

# git and github

a primer

carniello  
march 14 2016

# what you will learn in the next several minutes ...

- git
- github
- how we are using github to host our repository

## executive summary

- needed new home, googlecode closed
- github used by many many open source projects
- pretty easy to see what's there, what we offer programming community
- some hurdles in understanding how to best work with it
- that's what we'll cover now, later today, and tonight!

# git

- drafted by linus torvalds to be the source control system for the linux kernel
- command line interface
- content tracking system : not “version control”

# what is git?

- it's a repository to hold files and track changes to those files
- has the ability to go back in time and recreate what was
- it's different than “standard” version control approaches
  - it does not track the changes ...
- it stores every version of every file - not differences !

# git is a repository

- starts empty upon creation - then one adds stuff to it
- this repository has two parts
  - object store
  - index

# objects

- BLOBs - binary large objects (“stuff”)
- trees - directories
- commits - metadata of change from one state to another
- tags - names of the object

# index

- temporary and dynamic file which describes the directory at some given point in time



objects are just copies of this stuff ?

NO !!!

the SHA1 hashing algorithm is used to compute the hash value of an incoming object, giving it a unique name and identifier

## here's the idea

the object store grows as new objects are added, or existing ones are modified or deleted

each state of each object is stored - different hash id's, then

the object store is based on the hashed computation of the contents of its objects - not file nor directory names

when a human looks at this - the view shows a file having revisions, a progression through states - but that's an artifact, it's the way that git renders to us primates

# object organization

a BLOB is at the bottom - it's a file

a tree (directory) points to a BLOB, and maybe other trees

a commit points to one tree, tracks changes

a tag is the object name, taken from its hash

## git is command-line-based

but that's too unwieldy for us, we think

need a place to put the global repository

need local repository to sync with the global repository

lots of issues with local pc admin permissions and all that

can git exist solely in the cloud ?

