



INTRODUCTION TO DATA SCIENCE

TA class I – Git & GitHub

TA : Lee Chi-Hsuan



Git & GitHub






VCS



Version Control System


A system that records changes to files over time so that you can recall specific versions later.


Git


 **git** --local-branching-on-the-cheap


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.


Git is **easy to learn** and has a **tiny footprint with lightning fast performance**. It outclasses SCM tools like Subversion, CVS, Perforce, and ClearCase with features like **cheap local branching**, convenient staging areas, and **multiple workflows**.




**About**
The advantages of Git compared to other source control systems.

**Documentation**
Command reference pages, Pro Git book content, videos and other material.

**Downloads**
GUI clients and binary releases for all major platforms.

**Community**
Get involved! Bug reporting, mailing list, chat, development and more.

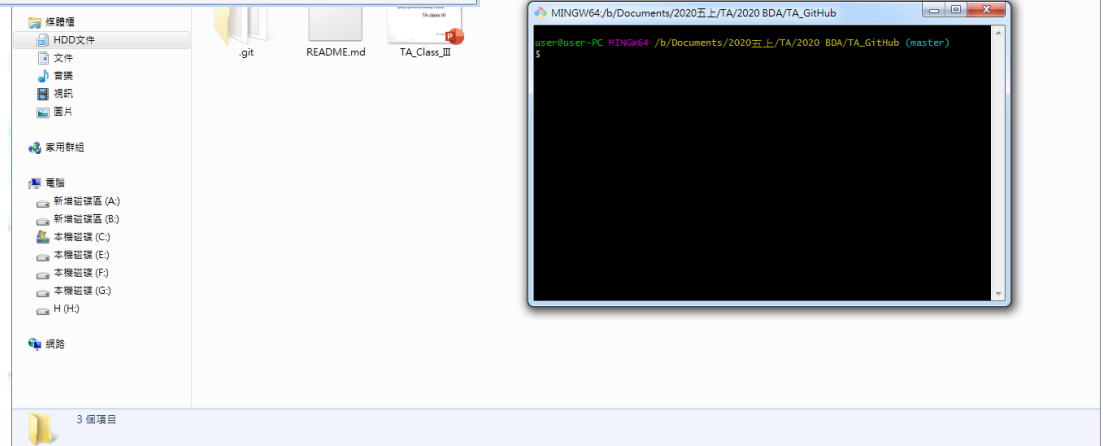


Latest source Release
2.28.0
[Release Notes \(2020-07-27\)](#)
[Download 2.28.0 for Windows](#)

TRIVIAL



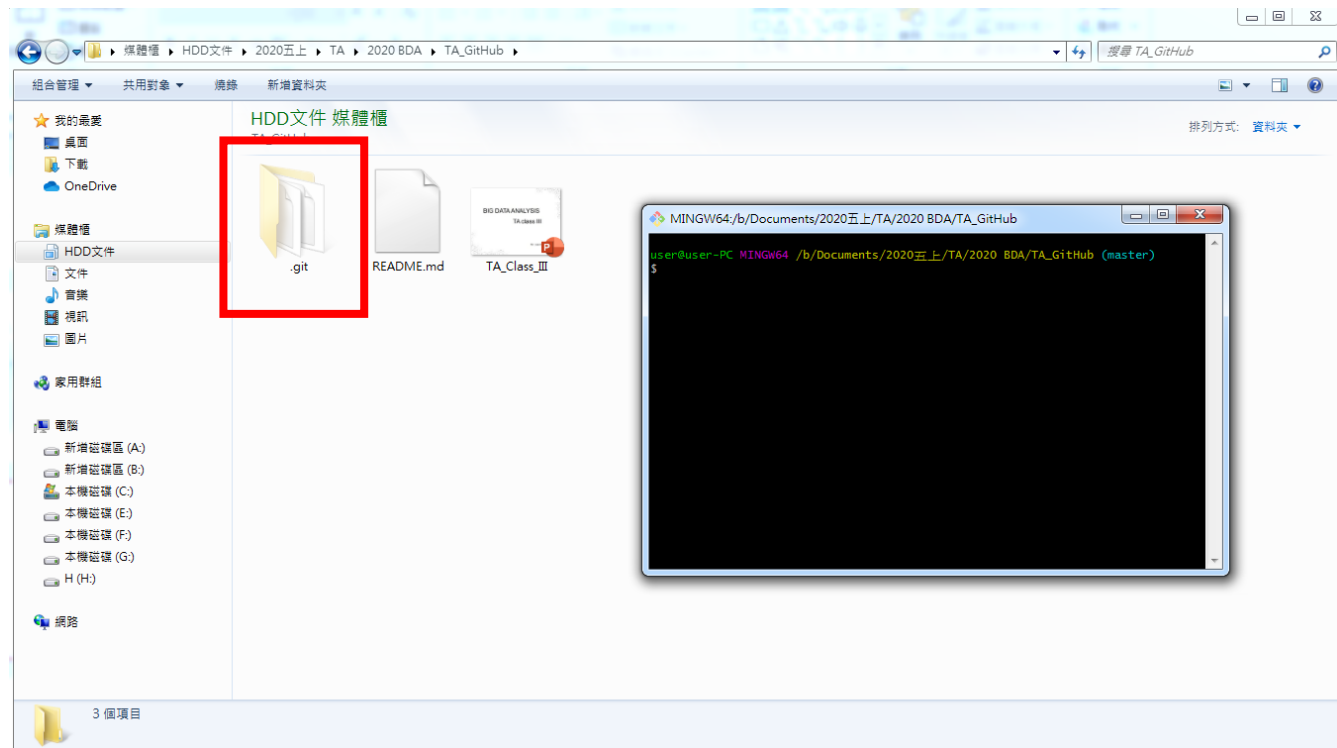
First Steps with Git



Your Identity

- `git config --global user.email "your_email_account.gmail.com"`
- `git config --global user.name "your name"`

- git init



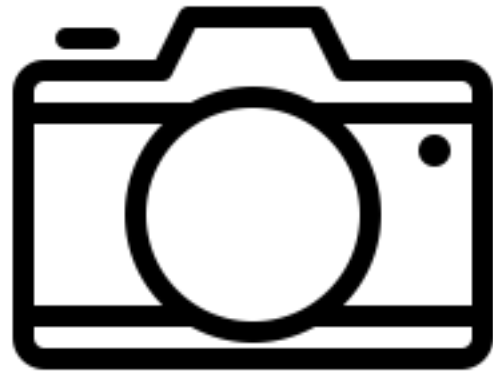
Three state of tracking files :

