

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 실행 (1/8)

아래 링크로 들어가 안내 드린 사용자 정보로 접속 ※ 비밀번호 변경 권장, 분실 시 문의

■ <https://multi-k3.signin.aws.amazon.com/console>

■ 계정 ID(12자리) 또는 계정 별칭 : multi-k3 [공통]



The first screenshot shows the AWS IAM login page. It features the AWS logo at the top left. Below it, the text 'IAM 사용자로 로그인' (Log in with IAM user) is displayed. A blue hexagonal icon with the number '1' is positioned to the left of the input fields. The input fields are labeled '계정 ID(12자리) 또는 계정 별칭' (Account ID (12 digits) or account alias), '사용자 이름:' (User name:), and '암호:' (Password:). The first field contains 'multi-k3', the second contains 'gscampus', and the third contains a masked password. A blue '로그인' (Log in) button is at the bottom. Below the button, there are links for '루트 사용자 이메일을 사용하여 로그인' (Log in with root user email) and '암호 찾기' (Forgot password).

aws

IAM 사용자로 로그인

계정 ID(12자리) 또는 계정 별칭

multi-k3

사용자 이름:

gscampus

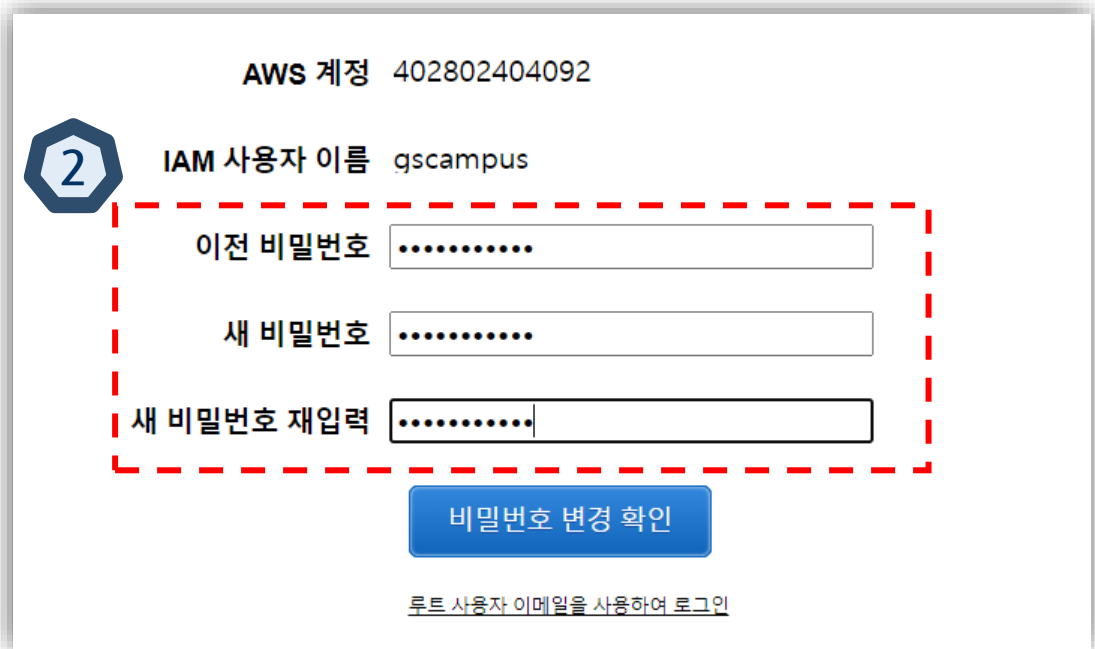
암호:

.....

로그인

루트 사용자 이메일을 사용하여 로그인

암호 찾기



The second screenshot shows the AWS IAM password change page. It features the text 'AWS 계정 402802404092' at the top. Below it, the text 'IAM 사용자 이름 gscampus' is displayed. A blue hexagonal icon with the number '2' is positioned to the left of the input fields. The input fields are labeled '이전 비밀번호' (Previous password), '새 비밀번호' (New password), and '새 비밀번호 재입력' (Re-enter new password). Each field contains a masked password. A blue '비밀번호 변경 확인' (Confirm password change) button is at the bottom. Below the button, there is a link for '루트 사용자 이메일을 사용하여 로그인' (Log in with root user email).

AWS 계정 402802404092

IAM 사용자 이름 gscampus

이전 비밀번호

.....

새 비밀번호

.....

새 비밀번호 재입력

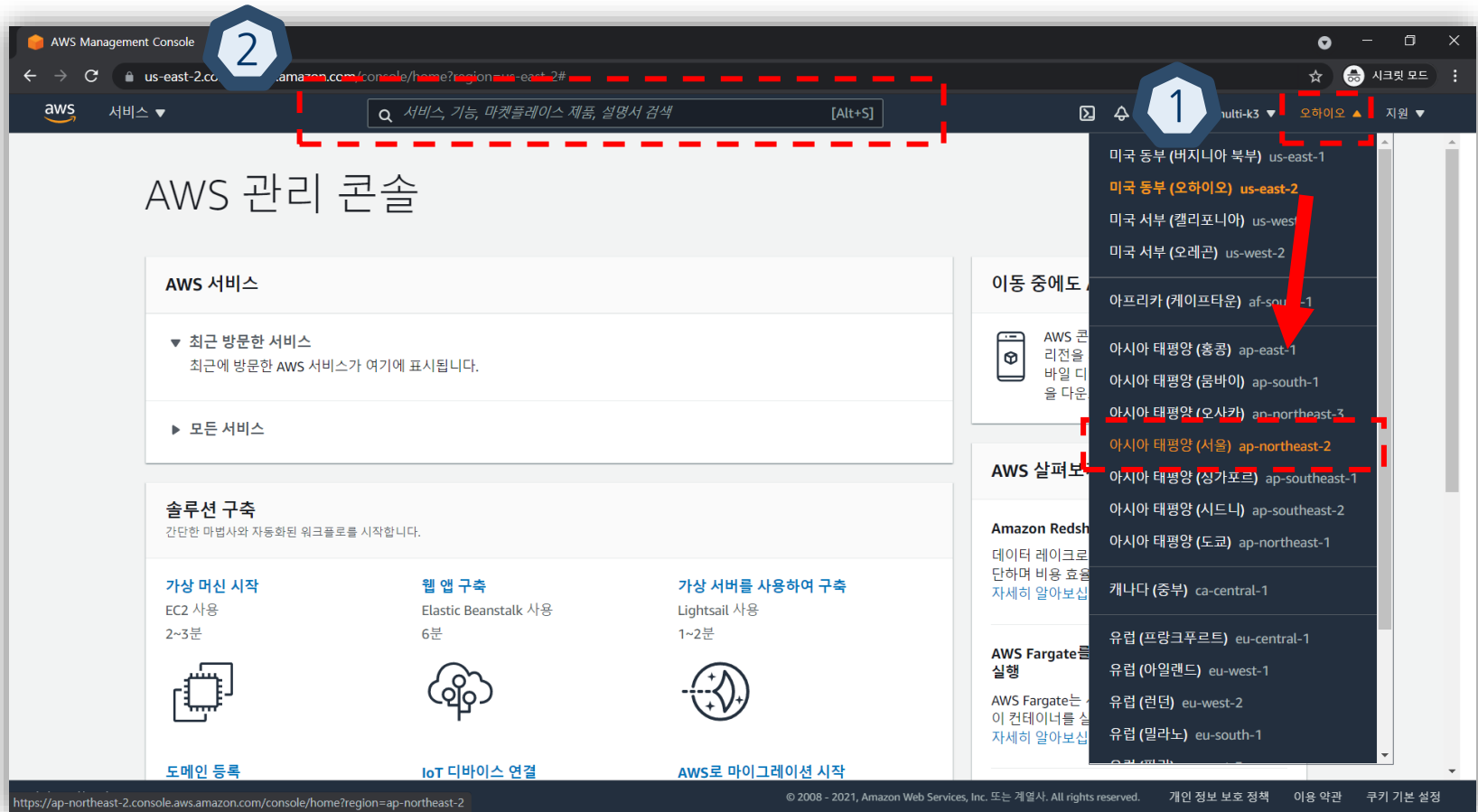
.....

비밀번호 변경 확인

루트 사용자 이메일을 사용하여 로그인

실습용 서버 실행 (2/8)

우측 상단 리전 정보를 서울로 변경 후 서비스 검색창에서 EC2 검색



실습용 서버 실행 (3/8)

서비스창에서 EC2(클라우드의 가상 서버) 클릭



실습용 서버 실행 (4/8)

현재 사용 가능한 서버 정보 확인 ※ 사용자별 배정받은 서버의 최소 사용 권한만 부여

■ 인스턴스 클릭 ※ 주의 : 인스턴스(실행중) 아님

The screenshot displays the AWS Management Console for the 'ap-northeast-2' region. The main content area shows the '리소스' (Resources) section, which lists various AWS EC2 resources. A red dashed box highlights the '인스턴스' (Instances) row, which has a count of 29. A red arrow points from the left sidebar to this row. The left sidebar shows the '인스턴스' menu item selected. The right sidebar shows the '계정 속성' (Account Properties) section, which includes VPC, EBS, and IAM settings. The bottom of the console shows the '인스턴스 시작' (Start Instance) button and the '서비스 상태' (Service Status) section, which indicates that the service is operational.

리소스	개수
인스턴스(실행 중)	2
배치 그룹	0
볼륨	29
인스턴스	29
키 페어	24
로드 밸런서	0
보안 그룹	2
스냅샷	6
전용 호스트	0
탄력적 IP	29

