

Laporan Praktikum Pertemuan 1

Data Science Lanjut

Dibuat Oleh

Nama : Muhamad Faisal Halim
NIM : 19.240.0163
Kelas : -
Mata Kuliah : Data Science Lanjut

Mahasiswa Pertukaran Mahasiswa.
Universitas Muhammadiyah Kalimantan Timur
~ STMIK Widya Pratama Pekalongan

Tools

Dalam praktikum ini saya menggunakan tools-tools dibawah ini.

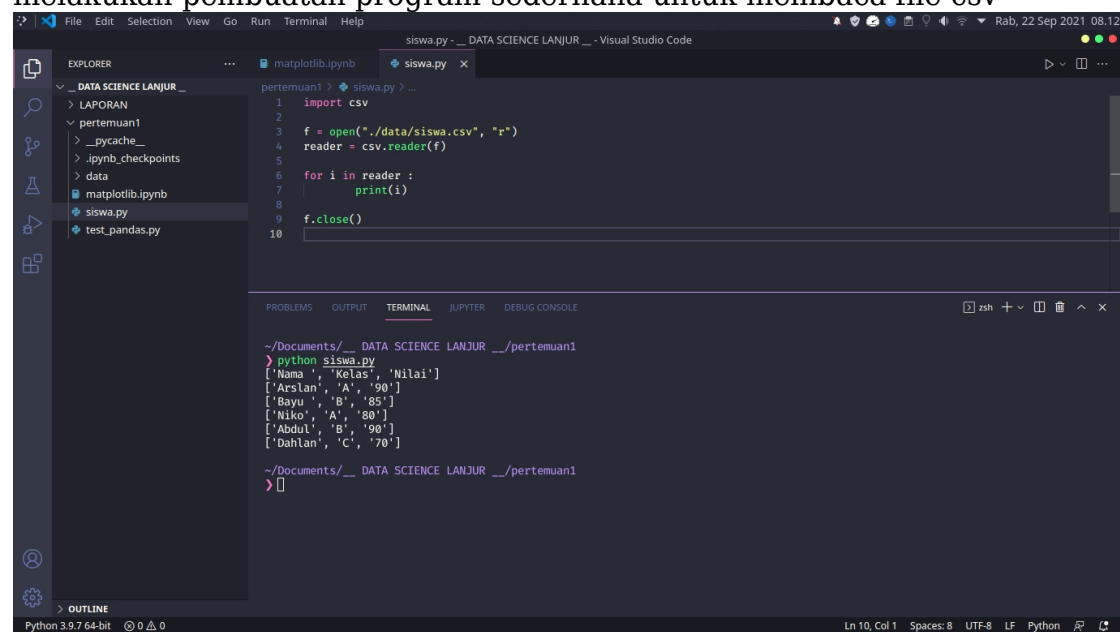
1. VSCODE dengan extensi
 - a) Python
 - b) Jupyter
2. BASH / ZSH
3. PYTHON & Library-nya
 - a) pandas
 - b) numpy
 - c) matplotlib

Praktikum

karena saya menggunakan VSCODE jadi saya membutuhkan ekstensi python dan Jupyter. setelah diinstall selanjutnya buat file dengan ekstensi .ipynb pada kasus ini, saya menamainya denagn matplotlib.ipynb.

Praktik 1

melakukan pembuatan program sederhana untuk membaca file csv



```
1 import csv
2
3 f = open("../data/siswa.csv", "r")
4 reader = csv.reader(f)
5
6 for i in reader:
7     print(i)
8
9 f.close()
10
```

```
~/Documents/_ DATA SCIENCE LANJUR _/pertemuan1
> python siswa.py
['Nama', 'Kelas', 'Nilai']
['Arslan', 'A', '90']
['Bayu', 'B', '85']
['Niko', 'A', '80']
['Abdul', 'B', '90']
['Dahlan', 'C', '70']

~/Documents/_ DATA SCIENCE LANJUR _/pertemuan1
>
```

pada praktikum ini hanya menggunakan library bawaan dari python, yaitu csv. library ini akan otomatis ada ketika kita install python di komputer kita. jika dilihat dari hasil output di terminalnya, kita mendapatkan data array.

