Git & GitHub Overview

Data Science and Visualization Institute for Librarians

Heidi Tebbe | Bret Davidson

bretdavidson.github.io/dsvil-github

The Plan

- Why use version control systems?
- What is Git?
- What is GitHub?
- Evaluating software projects on GitHub.
- Using GitHub without the command line.
- Non-code uses for Git & GitHub.

Why use version control?

How to save documents!

- annual.docx
- annual_final.docx
- annual_final_JC.docx
- annual_final_JC_2.docx
- annual_final_final.docx
- annual_final_final_corrected.docx
- FINAL_annual_final.docx

What does version control do?

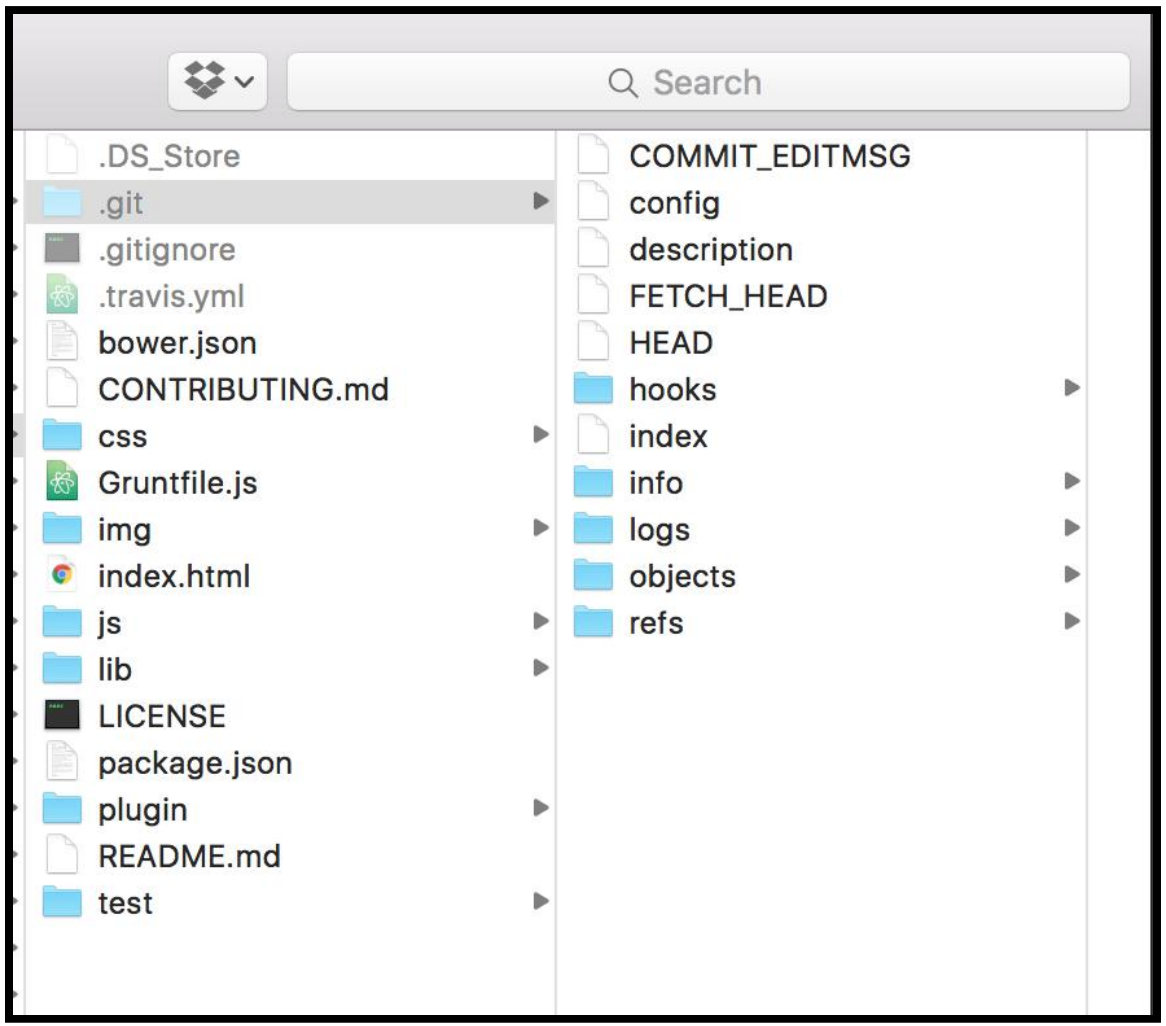
- Records the history of changes to a set of related files inside of a "repository".
- Enables navigation of that change history.
- Supports sharing of the repository among a group of collaborators.

