

# IRIS FLOWER CLASSIFICATION



```

[2] import numpy as np
import matplotlib.pyplot as plt
import seaborn as sns
import pandas as pd
%matplotlib inline

```

```

columns = ['sepal_length', 'sepal_width', 'petal_length', 'petal_width', 'species']
#Load the data
df = pd.read_csv('iris.data', names=columns)
df.head()

```



	sepal_length	sepal_width	petal_length	petal_width	species
0	sepal_length	sepal_width	petal_length	petal_width	species
1	5.1	3.5	1.4	0.2	Iris-setosa
2	4.9	3	1.4	0.2	Iris-setosa
3	4.7	3.2	1.3	0.2	Iris-setosa
4	4.6	3.1	1.5	0.2	Iris-setosa

Next steps:

Generate code with df

View recommended plots

iris.data X

```

1 sepal_length,
2 5.1,3.5,1.4,0.
3 4.9,3,1.4,0.2
4 4.7,3.2,1.3,0.
5 4.6,3.1,1.5,0.
6 5,3.6,1.4,0.2
7 5.4,3.9,1.7,0.
8 4.6,3.4,1.4,0.
9 5,3.4,1.5,0.2
10 4.4,2.9,1.4,0.
11 4.9,3.1,1.5,0.
12 5.4,3.7,1.5,0.
13 4.8,3.4,1.6,0.
14 4.8,3,1.4,0.1
15 4.3,3,1.1,0.1
16 5.8,4,1.2,0.2
17 5.7,4.4,1.5,0.4
18 5.4,3.9,1.3,0.4
19 5.1,3.5,1.4,0.3
20 5.7,3.8,1.7,0.3
21 5.1,3.8,1.5,0.3
22 5.4,3.4,1.7,0.2
23 5.1,3.7,1.5,0.4
24 4.6,3.6,1,0.2,I
25 5.1,3.3,1.7,0.5
26 4.8,3.4,1.9,0.2

```

File Tools Help Cannot save changes



+ Code + Text

Copy to Drive

```
[5] import numpy as np
import matplotlib.pyplot as plt
import seaborn as sns
import pandas as pd
%matplotlib inline
```

```
columns = ['sepal_length', 'sepal_width', 'petal_length', 'petal_width', 'species']
#load the data
df = pd.read_csv('/content/IRIS.csv', names=columns)
df.head(150)
```

	sepal_length	sepal_width	petal_length	petal_width	species
0	sepal_length	sepal_width	petal_length	petal_width	species
1	5.1	3.5	1.4	0.2	iris-setosa
2	4.9	3	1.4	0.2	iris-setosa

IRIS.csv

sepal\_length

5.1

4.9

4.7

4.6

5

5.4

4.8

5



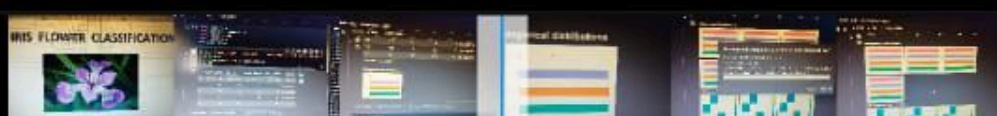
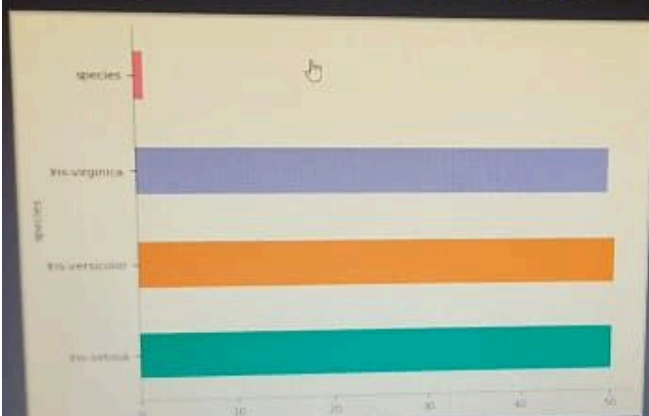


