

파이썬을 활용한 실전 웹크롤링 CAMP

2018. 04. 21
실습환경 가이드 라인

PYTHON설치

<https://www.python.org/downloads/>

Version = 3.6.2

Python Software Foundation [US] | <https://www.python.org/downloads/>

Wondering which version to use? [Here's more about the difference between Python 2 and 3.](#)

Looking for Python with a different OS? Python for [Windows](#), [Linux/UNIX](#), [Mac OS X](#), [Other](#)

Want to help test development versions of Python? [Pre-releases](#)

Looking for a specific release?

Python releases by version number:

Release version	Release date	Click for more	
Python 3.6.3	2017-10-03	Download	Release Notes
Python 3.3.7	2017-09-19	Download	Release Notes
Python 2.7.14	2017-09-16	Download	Release Notes
Python 3.4.7	2017-08-09	Download	Release Notes
Python 3.5.4	2017-08-08	Download	Release Notes
Python 3.6.2	2017-07-17	Download	Release Notes
Python 3.6.1	2017-03-21	Download	Release Notes
Python 3.4.6	2017-01-17	Download	Release Notes

[View older releases](#)

PYTHON설치

왼쪽 그림과 같이
파이썬 3.6을 패스에 추가

PATH란?

내 컴퓨터에서
PYTHON이라는 언어를 사용할때
기본적으로 사용될 경로!!

앞으로 PYTHON을 실행하면 지금 설치하는
PYTHON의 설치폴더로 자동적으로 연결해 줌!

혹시 급하게 설치하느라 체크하지 못했다면
강사에게 개인적으로 연락주세요!



다음부터는 MAC OSX사용자와 WINDOW 사용자의 명령어가 다른 이유로
슬라이드가 나누어져 있습니다.

MAC사용자는 바로 6페이지로 넘어가주세요!

윈도우 사용자용

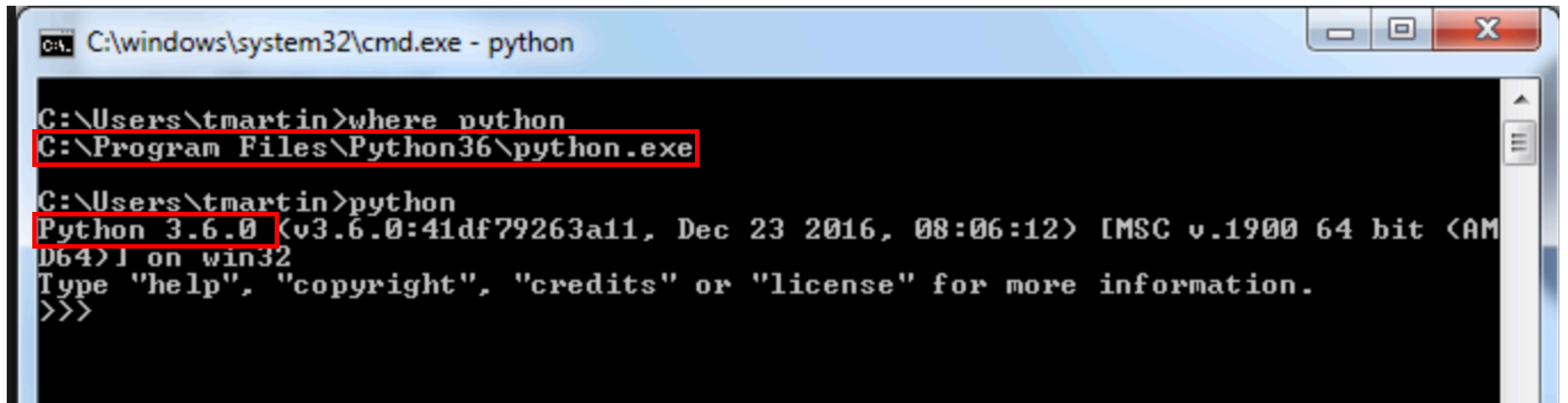
시작 - CMD 검색 후 실행 - 까만 바탕의 cmd가 실행이 되었다면

where python 입력 후 엔터

아래의 콘솔창과 같이 경로가 하나만 나온다면 OK!

