

# ADVANCED TEXT MINING

---

by FINGEREDMAN ([fingeredman@gmail.com](mailto:fingeredman@gmail.com))

APPENDIX 01

# Lecture Guide

---

## Participation

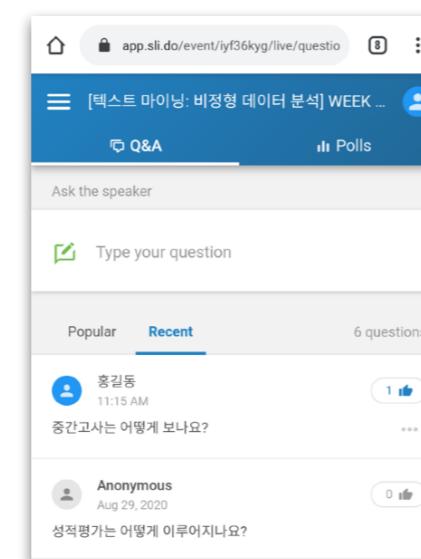
Q&A, Survey, Quiz

with SLIDO

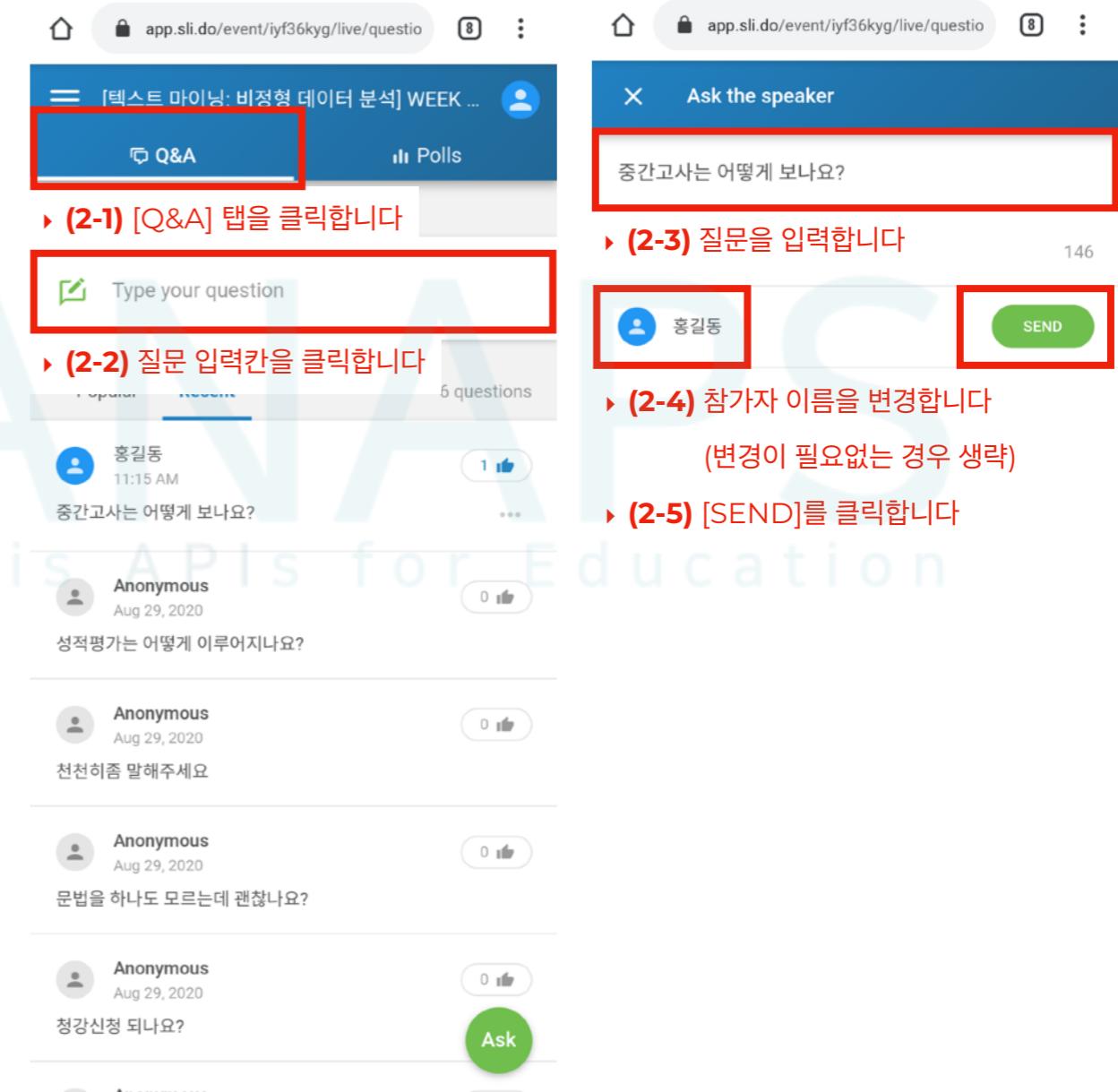
<https://sli.do>

# Q&A, Quiz, Survey with Slido

## Slido 접속하기 (PC, 모바일 모두 가능)



## Q&A 참여하기



# Q&A, Quiz, Survey with Slido

## Survey 참여하기

The screenshot shows a poll titled "Python 언어를 접해본 경험이 있나요?" with two options: "있다." (Yes) and "없다." (No). The "No" option is selected. A red box highlights the "Polls" tab at the top of the screen. Below the poll, a message says "▶ (3-1) [Polls] 탭을 클릭합니다".

Live poll

▶ (3-1) [Polls] 탭을 클릭합니다

Python 언어를 접해본 경험이 있나요?

있다.

없다.

SEND

▶ (3-2) 현재 진행중인 설문에 답변을 선택합니다

▶ (3-3) [SEND]를 클릭해 답변을 제출합니다

▶ (3-4) 답변 수정이 필요한 경우 [EDIT RESPONSE]를 클릭합니다

## Quiz 참여하기

The screenshot shows a poll titled "데이터 분석을 해본 경험이 있나요?" with two options: "없다." (No) and "있다.". The "No" option is selected. A red box highlights the "Polls" tab at the top of the screen. Below the poll, a message says "▶ (4-1) [Polls] 탭을 클릭합니다".

[퀴즈] 1주차

▶ (4-1) [Polls] 탭을 클릭합니다

홍길동

JOIN

▶ (4-2) 참가자 이름을 변경합니다  
(변경이 필요없는 경우 생략)

▶ (4-3) [JOIN]를 클릭합니다

▶ (4-4) 진행자가 퀴즈를 시작할때  
까지 대기합니다

Get ready, 홍길동!  
1 user has already joined

# Q&A, Quiz, Survey with Slido

# Quiz 참여하기 - 정답체크 및 결과확인

The screenshots illustrate the following steps:

- (4-5) 현재 진행중인 설문에 답변을 선택합니다**
- (4-6) [SEND]를 클릭해 답변을 제출합니다**
- (4-7) 정답을 확인합니다**
- (4-8) 퀴즈 결과를 확인합니다.**
- (4-9) [View correct answers]를 클릭해 전체 퀴즈 결과를 확인합니다**

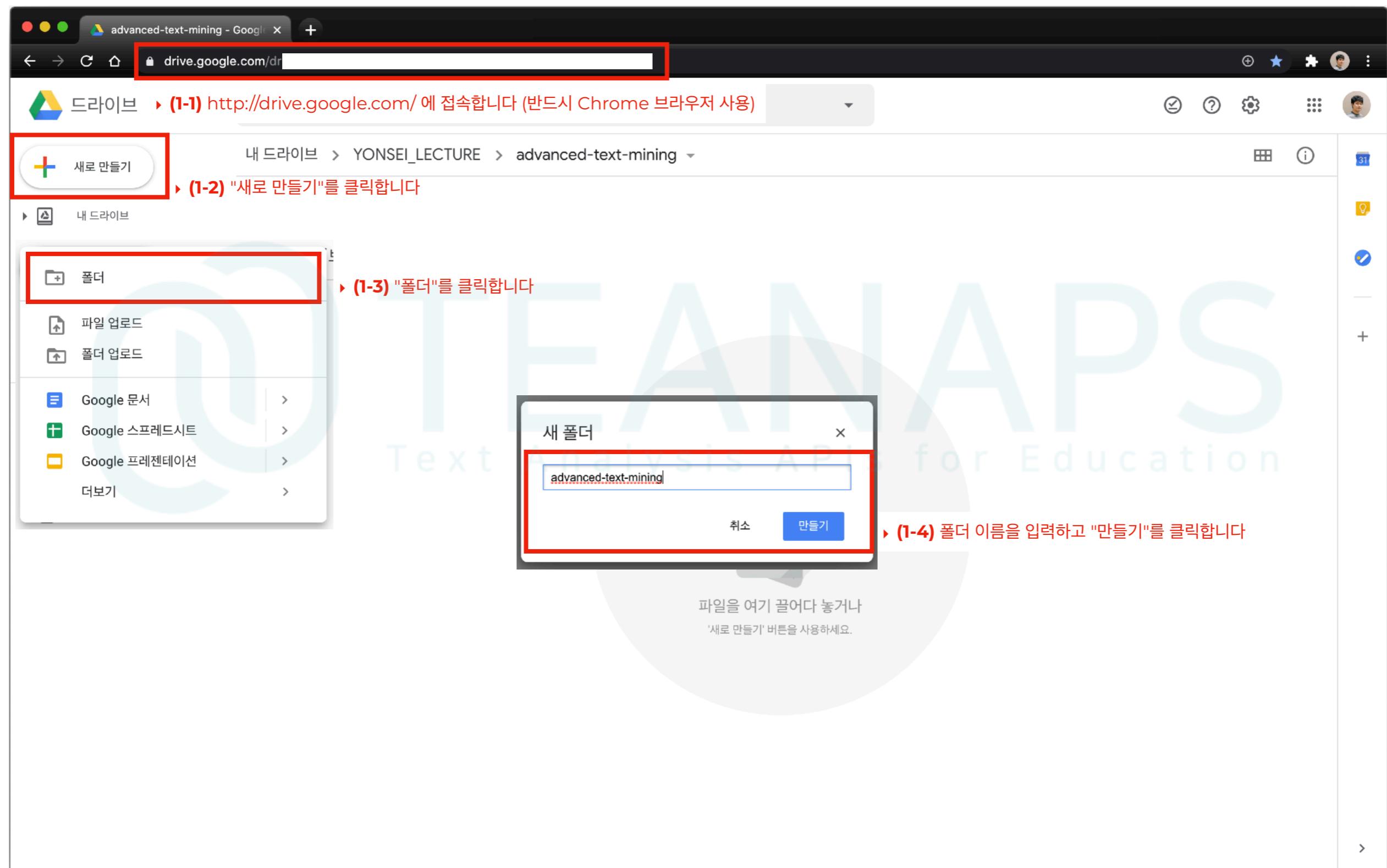
In the screenshots, the first poll shows a question with two options: "이거" (checked) and "저거". The second poll shows the results: "이거" is at 100% and "저거" is at 0%. The third poll shows the user finished 1st place with 1 correct answer out of 1, with a voting time of 0:10. The fourth poll shows the same results for the user.

