

Pengolahan Data kasus Covid 19 di Jakarta dengan menggunakan numpy dan matplotlib pada Python

Akhmad Sofwan

sofwan@sofwan.net

<http://www.sofwan.net>

Lisensi Dokumen:

Copyright © 2003-2007 IlmuKomputer.Com

Seluruh dokumen di IlmuKomputer.Com dapat digunakan, dimodifikasi dan disebarkan secara bebas untuk tujuan bukan komersial (nonprofit), dengan syarat tidak menghapus atau merubah atribut penulis dan pernyataan copyright yang disertakan dalam setiap dokumen. Tidak diperbolehkan melakukan penulisan ulang, kecuali mendapatkan ijin terlebih dahulu dari IlmuKomputer.Com.

Pendahuluan

Saat ini Dunia sedang menghadapi pandemik Covid 19 yang disebabkan oleh virus Sars cov-2, termasuk Indonesia. DKI Jakarta merupakan penyumbang kasus Covid 19 terbesar di Indonesia yang penambahan kasus harian yang terdiri dari elemen kasus baru, pasien yang meninggal dan pasien yang sembuh di data setiap hari dan ditampilkan kepada publik oleh Pemda DKI Jakarta lewat situs Korona Jakarta <https://corona.jakarta.go.id> [1] dalam berbagai bentuk tampilan data, termasuk grafik garis.

Saya mencoba untuk memvisualisasikan data kasus harian tersebut dengan menggunakan library numpy dan matplotlib di Python dalam bentuk grafik garis dengan mengambil data dari situs Korona Jakarta. Visualisasi dengan cara ini diperlukan langkah-langkah sebagai berikut:

1. Membuat dan membaca data CSV kasus baru harian, pasien meninggal dan pasien sembuh.
2. Menampilkan data dengan grafik garis.

Maksud dari menampilkan data kasus harian Covid 19 DKI Jakarta ini, yakni :

1. Memiliki data CSV data kasus harian yang diambil dari situs Covid 19 DKI Jakarta [1]
2. Menampilkan dan mengolah serta menganalisa data CSV tersebut.

Dengan data harian yang pada saat tulisan ini dibuat, data dari tanggal 2 Maret – 2 Mei 2020 yang diambil dari situs Corona Jakarta [1], kita dapat melakukan perhitungan statistika, seperti mengetahui jumlah kasus per hari, standard deviasi, nilai minimum dan maksimum dari masing-masing group data, yakni jumlah kasus baru, jumlah pasien meninggal dan jumlah pasien sembuh. Disamping itu kita dapat menampilkan data berupa grafik garis. Selain itu, data dapat juga diolah dan dianalisa lebih lanjut untuk tujuan tertentu seperti untuk prediksi dengan teknik tertentu pada Python.

Pada artikel ini, saya akan memaparkan petunjuk teknis cara pengolahan data dan visualisasi kasus Covid 19 harian di DKI Jakarta dengan menggunakan numpy dan matplotlib di Python, termasuk instalasi Python, matplotlib, numpy dan IDE Spider.

Pada artikel ini, penulisan source code dan perintah python dan instalasi menggunakan warna font biru.

Pengenalan Python, numpy dan matplotlib

Python adalah sebuah bahasa pemrograman *General Purpose* atau multi fungsi yang dapat digunakan dalam pemrograman desktop ,web dan microcontroller. Python sangat banyak digunakan di beberapa domain, yaitu pengolahan data dan keuangan. Python saat ini tersedia untuk Python versi 2.x dan Python versi 3.x. Di dalam situs www.python.org untuk Python versi 2.x tersedia Python 2.7 dan untuk Python versi 3.x, tersedia Python 3.5, Python 3.6, Python 3.7 dan yang terbaru saat artikel ini dibuat, Python 3.8 [2]

Numpy adalah paket fundamental untuk komputasi saintifik dengan Python, yang memiliki beberapa fitur, diantaranya adalah :

- Objek array N-dimensi
- Teknik untuk mengintegrasikan dengan kode C/C++ dan Fortran
- Kemampuan untuk kalkulasi Aljabar Linear, Fourier Transform dan angka random. [3]

Dalam artikel ini, Numpy digunakan untuk membaca data CSV dan melakukan summary data. Summary berisi jumlah data, mean, standard deviasi, nilai minimum, Quartile 25%, Rata-rata, Quartile 75% dan nilai maksimum dari masing-masing group data.

