

Action Items

- Group deadline is today at 11:59:59 PM
 - See Lecture 01 Notes on steps to take
 - 4-5 Students per group
 - Email necessary items to me
- After deadline no additional group formations will be accepted
- Only 18/38 students, Filled out the GitHub Form
 - fill out, so your assignment will be graded
- HW1 Assignment is out
 - See Syllabus on how to submit
- You all have access to the course repo

Grader

- Chandan Kumar Chunduru
- GitHub ID: Chandan-Chunduru
- **Add to your assignment repo**

Tools

CSC 405/605

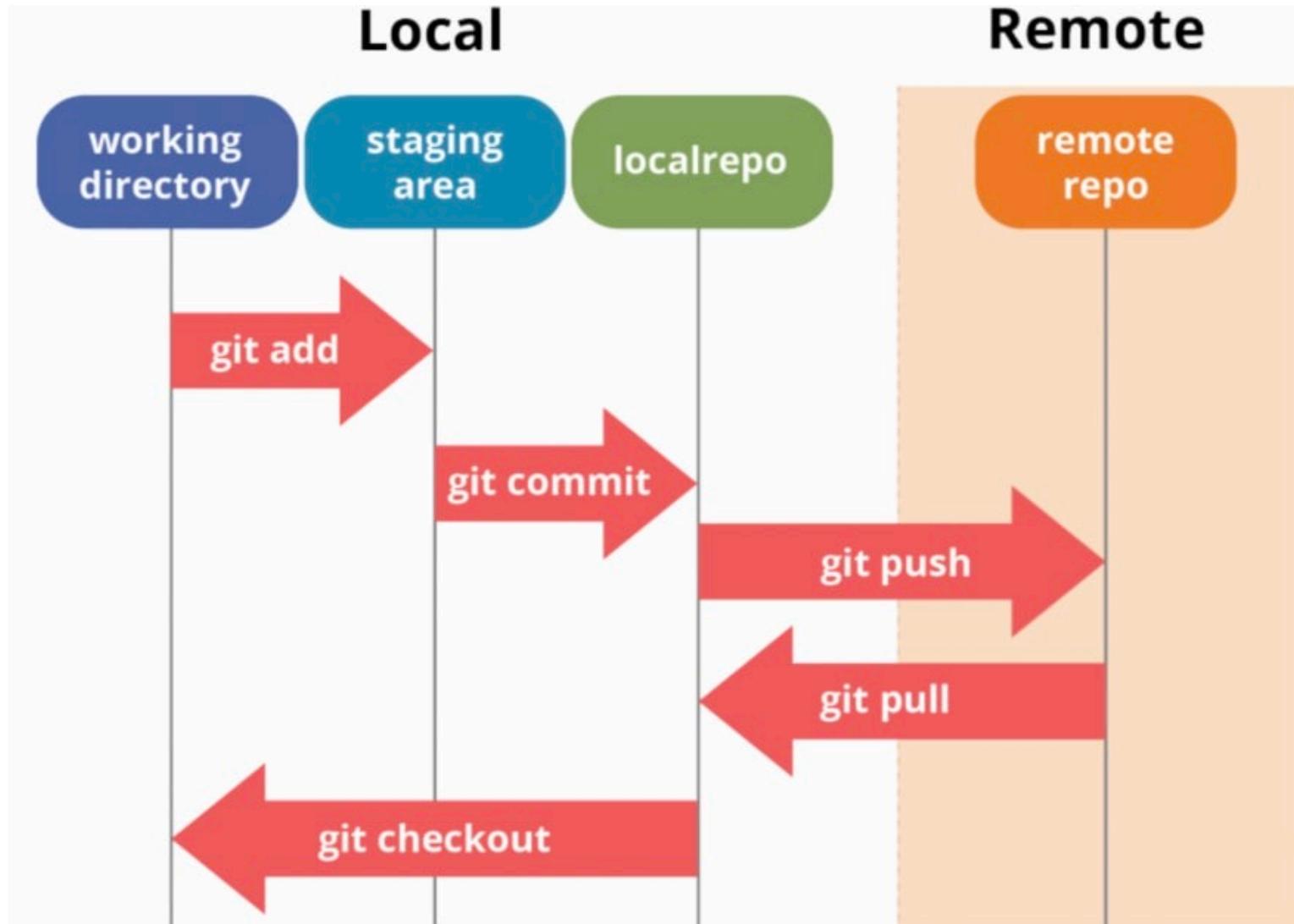
Version Control

- Practice of tracking and managing changes to files
- Useful in Data Science
 - Make Backups
 - Keep a history of changes
 - View changes
 - Experiment
 - Collaborate
- Version Control software
 - git
- git is useful in Industry