- modified

- staged

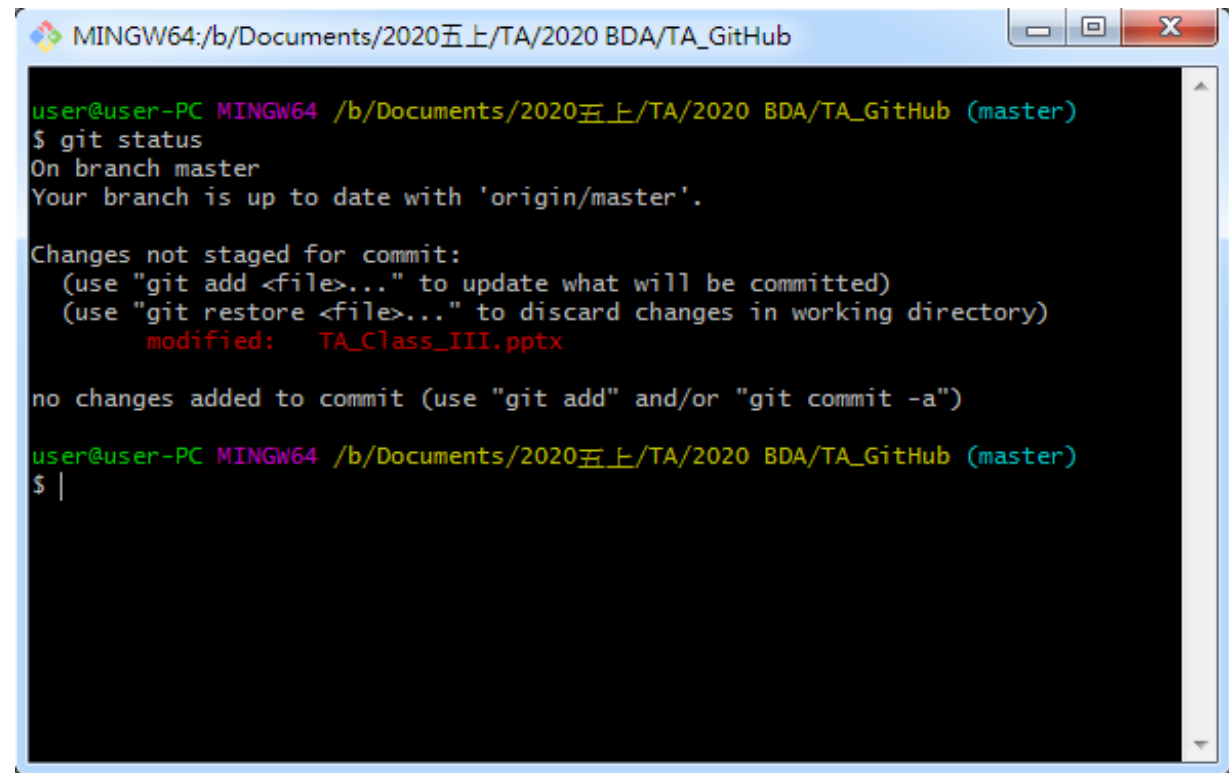
- committed

} state



- Snapshotting your repository

- git status



```
MINGW64:/b/Documents/2020五上/TA/2020 BDA/TA_GitHub

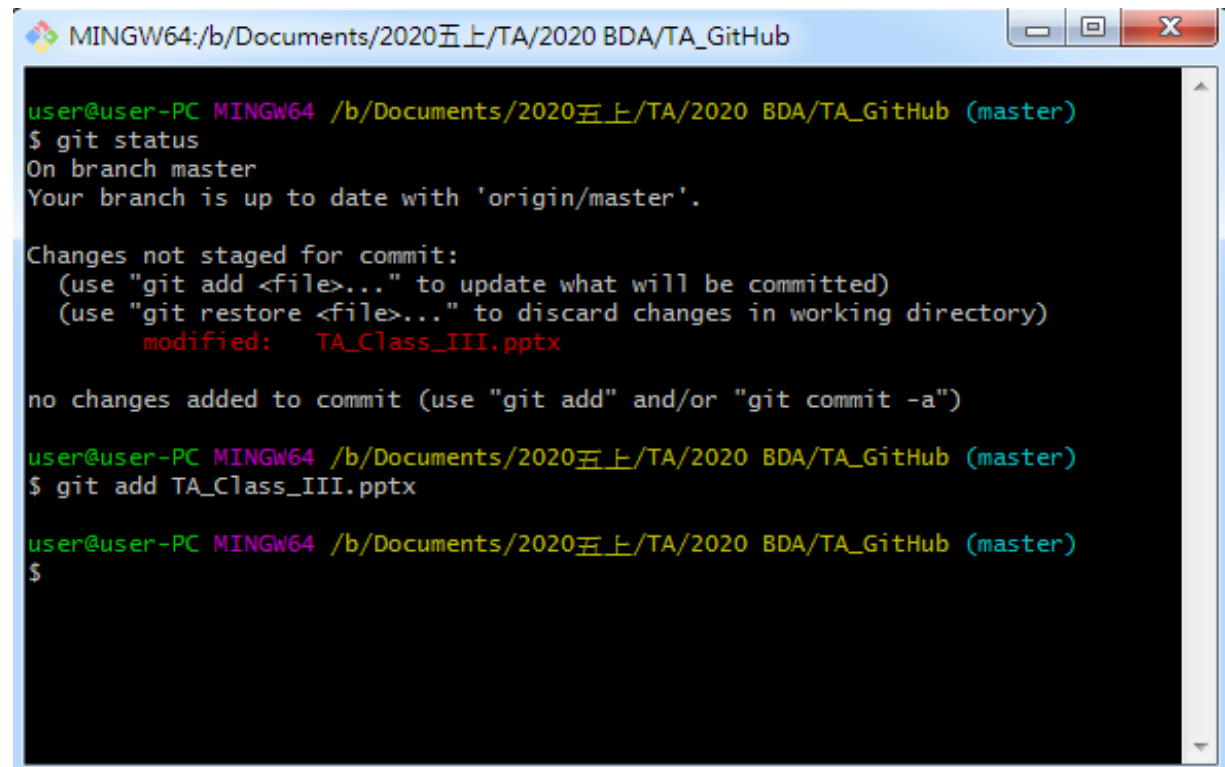
user@user-PC MINGW64 /b/Documents/2020五上/TA/2020 BDA/TA_GitHub (master)
$ git status
On branch master
Your branch is up to date with 'origin/master'.

Changes not staged for commit:
  (use "git add <file>..." to update what will be committed)
  (use "git restore <file>..." to discard changes in working directory)
        modified:   TA_Class_III.pptx

no changes added to commit (use "git add" and/or "git commit -a")

user@user-PC MINGW64 /b/Documents/2020五上/TA/2020 BDA/TA_GitHub (master)
$ |
```

