

Introduction to Git and GitHub

General Assembly – Data Science

Agenda

- I. Introduction
- II. Exploring GitHub
- III. Using Git with GitHub
- IV. Contributing on GitHub
- V. Bonus Content

I. Introduction

Why learn version control?

- Version control is useful when you write code, and data scientists write code
- Enables teams to easily collaborate on the same codebase
- Enables you to contribute to open source projects
- Attractive skill for employment

What is Git?

- Version control system that allows you to track files and file changes in a repository (“repo”)
- Primarily used by software developers
- Most widely used version control system
 - Alternatives: Mercurial, Subversion, CVS
- Runs from the command line (usually)
- Can be used alone or in a team

What is GitHub?

- A website, not a version control system
- Allows you to put your Git repos online
 - Largest code host in the world
 - Alternative: Bitbucket
- Benefits of GitHub:
 - Backup of files
 - Visual interface for navigating repos
 - Makes repo collaboration easy
- “GitHub is just Dropbox for Git”
- Note: Git does not require GitHub

Git can be challenging to learn

- Designed (by programmers) for power and flexibility over simplicity
- Hard to know if what you did was right
- Hard to explore since most actions are “permanent” (in a sense) and can have serious consequences
- We’ll focus on the most important 10% of Git

II. Exploring GitHub

GitHub setup

- Create an account at github.com
- There's nothing to install
 - “GitHub for Windows” & “GitHub for Mac” are GUI clients (alternatives to command line)

Navigating a GitHub repo (1 of 2)

- Example repo: http://github.com/sinanuozdemir/SF_DAT_17
- Account name, repo name, description
- Folder structure
- Viewing files:
 - Rendered view (with syntax highlighting)
 - Raw view
- README.md:
 - Describes a repo
 - Automatically displayed
 - Written in Markdown

Navigating a GitHub repo (2 of 2)

- Commits:
 - One or more changes to one or more files
 - Revision highlighting
 - Commit comments are required
 - Most recent commit comment shown by filename
- Profile page

Creating a profile

- Click on the signup button on the top-right
- Choose a plan (one of them is free)

GitHub

[Explore](#) [Features](#) [Enterprise](#) [Blog](#)

[Sign up](#)

[Sign in](#)

- **Remember your email and password!!!! You will need it again soon!!!!**

Basic Markdown

- Easy-to-read, easy-to-write markup language
- Usually (always?) rendered as HTML
- Many implementations (aka “flavors”)
- Let’s edit README.md using GitHub!
- Common syntax:
 - `##` Header size 2
 - `*italics*` and `**bold**`
 - `[link to GitHub](https://github.com)`
 - `*` bullet
 - ``inline code`` and ```code blocks```
- Valid HTML can also be used within Markdown

III. Using Git with GitHub

Cloning a GitHub repo

- Cloning == copying to your local computer
 - Like copying your Dropbox files to a new machine
- First, change your working directory to where you want the repo you created to be stored: `cd`
- Then, clone the repo: `git clone <URL>`
 - Get HTTPS or SSH URL from GitHub (ends in .git)
 - Clones to a subdirectory of the working directory
 - No visual feedback when you type your password
- Navigate to the repo (`cd`) then list the files (`ls`)

The url is on the repo page

Chrome | File | Edit | View | History | Bookmarks | People | Window | Help | 36% | Wed 5:15 PM

Netflix | sinanuozdemir/SF_DAT_17 | FiveThirtyEight | Data science - Wikipedia, the free encyclopedia | https://raw.githubusercontent.com/sinanuozdemir/SF_DAT_17/master/data/

GitHub, Inc. [US] | https://github.com/sinanuozdemir/sf_dat_17

Apps | Education | Programming | Math | APIs | Google Calendar | (29) The Big Boy League | Development Task Tracker | FOOTBALL | SF_DAT_17 | Other Bookmarks

This repository | Search | Pull requests | Issues | Gist

sinanuozdemir / SF_DAT_17 | Unwatch 3 | Unstar 4 | Fork 0

Description | **Website**

Short description of this repository | Website for this repository (optional) | Save or Cancel

