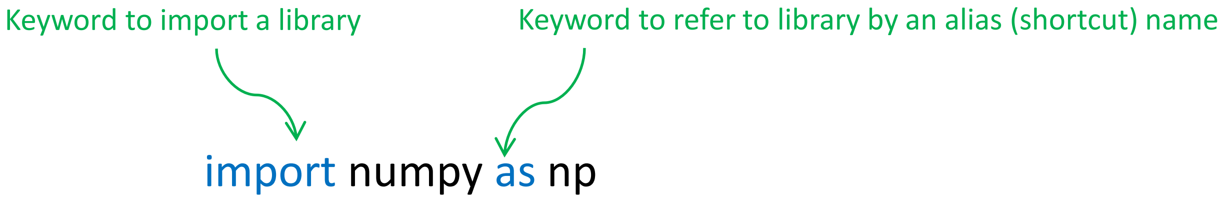
Libraries - Numpy

- **A popular math library in Python for Machine Learning is ‘numpy’.**



Keyword to import a library

Keyword to refer to library by an alias (shortcut) name

```
import numpy as np
```

Numpy.org : NumPy is the fundamental package for scientific computing with Python.

- a powerful N-dimensional array object
- sophisticated (broadcasting) functions
- tools for integrating C/C++ and Fortran code
- useful linear algebra, Fourier transform, and random number capabilities

Libraries - Numpy

http://www.physics.nyu.edu/pine/pymanual/html/chap3/chap3_arrays.html

The most import data structure for scientific computing in Python is the **NumPy array**. NumPy arrays are used to store lists of numerical data and to represent vectors, matrices, and even tensors.

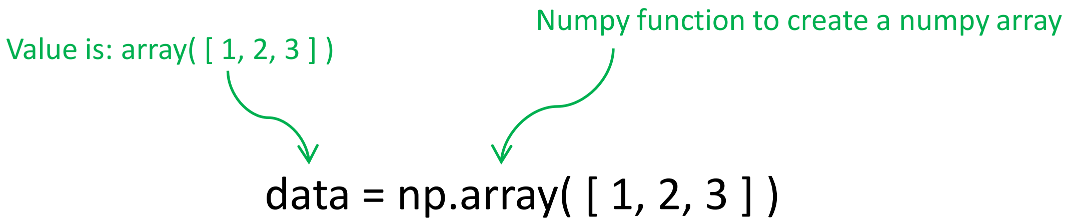
NumPy arrays are designed to handle large data sets efficiently and with a minimum of fuss. The NumPy library has a large set of routines for creating, manipulating, and transforming NumPy arrays.

Core Python has an array data structure, but it's not nearly as versatile, efficient, or useful as the NumPy array.

Numpy – Multidimensional Arrays

- **Numpy's main object is a multi-dimensional array.**

- **Creating a Numpy Array as a Vector:**

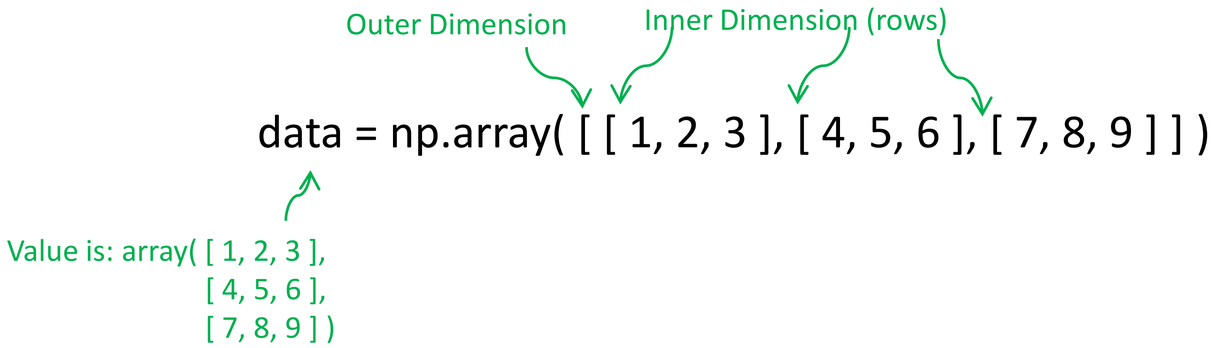


Value is: array([1, 2, 3])

