UCL Data Science Society - Term 2 Workshop 6: Introduction to *Git/Github*

Created by Zhaoxuan "Tony" Wu, First-Year Representative of Science Divison, *UCL Data Science Society*

[See me on *Github*💁](#)

Table of Contents

Prerequisition

Git

You should download *Git*. Feel free to check the *official documentation* if you are interested in. We will be using the *Terminal*, so it might be helpful to get familiar with the commands.

Recommended Terminal for MacOS: *iTerm2*

Recommended multi-lang IDE: *Atom*, *Visual Studio*, *Sublime Text*, or *Vim*, which is pre-installed on MacOS

Github

Get yourself a *Github* account, it's free! If you prefer graphical interface, *Github Desktop*. But we are going to learn *Git* and *Github* using command line today. Also, remember to claim your *Github Student Package*📦

What is *Git*, and Why?

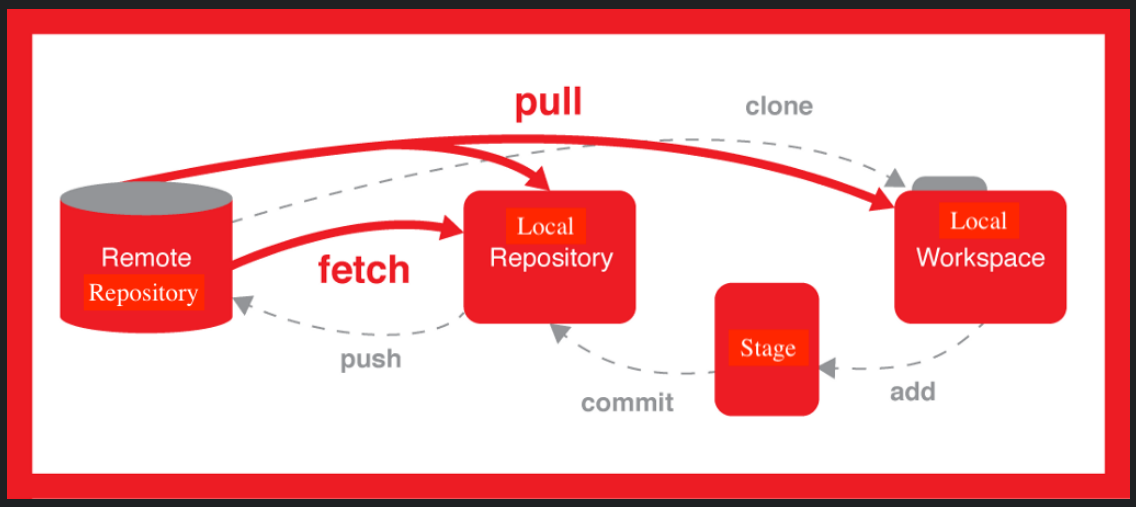
> *Git is a free and open source* ***distributed version control system*** *designed to handle everything from small to very large projects with speed and efficiency.*

- Industrial production standard
- Hackathon
- Writing a book
- Keeping lecture nots
- ...

How is it different from *Github*?

Git is a ***system*** that manages the version control of a project, while *Github* is a ***remote platform*** that hosts that project using *Git*. For instance: `git push` uploads your current local repository to *Github*

Relationship



> ***Repository*** ("Repo") : a receptacle or place where things are deposited or stored

Your first *Git* repo

Do the following on your own *Terminal*

Initialisation of project

Create a *directory* (folder) for your project:

```
mkdir myapp
```

Go to that directory:

```
cd myapp
```

Initialise the project with *Git* so that *Git* manages the version control of the project:

```
git init
```

Now you should have a `.git` folder in your directory, which is invisible currently. Also, your current project is called the *local workspace*.

Make changes

Now, create an `app.py` *Python* file in this folder and write a line of code that outputs `"Welcome to DSS Workshop 6"`

```
key = ""

def getKey():
     f = open("key.txt","r")
     key = f.read()
     f.close()

def login(account,password, ):
     if account ==
```

You can do it with your IDE or with command line. The method for doing that using command line is as the following:

Create `app.py` :

```
touch app.py
```

Edit the file using the built-in editor *Vim*, or use IDE:

```
vim app.py
```

```
key = ""

def getKey():
     f = open("key.txt","r")
     key = f.read()
     f.close()

def login(password):
     if account == key:
          print("successful")
     else:
          print("denied")
```

Save the file and quit:

```
:wq
```

Stage the changes

Before you commit your changes, you should *add* your changes to the stage, or "*stage*" it.

```
git add .
```

> ***Note that***: `git add` takes in a parameter, which is the filename. For example, you can do `git add app.py` to stage `app.py`, `git add .` is the wildcard mode that stages ***every*** file that has been changed

To check your *stage*:

```
1  git status
```

Commit the changes

Now, commit your changes:

```
1  git commit -m "first commit"
```

> ***Note that***: `git commit` takes in commit comment using the `-m` flag, followed by your comment string. It can be anything. Here we use `"first commit"` as example, which is usually what you do for your first commit literally

Congrats! 🥳 You just commited your first contribution to the project! Now this version of commit is officially in your *local repository* (*local repo*).

Create another file

Now create another file `key.txt` with whichever method you like and write the following line:

```
1  "ucldssistheBEST:)"
```

> Don't stage it yet

Security, security, security, and `.gitignore`

You might notice that it is quite dangerous to commit a copy of your `key.txt` to a *repo* or a *remote repo*.

- API keys
- Database password
- Credentials
- Biometrics data
- Data under NDA
- `.DS_Store`
- `/node_modules`
- ...

By using `.gitignore`, you can prevent certain files to be commited to a *repo*

```
1  touch .gitignore
```

Directly add the name of the files you want to hide to that `.gitignore` file:

```
vim .gitignore

.DS_Store
key.txt
```

Save and quit:

```
:wq
```

Now, *stage* and *commit* everything and see what happen.

And here comes a real story...

Remote repo and *Github*

Now you might want to share your code with other developer, you can do this by putting your project on a remote repo. Do this by call the following function:

```
seeTonyForLiveDemo()
```

Branch, and uploading changes to *remote repo*

For instance, you want to create an *HTML* for your app. Create and switch to a new branch called `html`:

```
git checkout -b html
```

> ***Note that***: `-b` flag is for creating a new branch. If you want to see all available branches, use `git branch`, if you want to switch to an existing branch, use `git checkout <branch>` where `<branch>` is the branch name

Craete your HTML:

```
<html>
     <head>
          <title>UCL DSS</title>
     </head>
     <body>
          <h1> My heading</h1>
          <p> Your first programme/page for any languages should always be: Hello World! </p>
     </body>
 </html>
```

Stage it, check the *stage*, *commit* it. Now, upload it:

```
git push -u origin html
```

Go to your *Github* repo to see what happened.

Collaborative Coding: 101

If that's not your own project:

Obtain the repo

- Fork an existing project
- Clone the *your remote repo* to local

Make changes

- Make changes to *local repo*
- Add *remote forked repo* to `remote` :

```
git remote -v
git remote add upstream <upstream_url>
```

Push to remote

- `push` to *your remote repo*
- Make a pull request

Sync with forked repo

EITHER

- Keep your local *repo* synced with the *remote forked repo* and push to *your remote repo*

```
git fetch upstream
git merge upstream/master
git push -u origin upstream
```

OR

- Sync *your remote repo* to *remote forked repo* on *Github*
- `pull` from *your remote_repo to keep your _local repo* synced

If that's your own project:

- Manage pull request
- Review changes, make comments, reject or `merge` pull request

Practice: Collaboration in Pairs

Find a partner. designate one as the *project manager* and the other as the *contributor*.

Project Manager

- Initialise a *local repo* and a *remote repo*
- Review your peer's pull requests

Contributor

- Fork your project manager's *remote repo*
- `clone` and make changes locally

- Make a pull request

Practice: Signing an attendance sheet

Let's sign an attendance sheet collaboratively!

The main repo for you to work on

```
signAttendanceSheet(dssWorkshopAttendee)
```

To Wrap Up

Git/Github is massive, I haven't figure out all of it as well. This is a brief introduction to the tip of this iceberg.

Official Documentation 📄

Thank you all for joining our journey to *data science*, *machine learning*,*neural nets*, *Python programming* and *Git/Github*. See you next week and hopefully ***next academic year*** as well!

Checkout my Mathematics for Machine Learning notes 🤖

Some of my past hackathon projects you might find inspiring 💡

Most importantly:

Stay tune for everything about data science

Subscribe to our offical IG if you haven't do so ☑️

And our FB! 🍾

Next Week

What: ***Introduction to Object-Orientated Programming (OOP)***

When: Monday, 16 Mar

Who: *Shirui "Eric" Lyu*

Where: 20 Bedford Way (IOE), w2.05