Lp1Da1

December 10, 2021

[59]:
```
import numpy as np
import pandas as pd
%matplotlib inline


import matplotlib.pyplot as plt
import seaborn as sns
```

[31]:
```
dat=pd.read_csv('Iris.csv')
```

[32]:
```
dat[0:10]
```

[32]:

	x1	x2	x3	x4	class
0	5.1	3.5	1.4	0.2	Iris-setosa
1	4.9	3.0	1.4	0.2	Iris-setosa
2	4.7	3.2	1.3	0.2	Iris-setosa
3	4.6	3.1	1.5	0.2	Iris-setosa
4	5.0	3.6	1.4	0.2	Iris-setosa
5	5.4	3.9	1.7	0.4	Iris-setosa
6	4.6	3.4	1.4	0.3	Iris-setosa
7	5.0	3.4	1.5	0.2	Iris-setosa
8	4.4	2.9	1.4	0.2	Iris-setosa
9	4.9	3.1	1.5	0.1	Iris-setosa

[38]:
```
dat.shape  #######how many features are there
list(dat.columns)
```

[38]:
```
['x1', 'x2', 'x3', 'x4', 'class']
```

[37]:
```
dat.dtypes  ##what are their types
```

[37]:
```
x1       float64
x2       float64
x3       float64
x4       float64
class     object
dtype: object
```

[39]:
```
dat['x1'].describe()    ###########statistics description for columns
```

[39]:
```
count    150.000000
mean       5.843333
std        0.828066
min        4.300000
25%        5.100000
50%        5.800000
75%        6.400000
max        7.900000
Name: x1, dtype: float64
```

[40]:
```
dat['x2'].describe()
```

[40]:
```
count    150.000000
mean       3.054000
std        0.433594
min        2.000000
25%        2.800000
50%        3.000000
75%        3.300000
max        4.400000
Name: x2, dtype: float64
```

[41]:
```
dat['x3'].describe()
```

[41]:
```
count    150.000000
mean       3.758667
std        1.764420
min        1.000000
25%        1.600000
50%        4.350000
75%        5.100000
max        6.900000
Name: x3, dtype: float64
```

