Nama: Jonathan Bob Dylan Mongisidi

> NIM: 064102400013

Hari/Tanggal: Senin,10 Februari 2024



Praktikum Statistika

MODUL 2

Nama Dosen: Dr. Dedy Sugiarto, S.Si, M.Kom

Nama Asisten Labratorium:

- Michael Briant (064002300004)
- Monica Sicilia
 Simanjuntak
 (065002300030)

Tipe Data, Filter Data & Koneksi ke Database

1. Teori Singkat

Python memiliki beberapa tipe data dasar, di antaranya:

- Integer (int): Representasi bilangan bulat, misalnya: 5, -3, 100.
- Float (float): Representasi bilangan pecahan, misalnya: 3.14, 2.718.
- String (str): Urutan karakter, misalnya: 'hello', "world".
- Boolean (bool): Representasi nilai kebenaran, yaitu True atau False.
- List (list): Kumpulan elemen yang terurut dan dapat diubah, misalnya: [1, 2, 3, 4], ['apple', 'banana', 'cherry'].
- Tuple (tuple): Kumpulan elemen yang terurut dan tidak dapat diubah, misalnya: (1, 2, 3), ('red', 'green', 'blue').
- Dictionary (dict): Kumpulan pasangan kunci-nilai yang tidak terurut, misalnya: {'name': 'John', 'age': 30}.
- Set (set): Kumpulan elemen yang unik dan tidak terurut, misalnya: {1, 2, 3, 4}.

Filter Data dalam Python:

Untuk melakukan filter data dalam Python, Anda dapat menggunakan berbagai cara, tergantung pada struktur data yang Anda gunakan. Dalam konteks DataFrame, seperti yang digunakan dalam Pandas, Anda dapat menggunakan metode query() atau pengindeksan boolean.

2. Alat dan Bahan

Hardware : Laptop/PC Software : R Studio

3. Elemen Kompetensi

Terdapat beberapa tipe data di Jupyter antara lain vektor, matriks dan data frame. Cantumkan setiap output yang dihasilkan dari console Jupyter, ke kolom yang sudah disediakan.

a. Latihan pertama – Vektor

1. Tuliskan Perintah berikut ini di jupyter notebook

```
a = [1, 2, -5, 0.3, 6, -2, 4] # numeric vector
b = ["one", "two", "three"] # character vector
c = [True, True, True, False, True] # logical vector
print(a)
print(b)
print(c)
```

Output:

```
a = [1,2,-5,0.3,6,-2,4]
b = ["one", "two", "three"]
c = [True, True, False, True]
print(a)
print(b)
print(c)

[1, 2, -5, 0.3, 6, -2, 4]
['one', 'two', 'three']
[True, True, False, True]
```

b. Latihan Kedua – Matriks

1. Seluruh kolom dalam sebuah matriks harus memiliki tipe yang sama (numerik semua, karakter semua, dll) dan memiliki panjang yang sama.

^{*}gunakan nama variable dengan nama anda masing-masing

```
#MATRIKS
import numpy as np
cells = [3, 15, -27, 38]
r_nama = ["R1", "R2"]
c_nama = ["C1", "C2"]
nama_matrix = np.matrix(cells).reshape(2, 2)
print(nama_matrix)
```

- c. Latihan Ketiga Data Frame
 - 1. Mengubah data input menjadi data frame
 - *gunakan nama variable dengan nama anda masing-masing

```
import pandas as pd
import numpy as np

nama1 = [1, 2, 3, 4]
nama2 = ["red", "white", "red", np.nan] # Menggunakan np.nan untuk
merepresentasikan NA
nama3 = [True, True, True, False]

dataku = pd.DataFrame({'ID': nama1, 'Color': nama2, 'Passed': nama3})
print(dataku)
```

```
import pandas as pd
 import numpy as np
nama1 = [1, 2, 3, 4]
nama2 = ["red", "white", "red", np.nan]
nama3 = [True, True, True, False]
dataku = pd.DataFrame({'ID': nama1, 'Color': nama2, 'Passed': nama3})
print(dataku)
   ID Color Passed
         red True
0 1
1
   2 white
               True
   3
        red
               True
         NaN False
```

