

Python for Data Visualization

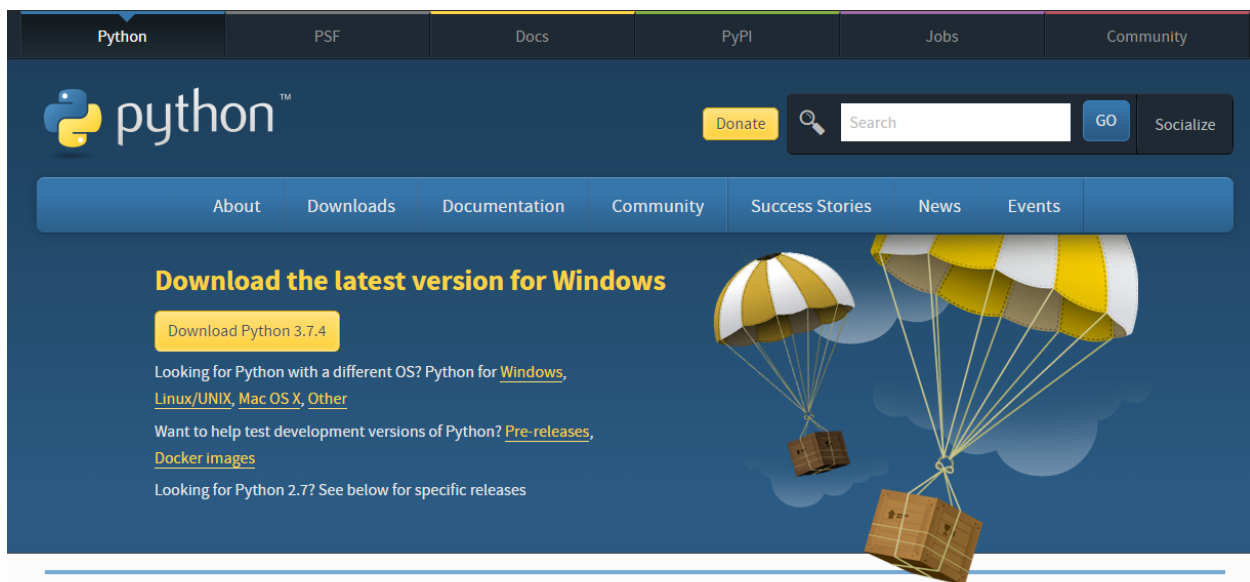
Preparing the Tools

Python menjadi salah satu bahasa pemrograman yang sangat digemari untuk proses data science termasuk data visualization dikarenakan kemudahan penggunaan, ketersediaan library yang lengkap dan komunitas yang baik.

Python

Python bersifat *open source*, Anda dapat dengan mudah mengunduhnya di

<https://www.python.org/downloads/>



Looking for a specific release?

Python releases by version number:

Release version	Release date	Click for more	
Python 3.7.4	July 8, 2019	Download	Release Notes
Python 3.6.9	July 2, 2019	Download	Release Notes
Python 3.7.3	March 25, 2019	Download	Release Notes
Python 3.4.10	March 18, 2019	Download	Release Notes
Python 3.5.7	March 18, 2019	Download	Release Notes
Python 2.7.16	March 4, 2019	Download	Release Notes

Download : Pada kesempatan ini kita akan menggunakan python versi **3.6.8**. Pada bagian Release Version, scroll ke bawah, hingga Anda menemukan Python 3.6.8., kemudian klik tautan download.

Files

Version	Operating System	Description	MD5 Sum	File Size	GPG
Gzipped source tarball	Source release		48f393a04c2e66c77bfc114e589ec630	23010188	SIG
XZ compressed source tarball	Source release		51aac91bdf8be95ec0a62d174890821a	17212420	SIG
macOS 64-bit/32-bit installer	Mac OS X	for Mac OS X 10.6 and later	eb1a23d762946329c2aa3448d256d421	33258809	SIG
macOS 64-bit installer	Mac OS X	for OS X 10.9 and later	786c4d9183c754f58751d52f509bc971	27073838	SIG
Windows help file	Windows		0b04278f5bdb8ee85ae5ae66af0430b2	7868305	SIG
Windows x86-64 embeddable zip file	Windows	for AMD64/EM64T/x64	73df7cb2f1500ff36d7dbeac3968711	7276004	SIG
Windows x86-64 executable installer	Windows	for AMD64/EM64T/x64	72f37686b7ab240ef70fdb931bdf3cb5	31830944	SIG
Windows x86-64 web-based installer	Windows	for AMD64/EM64T/x64	39dde5f535c16d642e84fc7a69f43e05	1331744	SIG
Windows x86 embeddable zip file	Windows		60470b4cceb5a52094121d43cd3f6ce3a	6560373	SIG
Windows x86 executable installer	Windows		9c7b1ebdd3a8df0eebfda2f107f1742c	30807656	SIG
Windows x86 web-based installer	Windows		80de96338691698e10a935ecd0bdaacb	1296064	SIG

Pilih installer sesuai operating sistem Anda, perhatikan tipe bit installer.