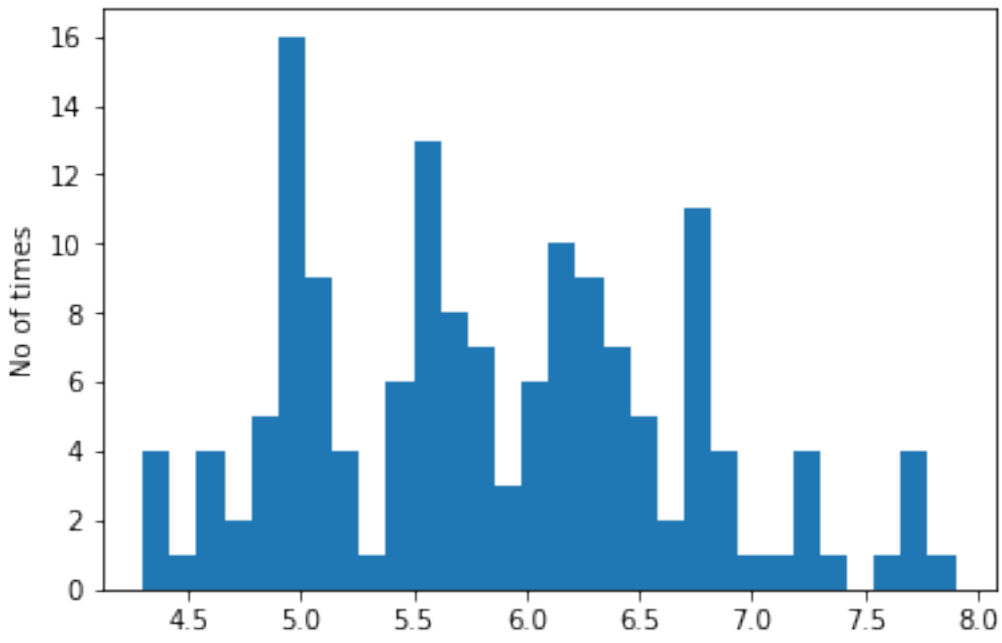
[42]:
```
dat['x4'].describe()
```

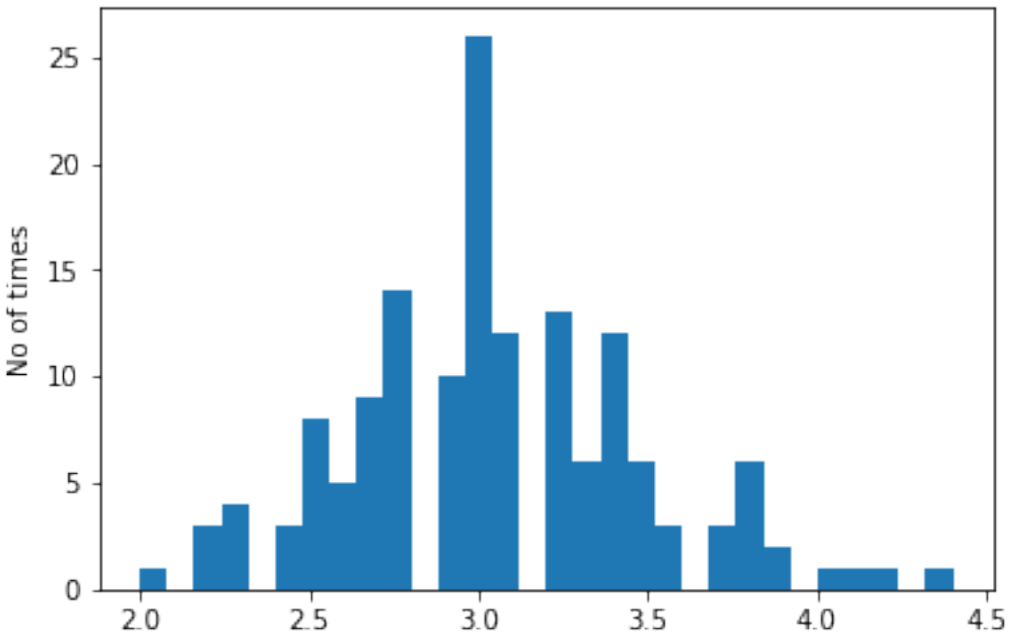
[42]:
```
count    150.000000
mean       1.198667
std        0.763161
min        0.100000
25%        0.300000
50%        1.300000
75%        1.800000
max        2.500000
Name: x4, dtype: float64
```

