Week3

July 6, 2020

1 Subplots

```
In [1]: %matplotlib notebook

        import matplotlib.pyplot as plt
        import numpy as np

        plt.subplot?

In [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>


Out[2]: [<matplotlib.lines.Line2D at 0x7f27de34ffd0>]

In [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')

Out[3]: [<matplotlib.lines.Line2D at 0x7f27de0cf7f0>]

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

Out[4]: [<matplotlib.lines.Line2D at 0x7f27de0cfa58>]
```

```
In [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>


Out[5]: [<matplotlib.lines.Line2D at 0x7f27ddf5f048>]
```

```
In [10]: 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>


Out[10]: True
```

```
In [21]: # create a 3x3 grid of subplots
         fig, ((ax1,ax2,ax3), (ax4,ax5,ax6), (ax7,ax8,ax9)) = plt.subplots(3, 3, sh
         # plot the linear_data on the 5th subplot axes
         ax5.plot(linear_data, '-')
```

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


<IPython.core.display.HTML object>


Out[21]: [<matplotlib.lines.Line2D at 0x7f27dc37cb70>]
```

```
In [44]: # 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 [45]: # necessary on some systems to update the plot
         plt.gcf().canvas.draw()
```

2 Histograms

```
In [46]: # 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 a
         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>
```

```
In [47]: # 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>
```

```
In [48]: plt.figure()
         Y = np.random.normal(loc=0.0, scale=1.0, size=10000)
         X = np.random.random(size=10000)
         plt.scatter(X,Y)
```

```
/opt/conda/lib/python3.6/site-packages/matplotlib/pyplot.py:524: RuntimeWarning: Mo
  max_open_warning, RuntimeWarning)
```

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

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

```
Out[48]: <matplotlib.collections.PathCollection at 0x7f27d230add8>
```

```
In [49]: # 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:])
```

```
/opt/conda/lib/python3.6/site-packages/matplotlib/pyplot.py:524: RuntimeWarning: Mo
  max_open_warning, RuntimeWarning)
```

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

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

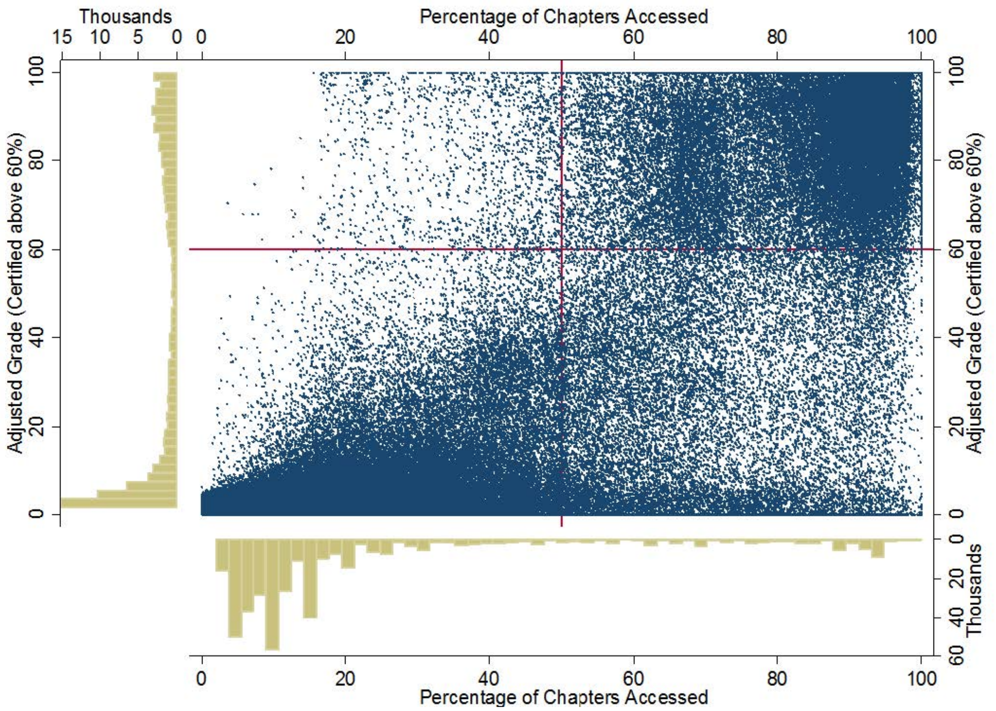
```
In [50]: Y = np.random.normal(loc=0.0, scale=1.0, size=10000)
         X = np.random.random(size=10000)
         lower_right.scatter(X, Y)
         top_histogram.hist(X, bins=100)
         s = side_histogram.hist(Y, bins=100, orientation='horizontal')
```

```
In [51]: # 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 [52]: # 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)
```

3 Box and Whisker Plots

```
In [53]: 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)
```



MOOC DATA

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

