

## Gaussian Blob

```
from sklearn.datasets.samples_generator import make_blobs
centers = [[2, 2], [4, 4], [6, 6], [8,8]]
X, y = make_blobs(n_samples=20, n_features=2, centers=centers, cluster_std=0.5, center_box=(1,
10.0), shuffle=True, random_state=0)

import matplotlib.pyplot as plt
# Plot the training points
plt.scatter(X[:, 0], X[:, 1])
plt.xlabel('X axis')
plt.ylabel('Y axis')

plt.show()

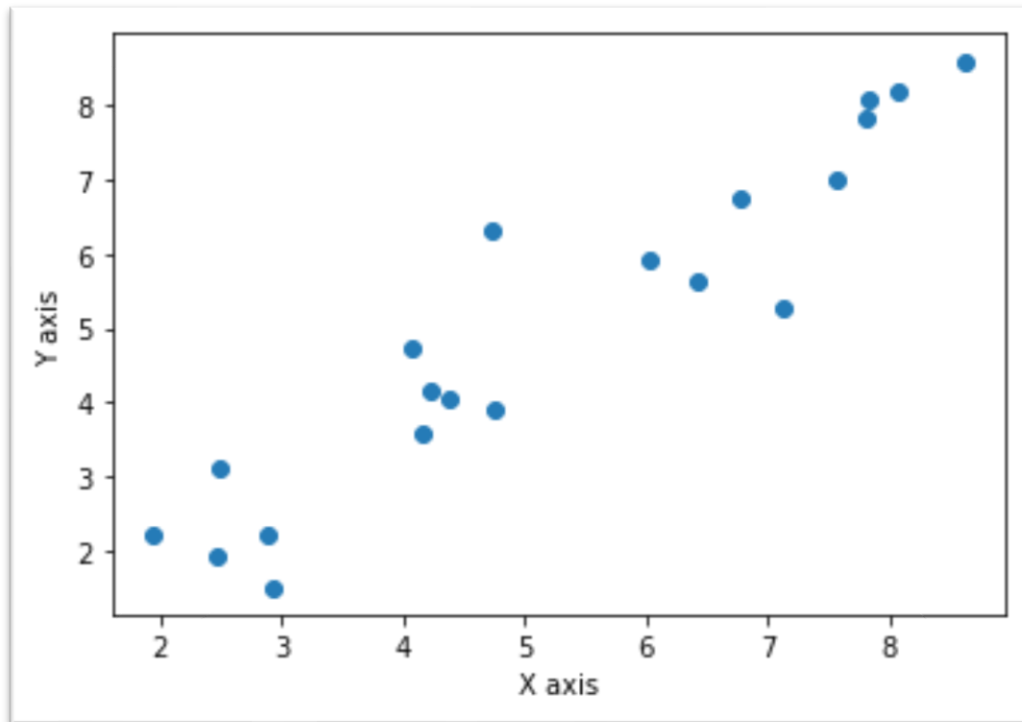
import seaborn as sns
sns.boxplot(X[:,0])
plt.show()

import numpy as np
import pandas as pd

df=pd.DataFrame(X)
df.describe()

sns.distplot(X[:,0])
plt.show()
sns.distplot(X[:,1])
plt.show()
```