# git in the cloud: github

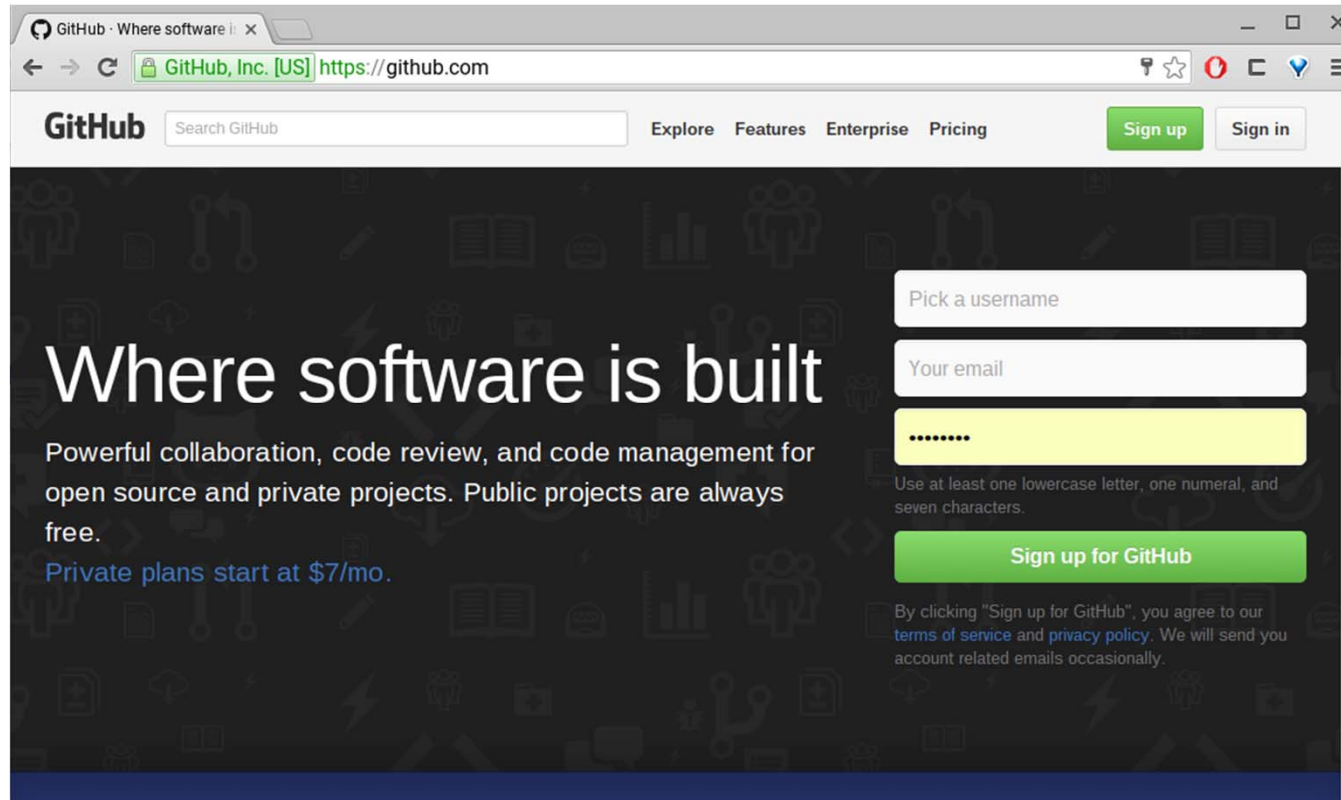
an environment for hosting a repository

we'll use github as an interface to OUR git repository -

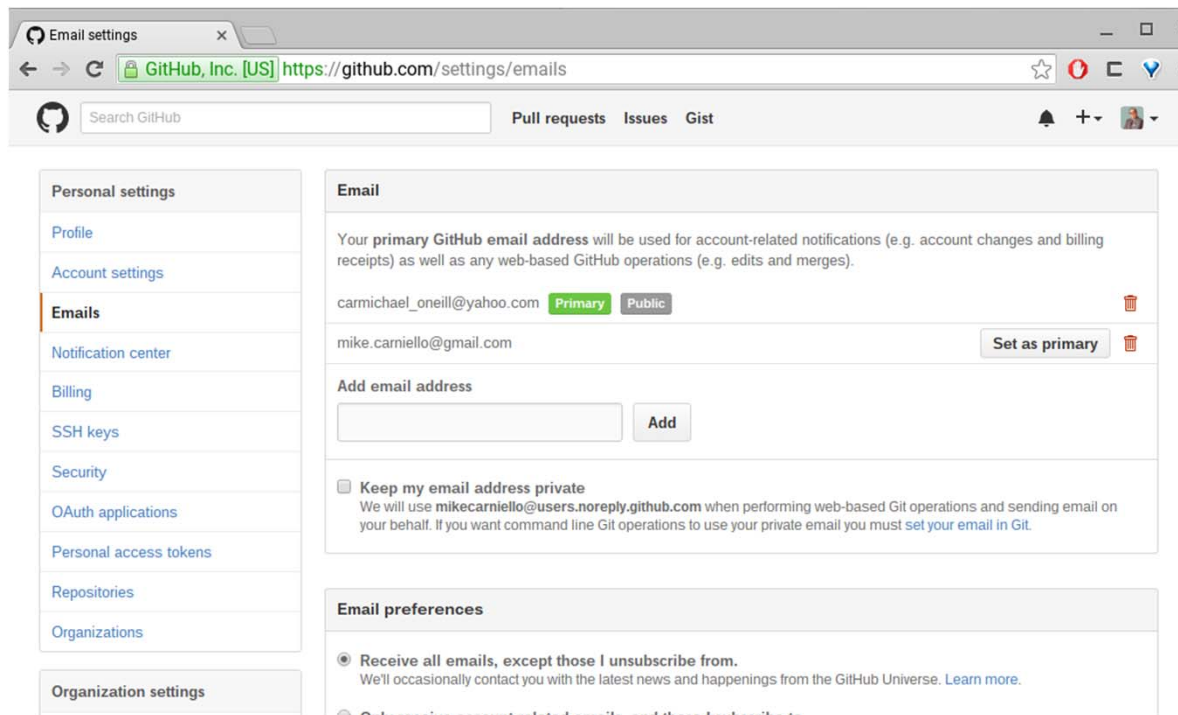
**phuse-org/phuse-scripts**

today we'll discuss how we will use github ...

