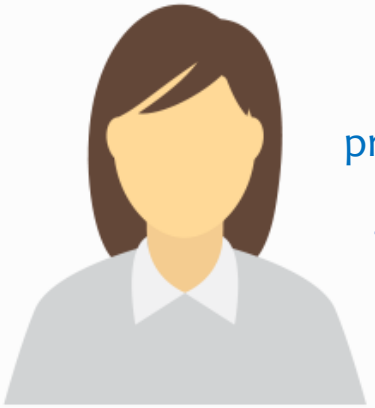
Financial Bigdata and Python

1. Introduction

Programming

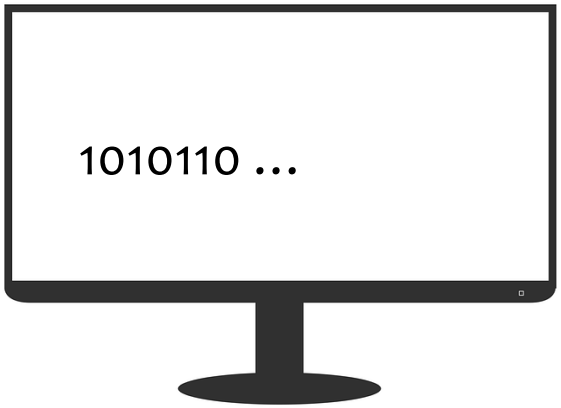
Act of writing a command algorithm in a language that allows it to communicate with a computer



programming language

compiler

machine language



storage

memory

CPU

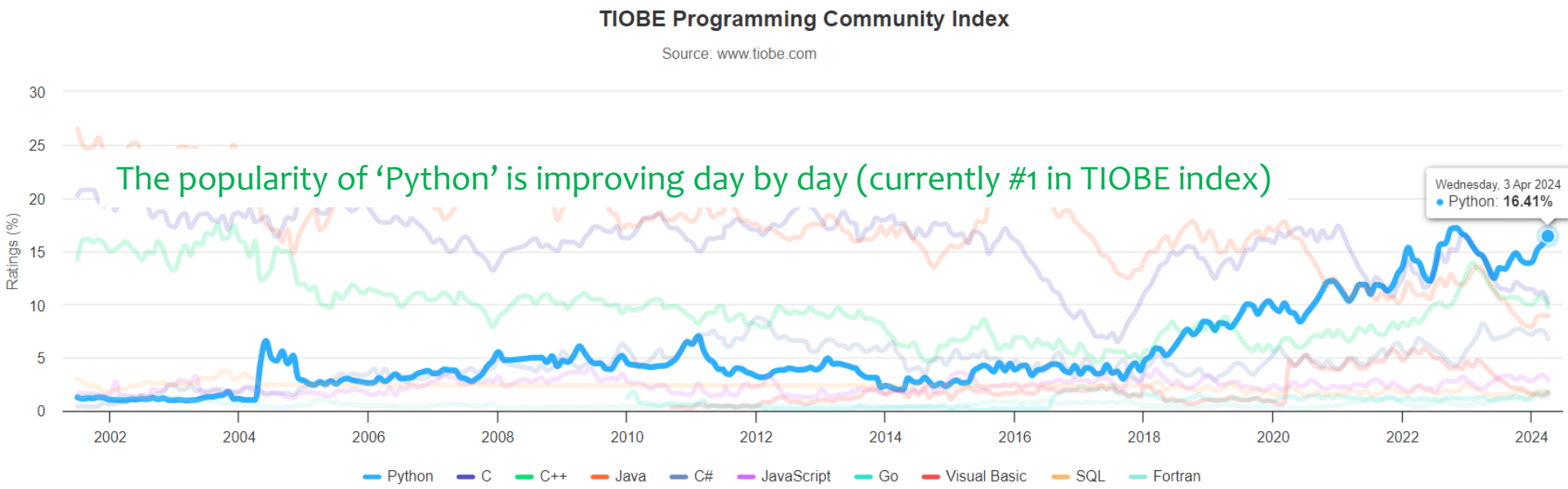
stores code and data

Data and code are allocated

processes instructions

Language	Difficulty	Performance	Versatility	Fitness for finance	Development year	Note
C	B	A	B	C	1972	simple but high performance
C++	C	A	A	A	1983	object-oriented language (main language for finance)
JAVA	B	B	A	C	1991	garbage collector (developers' main language)
Python	A	C	A	A	1989	high-productivity and general-purpose language
R	A	C	C	B	1990	specialization in statistics
MATLAB	A	C	C	B	1984	specialization in math/engineering

Recommended programming language for finance : Python, C++



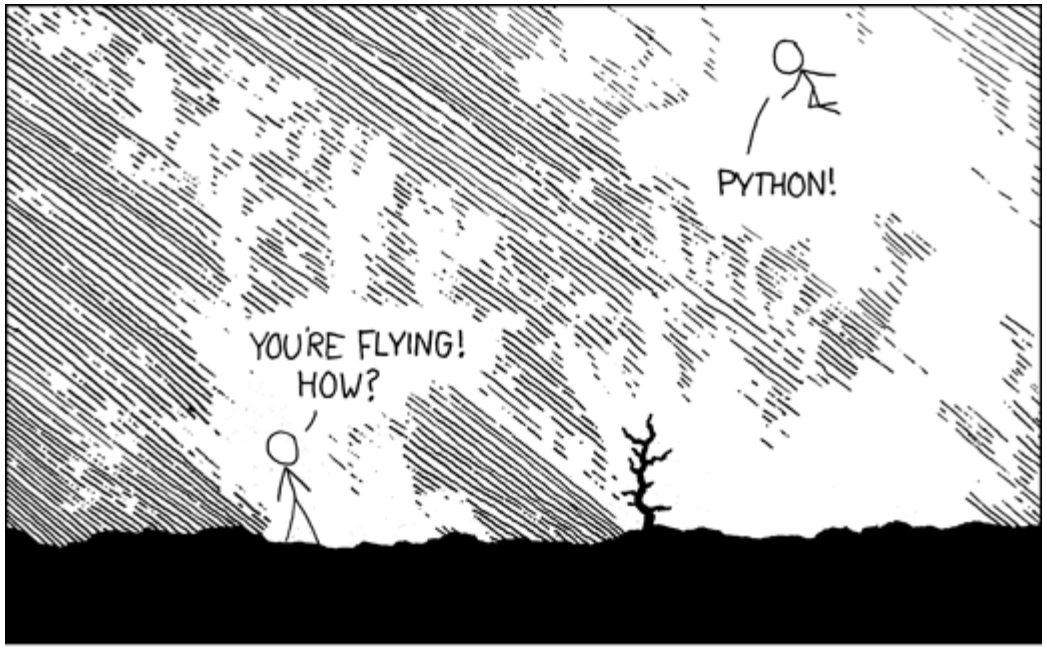
Rank	Language	Ratings
1	Python	16.41%
2	C	10.21%
3	C++	9.76%
16	MATLAB	1.11%

- #1 by PYPL (#2: Java)
- #2 by RedMonk (#1: JavaScript)

『Apr 2024』

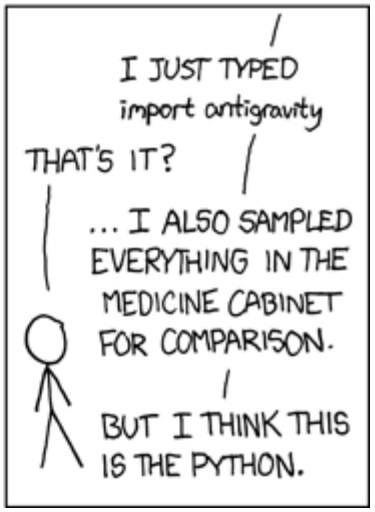
Article Title	Press	Date
Python craze in recruitment and vocational training	Apple Economy	2020.2.25
How long will **Python** be popular?	Byline Network	2020.3.17
Why **Python** is attracting attention as a next-generation language	Coding World	2019.7.26
Countdown to the end of life of **Python** 2 that dominates the AI ecosystem	GDNet Korea	2020.1.20
Julia vs **Python**... What is the best data language?	CIO Korea	2019.7.12
Python is the most preferred programming language for engineers.	GDNet Korea	2019.9.9
Unbelievable growth, about the innovation that **Python** will lead in the future	CIO Korea	2019.5.18
Math and **Python** data analysis you must know	IT news	2019.12.27
Python's popularity is at an all-time high... Overtake Java and C in 3 or 4 years	CIO Korea	2019.6.10
Anyone can analyze data with **Python**	UPI news	2019.6.14

The three most popular deep learning frameworks are all based on **Python.**

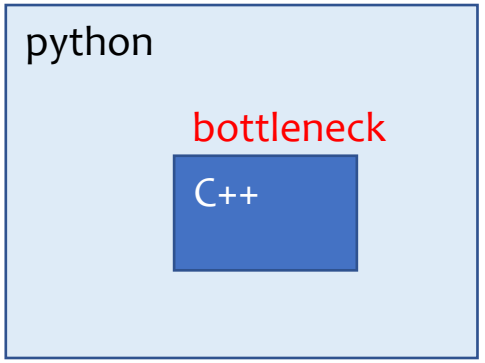


Life is short, You need Python.

No other language is as easy to learn,
yet as versatile and productive as Python.



Program development trend



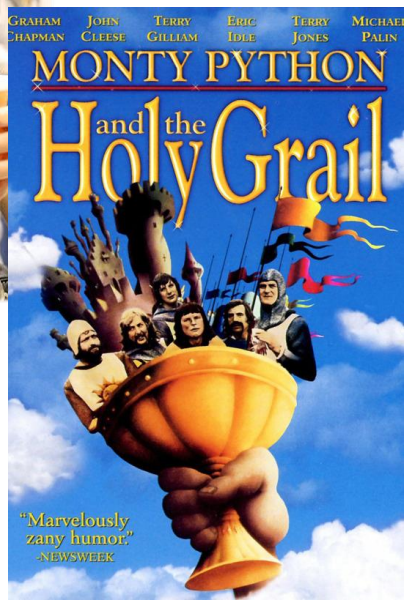
Guido van Rossum

- **How was 'Python' created?**

It was made by Guido van Rossum (1956-)
because he was just bored (!) on Christmas in 1989.

- Where did the name Python come from?

If you look up an English dictionary, it says it's a kind of snake.
But, according to the creator, regardless of this,
it was just named after his favorite British comedian group (Monty Python).
Nevertheless, the mark of Python comes from the shape of a snake.