Praktik 2

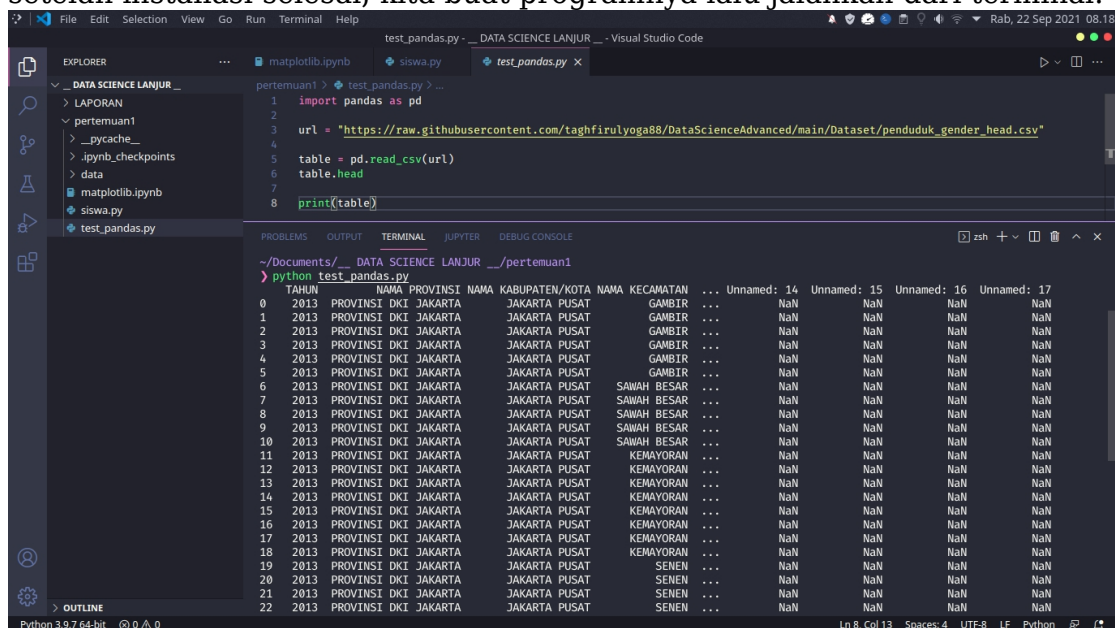
dalam praktikum ini sama halnya dengan paktik sebelumnya, hanya saja file csv yang akan kita buka terdapat diinternet. jadi pastikan kita terhubung ke internet. dan pastikan juga kita install library pandas.

```
pertemuan1: zsh — Konsole

~/Documents/___ DATA SCIENCE LANJUR ___/pertemuan1
> pip install pandas
Defaulting to user installation because normal site-packages is not writeable
Collecting pandas
  Downloading pandas-1.3.3-cp39-cp39-manylinux_2_17_x86_64.manylinux2014_x86_64.whl (11.5 MB)
    | 11.5 MB 387 kB/s
Collecting numpy>=1.17.3
  Downloading numpy-1.21.2-cp39-cp39-manylinux_2_12_x86_64.manylinux2010_x86_64.whl (15.8 MB)
    | 15.8 MB 53 kB/s
Requirement already satisfied: python-dateutil>=2.7.3 in /home/halim/.local/lib/python3.9/site-packages (from
pandas) (2.8.1)
Collecting pytz>=2017.3
  Downloading pytz-2021.1-py2.py3-none-any.whl (510 kB)
    | 510 kB 106 kB/s
Requirement already satisfied: six>=1.5 in /usr/lib/python3.9/site-packages (from python-dateutil>=2.7.3->pan
das) (1.16.0)
Installing collected packages: pytz, numpy, pandas
Successfully installed numpy-1.21.2 pandas-1.3.3 pytz-2021.1

~/Documents/___ DATA SCIENCE LANJUR ___/pertemuan1 2m 22s
>
```

setelah instalasi selesai, kita buat programnya lalu jalankan dari terminal.



