

「 포털 리뷰 기반 영화 추천 시스템

필름 이너사이드

개발환경 구축 가이드

ver1.0

」

빅데이터 분석 기반 AI 커머스 서비스 개발자 양성과정

2팀 - 인사이드아웃

고정원, 고내리, 권지민, 이호제

변경 이력

No.	변경 내역	작성		검토	
		일자	작성자	일자	담당자
1	초안 작성(ver0.1)	2022.10.26	고내리	2022.10.27	이진영 강사
2	DB 테이블 구조 추가(ver0.2)	2022.10.27	고내리	2022.10.28	이진영 강사
3	최종확인(ver1.0)	2022.10.28	고내리	2022.10.28	이진영 강사
4					
5					
6					

Contents

프로그램 설치

1.1 파이썬	4p
1.2 Visual Studio Code	8p
1.3 Java JDK	12p
1.4 MySQL	14p
1.5 CUDA / Cudnn	24p

환경설정

2.1 VS Code 확장프로그램	31p
2.2 Java 환경변수	34p
2.3 MySQL 환경변수	37p
2.4 MySQL Data импорт	38p
2.5 DB 테이블 내용확인	41p

프로젝트 구조

3.1 프로젝트 구조	45p
-------------	-------	-----

프로젝트 실행

4.1 VS Code 인터프리터 선택	47p
4.2 실행 메뉴 확인	48p

1.1 파이썬 ver3.10.6

프로그램 설치

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

[설치 파일] Environment/python-3.10.6-amd64.exe

Looking for a specific release?

Python releases by version number:

Release version	Release date		Click for more
Python 3.7.15	Oct. 11, 2022	Download	Release Notes
Python 3.7.14	Sept. 6, 2022	Download	Release Notes
Python 3.8.14	Sept. 6, 2022	Download	Release Notes
Python 3.9.14	Sept. 6, 2022	Download	Release Notes
Python 3.10.7	Sept. 6, 2022	Download	Release Notes
Python 3.10.6	Aug. 2, 2022	Download	Release Notes
Python 3.10.5	June 6, 2022	Download	Release Notes
-			

[View older releases](#)

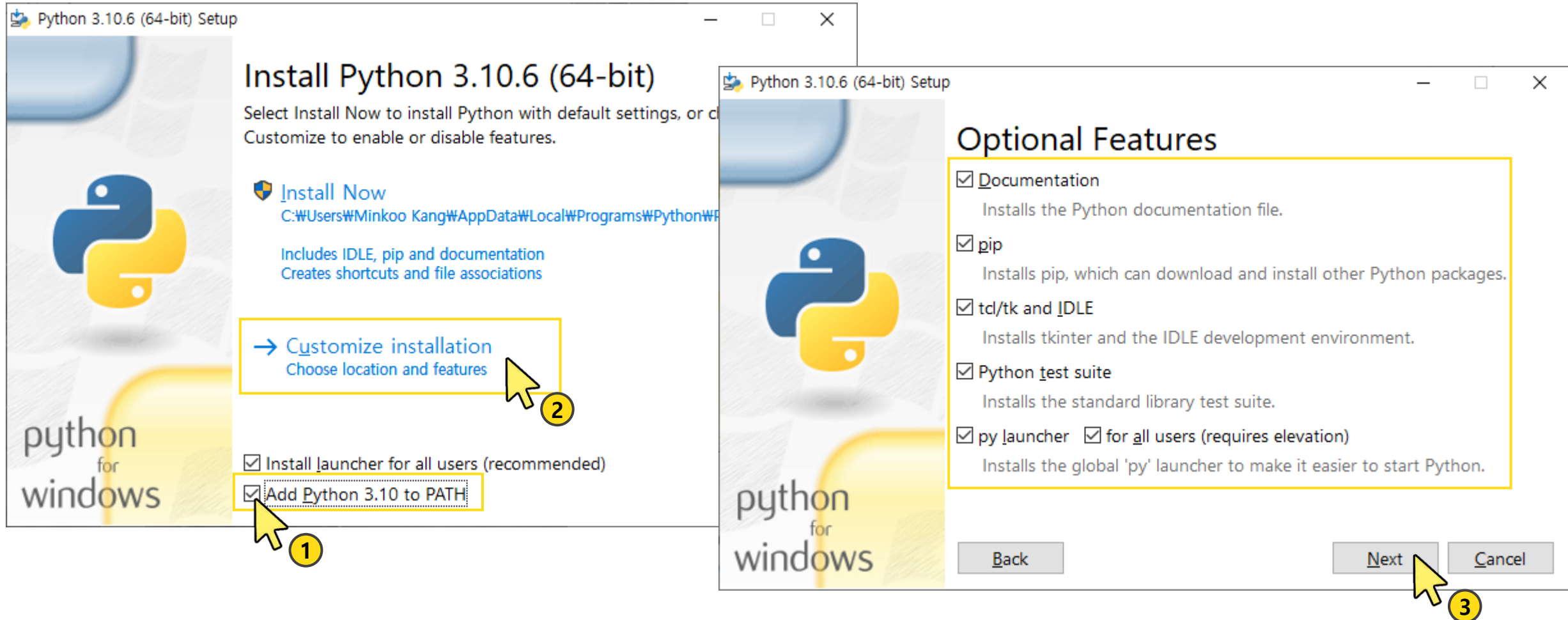
Files

Version	Operating System	Description	MD5 Sum
Gzipped source tarball	Source release		d7663
XZ compressed source tarball	Source release		afc7e
macOS 64-bit universal2 installer	macOS	for macOS 10.9 and later	2ce68
Windows embeddable package (32-bit)	Windows		a62cc
Windows embeddable package (64-bit)	Windows		37303
Windows help file	Windows		0aee6
Windows installer (32-bit)	Windows		c4aa2
Windows installer (64-bit)	Windows	Recommended	8f464

1.1 파이썬 ver3.10.6

프로그램 설치

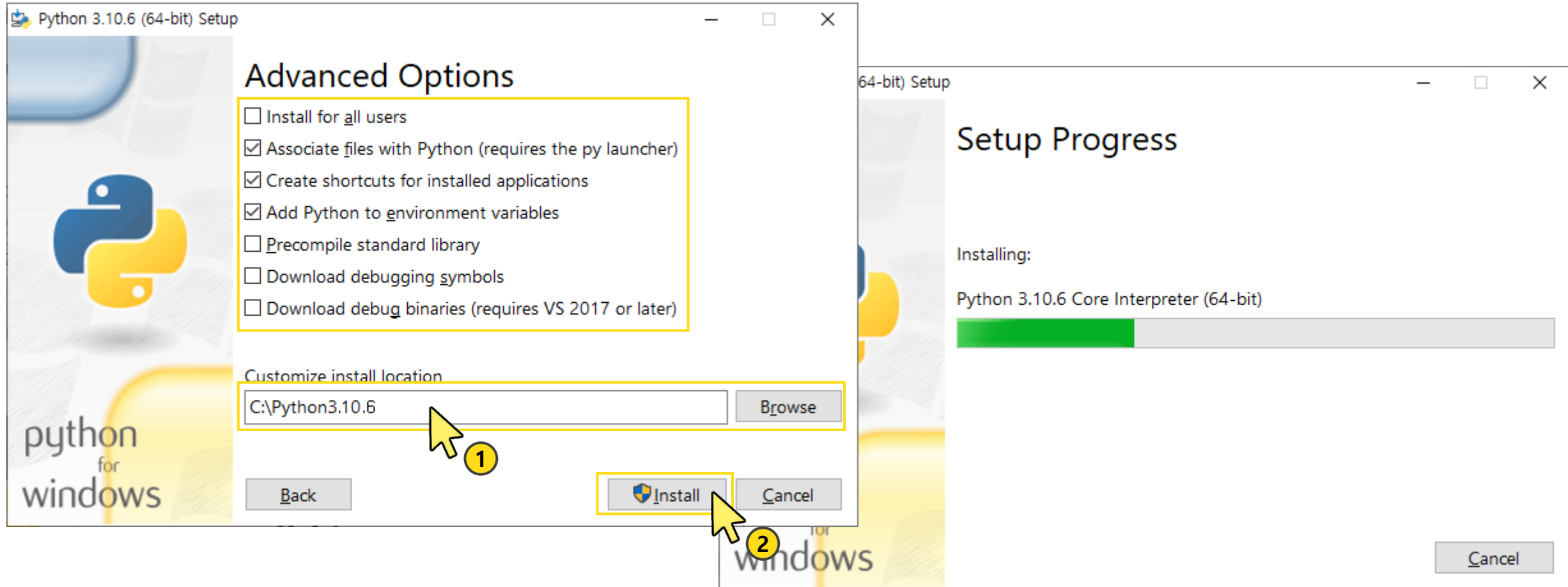
설치 시작



1.1 파이썬 ver3.10.6

프로그램 설치

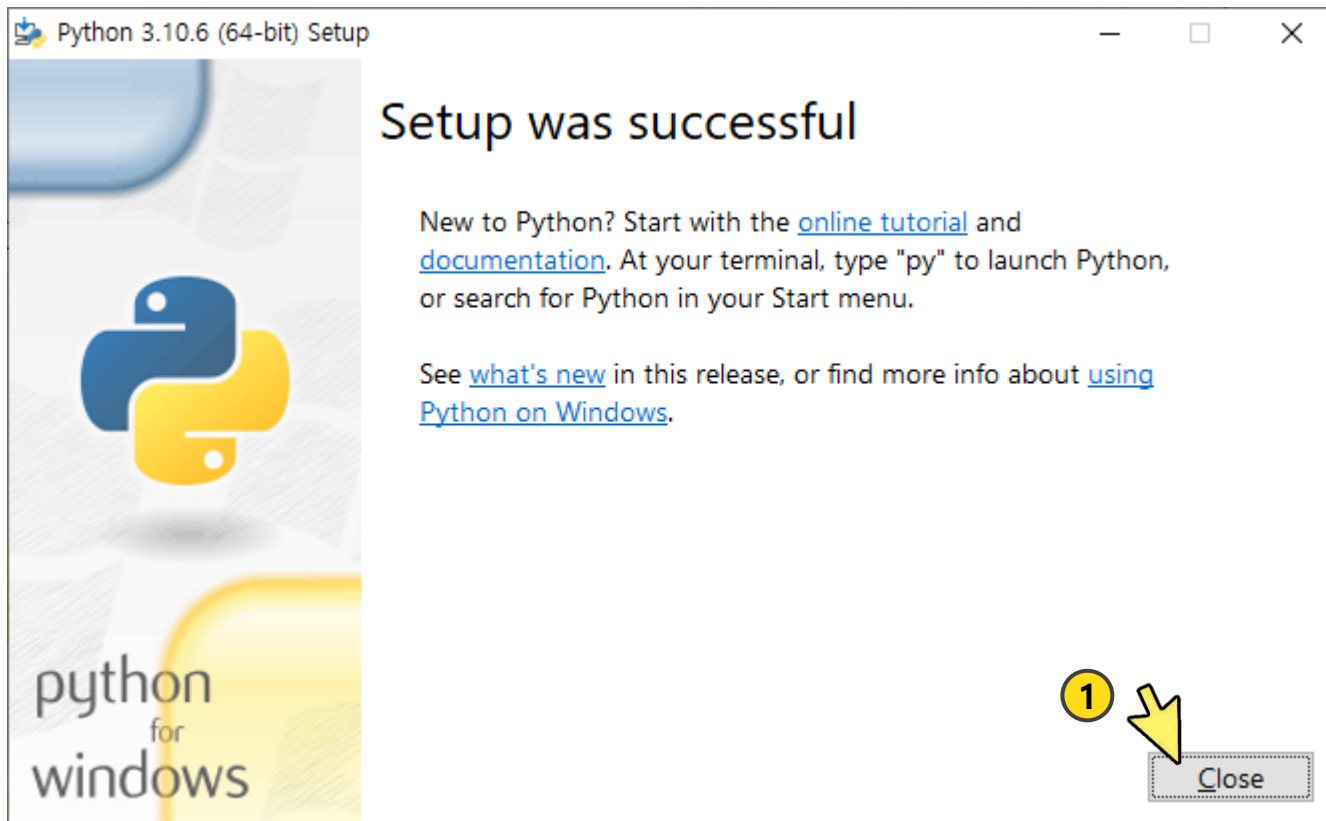
설치 시작



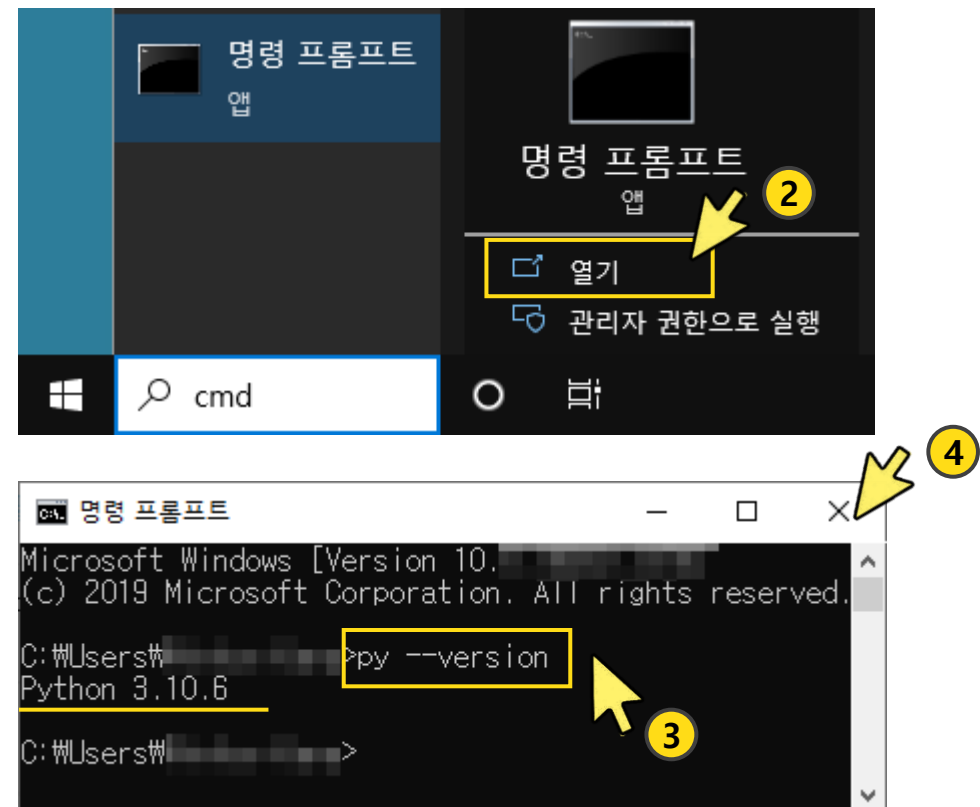
1.1 파이썬 ver3.10.6

프로그램 설치

설치 완료 및 확인

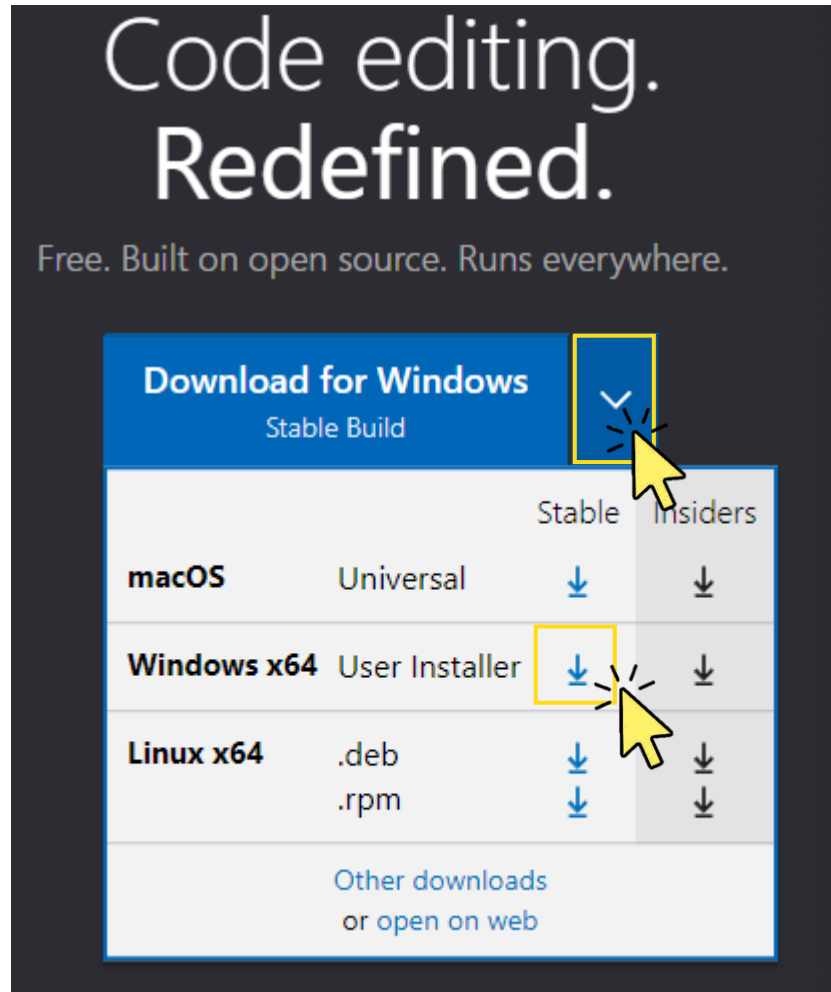


1. 윈도우 키 옆 검색창에 cmd를 입력합니다.
2. 명령 프롬프트를 실행하여 버전을 확인합니다.



1.2 Visual Studio Code

- 다운로드1 : <https://code.visualstudio.com>



- 다운로드2 : https://code.visualstudio.com/updates/v1_69

다운로드 1에서 ver1.69.2를 다운로드 받지 못한 경우에는 다운로드 2에서 설치파일을 다운로드 받으시면 됩니다. (권장버전: ver1.69.2)

June 2022 (version 1.69)

Update 1.69.1: The update addresses these [issues](#).

Update 1.69.2: The update addresses these [issues](#).

Downloads: Windows: [User System ARM](#) | Mac: [Universal 64 bit Arm64](#) | Linux: [deb rpm tarball ARM snap](#)

