

LAPORAN TUGAS PHYTON 2  
MENGOLAH DATA MENJADI GRAFIK



Disusun Oleh:

Joeliardi Tirta Hadi

03411940000005

INSTITUT TEKNOLOGI SEPULUH NOPEMBER

2020

## I. Data Asli

Pada tugas kali ini, saya menggunakan Dasar Harga Berlaku Menurut Lapangan Usaha Kab Mojokerto.

3	Kategori (SNA 2008)	PDRB Kabupaten Moioikerto Atas Dasar Harga Berlaku Menurut Lapangan Usaha (Juta Rupiah)									
		2010	2011	2012	2013	2014	2015	2016	2017	2018	2019
4	Pertanian, Kehutanan, dan Perikanan	3140818.09	3549020.68	4028695.15	4467951.52	4947284.44	5429556.66	5925564.68	6101986.93	6234214.66	6417782.09
5	Pertambangan dan Penggalian	379981.01	430241.58	459433.17	481927.17	548227.98	593571.91	636147.93	680601.05	741424.85	777514.55
6	Industri Pengolahan	18209046.48	20569307.46	22838763.66	24860898.74	27823078.06	31149113.85	34129479.39	37566587.72	41007107.19	44349866.05
7	Pengadaan Listrik dan Gas	24605.45	25817.44	26936.02	26423.36	29926.85	34316.86	37317.27	44799.54	49458.41	53307.39

## II. Pengolahan Data Pada Python

- `import pandas as pd`
- `iris =`  
`pd.read_csv(https://raw.githubusercontent.com/joeliardi/tugas02/master/data%20fix%202.csv )`
- `iris.head()`

```
In [14]: 1 iris.head()
```

Out [14]:

	Kategori (SNA 2008)	2010	2011	2012	2013	2014	2015	2016	2017	2018	2019
0	Pertanian Kehutanan dan Perikanan	3140818.09	3549020.68	4028695.15	4467951.52	4947284.44	5429556.66	5925564.68	6101986.93	6234214.66	6417782.09
1	Pertambangan dan Penggalian	379981.01	430241.58	459433.17	481927.17	548227.98	593571.91	636147.93	680601.05	741424.85	777514.55
2	Industri Pengolahan	18209046.48	20569307.46	22838763.66	24860898.74	27823078.06	31149113.85	34129479.39	37566587.72	41007107.19	44349866.05
3	Pengadaan Listrik dan Gas	24605.45	25817.44	26936.02	26423.36	29926.85	34316.86	37317.27	44799.54	49458.41	53307.39

- `iris.tail()`

```
In [15]: 1 iris.tail()
```

Out [15]:

	Kategori (SNA 2008)	2010	2011	2012	2013	2014	2015	2016	2017	2018	2019
0	Pertanian Kehutanan dan Perikanan	3140818.09	3549020.68	4028695.15	4467951.52	4947284.44	5429556.66	5925564.68	6101986.93	6234214.66	6417782.09
1	Pertambangan dan Penggalian	379981.01	430241.58	459433.17	481927.17	548227.98	593571.91	636147.93	680601.05	741424.85	777514.55
2	Industri Pengolahan	18209046.48	20569307.46	22838763.66	24860898.74	27823078.06	31149113.85	34129479.39	37566587.72	41007107.19	44349866.05
3	Pengadaan Listrik dan Gas	24605.45	25817.44	26936.02	26423.36	29926.85	34316.86	37317.27	44799.54	49458.41	53307.39

- iris.info()

Output of `iris.info()`:

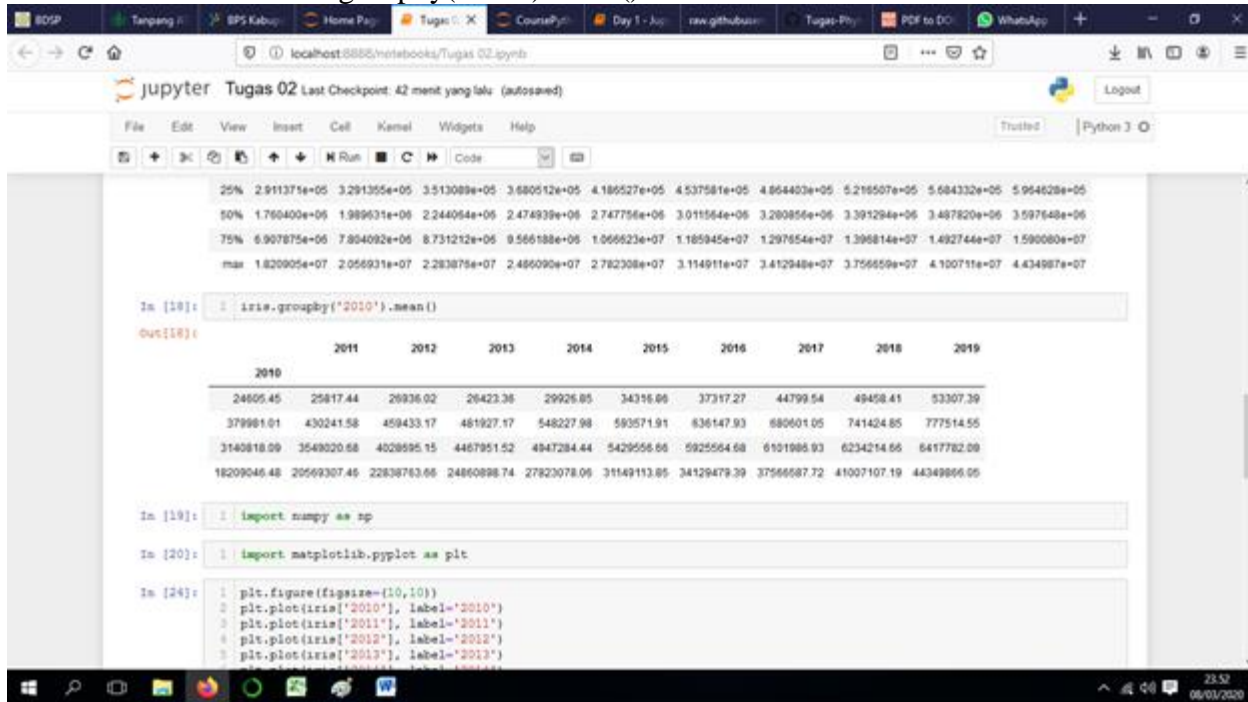
```
<class 'pandas.core.frame.DataFrame'>
RangeIndex: 4 entries, 0 to 3
Data columns (total 11 columns):
Kategori (SNA 2008)    4 non-null object
2010                  4 non-null float64
2011                  4 non-null float64
2012                  4 non-null float64
2013                  4 non-null float64
2014                  4 non-null float64
2015                  4 non-null float64
2016                  4 non-null float64
2017                  4 non-null float64
2018                  4 non-null float64
2019                  4 non-null float64
dtypes: float64(10), object(1)
memory usage: 480.0+ bytes
```

- iris.describe()

Output of `iris.describe()`:

	2010	2011	2012	2013	2014	2015
count	4.000000e+00	4.000000e+00	4.000000e+00	4.000000e+00	4.000000e+00	4.000000e+00
mean	5.438613e+06	6.143597e+06	6.838457e+06	7.459300e+06	8.337129e+06	9.301640e+06
std	8.626801e+06	9.745129e+06	1.081655e+07	1.177137e+07	1.317662e+07	1.476504e+07
min	2.460545e+04	2.581744e+04	2.693602e+04	2.642336e+04	2.992685e+04	3.431686e+04
25%	2.911371e+05	3.291355e+05	3.513089e+05	3.680512e+05	4.186527e+05	4.537581e+05
50%	1.760400e+06	1.989631e+06	2.244064e+06	2.474939e+06	2.747756e+06	3.011564e+06
75%	6.907875e+06	7.804092e+06	8.731212e+06	9.566188e+06	1.066623e+07	1.185945e+07
max	1.820905e+07	2.056931e+07	2.283876e+07	2.486090e+07	2.782308e+07	3.114911e+07

- iris.groupby('2010').mean()