git

- What is git?
 - Distributed Version Control System
 - Created by Linus for the development of the Linux Kernel
 - Can Use for Large Projects with many collaborators
 - Decentralized, Fast, and Flexible
- Many Git Servers
 - We will use GitHub (<https://education.github.com/pack>)
- git Tools
 - Command Line git (Recommend that you learn this)
 - GitHub Desktop (<https://desktop.github.com/>)

git



git Setup

- Windows
 - <https://gitforwindows.org/>
- Mac
 - <https://git-scm.com/download/mac>
- Linux
 - <https://git-scm.com/download/linux>
- Guides:
 - <https://www.atlassian.com/git/tutorials/what-is-version-control>
 - <http://rogerdudler.github.io/git-guide/>
 - http://slides.com/moji3000/git_presentation/fullscreen#/

git Setup

- Configuration
 - Add your user name and email
 - Replace my name/email with yours

```
(base) jarvis@Kelechis-MBP ~ % git config --global user.name "Kelechi Ikegwu"  
(base) jarvis@Kelechis-MBP ~ % git config --global user.email " "
```

- Double Check that it is correct with
- git config -l

```
(base) jarvis@Kelechis-MBP ~ % git config -l  
credential.helper=osxkeychain  
filter.lfs.clean=git-lfs clean -- %f  
filter.lfs.smudge=git-lfs smudge -- %f  
filter.lfs.process=git-lfs filter-process  
filter.lfs.required=true  
user.name=Kelechi Ikegwu  
user.email=kelechi.ikegwu@gmail.com  
(base) jarvis@Kelechis-MBP ~ %
```

git Setup

- Create a remote repository called CSC-Test
- Important
 - README
 - .gitignore
 - License
 - Leave them unchecked for now

Create a new repository

A repository contains all project files, including the revision history. Already have a project repository elsewhere? [Import a repository](#).

Owner * Repository name *

 ikegwukc / CSC-Test ✓

Great repository names are short and memorable. Need inspiration? How about [bug-free-fiesta](#)?

Description (optional)

 **Public**
Anyone on the internet can see this repository. You choose who can commit.

 **Private**
You choose who can see and commit to this repository.

Initialize this repository with:

Skip this step if you're importing an existing repository.

Add a README file
This is where you can write a long description for your project. [Learn more](#).

Add .gitignore
Choose which files not to track from a list of templates. [Learn more](#).

Choose a license
A license tells others what they can and can't do with your code. [Learn more](#).

Create repository

git

- git init
 - Create a new repository locally

```
(base) jarvis@Kelechis-MBP ~ % mkdir CSC-DataScienceExample
(base) jarvis@Kelechis-MBP ~ % cd CSC-DataScienceExample
(base) jarvis@Kelechis-MBP CSC-DataScienceExample % git init
hint: Using 'master' as the name for the initial branch. This default branch name
hint: is subject to change. To configure the initial branch name to use in all
hint: of your new repositories, which will suppress this warning, call:
hint:
hint:   git config --global init.defaultBranch <name>
hint:
hint: Names commonly chosen instead of 'master' are 'main', 'trunk' and
hint: 'development'. The just-created branch can be renamed via this command:
hint:
hint:   git branch -m <name>
Initialized empty Git repository in /Users/jarvis/CSC-DataScienceExample/.git/
```

git

- git init

```
(base) jarvis@Kelechis-MBP CSC-DataScienceExample % ls
(base) jarvis@Kelechis-MBP CSC-DataScienceExample % ls -a
.  ..  .git
(base) jarvis@Kelechis-MBP CSC-DataScienceExample % tree .git
.git
├── HEAD
├── config
├── description
└── hooks
    ├── applypatch-msg.sample
    ├── commit-msg.sample
    ├── fsmonitor-watchman.sample
    ├── post-update.sample
    ├── pre-applypatch.sample
    ├── pre-commit.sample
    ├── pre-merge-commit.sample
    ├── pre-push.sample
    ├── pre-rebase.sample
    ├── pre-receive.sample
    ├── prepare-commit-msg.sample
    ├── push-to-checkout.sample
    └── update.sample
└── info
    └── exclude
└── objects
    ├── info
    └── pack
└── refs
    ├── heads
    └── tags

8 directories, 17 files
```

