


<p><b>Nama:</b> Aini Rihhadatul Aisy</p> <p><b>NIM:</b> 064102400024</p>	 <p><b>Praktikum Algoritma &amp; Pemrograman</b></p>	<p><b>MODUL 12</b></p> <p><b>Nama Dosen:</b> Binti solihah, S.T, M.KOM</p>
<p><b>Hari/Tanggal:</b> Jumat, 06 Desember 2024</p>		<p><b>Nama Asisten Labratorium:</b></p> <ol style="list-style-type: none"> <li>1. Yustianas Rombon - 064002300015</li> <li>2. Vira Aditya Kurniawan - 065002300012</li> </ol>

## Pengantar Data Science

### 1. Teori Singkat

Data Science adalah suatu disiplin ilmu yang khusus mempelajari data, khususnya data kuantitatif, baik yang terstruktur maupun tidak terstruktur. Banyak Bahasa pemrograman yang dimana dia mendukung untuk melakukan pengolahan data, diantaranya yaitu Bahasa Pemrograman R, Python, SQL, dan JavaScript, dan lain sebagainya. Python sendiri merupakan salah satu bahasa yang mendukung untuk melakukan pengolahan data, bahkan python sendiri menyediakan library untuk pengolahan data itu sendiri, salah satunya adalah library *pandas*. Untuk melakukan pengolahan data python sendiri merekomendasikan untuk menggunakan IDE yang disediakan untuk melakukan pengolahan data yaitu adalah Jupyter Notebook.

#### Pandas Data Frame

Struktur data dasar pandas dinamakan DataFrame, yaitu sebuah koleksi kolom berurutan dengan nama dan jenis, dengan demikian merupakan sebuah tabel yang tampak seperti database dimana sebuah baris tunggal mewakili sebuah contoh tunggal dan kolom mewakili atribut tertentu. Pandas data frame juga dapat disebut sebagai dictionary of list karena bentuknya seperti list yang memiliki identifikasi *key-value* untuk mengidentifikasi setiap datanya.

### Contoh Program Data Science

```
import pandas as pd

data = {"Negara": ["Indonesia", "Jepang", "India", "China", "Amerika Serikat", "Brazil"],
        "Ibu Kota": ["Jakarta", "Tokyo", "New Delhi", "Beijing", "Washington, D.C.", "Brazillia"],
        "Benua": ["Asia", "Asia", "Asia", "Asia", "Amerika", "Amerika"],
        "Luas": [1905, 377, 3287, 9597, 9834, 8515],
        "Populasi": [264, 143, 1252, 1357, 329, 210] }

df = pd.DataFrame(data)
mean = df.groupby(['Benua']).mean()
std = df.groupby(['Benua']).std()

print(df)
print(mean)
print(std)
```

### Output



	Negara	Ibu Kota	Benua	Luas	Populasi
0	Indonesia	Jakarta	Asia	1905	264
1	Jepang	Tokyo	Asia	377	143
2	India	New Delhi	Asia	3287	1252
3	China	Beijing	Asia	9597	1357
4	Amerika Serikat	Washington, D.C.	Amerika	9834	329
5	Brazil	Brazilia	Amerika	8515	210
	Luas	Populasi			
	Benua				
	Amerika	9174.5	269.5		
	Asia	3791.5	754.0		
		Luas	Populasi		
	Benua				
	Amerika	932.673844	84.145707		
	Asia	4048.705266	639.018518		