인스턴스 시작

서비스 상태

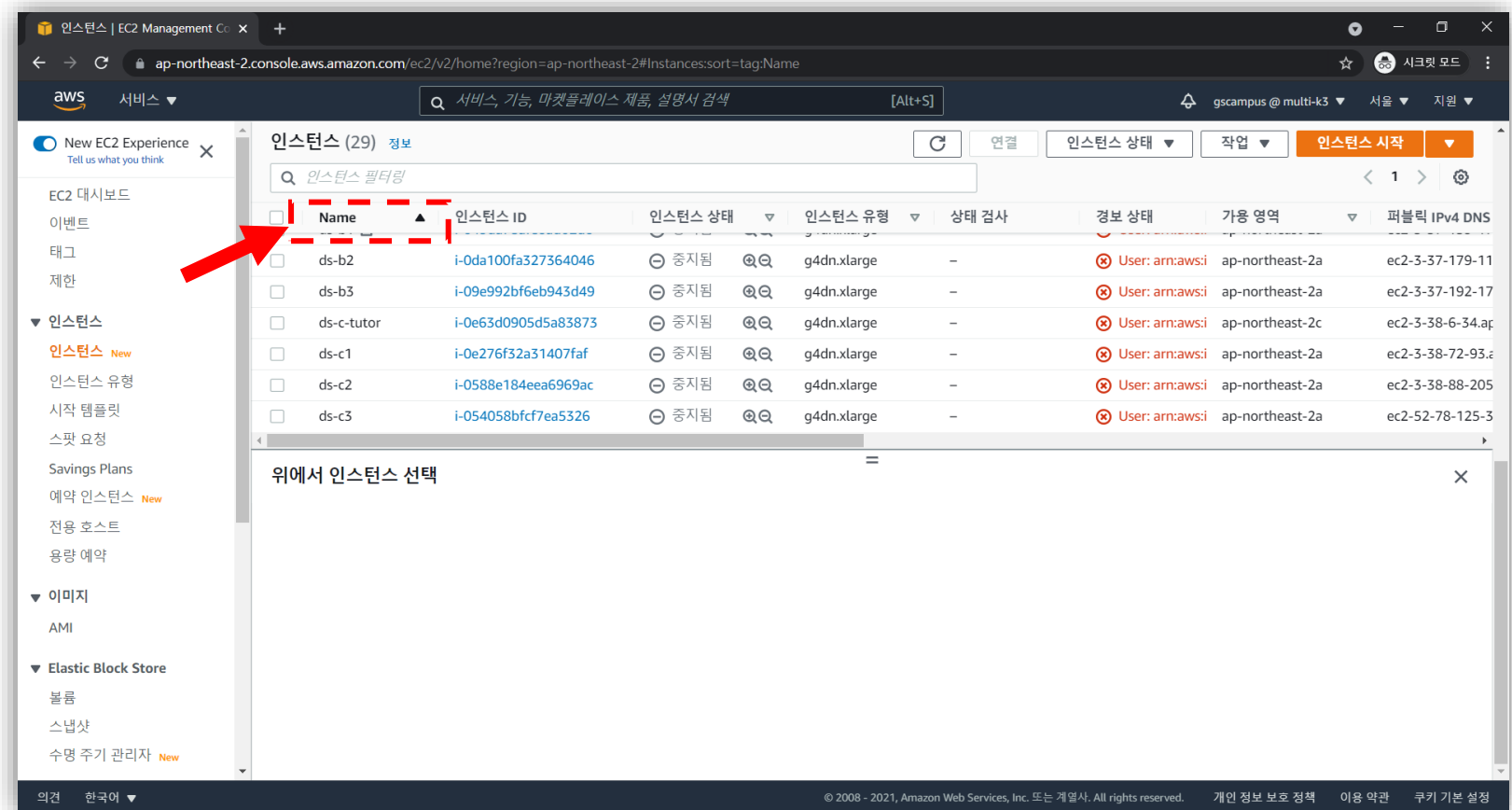
리전: 아시아 태평양 (서울)

상태: 이 서비스가 정상적으로 작동 중입니다.

실습용 서버 실행 (5/8)

현재 사용 가능한 서버 정보 확인

■ 이름(Name)순으로 정렬 후 할당된 서버 선택



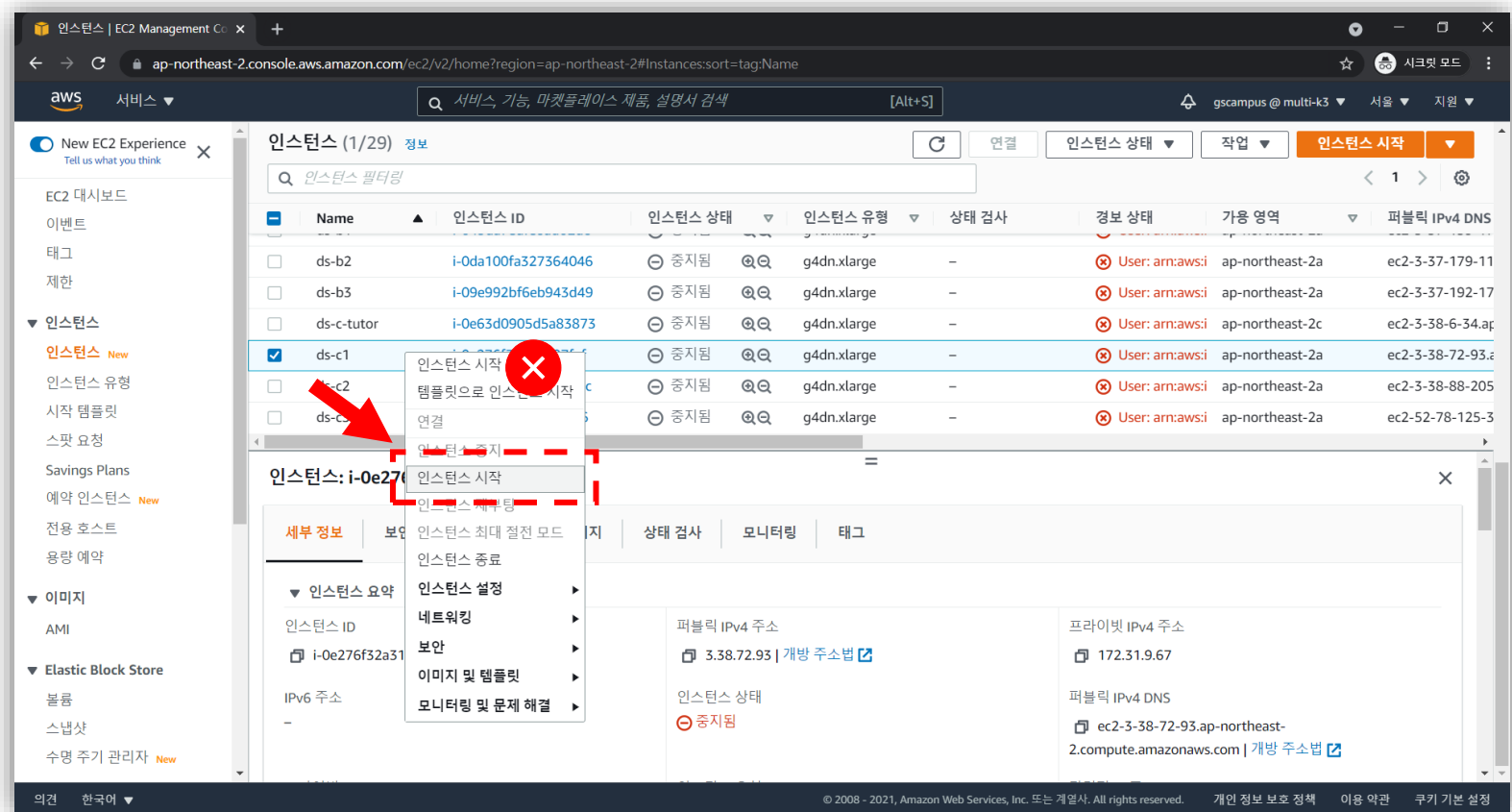
The screenshot shows the AWS Management Console for the 'ap-northeast-2' region. The 'Instances' page is displayed, showing a list of 29 instances. The instances are sorted by Name. A red arrow points to the 'Name' column header, indicating the sorting criteria. The table lists instances with names like 'ds-b2', 'ds-b3', 'ds-c-tutor', 'ds-c1', 'ds-c2', and 'ds-c3'. Each row includes columns for Name, Instance ID, State, Instance Type, State Checks, Health Status, Availability Zone, and Public IPv4 DNS.

Name	인스턴스 ID	인스턴스 상태	인스턴스 유형	상태 검사	경보 상태	가용 영역	퍼블릭 IPv4 DNS
ds-b2	i-0da100fa327364046	중지됨	g4dn.xlarge	-	User: arn:aws:iam::123456789012:user:awscli	ap-northeast-2a	ec2-3-37-179-11
ds-b3	i-09e992bf6eb943d49	중지됨	g4dn.xlarge	-	User: arn:aws:iam::123456789012:user:awscli	ap-northeast-2a	ec2-3-37-192-17
ds-c-tutor	i-0e63d0905d5a83873	중지됨	g4dn.xlarge	-	User: arn:aws:iam::123456789012:user:awscli	ap-northeast-2c	ec2-3-38-6-34.ap
ds-c1	i-0e276f32a31407faf	중지됨	g4dn.xlarge	-	User: arn:aws:iam::123456789012:user:awscli	ap-northeast-2a	ec2-3-38-72-93.2
ds-c2	i-0588e184eea6969ac	중지됨	g4dn.xlarge	-	User: arn:aws:iam::123456789012:user:awscli	ap-northeast-2a	ec2-3-38-88-205
ds-c3	i-054058bfcf7ea5326	중지됨	g4dn.xlarge	-	User: arn:aws:iam::123456789012:user:awscli	ap-northeast-2a	ec2-52-78-125-3

실습용 서버 실행 (6/8)

실습용 서버 동작 확인 및 서버 실행 ※ 사용자별로 서버 실행권한 별도부여

■ 사용할 서버 위에서 마우스 우클릭, 아래쪽 인스턴스 시작버튼 클릭 ※ 위쪽 아님



실습용 서버 실행 (7/8)

실습용 서버 동작 확인 및 서버 실행

■ 인스턴스 실행 확인을 위해 새로 고침 버튼 클릭 ※ 초기화 상태에서도 접속 가능

New EC2 Experience

인스턴스 (1/29) 정보

새로 고침 연결 인스턴스 상태 작업 인스턴스 시작

* 1분 후 새로 고침 클릭

Name	인스턴스 ID	인스턴스 상태	인스턴스 유형	상태 검사	경보 상태	가용 영역	퍼블릭 IPv4 DNS
ds-b-tutor	i-0f8a16c45fe1f55d5	중지됨	g4dn.xlarge	-	User: arn:aws:iam::111111111111:user	ap-northeast-2a	ec2-3-36-31-136
ds-b1	i-049da78afe6ad02d6	중지됨	g4dn.xlarge	-	User: arn:aws:iam::111111111111:user	ap-northeast-2a	ec2-3-37-159-17
ds-b2	i-0da100fa327364046	중지됨	g4dn.xlarge	-	User: arn:aws:iam::111111111111:user	ap-northeast-2a	ec2-3-37-179-11
ds-b3	i-09e992bf6eb943d49	중지됨	g4dn.xlarge	-	User: arn:aws:iam::111111111111:user	ap-northeast-2a	ec2-3-37-192-17
ds-c-tutor	i-0e63d0905d5a83873	중지됨	g4dn.xlarge	-	User: arn:aws:iam::111111111111:user	ap-northeast-2c	ec2-3-38-6-34.ap
ds-c1	i-0e276f32a31407faf	대기 중	g4dn.xlarge	-	User: arn:aws:iam::111111111111:user	ap-northeast-2a	ec2-3-38-72-93.a
ds-c2	i-0588e184eea6969ac	중지됨	g4dn.xlarge	-	User: arn:aws:iam::111111111111:user	ap-northeast-2a	ec2-3-38-88-205

New EC2 Experience

인스턴스 (29) 정보

새로 고침 연결 인스턴스 상태 작업 인스턴스 시작

Name	인스턴스 ID	인스턴스 상태	인스턴스 유형	상태 검사	경보 상태	가용 영역	퍼블릭 IPv4 DNS
ds-b1	i-0f8a16c45fe1f55d5	중지됨	g4dn.xlarge	-	User: arn:aws:iam::111111111111:user	ap-northeast-2a	ec2-3-37-159-17
ds-b2	i-0da100fa327364046	중지됨	g4dn.xlarge	-	User: arn:aws:iam::111111111111:user	ap-northeast-2a	ec2-3-37-179-11
ds-b3	i-09e992bf6eb943d49	중지됨	g4dn.xlarge	-	User: arn:aws:iam::111111111111:user	ap-northeast-2a	ec2-3-37-192-17
ds-c-tutor	i-0e63d0905d5a83873	중지됨	g4dn.xlarge	-	User: arn:aws:iam::111111111111:user	ap-northeast-2c	ec2-3-38-6-34.ap
ds-c1	i-0e276f32a31407faf	실행 중	g4dn.xlarge	-	User: arn:aws:iam::111111111111:user	ap-northeast-2a	ec2-3-38-72-93.a
ds-c2	i-0588e184eea6969ac	중지됨	g4dn.xlarge	-	User: arn:aws:iam::111111111111:user	ap-northeast-2a	ec2-3-38-88-205
ds-c3	i-054058bfcf7ea5326	중지됨	g4dn.xlarge	-	User: arn:aws:iam::111111111111:user	ap-northeast-2a	ec2-52-78-125-3

실습용 서버 실행 (8/8)

실습용 서버 동작 확인 및 서버 실행

- 인스턴스 최종 실행 확인을 위해 새로 고침 재 클릭 ※ 2/2개 검사 통과 시 서버 완전 실행

인스턴스 (1/29) 정보

Name	인스턴스 ID	인스턴스 상태	인스턴스 유형	상태 검사	경보 상태	가용 영역	퍼블릭 IPv4 DNS
ds-b2	i-0da100fa327364046	중지됨	g4dn.xlarge	-	User: arn:aws:iam::123456789012:user:root	ap-northeast-2a	ec2-3-37-179-11
ds-b3	i-09e992bf6eb943d49	중지됨	g4dn.xlarge	-	User: arn:aws:iam::123456789012:user:root	ap-northeast-2a	ec2-3-37-192-17
ds-c-tutor	i-0e63d0905d5a83873	중지됨	g4dn.xlarge	-	User: arn:aws:iam::123456789012:user:root	ap-northeast-2c	ec2-3-38-6-34.ap
ds-c1	i-0e276f32a31407faf	실행 중	g4dn.xlarge	2/2개 검사 통과...	User: arn:aws:iam::123456789012:user:root	ap-northeast-2a	ec2-3-38-72-93.a
ds-c2	i-0588e184eea6969ac	중지됨	g4dn.xlarge	-	User: arn:aws:iam::123456789012:user:root	ap-northeast-2a	ec2-3-38-88-205
ds-c3	i-054058bfcf7ea5326	중지됨	g4dn.xlarge	-	User: arn:aws:iam::123456789012:user:root	ap-northeast-2a	ec2-52-78-125-3

인스턴스: i-0e276f32a31407faf(ds-c1)

세부 정보 | 보안 | 네트워킹 | 스토리지 | 상태 검사 | 모니터링 | 태그

인스턴스 요약 정보

인스턴스 ID	퍼블릭 IPv4 주소	프라이빗 IPv4 주소
i-0e276f32a31407faf (ds-c1)	3.38.72.93 개방 주소법	172.31.9.67

IPv6 주소	인스턴스 상태	퍼블릭 IPv4 DNS
-	실행 중	ec2-3-38-72-93.ap-northeast-2.compute.amazonaws.com 개방 주소법

원격 접속 서버 정보 (사이언스)

■ 서버 사양 :

※ HW

- NVIDIA T4 Tensor 코어 급 GPU(단일 GPU VM)
- 1vGPU / 4vCPU / 16GB RAM / 16GB GPU 급 RAM / 1000GB SSD
- Ubuntu 18.04 OS / NVIDIA CUDA, cuDNN 既설치

※ SW

- Anaconda3 / python 3.7.6
- spark-3.1.2-bin-hadoop3.2(단일 머신) 설치

■ 서버 운영시간 : (평일) ~18:30(18시 30분 이후 서버 자동 종료 및 계정 잠금)

(주말) 추석 연휴, 멘토링 일정 없는 토요일과 일요일은 사용 불가

원격 접속 서버 정보 (엔지니어링)

■ 서버 사양 :

※ HW

- 고주파수 인텔 제온 E5-2686 v4(Broadwell)급 프로세서, DDR4 메모리
- 4vCPU / 30.5GB 급 RAM / 1000GB SSD
- Ubuntu 18.04 OS

※ SW

- Anaconda3 / python 3.7.6
- spark-3.1.2-bin-hadoop3.2(단일 머신) 설치

■ 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 (1/4)

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



- d*~** .ppk 암호화 키 보관
- www.putty.org 사이트를 통해 PuTTY 64-bit 다운로드



Download PuTTY

PuTTY is an SSH and telnet client, developed originally by Simon Tatham for the Windows platform. PuTTY is open source software that is available with source code and is developed and supported by a group of volunteers.

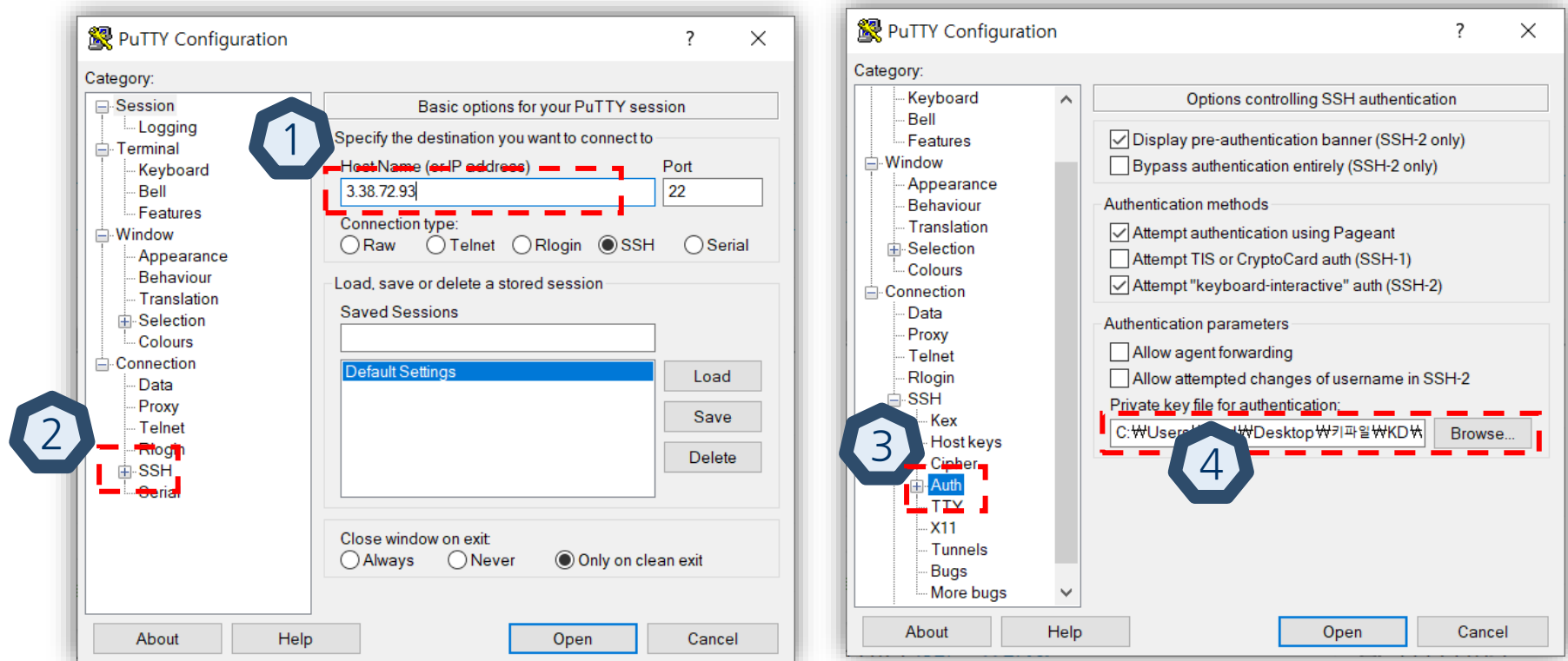
You can download PuTTY [here](http://www.putty.org).

Below suggestions are independent of the authors of PuTTY. They are *not* to be seen as endorsements by the PuTTY project.

서버 접속을 위한 사전 준비 (2/4)

사용자별로 부여 받은 서버 IP 입력(1) 및 암호화 키 업로드(2~4)

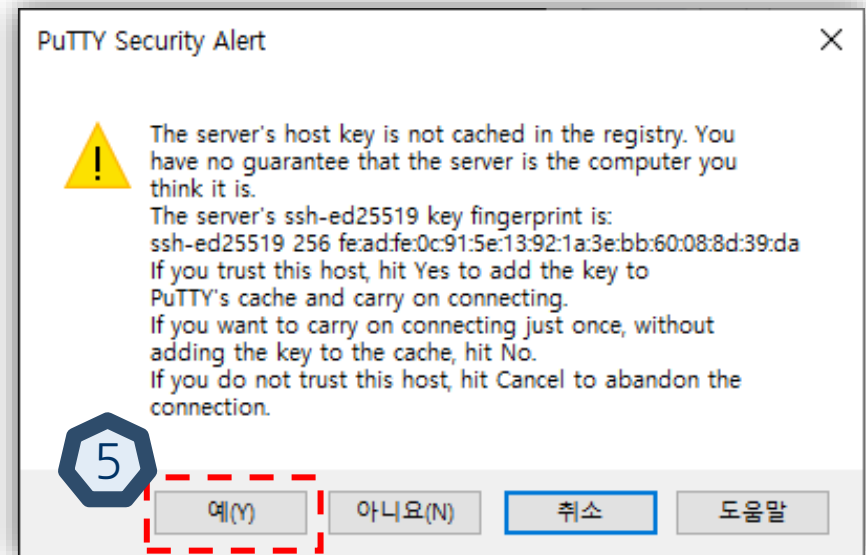
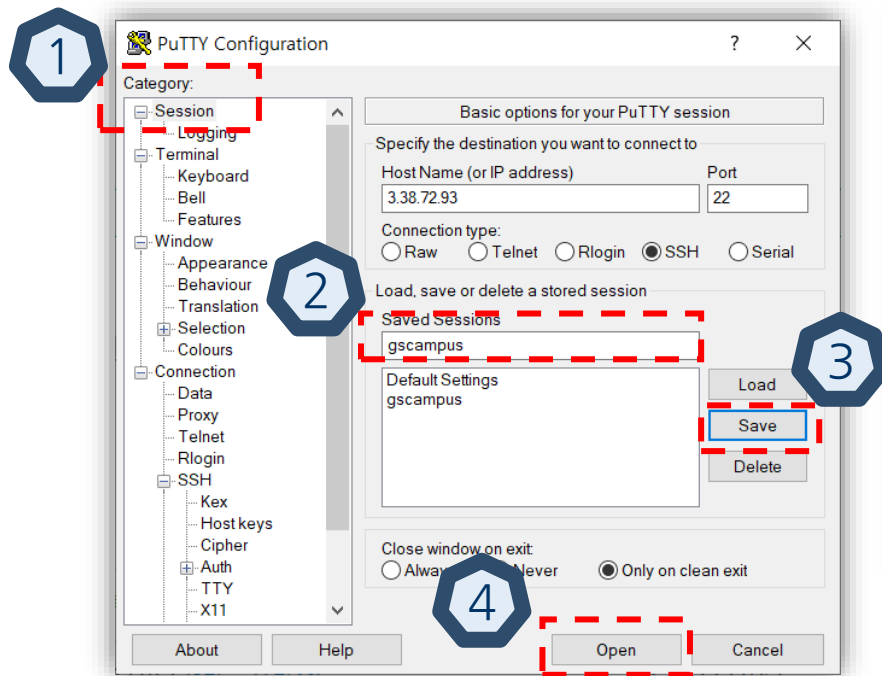
- HostName에 서버 IP 입력 후 좌측 SSH 메뉴 확장 후 Auth에서 암호화 키 업로드