Install :

1. Pada saat instalasi beri tanda centang pada **Add Python 3.6 to PATH**. Kemudian klik **Install Now**



2. Saat instalasi berakhir, **jika** muncul pesan **Disable path length limit**, pastikan anda memilihnya dengan cara klik.



Buat Environment (Windows)

Untuk menghindari konflik versi yang berbeda pada setiap library, Anda memerlukan membuat environment Python. Langkah-langkah membuat environment dalam Python adalah sebagai berikut :

1. Buat folder baru, misalnya dalam contoh kali ini adalah di **E:\DataViz**
2. Buka command prompt, masuk ke direktori folder yang baru Anda buat. Misalnya jika di direktori E, ketik **E:** dan tekan Enter.

```
C:\WINDOWS\system32\cmd.exe
Microsoft Windows [Version 10.0.18362.356]
(c) 2019 Microsoft Corporation. All rights reserved.

C:\Users\Satria>E:

E:\>
```

3. Setelah masuk ke direktori, ketik **python -m venv E:\DataViz**

```
C:\WINDOWS\system32\cmd.exe
Microsoft Windows [Version 10.0.18362.356]
(c) 2019 Microsoft Corporation. All rights reserved.

C:\Users\Satria>E:

E:\>python -m venv E:\DataMining

E:\>
```

4. Setelah installasi berhasil, maka folder yang Anda buat akan berisi data seperti di bawah ini :

> This PC > Data (E:) > DataMining				
Name		Date modified	Type	Size
✦	Include	30/09/2019 13:10	File folder	
	Lib	30/09/2019 13:10	File folder	
	Scripts	30/09/2019 13:10	File folder	
	pyvenv.cfg	30/09/2019 13:10	CFG File	1 KB

5. Selamat! Anda telah berhasil membuat environment di Python

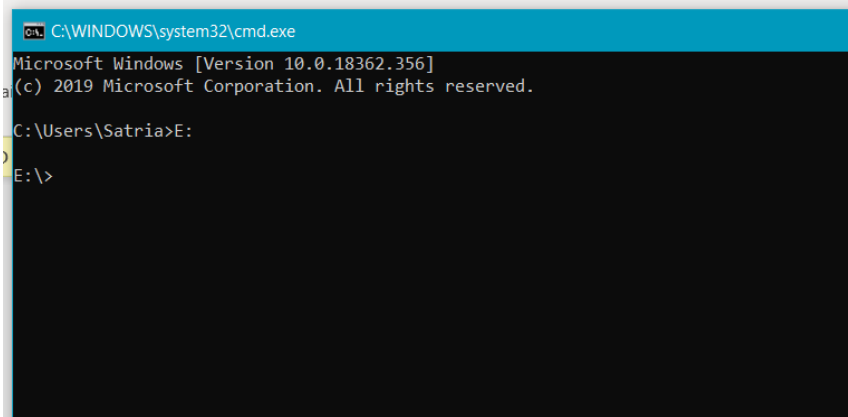
Library di Python

Alasan utama Python menjadi tool yang paling sering digunakan oleh data scientist adalah banyaknya library *machine learning/artificial intelligence* terkini. Anda dapat dengan mudah menggunakannya dengan **Python Package Indexing (PIP)**.

Beberapa library yang akan kita gunakan adalah Numpy (pemrosesan matriks, aljabar linier), Pandas (pemrosesan dataset), Matplotlib (visualisasi data), Seaborn (wrapping dari Matplotlib, visualisasi data dengan kode lebih ringkas), Scikit-learn (kumpulan implementasi algoritma data mining) dan Jupyter Notebook (web application untuk python code editor).

Adapun langkah-langkahnya adalah sebagai berikut :

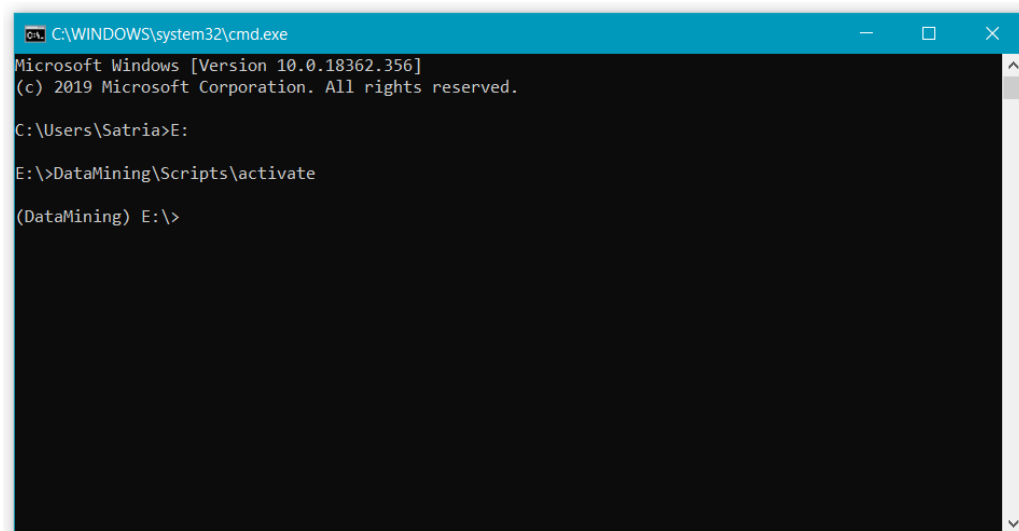
1. Buka command prompt, pindah ke direktori environment Anda.



```
C:\WINDOWS\system32\cmd.exe
Microsoft Windows [Version 10.0.18362.356]
(c) 2019 Microsoft Corporation. All rights reserved.

C:\Users\Satria>E:
E:\>
```