또한 python을 입력하고 엔터를 눌렀을 때 버전이 설치한 버전과 같은지도 확인!

BUT! 만약 경로가 여러개가 나온다면 강사에게 문의!



```
C:\windows\system32\cmd.exe - python

C:\Users\tmartin>where python
C:\Program Files\Python36\python.exe

C:\Users\tmartin>python
Python 3.6.0 (v3.6.0:41df79263a11, Dec 23 2016, 08:06:12) [MSC v.1900 64 bit (AMD64)] on win32
Type "help", "copyright", "credits" or "license" for more information.
>>>
```

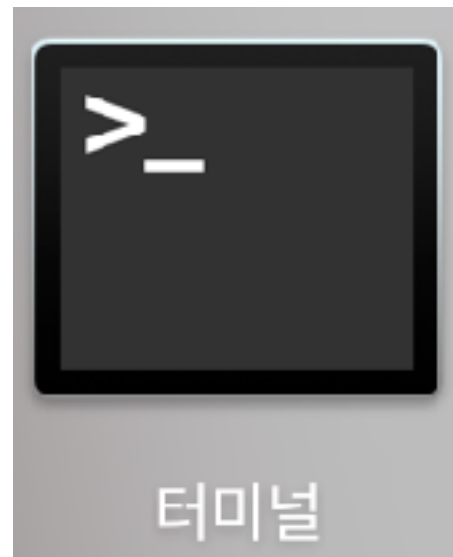
Mac 사용자용



MAC은 기본적으로 파이썬 2버전은 탑재한 상태
따라서 기본적으로 python이라는 PATH는 2버전에 연결이 되어있다.
따라서 python3라는 PATH에 우리가 설치한 버전이 연결된다.
아래의 콘솔창과 같이 경로가 하나만 나온다면 OK!

또한 python3을 입력하고 엔터를 눌렀을 때 버전이 설치한 버전과 같은지도 확인!
BUT! 만약 경로가 여러개가 나오거나 / 나오지 않는다면 강사에게 문의!

Mac 사용자용



```
joono — Python — 80x23
Last login: Wed Apr 11 18:14:42 on ttys001
[joonoui-MacBook-Pro:~ joono$ which python3
/Library/Frameworks/Python.framework/Versions/3.6/bin/python3
[joonoui-MacBook-Pro:~ joono$ which python
/usr/bin/python
[joonoui-MacBook-Pro:~ joono$ python
Python 2.7.10 (default, Oct 6 2017, 22:29:07)
[GCC 4.2.1 Compatible Apple LLVM 9.0.0 (clang-900.0.31)] on darwin
Type "help", "copyright", "credits" or "license" for more information.
[>>> exit()
[joonoui-MacBook-Pro:~ joono$ python3
Python 3.6.2 (v3.6.2:5fd33b5926, Jul 16 2017, 20:11:06)
[GCC 4.2.1 (Apple Inc. build 5666) (dot 3)] on darwin
Type "help", "copyright", "credits" or "license" for more information.
>>>
```

PYCHARM 설치

<https://www.jetbrains.com/pycharm/download/>

Download PyCharm

Windows

macOS

Linux

Professional

Full-featured IDE
for Python & Web
development

DOWNLOAD

Free trial

Community

Lightweight IDE
for Python & Scientific
development

DOWNLOAD

Free, open-source

Community 버전을 설치합니다!

설치한 후 프로젝트 생성은 수업시간에
같이 진행할 예정이니
설치만 하시면 됩니다!

Window 가상환경 설정

Window CMD를 열고 다음 명령어를 입력

파이썬 설치 경로(where python으로 확인 가능!) -m venv 가상환경 이름

Ex) C:\Python36\python -m venv gasang

파이썬 설치 경로는 각자의 컴퓨터마다 다를 수 있으니 꼭! 확인하고 입력해주세요
또한 가상환경 이름은 영어로 하되, 가능하면 짧게 입력하는 것이 좋습니다!

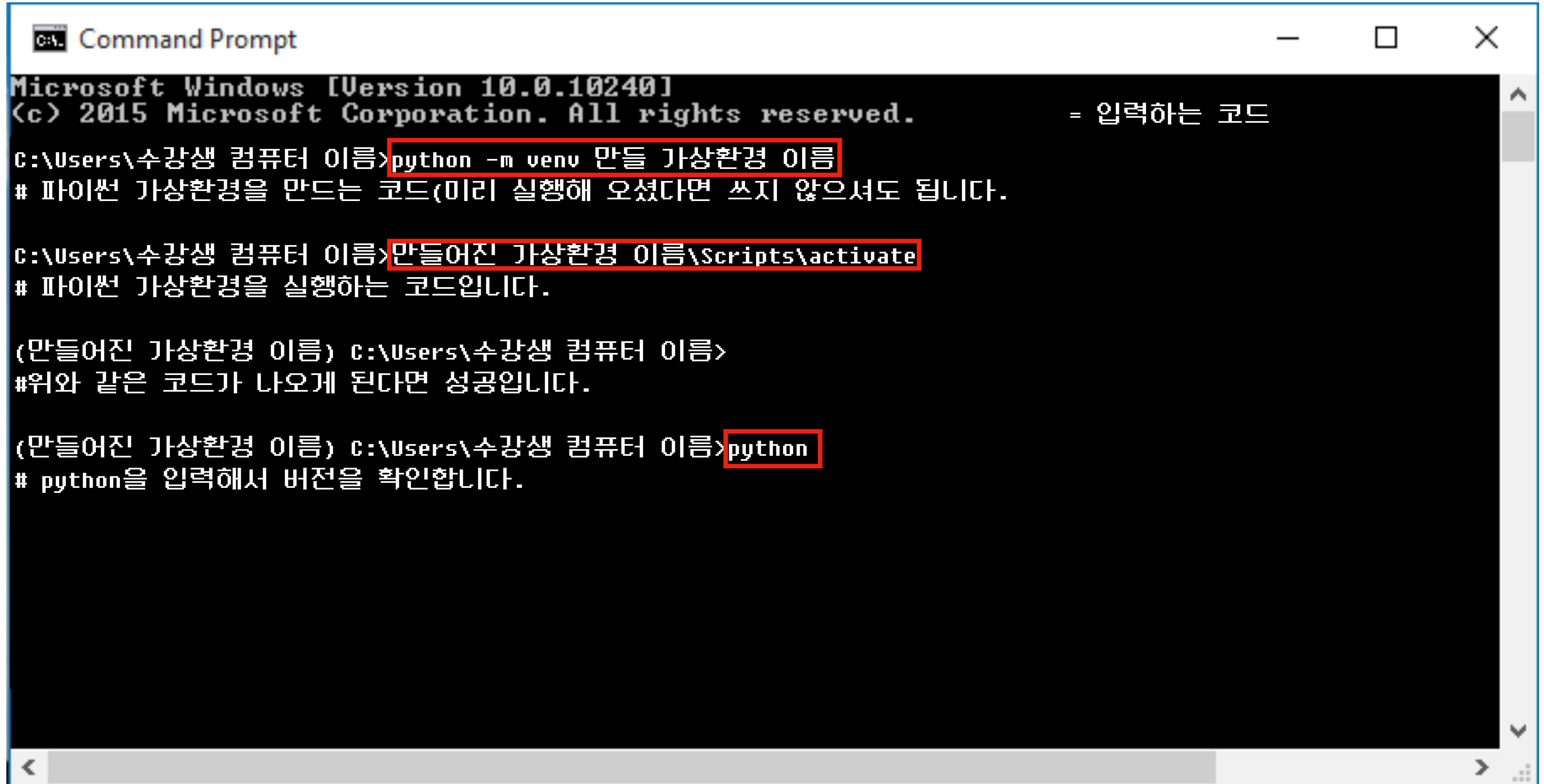
가상환경을 만들었다면 다음 명령어를 통해 가상환경을 활성화 합니다

가상환경 이름\Scripts\activate

이상태에서 이제 다음장으로 넘어가서 패키지 설치를 진행합니다.

(가상환경이름)이 CMD명령어 줄 맨 앞에 나와있는지를 꼭 확인해주세요!

Window 가상환경 설정



```
C:\> Command Prompt

Microsoft Windows [Version 10.0.10240]
(c) 2015 Microsoft Corporation. All rights reserved.

C:\Users\수강생 컴퓨터 이름>python -m venv 만들 가상환경 이름
# 파이썬 가상환경을 만드는 코드(미리 실행해 오셨다면 쓰지 않으셔도 됩니다.)

C:\Users\수강생 컴퓨터 이름>만들어진 가상환경 이름\Scripts\activate
# 파이썬 가상환경을 실행하는 코드입니다.

(만들어진 가상환경 이름) C:\Users\수강생 컴퓨터 이름>
#위와 같은 코드가 나오게 된다면 성공입니다.

(만들어진 가상환경 이름) C:\Users\수강생 컴퓨터 이름>python
# python을 입력해서 버전을 확인합니다.
```

= 입력하는 코드

MAC 가상환경 설정

터미널을 열고 다음 명령어를 입력

python3 -m venv 가상환경 이름

Ex) python3 -m venv gasang

가상환경 이름은 영어로 하되, 가능하면 짧게 입력하는 것이 좋습니다!

가상환경을 만들었다면 다음 명령어를 통해 가상환경을 활성화 합니다

source 가상환경 이름\bin\activate 왼쪽의 명령어가 오류가 난다면

. 가상환경 이름\bin\activate 을 입력합니다.

이상태에서 이제 다음장으로 넘어가서 패키지 설치를 진행합니다.

(가상환경이름)이 CMD명령어 줄 맨 앞에 나와있는지를 꼭 확인해주세요!

MAC 가상환경 설정

```
joono — Python — 121x22
Last login: Tue Apr 17 21:47:53 on ttys001
[joonoui-MacBook-Pro:~ joono$ python3 -m venv fastcampus
[joonoui-MacBook-Pro:~ joono$ . fastcampus/bin/activate
(fastcampus) joonoui-MacBook-Pro:~ joono$ python
Python 3.6.2 (v3.6.2:5fd33b5926, Jul 16 2017, 20:11:06)
[GCC 4.2.1 (Apple Inc. build 5666) (dot 3)] on darwin
Type "help", "copyright", "credits" or "license" for more information.
>>>
```

python3 -m venv fastcampus = 입력하는 코드

. fastcampus/bin/activate

python

패키지 설치

```
C:\> Command Prompt

Microsoft Windows [Version 10.0.10240]
(c) 2015 Microsoft Corporation. All rights reserved.

(만들어진 가상환경 이름) C:\Users\수강생 컴퓨터 이름>pip install pandas==0.20.3
# pandas를 설치해보겠습니다.
Collecting pandas
  Downloading https://files.pythonhosted.org/packages/31/
b4/38aa707d65f0e051abac7845280557919f4be77a8b5b1bd7e42ecd90ca6b/pandas-0.22.0-cp36-cp36m-
macosx_10_6_intel.macosx_10_9_intel.macosx_10_9_x86_64.macosx_10_10_intel.macosx_10_10_x86_64.whl
(14.9MB)
    100% |████████████████████████████████████████| 14.9MB 98kB/s
.....(중간에 설치과정이 길어서 생략했습니다.)
Installing collected packages: numpy, pytz, six, python-dateutil, pandas
Successfully installed numpy-1.14.2 pandas-0.22.0 python-dateutil-2.7.2 pytz-2018.4 six-1.11.0
# 마지막줄이 성공적으로 설치되었다는 Successfully installed 가 나오면 됩니다.
```

패키지 설치

MAC버전

```
(fastcampus) joonoui-MacBook-Pro:~ joono$ pip3 install pandas==0.20.3
Collecting pandas
  Downloading https://files.pythonhosted.org/packages/31/b4/38aa707d65f0e051abac7845280557919f4be77a8b5b1bd7e42ecd90ca6b/pandas-0.22.0-cp36-cp36m-macosx_10_6_intel.macosx_10_9_intel.macosx_10_9_x86_64.macosx_10_10_intel.macosx_10_10_x86_64.whl (14.9MB)
    100% |#####| 14.9MB 98kB/s
.....(중간에 설치과정이 길어서 생략했습니다.)
Installing collected packages: numpy, pytz, six, python-dateutil, pandas
Successfully installed numpy-1.14.2 pandas-0.22.0 python-dateutil-2.7.2 pytz-2018.4 six-1.11.0
# 마지막줄이 성공적으로 설치되었다는 Successfully installed 가 나오면 됩니다.
```

필요 패키지

Selenium 3.4.3

bs4 0.0.1

pandas 0.20.3

PyMySQL 0.8.0

browsermob-proxy 0.8.0






Window Mac 모두 위의 패키지를 예시와 같이 설치해 주시면 됩니다.

pip install 패키지명==버전 과 같이 설치를 해주시면 됩니다.

Web driver 설치

<https://chromedriver.storage.googleapis.com/index.html?path=2.37/>

Index of /2.37/

	<u>Name</u>	Last modified	Size	ETag
	Parent Directory		-	
	chromedriver_linux64.zip	2018-03-20 15:22:39	3.55MB	f4fbbf50eaec2b9c9e21536e39b9999e
	chromedriver_mac64.zip	2018-03-20 18:30:06	5.26MB	5e68436af99f9348d67127e5ab8d5eed
	chromedriver_win32.zip	2018-03-20 15:12:34	3.19MB	00c991baa1e51abf6d582aediae91291f
	notes.txt	2018-03-16 21:35:44	0.01MB	144b47991f471631ce010cd83d4c6a6c

본인의 운영체제에 맞는 파일을 다운받아서
홈 폴더(사용자 or Users 바로 밑의 내 컴퓨터 로그인 이름으로 된 폴더)에
저장해 놓으시면 됩니다!

Slack 가입하기

fc_wc ▾

● jjjjooonno

🔔

All Threads

Channels

+

questions

general

Direct Messages

+

♥ slackbot

● jjjjooonno (you)

+ Invite People

Apps

+

#questions

☆ | 👤 1 | 📌 0 | ✎ Add a topic

📞 ⓘ ⚙️ 🔍 Search @ ☆ ⋮

questions

You created this channel today. This is the very beginning of the #questions channel.
Purpose: 질문을 위한 채팅방입니다! (edit)
[+ Add an app](#) [👤 Invite others to this channel](#)

Today

jjjjooonno

12:43 AM

joined #questions.

jjjjooonno

12:43 AM

set the channel purpose: 질문을 위한 채팅방입니다!

jjjjooonno

7:43 PM

added this Python snippet: [question example](#) ▾

1 hi = 'hello'

2 print(hi1)

jjjjooonno

7:44 PM

“ 프린트가 안되네요 ㅠㅠ

+

Message #questions

@ 😊

< Files

×

jjjjooonno

question example ☆

✎ ⋮

1 hi = 'hello'

2 print(hi1)

✓ Wrap • Today at 7:43 PM • 24 B Python snippet •

Shared in [#questions](#) 🗨️

jjjjooonno

Today at 7:44 PM ☆

프린트가 안되네요 ㅠㅠ

Commenting in [#questions](#)

Add Comment

Slack 가입하기

슬랙은 위 슬라이드의 사진과 같이 질문을 받는 용도로 사용할 채팅방이라고 생각하시면 됩니다.

카톡이나 페이스북 페이지와 다른점은 질문에 코드를 삽입할 수 있다는 점입니다. 코드 실행시 오류가 나는 경우에 해결이 어렵거나, 혹은 기타 질문이 있을 경우에 질문글을 작성해 주시면 됩니다.

또한 이미지도 첨부가 가능하기 때문에 캡처한 이미지도 첨부하여 질문을 주시면 더욱 빠른 답변이 가능합니다.

https://join.slack.com/t/fcwcgroup/shared_invite/enQtMzQ0NjMyNDA2MDIwLTk5OGU3ZjBhYzdiMThmYTA3OGViYTE1ZDQxYTY5ZGM3ODVhYTY2OWJiMmM2ZDIyNTUzNTVINmI0MGExOWNiYTM

위 링크를 통해 slack에 가입하고 패스트캠퍼스 웹크롤링 강의를 위한 채널에 접근할 수 있습니다.

github에서 강의자료 받기

https://github.com/jjjjoonno/fc_wc

강의를 위한 자료는 모두 깃헙에 업로드할 예정입니다.

강의자료에는 PPT파일과 기타 공유자료들을 업로드
코드에는 실습 자료들과 과제가 업로드 될 예정입니다!