서버 접속을 위한 사전 준비 (3/4)

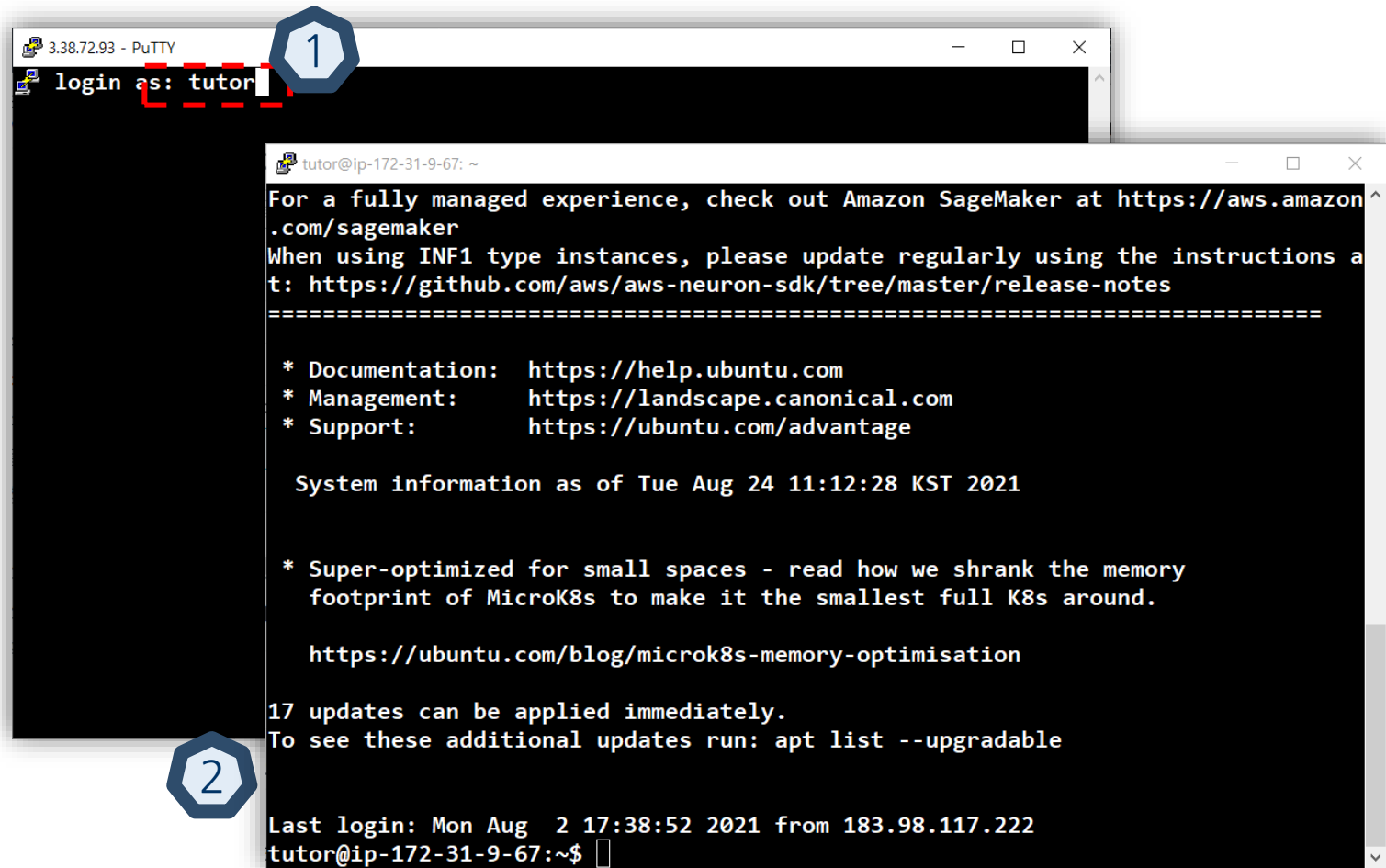
바로 Open 버튼을 누르지 말고 Session으로 돌아와 설정 내용을 저장

- 설정 내용을 저장해 놓으면 프로그램을 열 때마다 재설정할 필요 없음(1~3)
- Open 버튼 클릭 후 최초 접속 시 예(Y) 클릭(4~5)



서버 접속을 위한 사전 준비 (4/4)

부여 받은 ID를 입력해 서버에 접속(별도의 PW필요 없음)



```
3.38.72.93 - PuTTY
login as: tutor

tutor@ip-172-31-9-67: ~
For a fully managed experience, check out Amazon SageMaker at https://aws.amazon.com/sagemaker
When using INF1 type instances, please update regularly using the instructions at: https://github.com/aws/aws-neuron-sdk/tree/master/release-notes
=====
* Documentation: https://help.ubuntu.com
* Management: https://landscape.canonical.com
* Support: https://ubuntu.com/advantage

System information as of Tue Aug 24 11:12:28 KST 2021

* Super-optimized for small spaces - read how we shrank the memory footprint of MicroK8s to make it the smallest full K8s around.

https://ubuntu.com/blog/microk8s-memory-optimisation

17 updates can be applied immediately.
To see these additional updates run: apt list --upgradable

Last login: Mon Aug 2 17:38:52 2021 from 183.98.117.222
tutor@ip-172-31-9-67:~$
```

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 tutor@3.34.111.86 -p 22

맥북 내 암호키 저장 경로

부여 받은 ID@IP ssh접속을 위한 port 번호

※ 서버 최초 접속 후 Host key verification failed 오류 발생 시 키 값을 초기화

: \$ ssh-keygen -R 3.34.111.86(부여 받은 IP)

데이터 분석용 주피터노트북 사용 (1/4)

기 설치된 아나콘다 가상 환경 목록 검색 및 실행

■ `conda info --envs`

■ `conda activate [가상 환경 이름]` ※ **python3 사용 권장**

```
tutor@ip-172-31-9-67: ~  
17 updates can be applied immediately.  
To see these additional updates run: apt list --upgradable  
  
Last login: Mon Aug 2 17:38:52 2021 from 183.98.117.222  
tutor@ip-172-31-9-67:~$ conda info --envs  
# conda environments:  
#  
base * /home/ubuntu/anaconda3  
amazoni_mxnet_p36 /home/ubuntu/anaconda3/envs/amazoni_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  
mxnet_p36 /home/ubuntu/anaconda3/envs/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_p37 /home/ubuntu/anaconda3/envs/tensorflow2_p37  
tensorflow_p37 /home/ubuntu/anaconda3/envs/tensorflow_p37  
  
tutor@ip-172-31-9-67:~$
```

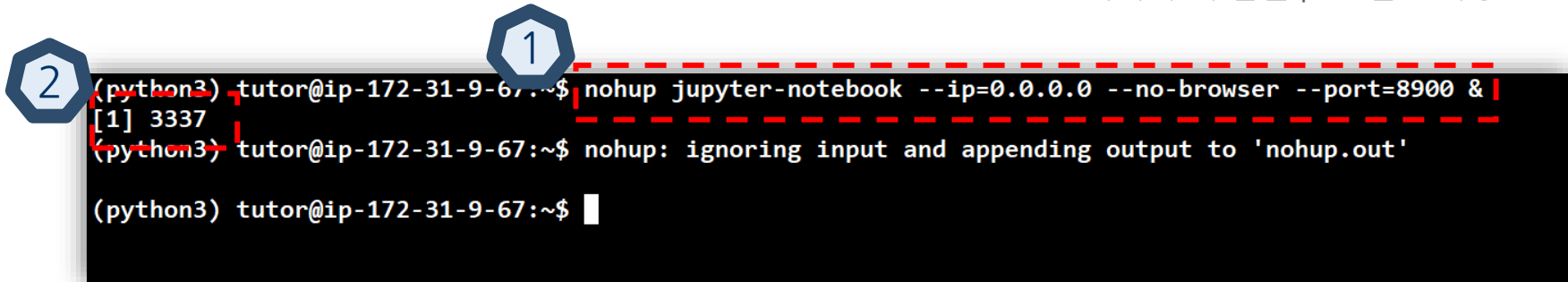
```
tutor@ip-172-31-9-67: ~  
To see these additional updates run: apt list --upgradable  
  
Last login: Mon Aug 2 17:38:52 2021 from 183.98.117.222  
tutor@ip-172-31-9-67:~$ conda info --envs  
# conda environments:  
#  
base * /home/ubuntu/anaconda3  
amazoni_mxnet_p36 /home/ubuntu/anaconda3/envs/amazoni_mxnet_p36  
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  
mxnet_p36 /home/ubuntu/anaconda3/envs/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_p37 /home/ubuntu/anaconda3/envs/tensorflow2_p37  
tensorflow_p37 /home/ubuntu/anaconda3/envs/tensorflow_p37  
  
tutor@ip-172-31-9-67:~$ conda activate python3  
(python3) tutor@ip-172-31-9-67:~$
```

데이터 분석용 주피터노트북 사용 (2/4)