git

- git status
 - Status of the repo

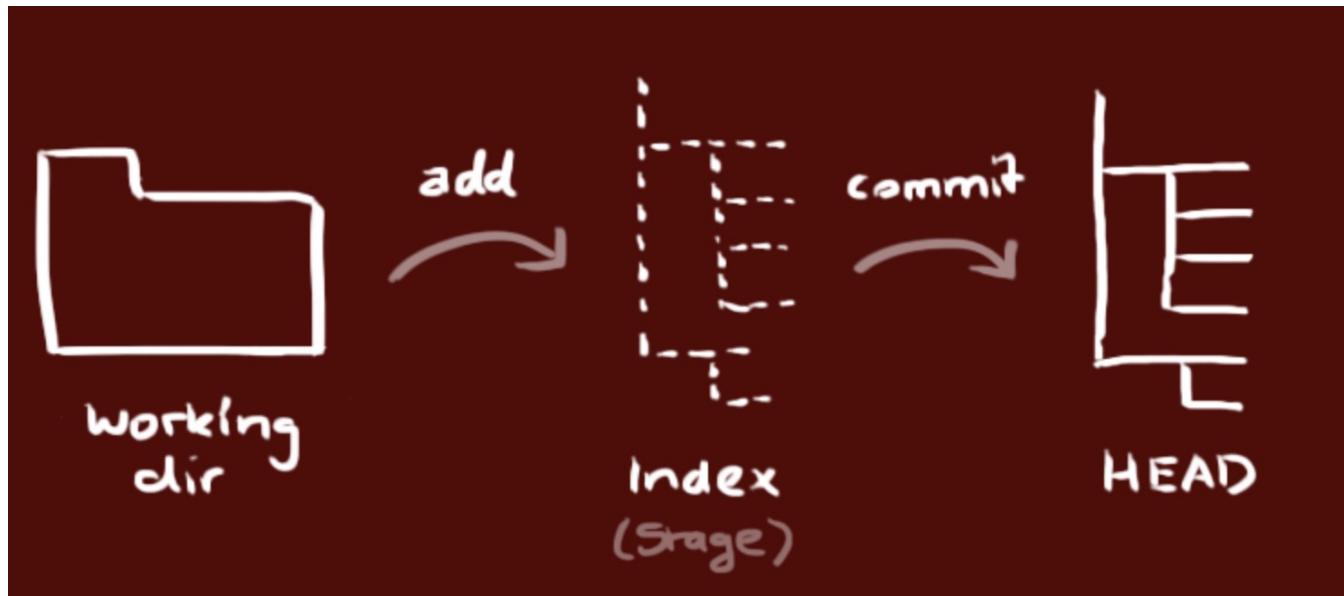
```
(base) jarvis@Kelechis-MBP CSC-DataScienceExample % git status
On branch master

No commits yet

nothing to commit (create/copy files and use "git add" to track)
```

git workflow

- Working Directory – Actual Files
- Index – Staging Area
 - Think of as rough drafts
- HEAD
 - Points to last commit you've made



git

- git add
 - Add a file to your working directory called “README.md”
 - (How you add the file is not important)
 - Add some text to the file and save it
 - (How you add text to the file is not important)

```
(base) jarvis@Kelechis-MBP CSC-DataScienceExample % touch README.md
(base) jarvis@Kelechis-MBP CSC-DataScienceExample % echo '# Welcome to CSC 405 605' > README.md
(base) jarvis@Kelechis-MBP CSC-DataScienceExample % cat README.md
# Welcome to CSC 405 605
```

git

- git add
 - Add README.md to Staging Area

```
(base) jarvis@Kelechis-MBP CSC-DataScienceExample % git add README.md
(base) jarvis@Kelechis-MBP CSC-DataScienceExample % git status
On branch master

No commits yet

Changes to be committed:
  (use "git rm --cached <file>..." to unstage)
    new file:   README.md
```

git

- git commit
 - Commit changes made to local repo

```
(base) jarvis@Kelechis-MBP CSC-DataScienceExample % git commit -m "Example for CSC405605"
[master (root-commit) f1e78cc] Example for CSC405605
 1 file changed, 1 insertion(+)
 create mode 100644 README.md
```

git

- Git push

```
(base) jarvis@Kelechis-MBP CSC-DataScienceExample % git push  
fatal: No configured push destination.
```

