



PYTHON FOR R USERS

Plotting directly using pandas

Daniel Chen
Instructor



Plotting in Python

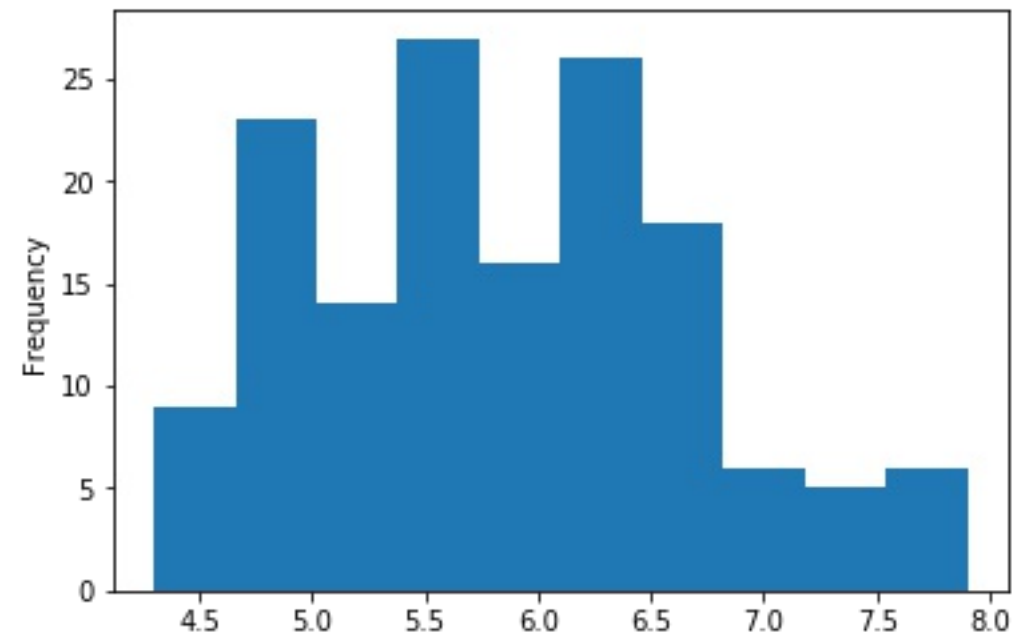
- Quickly show data patterns
- Plotting methods in Python:
 - Pandas
 - Seaborn
 - Matplotlib

Pandas plot method

- `plot()` method
- Works on the pandas DataFrame and Series objects
- Pass plot the kind argument
- kind of plots:
 - 'line' : line plot (default)
 - 'bar' : vertical bar plot
 - 'barh' : horizontal bar plot
 - 'hist' : histogram
 - 'box' : boxplot
 - 'kde' : Kernel Density Estimation plot
 - 'density' : same as 'kde'
 - 'area' : area plot
 - 'pie' : pie plot
 - 'scatter' : scatter plot
 - 'hexbin' : hexbin plot

Univariate: Histogram

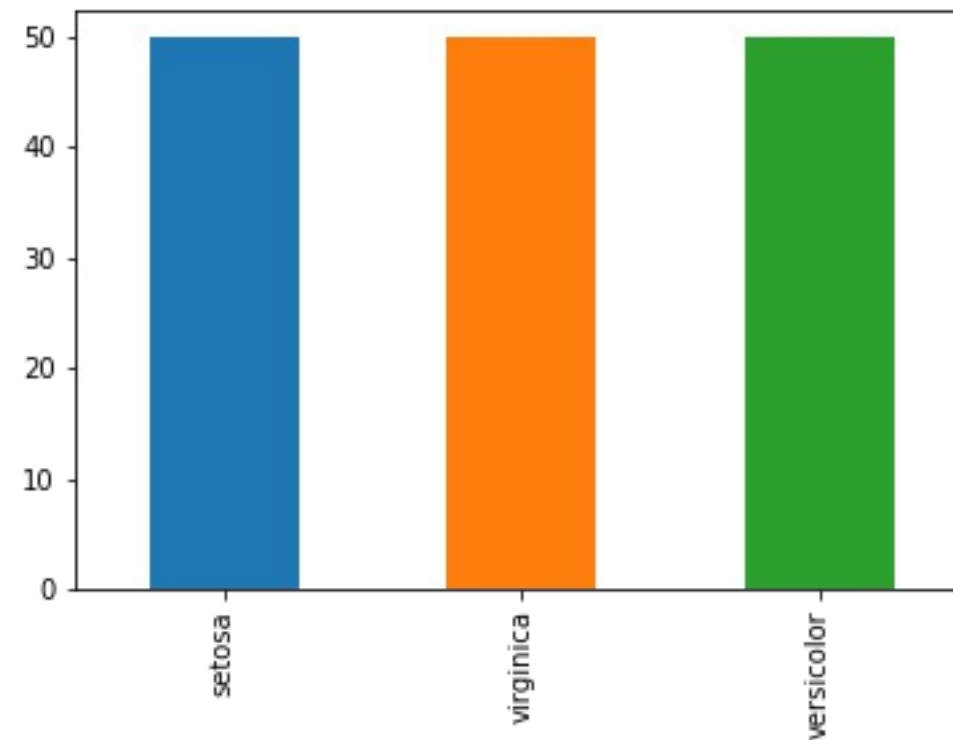
```
In [4]: import matplotlib.pyplot as plt
In [5]: iris['sepal_length'].plot(kind='hist')
...: plt.show()
...:
```





