

In [1]:

```
# TODO:Read data  
import pandas as pd  
TWITTER = pd.read_csv('/Users/khaladdin/Desktop/Twitter Project/TWITTER.csv')
```

In [2]:

```
# TODO:Create separate datasets for each stock  
ADI = TWITTER.loc[TWITTER['#RIC'] == 'ADI.OQ']  
FID = TWITTER.loc[TWITTER['#RIC'] == 'FIS.N']  
FIS = TWITTER.loc[TWITTER['#RIC'] == 'FISV.OQ']  
GPN = TWITTER.loc[TWITTER['#RIC'] == 'GPN.N']  
JUN = TWITTER.loc[TWITTER['#RIC'] == 'JNPR.N']
```

In [3]:

```
# TODO>Delete Trade data  
ADIA = ADI[ADI.Type != 'Trade']  
FIDA = FID[FID.Type != 'Trade']  
FISA = FIS[FIS.Type != 'Trade']  
GPNA = GPN[GPN.Type != 'Trade']  
JUNA = JUN[JUN.Type != 'Trade']
```

In [4]:

```
# TODO:Change the format  
ADIA['Date'] = pd.to_datetime(ADIA['Date-Time'])  
FIDA['Date'] = pd.to_datetime(FIDA['Date-Time'])  
FISA['Date'] = pd.to_datetime(FISA['Date-Time'])  
GPNA['Date'] = pd.to_datetime(GPNA['Date-Time'])  
JUNA['Date'] = pd.to_datetime(JUNA['Date-Time'])
```

```

/anaconda3/lib/python3.6/site-packages/ipykernel_launcher.py:1: Se
ttingWithCopyWarning:
A value is trying to be set on a copy of a slice from a DataFrame.
Try using .loc[row_indexer,col_indexer] = value instead

See the caveats in the documentation: http://pandas.pydata.org/pan
das-docs/stable/indexing.html#indexing-view-versus-copy
    """Entry point for launching an IPython kernel.
/anaconda3/lib/python3.6/site-packages/ipykernel_launcher.py:2: Se
ttingWithCopyWarning:
A value is trying to be set on a copy of a slice from a DataFrame.
Try using .loc[row_indexer,col_indexer] = value instead

See the caveats in the documentation: http://pandas.pydata.org/pan
das-docs/stable/indexing.html#indexing-view-versus-copy

/anaconda3/lib/python3.6/site-packages/ipykernel_launcher.py:3: Se
ttingWithCopyWarning:
A value is trying to be set on a copy of a slice from a DataFrame.
Try using .loc[row_indexer,col_indexer] = value instead

See the caveats in the documentation: http://pandas.pydata.org/pan
das-docs/stable/indexing.html#indexing-view-versus-copy
    This is separate from the ipykernel package so we can avoid doin
g imports until
/anaconda3/lib/python3.6/site-packages/ipykernel_launcher.py:4: Se
ttingWithCopyWarning:
A value is trying to be set on a copy of a slice from a DataFrame.
Try using .loc[row_indexer,col_indexer] = value instead

See the caveats in the documentation: http://pandas.pydata.org/pan
das-docs/stable/indexing.html#indexing-view-versus-copy
    after removing the cwd from sys.path.
/anaconda3/lib/python3.6/site-packages/ipykernel_launcher.py:5: Se
ttingWithCopyWarning:
A value is trying to be set on a copy of a slice from a DataFrame.
Try using .loc[row_indexer,col_indexer] = value instead

See the caveats in the documentation: http://pandas.pydata.org/pan
das-docs/stable/indexing.html#indexing-view-versus-copy
    """

```

In [5]:

```

# TODO:Change the format
ADIA['Date1'] = ADIA['Date'].dt.date
FIDA['Date1'] = FIDA['Date'].dt.date
FISA['Date1'] = FISA['Date'].dt.date
GPNA['Date1'] = GPNA['Date'].dt.date
JUNA['Date1'] = JUNA['Date'].dt.date

```

```

/anaconda3/lib/python3.6/site-packages/ipykernel_launcher.py:1: Se
ttingWithCopyWarning:
A value is trying to be set on a copy of a slice from a DataFrame.
Try using .loc[row_indexer,col_indexer] = value instead

See the caveats in the documentation: http://pandas.pydata.org/pan
das-docs/stable/indexing.html#indexing-view-versus-copy
    """Entry point for launching an IPython kernel.
/anaconda3/lib/python3.6/site-packages/ipykernel_launcher.py:2: Se
ttingWithCopyWarning:
A value is trying to be set on a copy of a slice from a DataFrame.
Try using .loc[row_indexer,col_indexer] = value instead

See the caveats in the documentation: http://pandas.pydata.org/pan
das-docs/stable/indexing.html#indexing-view-versus-copy

/anaconda3/lib/python3.6/site-packages/ipykernel_launcher.py:3: Se
ttingWithCopyWarning:
A value is trying to be set on a copy of a slice from a DataFrame.
Try using .loc[row_indexer,col_indexer] = value instead

See the caveats in the documentation: http://pandas.pydata.org/pan
das-docs/stable/indexing.html#indexing-view-versus-copy
    This is separate from the ipykernel package so we can avoid doin
g imports until
/anaconda3/lib/python3.6/site-packages/ipykernel_launcher.py:4: Se
ttingWithCopyWarning:
A value is trying to be set on a copy of a slice from a DataFrame.
Try using .loc[row_indexer,col_indexer] = value instead

See the caveats in the documentation: http://pandas.pydata.org/pan
das-docs/stable/indexing.html#indexing-view-versus-copy
    after removing the cwd from sys.path.
/anaconda3/lib/python3.6/site-packages/ipykernel_launcher.py:5: Se
ttingWithCopyWarning:
A value is trying to be set on a copy of a slice from a DataFrame.
Try using .loc[row_indexer,col_indexer] = value instead

See the caveats in the documentation: http://pandas.pydata.org/pan
das-docs/stable/indexing.html#indexing-view-versus-copy
    """

```

In [6]:

```
# TODO:Change the format
ADIA['Time1'] = ADIA['Date'].dt.time
FIDa['Time1'] = FIDa['Date'].dt.time
FISa['Time1'] = FISa['Date'].dt.time
GPNa['Time1'] = GPNa['Date'].dt.time
JUNa['Time1'] = JUNa['Date'].dt.time
```

/anaconda3/lib/python3.6/site-packages/ipykernel\_launcher.py:1: SettingWithCopyWarning:

A value is trying to be set on a copy of a slice from a DataFrame.  
Try using .loc[row\_indexer,col\_indexer] = value instead

See the caveats in the documentation: <http://pandas.pydata.org/pandas-docs/stable/indexing.html#indexing-view-versus-copy>

"""Entry point for launching an IPython kernel.

/anaconda3/lib/python3.6/site-packages/ipykernel\_launcher.py:2: SettingWithCopyWarning:

A value is trying to be set on a copy of a slice from a DataFrame.  
Try using .loc[row\_indexer,col\_indexer] = value instead

See the caveats in the documentation: <http://pandas.pydata.org/pandas-docs/stable/indexing.html#indexing-view-versus-copy>

/anaconda3/lib/python3.6/site-packages/ipykernel\_launcher.py:3: SettingWithCopyWarning:

A value is trying to be set on a copy of a slice from a DataFrame.  
Try using .loc[row\_indexer,col\_indexer] = value instead

See the caveats in the documentation: <http://pandas.pydata.org/pandas-docs/stable/indexing.html#indexing-view-versus-copy>

This is separate from the ipykernel package so we can avoid doing imports until

/anaconda3/lib/python3.6/site-packages/ipykernel\_launcher.py:4: SettingWithCopyWarning:

A value is trying to be set on a copy of a slice from a DataFrame.  
Try using .loc[row\_indexer,col\_indexer] = value instead

See the caveats in the documentation: <http://pandas.pydata.org/pandas-docs/stable/indexing.html#indexing-view-versus-copy>

after removing the cwd from sys.path.

/anaconda3/lib/python3.6/site-packages/ipykernel\_launcher.py:5: SettingWithCopyWarning:

A value is trying to be set on a copy of a slice from a DataFrame.  
Try using .loc[row\_indexer,col\_indexer] = value instead

See the caveats in the documentation: <http://pandas.pydata.org/pandas-docs/stable/indexing.html#indexing-view-versus-copy>

"""

In [7]:

```
# TODO:Delete milliseconds
ADIA['Time1'] = ADIA['Time1'].apply(lambda x: x.replace(microsecond=0))
FIDa['Time1'] = FIDa['Time1'].apply(lambda x: x.replace(microsecond=0))
FISa['Time1'] = FISa['Time1'].apply(lambda x: x.replace(microsecond=0))
GPNa['Time1'] = GPNa['Time1'].apply(lambda x: x.replace(microsecond=0))
JUNa['Time1'] = JUNa['Time1'].apply(lambda x: x.replace(microsecond=0))
```

/anaconda3/lib/python3.6/site-packages/ipykernel\_launcher.py:1: SettingWithCopyWarning:  
A value is trying to be set on a copy of a slice from a DataFrame.  
Try using .loc[row\_indexer,col\_indexer] = value instead