Either specify the URL from the command-line or configure a remote repository using

```
git remote add <name> <url>
```

and then push using the remote name

```
git push <name>
```

git

- git push

```
(base) jarvis@Kelechis-MBP CSC-DataScienceExample % git remote add example https://github.com/ikegwukc/CSC-Test
(base) jarvis@Kelechis-MBP CSC-DataScienceExample % git push --set-upstream example master
Username for 'https://github.com': ikegwukc
Password for 'https://ikegwukc@github.com':
remote: Support for password authentication was removed on August 13, 2021. Please use a personal access token instead.
remote: Please see https://github.blog/2020-12-15-token-authentication-requirements-for-git-operations/ for more information.
fatal: Authentication failed for 'https://github.com/ikegwukc/CSC-Test/'
```

- <https://docs.github.com/en/authentication/keeping-your-account-and-data-secure/creating-a-personal-access-token>

git

- git push
 - Add Personal Authentication Token and try again

```
(base) jarvis@Kelechis-MBP CSC-DataScienceExample % git push --set-upstream example master
Username for 'https://github.com': ikegwukc
Password for 'https://ikegwukc@github.com':
Enumerating objects: 3, done.
Counting objects: 100% (3/3), done.
Writing objects: 100% (3/3), 253 bytes | 253.00 KiB/s, done.
Total 3 (delta 0), reused 0 (delta 0), pack-reused 0
To https://github.com/ikegwukc/CSC-Test
 * [new branch]      master -> master
Branch 'master' set up to track remote branch 'master' from 'example'.
(base) jarvis@Kelechis-MBP CSC-DataScienceExample % █
```

git

ikegwukc / CSC-Test Public

Unwatch 1 ▾ Fork 0 Star 0 ▾

< Code Issues Pull requests Actions Projects Wiki Security Insights Settings

master ▾ 1 branch 0 tags Go to file Add file ▾ Code ▾

ikegwukc Example for CSC405605 f1e78cc 22 minutes ago 1 commit

README.md Example for CSC405605 22 minutes ago

README.md

Welcome to CSC 405 605

About

No description, website, or topics provided.

