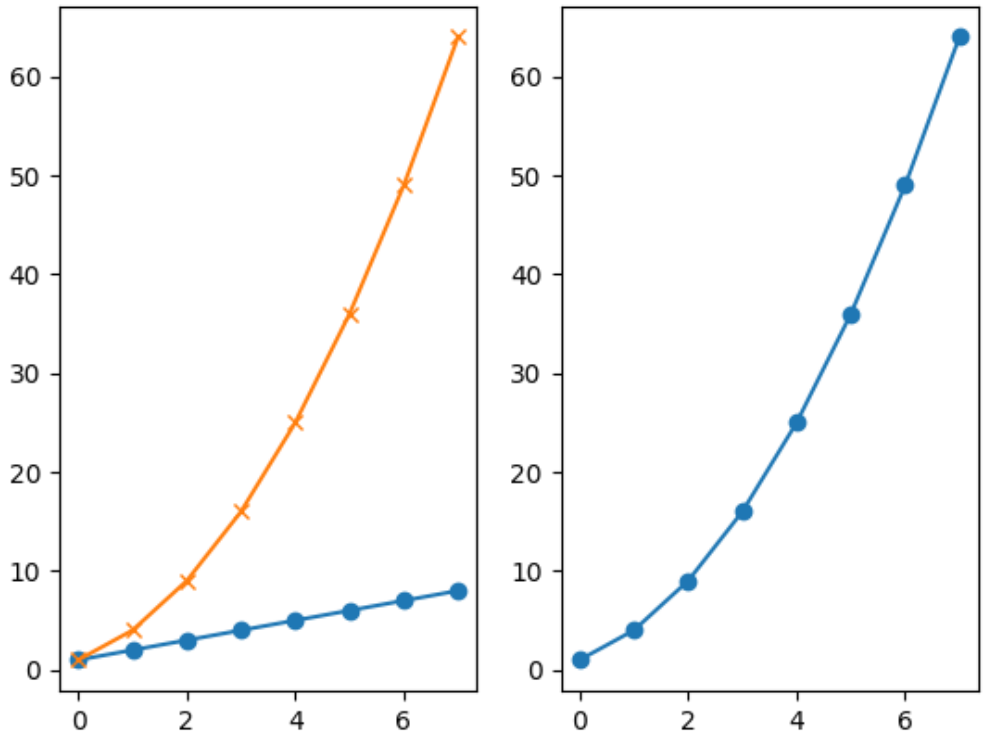
Subplots

In [3]:

```
%matplotlib notebook

import matplotlib.pyplot as plt
import numpy as np

plt.subplot?
```

In [9]:

```
plt.figure()
# subplot with 1 row, 2 columns, and current axis is 1st subplot axes
plt.subplot(1,2,1)
#plt.subplot(1,2,1,facecolor='y')
linear_data = np.array([1,2,3,4,5,6,7,8])

plt.plot(linear_data, '-o')
```



Out[9]:

```
[<matplotlib.lines.Line2D at 0x7f21204d5588>]
```

In [10]:

```
exponential_data = linear_data**2

# subplot with 1 row, 2 columns, and current axis is 2nd subplot axes
plt.subplot(1,2,2)
plt.plot(exponential_data, '-o')
```

Out[10]:

```
[<matplotlib.lines.Line2D at 0x7f20fc3cb898>]
```

In [11]:

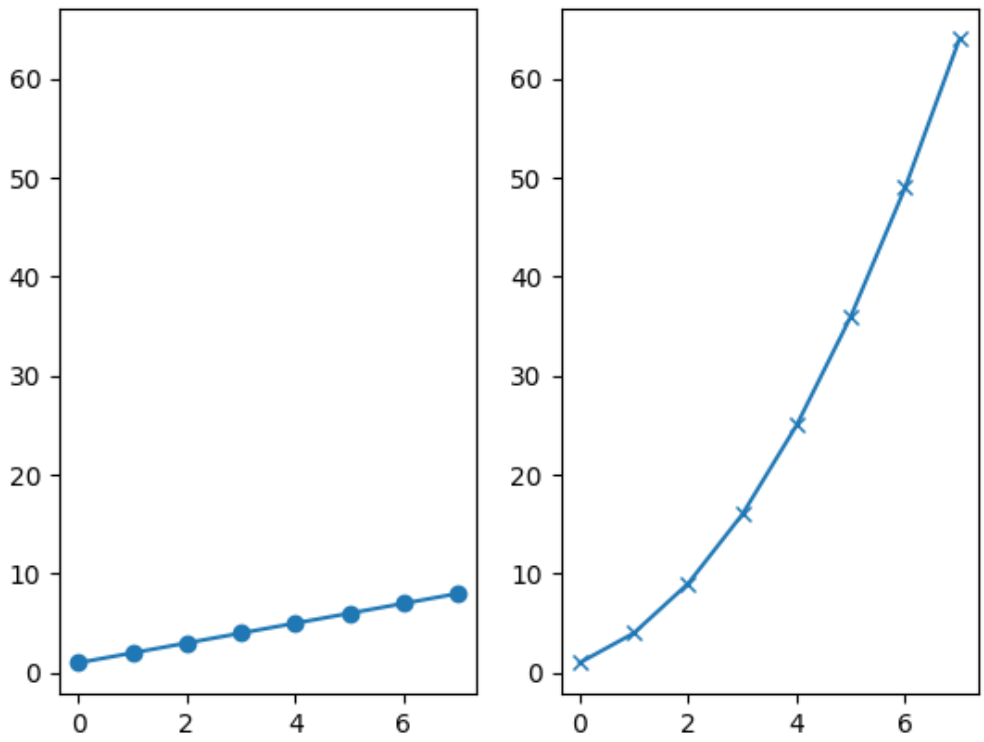
```
# plot exponential data on 1st subplot axes
plt.subplot(1, 2, 1)
plt.plot(exponential_data, '-x')
```

Out[11]:

```
[<matplotlib.lines.Line2D at 0x7f20fc3a12b0>]
```

In [12]:

```
plt.figure()
ax1 = plt.subplot(1, 2, 1)
plt.plot(linear_data, '-o')
# pass sharey=ax1 to ensure the two subplots share the same y axis
ax2 = plt.subplot(1, 2, 2, sharey=ax1)
plt.plot(exponential_data, '-x')
```

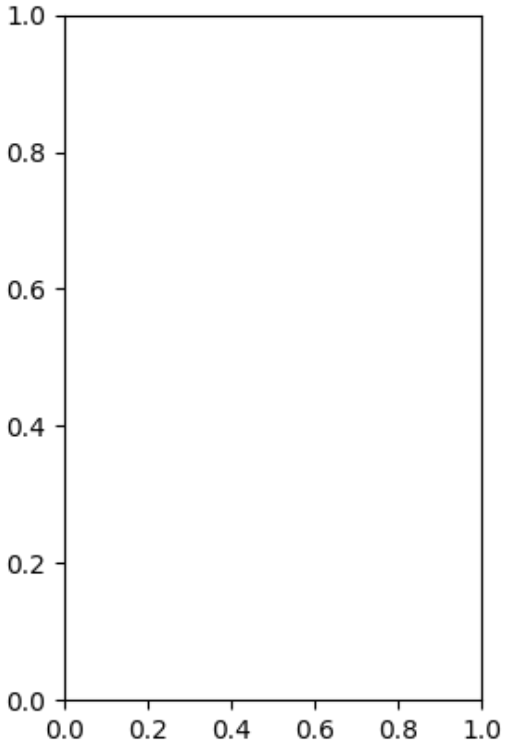


Out[12]:

```
[<matplotlib.lines.Line2D at 0x7f20fc2f97f0>]
```

In [15]:

```
plt.figure()
# the right hand side is equivalent shorthand syntax
plt.subplot(1,2,1) == plt.subplot(121)
```