See the caveats in the documentation: <http://pandas.pydata.org/pandas-docs/stable/indexing.html#indexing-view-versus-copy>

```
"""Entry point for launching an IPython kernel.
/anaconda3/lib/python3.6/site-packages/ipykernel_launcher.py:2: SettingWithCopyWarning:
A value is trying to be set on a copy of a slice from a DataFrame.
Try using .loc[row_indexer,col_indexer] = value instead
```

See the caveats in the documentation: <http://pandas.pydata.org/pandas-docs/stable/indexing.html#indexing-view-versus-copy>

```
/anaconda3/lib/python3.6/site-packages/ipykernel_launcher.py:3: SettingWithCopyWarning:
A value is trying to be set on a copy of a slice from a DataFrame.
Try using .loc[row_indexer,col_indexer] = value instead
```

See the caveats in the documentation: <http://pandas.pydata.org/pandas-docs/stable/indexing.html#indexing-view-versus-copy>

This is separate from the ipykernel package so we can avoid doing imports until

```
/anaconda3/lib/python3.6/site-packages/ipykernel_launcher.py:4: SettingWithCopyWarning:
A value is trying to be set on a copy of a slice from a DataFrame.
Try using .loc[row_indexer,col_indexer] = value instead
```

See the caveats in the documentation: <http://pandas.pydata.org/pandas-docs/stable/indexing.html#indexing-view-versus-copy>

```
after removing the cwd from sys.path.
/anaconda3/lib/python3.6/site-packages/ipykernel_launcher.py:5: SettingWithCopyWarning:
A value is trying to be set on a copy of a slice from a DataFrame.
Try using .loc[row_indexer,col_indexer] = value instead
```

See the caveats in the documentation: <http://pandas.pydata.org/pandas-docs/stable/indexing.html#indexing-view-versus-copy>

```
"""
```

In [8]:

```
# TODO:Delete seconds
ADIA['Time1'] = ADIA['Time1'].apply(lambda x: x.replace(second=0))
FIDa['Time1'] = FIDa['Time1'].apply(lambda x: x.replace(second=0))
FISa['Time1'] = FISa['Time1'].apply(lambda x: x.replace(second=0))
GPNa['Time1'] = GPNa['Time1'].apply(lambda x: x.replace(second=0))
JUNa['Time1'] = JUNa['Time1'].apply(lambda x: x.replace(second=0))
```

/anaconda3/lib/python3.6/site-packages/ipykernel\_launcher.py:1: SettingWithCopyWarning:

A value is trying to be set on a copy of a slice from a DataFrame.  
Try using .loc[row\_indexer,col\_indexer] = value instead

See the caveats in the documentation: <http://pandas.pydata.org/pandas-docs/stable/indexing.html#indexing-view-versus-copy>

"""Entry point for launching an IPython kernel.

/anaconda3/lib/python3.6/site-packages/ipykernel\_launcher.py:2: SettingWithCopyWarning:

A value is trying to be set on a copy of a slice from a DataFrame.  
Try using .loc[row\_indexer,col\_indexer] = value instead

See the caveats in the documentation: <http://pandas.pydata.org/pandas-docs/stable/indexing.html#indexing-view-versus-copy>

/anaconda3/lib/python3.6/site-packages/ipykernel\_launcher.py:3: SettingWithCopyWarning:

A value is trying to be set on a copy of a slice from a DataFrame.  
Try using .loc[row\_indexer,col\_indexer] = value instead

See the caveats in the documentation: <http://pandas.pydata.org/pandas-docs/stable/indexing.html#indexing-view-versus-copy>

This is separate from the ipykernel package so we can avoid doing imports until

/anaconda3/lib/python3.6/site-packages/ipykernel\_launcher.py:4: SettingWithCopyWarning:

A value is trying to be set on a copy of a slice from a DataFrame.  
Try using .loc[row\_indexer,col\_indexer] = value instead

See the caveats in the documentation: <http://pandas.pydata.org/pandas-docs/stable/indexing.html#indexing-view-versus-copy>

after removing the cwd from sys.path.

/anaconda3/lib/python3.6/site-packages/ipykernel\_launcher.py:5: SettingWithCopyWarning:

A value is trying to be set on a copy of a slice from a DataFrame.  
Try using .loc[row\_indexer,col\_indexer] = value instead

See the caveats in the documentation: <http://pandas.pydata.org/pandas-docs/stable/indexing.html#indexing-view-versus-copy>

"""

In [9]:

```
# TODO: Datetime group
ADIA.loc[:, 'Datetime'] = pd.to_datetime(ADIA.Date1.astype(str)+' '+ADIA.Time1.
astype(str))
FIDa.loc[:, 'Datetime'] = pd.to_datetime(FIDa.Date1.astype(str)+' '+FIDa.Time1.
astype(str))
FISa.loc[:, 'Datetime'] = pd.to_datetime(FISa.Date1.astype(str)+' '+FISa.Time1.
astype(str))
GPNa.loc[:, 'Datetime'] = pd.to_datetime(GPNa.Date1.astype(str)+' '+GPNa.Time1.
astype(str))
JUNa.loc[:, 'Datetime'] = pd.to_datetime(JUNa.Date1.astype(str)+' '+JUNa.Time1.
astype(str))
```

/anaconda3/lib/python3.6/site-packages/pandas/core/indexing.py:357

: SettingWithCopyWarning:

A value is trying to be set on a copy of a slice from a DataFrame.

Try using .loc[row\_indexer,col\_indexer] = value instead

See the caveats in the documentation: <http://pandas.pydata.org/pandas-docs/stable/indexing.html#indexing-view-versus-copy>

```
self.obj[key] = _infer_fill_value(value)
```

/anaconda3/lib/python3.6/site-packages/pandas/core/indexing.py:537

: SettingWithCopyWarning:

A value is trying to be set on a copy of a slice from a DataFrame.

Try using .loc[row\_indexer,col\_indexer] = value instead

See the caveats in the documentation: <http://pandas.pydata.org/pandas-docs/stable/indexing.html#indexing-view-versus-copy>

```
self.obj[item] = s
```

In [10]:

```
# TODO: Compute midprice and spread
ADIA['midprice'] = (ADIA['Bid Price']+ADIA['Ask Price'])/2
ADIA['spread'] = ADIA['Ask Price']-ADIA['Bid Price']
FIDa['midprice'] = (FIDa['Bid Price']+FIDa['Ask Price'])/2
FIDa['spread'] = FIDa['Ask Price']-FIDa['Bid Price']
FISa['midprice'] = (FISa['Bid Price']+FISa['Ask Price'])/2
FISa['spread'] = FISa['Ask Price']-FISa['Bid Price']
GPNa['midprice'] = (GPNa['Bid Price']+GPNa['Ask Price'])/2
GPNa['spread'] = GPNa['Ask Price']-GPNa['Bid Price']
JUNa['midprice'] = (JUNa['Bid Price']+JUNa['Ask Price'])/2
JUNa['spread'] = JUNa['Ask Price']-JUNa['Bid Price']
```

/anaconda3/lib/python3.6/site-packages/ipykernel\_launcher.py:1: Se

ttingWithCopyWarning:

A value is trying to be set on a copy of a slice from a DataFrame.

Try using .loc[row\_indexer,col\_indexer] = value instead

See the caveats in the documentation: <http://pandas.pydata.org/pandas-docs/stable/indexing.html#indexing-view-versus-copy>

```
"""Entry point for launching an IPython kernel.
```

/anaconda3/lib/python3.6/site-packages/ipykernel\_launcher.py:2: Se

ttingWithCopyWarning:

A value is trying to be set on a copy of a slice from a DataFrame.

Try using `.loc[row_indexer,col_indexer] = value` instead

See the caveats in the documentation: <http://pandas.pydata.org/pandas-docs/stable/indexing.html#indexing-view-versus-copy>

/anaconda3/lib/python3.6/site-packages/ipykernel\_launcher.py:3: SettingWithCopyWarning:  
A value is trying to be set on a copy of a slice from a DataFrame.  
Try using `.loc[row_indexer,col_indexer] = value` instead

See the caveats in the documentation: <http://pandas.pydata.org/pandas-docs/stable/indexing.html#indexing-view-versus-copy>

This is separate from the ipykernel package so we can avoid doing imports until

/anaconda3/lib/python3.6/site-packages/ipykernel\_launcher.py:4: SettingWithCopyWarning:  
A value is trying to be set on a copy of a slice from a DataFrame.  
Try using `.loc[row_indexer,col_indexer] = value` instead

See the caveats in the documentation: <http://pandas.pydata.org/pandas-docs/stable/indexing.html#indexing-view-versus-copy>

after removing the cwd from sys.path.  
/anaconda3/lib/python3.6/site-packages/ipykernel\_launcher.py:5: SettingWithCopyWarning:  
A value is trying to be set on a copy of a slice from a DataFrame.  
Try using `.loc[row_indexer,col_indexer] = value` instead

See the caveats in the documentation: <http://pandas.pydata.org/pandas-docs/stable/indexing.html#indexing-view-versus-copy>  
"""

/anaconda3/lib/python3.6/site-packages/ipykernel\_launcher.py:6: SettingWithCopyWarning:  
A value is trying to be set on a copy of a slice from a DataFrame.  
Try using `.loc[row_indexer,col_indexer] = value` instead

See the caveats in the documentation: <http://pandas.pydata.org/pandas-docs/stable/indexing.html#indexing-view-versus-copy>

/anaconda3/lib/python3.6/site-packages/ipykernel\_launcher.py:7: SettingWithCopyWarning:  
A value is trying to be set on a copy of a slice from a DataFrame.  
Try using `.loc[row_indexer,col_indexer] = value` instead

See the caveats in the documentation: <http://pandas.pydata.org/pandas-docs/stable/indexing.html#indexing-view-versus-copy>

import sys  
/anaconda3/lib/python3.6/site-packages/ipykernel\_launcher.py:8: SettingWithCopyWarning:  
A value is trying to be set on a copy of a slice from a DataFrame.  
Try using `.loc[row_indexer,col_indexer] = value` instead

See the caveats in the documentation: <http://pandas.pydata.org/pandas-docs/stable/indexing.html#indexing-view-versus-copy>

/anaconda3/lib/python3.6/site-packages/ipykernel\_launcher.py:9: SettingWithCopyWarning:  
A value is trying to be set on a copy of a slice from a DataFrame.