# Colaboratory

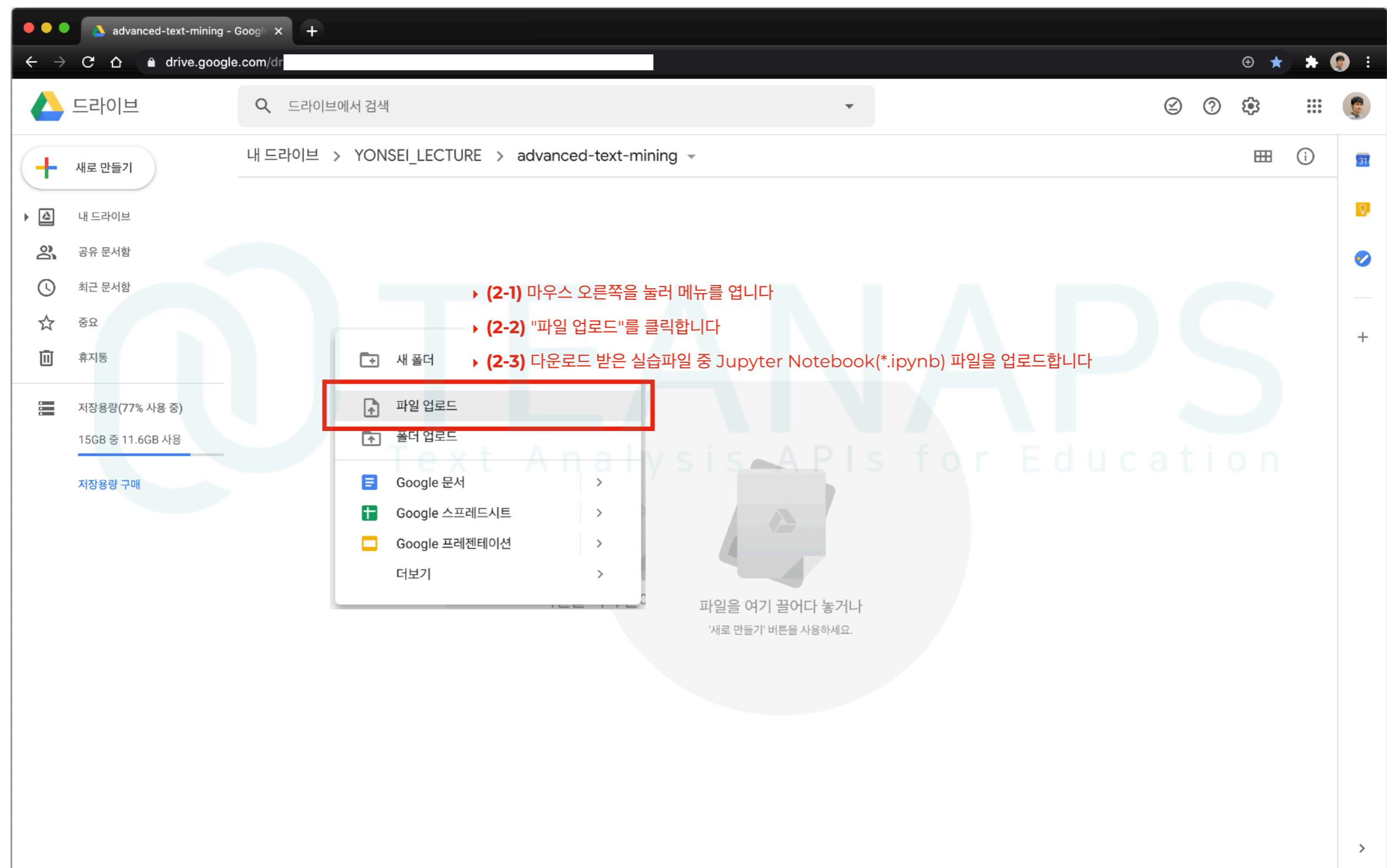
# Google Chrome + Colaboratory

---

# Google Colaboratory



# Google Colaboratory



# Google Colaboratory

The image shows a split-screen view. On the left is a screenshot of the Google Drive web interface. A red box highlights the file 'New\_Jupyter\_Notebook\_File.ipynb' in the 'advanced-text-mining' folder. Below the drive interface, two red bullet points provide instructions:

- ▶ (3-1) 업로드한 파일을 실행합니다
- ▶ (3-2) Google Colaboratory 확장프로그램 설치 안내창이 나오는 경우 동의 및 설치한 후 실행합니다

On the right is a screenshot of the Google Colaboratory environment. It shows a Jupyter Notebook cell with the Python code 'print("hello world!")' and the output 'hello world!'. A red box contains the following text:

▶ (3-3) Google Colaboratory 환경에서  
Jupyter Notebook 파일을 동일하게 활용 가능합니다

## WHAT THE ...

시작은 생각보다  
힘들고 고된 작업일 수 있습니다

# 실습환경 설치

## 실습환경 설치

- 본 강의 실습 진행을 개인 노트북에서 진행하기 위해 아래 1~3 과정을 메뉴얼에 따라 설치해야함
  - 1) 메뉴얼은 Windows 운영체제를 기준으로 제작되었으며, MAC 환경에서 필요한 부분은 별도로 표시함
  - 2) 모든 설치과정 진행 시 약 50 ~70분이 소요됨
  - 3) 설치에 문제가 있는 경우 PC 재부팅 후 다시 시도

실습환경	내용
Python	Python 3.6 Anaconda 설치 (최신 버전 통일을 위해 기존 설치 삭제 후 재설치 권장)
JAVA	Java jdk 설치
환경변수 설정	환경변수 설정 (Windows 운영체제만 해당함) - Python 환경변수 - JAVA 환경변수
Build Tool	Microsoft Visual C++ Build Tools 설치 (Windows 운영체제만 해당함)
Python Lib	Python Package 설치

# Python 3.6 Anaconda

The screenshot shows a web browser window with the URL <https://www.anaconda.com/download/> highlighted by a red box. The page title is "Downloads - Anaconda". Below the title, there is a message in Korean: "ANACONDA. ▶ (1-1) http://www.anaconda.com/download/에 접속합니다 (Internet Explorer, Chrome 등 모든 브라우저 가능)". On the right side of the header, there are links for Documentation, Blog, Contact, What Is Anaconda?, Products, Support, Resources, About, and Downloads.

The main content area has a green background with a geometric pattern. The title "Download Anaconda Distribution" is centered at the top. Below it, it says "Version 5.2 | Release Date: May 30, 2018". There are download links for Windows, macOS, and Linux. The Windows link is highlighted with a red box.

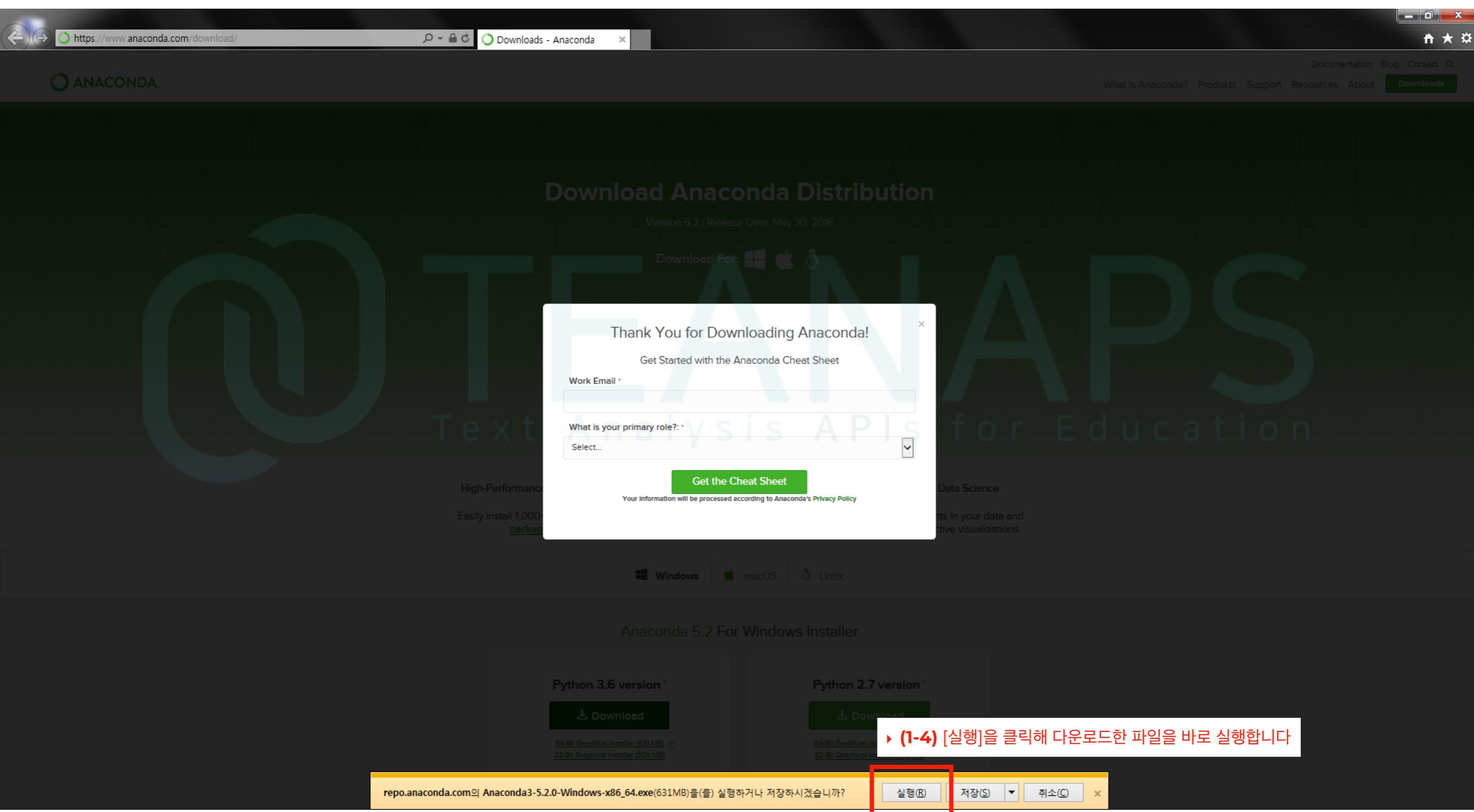
Below the download links, there are three sections: "High-Performance Distribution", "Package Management", and "Portal to Data Science". Each section has a brief description and a corresponding icon.

At the bottom, there are download buttons for "Anaconda 5.2 For Windows Installer" (Windows, macOS, Linux) and "Python 3.6 version" (Windows, macOS, Linux). The "Python 3.6 version" button is also highlighted with a red box. To the right of the Python 3.6 section, there is a note: "(예시 이미지는 Windows 운영체제, Internet Explorer)".

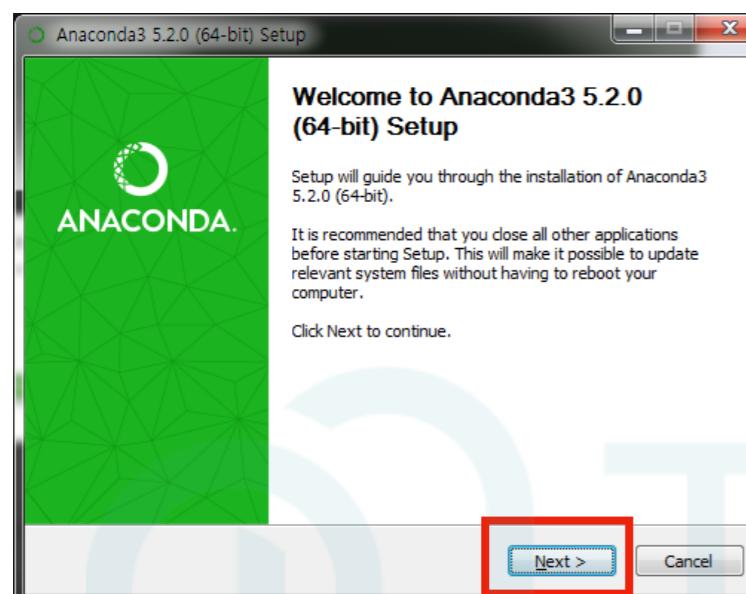
Red annotations in Korean are overlaid on the page:

- ▶ (1-2) 접속한 웹페이지에는 자신의 PC 운영체제에 맞는 다운로드 버튼이 생성됩니다  
(예시 이미지는 Windows 운영체제, Internet Explorer)
- ▶ (1-3) "Python 3.6 version" 아래 [Download] 버튼을 클릭합니다  
(자신의 운영체제에 맞는 설치파일이 자동으로 다운로드)