Matplotlib adalah salah satu library di Python yang komprehensif untuk membuat visualisasi yang statis, animatif dan interaktif [4]. Dalam artikel ini, matplotlib digunakan untuk membuat grafik garis Data harian Covid 19 di Jakarta.

Untuk perangkat menulis source code, kita gunakan IDE Spyder (www.spyder-ide.org).

Instalasi Python, numpy dan matplotlib

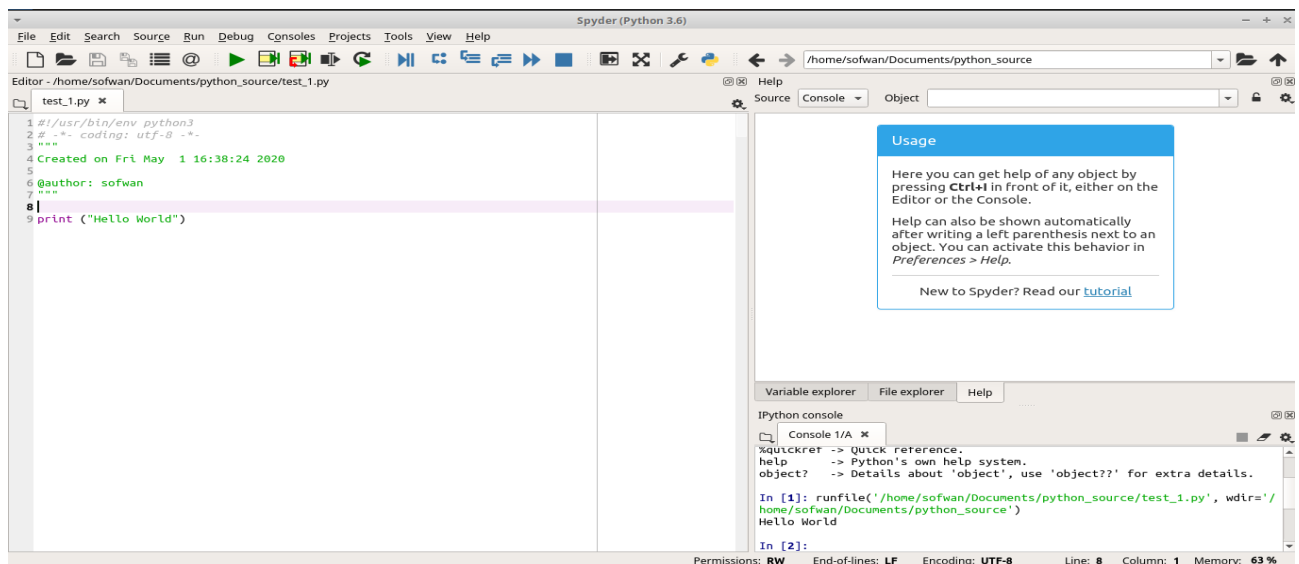
Kita dapat menggunakan paket anaconda untuk menginstall Python, library-library Python penting termasuk numpy dan matplotlib dan IDE spyder serta aplikasi lainnya. Anda dapat mendownload file installer anaconda pada halaman

<https://www.anaconda.com/products/individual> dan menjalankannya. Saya menyarankan, anda menginstall aplikasi-aplikasi yang dibutuhkan dengan menggunakan cara ini, karena lebih mudah.