Univariate: Bar plot

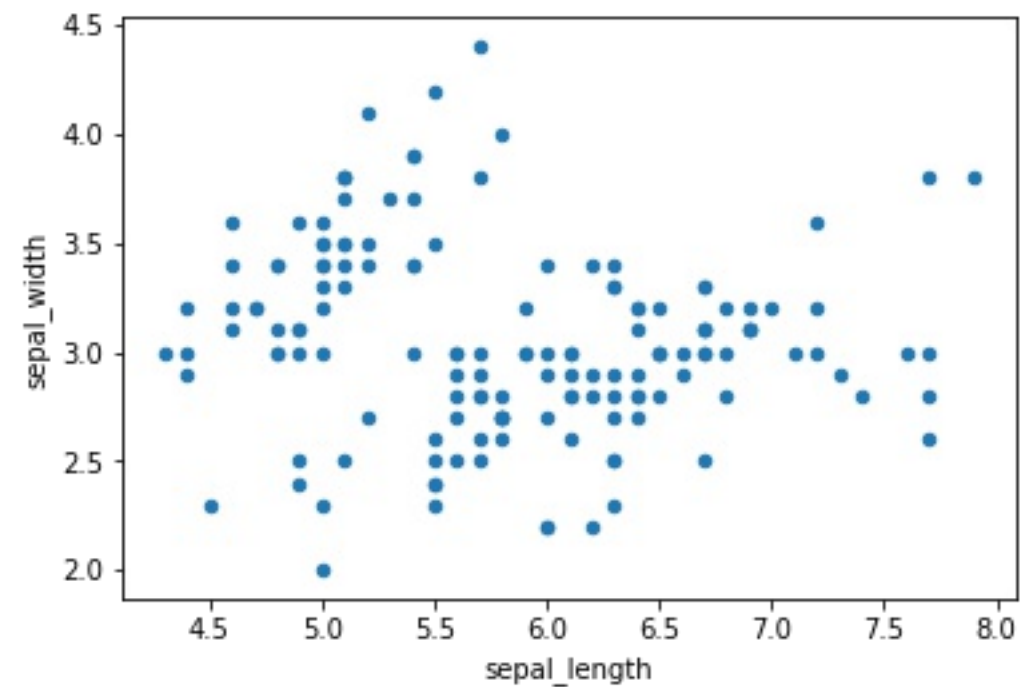
```
In [6]: cts = iris['species'].value_counts()  
...: cts.plot(kind='bar')  
...: plt.show()  
...:
```





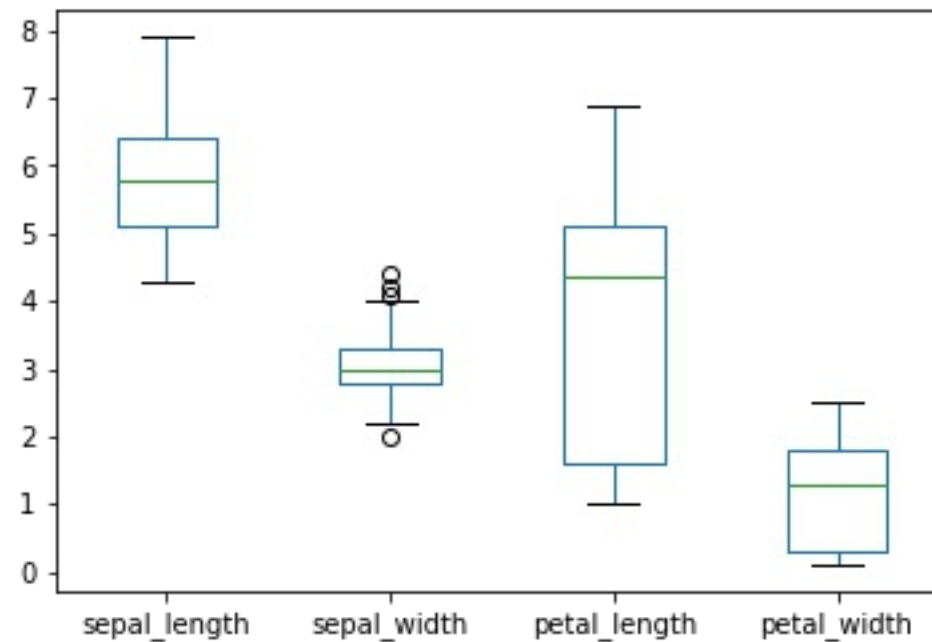
Bivariate: Scatter plot

```
In [8]: iris.plot(kind='scatter', x='Sepal.Length', y='Sepal.Width')  
...: plt.show()
```



Bivariate: Boxplots

```
In [7]: iris.plot(kind='box')  
...: plt.show()  
...:
```





Bivariate: Boxplots

```
In [10]: iris.boxplot(by='Species', column='Sepal.Length')  
...: plt.show()  
...:
```





PYTHON FOR R USERS

Let's practice!



PYTHON FOR R USERS

Seaborn

Daniel Chen
Instructor



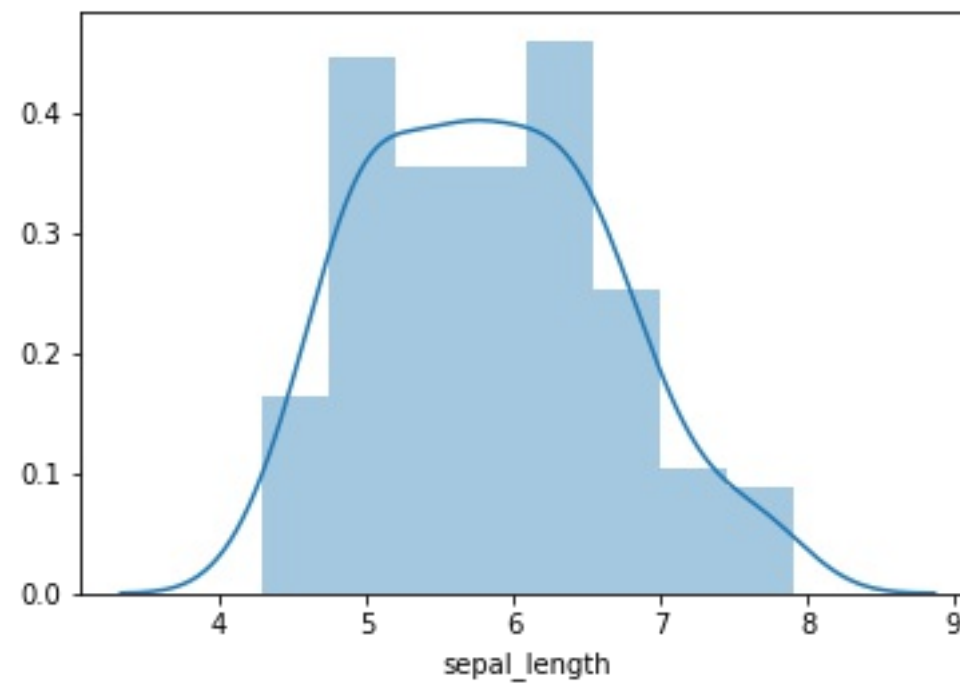
Seaborn

- barplots: barplot
- histograms: displot
- boxplot: boxplot
- scatter plot: regplot
- Seaborn benefits
 - colored points by data
 - facet plots by data