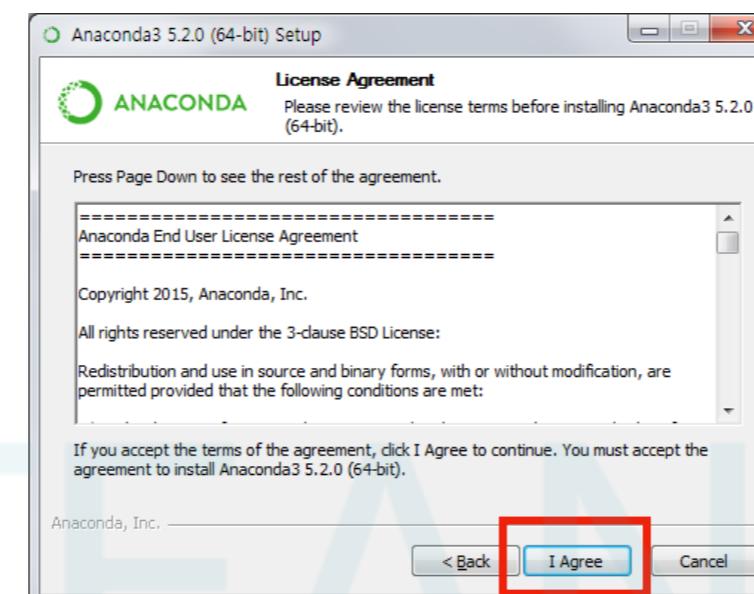
# Python 3.6 Anaconda



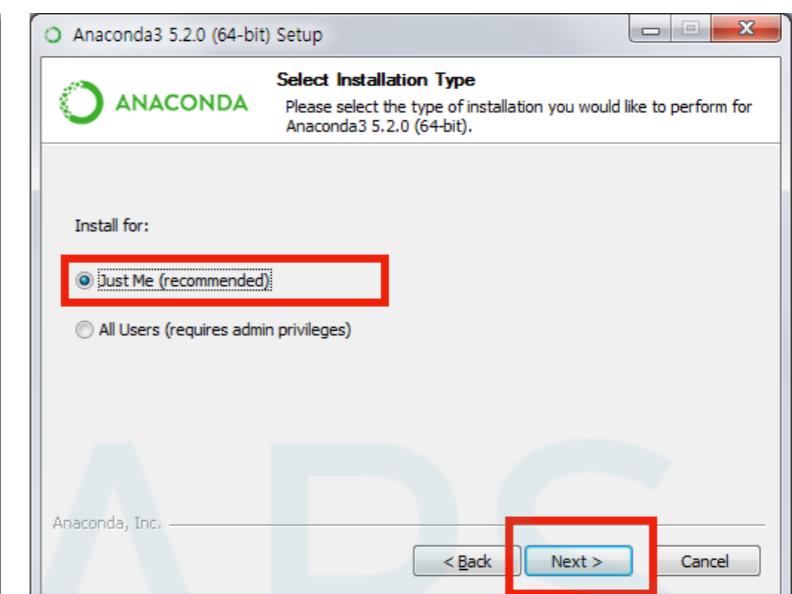
# Python 3.6 Anaconda



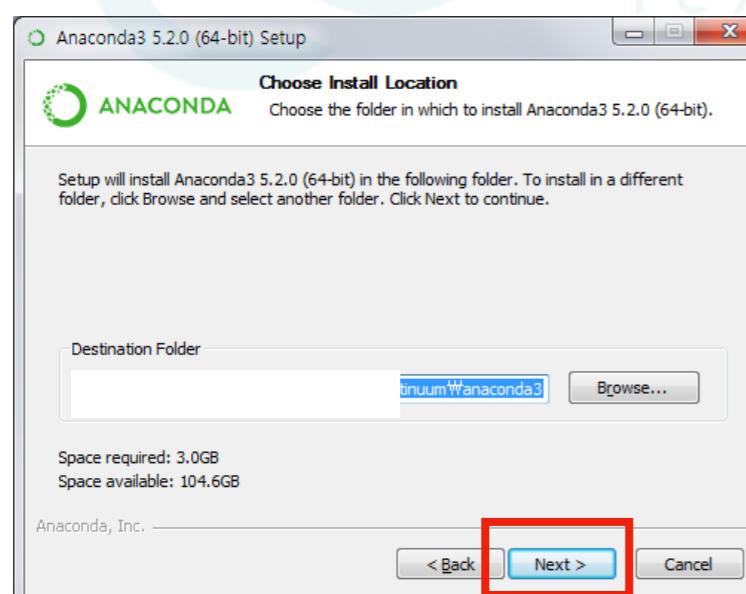
▶ (2-1) [Next >]를 클릭해 설치를 시작합니다



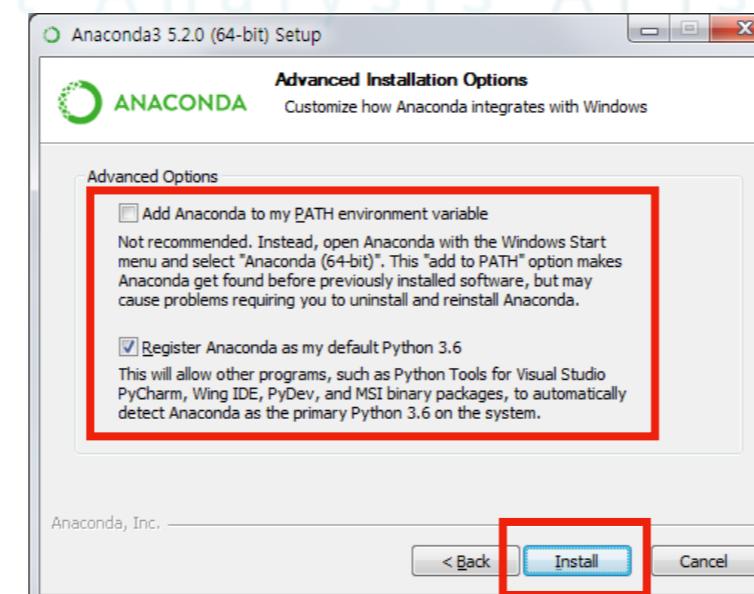
▶ (2-2) [I Agree]를 클릭해 약관에 동의합니다



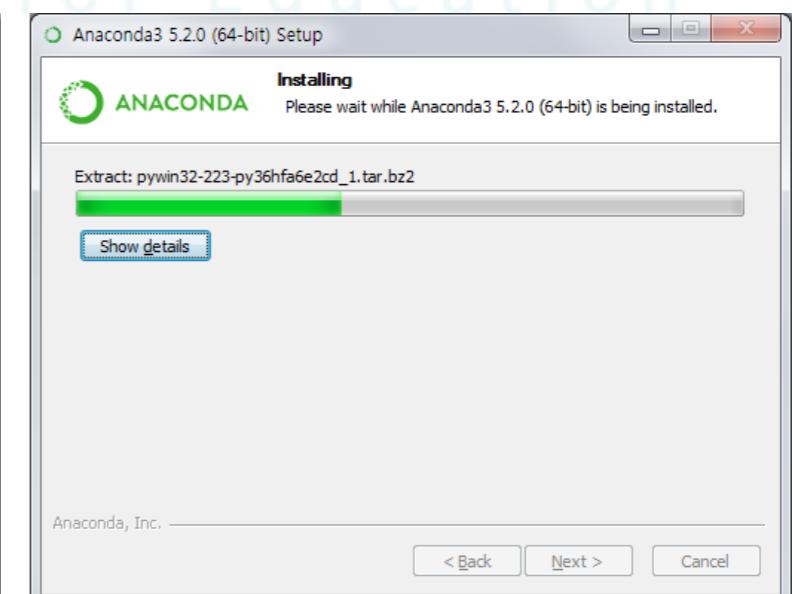
▶ (2-3) “Just Me (recommended)”를 선택하고  
[Next >]를 클릭합니다



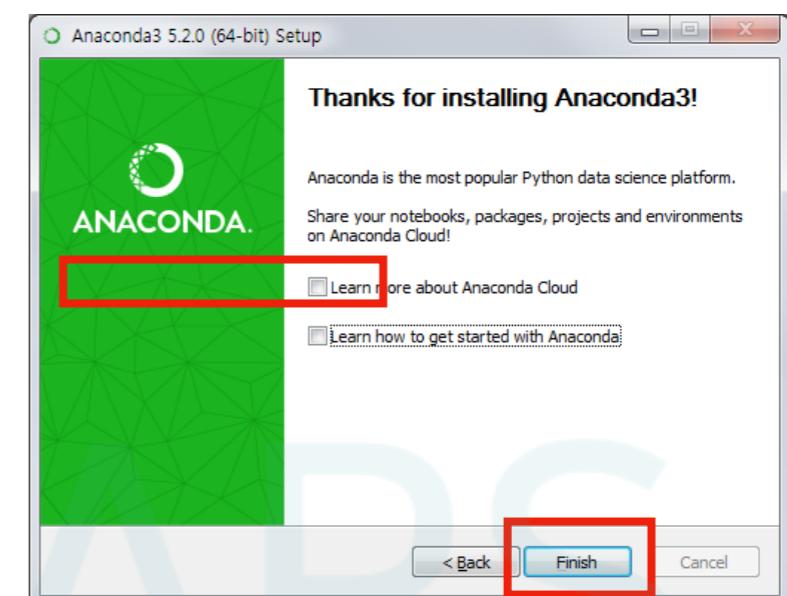
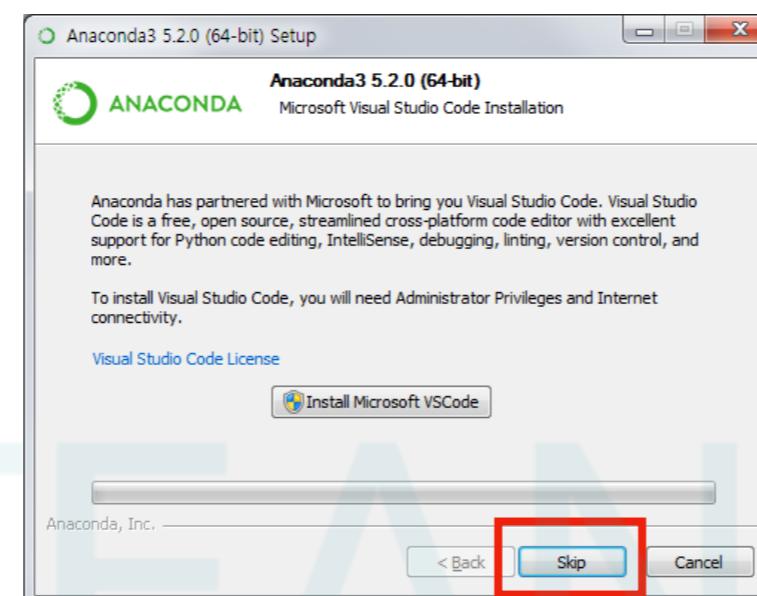
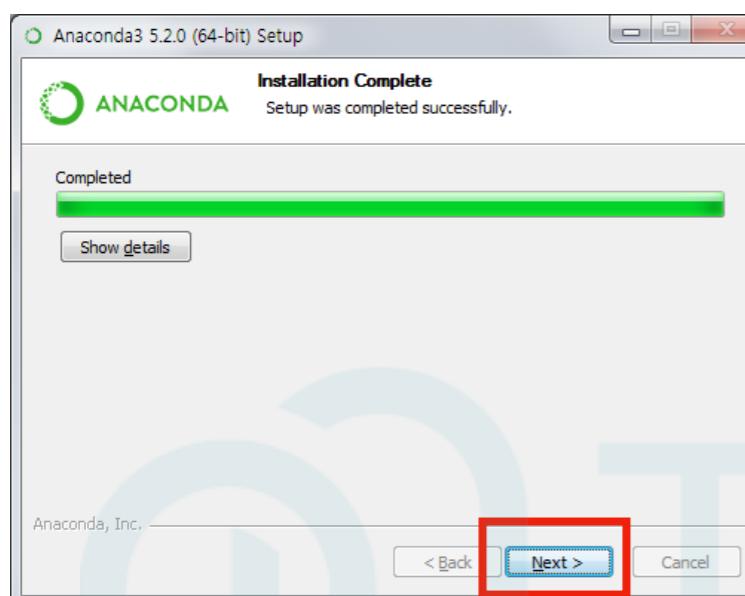
▶ (2-4) 설치경로를 그대로 두고 [Next >]를 클릭합니다



▶ (2-5) 옵션을 그대로 두고 [Install]을 클릭해 설치를 시작합니다 (설치에 약 5 ~ 15 분 소요)