Monty Python

Python (a kind of snake)

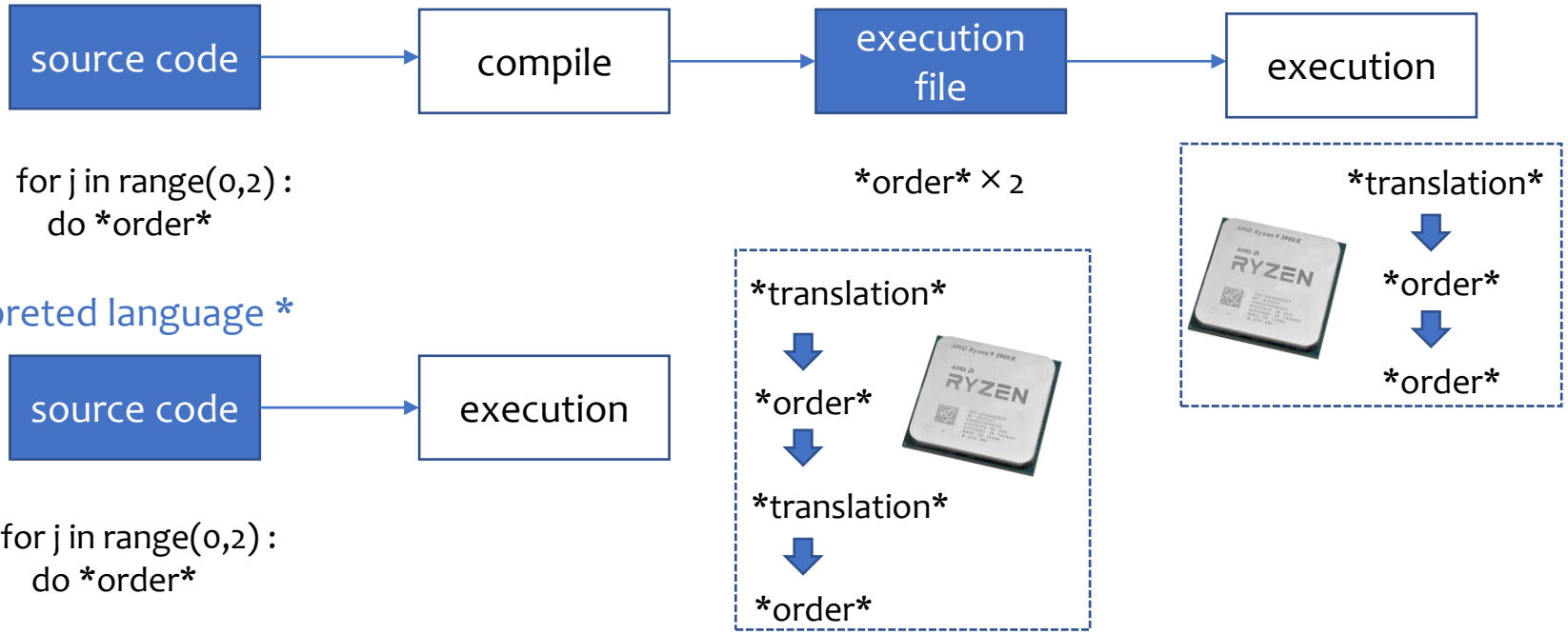
Python logo

Features of Python

1) Interpreter Language

One can run your code immediately without compiling it.
On the other hand, languages such as C and C++ require a compilation process.
Because it does not require compilation, it is highly flexible, such as trying various tests easily.
The disadvantage is that the execution speed is relatively slow compared to other languages.

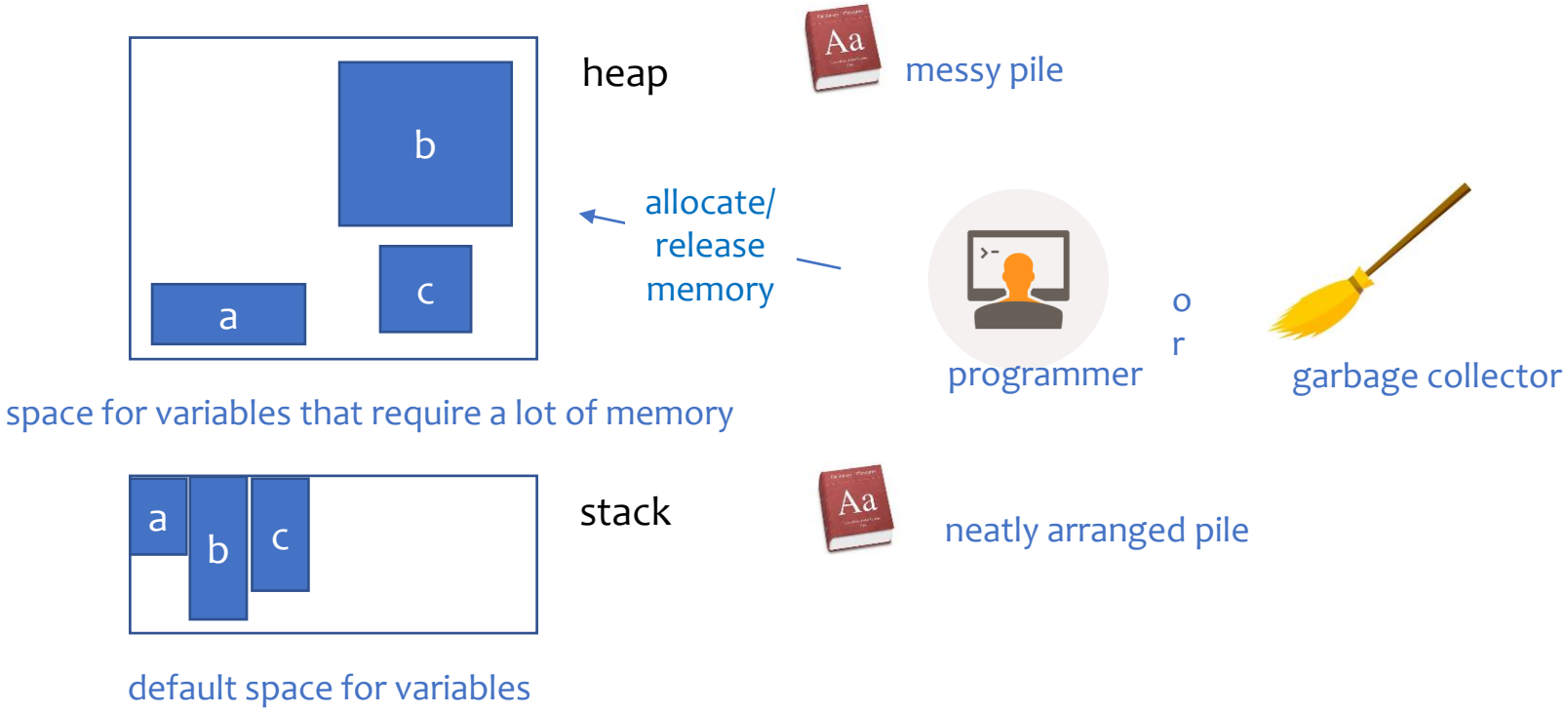
* compiled language *



Features of Python

2) automatic memory management

Computer memory is broadly divided into heap and stack. By default, data and code are allocated on the stack, but data that requires a lot of memory are allocated on the heap.
Memory allocated on the heap must be returned for the next operation when the variable is no longer needed. Languages such as C and C++ require the user to perform this series of processes themselves. In languages such as Python and Java, the garbage collector performs memory management on behalf of the user.



Features of Python

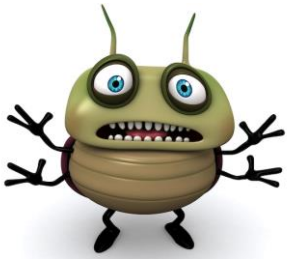
3) dynamic type language

In static type languages such as C, C++, Java, when declaring a variable, the data type must be written as shown below.

int a = 0; float b = 3.414; string c = “hello! world”;

However, in a dynamic type language like Python, there is no need to specify the data type.

a = 0; b = 3.414; c = “hello! world”;



* type error

a = c
static type: An error occurred because an attempt was made to assign a string to an integer variable.
dynamic type: The variable a is automatically converted from an integer variable to a string variable.

a + c
static type: An error occurred because an attempt was made to add a string to an integer variable.
dynamic type: Same as for static type.

Dynamic type languages are convenient because you don't have to worry about the types of variables, but always be aware that bugs can occur due to type errors!

Features of Python

4) A powerful ecosystem for data science

pandas

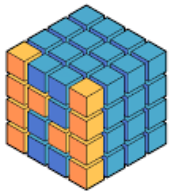
Data processing library

statsmodels

Statistics library

scikit-learn

Machine Learning
(Deep Learning x) library



numpy

Linear algebra library

tensorflow

Google's deep
learning library

scipy

Numerical analysis library

keras

A good deep learning
library for beginners

jupyter

Development environment
in a web browser

matplotlib

Data visualization library

pytorch

Facebook's deep
learning library

How do I install Python?

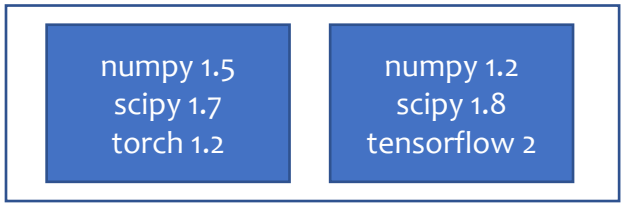
1) Install each package after installing Python (python.org)

- After installing only Python, packages are each installed using pip.
- Python 2 is supported until 2020. You must use Python 3.
- Most computers are 64-bit systems
  (32-bit and 64-bit refer to the amount of space for memory addresses)
- So, if you are not using Mac or Linux, you can download the 64-bit Windows version of Python 3.