Readme 0 stars 1 watching 0 forks

Releases

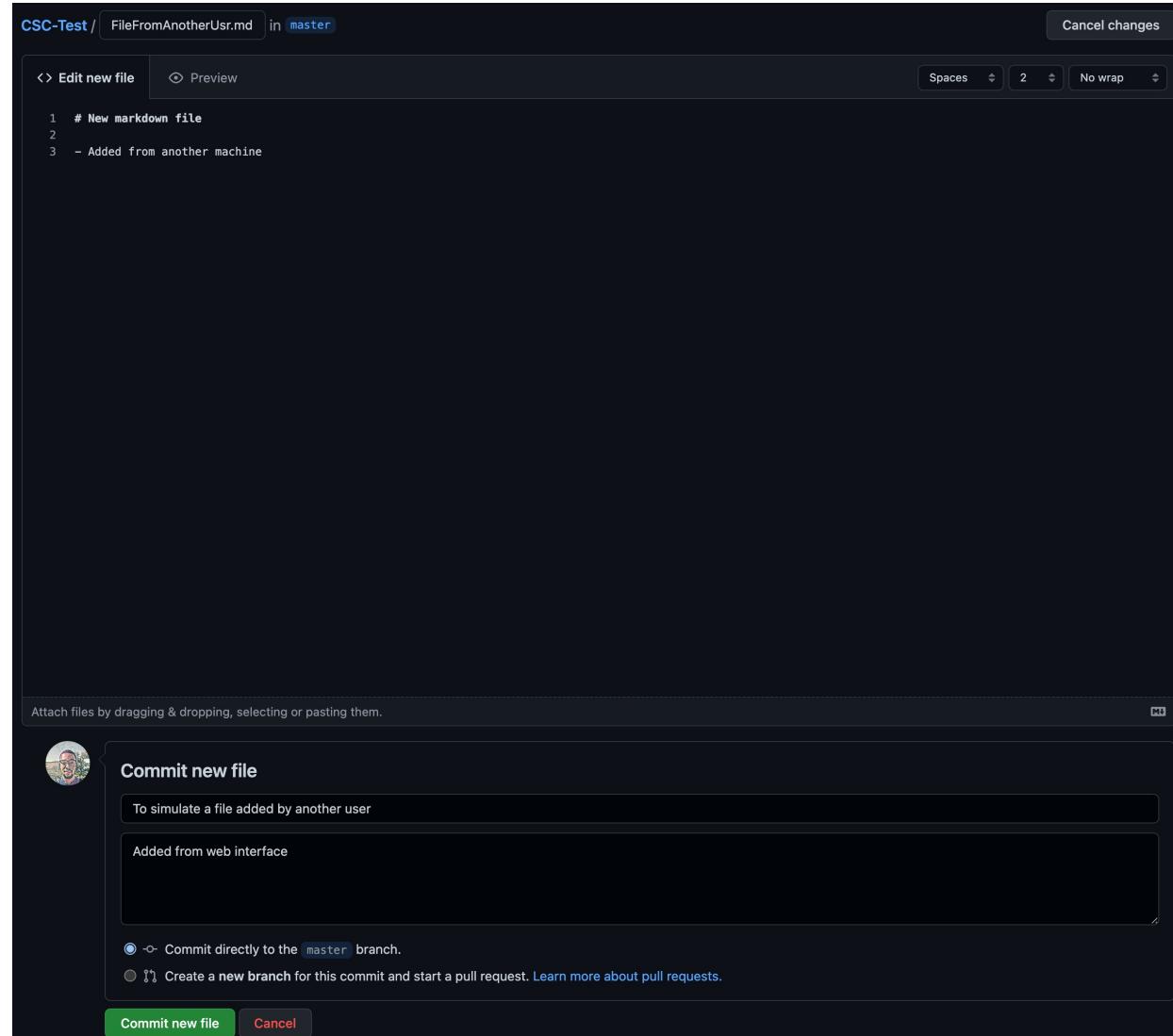
No releases published Create a new release

Packages

No packages published
Publish your first package

git

- Simulate another user adding files to GitHub repo



The screenshot shows a GitHub commit dialog for a file named "FileFromAnotherUsr.md" in the "master" branch of a repository named "CSC-Test". The file content is:

```
1 # New markdown file
2
3 - Added from another machine
```

The dialog includes options for "Edit new file" and "Preview". At the bottom, there is a "Commit new file" button and a "Cancel" button.