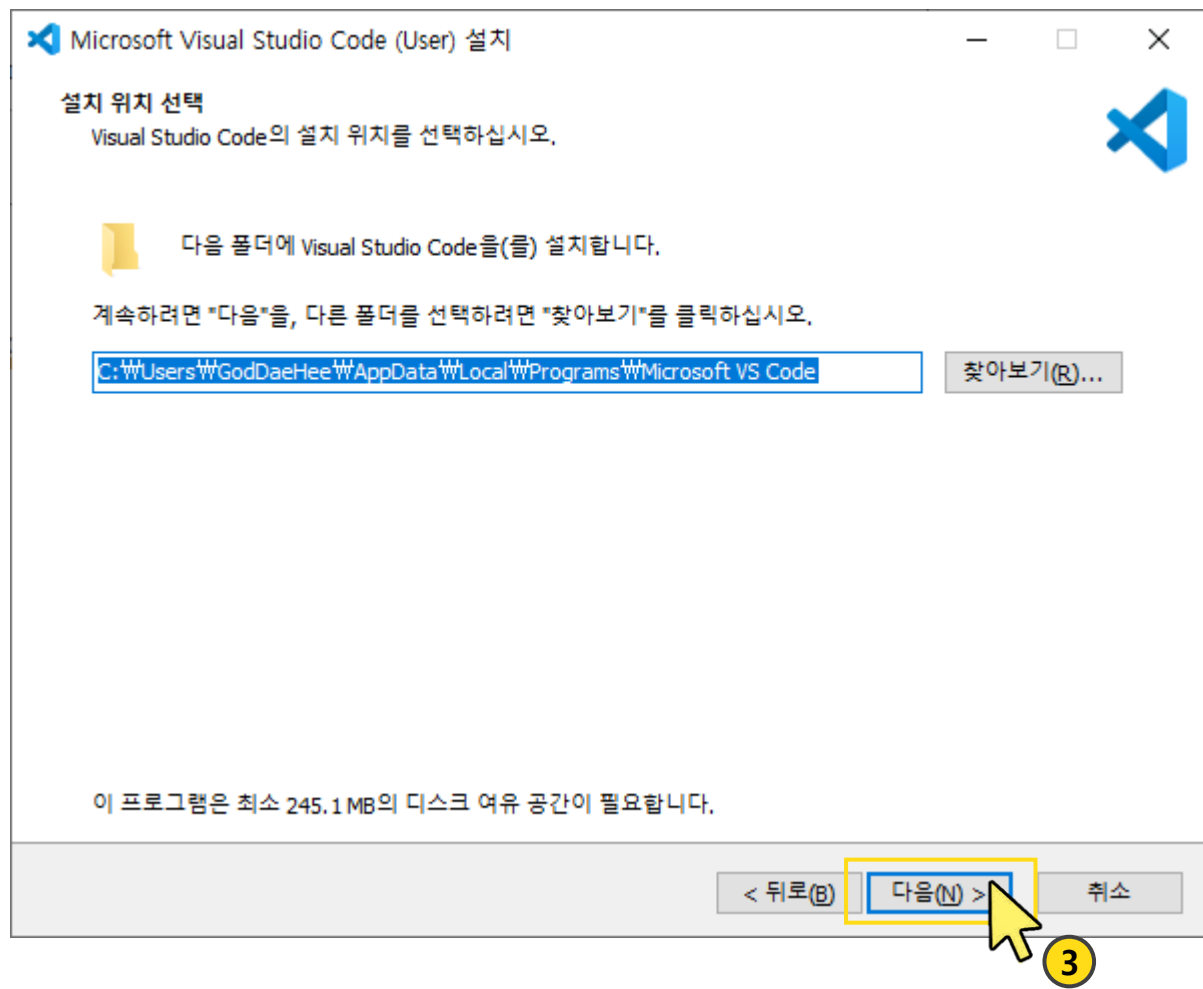
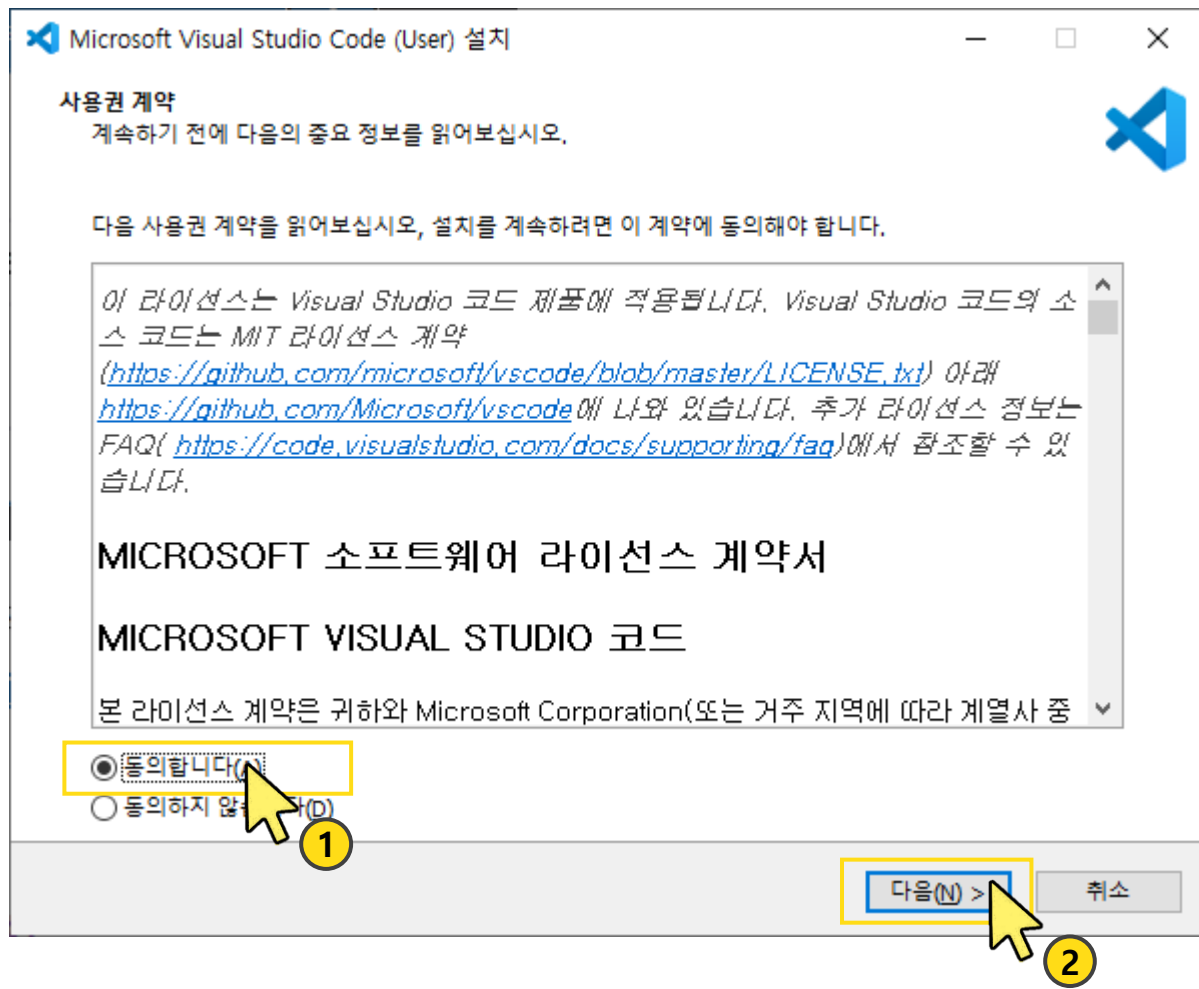
“ 컴퓨터의 OS에 따라서 해당하는 설치 프로그램을 다운로드 합니다. “

[설치 파일] Environment/VSCodeUserSetup-x64-1.69.2.exe

1.2 Visual Studio Code

프로그램 설치

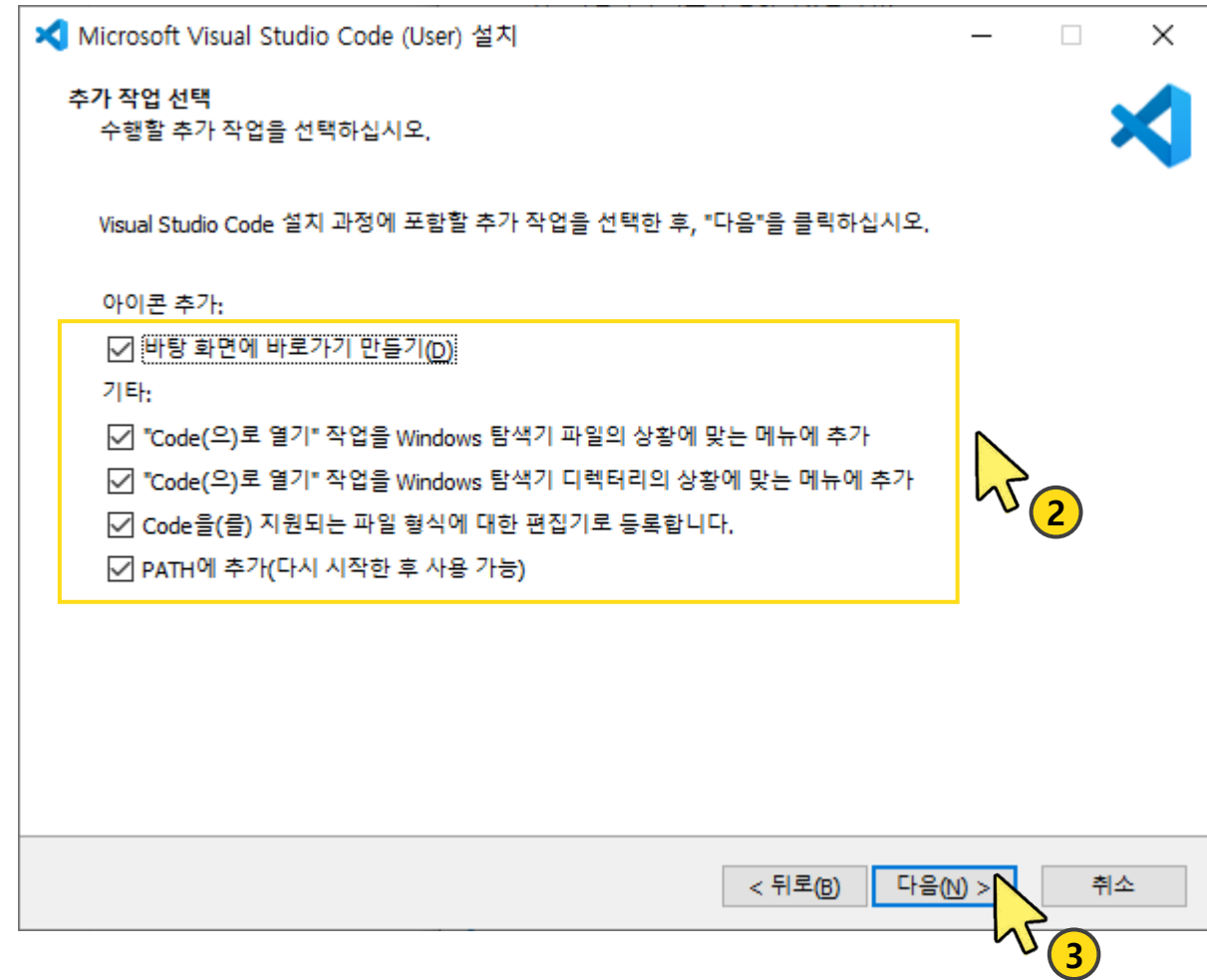
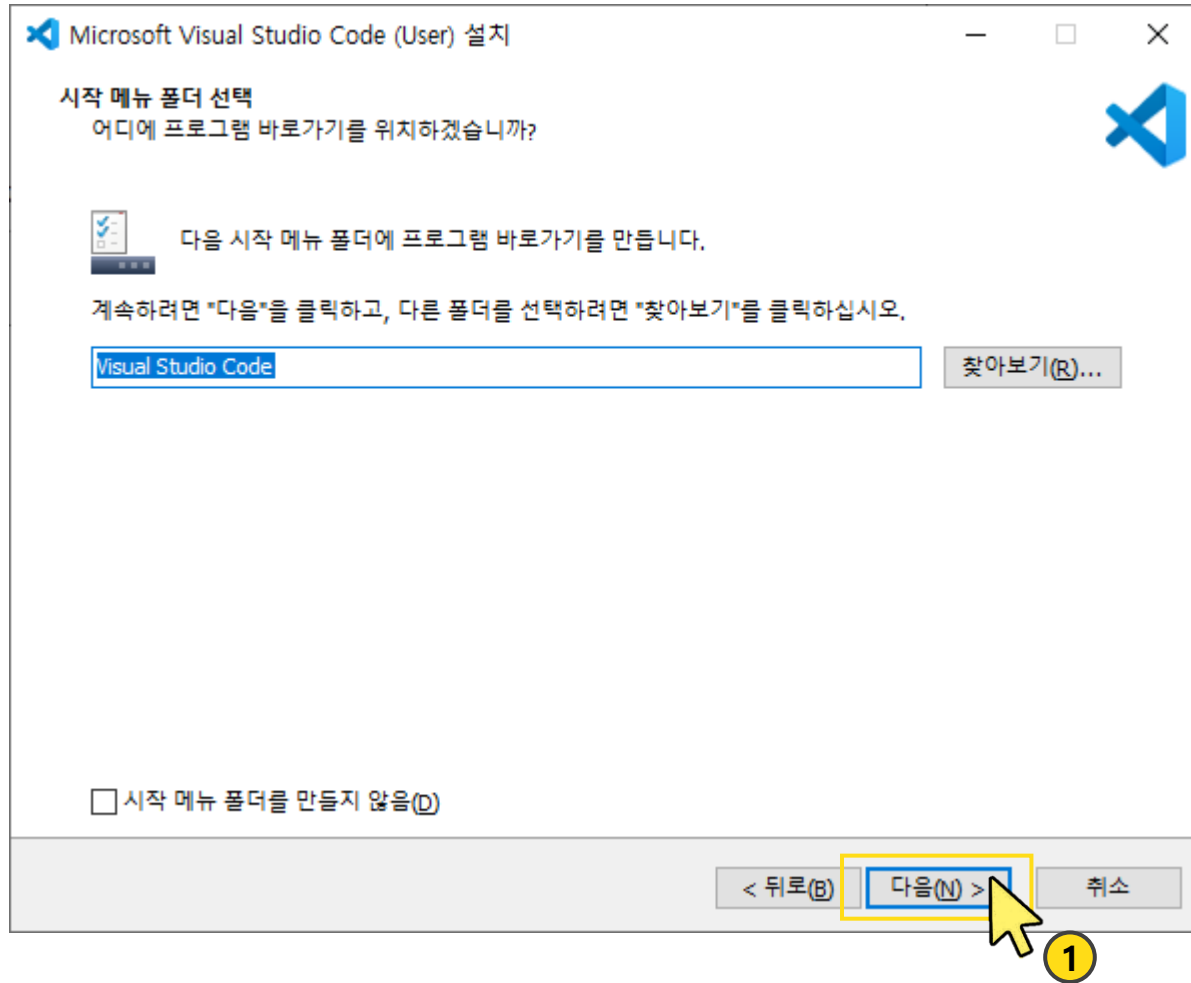
설치 설정 -1



1.2 Visual Studio Code

프로그램 설치

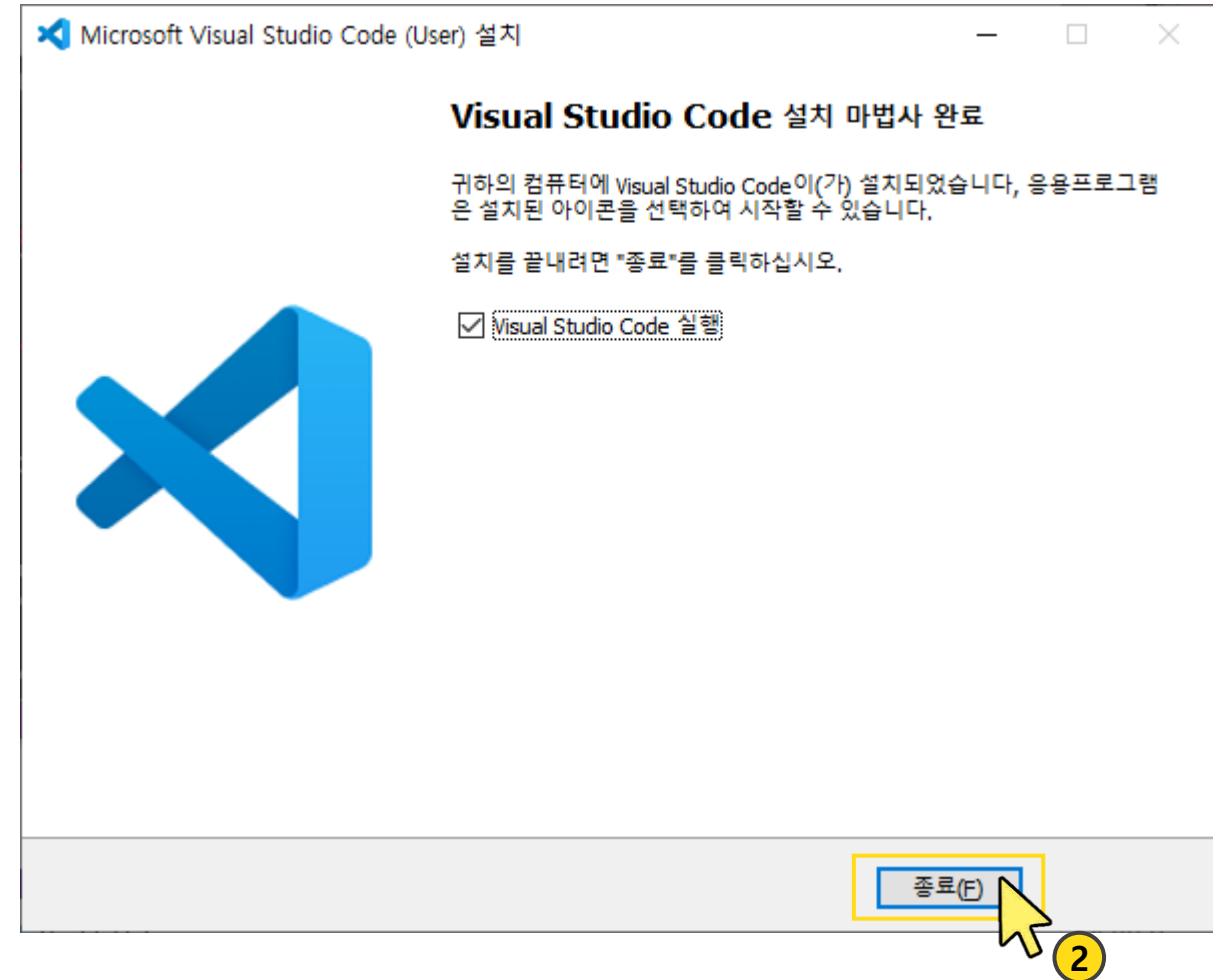
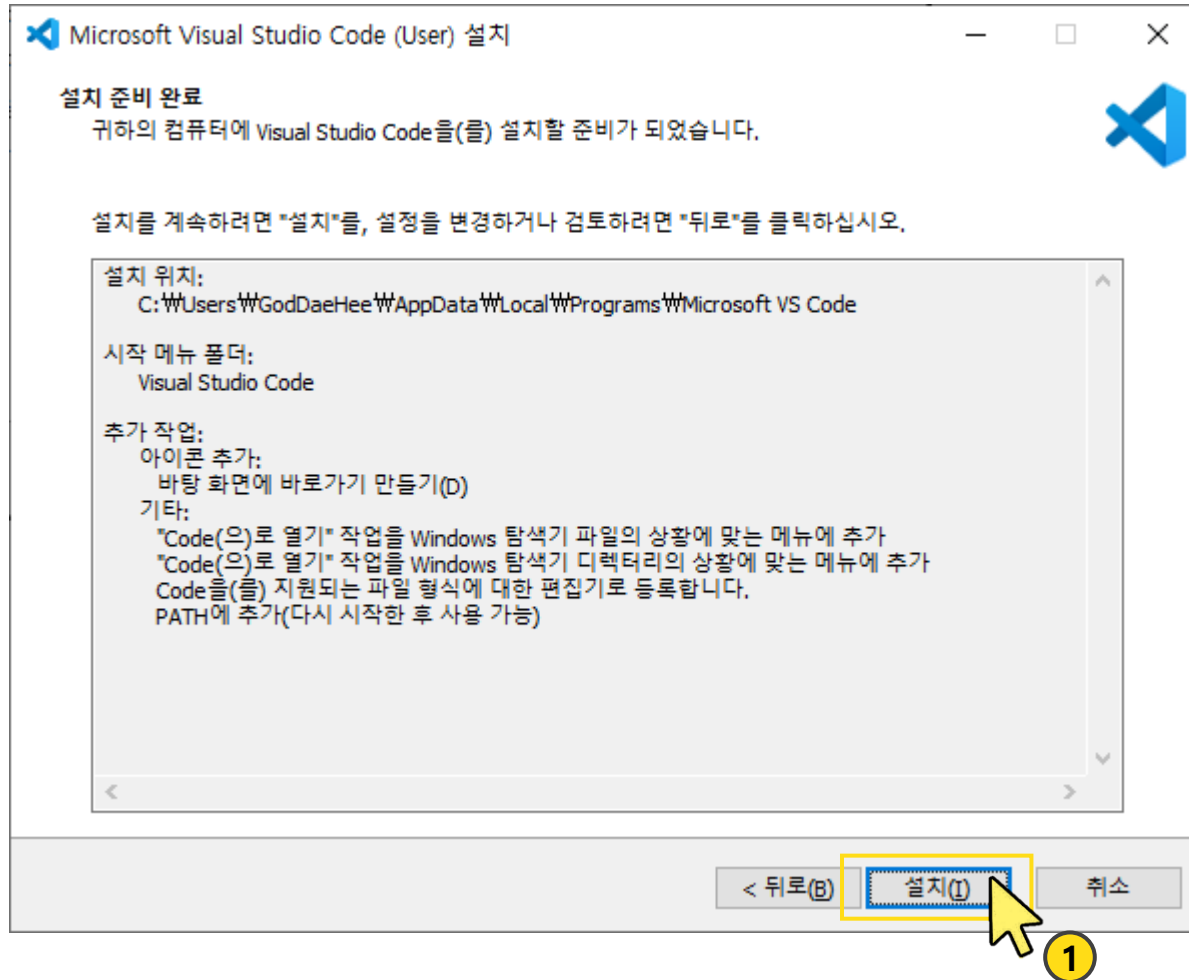
설치 설정 -2



1.2 Visual Studio Code

프로그램 설치

설치 시작 & 완료



1.3 Java JDK ver19

- 다운로드 : www.oracle.com/java/technologies/javase-downloads.html

[설치 파일] Environment/jdk-19_windows-x64_bin.exe

[Java 19](#) [Java 17](#)

Java SE Development Kit 19.0.1 downloads

Thank you for downloading this release of the Java™ Platform, Standard Edition Development Kit (JDK™). The JDK is a development environment for writing and running Java programs using the Java programming language.

The JDK includes tools for developing and testing programs written in the Java programming language and running on the Java platform.

