

Software Lab for Basic Scientific Problem Solving

Lab 2: Introduction to Anaconda and Jupyter Notebook

August 20, 2021

OBJECTIVE

This lab aims to describe the basic components included in Anaconda. Students will explore several tools available in the package and understand the purpose of each tool in order to use it for specific purpose. Especially, Jupyter notebook is a fancy and popular tool among scientists that use python as a programming language to perform their experiments (e.g. data analysis, data manipulation, and statistical calculation).

Note that:

- Each lab consists of a number of practical exercises, all of which are required for marks to be attained. In particular, you should try to accomplish the 3 labelled milestones.
- Attendance is compulsory and will be monitored on Mycourse. If you are unable to attend a session you must inform a tutor in advance.

During the assessments, your understandings will be tested, and the tutor will not give the marks if you do not display an understanding of the problem and its solution. The tutors are there to help you think through any issues — but they won't write the programs for you!

The normal rules of assessment apply: if you can't make a laboratory or miss one, you must have a good reason. If the problem is medical then a doctor's note must be presented as evidence. You should attempt to make up for lab absences in your own time.

Finally, you will be allowed to submit work in the lab a week after it was given. The work for this lab (lab2) can be submitted at the start of the lab next week (lab3). However, the marks will not be awarded beyond the next lab, meaning lab2 will not be marked during lab 4, 5 and so on.

WHAT IS ANACONDA?

Actually, the name "Anaconda" is used by two products. One is the installation program used by Fedora, Red Hat, and some other Linux distributions. Anaconda is also known as the leading open data science platform powered by Python the

fastest growing open data science and scientific programming language. We thus refer to later meaning of “Anaconda”. In other words, Anaconda is a Python distribution that is particularly popular for data analysis and scientific computing. Anaconda is a free, enterprise-ready Python distribution for data analytics, processing, and scientific computing. Anaconda comes with Python 2.7 or Python 3.8 and 100+ cross-platform tested and optimized Python packages. All of the usual Python ecosystem tools work with Anaconda. Additionally, Anaconda can create custom environments that mix and match different Python versions (2.6, 2.7, 3.3 or 3.8) and other packages into isolated environments and easily switch between them.

EXERCISE 0: PYTHON VERSION

Visit <https://www.anaconda.com/>.

What is the current version of Python in Anaconda available on the website?

Answer:

EXERCISE 1: INSTALL ANACONDA

Download Anaconda Installer from

<https://www.anaconda.com/products/individual#windows>.

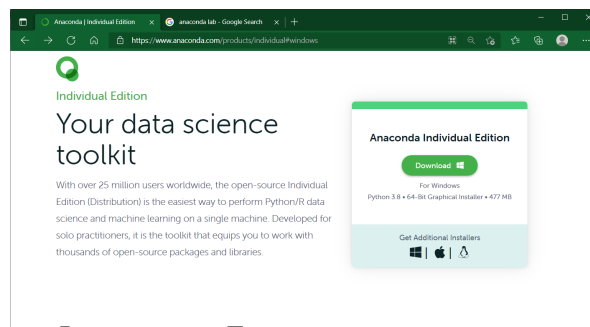


Figure 1: Down Anaconda Individual Edition.

Step I: start installation

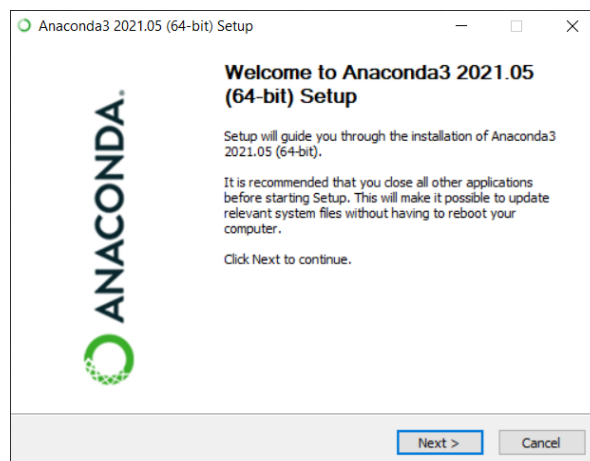


Figure 2: Begin to install Anaconda.