Two Kinds of Version Control

- Centralized
  - Examples are CVS and Subversion
- Distributed
  - Examples are Git and Mercurial

What is Git?

- Git is a distributed version control system.
- Original author is Linus Torvalds, of Linux fame.
- "I'm egotistical, and I name all my projects after myself."
- git means "unpleasant person" in British slang.


Search
.DS_Store
.git
.gitignore
.travis.yml
bower.json
CONTRIBUTING.md
css
Gruntfile.js
img
index.html
js
lib
LICENSE
package.json
plugin
README.md
test
COMMIT_EDITMSG
config
description
FETCH_HEAD
HEAD
hooks
index
info
logs
objects
refs

What is GitHub?

- GitHub is a Git repository hosting service.
- Adds services on top of Git like issues, wikis, and social networking.

Common Commands

checkout/clone

obtain a copy of a repository

pull/sync

update your local copy of repo to the latest version

branch

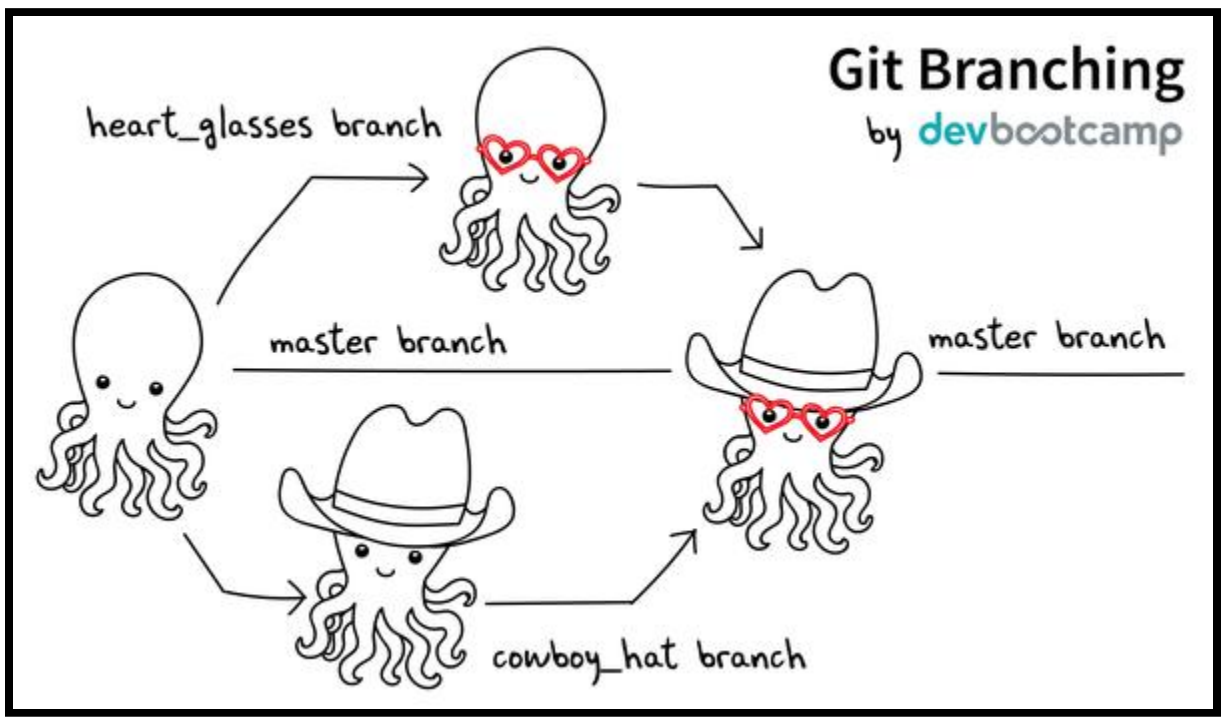
create a separate version of the files in the repo

