

Modern Workflow

Department of Sociology | University of Texas at Austin

Casey Breen

Session goals

High-level introduction to three tools:

1. Git + GitHub
2. Latex / Overleaf / Markdown (Quarto)
3. Zotero for managing references

Why invest in improving your workflow?

- Can help you become a more efficient researcher
- Facilitates collaboration
- Open Science + improve replicability of findings

Git and GitHub

- Git: version-control system for tracking changes in code
- GitHub: web-based hosting service for version control using Git



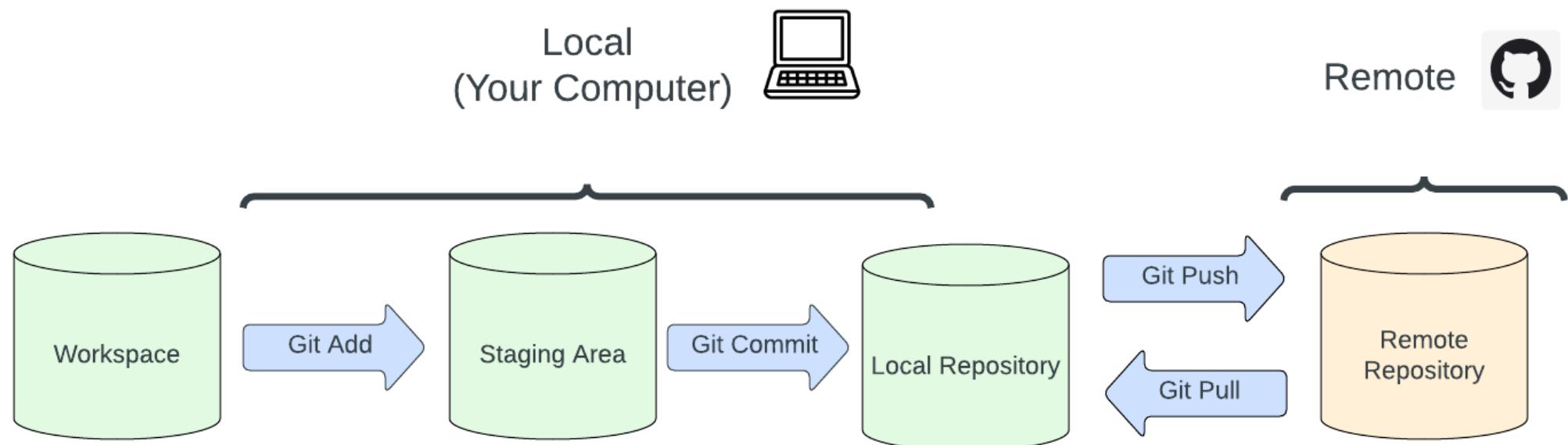
Key advantages of GitHub

- Version control helps track changes to code; easy to revert if you made a mistake
- Work across computers / servers
- Helpful for putting together replication packages

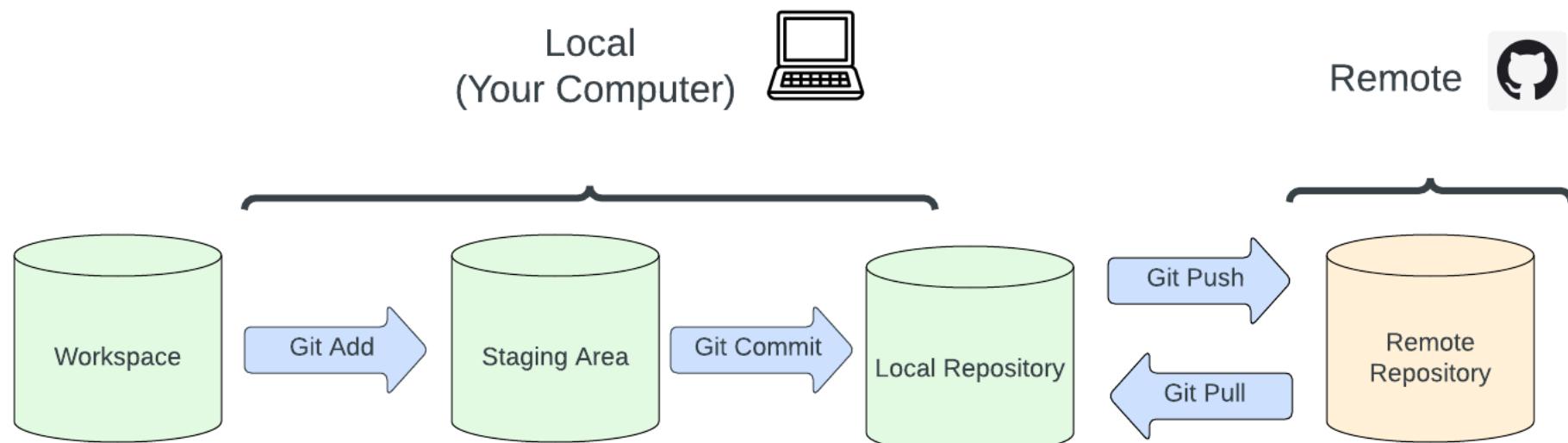
GitHub Project Management

- All files and code organized into a repository
- Repository can have different files and subfolders
 - Readme file, code folder, survey instrument folder,
 - Not good for storing data!

GitHub Workflow



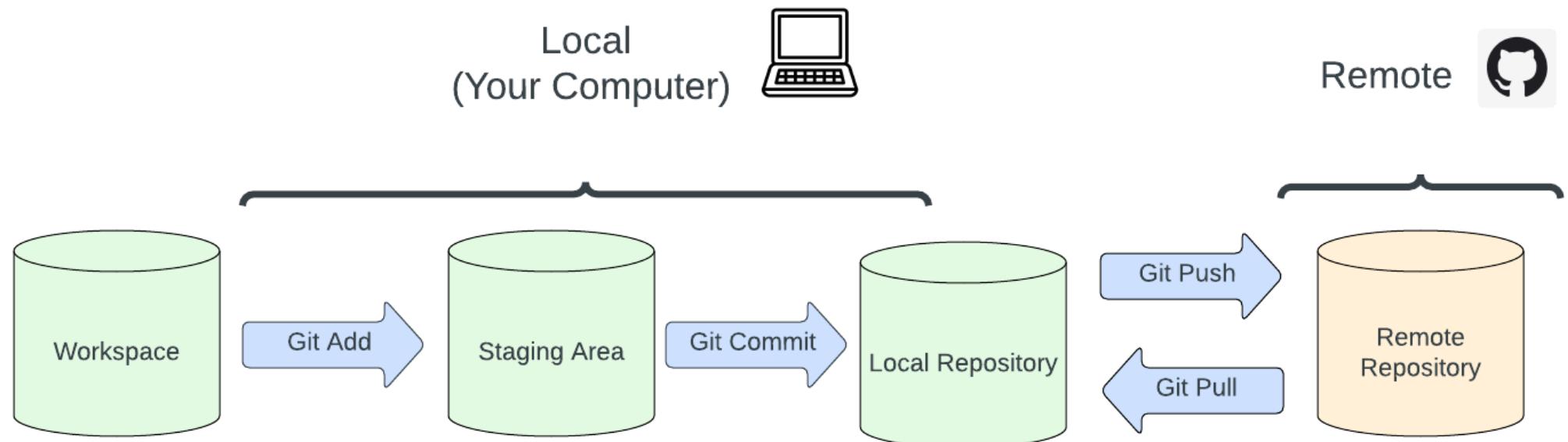
GitHub Workflow – Git Add



To add files to staging area:

```
1 ## add all changed files
2 git add -A
3
4 ## add specific file
5 git add /path/to/file/file.txt
```

