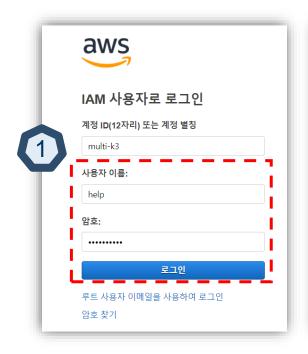
V-Lab (Group/EC2, DB, Network) 사용자 가이드

본 교안은 K-digital Training 멀티캠퍼스 국비지원교육 프로그램 데이터 사이언스/엔지니어링 전문가 과정을 위해 제작되었으며, 본 교육 外 배포/게시/공개를 금합니다.

실습용 서버 정보 (1/6)

아래 링크로 들어가 안내 드린 사용자 정보로 접속(비밀번호 변경 권장)

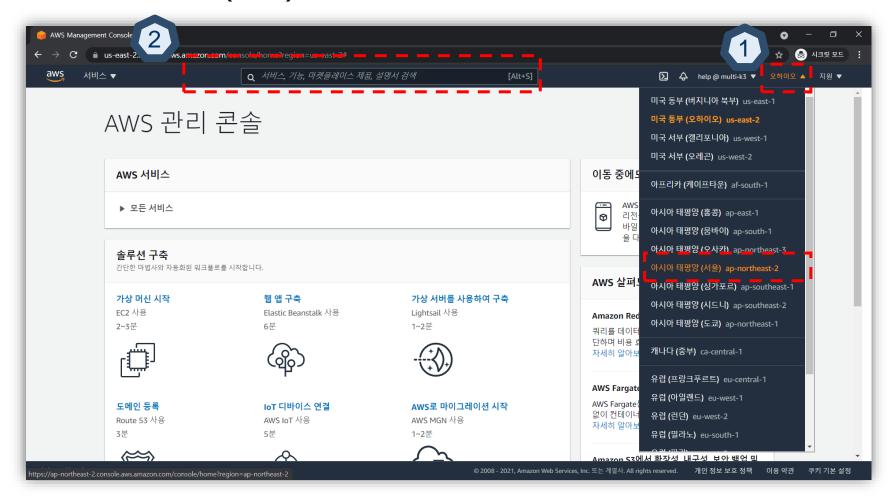
- https://multi-k3.signin.aws.amazon.com/console
- 계정 ID(12자리) 또는 계정 별칭 : multi-k3 [공통]





실습용 서버 정보 (2/6)

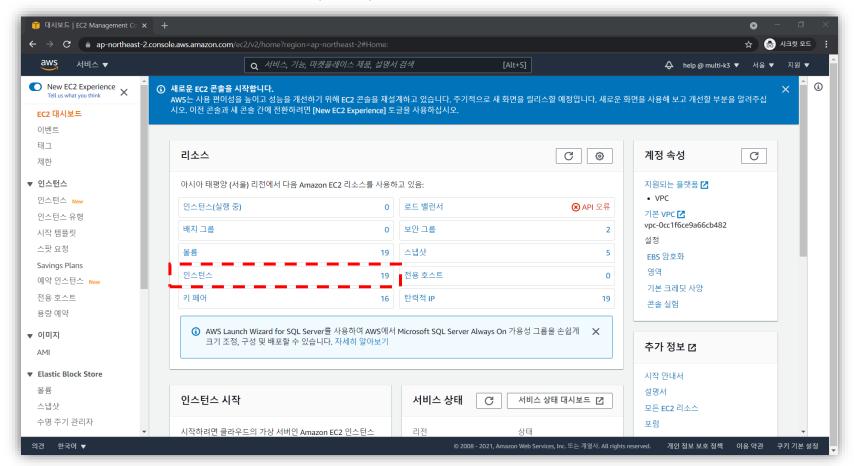
우측 상단 리전 정보(서울) 변경 후 서비스 검색창에서 EC2 입력 후 클릭



실습용 서버 정보 (3/6)

현재 사용 가능한 서버 정보 확인(최소 권한 부여)