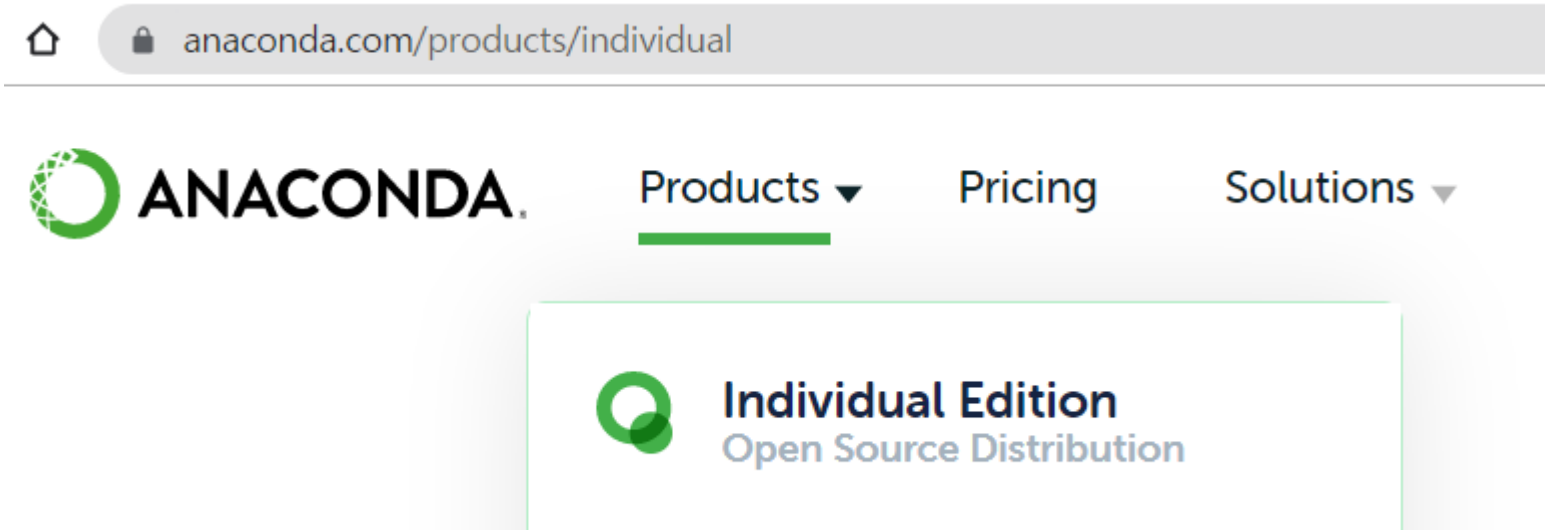
python.org

2) Install python and packages at once with anaconda (anaconda.com)

- Install python and required packages at once with anaconda
- Not all required packages are installed (eg, PyTorch)
- If you're not using a Mac or Linux, you can also get Anaconda for the 64-bit Windows version of Python 3.
- Anaconda allows you to create multiple Python virtual environments on a single computer. It is growing in popularity in line with the recent trend of utilizing high-end servers.



anaconda.com



Anaconda Installers

Windows

Python 3.8

64-Bit Graphical Installer (477 MB)

32-Bit Graphical Installer (409 MB)

MacOS

Python 3.8

64-Bit Graphical Installer (440 MB)

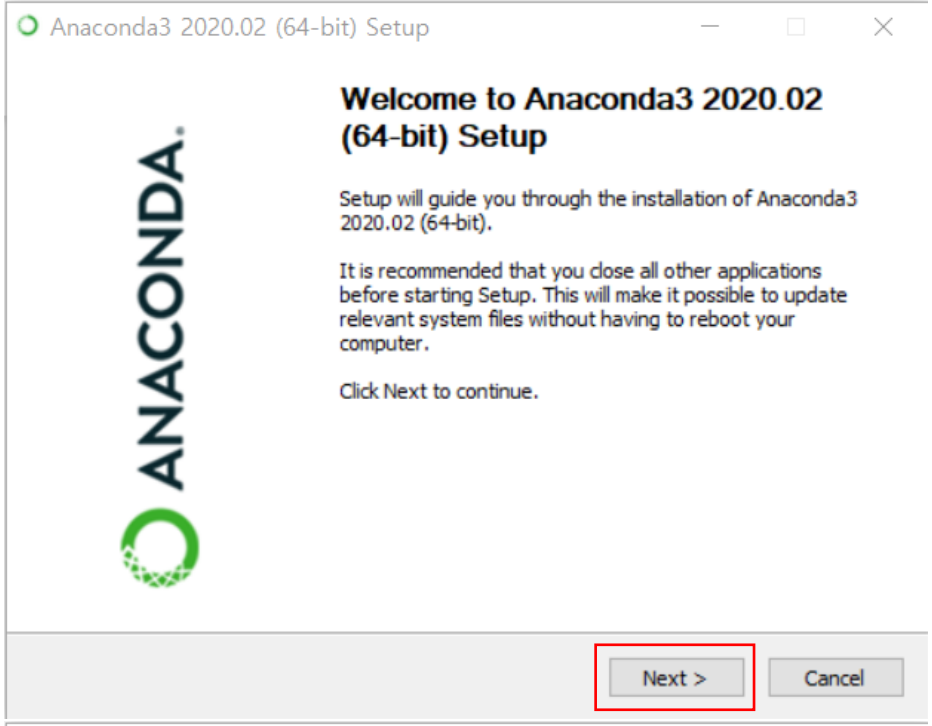
64-Bit Command Line Installer (433 MB)

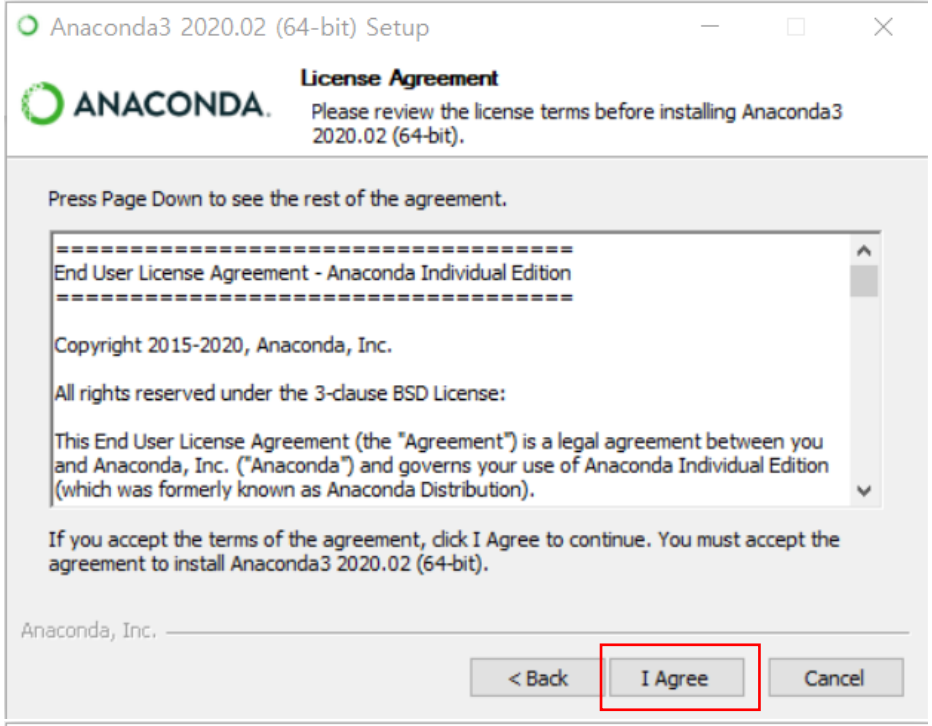
Linux

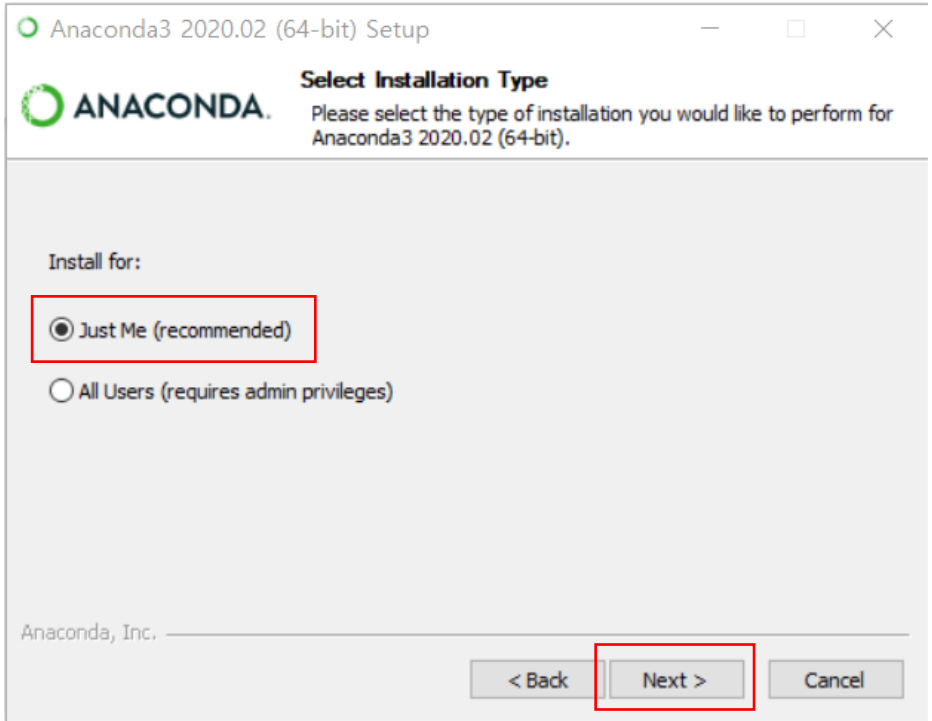
Python 3.8

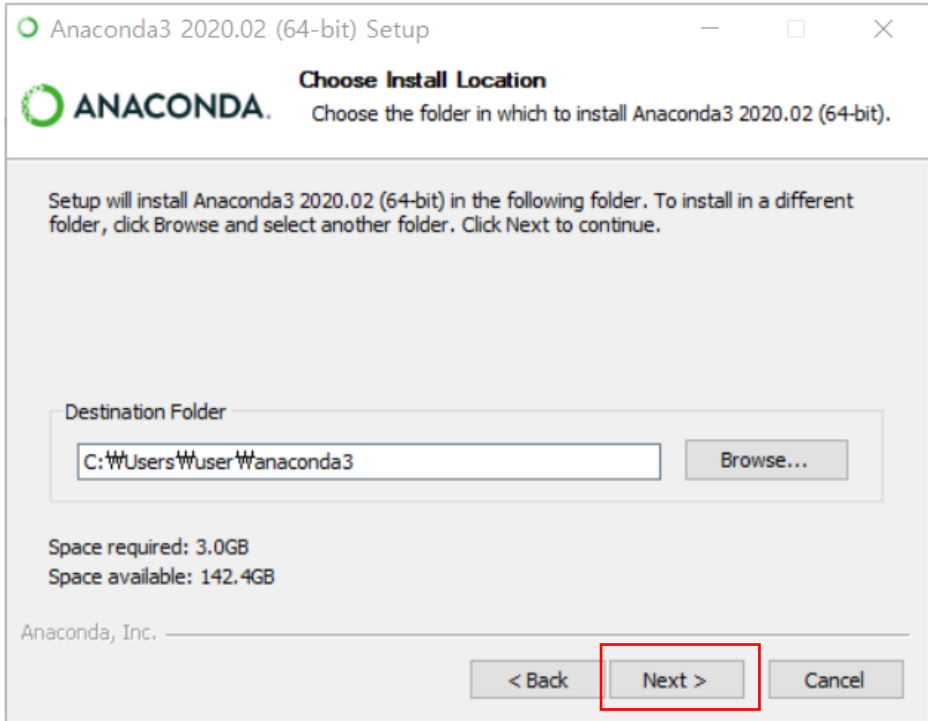
64-Bit (x86) Installer (544 MB)

