

Bioinformatics

wk01 TA hours

Introduction

2020/09/16

M.Y. HO

We have
a protein classification competition
at the end of the class!

Syllabus

Week 01 09/23	Introduction Python programming I	HW1	
Week 02 09/30	Python programming II	HW2	HW1 due
Week 03 10/07	Python package I Machine Learning	HW3 Final project	HW2 due
Week 04 10/21	Python package II	HW4(?)	HW3 due
Week 05 11/04	How to improve your model		HW4 due
Week 06 11/11	More information Student presentation		Final project

CONTENT

1 為何要使用Python

2 環境建置

3 Python programming I

4 HW1

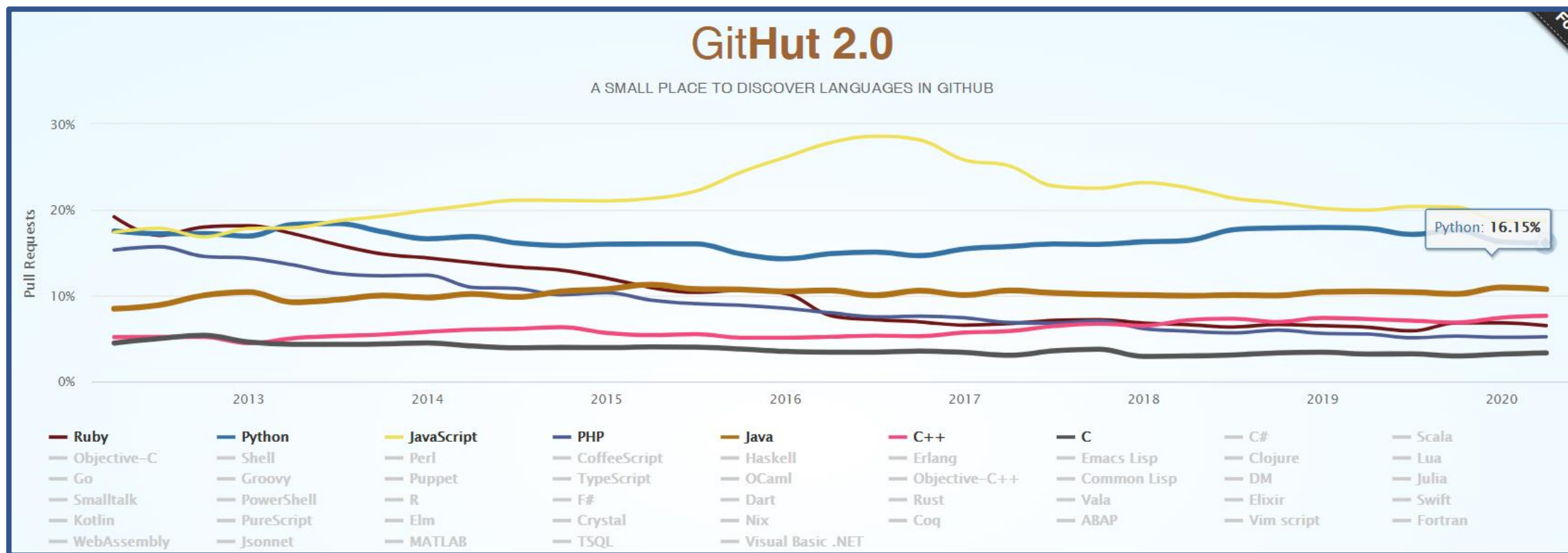
有無Programming經驗(any language)

家中電腦/筆電 OS

Part 1

為何要使用Python

為何要使用Python Popularity



為何要使用Python Popularity

Year		Quarter	
2020		2	
# Ranking	Programming Language	Percentage (Change)	Trend
1	JavaScript	18.789% (-1.133%)	
2	Python	16.108% (-1.694%)	
3	Java	10.731% (+0.250%)	
4	Go	8.922% (+1.006%)	
5	C++	7.636% (+0.383%)	
6	TypeScript	7.334% (+1.919%)	^
7	Ruby	6.492% (+0.196%)	v
8	PHP	5.198% (-0.318%)	v
9	C#	3.797% (-0.204%)	
10	C	3.320% (+0.130%)	

