

Tugas Matakuliah Topik Dalam Pengenalan Pola
Dosen : Dr. Toto Haryanto, S.Kom., M.Si



Oleh:

NIM : G6601231011
Nama : Irmawati Carolina

PROGRAM STUDI DOKTORAL ILMU KOMPUTER
SEKOLAH PASCASARJANA
INSTITUT PERTANIAN BOGOR
2023

PERTEMUAN 1 & 2

REPRESENTASI DATA DALAM PYTHON SEBAGAI PENDAHULUAN PADA PENGENALAN POLA

TUJUAN PRAKTIKUM

Mahasiswa mampu menterjemahkan representasi data dalam Python dan manipulasinya sehingga bisa menjadi input bagi sistem pengenalan Pola

Review – Representasi Data dalam Python

1. Vektor

Deklarasi variabel diawali dengan

Vektor dapat dituliskan dengan dan tanpa bantuan library

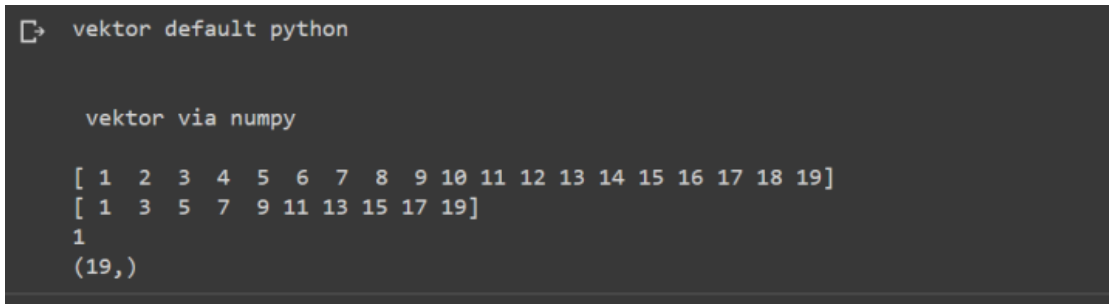
```
# vektor python numpy with range value
import numpy as np
print("vektor default python\n")
a = np.arange(1,20,1)
b = np.arange(1,20,2)
```

```
import numpy as np
print (" \n vektor via numpy \n")
```

```
# vektor via numpy
c = np.array ([1,2,3,4,5])
d = np.array ([1.5, 2.5, 5, 6, 7])
```

```
print(a)
print(b)
print(a.ndim)
print(a.shape)
```

Penyelesaian:



```
vektor default python

vektor via numpy

[ 1  2  3  4  5  6  7  8  9 10 11 12 13 14 15 16 17 18 19]
[ 1  3  5  7  9 11 13 15 17 19]
1
(19,)
```

Penjelasan :

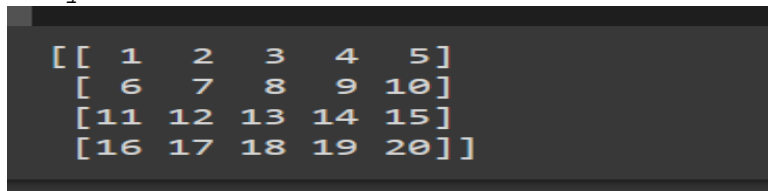
- Dalam bagian pertama, Anda menggunakan `np.arange()` untuk membuat vektor `a` dan `b`. Vektor `a` memiliki nilai dari 1 hingga 19 dengan kenaikan 1, sedangkan vektor `b` memiliki nilai dari 1 hingga 19 dengan kenaikan 2.
- Pada bagian kedua, Anda menggunakan `np.array()` untuk membuat vektor `c` dan `d`. Vektor `c` memiliki elemen bilangan bulat dari 1 hingga 5, sedangkan vektor `d` memiliki elemen float dari 1.5 hingga 7.
- Fungsi `print()` digunakan untuk mencetak nilai dari setiap vektor dan beberapa informasi tambahan seperti dimensi (`ndim`) dan bentuk (`shape`) dari vektor `a`

2. Matrix

Matriks adalah basic 2D table dari data dan dapat berisi nilai numerik dan/atau karakter. Dapat dibuat dengan cara sederhana dengan membuat urutan dari vektor, mengubah dari vektor atau membaca dari file

```
# mengubah dari 1D menjadi matrik 2D
a = np.arange(1,21,1)
c = a.reshape((4,5))
print(c)
```

Penyelesaian:



```
[[ 1  2  3  4  5]
 [ 6  7  8  9 10]
 [11 12 13 14 15]
 [16 17 18 19 20]]
```

Penjelasan:

- Anda membuat array 1D `a` dengan elemen dari 1 hingga 20.
- Kemudian, Anda menggunakan fungsi `reshape()` untuk mengubah array 1D `a` menjadi matriks 2D dengan ukuran (4, 5), artinya matriks ini memiliki 4 baris dan 5 kolom.
- Output menunjukkan matriks 2D yang dibentuk dari array `a` dengan susunan elemen seperti yang terlihat di atas.

3. List

List merupakan representasi struktur data yang dapat menyimpan data dengan nilai numerik, karakter, dan lain-lain secara bersamaan

```
list1 = ["apple", "banana", "cherry"]
list2 = [1, 5, 7, 9, 3]
list3 = [True, False, False]
list4 = ["abc", 34, True, 40, "male"]
```

```
print(list1);
```

Penyelesaian:

```
> ['apple', 'banana', 'cherry']
```

Penjelasan:

- Anda memiliki beberapa variabel yang berisi daftar (list) dengan tipe data yang berbeda-beda.
- Pada baris terakhir, Anda menggunakan perintah `print(list1)` untuk mencetak isi dari variabel `list1`.
- Output menunjukkan isi dari list `list1`, yaitu tiga elemen string: "apple", "banana", dan "cherry".

4. data.Frame

Data frame adalah spesialisasi dari tipe list untuk menyimpan vektor ke dalam bentuk frame (menyerupai basis data). Kelebihan dibandingkan matriks adalah dapat dimanipulasi dalam berbagai bentuk dan cara.

```
#data frame
import pandas as pd
df = pd.DataFrame(np.array([[1, 2, 3], [4, 5, 6], [7, 8, 9]]),
                  columns=['a', 'b', 'c'])

print(df)
```

Penyelesaian:

```
>      a  b  c
0     1  2  3
1     4  5  6
2     7  8  9
```

Import dan Export data serta manipulasinya

1. Import Data (loading data in Python: csv, text, XML, image)

Buat file dengan header, misalkan terdiri dari Nama Pasisen, Umur, Penyakit, Gender seperti pada tabel berikut dan simpan kedalam format CSV dan Text (Tab delimiter)

Nama Pasien	Umur	Gender	Diagnosa Sakit
Anto	24	L	Tidak
Budi	35	L	Ya
Adi	55	L	Ya
Delima	32	P	Ya
Dodi	21	L	Tidak
Tukiyem	19	P	Tidak
Rama	23	L	Tidak
Santi	35	P	Tidak
Mery	44	P	Ya
Yanti	27	P	Tidak
Parto	43	L	Ya
Dea	24	P	Tidak

Kemudian lakukan pembacaan untuk file tersebut:

```
# Membaca data dari file dengan format CSV
import pandas as pd
data = pd.read_csv("Data.csv", sep=";")
print(data)

# Membaca data dari file dengan format text (delimiter)
print("\n read text data with tab delimiter")
with open ('Data.txt') as data:
    print(data.read())

# Membaca data dari URL
import pandas as pd
f = pd.read_csv('http://www.exploredata.net/ftp/Spellman.csv')
print(f)
```

```

    Nama Pasien  Umur  Gender  Diagnosa  Sakit
0         Anto    24        L      Tidak
1         Budi    35        L         Ya
2         Adi    55        L         Ya
3        Delima    32        P         Ya
4         Dodi    21        L      Tidak
5        Tukiyem    19        P      Tidak
6         Rama    23        L      Tidak
7         Santi    35        P      Tidak
8         Mery    44        P         Ya
9         Yanti    27        P      Tidak
10        Parto    43        L         Ya
11         Dea    24        P      Tidak

read text data with tab delimiter
Nama Pasien,Umur,Gender,Diagnosa Sakit
Anto,24,L,Tidak
Budi,35,L,Ya
Adi,55,L,Ya
Delima,32,P,Ya
Dodi,21,L,Tidak
Tukiyem,19,P,Tidak
Rama,23,L,Tidak
Santi,35,P,Tidak
Mery,44,P,Ya
Yanti:27:P:Tidak
```

```

Parto,43,L,Ya
Dea,24,P,Tidak

time 40 50 60 70 80 90 100 110 120 ... \
0 YAL001C -0.070 -0.23 -0.100 0.03 -0.04 -0.12 -0.28 -0.44 -0.09 ...
1 YAL014C 0.215 0.09 0.025 -0.04 -0.04 -0.02 -0.51 -0.08 0.00 ...
2 YAL016W 0.150 0.15 0.220 0.29 -0.10 0.15 -0.73 0.19 -0.15 ...
3 YAL020C -0.350 -0.28 -0.215 -0.15 0.16 -0.12 0.26 0.00 0.13 ...
4 YAL022C -0.415 -0.59 -0.580 -0.57 -0.09 -0.34 0.49 0.32 1.15 ...
...
4376 YPR198W -0.060 0.08 0.210 0.34 0.65 -0.26 0.14 -0.33 0.53 ...
4377 YPR199C 0.155 0.19 0.235 0.28 -0.26 0.21 -0.40 0.34 -0.80 ...
4378 YPR201W -0.255 -0.36 -0.300 -0.24 1.30 -0.07 0.29 -0.20 0.25 ...
4379 YPR203W 0.570 0.12 -0.070 -0.26 -0.44 -0.21 -1.08 0.39 -0.17 ...
4380 YPR204W 0.405 0.17 -0.045 -0.26 -0.60 -0.09 -0.85 0.17 -0.05 ...

time 170 180 190 200 210 220 230 240 250 260
0 0.59 0.34 -0.28 -0.09 -0.44 0.31 0.03 0.57 0.00 0.010
1 -0.30 -0.38 0.07 -0.04 0.13 -0.06 -0.26 -0.10 0.27 0.235
2 0.12 -0.17 0.11 -0.15 0.03 -0.26 -0.34 -0.34 0.25 0.190
3 0.07 0.61 -0.20 0.49 -0.43 0.80 -0.47 1.01 -0.36 -0.405
4 -0.48 -0.40 -0.59 0.54 -0.09 1.03 0.08 0.57 -0.26 -0.310
...
4376 0.14 -0.64 -0.26 0.53 -0.17 0.59 -0.96 0.40 -0.23 -0.325
4377 0.34 0.15 0.30 -0.06 0.13 -0.44 -1.03 0.14 0.30 0.250

```

Membaca file dan menyajikan dalam bentuk grafik.

Untuk dapat menyajikan dalam bentuk grafik terlebih dahulu melakukan instalasi paket matplotlib.

```

import numpy as np

traffic = sp.genfromtxt("web_traffic.tsv",delimiter='\t')
print(traffic[:10])
print(traffic.shape)

x = traffic[:,0]
y = traffic[:,1]

x = x[~sp.isnan(y)]
y = y[~sp.isnan(y)]

import matplotlib.pyplot as plt
plt.scatter(x,y)
plt.title("Web traffic last month")
plt.xlabel("Time")