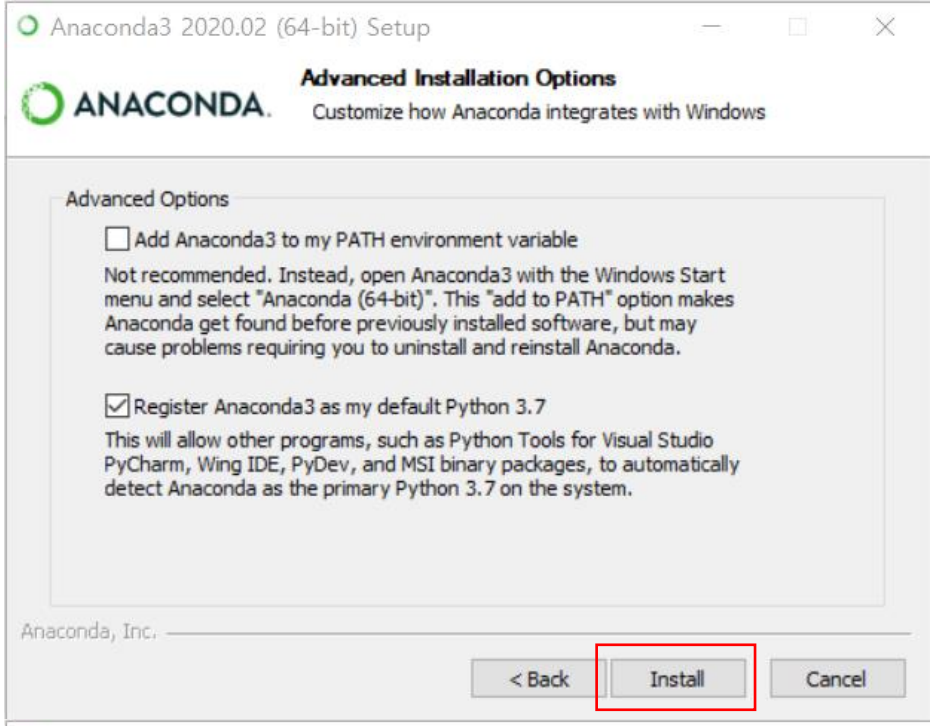
64-Bit (Power8 and Power9) Installer (285 MB)

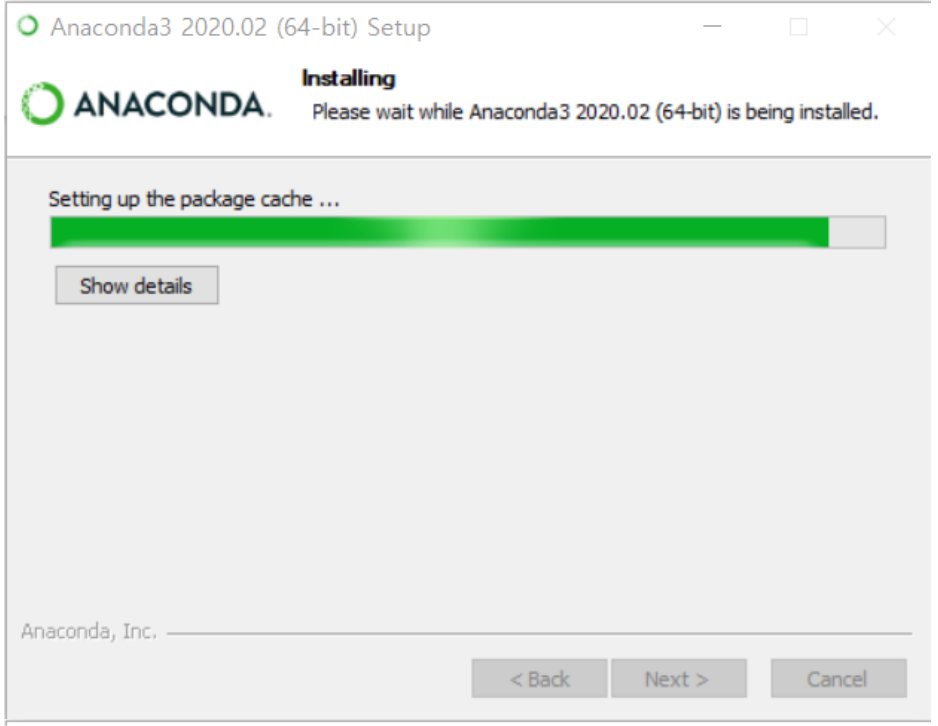

Anaconda3 2020.02 (64-bit) Setup
ANACONDA.
Welcome to Anaconda3 2020.02 (64-bit) Setup
Setup will guide you through the installation of Anaconda3 2020.02 (64-bit).
It is recommended that you close all other applications before starting Setup. This will make it possible to update relevant system files without having to reboot your computer.
Click Next to continue.
Next >
Cancel

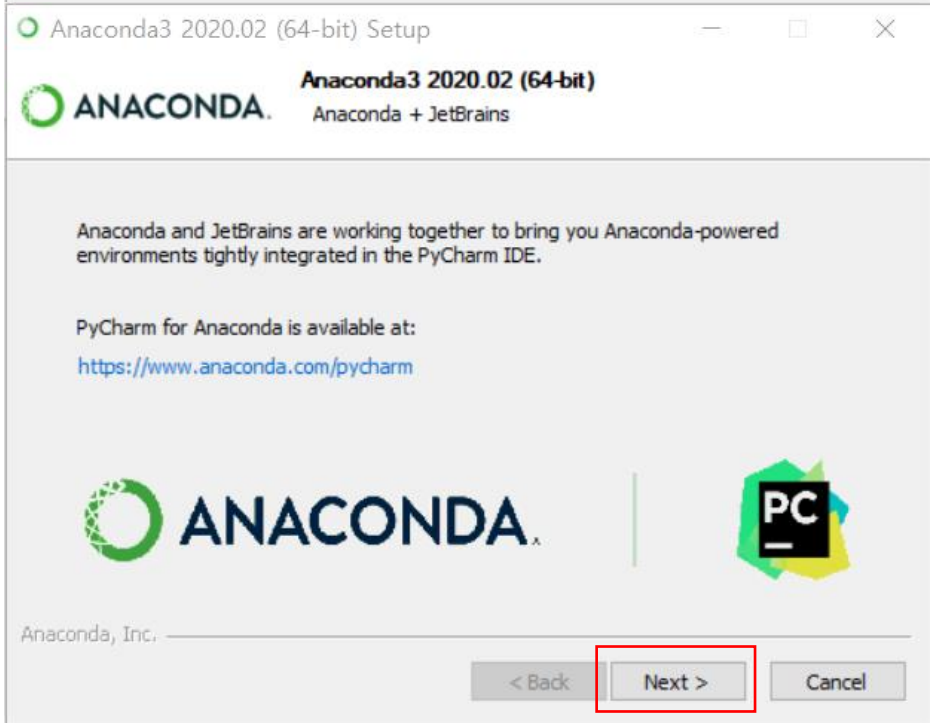

Anaconda3 2020.02 (64-bit) Setup
ANACONDA.
License Agreement
Please review the license terms before installing Anaconda3 2020.02 (64-bit).
Press Page Down to see the rest of the agreement.
End User License Agreement - Anaconda Individual Edition
Copyright 2015-2020, Anaconda, Inc.
All rights reserved under the 3-clause BSD License:
This End User License Agreement (the "Agreement") is a legal agreement between you and Anaconda, Inc. ("Anaconda") and governs your use of Anaconda Individual Edition (which was formerly known as Anaconda Distribution).
If you accept the terms of the agreement, click I Agree to continue. You must accept the agreement to install Anaconda3 2020.02 (64-bit).
Anaconda, Inc.
< Back
I Agree
Cancel

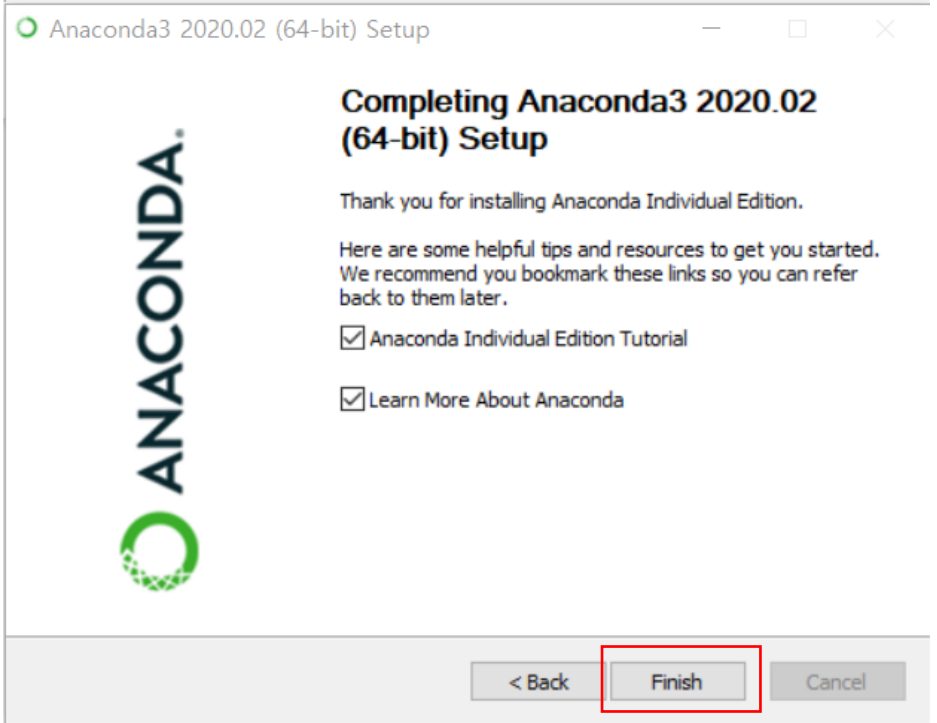

Anaconda3 2020.02 (64-bit) Setup
ANACONDA.
Select Installation Type
Please select the type of installation you would like to perform for Anaconda3 2020.02 (64-bit).
Install for:
Just Me (recommended)
All Users (requires admin privileges)
Anaconda, Inc.
< Back
Next >
Cancel