為何要使用Python

Versatility

hacking
web

🚩 PWNTOOLS

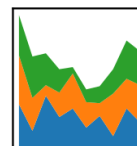
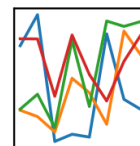
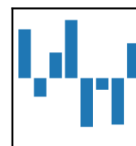
django

數據處理



NumPy

pandas
 $y_{it} = \beta' x_{it} + \mu_i + \epsilon_{it}$



視覺化

matplotlib Seaborn

生物資訊



PyMOL

爬蟲

BeautifulSoup



Scrapy

機器學習



PyTorch



TensorFlow

Part 2

環境建置

環境建置

Anaconda



ANACONDA®



環境建置

Anaconda

Anaconda Installers

Windows 

Python 3.8

64-Bit Graphical Installer (466 MB)


32-Bit Graphical Installer (397 MB)

MacOS 

Python 3.8

64-Bit Graphical Installer (462 MB)

64-Bit Command Line Installer (454 MB)

Linux 

Python 3.8

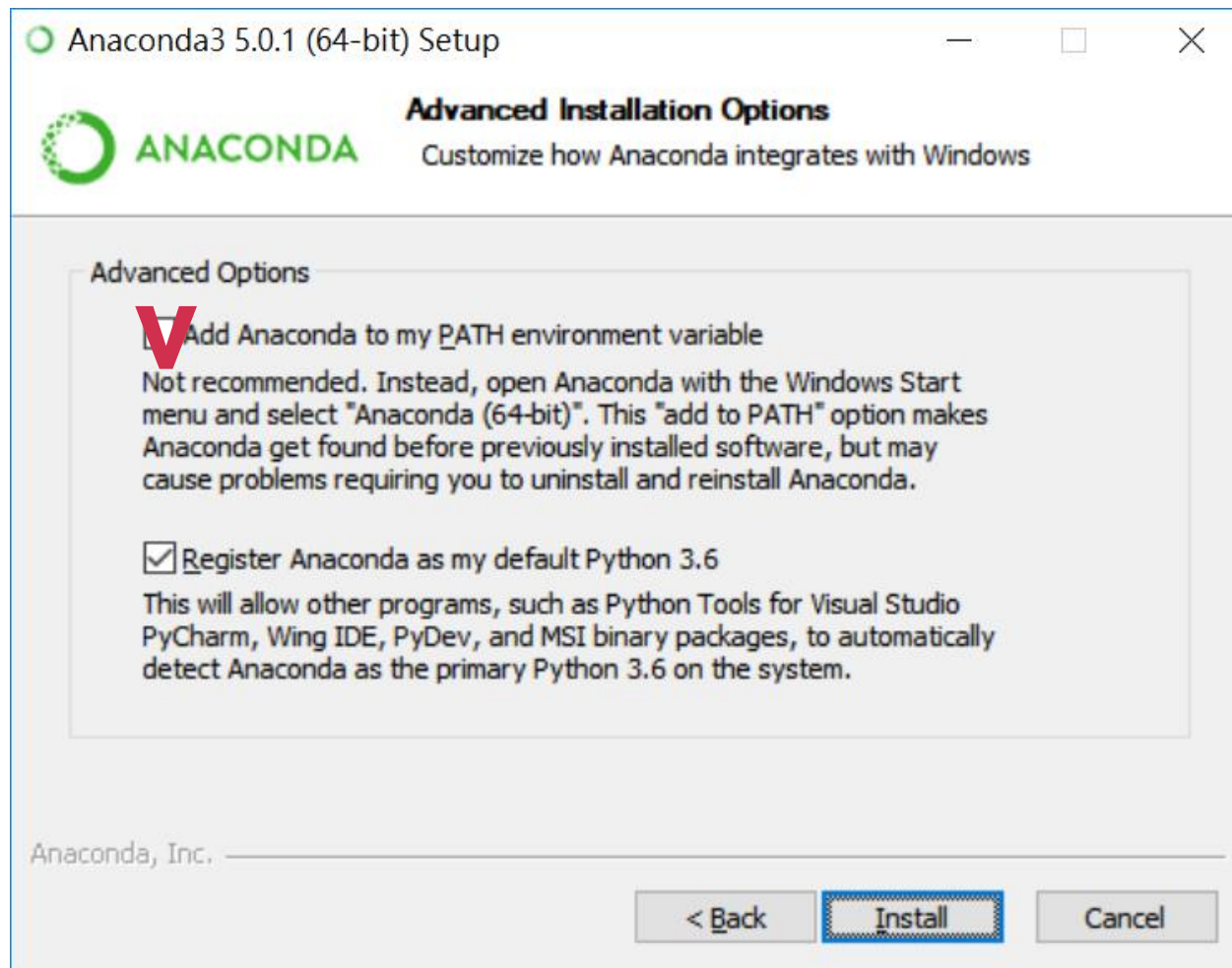
64-Bit (x86) Installer (550 MB)