Try using `.loc[row_indexer,col_indexer] = value` instead

See the caveats in the documentation: <http://pandas.pydata.org/pandas-docs/stable/indexing.html#indexing-view-versus-copy>

```
if __name__ == '__main__':  
/anaconda3/lib/python3.6/site-packages/ipykernel_launcher.py:10: SettingWithCopyWarning:  
A value is trying to be set on a copy of a slice from a DataFrame.  
Try using .loc[row_indexer,col_indexer] = value instead
```

See the caveats in the documentation: <http://pandas.pydata.org/pandas-docs/stable/indexing.html#indexing-view-versus-copy>

```
# Remove the CWD from sys.path while we load stuff.
```

In [11]:

```
# TODO:Logmidprice  
import numpy as np  
ADIA['logmid'] = np.log(ADIA['midprice'])  
FIDA['logmid'] = np.log(FIDA['midprice'])  
FISA['logmid'] = np.log(FISA['midprice'])  
GPNa['logmid'] = np.log(GPNa['midprice'])  
JUNa['logmid'] = np.log(JUNa['midprice'])
```

/anaconda3/lib/python3.6/site-packages/ipykernel\_launcher.py:2: RuntimeWarning: divide by zero encountered in log

/anaconda3/lib/python3.6/site-packages/ipykernel\_launcher.py:2: SettingWithCopyWarning:  
A value is trying to be set on a copy of a slice from a DataFrame.  
Try using `.loc[row_indexer,col_indexer] = value` instead

See the caveats in the documentation: <http://pandas.pydata.org/pandas-docs/stable/indexing.html#indexing-view-versus-copy>

/anaconda3/lib/python3.6/site-packages/ipykernel\_launcher.py:3: SettingWithCopyWarning:  
A value is trying to be set on a copy of a slice from a DataFrame.  
Try using `.loc[row_indexer,col_indexer] = value` instead

See the caveats in the documentation: <http://pandas.pydata.org/pandas-docs/stable/indexing.html#indexing-view-versus-copy>

This is separate from the ipykernel package so we can avoid doing imports until

/anaconda3/lib/python3.6/site-packages/ipykernel\_launcher.py:4: RuntimeWarning: divide by zero encountered in log

after removing the cwd from sys.path.

/anaconda3/lib/python3.6/site-packages/ipykernel\_launcher.py:4: SettingWithCopyWarning:  
A value is trying to be set on a copy of a slice from a DataFrame.  
Try using `.loc[row_indexer,col_indexer] = value` instead

See the caveats in the documentation: <http://pandas.pydata.org/pandas-docs/stable/indexing.html#indexing-view-versus-copy>

after removing the cwd from sys.path.

/anaconda3/lib/python3.6/site-packages/ipykernel\_launcher.py:5: SettingWithCopyWarning:  
A value is trying to be set on a copy of a slice from a DataFrame.  
Try using `.loc[row_indexer,col_indexer] = value` instead

See the caveats in the documentation: <http://pandas.pydata.org/pandas-docs/stable/indexing.html#indexing-view-versus-copy>

"""

/anaconda3/lib/python3.6/site-packages/ipykernel\_launcher.py:6: SettingWithCopyWarning:  
A value is trying to be set on a copy of a slice from a DataFrame.  
Try using `.loc[row_indexer,col_indexer] = value` instead

See the caveats in the documentation: <http://pandas.pydata.org/pandas-docs/stable/indexing.html#indexing-view-versus-copy>

In [12]:

```
# TODO: Cleaning
ADIA = ADIA[~(ADIA['spread'] > 400)]
FIDA = FIDA[~(FIDA['spread'] > 400)]
FISA = FISA[~(FISA['spread'] > 400)]
GPNA = GPNA[~(GPNA['spread'] > 400)]
JUNA = JUNA[~(JUNA['spread'] > 400)]
```

In [13]:

```
# TODO: CleaningADIA = ADIA[~(ADIA['spread'] < 0)]
FIDA = FIDA[~(FIDA['spread'] < 0)]
FISA = FISA[~(FISA['spread'] < 0)]
GPNA = GPNA[~(GPNA['spread'] < 0)]
JUNA = JUNA[~(JUNA['spread'] < 0)]
```

In [14]:

```
# TODO: Save data for state space modelling. I am using SAS for state space es
timiation, since Python is not able to provide what we want.
# TODO: I am working on python state space
ADIA.to_csv('/Users/khaladdin/SASUniversityEdition/myfolders/ADIA.csv')
FIDA.to_csv('/Users/khaladdin/SASUniversityEdition/myfolders/FIDA.csv')
FISA.to_csv('/Users/khaladdin/SASUniversityEdition/myfolders/FISA.csv')
GPNA.to_csv('/Users/khaladdin/SASUniversityEdition/myfolders/GPNA.csv')
JUNA.to_csv('/Users/khaladdin/SASUniversityEdition/myfolders/JUNA.csv')
```

In [15]:

```
# TODO: Read state space estimation from SAS folder
ADIdstate = pd.read_csv('/Users/khaladdin/SASUniversityEdition/myfolders/ADId.c
sv')
FIDstate = pd.read_csv('/Users/khaladdin/SASUniversityEdition/myfolders/FIDd.c
sv')
FISstate = pd.read_csv('/Users/khaladdin/SASUniversityEdition/myfolders/FISd.c
sv')
GPNstate = pd.read_csv('/Users/khaladdin/SASUniversityEdition/myfolders/GPNd.c
sv')
JUNstate = pd.read_csv('/Users/khaladdin/SASUniversityEdition/myfolders/JUNd.c
sv')
```

In [16]:

```
# TODO: Formatting
ADIdstate['Date'] = ADIdstate['Datetime'].astype(str).str[:7]
ADIdstate['Time'] = ADIdstate['Datetime'].astype(str).str[-8:]
FIDstate['Date'] = FIDstate['Datetime'].astype(str).str[:7]
FIDstate['Time'] = FIDstate['Datetime'].astype(str).str[-8:]
FISstate['Date'] = FISstate['Datetime'].astype(str).str[:7]
FISstate['Time'] = FISstate['Datetime'].astype(str).str[-8:]
GPNstate['Date'] = GPNstate['Datetime'].astype(str).str[:7]
GPNstate['Time'] = GPNstate['Datetime'].astype(str).str[-8:]
JUNstate['Date'] = JUNstate['Datetime'].astype(str).str[:7]
JUNstate['Time'] = JUNstate['Datetime'].astype(str).str[-8:]
```

In [17]:

```
# TODO: Formatting
ADIstate['Date'] = ADIstate['Date'].str.replace('JUL', '-JUL-')
FIDstate['Date'] = FIDstate['Date'].str.replace('JUL', '-JUL-')
FISstate['Date'] = FISstate['Date'].str.replace('JUL', '-JUL-')
GPNstate['Date'] = GPNstate['Date'].str.replace('JUL', '-JUL-')
JUNstate['Date'] = JUNstate['Date'].str.replace('JUL', '-JUL-')
```

In [19]:

```
# TODO: Formatting
ADIstate['Date'] = pd.to_datetime(ADIstate['Date'])
FIDstate['Date'] = pd.to_datetime(FIDstate['Date'])
FISstate['Date'] = pd.to_datetime(FISstate['Date'])
GPNstate['Date'] = pd.to_datetime(GPNstate['Date'])
JUNstate['Date'] = pd.to_datetime(JUNstate['Date'])
```

In [20]:

```
# TODO: Keep some columns.
ADIstatea = ADIstate[['Date', 'Time', 'Irregular', 'Level']]
FIDstatea = FIDstate[['Date', 'Time', 'Irregular', 'Level']]
FISstatea = FISstate[['Date', 'Time', 'Irregular', 'Level']]
GPNstatea = GPNstate[['Date', 'Time', 'Irregular', 'Level']]
JUNstatea = JUNstate[['Date', 'Time', 'Irregular', 'Level']]
```

In [21]:

```
# TODO: Change the format of date
ADIstatea['Date1'] = ADIstatea['Date'].dt.date
FIDstatea['Date1'] = FIDstatea['Date'].dt.date
FISstatea['Date1'] = FISstatea['Date'].dt.date
GPNstatea['Date1'] = GPNstatea['Date'].dt.date
JUNstatea['Date1'] = JUNstatea['Date'].dt.date
```

```
/anaconda3/lib/python3.6/site-packages/ipykernel_launcher.py:2: SettingWithCopyWarning:
A value is trying to be set on a copy of a slice from a DataFrame.
Try using .loc[row_indexer,col_indexer] = value instead
```