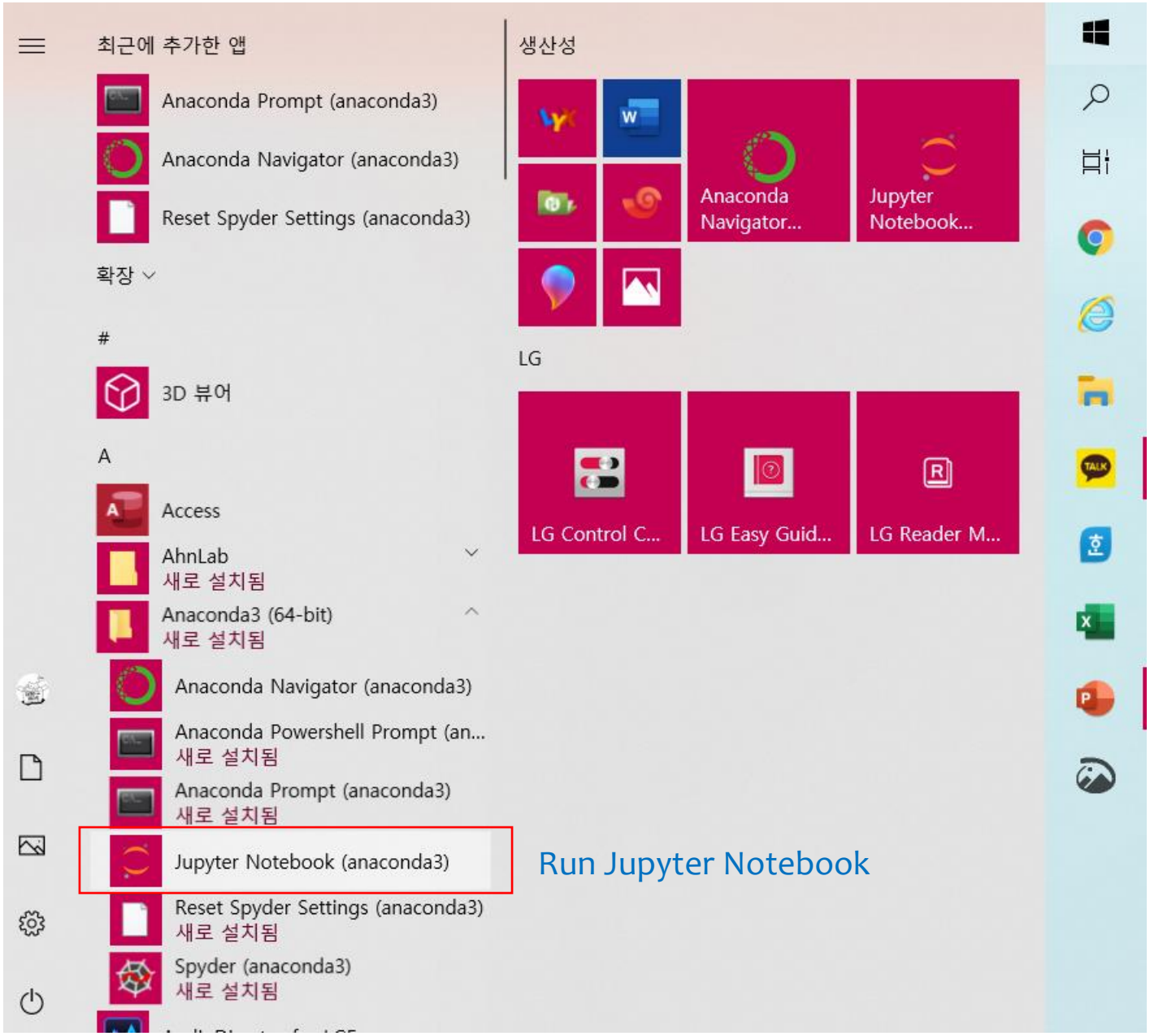

Anaconda3 2020.02 (64-bit) Setup
ANACONDA.
Choose Install Location
Choose the folder in which to install Anaconda3 2020.02 (64-bit).
Setup will install Anaconda3 2020.02 (64-bit) in the following folder. To install in a different folder, click Browse and select another folder. Click Next to continue.
Destination Folder
C:₩Users₩user₩anaconda3
Browse...
Space required: 3.0GB
Space available: 142.4GB
Anaconda, Inc.
< Back
Next >
Cancel


Anaconda3 2020.02 (64-bit) Setup
ANACONDA.
Advanced Installation Options
Customize how Anaconda integrates with Windows
Advanced Options
Add Anaconda3 to my PATH environment variable
Not recommended. Instead, open Anaconda3 with the Windows Start menu and select "Anaconda (64-bit)". This "add to PATH" option makes Anaconda get found before previously installed software, but may cause problems requiring you to uninstall and reinstall Anaconda.
Register Anaconda3 as my default Python 3.7
This will allow other programs, such as Python Tools for Visual Studio PyCharm, Wing IDE, PyDev, and MSI binary packages, to automatically detect Anaconda as the primary Python 3.7 on the system.
Anaconda, Inc.
< Back
Install
Cancel

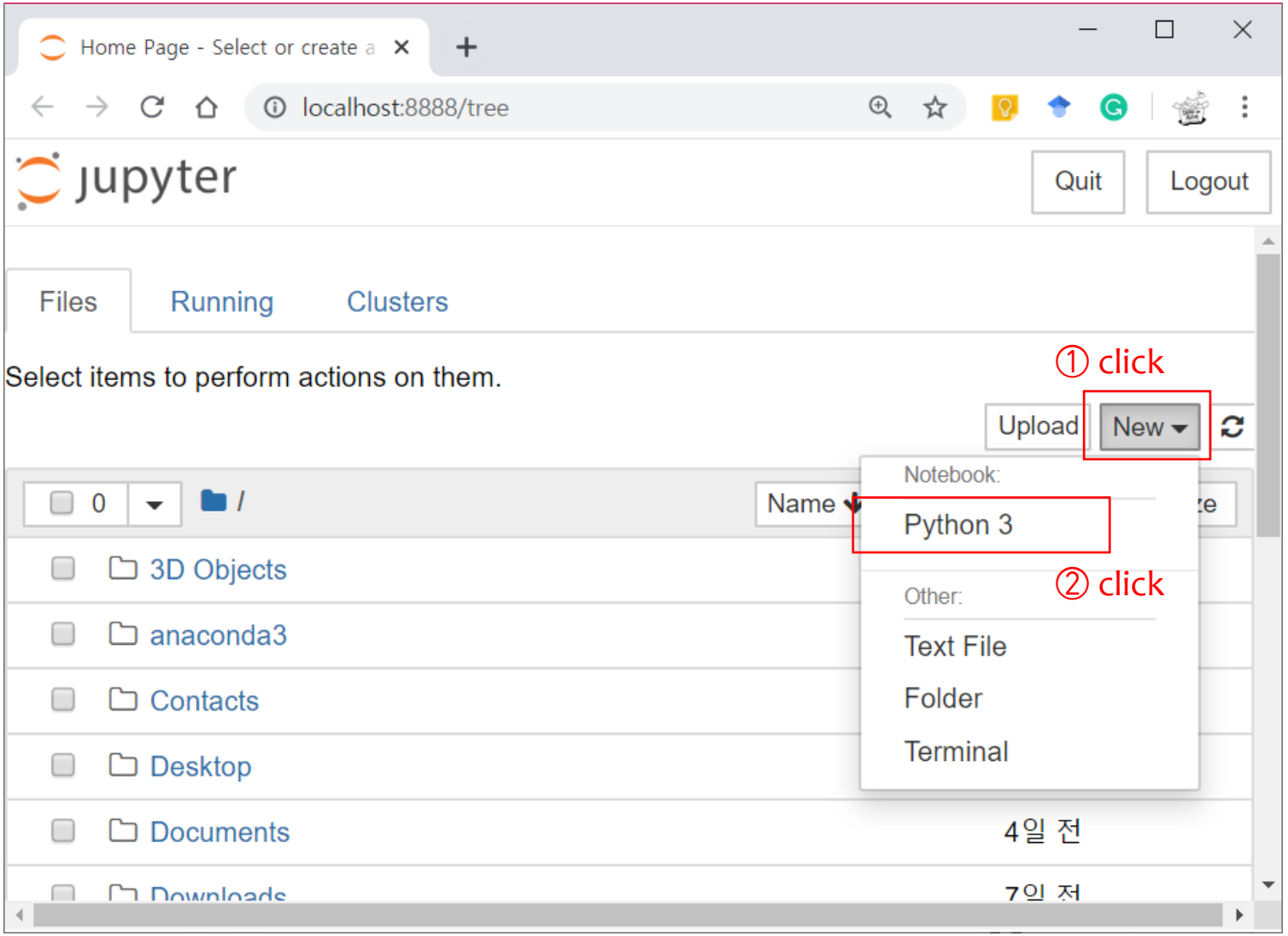

Anaconda3 2020.02 (64-bit) Setup
ANACONDA.
Installing
Please wait while Anaconda3 2020.02 (64-bit) is being installed.
Setting up the package cache ...
Show details
Anaconda, Inc.
< Back
Next >
Cancel

