

LAPORAN OPERASI PYTHON MENGGUNAKAN DATA ANALYSIS LIBRARY

**KECEPATAN ANGIN MENURUT BULAN DI KOTA BANDUNG PADA TAHUN
2019 MENGGUNAKAN FITUR GRAFIK PADA PYTHON**



Oleh :

Nama : Rizka Angelina

NRP : 03411940000007

DEPARTEMEN TEKNIK GEOFISIKA

FAKULTAS TEKNIK SIPIL, PERENCANAAN, DAN KEBUMIHAN

INSTITUT TEKNOLOGI SEPULUH NOPEMBER

Operasi Python menggunakan Data Analysis Library (Pandas)

Python memiliki 5 library dalam pengolahan data science. Dalam tugas pengolahan data kali ini hanya menggunakan 2 library, yaitu pandas dan matplotlib. Berikut ini penjelasan mengenai pandas dan matplotlib:

1. Pandas

Dengan menggunakan sistem dataframe, kamu dapat memuat sebuah file ke dalam tabel virtual ala spreadsheet dengan menggunakan Pandas. Dengan menggunakan Pandas, kamu dapat mengolah suatu data dan mengolahnya seperti join, distinct, group by, agregasi, dan teknik seperti pada SQL. Hanya saja dilakukan pada tabel yang dimuat dari file ke RAM.

Pandas juga dapat membaca file dari berbagai format seperti .txt, .csv, .tsv, dan lainnya. Anggap saja Pandas adalah spreadsheet namun tidak memiliki GUI dan punya fitur seperti SQL.

2. Matplotlib

Data yang kita olah tentu tidak elok apabila ditampilkan begitu saja dengan tabel hitam saja kepada investor atau manajemen. Bila ditampilkan dengan sejumlah grafik berwarna pasti mereka akan lebih tertarik melihatnya. Matplotlib membantu kamu untuk memvisualisasikan data dengan lebih indah dan rapi.

Ada plot untuk menampilkan data secara 2D atau 3D. Sehingga kamu dapat menampilkan data yang telah kamu olah sesuai kebutuhan. Matplotlib pun terintegrasi dengan iPython Notebook atau Jupyter dimana kamu dapat membuat sebuah buku interaktif yang dapat diberi penjelasan dan kode yang disisipkan begitupun hasil plottingnya.

Matplotlib adalah library paling banyak digunakan oleh data science untuk menyajikan datanya ke dalam visual yang lebih baik.

Lanjut ke praktik menggunakan python. Hal pertama yang harus kita lakukan adalah mencari data di internet dengan format .txt, .csv, .tsv, namun kali ini saya menggunakan data dengan format .csv yang saya ambil dari halaman “Data Kota Bandung”. Setelah mendapat data yang

diinginkan simpan link data tersebut. Setelah itu buka python. Karena kita menggunakan Pandas. Maka ketikkan “import pandas as pd”.

```
In [4]: import pandas as pd
```

Lalu tekan *Enter*, dan ketik judul data yang tadi sudah dipilih. Untuk judul yang panjang dan menggunakan spasi, spasi digantikan dengan *underscore* (`_`). Lalu ketikkan “`=pd.read_csv`”. Untuk yang memakai format selain `.csv` maka setelah *underscore* “`csv`” diubah menjadi format yang dipilih. Lalu, *copy paste* kan link sudah disimpan menggunakan kurung dan tanda petik.

```
In [5]: kecepatan_angin=pd.read_csv('http://data.bandung.go.id/dataset/03b7fc16-269d-48e3-aa78-e1b9c3ee2ab4/resource/013d65e7-6150-4560-9f5d-4a8b8b8b8b8b')
```

Setelah itu, cari dataset menggunakan keterangan `head`, `tail`, `info`, dan `describe`. Lalu satu persatu di klik *run*.

```
In [6]: kecepatan_angin.head()
```

```
Out[6]:
```

	No	Bulan	Kecepatan Angin Rata-rata (km/jam)	Kecepatan Angin Terbesar (km/jam)
0	1	Januari	6	17
1	2	Februari	4	12
2	3	Maret	5	15
3	4	April	4	13
4	5	Mei	5	13

```
In [7]: kecepatan_angin.tail()
```

```
Out[7]:
```

	No	Bulan	Kecepatan Angin Rata-rata (km/jam)	Kecepatan Angin Terbesar (km/jam)
7	8	Agustus	4	15
8	9	September	5	17
9	10	Oktober	5	17
10	11	November	3	17
11	12	Desember	3	13

```
In [6]: kecepatan_angin.info()
```

```
<class 'pandas.core.frame.DataFrame'>
RangeIndex: 12 entries, 0 to 11
Data columns (total 4 columns):
No                12 non-null int64
Bulan             12 non-null object
Kecepatan Angin Rata-rata (km/jam)  12 non-null int64
Kecepatan Angin Terbesar (km/jam)  12 non-null int64
dtypes: int64(3), object(1)
memory usage: 512.0+ bytes
```

```
In [7]: kecepatan_angin.describe()
```

```
Out[7]:
```

