

# **Foundations of Data Science using Python**

## **Python and Libraries for Data Science Installation**

Why Python?

Python offers all the skillsets that are required for Machine Learning.

- Simple and Consistent
- Libraries and Frameworks
- Platform Independence
- Great Community Base

### **Simple and Consistent**

Python's simple syntax allows developers to write code that is reliable, concise and readable. The productivity is increased during the overall development of the process. Prototype can be built faster as the complex machine learning tasks and testing process can be done quickly.

### **Libraries and Frameworks**

Another reason to learn python for machine learning is the availability of amazing libraries and frameworks of python. It is nothing but the pre-written code that can be used to solve common programming problem. Python has rich bank of libraries for machine learning, some of them are Tensor-Flow, Keras, Scikit-learn, Numpy, Scipy, Pandas, Seaborn etc. Numpy and Scipy are specifically for scientific and advance computing respectively. Pandas is generally used for Data analysis and Seaborn is specifically for Data visualization.

### **Platform Independence**

It basically means one can freely shift from one machine to another without making changes to the actual code (on with minimal changes). It supports many platforms such as Windows, macOS, and Linux.

## Great Community Base

The python's community has grown across the globe, and especially in the world of machine learning and data science. There are active communities that contribute to the large exchange of information which involves solutions to the problems. For any problem you come across, chances are very high that someone out there has already gone through that same problem and solved it successfully. Hence you can find guidance and bits of advice at any level of doubt. You won't be the only one who went through it. Also, you may come to know some of the best results according to your specifications, all you need to do it to turn to the huge python community.

### Python Version 3.8.2 Installation Steps

1. Go to [www.python.org/downloads/](https://www.python.org/downloads/) for downloading the latest version of python



2. For python documentation, visit <https://www.python.org/doc/>
3. Select the file as per your machine.

A screenshot of the Python 3.8.2 release page. The browser's address bar shows 'python.org/downloads/release/python-382/'. The page title is 'Files'. Below it is a table with columns: Version, Operating System, Description, MD5 Sum, File Size, and GPG. The table lists various download options for different operating systems. The row for 'Windows x86 executable installer' is highlighted with a red box.

Version	Operating System	Description	MD5 Sum	File Size	GPG
<a href="#">Gzipped source tarball</a>	Source release		f9f3768f757e34b342dbc06b41cbc844	24007411	SIG
<a href="#">XZ compressed source tarball</a>	Source release		e9d6ebc92183a177b8e8a58cad5b8d67	17869888	SIG
<a href="#">macOS 64-bit installer</a>	Mac OS X	for OS X 10.9 and later	f12203128b5c639dc08e5a43a2812cc7	30023420	SIG
<a href="#">Windows help file</a>	Windows		7506675dcb9a1569b54e600ae66c9fb	8507261	SIG
<a href="#">Windows x86-64 embeddable zip file</a>	Windows	for AMD64/EM64T/x64	1a98565285491c0ea65450e78afe6f8d	8017771	SIG
<a href="#">Windows x86-64 executable installer</a>	Windows	for AMD64/EM64T/x64	b5df1cbb2bc152cd70c3da9151cb510b	27586384	SIG
<a href="#">Windows x86-64 web-based installer</a>	Windows	for AMD64/EM64T/x64	2586cdad1a363d1a8abb5fc102b2d418	1363760	SIG
<a href="#">Windows x86 embeddable zip file</a>	Windows		1b1f0fc5ee8601f160cfad5b560e3a7	7147713	SIG
<a href="#">Windows x86 executable installer</a>	Windows		6f0ba59c7db8a7bb0ee21682fe39748	26481424	SIG
<a href="#">Windows x86 web-based installer</a>	Windows		04d97979534f4bd33752c183fc4ce680	1325416	SIG