Numpy function to create a numpy array

```
data = np.array( [ 1, 2, 3 ] )
```

- **Creating a Numpy Array as a Matrix:**



Outer Dimension

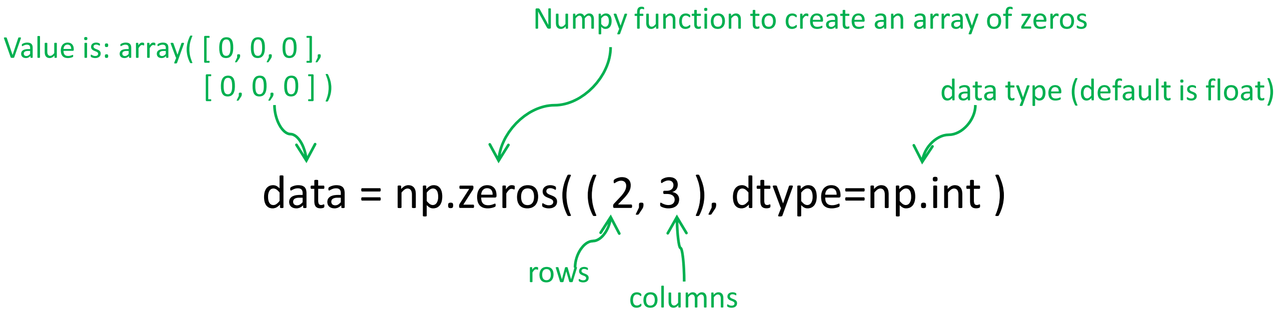
Inner Dimension (rows)

```
data = np.array( [ [ 1, 2, 3 ], [ 4, 5, 6 ], [ 7, 8, 9 ] ] )
```

Value is: array([1, 2, 3],
[4, 5, 6],
[7, 8, 9])

Numpy – Multidimensional Arrays

- **Creating an array of Zeros:**



Numpy function to create an array of zeros

Value is: array([0, 0, 0],
[0, 0, 0])

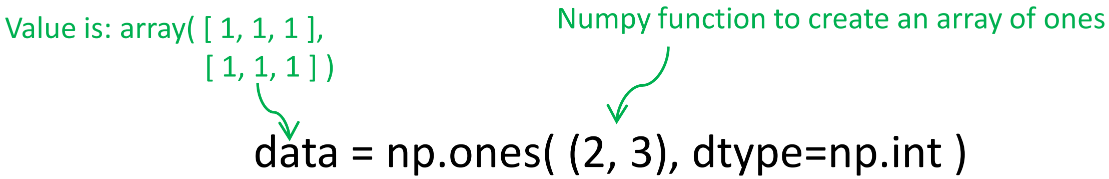
data type (default is float)

```
data = np.zeros( ( 2, 3 ), dtype=np.int )
```

rows

columns

- **Creating an array of Ones:**



Value is: array([1, 1, 1],
[1, 1, 1])

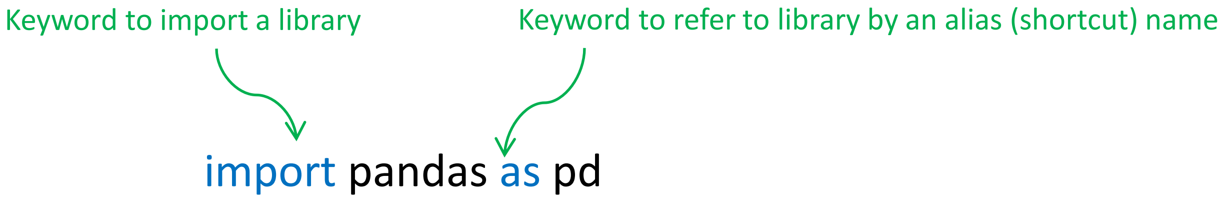
Numpy function to create an array of ones

```
data = np.ones( (2, 3), dtype=np.int )
```

