URBAN
INSTITUTE · ELEVATE · THE · DEBATE

Git & GitHub for Social Science Research

Alyssa Harris
aharris@urban.org

Git is different than GitHub

Git is a free and open-source software for version control (does other nice things too).

GitHub is a private website, run for-profit, but has generously given Urban free accounts:

- Private: https://github.com/UI-Research/
- Public: https://github.com/UrbanInstitute/

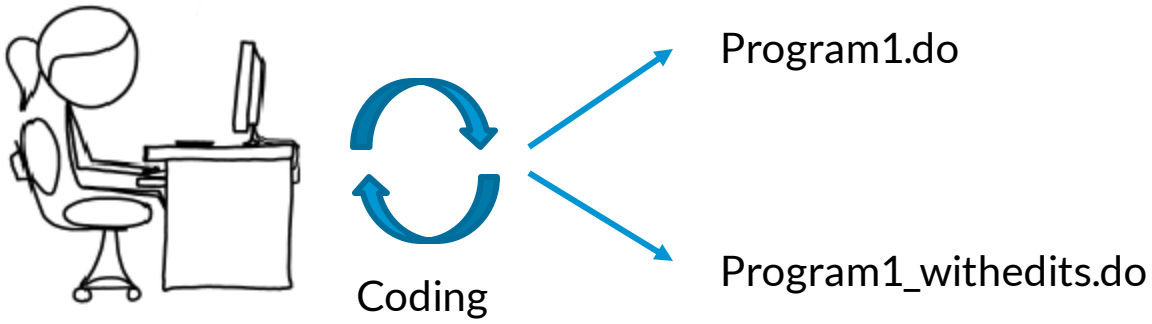
Git is different than GitHub

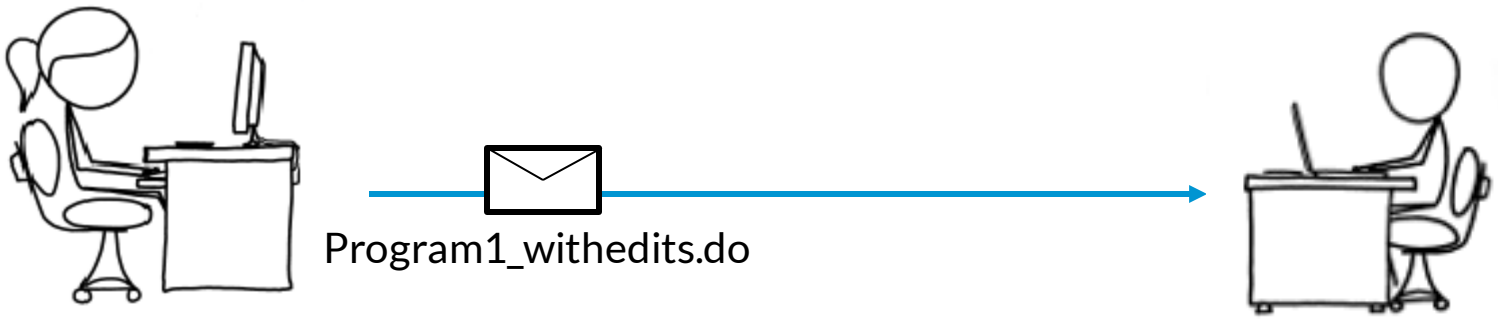
Git is the software.
Can't do anything without it.

