README

For my Capstone project I used Jupyter notebook and Python 3 through Anaconda.

The libraries I used were :

sklearn: Features various classification, regression and clustering algorithm
numpy : Adding support for large, multidimensional arrays and matrices
Pandas: Provides fast, flexible and expressive data structures
matplotlib:plotting library for Python and its numerical extension Numpy
datetime :A module that supplies classes for manipulating dates and time
*seaborn:*A visualization library , provides high-level interface for drawing attractive statistical graphics
math : Provides standard mathematical functions for Python
Statsmodel: Provides statistical support in Python

sklearn

The dataset I used was retrieved from Yahoo! finance . I access this by web.DataReader function. In my proposal I had intended to use the FTSE Russell 3000 index as the instrument but I could not obtain volume information so I settled to use General Electric (GE) as the stock to analyze for this model. Any stock that is traded and its data is archived by Yahoo! Finance can be used instead of GE.