The screenshot shows the Visual Studio Code interface. The Explorer panel on the left shows the file structure of the project. The main editor area displays a Python script named `test_pandas.py`. The script imports `pandas` as `pd`, defines a URL to a CSV file, reads the CSV into a DataFrame, and prints the first few rows. The Output panel at the bottom shows the execution of the script, displaying the first 22 rows of the dataset. The dataset has columns for year, name, province, district, and sub-district, followed by several unnamed columns containing NaN values.

```
test_pandas.py - DATA SCIENCE LANJUR - Visual Studio Code

EXPLORER
  DATA SCIENCE LANJUR
    LAPORAN
    pertemuan1
    __pycache__
    .ipynb_checkpoints
    matplotlib.ipynb
    siswa.py
    test_pandas.py

test_pandas.py
1 import pandas as pd
2
3 url = "https://raw.githubusercontent.com/tagfirulyoga88/DataScienceAdvanced/main/Dataset/penduduk_gender_head.csv"
4
5 table = pd.read_csv(url)
6 table.head()
7
8 print(table)

PROBLEMS OUTPUT TERMINAL JUPYTER DEBUG CONSOLE

~/Documents/___ DATA SCIENCE LANJUR ___/pertemuan1
> python test_pandas.py
TAHUN  NAMA  PROVINSI  NAMA  KABUPATEN/KOTA  NAMA  KECAMATAN  ...  Unnamed: 14  Unnamed: 15  Unnamed: 16  Unnamed: 17
0  2013  PROVINSI DKI JAKARTA  JAKARTA  PUSAT  GAMBIR  ...  NaN  NaN  NaN  NaN
1  2013  PROVINSI DKI JAKARTA  JAKARTA  PUSAT  GAMBIR  ...  NaN  NaN  NaN  NaN
2  2013  PROVINSI DKI JAKARTA  JAKARTA  PUSAT  GAMBIR  ...  NaN  NaN  NaN  NaN
3  2013  PROVINSI DKI JAKARTA  JAKARTA  PUSAT  GAMBIR  ...  NaN  NaN  NaN  NaN
4  2013  PROVINSI DKI JAKARTA  JAKARTA  PUSAT  GAMBIR  ...  NaN  NaN  NaN  NaN
5  2013  PROVINSI DKI JAKARTA  JAKARTA  PUSAT  GAMBIR  ...  NaN  NaN  NaN  NaN
6  2013  PROVINSI DKI JAKARTA  JAKARTA  PUSAT  SAWAH BESAR  ...  NaN  NaN  NaN  NaN
7  2013  PROVINSI DKI JAKARTA  JAKARTA  PUSAT  SAWAH BESAR  ...  NaN  NaN  NaN  NaN
8  2013  PROVINSI DKI JAKARTA  JAKARTA  PUSAT  SAWAH BESAR  ...  NaN  NaN  NaN  NaN
9  2013  PROVINSI DKI JAKARTA  JAKARTA  PUSAT  SAWAH BESAR  ...  NaN  NaN  NaN  NaN
10 2013  PROVINSI DKI JAKARTA  JAKARTA  PUSAT  SAWAH BESAR  ...  NaN  NaN  NaN  NaN
11 2013  PROVINSI DKI JAKARTA  JAKARTA  PUSAT  KEMAYORAN  ...  NaN  NaN  NaN  NaN
12 2013  PROVINSI DKI JAKARTA  JAKARTA  PUSAT  KEMAYORAN  ...  NaN  NaN  NaN  NaN
13 2013  PROVINSI DKI JAKARTA  JAKARTA  PUSAT  KEMAYORAN  ...  NaN  NaN  NaN  NaN
14 2013  PROVINSI DKI JAKARTA  JAKARTA  PUSAT  KEMAYORAN  ...  NaN  NaN  NaN  NaN
15 2013  PROVINSI DKI JAKARTA  JAKARTA  PUSAT  KEMAYORAN  ...  NaN  NaN  NaN  NaN
16 2013  PROVINSI DKI JAKARTA  JAKARTA  PUSAT  KEMAYORAN  ...  NaN  NaN  NaN  NaN
17 2013  PROVINSI DKI JAKARTA  JAKARTA  PUSAT  KEMAYORAN  ...  NaN  NaN  NaN  NaN
18 2013  PROVINSI DKI JAKARTA  JAKARTA  PUSAT  KEMAYORAN  ...  NaN  NaN  NaN  NaN
19 2013  PROVINSI DKI JAKARTA  JAKARTA  PUSAT  SENEN  ...  NaN  NaN  NaN  NaN
20 2013  PROVINSI DKI JAKARTA  JAKARTA  PUSAT  SENEN  ...  NaN  NaN  NaN  NaN
21 2013  PROVINSI DKI JAKARTA  JAKARTA  PUSAT  SENEN  ...  NaN  NaN  NaN  NaN
22 2013  PROVINSI DKI JAKARTA  JAKARTA  PUSAT  SENEN  ...  NaN  NaN  NaN  NaN
```