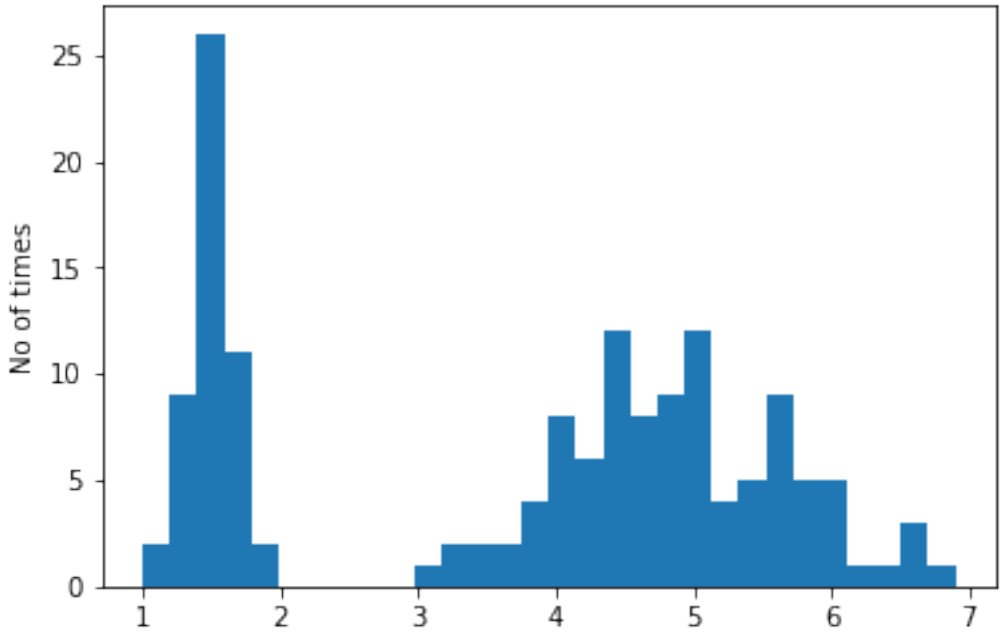
[43]:
```
dat['class'].describe()
```

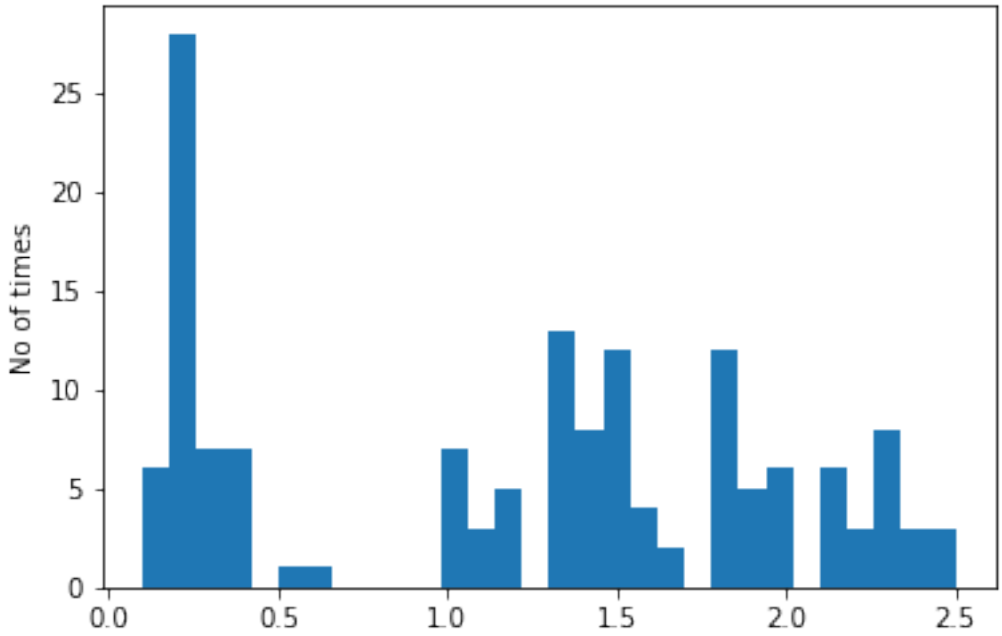
[43]:
```
count                150
unique                 3
top       Iris-virginica
freq                  50
Name: class, dtype: object
```

[58]:
```
plt.hist(dat['x1'],bins=30)            ###############plot histogram
plt.ylabel('No of times')
plt.show()


plt.hist(dat['x2'],bins=30)            ###############plot histogram
plt.ylabel('No of times')
plt.show()


plt.hist(dat['x3'],bins=30)            ###############plot histogram
plt.ylabel('No of times')
plt.show()


plt.hist(dat['x4'],bins=30)            ###############plot histogram
plt.ylabel('No of times')
plt.show()
```



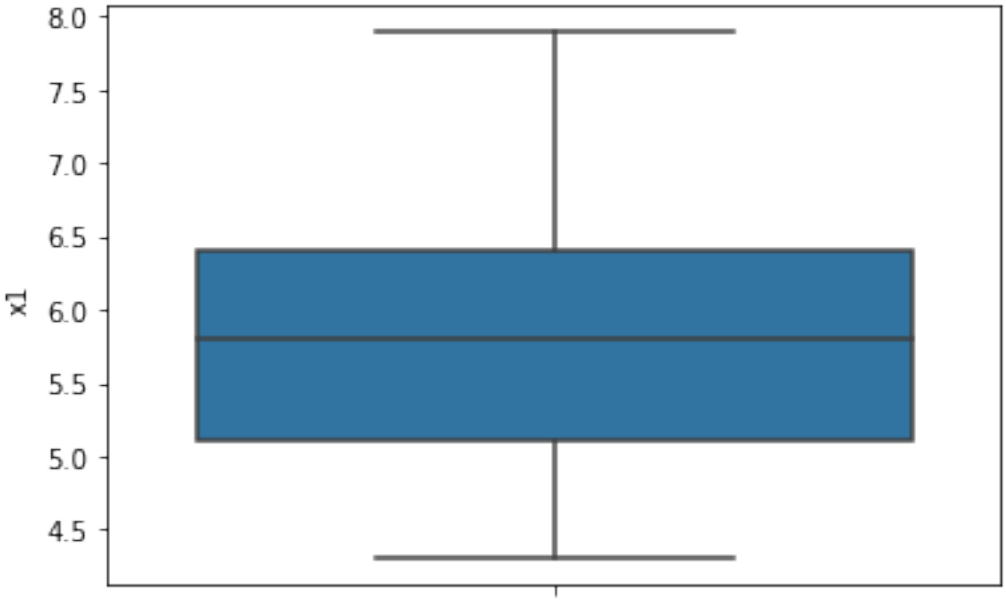

No of times
0
5
10
15
20
25
2.0
2.5
3.0
3.5
4.0
4.5


No of times
0
5
10
15
20
25
1
2
3
4
5
6
7



[65]:
```
########################box plot for single feature  same for rest

sns.boxplot(y=dat['x1'])
```