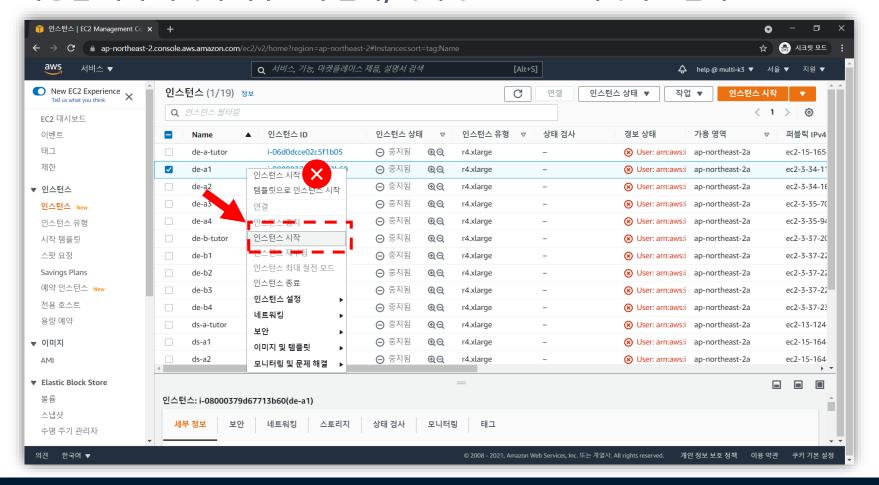
■ 인스턴스 클릭 ※주의: 인스턴스(실행중) 아님



실습용 서버 정보 (4/6)

실습용 서버 동작 확인 및 서버 실행(사용자별로 서버 실행권한 별도부여)

■ 사용할 서버 위에서 마우스 우클릭, 아래쪽 인스턴스 시작버튼 클릭

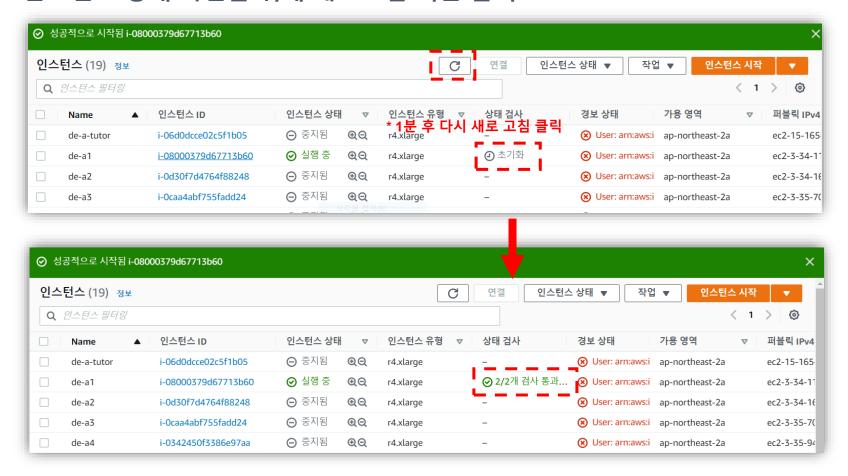


Copyright © 2013 Global Software Campus Co, Ltd. all rights reserved

실습용 서버 정보 (5/6)

실습용 서버 동작 확인 및 서버 실행(사용자별로 서버 실행권한 별도부여)

■ 인스턴스 상태 확인을 위해 새로고침 버튼 클릭



원격 접속 서버 정보

■ 서버 사양:

*** HW**

- 고주파수 인텔 제온 E5-2686 v4(Broadwell)급 프로세서, DDR4 메모리
- 4vCPU / 30.5GB 급 RAM / 1000GB SSD
- Ubuntu 18.04 0S
- mxnet, tensorflow, caffe2, chainer, cntk, pytorch 등으로 이루어진 가상환경 有

***** SW

- Anaconda3 / python 3.7.6
- spark-3.1.2-bin-hadoop3.2(단일 머신) 설치
- 기타 : (base) Anaconda 환경에서 TensorFlow, Keras 등 머신러닝, 딥러닝 관련 패키지 및 라이브러리 추가 설치 가능
- 서버 운영시간 : (평일) ~18:30(18시 30분 이후 서버 자동 종료 및 계정 잠금)