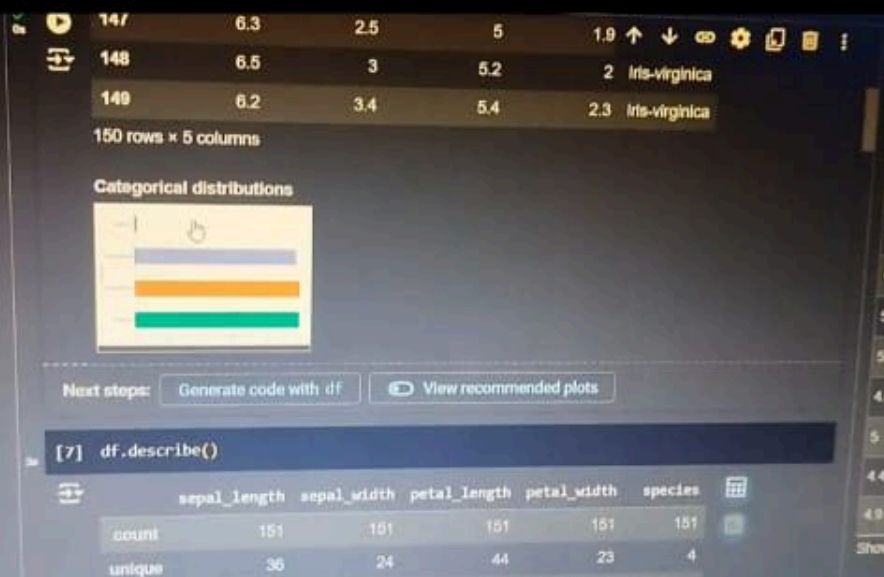
July 13, 2024

21:22



## Categorical distributions







July 13, 2024

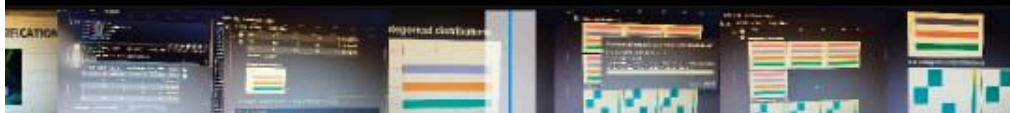
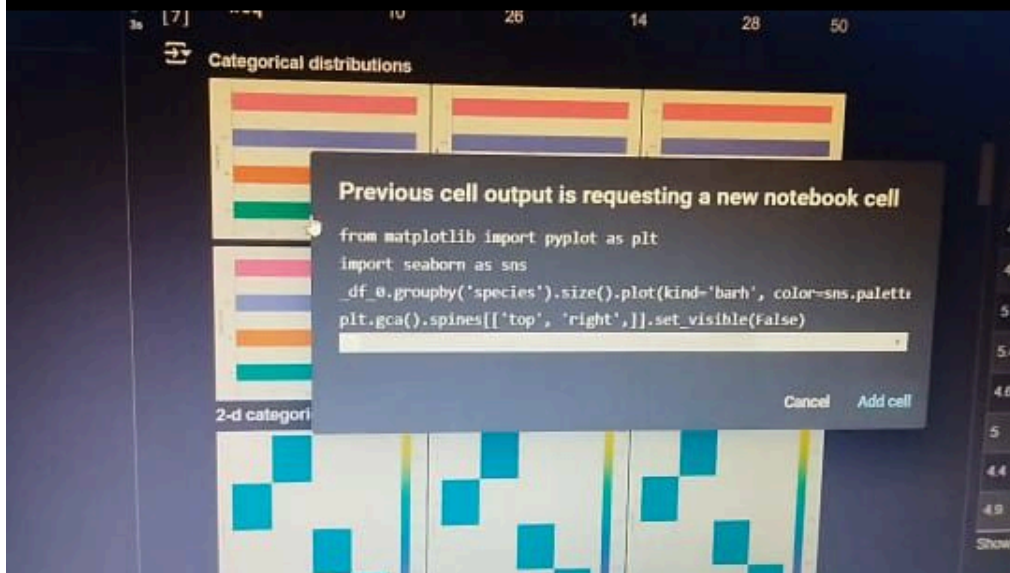
21:22



