

<p>Nama: Dimas Alip Priyono</p> <p>NIM: 064102400032</p>	 <p>Praktikum Algoritma & Pemrograman</p>	<h2>MODUL 12</h2> <p>Nama Dosen: Binti solihah, S.T, M.KOM</p> <p>Nama Asisten Labratorium:</p> <ol style="list-style-type: none">1. Yustianas Rombon - 0640023000152. Vira Aditya Kurniawan - 065002300012
<p>Hari/Tanggal: Jumat, 6 Desember 2024</p>		

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.", "Brazilia"],
        "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
```

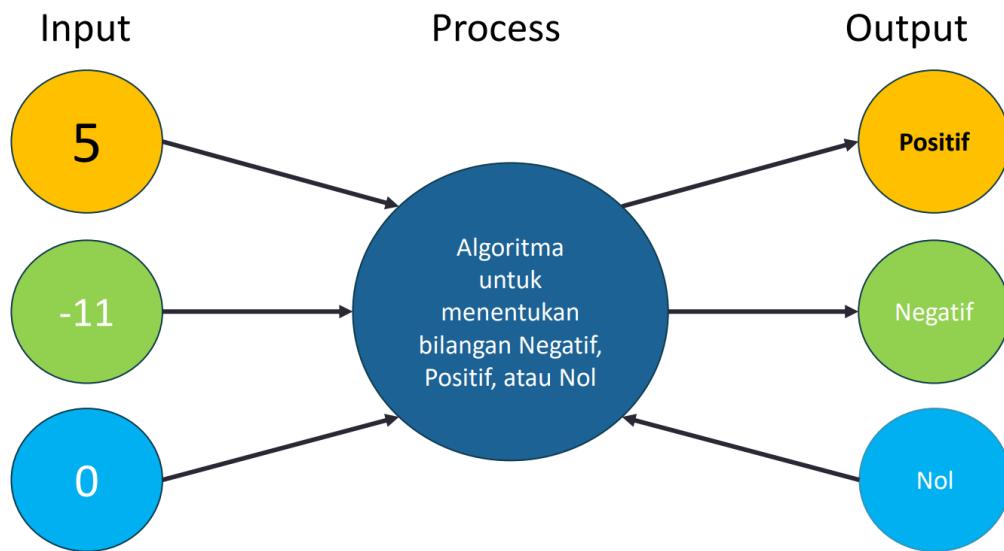
IPO (Input Process Output)

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 langkah-langkah 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.

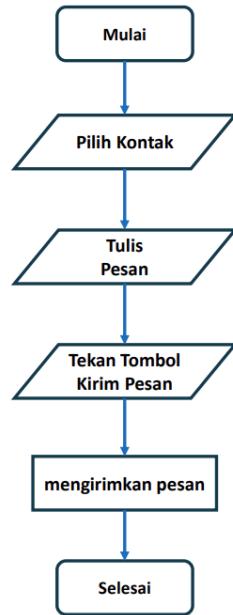
Simbol-simbol pada Flowchart