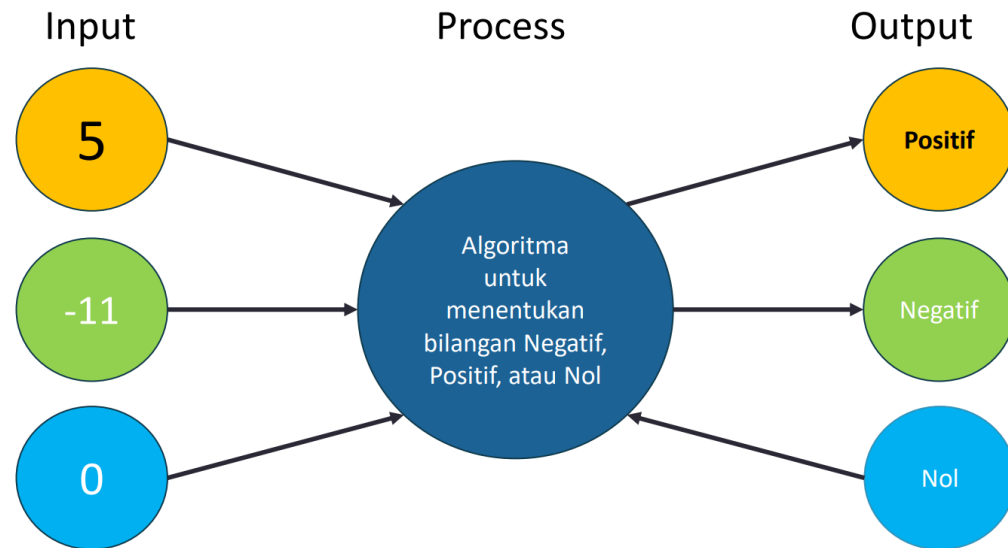
## IPO (Input Process Output)

Konsep Dasar Input, Process, dan Output (IPO)

- Konsep input, process, dan output adalah prinsip dasar dalam pemrograman dan pengembangan algoritma.
- Setiap algoritma melibatkan tiga tahap utama: mengambil data masukan (input), melakukan operasi atau pengolahan data (process), dan menghasilkan hasil akhir (output).
- Konsep ini menggambarkan bagaimana algoritma beroperasi untuk memproses informasi.



## Gambaran IPO (Menentukan Bilangan)



### Notasi Algoritma Flowchart

1. Flowchart adalah representasi visual atau diagram alir yang digunakan untuk menggambarkan langkahlangkah dan urutan proses suatu algoritma atau program.
2. Flowchart menyajikan langkah-langkah dalam bentuk simbol-simbol grafis yang saling terhubung, membantu dalam memvisualisasikan bagaimana informasi mengalir dan bagaimana proses dilakukan.
3. Dalam kaitannya dengan notasi deskriptif, notasi algoritma yang menggunakan flowchart dapat lebih cepat dibaca dan dilihat alur dan hubungannya.

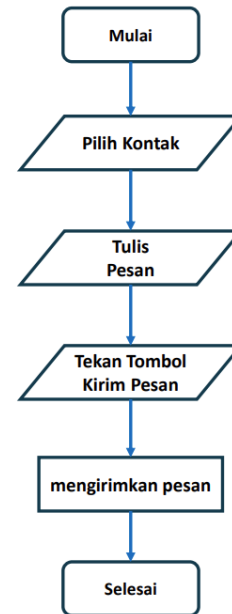
### Simbol-simbol pada Flowchart

1. Setiap elemen flowchart dihubungkan oleh garis aliran bertanda panah
2. Garis aliran dimulai dari atas symbol dan keluar dari bagian bawah, kecuali symbol keputusan yang alirannya keluar dari bawah atau samping
3. Aliran bergerak dari atas ke bawah
4. Proses awal dan akhir menggunakan symbol terminal.



Contoh sederhana  
Penggunaan *flowchart*  
untuk menunjukan algoritma

Kasus/Aliran:  
Mengirim pesan WhatsApp



## 2. Alat dan Bahan

Hardware : Laptop/PC

Software : Spyder (Anaconda Python)

## 3. Elemen Kompetensi

### a. Latihan pertama

Buatlah sebuah program yang dapat membaca data frame yang datanya diambil dari file CSV, buatlah minimal 10 data negara dan tampilkan Mean(Rata-rata) dan Standar Deviasinya.