Seaborn histograms

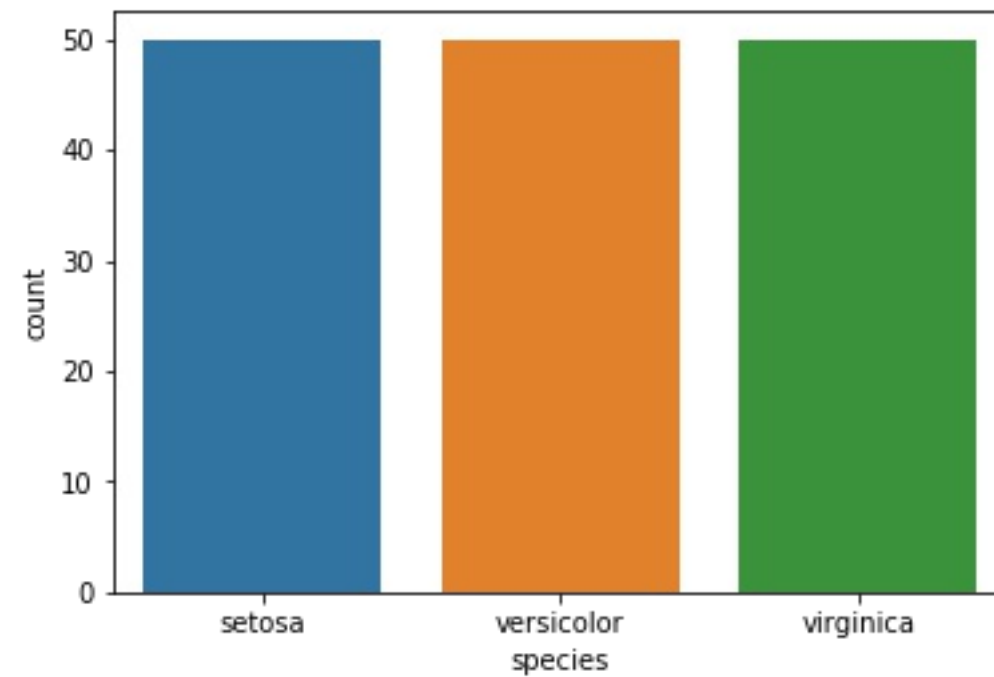
```
In [13]: import seaborn as sns
....: import matplotlib.pyplot as plt
....: sns.distplot(iris['sepal_length'])
....: plt.show()
....:
```





Seaborn count plot

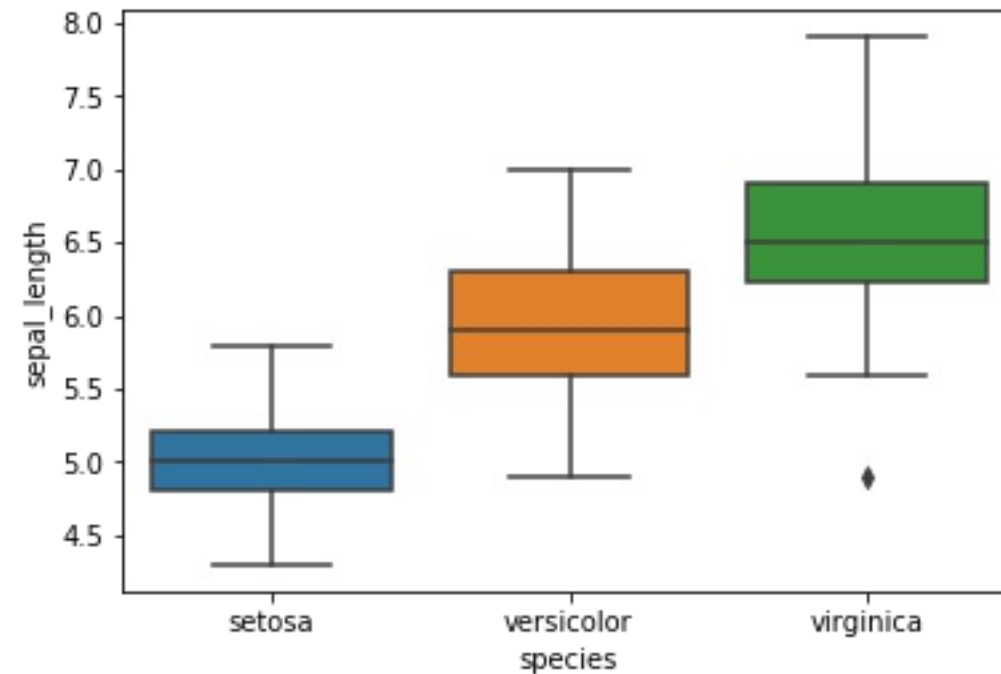
```
In [12]: sns.countplot('species', data=iris)
....: plt.show()
```