# Python 3.6 Anaconda

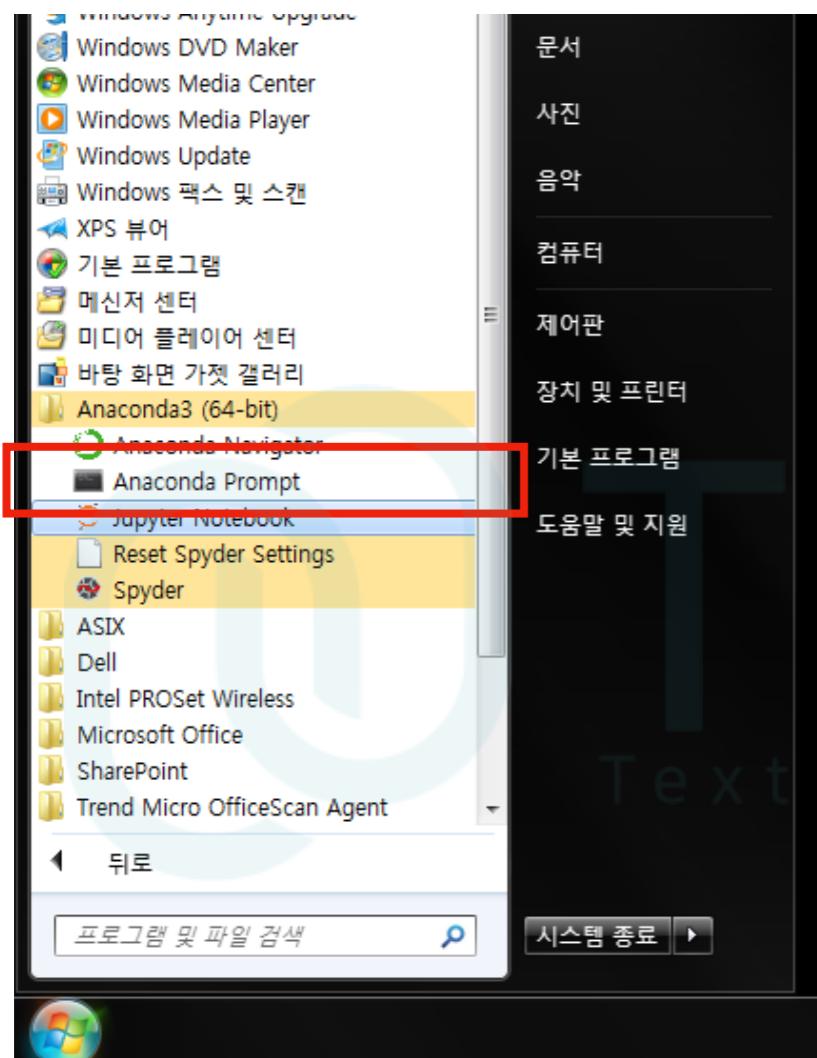


▶ (2-6) 설치가 완료되면 [Next >]를 클릭합니다

▶ (2-7) [Skip]을 클릭합니다

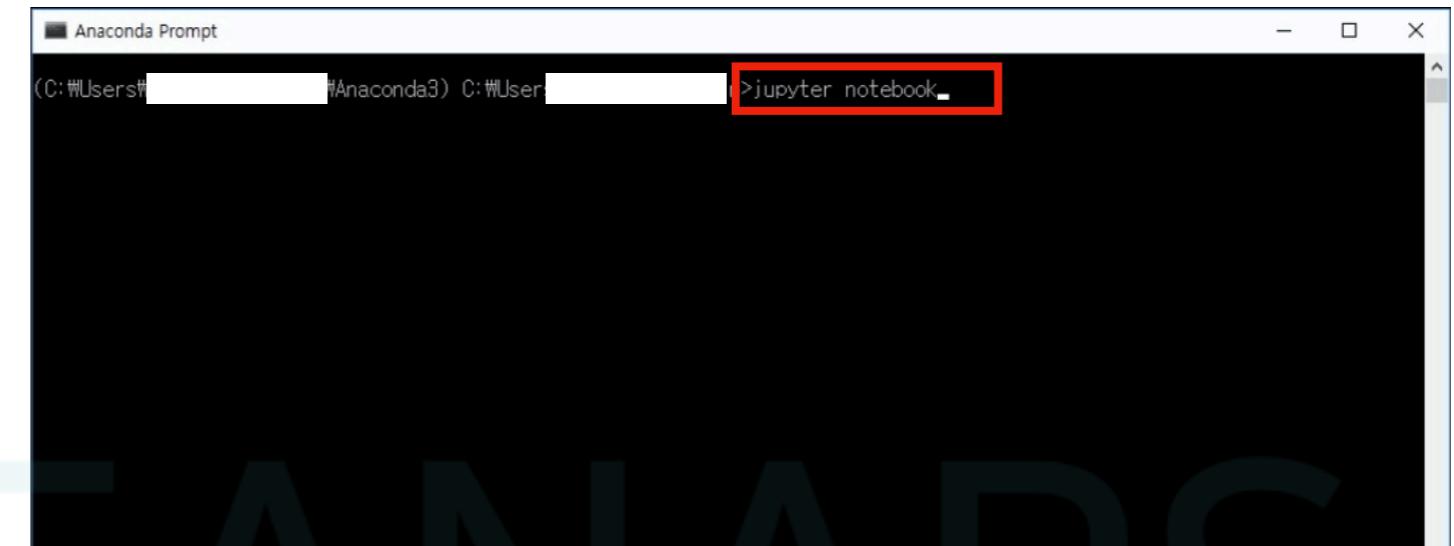
▶ (2-8) 체크박스를 모두 해제하고 [Finish]를 클릭해  
설치를 마칩니다

# Python 3.6 Anaconda

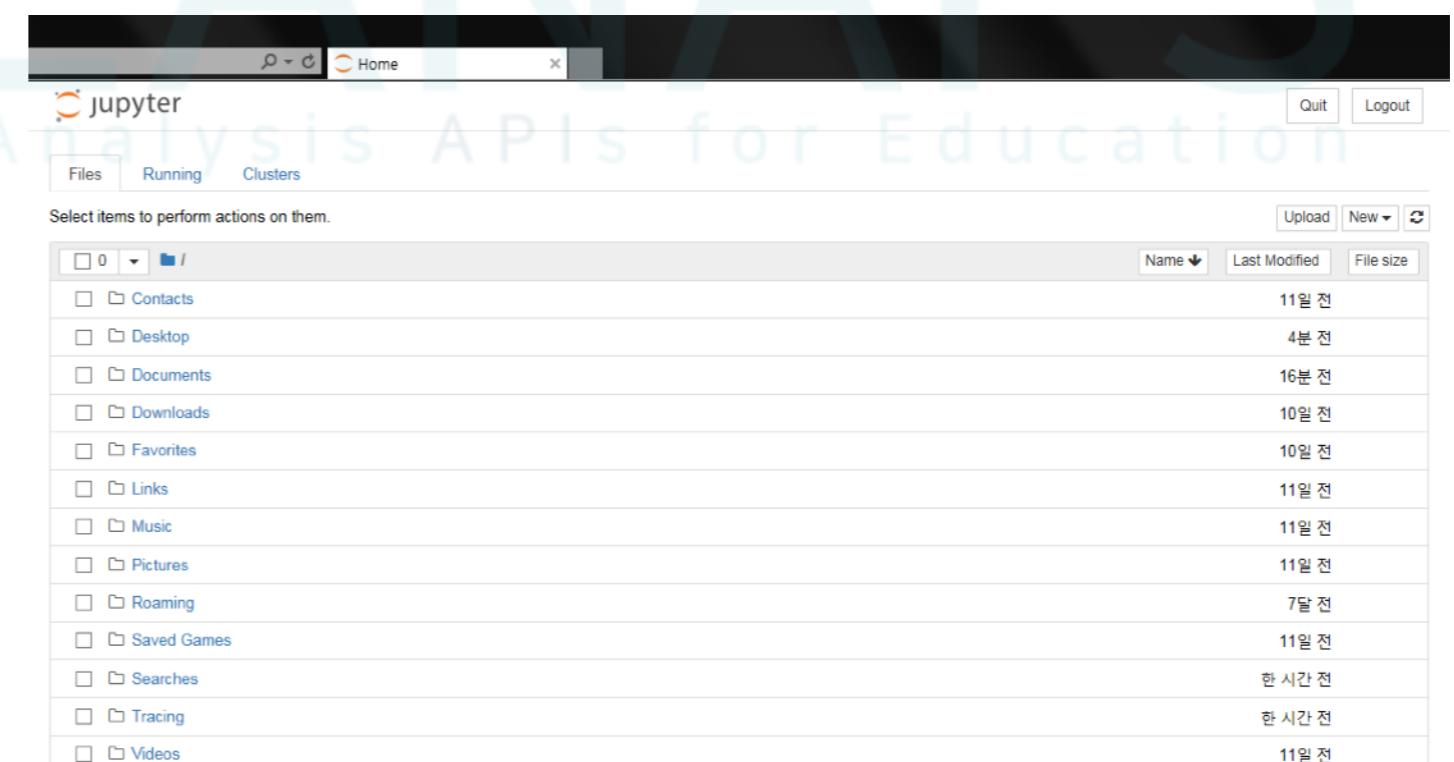


▶ (3-1) [시작] > [모든 프로그램] > [Anaconda3 (32/64-bit)]

> [Anaconda Prompt]를 클릭합니다



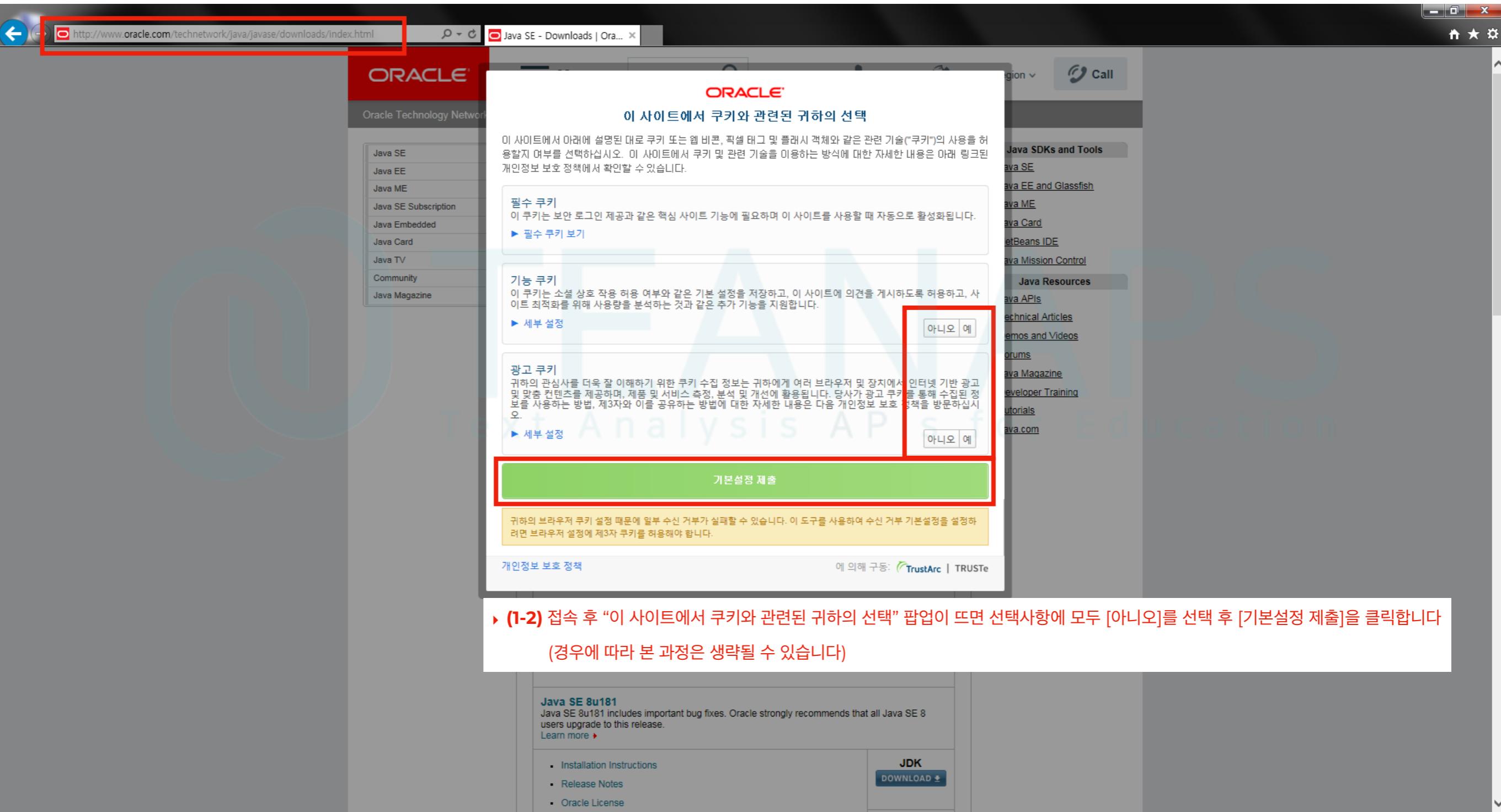
▶ (3-2) “jupyter notebook”을 입력하고 [엔터]를 입력합니다



