STAT222

Unknown Author

March 5, 2014

In [70]:
```
import os
print os.getcwd()
os.chdir("/Users/edwsurewin/Dropbox/Berkeley MA/222/ipython")

from IPython.display import HTML
from numpy import nan
from numpy.random import randn
import numpy as np
np.set_printoptions(precision=2)
import pandas as pd
from pandas.core.common import adjoin
from pandas.io.data import DataReader
import pandas.util.testing as tm
tm.N = 10
import statsmodels.api as sm
import matplotlib.pyplot as plt
from matplotlib import cm
import matplotlib as mpl
import scipy.stats as stats
pd.options.display.max_columns=80
from StringIO import StringIO
import requests
from statsmodels.graphics.api import qqplot
```

```
/Users/edwsurewin/Dropbox/Berkeley MA/222/ipython
```

In [71]:
```
from pandas import DataFrame, read_csv
import pandas as pd #only needed to determine version number
print 'Pandas version ' + pd.__version__
```

```
Pandas version 0.13.1
```

In [96]:
```
import csv
location = requests.get('http://ichart.finance.yahoo.com/table.csv?s=%5EGSPC&d=2&e=6&f
data = location.content
df = pd.read_csv(StringIO(data),index_col=0,parse_dates=['Date']) #converts to numbers
df
```

Out [96]:

Date	Open	High	Low	Close	Volume	Adj Close
2014-03-05	1874.05	1876.53	1871.11	1873.81	3392990000	1873.81
2014-03-04	1849.23	1876.23	1849.23	1873.91	3765770000	1873.91
2014-03-03	1857.68	1857.68	1834.44	1845.73	3428220000	1845.73
2014-02-28	1855.12	1867.92	1847.67	1859.45	3917450000	1859.45
2014-02-27	1844.90	1854.53	1841.13	1854.29	3547460000	1854.29

2014-02-26	1845.79	1852.65	1840.66	1845.16	3716730000	1845.16
2014-02-25	1847.66	1852.91	1840.19	1845.12	3515560000	1845.12
2014-02-24	1836.78	1858.71	1836.78	1847.61	4014530000	1847.61
2014-02-21	1841.07	1846.13	1835.60	1836.25	3403880000	1836.25
2014-02-20	1829.24	1842.79	1824.58	1839.78	3404980000	1839.78
2014-02-19	1838.90	1847.50	1826.99	1828.75	3661570000	1828.75
2014-02-18	1839.03	1842.87	1835.01	1840.76	3421110000	1840.76
2014-02-14	1828.46	1841.65	1825.59	1838.63	3114750000	1838.63
2014-02-13	1814.82	1830.25	1809.22	1829.83	3289510000	1829.83
2014-02-12	1820.12	1826.55	1815.97	1819.26	3326380000	1819.26
2014-02-11	1800.45	1823.54	1800.41	1819.75	3699380000	1819.75
2014-02-10	1796.20	1799.94	1791.83	1799.84	3312160000	1799.84
2014-02-07	1776.01	1798.03	1776.01	1797.02	3775990000	1797.02
2014-02-06	1752.99	1774.06	1752.99	1773.43	3825410000	1773.43
2014-02-05	1753.38	1755.79	1737.92	1751.64	3984290000	1751.64
2014-02-04	1743.82	1758.73	1743.82	1755.20	4068410000	1755.20
2014-02-03	1782.68	1784.83	1739.66	1741.89	4726040000	1741.89
2014-01-31	1790.88	1793.88	1772.26	1782.59	4059690000	1782.59
2014-01-30	1777.17	1798.77	1777.17	1794.19	3547510000	1794.19
2014-01-29	1790.15	1790.15	1770.45	1774.20	3964020000	1774.20
2014-01-28	1783.00	1793.87	1779.49	1792.50	3437830000	1792.50
2014-01-27	1791.03	1795.98	1772.88	1781.56	4045200000	1781.56
2014-01-24	1826.96	1826.96	1790.29	1790.29	4618450000	1790.29
2014-01-23	1842.29	1842.29	1820.06	1828.46	3972250000	1828.46
2014-01-22	1844.71	1846.87	1840.88	1844.86	3374170000	1844.86
2014-01-21	1841.05	1849.31	1832.38	1843.80	3782470000	1843.80
2014-01-17	1844.23	1846.04	1835.23	1838.70	3626120000	1838.70
2014-01-16	1847.99	1847.99	1840.30	1845.89	3491310000	1845.89
2014-01-15	1840.52	1850.84	1840.52	1848.38	3777800000	1848.38
2014-01-14	1821.36	1839.26	1821.36	1838.88	3353270000	1838.88
2014-01-13	1841.26	1843.45	1815.52	1819.20	3591350000	1819.20
2014-01-10	1840.06	1843.15	1832.43	1842.37	3335710000	1842.37
2014-01-09	1839.00	1843.23	1830.38	1838.13	3581150000	1838.13
2014-01-08	1837.90	1840.02	1831.40	1837.49	3652140000	1837.49
2014-01-07	1828.71	1840.10	1828.71	1837.88	3511750000	1837.88
2014-01-06	1832.31	1837.16	1823.73	1826.77	3294850000	1826.77
2014-01-03	1833.21	1838.24	1829.13	1831.37	2774270000	1831.37
2014-01-02	1845.86	1845.86	1827.74	1831.98	3080600000	1831.98
2013-12-31	1842.61	1849.44	1842.41	1848.36	2312840000	1848.36
2013-12-30	1841.47	1842.47	1838.77	1841.07	0	1841.07
2013-12-27	1842.97	1844.89	1839.81	1841.40	2052920000	1841.40
2013-12-26	1834.96	1842.84	1834.96	1842.02	1982270000	1842.02
2013-12-24	1828.02	1833.32	1828.02	1833.32	1307630000	1833.32
2013-12-23	1822.92	1829.75	1822.92	1827.99	2851540000	1827.99
2013-12-20	1810.39	1823.75	1810.25	1818.32	5097700000	1818.32
2013-12-19	1809.00	1810.88	1801.35	1809.60	3497210000	1809.60
2013-12-18	1781.46	1811.08	1767.99	1810.65	4327770000	1810.65
2013-12-17	1786.47	1786.77	1777.05	1781.00	3270030000	1781.00
2013-12-16	1777.48	1792.22	1777.48	1786.54	3209890000	1786.54
2013-12-13	1777.98	1780.92	1772.45	1775.32	3061070000	1775.32
2013-12-12	1781.71	1782.99	1772.28	1775.50	3306640000	1775.50
2013-12-11	1802.76	1802.97	1780.09	1782.22	3472240000	1782.22
2013-12-10	1807.60	1808.52	1801.75	1802.62	3117150000	1802.62
2013-12-09	1806.21	1811.52	1806.21	1808.37	3129500000	1808.37

```
2013-12-06  1788.36  1806.04  1788.36  1805.09  3150030000    1805.09
                ...      ...      ...      ...         ...        ...

[16146 rows x 6 columns]
```