See the caveats in the documentation: <http://pandas.pydata.org/pandas-docs/stable/indexing.html#indexing-view-versus-copy>

```
/anaconda3/lib/python3.6/site-packages/ipykernel_launcher.py:3: SettingWithCopyWarning:
A value is trying to be set on a copy of a slice from a DataFrame.
Try using .loc[row_indexer,col_indexer] = value instead
```

See the caveats in the documentation: <http://pandas.pydata.org/pandas-docs/stable/indexing.html#indexing-view-versus-copy>

This is separate from the ipykernel package so we can avoid doing imports until

```
/anaconda3/lib/python3.6/site-packages/ipykernel_launcher.py:4: SettingWithCopyWarning:
A value is trying to be set on a copy of a slice from a DataFrame.
Try using .loc[row_indexer,col_indexer] = value instead
```

See the caveats in the documentation: <http://pandas.pydata.org/pandas-docs/stable/indexing.html#indexing-view-versus-copy>

after removing the cwd from sys.path.

```
/anaconda3/lib/python3.6/site-packages/ipykernel_launcher.py:5: SettingWithCopyWarning:
A value is trying to be set on a copy of a slice from a DataFrame.
Try using .loc[row_indexer,col_indexer] = value instead
```

See the caveats in the documentation: <http://pandas.pydata.org/pandas-docs/stable/indexing.html#indexing-view-versus-copy>

"""

```
/anaconda3/lib/python3.6/site-packages/ipykernel_launcher.py:6: SettingWithCopyWarning:
A value is trying to be set on a copy of a slice from a DataFrame.
Try using .loc[row_indexer,col_indexer] = value instead
```

See the caveats in the documentation: <http://pandas.pydata.org/pandas-docs/stable/indexing.html#indexing-view-versus-copy>

In [22]:

```
# TODO:Change format of columns
ADIstatea['Time'] = pd.to_datetime(ADIstatea['Time'])
ADIstatea['Time1'] = ADIstatea['Time'].dt.time
FIDstatea['Time'] = pd.to_datetime(FIDstatea['Time'])
FIDstatea['Time1'] = FIDstatea['Time'].dt.time
FISstatea['Time'] = pd.to_datetime(FISstatea['Time'])
FISstatea['Time1'] = FISstatea['Time'].dt.time
GPNstatea['Time'] = pd.to_datetime(GPNstatea['Time'])
GPNstatea['Time1'] = GPNstatea['Time'].dt.time
JUNstatea['Time'] = pd.to_datetime(JUNstatea['Time'])
JUNstatea['Time1'] = JUNstatea['Time'].dt.time
```

```
/anaconda3/lib/python3.6/site-packages/ipykernel_launcher.py:2: Se
ttingWithCopyWarning:
A value is trying to be set on a copy of a slice from a DataFrame.
Try using .loc[row_indexer,col_indexer] = value instead

See the caveats in the documentation: http://pandas.pydata.org/pan
das-docs/stable/indexing.html#indexing-view-versus-copy

/anaconda3/lib/python3.6/site-packages/ipykernel_launcher.py:3: Se
ttingWithCopyWarning:
A value is trying to be set on a copy of a slice from a DataFrame.
Try using .loc[row_indexer,col_indexer] = value instead

See the caveats in the documentation: http://pandas.pydata.org/pan
das-docs/stable/indexing.html#indexing-view-versus-copy
    This is separate from the ipykernel package so we can avoid doin
g imports until
/anaconda3/lib/python3.6/site-packages/ipykernel_launcher.py:4: Se
ttingWithCopyWarning:
A value is trying to be set on a copy of a slice from a DataFrame.
Try using .loc[row_indexer,col_indexer] = value instead

See the caveats in the documentation: http://pandas.pydata.org/pan
das-docs/stable/indexing.html#indexing-view-versus-copy
    after removing the cwd from sys.path.
/anaconda3/lib/python3.6/site-packages/ipykernel_launcher.py:5: Se
ttingWithCopyWarning:
A value is trying to be set on a copy of a slice from a DataFrame.
Try using .loc[row_indexer,col_indexer] = value instead

See the caveats in the documentation: http://pandas.pydata.org/pan
das-docs/stable/indexing.html#indexing-view-versus-copy
    """
/anaconda3/lib/python3.6/site-packages/ipykernel_launcher.py:6: Se
ttingWithCopyWarning:
A value is trying to be set on a copy of a slice from a DataFrame.
Try using .loc[row_indexer,col_indexer] = value instead

See the caveats in the documentation: http://pandas.pydata.org/pan
das-docs/stable/indexing.html#indexing-view-versus-copy

/anaconda3/lib/python3.6/site-packages/ipykernel_launcher.py:7: Se
ttingWithCopyWarning:
A value is trying to be set on a copy of a slice from a DataFrame.
Try using .loc[row_indexer,col_indexer] = value instead

See the caveats in the documentation: http://pandas.pydata.org/pan
das-docs/stable/indexing.html#indexing-view-versus-copy
    import sys
/anaconda3/lib/python3.6/site-packages/ipykernel_launcher.py:8: Se
ttingWithCopyWarning:
A value is trying to be set on a copy of a slice from a DataFrame.
Try using .loc[row_indexer,col_indexer] = value instead

See the caveats in the documentation: http://pandas.pydata.org/pan
das-docs/stable/indexing.html#indexing-view-versus-copy
```

```
/anaconda3/lib/python3.6/site-packages/ipykernel_launcher.py:9: SettingWithCopyWarning:
```

```
A value is trying to be set on a copy of a slice from a DataFrame.  
Try using .loc[row_indexer,col_indexer] = value instead
```

See the caveats in the documentation: <http://pandas.pydata.org/pandas-docs/stable/indexing.html#indexing-view-versus-copy>

```
if __name__ == '__main__':  
/anaconda3/lib/python3.6/site-packages/ipykernel_launcher.py:10: SettingWithCopyWarning:
```

```
A value is trying to be set on a copy of a slice from a DataFrame.  
Try using .loc[row_indexer,col_indexer] = value instead
```

See the caveats in the documentation: <http://pandas.pydata.org/pandas-docs/stable/indexing.html#indexing-view-versus-copy>

```
# Remove the CWD from sys.path while we load stuff.  
/anaconda3/lib/python3.6/site-packages/ipykernel_launcher.py:11: SettingWithCopyWarning:
```

```
A value is trying to be set on a copy of a slice from a DataFrame.  
Try using .loc[row_indexer,col_indexer] = value instead
```

See the caveats in the documentation: <http://pandas.pydata.org/pandas-docs/stable/indexing.html#indexing-view-versus-copy>

```
# This is added back by InteractiveShellApp.init_path()
```

In [23]:

```
# TODO: Import Twit data for selected stock.  
TWITTERADI = pd.read_csv('/Users/khaladdin/Desktop/Twitter Project/ADI.csv')  
TWITTERFID = pd.read_csv('/Users/khaladdin/Desktop/Twitter Project/Fidelity.csv')  
TWITTERFIS = pd.read_csv('/Users/khaladdin/Desktop/Twitter Project/Fiserv.csv')  
TWITTERGPN = pd.read_csv('/Users/khaladdin/Desktop/Twitter Project/Global.csv')  
TWITTERJUN = pd.read_csv('/Users/khaladdin/Desktop/Twitter Project/Juniper.csv')
```

In [25]:

```
# TODO: Keep some columns.  
TWITTERADIa = TWITTERADI[['Row ID', 'Date', 'Time', 'Number of Tweets']]  
TWITTERFIDa = TWITTERFID[['Row ID', 'Date', 'Time', 'Number of Tweets']]  
TWITTERFISa = TWITTERFIS[['Row ID', 'Date', 'Time', 'Number of Tweets']]  
TWITTERGPNa = TWITTERGPN[['Row ID', 'Date', 'Time', 'Number of Tweets']]  
TWITTERJUNa = TWITTERJUN[['Row ID', 'Date', 'Time', 'Number of Tweets']]
```

In [26]:

```
# TODO: Formatting
TWITTERADIA['Date'] = pd.to_datetime(TWITTERADIA['Date'], format='%d/%m/%Y')
TWITTERFIDA['Date'] = pd.to_datetime(TWITTERFIDA['Date'], format='%d/%m/%Y')
TWITTERFISA['Date'] = pd.to_datetime(TWITTERFISA['Date'], format='%d/%m/%Y')
TWITTERGPNA['Date'] = pd.to_datetime(TWITTERGPNA['Date'], format='%d/%m/%Y')
TWITTERJUNA['Date'] = pd.to_datetime(TWITTERJUNA['Date'], format='%d/%m/%Y')
```

/anaconda3/lib/python3.6/site-packages/ipykernel\_launcher.py:1: SettingWithCopyWarning:  
A value is trying to be set on a copy of a slice from a DataFrame.  
Try using .loc[row\_indexer,col\_indexer] = value instead

See the caveats in the documentation: <http://pandas.pydata.org/pandas-docs/stable/indexing.html#indexing-view-versus-copy>

"""Entry point for launching an IPython kernel.