The name of the executable file is Python-3.8.2.exe

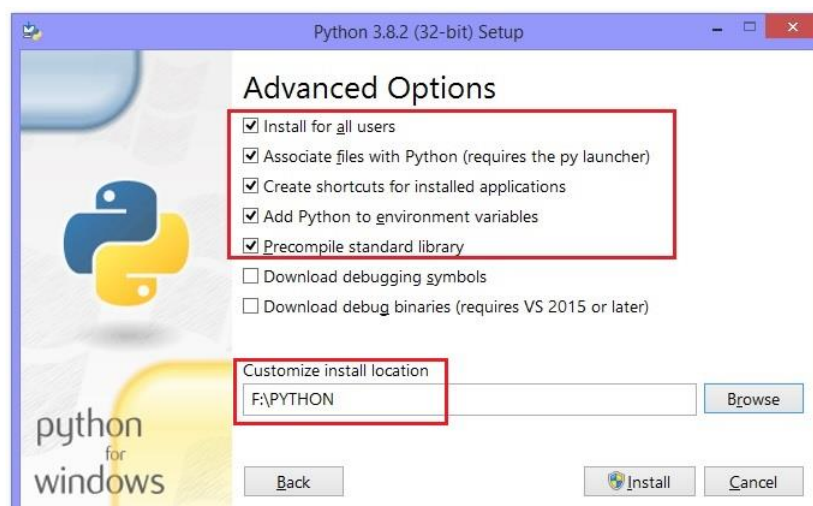
4. Launch the python installer and select the options of Custom installation for choosing the location and other features. Select the options as highlighted below:

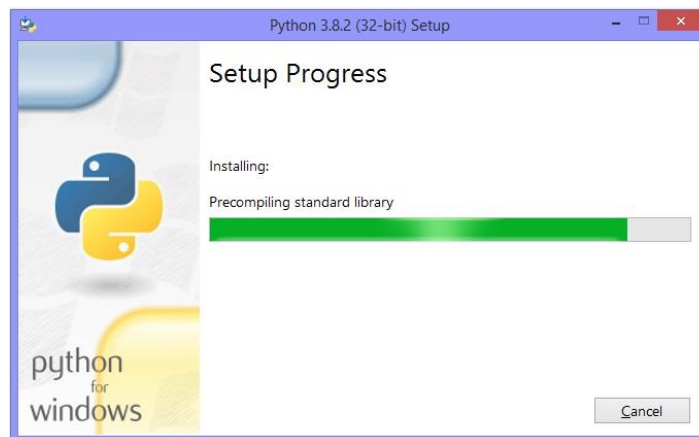
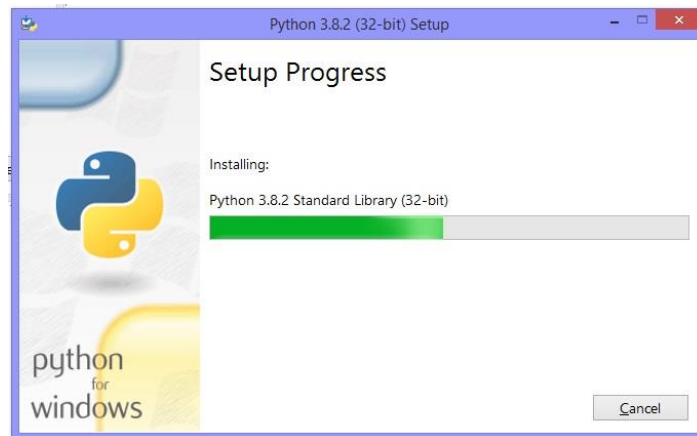


5. Select all the optional features and click 'Next'



6. Select the advanced options, for example, add python to environment variables, custom location etc., as highlighted below and click 'Install', the Setup successful window appears once the installation is completed.