(주말) 추석 연휴, 멘토링 일정 없는 토요일과 일요일은 사용 불가

서버 접속을 위한 사전 준비

원격 서버 접속을 위해 암호화 키를 개인 보관하고 원격 콘솔프로그램인 PuTTY를 다운로드



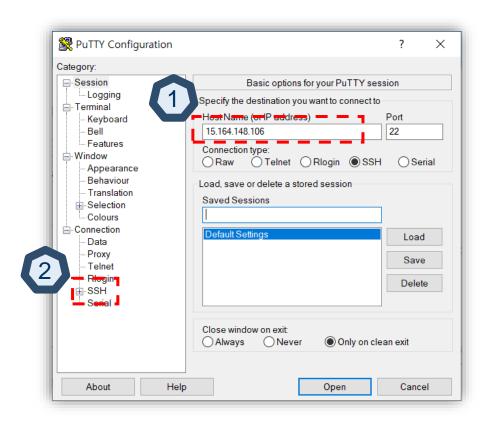
- d*-**.ppk 암호화 키 보관
- <u>www.putty.org</u> 사이트를 통해 PuTTY 64-bit 다운로드

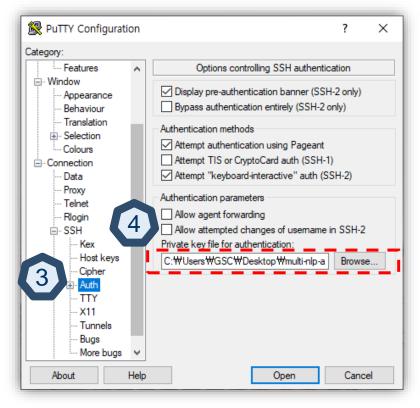


PuTTY 프로그램 실행 및 설정 (1/2)

사용자별로 부여 받은 서버 IP 입력 및 암호화 키 업로드

(암호화 키 : d*-**.ppk)

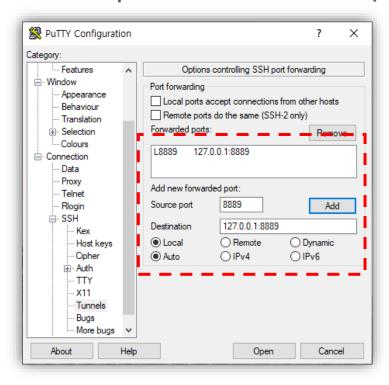




PuTTY 프로그램 실행 및 설정 (2/2)

원격 서버에서 다른 사용자들과 Jupyter-Notebook을 동시 사용시 접속 충돌 방지를 위해 putty 설정에서 터널링 작업 추가 진행 후 저장

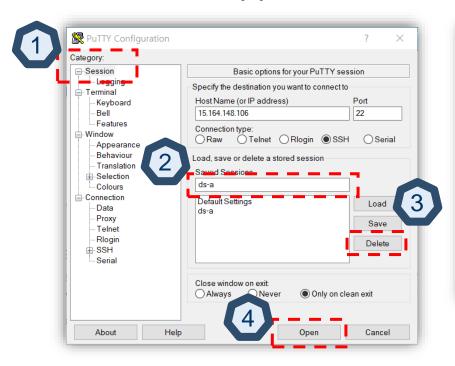
- Source port(8xxx), Destination(127.0.0.1:8xxx) 입력
- 각자 부여된 Jupyter notebook port번호로 터널링 설정(8889 ~ 9000)