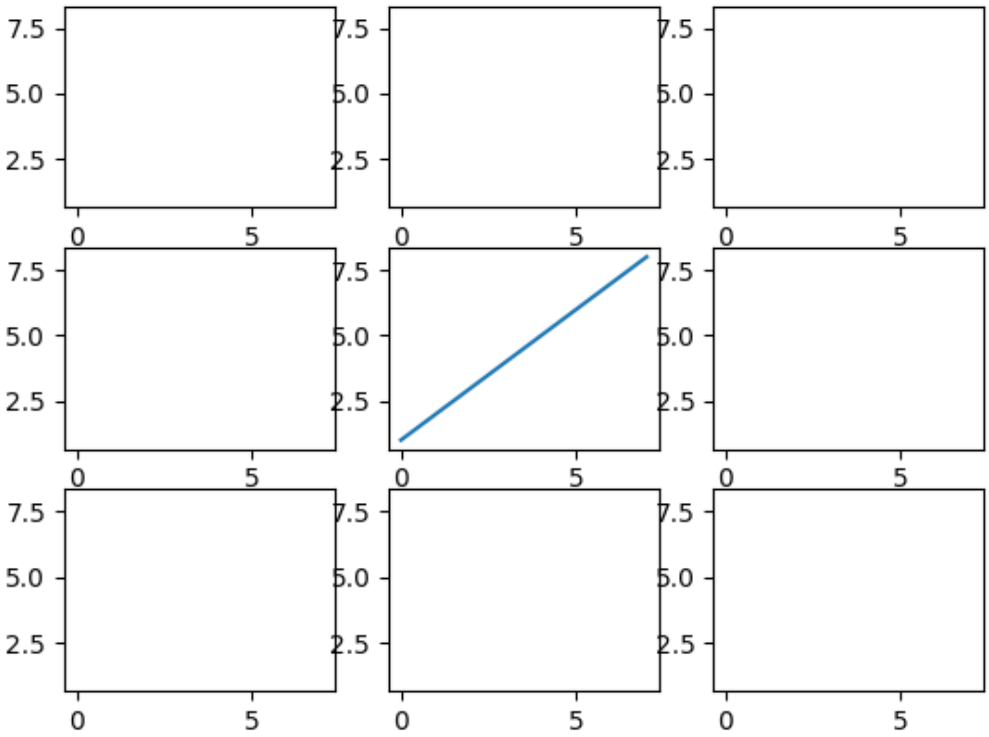


Out[15]:

True

In [16]:

```
# create a 3x3 grid of subplots
fig, ((ax1,ax2,ax3), (ax4,ax5,ax6), (ax7,ax8,ax9)) = plt.subplots(3, 3, sharex=True, sharey
# plot the linear_data on the 5th subplot axes
ax5.plot(linear_data, '-')
```



Out[16]:

```
[<matplotlib.lines.Line2D at 0x7f20fc1835c0>]
```

In [19]:

```
# set inside tick labels to visible
for ax in plt.gcf().get_axes():
    for label in ax.get_xticklabels() + ax.get_yticklabels():
        label.set_visible(True)
```

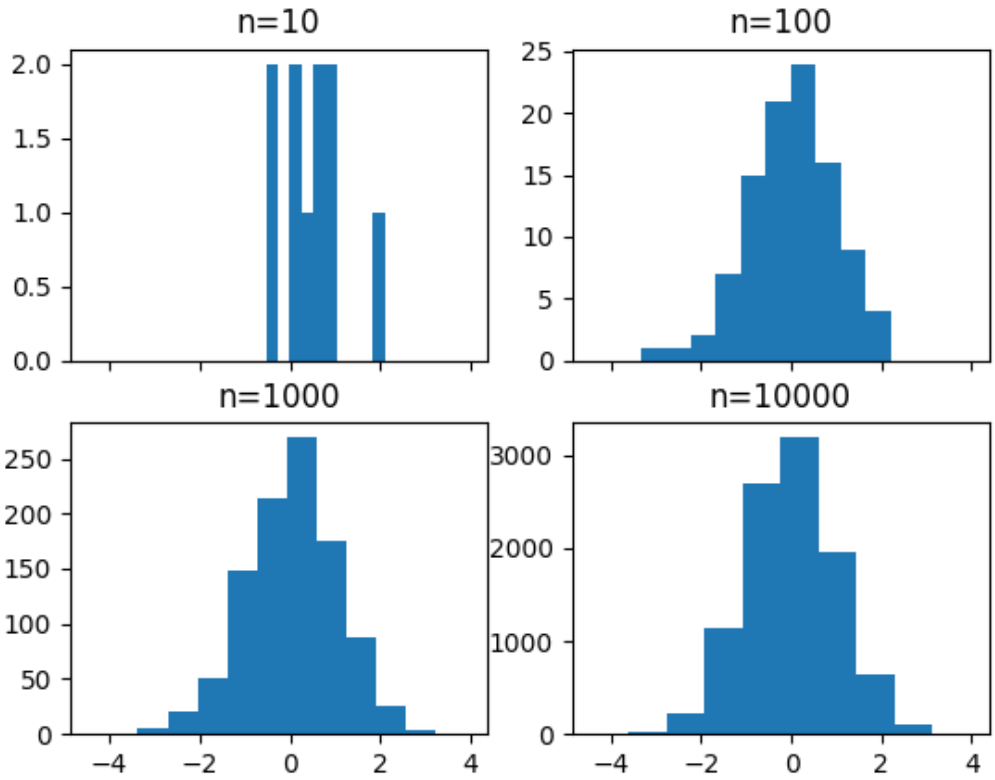
In [20]:

```
# necessary on some systems to update the plot
plt.gcf().canvas.draw()
```

Histograms

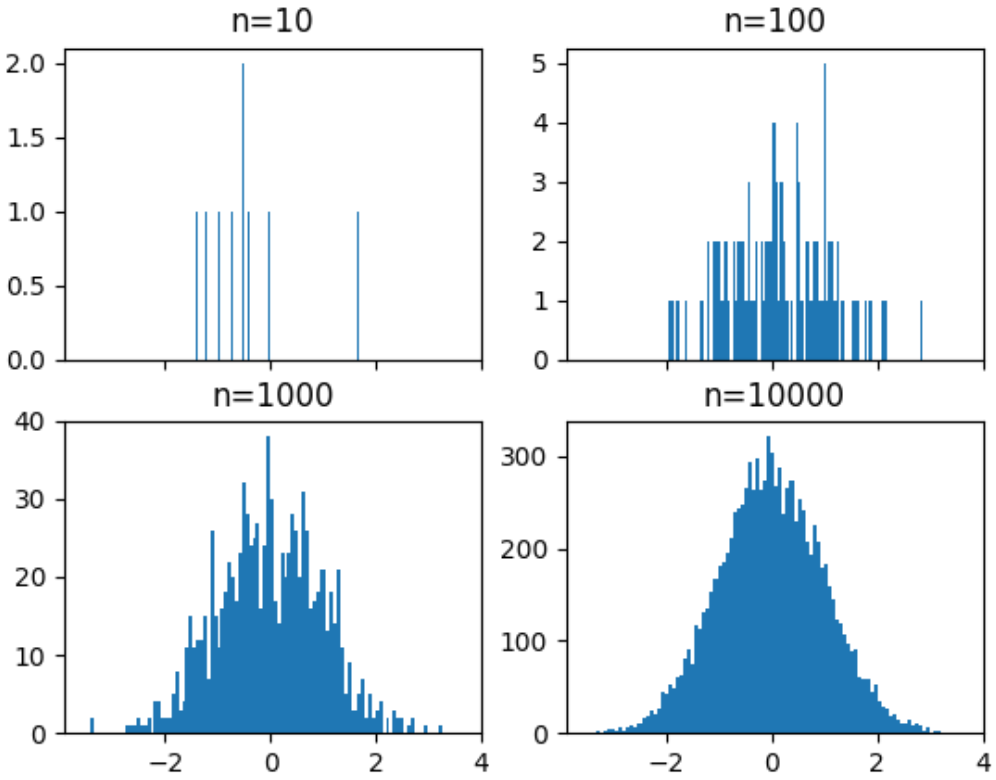
In [21]:

```
# create 2x2 grid of axis subplots
fig, ((ax1, ax2), (ax3, ax4)) = plt.subplots(2, 2, sharex=True)
axs = [ax1,ax2,ax3,ax4]

# draw n = 10, 100, 1000, and 10000 samples from the normal distribution and plot correspor
for n in range(0,len(axs)):
    sample_size = 10**(n+1)
    sample = np.random.normal(loc=0.0, scale=1.0, size=sample_size)
    axs[n].hist(sample)
    axs[n].set_title('n={}'.format(sample_size))
```