/anaconda3/lib/python3.6/site-packages/ipykernel\_launcher.py:2: SettingWithCopyWarning:  
A value is trying to be set on a copy of a slice from a DataFrame.  
Try using .loc[row\_indexer,col\_indexer] = value instead

See the caveats in the documentation: <http://pandas.pydata.org/pandas-docs/stable/indexing.html#indexing-view-versus-copy>

/anaconda3/lib/python3.6/site-packages/ipykernel\_launcher.py:3: SettingWithCopyWarning:  
A value is trying to be set on a copy of a slice from a DataFrame.  
Try using .loc[row\_indexer,col\_indexer] = value instead

See the caveats in the documentation: <http://pandas.pydata.org/pandas-docs/stable/indexing.html#indexing-view-versus-copy>

This is separate from the ipykernel package so we can avoid doing imports until

/anaconda3/lib/python3.6/site-packages/ipykernel\_launcher.py:4: SettingWithCopyWarning:  
A value is trying to be set on a copy of a slice from a DataFrame.  
Try using .loc[row\_indexer,col\_indexer] = value instead

See the caveats in the documentation: <http://pandas.pydata.org/pandas-docs/stable/indexing.html#indexing-view-versus-copy>

after removing the cwd from sys.path.

/anaconda3/lib/python3.6/site-packages/ipykernel\_launcher.py:5: SettingWithCopyWarning:  
A value is trying to be set on a copy of a slice from a DataFrame.  
Try using .loc[row\_indexer,col\_indexer] = value instead

See the caveats in the documentation: <http://pandas.pydata.org/pandas-docs/stable/indexing.html#indexing-view-versus-copy>

"""



In [27]:

```
# TODO: Formatting
TWITTERADIA['Date1'] = TWITTERADIA['Date'].dt.date
TWITTERFIDA['Date1'] = TWITTERFIDA['Date'].dt.date
TWITTERFISA['Date1'] = TWITTERFISA['Date'].dt.date
TWITTERGPNa['Date1'] = TWITTERGPNa['Date'].dt.date
TWITTERJUNa['Date1'] = TWITTERJUNa['Date'].dt.date
```

/anaconda3/lib/python3.6/site-packages/ipykernel\_launcher.py:1: SettingWithCopyWarning:

A value is trying to be set on a copy of a slice from a DataFrame.  
Try using .loc[row\_indexer,col\_indexer] = value instead

See the caveats in the documentation: <http://pandas.pydata.org/pandas-docs/stable/indexing.html#indexing-view-versus-copy>

"""Entry point for launching an IPython kernel.

/anaconda3/lib/python3.6/site-packages/ipykernel\_launcher.py:2: SettingWithCopyWarning:

A value is trying to be set on a copy of a slice from a DataFrame.  
Try using .loc[row\_indexer,col\_indexer] = value instead

See the caveats in the documentation: <http://pandas.pydata.org/pandas-docs/stable/indexing.html#indexing-view-versus-copy>

/anaconda3/lib/python3.6/site-packages/ipykernel\_launcher.py:3: SettingWithCopyWarning:

A value is trying to be set on a copy of a slice from a DataFrame.  
Try using .loc[row\_indexer,col\_indexer] = value instead

See the caveats in the documentation: <http://pandas.pydata.org/pandas-docs/stable/indexing.html#indexing-view-versus-copy>

This is separate from the ipykernel package so we can avoid doing imports until

/anaconda3/lib/python3.6/site-packages/ipykernel\_launcher.py:4: SettingWithCopyWarning:

A value is trying to be set on a copy of a slice from a DataFrame.  
Try using .loc[row\_indexer,col\_indexer] = value instead

See the caveats in the documentation: <http://pandas.pydata.org/pandas-docs/stable/indexing.html#indexing-view-versus-copy>

after removing the cwd from sys.path.

/anaconda3/lib/python3.6/site-packages/ipykernel\_launcher.py:5: SettingWithCopyWarning:

A value is trying to be set on a copy of a slice from a DataFrame.  
Try using .loc[row\_indexer,col\_indexer] = value instead

See the caveats in the documentation: <http://pandas.pydata.org/pandas-docs/stable/indexing.html#indexing-view-versus-copy>

"""

In [28]:

```
# TODO: Formatting
TWITTERADIA['Time'] = pd.to_datetime(TWITTERADIA['Time'])
TWITTERFIDA['Time'] = pd.to_datetime(TWITTERFIDA['Time'])
TWITTERFISA['Time'] = pd.to_datetime(TWITTERFISA['Time'])
TWITTERGPNa['Time'] = pd.to_datetime(TWITTERGPNa['Time'])
TWITTERJUNa['Time'] = pd.to_datetime(TWITTERJUNa['Time'])
```

/anaconda3/lib/python3.6/site-packages/ipykernel\_launcher.py:1: SettingWithCopyWarning:

A value is trying to be set on a copy of a slice from a DataFrame.  
Try using .loc[row\_indexer,col\_indexer] = value instead

See the caveats in the documentation: <http://pandas.pydata.org/pandas-docs/stable/indexing.html#indexing-view-versus-copy>

"""Entry point for launching an IPython kernel.

/anaconda3/lib/python3.6/site-packages/ipykernel\_launcher.py:2: SettingWithCopyWarning:

A value is trying to be set on a copy of a slice from a DataFrame.  
Try using .loc[row\_indexer,col\_indexer] = value instead

See the caveats in the documentation: <http://pandas.pydata.org/pandas-docs/stable/indexing.html#indexing-view-versus-copy>

/anaconda3/lib/python3.6/site-packages/ipykernel\_launcher.py:3: SettingWithCopyWarning:

A value is trying to be set on a copy of a slice from a DataFrame.  
Try using .loc[row\_indexer,col\_indexer] = value instead

See the caveats in the documentation: <http://pandas.pydata.org/pandas-docs/stable/indexing.html#indexing-view-versus-copy>

This is separate from the ipykernel package so we can avoid doing imports until

/anaconda3/lib/python3.6/site-packages/ipykernel\_launcher.py:4: SettingWithCopyWarning:

A value is trying to be set on a copy of a slice from a DataFrame.  
Try using .loc[row\_indexer,col\_indexer] = value instead

See the caveats in the documentation: <http://pandas.pydata.org/pandas-docs/stable/indexing.html#indexing-view-versus-copy>

after removing the cwd from sys.path.

/anaconda3/lib/python3.6/site-packages/ipykernel\_launcher.py:5: SettingWithCopyWarning:

A value is trying to be set on a copy of a slice from a DataFrame.  
Try using .loc[row\_indexer,col\_indexer] = value instead

See the caveats in the documentation: <http://pandas.pydata.org/pandas-docs/stable/indexing.html#indexing-view-versus-copy>

"""

In [29]:

```
# TODO: Formatting
TWITTERADIA['Time1'] = TWITTERADIA['Time'].dt.time
TWITTERFIDA['Time1'] = TWITTERFIDA['Time'].dt.time
TWITTERFISA['Time1'] = TWITTERFISA['Time'].dt.time
TWITTERGPNa['Time1'] = TWITTERGPNa['Time'].dt.time
TWITTERJUNa['Time1'] = TWITTERJUNa['Time'].dt.time
```

/anaconda3/lib/python3.6/site-packages/ipykernel\_launcher.py:1: SettingWithCopyWarning:

A value is trying to be set on a copy of a slice from a DataFrame.  
Try using .loc[row\_indexer,col\_indexer] = value instead

See the caveats in the documentation: <http://pandas.pydata.org/pandas-docs/stable/indexing.html#indexing-view-versus-copy>

"""Entry point for launching an IPython kernel.

/anaconda3/lib/python3.6/site-packages/ipykernel\_launcher.py:2: SettingWithCopyWarning:

A value is trying to be set on a copy of a slice from a DataFrame.  
Try using .loc[row\_indexer,col\_indexer] = value instead

See the caveats in the documentation: <http://pandas.pydata.org/pandas-docs/stable/indexing.html#indexing-view-versus-copy>

/anaconda3/lib/python3.6/site-packages/ipykernel\_launcher.py:3: SettingWithCopyWarning:

A value is trying to be set on a copy of a slice from a DataFrame.  
Try using .loc[row\_indexer,col\_indexer] = value instead

See the caveats in the documentation: <http://pandas.pydata.org/pandas-docs/stable/indexing.html#indexing-view-versus-copy>

This is separate from the ipykernel package so we can avoid doing imports until

/anaconda3/lib/python3.6/site-packages/ipykernel\_launcher.py:4: SettingWithCopyWarning:

A value is trying to be set on a copy of a slice from a DataFrame.  
Try using .loc[row\_indexer,col\_indexer] = value instead

See the caveats in the documentation: <http://pandas.pydata.org/pandas-docs/stable/indexing.html#indexing-view-versus-copy>

after removing the cwd from sys.path.