7. From the command prompt, launch python with the commands 'py' or 'python'. If python is successfully installed, the python prompt '>>>' will appear on the command prompt.

```
C:\Windows\system32\cmd.exe

F:\>py
Python 3.8.2 (tags/v3.8.2:7b3ab59, Feb 25 2020, 22:45:29) [MSC v.1916 32 bit (Intel)] on win32
Type "help", "copyright", "credits" or "license" for more information.
>>> exit()

F:\>python
Python 3.8.2 (tags/v3.8.2:7b3ab59, Feb 25 2020, 22:45:29) [MSC v.1916 32 bit (Intel)] on win32
Type "help", "copyright", "credits" or "license" for more information.
>>> exit()

F:\>|
```

## Python Libraries Installation Steps

8. For installing any python package, from the python prompt '>>>', execute the command 'pip install <package name>'. For example, install Numpy by executing the command 'pip install numpy'

```
C:\Windows\system32\cmd.exe
F:\>pip install numpy
Collecting numpy
  Downloading https://files.pythonhosted.org/packages/52/2c/bf86d762ae6550dc8a7ab8381ba610bb69af6db619b3755f2b73052c6b9/numpy-1.18.4-cp38-wi
  | 10.8MB 312kB/s
Installing collected packages: numpy
Successfully installed numpy-1.18.4
WARNING: You are using pip version 19.2.3, however version 20.1 is available.
You should consider upgrading via the 'python -m pip install --upgrade pip' command.
```

9. In case of upgrading the pip, execute the command 'python -m pip install --upgrade'

```
F:\>python -m pip install --upgrade pip
Collecting pip
  Downloading https://files.pythonhosted.org/packages/54/2e/df11ea7e23e7e761d484ed3740285a34e38548cf2bad2bed3dd5768ec8b9/pip-20.1-py2.py3-none-any
  | 1.5MB 261kB/s
Installing collected packages: pip
Found existing installation: pip 19.2.3
Uninstalling pip-19.2.3:
  Successfully uninstalled pip-19.2.3
Successfully installed pip-20.1
```

10. Install all the essential packages like Scipy, Pandas, Matplotlib, Scikit-learn, Seaborn using pip install command

```
F:\>pip install scipy
Collecting scipy
  Downloading scipy-1.4.1-cp38-cp38-win32.whl (27.9 MB)
  | 27.9 MB 125 kB/s
Requirement already satisfied: numpy>=1.13.3 in f:\python\lib\site-packages (from scipy) (1.18.4)
Could not build wheels for numpy, since package 'wheel' is not installed.
Installing collected packages: scipy
Successfully installed scipy-1.4.1
```

```
F:\>pip install wheel
Collecting wheel
  Downloading wheel-0.34.2-py2.py3-none-any.whl (26 kB)
Installing collected packages: wheel
Successfully installed wheel-0.34.2
```