dari output yang ada sudah berupa tabel lagi, jadi secara tampilah lebih mudah dipahami dari sebelumnya.

Praktik 3

pada praktik ini karena akan menampilkan cart, oleh karena itu kita membuatnya menggunakan jupyter, atau file ipynb. atau kita juga bisa membuatnya di google colab untuk mempermudah. disini saya mengguankan VSCODE jadi membuat file .ipynb

sebelum memulai membuat codenya, pastikan kita sudah menginstall numpy dan matplotlib. matplotlib disini digunakan untuk mengenerate cart-nya

```
pertemuan1: zsh — Konsole

~/Documents/___ DATA SCIENCE LANJUR ___/pertemuan1
> pip install matplotlib
Defaulting to user installation because normal site-packages is not writeable
Collecting matplotlib
  Downloading matplotlib-3.4.3-cp39-cp39-manylinux1_x86_64.whl (10.3 MB)
    | 10.3 MB 170 kB/s
Requirement already satisfied: pyparsing>=2.2.1 in /usr/lib/python3.9/site-packages (from matplotlib) (2.4.7)
Collecting kiwisolver>=1.0.1
  Downloading kiwisolver-1.3.2-cp39-cp39-manylinux_2_12_x86_64.manylinux2010_x86_64.whl (1.6 MB)
    | 1.6 MB 110 kB/s
Requirement already satisfied: pillow>=6.2.0 in /usr/lib/python3.9/site-packages (from matplotlib) (8.3.2)
Requirement already satisfied: numpy>=1.16 in /home/halim/.local/lib/python3.9/site-packages (from matplotlib) (1.21.2)
Collecting cycler>=0.10
  Downloading cycler-0.10.0-py2.py3-none-any.whl (6.5 kB)
Requirement already satisfied: python-dateutil>=2.7 in /home/halim/.local/lib/python3.9/site-packages (from m
atplotlib) (2.8.1)
Requirement already satisfied: six in /usr/lib/python3.9/site-packages (from cycler>=0.10->matplotlib) (1.16.
0)
Installing collected packages: kiwisolver, cycler, matplotlib
Successfully installed cycler-0.10.0 kiwisolver-1.3.2 matplotlib-3.4.3

~/Documents/___ DATA SCIENCE LANJUR ___/pertemuan1 2m 46s
> pip install numpy
Defaulting to user installation because normal site-packages is not writeable
Requirement already satisfied: numpy in /home/halim/.local/lib/python3.9/site-packages (1.21.2)

~/Documents/___ DATA SCIENCE LANJUR ___/pertemuan1
> |
```