[Linux](#) [macOS](#) [Windows](#)

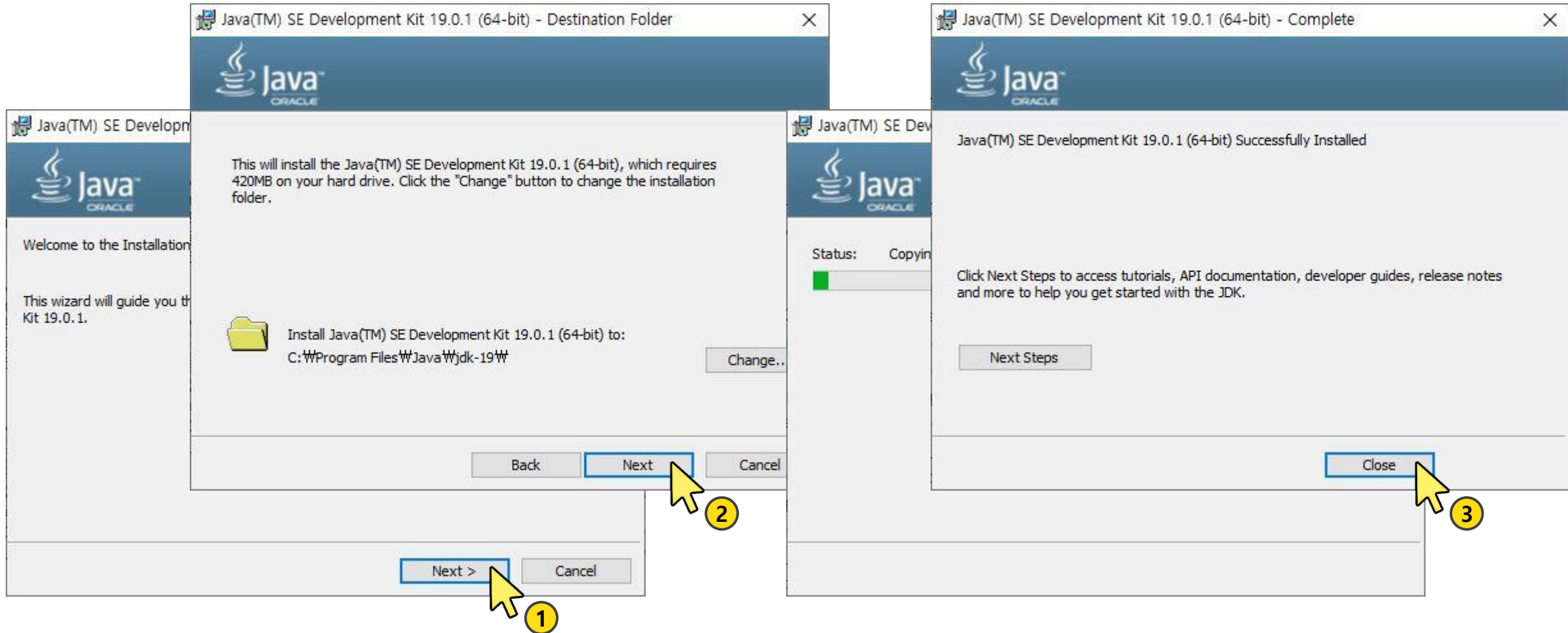
Product/file description	File size	Download
x64 Compressed Archive	179.13 MB	https://download.oracle.com/java/19/latest/jdk-19_windows-x64_bin.zip (sha256)
x64 Installer	158.91 MB	https://download.oracle.com/java/19/latest/jdk-19_windows-x64_bin.exe (sha256)
x64 MSI Installer	157.76 MB	https://download.oracle.com/java/19/latest/jdk-19_windows-x64_bin.msi (sha256)

컴퓨터의 환경에 맞게
설치 프로그램 다운로드

1.3 Java JDK ver19

프로그램 설치

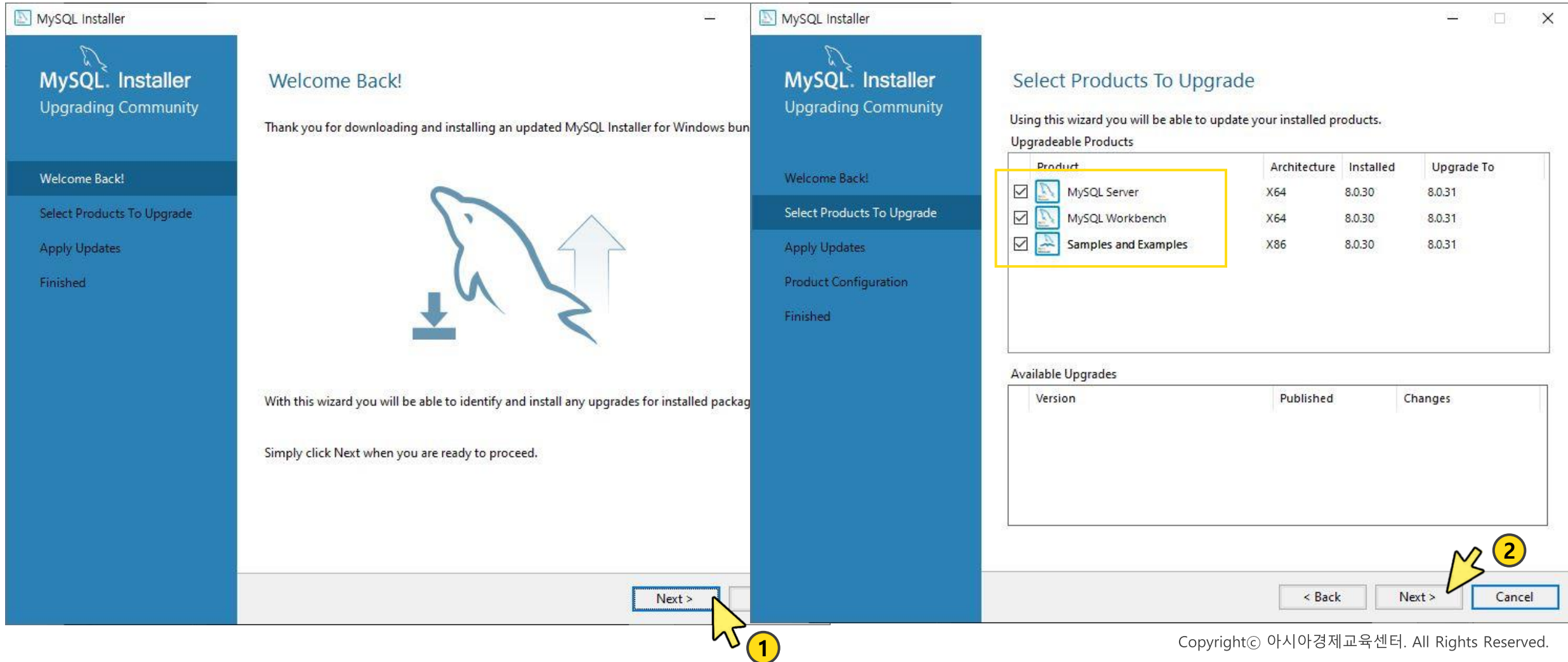
- 설치 시작 및 완료



1.4 MySQL

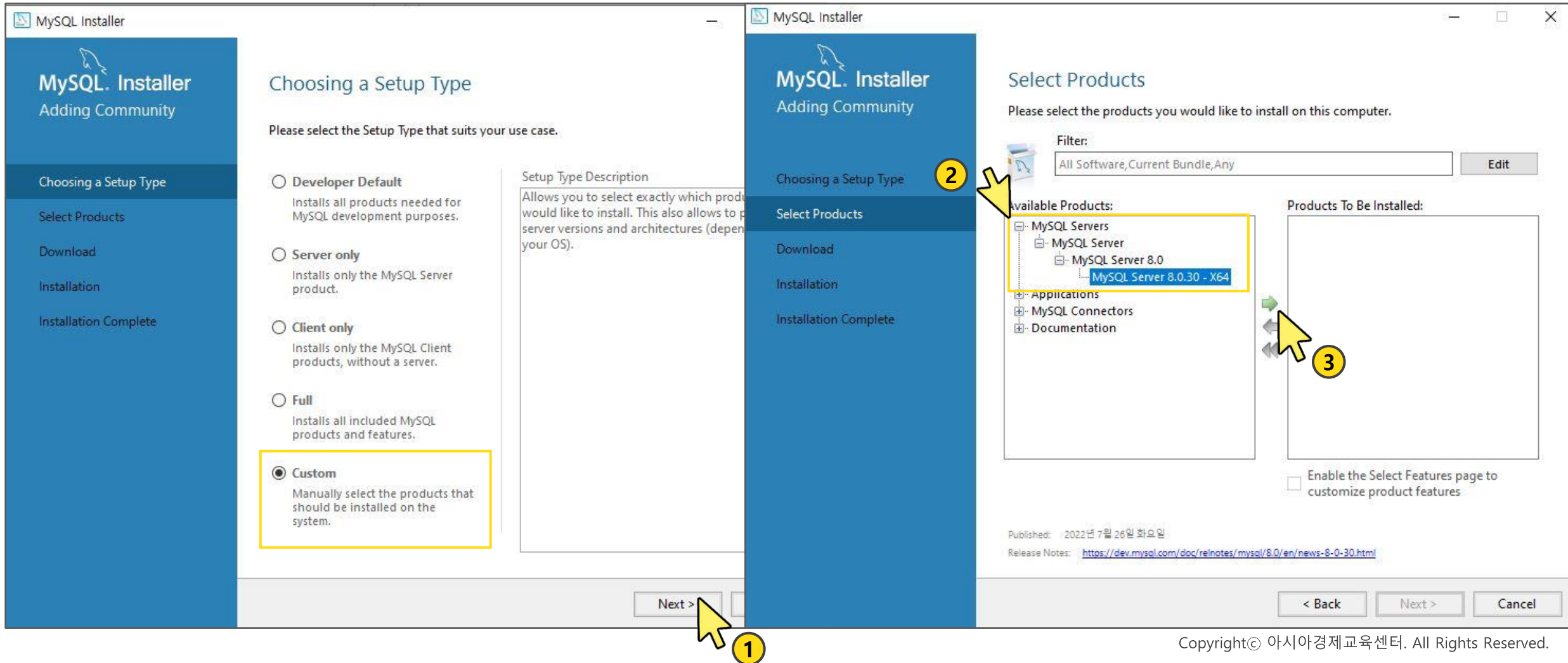
프로그램 설치

[설치 파일] Environment/mysql-installer-community-8.0.30.0.msi



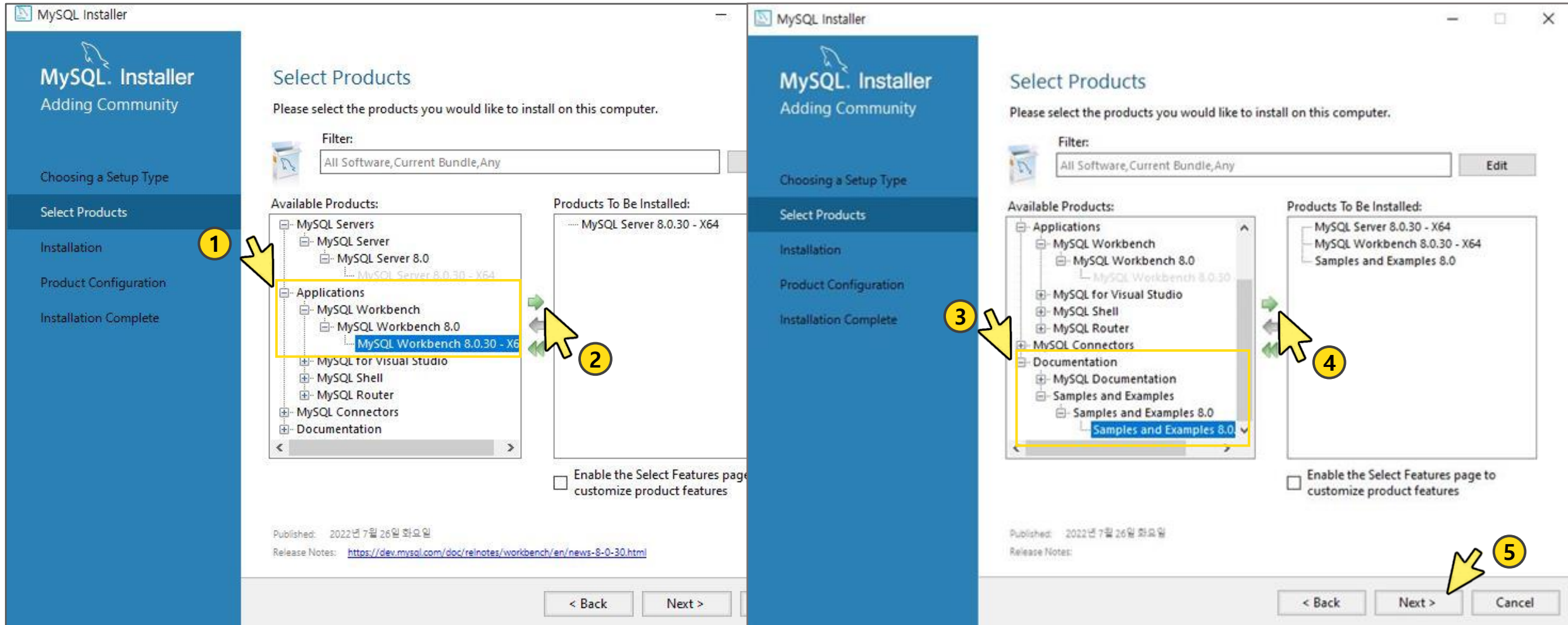
1.4 MySQL

프로그램 설치



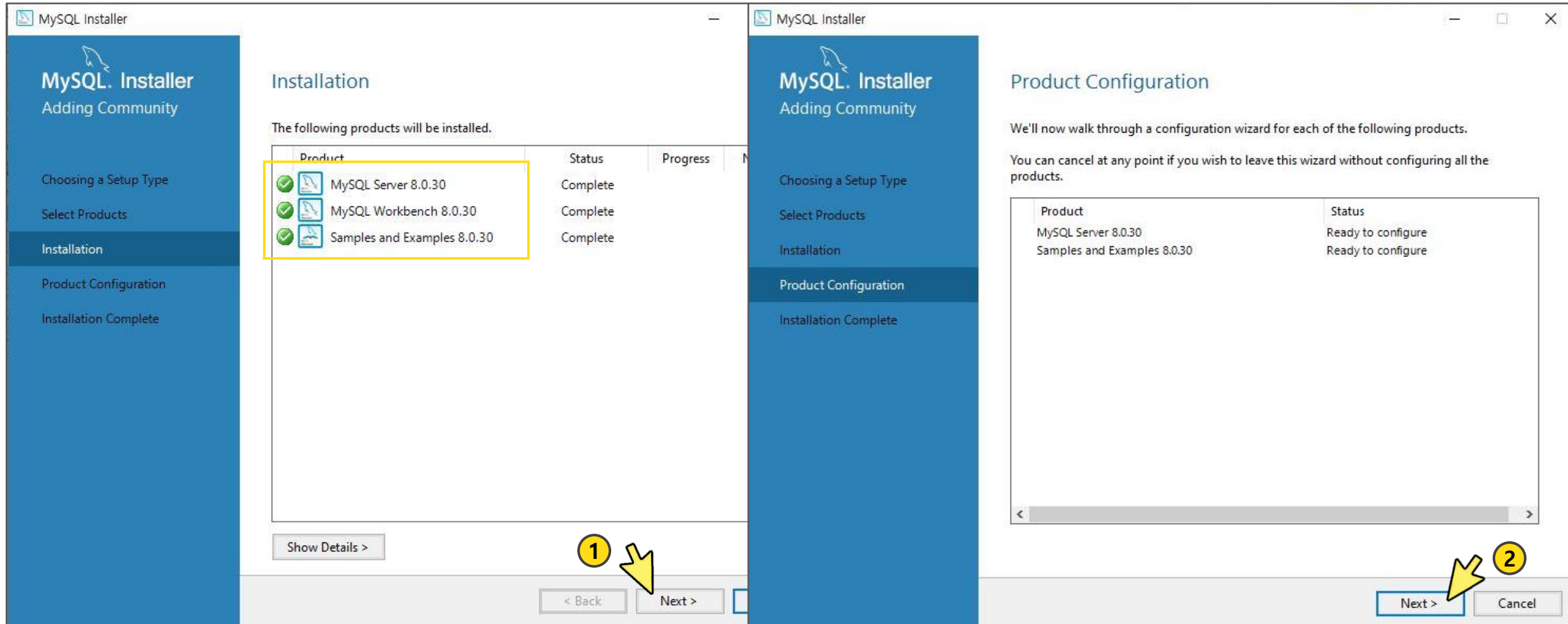
1.4 MySQL

프로그램 설치



1.4 MySQL

프로그램 설치



1.4 MySQL

MySQL Installer
MySQL Server 8.0.30

Type and Networking

Service Configuration Type
1. Select the correct server configuration type for this MySQL Server installation. This defines how much system resources are assigned to the MySQL Server instance.

Config Type: Development Computer

Connectivity
Use the following controls to select how you would like to connect to this server.

☒ TCP/IP Port: 3306 X Protocol Port:
2. ☐ Machine Name Pipe Name: MYSQL 3.
☐ Shared Memory Memory Name: MYSQL

Advanced Configuration
Select the check box below to get additional configuration pages where you can set and logging options for this server instance.

☐ Show Advanced and Logging Options

4. Next >

Authentication Method

☒ Use Strong Password Encryption for Authentication (RECOMMENDED)
5. MySQL 8 supports a new authentication based on improved stronger SHA256 based password methods. It is recommended that all new MySQL Server installations use this method going forward.

! Attention: This new authentication plugin on the server side requires new versions of connectors and clients which add support for this new 8.0 default authentication (caching_sha2_password authentication).

Currently MySQL 8.0 Connectors and community drivers which use libmysqlclient 8.0 support this new method. If clients and applications cannot be updated to support this new authentication method, the MySQL 8.0 Server can be configured to use the legacy MySQL Authentication Method below.

☐ Use Legacy Authentication Method (Retain MySQL 5.x Compatibility)
Using the old MySQL 5.x legacy authentication method should only be considered in the following cases:

- If applications cannot be updated to use MySQL 8 enabled Connectors and drivers.
- For cases where re-compilation of an existing application is not feasible.
- An updated, language specific connector or driver is not yet available.

Security Guidance: When possible, we highly recommend taking needed steps towards upgrading your applications, libraries, and database servers to the new stronger authentication. This new method will significantly improve your security.

6. < Back Next > Cancel

1.4 MySQL

프로그램 설치

MySQL Installer
MySQL Server 8.0.30

Type and Networking
Authentication Method
Accounts and Roles
Windows Service
Apply Configuration

Accounts and Roles

Root Account Password
Enter the password for the root account. Please remember to store this password in a safe place.

MySQL Root Password:

Repeat Password:

Password strength: **Weak**

MySQL User Accounts
Create MySQL user accounts for your users and applications. Assign a role to the user. A user role consists of a set of privileges.

MySQL User Name	Host	User Role
-----------------	------	-----------

< Back Next >

MySQL Installer
MySQL Server 8.0.30

Type and Networking
Authentication Method
Accounts and Roles
Windows Service
Apply Configuration

Windows Service

☒ **Configure MySQL Server as a Windows Service**

Windows Service Details
Please specify a Windows Service name to be used for this MySQL Server instance. A unique name is required for each instance.

Windows Service Name:

☒ **Start the MySQL Server at System Startup**

Run Windows Service as ...
The MySQL Server needs to run under a given user account. Based on the security requirements of your system you need to pick one of the options below.

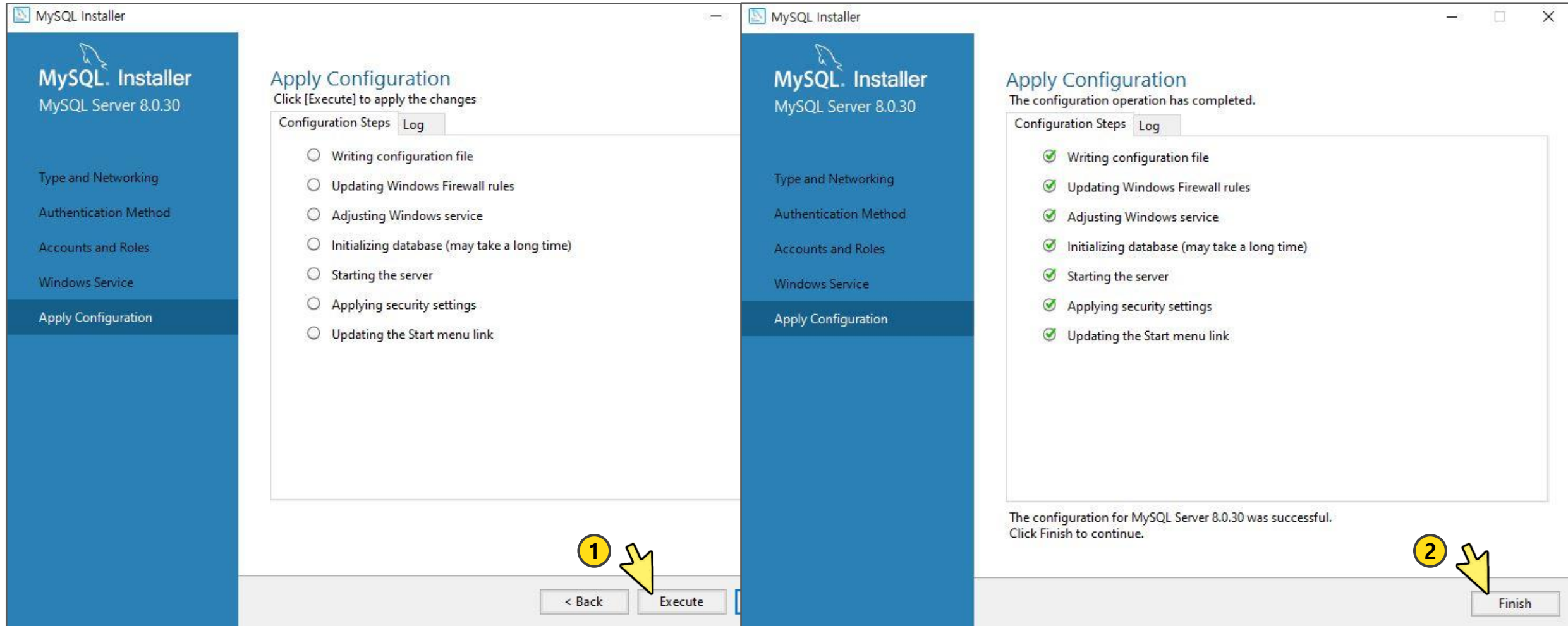
☒ **Standard System Account**
Recommended for most scenarios.

☐ **Custom User**
An existing user account can be selected for advanced scenarios.