And many more functions: size, ndim, reshape, arange, …

Libraries - Pandas

- **A popular library for importing and managing datasets in Python for Machine Learning is ‘pandas’.**



Keyword to import a library

Keyword to refer to library by an alias (shortcut) name

```
import pandas as pd
```

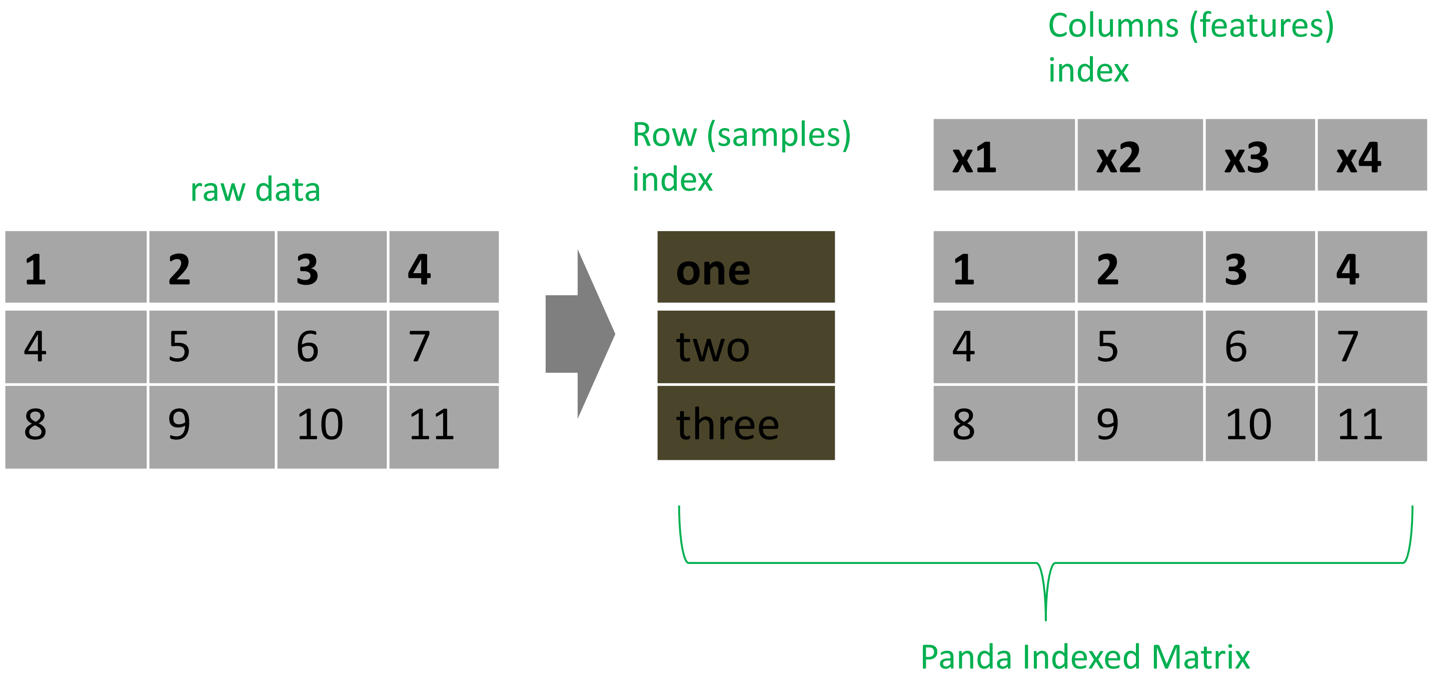
Used for:

- **Data Analysis**
- **Data Manipulation**
- **Data Visualization**

PyData.org : high-performance, easy-to-use data structures and data analysis tools for the Python programming language.

