Week-3

July 3, 2020

1 Subplots

[1]:
```
%matplotlib notebook

import matplotlib.pyplot as plt
import numpy as np

plt.subplot?
```

[2]:
```
plt.figure()
# subplot with 1 row, 2 columns, and current axis is 1st subplot axes
plt.subplot(1, 2, 1)

linear_data = np.array([1,2,3,4,5,6,7,8])

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

```
<IPython.core.display.Javascript object>
```

```
<IPython.core.display.HTML object>
```

[2]:
```
[<matplotlib.lines.Line2D at 0x7f90b2de8668>]
```

[3]:
```
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')
```

[3]:
```
[<matplotlib.lines.Line2D at 0x7f90b2dc9710>]
```

[4]:
```
# plot exponential data on 1st subplot axes
plt.subplot(1, 2, 1)
plt.plot(exponential_data, '-x')
```

[4]:
```
[<matplotlib.lines.Line2D at 0x7f90b2dc9630>]
```

[5]:
```
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')
```

```
<IPython.core.display.Javascript object>
```

```
<IPython.core.display.HTML object>
```

[5]:
```
[<matplotlib.lines.Line2D at 0x7f90b0cb04a8>]
```

[6]:
```
plt.figure()
# the right hand side is equivalent shorthand syntax
plt.subplot(1,2,1) == plt.subplot(121)
```

```
<IPython.core.display.Javascript object>
```

```
<IPython.core.display.HTML object>
```

[6]:
```
True
```

[7]:
```
# create a 3x3 grid of subplots
fig, ((ax1,ax2,ax3), (ax4,ax5,ax6), (ax7,ax8,ax9)) = plt.subplots(3, 3,␣
 ↪sharex=True, sharey=True)
# plot the linear_data on the 5th subplot axes
ax5.plot(linear_data, '-')
```

```
<IPython.core.display.Javascript object>
```

```
<IPython.core.display.HTML object>
```

[7]:
```
[<matplotlib.lines.Line2D at 0x7f90b0b597f0>]
```

[8]:
```
# 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)
```

[9]:
```
# necessary on some systems to update the plot
plt.gcf().canvas.draw()
```

2 Histograms

[10]:
```
# 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 corresponding histograms
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))
```

```
<IPython.core.display.Javascript object>
```

```
<IPython.core.display.HTML object>
```

[11]:
```
# 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))
```

```
<IPython.core.display.Javascript object>
```

```
<IPython.core.display.HTML object>
```

[12]:
```
plt.figure()
Y = np.random.normal(loc=0.0, scale=1.0, size=10000)
X = np.random.random(size=10000)
plt.scatter(X,Y)
```

```
<IPython.core.display.Javascript object>
```

```
<IPython.core.display.HTML object>
```

[12]:
```
<matplotlib.collections.PathCollection at 0x7f90af965e80>
```

[13]:
```
# 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:])
```

```
<IPython.core.display.Javascript object>
```

```
<IPython.core.display.HTML object>
```

[14]:
```
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')
```

[15]:
```
# 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()
```

[16]:
```
# 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)
```

[17]:
```
%%HTML
<img src='http://educationxpress.mit.edu/sites/default/files/journal/WP1-Fig13.
 ↪jpg' />
```

```
<IPython.core.display.HTML object>
```

3 Box and Whisker Plots

[18]:
```
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})
```

[19]:
```
df.describe()
```

[19]:

	gamma	normal	random
count	10000.000000	10000.000000	10000.000000
mean	1.991803	0.016627	0.502057
std	1.402459	1.002956	0.289292
min	0.010887	-4.144426	0.000064
25%	0.961823	-0.662630	0.251931
50%	1.670746	0.002558	0.506380
75%	2.696845	0.694670	0.752469
max	12.016970	3.746622	0.999915

[20]:
```
plt.figure()
# create a boxplot of the normal data, assign the output to a variable to␣
 ↪supress output
_ = plt.boxplot(df['normal'], whis='range')
```

```
<IPython.core.display.Javascript object>
```

```
<IPython.core.display.HTML object>
```

[21]:
```
# 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')
```

[22]:
```
plt.figure()
_ = plt.hist(df['gamma'], bins=100)
```

```
<IPython.core.display.Javascript object>
```

```
<IPython.core.display.HTML object>
```

[23]:
```
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)
```

```
<IPython.core.display.Javascript object>
```

```
<IPython.core.display.HTML object>
```

[24]:
```
# switch the y axis ticks for ax2 to the right side
ax2.yaxis.tick_right()
```

[25]:
```
# if `whis` argument isn't passed, boxplot defaults to showing 1.
 ↪5*interquartile (IQR) whiskers with outliers
plt.figure()
_ = plt.boxplot([ df['normal'], df['random'], df['gamma'] ] )
```

```
<IPython.core.display.Javascript object>
```

```
<IPython.core.display.HTML object>
```

4 Heatmaps

[26]:
```
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)
```

```
<IPython.core.display.Javascript object>
```

```
<IPython.core.display.HTML object>
```

[27]:
```
plt.figure()
_ = plt.hist2d(X, Y, bins=100)
```

```
<IPython.core.display.Javascript object>
```

```
<IPython.core.display.HTML object>
```

[28]:
```
# add a colorbar legend
plt.colorbar()
```

[28]:
```
<matplotlib.colorbar.Colorbar at 0x7f90eede1940>
```

5 Animations

[29]:
```
import matplotlib.animation as animation

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

[30]:
```
# 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])
```

[31]:
```
fig = plt.figure()
a = animation.FuncAnimation(fig, update, interval=100)
```

```
<IPython.core.display.Javascript object>
```

```
<IPython.core.display.HTML object>
```

6 Interactivity

[32]:
```
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, event.xdata, event.ydata))

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

```
<IPython.core.display.Javascript object>
```

```
<IPython.core.display.HTML object>
```

[32]:
```
7
```

[33]:
```
from random import shuffle
origins = ['China', 'Brazil', 'India', 'USA', 'Canada', 'UK', 'Germany',␣
 ↪'Iraq', 'Chile', 'Mexico']

shuffle(origins)

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

[33]:

	height	origin	weight
0	0.604113	India	0.925326
1	0.728469	China	0.193065
2	0.645815	Germany	0.276170
3	0.718207	USA	0.748724
4	0.326398	Iraq	0.105455
5	0.954235	Brazil	0.258067
6	0.681956	Mexico	0.938156
7	0.402718	Chile	0.641646
8	0.048826	Canada	0.480128
9	0.012536	UK	0.156027

[34]:
```
plt.figure()
# picker=5 means the mouse doesn't have to click directly on an event, but can␣
 ↪be up to 5 pixels away
plt.scatter(df['height'], df['weight'], picker=5)
plt.gca().set_ylabel('Weight')
plt.gca().set_xlabel('Height')
```

```
<IPython.core.display.Javascript object>
```

```
<IPython.core.display.HTML object>
```

[34]:
```
<matplotlib.text.Text at 0x7f90eed30ac8>
```

[35]:
```
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)
```

[35]:
```
7
```