< Back Next > Cancel

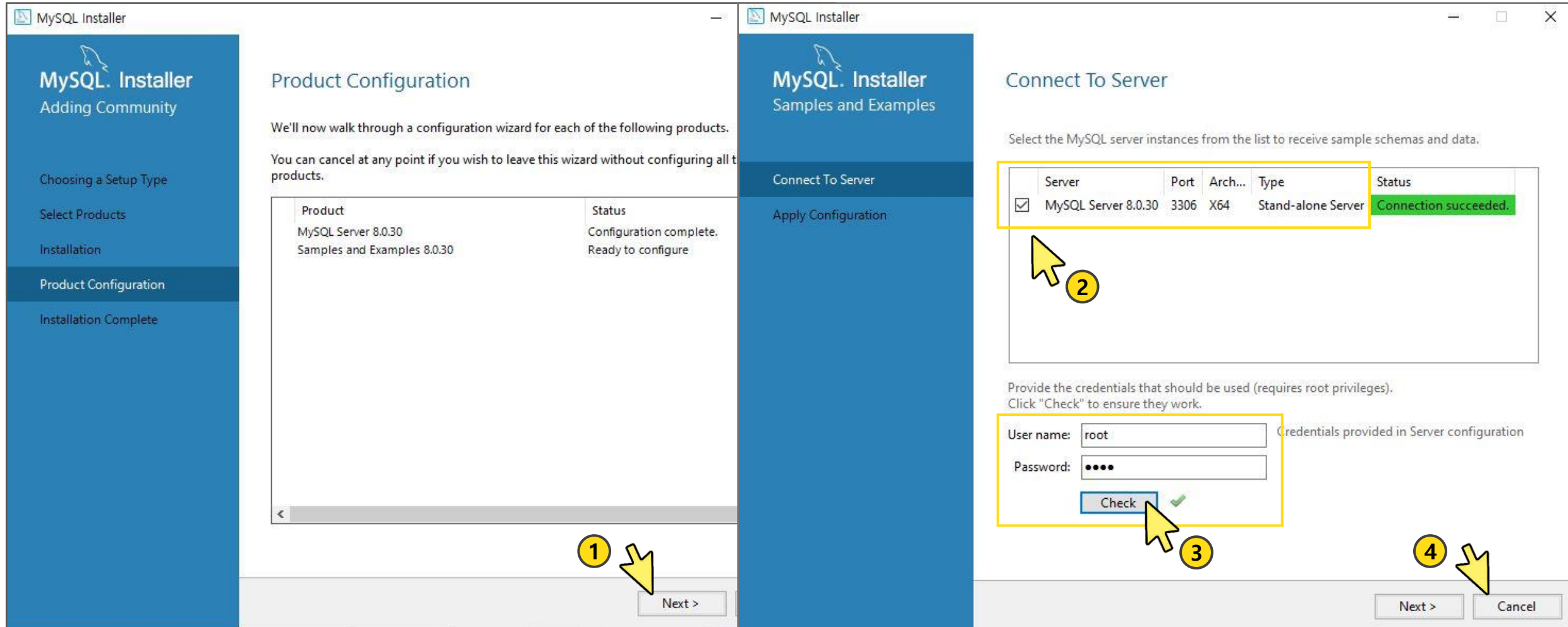
1.4 MySQL

프로그램 설치



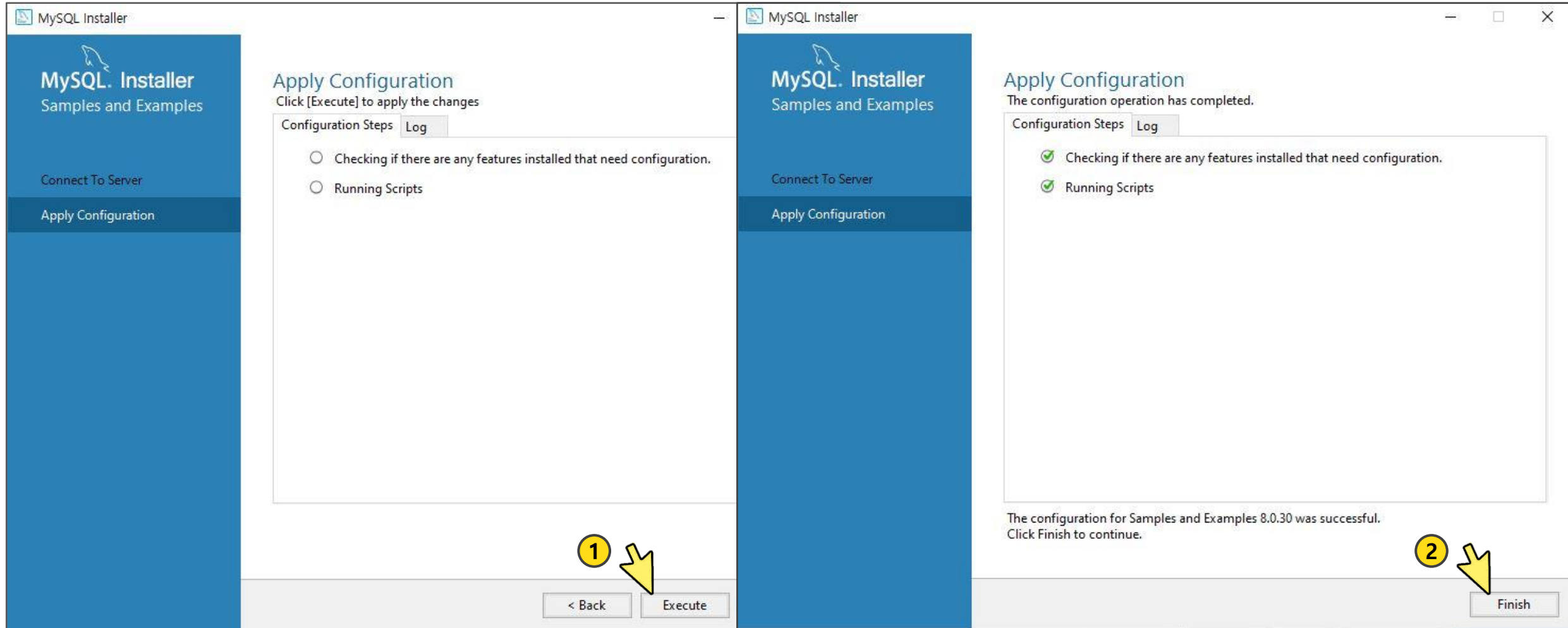
1.4 MySQL

프로그램 설치



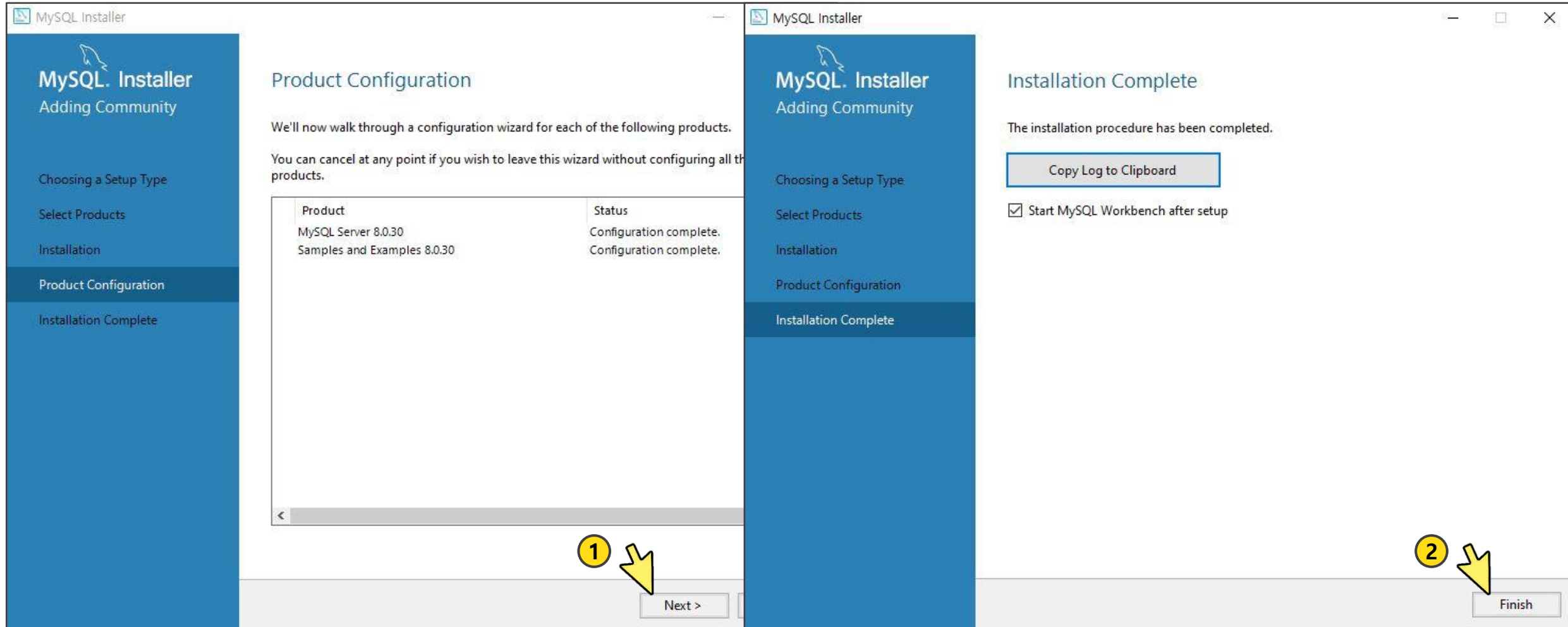
1.4 MySQL

프로그램 설치



1.4 MySQL

프로그램 설치



1.5 CUDA / Cudnn

- NVIDIA 그래픽카드 드라이버 다운로드 : <https://www.nvidia.co.kr/Download/index.aspx?lang=kr>

NVIDIA 드라이버 다운로드

아래의 드롭다운 목록에서 자신의 NVIDIA 제품에 알맞은 드라이버를 선택하세요. 1

제품 유형: GeForce

제품 시리즈: GeForce RTX 40 Series

제품 계열: NVIDIA GeForce RTX 4090

운영 체제: Windows 10 64-bit

다운로드 타입: Game Ready 드라이버(GRD) ?

언어: Korean

검색

GeForce Game Ready 드라이버

버전:	526.47 WHQL
배포 날짜:	2023.10.27
운영 체제:	Windows 10 64-bit, Windows 11
언어:	Korean
파일 크기:	812.32 MB

다운로드

컴퓨터에 NVIDIA 그래픽카드를 새로 장착한 경우, 컴퓨터가 정상인식을 하도록 해당하는 모델을 선택하여 그래픽카드 드라이버를 다운로드 받아서 설치해야 합니다.

1.5 CUDA ver11.8 / Cudnn

프로그램 설치

- 다운로드 : https://developer.nvidia.com/cuda-downloads?target_os=Windows&target_arch=x86_64

[설치 파일] Environment/cuda_11.8.0_windows_network.exe

CUDA Toolkit 11.8 Downloads

Select Target Platform

Click on the green buttons that describe your target platform. Only supported platforms will be shown. By downloading and using the software, you agree to fully comply with the terms and conditions of the [CUDA EULA](#).

Operating System: Linux, Windows

Architecture: x86_64

Version: 10, 11, Server 2016, Server 2019, Server 2022

exe (local), exe (network)

Download Installer for Windows 10 x86_64

The base installer is available for download below.

> **Base Installer** [Download \(29.1 MB\)](#)

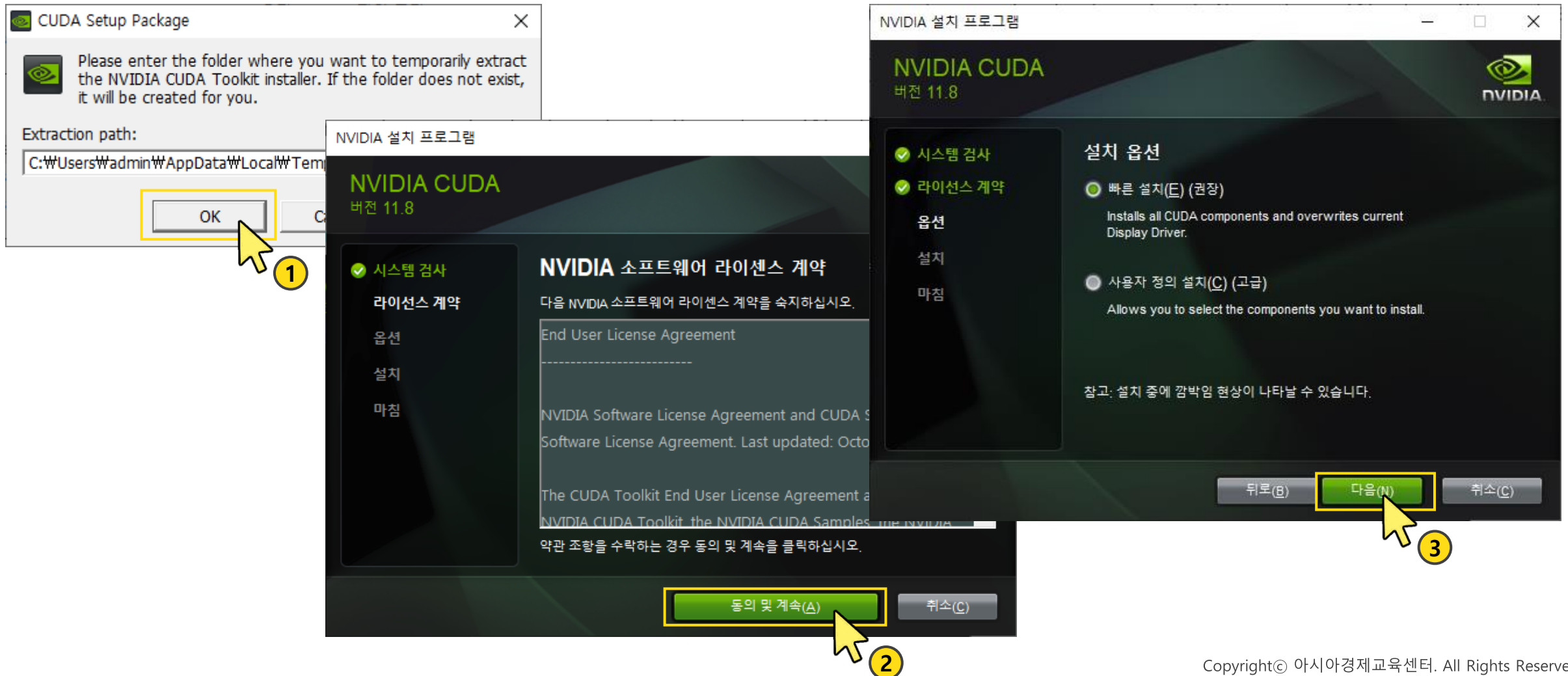
Installation Instructions:

1. Double click cuda_11.8.0_windows_network.exe
2. Follow on-screen prompts

1.5 CUDA ver11.8 / Cudnn

프로그램 설치

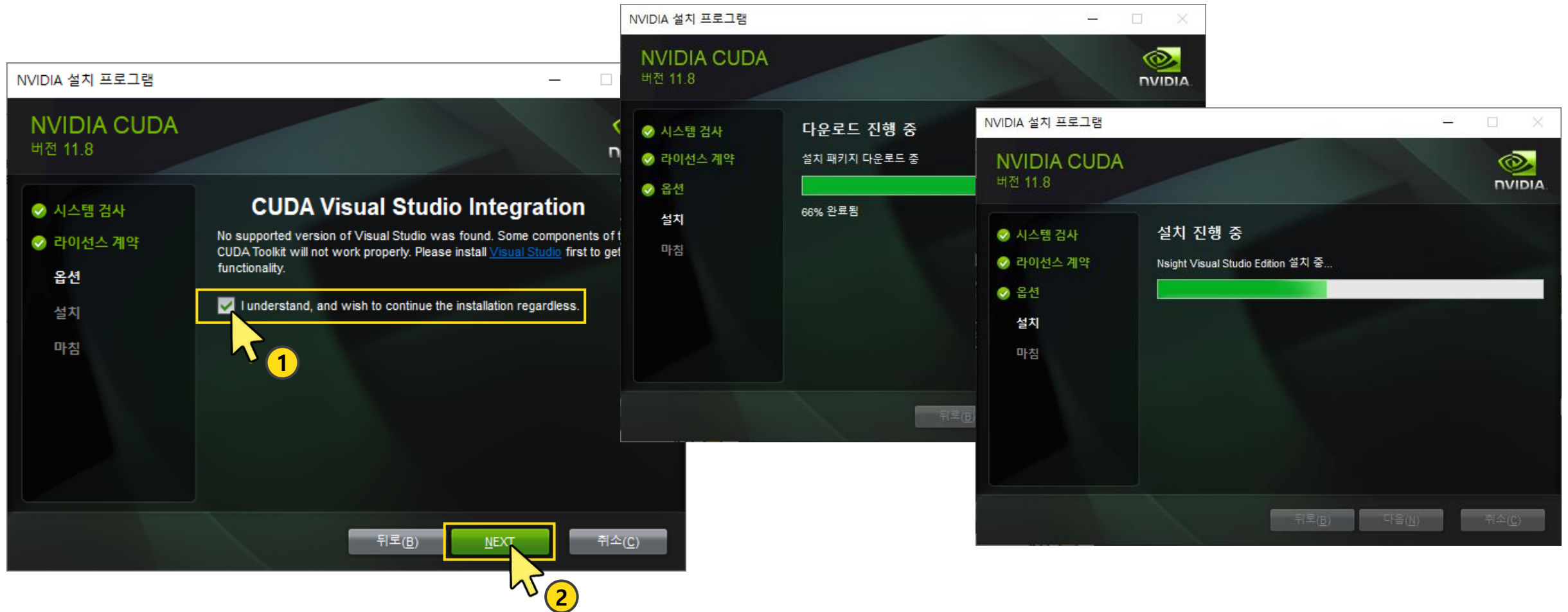
■ CUDA ver11.8 : 설치 -1



1.5 CUDA ver11.8 / Cudnn

프로그램 설치

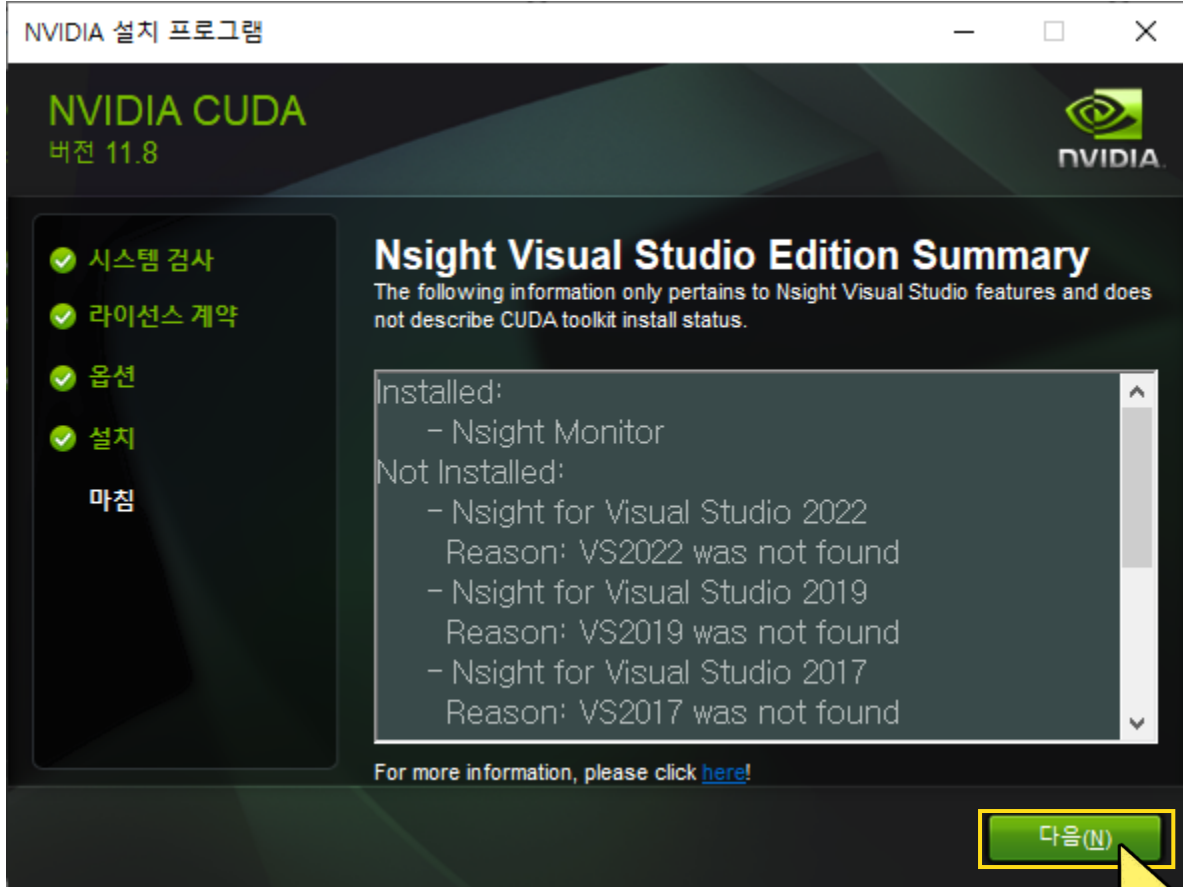
■ CUDA ver11.8 : 설치 -2



1.5 CUDA ver11.8 / Cudnn

프로그램 설치

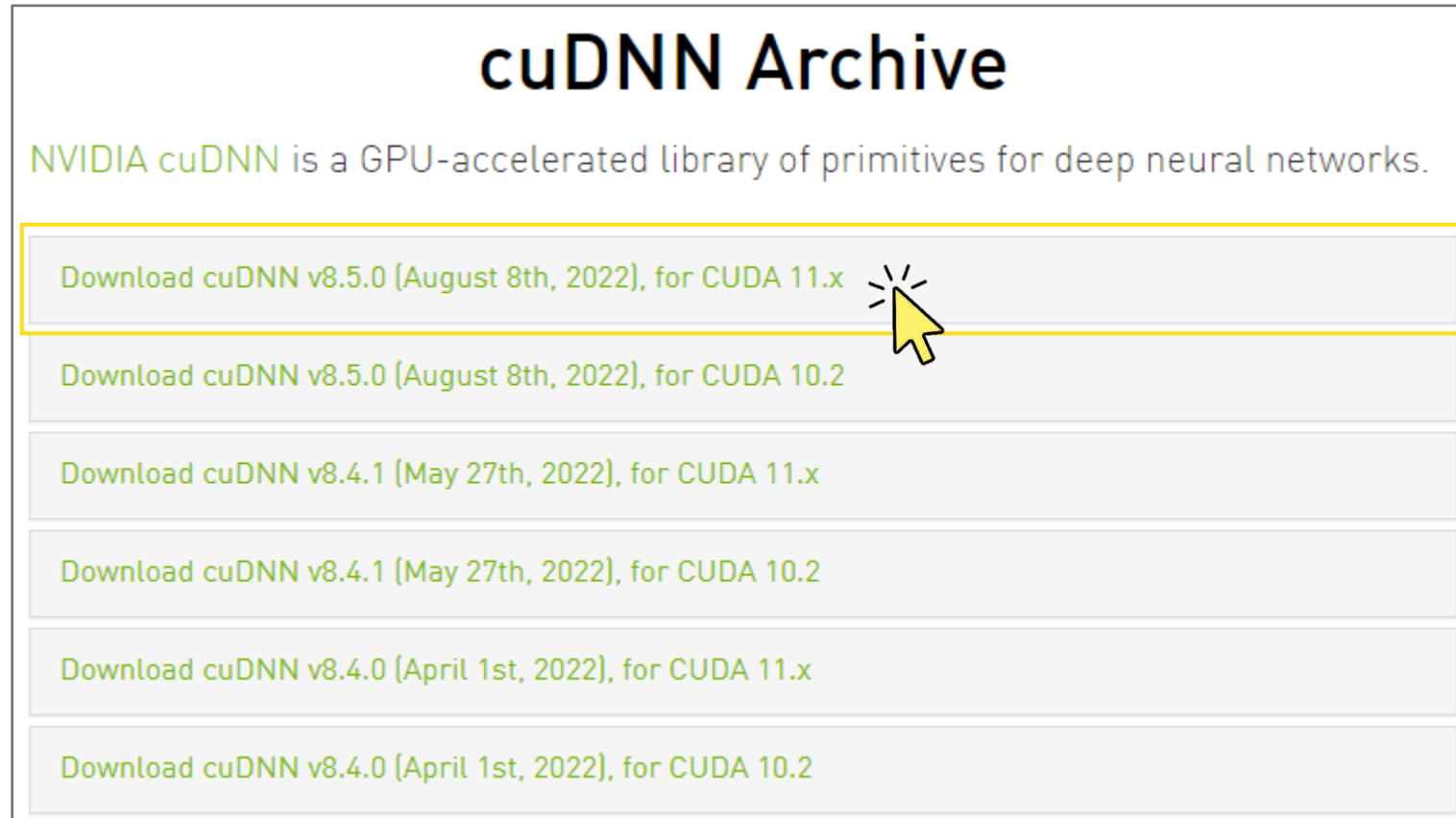
■ CUDA ver11.8 : 설치 -3



1.5 CUDA / Cudnn ver8.5.0

- 다운로드 : <https://developer.nvidia.com/rdp/cudnn-archive>

[설치 파일] Environment/cudnn-windows-x86_64-8.5.0.96_cuda11-archive.zip



1.5 CUDA / Cudnn ver8.5.0

프로그램 설치

1 압축 풀기

2 복사 Ctrl+C

3 붙여넣기 Ctrl+V

로컬 디스크 (C:) > Program Files > NVIDIA GPU Computing Toolkit > CUDA > v11.8

이름	유형	크기
bin	파일 폴더	
compute-sanitizer	파일 폴더	
extras	파일 폴더	
include	파일 폴더	
lib	파일 폴더	
libnvvp	파일 폴더	
nvml	파일 폴더	
nvvm	파일 폴더	
src	파일 폴더	
tools	파일 폴더	
CUDA_Toolkit_Release_Notes.txt	텍스트 문서	79KB
DOCS	파일	1KB
EULA.txt	텍스트 문서	61KB
README	파일	1KB
version.json	JSON 원본 파일	3KB

1. 설치파일의 압축을 해제합니다.
2. 안에 폴더 3개를 복사합니다.
3. 복사한 폴더를 아래의 경로에 붙여넣는다.

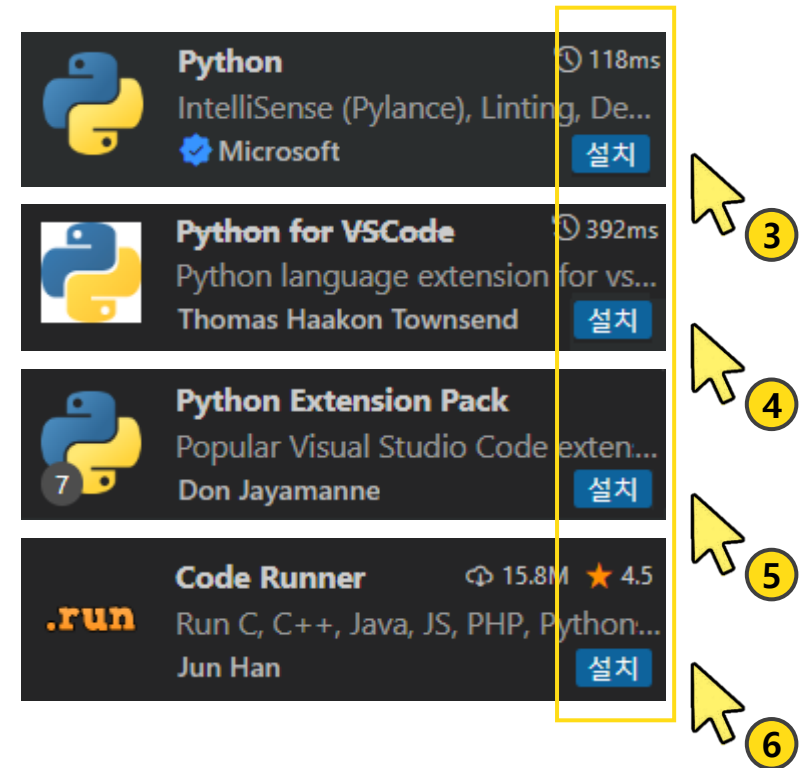
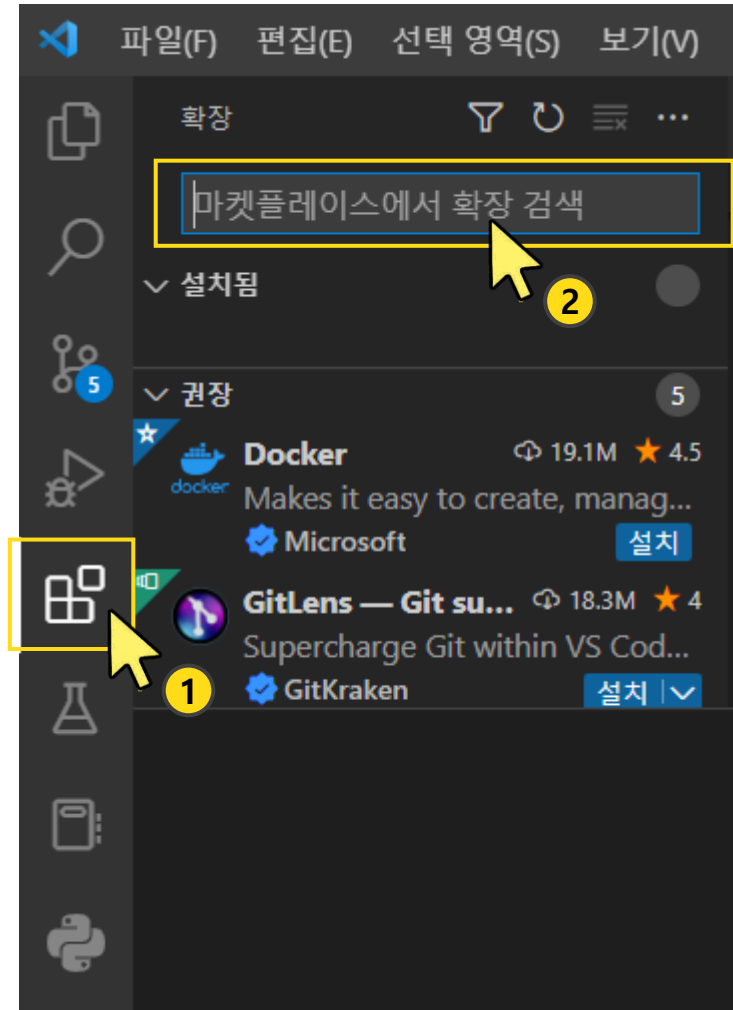
C:\Program Files\NVIDIA GPU Computing Toolkit\CUDA\v11.8

2.1 VS Code 확장프로그램 - [1]

환경설정

- 확장 프로그램 설치
 - python
 - python for VS Code
 - Python Extension Pack
 - Code Runner

1. 확장 탭을 클릭한다.
2. 설치할 확장 프로그램을 하나씩 검색하여 설치한다.