주피터 노트북 사용을 위한 명령어 입력 ※ 반드시 포트번호 뒤 '&' 까지 작성

■ `nohup jupyter-notebook --ip=0.0.0.0 --no-browser --port=8xxx &`

각자 부여 받은 port번호 사용



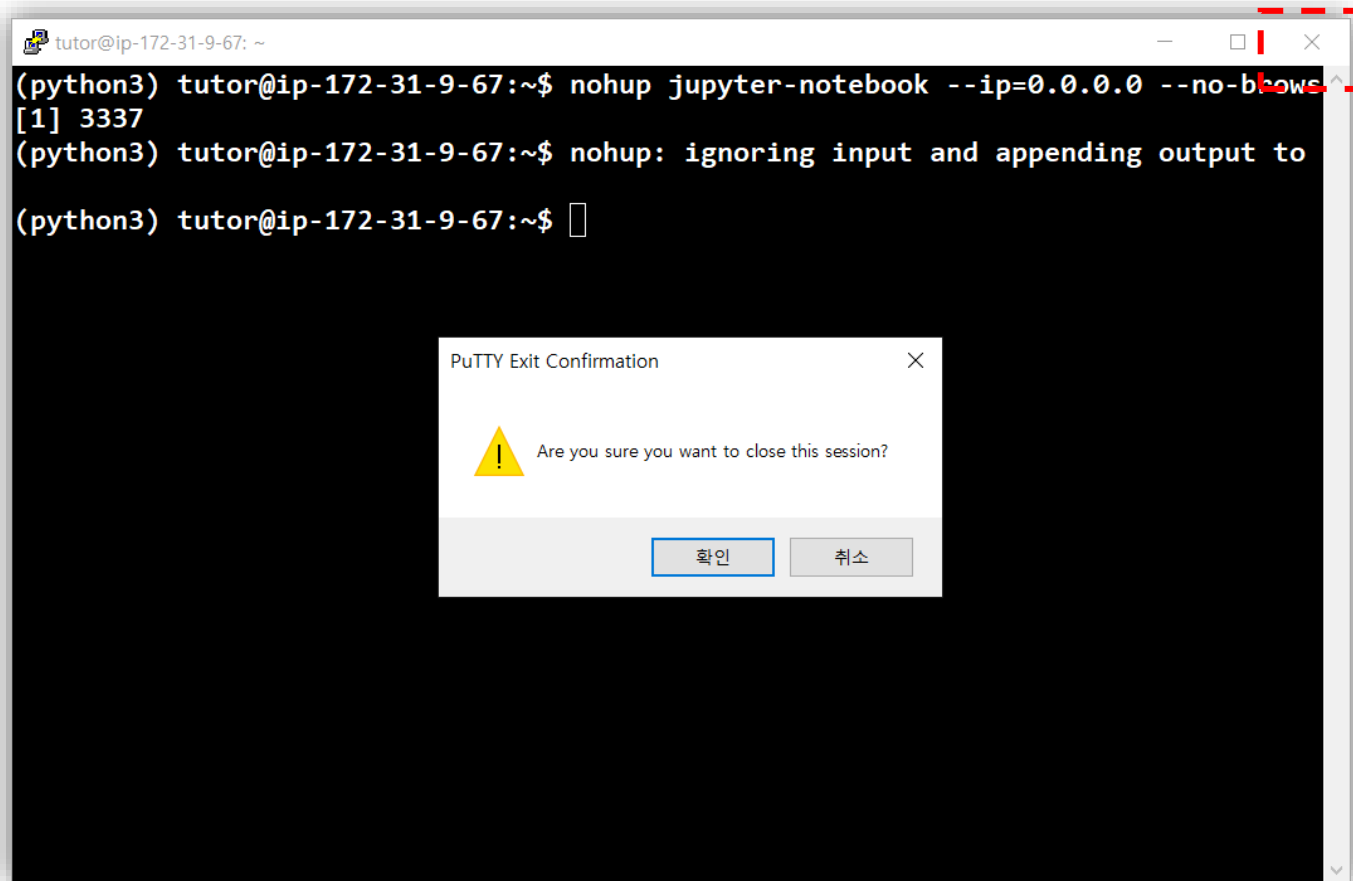
```
(python3) tutor@ip-172-31-9-67:~$ nohup jupyter-notebook --ip=0.0.0.0 --no-browser --port=8900 &
[1] 3337
(python3) tutor@ip-172-31-9-67:~$ nohup: ignoring input and appending output to 'nohup.out'

(python3) tutor@ip-172-31-9-67:~$
```

■ [1] 3337(번호는 다름) 와 같이 프로세스 번호가 나오면 정상 실행

데이터 분석용 주피터노트북 사용 (3/4)

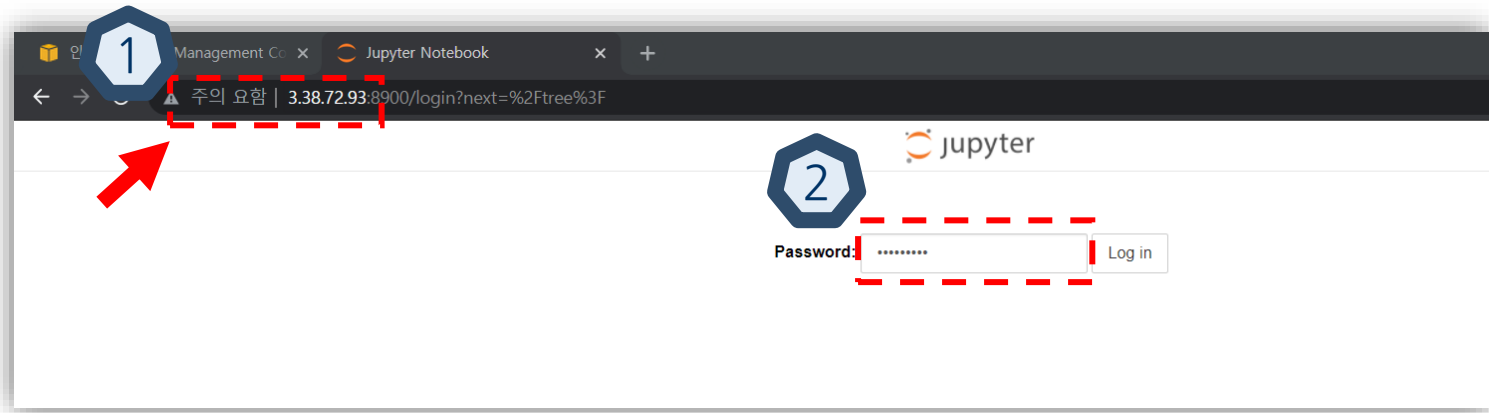
PuTTY를 종료하고 웹브라우저(크롬 등) 실행 ※ PuTTY를 종료해도 주피터 실행가능



데이터 분석용 주피터노트북 사용 (4/4)

웹 브라우저 창에서 부여 받은 IP와 포트번호 입력

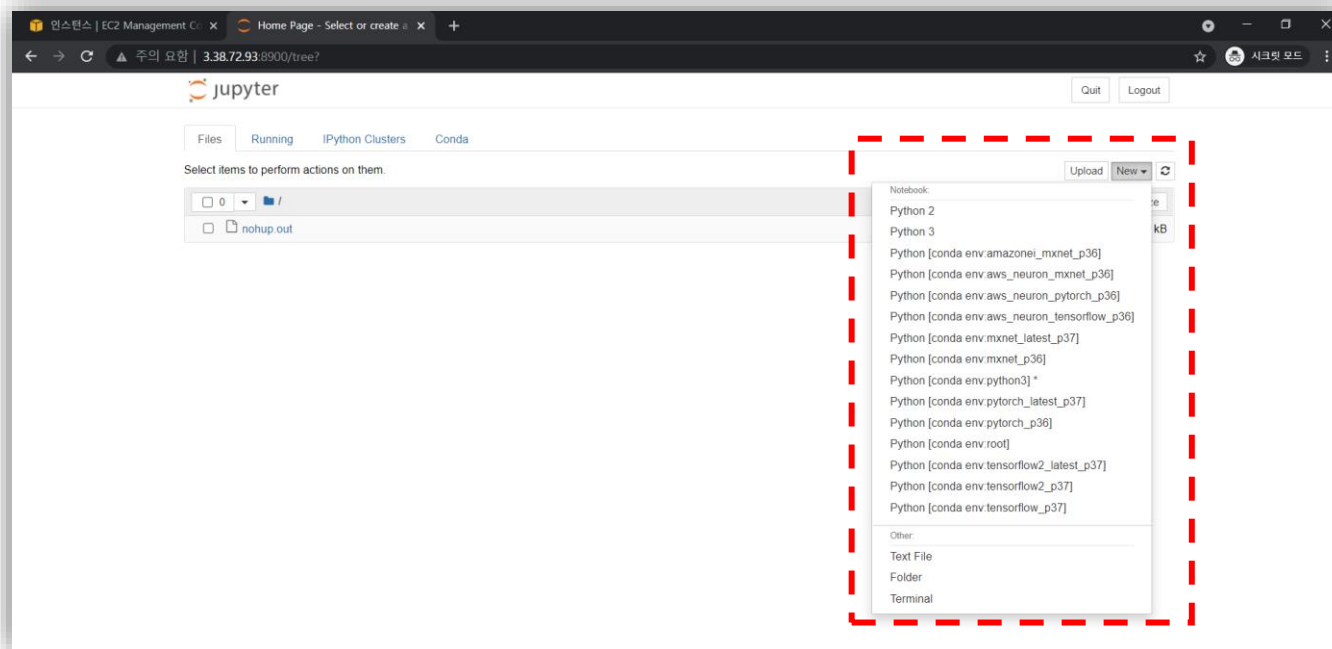
- `http://3.38.72.93:8900` ※ 포트번호는 89다음에 개인 PuTTY접속 번호와 동일 (예)lab01 : 8901
- (공통)Password : multi1234! ※ Password 변경 희망 시 문의



분석을 위한 주피터노트북 활용 (1/7)

개인 로컬 PC와 동일한 방식으로 주피터노트북 사용 가능

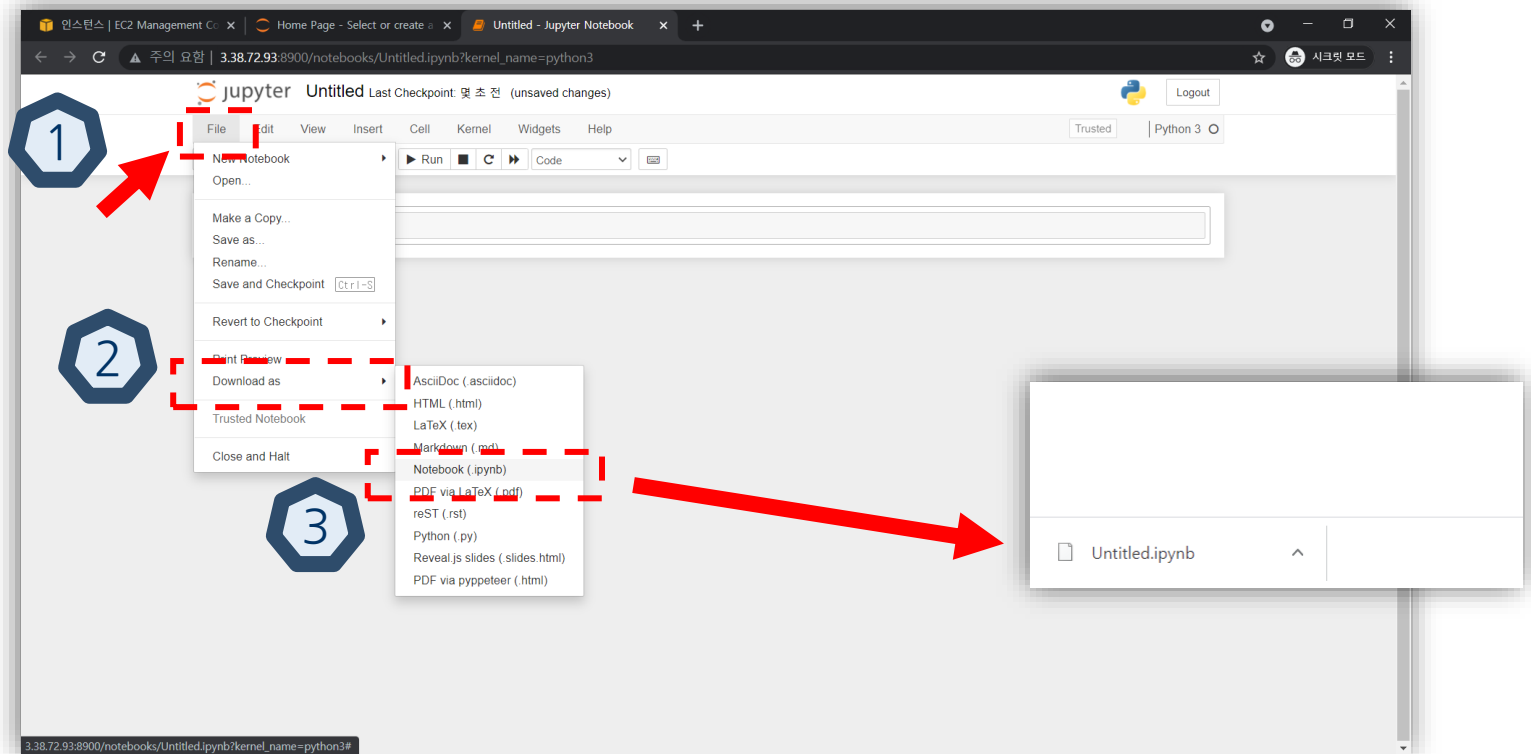
- Upload : 예제코드 및 데이터 파일 일괄 업로드 가능 ※ 폴더 업로드는 불가(따로 생성)
- New : 기 설정된 가상환경의 주피터 노트북 파일(.ipynb) 및 폴더 생성
- Python3 사용 권장