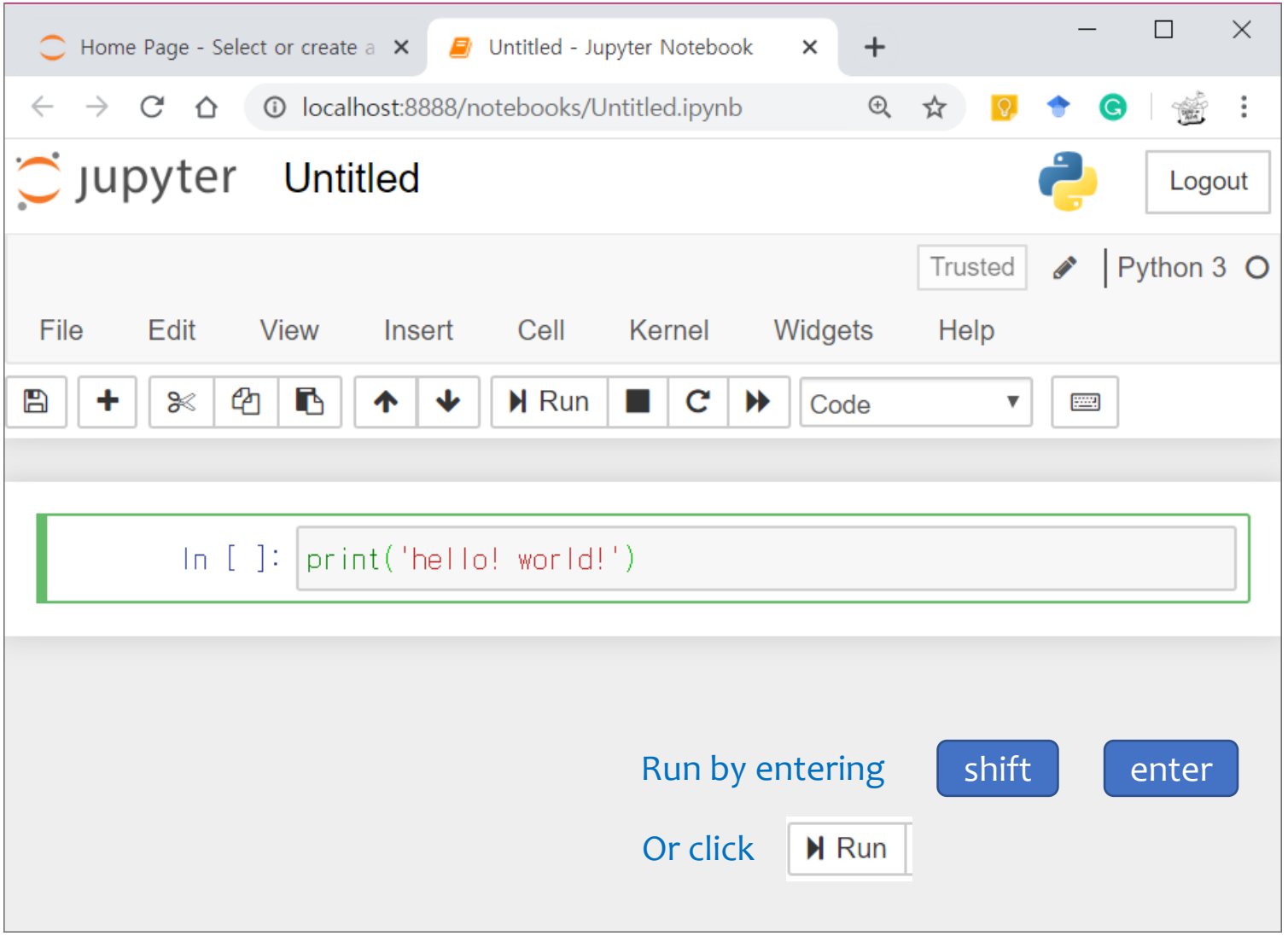

Anaconda3 2020.02 (64-bit) Setup
ANACONDA.
Anaconda3 2020.02 (64-bit)
Anaconda + JetBrains
Anaconda and JetBrains are working together to bring you Anaconda-powered environments tightly integrated in the PyCharm IDE.
PyCharm for Anaconda is available at:
https://www.anaconda.com/pycharm
ANACONDA.
PC
Anaconda, Inc.
< Back
Next >
Cancel

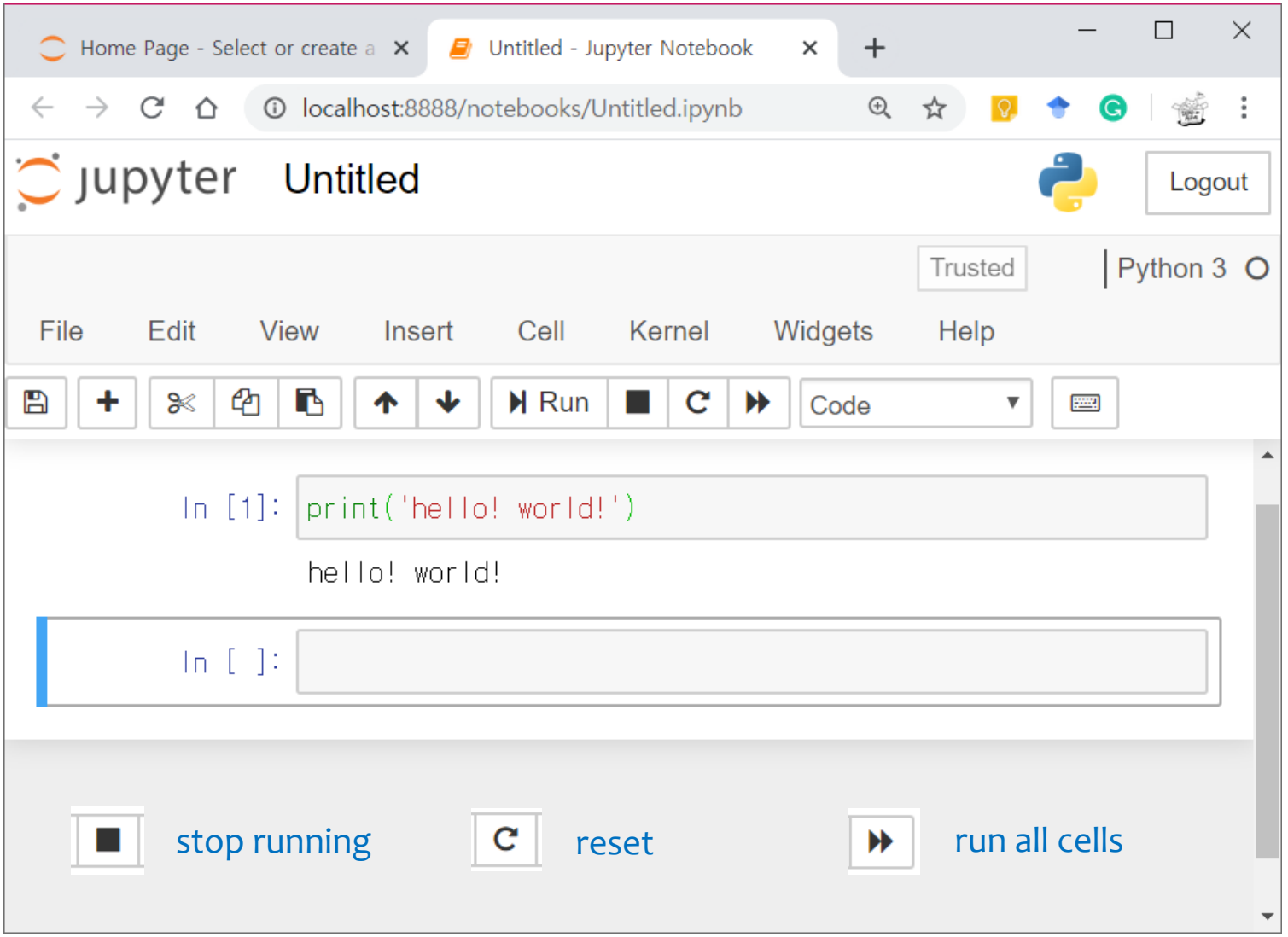

Anaconda3 2020.02 (64-bit) Setup
ANACONDA.
Completing Anaconda3 2020.02 (64-bit) Setup
Thank you for installing Anaconda Individual Edition.
Here are some helpful tips and resources to get you started. We recommend you bookmark these links so you can refer back to them later.
Anaconda Individual Edition Tutorial
Learn More About Anaconda
< Back
Finish
Cancel