▶ 그림과 같이 브라우저가 실행되면 실습을 위한 Python 3.6 Anaconda 환경구축이 완료됩니다

# JAVA jdk 설치

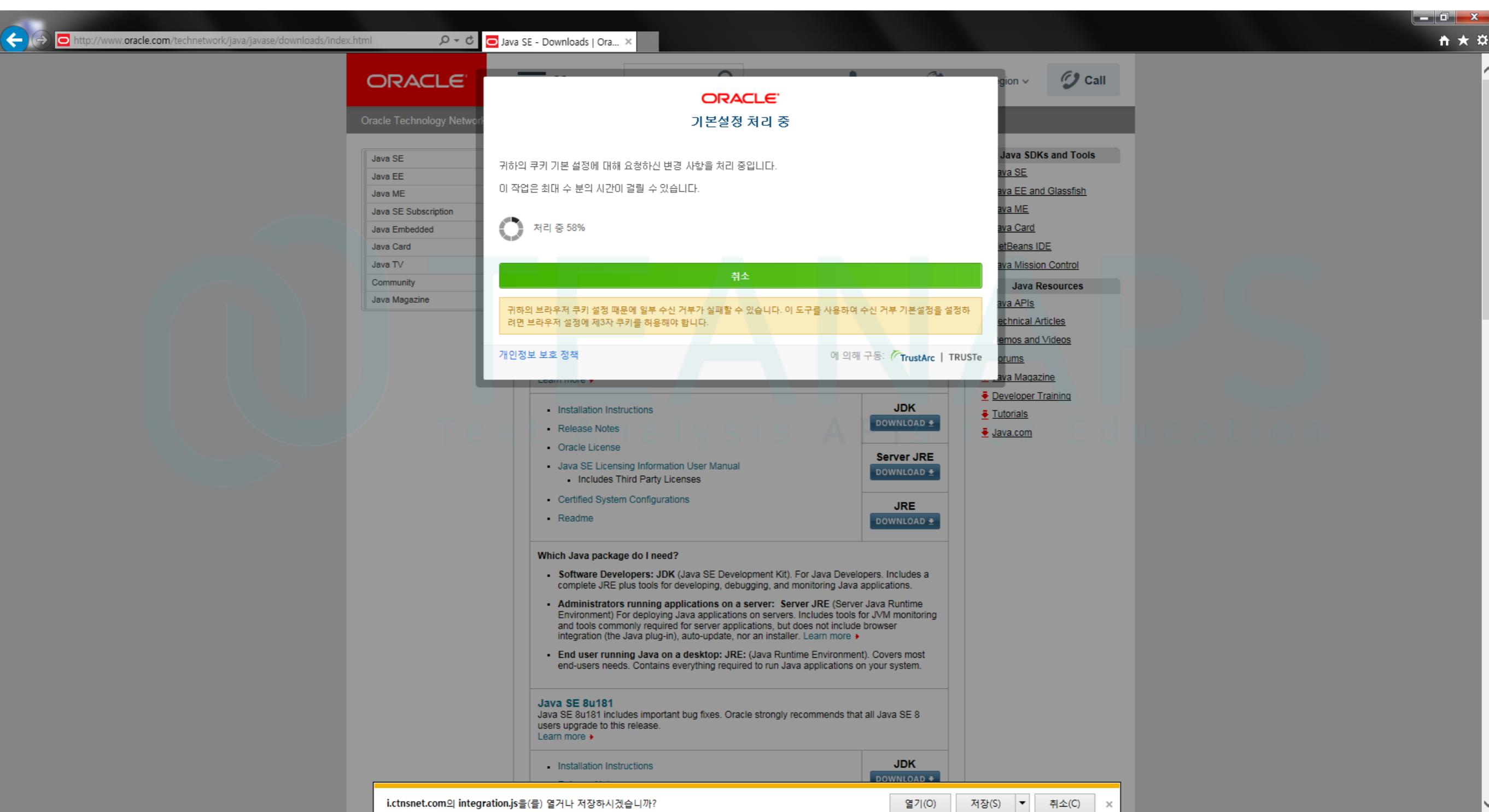
- ▶ (1-1) <https://www.oracle.com/technetwork/java/javase/downloads/index.html>에 접속합니다



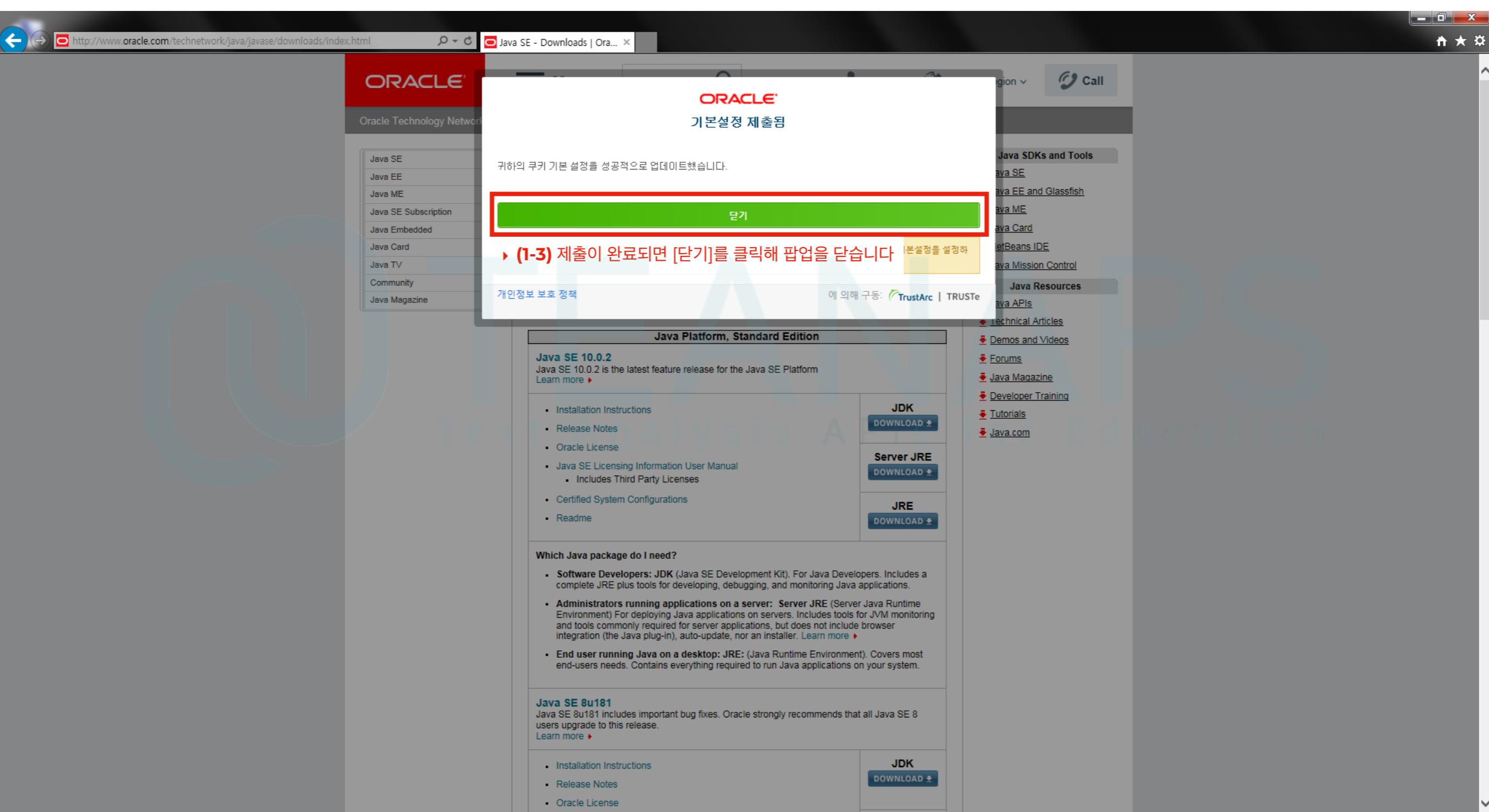
- ▶ (1-2) 접속 후 “이 사이트에서 쿠키와 관련된 귀하의 선택” 팝업이 뜨면 선택사항에 모두 [아니오]를 선택 후 [기본설정 제출]을 클릭합니다

(경우에 따라 본 과정은 생략될 수 있습니다)

# JAVA jdk 설치



# JAVA jdk 설치



# JAVA jdk 설치

The screenshot shows the Oracle Java SE Downloads page. At the top, there's a navigation bar with links for Overview, Downloads (which is selected), Documentation, Community, Technologies, and Training. On the left, a sidebar lists categories like Java SE, Java EE, Java ME, etc. The main content area is titled "Java SE Downloads" and features a large button with the Java logo and the text "Java Platform (JDK) 10". This button is highlighted with a red box. Below it, there's information about Java SE 10.0.2 and links for Installation Instructions, Release Notes, Oracle License, Java SE Licensing Information User Manual, Certified System Configurations, and Readme. To the right of these links are three download buttons: "JDK DOWNLOAD", "Server JRE DOWNLOAD", and "JRE DOWNLOAD". Further down, there's a section titled "Which Java package do I need?" with descriptions for Software Developers, Administrators running applications on a server, and End user running Java on a desktop. At the bottom, there's another section for Java SE 8u181 with similar download links.

# JAVA jdk 설치

The screenshot shows the Oracle Java SE Development Kit 10 Downloads page. At the top, there's a navigation bar with links for Overview, Downloads (which is selected), Documentation, Community, Technologies, and Training. On the left, there's a sidebar with links for Java SE, Java EE, Java ME, Java SE Subscription, Java Embedded, Java Card, Java TV, Community, and Java Magazine. The main content area is titled "Java SE Development Kit 10 Downloads" and contains text about the JDK being a development environment for building applications. It also mentions that the JDK includes tools for developing and testing programs written in Java. Below this, there's a section for "JDK 10.0.2 checksum". Further down, there's a table of download links for various operating systems. At the bottom of the table, there are two radio buttons: "Accept License Agreement" (unchecked) and "Decline License Agreement" (checked). A red box highlights the "Accept License Agreement" button. To the right of the table, there's a red arrow pointing right with the text "(1-5) 화면 아래 [Accept License Agreement]를 선택하여 라이선스 정책에 동의합니다" (Select the [Accept License Agreement] option below to agree to the license terms).

Java SE Development Kit 10 Downloads

You must accept the [Oracle Binary Code License Agreement](#) for Java SE to download this software.

Accept License Agreement  Decline License Agreement

Product / File Description	File Size	Download
Linux	306 MB	<a href="#">jdk-10.0.2_linux-x64_bin.rpm</a>
Linux	338.43 MB	<a href="#">jdk-10.0.2_linux-x64_bin.tar.gz</a>
macOS	395.46 MB	<a href="#">jdk-10.0.2_osx-x64_bin.dmg</a>
Solaris SPARC	207.07 MB	<a href="#">jdk-10.0.2_solaris-sparcv9_bin.tar.gz</a>
Windows	390.25 MB	<a href="#">jdk-10.0.2_windows-x64_bin.exe</a>

E-mail this page Printer View

Contact Us

US Sales: +1.800.633.0738  
Global Contacts  
Support Directory  
Subscribe to Emails

About Oracle

Careers  
Company Information  
Social Responsibility  
Communities

Downloads and Trials

Java Runtime Download  
Java for Developers  
Software Downloads  
Try Oracle Cloud Free

News and Events

Acquisitions  
Blogs  
Events  
Newsroom

ORACLE Integrated Cloud Applications & Platform Services

f t in G+ y r

# JAVA jdk 설치

The screenshot shows a web browser displaying the Oracle Java SE Development Kit 10 Downloads page. The URL in the address bar is <http://www.oracle.com/technetwork/java/javase/downloads/jdk10-downloads-441664.html>. The page has a red header with the Oracle logo. The main content area is titled "Java SE Development Kit 10 Downloads". It includes a sidebar with links to Java SE, Java EE, Java ME, Java Subscription, Java Embedded, Java Card, Java TV, Community, and Java Magazine. The "Downloads" tab is selected in the top navigation bar. A table lists download links for various operating systems: Linux (306 MB), Linux (338.43 MB), macOS (395.46 MB), Solaris SPARC (207.07 MB), and Windows (390.25 MB). The Windows row is highlighted with a red border. Below the table, there is a note about accepting the Oracle Binary Code License Agreement. At the bottom of the page, there are links for "Contact Us", "About Oracle", "Downloads and Trials", and "News and Events". Social media icons for Facebook, Twitter, LinkedIn, Google+, YouTube, and RSS are also present.