IPO (Input Process Output)

#### Input:

1. File CSV: Aini Rihhadatul Aisy - Negara.csv
2. Kolom dalam file:
  - o **Negara:** Nama negara
  - o **Populasi:** Data populasi negara



**Process:**

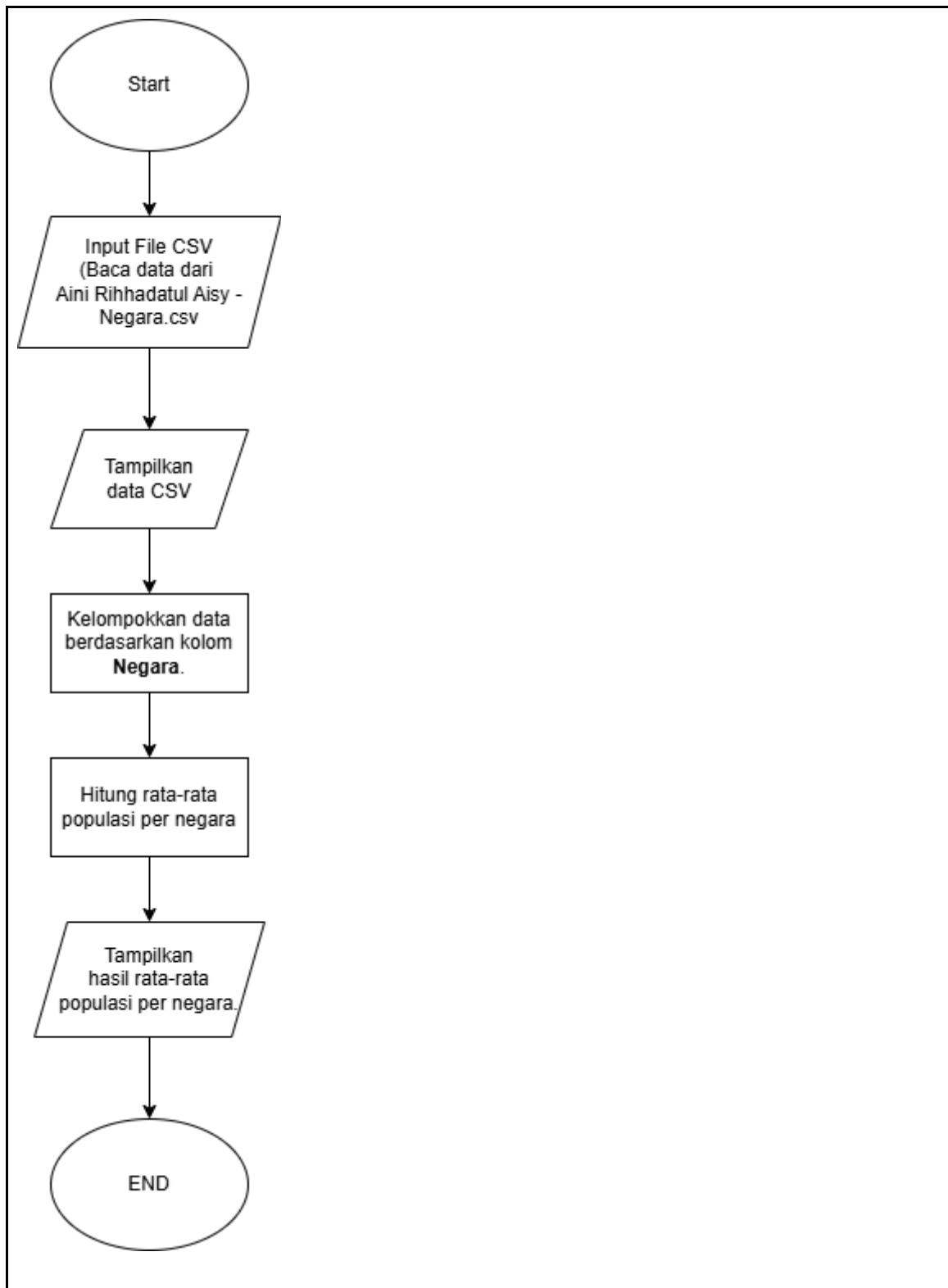
1. Membaca file CSV menggunakan `pandas.read_csv`.
2. Kelompokkan data berdasarkan kolom **Negara**.
3. Hitung rata-rata kolom **Populasi** untuk setiap **Negara** menggunakan `.groupby()` dan `.mean()`.

**Output:**

1. Menampilkan isi dataset (dp) secara keseluruhan.
2. Menampilkan hasil rata-rata populasi tiap negara (`mean_population`).