F:\>

```
C:\Windows\system32\cmd.exe
F:\>pip install pandas
Collecting pandas
  Downloading pandas-1.0.3-cp38-cp38-win32.whl (7.6 MB)
  | 7.6 MB 63 kB/s
Collecting python-dateutil>=2.6.1
  Downloading python_dateutil-2.8.1-py2.py3-none-any.whl (227 kB)
  | 227 kB 41 kB/s
Requirement already satisfied: numpy>=1.13.3 in f:\python\lib\site-packages (from pandas) (1.18.4)
Collecting pytz>=2017.2
  Downloading pytz-2020.1-py2.py3-none-any.whl (510 kB)
  | 510 kB 41 kB/s
Collecting six>=1.5
  Downloading six-1.14.0-py2.py3-none-any.whl (10 kB)
Installing collected packages: six, python-dateutil, pytz, pandas
Successfully installed pandas-1.0.3 python-dateutil-2.8.1 pytz-2020.1 six-1.14.0

F:\>
```

```
F:\>pip install matplotlib
Collecting matplotlib
  Downloading matplotlib-3.2.1-cp38-cp38-win32.whl (9.0 MB)
    |████████████████████████████████████████| 9.0 MB 154 kB/s
Collecting kiwisolver>=1.0.1
  Downloading kiwisolver-1.2.0-cp38-none-win32.whl (43 kB)
    |████████████████████████████████████████| 43 kB 128 kB/s
Collecting cycler>=0.10
  Downloading cycler-0.10.0-py2.py3-none-any.whl (6.5 kB)
Requirement already satisfied: python-dateutil>=2.1 in f:\python\lib\site-packages (from matplotlib) (2.8.1)
Requirement already satisfied: numpy>=1.11 in f:\python\lib\site-packages (from matplotlib) (1.18.4)
Collecting pyparsing!=2.0.4,!=2.1.2,!=2.1.6,>=2.0.1
  Downloading pyparsing-2.4.7-py2.py3-none-any.whl (67 kB)
    |████████████████████████████████████████| 67 kB 120 kB/s
Requirement already satisfied: six in f:\python\lib\site-packages (from cycler>=0.10->matplotlib) (1.14.0)
Installing collected packages: kiwisolver, cycler, pyparsing, matplotlib
Successfully installed cycler-0.10.0 kiwisolver-1.2.0 matplotlib-3.2.1 pyparsing-2.4.7
```

F:\>█

```
F:\>pip install scikit-learn
Collecting scikit-learn
  Downloading scikit_learn-0.22.2.post1-cp38-cp38-win32.whl (5.7 MB)
    |████████████████████████████████████████| 5.7 MB 42 kB/s
Requirement already satisfied: scipy>=0.17.0 in f:\python\lib\site-packages (from scikit-learn) (1.4.1)
Requirement already satisfied: numpy>=1.11.0 in f:\python\lib\site-packages (from scikit-learn) (1.18.4)
Collecting joblib>=0.11
  Downloading joblib-0.14.1-py2.py3-none-any.whl (294 kB)
    |████████████████████████████████████████| 294 kB 94 kB/s
Installing collected packages: joblib, scikit-learn
Successfully installed joblib-0.14.1 scikit-learn-0.22.2.post1
```

F:\>█

```
F:\>pip install seaborn
Collecting seaborn
  Downloading seaborn-0.10.1-py3-none-any.whl (215 kB)
    |████████████████████████████████████████| 215 kB 32 kB/s
Requirement already satisfied: numpy>=1.13.3 in f:\python\lib\site-packages (from seaborn) (1.18.4)
Requirement already satisfied: pandas>=0.22.0 in f:\python\lib\site-packages (from seaborn) (1.0.3)
Requirement already satisfied: scipy>=1.0.1 in f:\python\lib\site-packages (from seaborn) (1.4.1)
Requirement already satisfied: matplotlib>=2.1.2 in f:\python\lib\site-packages (from seaborn) (3.2.1)
Requirement already satisfied: python-dateutil>=2.6.1 in f:\python\lib\site-packages (from pandas>=0.22.0->seaborn) (2.8.1)
Requirement already satisfied: pytz>=2017.2 in f:\python\lib\site-packages (from pandas>=0.22.0->seaborn) (2020.1)
Requirement already satisfied: kiwisolver>=1.0.1 in f:\python\lib\site-packages (from matplotlib>=2.1.2->seaborn) (1.2.0)
Requirement already satisfied: pyparsing!=2.0.4,!=2.1.2,!=2.1.6,>=2.0.1 in f:\python\lib\site-packages (from matplotlib>=2.1.2->seaborn) (2.4.7)
Requirement already satisfied: cycler>=0.10 in f:\python\lib\site-packages (from matplotlib>=2.1.2->seaborn) (0.10.0)
Requirement already satisfied: six>=1.5 in f:\python\lib\site-packages (from python-dateutil>=2.6.1->pandas>=0.22.0->seaborn) (1.14.0)
Installing collected packages: seaborn
Successfully installed seaborn-0.10.1
F:\>
```