- git add



```
MINGW64:/b/Documents/2020五上/TA/2020 BDA/TA_GitHub

user@user-PC MINGW64 /b/Documents/2020五上/TA/2020 BDA/TA_GitHub (master)
$ git status
On branch master
Your branch is up to date with 'origin/master'.

Changes not staged for commit:
  (use "git add <file>..." to update what will be committed)
  (use "git restore <file>..." to discard changes in working directory)
        modified:   TA_Class_III.pptx

no changes added to commit (use "git add" and/or "git commit -a")

user@user-PC MINGW64 /b/Documents/2020五上/TA/2020 BDA/TA_GitHub (master)
$ git add TA_Class_III.pptx

user@user-PC MINGW64 /b/Documents/2020五上/TA/2020 BDA/TA_GitHub (master)
$
```

- `git commit -m '...'`

```
MINGW64:/b/Documents/2020五上/TA/2020 BDA/TA_GitHub

user@user-PC MINGW64 /b/Documents/2020五上/TA/2020 BDA/TA_GitHub (master)
$ git status
On branch master
Your branch is up to date with 'origin/master'.

Changes not staged for commit:
  (use "git add <file>..." to update what will be committed)
  (use "git restore <file>..." to discard changes in working directory)
        modified:   TA_Class_III.pptx

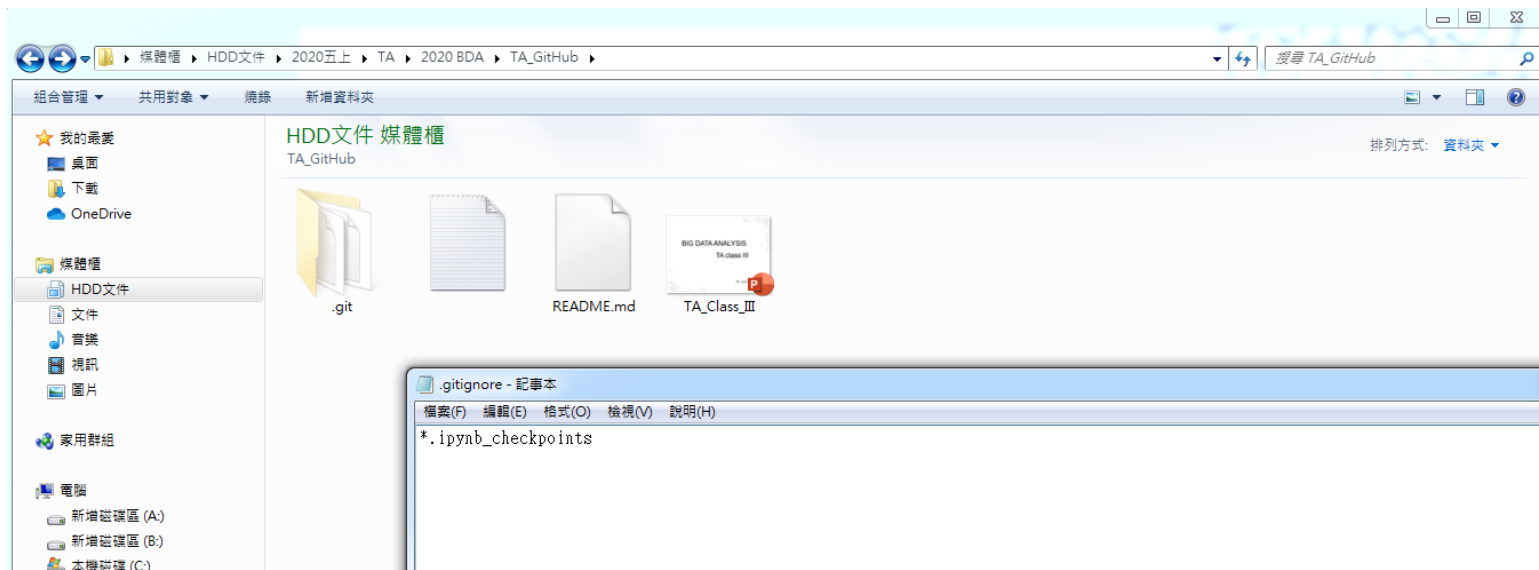
no changes added to commit (use "git add" and/or "git commit -a")

user@user-PC MINGW64 /b/Documents/2020五上/TA/2020 BDA/TA_GitHub (master)
$ git add TA_Class_III.pptx

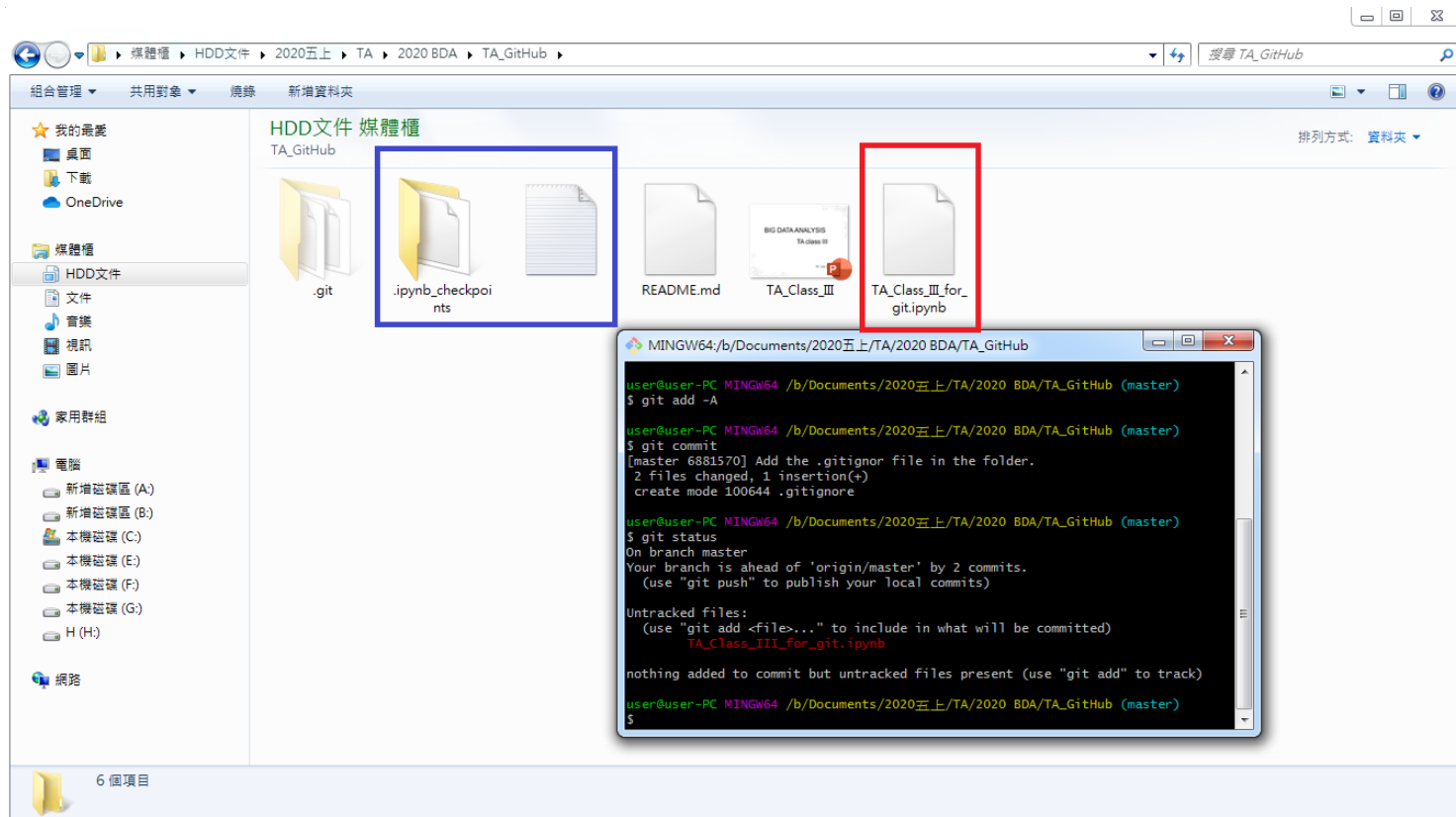
user@user-PC MINGW64 /b/Documents/2020五上/TA/2020 BDA/TA_GitHub (master)
$ git commit -m 'Modify the PPT'
[master 283eeeb] Modify the PPT
 1 file changed, 0 insertions(+), 0 deletions(-)

user@user-PC MINGW64 /b/Documents/2020五上/TA/2020 BDA/TA_GitHub (master)
$
```

