How to install python

July 22, 2020

Introduction:

What are packages?

If you are familiar with programming, packages are like libraries. They are add-ons that allow you to use python to perform advanced functions without having to code everything from scratch. The most common two are Pandas and NumPy. The former is for data management (like excel), and the latter is for math calculations.

Installing python:

What is Anaconda?

Without getting too specific, anaconda is just a compiler with a few default packages installed. Realistically it is possible to code everything in a .txt file, and when you're done change the extension to .py and run in the command line, but that is really tedious. Anaconda also provides you with several ways to write code such as Spyder and Jupyter.

How to install python (Anaconda):

1) Go to the Anaconda website: https://www.anaconda.com/. It should look something like this:



2) Click on "Get Started" \rightarrow "Install Anaconda Individual Edition" \rightarrow "Download". You should get to the following menu:

	Anaconda Installer	S
Windows 📲	MacOS 🗉	Linux 🛆
Python 3.7	Python 3.7	Python 3.7
64-Bit Graphical Installer (466 MB)	64-Bit Graphical Installer (442 MB)	64-Bit (x86) Installer (522 MB)
32-Bit Graphical Installer (423 MB)	64-Bit Command Line Installer (430 MB)	64-Bit (Power8 and Power9) Installer (276 MB)
Python 2.7	Python 2.7	
64-Bit Graphical Installer (413 MB)	64-Bit Graphical Installer (637 MB)	Python 2.7
32-Bit Graphical Installer (356 MB)	64-Bit Command Line Installer (409 MB)	64-Bit (x86) Installer (477 MB)
		64-Bit (Power8 and Power9) Installer (295 MB)

3) Choose your Operating System, and download the graphical installer under for your respective Operating System

** Note if you have windows: it comes in two versions, one for 64-bit and one for 32-bit. To check which version of windows you have, follow the directions in this link: https://support.microsoft.com/en-us/help/ 13443/windows-which-version-am-i-running

4) The file is very big (466Mb), wait for it to download and run the ".exe" file

a) Click "Next"



b) Agree to the EULA

	License Agre	ement			
ANACONDA.	Please review 2020.02 (64-	the license terms bit).	before installing A	Anaconda3	
Press Page Down to see th	ne rest of the agr	reement.			
					^
End User License Agreeme	ent - Anaconda I	ndividual Edition			
Copyright 2015-2020, And	aconda, Inc.	licenser			
Air fights reserved under t	TIE S-Clause DSD	LICENSE.			
This End User License Agr and Anaconda, Inc. ("Ana	eement (the "Ag conda") and gov	reement") is a legal	agreement betw naconda Individu	een you al Edition	
(which was formerly know	n as Anaconda D	istribution).			~
If you accept the terms of agreement to install Anaco	the agreement, nda3 2020.02 (6	click I Agree to cor 54-bit).	tinue. You must a	accept the	
aconda Inc					
aconoa, me					

c) Choose whether to install it for "All users" or "Just me"

ANACONDA.	Select Installation Type Please select the type of ins Anaconda3 2020.02 (64-bit)	tallation you would	d like to perfo	orm for
Install for:				
 Just Me (recommended) 	i -			
O All Users (requires admi	n privileges)			

d) Select a folder to install Anaconda (There must be at least 3.0Gb of space free in the folder)

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.	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\Your_name\anaconda3 Browse Space required: 3.0GB	Setup will instal 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\Your_name\anaconda3 Space required: 3.0GB Space available: 116.3GB	ANACONDA.	Choose Install Location Choose the folder in which to install A	naconda3 2	020.02 (64	1-bit).
Destination Folder	Destination Folder C:\Users\Your_name\anaconda3 Browse Space required: 3.0GB	Destination Folder C:\Users\Your_name\anaconda3 Browse Space required: 3.0GB Space available: 116.3GB	Setup will install Anaconda folder, click Browse and se	3 2020.02 (64-bit) in the following folder. ect another folder. Click Next to continue	. To install in e.	a differen	it
C: Jusers (rour_name variacoridas	Space required: 3.0GB	Space required: 3.0GB Space available: 116.3GB					

e) Click "Install"



The installation process is finished at this point. Check that Anaconda can be searched from windows search.