### Categorical distributions



← July 13, 2024  
21:22





July 13, 2024

21:22



# Separate features and target

data = df.values

X = data[:,0:4]

Y = data[:,4]

print(Y)



```
['species' 'Iris-setosa' 'Iris-setosa' 'Iris-setosa'
'Iris-setosa' 'Iris-setosa' 'Iris-setosa' 'Iris-setosa'
'Iris-setosa' 'Iris-setosa' 'Iris-setosa' 'Iris-setosa'
'Iris-setosa' 'Iris-setosa' 'Iris-setosa' 'Iris-setosa']
```





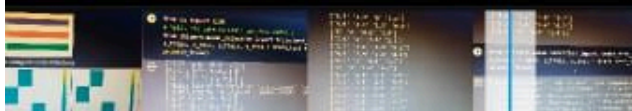


July 13, 2024

21:22



4:7	3:2	1:6	0:2
5:5	2:4	3:8	1:1
6:6	2:9	4:6	1:3
5:1	3:8	1:6	0:2
6:7	3:1	5:6	2:4
6:1	2:9	4:7	1:4
7:1	3:5	0:2	1:1
4:4	3:2	1:3	0:2
5:3	1:6	0:2	1:1
6:5	3:5	2:2	1:1
5:1	3:4	1:5	0:2
4:6	3:4	1:4	0:3
6:3	4:4	1:6	1:1
5:1	3:5	1:4	0:2





July 13, 2024

21:22

✓  
2s

```
from os import X_OK
# Split the data to train and test dataset
from sklearn.model_selection import train_test_split
x_train, x_test, y_train, y_test = train_test_split(
print(x_train)
```



```
['6' '2.2' '5' '1.5']
['6.3' '2.8' '5.1' '1.5']
['5.7' '2.8' '4.5' '1.3']
['6.3' '3.3' '4.7' '1.6']
['5.7' '2.8' '4.1' '1.3']
['6' '3.4' '4.5' '1.6']
['6.7' '3' '5' '1.7']
['7.7' '3.8' '6.7' '2.2']
```





July 13, 2024

21:22



```
['4.6' '3.1' '1.5' '0.2']  
['5.2' '3.4' '1.4' '0.2']  
['6.6' '3' '4.4' '1.4']]
```

18



```
from sklearn.model_selection import train_test_split  
x_train, x_test, y_train, y_test = train_test_split(X,  
print(y_train)
```



```
['Iris-versicolor' 'Iris-setosa' 'Iris-virginica' 'Iris-versicolor'  
'Iris-versicolor' 'Iris-versicolor' 'Iris-versicolor'  
'Iris-virginica' 'Iris-setosa' 'Iris-versicolor' 'Iris-versicolor'  
'Iris-virginica' 'Iris-versicolor' 'Iris-versicolor' 'Iris-versicolor'  
'Iris-setosa' 'Iris-virginica' 'Iris-setosa' 'Iris-setosa'  
'Iris-versicolor' 'Iris-setosa' 'Iris-setosa' 'Iris-versicolor']
```





July 13, 2024

21:22



```
[ 0. '3.4' '4.5' '1.6']  
['5.1' '3.5' '1.4' '0.2']]
```

0s

`df.isna().sum()`

```
sepal_length    0  
sepal_width     0  
petal_length    0  
petal_width     0  
species         0  
dtype: int64
```