2. Selanjutnya ketikkan perintah dibawah ini

```
import pandas as pd

data_nama = pd.DataFrame(('id': list('abcdefghij'), 'x': list(range(1, 11)), 'y': list(range(11, 21))))
print(data_nama)
```

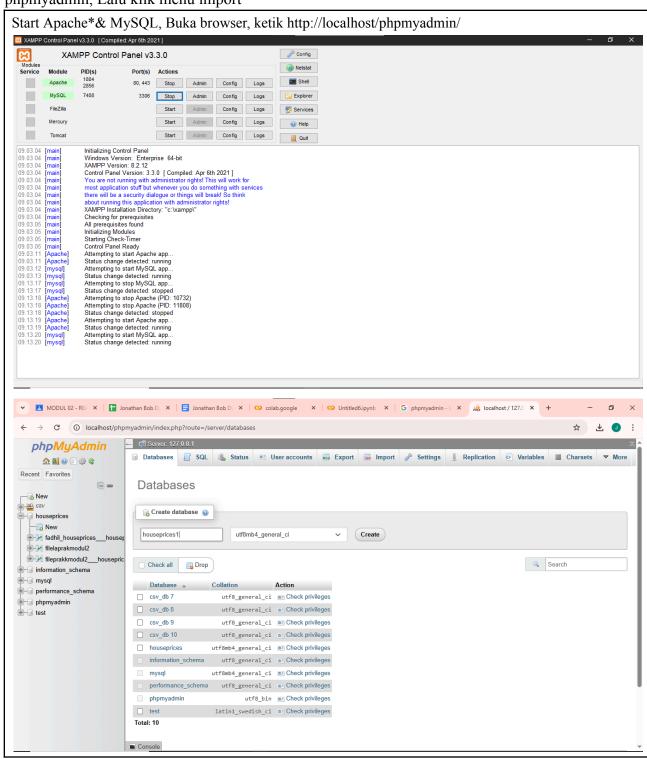
```
import pandas as pd

data_nama = pd.DataFrame({'id': list('abcdefghij'), 'x': list(range(1, 11)), 'y': list(range(11, 21))})

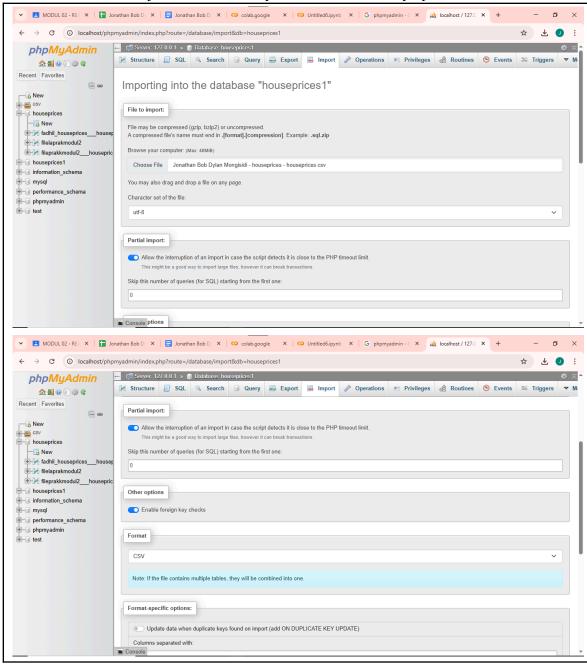
id x y
0 a 1 11
1 b 2 12
2 c 3 13
3 d 4 14
4 e 5 15
5 f 6 16
6 g 7 17
7 h 8 18
8 i 9 19
9 j 10 20
```