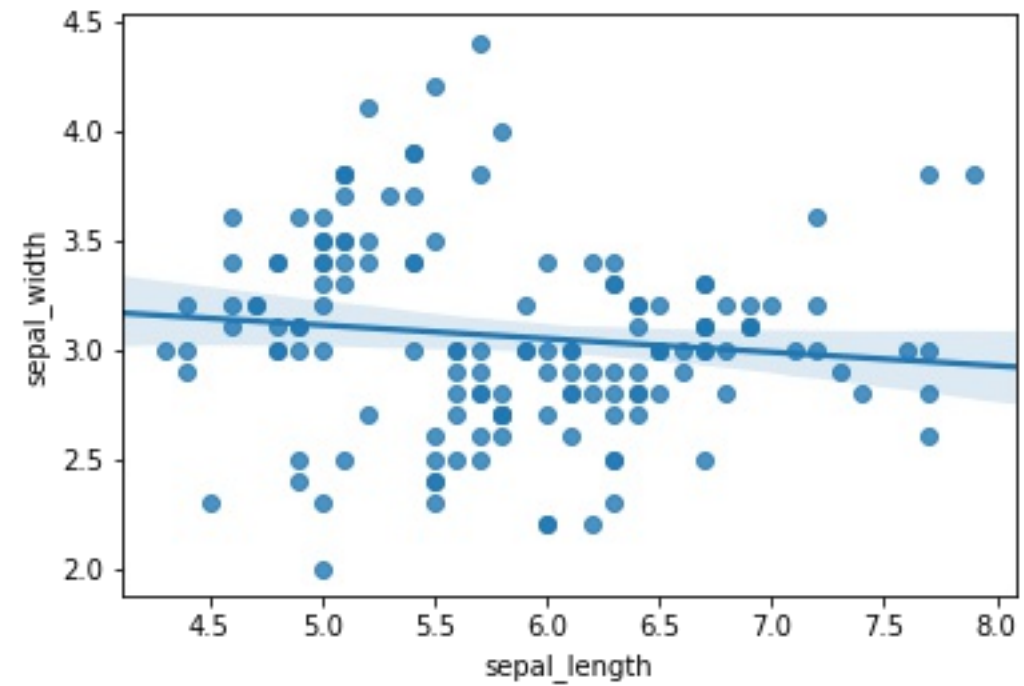
Seaborn boxplots

```
In [14]: sns.boxplot(x='species', y='sepal_length', data=iris)
...: plt.show()
...:
```



Seaborn scatterplots

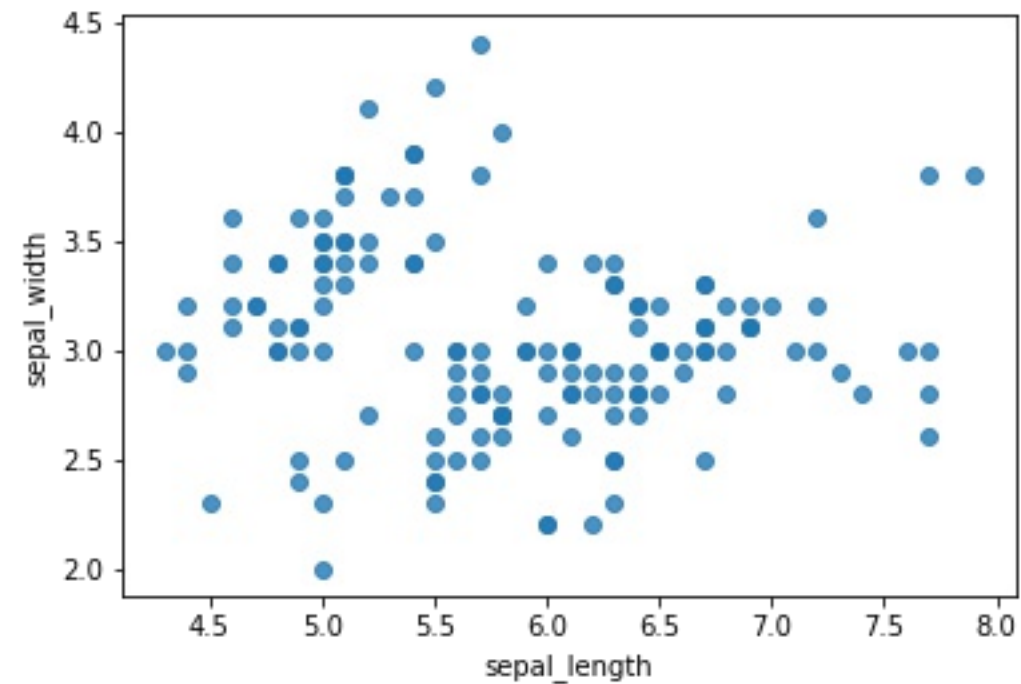
```
In [15]: sns.regplot(x='sepal_length', y='sepal_width', data=iris)
....: plt.show()
....:
```





Seaborn scatterplots w/out regression line

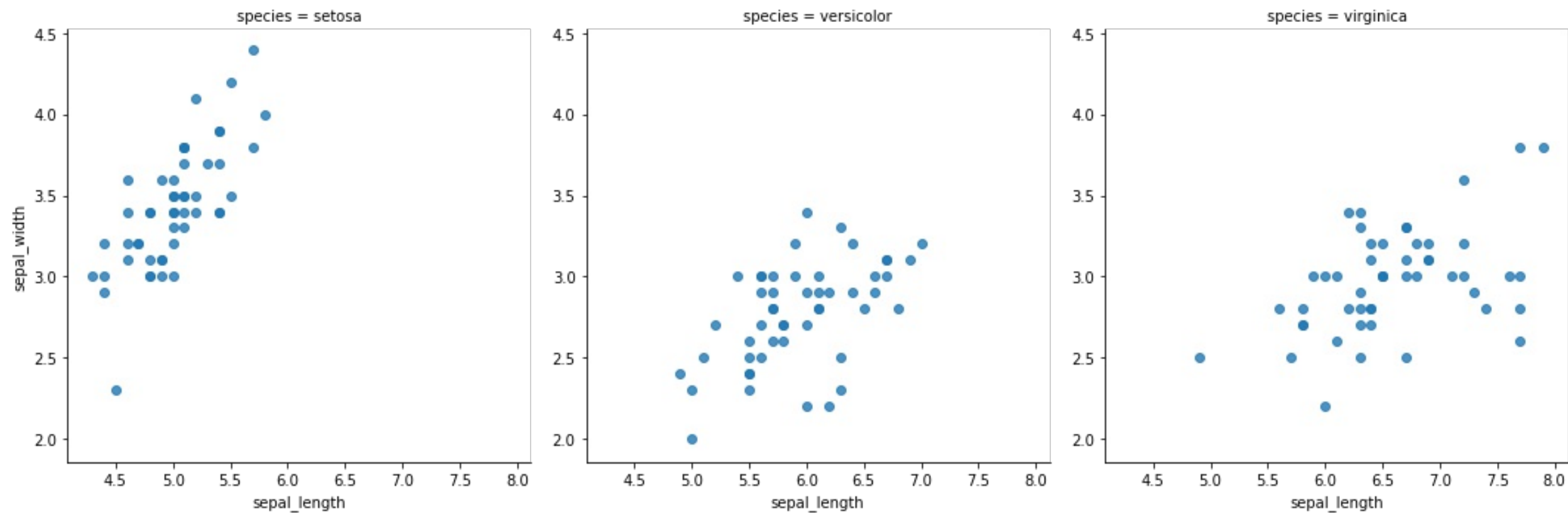
```
In [15]: sns.regplot(x='sepal_length', y='sepal_width', data=iris,  
....:               fit_reg=False)  
....: plt.show()  
....:
```