2.1 VS Code 확장프로그램 - [2]

환경설정

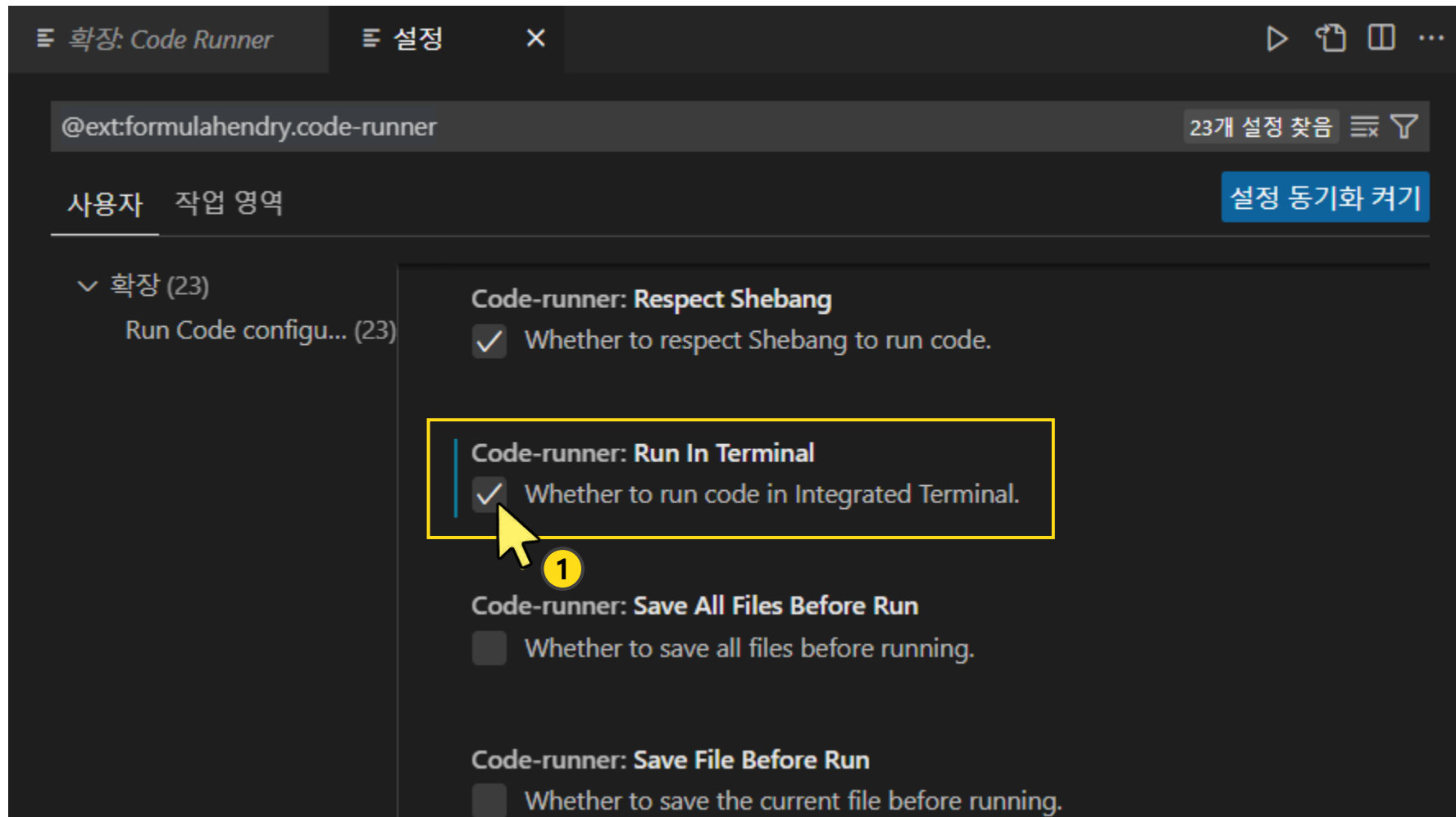
- 확장 프로그램 : Code Runner 설정 변경 -1



2.1 VS Code 확장프로그램 - [3]

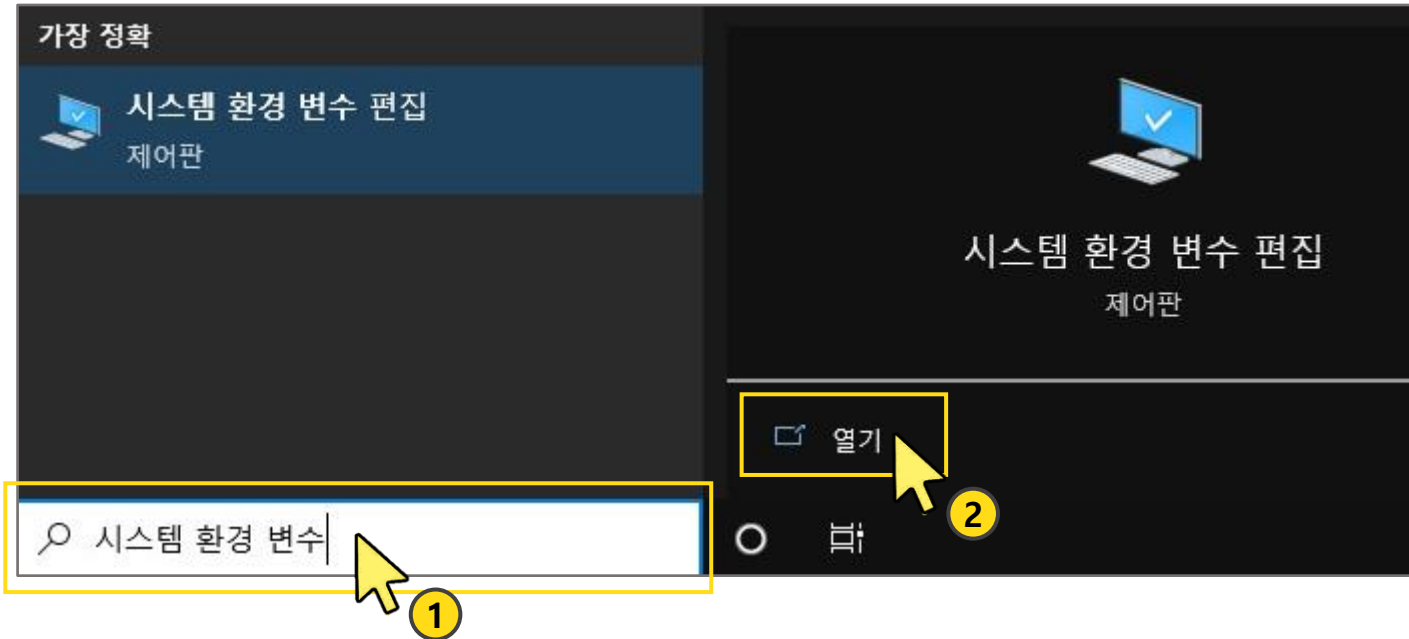
환경설정

- 확장 프로그램 : Code Runner 설정 변경 -2

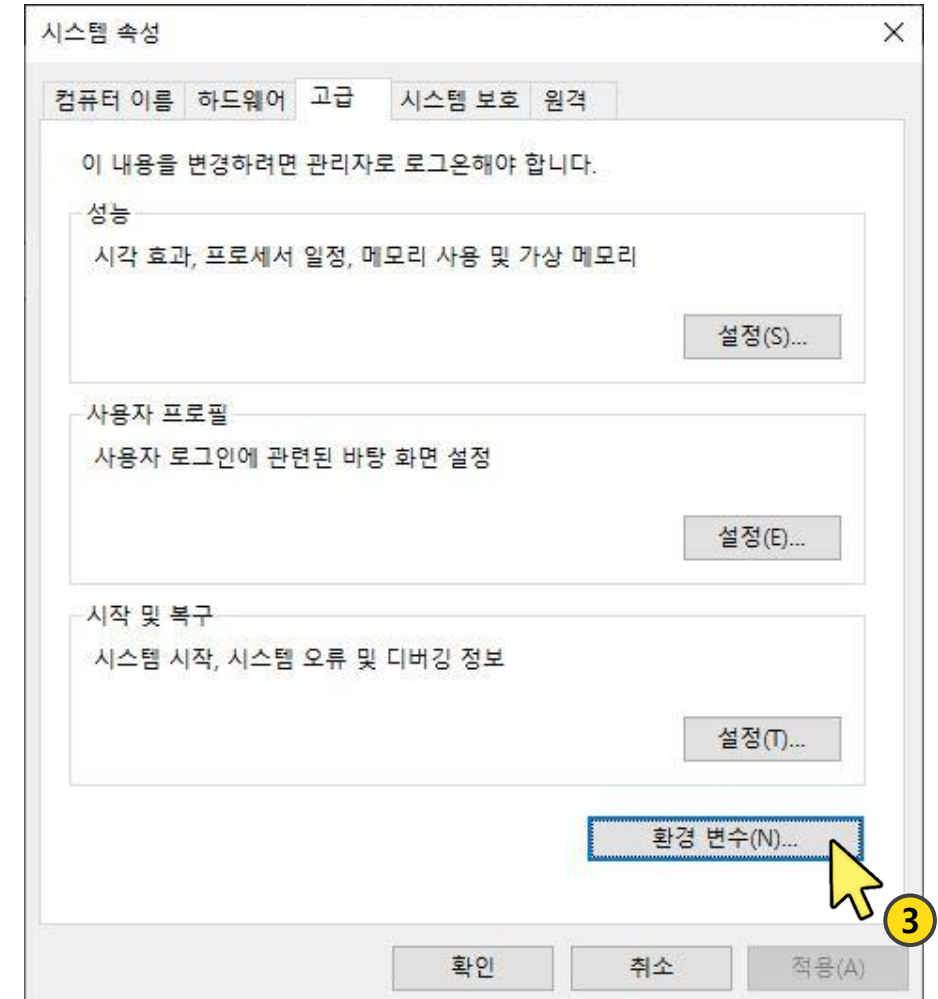


2.2 Java 환경변수 - [1]

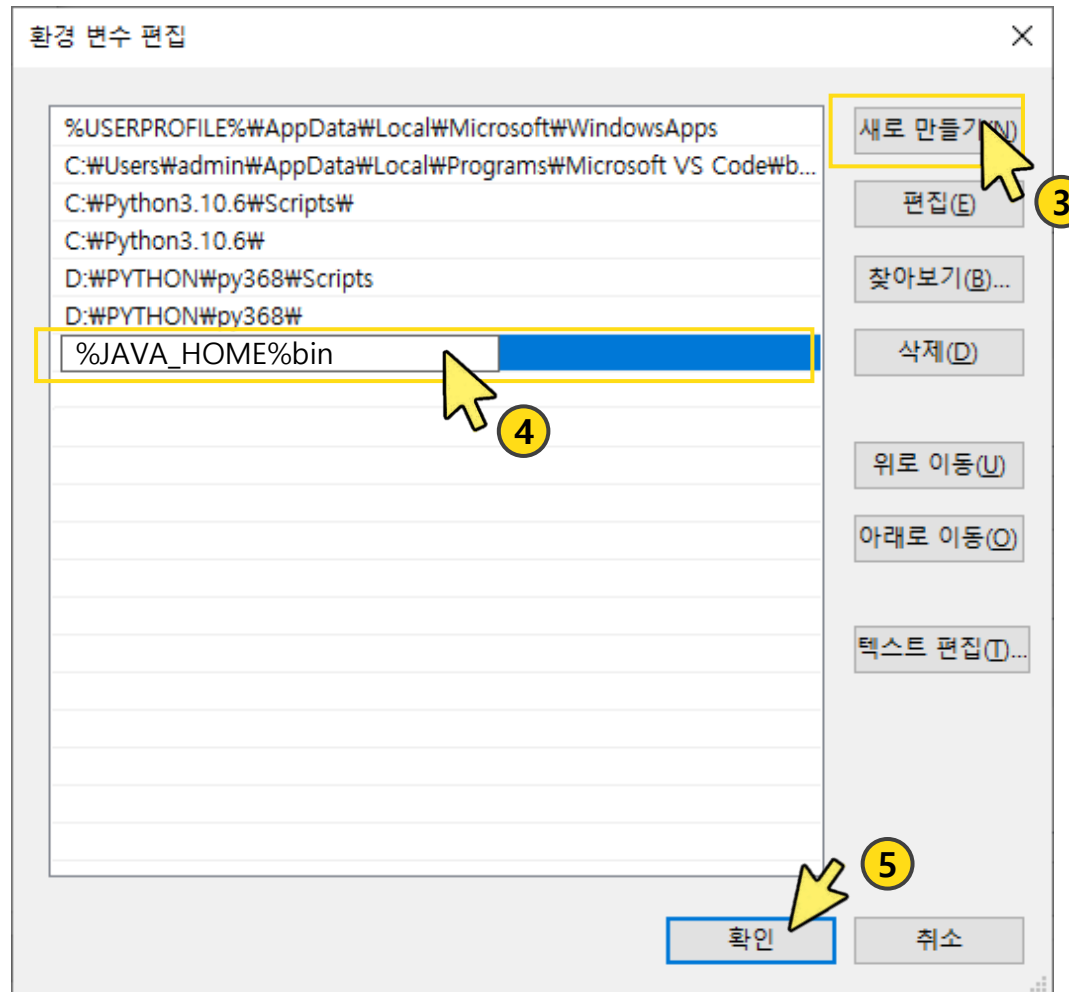
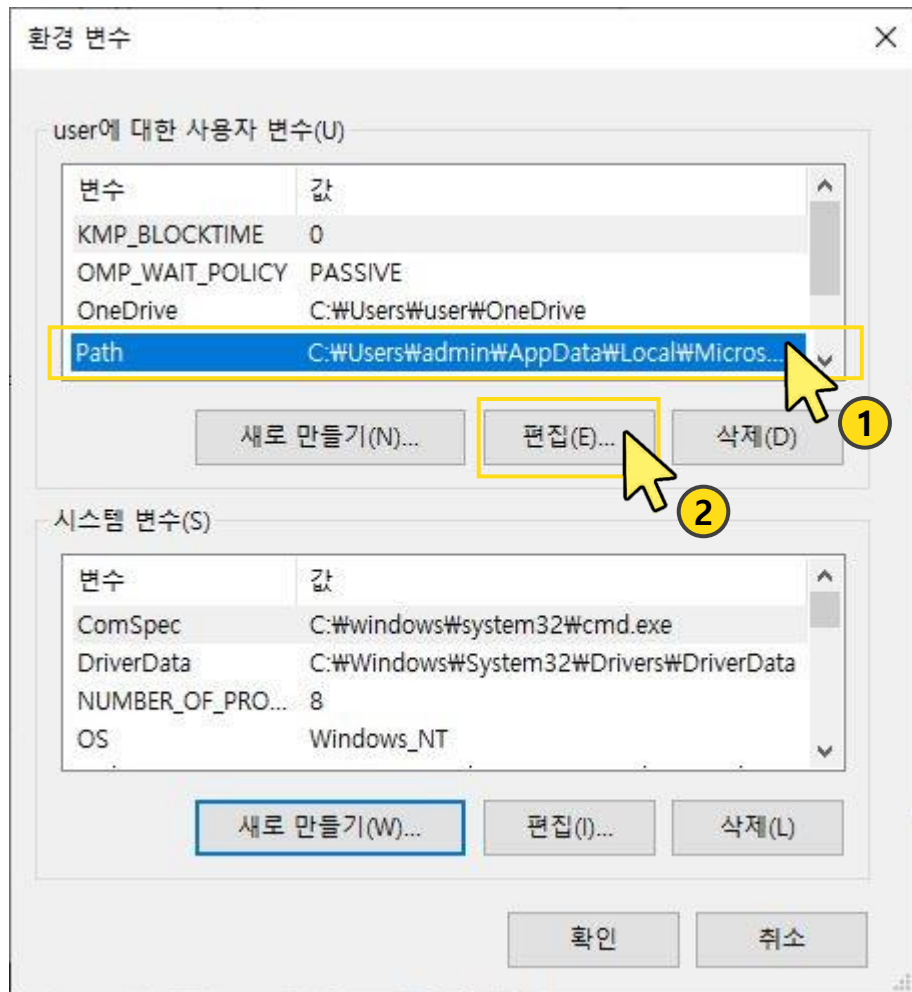
환경설정



1. 시작 옆의 검색창에서 '시스템 환경 변수' 를 검색합니다.
2. [시스템 환경 변수 편집]을 실행합니다.
3. [환경 변수]를 클릭합니다.



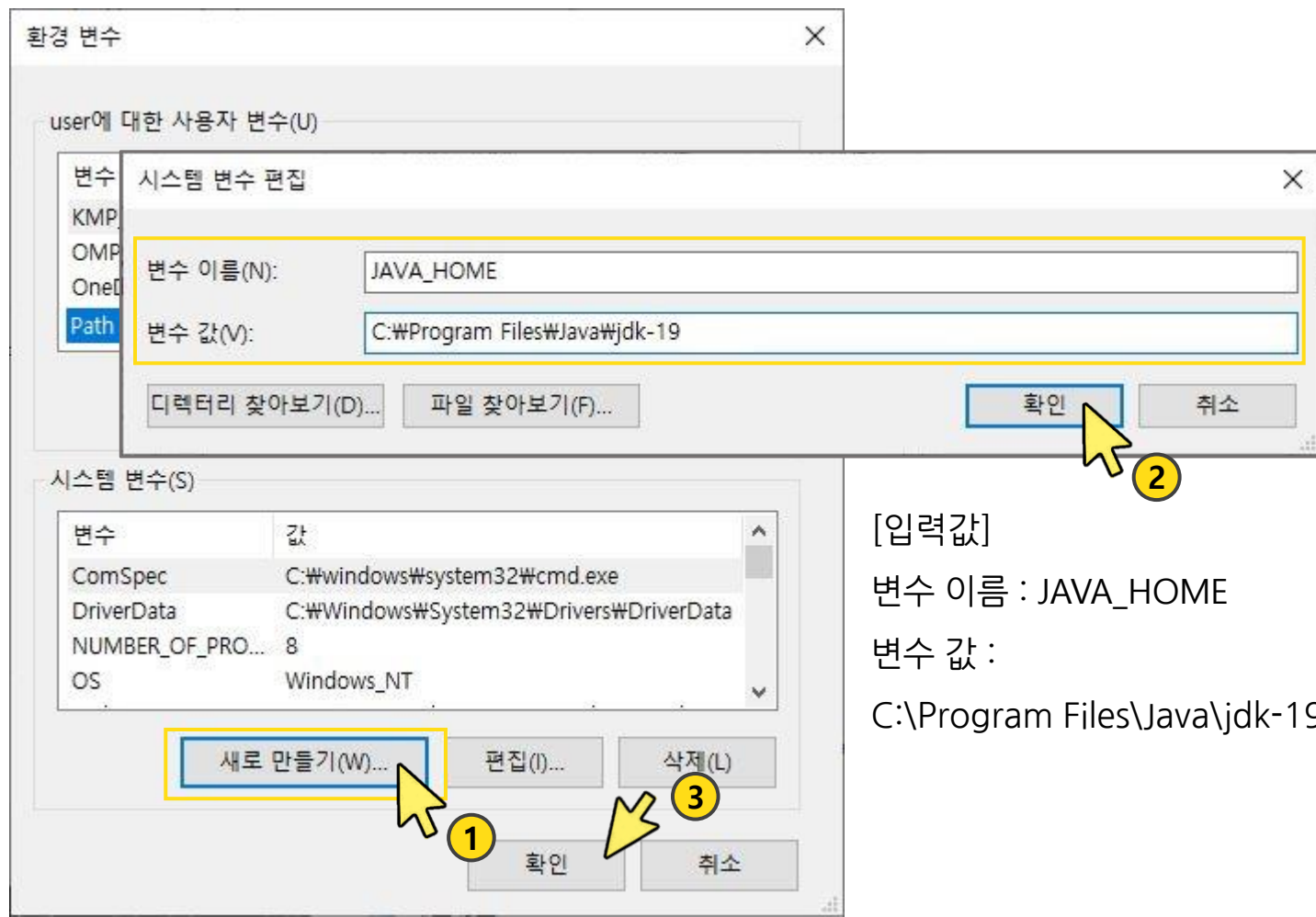
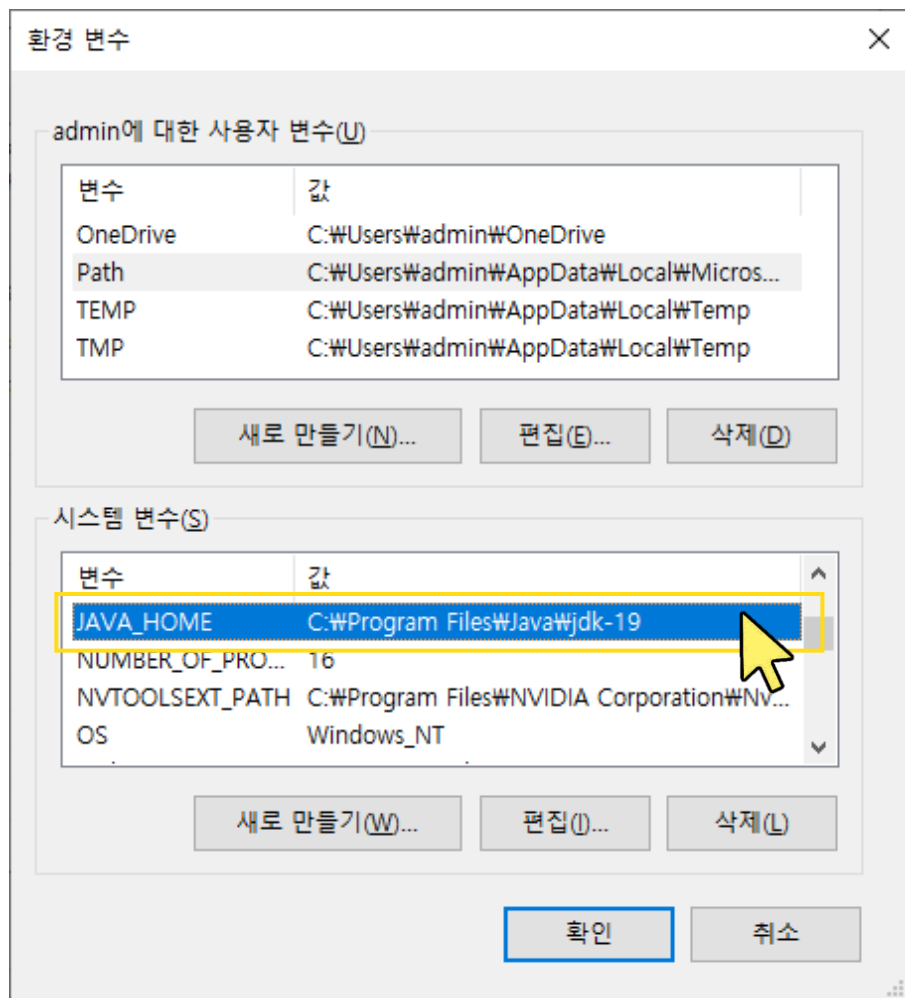
2.2 Java 환경변수 - [2]



입력값 : %JAVA_HOME%bin

2.2 Java 환경변수 - [3]

- JAVA_HOME 변수가 있는지 확인합니다. 만약에 없다면 새로 만들기로 추가합니다.



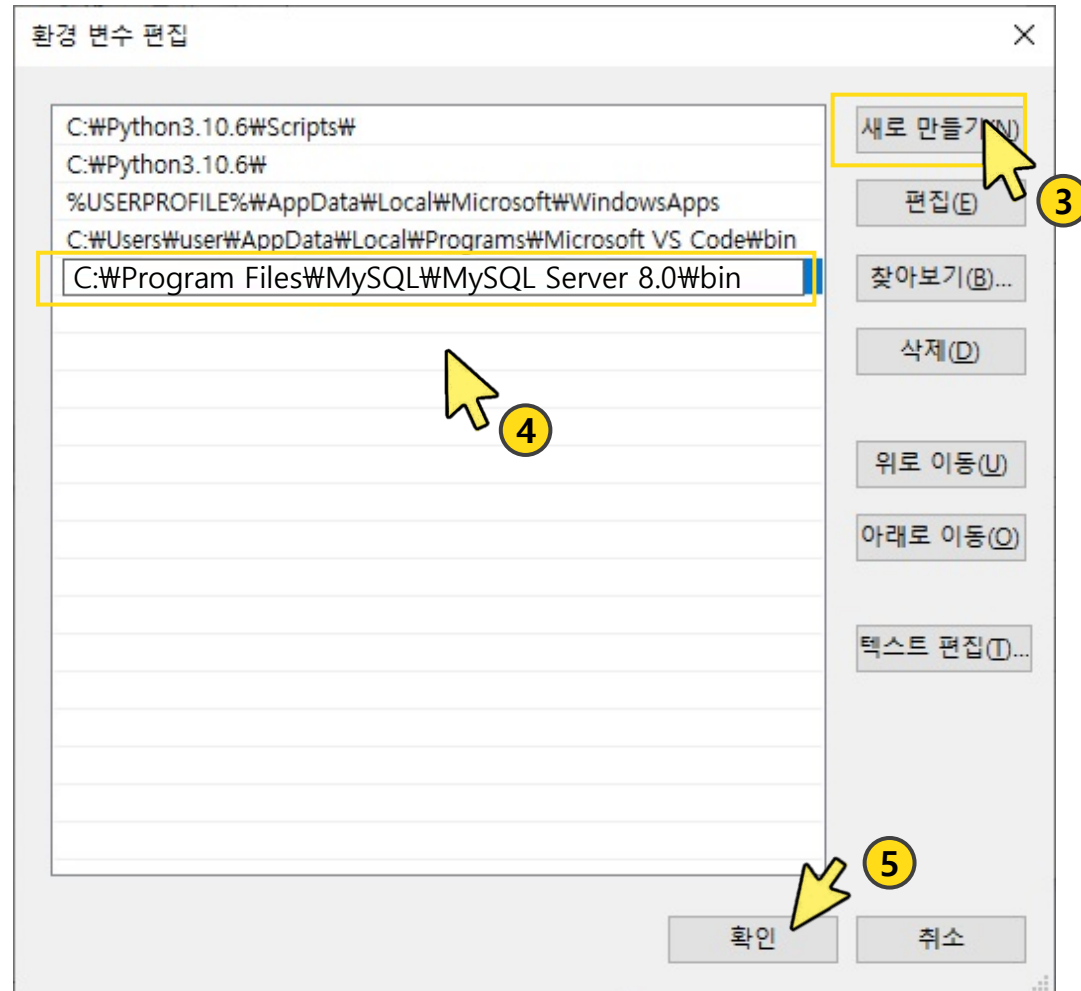
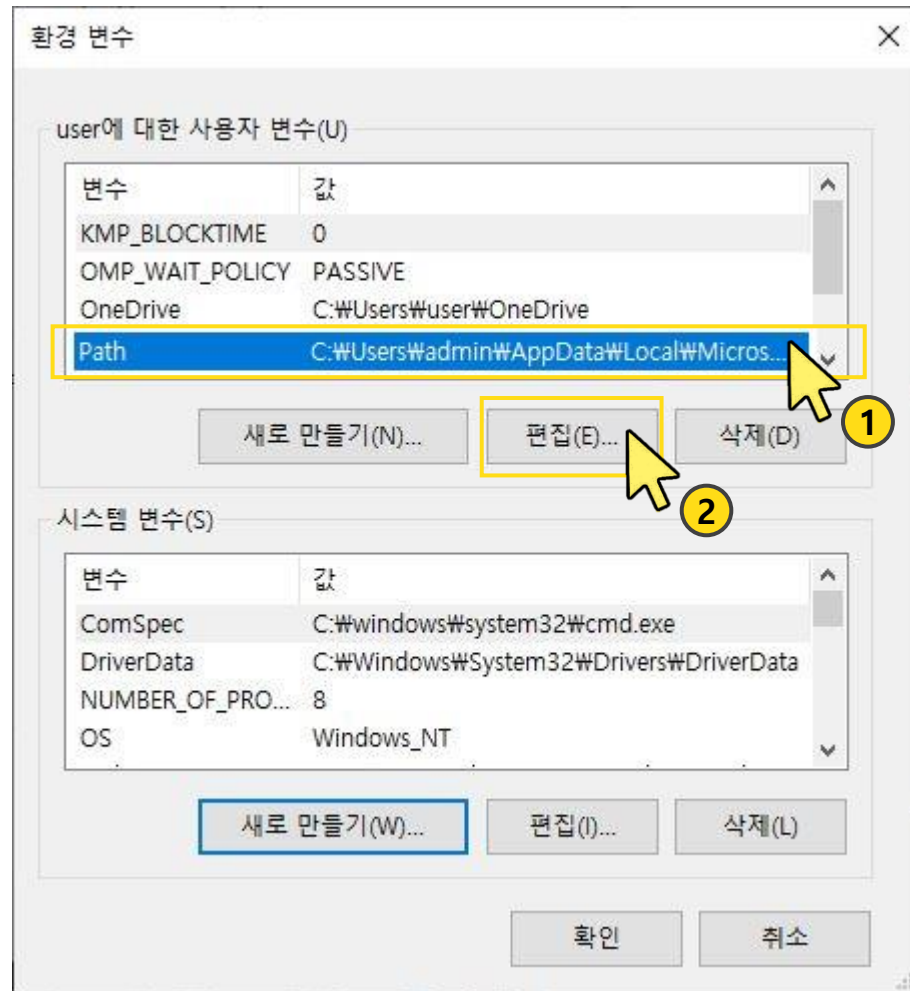
[입력값]

변수 이름 : JAVA_HOME

변수 값 :

C:\Program Files\Java\jdk-19

2.3 MySQL 환경변수

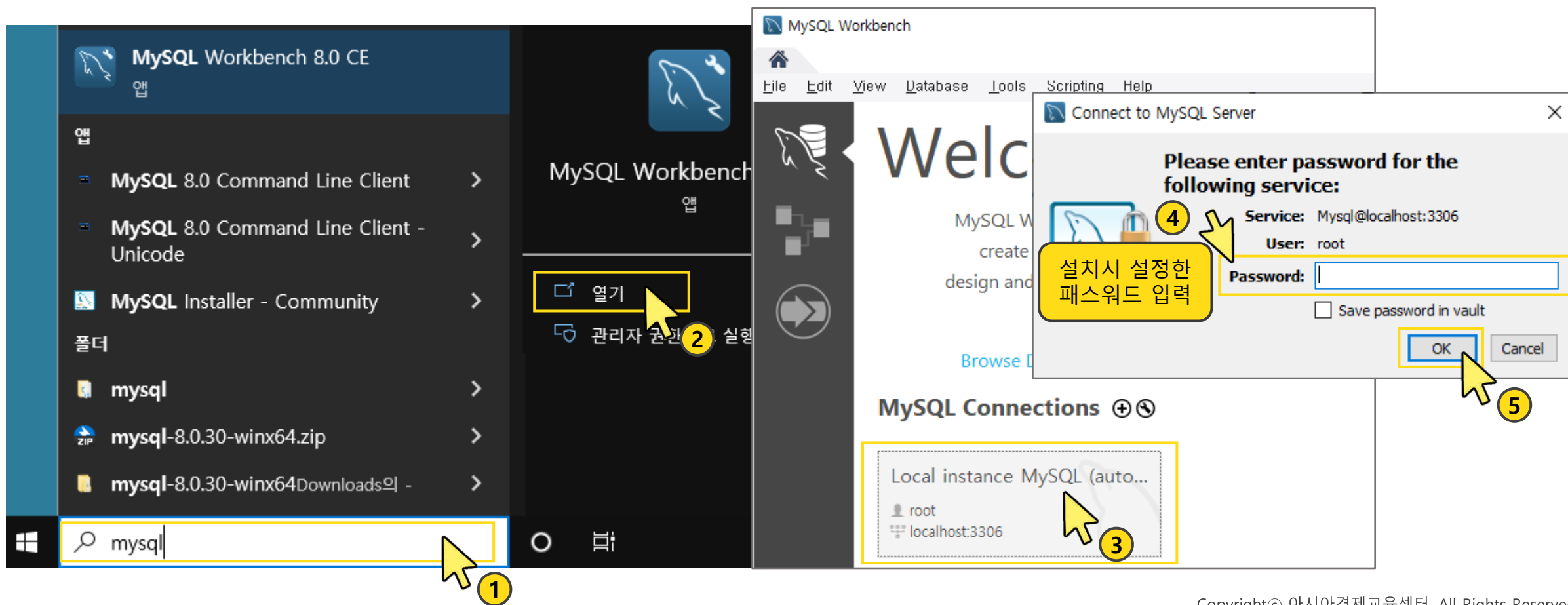


입력값 : C:\Program Files\MySQL\MySQL Server 8.0\bin

2.4 MySQL Data импорт - [1]

환경설정

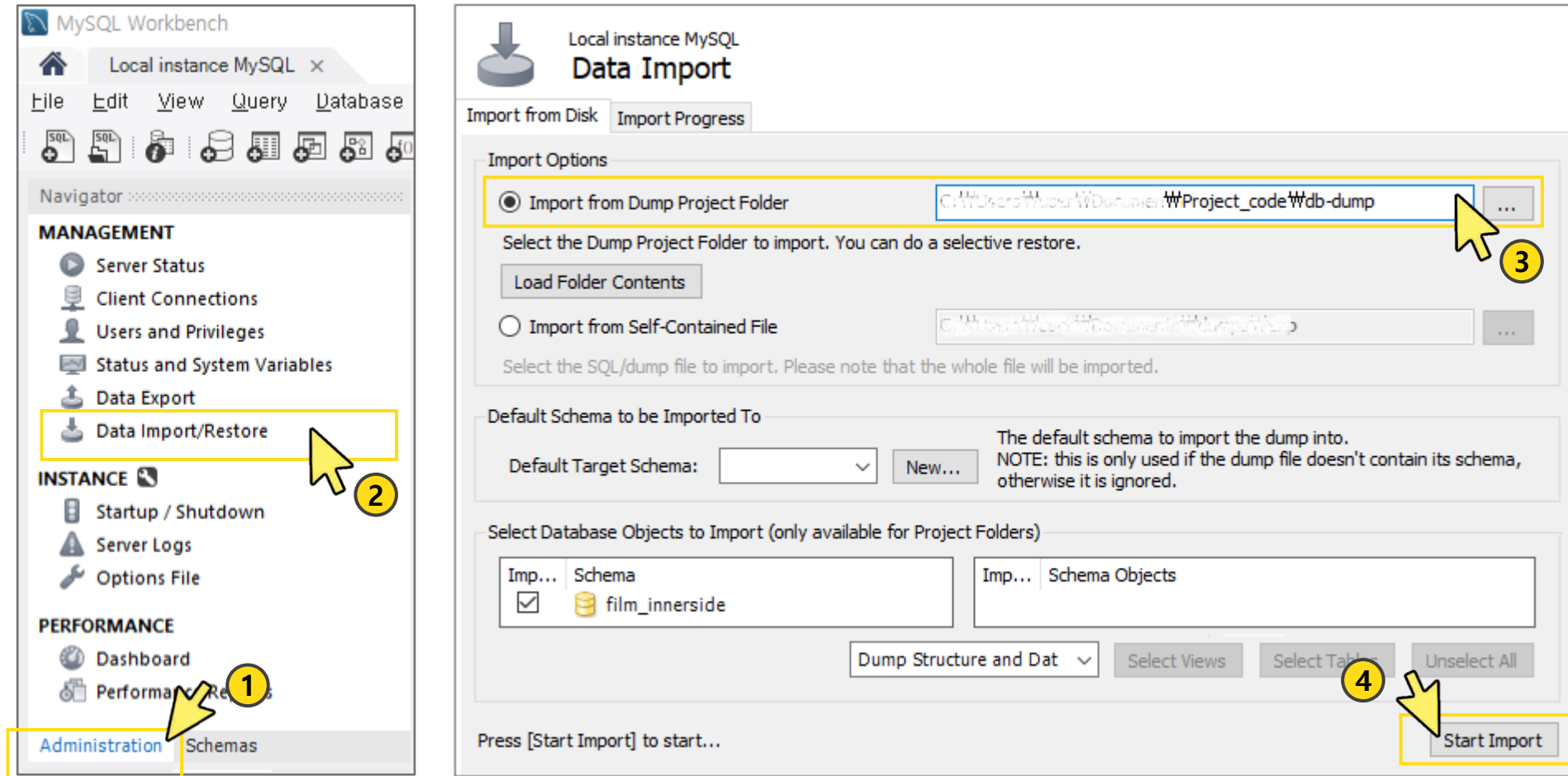
- MySQL Data 파일위치 : Project_code\db-dump 폴더
- 데이터 수집을 일부만 진행한 경우, 초기 데이터베이스를 복원하여 '네이버 영화리뷰의 감성분석' 을 진행하실 수 있습니다.