	No	Kecepatan Angin Rata-rata (km/jam)	Kecepatan Angin Terbesar (km/jam)
count	12.000000	12.000000	12.000000
mean	6.500000	4.333333	14.583333
std	3.605551	0.984732	1.975225
min	1.000000	3.000000	12.000000
25%	3.750000	3.750000	13.000000
50%	6.500000	4.500000	14.000000
75%	9.250000	5.000000	17.000000
max	12.000000	6.000000	17.000000

Setelah di *Run* akan muncul dataset seperti di atas. Setelah itu ketik “import matplotlib.pyplot as plt”. Library ini digunakan untuk membuat grafik pada data yang sudah ada di atas. Lalu klik *Run*.

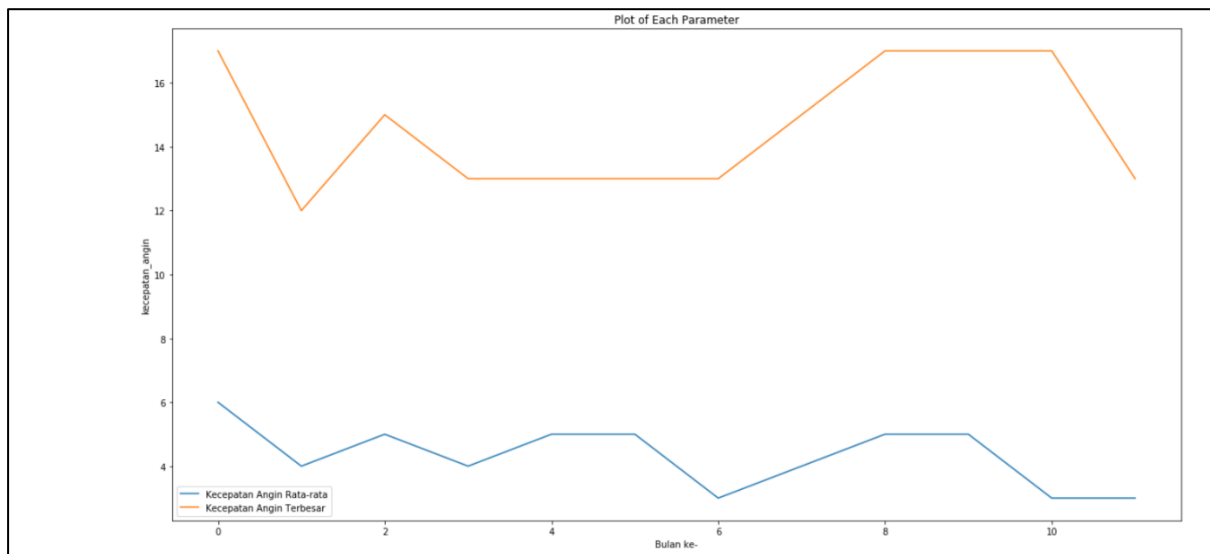
```
In [8]: import matplotlib.pyplot as plt
```

Lalu ketik seperti di bawah ini:

```
In [9]: plt.figure(figsize=(20,10))
plt.plot(kecepatan_angin['Kecepatan Angin Rata-rata (km/jam)'], label='Kecepatan Angin Rata-rata')
plt.plot(kecepatan_angin['Kecepatan Angin Terbesar (km/jam)'], label='Kecepatan Angin Terbesar')
plt.xlabel('Bulan ke-')
plt.ylabel('kecepatan_angin')
plt.title('Plot of Each Parameter')
plt.legend()
plt.show()
```

1. Untuk *figsize* yaitu ukuran figure atau ukuran grafik, saya menggunakan 20x10.
2. Untuk plot, dapat di lihat di dataset info, tentang apa aja yang harus diplot.
3. Label pada sumbu x saya gunakan untuk bulan ke x.
4. Label pada sumbu y saya gunakan untuk data kecepatan angin.
5. Untuk judul saya gunakan plot parameter

Setelah itu klik *Run*, dan akan muncul grafik pada gambar di bawah.



Pada sumbu x menampilkan bulan pada 1 tahun. Bulan 0 sampai bulan 11. Bulan 0 adalah bulan januari, bulan 1 adalah bulan februari, begitu seterusnya sampai bulan 11. Pada sumbu y menampilkan data kecepatan angin rata-rata dan kecepatan angina terbesar. Lalu ada dua garis berbeda warna, dan sudah ada keterangannya di dalam grafik.

Berikut hasil foto bila digabungkan:

