

Tugas LAB WEEK 5 ASYNCHRON Christopher Darren

Jupyter 00000054804_Christopher Darren_Week5 Last Checkpoint: 13 minutes ago (unsaved changes) Logout

File Edit View Insert Cell Kernel Widgets Help Trusted Python 3 (ipykernel) C

Statistic for Data Science ¶

Scikit-Learn

```
In [1]: 1 #Import Libraries
        2
        3 import numpy as np
        4 import pandas as pd

In [2]: 1 #import dataset iris
        2 from sklearn.datasets import load_iris

In [3]: 1 #read data
        2 iris = load_iris()
        3 print(iris.data.shape)

(150, 4)

In [4]: 1 #jumlah baris dan kolomnya: 150, dan 4

In [5]: 1 #mengubah data frame agar lebih mudah melakukan memproses data
        2
        3 dataset_iris = load_iris(as_frame=True)

In [6]: 1 dataset_iris.data.info()

<class 'pandas.core.frame.DataFrame'>
RangeIndex: 150 entries, 0 to 149
Data columns (total 4 columns):
```

Gambar 1.

```
# Column Non-Null Count Dtype
---
0 sepal length (cm) 150 non-null float64
1 sepal width (cm) 150 non-null float64
2 petal length (cm) 150 non-null float64
3 petal width (cm) 150 non-null float64
dtypes: float64(4)
memory usage: 4.8 KB

In [7]: 1 #melihat description
        2 dataset_iris.data.describe()

Out[7]:
```

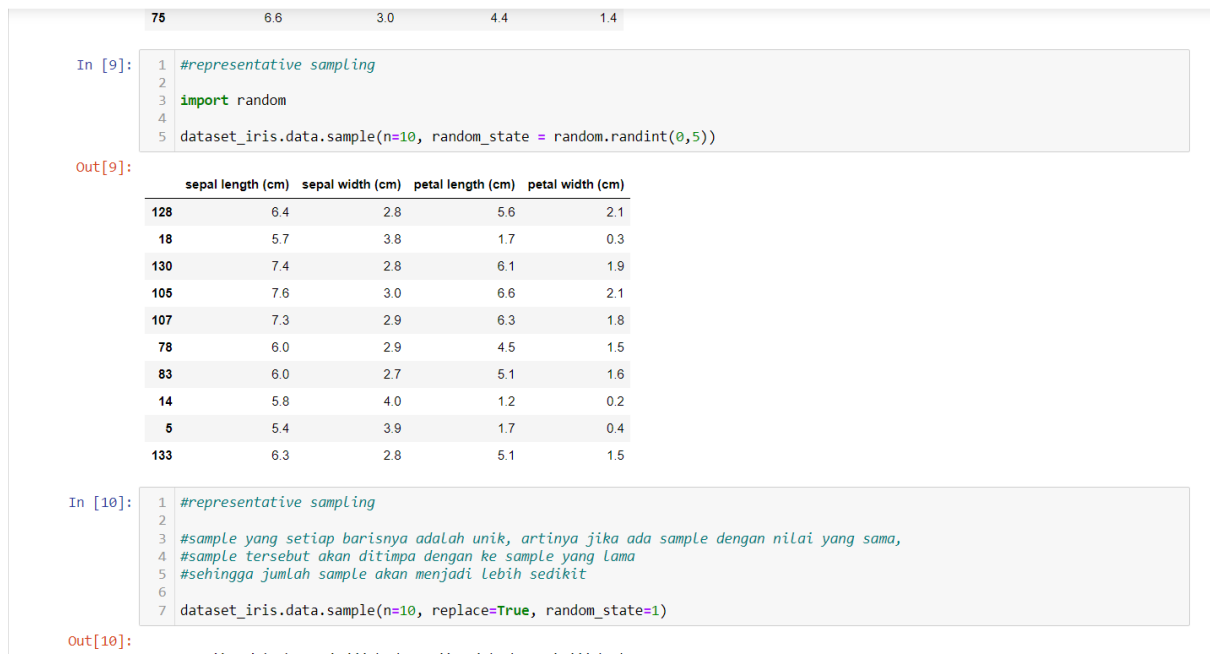
	sepal length (cm)	sepal width (cm)	petal length (cm)	petal width (cm)
count	150.000000	150.000000	150.000000	150.000000
mean	5.843333	3.057333	3.758000	1.199333
std	0.828066	0.435866	1.765298	0.762238
min	4.300000	2.000000	1.000000	0.100000
25%	5.100000	2.800000	1.600000	0.300000
50%	5.800000	3.000000	4.350000	1.300000
75%	6.400000	3.300000	5.100000	1.800000
max	7.900000	4.400000	6.900000	2.500000

```
In [8]: 1 #random sampling
        2 dataset_iris.data.sample(n=3, random_state=1)

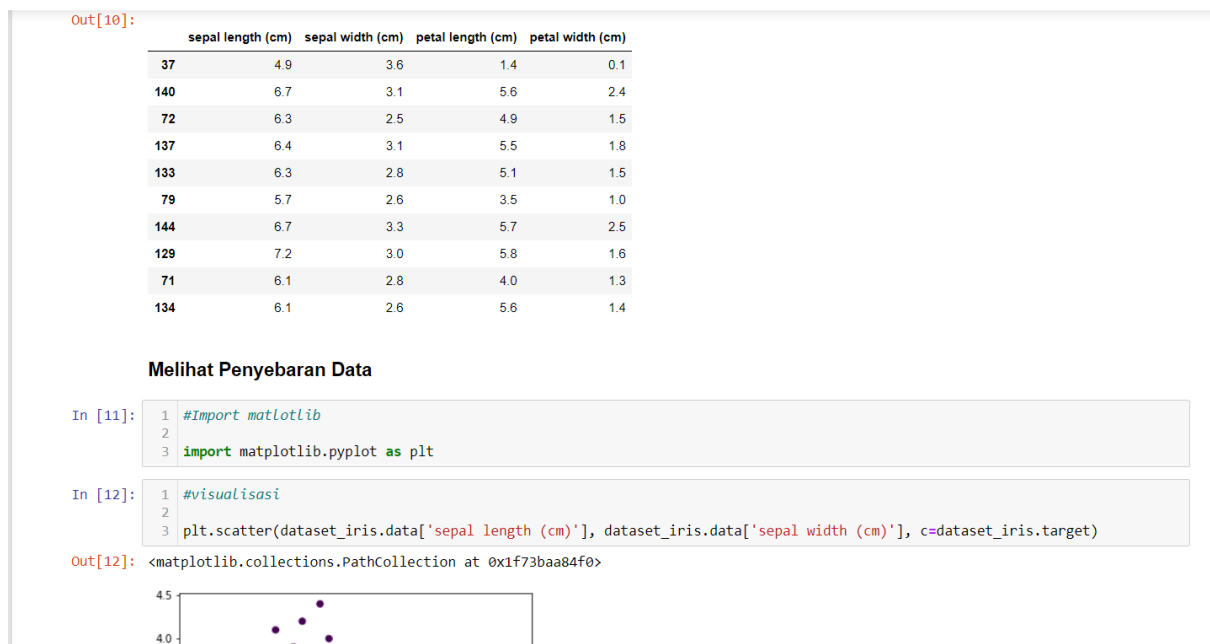
Out[8]:
```

	sepal length (cm)	sepal width (cm)	petal length (cm)	petal width (cm)
14	5.8	4.0	1.2	0.2
98	5.1	2.5	3.0	1.1

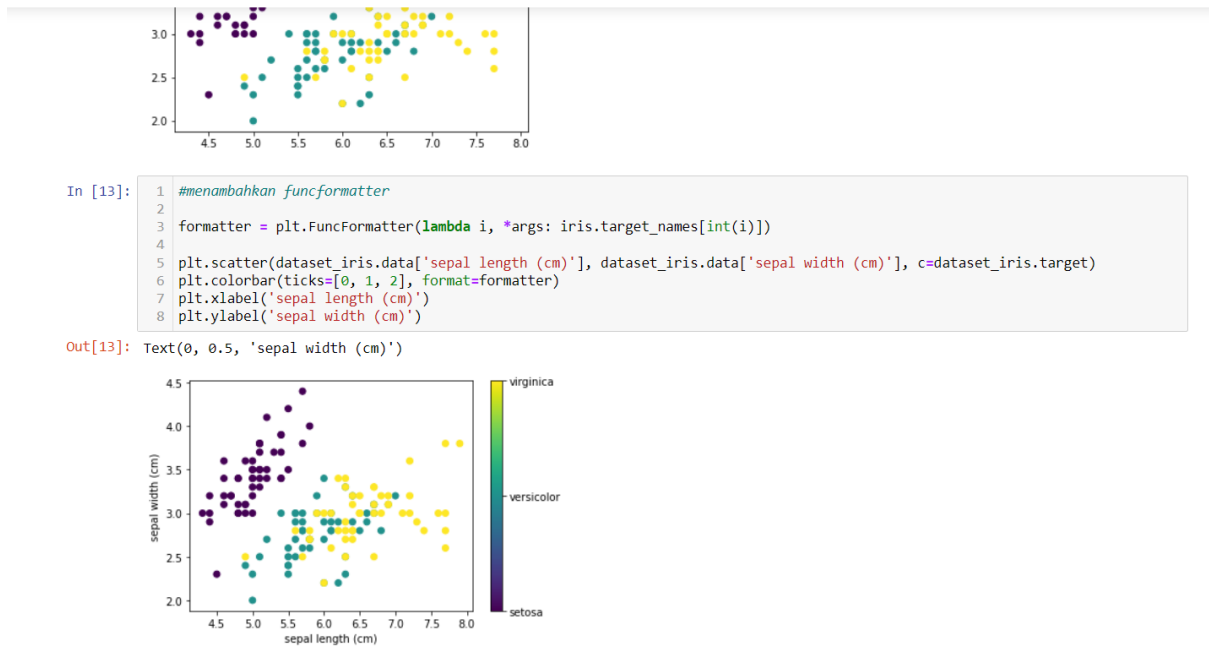
Gambar 2.