Java SE Development Kit 10 Downloads

You must accept the [Oracle Binary Code License Agreement](#) for Java SE to download this software.

Product / File Description      File Size      Download

Linux	306 MB	<a href="#">jdk-10.0.2_linux-x64_bin.rpm</a>
Linux	338.43 MB	<a href="#">jdk-10.0.2_linux-x64_bin.tar.gz</a>
macOS	395.46 MB	<a href="#">jdk-10.0.2_osx-x64_bin.dmg</a>
Solaris SPARC	207.07 MB	<a href="#">jdk-10.0.2_solaris-sparcv9_bin.tar.gz</a>
Windows	390.25 MB	<a href="#">jdk-10.0.2_windows-x64_bin.exe</a>

(1-6) 자신의 PC 운영체제에 해당하는 설치파일 링크를 클릭해 다운로드합니다  
(Windows PC는 [jdk-(version)\_windows-x64\_bin.exe], Mac OS는 [jdk-(version)-osx-x64\_bin.dmg]를 클릭)

# JAVA jdk 설치

The screenshot shows the Oracle Java SE Development Kit 10 Downloads page. The main content area displays the following information:

**Java SE Development Kit 10 Downloads**  
Thank you for downloading this release of the Java™ Platform, Standard Edition Development Kit (JDK™). The JDK is a development environment for building applications, and components using the Java programming language.

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

See also:

- Java Developer Newsletter: From your Oracle account, select Subscriptions, expand Technology, and subscribe to Java.
- Java Developer Day hands-on workshops (free) and other events
- Java Magazine

JDK 10.0.2 checksum

**Java SE Development Kit 10.0.2**  
You must accept the Oracle Binary Code License Agreement for Java SE to download this software.  
Thank you for accepting the Oracle Binary Code License Agreement for Java SE; you may now download this software.

Product / File Description	File Size	Download
Linux	306 MB	<a href="#">jdk-10.0.2_linux-x64_bin.rpm</a>
Linux	338.43 MB	<a href="#">jdk-10.0.2_linux-x64_bin.tar.gz</a>
macOS	395.46 MB	<a href="#">jdk-10.0.2_osx-x64_bin.dmg</a>
Solaris SPARC	207.07 MB	<a href="#">jdk-10.0.2_solaris-sparcv9_bin.tar.gz</a>
Windows	390.25 MB	<a href="#">jdk-10.0.2_windows-x64_bin.exe</a>

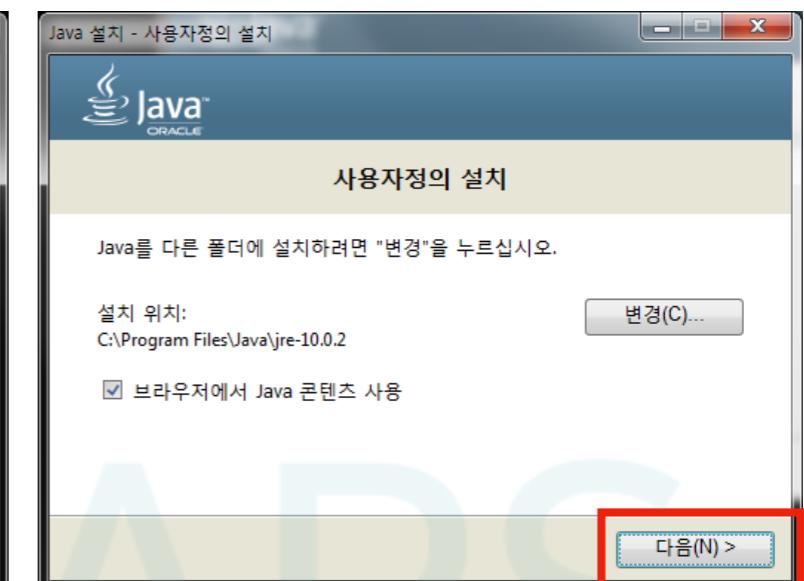
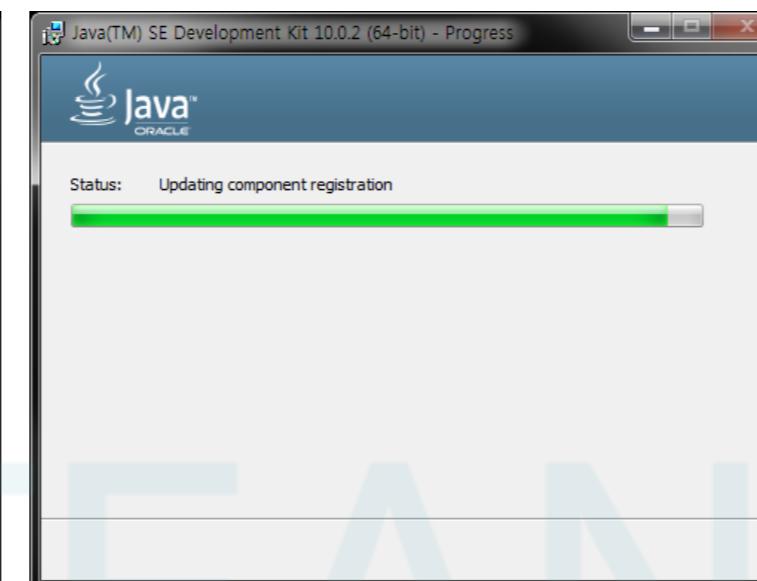
At the bottom of the page, there is a note in Korean: ▶ (1-7) [실행]을 클릭해 다운로드한 파일을 바로 실행합니다 (▶ (1-7) Click [Run] to execute the downloaded file immediately).

The browser address bar shows the URL: http://www.oracle.com/technetwork/java/javase/downloads/jdk10-downloads-441664.

# JAVA jdk 설치



▶ (2-1) [Next >]를 클릭해 설치를 시작합니다



▶ (2-2) 설치경로를 그대로 두고 [Next >]를 클릭합니다

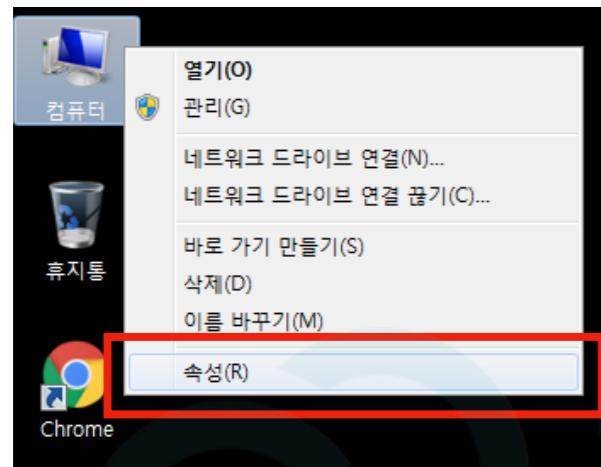


▶ (2-3) “사용자정의 설치” 창이 뜨면 경로를 그대로 두고  
[다음(N) >]을 클릭합니다



▶ (2-4) 설치가 완료되면 [close]를 클릭해 설치를 마칩니다

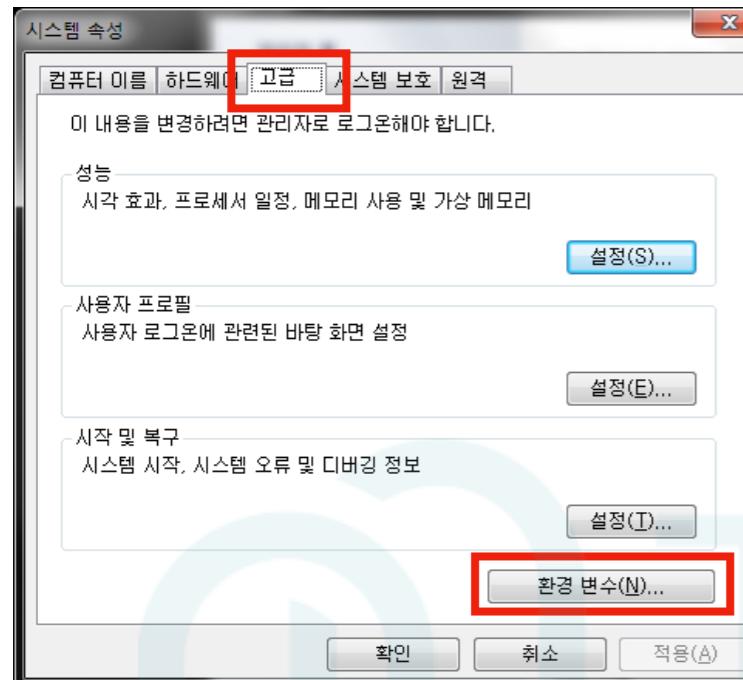
# 환경변수 설정



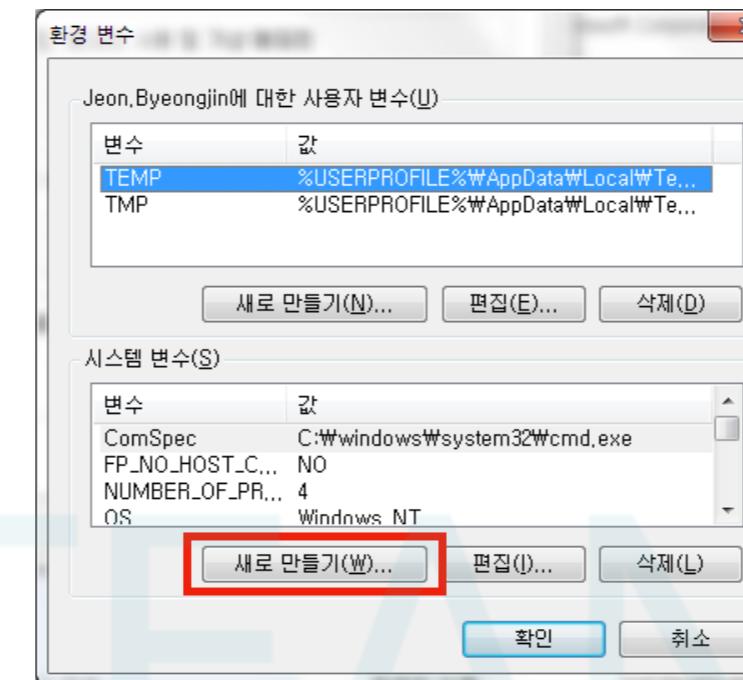
▶ (1-1) 바탕화면의 “내컴퓨터” 아이콘에서 마우스 오른쪽을 클릭한 후 [속성]을 클릭합니다



