overlay-histogram-example

September 10, 2021

1 Overlay Histogram Example

source: modified from https://stackoverflow.com/questions/6871201/plot-two-histograms-on-single-chart-with-matplotlib

1.1 Import modules

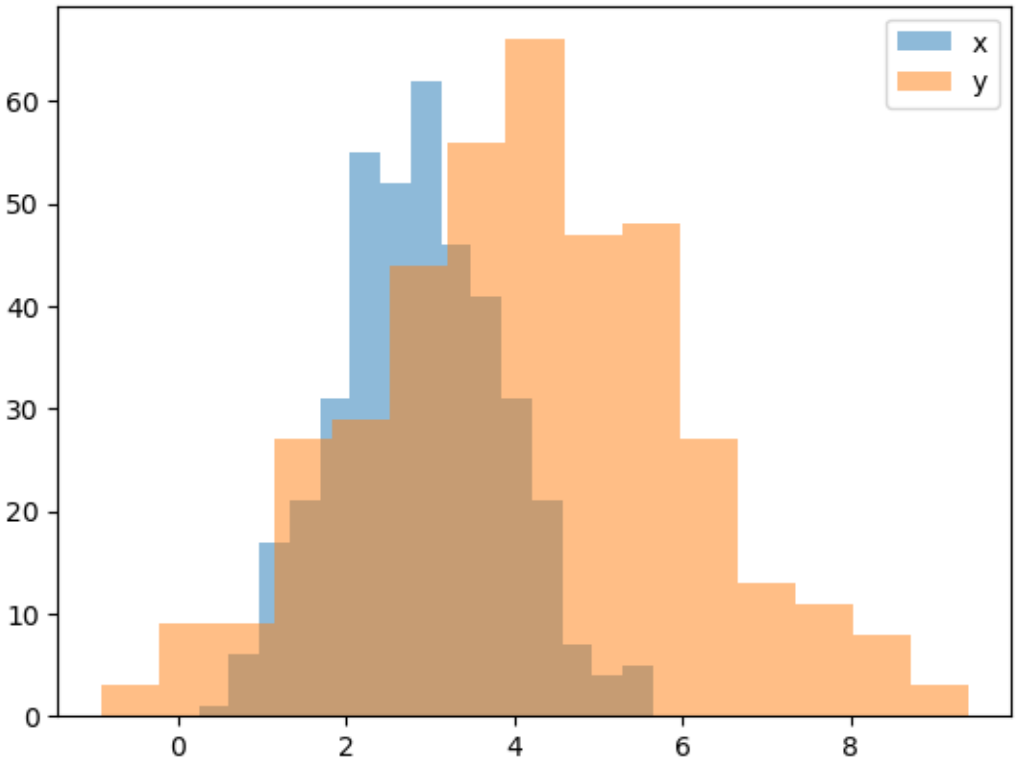
[5]:
```
import random
import numpy
from matplotlib.pyplot import hist, legend, show
import numpy as np
```

1.2 Generate test data

[ ]:
```
x = [random.gauss(3,1) for _ in range(400)]
y = [random.gauss(4,2) for _ in range(400)]
```

1.3 Histogram overlay

[11]:
```
hist(x, bins="auto", alpha=0.5, label='x')
hist(y, bins="auto", alpha=0.5, label='y')
legend(loc='upper right')
show()
```