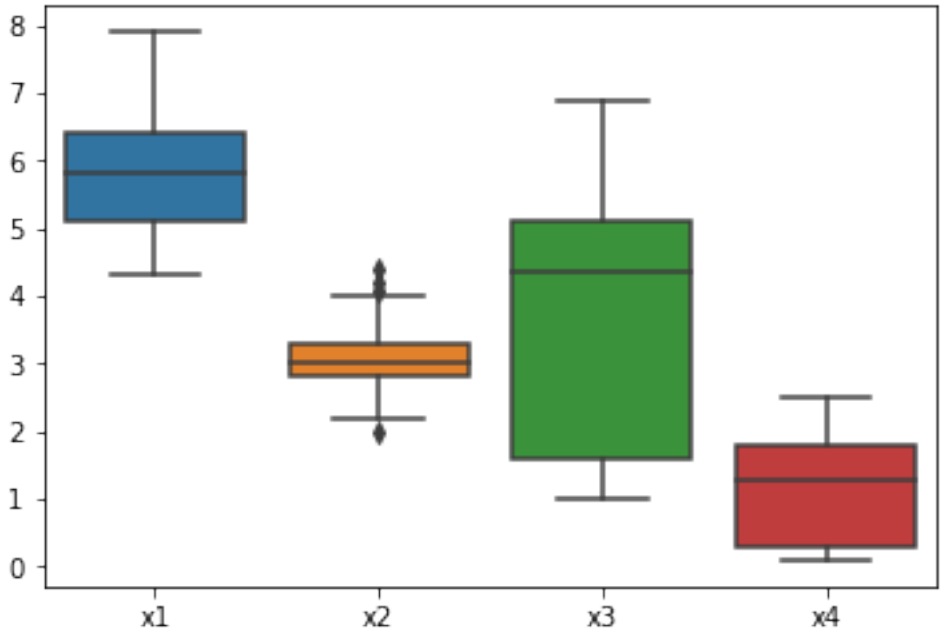
[65]: <matplotlib.axes._subplots.AxesSubplot at 0x7f2620394860>



[ ]:
```
sns.boxplot(x='class',y=dat['x2'])
```

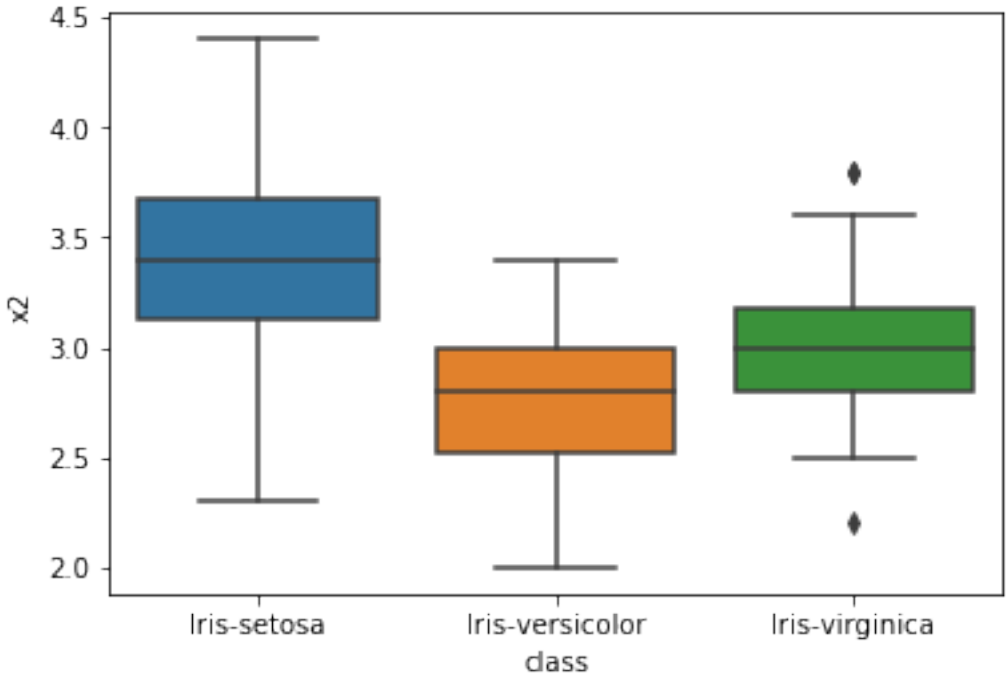
[67]:
```
sns.boxplot(data=dat.ix[:,0:4])    ##################for multiple
```

[67]:
```
<matplotlib.axes._subplots.AxesSubplot at 0x7f26203c4e80>
```



[70]:
```
sns.boxplot(x=dat['class'],y=dat['x2'])  #############one vs all
```

[70]:
```
<matplotlib.axes._subplots.AxesSubplot at 0x7f260f40d710>
```


4.5
4.0
3.5
3.0
2.5
2.0
x2
Iris-setosa
Iris-versicolor
Iris-virginica
class