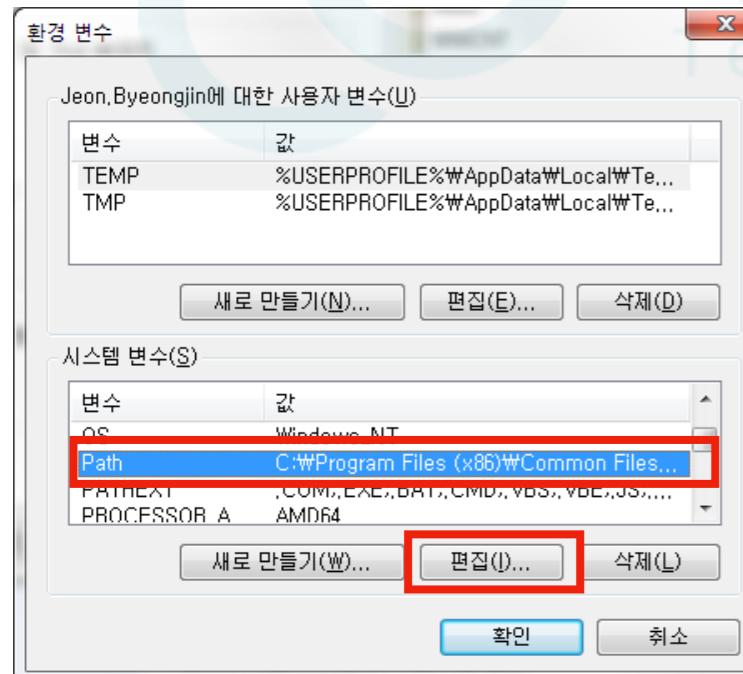
# 환경변수 설정



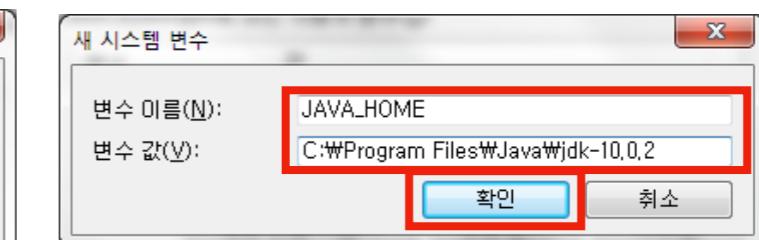
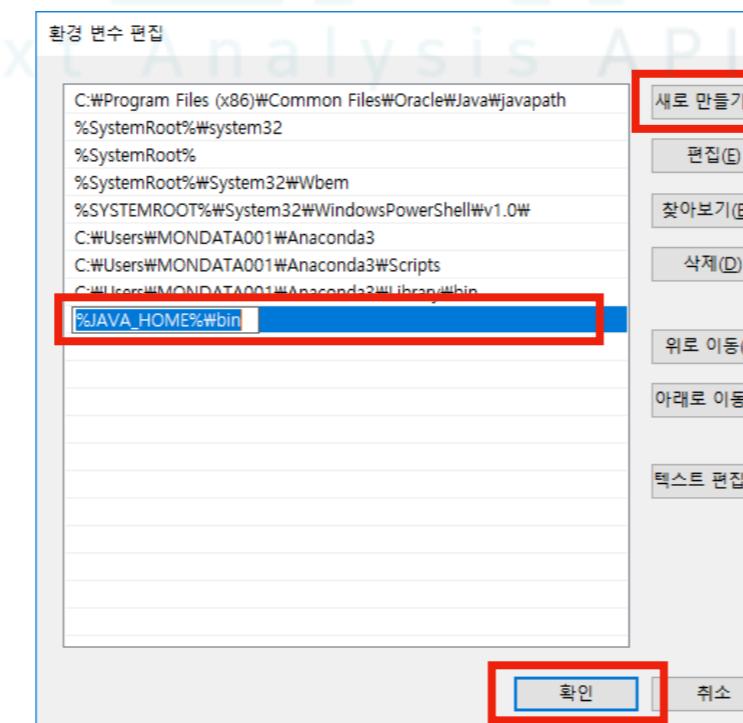
▶ (1-3) [고급] 탭을 선택하고 아래 [환경변수]를 클릭합니다



▶ (1-4) 아래 두번째 [새로 만들기(W)...]를 클릭합니다



▶ (1-6) 아래 “시스템 변수(S)” 목록 중 변수 이름이 “Path”인 항목을 선택하고 [편집(I)...]을 클릭합니다



▶ (1-5) “변수 이름(N)”에 “JAVA\_HOME”, “변수 값(V)”에 “C:\Program Files\Java\jdk-(version)”를 입력한 후 [확인]을 클릭합니다  
(version은 JAVA jdk 경로 C:\Program Files\Java\... 를 찾아가 확인, Window 10 사용자는 오른쪽 [새로 만들기(N)]를 클릭해 추가)

▶ (1-7) [새로 만들기(N)]을 클릭하고 “%JAVA\_HOME%\bin”을 입력 후 [확인]을 클릭합니다

▶ (1-8) 계속 [확인]을 클릭해 모든 열었던 창을 모두 닫습니다

# 환경변수 설정



▶ (2-1) 바탕화면에서 [시작]을 클릭하고 “cmd”를  
입력해 “cmd” 명령창을 실행합니다

```
C:\> C:\Windows\system32\cmd.exe
Microsoft Windows [Version 6.1.7601]
Copyright <c> 2009 Microsoft Corporation. All rights reserved.

C:\> [REDACTED]>java --version
```

▶ (2-2) “java -version”을 입력하고 [엔터]를 입력합니다

```
C:\> C:\Windows\system32\cmd.exe
Microsoft Windows [Version 6.1.7601]
Copyright <c> 2009 Microsoft Corporation. All rights reserved.

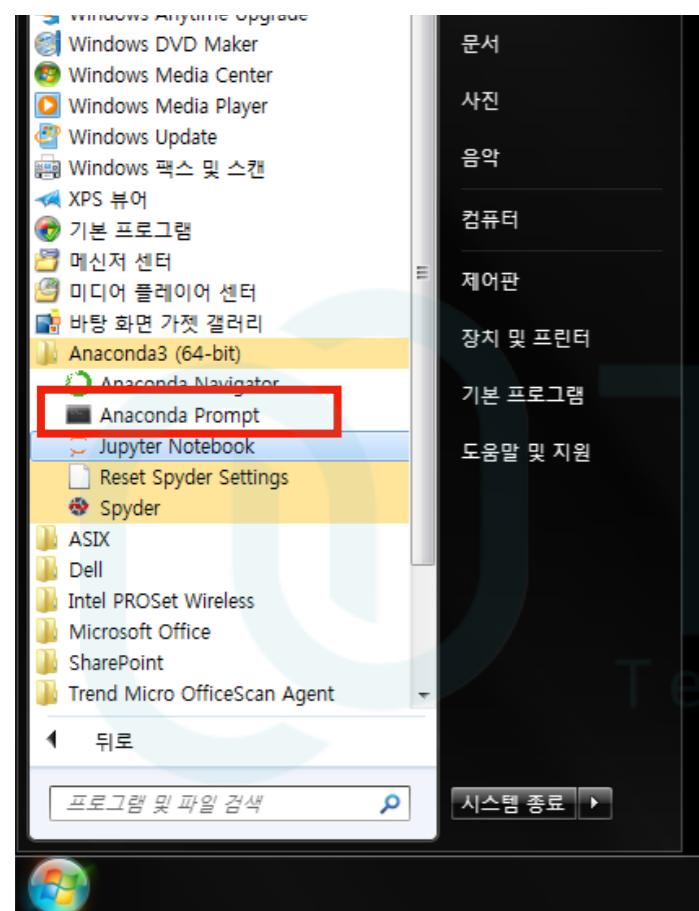
C:\> [REDACTED]>java --version
java 10.0.2 2018-07-17
Java(TM) SE Runtime Environment 18.3 (build 10.0.2+13)
Java HotSpot(TM) 64-Bit Server VM 18.3 (build 10.0.2+13, mixed mode)

C:\> [REDACTED]>
```

▶ (2-3) “java (version) (날짜) ...”가 출력되면 환경변수 설정이 완료됩니다

(설치된 버전은 유저마다 다를 수 있습니다)

# Microsoft Visual C++ Build Tools



▶ (1-1) [시작] > [모든 프로그램]

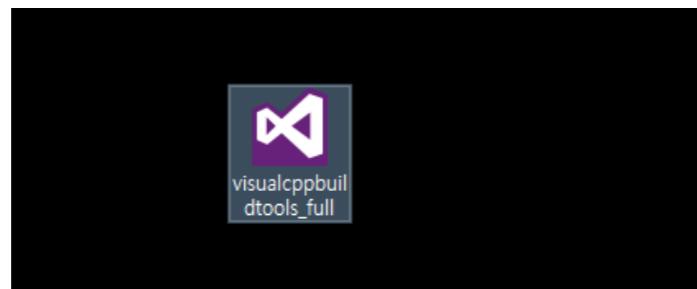
> [Anaconda3 (32/64-bit)]

> [Anaconda Prompt]를 클릭합니다

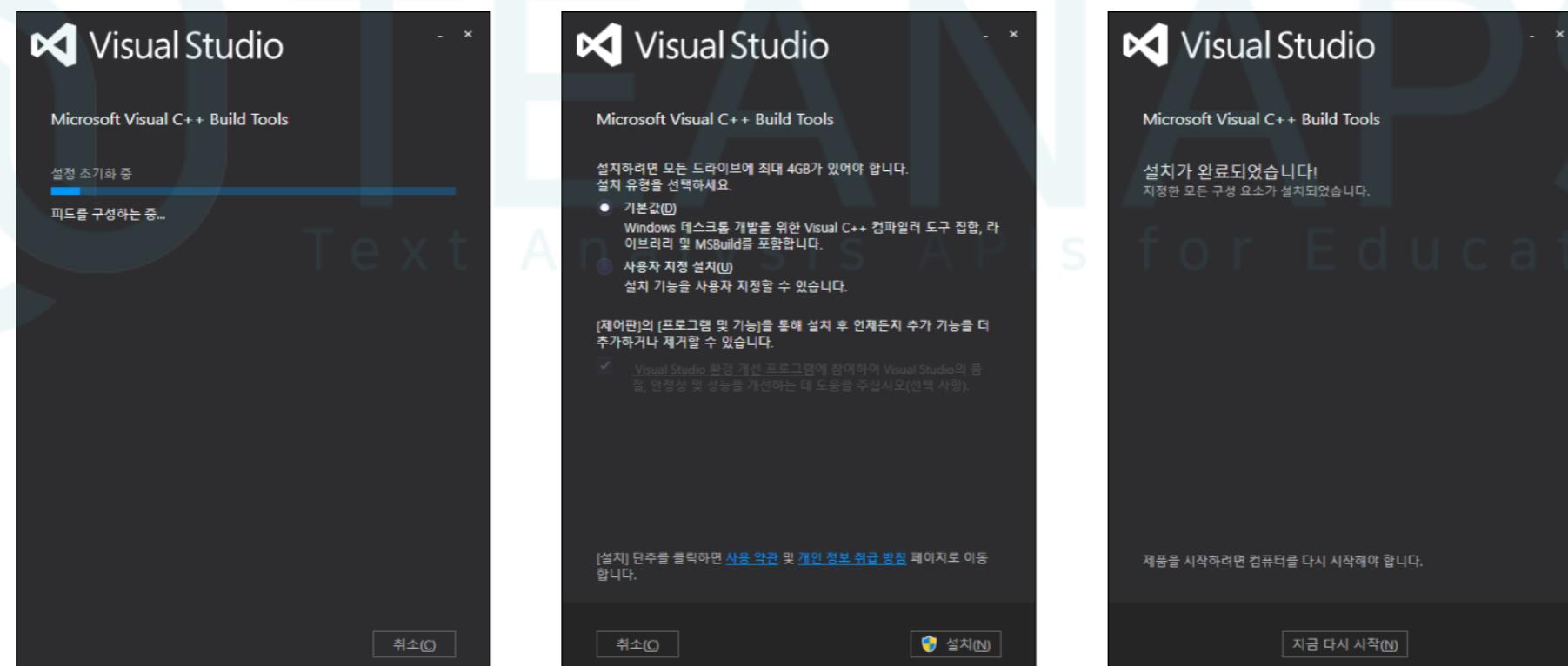
The image shows an Anaconda Prompt window titled 'Anaconda Prompt'. The command line shows '(base) C:\Users\...>python -m pip install --upgrade pip', with the entire command highlighted by a red box. The background of the slide features a large, semi-transparent watermark reading 'TEANAPS Text Analysis APIs for Education'.