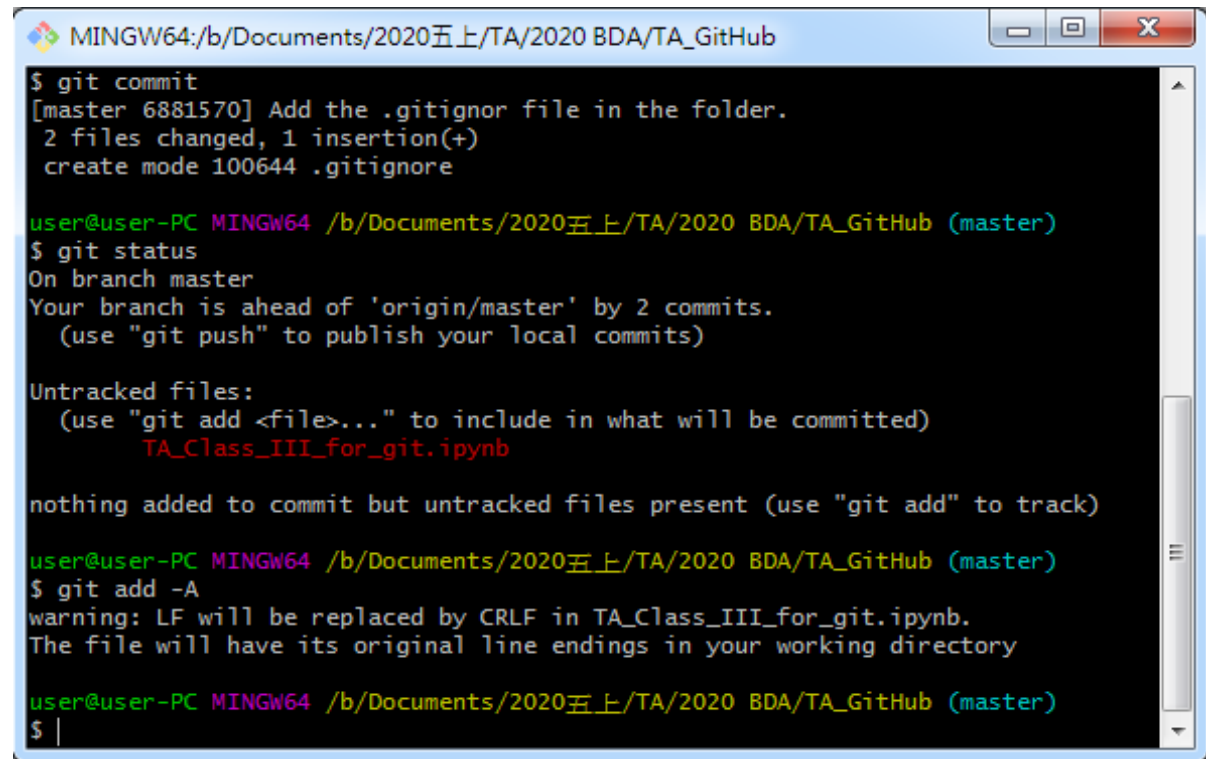
.gitignor



...after adding an ipynb file
(the ipynb_checkpoints is ignored)