Pandas – Indexed Arrays

- **Pandas are used to build indexed arrays (1D) and matrices (2D), where columns and rows are labeled (named) and can be accessed via the labels (names).**



raw data

1	**2**	**3**	**4**
4	5	6	7
8	9	10	11

Row (samples) index

Columns (features) index

	x1	**x2**	**x3**	**x4**
one	**1**	**2**	**3**	**4**
two	4	5	6	7
three	8	9	10	11

Panda Indexed Matrix

Pandas – Series and Data Frames

- Pandas Indexed Arrays are referred to as **Series** (1D) and **Data Frames** (2D).

- **Series** is a 1D labeled (indexed) array and can hold any data type, and mix of data types.

Series Raw data Column Index Labels

```
s = pd.Series( data, index=[ ‘x1’, ‘x2’, ‘x3’, ‘x4’ ] )
```

- **Data Frame** is a 2D labeled (indexed) matrix and can hold any data type, and mix of data types.

Data Frame Row Index Labels Column Index Labels

```
df = pd.DataFrame( data, index=[‘one’, ‘two’], columns=[ ‘x1’, ‘x2’, ‘x3’, ‘x4’ ] )
```

Pandas – Selecting

- **Selecting One Column**

Selects column labeled x1 for all rows

```
x1 = df[ ‘x1’ ]
```

1
4
8

- **Selecting Multiple Columns**

Note: df[‘x1’:’x3’] this python syntax does not work!

Selects columns labeled x1 and x3 for all rows

```
x1 = df[ [ ‘x1’, ‘x3’ ] ]
```

1	**3**
4	6
8	10

Selects columns labeled x1 through x3 for all rows

```
x1 = df.ix[ :, ‘x1’:’x3’ ]
```

rows (all)

columns

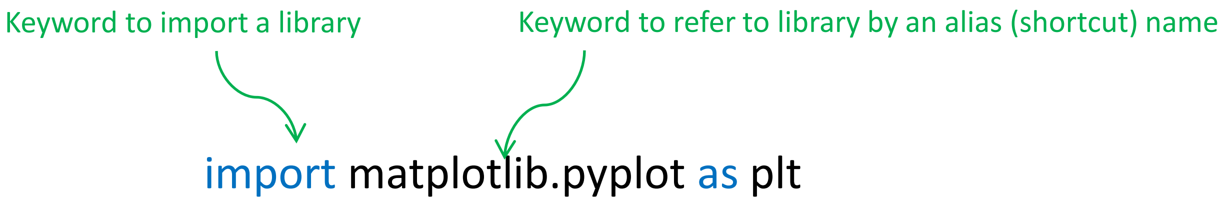
Slicing function

1	**2**	**3**
4	5	6
8	9	10

And many more functions: merge, concat, stack, …

Libraries - Matplotlib

- **A popular library for plotting and visualizing data in Python**



Keyword to import a library

Keyword to refer to library by an alias (shortcut) name

```
import matplotlib.pyplot as plt
```

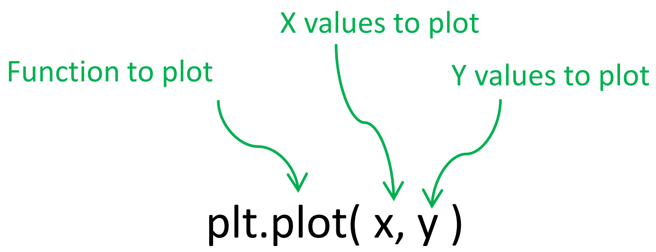
Used for:

- **Plots**
- **Histograms**
- **Bar Charts**
- **Scatter Plots**
- **etc**

matplotlib.org: Matplotlib is a Python 2D plotting library which produces publication quality figures in a variety of hardcopy formats and interactive environments across platforms.