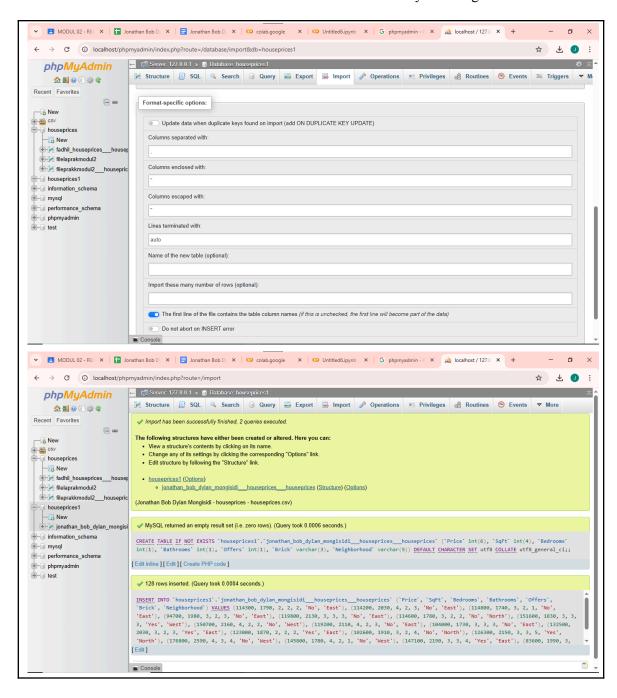
https://pandas.pydata.org/docs/reference/api/pandas.DataFrame.htm

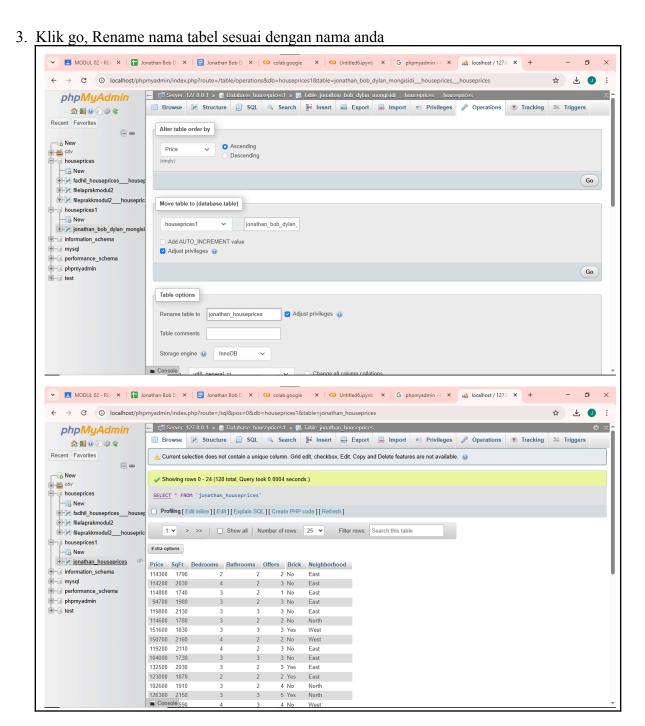
- d. Latihan Keempat Koneksi ke Database
 - 1. Buat sebuah nama database terlebih dahulu dengan nama houseprices di phpmyadmin, Lalu klik menu import



2. Pilih file yang ingin di import ke database (untuk file nama_excel.csv), Ceklis the first line of the file contains the table column name untuk membuat baris pertama pada file excel tersebut menjadi nama atributnya atau nama kolomnya pada database.







4. Kembali ke jupyter notebook, lalu instal dahulu library yang dibutuhkan pada python. Jika belum tersedia, maka lakukan instruksi:

```
pip install mysql-connector-python
Windows PowerShell
Copyright (C) Microsoft Corporation. All rights reserved.

Try the new cross-platform PowerShell https://aka.ms/pscore6

PS C:\Users\User\Documents> pip install mysql-connector-python
Requirement already satisfied: mysql-connector-python in c:\users\user\anaconda3\lib\site-packages (9.2.0)
PS C:\Users\User\Documents>
```