github



# landing github page for a github user



# landing github page for a github user

The screenshot shows a web browser window displaying the GitHub profile page for user 'mikecarniello'. The browser's address bar shows the URL 'https://github.com/mikecarniello'. The page features a profile picture of Mike Carniello, a man in a suit and glasses. Below the picture, the name 'Mike Carniello' and the username 'mikecarniello' are displayed, along with the text 'Joined on Mar 30, 2012'. There are three statistics: '0 Followers', '0 Starred', and '0 Following'. The main content area includes tabs for 'Contributions', 'Repositories', and 'Public activity', with an 'Edit profile' button. Under 'Popular repositories', the repository 'css2016' is listed with 0 stars. Under 'Repositories contributed to', 'phuse-org/phuse-scripts' is listed with 4 stars. A 'Contributions' section shows a calendar grid with green squares indicating contributions on specific days. At the bottom, three summary boxes show: 'Contributions in the last year: 5 total (Feb 26, 2015 - Feb 26, 2016)', 'Longest streak: 1 day (April 14 - April 14)', and 'Current streak: 0 days (Last contributed 8 days ago)'.

Contributions in the last year: 5 total (Feb 26, 2015 - Feb 26, 2016)

Longest streak: 1 day (April 14 - April 14)

Current streak: 0 days (Last contributed 8 days ago)