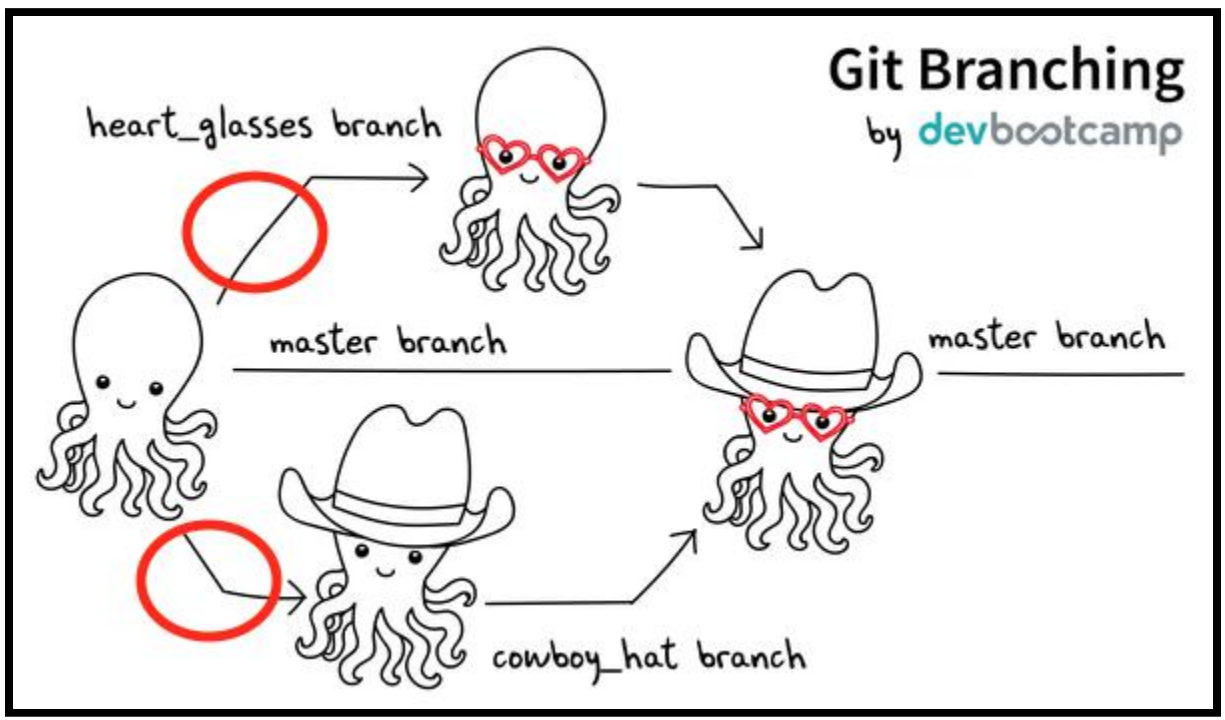
commit

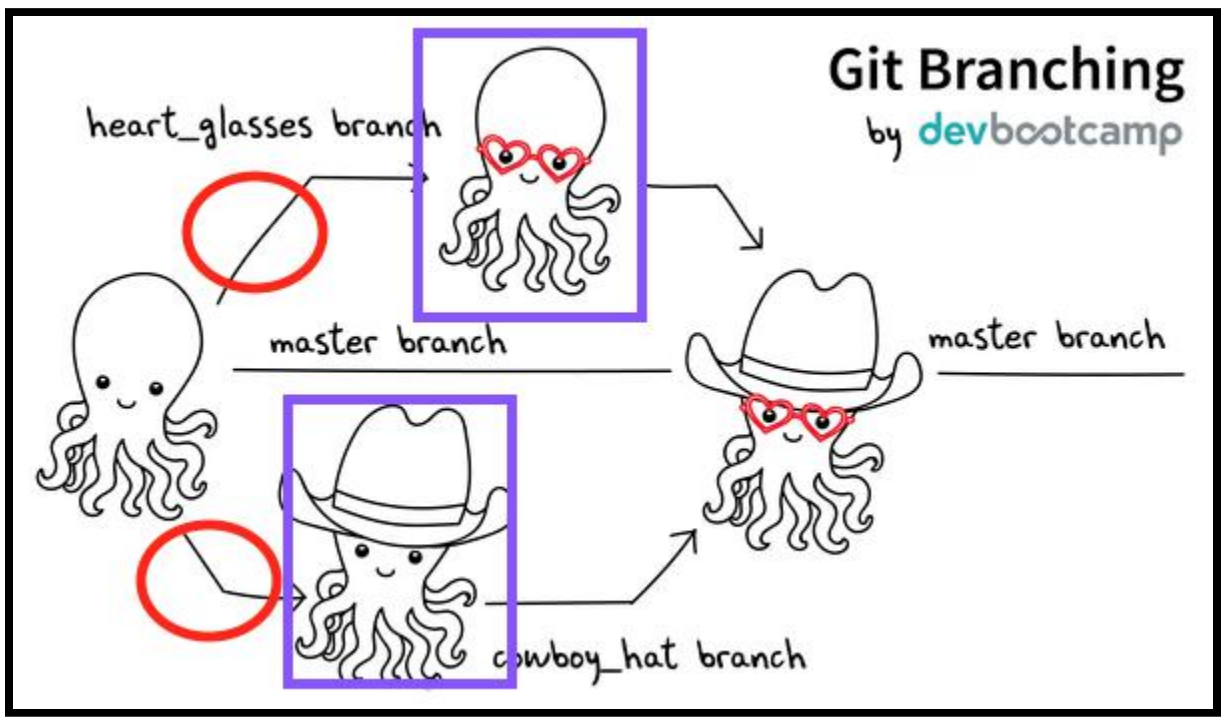
record a set of changes

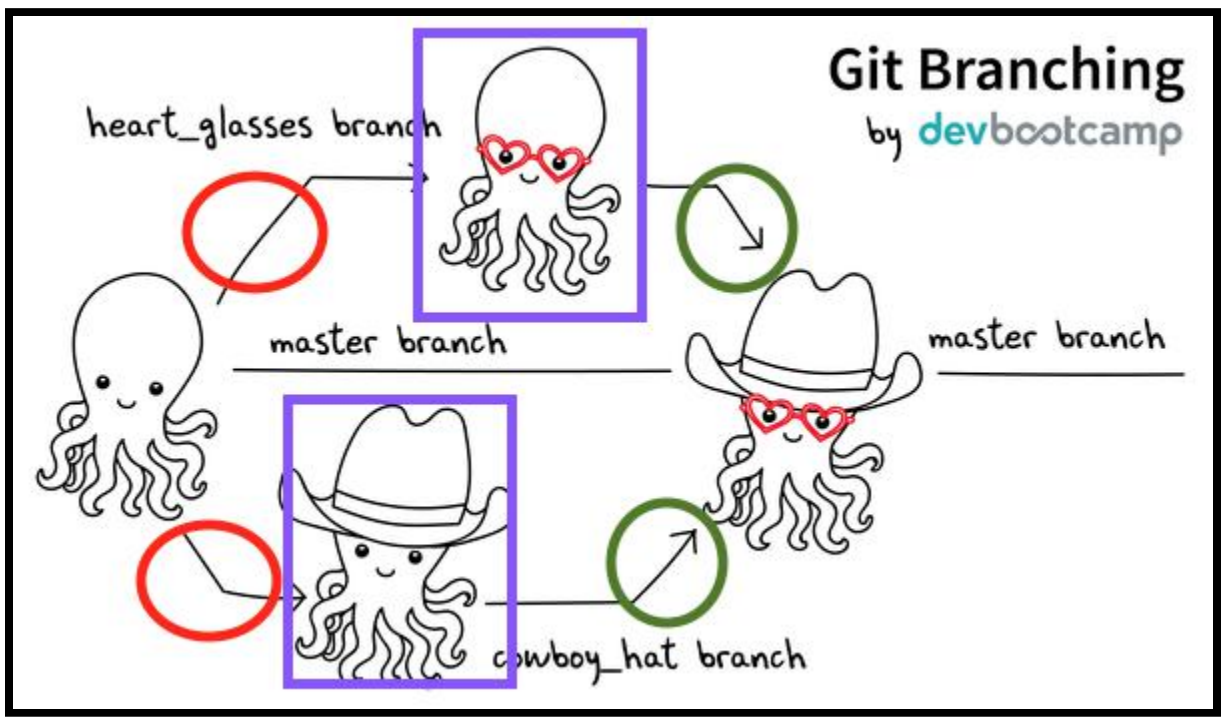
merge

add your changes to 'master' branch


Git Branching
by devbootcamp
heart_glasses branch
master branch
master branch
cowboy_hat branch