Seaborn Facets

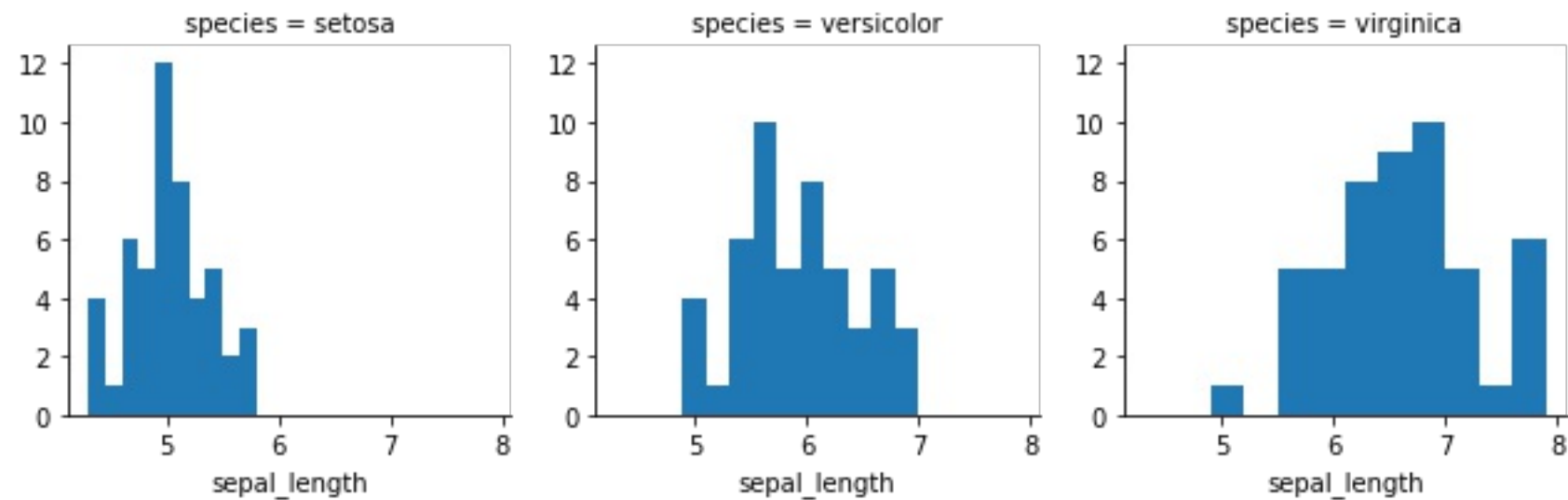
```
In [16]: sns.lmplot(x='sepal_length', y='sepal_width', data=iris,  
....:               fit_reg=False,  
....:               col='species')  
....: plt.show()  
....:
```





Seaborn FacetGrid

```
In [21]: g = sns.FacetGrid(iris, col="species")  
....: g = g.map(plt.hist, "sepal_length")  
....: plt.show()
```





PYTHON FOR R USERS

Let's practice!