Step II: Choice to install only you

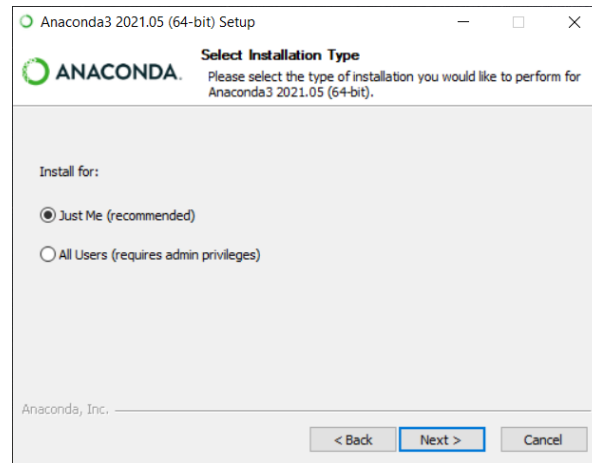


Figure 3: Play save to install for only one user

Step III: Browse to folder that you want to install Anaconda

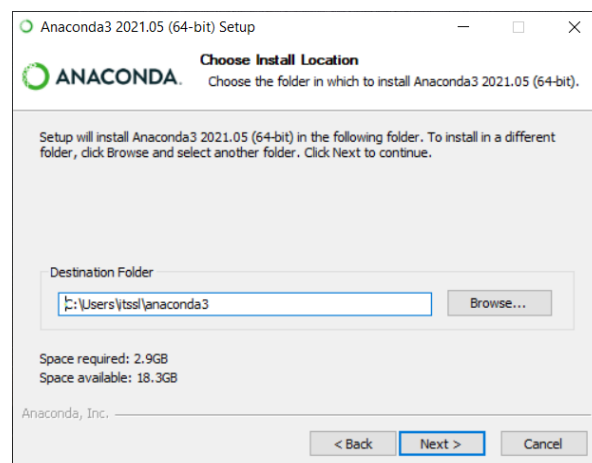


Figure 4: default path in your home directory

Step IV: Python Environment setup: default

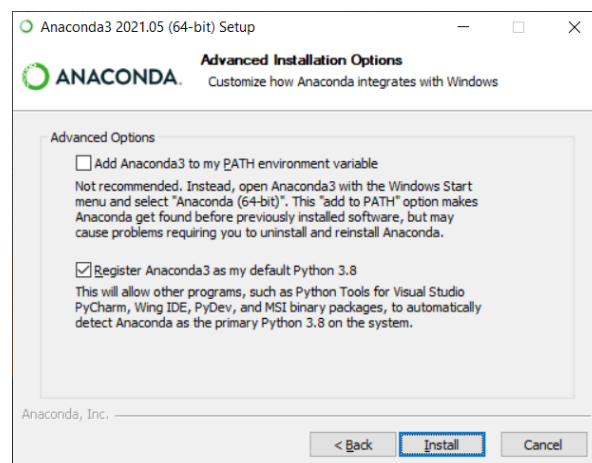


Figure 5: Insolated Anaconda from Pyton windows version: Recommnd selection

EXERCISE 2: CHECK THE ANACONDA NAVIGATOR

Now, you have to check whether Anaconda is installed in your machine.

Click “start” and look for the Anaconda folder, and click “Anaconda Navigator”. At the Anaconda Navigator, click at the “Home” tab. How many applications available in Anaconda? List all of them here.

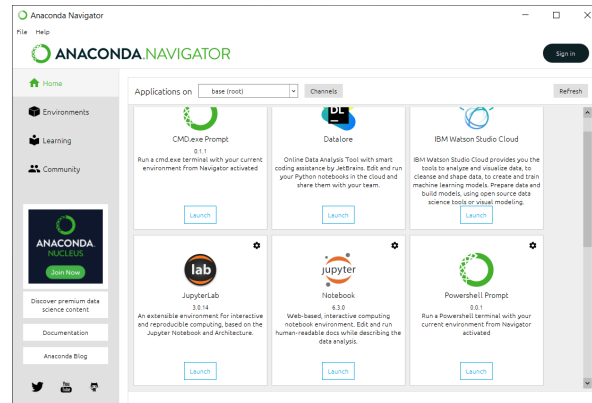


Figure 6: Anaconda Navigator

Answer:

On left hand side of Navigator Click the “Environments” tab, how many packages are installed in the Anaconda package installed in your machine?

Answer:

How many packages are available in the Anaconda ? Those are ready for you to install.

Answer:

EXERCISE 3: PLAY WITH ANACONDA CONSOLE

Basically, using command line is the best way to interact with Python and Anaconda.

Click “start” and select “Anaconda Prompt” in the Anaconda folder. You will see the Anaconda command prompt ready for your command.

Now you have to check the version of Anaconda installed in your machine by using command line mode.

Type `conda list anaconda$`

What is the output? and what is the current version of Anaconda that is installed in your machine?

Answer:

Next, you need to check the installed packages by using the following command

```
conda list
```

Is the list that appear in the command line mode same with the list of installed packages that you saw in the Anaconda navigation?