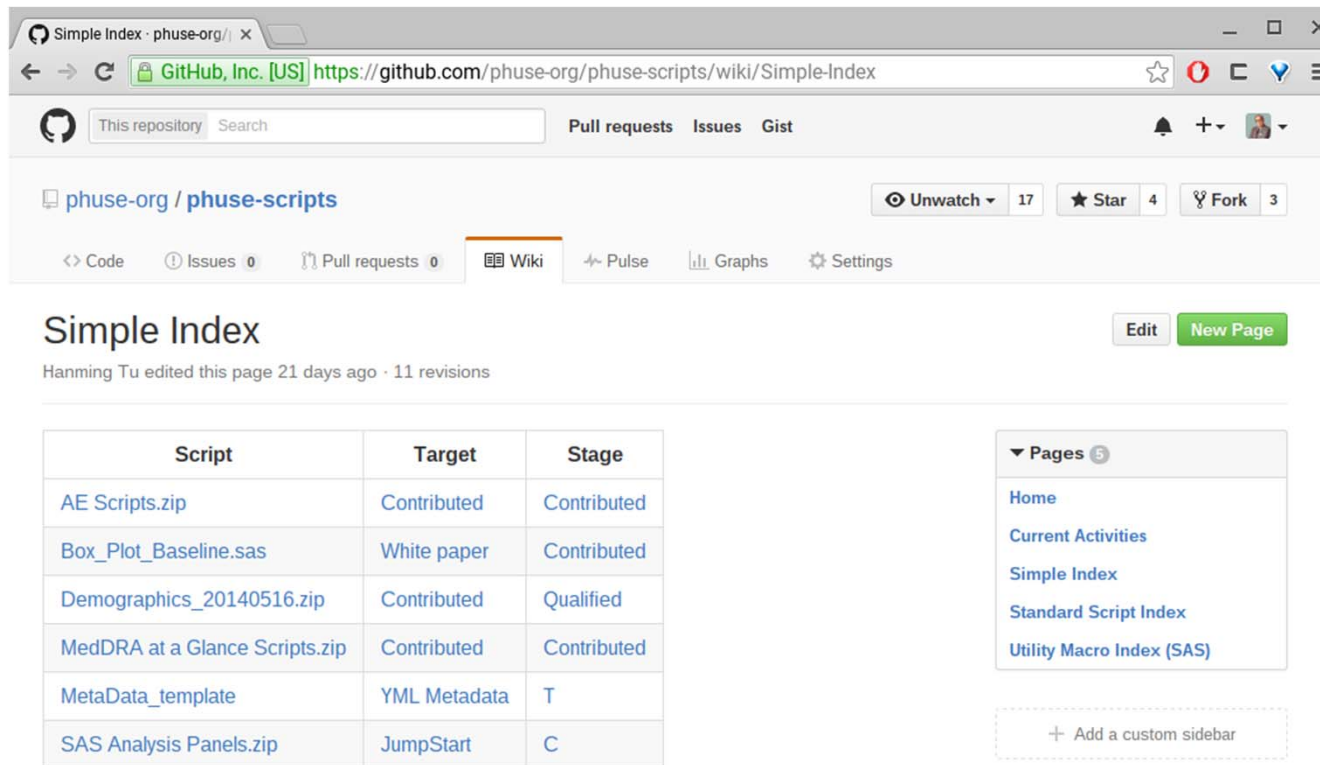


# our repository main page

The screenshot shows the GitHub repository page for `phuse-org/phuse-scripts`. The browser address bar shows the URL `https://github.com/phuse-org/phuse-scripts`. The repository name is `phuse-org / phuse-scripts`. The page includes navigation tabs for `Code`, `Issues`, `Pull requests`, `Wiki`, `Pulse`, `Graphs`, and `Settings`. A description of the repository is provided: "Delivery standard industry analyses, built upon CDISC standards for analysis data [http://www.phusewiki.org/wiki/index.php?title=Standard\\_Scripts](http://www.phusewiki.org/wiki/index.php?title=Standard_Scripts) — Edit". Below the description, statistics are shown: 600 commits, 1 branch, 0 releases, and 6 contributors. A horizontal progress bar is visible. The page also features a "Branch: master" dropdown, a "New pull request" button, and buttons for "New file", "Upload files", "Find file", "HTTPS", and "Download ZIP". A commit history table is displayed, showing the latest commit by DanteDT and a list of previous commits with their messages and dates.

Commit	Message	Time
DanteDT	Update TODO.md	7 hours ago
contributed	No commit message	21 days ago
data	Rename EADME.md to README.md	10 days ago
development	Create foo.txt	5 months ago
docs	Update MarkdownTips.md	8 days ago
lang	Update SAS2XLXS macro	3 months ago

# what we're presenting to the community

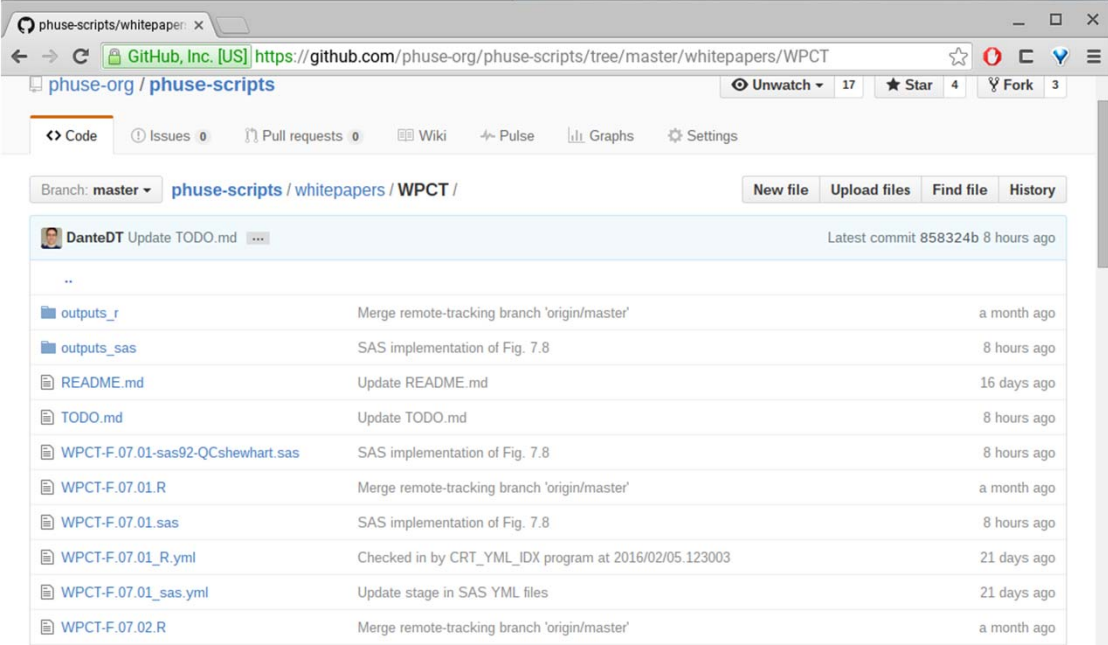


The screenshot shows a GitHub Wiki page for the repository 'phuse-org / phuse-scripts'. The page title is 'Simple Index', and it was last edited by Hanming Tu 21 days ago with 11 revisions. The page contains a table with three columns: 'Script', 'Target', and 'Stage'. The table lists six scripts: 'AE Scripts.zip', 'Box\_Plot\_Baseline.sas', 'Demographics\_20140516.zip', 'MedDRA at a Glance Scripts.zip', 'MetaData\_template', and 'SAS Analysis Panels.zip'. A sidebar on the right shows a list of pages: 'Home', 'Current Activities', 'Simple Index', 'Standard Script Index', and 'Utility Macro Index (SAS)'. There are also buttons for 'Edit' and 'New Page'.