/anaconda3/lib/python3.6/site-packages/ipykernel\_launcher.py:5: SettingWithCopyWarning:

A value is trying to be set on a copy of a slice from a DataFrame.  
Try using .loc[row\_indexer,col\_indexer] = value instead

See the caveats in the documentation: <http://pandas.pydata.org/pandas-docs/stable/indexing.html#indexing-view-versus-copy>

"""

In [31]:

```
# TODO: Create datetime column
ADIstatea.loc[:, 'Datetime'] = pd.to_datetime(ADIstatea.Date1.astype(str)+' '+A
DIstatea.Time1.astype(str))
FIDstatea.loc[:, 'Datetime'] = pd.to_datetime(FIDstatea.Date1.astype(str)+' '+F
IDstatea.Time1.astype(str))
FISstatea.loc[:, 'Datetime'] = pd.to_datetime(FISstatea.Date1.astype(str)+' '+F
ISstatea.Time1.astype(str))
GPNstatea.loc[:, 'Datetime'] = pd.to_datetime(GPNstatea.Date1.astype(str)+' '+G
PNstatea.Time1.astype(str))
JUNstatea.loc[:, 'Datetime'] = pd.to_datetime(JUNstatea.Date1.astype(str)+' '+J
UNstatea.Time1.astype(str))
```

In [32]:

```
# TODO: Create datetime column
TWITTERADIA.loc[:, 'Datetime'] = pd.to_datetime(TWITTERADIA.Date1.astype(str)+'
'+TWITTERADIA.Time1.astype(str))
TWITTERFIDA.loc[:, 'Datetime'] = pd.to_datetime(TWITTERFIDA.Date1.astype(str)+'
'+TWITTERFIDA.Time1.astype(str))
TWITTERFISA.loc[:, 'Datetime'] = pd.to_datetime(TWITTERFISA.Date1.astype(str)+'
'+TWITTERFISA.Time1.astype(str))
TWITTERGPNa.loc[:, 'Datetime'] = pd.to_datetime(TWITTERGPNa.Date1.astype(str)+'
'+TWITTERGPNa.Time1.astype(str))
TWITTERJUNa.loc[:, 'Datetime'] = pd.to_datetime(TWITTERJUNa.Date1.astype(str)+'
'+TWITTERJUNa.Time1.astype(str))
```

/anaconda3/lib/python3.6/site-packages/pandas/core/indexing.py:357  
: SettingWithCopyWarning:  
A value is trying to be set on a copy of a slice from a DataFrame.  
Try using .loc[row\_indexer,col\_indexer] = value instead

See the caveats in the documentation: <http://pandas.pydata.org/pandas-docs/stable/indexing.html#indexing-view-versus-copy>

```
self.obj[key] = _infer_fill_value(value)
/anaconda3/lib/python3.6/site-packages/pandas/core/indexing.py:537  
: SettingWithCopyWarning:  
A value is trying to be set on a copy of a slice from a DataFrame.  
Try using .loc[row_indexer,col_indexer] = value instead
```

See the caveats in the documentation: <http://pandas.pydata.org/pandas-docs/stable/indexing.html#indexing-view-versus-copy>

```
self.obj[item] = s
```

In [33]:

```
# TODO: Create twit dummy
TWITTERADIA['dummy'] = 1
TWITTERFIDA['dummy'] = 1
TWITTERFISA['dummy'] = 1
TWITTERGPNa['dummy'] = 1
TWITTERJUNa['dummy'] = 1
```

/anaconda3/lib/python3.6/site-packages/ipykernel\_launcher.py:2: SettingWithCopyWarning:

A value is trying to be set on a copy of a slice from a DataFrame.  
Try using `.loc[row_indexer,col_indexer] = value` instead

See the caveats in the documentation: <http://pandas.pydata.org/pandas-docs/stable/indexing.html#indexing-view-versus-copy>

/anaconda3/lib/python3.6/site-packages/ipykernel\_launcher.py:3: SettingWithCopyWarning:

A value is trying to be set on a copy of a slice from a DataFrame.  
Try using `.loc[row_indexer,col_indexer] = value` instead

See the caveats in the documentation: <http://pandas.pydata.org/pandas-docs/stable/indexing.html#indexing-view-versus-copy>

This is separate from the ipykernel package so we can avoid doing imports until

/anaconda3/lib/python3.6/site-packages/ipykernel\_launcher.py:4: SettingWithCopyWarning:

A value is trying to be set on a copy of a slice from a DataFrame.  
Try using `.loc[row_indexer,col_indexer] = value` instead

See the caveats in the documentation: <http://pandas.pydata.org/pandas-docs/stable/indexing.html#indexing-view-versus-copy>

after removing the cwd from sys.path.

/anaconda3/lib/python3.6/site-packages/ipykernel\_launcher.py:5: SettingWithCopyWarning:

A value is trying to be set on a copy of a slice from a DataFrame.  
Try using `.loc[row_indexer,col_indexer] = value` instead

See the caveats in the documentation: <http://pandas.pydata.org/pandas-docs/stable/indexing.html#indexing-view-versus-copy>

"""

In [35]:

```
# TODO: Delete second part
TWITTERADIA['Datetime'] = TWITTERADIA['Datetime'].apply(lambda x: x.replace(second=0))
TWITTERFIDA['Datetime'] = TWITTERFIDA['Datetime'].apply(lambda x: x.replace(second=0))
TWITTERFISA['Datetime'] = TWITTERFISA['Datetime'].apply(lambda x: x.replace(second=0))
TWITTERGPNA['Datetime'] = TWITTERGPNA['Datetime'].apply(lambda x: x.replace(second=0))
TWITTERJUNA['Datetime'] = TWITTERJUNA['Datetime'].apply(lambda x: x.replace(second=0))
```

/anaconda3/lib/python3.6/site-packages/ipykernel\_launcher.py:2: SettingWithCopyWarning:  
A value is trying to be set on a copy of a slice from a DataFrame.  
Try using .loc[row\_indexer,col\_indexer] = value instead

See the caveats in the documentation: <http://pandas.pydata.org/pandas-docs/stable/indexing.html#indexing-view-versus-copy>

/anaconda3/lib/python3.6/site-packages/ipykernel\_launcher.py:3: SettingWithCopyWarning:  
A value is trying to be set on a copy of a slice from a DataFrame.  
Try using .loc[row\_indexer,col\_indexer] = value instead

See the caveats in the documentation: <http://pandas.pydata.org/pandas-docs/stable/indexing.html#indexing-view-versus-copy>

This is separate from the ipykernel package so we can avoid doing imports until

/anaconda3/lib/python3.6/site-packages/ipykernel\_launcher.py:4: SettingWithCopyWarning:  
A value is trying to be set on a copy of a slice from a DataFrame.  
Try using .loc[row\_indexer,col\_indexer] = value instead

See the caveats in the documentation: <http://pandas.pydata.org/pandas-docs/stable/indexing.html#indexing-view-versus-copy>

after removing the cwd from sys.path.

/anaconda3/lib/python3.6/site-packages/ipykernel\_launcher.py:5: SettingWithCopyWarning:  
A value is trying to be set on a copy of a slice from a DataFrame.  
Try using .loc[row\_indexer,col\_indexer] = value instead

See the caveats in the documentation: <http://pandas.pydata.org/pandas-docs/stable/indexing.html#indexing-view-versus-copy>

"""

In [36]:

```
# TODO: Keep some columns.
TWITTERADIA = TWITTERADIA[['Datetime', 'Date1', 'Time1', 'dummy']]
TWITTERFIDA = TWITTERFIDA[['Datetime', 'Date1', 'Time1', 'dummy']]
TWITTERFISA = TWITTERFISA[['Datetime', 'Date1', 'Time1', 'dummy']]
TWITTERGPNA = TWITTERGPNA[['Datetime', 'Date1', 'Time1', 'dummy']]
TWITTERJUNA = TWITTERJUNA[['Datetime', 'Date1', 'Time1', 'dummy']]
```

In [37]:

```
# TODO: Merge datasets.
Task5ADI = pd.merge(ADISTATEA, TWITTERADIA, on='Datetime', how='outer')
Task5FID = pd.merge(FIDSTATEA, TWITTERFIDA, on='Datetime', how='outer')
Task5FIS = pd.merge(FISSTATEA, TWITTERFISA, on='Datetime', how='outer')
Task5GPN = pd.merge(GPNSTATEA, TWITTERGPNA, on='Datetime', how='outer')
Task5JUN = pd.merge(JUNSTATEA, TWITTERJUNA, on='Datetime', how='outer')
```

In [38]:

```
# TODO: Delete missing.
Task5ADI = Task5ADI[np.isfinite(Task5ADI['Irregular'])]
Task5FID = Task5FID[np.isfinite(Task5FID['Irregular'])]
Task5FIS = Task5FIS[np.isfinite(Task5FIS['Irregular'])]
Task5GPN = Task5GPN[np.isfinite(Task5GPN['Irregular'])]
Task5JUN = Task5JUN[np.isfinite(Task5JUN['Irregular'])]
```

In [39]:

```
# TODO: 0s and 1s for dummy
Task5ADI['dummy'].fillna(0, inplace=True)
Task5FID['dummy'].fillna(0, inplace=True)
Task5FIS['dummy'].fillna(0, inplace=True)
Task5GPN['dummy'].fillna(0, inplace=True)
Task5JUN['dummy'].fillna(0, inplace=True)
```

In [41]:

```
# TODO: Create separate datasets for each stocks
ADITRade = TWITTER.loc[TWITTER['#RIC'] == 'ADI.OQ']
FIDTRade = TWITTER.loc[TWITTER['#RIC'] == 'FIS.N']
FISTRade = TWITTER.loc[TWITTER['#RIC'] == 'FISV.OQ']
GPNTRade = TWITTER.loc[TWITTER['#RIC'] == 'GPN.N']
JUNTRade = TWITTER.loc[TWITTER['#RIC'] == 'JNPR.N']
```

In [42]:

```
# TODO: Delete quote data
ADITRadea = ADITRade[ADITRade.Type != 'Quote']
FIDTRadea = FIDTRade[FIDTRade.Type != 'Quote']
FISTRadea = FISTRade[FISTRade.Type != 'Quote']
GPNTRadea = GPNTRade[GPNTRade.Type != 'Quote']
JUNTRadea = JUNTRade[JUNTRade.Type != 'Quote']
```

In [43]:

```
# TODO: Formatting
ADITradea['Date'] = pd.to_datetime(ADITradea['Date-Time'])
FIDTradea['Date'] = pd.to_datetime(FIDTradea['Date-Time'])
FISTradea['Date'] = pd.to_datetime(FISTradea['Date-Time'])
GPNTradea['Date'] = pd.to_datetime(GPNTradea['Date-Time'])
JUNTradea['Date'] = pd.to_datetime(JUNTradea['Date-Time'])
```

/anaconda3/lib/python3.6/site-packages/ipykernel\_launcher.py:1: SettingWithCopyWarning:

A value is trying to be set on a copy of a slice from a DataFrame.  
Try using .loc[row\_indexer,col\_indexer] = value instead

See the caveats in the documentation: <http://pandas.pydata.org/pandas-docs/stable/indexing.html#indexing-view-versus-copy>

"""Entry point for launching an IPython kernel.