Flowchart





## Source Code

```
import pandas as pd

dp = pd.read_csv('Aini Rihhadatul Aisy - Negara.csv')
mean_population = dp.groupby(['Negara'])['Populasi'].mean()

print(dp)
print()
print(mean_population)
```

## Output

	Negara	Ibu Kota	Benua	Luas	Populasi
0	Indonesia	Jakarta	Asia	1905	264
1	Jepang	Tokyo	Asia	377	143
2	India	New Delhi	Asia	3287	1252
3	China	Beijing	Asia	9597	1357
4	Amerika Serikat	Washington DC	Amerika	9834	329
5	Brazil	Brazilia	Amerika	8515	210
6	Rusia	Moskow	Asia	17098	146
7	Meksiko	Meksiko City	Amerika	1964	126
8	Nigeria	Abuja	Afrika	923	200
9	Jerman	Berlin	Erupa	357	83
10	Aljazair	Aljazair	Afrika	2381	43
11	Inggris	London	Erupa	242	66





```

Negara
Aljazair          43.0
Amerika Serikat   329.0
Brazil            210.0
China             1357.0
India             1252.0
Indonesia         264.0
Inggris           66.0
Jepang            143.0
Jerman            83.0
Meksiko           126.0
Nigeria          200.0
Rusia             146.0
Name: Populasi, dtype: float64

```

b. Latihan Kedua

Buatlah sebuah program yang dapat menulis file CSV yang berisi data mean dan standar deviasi dari hasil keluaran data frame Negara, Luas Area dan Total Populasi pada Latihan sebelumnya. Terdapat dua File output yaitu NegaraStandarDeviasi.csv dan juga NegaraMean.csv

IPO

**Input**

- Data negara: nama negara, luas wilayah, dan populasi.

**Process**

1. Membuat DataFrame dari data negara.
2. Menghitung rata-rata (mean) kolom Luas dan Populasi.
3. Menghitung standar deviasi (std) kolom Luas dan Populasi.
4. Membuat dua DataFrame:
  - Satu untuk rata-rata (mean).
  - Satu untuk standar deviasi.

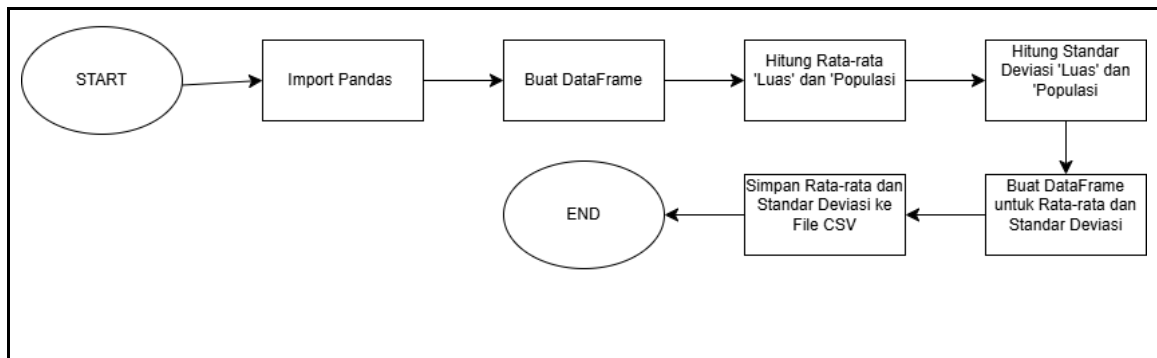


5. Menyimpan DataFrame ke file CSV:
  - NegaraMean.csv untuk rata-rata.
  - NegaraStandarDeviasi.csv untuk standar deviasi.
6. Menampilkan data secara rapi di layar.

#### Output

1. File CSV:
  - NegaraMean.csv
  - NegaraStandarDeviasi.csv
2. Tabel di layar:
  - Rata-rata (mean).
  - Standar deviasi.