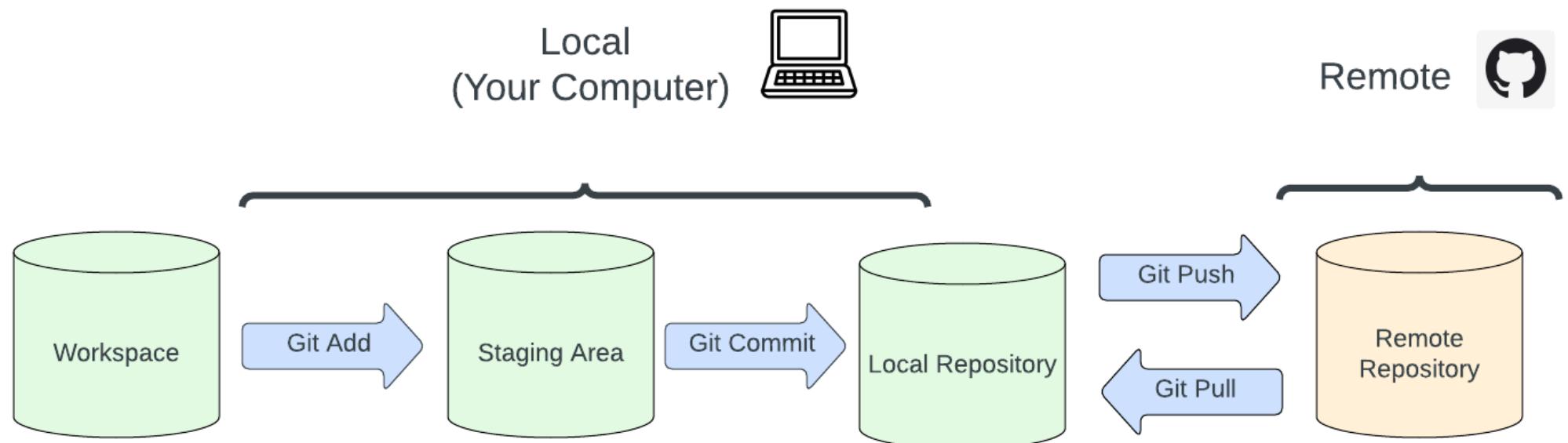
GitHub Workflow – Git Commit



Commit changes

```
1 ## commit changes
2 git commit -m "add an informative message here"
```

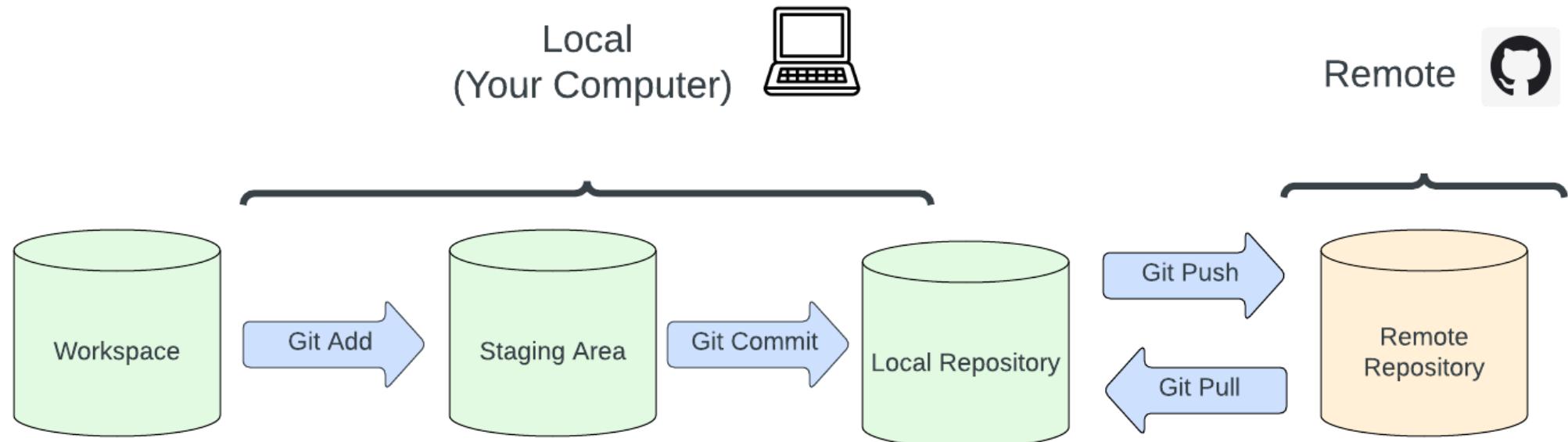
GitHub Workflow – Git Push



Send (“push”) your changes up to remote repository (where collaborators / public can see)

```
1 ## add all changed files  
2 git push
```

GitHub Workflow – Git Pull



Pull down changes from repository (e.g., changes your collaborator made)

```
1 git pull
```

Other features

- **Git Clone:** Clone a remote repository to your local machine

```
1 ## general command  
2 git clone https://github.com/username/repository-name.git  
3  
4 ## clone the repository for this class  
5 git clone https://github.com/username/repository-name.git
```

- **Git fork:** Copy a (public) repository to have your own version
- **Branches:** Isolated environment for development, merge back in with main codebase when you're ready

Link to Open Science Framework repository

- Create public GitHub repository with replication code + instructions for a manuscript
- Link that repository to an Open Science Framework Repository
 - Upload data etc.
- Share DOI (permanent identifier) in paper

Live Demo

- Adding new file to GitHub repository
- OSF / GitHub replication package

Exercises

1. Create a new Github repository
2. Clone this GitHub repository to your local machine
3. ‘Push’ an R markdown script to your GitHub Repository

LaTeX/Overleaf

LaTex

- Latex is a high-quality typesetting system for writing scientific documents
- Free, open source — customizable, extendable
- Learning curve!

