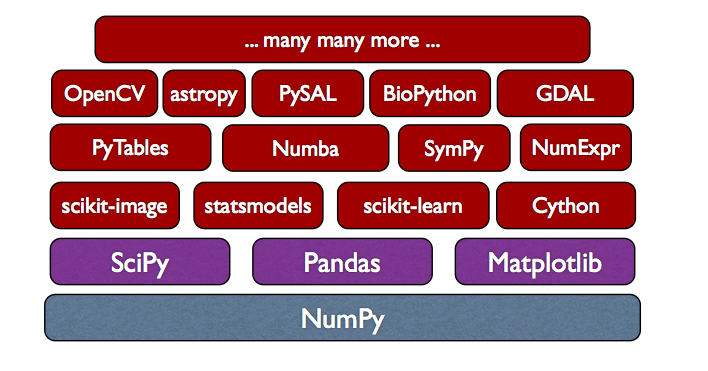
<https://scipy.org/>

<https://docs.python.org/3/tutorial/>

<https://www.tensorflow.org/>

<http://www.nltk.org/>



**FOR DATA MANUPULATION**

So, when do you use each? If you want to understand the data you have, use Pandas. If you want to use this data to train a machine-learning algorithm, use scikit-learn. If you are doing some other scientific of engineering calculation, use SciPy. If you want to plot data, use matplotlib. All of these use Numpy, so you will need to set up your data structures using Numpy.

**NumPy** is a library for efficient array computations, modeled after Matlab. Arrays differ from plain Python lists in the way they are stored and handled. Array elements stay together in memory, so they can be quickly accessed. NumPy also supports quick subindexing, e.g., a[0, :, 2] gives you all array elements whose first index is 0 and third index is 2.Furthermore, NumPy provides vectorized mathematical functions. When, e.g., you call numpy.sin(a), the sine function is applied on every element of array a. This is done using compiled C code, so it works much faster than a Python for loop, even faster than list comprehensions.

Has funcs like:

np.arange(9).reshape(3,3)

np.random.rand(3)

np.linspace(0,51,20)

**Scipy** provides low-level mathematical and scientific functions. SciPy provides a large menu of libraries for scientific computation, such as integration, interpolation, signal processing, linear algebra, statistics, etc. It is built upon the infrastructure of Numpy.

Has funcs like:

linalg.det(A)

EW, EV = linalg.eig(A)

**Pandas.**Software library written for data manipulation and analysis in Python. Offers data structures and operations for manipulating numerical tables and time series.  It's able to describe the data for you. It can do grouping and pivot tables on larger data

sal=pd.read\_csv('Salaries.csv')

sal.groupby('Year')['BasePay'].mean()

sal.iloc[sal['TotalPayBenefits'].idxmax()]

sum(sal[sal['Year']==2013]['JobTitle'].value\_counts() == 1)

**FOR DATA VISUALIZATION**

**Matplotlib**

Matplotlib is a Python 2D plotting library which produces publication quality figures in a variety of hardcopy formats and interactive environments across platforms For simple plotting the pyplot module provides a MATLAB-like interface, particularly when combined with IPython.

Has funcs like:

fig, axes = plt.subplots(nrows=1, ncols=2,figsize=(12,2))

axes[1].plot(x,z,color="red", lw=3, ls='--')

**Seaborn** is a Python visualization library based on matplotlib. It provides a high-level interface for drawing attractive statistical graphics.

sns.jointplot(x='fare',y='age',data=titanic)

sns.distplot(titanic['fare'],bins=30,kde=False,color='red')

sns.swarmplot(x='class',y='age',data=titanic,palette='Set2')

sns.countplot(x='sex',data=titanic)

IMP func

sns.lmplot(x='Month',y='twp',data=byMonth.reset\_index())

sns.heatmap(titanic.corr(),cmap='coolwarm')

g = sns.FacetGrid(data=titanic,col='sex') ; g.map(plt.hist,'age')

**Plotly** (cufflinks)is a library that allows you to create interactive plots that you can use in dashboards or websites (you can save them as html files or static images).Also offline plotly is used for plottingfigures on worldmap .check the ipynb

Has problem in connectivity

, has fuctions like

df.iplot(kind='scatter',x='A',y='B',mode='markers',size=10)# Similar to sns.pairplot()

df.iplot(kind='box')

rest refer the notebook

**Pandas in built**

**FOR CORE MACHINE LEARNING**

**scikit**-learn provides high-level machine learning functions.

It also is built upon Numpy and SciPy.

Has funcs like

from sklearn.model\_selection import train\_test\_split

from sklearn.metrics import classification\_report,confusion\_matrix

from sklearn.linear\_model import LogisticRegression

from sklearn.preprocessing import StandardScaler #Feature scaling

from sklearn.svm import SVC

from sklearn.decomposition import PCA

from sklearn.cluster import KMeans

from sklearn.neighbors import KNeighborsClassifier

**Natural Language Toolkit**, or more commonly **NLTK**, is a suite of [libraries](https://en.wikipedia.org/wiki/Library_(computer_science)) and programs for symbolic and statistical [natural language processing](https://en.wikipedia.org/wiki/Natural_language_processing) (NLP) for English written in the [Python programming language](https://en.wikipedia.org/wiki/Python_(programming_language)).

Has funcs like

from nltk.corpus import stopwords

**tensorflow-**