64-Bit (Power8 and Power9) Installer (290 MB)

<https://www.anaconda.com/products/individual>

環境建置

Anaconda (windows)



環境建置

Anaconda (windows)



命令提示字元

```
Microsoft Windows [版本 10.0.18362.1082]  
(c) 2019 Microsoft Corporation. 著作權所有，並保留一切權利。  
C:\Users\User>jupyter notebook
```

環境建置

Anaconda (linux)

<https://repo.anaconda.com/archive/>

Anaconda installer archive

Filename	Size	Last Modified	MD5
Anaconda3-2020.07-Linux-ppc64le.sh	290.4M	2020-07-23 12:16:47	daf3de1185a390f435ab80b3c2212205
Anaconda3-2020.07-Linux-x86_64.sh	550.1M	2020-07-23 12:16:50	1046c40a314ab2531e4c099741530ada
Anaconda3-2020.07-MacOSX-x86_64.pkg	462.3M	2020-07-23 12:16:42	2941ddbaf0cdb49b342c18cde51fee43
Anaconda3-2020.07-MacOSX-x86_64.sh	454.1M	2020-07-23 12:16:44	50f20c90b8b5bfdc09759c09e32dce68
Anaconda3-2020.07-Windows-x86.exe	397.3M	2020-07-23 12:16:51	aa7dcf4d02baa25d14baf5728e29d067

```
$ curl -O https://repo.anaconda.com/archive/Anaconda3-2020.07-Linux-x86_64.sh
```

```
$ sha256sum Anaconda3-2019.03-Linux-x86_64.sh
```

 (Option)

```
$ bash Anaconda3-2020.07-Linux-x86_64.sh
```


```
$ source ~/.bashrc
```

```
$ nohup jupyter notebook --port 8888 --ip=0.0.0.0 --allow-root &
```

localhost:8888/tree

環境建置

Jupyter notebook

 jupyter Basic python programming | Last Checkpoint: 幾秒前 (unsaved changes)






Logout

File Edit View Insert Cell Kernel Widgets Help

Trusted

Python 3 

        Run    Code  

In []:

1

環境建置

Visual Studio Code

The image shows the Visual Studio Code website on the left and a screenshot of the Visual Studio Code IDE interface on the right.

Visual Studio Code Website:

- Navigation: Visual Studio Code, Docs, Updates, Blog, API, Extensions, FAQ, Search Docs, Download.
- Message: Version 1.49 is now available! Read about the new features and fixes from August.
- Text: Code editing. Redefined.
- Text: Free. Built on open source. Runs everywhere.
- Buttons: Download for Windows (Stable Build), Other platforms and Insiders Edition.
- Text: By using VS Code, you agree to its [license and privacy statement](#).