2.4 MySQL Data импорт - [2]

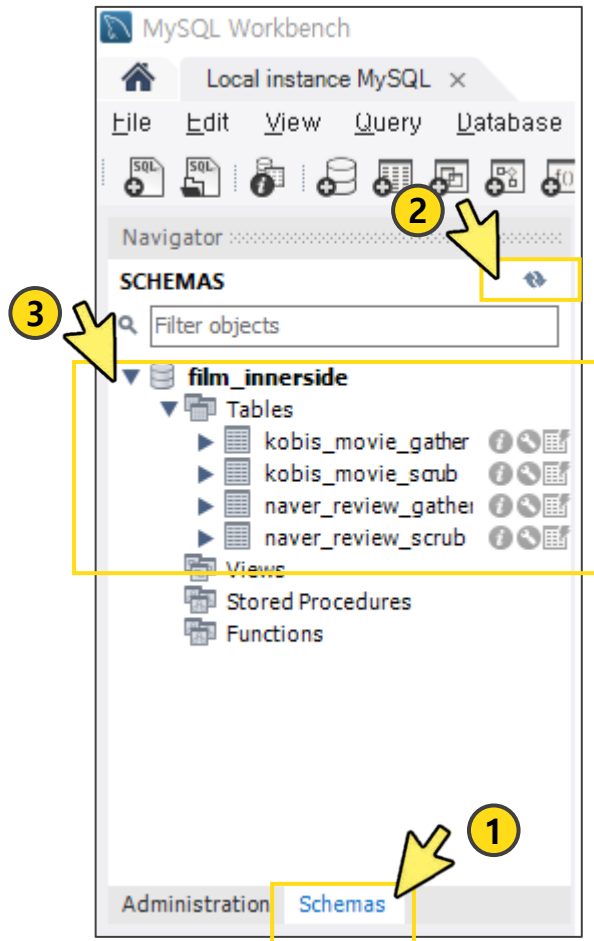
환경설정

- MySQL Data 파일위치 : Project_code\db-dump 폴더
- 데이터 수집을 일부만 진행한 경우, 초기 데이터베이스를 복원하여 '네이버 영화리뷰의 감성분석' 을 진행하실 수 있습니다.



2.4 MySQL Data импорт - [3]

- Data가 정상적으로 импорт되면, 아래와 같이 “film_innerside” DB에 4개의 테이블과 데이터를 확인하실 수 있습니다.



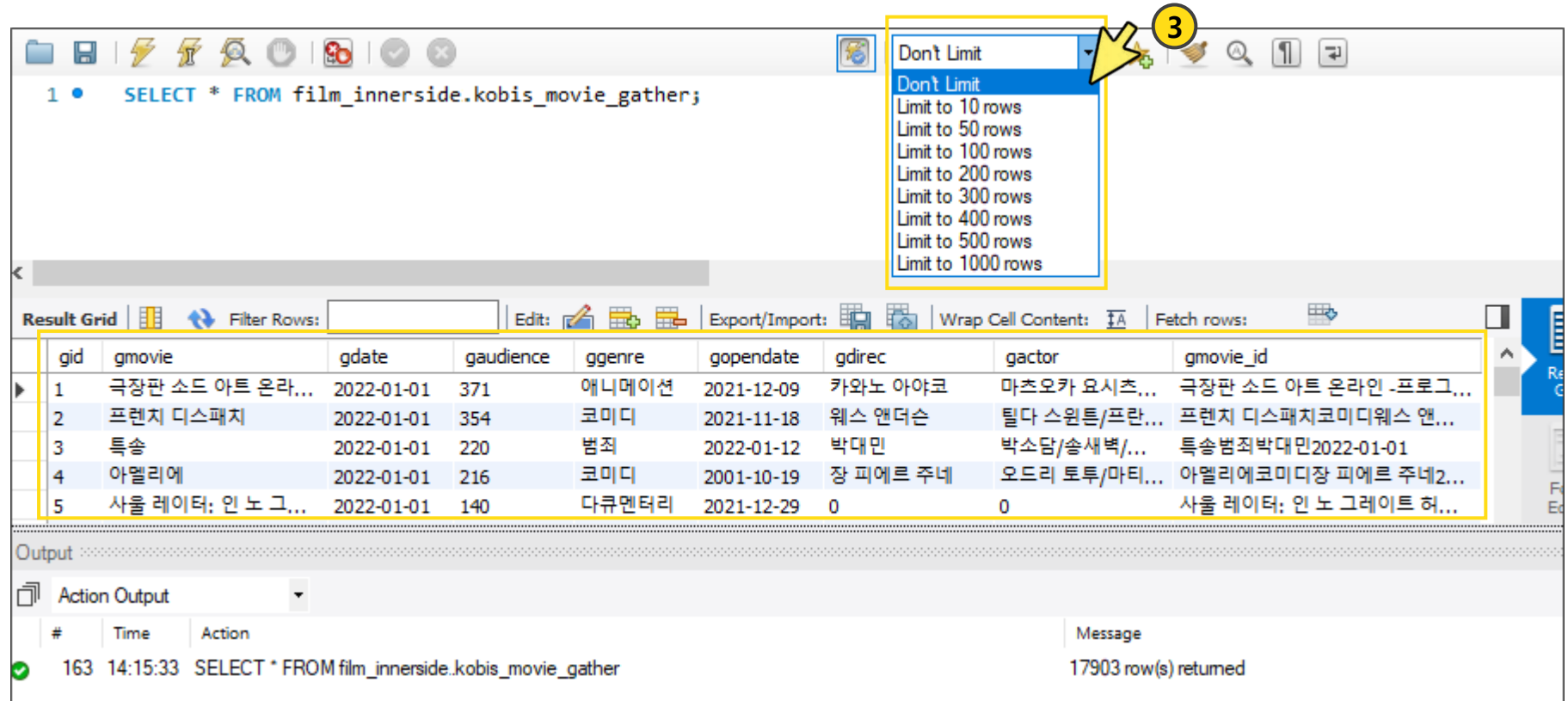
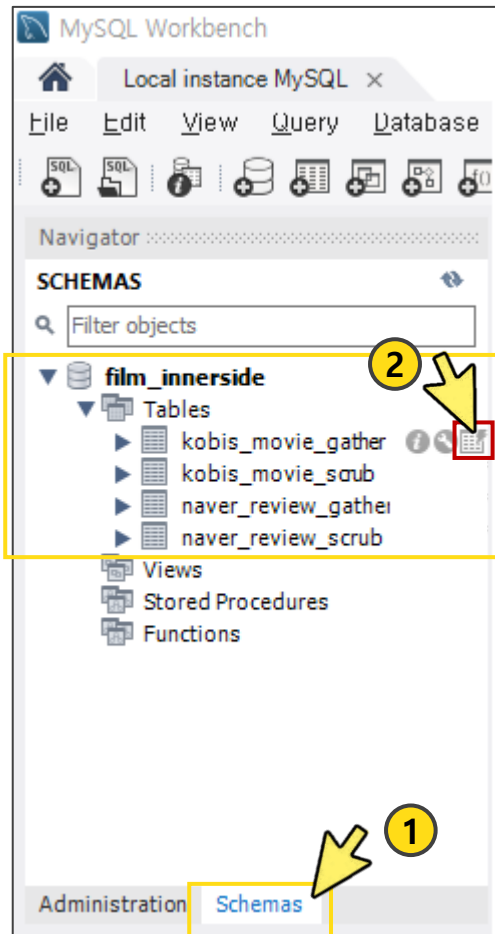
- ✓ DB명: **film_innerside**
- ✓ 테이블명 : **kobis_movie_gather** (KOBIS 수집 데이터)
데이터 : 17,903row
- ✓ 테이블명 : **kobis_movie_scrub** (KOBIS 전처리 데이터)
데이터 : 17,449row
- ✓ 테이블명 : **kobis_review_gather** (NAVER 영화리뷰 수집 데이터)
데이터 : 30,000row
- ✓ 테이블명 : **kobis_review_scrub** (NAVER 영화리뷰 전처리 데이터)
데이터 : 25,000row

2.5 DB 테이블 내용확인 - [1]

프로젝트 구조

- 테이블명 : kobis_movie_gather

설명: KOBIS 수집 데이터 / 데이터 : 17,903row

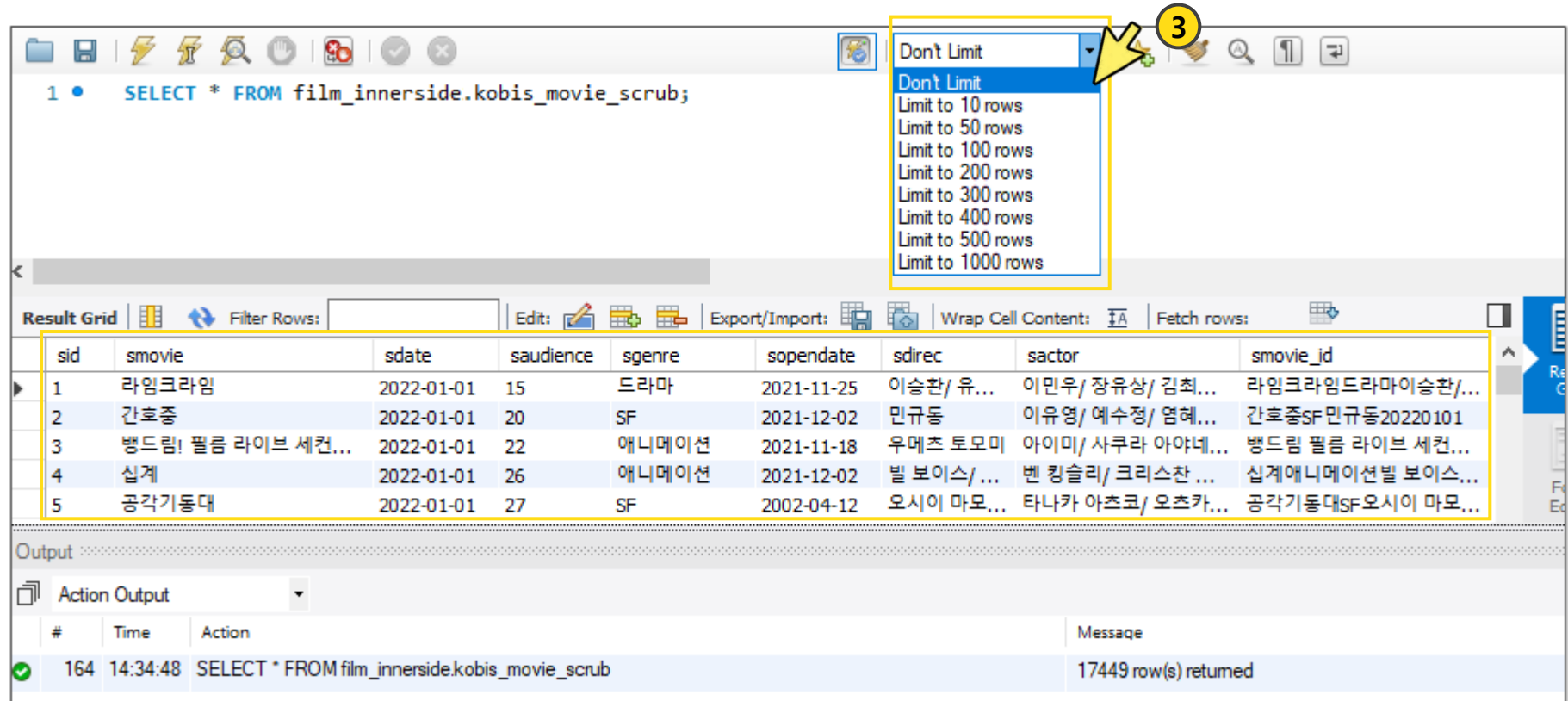
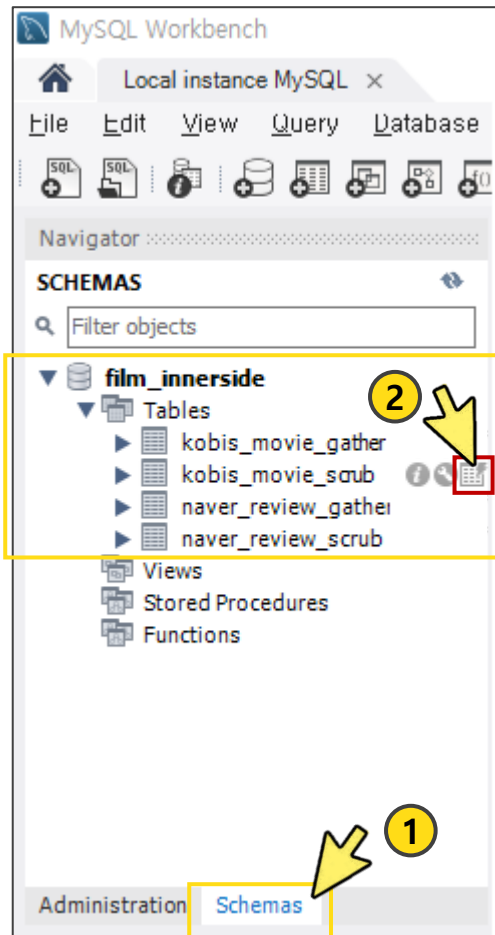


2.5 DB 테이블 내용확인 - [2]

프로젝트 구조

- 테이블명 : kobis_movie_scrub

설명: KOBIS 전처리 데이터 / 데이터 : 17,449row

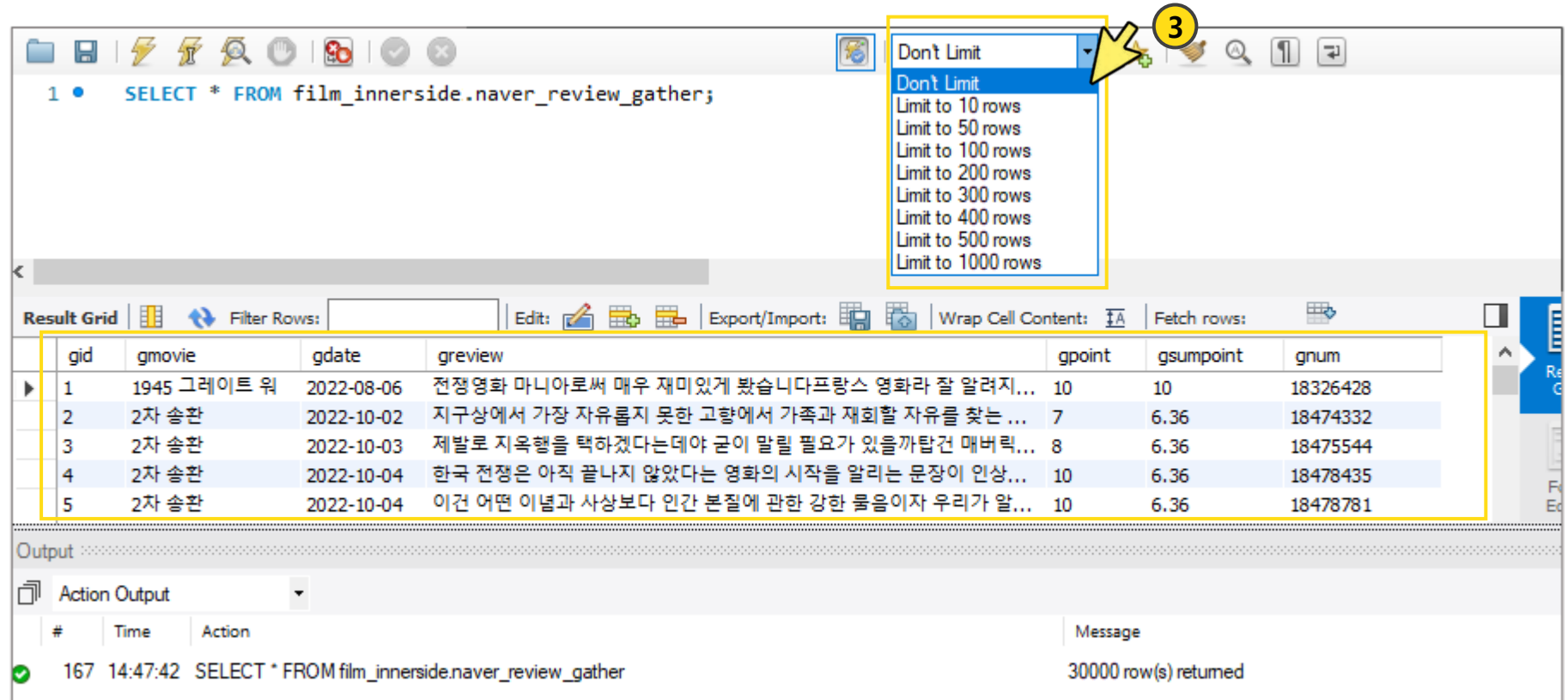
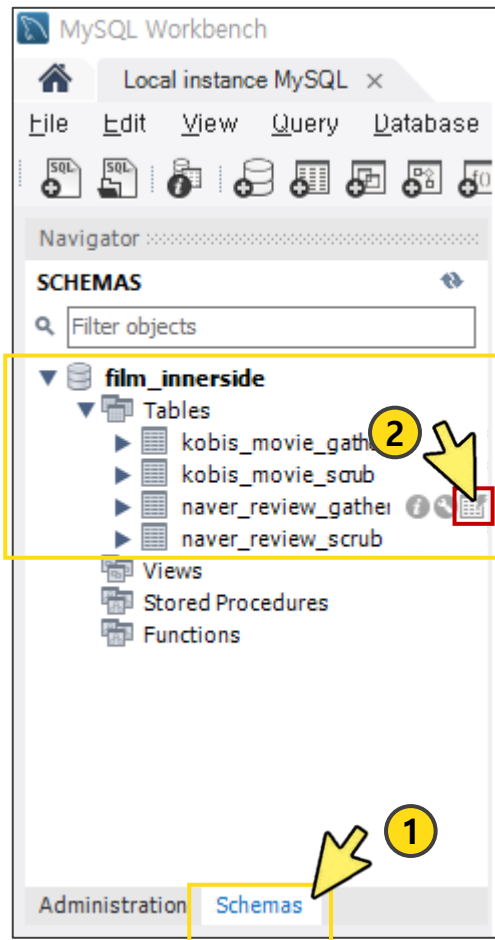


2.5 DB 테이블 내용확인 - [3]

프로젝트 구조

- 테이블명 : kobis_review_gather

설명: NAVER 영화리뷰 수집 데이터 / 데이터 : 30,000row



2.5 DB 테이블 내용확인 - [4]

프로젝트 구조

- 테이블명 : kobis_review_scrub

설명: NAVER 영화리뷰 전처리 데이터 / 데이터 : 25,000row










The screenshot shows the MySQL Workbench interface. On the left, the 'SCHEMAS' pane displays the 'film_innerside' database with its tables. A yellow box highlights the 'Tables' section, and a yellow arrow points to the 'naver_review_scrub' table, labeled with a yellow circle '2'. At the bottom, the 'Administration' pane shows the 'Schemas' tab selected, with a yellow arrow pointing to it, labeled with a yellow circle '1'. The main query editor shows the SQL query: `SELECT * FROM film_innerside.naver_review_scrub;`. A dropdown menu is open, showing options to limit the number of rows returned, with 'Don't Limit' selected. A yellow arrow points to this menu, labeled with a yellow circle '3'. Below the query editor, the 'Result Grid' displays the first five rows of the table. The 'Output' pane at the bottom shows the execution result: '25000 row(s) returned'.

sid	smovie	sdate	sreview	spoint	ssumpoint	snum	label	positive
1	3.5교시	2022-04-19	시계/하나로/지배/하다/수/좋다/타임스/트/래/치/처음/들다/보다/용다/...	10	5.89	18034334	1	0.999731
2	3.5교시	2022-04-17	여주/연기/너무/부자연스럽다/스토리/집중/도/안되다	1	5.89	18032266	0	0.140458
3	3.5교시	2022-04-17	하다/연기	3	5.89	18031994	0	0.667555
4	3.5교시	2022-04-16	조아/연기/좋다/예쁘다/작/기대/되다/조아/땀/별개	9	5.89	18030731	1	0.991799
5	3.5교시	2022-04-15	교시	3	5.89	18029162	0	0

#	Time	Action	Message
✓ 176	14:53:05	SELECT * FROM film_innerside.naver_review_scrub	25000 row(s) returned

3.1 프로젝트 구조 - [1]

프로그램 구조

폴더명 / 파일명	설명
 Document	프로젝트 관련 기획 문서
 note	회의록
 프로젝트 계획서	
 시스템구조 설계서	
 요구사항 정의서	
 테이블 리스트	
 테이블 정의서	
 개발환경 구축 가이드	
 Environment	설치파일
python-3.10.6-amd64.exe	
jdk-19_windows-x64_bin.exe	
mysql-installer-community-8.0.30.0.msi	
VSCodeUserSetup-x64-1.69.2.exe	
cuda_11.8.0_windows_network.exe	
cudnn-windows-x86_64-8.5.0.96_cuda11-archi...	