분석을 위한 주피터노트북 활용 (2/7)

작업한 파일을 로컬 PC로 다운로드 하는 방법

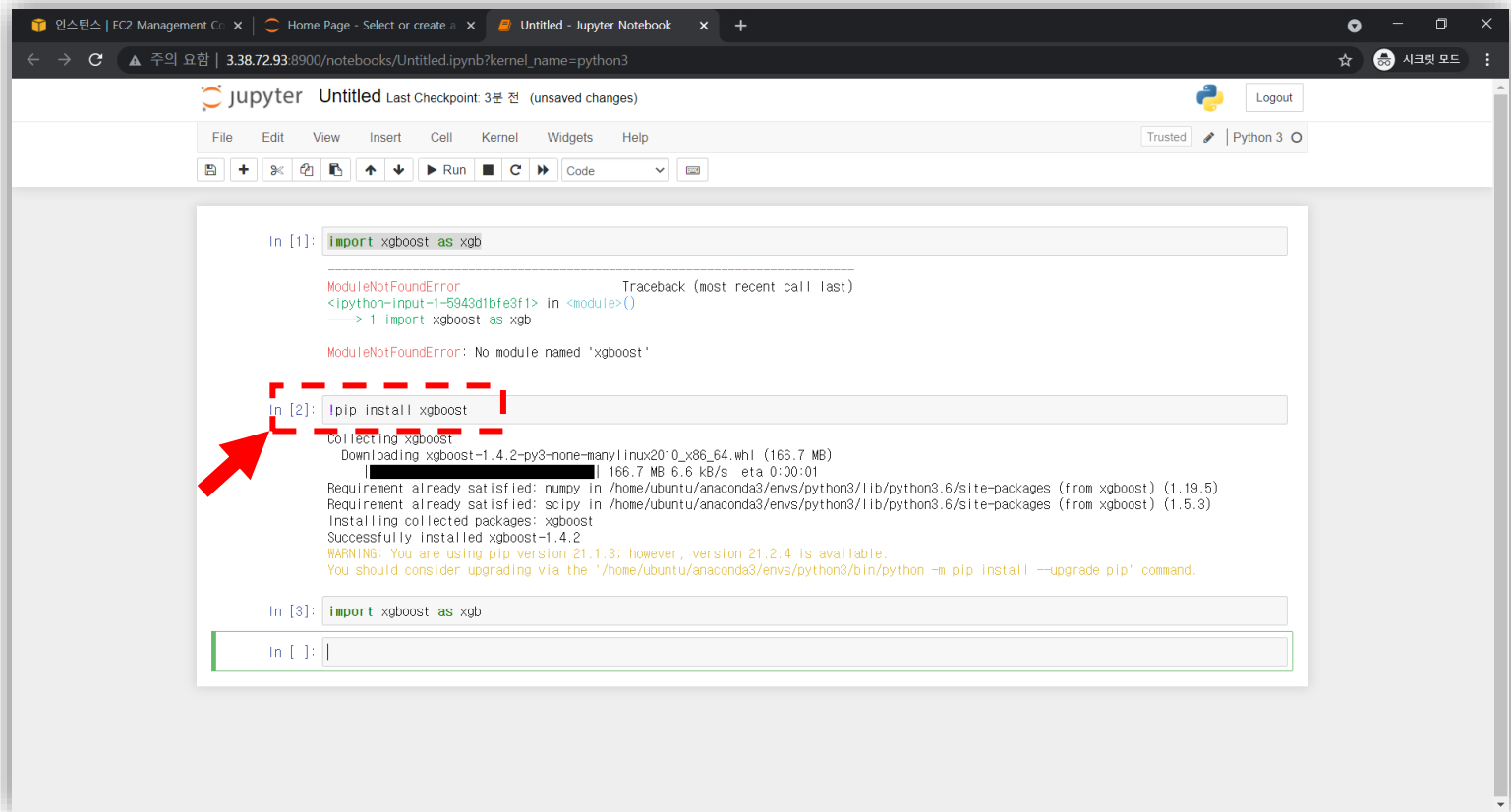
- 업로드한 파일의 저장 위치는 서버에 위치한 홈 폴더(/home/lab**)
- 우측 상단 File > Download as > Notebook(.ipynb) 클릭 시 아래와 같이 로컬 다운



분석을 위한 주피터노트북 활용 (3/7)

원하는 패키지는 각자 설치

■ !pip install [패키지명]



The screenshot shows a Jupyter Notebook window with the following content:

```
In [1]: import xgboost as xgb
```

ModuleNotFoundError: Traceback (most recent call last)
<ipython-input-1-5943d1bfe3f1> in <module>()
----> 1 import xgboost as xgb
ModuleNotFoundError: No module named 'xgboost'

In [2]: !pip install xgboost

Collecting xgboost
 Downloading xgboost-1.4.2-py3-none-manylinux2010_x86_64.whl (166.7 MB)
 |#####| 166.7 MB 6.6 kB/s eta 0:00:01
Requirement already satisfied: numpy in /home/ubuntu/anaconda3/envs/python3/lib/python3.6/site-packages (from xgboost) (1.19.5)
Requirement already satisfied: scipy in /home/ubuntu/anaconda3/envs/python3/lib/python3.6/site-packages (from xgboost) (1.5.3)
Installing collected packages: xgboost
Successfully installed xgboost-1.4.2
WARNING: You are using pip version 21.1.3; however, version 21.2.4 is available.
You should consider upgrading via the '/home/ubuntu/anaconda3/envs/python3/bin/python -m pip install --upgrade pip' command.

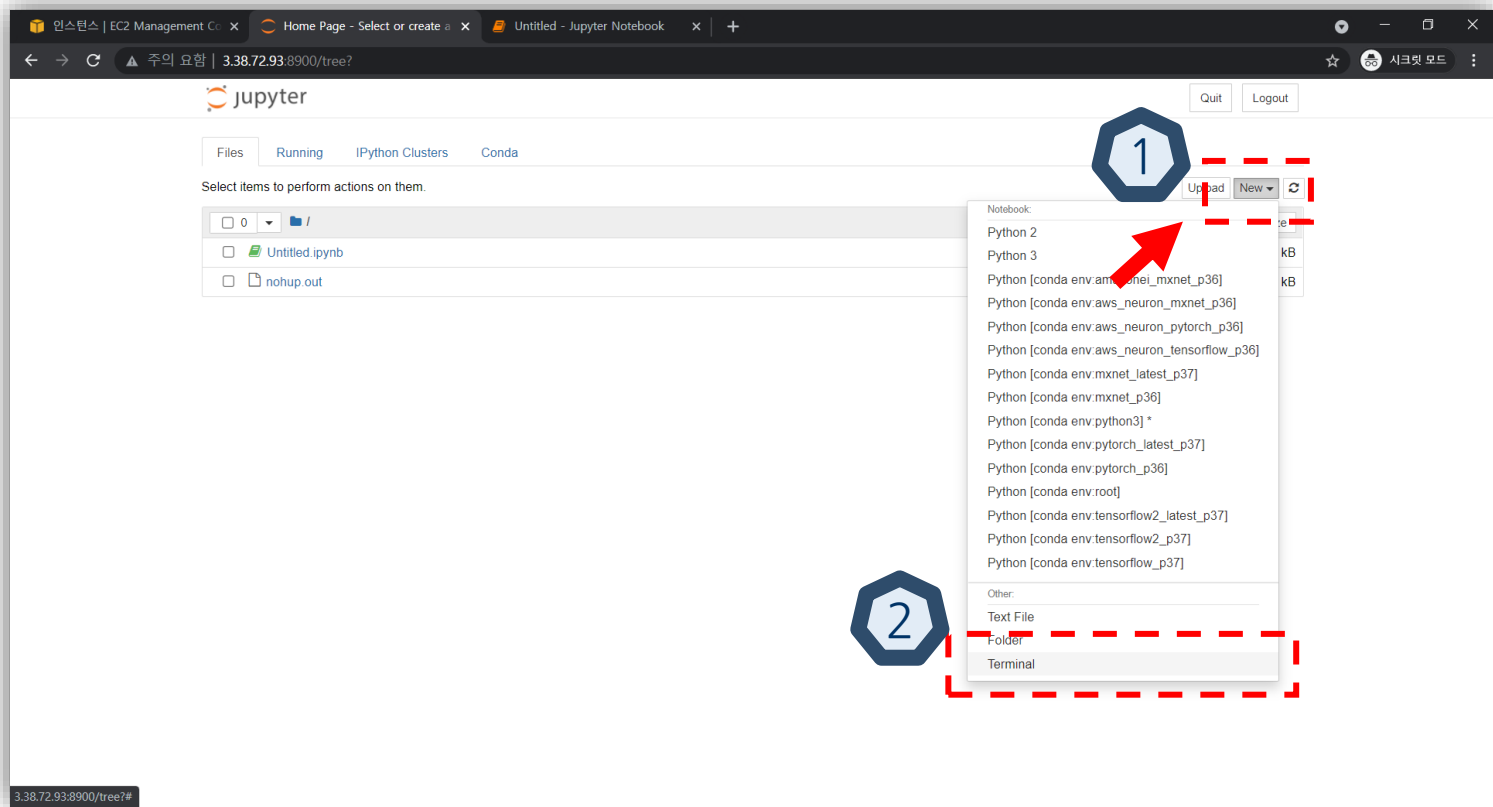
```
In [3]: import xgboost as xgb
```

```
In [ ]: |
```

분석을 위한 주피터노트북 활용 (4/7)