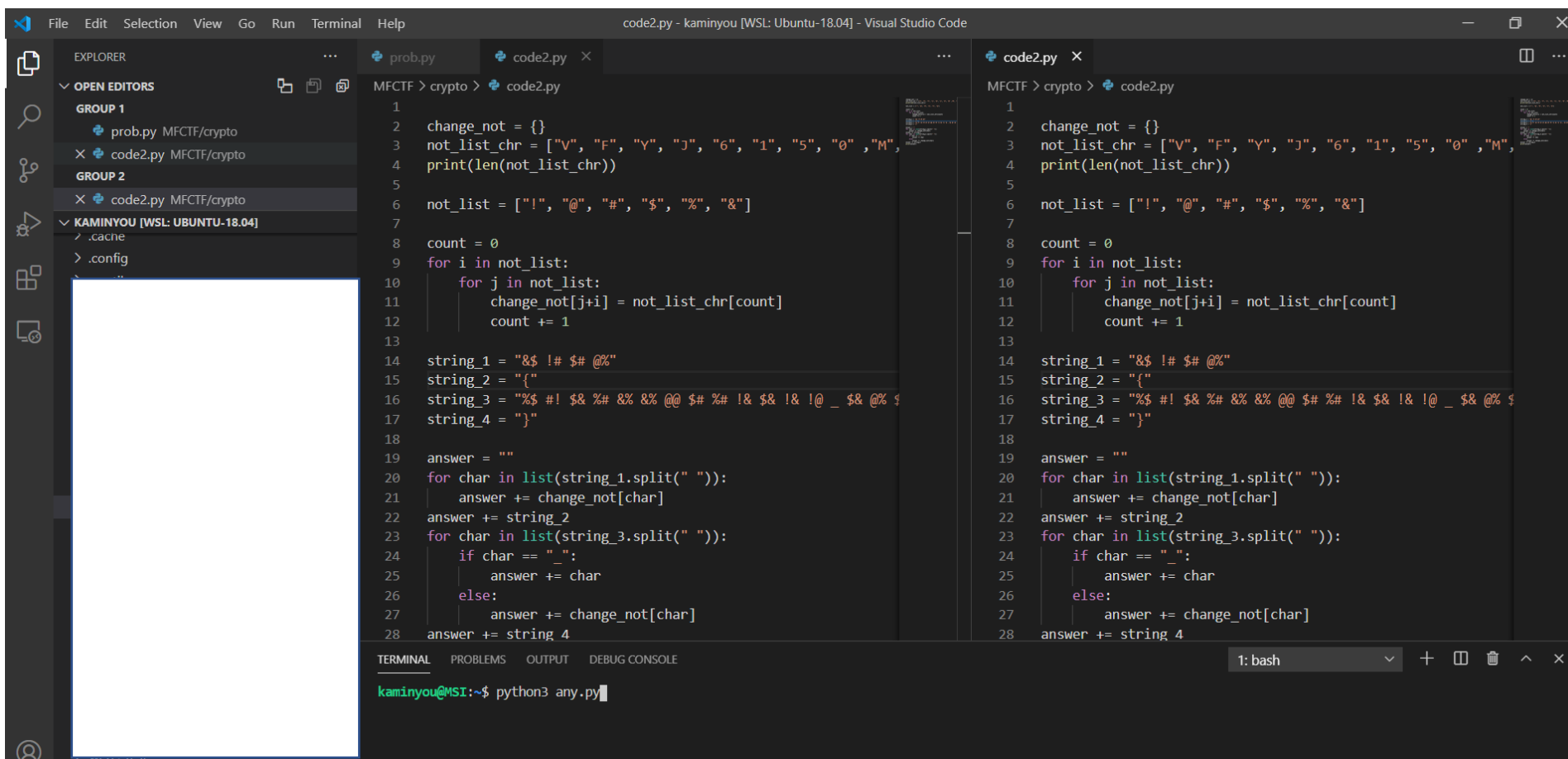
Visual Studio Code IDE Screenshot:

- File Explorer: Extensions: Marketplace.
- Search: @sortinstalls.
- Extensions List:
 - Python 2019.6.24221 54.9M 4.5 (Linting, Debugging (multi-threaded, ... Microsoft) Install)
 - GitLens — Git sup... 9.8.5 23.1M 5 (Supercharge the Git capabilities built... Eric Amodio Install)
 - C/C++ 0.24.0 23M 3.5 (C/C++ IntelliSense, debugging, and ... Microsoft Install)
 - ESLint 1.9.0 21.9M 4.5 (Integrates ESLint JavaScript into VS ... Dirk Baumer Install)
 - Debugger for Ch... 4.11.6 20.6M 4 (Debug your JavaScript code in the C... Microsoft Install)
 - Language Supp... 0.47.0 18.7M 4.5 (Java Linting, Intellisense, formatting, ... Red Hat Install)
 - vscode-icons 8.8.0 17.2M 5 (Icons for Visual Studio Code VSCode Icons Team Install)
 - Vetur 0.21.1 17M 4.5 (Vue tooling for VS Code Pine Wu Install)
- Code Editor: JS serviceWorker.js. Content:

```
src > JS serviceWorker.js > register > window.addEventListener("load") callback
39
40
41 checkValidServiceWorker(swUrl, config);
42
43 // Add some additional logging to localhost, p...
44 // service worker/PWA documentation.
45 navigator.serviceWorker.ready.then(() => {
46
47   product
48   productSub
49   removeSiteSpecificTrackingException
50   removeWebWideTrackingException
51   requestMediaKeySystemAccess
52   sendBeacon
53   serviceWorker (property) Navigator.serviceWorke...
54   storage
55   storeSiteSpecificTrackingException
56   storeWebWideTrackingException
57   userAgent
58   vendor
59
60 function registerValidSW(swUrl, config) {
61   navigator.serviceWorker
62     .register(swUrl)
63     .then(registration => {
```
- Terminal: 1: node. Message: You can now view create-react-app in the browser.

環境建置

Visual Studio Code



環境建置

記事本



*未命名 - 記事本

檔案(F) 編輯(E) 格式(O) 檢視(V) 說明

```
if __name__ == "__main__":  
    print("Hello world")
```

```
$ python3 helloworld.py
```

環境建置 Colab

 GoogleDrive_Tutorial.ipynb ☆

檔案 編輯 檢視畫面 插入 執行階段 工具 說明 系統不會儲存變更

共用 設定 明洋

+ 程式碼 + 文字 複製到雲端硬碟

連線 編輯

Google Drive 檔案存取教學

作者：黃健祐 Chien-yu Huang

若有任何問題，歡迎來信至助教信箱 ntu-ml-2020spring-ta@googlegroups.com

Colab 可以透過兩種方式存取 Google Drive 的檔案：

1. 透過檔案共用連結

一般來說，Google Drive 的檔案連結會是類似下面的結構：

https://drive.google.com/open?id=1duQU7xqXRSPYeOR0zLiSA8g_LCFzoV

其中 "open?id=" 後面的那一串亂碼稱為 **file_id**

在 colab 中，我們可以直接用 **file_id** 下載檔案，優點是這個過程不需要再手動輸入任何訊息，但缺點是下載下來的檔案在程式執行完畢後就會消失，也就是每一次都必須重新下載，使得執行時間增加。

```
[ ] # 下載 file_id 為 "1duQU7xqXRSPYeOR0zLiSA8g_LCFzoV" 的檔案，並將它命名為 Minori.jpg
!gdown --id '1duQU7xqXRSPYeOR0zLiSA8g_LCFzoV' --output Minori.jpg
# 列出目前目錄下所有的檔案
!ls
```

Downloading...
From: https://drive.google.com/uc?id=1duQU7xqXRSPYeOR0zLiSA8g_LCFzoV
To: /content/Minori.jpg

Jupyter notebook -> .ipynb

Python -> .py

Part 3

Python programming I

Part 4

HW1

HW1

建好你的環境

請確保你可以進行
python programming

HW1

Hello bioinformatics

請寫一個程式，可以輸出以下結果

```
kaminyou@MSI:~/bioinformatiics$ python3 yourid_hw1.py
How many times: 5
Counts:1 || Hello bioinformatics!
Counts:2 || Hello bioinformatics!
Counts:3 || Hello bioinformatics!
Counts:4 || Hello bioinformatics!
Counts:5 || Hello bioinformatics!
```

```
kaminyou@MSI:~/bioinformatiics$ python3 yourid_hw1.py
How many times: 21
ERROR: MAX=20
```

```
kaminyou@MSI:~/bioinformatiics$ python3 yourid_hw1.py
How many times: -1
ERROR: negative value is not allowed
```

Thank you

r08945027@ntu.edu.tw