Selain menggunakan paket Anconda, kita juga dapat menginstall Python, matplotlib, numpy dan IDE Spider secara terpisah, yang masing-masing file instalasinya dapat anda download pada situsnya masing-masing.

Menguji Python dengan IDE spyder

Setelah kita berhasil menginstall anaconda atau menginstall satu persatu Python, numpy, matplotlib dan Spyder, maka kita sudah memiliki perangkat untuk mengolah data dan memvisualisasikan data Covid 19. Anda dapat membuka IDE spyder dan mempelajari tampilannya serta melakukan uji source code. Anda dapat mengetikkan perintah : `print ("Hello World")`. Untuk menjalankannya, tekan F5 atau pilih menu Run → Run.



Gambar 1 : Tampilan IDE Spyder

Menjalankan matplotlib

Dengan melakukan instalasi menggunakan anaconda, anda sudah memiliki matplotlib, sehingga anda tidak perlu menginstall matplotlib. Jika anda belum memiliki matplotlib karena menginstall masing-masing aplikasi, anda perlu melakukan instalasi matplotlib dengan perintah [2] :

```
python -m pip install -U pip  
python -m pip install -U matplotlib
```

Untuk mengetahui apakah anda sudah memiliki matplotlib atau belum, anda dapat menuliskan perintah berikut ini pada spyder: `import matplotlib`.
Jika dirunning tidak ditemukan pesan error, maka system anda sudah memiliki matplotlib.

Menjalankan numpy

Dengan melakukan instalasi dengan menggunakan anaconda, anda sudah memiliki numpy, sehingga anda tidak perlu menginstall numpy. Jika anda belum memiliki numpy, anda dapat melakukan instalasi numpy dengan mengetikkan perintah : `pip install numpy`

Untuk mengetahui apakah anda sudah memiliki numpy di dalam sistem anda atau belum, anda dapat mengetikkan di spyder perintah : `import numpy`
Jika setelah di running, tidak ada error, berarti anda sudah memiliki numpy.

Membaca data CSV dengan numpy

Kita akan membaca data korona Jakarta yang berbentuk csv. Data csv dan berikut keterangan data serta source code dapat anda peroleh di github sofwanbl https://github.com/sofwanbl/matplotlib_korona .

Pada data korona di file csv yang anda peroleh, terdiri dari 4 kolom, dengan keterangannya sebagai berikut :

- kolom 1 : tanggal kejadian, dengan susunan tanggal, bulan dan tahun
- kolom 2 : Jumlah kasus baru
- kolom 3 : Jumlah meninggal
- kolom 4 : Jumlah sembuh

Contoh : 2-03-20, 0,1,0

Data tersebut akan kita baca, tampilkan dan tampilkan kesimpulan statistiknya.

Source code nya adalah :

```
1. import pandas as pd
2. import matplotlib
3. import matplotlib.pyplot as plt
4. import csv
5. import numpy as np
6.
7. data=pd.read_csv("data_korona_dki_x.txt",sep=",")
8. print(data)
9. print("Describe :",data.describe())
```

Output :

```

    2-03-20    0    1    0.0
0  3-03-20    3    1    0
1  4-03-20    0    0    0
2  5-03-20    4    2    0
3  6-03-20    0    0    0
4  7-03-20    0    0    0
..  ...    ..  ...
56 28-04-20  118   4    3
57 29-04-20   83   2   71
58 30-04-20  105   0    0
59 1-05-20   145  12   15
60 2-05-20   72   7  145

[61 rows x 4 columns]
Describe :
count    61.000000    61.000000    61.000000
mean     71.393443    6.540984     9.377049
std      54.555256    7.817872    22.951517
min       0.000000    0.000000     0.000000
25%      26.000000    2.000000     0.000000
50%      75.000000    4.000000     1.000000
75%     105.000000    8.000000     5.000000
max     223.000000   35.000000   145.000000
```

Gambar 2 : Output pengolahan data dengan numpy

Output berupa tampilan data 5 besar bagian atas dan 5 besar bagian bawah serta kesimpulan statistik.