Script	Target	Stage
AE Scripts.zip	Contributed	Contributed
Box_Plot_Baseline.sas	White paper	Contributed
Demographics_20140516.zip	Contributed	Qualified
MedDRA at a Glance Scripts.zip	Contributed	Contributed
MetaData_template	YML Metadata	T
SAS Analysis Panels.zip	JumpStart	C

# working with our repository

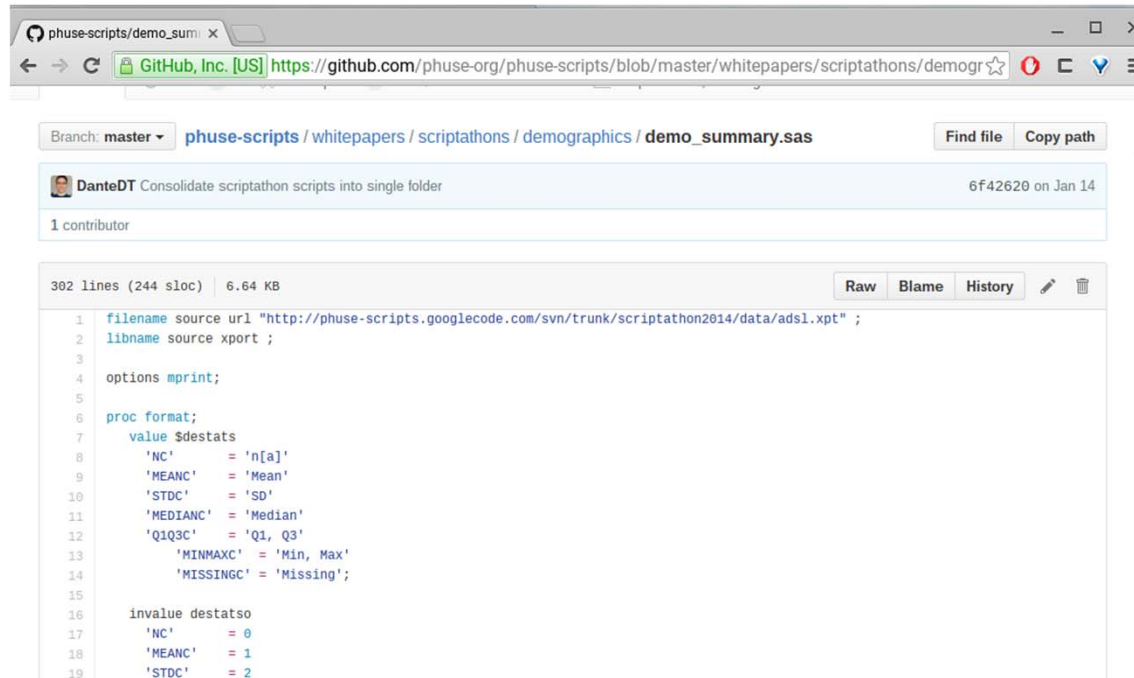
## -- browse contents



The screenshot shows a web browser window displaying the GitHub repository page for 'phuse-org / phuse-scripts'. The browser's address bar shows the URL 'https://github.com/phuse-org/phuse-scripts/tree/master/whitepapers/WPCT'. The repository page includes navigation tabs for 'Code', 'Issues', 'Pull requests', 'Wiki', 'Pulse', 'Graphs', and 'Settings'. Below the navigation, the current branch is 'master' and the path is 'phuse-scripts / whitepapers / WPCT /'. There are buttons for 'New file', 'Upload files', 'Find file', and 'History'. The main content area shows a list of files and directories with their commit details:

File/Directory	Commit Message	Time Ago
..		
outputs_r	Merge remote-tracking branch 'origin/master'	a month ago
outputs_sas	SAS implementation of Fig. 7.8	8 hours ago
README.md	Update README.md	16 days ago
TODO.md	Update TODO.md	8 hours ago
WPCT-F.07.01-sas92-QCshewhart.sas	SAS implementation of Fig. 7.8	8 hours ago
WPCT-F.07.01.R	Merge remote-tracking branch 'origin/master'	a month ago
WPCT-F.07.01.sas	SAS implementation of Fig. 7.8	8 hours ago
WPCT-F.07.01_R.yml	Checked in by CRT_YML_IDX program at 2016/02/05.123003	21 days ago
WPCT-F.07.01_sas.yml	Update stage in SAS YML files	21 days ago
WPCT-F.07.02.R	Merge remote-tracking branch 'origin/master'	a month ago

working with our repository  
-- grab a script for your own use



The screenshot shows a web browser window displaying a GitHub repository page. The browser's address bar shows the URL: [https://github.com/phuse-org/phuse-scripts/blob/master/whitepapers/scriptathons/demographics/demo\\_summary.sas](https://github.com/phuse-org/phuse-scripts/blob/master/whitepapers/scriptathons/demographics/demo_summary.sas). The page header indicates the branch is 'master' and the file path is 'phuse-scripts / whitepapers / scriptathons / demographics / demo\_summary.sas'. A commit message by 'DanteDT' is visible: 'Consolidate scriptathon scripts into single folder', dated '6f42620 on Jan 14'. The file size is listed as '302 lines (244 sloc) | 6.64 KB'. The main content area displays the SAS code for the script, which includes source file definitions, options, and a PROC FORMAT statement for statistical output.

```
1 filename source url "http://phuse-scripts.googlecode.com/svn/trunk/scriptathon2014/data/ads1.xpt" ;
2 libname source xport ;
3
4 options mprint;
5
6 proc format;
7   value $destats
8     'NC'      = 'n[a]'
9     'MEANC'   = 'Mean'
10    'STDC'    = 'SD'
11    'MEDIANC' = 'Median'
12    'Q1Q3C'   = 'Q1, Q3'
13    'MINMAXC' = 'Min, Max'
14    'MISSINGC' = 'Missing';
15
16 invalue destatso
17   'NC'      = 0
18   'MEANC'   = 1
19   'STDC'    = 2
```