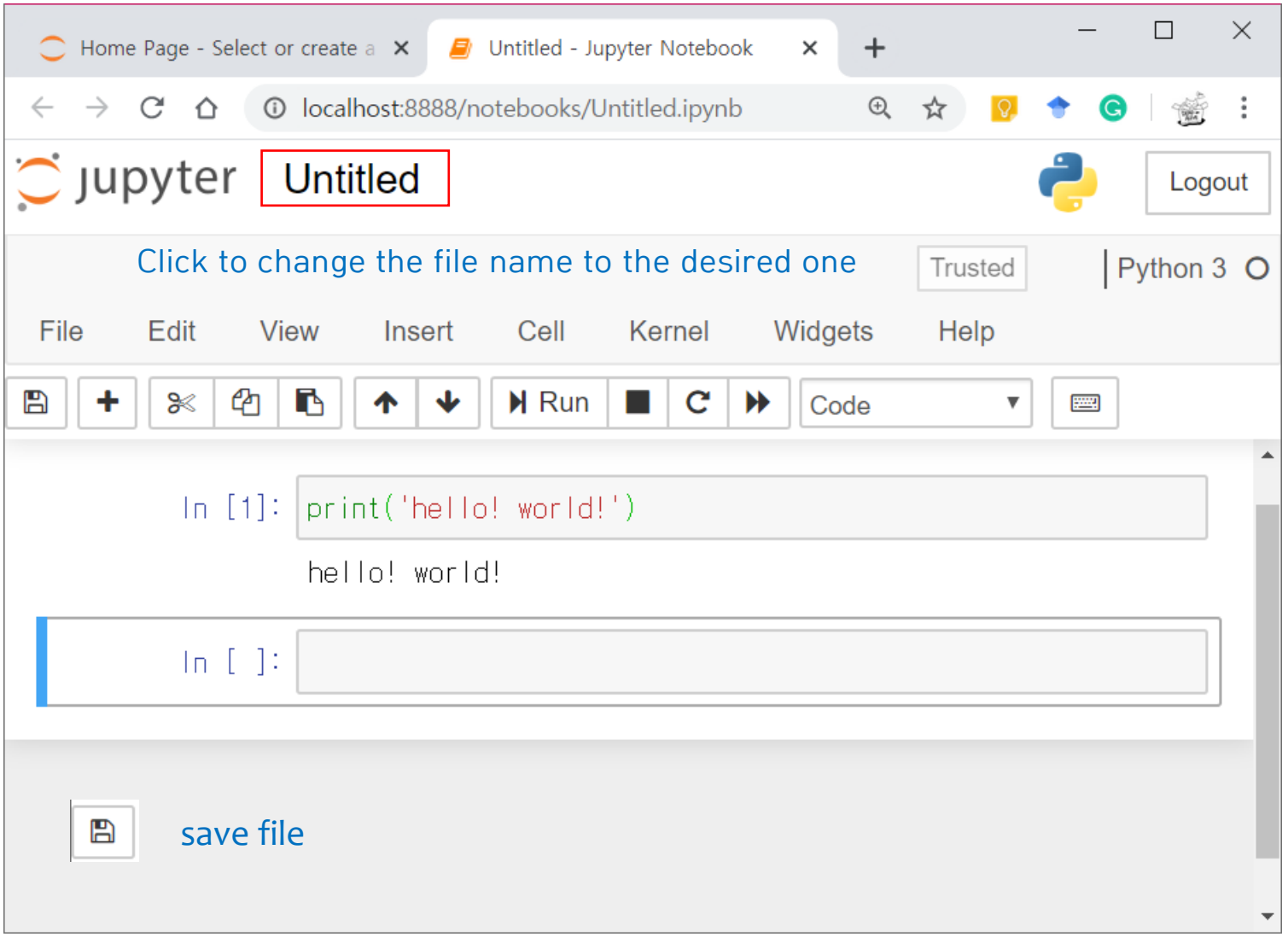


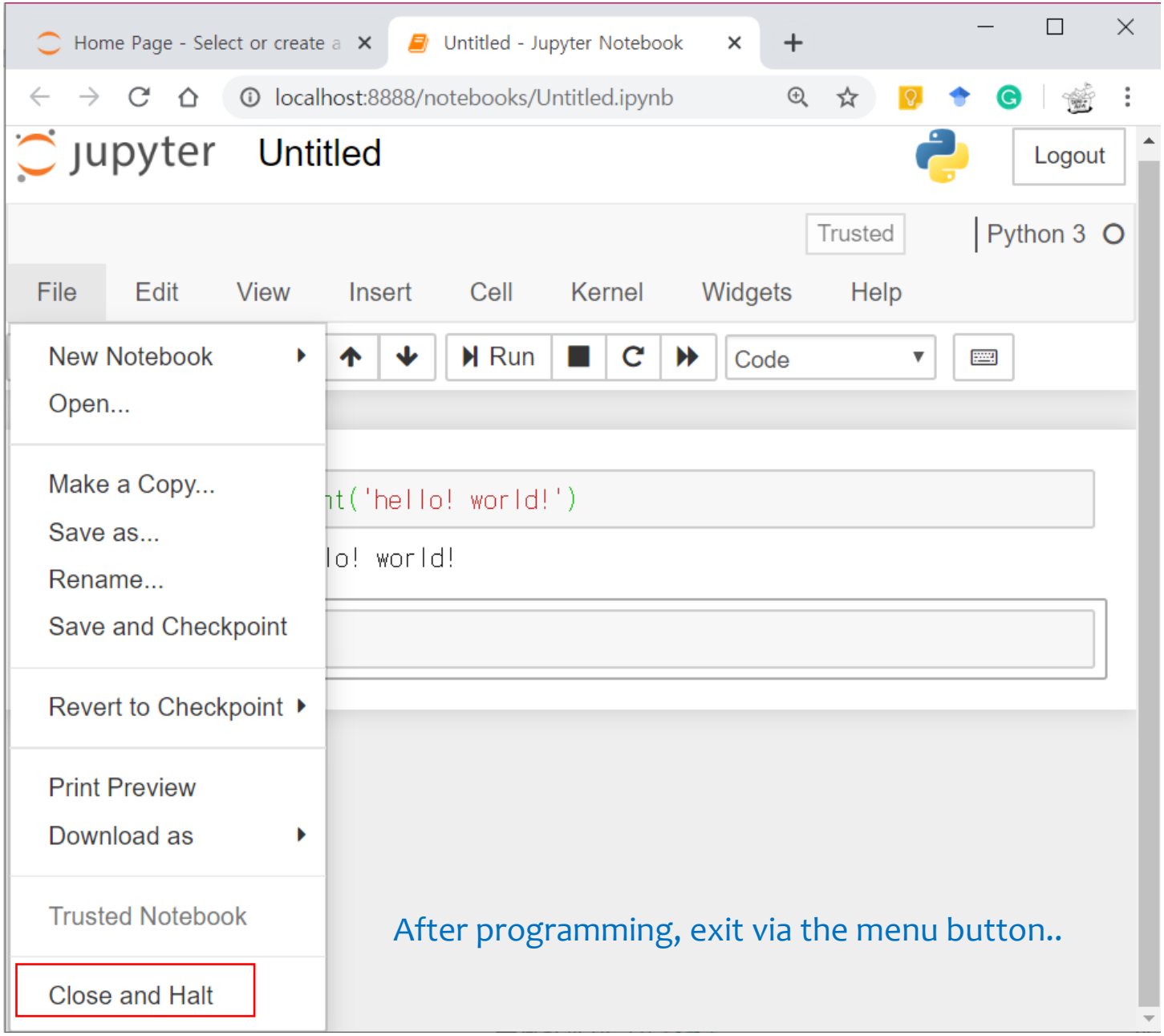
Run Jupyter Notebook


Home Page - Select or create a
localhost:8888/tree
jupyter
Quit
Logout
Files
Running
Clusters
Select items to perform actions on them.
① click
Upload
New
Notebook:
Python 3
② click
Other:
Text File
Folder
Terminal
0
/
Name
3D Objects
anaconda3
Contacts
Desktop
Documents
4일 전
Downloads
7일 전


Home Page - Select or create a
Untitled - Jupyter Notebook
localhost:8888/notebooks/Untitled.ipynb
jupyter
Untitled
Logout
Trusted
Python 3
File
Edit
View
Insert
Cell
Kernel
Widgets
Help
Run
Code
In [ ]:
print('hello! world!')
Run by entering
shift
enter
Or click
Run


Home Page - Select or create a
Untitled - Jupyter Notebook
localhost:8888/notebooks/Untitled.ipynb
jupyter
Untitled
Logout
Trusted
Python 3
File
Edit
View
Insert
Cell
Kernel
Widgets
Help
Run
Code
In [1]:
print('hello! world!')
hello! world!
In [ ]:
stop running
reset
run all cells


Home Page - Select or create a
Untitled - Jupyter Notebook
localhost:8888/notebooks/Untitled.ipynb
jupyter
Untitled
Logout
Click to change the file name to the desired one
Trusted
Python 3
File
Edit
View
Insert
Cell
Kernel
Widgets
Help
Run
Code
In [1]:
print('hello! world!')
hello! world!
In [ ]:
save file


Home Page - Select or create a
Untitled - Jupyter Notebook
localhost:8888/notebooks/Untitled.ipynb
jupyter
Untitled
Logout
Trusted
Python 3
File
Edit
View
Insert
Cell
Kernel
Widgets
Help
New Notebook
Open...
Make a Copy...
Save as...
Rename...
Save and Checkpoint
Revert to Checkpoint
Print Preview
Download as
Trusted Notebook
Close and Halt
Run
Code
nt('hello! world!')
lo! world!
After programming, exit via the menu button..

It provides a Python development environment that is almost like that of Jupyter Notebook.
As a result, data analysis using Python is possible even if anaconda installation fails.

colab

Google CoLab https://colab.research.google.com (Google ID required)

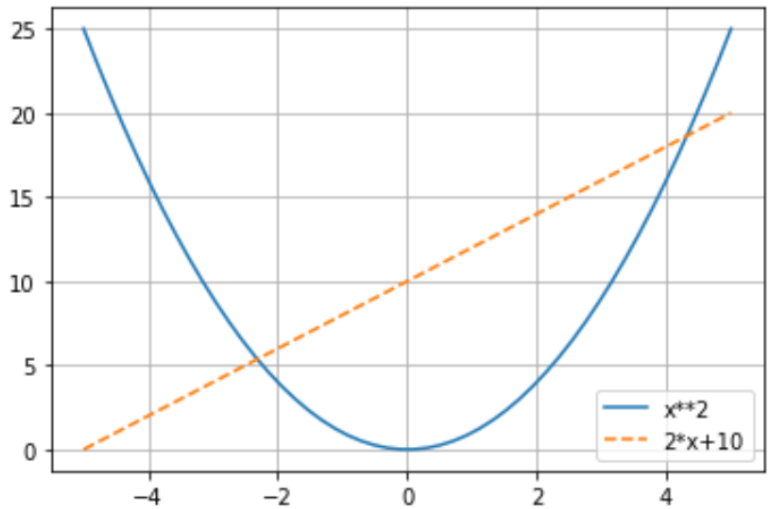
chonnamipynb ☆

파일 수정 보기 삽입 런타임 도구 도움말 저장 중...

+ 코드 + 텍스트

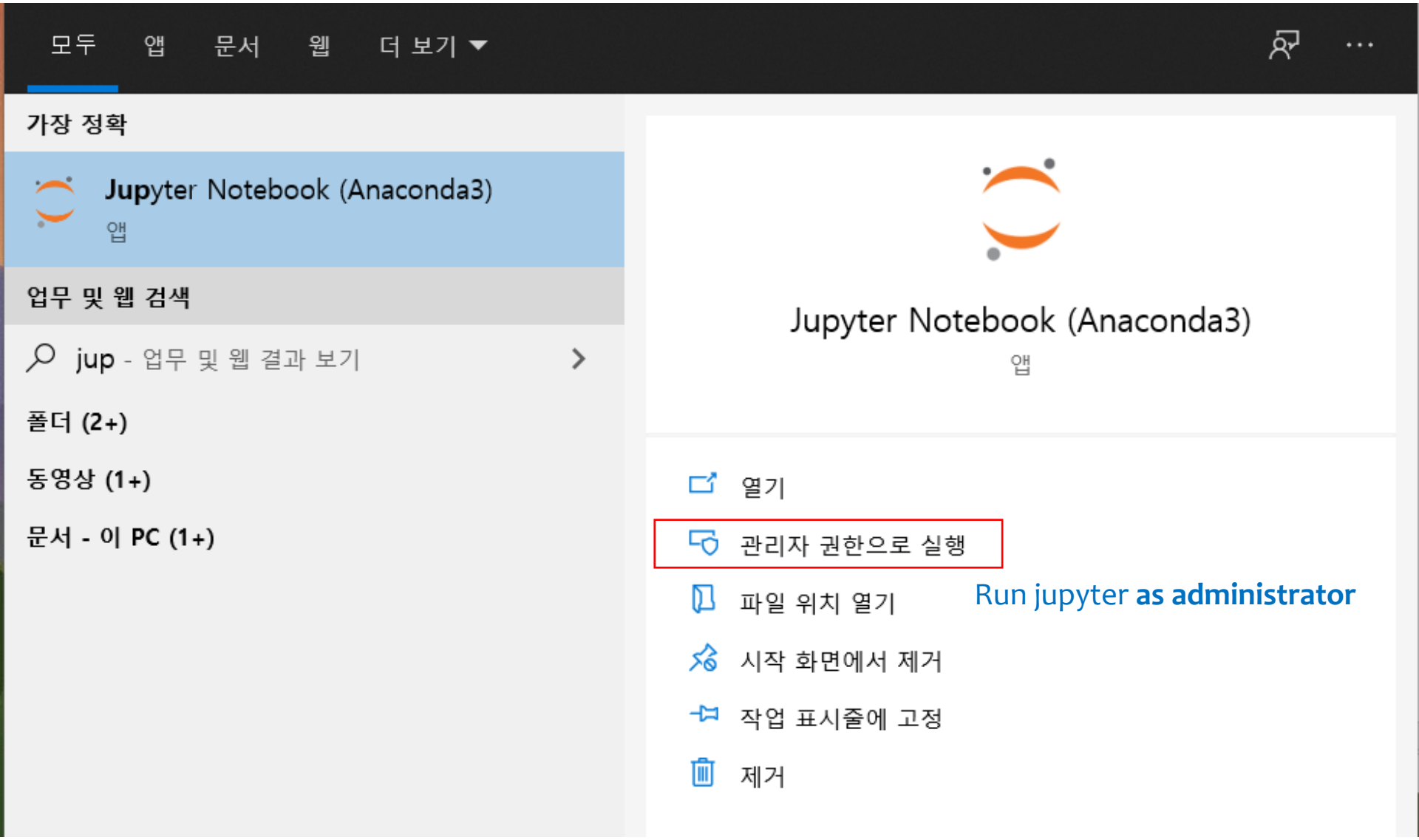
```
[1] import numpy as np
    import matplotlib.pyplot as plt
```