/anaconda3/lib/python3.6/site-packages/ipykernel\_launcher.py:2: SettingWithCopyWarning:

A value is trying to be set on a copy of a slice from a DataFrame.  
Try using .loc[row\_indexer,col\_indexer] = value instead

See the caveats in the documentation: <http://pandas.pydata.org/pandas-docs/stable/indexing.html#indexing-view-versus-copy>

/anaconda3/lib/python3.6/site-packages/ipykernel\_launcher.py:3: SettingWithCopyWarning:

A value is trying to be set on a copy of a slice from a DataFrame.  
Try using .loc[row\_indexer,col\_indexer] = value instead

See the caveats in the documentation: <http://pandas.pydata.org/pandas-docs/stable/indexing.html#indexing-view-versus-copy>

This is separate from the ipykernel package so we can avoid doing imports until

/anaconda3/lib/python3.6/site-packages/ipykernel\_launcher.py:4: SettingWithCopyWarning:

A value is trying to be set on a copy of a slice from a DataFrame.  
Try using .loc[row\_indexer,col\_indexer] = value instead

See the caveats in the documentation: <http://pandas.pydata.org/pandas-docs/stable/indexing.html#indexing-view-versus-copy>

after removing the cwd from sys.path.

/anaconda3/lib/python3.6/site-packages/ipykernel\_launcher.py:5: SettingWithCopyWarning:

A value is trying to be set on a copy of a slice from a DataFrame.  
Try using .loc[row\_indexer,col\_indexer] = value instead

See the caveats in the documentation: <http://pandas.pydata.org/pandas-docs/stable/indexing.html#indexing-view-versus-copy>

"""



In [44]:

```
# TODO: Formatting
ADITradea['Date1'] = ADITradea['Date'].dt.date
FIDTradea['Date1'] = FIDTradea['Date'].dt.date
FISTradea['Date1'] = FISTradea['Date'].dt.date
GPNTradea['Date1'] = GPNTradea['Date'].dt.date
JUNTradea['Date1'] = JUNTradea['Date'].dt.date
```

/anaconda3/lib/python3.6/site-packages/ipykernel\_launcher.py:1: SettingWithCopyWarning:

A value is trying to be set on a copy of a slice from a DataFrame.  
Try using .loc[row\_indexer,col\_indexer] = value instead

See the caveats in the documentation: <http://pandas.pydata.org/pandas-docs/stable/indexing.html#indexing-view-versus-copy>

"""Entry point for launching an IPython kernel.

/anaconda3/lib/python3.6/site-packages/ipykernel\_launcher.py:2: SettingWithCopyWarning:

A value is trying to be set on a copy of a slice from a DataFrame.  
Try using .loc[row\_indexer,col\_indexer] = value instead

See the caveats in the documentation: <http://pandas.pydata.org/pandas-docs/stable/indexing.html#indexing-view-versus-copy>

/anaconda3/lib/python3.6/site-packages/ipykernel\_launcher.py:3: SettingWithCopyWarning:

A value is trying to be set on a copy of a slice from a DataFrame.  
Try using .loc[row\_indexer,col\_indexer] = value instead

See the caveats in the documentation: <http://pandas.pydata.org/pandas-docs/stable/indexing.html#indexing-view-versus-copy>

This is separate from the ipykernel package so we can avoid doing imports until

/anaconda3/lib/python3.6/site-packages/ipykernel\_launcher.py:4: SettingWithCopyWarning:

A value is trying to be set on a copy of a slice from a DataFrame.  
Try using .loc[row\_indexer,col\_indexer] = value instead

See the caveats in the documentation: <http://pandas.pydata.org/pandas-docs/stable/indexing.html#indexing-view-versus-copy>

after removing the cwd from sys.path.

/anaconda3/lib/python3.6/site-packages/ipykernel\_launcher.py:5: SettingWithCopyWarning:

A value is trying to be set on a copy of a slice from a DataFrame.  
Try using .loc[row\_indexer,col\_indexer] = value instead

See the caveats in the documentation: <http://pandas.pydata.org/pandas-docs/stable/indexing.html#indexing-view-versus-copy>

"""

In [45]:

```
# TODO: Formatting
ADITradea['Time1'] = ADITradea['Date'].dt.time
FIDTradea['Time1'] = FIDTradea['Date'].dt.time
FISTradea['Time1'] = FISTradea['Date'].dt.time
GPNTradea['Time1'] = GPNTradea['Date'].dt.time
JUNTradea['Time1'] = JUNTradea['Date'].dt.time
```

```
/anaconda3/lib/python3.6/site-packages/ipykernel_launcher.py:1: SettingWithCopyWarning:
```

A value is trying to be set on a copy of a slice from a DataFrame.  
Try using `.loc[row_indexer,col_indexer] = value` instead

See the caveats in the documentation: <http://pandas.pydata.org/pandas-docs/stable/indexing.html#indexing-view-versus-copy>

```
"""Entry point for launching an IPython kernel.
```

```
/anaconda3/lib/python3.6/site-packages/ipykernel_launcher.py:2: SettingWithCopyWarning:
```

A value is trying to be set on a copy of a slice from a DataFrame.  
Try using `.loc[row_indexer,col_indexer] = value` instead

See the caveats in the documentation: <http://pandas.pydata.org/pandas-docs/stable/indexing.html#indexing-view-versus-copy>

```
/anaconda3/lib/python3.6/site-packages/ipykernel_launcher.py:3: SettingWithCopyWarning:
```

A value is trying to be set on a copy of a slice from a DataFrame.  
Try using `.loc[row_indexer,col_indexer] = value` instead

See the caveats in the documentation: <http://pandas.pydata.org/pandas-docs/stable/indexing.html#indexing-view-versus-copy>

This is separate from the ipykernel package so we can avoid doing imports until

```
/anaconda3/lib/python3.6/site-packages/ipykernel_launcher.py:4: SettingWithCopyWarning:
```

A value is trying to be set on a copy of a slice from a DataFrame.  
Try using `.loc[row_indexer,col_indexer] = value` instead

See the caveats in the documentation: <http://pandas.pydata.org/pandas-docs/stable/indexing.html#indexing-view-versus-copy>

```
after removing the cwd from sys.path.
```

```
/anaconda3/lib/python3.6/site-packages/ipykernel_launcher.py:5: SettingWithCopyWarning:
```

A value is trying to be set on a copy of a slice from a DataFrame.  
Try using `.loc[row_indexer,col_indexer] = value` instead

See the caveats in the documentation: <http://pandas.pydata.org/pandas-docs/stable/indexing.html#indexing-view-versus-copy>

```
"""
```

In [46]:

```
# TODO: Clean microsecond
ADITradea['Time1'] = ADITradea['Time1'].apply(lambda x: x.replace(microsecond=
0))
FIDTradea['Time1'] = FIDTradea['Time1'].apply(lambda x: x.replace(microsecond=
0))
FISTradea['Time1'] = FISTradea['Time1'].apply(lambda x: x.replace(microsecond=
0))
GPNTradea['Time1'] = GPNTradea['Time1'].apply(lambda x: x.replace(microsecond=
0))
JUNTradea['Time1'] = JUNTradea['Time1'].apply(lambda x: x.replace(microsecond=
0))
```

/anaconda3/lib/python3.6/site-packages/ipykernel\_launcher.py:1: Se  
ttingWithCopyWarning:  
A value is trying to be set on a copy of a slice from a DataFrame.  
Try using .loc[row\_indexer,col\_indexer] = value instead

See the caveats in the documentation: <http://pandas.pydata.org/pandas-docs/stable/indexing.html#indexing-view-versus-copy>

```
"""Entry point for launching an IPython kernel.
/anaconda3/lib/python3.6/site-packages/ipykernel_launcher.py:2: Se
ttingWithCopyWarning:
A value is trying to be set on a copy of a slice from a DataFrame.
Try using .loc[row_indexer,col_indexer] = value instead
```

See the caveats in the documentation: <http://pandas.pydata.org/pandas-docs/stable/indexing.html#indexing-view-versus-copy>

```
/anaconda3/lib/python3.6/site-packages/ipykernel_launcher.py:3: Se
ttingWithCopyWarning:
A value is trying to be set on a copy of a slice from a DataFrame.
Try using .loc[row_indexer,col_indexer] = value instead
```