PYTHON FOR R USERS

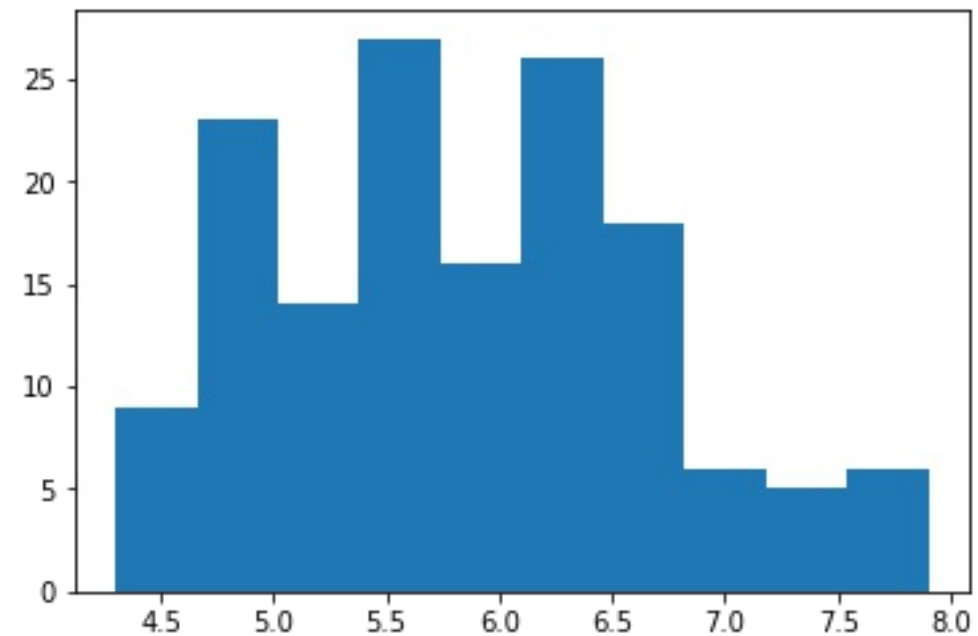
Matplotlib

Daniel Chen
Instructor



Matplotlib plots

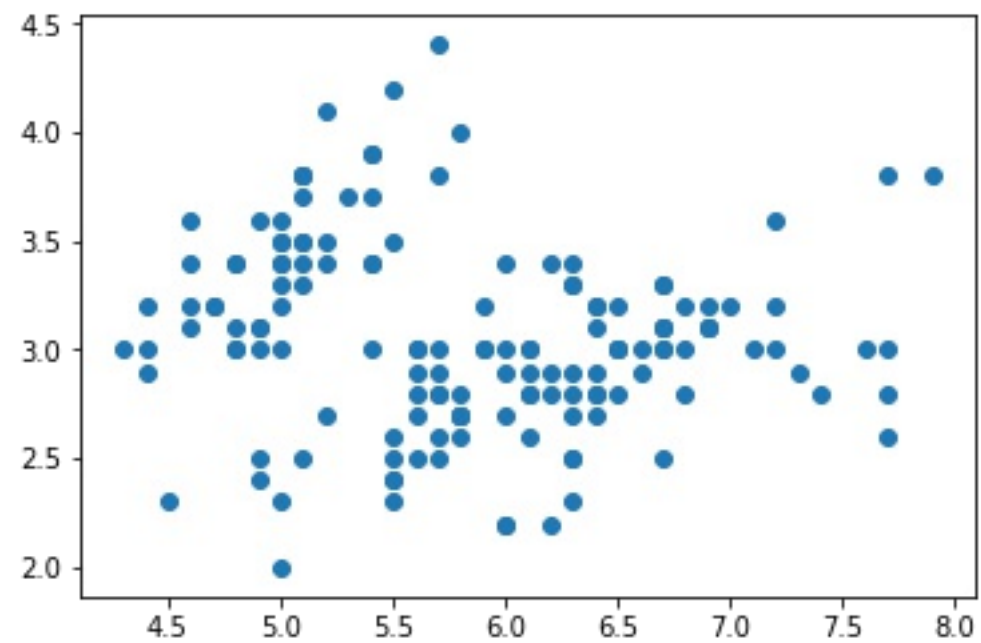
```
In [12]: import matplotlib.pyplot as plt  
....: plt.hist(iris['sepal_length'])  
....: plt.show()
```





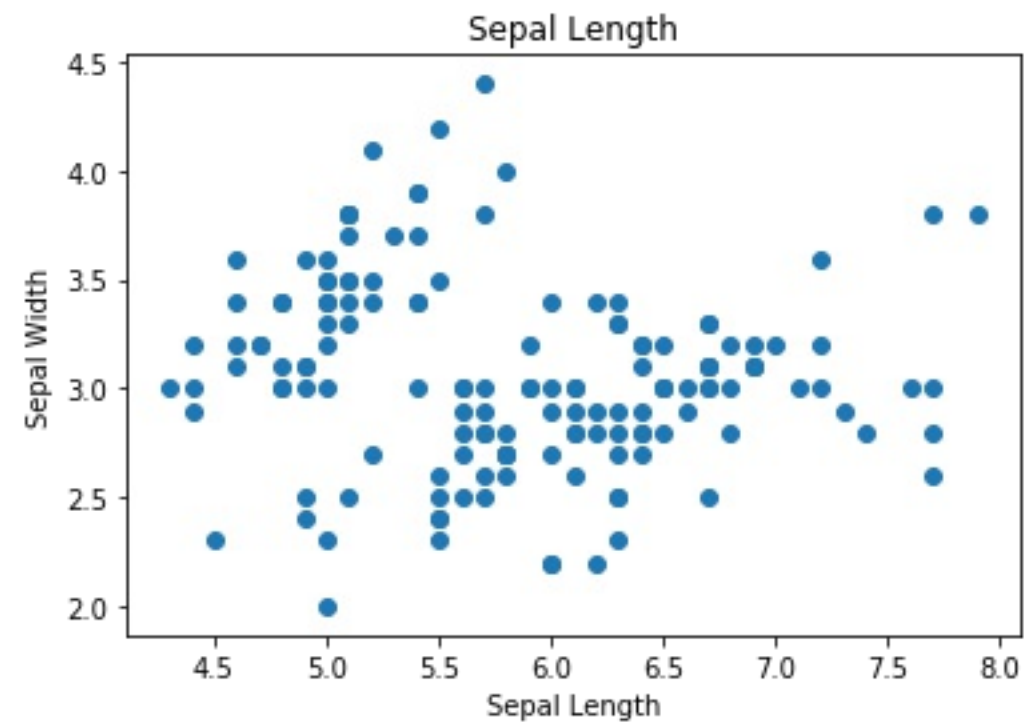
Matplotlib scatter

```
In [13]: plt.scatter(iris['sepal_length'], iris['sepal_width'])  
....: plt.show()
```



Polishing up the figure

```
In [7]: fig, ax = plt.subplots()
....: ax.scatter(iris['sepal_length'], iris['sepal_width'])
....: ax.set_title('Sepal Length')
....: ax.set_xlabel('Sepal Length')
....: ax.set_ylabel('Sepal Width')
....: plt.show()
```