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

Data berupa dictionary yang terdiri dari:

- Nama negara
- Nama ibu kota
- Nama benua
- Luas wilayah dalam ribuan kilometer persegi
- Populasi dalam jutaan



Process

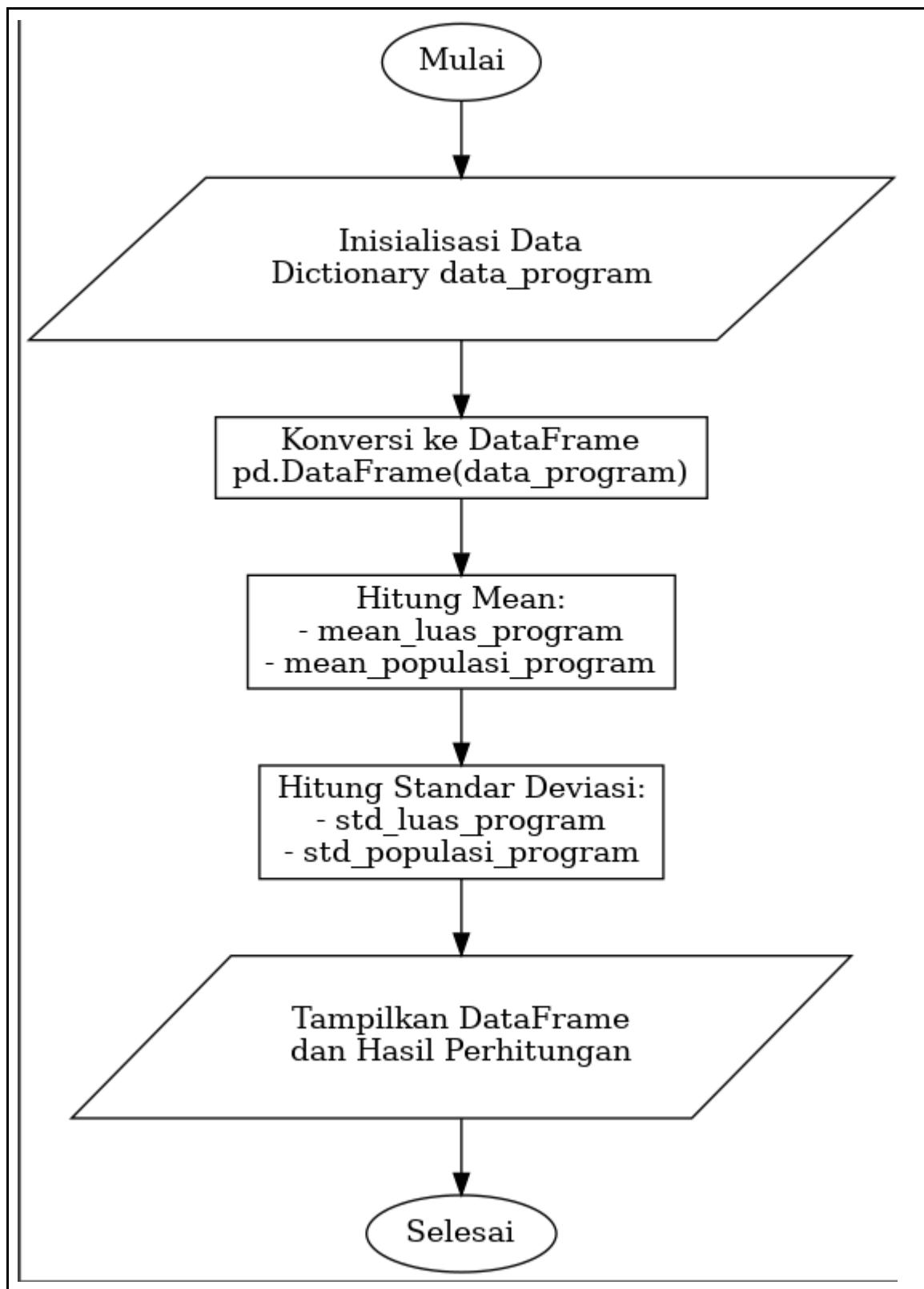
1. Membuat DataFrame menggunakan Pandas berdasarkan data input.
2. Menghitung:
 - o Rata-rata (mean) dari luas wilayah.
 - o Standar deviasi (standard deviation) dari luas wilayah.
 - o Rata-rata (mean) dari populasi.
 - o Standar deviasi (standard deviation) dari populasi.
3. Menampilkan DataFrame dan hasil perhitungan.

Output

1. DataFrame yang merepresentasikan tabel negara, ibu kota, benua, luas wilayah, dan populasi.
2. Nilai rata-rata dan standar deviasi untuk:
 - o Luas wilayah
 - o Populasi

Flowchart





Source Code