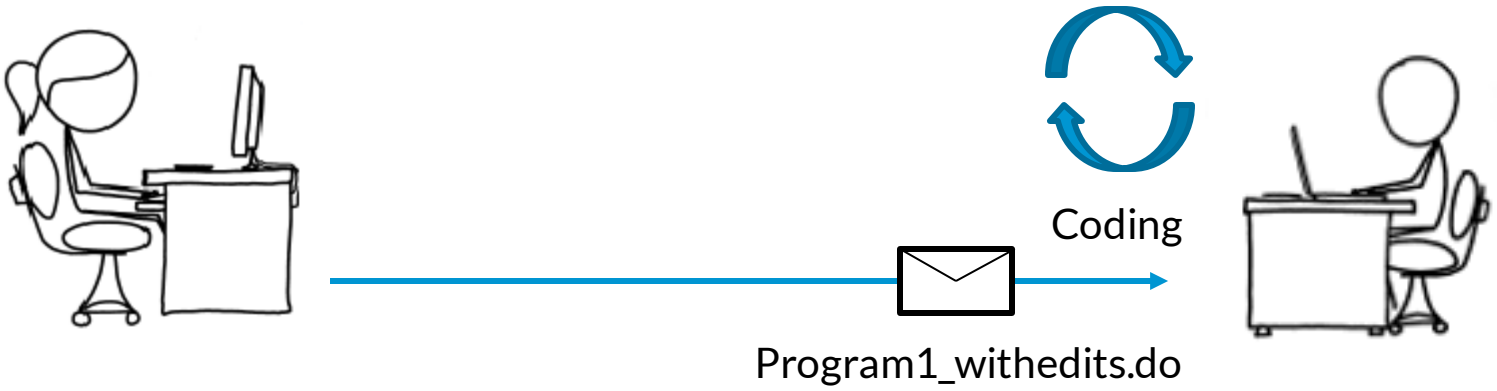
GitHub adds functionality and makes working with Git easier, but can't do anything on its own.

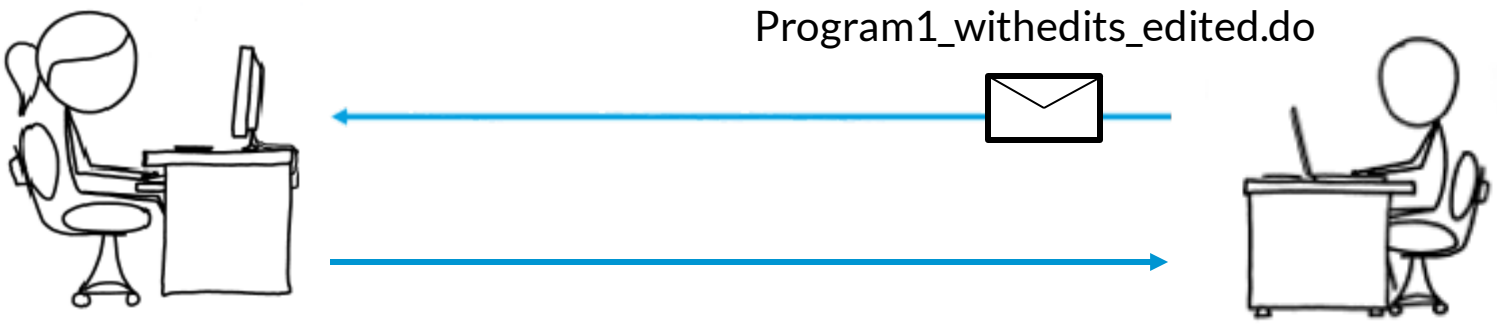
Slide prepared by Alex Engler and Jessica Kelly

Traditional Workflow


Program1.do
Program1_withedits.do
Coding


Program1_withedits.do


Coding
Program1_withedits.do


Program1_withedits_edited.do

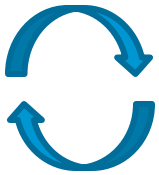
Program1.do

Program1_withedits.do

Program1_withedits_edited.do

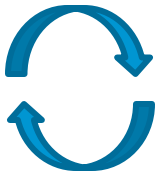
Using GitHub

Program1.do This is my first attempt



Coding

Program1.do Limit the dataset to only the relevant observations
Program1.do This is my first attempt



Coding

GitHub

Push

Program1.do Limit the dataset to only the relevant observations

Program1.do This is my first attempt

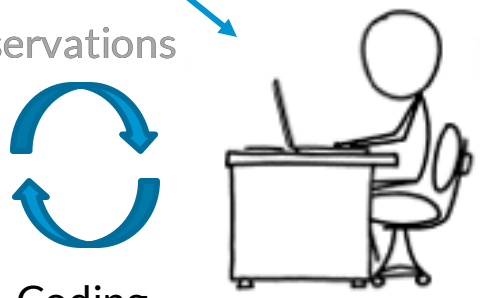
GitHub

Program1.do Limit the dataset to only the relevant observations
Program1.do This is my first attempt