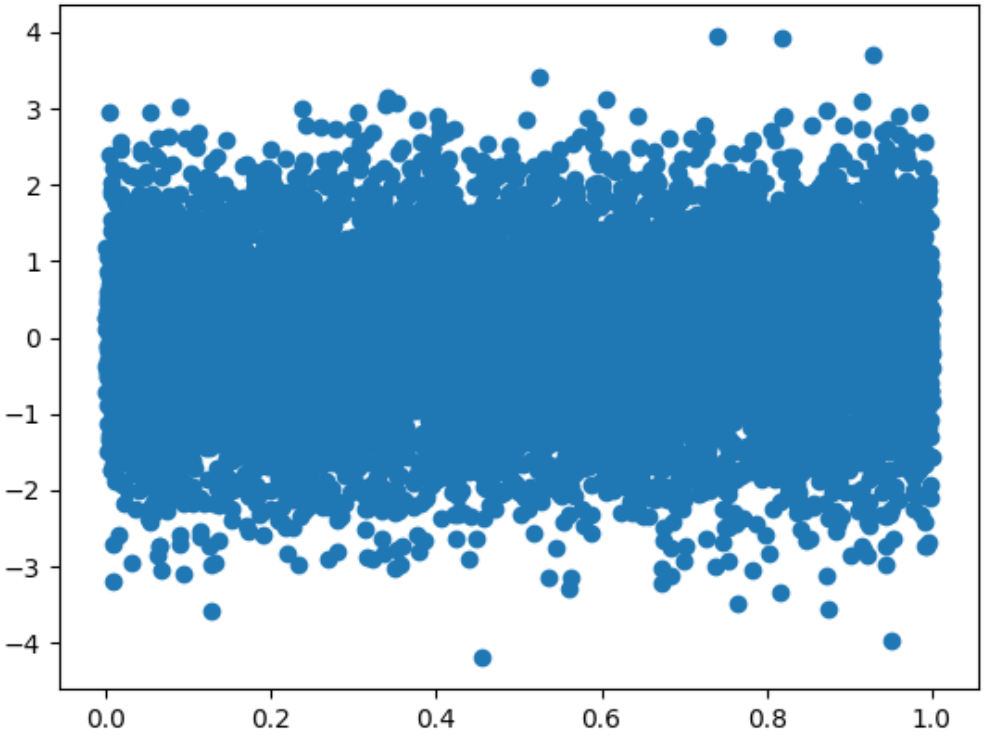
In [22]:

```
# repeat with number of bins set to 100
fig, ((ax1, ax2), (ax3, ax4)) = plt.subplots(2, 2, sharex=True)
axs = [ax1,ax2,ax3,ax4]

for n in range(0,len(axs)):
    sample_size = 10**(n+1)
    sample = np.random.normal(loc=0.0, scale=1.0, size=sample_size)
    axs[n].hist(sample, bins=100)
    axs[n].set_title('n={}'.format(sample_size))
```



In [23]:

```
plt.figure()
Y = np.random.normal(loc=0.0, scale=1.0, size=10000)
X = np.random.random(size=10000)
plt.scatter(X,Y)
```

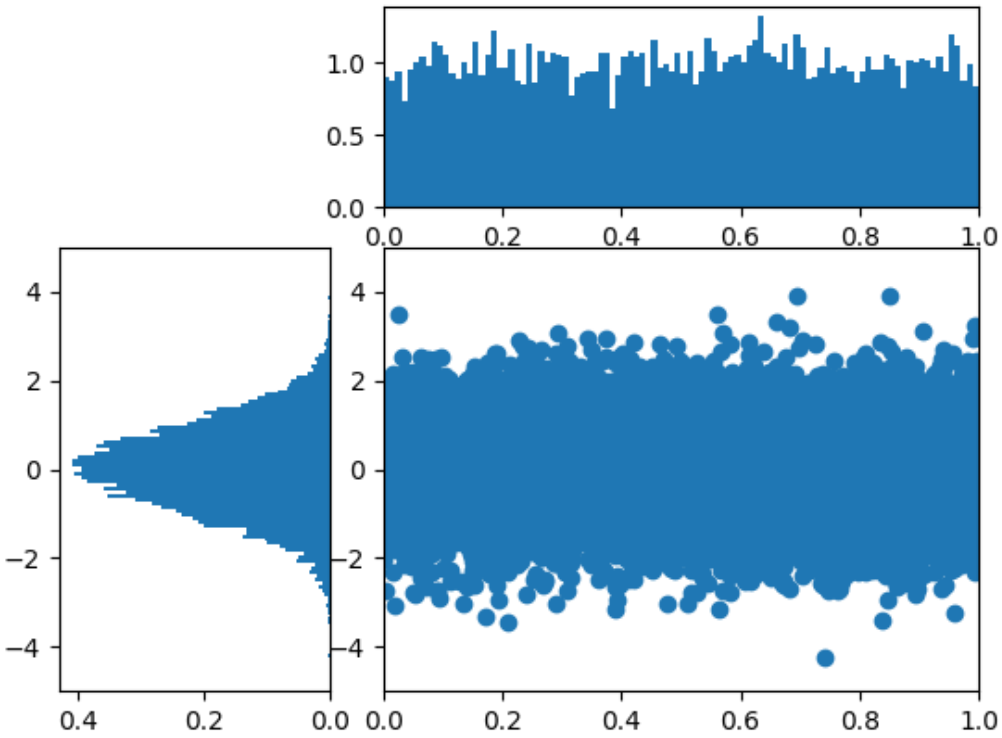


Out[23]:

```
<matplotlib.collections.PathCollection at 0x7f20f760eb38>
```

In [24]:

```
# use gridspec to partition the figure into subplots
import matplotlib.gridspec as gridspec

plt.figure()
gspec = gridspec.GridSpec(3, 3)

top_histogram = plt.subplot(gspec[0, 1:])
side_histogram = plt.subplot(gspec[1:, 0])
lower_right = plt.subplot(gspec[1:, 1:])
```



In [25]:

```
Y = np.random.normal(loc=0.0, scale=1.0, size=10000)
X = np.random.random(size=10000)
lower_right.scatter(X, Y)
top_histogram.hist(X, bins=100)
s = side_histogram.hist(Y, bins=100, orientation='horizontal')
```

In [26]:

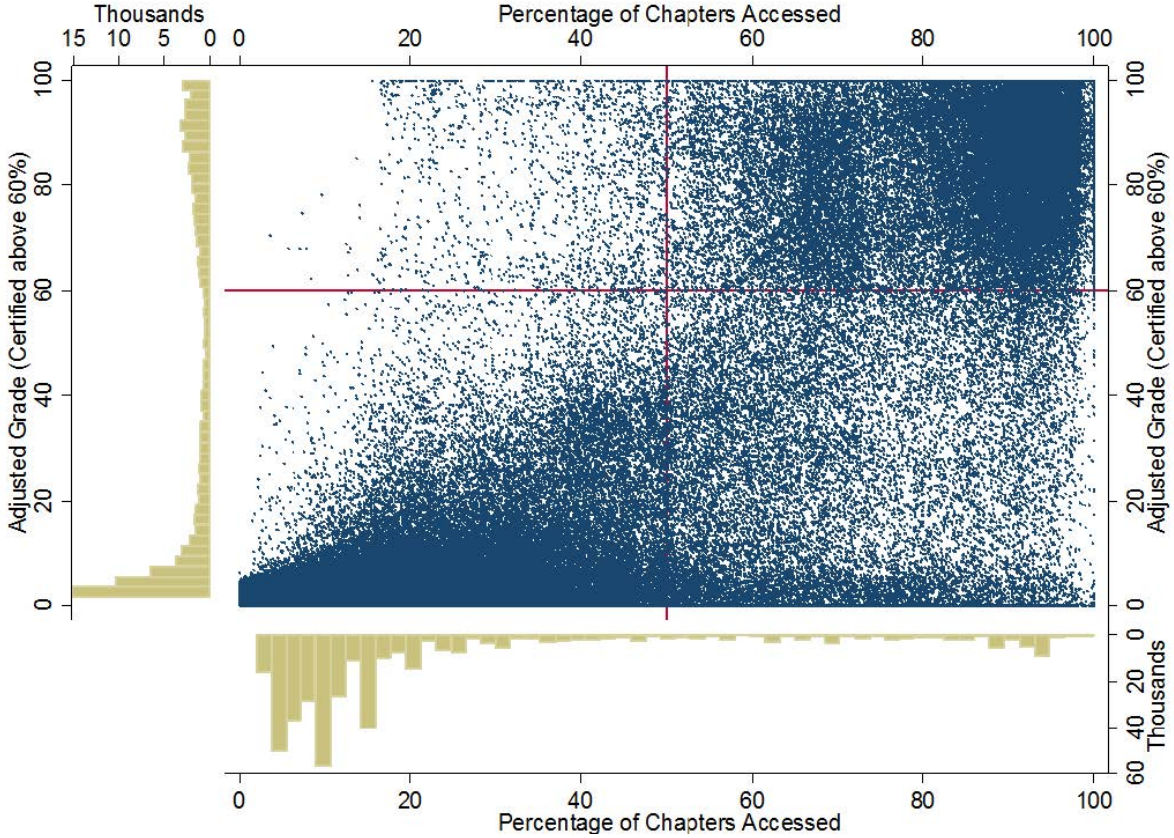
```
# clear the histograms and plot normed histograms
top_histogram.clear()
top_histogram.hist(X, bins=100, normed=True)
side_histogram.clear()
side_histogram.hist(Y, bins=100, orientation='horizontal', normed=True)
# flip the side histogram's x axis
side_histogram.invert_xaxis()
```

In [28]:

```
# change axes limits
for ax in [top_histogram, lower_right]:
    ax.set_xlim(0, 1)
for ax in [side_histogram, lower_right]:
    ax.set_ylim(-5, 5)
```

In [29]:

```
%%HTML
<img src='http://educationxpress.mit.edu/sites/default/files/journal/WP1-Fig13.jpg' />
```



Box and Whisker Plots

In [38]:

```
import pandas as pd
normal_sample = np.random.normal(loc=0.0, scale=1.0, size=10000)
random_sample = np.random.random(size=10000)
gamma_sample = np.random.gamma(2, size=10000)

df = pd.DataFrame({'normal': normal_sample,
                   'random': random_sample,
                   'gamma': gamma_sample})
```

In [33]:

```
type(normal_sample)
```

Out[33]:

```
numpy.ndarray
```

In [32]:

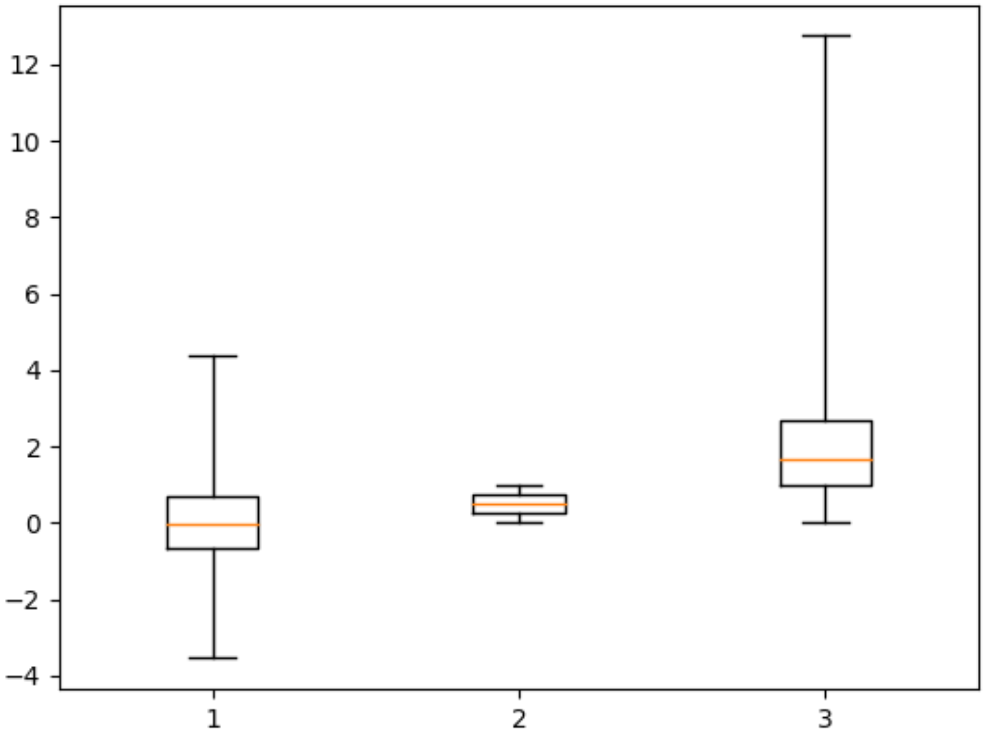
```
df.describe()
```

Out[32]:

	gamma	normal	random
count	10000.000000	10000.000000	10000.000000
mean	2.002029	0.001172	0.498940
std	1.417888	1.004378	0.289626
min	0.016312	-3.537719	0.000031
25%	0.967389	-0.679215	0.250644
50%	1.682233	-0.010133	0.491118
75%	2.700730	0.684953	0.751001
max	12.749017	4.387283	0.999939

In [34]:

```
plt.figure()
# create a boxplot of the normal data, assign the output to a variable to supress output
_ = plt.boxplot(df['normal'], whis='range')
```

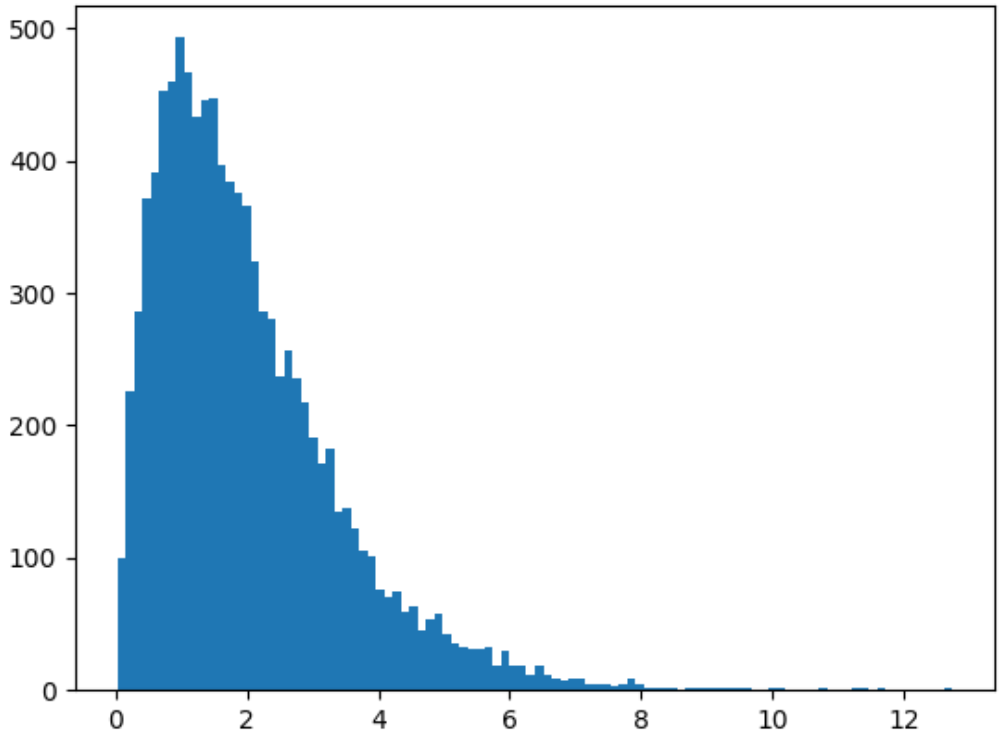


In [35]:

```
# clear the current figure
plt.clf()
# plot boxplots for all three of df's columns
_ = plt.boxplot([ df['normal'], df['random'], df['gamma'] ], whis='range')
```

In [36]:

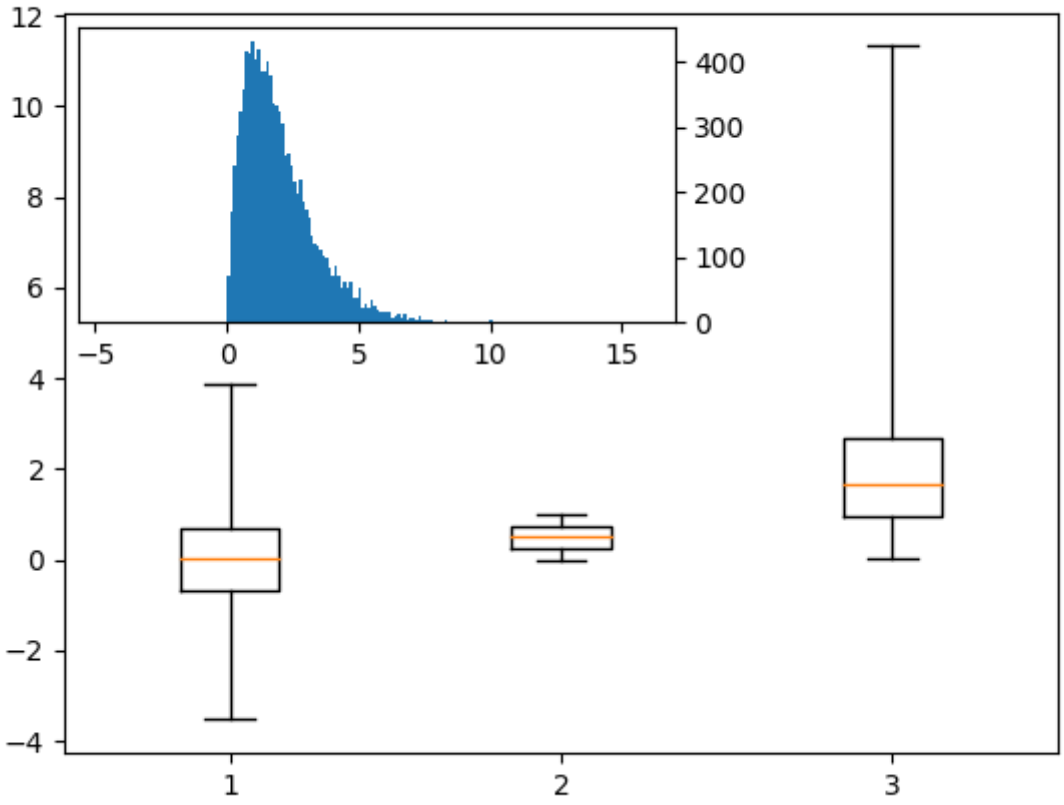
```
plt.figure()
_ = plt.hist(df['gamma'], bins=100)
```



In [41]:

```
import mpl_toolkits.axes_grid1.inset_locator as mpl_il

plt.figure()
plt.boxplot([ df['normal'], df['random'], df['gamma'] ], whis='range')
# overlay axis on top of another
ax2 = mpl_il.inset_axes(plt.gca(), width='60%', height='40%', loc=2)
ax2.hist(df['gamma'], bins=100)
ax2.margins(x=0.5)
```

Figure 12



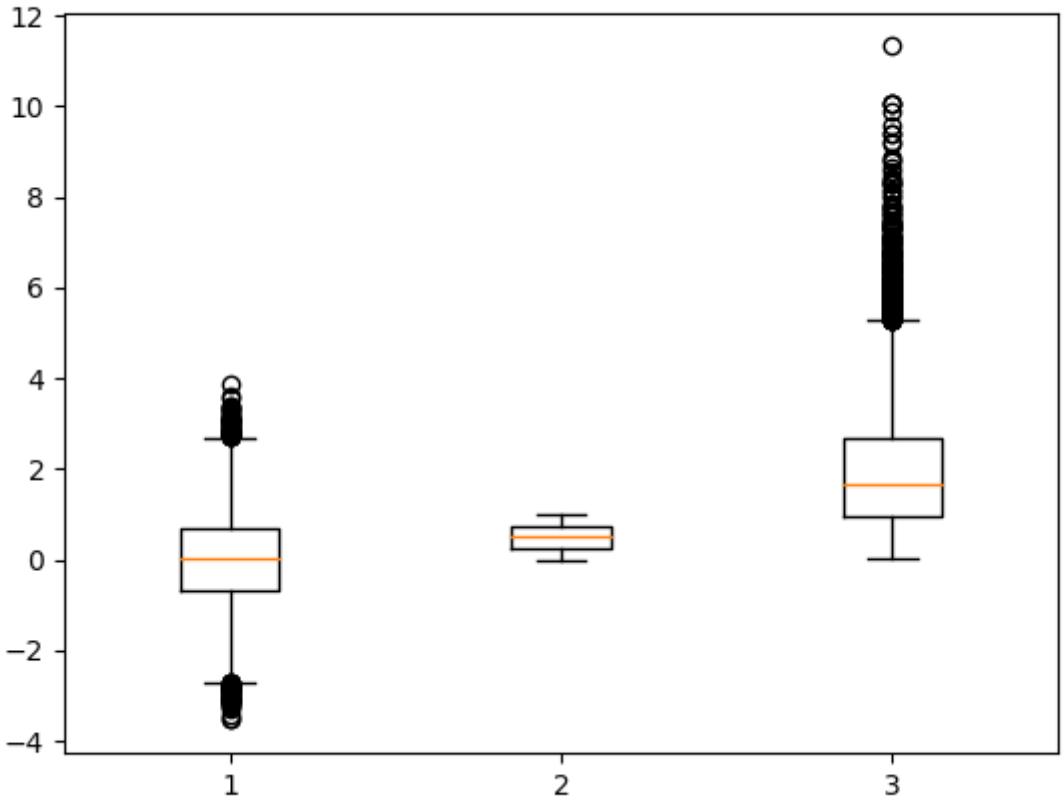
In [42]:

```
# switch the y axis ticks for ax2 to the right side
ax2.yaxis.tick_right()
```

In [43]:

```
# if `whis` argument isn't passed, boxplot defaults to showing 1.5*interquartile (IQR) whis
plt.figure()
_ = plt.boxplot([ df['normal'], df['random'], df['gamma'] ] )
```

Figure 13

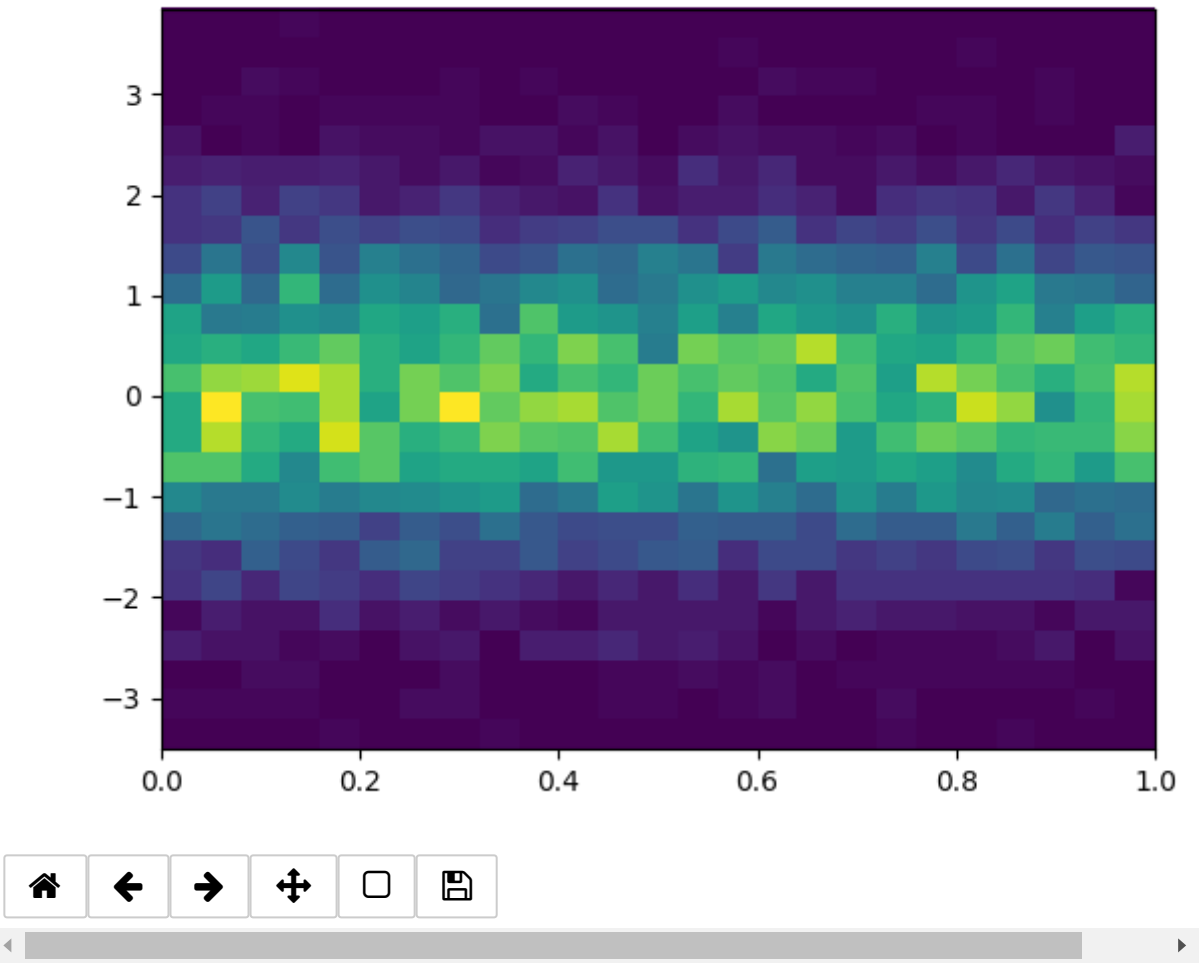


Heatmaps

In [44]:

```
plt.figure()

Y = np.random.normal(loc=0.0, scale=1.0, size=10000)
X = np.random.random(size=10000)
_ = plt.hist2d(X, Y, bins=25)
```

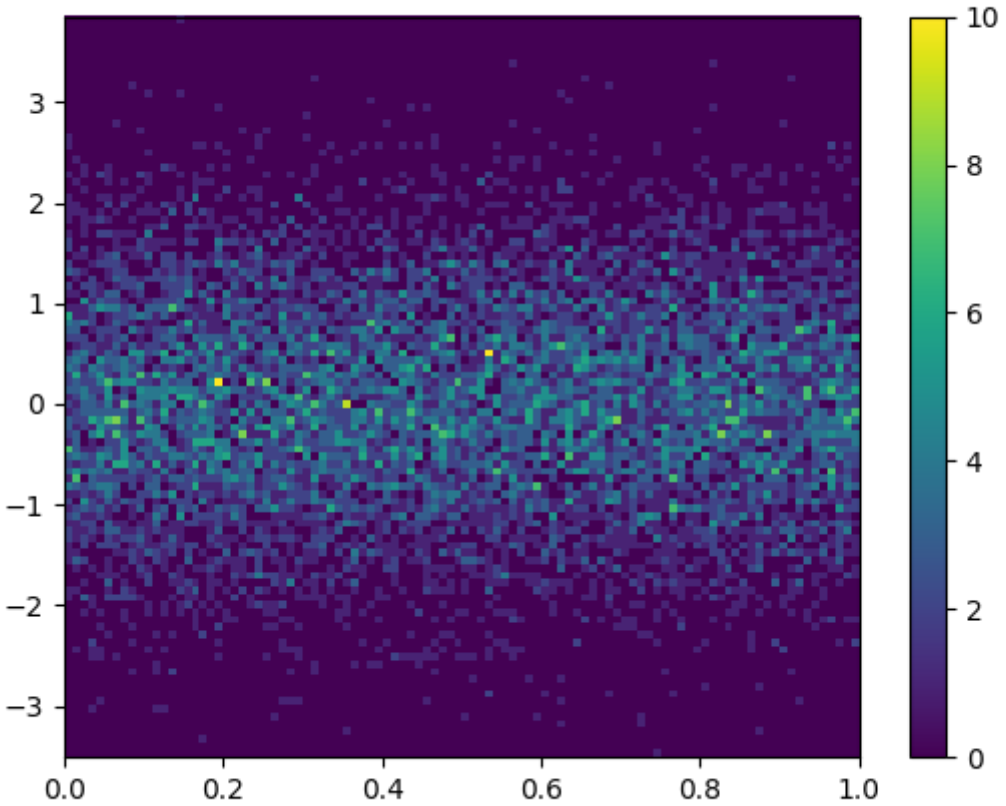
Figure 14



In [45]:

```
plt.figure()
_ = plt.hist2d(X, Y, bins=100)
```

Figure 15



Downlo

In [46]:

```
# add a colorbar legend
plt.colorbar()
```

Out[46]:

```
<matplotlib.colorbar.Colorbar at 0x7f20eb09ea90>
```

Animations

In [47]:

```
import matplotlib.animation as animation

n = 100
x = np.random.randn(n)
```

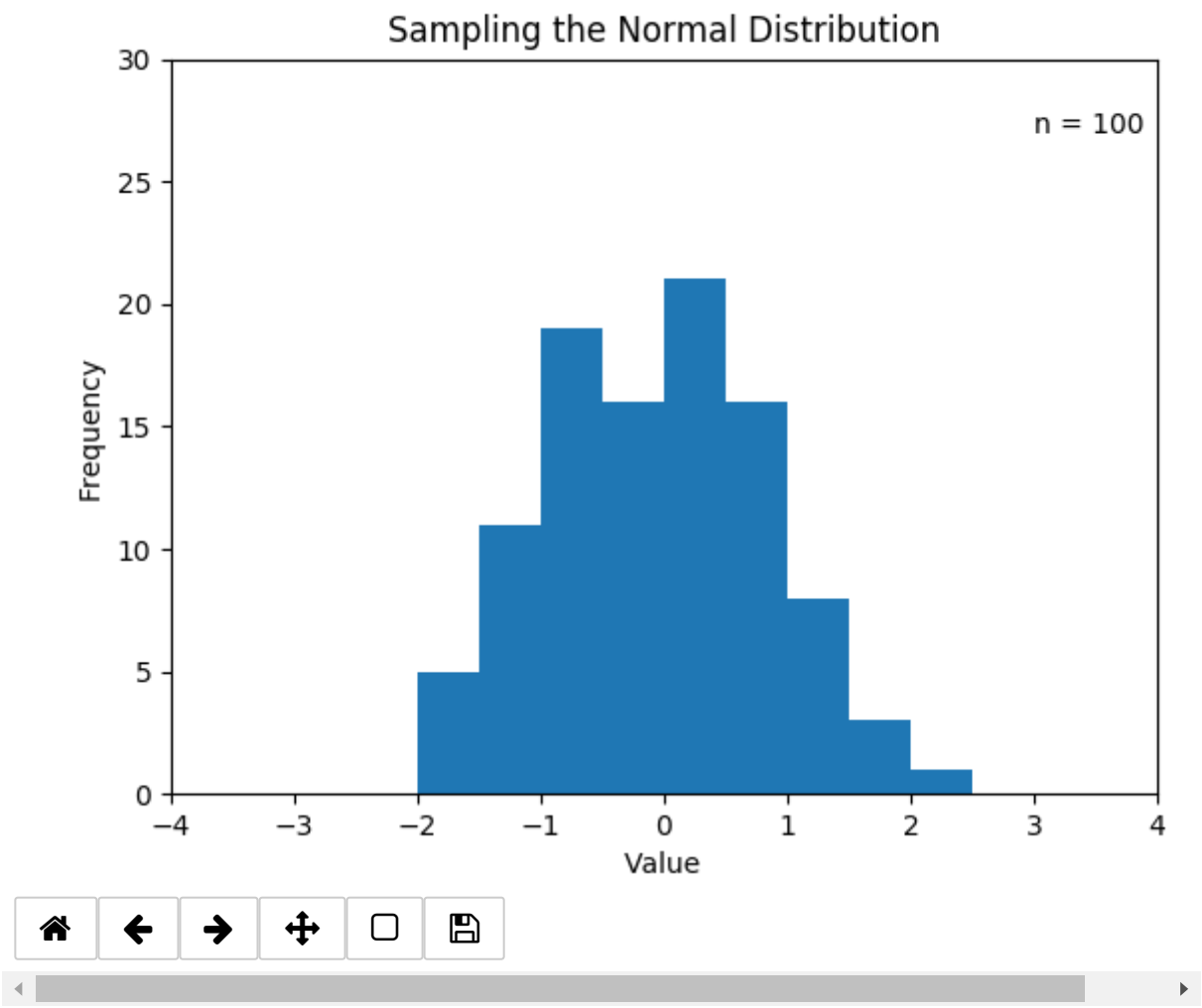
In [57]:

```
# create the function that will do the plotting, where curr is the current frame
def update(curr):
    # check if animation is at the last frame, and if so, stop the animation a
    if curr == n:
        a.event_source.stop()
    plt.cla()
    bins = np.arange(-4, 4, 0.5)
    plt.hist(x[:curr], bins=bins)
    plt.axis([-4,4,0,30])
    plt.gca().set_title('Sampling the Normal Distribution')
    plt.gca().set_ylabel('Frequency')
    plt.gca().set_xlabel('Value')
    plt.annotate('n = {}'.format(curr), [3,27])  # places text on location [3,27]
```

In [58]:

```
fig = plt.figure()
a = animation.FuncAnimation(fig, update, interval=100)
```

Figure 16

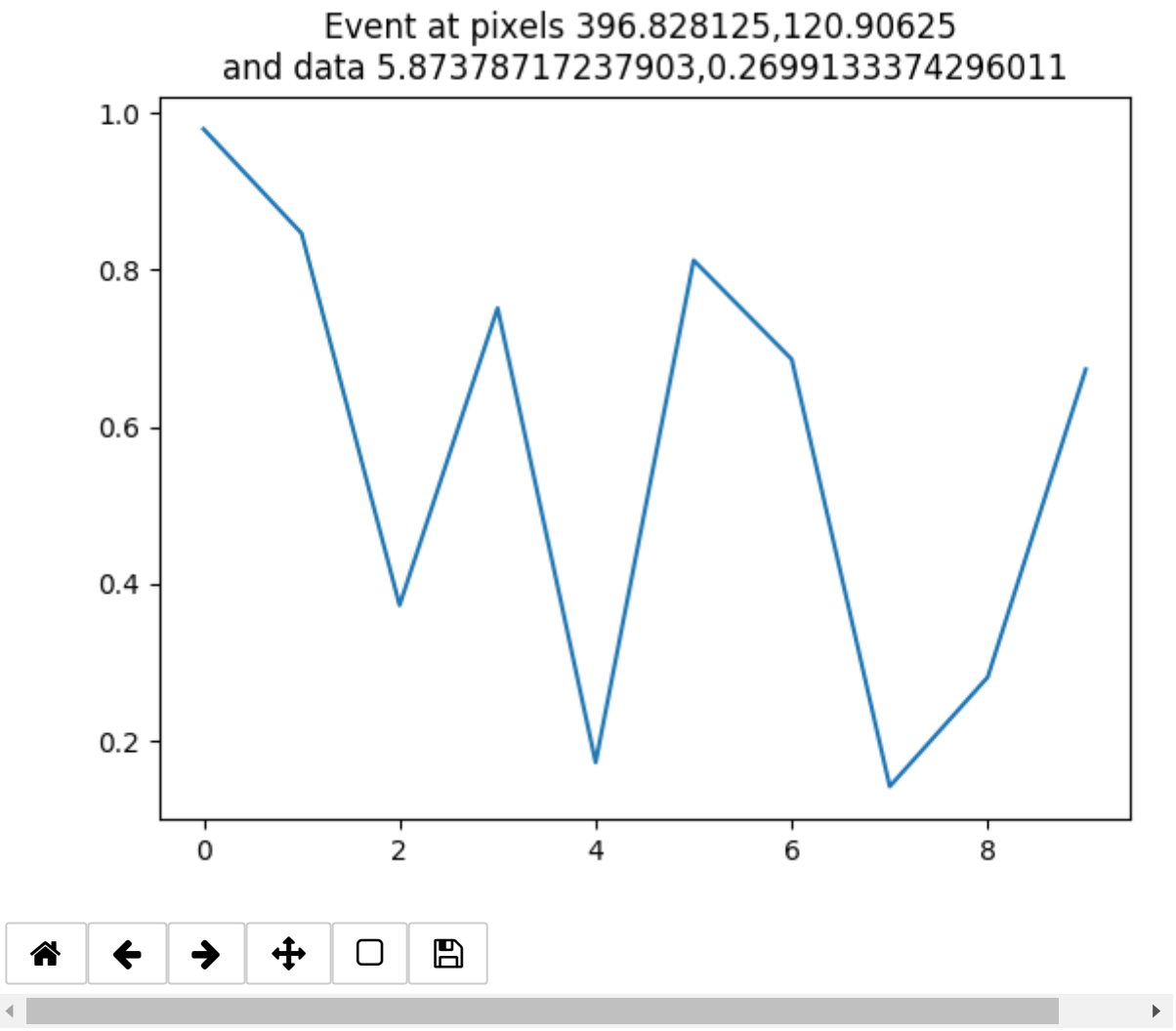


Interactivity

In [61]:

```
plt.figure()
data = np.random.rand(10)
plt.plot(data)

def onclick(event):
    plt.cla()
    plt.plot(data)
    plt.gca().set_title('Event at pixels {},{} \nand data {},{}'.format(event.x, event.y, e

# tell mpl_connect we want to pass a 'button_press_event' into onclick when the event is de
plt.gcf().canvas.mpl_connect('button_press_event', onclick)
```

Figure 17



Out[61]:

7

In [62]:

```
from random import shuffle
origins = ['China', 'Brazil', 'India', 'USA', 'Canada', 'UK', 'Germany', 'Iraq', 'Chile', '

shuffle(origins)

df = pd.DataFrame({'height': np.random.rand(10),
                   'weight': np.random.rand(10),
                   'origin': origins})
df
```

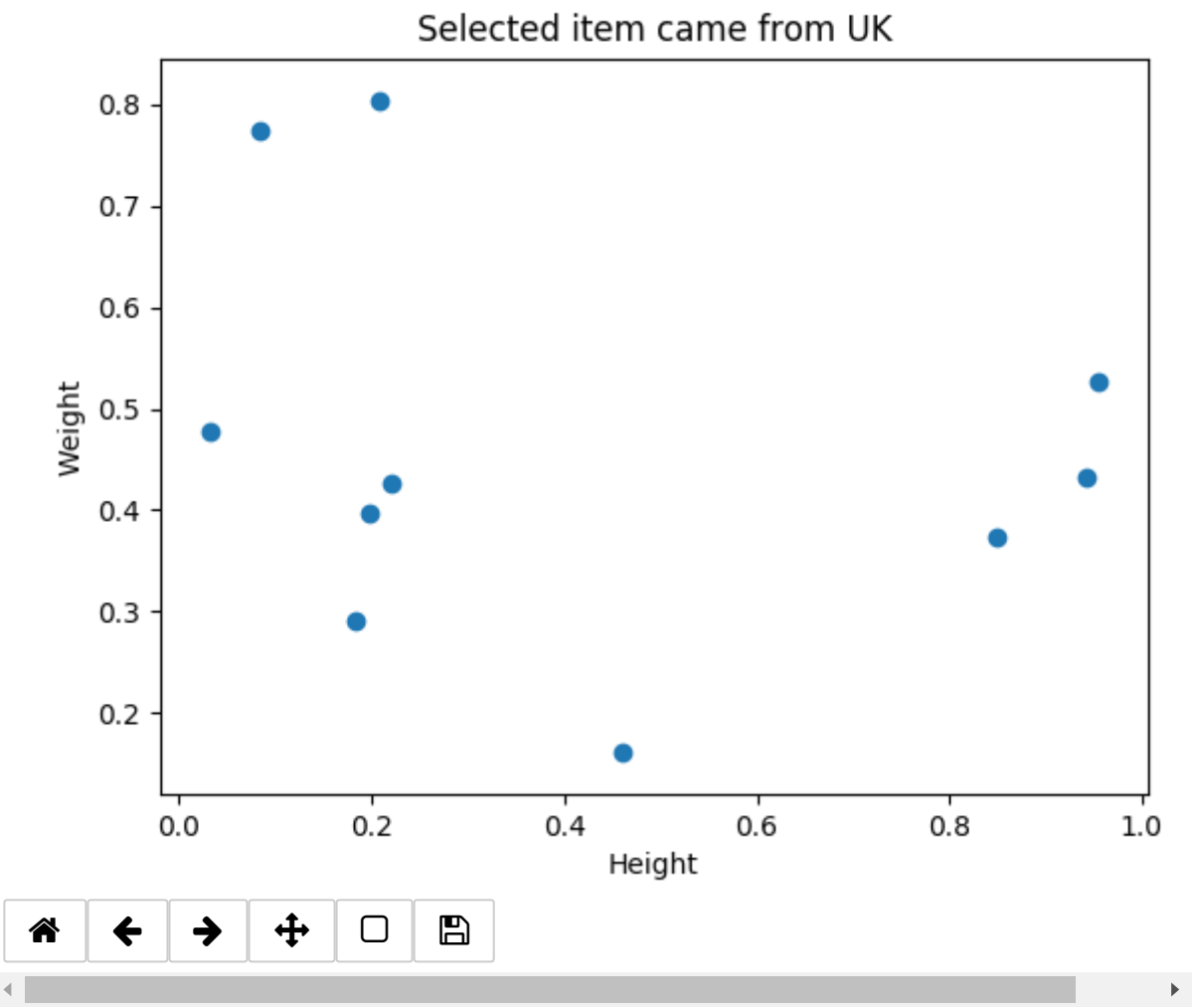
Out[62]:

	height	origin	weight
0	0.197960	Mexico	0.397571
1	0.848537	Germany	0.373720
2	0.943266	Canada	0.433174
3	0.953814	China	0.527327
4	0.183593	USA	0.290036
5	0.220233	Brazil	0.426418
6	0.033049	UK	0.477162
7	0.460844	Chile	0.160462
8	0.208287	India	0.804576
9	0.083083	Iraq	0.775061

In [63]:

```
plt.figure()
# picker=5 means the mouse doesn't have to click directly on an event, but can be up to 5 p
plt.scatter(df['height'], df['weight'], picker=5)
plt.gca().set_ylabel('Weight')
plt.gca().set_xlabel('Height')
```

Figure 18



Out[63]:

```
<matplotlib.text.Text at 0x7f20edeac7f0>
```

In [68]:

```
def onpick(event):
    origin = df.iloc[event.ind[0]]['origin']
    plt.gca().set_title('Selected item came from {}'.format(origin))

# tell mpl_connect we want to pass a 'pick_event' into onpick when the event is detected
plt.gcf().canvas.mpl_connect('pick_event', onpick)
```

Out[68]:

```
11
```