In [98]:
```
oopen =df[['Open']]
oopen
```

Out [98]:
```
               Open
Date
2014-03-05  1874.05
2014-03-04  1849.23
2014-03-03  1857.68
2014-02-28  1855.12
2014-02-27  1844.90
2014-02-26  1845.79
2014-02-25  1847.66
2014-02-24  1836.78
2014-02-21  1841.07
2014-02-20  1829.24
2014-02-19  1838.90
2014-02-18  1839.03
2014-02-14  1828.46
2014-02-13  1814.82
2014-02-12  1820.12
2014-02-11  1800.45
2014-02-10  1796.20
2014-02-07  1776.01
2014-02-06  1752.99
2014-02-05  1753.38
2014-02-04  1743.82
2014-02-03  1782.68
2014-01-31  1790.88
2014-01-30  1777.17
2014-01-29  1790.15
2014-01-28  1783.00
2014-01-27  1791.03
2014-01-24  1826.96
2014-01-23  1842.29
2014-01-22  1844.71
2014-01-21  1841.05
2014-01-17  1844.23
2014-01-16  1847.99
2014-01-15  1840.52
2014-01-14  1821.36
2014-01-13  1841.26
2014-01-10  1840.06
2014-01-09  1839.00
2014-01-08  1837.90
2014-01-07  1828.71
2014-01-06  1832.31
2014-01-03  1833.21
2014-01-02  1845.86
```

```
2013-12-31  1842.61
2013-12-30  1841.47
2013-12-27  1842.97
2013-12-26  1834.96
2013-12-24  1828.02
2013-12-23  1822.92
2013-12-20  1810.39
2013-12-19  1809.00
2013-12-18  1781.46
2013-12-17  1786.47
2013-12-16  1777.48
2013-12-13  1777.98
2013-12-12  1781.71
2013-12-11  1802.76
2013-12-10  1807.60
2013-12-09  1806.21
2013-12-06  1788.36
                ...

[16146 rows x 1 columns]
```

In [99]:

```
from statsmodels import tsa
p=1
d=1
q=1
model1=tsa.arima_model.ARIMA(oopen, [p, d, q],freq='M').fit() #creates model
```

In [100]:

```
model1.summary()
```

Out [100]:

```
<class 'statsmodels.iolib.summary.Summary'>
"""
                              ARIMA Model Results
======================================================================
========
Dep. Variable:                  D.Open   No. Observations:
16145
Model:                  ARIMA(1, 1, 1)   Log Likelihood
-55293.106
Method:                        css-mle   S.D. of innovations
7.432
Date:                 Wed, 05 Mar 2014   AIC
110594.211
Time:                         22:13:04   BIC
110624.969
Sample:                     03-04-2014   HQIC
110604.379
                          - 01-03-1950
======================================================================
==========
                   coef    std err          z      P>|z|      [95.0%
Conf. Int.]
----------------------------------------------------------------------
----------
```

```
const             -0.1148      0.050     -2.304      0.021       -0.212
-0.017
ar.L1.D.Open       0.6960      0.050     13.892      0.000        0.598
0.794
ma.L1.D.Open      -0.7410      0.047    -15.847      0.000       -0.833
-0.649
                                    Roots
=====================================================================
=======
                 Real          Imaginary           Modulus
Frequency
---------------------------------------------------------------------
-------
AR.1            1.4368           +0.0000j            1.4368
0.0000
MA.1            1.3496           +0.0000j            1.3496
0.0000
---------------------------------------------------------------------
-------
"""
```