Tutorial of Anaconda (Python, Jupyter Notebook) installation

Contents

1 Introduction 2
2 Install Anaconda 2
2.1 Download Anaconda 2
2.2 Install Anaconda 3
2.3 Anaconda Prompt 3
2.4 Verify Anaconda 4
2.5 Install completed 4
3 Run the Notebook 4
3.1 Change the default work directory 5
4 Install the required Python packages 6

1 Introduction

This is a tutorial on how to properly install Anaconda, Python and Jupyter Notebook, preparing for the Data Science and Machine learning course provided by EOM, SDU.

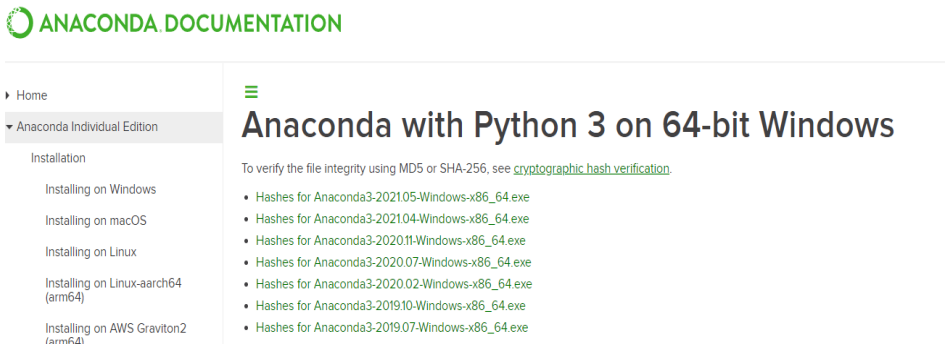
2 Install Anaconda

2.1 Download Anaconda

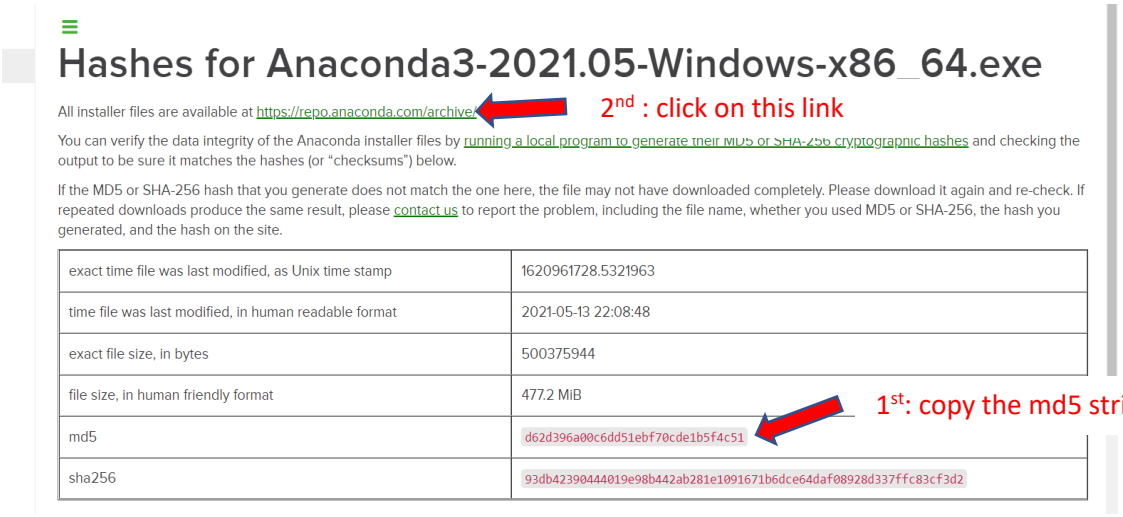
Download Anaconda from the following link (This file consists of different installation versions of Anaconda for any OS. **Python and Jupyter Notebook are built-in with this installation file, which means you do not need to download and install them separately**): https://docs.anaconda.com/anaconda/install/hashes/win-3-64/

The following steps show how you can properly download the desired file:

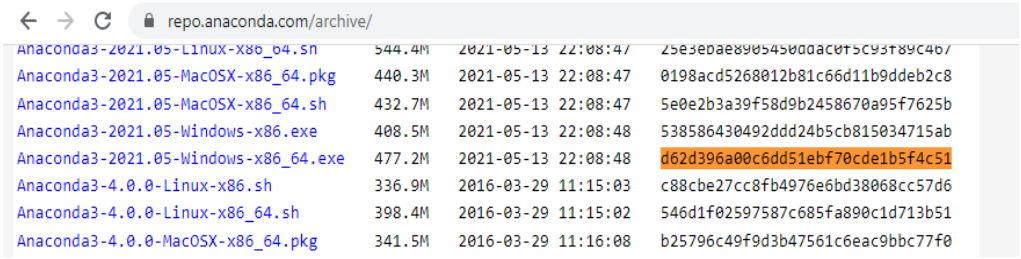
a) Click on the provided link, select one of the executive files based on your own OS system. Here I use the 64-bit window OS as an example. You will see the same webpage, as shown in the following image.



b) Select the latest version and go to the specific web page, copying the md5 string and then clicking on the installer files link, as shown below.



c) Look up the specific executive file by the copied md5, and download the file, as show below.