아나콘다 가상 환경 개인적으로 생성 가능 ※ 다소 시간 소요

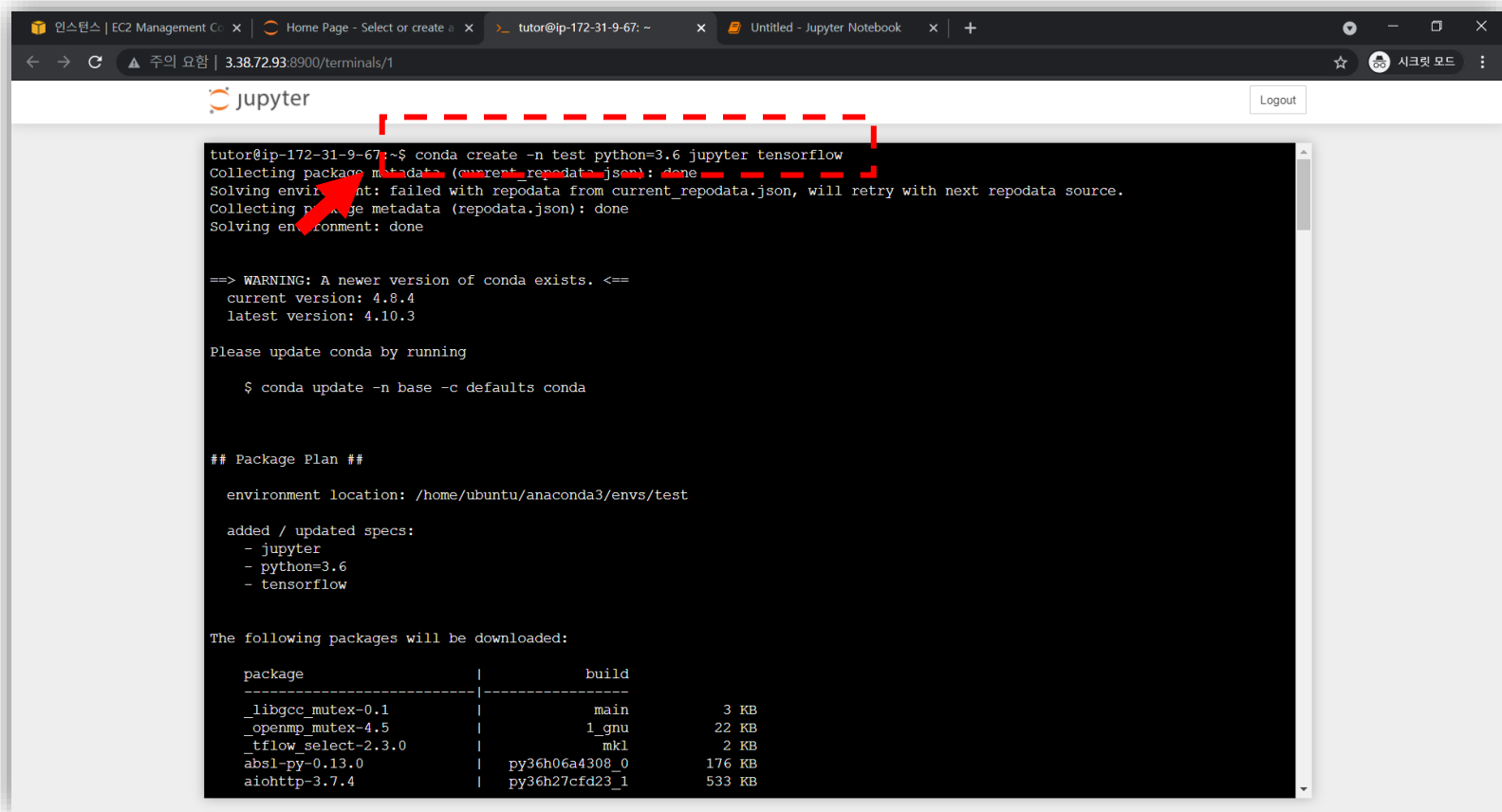
■ New > Terminal 클릭



분석을 위한 주피터노트북 활용 (5/7)

아나콘다 가상 환경 개인적으로 생성 가능

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



```
tutor@ip-172-31-9-67:~$ 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
- python=3.6
- tensorflow

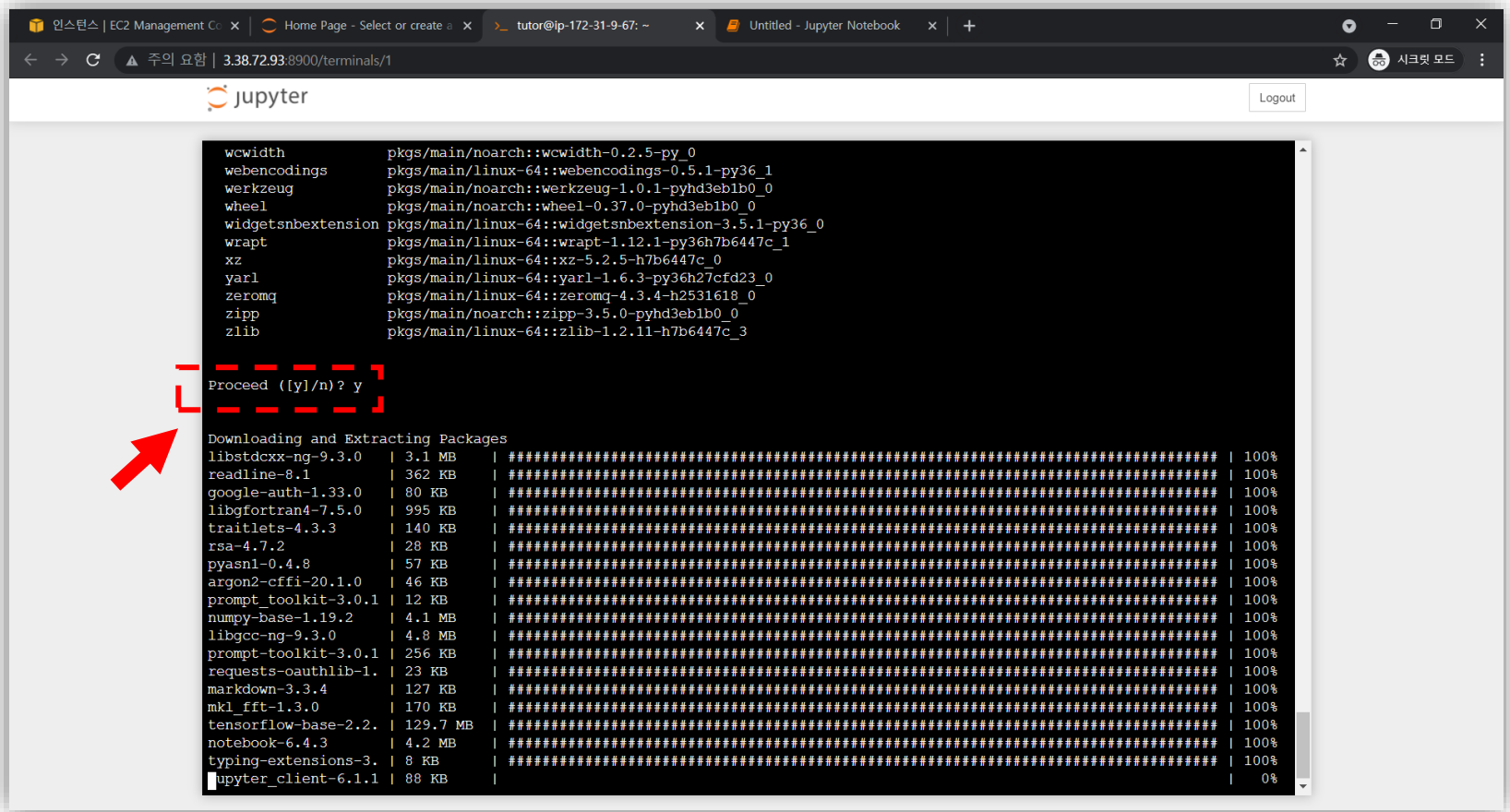
The following packages will be downloaded:
```

package	build	size
_libgcc_mutex-0.1	main	3 KB
_openmp_mutex-4.5	1_gnu	22 KB
tfselect-2.3.0	mk1	2 KB
absl-py-0.13.0	py36h06a4308_0	176 KB
aiohttp-3.7.4	py36h27cfd23_1	533 KB

분석을 위한 주피터노트북 활용 (6/7)

아나콘다 가상 환경 개인적으로 생성 가능(다소 시간 소요)

■ Proceed ([y]/n) ? y 입력



```
insights | EC2 Management Co x Home Page - Select or create a x tutor@ip-172-31-9-67: ~ x Untitled - Jupyter Notebook x +
< -> C 주의 요약 | 3.38.72.93:8900/terminals/1
jupyter Logout

wcwidth pkgs/main/noarch::wcwidth-0.2.5-py_0
webencodings pkgs/main/linux-64::webencodings-0.5.1-py36_1
werkzeug pkgs/main/noarch::werkzeug-1.0.1-pyhd3eb1b0_0
wheel pkgs/main/noarch::wheel-0.37.0-pyhd3eb1b0_0
widgetsnbextension pkgs/main/linux-64::widgetsnbextension-3.5.1-py36_0
wrapt pkgs/main/linux-64::wrapt-1.12.1-py36h7b6447c_1
xz pkgs/main/linux-64::xz-5.2.5-h7b6447c_0
yaml pkgs/main/linux-64::yaml-1.6.3-py36h27cfd23_0
zeromq pkgs/main/linux-64::zeromq-4.3.4-h2531618_0
zipp pkgs/main/noarch::zipp-3.5.0-pyhd3eb1b0_0
zlib pkgs/main/linux-64::zlib-1.2.11-h7b6447c_3

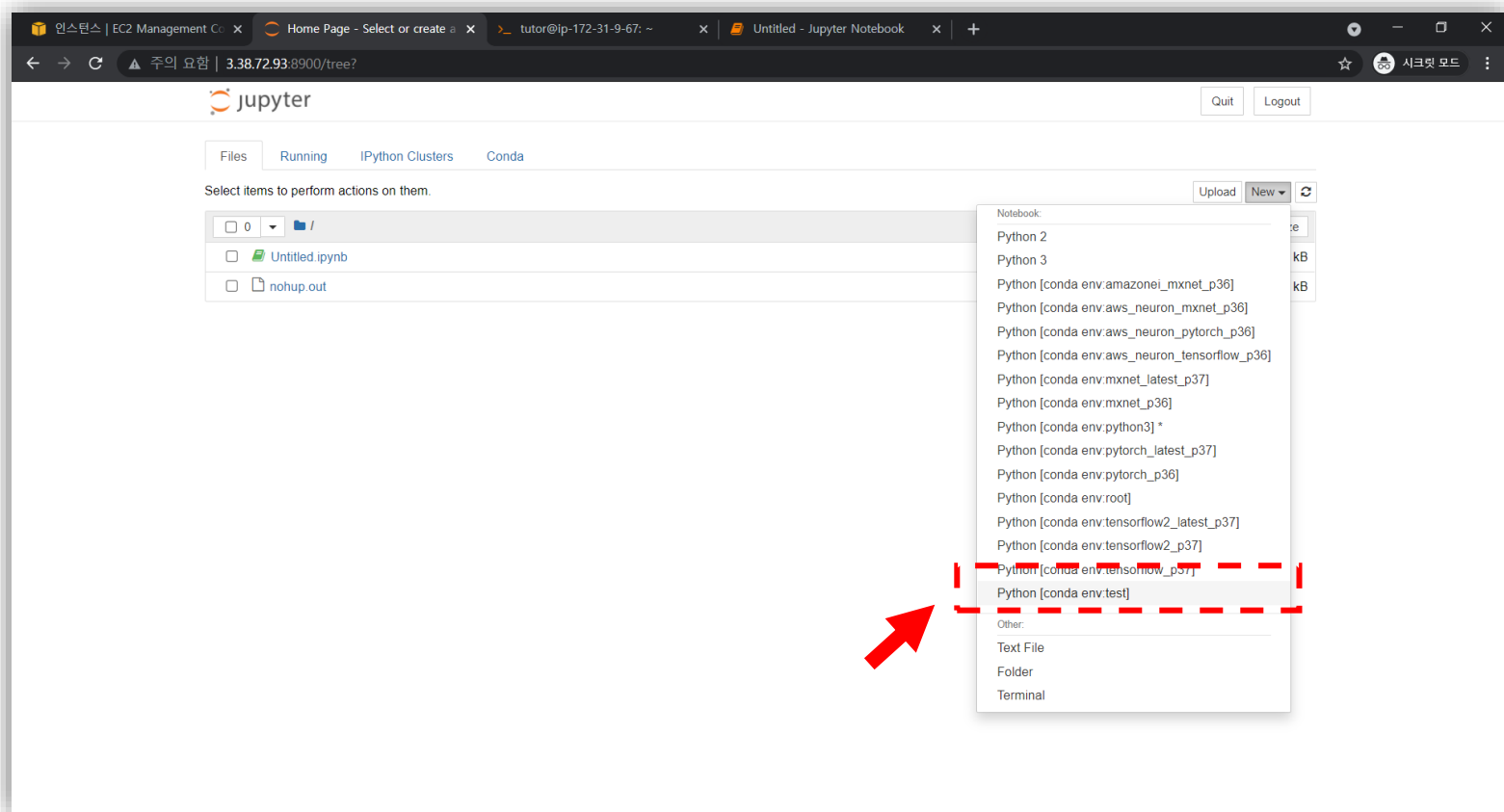
Proceed ([y]/n) ? y

Downloading and Extracting Packages
libstdcxx-ng-9.3.0 | 3.1 MB | ##### | 100%
readline-8.1 | 362 KB | ##### | 100%
google-auth-1.33.0 | 80 KB | ##### | 100%
libgfortran4-7.5.0 | 995 KB | ##### | 100%
traitlets-4.3.3 | 140 KB | ##### | 100%
rsa-4.7.2 | 28 KB | ##### | 100%
pyasn1-0.4.8 | 57 KB | ##### | 100%
argon2-cffi-20.1.0 | 46 KB | ##### | 100%
prompt_toolkit-3.0.1 | 12 KB | ##### | 100%
numpy-base-1.19.2 | 4.1 MB | ##### | 100%
libgcc-ng-9.3.0 | 4.8 MB | ##### | 100%
prompt-toolkit-3.0.1 | 256 KB | ##### | 100%
requests-oauthlib-1. | 23 KB | ##### | 100%
markdown-3.3.4 | 127 KB | ##### | 100%
mkl_fft-1.3.0 | 170 KB | ##### | 100%
tensorflow-base-2.2. | 129.7 MB | ##### | 100%
notebook-6.4.3 | 4.2 MB | ##### | 100%
typing-extensions-3. | 8 KB | ##### | 100%
jupyter_client-6.1.1 | 88 KB | | 0%
```