Pull

GitHub

Program1.do	Add in the first two regressions
Program1.do	Limit the dataset to only the relevant observations
Program1.do	This is my first attempt

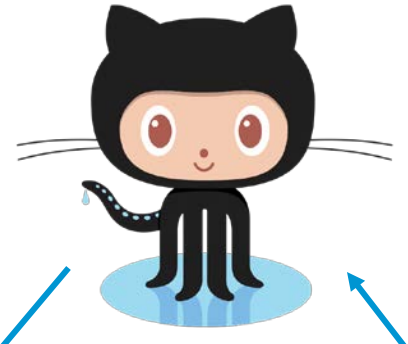


Coding

GitHub

Push

Program1.do	Add in the first two regressions
Program1.do	Limit the dataset to only the relevant observations
Program1.do	This is my first attempt



GitHub

Program1.do
Program1.do
Program1.do

But first: command line

Command line

Open the folder where you want to put your repository (for example, Documents)

Copy and paste the file path in the text box at the top

Leaving the folder open, type **cmd** into the search bar to open the command line window

cd

cd means "change directory"

Use it to switch folders in the command line window

In cmd, type >

cd D:\Documents

dir

dir will show you the contents of a directory in command line

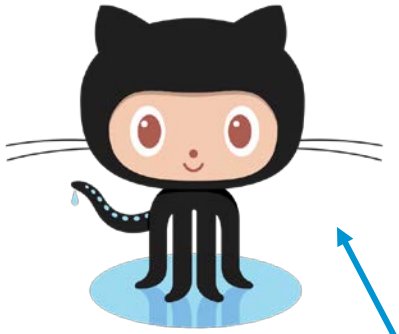
Use it to show the contents of a folder

In cmd, type >

dir

Now back to git

GitHub's naming convention



The **remote** repository on GitHub is called the **origin**

git clone

origin

```
Copyright © 2009 Microsoft Corporation

D:\Documents\> git clone
https://github.com/UI-Research/Github-
Training

Cloning into GitHub-Training…
```

git status

origin

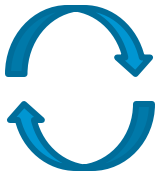
```
Copyright © 2009 Microsoft Corporation

D:\Documents\GitHub-Training> git status
On branch master

Initial commit

D:\Documents\GitHub-Training>
```

Let's create a new text file in the git repo:

MyTextFile.txt

`git add`

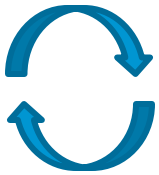
MyTextFile.txt



Coding

MyTextFile.txt

MyTextFile.txt



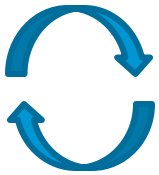
Coding

git add

MyTextFile.txt

\+ This is my text file

MyTextFile.txt



Coding

git commit

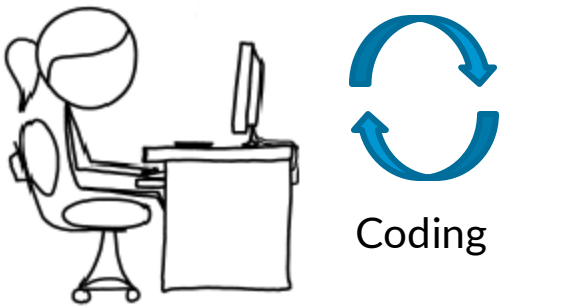
Commit message →

MyTextFile.txt

\+ This is my text file

My first commit!