```

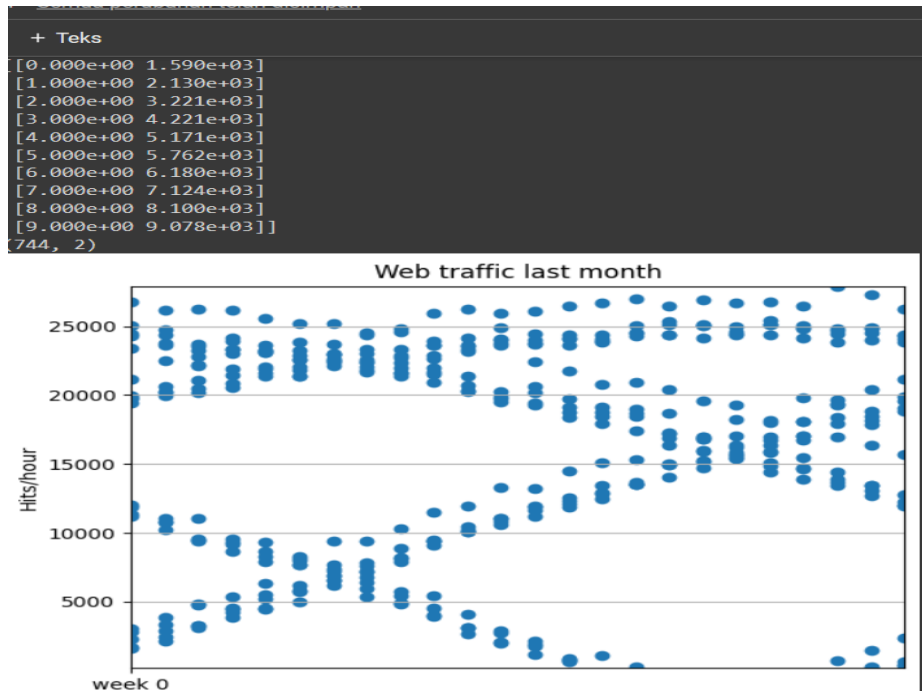
```
plt.ylabel("Hits/hour")

plt.xticks([w*7*24 for w in range(10)],['week %i' %w for w in
range(10)])

plt.autoscale(tight=True)

plt.grid()
```

lakukan perintah di atas dan perhatikan hasilnya kemudian tunjukkan pada kolom berikut



```
# Operasi image pada Python
# Instalasi paket open cv
pip install opencv-contrib-python

# pilihan load image (contoh logo ipb)
import matplotlib.pyplot as plt
import cv2
import numpy as np

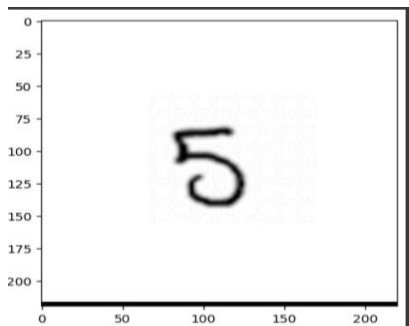
print("read images using opencv")
five = cv2.imread("5.png")
print(five.shape)
print(five.size)
plt.imshow(five)
cv2.waitKey(0)
```




```
# Operasi image pada Python
# Instalasi paket open cv pip install opencv-contrib-python

# pilihan load image (contoh angka 5)
import matplotlib.pyplot as plt
import cv2
import numpy as np

print("read images using opencv")
five = cv2.imread("5.png")
print(five.shape)
print(five.size)
plt.imshow(five)
cv2.waitKey(0)
```

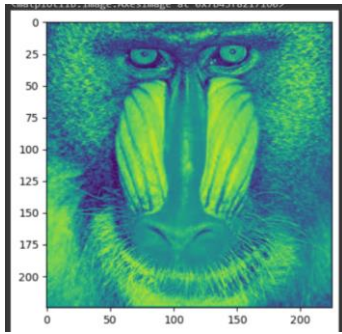


konversi image

```
import cv2

babon =
cv2.imread("babon.jpg")
babon_gray =
cv2.cvtColor(babon,
cv2.COLOR_BGR2GRAY)

plt.imshow(babon)
plt.imshow(babon_gray)
```



```
# mengambil nilai matriksnya
```

```
# acces pixel of images per postion  
pixels = five[100,100]  
print(pixels)
```

```
[ 229  229  229]
```

TUGAS PRAKTIKUM

Pelajari secara mandiri terkait dengan Python pengolahan dasar mulai dari text ataupun citra. Lakukan operasi-operasi dasar teks dan citra tersebut!
Buat dalam repostori github untuk masing-masing yang sudah Anda lakukan. Link github dapat dishare via newlms.

LINK Github Hasil dari pengerjaan:

<https://github.com/IrmawatiCarolina/Tugaspraktekpola>

DAFTAR PUSTAKA

1. Richert W & Coelho LP. *Buuildng Machine Learning System with Python*. 2013. Packt Publising. Birmingham, UK.