- git add -A



```
MINGW64:/b/Documents/2020五上/TA/2020 BDA/TA_GitHub
$ git commit
[master 6881570] Add the .gitignor file in the folder.
 2 files changed, 1 insertion(+)
 create mode 100644 .gitignore

user@user-PC MINGW64 /b/Documents/2020五上/TA/2020 BDA/TA_GitHub (master)
$ git status
On branch master
Your branch is ahead of 'origin/master' by 2 commits.
  (use "git push" to publish your local commits)

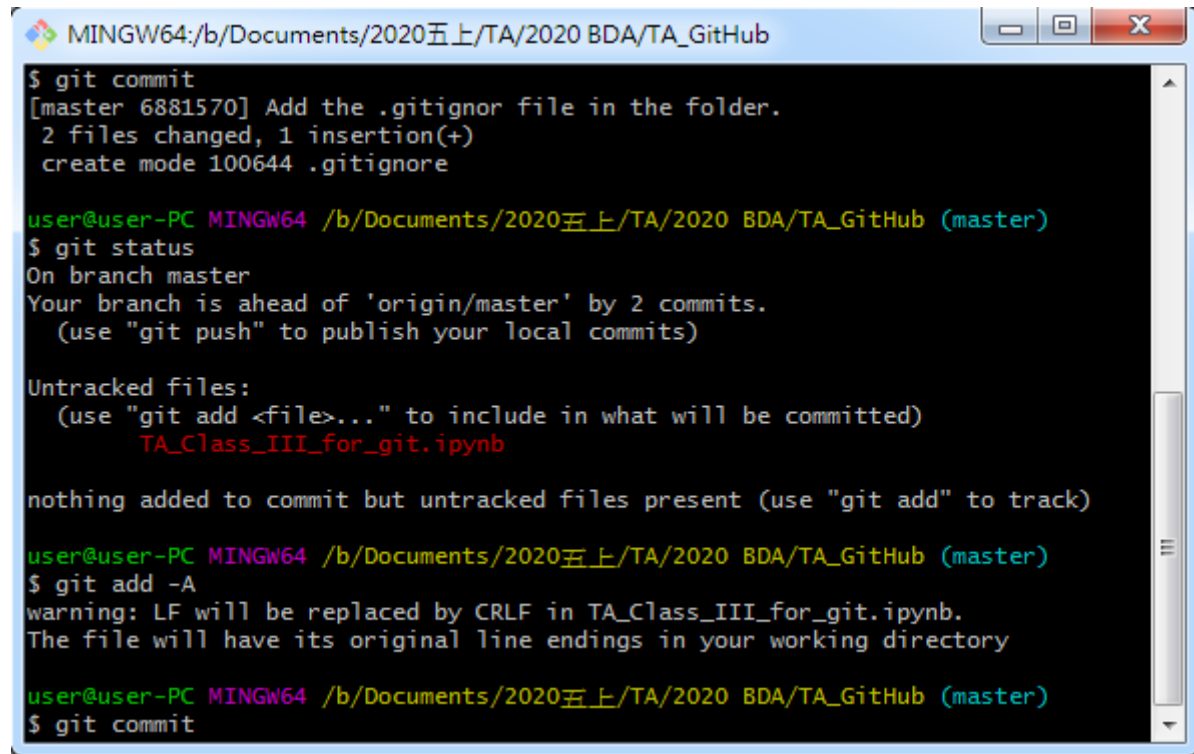
Untracked files:
  (use "git add <file>..." to include in what will be committed)
       TA_Class_III_for_git.ipynb

nothing added to commit but untracked files present (use "git add" to track)

user@user-PC MINGW64 /b/Documents/2020五上/TA/2020 BDA/TA_GitHub (master)
$ git add -A
warning: LF will be replaced by CRLF in TA_Class_III_for_git.ipynb.
The file will have its original line endings in your working directory

user@user-PC MINGW64 /b/Documents/2020五上/TA/2020 BDA/TA_GitHub (master)
$ |
```

- git commit



```
MINGW64:/b/Documents/2020五上/TA/2020 BDA/TA_GitHub
$ git commit
[master 6881570] Add the .gitignore file in the folder.
 2 files changed, 1 insertion(+)
 create mode 100644 .gitignore

user@user-PC MINGW64 /b/Documents/2020五上/TA/2020 BDA/TA_GitHub (master)
$ git status
On branch master
Your branch is ahead of 'origin/master' by 2 commits.
  (use "git push" to publish your local commits)

Untracked files:
  (use "git add <file>..." to include in what will be committed)
    TA_Class_III_for_git.ipynb

nothing added to commit but untracked files present (use "git add" to track)

user@user-PC MINGW64 /b/Documents/2020五上/TA/2020 BDA/TA_GitHub (master)
$ git add -A
warning: LF will be replaced by CRLF in TA_Class_III_for_git.ipynb.
The file will have its original line endings in your working directory

user@user-PC MINGW64 /b/Documents/2020五上/TA/2020 BDA/TA_GitHub (master)
$ git commit
```

