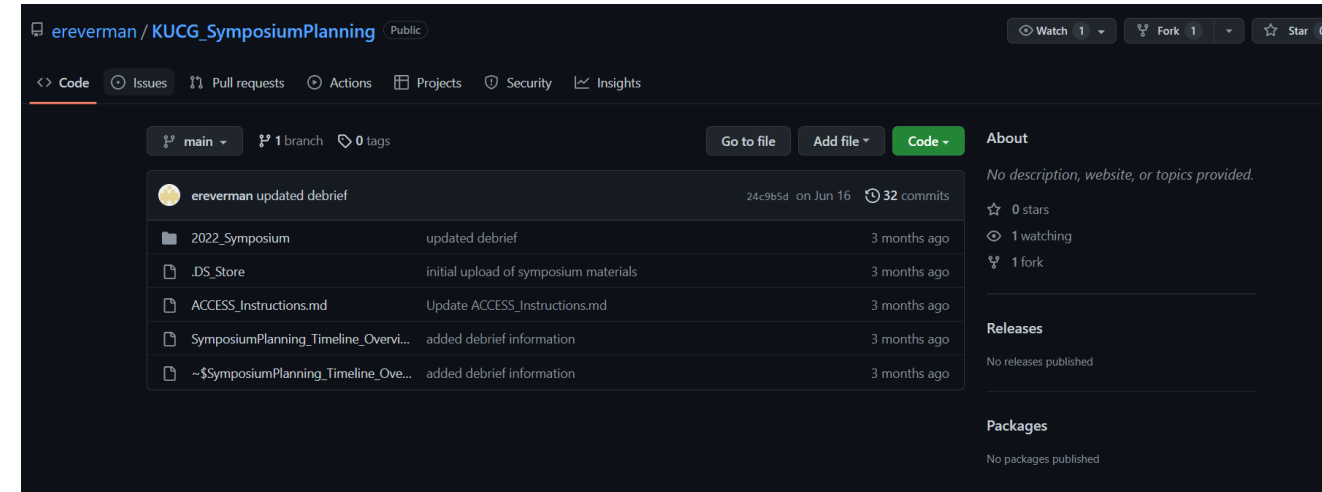
☒ **Copy the `main` branch only**
Contribute back to ereverman/KUCG_SymposiumPlanning by adding your own branch. [Learn more.](#)

 You are creating a fork in your personal account.

Create fork

Forks

- Now we clone the directory locally.
- Then we make our changes and contributions and commit and push them.



Open a pull request

ereverman / KUCG_SymposiumPlanningPublic

Watch1

Fork1

Star0

<> Code

Issues

Pull requests

Actions

Projects

Security

Insights

Comparing changes

Choose two branches to see what's changed or to start a new pull request. If you need to, you can also [compare across forks](#).

base repository: ereverman/KUCG_SymposiumP...▼

base: main▼

←

head repository: Kenizzer/KUCG_SymposiumPla...▼

compare: main▼

✓ **Able to merge.** These branches can be automatically merged.

Discuss and review the changes in this comparison with others. [Learn about pull requests](#)

Create pull request

1 commit

18 files changed

1 contributor

Commits on Sep 13, 2022

Add E's README, deleted broken files, added 2023 folder...

Kenizzer committed 3 minutes ago

06f3f48

<>

Showing 18 changed files with 0 additions and 0 deletions.

Split

Unified

▼


BIN -6 KB .DS_Store

Open a pull request

- Notice the message that we can auto merge?
- Why is that?

Open a pull request


Create a new pull request by comparing changes across two branches. If you need to, you can also [compare across forks](#).

 base repository: ereverman/KUCG_SymposiumP... base: main

←

head repository: Kenizzer/KUCG_SymposiumPla... compare: main









✓ **Able to merge.** These branches can be automatically merged.




Add E's README, deleted broken files, added 2023 folder


Write

Preview


H B I       @  

- Added 2023 folder with the same directory structure as 2022.
- Renamed the access instructions to README so that [github](#) will display it natively (Nice markdown/instructions E)!
- Removed just a couple [werid](#) files that I think crop up because of Mac/external [harddrives](#).
- Added minutes for 9-13-2022

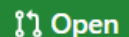
Attach files by dragging & dropping, selecting or pasting them. 

☒ Allow edits by maintainers 

Create pull request

 Remember, contributions to this repository should follow our [GitHub Community Guidelines](#).

Add E's README, deleted broken files, added 2023 folder #1

[Edit](#)[Code](#)

Kenizzer wants to merge 1 commit into `ereverman:main` from `Kenizzer:main`



Conversation 0



Commits 1



Checks 0



Files changed 18

+0 -0



Kenizzer commented 29 seconds ago



- Added 2023 folder with the same directory structure as 2022.
- Renamed the access instructions to README so that github will display it natively (Nice markdown/instructions E!)
- Removed just a couple weird files that I think crop up because of Mac/external harddrives.
- Added minutes for 9-13-2022



Add E's README, deleted broken files, added 2023 folder



06f3f48

Add more commits by pushing to the `main` branch on [Kenizzer/KUCG_SymposiumPlanning](#).



This branch has no conflicts with the base branch

Only those with [write access](#) to this repository can merge pull requests.



Write

Preview

H B I @

Leave a comment

Reviewers

No reviews

Still in progress? [Convert to draft](#)

Assignees

No one assigned

Labels

None yet

Projects

None yet

Milestone

No milestone

Development

Questions?



GitHub is old-fashioned.