5. Lalu jalankan perintah dibawah ini

Output:

```
import mysql.connector

connection = mysql.connector.connect(
    host="localhost",
    user="root",
    password="",
    database="houseprices!"
)

cursor = connection.cursor()

try:
    my_query= "SELECT * FROM jonathan_houseprices;"
    cursor.execute(my_query)

result = cursor.fetchall()

print("\hasil Kueri")
    for row in result:
        print(row)

finally:
    cursor.close()
    connection.close()
```

```
Hasil Kueri
(114300, 1790, 2, 2, 2, 'No', 'East')
(114200, 2030, 4, 2, 3, 'No', 'East')
(114800, 1740, 3, 2, 1, 'No', 'East')
(94700, 1980, 3, 2, 3, 'No', 'East')
(119800, 2130, 3, 3, 3, 'No', 'East')
(119800, 1780, 3, 2, 2, 'No', 'North')
(151600, 1830, 3, 3, 'No', 'East')
(119700, 2160, 4, 2, 2, 'No', 'West')
(119200, 2110, 4, 2, 3, 'No', 'East')
(104000, 1730, 3, 3, 3, 'No', 'East')
(123000, 2230, 3, 2, 3, 'Yes', 'East')
(123000, 1870, 2, 2, 2, 'Ves', 'East')
(123000, 1870, 2, 2, 2, 'Ves', 'East')
(125000, 1910, 3, 2, 4, 'No', 'North')
(126000, 1910, 3, 2, 4, 'No', 'North')
(126800, 2150, 3, 3, 5, 'Yes', 'North')
(176800, 2590, 4, 3, 4, 'No', 'Nest')
```

6. Jalankan perintah dibawah ini:

*Perintah ini akan menampilkan 86 baris data hasil filter.

```
In []: import pandas as pd
    # Mengonversi hasil kueri ke DataFrame Pandas
    df = pd.DataFrame(result, columns=[desc[0] for desc in cursor.description])

# Filter data berdasarkan kolom 'Brick' yang bernilai 'No'
    df_filtered = df[df['Brick'] == 'No']

# Menampilkan hasil filter
    print("\nHasil Filter:")
    print(df_filtered)
```

Output:

7. Jalankan perintah dibawah ini:

*Perintah ini akan menampilkan 105 baris data hasil filter.

```
In [ ]: import pandas as pd
# Mengonversi hasil kueri ke DataFrame Pandas
df = pd.DataFrame(result, columns=[desc[0] for desc in cursor.description])

# Filter data berdasarkan kondisi yang kompleks
df_filtered = df[df['Brick'] == 'No') | (df['Neighborhood'] == 'East')]

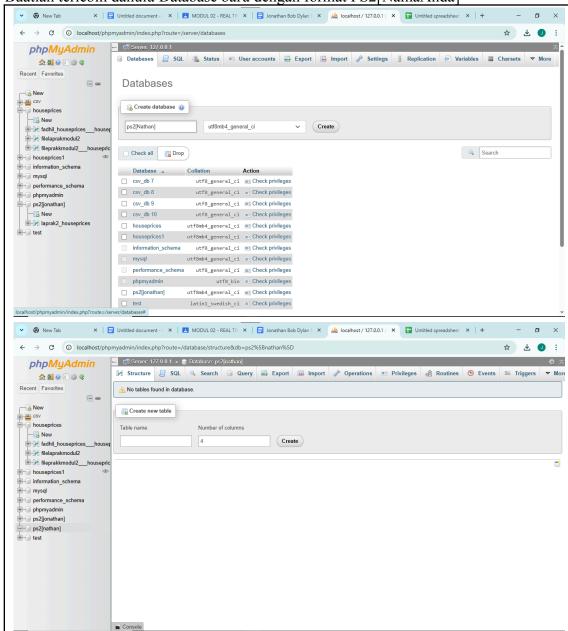
# Menampilkan hasil filter
print(df_filtered)
```

e. Latihan Keenam – Tugas

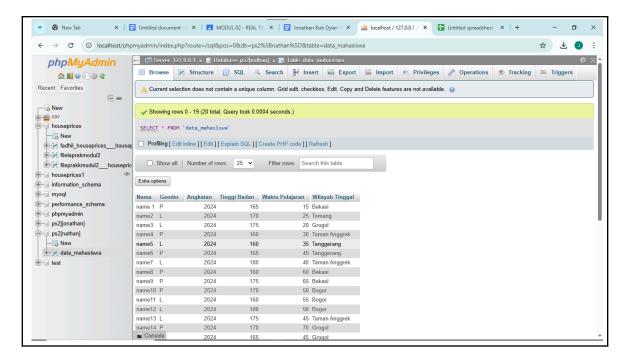
Buat sebuah database serta tabel di dalamnya (bisa gunakan data teman dipraktikum ke