폴더명 / 파일명	설명
 Project_code	소스 코드 파일
 scrub_main.py	*실행파일
 dataScrub.py	
 dbProc.py	
 dataGather.py	*실행파일
 kobis_crawling.py	*실행파일
 naver_review_crawling.py	*실행파일
 data	불용어 등 필요 자료
 db-dump	DB 백업 파일
film_innerside_routines.sql	
film_innerside_kobis_movie_gather.sql	
film_innerside_kobis_movie_scrub.sql	
film_innerside_naver_review_gather.sql	
film_innerside_naver_review_scrub.sql	
 chrometemp	크롤러 실행시 생성되는 임시 파일 폴더로 삭제가능

3.1 프로젝트 구조 - [2]

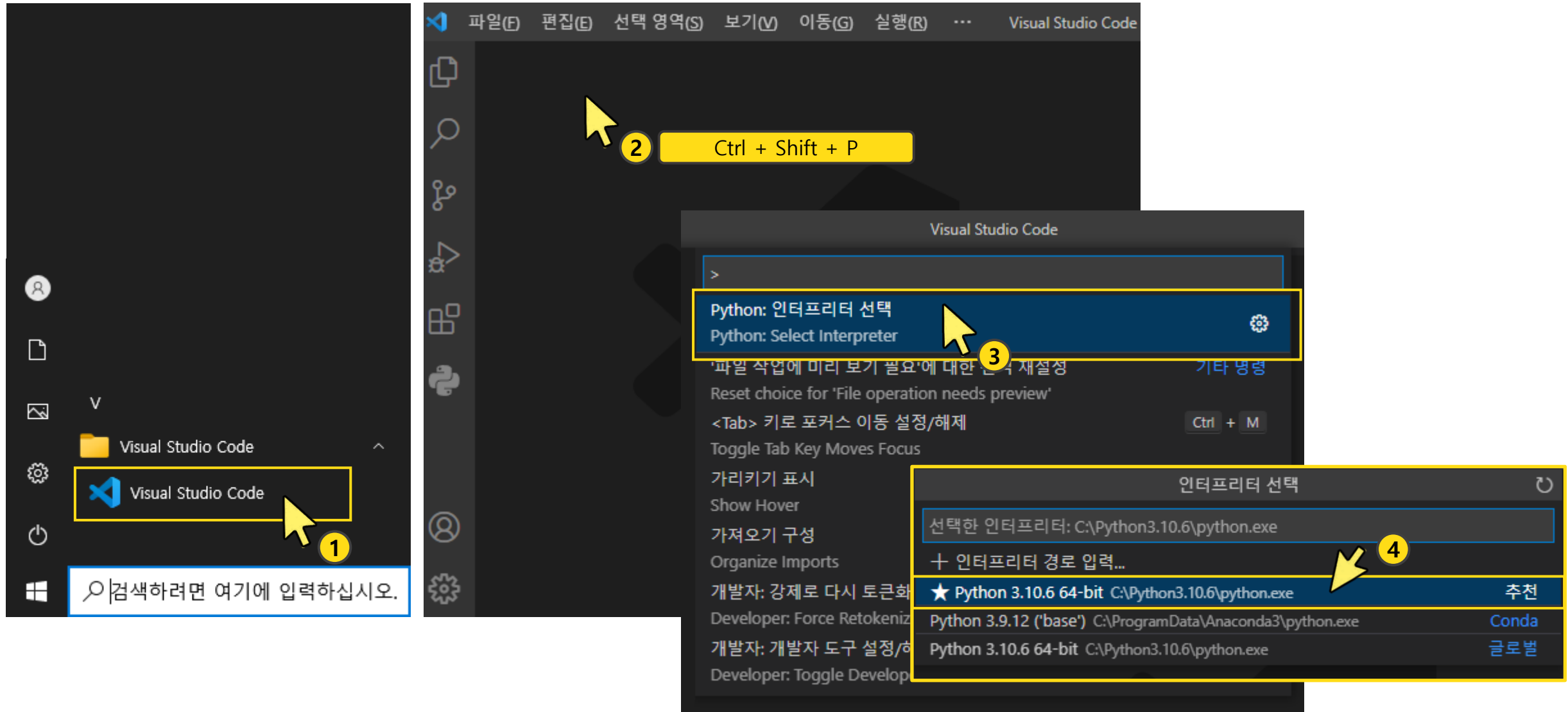
[소스 코드 파일의 기능 설명]

📁 Project_code

📄 scrub_main.py	데이터 전처리 및 감성분석 실행 메뉴 선택
📄 dataScrub.py	네이버 영화리뷰 데이터 전처리 / 감성분석 / 긍정지수 예측 / KOBIS 데이터 전처리
📄 dbProc.py	전처리 과정에서의 DB 연결 및 저장
📄 dataGather.py	데이터 수집 과정에서의 DB 연결 및 저장
📄 kobis_crawling.py	KOBIS 영화관 입장권 데이터 수집
📄 naver_review_crawling.py	네이버 영화 리뷰 데이터 수집

4.1 VS Code 인터프리터 선택

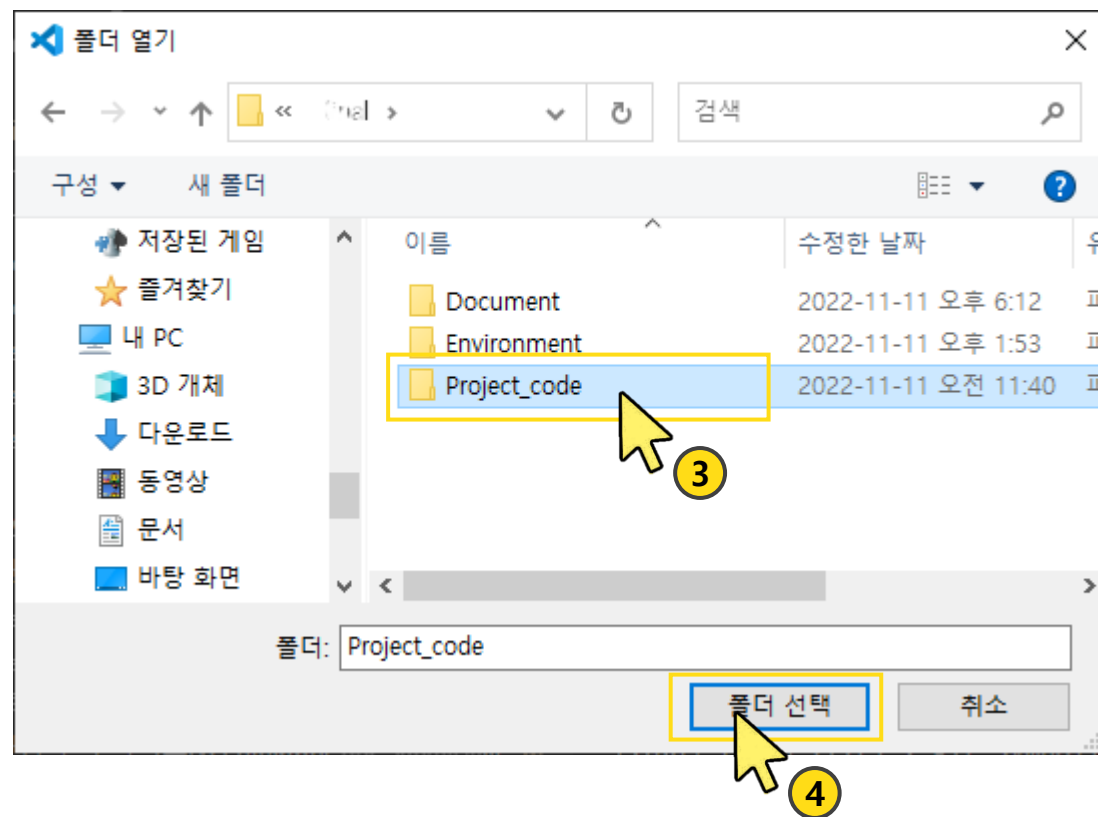
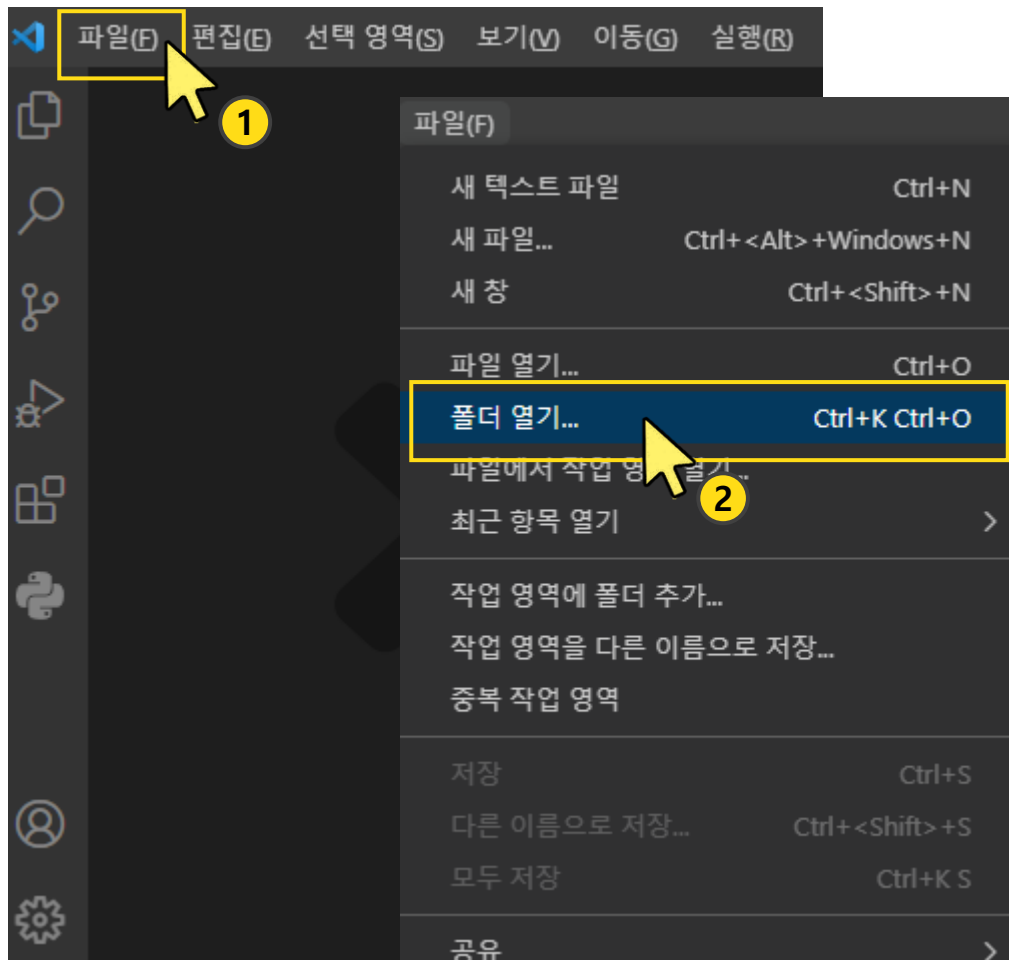
프로그램 실행



4.2 실행메뉴 확인 - [1]

프로그램 실행

- 폴더 열기 : Project_code



4.2 실행메뉴 확인 - [2]

프로그램 실행

[NAVER 영화 리뷰 데이터 수집 크롤러]

- 실행파일 : **naver_review_crawling.py**
- 기능 : 1. NAVER 영화 크롤링하여 데이터 수집
2. DB 저장 (테이블 : naver_review_gather)
3. CSV 파일 저장 (naver_review_gather.csv)

```
-----
총 14,366,977개의 평점이 있습니다.
<총 14,366,977 건 수집완료 되었습니다.>

<NAVER 리뷰 크롤링이 완료되었습니다.>

-----
D:\prj_work2
위 경로에 파일을 저장합니다.
```

[KOBIS 영화관 입장권 데이터 수집 크롤러]

- 실행파일 : **kobis_crawling.py**
- 기능 : 1. KOBIS 크롤링하여 데이터 수집
2. DB 저장 (테이블 : kobis_movie_gather)
3. CSV 파일 저장 (kobis_movie_gather.csv)

```
-----
<KOBIS 일별 박스오피스 크롤링을 시작합니다.>

<KOBIS 일별 박스오피스 크롤링이 완료되었습니다.>

-----
D:\prj_work2
위 경로에 파일을 저장합니다.
```

4.2 실행메뉴 확인 - [3]

[수집 데이터 DB 저장]

- 실행파일 : **dataGather.py**
- 기능 : 1. DB 연결
2. DB 저장
네이버 리뷰 (테이블: naver_review_gather)
KOBIS (테이블: kobis_movie_gather)

```
'film_innerside' 데이터베이스 생성  
'naver_review_gather' 테이블 생성  
'naver_review_scrub' 테이블 생성  
'kobis_movie_gather' 테이블 생성  
'kobis_movie_scrub' 테이블 생성
```

```
naver_review.csv  
👉 'naver_review_gather' 테이블에 저장
```

```
kobis_movie.csv  
kobis_rank.csv  
👉 'kobis_movie_gather' 테이블에 저장
```


4.2 실행메뉴 확인 - [5]

프로그램 실행

- 실행파일 : scrub_main.py
- ## NAVER 영화 리뷰 : 데이터 전처리 및 감성분석 ##
[STEP 2] 단어 토큰화 및 불용어 삭제
 - 기능 : 1. DB 접속후 데이터 로딩
 - 2. 단어 집합화
 - 3. 데이터 전처리
 - 4. 단어 시퀀스 정제 및 벡터화

```
*** [ STEP 2 ] 단어 토큰화 및 불용어 삭제 ***
STEP 2를 진행하시겠습니까? (진행:Y) y

>>>> Connecting... 데이터베이스 접속

>>>> Loading... 네이버 영화 리뷰 데이터
100%|████████████████████████████████████████| 29697/29697 [00:35<00:00, 840.33it/s]

>>>> Vocabulary... 단어 집합화

[*] 단어 집합(vocabulary)의 총 크기 : 16534
[*] 빈도가 2번 이하인 단어의 수: 9433
[*] 전체에서 희귀 단어의 빈도 비율 : 3.10%
[*] 사용할 단어 집합의 크기 : 7102

>>>> Vocabulary... 낮은 빈도 단어 정제
>>>> Integer Encoding... 단어의 정수 시퀀스 정제
>>>> Vocabulary... 낮은 빈도 단어의 관련 데이터 정제
>>>> Vectorizing... 정수의 벡터 정제
```

4.2 실행메뉴 확인 - [6]

프로그램 실행

- 실행파일 : scrub_main.py
- ## NAVER 영화 리뷰 : 데이터 전처리 및 감성분석 ##
[STEP 3] 모델 훈련 및 검증
 - 실행 예외 사항: STEP2를 실행하지 않으면,
STEP3도 실행되지 않습니다.
 - 기능 : 1. 모델 정의 및 모델 컴파일
 - 2. 훈련과 검증 데이터 셋팅 (8:2)
 - 3. 모델 훈련 및 검증 (Epoch 15)
 - 4. 결과를 손실율, 정확도 그래프로 확인
 - 5. 모델링 - 훈련 및 검증 재실행 (Epoch 2)
 - 6. 결과 확인 - 손실율, 정확도 수치
 - 7. 긍정지수 예측
 - 8. DB 저장

```
*** [ STEP 3 ] 모델 훈련 및 검증 ***

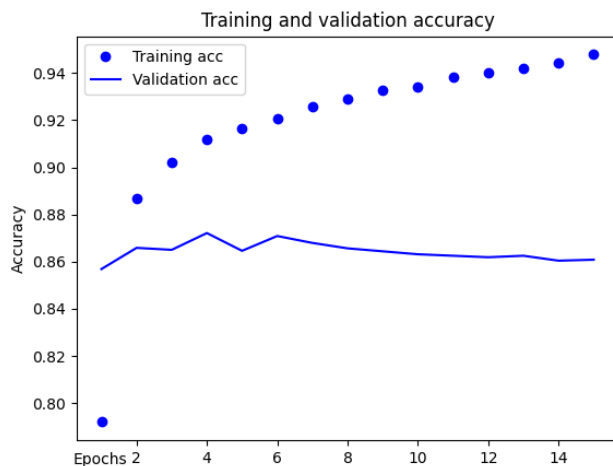
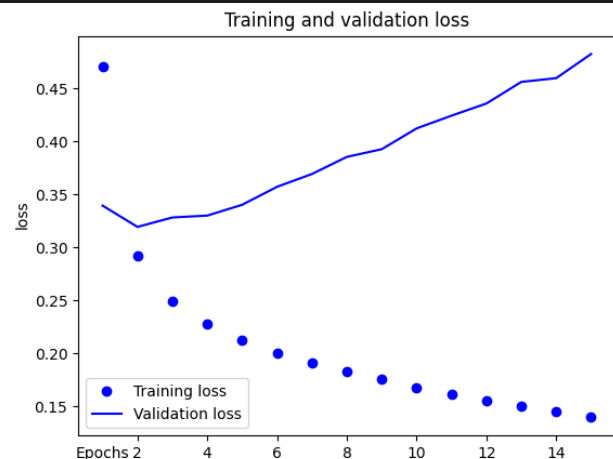
>>>> Modeling... 모델 정의
>>>> Modeling... 모델 컴파일
>>>> Modeling... 데이터 셋팅
훈련데이터 -리뷰 개수: 19115
검증데이터 -리뷰 개수: 4778

>>>> Modeling... 모델 훈련 및 검증
Epoch 1/15
155/155 [=====] - 3s 9ms/step - loss: 0.4707 - accu
racy: 0.7920 - val_loss: 0.3395 - val_accuracy: 0.8568
Epoch 2/15
155/155 [=====] - 1s 5ms/step - loss: 0.2922 - accu
racy: 0.8867 - val_loss: 0.3194 - val_accuracy: 0.8658
Epoch 3/15
155/155 [=====] - 1s 5ms/step - loss: 0.2493 - accu
racy: 0.9022 - val_loss: 0.3283 - val_accuracy: 0.8650
Epoch 4/15
155/155 [=====] - 1s 6ms/step - loss: 0.2281 - accu
racy: 0.9118 - val_loss: 0.3301 - val_accuracy: 0.8721
:
:
```

4.2 실행메뉴 확인 - [7]

프로그램 실행

>> 결과확인 : 손실과 정확도



>> 결과확인 : Overfitting (과대적합), 에크모를 2회로 조정 후 재실행

>>>> Modeling... 모델 정의

>>>> Modeling... 모델 컴파일

>>>> Modeling... 모델 훈련 및 검증

Epoch 1/2

38/38 [=====] - 1s 11ms/step - loss: 0.6134 - accuracy: 0.8026

Epoch 2/2

38/38 [=====] - 0s 9ms/step - loss: 0.4528 - accuracy: 0.8724

150/150 [=====] - 0s 1ms/step - loss: 0.3553 - accuracy: 0.8686

>> 결과확인 : 손실율 35.53%

>> 결과확인 : 정확도 86.86%

>>>> Uploading... DB 저장

100% | 23893/23893 [00:20<00:00, 1152.24it/s]

4.2 실행메뉴 확인 - [8]

프로그램 실행

- 실행파일 : scrub_main.py
- ## KOBIS 영화관 입장권 통합 : 데이터 전처리 ##
[STEP 1] 영화관 입장권 수집 데이터 전처리
- 기능 : 1. DB 접속후 데이터 로딩
 2. 데이터 전처리
 3. DB 저장

```
##### KOBIS 영화관 입장권 통합 : 데이터 전처리 #####

*** [ STEP 1 ] 영화관 입장권 수집 데이터 전처리 ***
STEP 1을 진행하시겠습니까? (진행:Y) Y

>>>> Connecting... 데이터베이스 접속

>>>> Loading... KOBIS 영화관 입장권 데이터
100%|████████████████████████████████████████| 17903/17903 [00:43<00:00, 415.06it/s]
>>>> Refinementing... 여백 정제
>>>> Refinementing... 결측치 정제
>>>> Uploading... DB 저장
100%|████████████████████████████████████████| 17903/17903 [00:12<00:00, 1379.64it/s]
```

QnA

질문이 있다면 말씀해주세요.

Thank You!