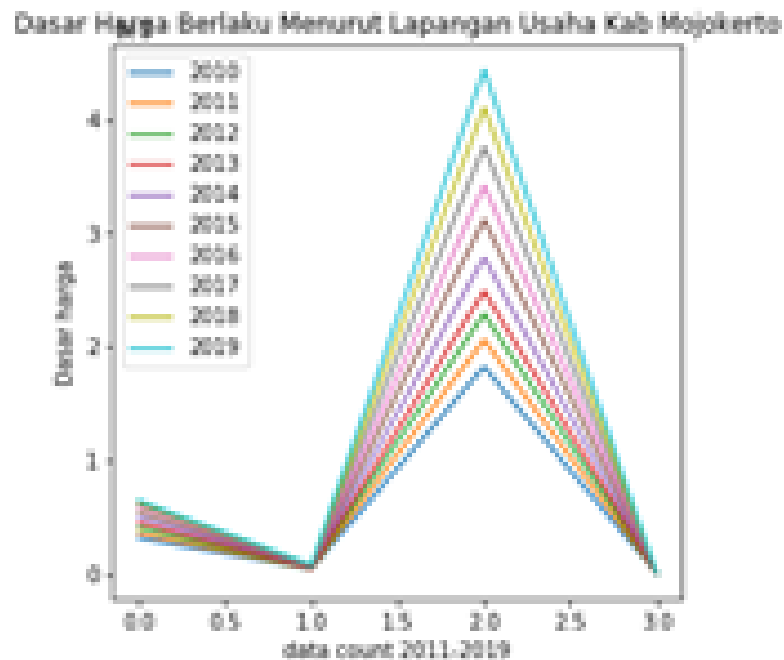


- import numpy as np
- import matplotlib.pyplot as plt
- plt.figure(figsize=(10,10))
  - plt.plot(iris['2010'], label='2010')
  - plt.plot(iris['2011'], label='2011')
  - plt.plot(iris['2012'], label='2012')
  - plt.plot(iris['2013'], label='2013')
  - plt.plot(iris['2014'], label='2014')

```

plt.plot(iris['2015'], label='2015')
plt.plot(iris['2016'], label='2016')
plt.plot(iris['2017'], label='2017')
plt.plot(iris['2018'], label='2018')
plt.plot(iris['2019'], label='2019')
plt.xlabel('data count 2011-2019')
plt.ylabel('Dasar harga')
plt.title('Dasar Harga Berlaku Menurut Lapangan Usaha Kab Mojokerto')
plt.legend()
plt.show()

```



### III. Penjelasan Fungsi Pada Script

1. `import pandas as pd` berfungsi untuk memanipulasi data, persiapan data, dan pembersihan data. Penggunaan *as* berarti menggantikan *pandas* dengan *prefix* *pd* untuk proses berikutnya.
2. `iris=pd.read_csv('https://raw.githubusercontent.com/joeliardi/tugas02/master/data%20fix%202.csv')` `iris` berfungsi sebagai nama file yang akan digunakan, `pd.read` berfungsi untuk membaca data dari file yang tertera dalam link tersebut, pembacaan data (tabel) menggunakan format `csv`.
3. `iris.head()` berfungsi untuk menampilkan data secara *default* untuk 5 data teratas.

4. `iris.tail()` berfungsi untuk menampilkan data secara *default* untuk 5 data terbawah.
5. `iris.info()` berfungsi untuk menampilkan informasi yang berkaitan dengan data yang ada.
6. `iris.describe()` berfungsi untuk mengetahui statistika data untuk data *numeric* seperti *count*, *mean*, *standard deviation*, *maximum*, *minimum*, dan *quartile*.
7. `iris.groupby('2010').mean()` berfungsi untuk mengelompokkan data dari tabel, pada data di atas dikelompokkan setelah dikalikan dengan rata-rata data.
8. `import numpy as np` berfungsi untuk memanggil numpy (membentuk objek N-dimensional array, mirip dengan list pada Python).
9. `import matplotlib.pyplot as plt` penggunaan `as` berfungsi untuk memanggil atau mengaktifkan fungsi, matplotlib berfungsi untuk membuat grafik dari data tabel yang tersedia.
10. `plt.figure(figsize=(10,10))` berfungsi untuk membuat grafik dengan ukuran yang sudah ditentukan, dalam grafik di atas menggunakan ukuran (10,10)
11. `plt.plot(iris['2010'], label='2010')` berfungsi untuk memasukkan data tabel ke dalam grafik yang akan dibuat, misalnya pada poin tersebut berarti memasukkan data pada tahun 2011 untuk diinput ke dalam grafik, kemudian diberi label atau nama (2010).
12. `plt.xlabel('data count 2010-2019')` berfungsi untuk memberi label atau nama pada sumbu-x grafik, pada grafik di atas diberi label (data count 2011-2020).
13. `plt.ylabel('Data harga')` berfungsi untuk memberi label atau nama pada sumbu-y grafik, pada grafik di atas diberi label (penduduk).
14. `plt.title(Dasar Harga Berlaku Menurut Lapangan Usaha Kab Mojokerto)` berfungsi untuk memberi judul pada grafik agar pembaca mudah memahami, grafik di atas diberi nama (Dasar Harga Berlaku Menurut Lapangan Usaha Kab Mojokerto).
15. `plt.legend()` berfungsi untuk menampilkan legenda dari grafik yang dibuat, meliputi judul, label, warna, dan keterangan lainnya.
16. `plt.show()` berfungsi untuk menampilkan grafik pada hasil script Python (Window).

#### IV. Penjelasan Grafik

Pada sumbu-x (nomor 0 sampai 3) adalah data tabel baris satu sampai tiga. Dalam grafik juga tertera keterangan warna yang menunjukkan pada tahun tertentu. Sumbu-y merupakan dasar harga.