8 commits | 1 branch | 0 releases | 1 contributor

Branch: master | SF_DAT_17 / +

added class 2 material


sinanuozdemir authored 27 minutes ago | latest commit c4414aadb2

GA perks	Added ga perks document	a day ago
code	added class 2 material	27 minutes ago
project-examples	first real commit	6 days ago
slides	added class 2 material	27 minutes ago
.gitignore	Update .gitignore	6 days ago
README.md	added class 2 material	27 minutes ago
peer_review.md	first real commit	6 days ago
project-examples.md	first real commit	6 days ago
project.md	first real commit	6 days ago

Issues 0 | Pull requests 0 | Wiki

Pulse | Graphs | Settings

HTTPS clone URL

https://github.com/: 

You can clone with [HTTPS](#), [SSH](#), or [GitHub CLI](#).

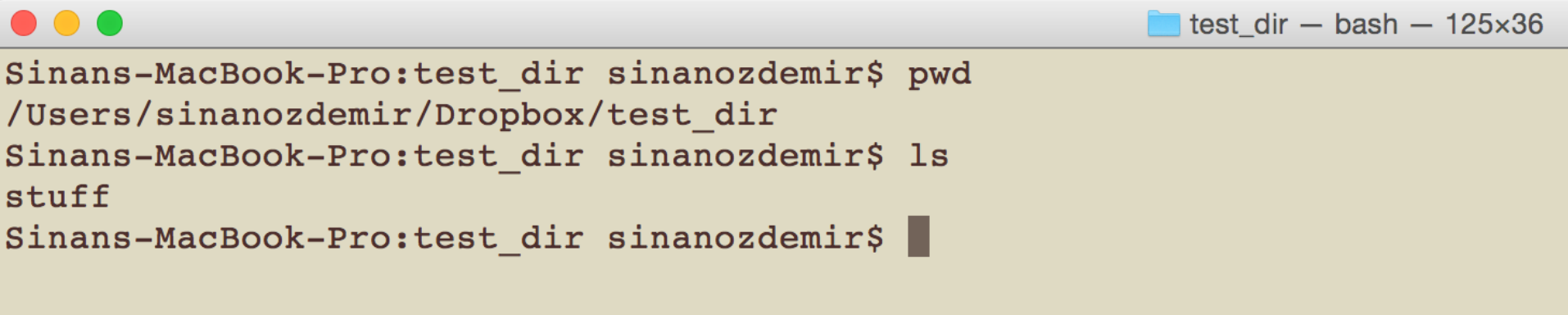
Clone in Desktop | Download ZIP

https://github.com/sinanuozdemir/SF_DAT_17/branches

First Clone

- First we will clone the main class repo!
- You will need to do this to stay up to date with all of class info

Before Cloning

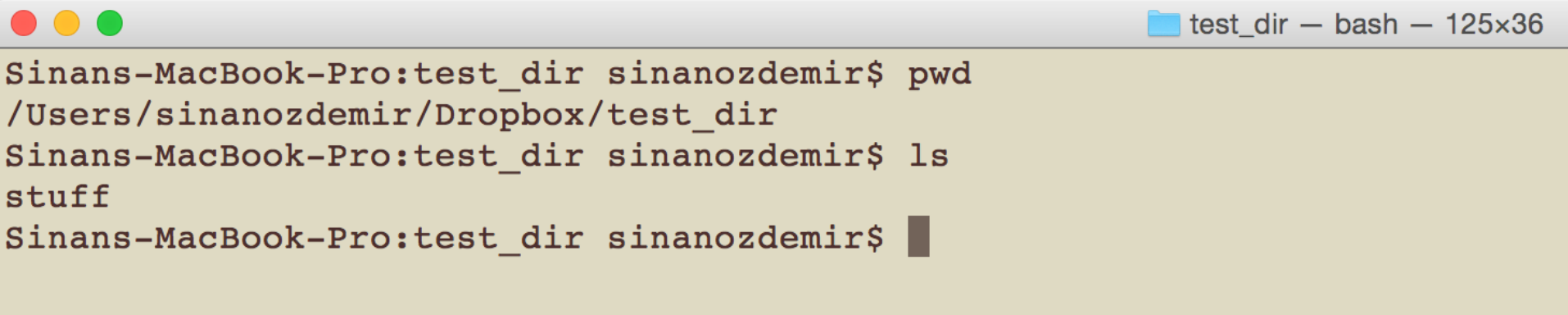


A terminal window with a title bar showing three colored window control buttons (red, yellow, green) on the left and a title 'test_dir — bash — 125x36' on the right. The terminal content shows a user at a MacPro prompt navigating into a directory and listing its contents.

```
Sinans-MacBook-Pro:test_dir sinanozdemir$ pwd
/Users/sinanozdemir/Dropbox/test_dir
Sinans-MacBook-Pro:test_dir sinanozdemir$ ls
stuff
Sinans-MacBook-Pro:test_dir sinanozdemir$
```

- Move into a Directory that you want to store the info for the next 11 weeks

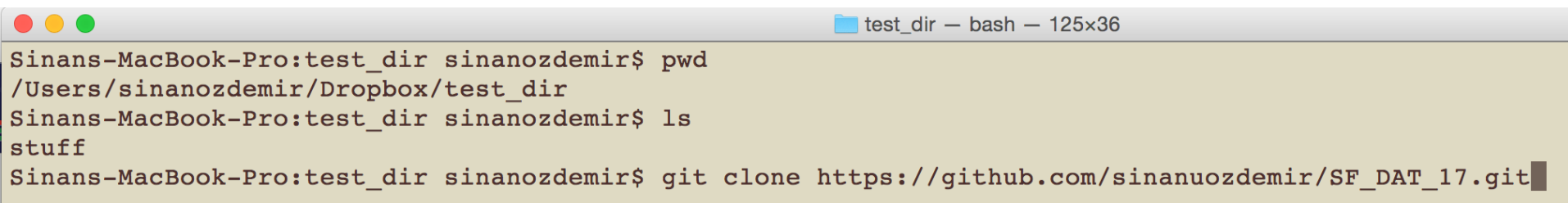
Before Cloning



A terminal window titled "test_dir — bash — 125x36" with three colored window control buttons (red, yellow, green) in the top-left corner. The terminal text shows the user's current directory and its contents.

```
Sinans-MacBook-Pro:test_dir sinanozdemir$ pwd
/Users/sinanozdemir/Dropbox/test_dir
Sinans-MacBook-Pro:test_dir sinanozdemir$ ls
stuff
Sinans-MacBook-Pro:test_dir sinanozdemir$
```

During Cloning

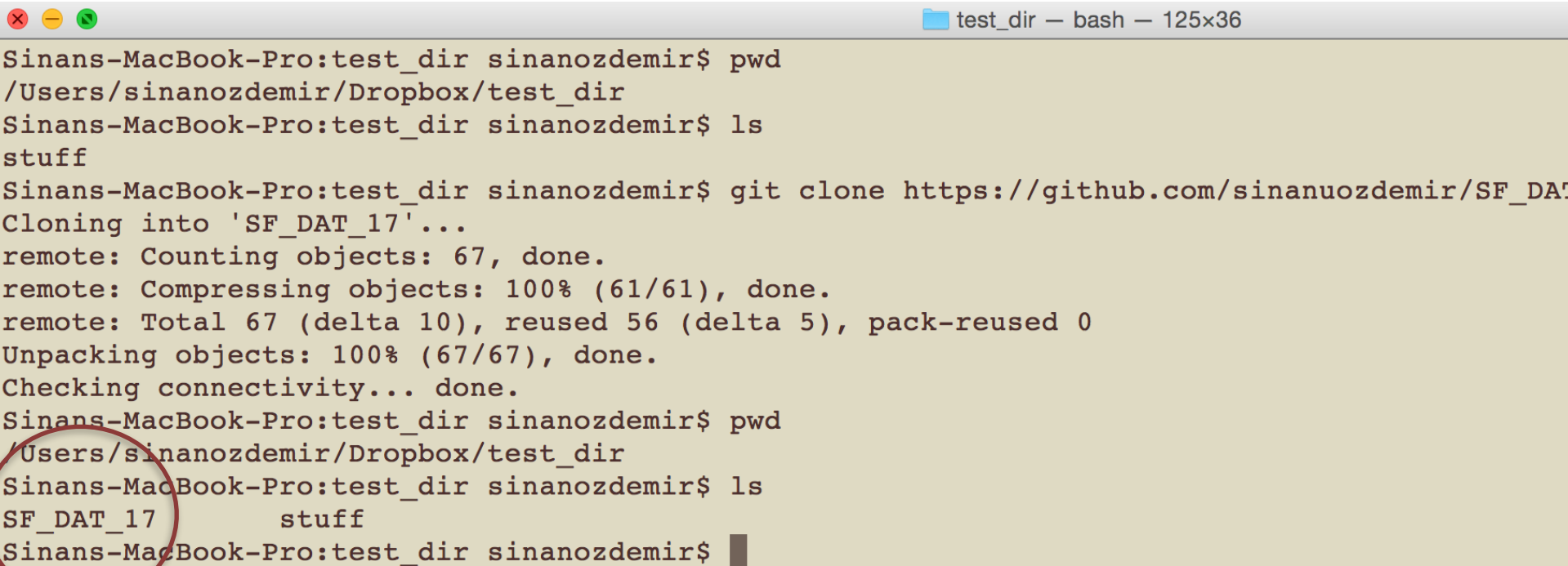


A terminal window titled "test_dir — bash — 125x36" with three colored window control buttons (red, yellow, green) in the top-left corner. The terminal text shows the user's current directory and the execution of a git clone command.

```
Sinans-MacBook-Pro:test_dir sinanozdemir$ pwd
/Users/sinanozdemir/Dropbox/test_dir
Sinans-MacBook-Pro:test_dir sinanozdemir$ ls
stuff
Sinans-MacBook-Pro:test_dir sinanozdemir$ git clone https://github.com/sinanuoazdemir/SF_DAT_17.git
```

- git clone https://github.com/sinanuoazdemir/SF_DAT_17.git

After Cloning



```
Sinans-MacBook-Pro:test_dir sinanozdemir$ pwd
/Users/sinanozdemir/Dropbox/test_dir
Sinans-MacBook-Pro:test_dir sinanozdemir$ ls
stuff
Sinans-MacBook-Pro:test_dir sinanozdemir$ git clone https://github.com/sinanuoazdemir/SF_DAT_17
Cloning into 'SF_DAT_17'...
remote: Counting objects: 67, done.
remote: Compressing objects: 100% (61/61), done.
remote: Total 67 (delta 10), reused 56 (delta 5), pack-reused 0
Unpacking objects: 100% (67/67), done.
Checking connectivity... done.
Sinans-MacBook-Pro:test_dir sinanozdemir$ pwd
/Users/sinanozdemir/Dropbox/test_dir
Sinans-MacBook-Pro:test_dir sinanozdemir$ ls
SF_DAT_17      stuff
Sinans-MacBook-Pro:test_dir sinanozdemir$
```

- You have a new folder!

Same as on Github!

```
Sinans-MacBook-Pro:test_dir sinanozdemir$ ls
SF_DAT_17      stuff
Sinans-MacBook-Pro:test_dir sinanozdemir$ cd SF_DAT_17/
Sinans-MacBook-Pro:SF_DAT_17 sinanozdemir$ ls
GA perks                peer_review.md          project-examples.md      public_data.md
README.md               project-examples         project.md                slides
Sinans-MacBook-Pro:SF_DAT_17 sinanozdemir$
```

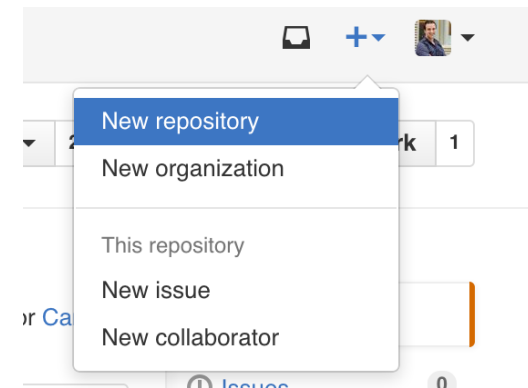
- **cd** into it and **ls**
- **Try this now! (take 5-10 minutes)**

Second Clone

- Now we will clone your new repo!
- First we have to make one

Creating a repo on GitHub

- Click “Create New” (plus sign) on your profile:
 - Define name, description, public or private
 - Initialize with README (if you’re going to clone)
 - Please call it **SF_DAT_17_WORK**
- Notes:
 - Nothing has happened to your local computer
 - This was done on GitHub, but GitHub add the README.md file



Preview of what you're about to do

- Copy (“**clone**”) your new GitHub repo to your computer
- Make some file changes locally
- Save those changes locally (“**commit**” them)
- Update your GitHub repo with those changes (“**push**”)

Preview of what you're about to do

- Copy (“**clone**”) your new GitHub repo to your computer
- Try this now! (take 5-10 minutes)
- SUPER IMPORTANT:
 - Make sure that you **LEAVE SF_DAT_17** before cloning the new repo
 - `cd ..`

Making changes, checking your status

- Making changes:
 - Modify README.md in any text editor
 - Create a new file: `touch <filename>`
- Check your status:
 - `git status`
- File statuses (possibly color-coded):
 - Untracked (red)
 - Tracked and modified (red)
 - Staged for committing (green)
 - Committed
- Try this now! (take 5-10 minutes)

Committing changes

- Stage changes for committing:
 - Add all “red” files: `git add .`
- Check your status:
 - `git status`
 - Red files have turned green
- Commit changes:
 - `git commit -m “message about commit”`
- Check your status again!
- Try this now! (take 5-10 minutes)

Pushing to GitHub

- Everything you've done to your cloned repo (so far) has been local
- You've been working in the “master” branch
- Push committed changes to GitHub:
 - Like syncing local file changes to Dropbox
 - `git push <remote> <branch>`
 - Often: `git push origin master`
- Refresh your GitHub repo to check!

Quick recap of what you've done

- Created a repo on GitHub
- Cloned repo to your local computer (**git clone**)
 - Automatically sets up your “origin” remote
- Made two file changes
- Staged changes for committing (**git add**)
- Committed changes (**git commit**)
- Pushed changes to GitHub (**git push**)
- Inspected along the way (**git remote**, **git status**, **git log**)

Let's do it again!

- Modify or add a file, then `git status`
- `git add .`, then `git status`
- `git commit -m "message"`
- `git push origin master`
- Refresh your GitHub repo
- Note if you want to add a folder, the folder must have an item in it or else github will not recognize it
- Try this now! (take 5-10 minutes)

Lab -- Before you leave

- Install [Git](#) on your machine
- Make a [Github](#) Profile on the web
- Create your own repo, call it “SF_DAT_17_WORK” and clone it to your machine
- Clone the [class repo](#)

IV. Bonus Content

Git installation and setup

- Installation: tiny.cc/installgit
- Open Git Bash (Windows) or Terminal (Mac/Linux):
 - `git config --global user.name "YOUR FULL NAME"`
 - `git config --global user.email "YOUR EMAIL"`
- Use the same email address you used with your GitHub account
- Generate SSH keys (optional): tiny.cc/gitssh
 - More secure than HTTPS
 - Only necessary if HTTPS doesn't work for you

Checking your remotes

- A “remote alias” is a reference to a repo not on your local computer
 - Like a connection to your Dropbox account
- View remotes: `git remote -v`
- “origin” remote was set up by “git clone”
- Note: Remotes are repo-specific

Two ways to initialize Git

- Initialize on GitHub:
 - Create a repo on GitHub (with README)
 - Clone to your local machine
- Initialize locally:
 - Initialize Git in existing local directory: `git init`
 - Create a repo on GitHub (without README)
 - Add remote: `git remote add origin <URL>`

Deleting or moving a repo

- Deleting a GitHub repo:
 - Settings, then Delete
- Deleting a local repo:
 - Just delete the folder!
- Moving a local repo:
 - Just move the folder!

Excluding files from a repo

- Create a “.gitignore” file in your repo:
`touch .gitignore`
- Specify exclusions, one per line:
 - Single files: `pip-log.txt`
 - All files with a matching extension: `*.pyc`
 - Directories: `env/`
- Templates: github.com/github/gitignore

Gists: lightweight repos

- You have access to Gist: gist.github.com
- Add one or more files
- Supports cloning, forking, commenting, committing
- Can be public or secret (not private)
- Useful for snippets, embedding, IPython nbviewer, etc.

Useful to learn next

- Working with branches
- Rolling back changes
- Resolving merge conflicts
- Fixing LF/CRLF issues