Git Branching
by devbootcamp
heart_glasses branch
master branch
master branch
cowboy_hat branch


Git Branching
by devbootcamp
heart_glasses branch
master branch
master branch
cowboy_hat branch


Git Branching
by devbootcamp
heart_glasses branch
master branch
master branch
cowboy_hat branch

push/sync

send your changes to the central or remote repo

Critical Features of GitHub

- Fork
- Pull Request

fork

create a copy (clone) in order to make your own changes

use someone else's project as a starting point

propose changes to the original

pull request

mechanism for submitting your changes to original project

project owner "pulls" in your changes

Web Client Overview

github.com/bretdavidson

github.com/suma-project/Suma

Activity

What is GitHub Desktop?

desktop.github.com

Non-code Uses

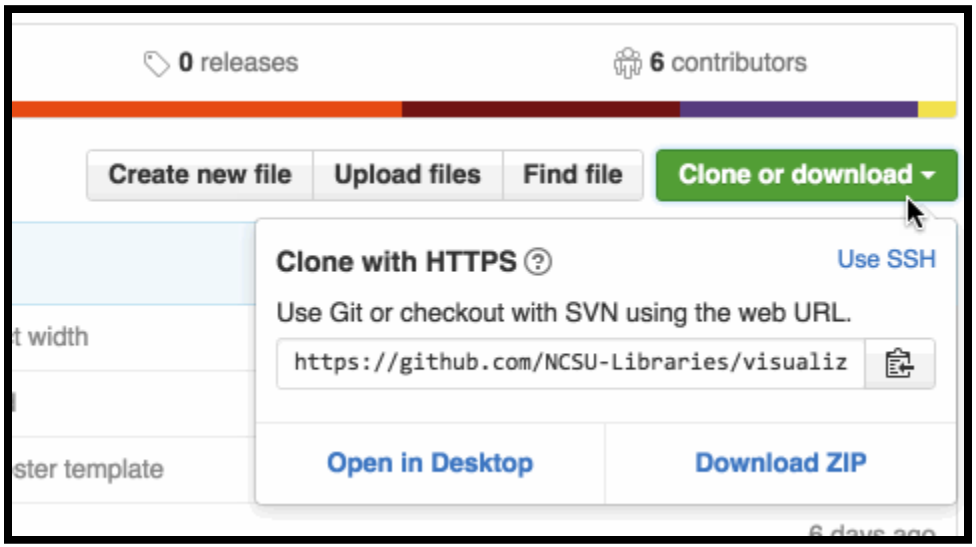
- Knowledge Base
- Data Management
- Communication

Knowledge Base

github.com/NCSU-Libraries/visualization_templates

Diffing Images

Caveats



- No ZIP files for folders
- Not a digital asset management system
- Licenses

Data Management

Making Your Code Citable

Communication

- Jupyter Notebooks
- GitHub Pages

Jupyter Notebooks

- interactive environments that can contain executable code, rich text, mathematics, plots
- support over 40 programming languages, including Python and R
- *Jupyter Notebook App* : server-client application that allows editing and running of notebook documents via a web browser
- *Jupyter Notebook Viewer* : free web service that allows you to share static versions of publicly available hosted notebook files
- render directly on GitHub

Pages

pages.github.com

suma-project.github.io/Suma

Resources for learning to create infographics

Data Visualization Workshops

Logistics

File and Repository Sizes

- git : *no file size or file number limits*
- GitHub repositories : *1 GB*
- files added to GitHub via browser : *25 MB limit*
- files added to GitHub via command line : *100 MB limit*
- files stored using Git Large File Storage (LFS) : *2 GB limit*

Public vs. Private

- git : *private*
- free GitHub account : *public*
- paid GitHub plans : *public* AND *private*
- students using GitHub : *public* AND *private*

Further Learning

Free Tutorials

- Try Git: www.codeschool.com/courses/try-git
- Pro Git Ebook: git-scm.com/book/en/v2
- GitHub Guides: guides.github.com

Documentation

- git Documentation: git-scm.com/documentation
- GitHub Help: help.github.com

Questions?

bretdavidson.github.io/dsvil-github