2. Kemudian aktifkan environment yang Anda buat dengan cara mengetikkan DataMining\Scripts\activate dan tekan Enter



```
C:\WINDOWS\system32\cmd.exe
Microsoft Windows [Version 10.0.18362.356]
(c) 2019 Microsoft Corporation. All rights reserved.

C:\Users\Satria>E:
E:\>DataMining\Scripts\activate
(DataMining) E:\>
```

3. Jika command prompt telah berubah seperti gambar di atas, maka environment Anda telah aktif. Kemudian Anda bisa menginstall library yang dibutuhkan satu persatu. Untuk

instalasi library, Anda harus terkoneksi ke internet. Ketik command di bawah ini satu persatu:

pip install numpy

pip install matplotlib

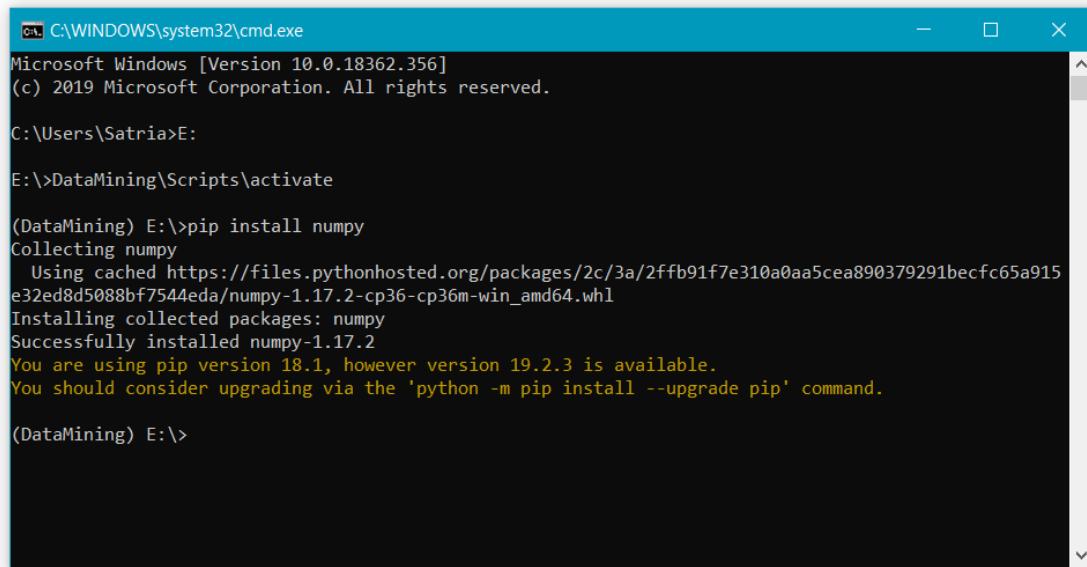
pip install pandas

pip install seaborn

pip install scikit-learn

pip install jupyterlab

Contoh instalasi package yang berhasil dapat dilihat pada gambar di bawah



```
C:\WINDOWS\system32\cmd.exe
Microsoft Windows [Version 10.0.18362.356]
(c) 2019 Microsoft Corporation. All rights reserved.

C:\Users\Satria>E:

E:\>DataMining\Scripts\activate

(DataMining) E:\>pip install numpy
Collecting numpy
  Using cached https://files.pythonhosted.org/packages/2c/3a/2ffb91f7e310a0aa5cea890379291becfc65a915e32ed8d5088bf7544eda/numpy-1.17.2-cp36-cp36m-win_amd64.whl
Installing collected packages: numpy
Successfully installed numpy-1.17.2
You are using pip version 18.1, however version 19.2.3 is available.
You should consider upgrading via the 'python -m pip install --upgrade pip' command.

(DataMining) E:\>
```

4. Jika Anda telah selesai menggunakan environment, ketik **deactivate** pada command prompt dan tekan Enter