## working with our repository

-- edit a script - indirectly

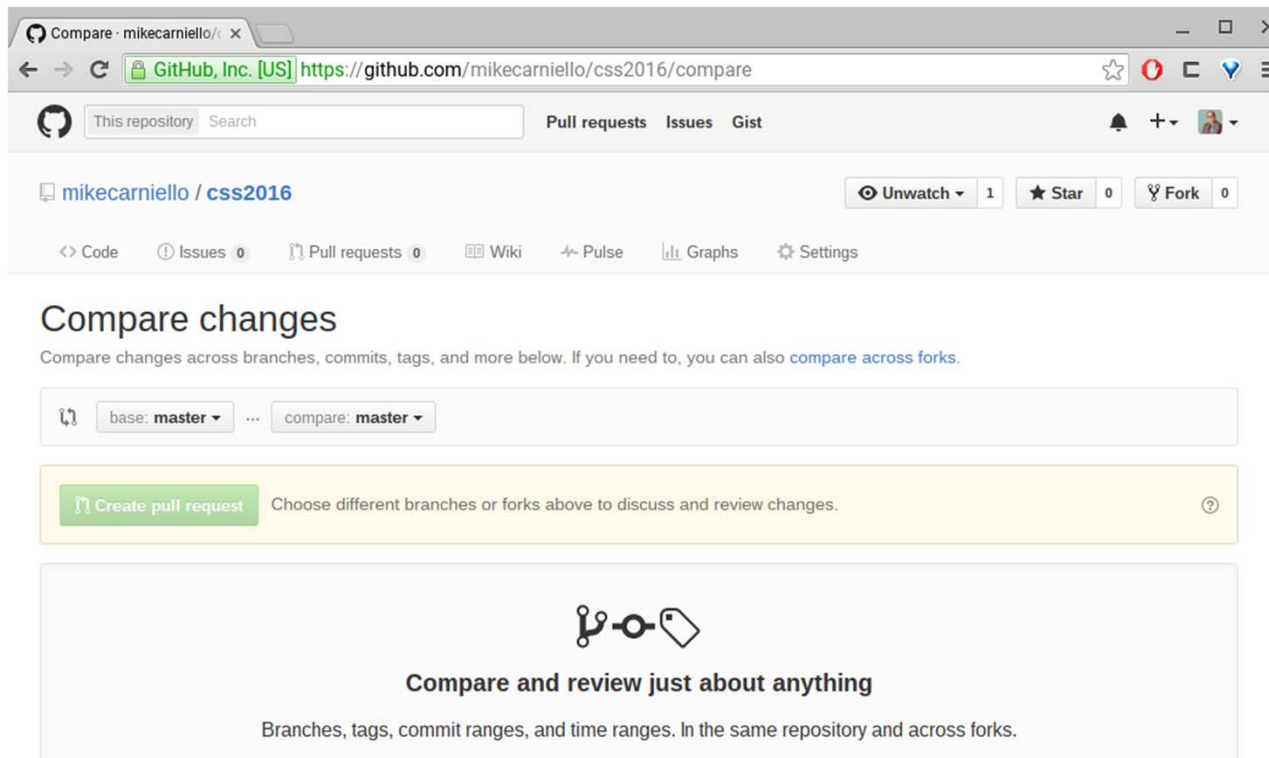
--- fork project

--- edit your local version

--- start a pull request

a **pull request** is a request **from** you, the developer, **to** a maintainer, to **pull** your changes into the master branch of the repository

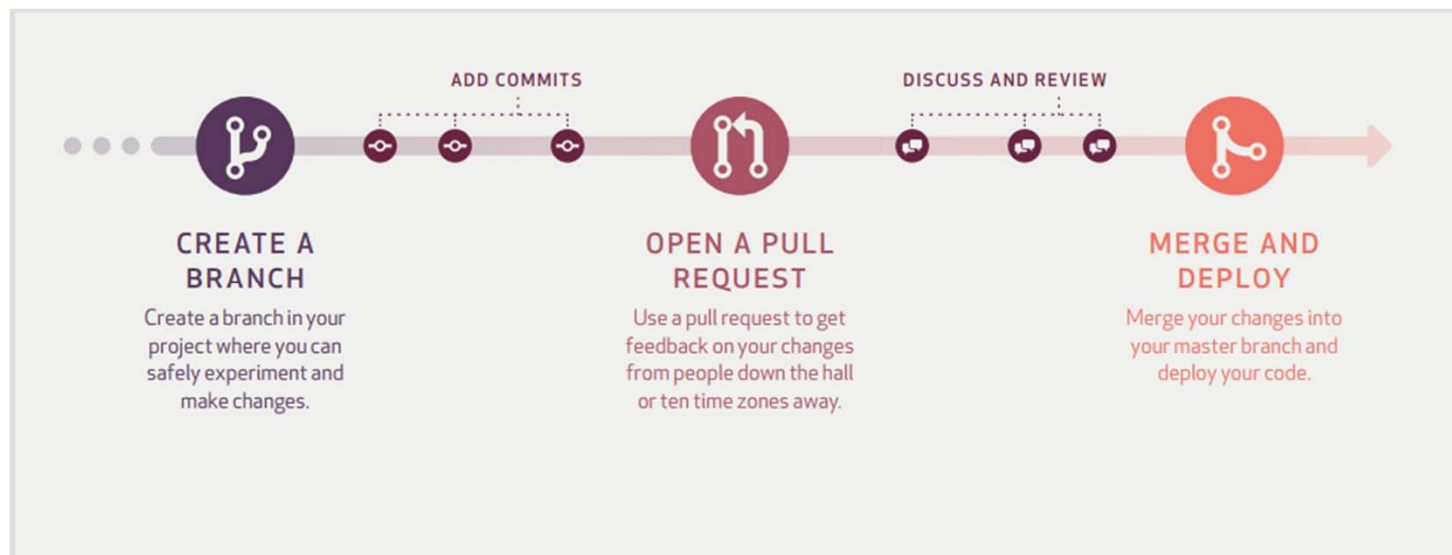
pull request: maintainer compares against “master”



# github philosophy

-- somebody has done a swell job of this already:

<https://guides.github.com/introduction/flow/>



# git and github references

pro git:

<https://git-scm.com/book/en/v2>

github guides

<https://guides.github.com/>

books

<http://www.it-ebooks.org/tag/git>