```
import pandas as pd

data_program = {
    "Negara": ["Indonesia", "Jepang", "India", "China", "Amerika Serikat", "Brazil", "Rusia", "Meksiko", "Nigeria", "Jerman"],
    "Ibu Kota": ["Jakarta", "Tokyo", "New Delhi", "Beijing", "Washington DC", "Brazilia", "Moskow", "Meksiko City", "Abuja", "Berlin"],
    "Benua": ["Asia", "Asia", "Asia", "Asia", "Amerika", "Amerika", "Asia", "Amerika", "Afrika", "Eropa"],
    "Luas (ribu km2)": [1905, 377, 3287, 9597, 9834, 8515, 17098, 1964, 923, 357],
    "Populasi (juta)": [264, 143, 1252, 1357, 329, 210, 146, 126, 200, 83]
}

df_program = pd.DataFrame(data_program)

mean_luas_program = df_program["Luas (ribu km2)"].mean()
std_luas_program = df_program["Luas (ribu km2)"].std()
mean_populasi_program = df_program["Populasi (juta)"].mean()
std_populasi_program = df_program["Populasi (juta)"].std()

print(df_program)
print(f"Mean Luas: {mean_luas_program}")
print(f"Standard Deviation Luas: {std_luas_program}")
print(f"Mean Populasi: {mean_populasi_program}")
print(f"Standard Deviation Populasi: {std_populasi_program}")
```

Output



	Negara	Ibu Kota	Benua	Luas (ribu km ²)	Populasi (juta)
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	Eropa	357	83
Mean Luas: 5385.7					
Standard Deviation Luas: 5607.647021899846					
Mean Populasi: 411.0					
Standard Deviation Populasi: 476.83097401257163					

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 berupa dictionary data yang berisi:
 - Nama negara.
 - Luas wilayah dalam ribuan kilometer persegi.
 - Populasi dalam jutaan.

Process

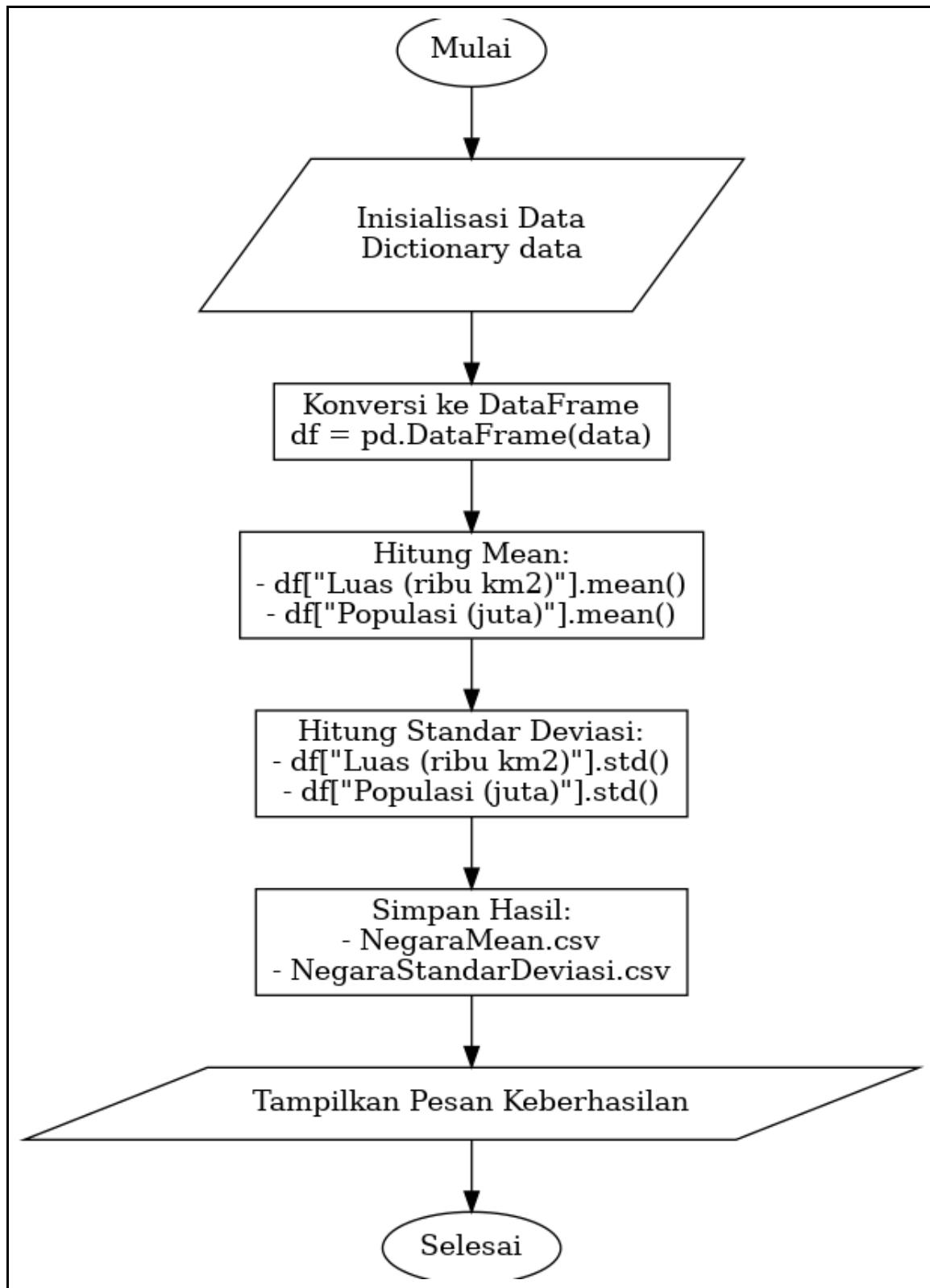
1. **Membuat DataFrame:**
 - Data dari dictionary diubah menjadi DataFrame menggunakan pd.DataFrame(data).
2. **Menghitung Mean:**
 - Rata-rata dari kolom "Luas (ribu km²)" dihitung dengan df["Luas (ribu km²)"].mean().



- Rata-rata dari kolom "Populasi (juta)" dihitung dengan df["Populasi (juta)"].mean().
- 3. **Menghitung Standar Deviasi:**
 - Standar deviasi dari kolom "Luas (ribu km2)" dihitung dengan df["Luas (ribu km2)"].std().
 - Standar deviasi dari kolom "Populasi (juta)" dihitung dengan df["Populasi (juta)"].std().
- 4. **Membuat DataFrame Baru:**
 - Data rata-rata disimpan dalam DataFrame df_mean.
 - Data standar deviasi disimpan dalam DataFrame df_std_dev.
- 5. **Menyimpan ke CSV:**
 - Data mean disimpan dalam file NegaraMean.csv.
 - Data standar deviasi disimpan dalam file NegaraStandarDeviasi.csv

Flowchart





Source Code