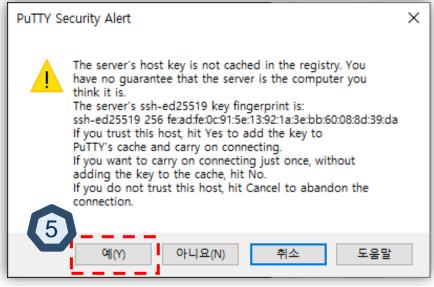


PuTTY 설정 내용 저장

반드시 Session으로 돌아와 설정 내용을 저장

- 설정 내용을 저장해 놓으면 프로그램을 열 때마다 재설정할 필요 없음
- 최초 접속 시 예(Y) 클릭





참고1 . MAC에서 서버 접속

맥북 사용자가 서버에 접속하는 방법

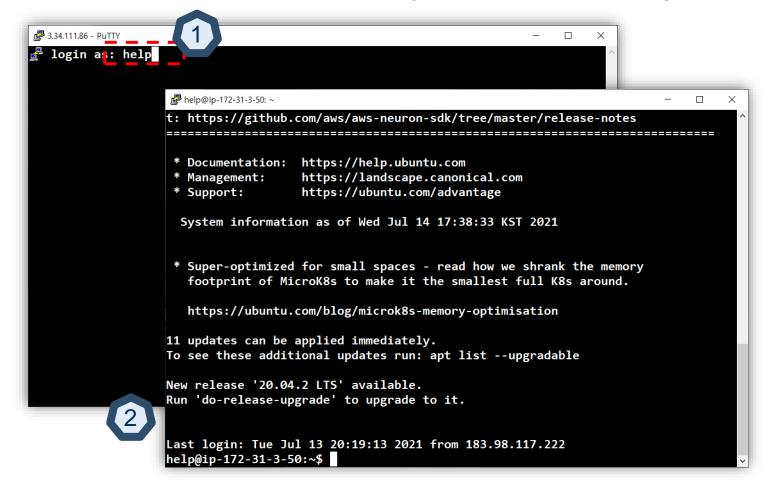
- 기본 터미널 프로그램 실행 후 암호화 키 저장위치로 이동 후 권한 변경 : \$ chmod 400 /암호화 키 저장 위치/d*-**.pem
- 터미널 창에서 실습 서버에 접속

```
: $ ssh -i ~/Downloads/d*-**.pem help@3.34.111.86 -p 22 -L 8xxx:127.0.0.1:8xxx

맥북 내 암호키 저장 경로 부여 받은 ID@IP (ssh접속을 위한 port 터널링)
```

서버 접속

부여 받은 ID를 입력해 서버에 접속(별도의 PW필요 없음)



데이터 분석을 위한 주피터노트북 사용 (1/4)

- 기 설치된 아나콘다 가상 환경 목록 검색 및 실행
- conda info --envs
- conda activate [가상 환경 이름] ([]미 입력시 base 환경으로 접속)

```
₽ help@ip-172-31-3-50: ~
                                                                         New release '20.04.2 LTS' available.
Run 'do-release√
                        to upgrade to it.
Last login: Tue
                      20:19:13 2021 from 183.98.117.222
help@ip-172-31-3-50:~$ conda info --envs
# conda environments:
base
                     * /home/ubuntu/anaconda3
                        /home/ubuntu/anaconda3/envs/amazonei mxnet p36
amazonei mxnet p36
aws neuron mxnet p36
                        /home/ubuntu/anaconda3/envs/aws neuron mxnet p36
aws_neuron_pytorch_p36
                          /home/ubuntu/anaconda3/envs/aws neuron pytorch p36
                             /home/ubuntu/anaconda3/envs/aws_neuron_tensorflow
aws neuron tensorflow p36
mxnet latest p37
                        /home/ubuntu/anaconda3/envs/mxnet latest p37
mxnet p36
                        /home/ubuntu/anaconda3/envs/mxnet p36
                        /home/ubuntu/anaconda3/envs/python3
python3
                        /home/ubuntu/anaconda3/envs/pytorch_latest_p37
pytorch_latest p37
pytorch_p36
                        /home/ubuntu/anaconda3/envs/pytorch_p36
                          /home/ubuntu/anaconda3/envs/tensorflow2 latest p37
tensorflow2 latest p37
tensorflow2_p36
                        /home/ubuntu/anaconda3/envs/tensorflow2 p36
tensorflow p37
                        /home/ubuntu/anaconda3/envs/tensorflow p37
help@ip-172-31-3-50:~$
```

```
# help@ip-172-31-3-50: ~
Run 'do-release-upgrade' to upgrade to it.
Last login: Tue Jul 13 20:19:13 2021 from 183.98.117.222
help@ip-172-31-3-50:~$ conda info --envs
# conda environments:
                      * /home/ubuntu/anaconda3
base
                         /home/ubuntu/anaconda3/envs/amazonei mxnet p36
amazonei mxnet p36
                         /home/ubuntu/anaconda3/envs/aws neuron mxnet p36
aws neuron mxnet p36
aws_neuron_pytorch_p36
                           /home/ubuntu/anaconda3/envs/aws_neuron_pytorch_p36
aws neuron tensorflow p36
                              /home/ubuntu/anaconda3/envs/aws neuron tensorflow
mxnet_latest_p37
                         /home/ubuntu/anaconda3/envs/mxnet latest p37
                         /home/ubuntu/anaconda3/envs/mxnet_p36
mxnet_p36
python3
                         /home/ubuntu/anaconda3/envs/python3
                         /home/ubuntu/anaconda3/envs/pytorch latest p37
pytorch_latest_p37
pytorch_p36
                         /home/ubuntu/anaconda3/envs/pytorch_p36
                           /home/ubuntu/anaconda3/envs/tensorflow2 latest p37
tensorflo
              est p37
                         /home/ubuntu/anaconda3/envs/tensorflow2 p36
tensorf]
                         /home/ubuntu/anaconda3/envs/tensorflow p37
tensorf
help@ip-17z-31-3-50:~$ conda activate
(base) help@ip-172-31-3-50:~$
```

