|  |  |  |
| --- | --- | --- |
| **Nama:**  **M.Satria Pratama**  **NIM:**  **065002200017** | C:\Users\RPL-SI 02\Pictures\288px-Trisakti_Logo.svg.png | **MODUL 12**  **Nama Dosen:**  **Anung B. Ariwibowo, M. Kom** |
| **Hari/Tanggal:**  **Jum’at, 9,12 2022** | **Praktikum Algoritma & Pemrograman** | **Nama Asisten Labratorium:**   1. **Azhar Rizki Zulma 065001900001** |

**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 *pandasi*. 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

|  |
| --- |
|  |

Output

|  |
| --- |
|  |

1. **Alat dan Bahan**

Hardware : Laptop/PC

Software : Spyder (Anaconda Python)

1. **Elemen Kompetensi**
   1. Latihan pertama

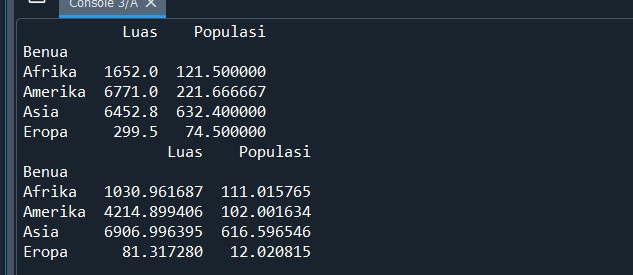
Buatlah sebuah program yang dapat membaca data frame yang datanya diambil dari file CSV, buatlah minimal 10 data negara dan tampilkan Mean(Rata-rata) dan Standar Deviasinya.

Source Code

|  |
| --- |
| import pandas as pd    data= pd.read\_csv("negara.csv")    df = pd.DataFrame(data) mean = df.groupby(['Benua']).mean() std = df.groupby(['Benua']).std()    print(data) print(mean)  print(std) |

Output

|  |
| --- |
| Output |



* 1. 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

Source Code

|  |
| --- |
|  |

Output

|  |
| --- |
|  |

1. **File Praktikum**

Github Repository:

|  |
| --- |
| https://github.com/satria461/-Tugas-Laporan-Praktikum-Algo-12 |

1. **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. Pandas adalah salah satu library yang wajib dipelajari pemula dalam belajar Python. Dengan berdasarkan sistem dataframe, modul ini dapat memuat sebuah file ke dalam tabel virtual menyerupai spreadsheet. Pandas juga berfungsi mengolah suatu data seperti teknik join, distinct, group by, agregasi, dan teknik lainnya seperti pada SQL. Bedanya, ini dilakukan pada tabel. Kelebihan dari library ini juga dapat membaca file dari berbagai format seperti .txt, .csv, dan .tsv. Dengan adanya fitur dataframe memudahkan untuk membaca sebuah file dan menjadikannya table, kita juga dapat mengolah suatu data dengan menggunakan operasi seperti join, distinct, group by, agregasi, dan teknik lainnya yang terdapat pada SQL. Banyak format file yang dapat dibaca menggunakan Pandas, seperti file .txt, .csv, .tsv dan lainnya.
2. Pada Latihan ke 2 sebelum mengeluarkan sebuah data dalam bentuk file csv di python, dengan memanggil sebuah library yaitu pandas untuk melakukan tugas yang di inginkan pada Latihan ke2. Selanjutnya alue code seperti pada Latihan ke1, hanya saja utuk mengeluarkan data mean dan data standart deviation, dipisah, lalu membuat file csv.
3. **Kesimpulan**
   1. Dalam pengerjaan program dengan bahasa pemrograman Python, kita harus benar-benar teliti dalam menginputkan suatu fungsi untuk menampilkan suatu keluaran pada layar dengan sesuai.
   2. Kita dapat mengetahui salah satu library pandas yang digunakan untuk manage data. Dalam manage data dengan adanya library ini memudahkan seseorang yang ingin menjadi data scient bisa lebih mudah untuk memproses data, serta memahami Struktur data dasar pandas dinamakan DataFrame. Pandas data frame juga dapat disebut sebagai dictionary of list karena bentuknya seperti list yang memiliki identifikasi key-value untuk mengidentifikasi setiap datanya

**Cek List (**✓**)**

|  |  |  |  |
| --- | --- | --- | --- |
| **No** | **Elemen Kompetensi** | **Penyelesaian** | |
| **Selesai** | **Tidak Selesai** |
| **1.** | Latihan Pertama | **(**✓**)** |  |
| **2.** | Latihan Kedua | **(**✓**)** |  |

1. **Formulir Umpan Balik**

|  |  |  |  |
| --- | --- | --- | --- |
| **No** | **Elemen Kompetensi** | **Waktu Pengerjaan** | **Kriteria** |
| **1.** | Latihan Pertama | 30 Menit | Menarik |
| **2.** | Latihan Kedua | 35 Menit | Menarik |

Keterangan:

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