1). Lakukan koneksi python ke database serta berikan beberapa filter data sesuai yang anda inginkan. Tampilkan data tersebut

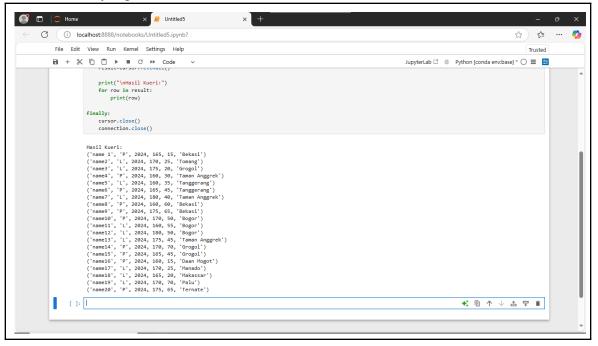
1. Buatlah terlebih dahulu Database baru dengan format PS2[NamaAnda]



2. Import file berformat csv yang telah anda buat sebelumnya di Praktikum 1, dimana file tersebut berisi 6 kolom dan 20 baris data.

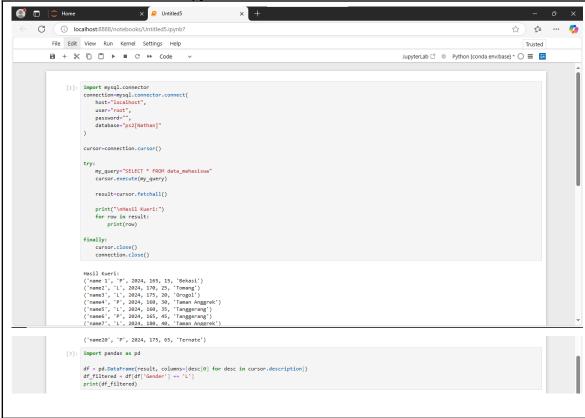


3. Koneksikan Python ke Database tersebut sebagaimana yang telah anda lakukan pada Elemen Kompetensi 1 dimodul kedua ini dengan menyesuaikan kembali nama Database baru yang sudah dibuat.



4. Lakukan filter data terhadap Kolom Gender, untuk melihat berapa baris data Pria/Wanita (Pilih salah 1).

- 5. Lampirkan Screenshot
- « Kode koneksi Jupyter ke Database
- « Kode serta hasil filter di Jupyter



Jonathan Bob Dylan Mongisidi - 064102400013

4.	File	Pral	ktikum
----	------	------	--------

Github Repository:		

5. Kesimpulan

Kesimpulan yang bisa saya ambil dari praktikum probabilitas dan statistika kali ini adalah dimana saya bisa mengerti tipe data, filter data, dan koneksi ke database dan juga mendapakan source code baru

6. Cek List (**✓**)

No	Elemen Kompetensi	Penyelesaian		
110	Elemen Rompetensi	Selesai	Tidak Selesai	
1.	Latihan Pertama	V		
2.	Latihan Kedua	V		
3.	Latihan Ketiga	V		
4.	Latihan Keempat	V		
5.	Latihan Kelima	V		
6.	Latihan Keenam	V		

7. Formulir Umpan Balik

No	Elemen Kompetensi	Waktu Pengerjaan	Kriteria
1.	Latihan Pertama	30 Menit	Menarik
2.	Latihan Kedua	30 Menit	Menarik
3.	Latihan Ketiga	30 Menit	Menarik
4.	Latihan Keempat	30 Menit	Menarik
5.	Latihan Kelima	30 Menit	Menarik
6.	Latihan Keenam	30 Menit	Menarik

Keterangan: 1. Menarik

- 2. Baik
- 3. Cukup4. Kurang

Jonathan Bob Dylan Mongisidi - 064102400013