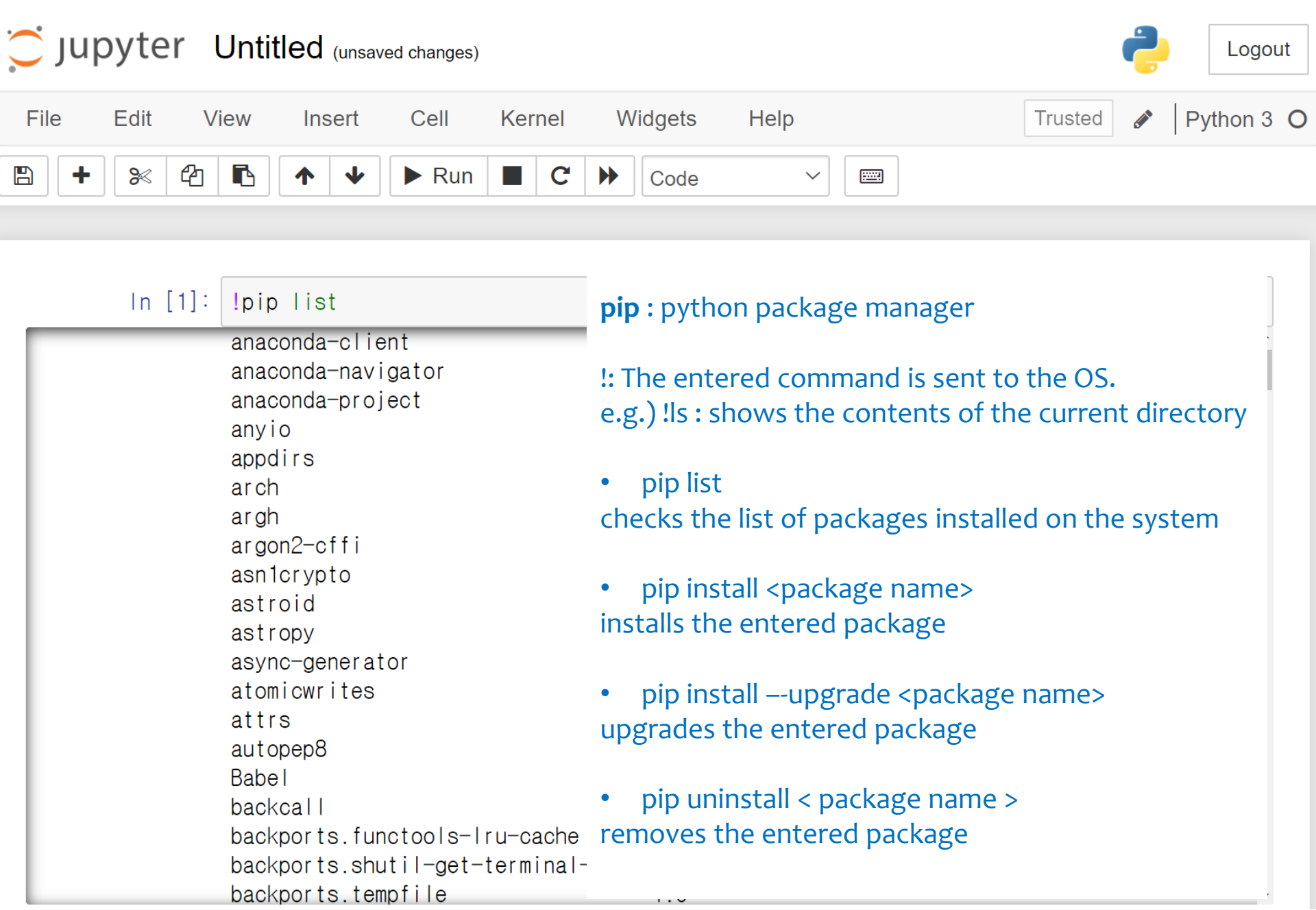
```
x = np.linspace(-5,5)
y1 = x**2
y2 = 2*x+10
plt.plot(x,y1)
plt.plot(x,y2,'--')
plt.legend(['x**2','2*x+10'])
plt.grid()
plt.show()
```



integrated with Google Drive

Some of Google's GPU resources are available
(The devices are suitable for deep learning)


모두
앱
문서
웹
더 보기
가장 정확
Jupyter Notebook (Anaconda3)
앱
업무 및 웹 검색
jup - 업무 및 웹 결과 보기
폴더 (2+)
동영상 (1+)
문서 - 이 PC (1+)
Jupyter Notebook (Anaconda3)
앱
열기
관리자 권한으로 실행
파일 위치 열기
시작 화면에서 제거
작업 표시줄에 고정
제거
Run jupyter as administrator



pip : python package manager

!: The entered command is sent to the OS.
e.g.) !ls : shows the contents of the current directory

- pip list

checks the list of packages installed on the system

- pip install <package name>

installs the entered package

- pip install –-upgrade <package name>

upgrades the entered package

- pip uninstall < package name >

removes the entered package

jupyter Untitled (unsaved changes)

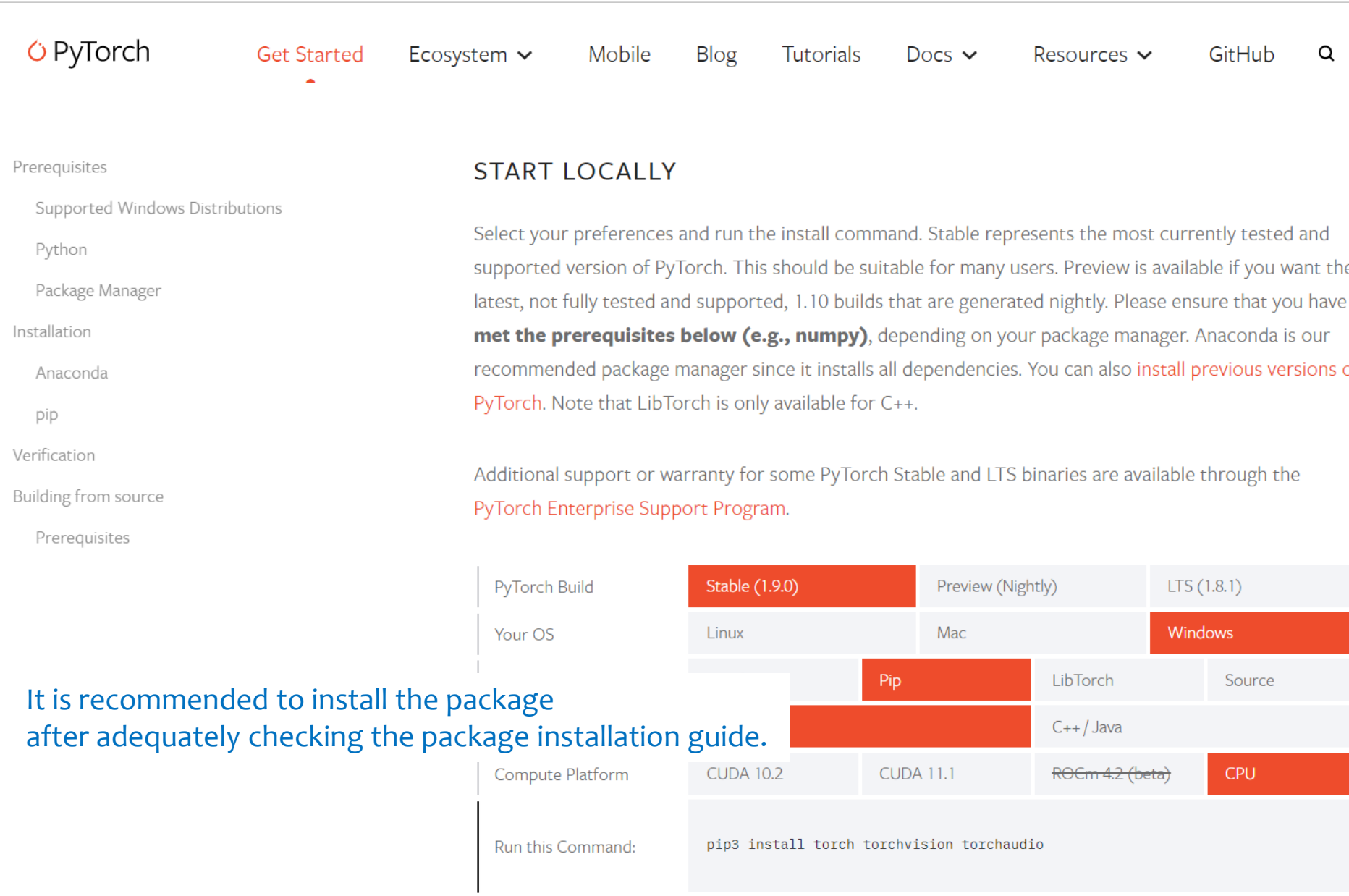
Logout

File Edit View Insert Cell Kernel Widgets Help Not Trusted Python 3

Run Code

In [1]:
```
!pip install autokeras
```

```
Requirement already satisfied: threadpoolctl>=2.0.0 in c:₩programdata₩anaconda3₩lib
₩site-packages (from scikit-learn->autokeras) (2.1.0)
Installing collected packages: numpy, markdown, grpcio, google-auth-oauthlib, absl-p
y, termcolor, tensorflow-estimator, tensorboard, opt-einsum, kt-legacy, keras-prepro
cessing, keras-nightly, h5py, google-pasta, gast, flatbuffers, astunparse, tensorflo
w, keras-tuner, autokeras
  Attempting uninstall: numpy
    Found existing installation: numpy 1.20.1
    Uninstalling numpy-1.20.1:
      Successfully uninstalled numpy-1.20.1
  Attempting uninstall: h5py
    Found existing installation: h5py 2.10.0
    Uninstalling h5py-2.10.0:
      Successfully uninstalled h5py-2.10.0
Successfully installed absl-py-0.13.0 astunparse-1.6.3 autokeras-1.0.15 flatbuffers-
1.12 gast-0.4.0 google-auth-oauthlib-0.4.5 google-pasta-0.2.0 grpcio-1.34.1 h5py-3.
1.0 keras-nightly-2.5.0.dev2021032900 keras-preprocessing-1.1.2 keras-tuner-1.0.3 kt
-legacy-1.0.3 markdown-3.3.4 numpy-1.19.5 opt-einsum-3.3.0 tensorboard-2.5.0 tensorf
low-2.5.0 tensorflow-estimator-2.5.0 termcolor-1.1.0
```



It is recommended to install the package
after adequately checking the package installation guide.