1.4 Mean and Standard Deviation

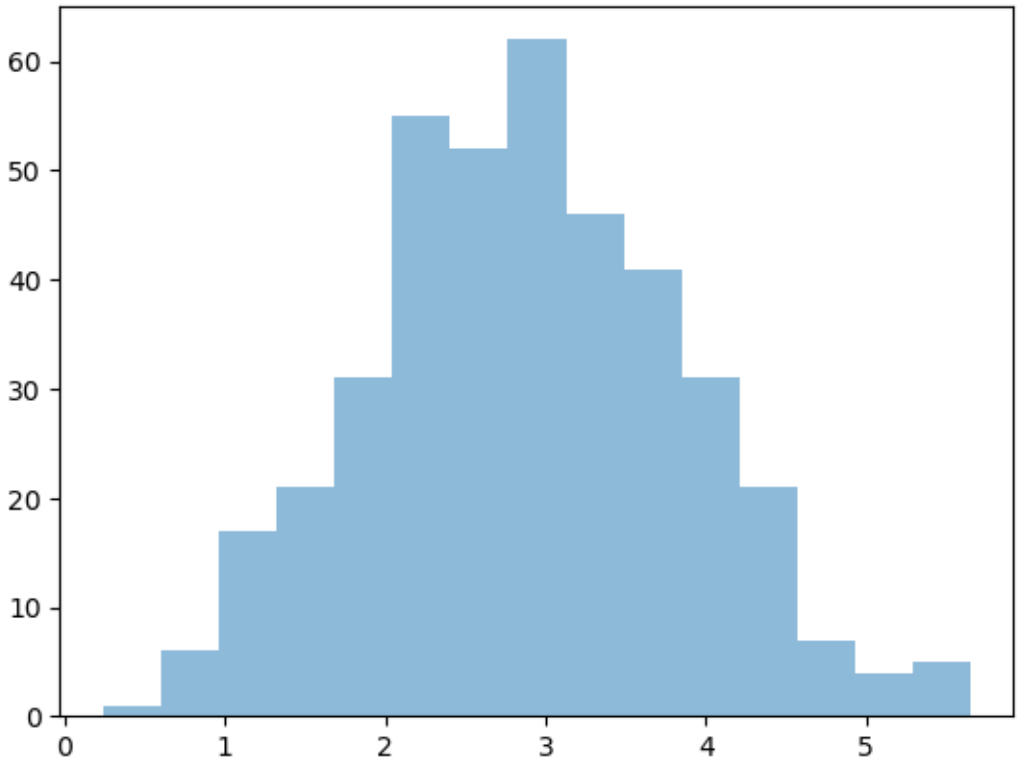
[10]:
```
print(np.mean(x), np.std(x))
print(np.mean(y), np.std(y))
```

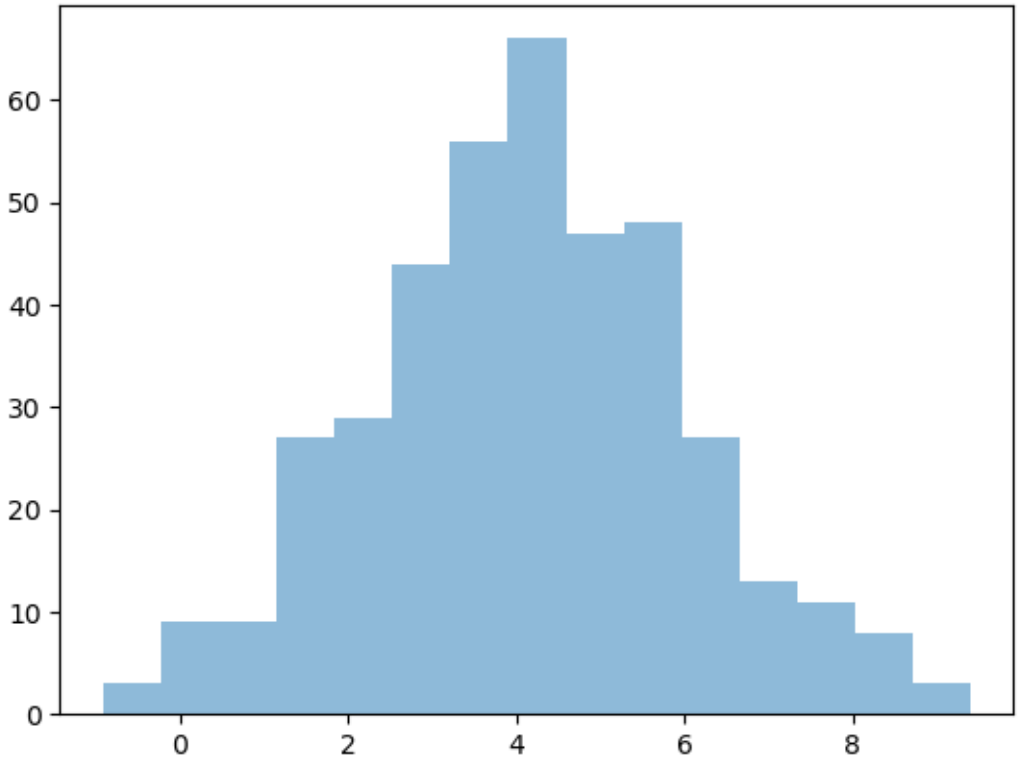
```
2.8880676273225014 0.9726233031755469
4.140335712151743 1.8909099366710311
```

1.4.1 Plot separately

Note that overlaying the two histograms makes it a lot easier to see the differences between the two distributions as opposed to plotting them separately.

[12]:
```
hist(x, bins="auto", alpha=0.5, label='x')
show()
hist(y, bins="auto", alpha=0.5, label='y')
show()
```


60
50
40
30
20
10
0
0
1
2
3
4
5


60
50
40
30
20
10
0
0
2
4
6
8

When plotted separately, you need to manually look at the min and max of the bins, try to figure out where the mean is, how the spread of the distribution compares, etc. In other words, it's a lot easier to just look at an overlay.