```
In [54]: df.describe()
```

```
Out[54]:                gamma        normal        random
         count   10000.000000  10000.000000  10000.000000
         mean        1.987223      0.004812      0.502268
         std         1.392530      1.010492      0.289917
         min         0.012532     -4.638756      0.000024
         25%         0.964329     -0.681653      0.250757
         50%         1.678512      0.005307      0.503273
         75%         2.696285      0.689283      0.756608
         max        11.520185      4.206962      0.999958
```

```
In [55]: plt.figure()
         # create a boxplot of the normal data, assign the output to a variable to
         _ = plt.boxplot(df['normal'], whis='range')
```

```
/opt/conda/lib/python3.6/site-packages/matplotlib/pyplot.py:524: RuntimeWarning: Mo
  max_open_warning, RuntimeWarning)
```

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

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

```
In [56]: # 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 [57]: plt.figure()
         _ = plt.hist(df['gamma'], bins=100)
```

```
/opt/conda/lib/python3.6/site-packages/matplotlib/pyplot.py:524: RuntimeWarning: Mo
  max_open_warning, RuntimeWarning)
```

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

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

```
In [58]: 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)
```

```
/opt/conda/lib/python3.6/site-packages/matplotlib/pyplot.py:524: RuntimeWarning: Mo
  max_open_warning, RuntimeWarning)
```

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

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

```
In [59]: # switch the y axis ticks for ax2 to the right side
         ax2.yaxis.tick_right()
```

```
In [60]: # if `whis` argument isn't passed, boxplot defaults to showing 1.5*interqu
         plt.figure()
         _ = plt.boxplot([ df['normal'], df['random'], df['gamma'] ] )
```

```
/opt/conda/lib/python3.6/site-packages/matplotlib/pyplot.py:524: RuntimeWarning: Mo
  max_open_warning, RuntimeWarning)
```

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

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

4 Heatmaps

```
In [61]: 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)
```

```
/opt/conda/lib/python3.6/site-packages/matplotlib/pyplot.py:524: RuntimeWarning: Mo
  max_open_warning, RuntimeWarning)
```

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

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

```
In [62]: plt.figure()
         _ = plt.hist2d(X, Y, bins=100)
```

```
/opt/conda/lib/python3.6/site-packages/matplotlib/pyplot.py:524: RuntimeWarning: Mo
  max_open_warning, RuntimeWarning)
```

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

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

```
In [63]: # add a colorbar legend
         plt.colorbar()

Out[63]: <matplotlib.colorbar.Colorbar at 0x7f27d14b6a20>
```

5 Animations

```
In [64]: import matplotlib.animation as animation

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

```
In [65]: # create the function that will do the plotting, where curr is the current
         def update(curr):
             # check if animation is at the last frame, and if so, stop the animati
             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])
```

```
In [71]: fig = plt.figure()
         a = animation.FuncAnimation(fig, update, interval=100)
```

```
/opt/conda/lib/python3.6/site-packages/matplotlib/pyplot.py:524: RuntimeWarning: Mo
  max_open_warning, RuntimeWarning)
```

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

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

6 Interactivity

```
In [67]: 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(ev

         # tell mpl_connect we want to pass a 'button_press_event' into onclick whe
         plt.gcf().canvas.mpl_connect('button_press_event', onclick)
```

```
/opt/conda/lib/python3.6/site-packages/matplotlib/pyplot.py:524: RuntimeWarning: Mo
  max_open_warning, RuntimeWarning)
```

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

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

Out[67]: 7

In [68]: from random import shuffle
         origins = ['China', 'Brazil', 'India', 'USA', 'Canada', 'UK', 'Germany', '

         shuffle(origins)

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

Out[68]:      height   origin    weight
         0  0.595661    Chile  0.846556
         1  0.562407     Iraq  0.399659
         2  0.367942       UK  0.904431
         3  0.832223   Brazil  0.517931
         4  0.186209   Canada  0.749684
         5  0.723429    India  0.735089
         6  0.863893      USA  0.617352
         7  0.839167  Germany  0.519539
         8  0.907858   Mexico  0.567233
         9  0.890180    China  0.719626

In [69]: plt.figure()
         # picker=5 means the mouse doesn't have to click directly on an event, but
         plt.scatter(df['height'], df['weight'], picker=5)
         plt.gca().set_ylabel('Weight')
         plt.gca().set_xlabel('Height')

/opt/conda/lib/python3.6/site-packages/matplotlib/pyplot.py:524: RuntimeWarning: M
  max_open_warning, RuntimeWarning)

<IPython.core.display.Javascript object>

<IPython.core.display.HTML object>

Out[69]: <matplotlib.text.Text at 0x7f27d1419860>

In [70]: 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 eve
         plt.gcf().canvas.mpl_connect('pick_event', onpick)

Out[70]: 7
```