Answer:

Now, you need to learn how to update Anaconda. First, it is better to update the package manager (conda).

Type `conda update conda`

Are there any packages need to be updated? What are they?

Answer:

You can check which version of Python that Anaconda is using, and also on which platform that it is running on, along with base paths for environment and packages, just type `conda info`

milestone 1) You have now completed the Milestone 1. You should get sign off by your lab assistant.

EXERCISE 4: INSTALL A NEW PACKAGE USING ANACONDA PROMPT

You need to know how to install new packages using Anaconda prompt in cases that the installed packages are not enough to solve your problem. We use Git package as an example. Git is one of the popular version control package that we will try it later.

Type `conda update git`

What is the output?

Answer:

Type `conda install git`

What is the output?

Answer:

Check the version of the Git package by using `git -version` command

What is the output?

Answer:

EXERCISE 5: PLAY WITH THE JUPYTER NOTEBOOK

Jupyter notebook is a set of tools for interactively developing and presenting Python programming. It makes a working with Python and data easier for scientists. It is a web-based application suitable for the process of solving a computational problem and presenting the results along with code, explanatory text and background. The Jupyter notebook combines two components:

1. The web application: a browser-based which combine text, mathematics, computations and their rich media output.
2. Notebook documents: a representation of all content visible in the web application. The Jupyter notebook file extension is .ipynb

Let's start, at the Anaconda navigator, click "Jupyter Notebook" (not Jupyterlab). You will see the Jupyter notebook on your browser with the list of files.

What is the URL of the Jupyter notebook that shows on your browser?

Answer:

Then, click "New", and select "Python3". You will see the file called "Untitled". The gray rectangle box with `in[]` is called "cell" where we can input both Python code and text using Mark Down tag.

Let's start with the coding, in the "code" cell type, you can start typing in Python code directly. Executing code in this cell can be done by either clicking on the run cell button or hitting Shift + Return keys.

Type `print('Hello World')` in the first cell and run. The output becomes visible right underneath the cell.

What is the output?

Answer:

Next, we will test how to create a figure on the Jupyter notebook. Type the following code in the second cell and run.

```
import numpy as np
import matplotlib.pyplot as plt
np.random.seed(999)
N = 50
x = np.random.rand(N)
y = np.random.rand(N)
colors = np.random.rand(N)
area = (30 * np.random.rand(N))**2
plt.scatter(x, y, s=area, c=colors, alpha=0.5)
plt.show()
```

What is the output?

Answer:

You have to modify the figure by adding 50 more circles. Please, try to edit the code above. What is your new code that add 50 more circles?

Next, the Jupyter notebook also supports displaying table objects. Type the following code in the next cell and run.

```
import pandas as pd
df = pd.read_csv("https://raw.githubusercontent.com/
plotly/datasets/master/school_earnings.csv")
```

```
df
```

How many data items contain in this dataset?

Answer:

In the next cell, type `print(df)` and run. What is the difference between `df` from the previous command and `print(df)`, and why it is different?

Answer:

milestone 2) You have now completed the Milestone 2. You should get sign off by your lab assistant.

EXERCISE 6: MARKDOWN AND CODE IN JUPYTER NOTEBOOK

You can change the cell type from “Code” to “Markdown” to include explanatory text in your notebook. To change the type, you can use the dropdown input control.

Change the next cell to “Markdown” and search the Internet for the Markdown tag. [Markdown in Jupyter Notebook](#)

Use Markdown to generate the following output on the current notebook.

Mahidol University

Faculty of Information and Communication Technology

We are studying ITCS159

Software Lab for Basic Scientific Problem Solving

Markdown can create lists, for example

- first
- second
- third

Markdown can also create inline code, for example

```
Sayyy, Say!!!, Say Ouu, Say Ahh, Say ICT
```

The Jupyter notebook also supports latex’s mathematical formula. Type the following code in the next cell, and run.

```
%\latex
\begin{align}
```

```

B'&=-\nabla \times E,\\
E'&=\nabla \times B - 4\pi j, \\
x &= \frac{-b\pm\sqrt{b^2-4ac}}{2a}
\end{align}

```

What is the output?

Answer:

Given the definite integral is the limit of the Riemann sum as follow

$$\int_a^b f(x) dx = \lim_{n \rightarrow +\infty} \sum_{i=1}^n \Delta x \cdot f(x)$$

Learn some Latex code in Jupyter Notebook [follow the guide of Latex](#)

Create derivative of equation in Jupyter Notebook.

Answer Code:

milestone 3) You have now completed the Milestone 3. You should get sign off by your lab assistant.