```
import pandas as pd
import numpy as np

data = {
    "Negara": ["Indonesia", "Jepang", "India", "China", "Amerika Serikat", "Brazil", "Rusia", "Meksiko", "Nigeria", "Jerman"],
    "Luas (ribu km2)": [1905, 377, 3287, 9597, 9834, 8515, 17098, 1964, 923, 357],
    "Populasi (juta)": [264, 143, 1252, 1357, 329, 210, 146, 126, 200, 83]
}
df = pd.DataFrame(data)

mean_values = {
    "Negara": ["Luas (ribu km2)", "Populasi (juta)"],
    "Mean": [df["Luas (ribu km2)"].mean(), df["Populasi (juta)"].mean()]
}

std_dev_values = {
    "Negara": ["Luas (ribu km2)", "Populasi (juta)"],
    "Standar Deviasi": [df["Luas (ribu km2)"].std(), df["Populasi (juta)"].std()]
}

df_mean = pd.DataFrame(mean_values)
df_std_dev = pd.DataFrame(std_dev_values)

df_mean.to_csv("NegaraMean.csv", index=False)
df_std_dev.to_csv("NegaraStandarDeviasi.csv", index=False)

print("File CSV telah dibuat:")
print("- NegaraMean.csv")
print("- NegaraStandarDeviasi.csv")
```

Output



→ File CSV telah dibuat:

- NegaraMean.csv
- NegaraStandarDeviasi.csv

1 to 2 of 2 entries

Negara	Mean
Luas (ribu km2)	5385.7
Populasi (juta)	411.0

Show per page

1 to 2 of 2 entries

Negara	Standar Deviasi
Luas (ribu km2)	5607.647021899846
Populasi (juta)	476.83097401257163

Show per page

4. File Praktikum

Github Repository:

<https://github.com/Alip1023/Prak12algo.git>

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 utama dalam library Pandas yang dirancang untuk mengolah data dalam bentuk tabel, mirip dengan spreadsheet atau tabel database.



2. • Program menginisialisasi data dalam bentuk dictionary dan mengonversinya ke DataFrame.
- Program menghitung rata-rata dan standar deviasi untuk kolom "Luas Wilayah" dan "Populasi".
- Hasil perhitungan disimpan ke dalam dua DataFrame terpisah (`df_mean` dan `df_std_dev`).
- Program menyimpan hasil tersebut dalam file CSV (NegaraMean.csv dan NegaraStandarDeviasi.csv).
- Pesan keberhasilan ditampilkan di konsol.

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 program ini dirancang untuk mengolah data dalam format CSV yang berisi informasi mengenai negara, luas area, dan total populasi. Pertama, program akan membaca data dari file CSV untuk membentuk data frame yang berisi minimal 10 data negara. Selanjutnya, program akan menghitung nilai rata-rata (mean) dan standar deviasi dari data ini untuk masing-masing kolom - negara, luas area, dan total populasi.

7. Cek List (✓)

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

8. Formulir Umpam Balik

No	Elemen Kompetensi	Waktu Pengerjaan	Kriteria
1.	Latihan Pertama	60 Menit	Cukup
2.	Latihan Kedua	60 Menit	Cukup

Keterangan:

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