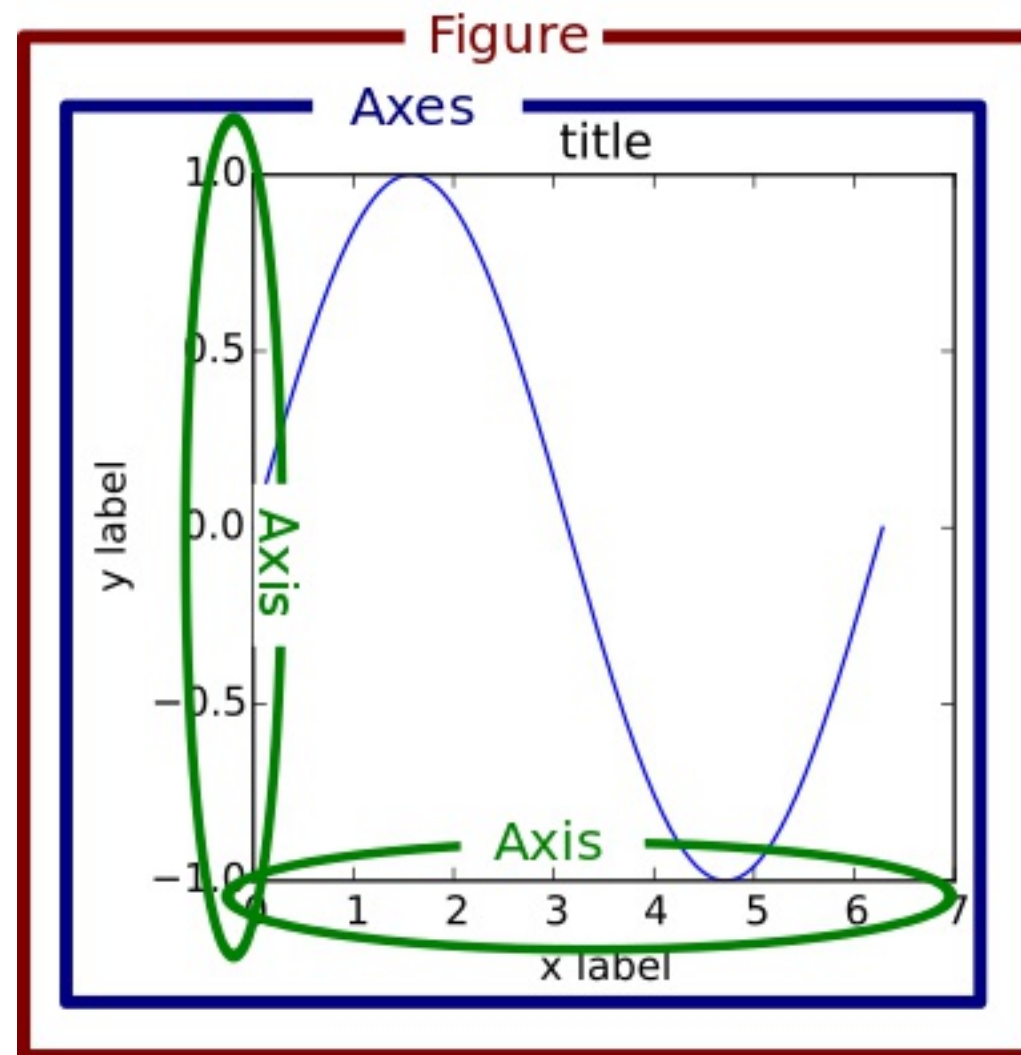


Rotating axis ticks

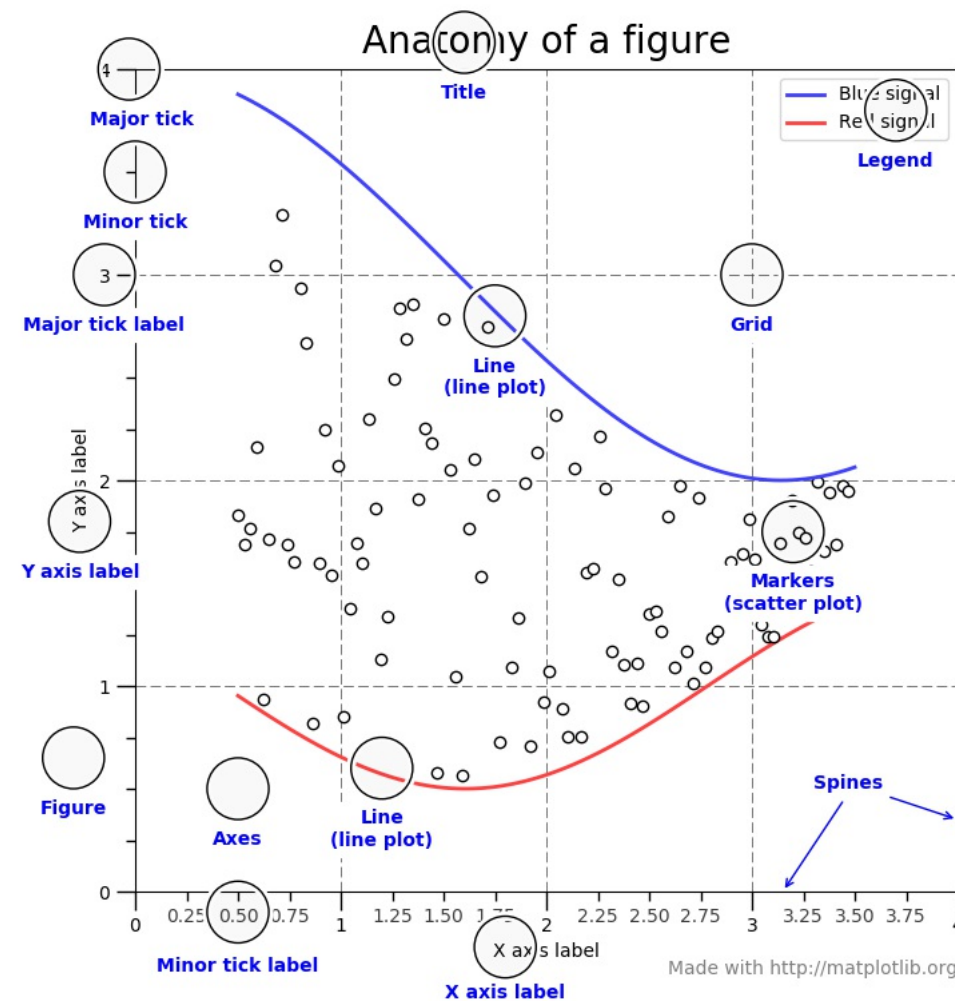
```
In [6]: fig, ax = plt.subplots()
...: ax.scatter(iris['sepal_length'], iris['sepal_width'])
...: ax.set_title('Sepal Length')
...: ax.set_xlabel('Sepal Length')
...: ax.set_ylabel('Sepal Width')
...: plt.xticks(rotation=45) # rotate the x-axis ticks
...: plt.show()
```



