

ASSIGNMENT
PYTHON PROJECT 2



Daneswara Widyanggastari

03411942000003

Department of Geophysical Engineering

2019

Sepuluh Nopember Institute of Technology

Surabaya

In the second week, we have been taught how to import three famous library on Python. The first library is numpy, then the second one is matplotlib, and the last one is pandas. We also learned about how to analyze data in a form CSV or XLS. Therefore, the following results of the script are the data processing:

```
In [3]: 1 import pandas as pd
```

```
In [17]: 1 import matplotlib.pyplot as plt
```

```
In [6]: 1 iris=pd.read_csv('http://data.jakarta.go.id/dataset/8282968f-671d-488e-9e91-4b54d7772425/resource/c6fdb43f-05a7-4db3-887b-fd
<
>
```

```
In [7]: 1 iris.head()
```

```
Out[7]:
```

	nama_titik	lokasi	latitude	longitude	h_ellipsoid_m
0	439B	Pluit Raya Jakut	-6.120881	106.794169	18.7033
1	ANBA	Ancol Baru Jakut	-6.122604	106.824513	18.9314
2	ANCL	Wisata Ancol Jakut	-6.121193	106.845625	19.8997
3	BM06	Jalan Jend. Sudirman Jakpus	-6.199965	106.823154	23.1547
4	BM07	Jalan Jend. Sudirman Jakpus	-6.212325	106.820099	32.6150


```
In [8]: 1 iris.tail()
```

```
Out[8]:
```

	nama_titik	lokasi	latitude	longitude	h_ellipsoid_m
49	PTRO	cempaka putih	-6.171890	106.885024	NaN
50	UKG4	sunter	-6.116603	106.886811	NaN
51	AW12	Kemayoran	-6.159925	106.846849	20.6340
52	CP08	Rawa Buaya	-6.159311	106.723726	20.8346
53	BBKN	Situ Babakan	-6.338573	106.826394	75.0480

Figure 1. Shows the processing data with head and tail syntax

```
In [9]: 1 iris.info()
```

```
<class 'pandas.core.frame.DataFrame'>
RangeIndex: 54 entries, 0 to 53
Data columns (total 5 columns):
nama_titik      54 non-null object
lokasi          54 non-null object
latitude        54 non-null float64
longitude       54 non-null float64
h_ellipsoid_m   44 non-null float64
dtypes: float64(3), object(2)
memory usage: 2.2+ KB
```



```
In [10]: 1 iris.describe()
```

```
Out[10]:
```

	latitude	longitude	h_ellipsoid_m
count	54.000000	54.000000	44.000000
mean	-6.169747	106.822206	27.861795
std	0.066968	0.056745	15.109389
min	-6.351533	106.688977	18.357800
25%	-6.195509	106.792399	20.046325
50%	-6.152156	106.823072	21.024000
75%	-6.121545	106.853759	30.399725
max	-6.086202	106.953421	80.660600

Figure 2. Data processing with info and describe syntax

Then after the data processing with a few of syntax, next is grouping. You see that in describe syntax you have got count, mean, std, max, min, etc. So, unfortunately the grouping data is too long. The first is longitude data and then I also add mean on the script. And then the latitude data, the last is h_ellipsoid_m all the describe syntax that I use was mean.

```
In [18]: 1 plt.figure(figsize=(8,6))
2 plt.plot(iris['latitude'], label='LT')
3 plt.plot(iris['longitude'], label='LG')
4 plt.plot(iris['h_ellipsoid_m'], label='HEM')
5
6 plt.xlabel('NUMBER OF DATA')
7 plt.ylabel('VALUE')
8 plt.title('Data Penurunan Tanah di DKI Jakarta Tahun 2014')
9 plt.legend()
10 plt.show()
```

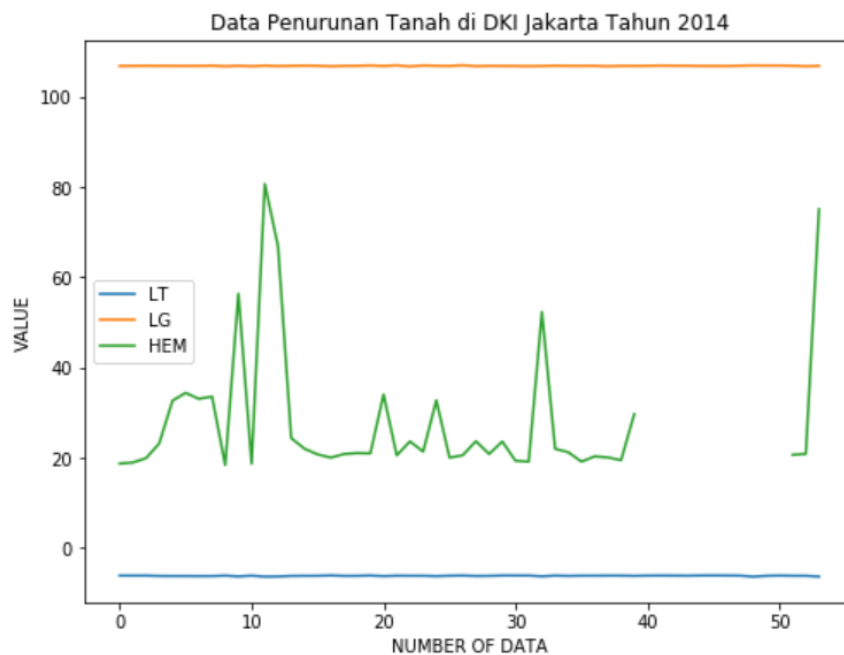


Figure 3. Graph figure of the data process

Based on the figure that shown above is about how to analyze the data from csv and then we make a graph using python. The data that I use is from data.jakarta.go.id because I want to know what is inside. There are a lot of data that shows us about population, traffic, flood, birth, death, etc. But I'm very interested about the data of land subsidence in Jakarta. Here is the explanation.

In the script figure above, there are several functions used, namely, the "head" function is inputted on the data to show the top 5 data, the tail function is inputted on that data to show the bottom 5 data, the "info" function is used on that data to show the index number along with its data type, the "describe" function to show statistical summaries such as averages, medians, and quartiles and much more in that column, the mean "groupby" function is inputted for grouping averages on the data inputted, plotting for graphs and also contained graphical interpretation of the data that I'm processing. In making graphs it uses a function in the form of "matplotlib.pyplot" which is imported as "plt". "Plt" here serves to plot a graph. After completing the plot, a graph interpretation is obtained, like the third image.