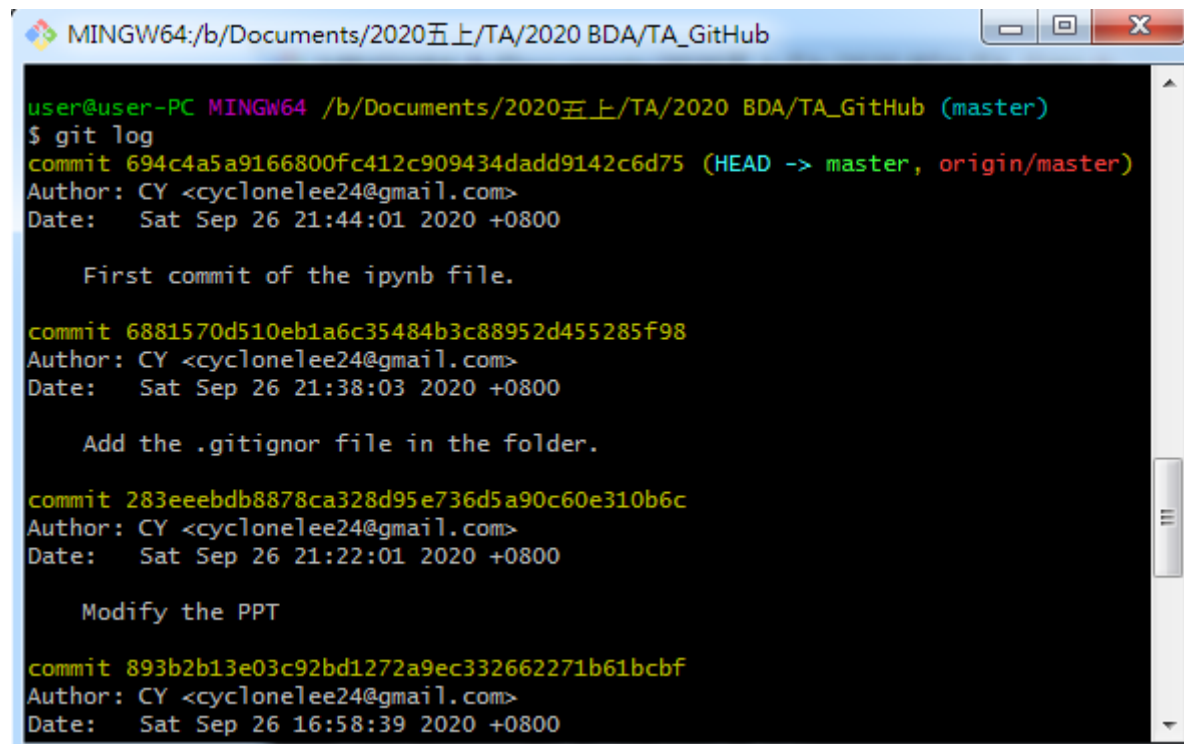



- vim

- press 'i'
- write your commit message
- press 'Esc'
- type ':wq'

[illegible]

- git log



```
MINGW64:/b/Documents/2020五上/TA/2020 BDA/TA_GitHub

user@user-PC MINGW64 /b/Documents/2020五上/TA/2020 BDA/TA_GitHub (master)
$ git log
commit 694c4a5a9166800fc412c909434dadd9142c6d75 (HEAD -> master, origin/master)
Author: CY <cyclonelee24@gmail.com>
Date: Sat Sep 26 21:44:01 2020 +0800

    First commit of the ipynb file.

commit 6881570d510eb1a6c35484b3c88952d455285f98
Author: CY <cyclonelee24@gmail.com>
Date: Sat Sep 26 21:38:03 2020 +0800

    Add the .gitignor file in the folder.

commit 283eeebdb8878ca328d95e736d5a90c60e310b6c
Author: CY <cyclonelee24@gmail.com>
Date: Sat Sep 26 21:22:01 2020 +0800

    Modify the PPT

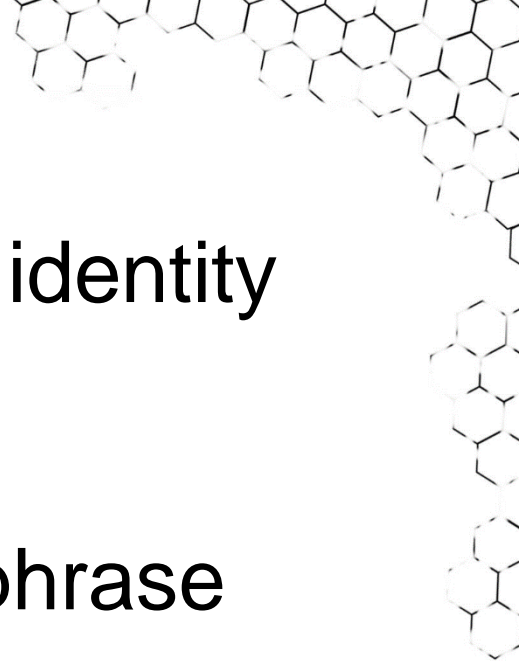
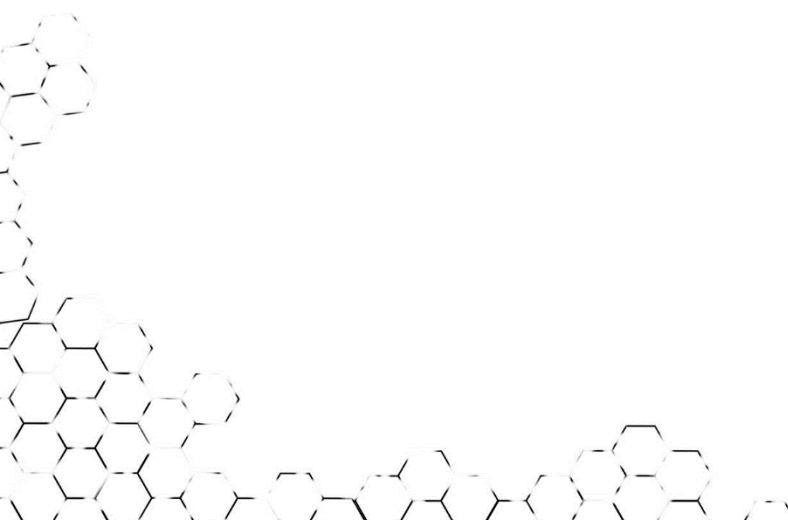
commit 893b2b13e03c92bd1272a9ec332662271b61bcbf
Author: CY <cyclonelee24@gmail.com>
Date: Sat Sep 26 16:58:39 2020 +0800
```