Keterangan Program :

```
1. import pandas as pd
2. import matplotlib
3. import matplotlib.pyplot as plt
4. import csv
5. import numpy as np
```

Keterangan :Load library pandas, numpy dan matplotlib serta meload modul csv.

```
7. data=pd.read_csv("data_korona_dki_x.txt",sep=",")
```

Keterangan : Membaca file csv dengan menggunakan metode read_csv.

Parameter sep="," digunakan, karena antar data dipisahkan dengan koma (,).

8. `print(data)`

Keterangan : Mencetak data. Data yang tercetak 5 besar paling atas dan 5 besar paling bawah.

9. `print("Describe :",data.describe())`

Keterangan : Mencetak kesimpulan data, yang menampilkan informasi berupa :

count : Jumlah baris

mean : Rata-rata

std : Standard Deviasi

min : Nilai Minimal

25% : Quartile 25%

50% : Quartile 50% atau median atau nilai tengah

75% : Quartile 75%

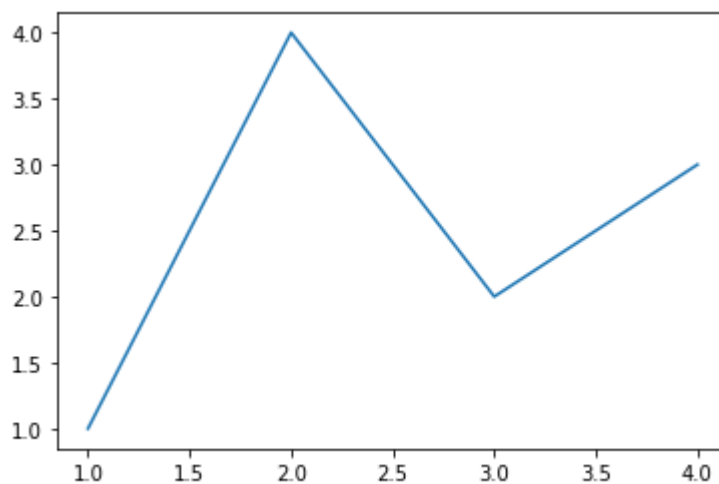
Max : Nilai maksimal

Membuat grafik garis dengan matplotlib

Untuk pembuatan grafik pertama kali, kita akan membuat grafik garis sederhana dahulu (untuk pemanasan dulu :-)) sebelum masuk membuat grafik garis kasus harian Covid 19 Jakarta. Grafik dibuat dengan data x dan y yang ditulis dalam source code. Source code pembuatan grafik sederhana diambil dari tutorial matplotlib [4]. Pada spyder, ketikkan perintah berikut (tanpa angka barisnya) :

1. `import matplotlib.pyplot as plt`
2. `fig,ax = plt.subplots()`
3. `ax.plot([1,2,3,4],[1,4,2,3])`

Setelah selesai penulisan kode, lalu run. Maka hasil yang didapat seperti ini:



Gambar 3 : Grafik garis sederhana dengan matplotlib

Grafik pada gambar 3 menunjukkan grafik garis dengan kordinat (1,1),(2,4),(3,2) dan (4,3)