분석을 위한 주피터노트북 활용 (7/7)

기존에 열어놓은 주피터노트북 창 화면 새로고침(F5) 후 New 버튼 클릭

■ 생성한 가상환경으로 실습 진행 가능



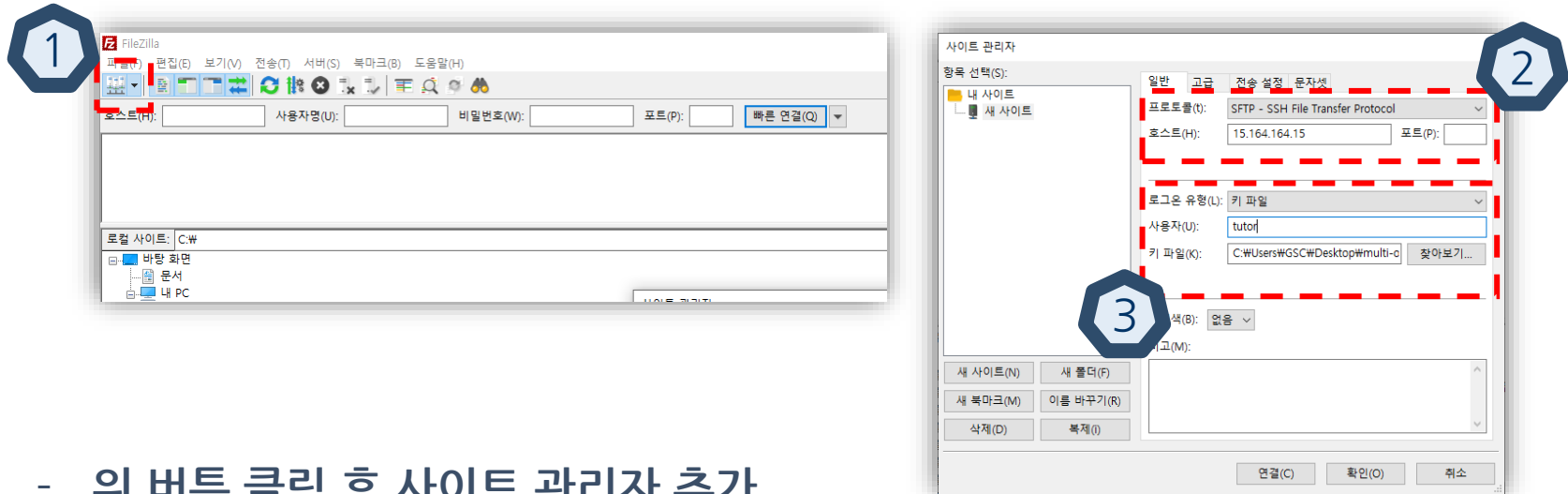
참고2. 원격 서버로 파일 전송

로컬 PC에서 원격 서버로 파일을 전송하는 방법(윈도우, Mac 모두 가능)

■ FileZilla Client 프로그램 다운로드

: <https://filezilla-project.org/download.php?platform=win64>

■ 프로그램 실행 후 아래 순서대로 진행

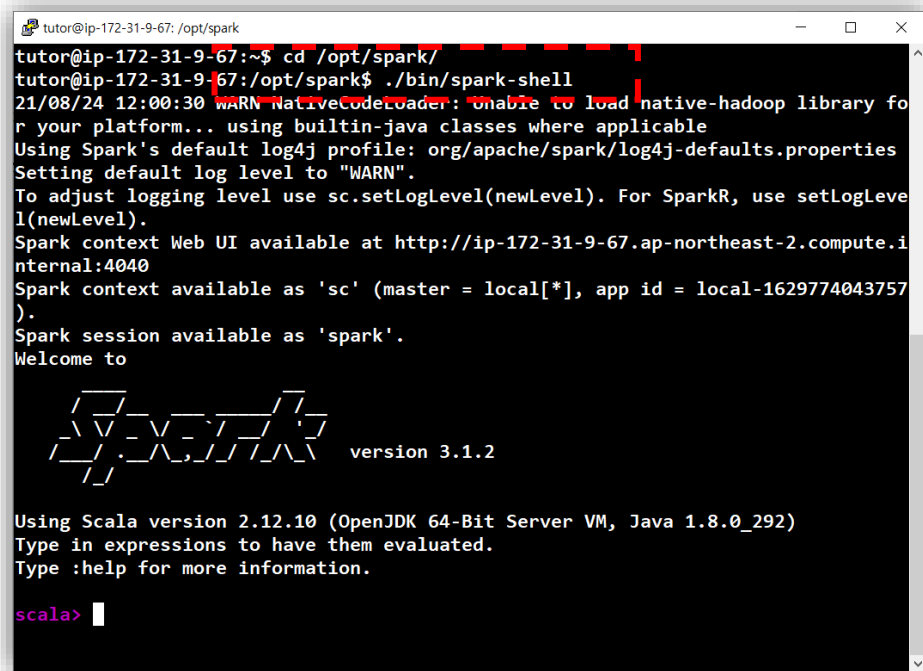


- 위 버튼 클릭 후 사이트 관리자 추가
- 프로토콜(SFTP) 선택 후 호스트에 IP주소 입력, 로그인 유형은 키파일로 선택
- 사용자에게 맞게 사용자명은 변경해서 키파일(.ppk or .pem) 업로드 후 파일 전송

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`



```
tutor@ip-172-31-9-67: /opt/spark
tutor@ip-172-31-9-67:~$ cd /opt/spark/
tutor@ip-172-31-9-67:/opt/spark$ ./bin/spark-shell
21/08/24 12:00:30 WARN NativeCodeLoader: Unable to load native-hadoop library for
your platform... using builtin-java classes where applicable
Using Spark's default log4j profile: org/apache/spark/log4j-defaults.properties
Setting default log level to "WARN".
To adjust logging level use sc.setLogLevel(newLevel). For SparkR, use setLogLevel(newLevel).
Spark context Web UI available at http://ip-172-31-9-67.ap-northeast-2.compute.i
nternal:4040
Spark context available as 'sc' (master = local[*], app id = local-1629774043757
).
Spark session available as 'spark'.
Welcome to

  ____
 /  __ \
/   /  \
/_____/

version 3.1.2

Using Scala version 2.12.10 (OpenJDK 64-Bit Server VM, Java 1.8.0_292)
Type in expressions to have them evaluated.
Type :help for more information.

scala> |
```

데이터 처리를 위한 SPARK 사용 (2/2)

Python3.x 버전의 pySpark 실행

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

```
tutor@ip-172-31-9-67: /opt/spark$ pyspark
Python 3.6.9 (default, Jan 26 2021, 15:33:00)
[GCC 8.4.0] on linux
Type "help", "copyright", "credits" or "license" for more information.
21/08/24 12:02:25 WARN NativeCodeLoader: Unable to load native-hadoop library fo
r your platform... using builtin-java classes where applicable
Using Spark's default log4j profile: org/apache/spark/log4j-defaults.properties
Setting default log level to "WARN".
To adjust logging level use sc.setLogLevel(newLevel). For SparkR, use setLogLeve
l(newLevel).
Welcome to

      _/_/_/_/_/_/_/_/_/_/_/_/_/_/_/_
     /_/_/_/_/_/_/_/_/_/_/_/_/_/_/_\
    /_/_/_/_/_/_/_/_/_/_/_/_/_/_\_/_
   /_/_/_/_/_/_/_/_/_/_/_/_/_/_\_/_\
  /_/_/_/_/_/_/_/_/_/_/_/_/_/_\_/__\
 /_/_/_/_/_/_/_/_/_/_/_/_/_/_\_/_/_\
/_/_/_/_/_/_/_/_/_/_/_/_/_/_\_/_/_/_\

version 3.1.2

Using Python version 3.6.9 (default, Jan 26 2021 15:33:00)
Spark context Web UI available at http://ip-172-31-9-67.ap-northeast-2.compute.i
nternal:4040
Spark context available as 'sc' (master = local[*], app id = local-1629774147692
).
SparkSession available as 'spark'.
>>> █
```