Run your first piece of code:

1) Run Anaconda Navigator

In Windows: it appears as "Anaconda Navigator (Anaconda)" in the search menu

In Mac OS: it appears as "Anaconda" when searched

2) Launch Jupyter: Click on the launch button below Jupyter in the menu



**Note: The Anaconda platform comes with other compilers such as Spyder. For processing data, usually Jupyter is used because of a few beneficial characteristics (like cell based statement execution).

3) You will arrive at a page like this in your web browser:

💭 Jupyter	Quit Logout
Files Running Clusters	
Select items to perform actions on them.	Upload New • 3
	Name 🔶 Last Modified File size
C 1. Finance	10 hours ago
2. Finance Web scraping	a month ago

You won't have the same folders, but it is highly suggested that you create a folder. To do so, hit the "New" drop-down menu near the top right. Then, click on folder. A new folder named "Untitled Folder" will appear near the bottom of the page. To rename it, check the checkbox to the left of it, and click rename near the top left of the page.

4) Click on "New" \rightarrow "Python 3". You will get to a page like this:

C Jupyter Tutorial Last Checkpoint a few seconds ago (unsaved changes)	4	Logout
File Edit View Insert Cell Kernel Widgets Help	Trusted	Python 3 O
E + » (2 E +) Run C > Code v		
In []:		

5) Edit the first rectangle and type the following (no double quotes): "print('Hello World')". Then, press shift + enter at the same time. The result should be:

[1]:	<pre>print('Hello World')</pre>
	Hello World
[]:	

Installing a Package, in this case WRDS:

1) Open the anaconda prompt:

On Windows: search anaconda prompt in the start menu



On Mac OS: just use the terminal

2) Install the WRDS package by using the command: "pip install WRDS". Then hit "Enter". You should get something similar to:

Anaconda Prompt (Anaconda)	-		×
(base) C:\Users\William>pip install WRDS			
Using cached https://files.pythonhosted.org/packages/1c/35/9d86097c36a8528a1f5bbc3815c161c2dfce7fffc28 Ø/wrds-3.0.8-py3-none-any.whl	e10cf	d7c9726	c1ac
<pre>Kequirement already satisfied: sqlalchemy in d:\anaconda\lib\site-packages (from WKDS) (1.3.9) Requirement already satisfied: mock in d:\anaconda\lib\site-packages (from WKDS) (3.6.5) Requirement already satisfied: psycopg2-binary in d:\anaconda\lib\site-packages (from WKDS) (3.6.5) Requirement already satisfied: pandas in d:\anaconda\lib\site-packages (from WKDS) (0.25.1) Requirement already satisfied: numpy=1.13.3 in d:\anaconda\lib\site-packages (from mock->WKDS) (1.12.0) Requirement already satisfied: numpy=1.13.3 in d:\anaconda\lib\site-packages (from moda>-WKDS) (1.12.0) Requirement already satisfied: pytz>=2017.2 in d:\anaconda\lib\site-packages (from pandas->WRDS) (2.2019.3 Requirement already satisfied: python-dateutil>=2.6.1 in d:\anaconda\lib\site-packages (from pandas->WRDS) Installing collected packages: WRDS SuccessFully installed WRDS-3.0.8</pre>	5)) S) (2	.8.1)	
(base) C:\Users\William>			

**Note: Anaconda supports conda install as well as pip install, you can use either just make sure that the package you want to install is supported by conda or pip. If you need to install any other package, just replace "WRDS" with the name of the package.

