California Teachers Study ETL Methodology

Emma Spielfogel

2024-06-28

Table of contents

Preface

This is a Quarto book for documenting the CTS' ETL methodology. This book is currently under development.

To learn more about Quarto books visit https://quarto.org/docs/books.

1 Introduction

This is a book created from markdown and executable code.

See Knuth (1984) for additional discussion of literate programming.

```
1 + 1
```

```
[1] 2
```

2 RStudio Setup

Purpose: This section will describe how to set up RStudio for development.

2.1 Using Git and GitHub

Follow the instructions here to set up Git and GitHub. As described in these instructions, using R Studio Projects makes working with Git and GitHub easier.

2.2 Git for Command Line

2.2.1 To clone a repository from GitHub to your RStudio session:

- Copy the link from the GitHub repository you would like to clone
  - Click the green "Code" button to see the link
- Go the Terminal tab in RStudio
- Make sure you are at the location you want to clone to—if not, change your directory using cd
- git clone "URL of repository from GitHub"

2.2.2 To add new files to a git repo then push to GitHub:

- Go the Terminal tab in RStudio
- Open the repository file path if you are not already in the project (use cd in the Terminal to change your directory)
- git pull
  - This will pull the repository your current location in order to ensure you are using the most up-to-date version. This is especially impo
- git status
  - This allows you to check the status—doing this often is a good check.
- git add + file name

- This adds files to git (not GitHub).
- git add -u adds anything that is tracked but has changed (the u stands for updated)

- git status
  - Checking the status again here can be helpful to double-check you have added all files and folders you would like to.
- git commit -m “Commit message”
  - This commits the changes you have made. Everything is easily reversible **until this step.** Once committed, changes are part of the hi
- git push
  - This pushes the committed changes to GitHub.

If you have messed up the branches, you can revert (do with mega caution):

- git reset –hard {insert commit ID}
- git log
- git status
- git push -f
- git status

References

Knuth, Donald E. 1984. “Literate Programming.” *Comput. J.* 27 (2): 97–111. https://doi.org/10.1093/comjnl/27.2.97.