Parts of a matplotlib figure



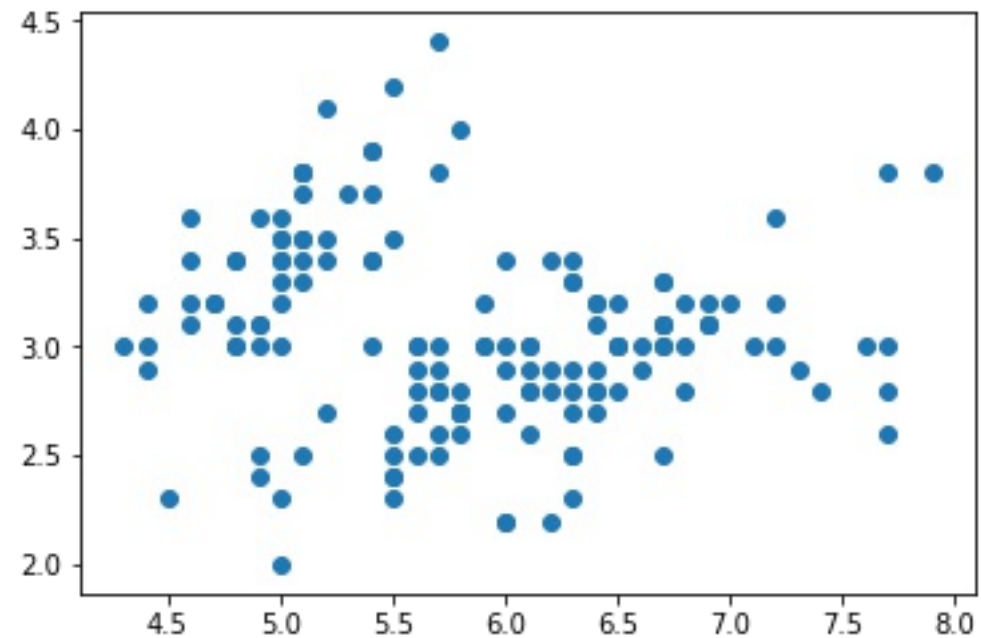
Parts of a matplotlib figure 2





Figures and axes

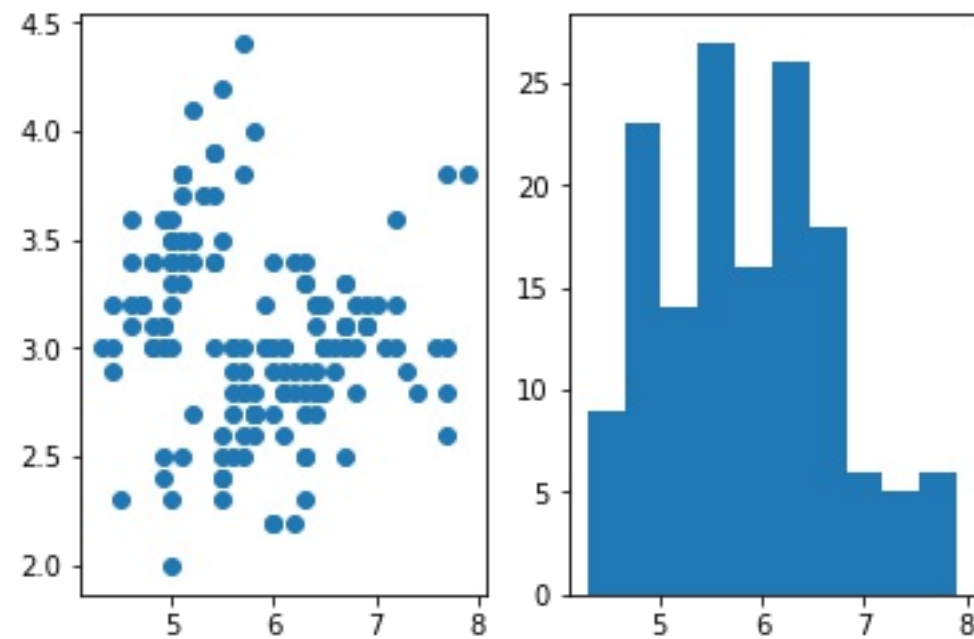
```
In [11]: fig, ax = plt.subplots()
....: ax.scatter(iris['sepal_length'], iris['sepal_width'])
....: plt.show()
```





Multiple Axes'

```
In [9]: fig, (ax1, ax2) = plt.subplots(1, 2)
...: ax1.scatter(iris['sepal_length'], iris['sepal_width'])
...: ax2.hist(iris['sepal_length'])
...: plt.show()
```





Remember

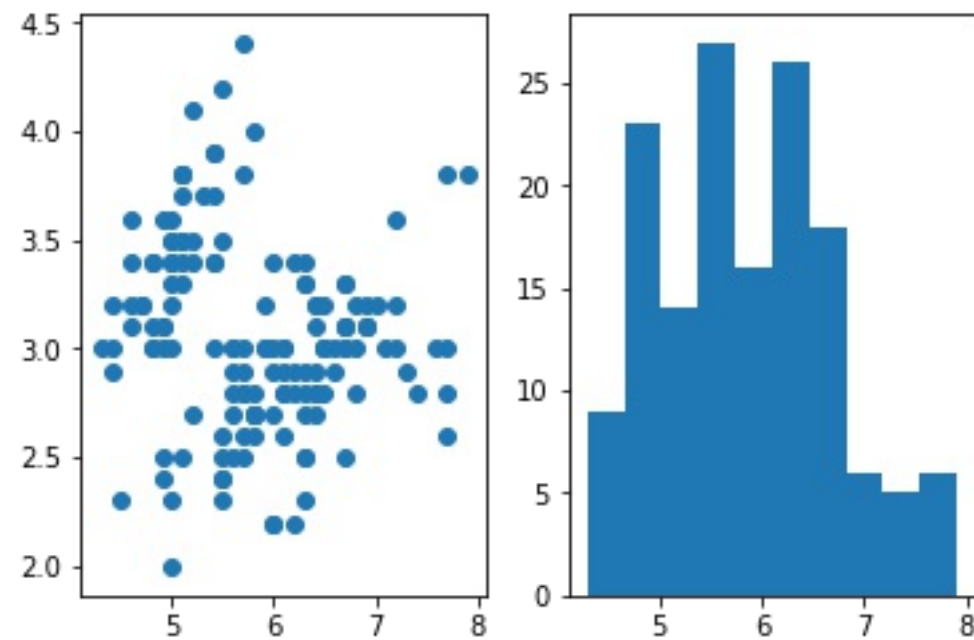
```
In [30]: fig, ax = plt.subplots()
....: sns.regplot(x='sepal_length', y='sepal_width',
....:             data=iris, fit_reg=False, ax=ax)
....: plt.show()
```





Clearing the Figure

```
In [29]: fig, (ax1, ax2) = plt.subplots(1, 2)
....: ax1.scatter(iris['sepal_length'], iris['sepal_width'])
....: ax2.hist(iris['sepal_length'])
....: plt.show()
....: plt.clf()
```





PYTHON FOR R USERS

Let's practice!