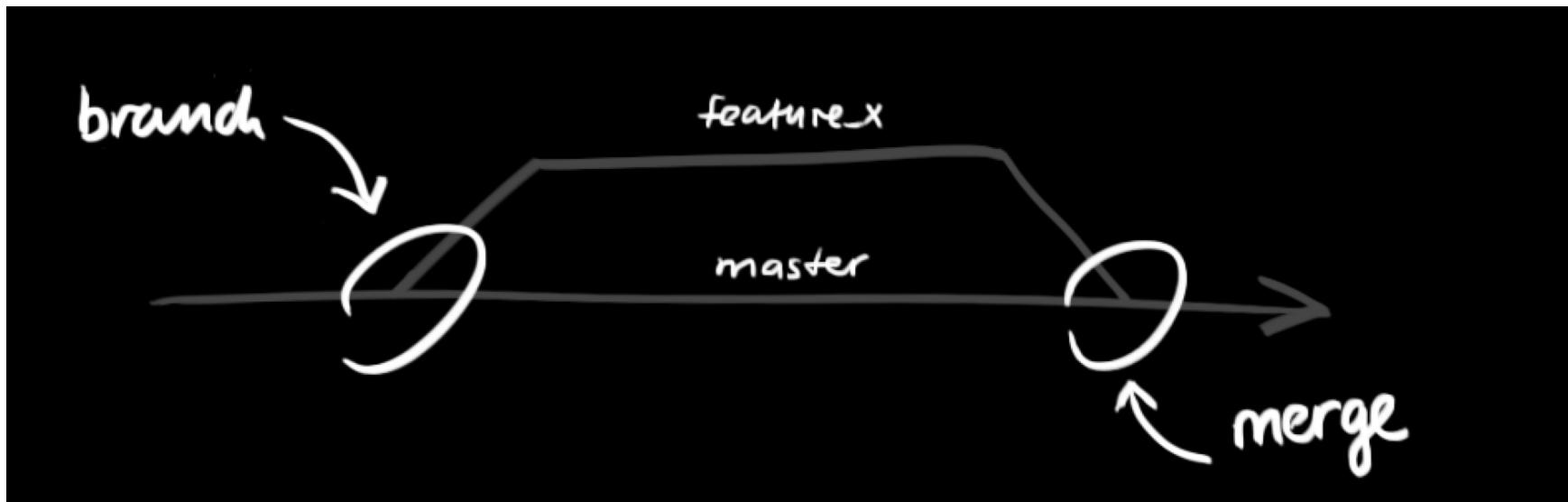
git

- git pull
 - Get new updates from git server

```
(base) jarvis@Kelechis-MBP CSC-DataScienceExample % git pull
remote: Enumerating objects: 4, done.
remote: Counting objects: 100% (4/4), done.
remote: Compressing objects: 100% (3/3), done.
remote: Total 3 (delta 0), reused 0 (delta 0), pack-reused 0
Unpacking objects: 100% (3/3), 746 bytes | 373.00 KiB/s, done.
From https://github.com/ikegwukc/CSC-Test
      f1e78cc..5d5a4ea  master      -> example/master
Updating f1e78cc..5d5a4ea
Fast-forward
  FileFromAnotherUsr.md | 3 +++
  1 file changed, 3 insertions(+)
  create mode 100644 FileFromAnotherUsr.md
(base) jarvis@Kelechis-MBP CSC-DataScienceExample %
```

git

- git branch
 - <https://www.atlassian.com/git/tutorials/using-branches>



git

- git branch

```
(base) jarvis@Kelechis-MBP CSC-DataScienceExample % git branch  
* master
```

- Create a new branch

```
(base) jarvis@Kelechis-MBP CSC-DataScienceExample % git checkout -b 'feature_x'  
Switched to a new branch 'feature_x'  
(base) jarvis@Kelechis-MBP CSC-DataScienceExample % git branch  
* feature_x  
  master
```