Pengolahan data dan visualisasi data kasus harian Covid 19 di Jakarta

Setelah berhasil membaca data csv dengan menggunakan numpy pada bagian “Membaca data csv dengan numpy”, selanjutnya kita mencoba untuk menampilkan grafik dari data tersebut. Terdapat 3 kolom untuk ditampilkan, yaitu jumlah kasus, meninggal dan sembuh, sehingga kita akan memiliki 3 garis dalam 1 grafik. Kode akan kita tambahkan dari bagian “Membaca data CSV dengan numpy”.

```
10. x=[]
11. y=[]
12. z=[]
13. xx=[]
14.
15. with open("data_korona_dki_x.txt","r") as csvfile:
16.     plots=csv.reader(csvfile,delimiter=",")
17.     for row in plots:
18.         x.append(row[0])
19.         y.append(int(row[1]))
20.         z.append(int(row[2]))
21.         xx.append(int(row[3]))
22.
23. fig,ax=plt.subplots()
24.
25. ax.plot(x,y,marker="o",label="Kasus Baru")
26. ax.plot(x,z,marker="x",label="Meninggal")
27. ax.plot(x,xx,marker="+",label="Sembuh")
28. plt.title("Data Korona Jakarta")
29. plt.xlabel("Tanggal")
30. plt.ylabel("Jumlah")
31. plt.xticks(rotation=90)
32. legend=ax.legend(loc="upper left", shadow=True)
33.
34. for i,j in zip(x,y):
35.     ax.annotate('%s' %j,xy=(i,j),xytext=(4,0), textcoords='offset points')
36.
37. for a,b in zip(x,z):
38.     ax.annotate('%s' %b,xy=(a,b),xytext=(4,0), textcoords='offset points')
39.
40. for d,e in zip(x,xx):
41.     ax.annotate('%s' %e,xy=(d,e),xytext=(4,0), textcoords='offset points')
42.
43. print(plt.show())
```

Keterangan kode :

```
10.x=[]
11.y=[]
12.z=[]
13.xx=[]
```

Keterangan : Menyatakan variabel x,y,z dan xx sebagai list

```
15. with open("data_korona_dki_x.txt","r") as csvfile:
16.     plots=csv.reader(csvfile,delimiter=",")
17.     for row in plots:
18.         x.append(row[0])
19.         y.append(int(row[1]))
```

```
20. z.append(int(row[2]))  
21. xx.append(int(row[3]))
```

Keterangan :

Membuka file data_korona_dki_x.txt, membaca isi filenya dengan cara iterasi, lalu menambahkan data pada list berikut :

- x di kolom pertama (kolom tanggal)
- y di kolom kedua (kolom kasus)
- z di kolom kedua (kolom jumlah meninggal)
- xx di kolom ketiga (kolom jumlah sembuh)

```
23. fig,ax=plt.subplots()
```

Keterangan : Memungkinkan 1 kanvas atau gambar terdiri dari banyak grafik, dalam hal ini 3 garis grafik (kasus, meninggal dan sembuh).

```
25. ax.plot(x,y,marker="o",label="Kasus Baru")
```

```
26. ax.plot(x,z,marker="x",label="Meninggal")
```

```
27. ax.plot(x,xx,marker="+",label="Sembuh")
```

Keterangan : Membuat 3 garis, dengan marker dan label masing-masing.

```
28.plt.title("Data Korona Jakarta")
```

```
29.plt.xlabel("Tanggal")
```

```
30 plt.ylabel("Jumlah")
```

Keterangan : Membuat judul , xlabel atau label di bawah grafik dan ylabel atau label di samping grafik.

```
31.plt.xticks(rotation=90)
```

Keterangan : Label Horizontal, ditulis vertikal.

```
32.legend=ax.legend(loc="upper left", shadow=True)
```

Keterangan : Penempatan Legenda di upper left dengan ada shadow atau bayangannya.

```
34.for i,j in zip(x,y):
```

```
35. ax.annotate('%s' %j,xy=(i,j),xytext=(4,0), textcoords='offset points')
```

Keterangan : Iterasi untuk meletakkan value pada masing-masing titik di garis kasus

```
37.for a,b in zip(x,z):
```

```
38. ax.annotate('%s' %b,xy=(a,b),xytext=(4,0), textcoords='offset points')
```

Keterangan : Iterasi untuk meletakkan value pada masing-masing titik di garis meninggal

```
40.for d,e in zip(x,xx):
```