#### Flowchart



#### Source Code



```
import pandas as pd
data = {
    "Negara": ["Indonesia", "Jepang", "India", "China", "Amerika Serikat", "Brazil", "Rusia", "Meksiko", "Nigeria", "Jerman", "Aljazair", "Inggris"],
    "Luas": [1905, 377, 3287, 9597, 9834, 8515, 17098, 1964, 923, 357, 2381, 242],
    "Populasi": [264, 143, 1252, 1357, 329, 210, 146, 126, 200, 83, 43, 66]
}
df = pd.DataFrame(data)
mean_data = df[["Luas", "Populasi"]].mean()
std_data = df[["Luas", "Populasi"]].std()
mean_df = pd.DataFrame({
    "Deskripsi": ["Mean"],
    "Luas": [f"{mean_data['Luas']:.2f}"],
    "Populasi": [f"{mean_data['Populasi']:.2f}"]
})
std_df = pd.DataFrame({
    "Deskripsi": ["Standard Deviation"],
    "Luas": [f"{std_data['Luas']:.2f}"],
    "Populasi": [f"{std_data['Populasi']:.2f}"]
})
```

```
mean_df.to_csv("NegaraMean.csv", index=False)
std_df.to_csv("NegaraStandarDeviiasi.csv", index=False)
print("File NegaraMean.csv dan NegaraStandarDeviiasi.csv telah dibuat!")
print("\nIsi file NegaraMean.csv:")
print(mean_df.to_markdown(index=False))
print("\nIsi file NegaraStandarDeviiasi.csv:")
print(std_df.to_markdown(index=False))
```

## Output

File NegaraMean.csv dan NegaraStandarDeviiasi.csv telah dibuat!

Isi file NegaraMean.csv:

Deskripsi	Luas	Populasi
Mean	4706.67	351.58

Isi file NegaraStandarDeviiasi.csv:

Deskripsi	Luas	Populasi
Standard Deviation	5333.98	453.11

## 4. File Praktikum

Github Repository:



## 5. Soal Latihan

Soal:

1. Apa kegunaan data frame dalam library pandas pada bahasa pemrograman python dan jelaskan mengapa sebelum melakukan pengolahan data kita diharuskan menggunakan data frame?
2. Deskripsikan serta narasikan jalannya alur source code program yang sebelumnya telah kalian buat pada Elemen Kompetensi Latihan Kedua!

Jawaban:

1. DataFrame adalah struktur data dua dimensi yang mempermudah pengelolaan dan analisis data terstruktur. Sebelum mengolah data, DataFrame digunakan untuk membuat data lebih terorganisir, mendukung berbagai format, dan menyediakan fungsi statistik bawaan.
2. Program dimulai dengan pembuatan DataFrame yang memuat data nama negara, luas wilayah, dan populasi. Kemudian, program menghitung rata-rata dan standar deviasi untuk kolom "Luas" dan "Populasi" menggunakan fungsi `.mean()` dan `.std()` dari Pandas. Hasil perhitungan ini disimpan dalam dua DataFrame terpisah: satu untuk rata-rata dan satu lagi untuk standar deviasi. Kedua DataFrame ini kemudian diekspor ke file CSV bernama "NegaraMean.csv" dan "NegaraStandarDeviasi.csv". Akhirnya, isi DataFrame ditampilkan ke layar untuk validasi hasil pengolahan.

## 6. Kesimpulan

- a. Dalam pengerjaan program dengan bahasa pemrograman Python, kita harus benar-benar teliti dalam menginputkan suatu fungsi untuk menampilkan suatu keluaran pada layar dengan sesuai.
- b. Kita dapat mengetahui bahwa penggunaan DataFrame dalam Pandas sangat membantu mengelola data terstruktur dan melakukan analisis seperti rata-rata dan standar deviasi. Menyimpan data ke file CSV mempermudah dokumentasi dan analisis di masa depan. Latihan ini juga mengajarkan pentingnya memahami alur kerja program melalui flowchart.

## 7. Cek List (✓)

No	Elemen Kompetensi	Penyelesaian	
		Selesai	Tidak Selesai
1.	Latihan Pertama	✓	



2.	Latihan Kedua	✓	
----	---------------	---	--

## 8. Formulir Umpan Balik

No	Elemen Kompetensi	Waktu Pengerjaan	Kriteria
1.	Latihan Pertama	20 Menit	Menarik
2.	Latihan Kedua	40 Menit	Cukup

Keterangan:

1. Menarik
2. Baik
3. Cukup
4. Kurang