참고2. 주피터노트북 활용

아나콘다 가상 환경 개인적으로 생성 가능

■ 원하는 패키지 활용 개인생성 예시

conda create -n [이름] python=[버전] jupyter(기본 설치 권장) [패키지 명]

```
₽ help@ip-172-31-3-50: ~
(base) help@ip-172-31-3-50:∼∮ conda create -n test python=3.6 jupyter tensorflow
Collecting package metadata (current repodata.json): done
Solving environment: failed with repodata from current repodata.json, will retry with next
repodata source.
Collecting package metadata (repodata.json): done
Solving environment: done
==> WARNING: A newer version of conda exists. <==
  current version: 4.8.4
  latest version: 4.10.3
Please update conda by running
    $ conda update -n base -c defaults conda
## Package Plan ##
  environment location: /home/ubuntu/anaconda3/envs/test
  added / updated specs:

    jupyter
```

데이터 분석을 위한 주피터노트북 사용 (2/4)

주피터 노트북 사용을 위한 명령어 입력

■ Jupyter-notebook --ip=0.0.0.0 --no-browser --port=8xxx

각자 부여 받은 port번호 사용

```
help@ip-172-31-3-50: ~
                                                                                      Last login: Tue Jul 13 20:19:13 2021 from 183.98.117.222
help@ip-172-31-3-50:~$ conda info --envs
# conda environments:
                      * /home/ubuntu/anaconda3
base
                         /home/ubuntu/anaconda3/envs/amazonei mxnet p36
amazonei mxnet p36
aws_neuron_mxnet_p36
                         /home/ubuntu/anaconda3/envs/aws neuron mxnet p36
aws_neuron_pytorch_p36
                           /home/ubuntu/anaconda3/envs/aws neuron pytorch p36
aws neuron tensorflow p36
                              /home/ubuntu/anaconda3/envs/aws neuron tensorflow
p36
mxnet latest p37
                         /home/ubuntu/anaconda3/envs/mxnet latest p37
                         /home/ubuntu/anaconda3/envs/mxnet p36
mxnet_p36
python3
                         /home/ubuntu/anaconda3/envs/python3
pytorch latest p37
                         /home/ubuntu/anaconda3/envs/pytorch latest p37
pytorch_p36
                         /home/ubuntu/anaconda3/envs/pytorch_p36
tensorflow2 latest p37
                           /home/ubuntu/anaconda3/envs/tensorflow2 latest p37
tensorflow2 p36
                         /home/ubuntu/anaconda3/envs/tensorflow2 p36
tensorflow_p37
                         /home/ubuntu/anaconda3/envs/tensorflow_p37
help@ip-172-31-3-50:~$ conda activate
(base) help@ip-172-31-3-50:~¶ jupyter-notebook --ip=0.0.0.0 --no<u>-browser --port=9000</u>
```

데이터 분석을 위한 주피터노트북 사용 (3/4)