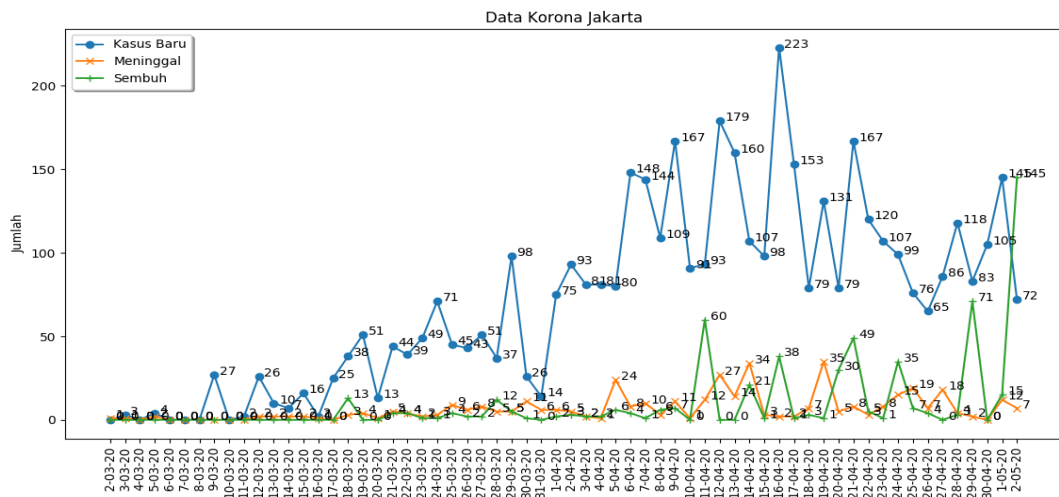
```
41. ax.annotate('%s' %e,xy=(d,e),xytext=(4,0), textcoords='offset points')
```

Keterangan : Iterasi untuk meletakkan value pada masing-masing titik di garis sembuh

```
43. print(plt.show())
```

Keterangan : Mencetak grafik

Output grafik akan seperti ini :



Gambar 4 :Output grafik garis data harian Covid 19 Jakarta

Penutup

Covid 19 adalah penyakit yang sedang menjadi pandemi saat tulisan ini dibuat. Semua orang berharap pandemi ini akan segera berakhir, termasuk orang-orang di Indonesia. Perkembangan harian Covid 19 ini diterangkan melalui website <https://corona.jakarta.go.id> oleh Pemerintah DKI Jakarta, salah satunya data berbentuk grafik garis. Penulis mencoba membuat mengolah dan membuat visualisasi data grafik garis kasus harian tersebut melalui numpy dan matplotlib di Python. Langkah selanjutnya, kita dapat mengolah lebih jauh data harian tersebut, seperti menganalisa perkembangan terkini dan membuat prediksi perkembangan kasus harian Covid 19 ini. Terakhir, penulis berharap semoga tulisan ini berguna.

Referensi

- [1] Data Pemantauan Covid 19, <https://corona.jakarta.go.id>, diakses : 2 Mei 2020
- [2] Python, <https://www.python.org/>, diakses : 1 Mei 2020
- [3] Numpy, <https://numpy.org>, diakses : 1 Mei 2020
- [4] matplotlib, <https://matplotlib.org>, akses : 1 Mei 2020

Biografi Penulis



Akhmad Sofwan - Menyelesaikan Sarjana di Teknik Informatika Universitas Budi Luhur , Jakarta tahun 2001 dan Magister Ilmu Komputer, Fakultas Ilmu Komputer ,Universitas Indonesia, tahun 2017. Penulis saat ini adalah seorang Freelance Software Developer. Selain itu, penulis juga mengajar di Program Vokasi, jurusan Manajemen Rekod dan Arsip (MRA) Universitas Indonesia. Penulis juga aktif di Pusat Kajian Biostatistik dan Informatika Kesehatan (PKBIK) Fakultas Kesehatan Masyarakat, Universitas Indonesia sebagai Software Developer. Bidang peminatan penulis adalah Data Mining dan

Machine Learning.