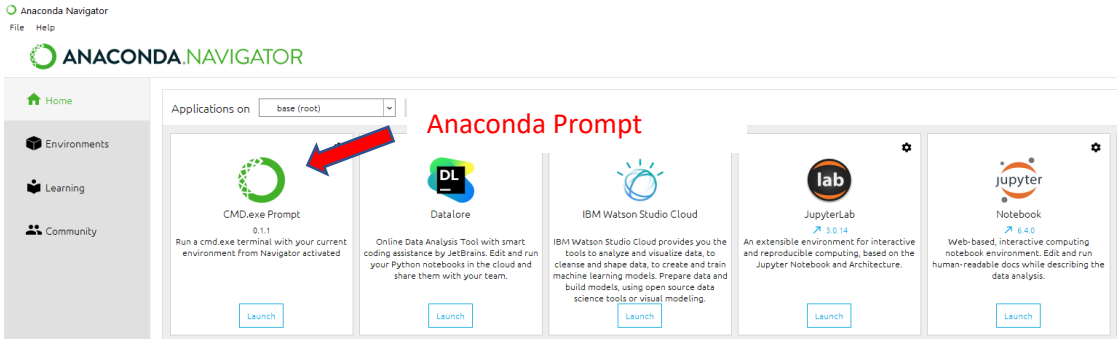


2.2 Install Anaconda

Install the version of Anaconda which you have downloaded, following the instructions on the download page. **Note that** the latest version of Python will be installed automatically together with Anaconda.

2.3 Anaconda Prompt

Open "Anaconda Prompt" in the "Anaconda Navigator".



2.4 Verify Anaconda

Type the command to verify if Anaconda and Python are installed.

```
(base) C:\Users\jiec>python --version
Python 3.8.10
```

```
(base) C:\Users\jiec>conda --version
conda 4.10.3
```

2.5 Install completed

Congratulations, you have installed Anaconda, Python and Jupyter Notebook successfully.

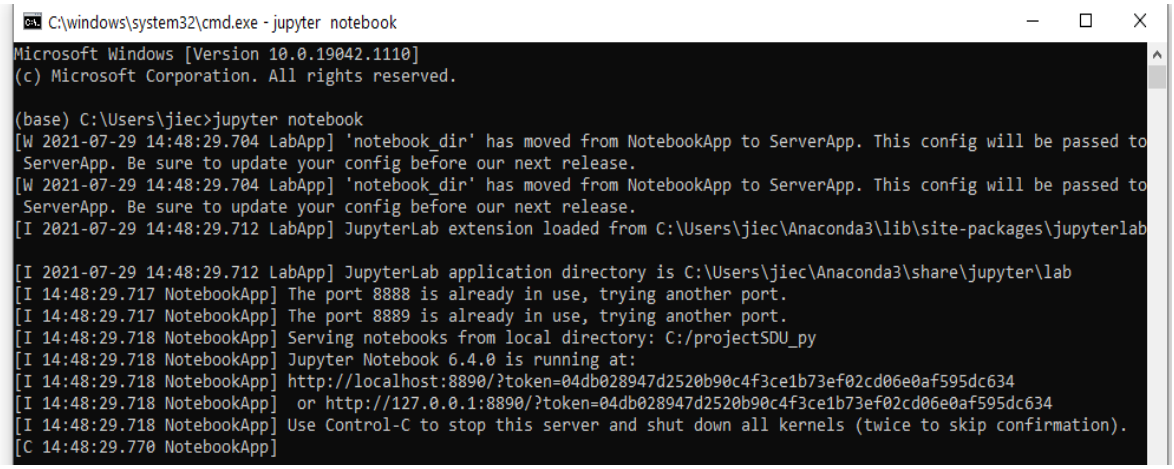
3 Run the Notebook

In this section, I will show you how to run the Notebook properly after installation. Basically, there are two options you can use to open the Notebook for using. You may choose either of them.