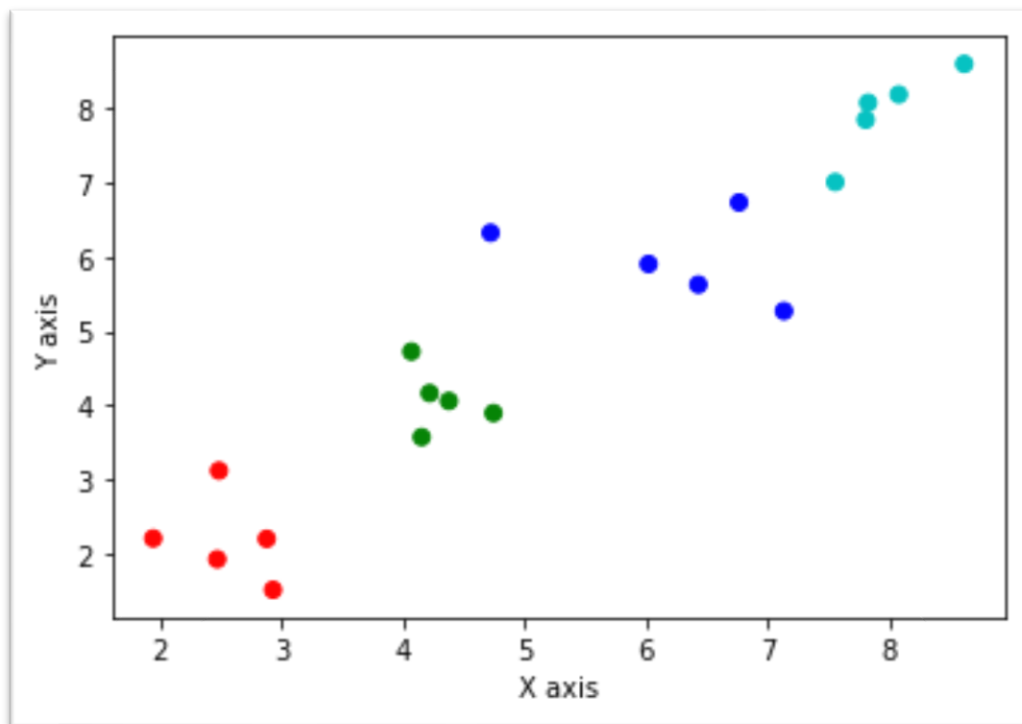
Name ^	Type	Size	Value
X	Array of float64	(20, 2)	[[7.82604393 8.07817448] [4.07202179 4.72713675]]
centers	list	4	[[2, 2], [4, 4], [6, 6], [8, 8]]
df	DataFrame	(20, 2)	Column names: 0, 1
y	Array of int32	(20,)	[3 1 0 ... 2 0 2]

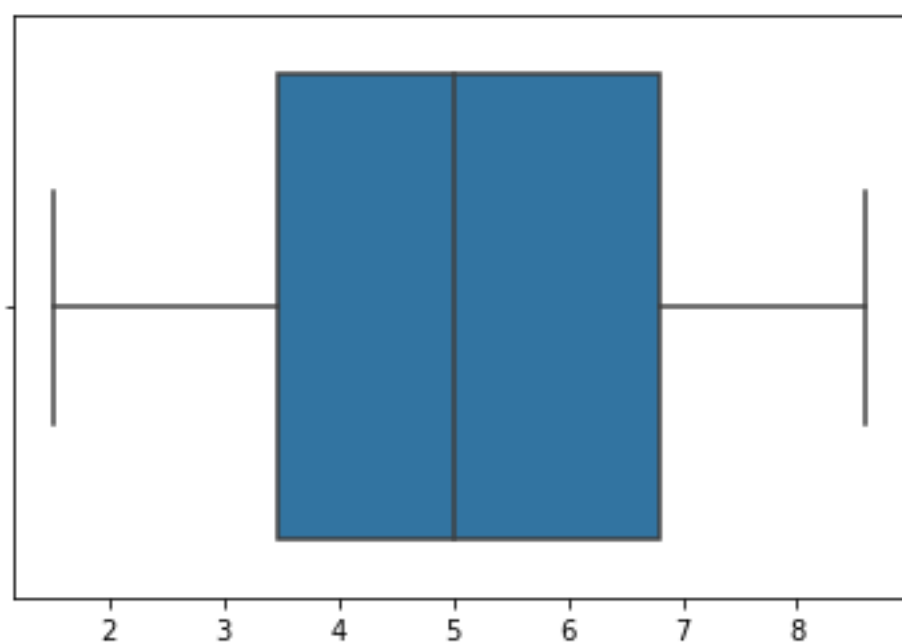
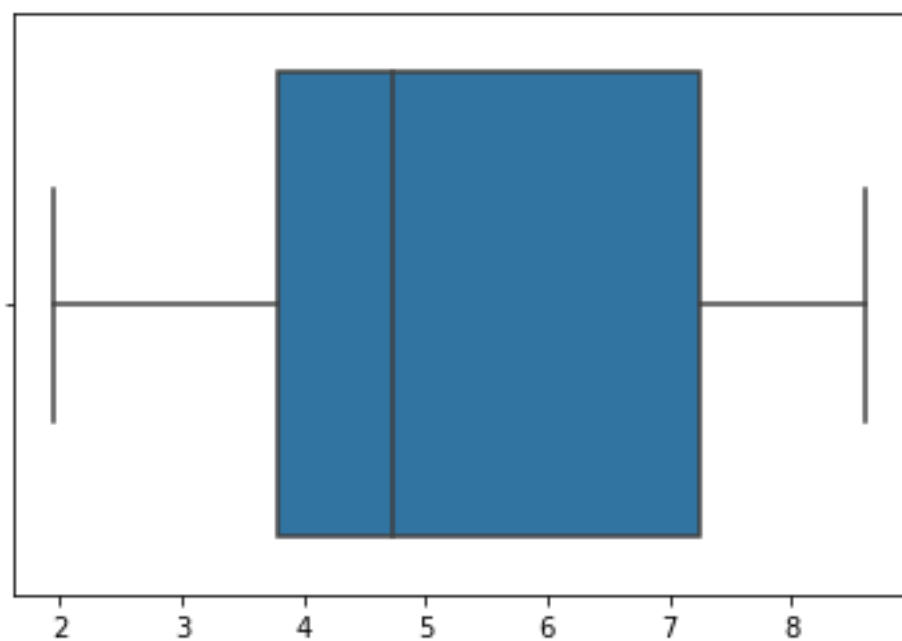


```
colors = ['r','g','b','c']  
c=[]  
for i in y:  
    c.append(colors[i])  
  
# Plot the training points  
plt.scatter(X[:, 0], X[:, 1],c=c)  
plt.xlabel('X axis')  
plt.ylabel('Y axis')  
  
plt.show()
```