GitHub

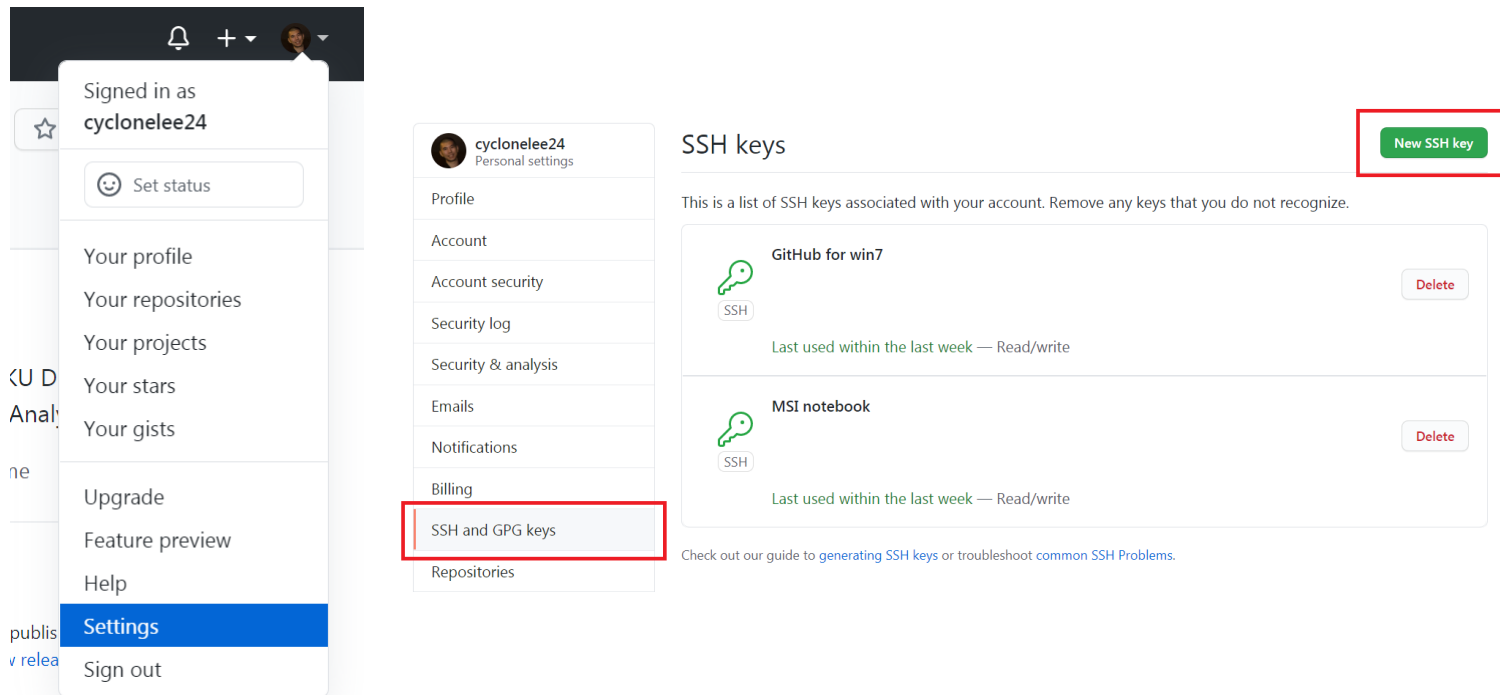


- A web-based Git repository hosting service.
- On top of Git's functions, it includes extra features.



- 
- 
- ssh-key : let GitHub know your identity
 - `$ ssh-keygen`
 - you can skip entering the passphrase
 - `$ cat /directory of your key/id_rsa.pub`

- copy from ssh-rsa... to ...@gmail.com
(or to ...@pc-name)



- Identify yourself with id_rsa.pub to GitHub

[SSH keys](#) / Add new

Title

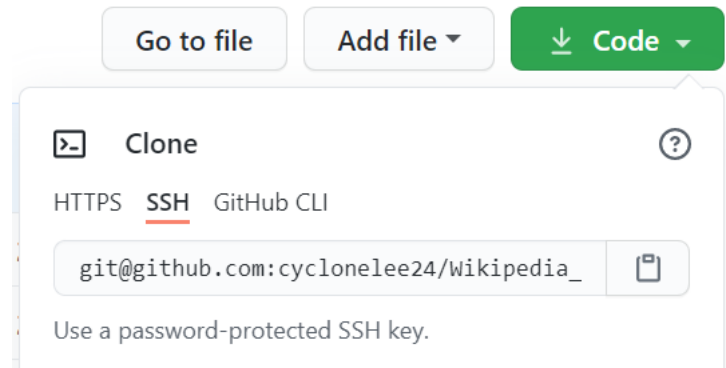
Key

Begins with 'ssh-rsa', 'ssh-ed25519', 'ecdsa-sha2-nistp256', 'ecdsa-sha2-nistp384', or 'ecdsa-sha2-nistp521'

Add SSH key

- 
- 
- Adding SSH key steps reference

- git clone



文件 > 2021_Spring > TA_2021 >				搜尋 TA_2021
名稱	修改日期	類型	大小	
HW_0	2021/2/22 下午 12:18	檔案資料夾		
Lecture_0	2021/2/22 下午 02:10	檔案資料夾		
TA_Git_GitHub	2021/3/3 下午 04:58	檔案資料夾		
Wikipedia_Wikidata_Alignment	2021/3/3 下午 05:09	檔案資料夾		

```
MINGW64:/c/Users/cyclo/Documents/2021_Spring/TA_2021

cyclo@MSI MINGW64 ~/Documents/2021_Spring/TA_2021
$ git clone git@github.com:cyclonelee24/Wikipedia_Wikidata_Alignment.git
Cloning into 'Wikipedia_Wikidata_Alignment'...
remote: Enumerating objects: 52, done.
remote: Counting objects: 100% (52/52), done.
remote: Compressing objects: 100% (34/34), done.
remote: Total 52 (delta 25), reused 38 (delta 14), pack-reused 0
Receiving objects: 100% (52/52), 30.32 MiB | 2.27 MiB/s, done.
Resolving deltas: 100% (25/25), done.

cyclo@MSI MINGW64 ~/Documents/2021_Spring/TA_2021
$ |
```

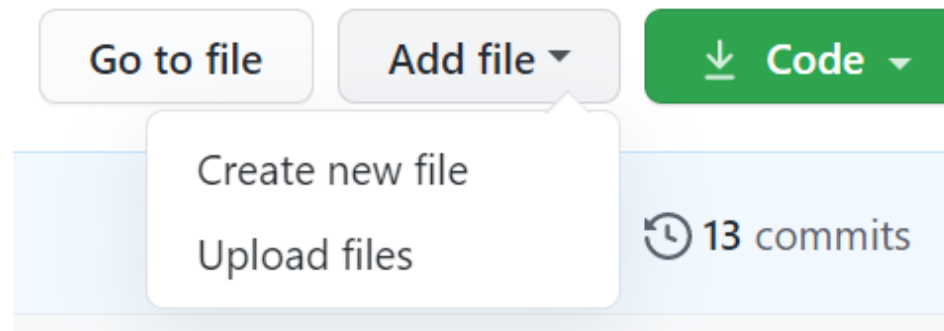
TA_2021 > Wikipedia_Wikidata_Alignment >