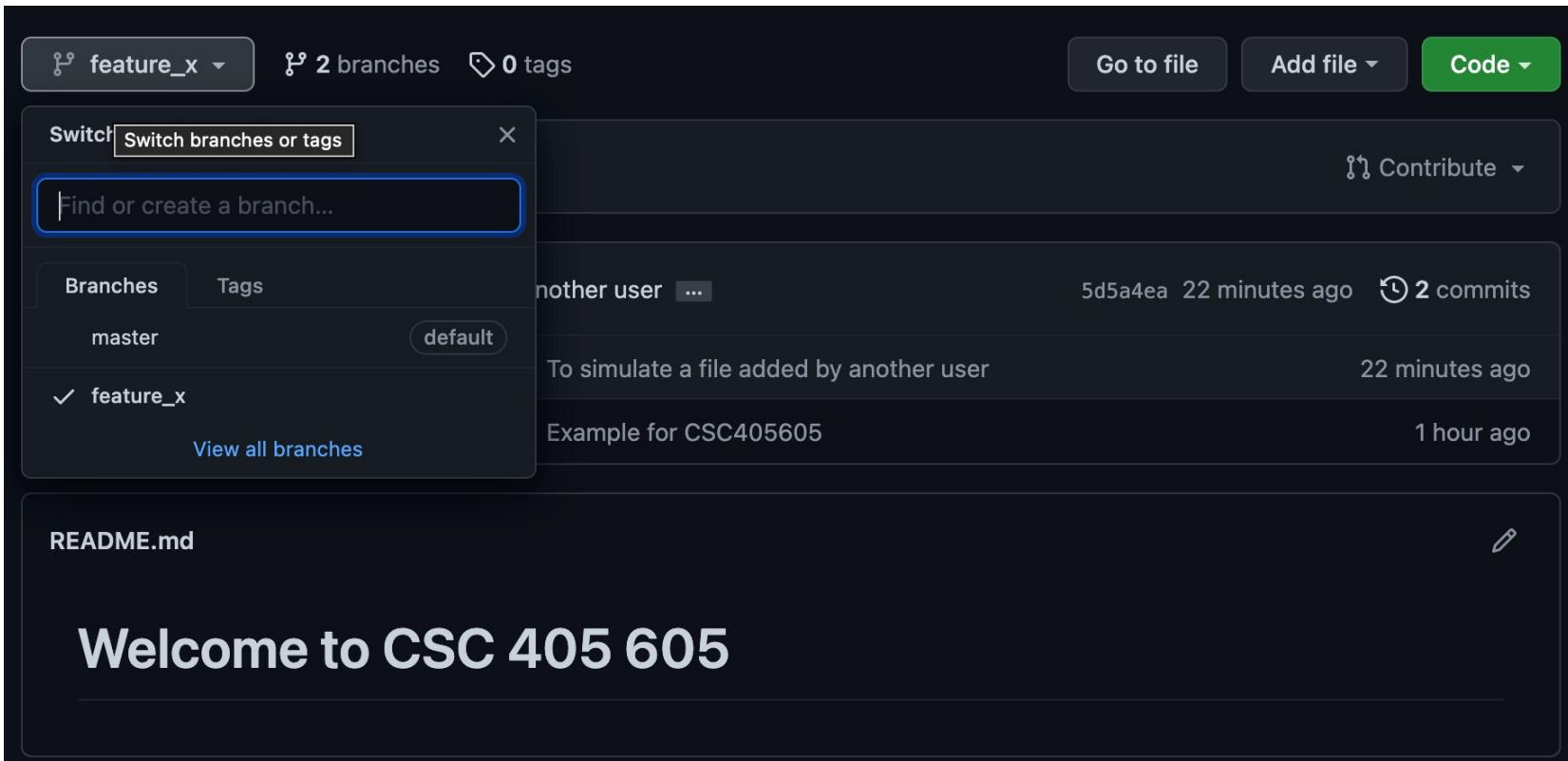
git

- git push feature_x branch

```
(base) jarvis@Kelechis-MBP CSC-DataScienceExample % git push example 'feature_x'
Total 0 (delta 0), reused 0 (delta 0), pack-reused 0
remote:
remote: Create a pull request for 'feature_x' on GitHub by visiting:
remote:     https://github.com/ikegwukc/CSC-Test/pull/new/feature_x
remote:
To https://github.com/ikegwukc/CSC-Test
 * [new branch]      feature_x -> feature_x
(base) jarvis@Kelechis-MBP CSC-DataScienceExample % █
```

git

- Feature_x branch on github



git

- Add and commit new file to `feature_x` branch

```
(base) jarvis@Kelechis-MBP CSC-DataScienceExample % git branch
* feature_x
  master
(base) jarvis@Kelechis-MBP CSC-DataScienceExample % touch FeatureX.md
(base) jarvis@Kelechis-MBP CSC-DataScienceExample % echo "Life Changing Feature" < FeatureX.md
Life Changing Feature
(base) jarvis@Kelechis-MBP CSC-DataScienceExample % git add FeatureX.md
(base) jarvis@Kelechis-MBP CSC-DataScienceExample % git commit -m "FeatureX file"
[feature_x f60948e] FeatureX file
 1 file changed, 0 insertions(+), 0 deletions(-)
 create mode 100644 FeatureX.md
```

git

- Push to feature_x branch on GitHub

```
(base) jarvis@Kelechis-MBP CSC-DataScienceExample % git push example 'feature_x'  
Enumerating objects: 4, done.  
Counting objects: 100% (4/4), done.  
Delta compression using up to 16 threads  
Compressing objects: 100% (2/2), done.  
Writing objects: 100% (3/3), 326 bytes | 326.00 KiB/s, done.  
Total 3 (delta 0), reused 0 (delta 0), pack-reused 0  
To https://github.com/ikegwukc/CSC-Test  
 5d5a4ea..f60948e feature_x -> feature_x
```