▶ (1-2) Jupyter Notebook에 “python -m pip install --upgrade pip”를 입력하고 [엔터]를 입력합니다

# Microsoft Visual C++ Build Tools



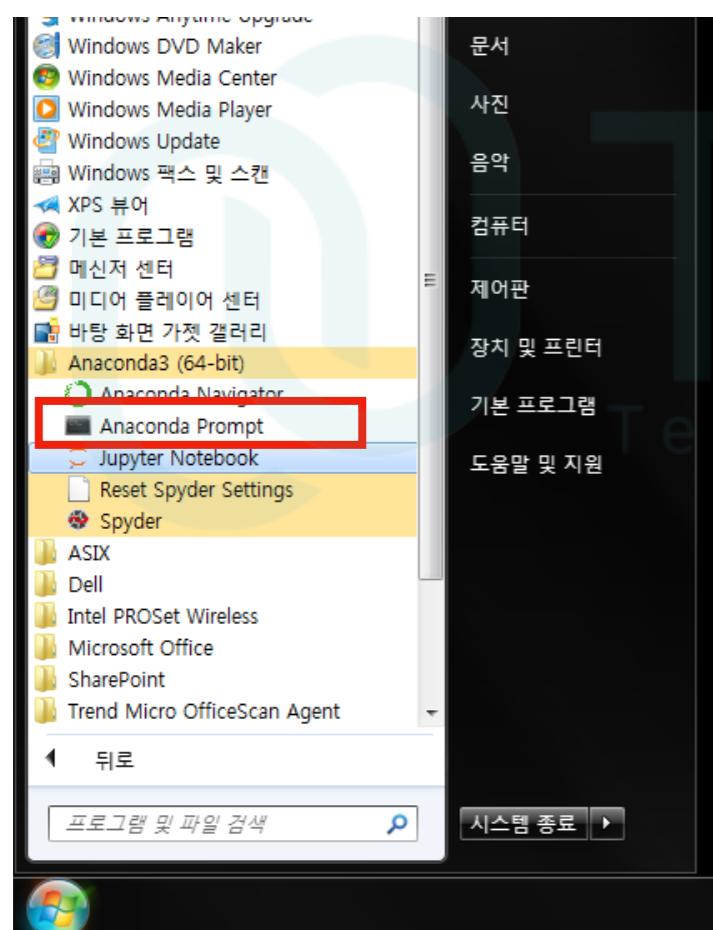
- ▶ (1-3) 링크에서 파일을 다운로드합니다 (<http://go.microsoft.com/fwlink/?LinkId=691126&fixForIE=.exe>.)
- ▶ (2-1) 다운로드한 설치파일을 실행합니다 (visualcppbuildtools\_full.exe)



- ▶ (2-2) [설치(N)]를 클릭하고 완료를 기다립니다 (약 5 ~ 6 분 소요)
- ▶ (2-3) 설치준비가 완료되면 [지금 다시 시작(N)]을 클릭하고 재부팅합니다

# Python Package 설치

- ◎ “jpype1” 패키지 설치 중 오류가 발생하는 경우 **JAVA 환경변수 설정 및 Microsoft Visual C++ Build Tools 설치여부를 확인해주세요**
- ◎ 또는 **Microsoft Visual C++ Build Tools**가 버전별로 중복설치되어있지는 않은지 확인해주세요



- ▶ (1-1) [시작] > [모든 프로그램] > [Anaconda3 (32/64-bit)]  
> [Anaconda Prompt]를 클릭합니다

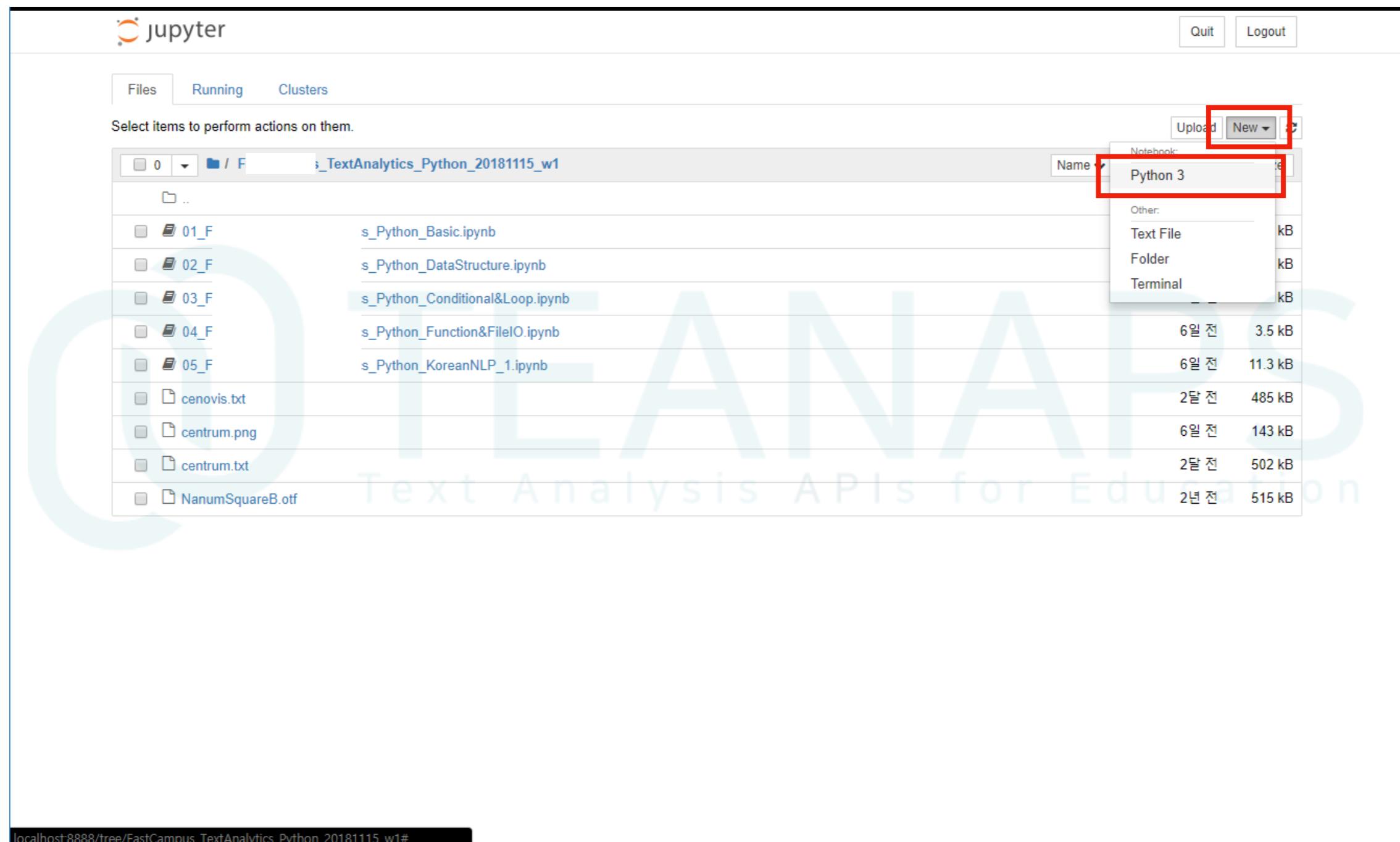
A screenshot of the Anaconda Prompt terminal window. The terminal output shows three commands being run:  
(base) C:\Users\...>pip install jpype1  
Collecting jpype1  
Installing collected packages: jpype1  
Successfully installed jpype1-0.6.3  
  
(base) C:\Users\...>pip install konlpy  
Collecting konlpy  
Using cached https://files.pythonhosted.org/packages/e5/3d/4e983cd98d87b50b2ab0387d73fa946f745aa8164e8888a714d5129f9765/konlpy-0.5.1-py2.py3-none-any.whl  
Requirement already satisfied: JPype1>=0.5.7 in c:\Users\mondata001\anaconda3\lib\site-packages (from konlpy) (0.6.3)  
Installing collected packages: konlpy  
Successfully installed konlpy-0.5.1  
  
(base) C:\Users\...>pip install wordcloud  
Collecting wordcloud  
Using cached https://files.pythonhosted.org/packages/dd/57/8fd38e31be7951f9164f4cb3a769159a0c2ef4af0f597ddf09a7e736593e/wordcloud-1.5.0-cp36-cp36m-win\_amd64.whl  
Requirement already satisfied: pillow in c:\Users\mondata001\anaconda3\lib\site-packages (from wordcloud) (5.1.0)  
Requirement already satisfied: numpy>=1.6.1 in c:\Users\mondata001\anaconda3\lib\site-packages (from wordcloud) (1.14.3)  
Installing collected packages: wordcloud  
Successfully installed wordcloud-1.5.0  
  
(base) C:\Users\...>

- ▶ (2-1) “pip install jpype1”을 입력하고 [엔터]를 입력합니다
- ▶ (2-2) “pip install konlpy”를 입력하고 [엔터]를 입력합니다
- ▶ (2-3) “pip install wordcloud”를 입력하고 [엔터]를 입력합니다

A screenshot of the Anaconda Prompt terminal window. The terminal output shows the successful installation of the wordcloud package:  
successfully installed wordcloud-1.5.0  
(base) C:\Users\...>jupyter notebook

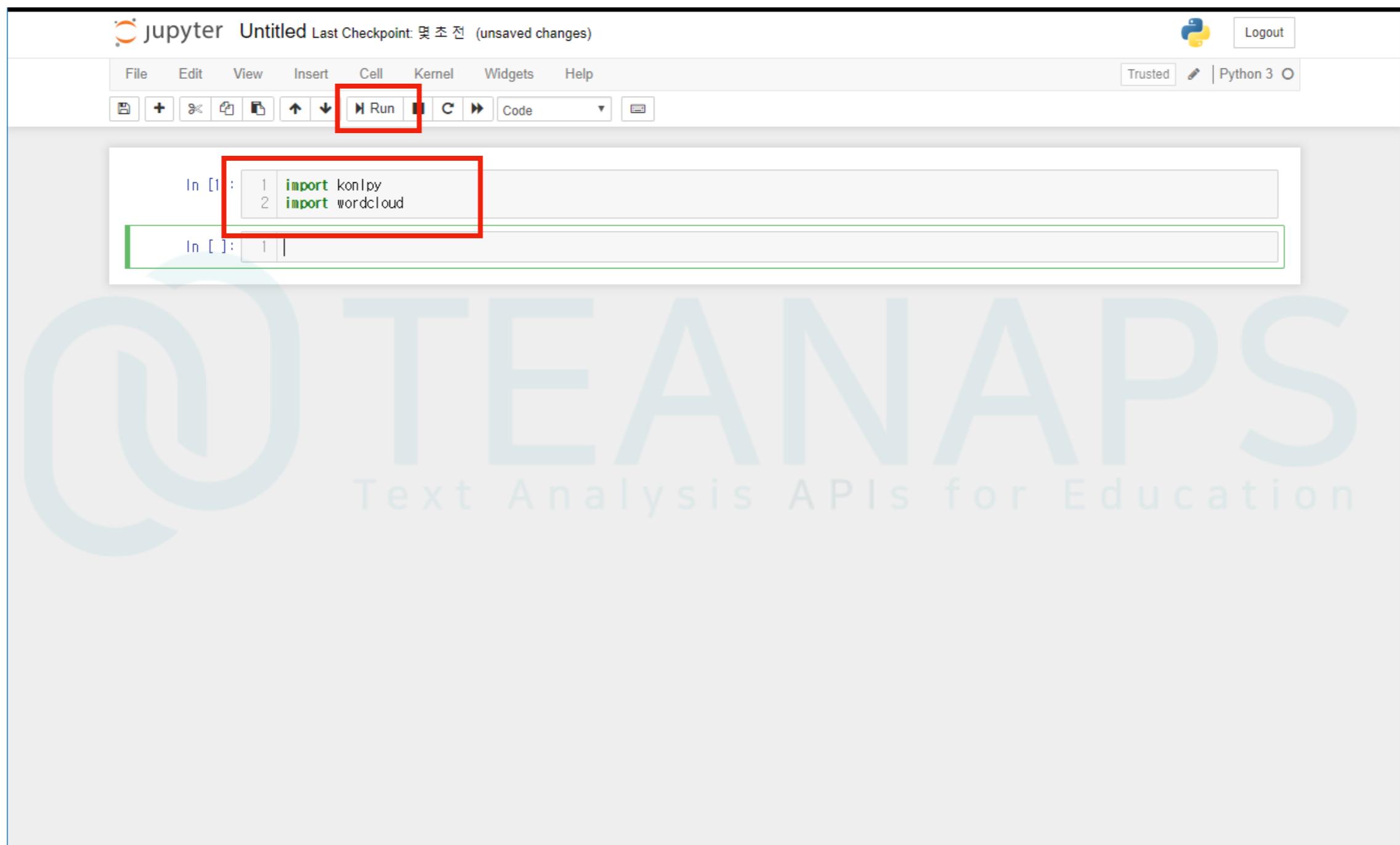
- ▶ (3-1) “jupyter notebook”을 입력하고 [엔터]를 입력합니다

# Python Package 설치



- ▶ (3-2) “jupyter notebook” 화면 우측 상단에 [New▼]를 클릭하고 [Python 3]를 선택합니다

# Python Package 설치



- ▶ (3-3) 화면에 “import konlpy”, “import wordcloud”를 입력하고 [▶Run] 버튼을 클릭합니다
- ▶ (3-4) 에러가 발생하지 않고 정상 실행되면 실습환경이 정상 설치되었습니다

# E.O.D

## Contact

-  <http://www.teanaps.com>
-  [fingeredman@gmail.com](mailto:fingeredman@gmail.com)