주피터 노트북 접속 URL 복사

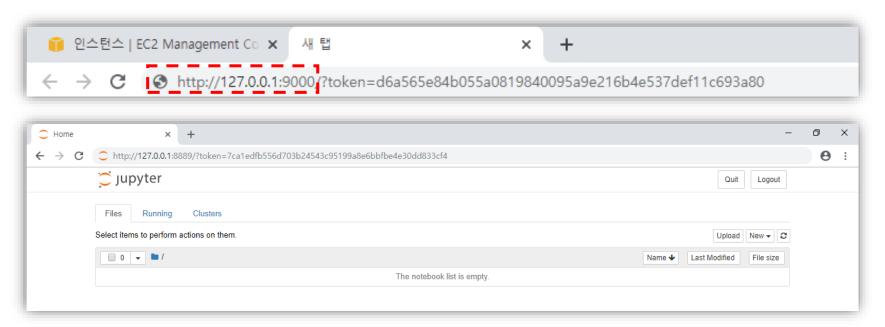
■ http://127.0.0.1로 시작하는 URL주소 드래그 & 마우스 우클릭(복사기능)

```
№ help@ip-172-31-3-50: ~
[W 2021-07-14 17:54:46.906 LabApp] 'port' has moved from NotebookApp to ServerApp. This con
fig will be passed to ServerApp. Be sure to update your config before our next release.
[W 2021-07-14 17:54:46.906 LabApp] 'port' has moved from NotebookApp to ServerApp. This con
fig will be passed to ServerApp. Be sure to update your config before our next release.
I 2021-07-14 17:54:46.938 LabApp JupyterLab extension loaded from /home/ubuntu/anaconda3/
lib/python3.7/site-packages/jupyterlab
[I 2021-07-14 17:54:46.938 LabApp] JupyterLab application directory is /home/ubuntu/anacond
a3/share/jupyter/lab
[I 17:54:46.943 NotebookApp] Serving notebooks from local directory: /home/help
[I 17:54:46.943 NotebookApp] Jupyter Notebook 6.2.0 is running at:
I 17:54:46.943 NotebookApp http://ip-172-31-3-50:9000/?token=f5928d36f7bde923ff3fa0d61785
a3870b19835b2db894a3
[I 17:54:46.943 NotebookApp]
                              or http://127.0.0.1:9000/?token=f5928d36f7bde923ff3fa0d61785a
3870b19835b2db894a3
[I 17:54:46.943 NotebookApp] Use Control-C to stop this server and shut down all kernels (t
wice to skip confirmation).
[C 17:54:46.948 NotebookApp]
   To access the notebook, open this file in a browser:
        file:///home/help/.local/share/jupyter/runtime/nbserver-1927-open.html
    Or copy and paste one of these URLs:
       http://ip-172-31-3-50:9000/?token=f5928d36f7bde923ff3fa0d61785a3870b19835b2db894l3
     or http://127.0.0.1:9000/?token=f5928d36f7bde923ff3fa0d61785a3870b19835b2db894a3
```

데이터 분석을 위한 주피터노트북 사용 (4/4)

개인 로컬 PC 웹브라우저에서 원격 서버에 있는 주피터 노트북 실행

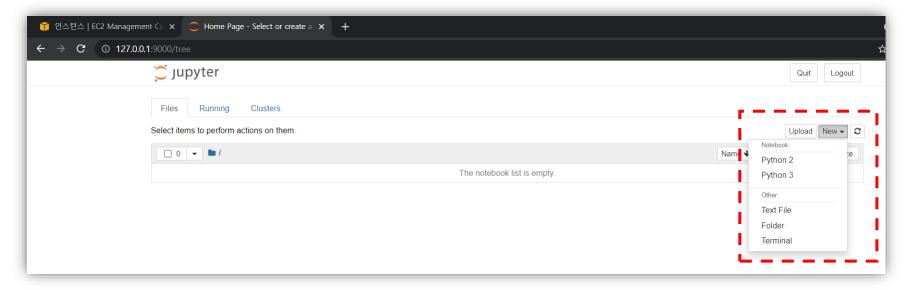
- (주의) 개인 로컬 PC에서 이미 실행중인 주피터노트북 종료 후 진행
- URL 주소는 127.0.0.1부터 시작해 달라지는 port번호에 유의(8889~9000)



참고2 . 주피터노트북 활용

개인 로컬 PC와 동일한 방식으로 주피터노트북 사용 가능

- Upload : 예제코드 및 데이터 파일 일괄 업로드 가능(폴더는 따로 생성)
- New : 사용 라이브러리 관련 주피터 노트북 파일(.ipynb) 및 폴더 생성
- 업로드한 파일의 저장 위치는 서버에 위치한 각 개인 홈 폴더(/home/사용자명)



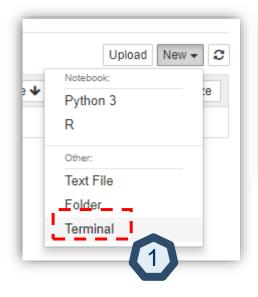
참고2 . 주피터노트북 활용

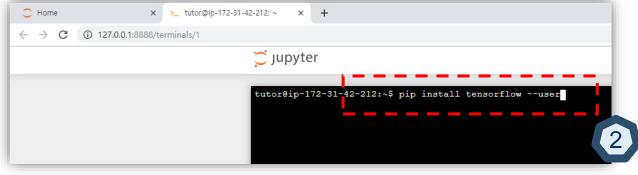
(base) 환경에서 원하는 패키지는 터미널 창을 통해 각자 설치

- New -> Terminal 을 통해 기본 라이브러리 이외의 패키지 설치
- Terminal에서 pip3 install tensorflow --user 등의 명령어를 통해 설치

(주의) 다른 사용자도 영향을 받기에 반드시 --user 옵션을 붙여 줘야함

설치 후에도 적용이 안될 시 커널 restart

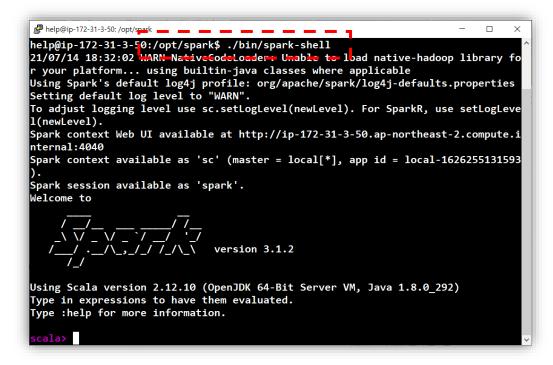




데이터 처리를 위한 SPARK 사용 (1/2)

Spark 설치 위치 및 실행방법

- SPARK_HOME=/opt/spark
- SPARK 설치 위치 이동을 위한 명령어 : cd /opt/spark/
- 원격 서버환경에서 spark-shell 실행 명령어 : ./bin/spark-shell



데이터 처리를 위한 SPARK 사용 (2/2)

(base)아나콘다 환경에서 Python3.x 버전의 pySpark 실행

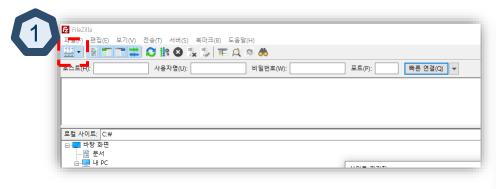
- conda activate
- pyspark 명령어 입력 시 python3.x버전의 pySpark 바로 사용 가능

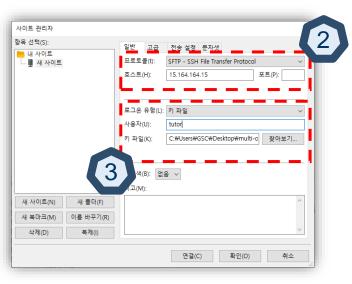
```
help@ip-172-31-3-50:~$ conda activate
(base) help@ip-172-31-3-50:~$ pyspark
Python 3.7.10 | packaged_by_conda_fonge | (default, Feb 19 2021, 16:07:37)
[GCC 9.3.0] on linux
Type "help", "copyright", "credits" or "license" for more information.
21/07/14 18:33:21 WARN NativeCodeLoader: Unable to load native-hadoop library fo
 your platform... using builtin-java classes where applicable
Using Spark's default log4j profile: org/apache/spark/log4j-<u>defaults.properties</u>
Setting default log level to "WARN".
To adjust logging level use sc.setLogLevel(newLevel). For SparkR, use setLogLeve
1(newLevel).
Welcome to
Using Python version 3.7.10 (default, Feb 19 2021 16:07:37)
Spark context Web UI available at http://ip-172-31-3-50.ap-northeast-2.compute.i
nternal:4040
Spark context available as 'sc' (master = local[*], app id = local-1626255203349
SparkSession available as 'spark'.
```

참고3. 원격 서버로 파일 전송

로컬 PC에서 원격 서버로 파일을 전송하는 방법(window 기준)

- FileZilla Client 프로그램 다운로드
 - : https://filezilla-project.org/download.php?platform=win64
- 프로그램 실행 후 아래 순서대로 진행





- 위 버튼 클릭 후 사이트 관리자 추가
- 프로토콜 선택 후 putty설정과 유사하게 호스트, 로그온 유형 진행
- 사용자에 맞게 사용자명은 변경해서 접속 후 파일 전송