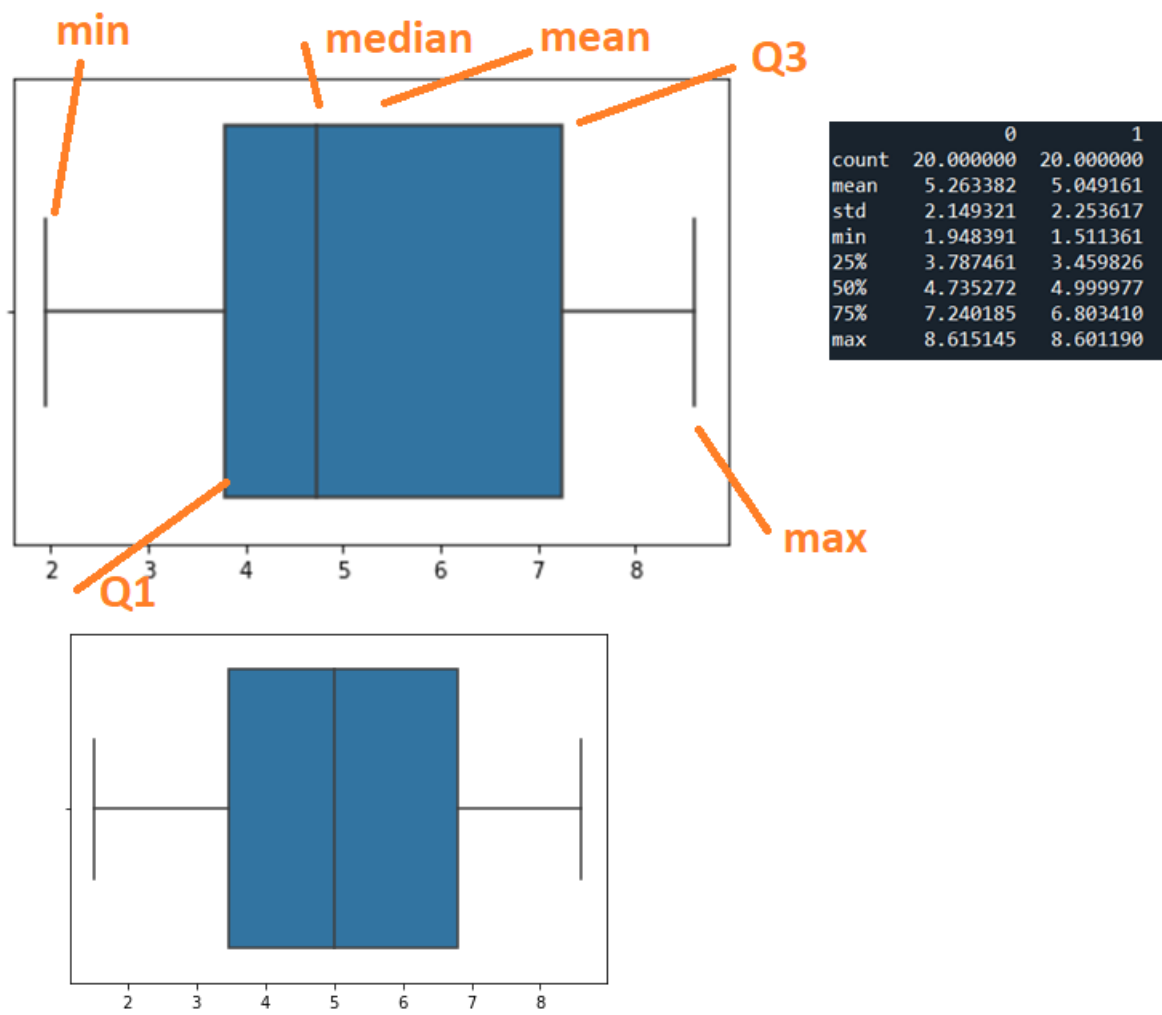
Y:

	0
4	1
5	1
6	0
7	1
8	3
9	3
10	3
11	2
12	1
13	3
14	2
15	0
16	2
17	2
18	0
19	2

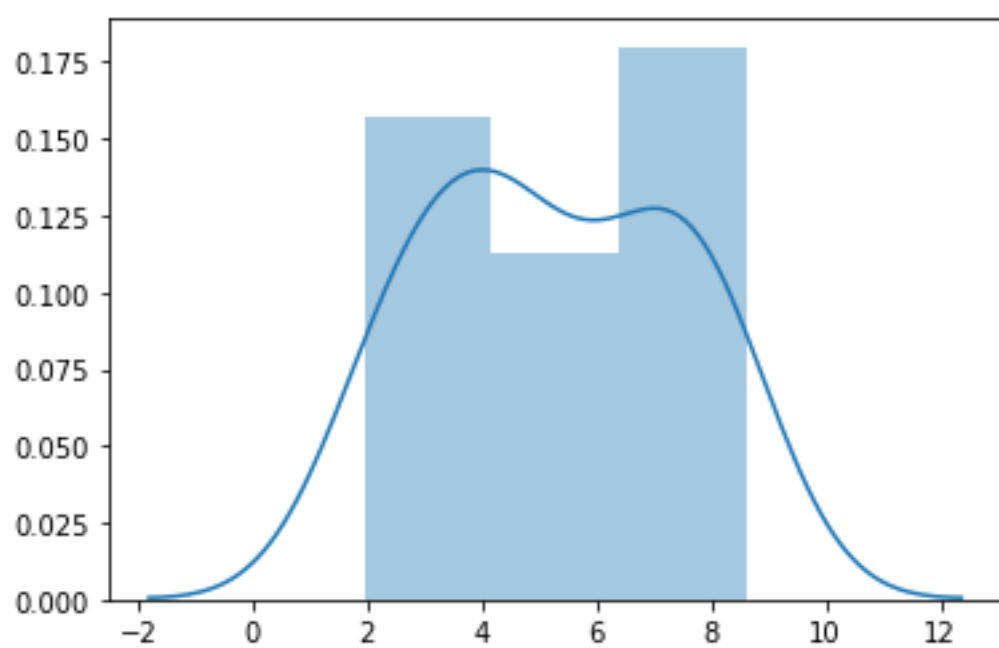




	0	1
count	20.000000	20.000000
mean	5.263382	5.049161
std	2.149321	2.253617
min	1.948391	1.511361
25%	3.787461	3.459826
50%	4.735272	4.999977
75%	7.240185	6.803410
max	8.615145	8.601190



X1



X2