lalu membuat kodenya

```
import pandas as pd
import numpy as np
import matplotlib.pyplot as plt

url =
"https://raw.githubusercontent.com/taghfirulyoga88/DataScienceAdvanced/main/Dataset/penduduk_gender_head.csv"

table = pd.read_csv(url)
table.head()

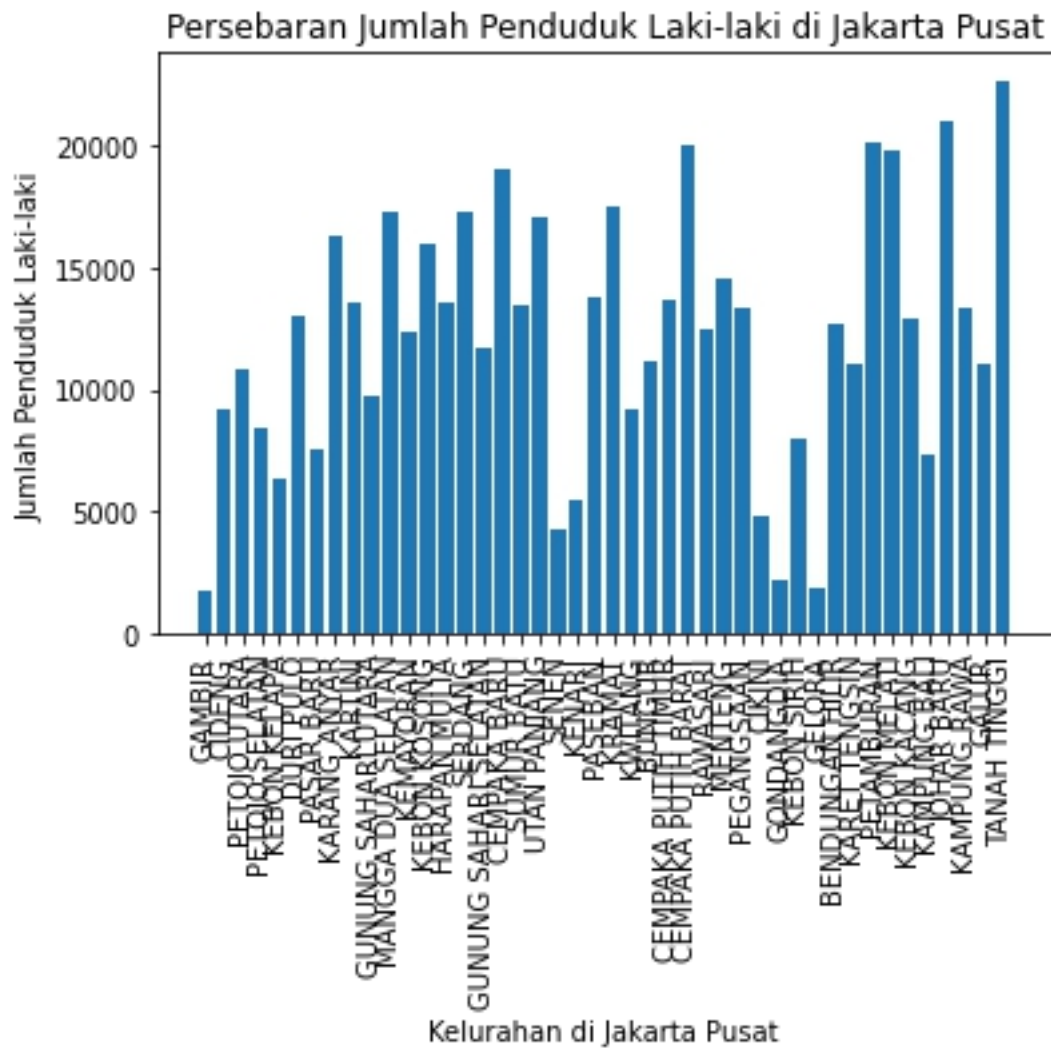
x_label = table['NAMA KELURAHAN']

plt.bar(x = np.arange(len(x_label)), height = table['LAKI-LAKI WNI'])
plt.xticks(np.arange(len(x_label)), table['NAMA KELURAHAN'], rotation = 90)
plt.xlabel("Kelurahan di Jakarta Pusat")
plt.ylabel("Jumlah Penduduk Laki-laki")
plt.title("Persebaran Jumlah Penduduk Laki-laki di Jakarta Pusat")

plt.show()
```

dari kode diatas kita akan mengambil list kelurahan dan jumlah laki-laki wni, lalu digenerate menjadi bar cart dan diberikan index nama kelurahan menggunakan .xticks, setelahnya memberikan label pada bagian bawah dan sampling, terakhir memberikan titlenya.

dari kode diatas akan mendapatkan hasil akhir seperti ini.



Cukup sekian lapran ini, kurang lebihnya mohon maaf.
terimakasih untuk materi dan perhatiannya.