git

- Check Status of Branches on GitHub

Master Branch

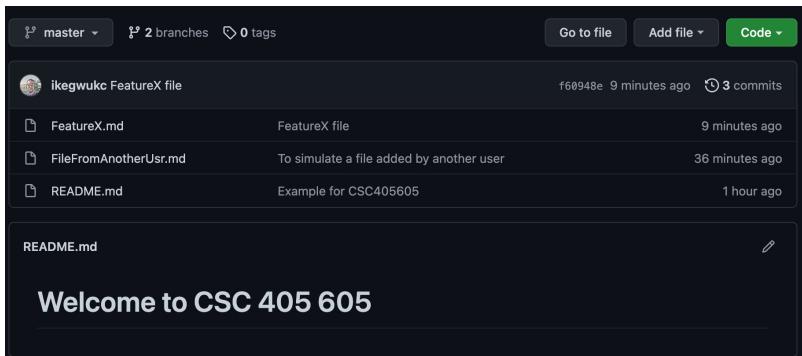
The screenshot shows the GitHub repository page for the 'master' branch. At the top, a message says "feature_x had recent pushes 3 minutes ago". Below it, there are buttons for "Compare & pull request", "Go to file", "Add file", and "Code". The branch dropdown shows "master". There are 2 branches and 0 tags. A commit by user "ikegwukc" is listed, adding a file "FileFromAnotherUsr.md" to simulate a file added by another user. Another commit by "ikegwukc" adds a "README.md" file for CSC405605. The repository summary at the bottom says "Welcome to CSC 405 605".

feature_x Branch

The screenshot shows the GitHub repository page for the "feature_x" branch. At the top, a message says "feature_x had recent pushes 2 minutes ago". Below it, there are buttons for "Compare & pull request", "Go to file", "Add file", and "Code". The branch dropdown shows "feature_x". There are 2 branches and 0 tags. A commit by user "ikegwukc" is listed, adding a file "FeatureX.md" to simulate a file added by another user. Another commit by "ikegwukc" adds a "README.md" file for CSC405605. The repository summary at the bottom says "Welcome to CSC 405 605".

git

- Switch back to master branch
- Merge feature_x
- Push to GitHub



A screenshot of a GitHub repository interface. The repository name is 'CSC-DataScienceExample'. The 'master' branch is selected. There are 2 branches and 0 tags. The commit history shows:

- iKEGWUKC FeatureX file (f60948e, 9 minutes ago)
- FeatureX.md (9 minutes ago)
- FileFromAnotherUsr.md (To simulate a file added by another user, 36 minutes ago)
- README.md (Example for CSC405605, 1 hour ago)

The README file content is: Welcome to CSC 405 605.

```
(base) jarvis@Kelechis-MBP CSC-DataScienceExample % git checkout master
Switched to branch 'master'
Your branch is up to date with 'example/master'.
(base) jarvis@Kelechis-MBP CSC-DataScienceExample % git merge feature_x
Updating 5d5a4ea..f60948e
Fast-forward
  FeatureX.md | 0
  1 file changed, 0 insertions(+), 0 deletions(-)
  create mode 100644 FeatureX.md
(base) jarvis@Kelechis-MBP CSC-DataScienceExample % git push example master
Total 0 (delta 0), reused 0 (delta 0), pack-reused 0
To https://github.com/ikegwukc/CSC-Test
  5d5a4ea..f60948e  master -> master
(base) jarvis@Kelechis-MBP CSC-DataScienceExample %
```

git

- git clone
 - Create a working copy from remote server
- GitHub Explore Page:
 - <https://github.com/explore>

git

- git clone
 - Create a working copy from remote server
 - Tensorflow
 - <https://github.com/tensorflow/tensorflow>

```
(base) jarvis@Kelechis-MBP CSC-DataScienceExample % cd ..  
(base) jarvis@Kelechis-MBP ~ % git clone https://github.com/tensorflow/tensorflow.git  
Cloning into 'tensorflow'...  
remote: Enumerating objects: 1285201, done.  
remote: Counting objects: 100% (18/18), done.  
remote: Compressing objects: 100% (18/18), done.  
remote: Total 1285201 (delta 0), reused 17 (delta 0), pack-reused 1285183  
Receiving objects: 100% (1285201/1285201), 806.84 MiB | 13.80 MiB/s, done.  
Resolving deltas: 100% (1044104/1044104), done.  
Updating files: 100% (25319/25319), done.  
(base) jarvis@Kelechis-MBP ~ % █
```

git

A Simple, Effective Git Workflow