搜尋 Wikipedia_Wikidata_Alignment

名稱	修改日期	類型	大小
.git	2021/3/3 下午 05:10	檔案資料夾	
data	2021/3/3 下午 05:10	檔案資料夾	
.gitignore	2021/3/3 下午 05:09	文字文件	1 KB
README.md	2021/3/3 下午 05:09	MD 檔案	1 KB
Wikidata_Coattention_NSMN.ipynb	2021/3/3 下午 05:09	IPYNB 檔案	46 KB
Wikidata_pretrained_BERT.ipynb	2021/3/3 下午 05:09	IPYNB 檔案	51 KB
Wikidata_pretrained_BERT_unlabel.ipynb	2021/3/3 下午 05:09	IPYNB 檔案	42 KB

- other method to upload your homework





Drag files here to add them to your repository

Or [choose your files](#)



Commit changes

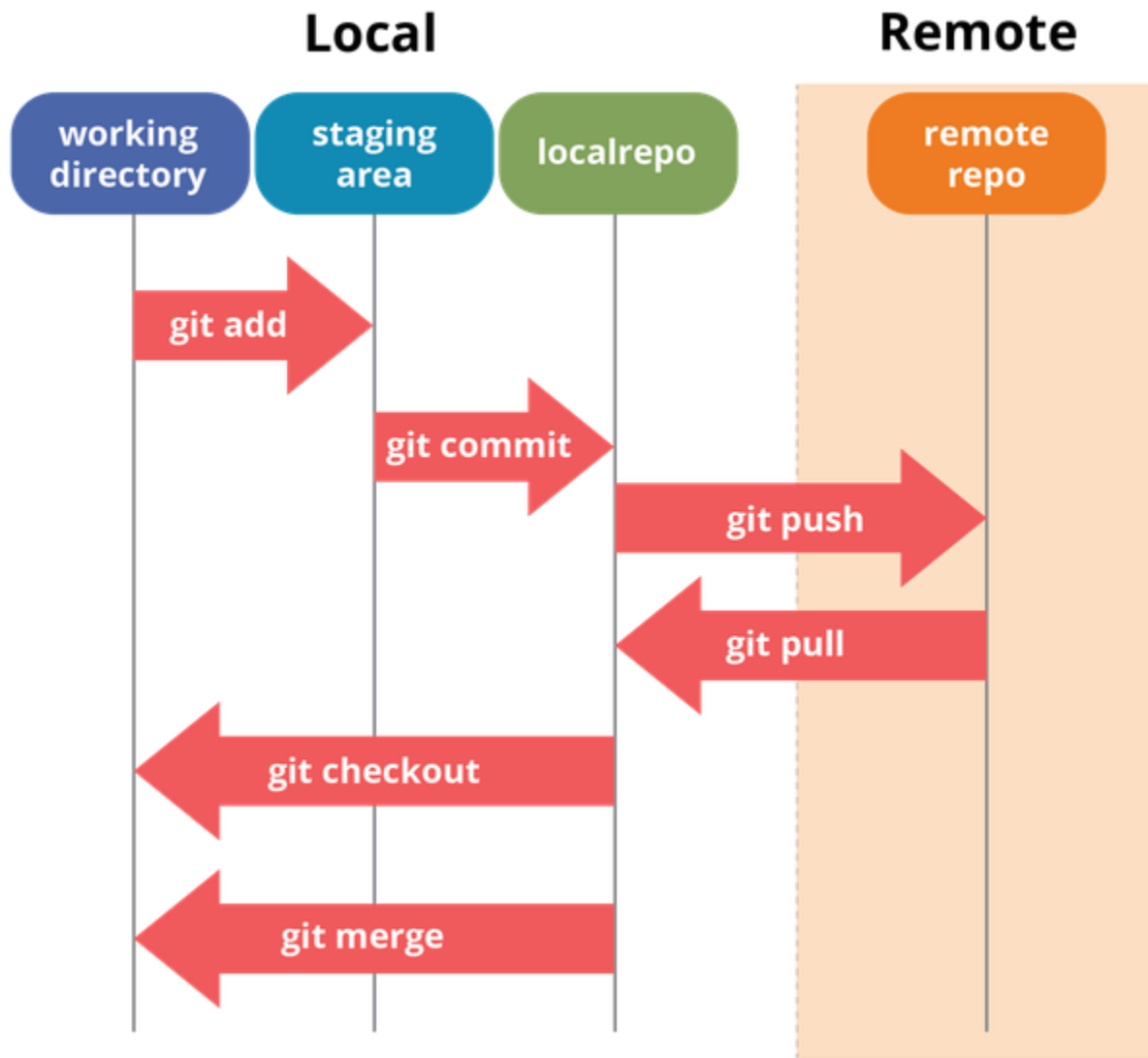
Add files via upload

Add an optional extended description...


- ☒ Commit directly to the `main` branch.
- ☐ Create a **new branch** for this commit and start a pull request. [Learn more about pull requests.](#)

Commit changes

Cancel



More questions ?


 **git** --local-branching-on-the-cheap


Search entire site...


About
Documentation
Reference
Book
Videos
External Links
Downloads
Community


Reference

Quick reference guides: [GitHub Cheat Sheet](#) | [Visual Git Cheat Sheet](#)


 [Complete list of all commands](#)

 **Setup and Config**
[git](#)
[config](#)
[help](#)
[bugreport](#)

 **Getting and Creating Projects**
[init](#)
[clone](#)

 **Basic Snapshotting**
[add](#)
[status](#)
[diff](#)
[commit](#)

Guides
[gitattributes](#)
[Command-line interface conventions](#)
[Everyday Git](#)
[Frequently Asked Questions \(FAQ\)](#)
[Glossary](#)
[githooks](#)
[gitignore](#)
[gitmodules](#)
[Revisions](#)
[Submodules](#)
[Tutorial](#)
[Workflows](#)

 **Email**