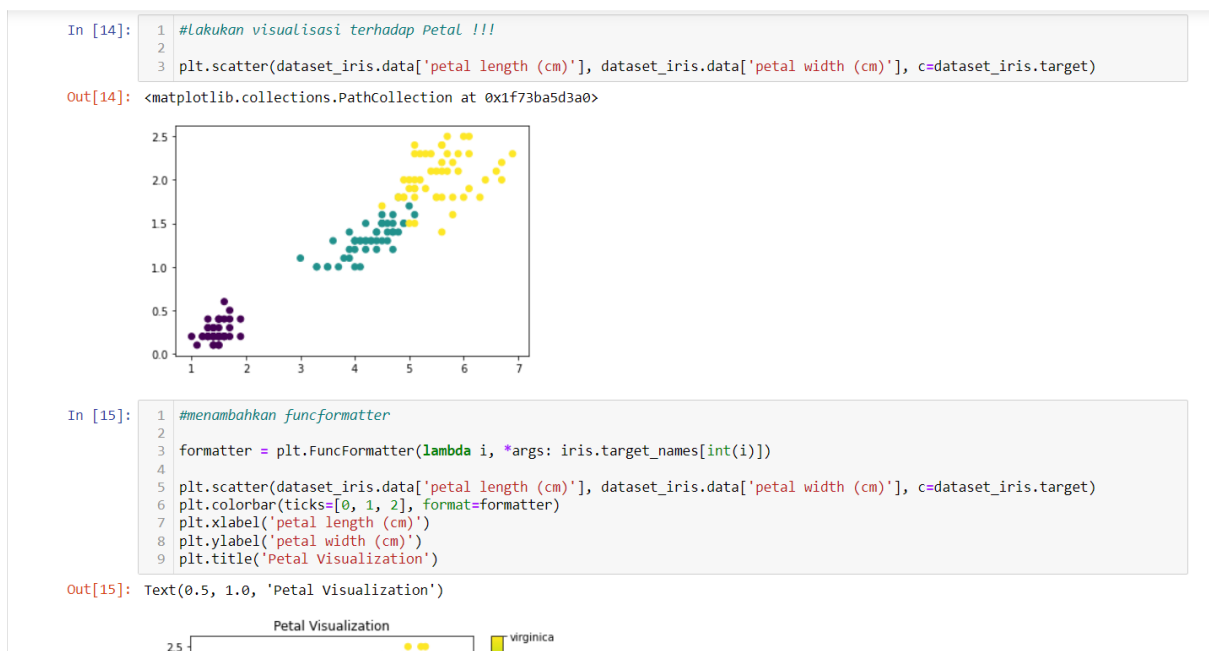
Gambar 3.



Gambar 4.



Gambar 5.



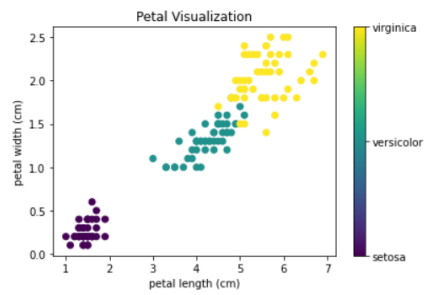
Gambar 6.

```

5 plt.scatter(dataset_iris.data['petal length (cm)'], dataset_iris.data['petal width (cm)'], c=dataset_iris.target)
6 plt.colorbar(ticks=[0, 1, 2], format=formatter)
7 plt.xlabel('petal length (cm)')
8 plt.ylabel('petal width (cm)')
9 plt.title('Petal Visualization')

```

Out[15]: Text(0.5, 1.0, 'Petal Visualization')



```

In [16]: 1 #testing data with Scikit-Learn
          2
          3 from sklearn.model_selection import train_test_split
          4 X_train, X_test, y_train, y_test = train_test_split(dataset_iris.data, dataset_iris.target, test_size=0.33, random_state=42)

```

Gambar 7.