Get data from WRDS:

1) Open up an anaconda python 3 project as we did in the previous section. Copy the code from below in the grey box and replace YourUsername with your WRDS username. It will prompt you for a password on the first time, but it is not shown here since my WRDS has already been set up. The password is saved for all later logins.

```
[1]: import wrds
db = wrds.Connection(wrds_username='YourUsername')
db.raw_sql('SELECT date,dji FROM djones.djdaily')
```

Loading library list... Done

[1]:		date	dji
	0	1896-05-26	40.94
	1	1896-05-27	40.58
	2	1896-05-28	40.20
	3	1896-05-29	40.63
	4	1896-06-01	40.60
	28068	2008-06-13	12307.35
	28069	2008-06-16	12269.08
	28070	2008-06-17	12160.30
	28071	2008-06-18	12029.06
	28072	2008-06-19	12063.09
	[28073	rows x 2 co	lumns]

[2]: db.close()

[]:

Let's go over what the code does:

import wrds: imports the wrds package so you can use it

db = wrds.Connection(wrds_username='YourUsername'): creates a db object which allows you to connect to the wrds database

db.raw_sql('SELECT date,dji FROM djones.djdaily'): gets the data from wrds. The line in the quotes is in SQL, which is a databasing language that allows you to get information from SQL databases. A beginners tutorial can be found at: https://www.w3schools.com/sql/

db.close(): closes the connection with the database

**Note: If you mess up, WRDS has an error with their python API where sometimes it will not allow you to attempt to login again. To fix this error, close the tab and go to the main Jupyter tab. Then, click on the box next to your project and hit shutdown on the top left. Then restart your project.

Common functions in WRDS:

Libraries:

The WRDS database is categorized into libraries based on the data source/ data provider. To get a list of all libraries in the database, use the following code:

```
[18]: pd.set_option('display.max_rows', None)
    db.list_libraries()
```

Let's go over what the code does:

pd.set_option('display.max_rows', None): allows pandas to display all rows of a list that is printed. This is not needed on all machines, but if your code only outputs the first 5 lines of a list or the last 5 lines of a list, this could be the solution.

db.list_libraries(): Lists all the libraries in the WRDS database

Tables:

Libraries contain tables that contain information. The next logical step after you figure out what library to use, it to find the tables that exist in that library. To do so, follow the following example (this is done with the "djones" library but you can change it to any library you would like:

```
[19]: db.list_tables(library="djones")
```

```
[19]: ['djdaily', 'djmonthly']
```

We need to know what is in each table to construct SQL statements, which we do with:

```
[21]: db.describe_table(library="djones", table="djdaily")
```

Approximately 28003 rows in djones.djdaily.

[21]:		name	nullable		type
	0	date	True		DATE
	1	djc	True	DOUBLE	PRECISION
	2	djct	True	DOUBLE	PRECISION
	3	dji	True	DOUBLE	PRECISION
	4	djit	True	DOUBLE	PRECISION
	5	djt	True	DOUBLE	PRECISION
	6	djtt	True	DOUBLE	PRECISION
	7	dju	True	DOUBLE	PRECISION
	8	djut	True	DOUBLE	PRECISION

Let's go over what the code does:

db.list_tables(library = "libraryName"): Lists all the tables in the library with the name libraryName

db.describe_tables(library = "libraryName", table = "tableName"): Lists all the features in the table with the name tableName in the library with the name libraryName

Getting actual data from the tables:

dii

There are two methods that could be used to obtain the tables from libraries. They do the same thing, but just have different ways of describing info to be gotten.

[15]:

		5
0	1896-05-26	40.94
1	1896-05-27	40.58
2	1896-05-28	40.20
3	1896-05-29	40.63
4	1896-06-01	40.60
5	1896-06-02	40.04
6	1896-06-03	39.77
7	1896-06-04	39.94
8	1896-06-05	40.32
9	1896-06-08	39.81

date

[16]:		date	dji
	0	1896-05-26	40.94
	1	1896-05-27	40.58
	2	1896-05-28	40.20
	3	1896-05-29	40.63
	4	1896-06-01	40.60
	5	1896-06-02	40.04
	6	1896-06-03	39.77
	7	1896-06-04	39.94
	8	1896-06-05	40.32
	9	1896-06-08	39.81

Let's go over what the code does:

db.get_table(): gets the table using library name, table name, and column title in array form. The obs parameter limits the number of search results.

db.raw_sql(): We have used this method before to get data from the database. It uses a SQL statement to query the database.