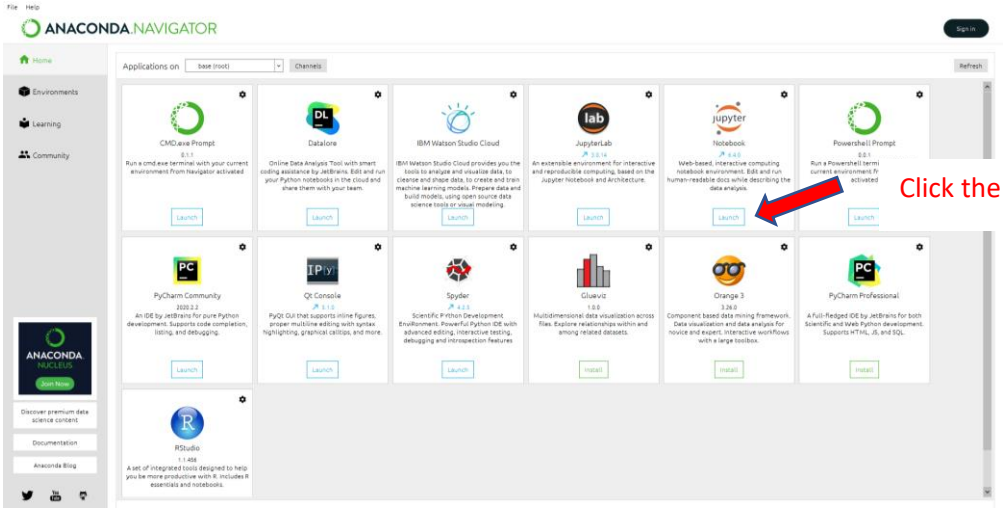
Option 1: Type the command to start the Jupyter Notebook (Open the "Anaconda Prompt" in the "Anaconda Navigator".)

```
(base) C:\Users\jiec>jupyter notebook
```

This will print some information about the Notebook server in your terminal, including the URL of the web application (by default, http://localhost:8888):



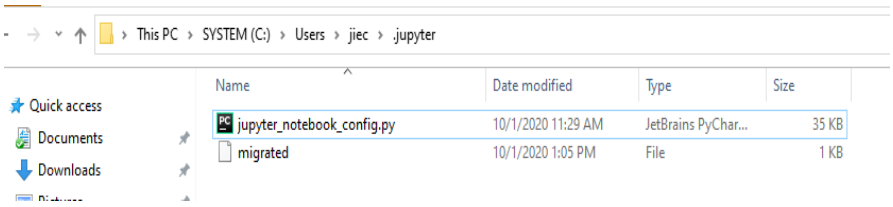
Option 2: Use the Anaconda navigator to start the Jupyter Notebook. The navigator is illustrated below.



3.1 Change the default work directory

Obviously, the default local directory when we start the Notebook is not what we want for our own project. Therefore, we need to change the default work directory. The procedure is shown in the following steps.

a) Open Anaconda Prompt and run: `jupyter notebook --generate-config`.
b) This writes a file to `C:\Users\username\.jupyter\jupyter_notebook_config.py`. An example is presented below:



c) Browse to the file location and open it in an Editor
d) Search for the following line in the file: `#c.NotebookApp.notebook_dir = ''`
e) Replace the line by the new work directory you want to have, as shown below

```
The extensions
#  will be loaded in alphabetical order.
#c.NotebookApp.nbserver_extensions = {}

## The directory to use for notebooks and kernels.
c.NotebookApp.notebook_dir = 'C:/projectSDU_py'

## Whether to open in a browser after starting. The specific
browser used is
#  platform dependent and determined by the python standard
```

The following links provide details so that you can go through them if you are interested in:

Install Jupyter Notebook

https://test-jupyter.readthedocs.io/en/latest/install.html

Run Jupyter Notebook

https://test-jupyter.readthedocs.io/en/latest/running.html#running

How to change the Jupyter start-up folder (Stack Overflow)

https://stackoverflow.com/questions/35254852/how-to-change-the-jupyter-start-up-folder#:~:text=In%20the%20start%20menu%2C%20right,new%20%22D%3A%5Cpath%22%20.

4 Install the required Python packages

As a data scientist, the popular Python packages you need to install include, but not limited to, **Pandas, Numpy, Matplotlib, Scikit-learn and Keras**, etc. When we adopt the method in Section 1 to install Anaconda, some default packages such as Numpy and Pandas are automatically installed. Thus, you can first check if these packages are available. By using this command: `Conda list`, you will display all the installed Python packages, as shown below.

```
(base) C:\Users\jiec>conda list
# packages in environment at C:\Users\jiec\Anaconda3:
#
# Name                    Version                   Build  Channel
_anaconda_depends         2020.07                  py38_0
_ipyw_jlab_nb_ext_conf    0.1.0                    py38_0
alabaster                 0.7.12             pyhd3eb1b0_0
altgraph                  0.17                     pypi_0    pypi
anaconda                  custom                   py38_1
anaconda-client           1.8.0            py38haa95532_0
anaconda-navigator        2.0.4                    py38_0
anaconda-project          0.10.1             pyhd3eb1b0_0
anyio                     2.2.0            py38haa95532_2
appdirs                   1.4.4                    pypi_0    pypi
argh                      0.26.2                   py38_0
argon2-cffi               20.1.0           py38h2bbff1b_1
asn1crypto                1.4.0                     py_0
```

Other useful Conda commands are listed here:

`conda install pandas` — Download and install the pandas package from the main conda channel.

`conda install -c conda-forge pandas` — Download and install the pandas package from the main conda-forge channel.

`conda update pandas` — Download and install the latest pandas package from the main conda channel.

`conda update all` — Download and install the the latest versions of all installed packages. Can be very slow if you have a lot of outdated packages.

`conda uninstall pandas` — Uninstall the pandas package in your conda environment.