Matplotlib - Plot

- **The function plot plots a 2D graph.**



Function to plot

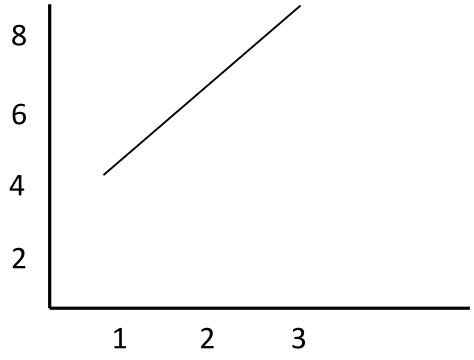
X values to plot

Y values to plot

```
plt.plot( x, y )
```

- **Example:**

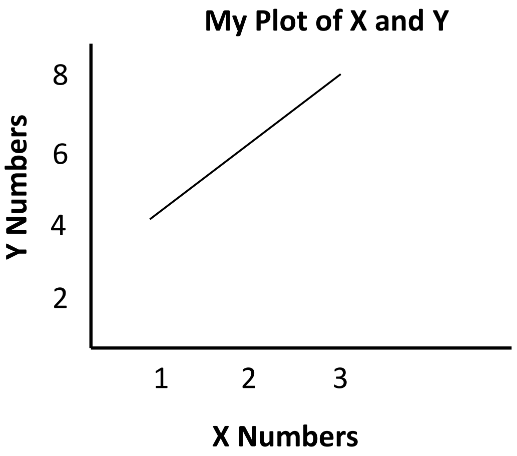
```
            X              Y
plt.plot( [ 1, 2, 3 ], [ 4, 6, 8 ] ) # Draws plot in the background
plt.show()                           # Displays the plot
```



Matplotlib – Plot Labels

- **Add Labels for X and Y Axis and Plot Title (caption)**

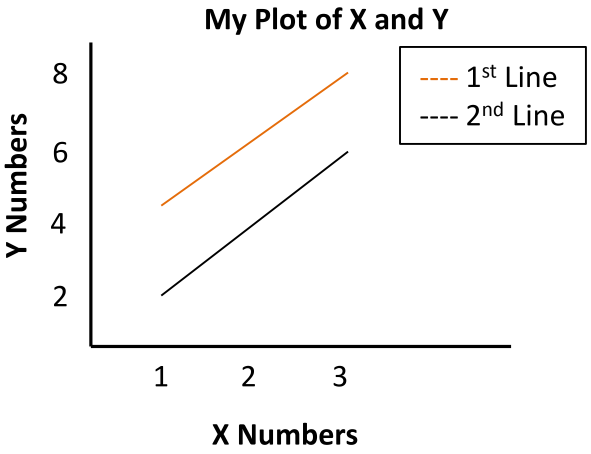
```
plt.plot( [ 1, 2, 3 ], [ 4, 6, 8 ] )
plt.xlabel( “X Numbers” )                  # Label on the X-axis
plt.ylabel( “Y Numbers” )                  # Label on the Y-axis
plt.title( “My Plot of X and Y”)           # Title for the Plot
plt.show()
```



Matplotlib – Multiple Plots and Legend

- **You can add multiple plots in a Graph**

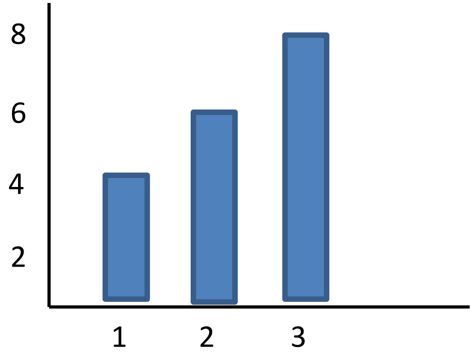
```
plt.plot( [ 1, 2, 3 ], [ 4, 6, 8 ], label=‘ 1st Line’ )     # Plot for 1st Line
plt.plot( [ 1, 2, 3 ], [ 2, 4, 6 ], label=‘2nd Line’ )      # Plot for 2nd Line
plt.xlabel( “X Numbers” )
plt.ylabel( “Y Numbers” )
plt.title( “My Plot of X and Y”)
plt.legend()                                             # Show Legend for the plots
plt.show()
```



Matplotlib – Bar Chart

- **The function bar plots a bar graph.**

```
plt.plot( [ 1, 2, 3 ], [ 4, 6, 8 ] )  # Plot for 1st Line
plt.bar()                             # Draw a bar chart
plt.show()
```



And many more functions: hist, scatter, …