

Nama : Aussie Anindya Nari Pinandita

NRP : 03411940000018

Tugas II Phyton

Pada tugas Phyton mengenai pengulasan data tabular, saya mengambil data dari sumber Satu Data Indonesia mengenai jumlah kunjungan mancanegara yang datang ke DKI Jakarta dan Indonesia pada tahun 2004-2012, seperti pada tabel berikut.

Kunjungan Wisatawan Mancanegara yang Datang ke DKI Jakarta dan Indonesia Tahun 20014-2012		
Tahun	DKI Jakarta	Indonesia
2004	1063910	4541165
2005	1235514	4074354
2006	1216132	4871351
2007	1216057	5505759
2008	1534785	6234497
2009	1451914	6323730
2010	1892866	7002944
2011	2003944	7649731
2012	2125513	8044462

Tabel 1.1 Data kunjungan wisatawan mancanegara yang datang ke DKI Jakarta dan Indonesia

Untuk mengolah data dari tabel tersebut pada Phyton, perlu dilakukan langkah-langkah seperti sebagai berikut.

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

In [85]: iris=pd.read_csv('https://raw.githubusercontent.com/aussieanindya/tugas-python2/master/Kunjungan-Wisatawan-Mancanegara-yar')

In [86]: iris.head()

Out[86]:
```

	tahun	dki_jakarta	indonesia
0	2004	1063910	4541165
1	2005	1235514	4074354
2	2006	1216132	4871351
3	2007	1216057	5505759
4	2008	1534785	6234497

```
In [87]: iris.tail()

Out[87]:
```

	tahun	dki_jakarta	indonesia
4	2008	1534785	6234497
5	2009	1451914	6323730
6	2010	1892866	7002944
7	2011	2003944	7649731
8	2012	2125513	8044462

1. Pertama, lakukan programming pada Phytom dengan menginput **import pandas as pd** yang berfungsi untuk memanipulasi, persiapan, dan pembersihan data yang akan diolah pada Phyton, kemudian jalankan program.

- Kedua, upload file data csv di github. Setelah selesai mengupload, buka raw pada file kemudian copy link raw.
- Paste link yang telah dicopy pada Phyton dengan format **iris=pd.read_csv('paste link')** kemudian run. Dengan **iris** berfungsi sebagai variabel atau nama file yang akan digunakan, dan **pd.read** berfungsi untuk membaca data pada link yang telah dimasukkan.
- Masukkan format **iris.head()** untuk menampilkan 5 data teratas dalam table.
- Masukkan format **iris.tail()** untuk menampilkan 5 data terakhir dalam table.

```
In [68]: iris.info()

<class 'pandas.core.frame.DataFrame'>
RangeIndex: 9 entries, 0 to 8
Data columns (total 3 columns):
tahun      9 non-null int64
dki_jakarta  9 non-null int64
indonesia   9 non-null int64
dtypes: int64(3)
memory usage: 344.9 bytes

In [69]: iris.describe()

Out[69]:
```

	tahun	dki_jakarta	indonesia
count	9.000000	9.000000e+00	9.000000e+00
mean	2008.000000	1.526731e+06	4.827555e+06
std	2.739613	1.961907e+05	1.388107e+06
min	2004.000000	1.062010e+06	4.074354e+06
25%	2008.000000	1.216132e+06	4.871351e+06
50%	2008.000000	1.411914e+06	6.234497e+06
75%	2010.000000	1.892066e+06	7.002944e+06
max	2012.000000	2.125513e+06	8.044462e+06

- Masukkan format **iris.info()** untuk menampilkan info dari data table terkait.
- Masukkan format **iris.describe()** untuk melihat statistik data termasuk count, mean, nilai min dan max serta beberapa persentase.

```
In [70]: iris.groupby('tahun').mean()

Out[70]:
```

	dki_jakarta	indonesia
tahun		
2004	1063010	4541165
2005	1235514	4074354
2006	1216132	4871351
2007	1216057	5605759
2008	1534785	6234497
2009	1451914	6323730
2010	1892066	7002944
2011	2003944	7649731
2012	2125513	8044462

```
In [71]: import numpy as np

In [72]: import matplotlib.pyplot as plt
```

- Masukkan **iris.groupby('tahun').mean()** untuk mengelompokkan data dari table setelah dikalikan dengan rata-rata.
- Masukkan import **numpy as np** untuk menyederhanakan numpy menjadi np yang berfungsi untuk membentuk N-dimensional array atau seperti list pada phyton.
- Masukkan import **matplotlib.pyplot as plt** untuk menyederhanakan panggilan menjadi plt, dan berfungsi untuk membuat grafik dari data table tersebut.

```

1 plt.figure(figsize=(10,10))
2 plt.plot(iris['dki_jakarta'], label='dki_jakarta')
3 plt.plot(iris['indonesia'], label='indonesia')
4
5
6 plt.xlabel('data count 2004-2012')
7 plt.ylabel('Jumlah Pengunjung')
8 plt.title('Kunjungan Wisatawan Mancanegara yang Datang ke DKI Jakarta dan Indonesia 2004-2012')
9 plt.legend()
10 plt.show()

```

11. Masukkan format **plt.figure(figsize=(10,10))** untuk membuat ukuran grafik sebesar 10x10
12. Masukkan format **plt.plot(iris['dki_jakarta'], label='dki_jakarta')** untuk memasukkan setiap data dalam table ke dalam grafik yang akan dibuat.
13. Masukkan format **plt.xlabel('data count 2004-2012')** untuk memberi label atau nama variabel pada sumbu x
14. Masukkan format **plt.ylabel('Jumlah Pengunjung')** untuk memberi nama variabel pada sumbu y.
15. Masukkan format **plt.title('Kunjungan Wisatawan Mancanegara yang Datang ke DKI Jakarta dan Indonesia 2004-2012')** untuk memberi judul pada grafik.
16. Masukkan format **plt.legend()** untuk menampilkan legenda dari grafik yang telah dibuat, seperti judul, label, warna, dan keterangan garis.
17. Masukkan format **plt.show()** untuk menampilkan grafik yang telah dibuat pada hasil script.

Setelah memasukkan script yang telah dibuat, akan muncul sebuah grafik yang diinginkan seperti gambar berikut.