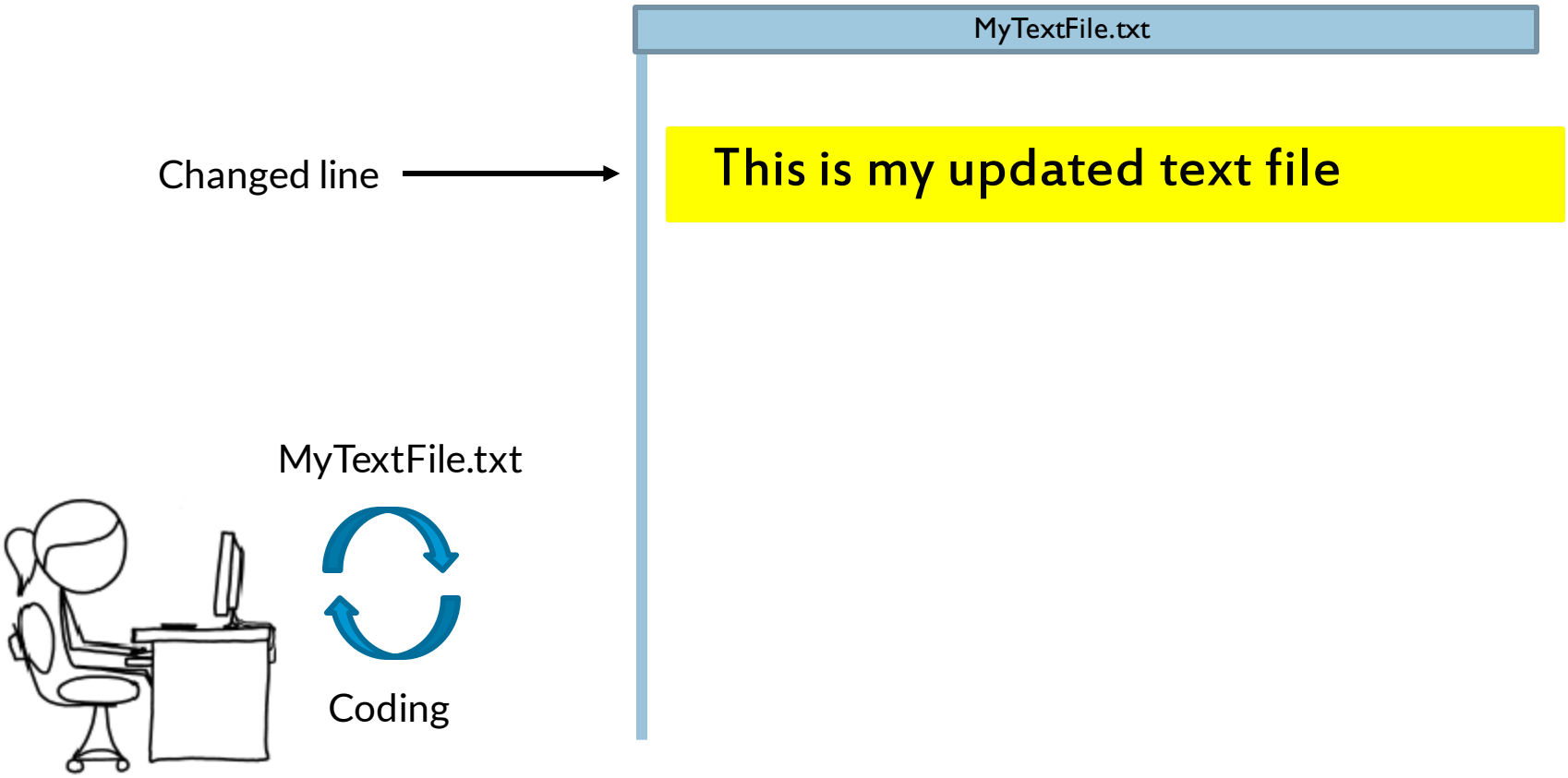
Let's change that line:



MyTextFile.txt

This is my updated text

Git sees this:



As this:

MyTextFile.txt

Deleted line → − This is my text file

Added line → + This is my updated text file

MyTextFile.txt

Coding

git commit

Commit message →

MyTextFile.txt

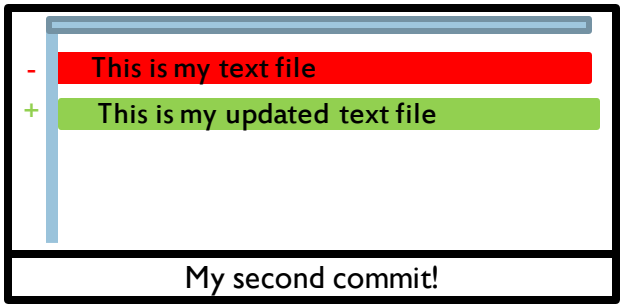
\- This is my text file

\+ This is my updated text file

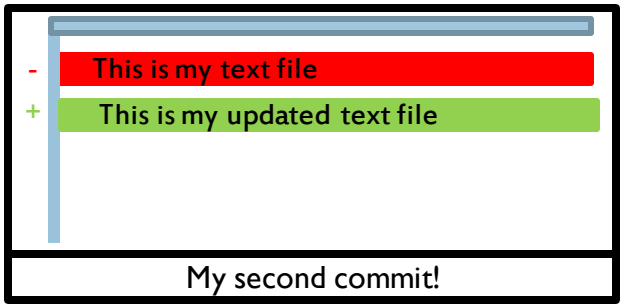
My second commit!

origin

Push


-
This is my text file
+
This is my updated text file
My second commit!

origin



`git pull`

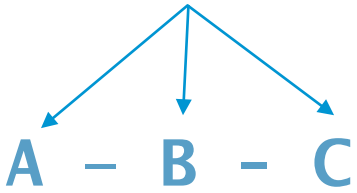
Pull

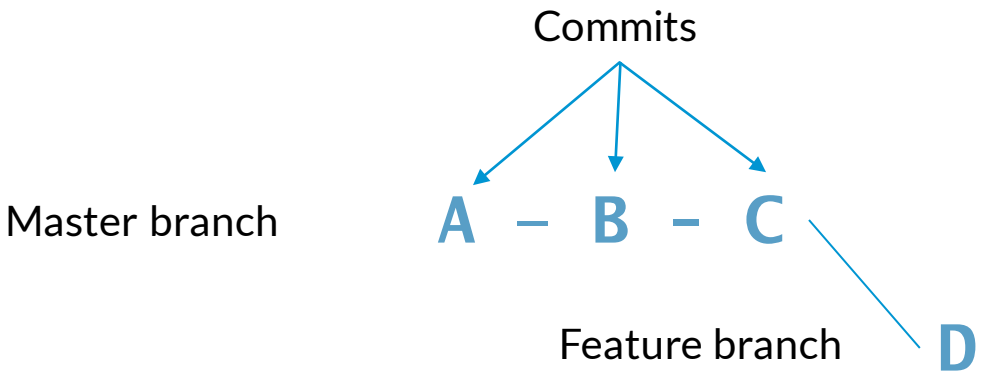
What if we want to have access to multiple snapshots at one time?

Using branches

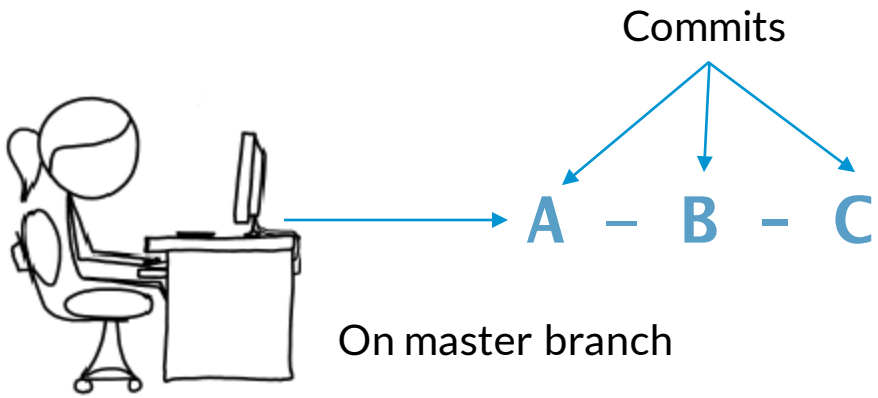
Commits

Master branch

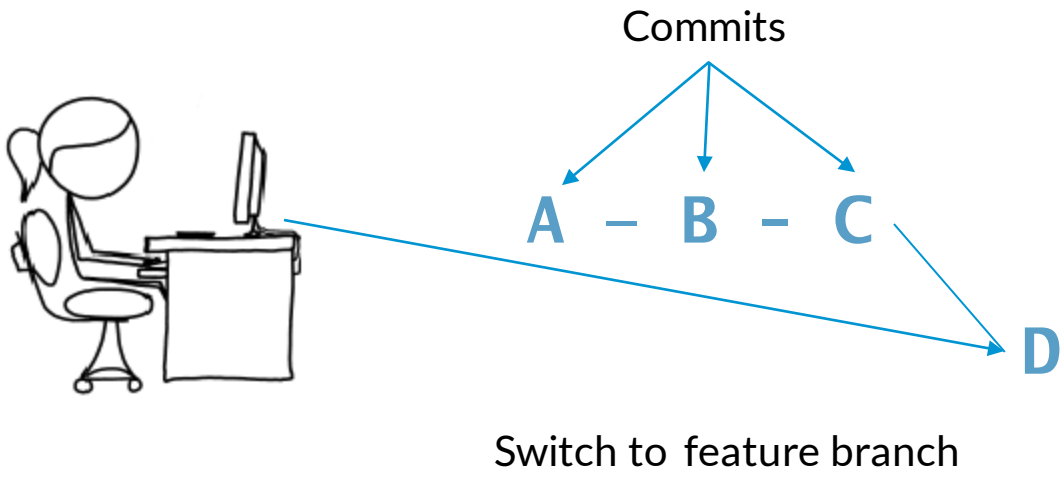



Commits
Master branch
A – B – C
Feature branch
D

git branch



git checkout



Switch to feature branch

`git checkout -b new-branch`

Create and switch to feature branch

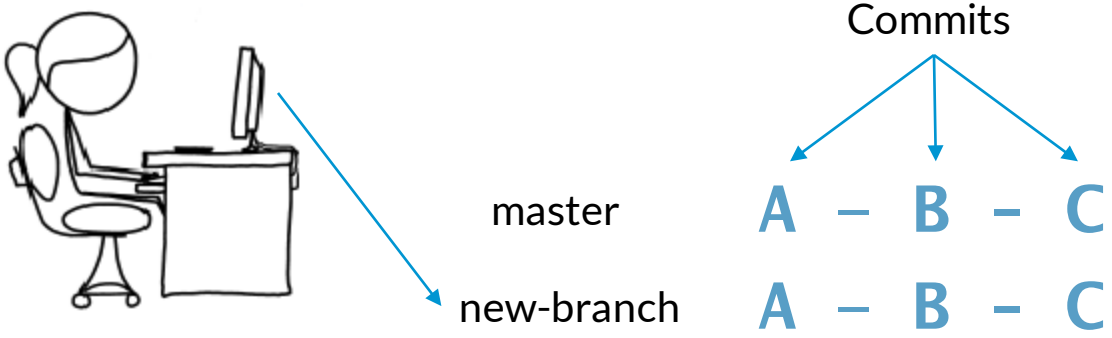
master

Commits

A – B – C

`git checkout -b new-branch`

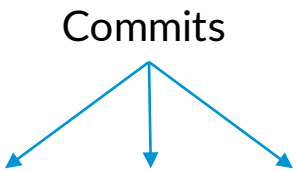
Create and switch to feature branch



`git checkout -b new-branch`

Create and switch to feature branch

Coding



master A – B – C

new-branch A – B – C – D

Still not convinced? Git also offers…

- Reproducible Research
  - Work is stored over the long term
  - Can easily be made public

- Stealing Code (in a good way)
  - Understanding Git & GitHub offers access to an incredible assortment of open source code and tools.

Slide prepared by Alex Engler and Jessica Kelly

Still not convinced? Git also offers…

- Distributed
  - Keeps copies of your work in several places
- Version Control
  - Keeps track of all your changes over time;
  - Makes sure you never lose code/progress;
- Tools for Collaboration
  - Allows for groups to work together on complex projects (e.g. TPC Model; Comms's Interactive Web Graphics)

Slide prepared by Alex Engler and Jessica Kelly