Insert equations

- Use \$ or \begin{equation} to enter equation mode
- $\$h(x) = e^{x}\$$

$$h(x) = ae^{\beta x}$$

Nicely handle references

- Create a .bib file from list of reference
- Automatically write out reference list
 - bibtex, natbib
- In-text citations:
 - \citep{citation_key} = (Gupta 2019)
 - \citet{citation_key} = Gupta (2019)

Simplify tables + figures management

- Include figures in text

```
1 \begin{figure}[!htb]
2   \centering
3   \includegraphics[width = .9\textwidth]{figs/tanganyika_map.png}
4   \caption{Tanganyika province with three Zone de Santes highlighted.}
5   \label{fig:map}
6 \end{figure}
```

- Reference figures and tables in text using label
 - `\cref{fig:map}` will render “Figure 4”

Overleaf

- Google docs for Latex
- See .tex code and compiled PDF at same time
- Great for collaborating

The screenshot shows a dual-pane interface of the Overleaf web editor. The left pane displays the LaTeX source code for a document titled 'drc_paa_abstract (pre-registration)'. The right pane shows a preview of the resulting PDF. The PDF page contains text about a study in Tanganyika Province, Democratic Republic of Congo (DRC), mentioning high death rates and insecurity. It also includes a figure and a table. The Overleaf interface includes standard browser navigation tools and a toolbar with various document management and collaboration features.

Abstract

Reliable estimates of death rates in complex humanitarian emergencies are critical for assessing the severity of a crisis and effectively allocating resources. However, in many humanitarian settings, logistical and security concerns make conventional methods for estimating death rates difficult. In this study, we develop and test a new strategy for estimating death rates in complex emergencies. Our strategy is based on the network method, a promising new approach that uses reports about deaths in survey respondents' social networks. To test our method, we collected original data in a setting where reliable estimates of death rates are in high demand: Tanganyika Province of the Democratic Republic of the Congo ($N = 2,651$). In Tanganyika, qualitative work suggested basing death rate estimates on respondents' reports about deaths among members of their personal networks. We conducted a household survey and a contemporaneous retrospective household mortality survey ($N = 2,970$). The results will allow us to decompose different sources of error in our network survival estimates. Through this systematic validation, we aim to demonstrate the promise of applying the network survival method in humanitarian disasters.

*Preliminary. Please do not cite or redistribute.
†Department of Demography, University of California, Berkeley. caseybreen@berkeley.edu.
‡Reach Initiative, Geneva, Switzerland
§School of Public Health, University of Khudza.
¶Department of Demography, University of California, Berkeley. feehan@berkeley.edu.

Live Demo + Casey templates

- Latex CV
- Beamer Slides
- Latex Manuscript

Exercises (at home)

1. Create a new overleaf project
2. Upload and insert a new figure
3. Write out a mathematical equation
4. Share the overleaf document with a friend

R users: Quarto (markdown)

- Designed for writing code
 - Interwoven text and code chunks
- Can also write papers, take notes, make slides
- Markdown is simpler than Latex, but fewer features / less resources for help

Zotero for reference management

Zotero

- Free tool for simplifying reference management
- Organize, annotate, and track references across devices

Simple Zotero workflow

1. Add papers to Zotero to “collection”
 - Collection = set of references for one project
1. Generate .bibtex file (or individual citations)
1. Move .bibtex file to paper repository

Zotero tips

- [Zotero connector](#): browser plug-in that automatically adds and pdf paper to a collection
- [Better Bibtex](#): help customize the different citation keys
- Create shared libraries when working with collaborators

Zotero live demo

- Creating new library
- Adding paper using connector (or drag-and-drop pdf)
- Generating bibtex file

Exercises (at home)

1. Create a new Zotero collection
2. Add 3 papers to the new collection (use Zotero connector or drag in PDFs)
3. Export the .bibtex file to Overleaf

Takeaways

- Lots of exciting resources for improving workflow
 - Many resources not mentioned
- Pick workflow that works best for you
 - Definitely worth investing in your workflow!

More resources

- Github: [Happy Git with R](#)
- LaTeX: [Learn Latex in 30 Minutes, Not-so-short introduction to LaTeX](#)
- Zotero: [Getting Started with Zotero](#)

Questions?

- Thanks for your attendance and participation
- Questions: casey.breen@demography.ox.ac.uk