See the caveats in the documentation: <http://pandas.pydata.org/pandas-docs/stable/indexing.html#indexing-view-versus-copy>

This is separate from the ipykernel package so we can avoid doing imports until

```
/anaconda3/lib/python3.6/site-packages/ipykernel_launcher.py:4: Se
ttingWithCopyWarning:
A value is trying to be set on a copy of a slice from a DataFrame.
Try using .loc[row_indexer,col_indexer] = value instead
```

See the caveats in the documentation: <http://pandas.pydata.org/pandas-docs/stable/indexing.html#indexing-view-versus-copy>

```
after removing the cwd from sys.path.
/anaconda3/lib/python3.6/site-packages/ipykernel_launcher.py:5: Se
ttingWithCopyWarning:
A value is trying to be set on a copy of a slice from a DataFrame.
Try using .loc[row_indexer,col_indexer] = value instead
```

See the caveats in the documentation: <http://pandas.pydata.org/pandas-docs/stable/indexing.html#indexing-view-versus-copy>

```
"""
```

In [47]:

```
# TODO: Clean second
ADITradea['Time1'] = ADITradea['Time1'].apply(lambda x: x.replace(second=0))
FIDTradea['Time1'] = FIDTradea['Time1'].apply(lambda x: x.replace(second=0))
FISTradea['Time1'] = FISTradea['Time1'].apply(lambda x: x.replace(second=0))
GPNTradea['Time1'] = GPNTradea['Time1'].apply(lambda x: x.replace(second=0))
JUNTradea['Time1'] = JUNTradea['Time1'].apply(lambda x: x.replace(second=0))
```

/anaconda3/lib/python3.6/site-packages/ipykernel\_launcher.py:1: SettingWithCopyWarning:

A value is trying to be set on a copy of a slice from a DataFrame.  
Try using .loc[row\_indexer,col\_indexer] = value instead

See the caveats in the documentation: <http://pandas.pydata.org/pandas-docs/stable/indexing.html#indexing-view-versus-copy>

"""Entry point for launching an IPython kernel.

/anaconda3/lib/python3.6/site-packages/ipykernel\_launcher.py:2: SettingWithCopyWarning:

A value is trying to be set on a copy of a slice from a DataFrame.  
Try using .loc[row\_indexer,col\_indexer] = value instead

See the caveats in the documentation: <http://pandas.pydata.org/pandas-docs/stable/indexing.html#indexing-view-versus-copy>

/anaconda3/lib/python3.6/site-packages/ipykernel\_launcher.py:3: SettingWithCopyWarning:

A value is trying to be set on a copy of a slice from a DataFrame.  
Try using .loc[row\_indexer,col\_indexer] = value instead

See the caveats in the documentation: <http://pandas.pydata.org/pandas-docs/stable/indexing.html#indexing-view-versus-copy>

This is separate from the ipykernel package so we can avoid doing imports until

/anaconda3/lib/python3.6/site-packages/ipykernel\_launcher.py:4: SettingWithCopyWarning:

A value is trying to be set on a copy of a slice from a DataFrame.  
Try using .loc[row\_indexer,col\_indexer] = value instead

See the caveats in the documentation: <http://pandas.pydata.org/pandas-docs/stable/indexing.html#indexing-view-versus-copy>

after removing the cwd from sys.path.

/anaconda3/lib/python3.6/site-packages/ipykernel\_launcher.py:5: SettingWithCopyWarning:

A value is trying to be set on a copy of a slice from a DataFrame.  
Try using .loc[row\_indexer,col\_indexer] = value instead

See the caveats in the documentation: <http://pandas.pydata.org/pandas-docs/stable/indexing.html#indexing-view-versus-copy>

"""

In [48]:

```
# TODO: Create datetime column
ADITradea.loc[:, 'Datetime'] = pd.to_datetime(ADITradea.Date1.astype(str)+' '+A
DITradea.Time1.astype(str))
FIDTradea.loc[:, 'Datetime'] = pd.to_datetime(FIDTradea.Date1.astype(str)+' '+F
IDTradea.Time1.astype(str))
FISTradea.loc[:, 'Datetime'] = pd.to_datetime(FISTradea.Date1.astype(str)+' '+F
ISTradea.Time1.astype(str))
GPNTradea.loc[:, 'Datetime'] = pd.to_datetime(GPNTradea.Date1.astype(str)+' '+G
PNTradea.Time1.astype(str))
JUNTradea.loc[:, 'Datetime'] = pd.to_datetime(JUNTradea.Date1.astype(str)+' '+J
UNTradea.Time1.astype(str))
```

```
/anaconda3/lib/python3.6/site-packages/pandas/core/indexing.py:357
: SettingWithCopyWarning:
A value is trying to be set on a copy of a slice from a DataFrame.
Try using .loc[row_indexer,col_indexer] = value instead
```

See the caveats in the documentation: <http://pandas.pydata.org/pandas-docs/stable/indexing.html#indexing-view-versus-copy>

```
self.obj[key] = _infer_fill_value(value)
/anaconda3/lib/python3.6/site-packages/pandas/core/indexing.py:537
: SettingWithCopyWarning:
A value is trying to be set on a copy of a slice from a DataFrame.
Try using .loc[row_indexer,col_indexer] = value instead
```

See the caveats in the documentation: <http://pandas.pydata.org/pandas-docs/stable/indexing.html#indexing-view-versus-copy>

```
self.obj[item] = s
```

In [49]:

```
# TODO: Create column which value is 1 to compute number of transactions for e
ach interval
ADITradea['trans'] = 1
FIDTradea['trans'] = 1
FISTradea['trans'] = 1
GPNTradea['trans'] = 1
JUNTradea['trans'] = 1
```

```

/anaconda3/lib/python3.6/site-packages/ipykernel_launcher.py:1: Se
ttingWithCopyWarning:
A value is trying to be set on a copy of a slice from a DataFrame.
Try using .loc[row_indexer,col_indexer] = value instead

See the caveats in the documentation: http://pandas.pydata.org/pan
das-docs/stable/indexing.html#indexing-view-versus-copy
    """Entry point for launching an IPython kernel.
/anaconda3/lib/python3.6/site-packages/ipykernel_launcher.py:2: Se
ttingWithCopyWarning:
A value is trying to be set on a copy of a slice from a DataFrame.
Try using .loc[row_indexer,col_indexer] = value instead

See the caveats in the documentation: http://pandas.pydata.org/pan
das-docs/stable/indexing.html#indexing-view-versus-copy

/anaconda3/lib/python3.6/site-packages/ipykernel_launcher.py:3: Se
ttingWithCopyWarning:
A value is trying to be set on a copy of a slice from a DataFrame.
Try using .loc[row_indexer,col_indexer] = value instead

See the caveats in the documentation: http://pandas.pydata.org/pan
das-docs/stable/indexing.html#indexing-view-versus-copy
    This is separate from the ipykernel package so we can avoid doin
g imports until
/anaconda3/lib/python3.6/site-packages/ipykernel_launcher.py:4: Se
ttingWithCopyWarning:
A value is trying to be set on a copy of a slice from a DataFrame.
Try using .loc[row_indexer,col_indexer] = value instead

See the caveats in the documentation: http://pandas.pydata.org/pan
das-docs/stable/indexing.html#indexing-view-versus-copy
    after removing the cwd from sys.path.
/anaconda3/lib/python3.6/site-packages/ipykernel_launcher.py:5: Se
ttingWithCopyWarning:
A value is trying to be set on a copy of a slice from a DataFrame.
Try using .loc[row_indexer,col_indexer] = value instead

See the caveats in the documentation: http://pandas.pydata.org/pan
das-docs/stable/indexing.html#indexing-view-versus-copy
    """

```

In [50]:

```
# TODO: Compute volume for minute interval
ADITradea['Volumecum'] = ADITradea.groupby(['Datetime'])['Volume'].apply(lambda x: x.cumsum())
FIDTradea['Volumecum'] = FIDTradea.groupby(['Datetime'])['Volume'].apply(lambda x: x.cumsum())
FISTradea['Volumecum'] = FISTradea.groupby(['Datetime'])['Volume'].apply(lambda x: x.cumsum())
GPNTradea['Volumecum'] = GPNTradea.groupby(['Datetime'])['Volume'].apply(lambda x: x.cumsum())
JUNTradea['Volumecum'] = JUNTradea.groupby(['Datetime'])['Volume'].apply(lambda x: x.cumsum())
```

/anaconda3/lib/python3.6/site-packages/ipykernel\_launcher.py:1: SettingWithCopyWarning:  
A value is trying to be set on a copy of a slice from a DataFrame.  
Try using .loc[row\_indexer,col\_indexer] = value instead

See the caveats in the documentation: <http://pandas.pydata.org/pandas-docs/stable/indexing.html#indexing-view-versus-copy>

```
"""Entry point for launching an IPython kernel.
/anaconda3/lib/python3.6/site-packages/ipykernel_launcher.py:2: SettingWithCopyWarning:
A value is trying to be set on a copy of a slice from a DataFrame.
Try using .loc[row_indexer,col_indexer] = value instead
```

See the caveats in the documentation: <http://pandas.pydata.org/pandas-docs/stable/indexing.html#indexing-view-versus-copy>

```
/anaconda3/lib/python3.6/site-packages/ipykernel_launcher.py:3: SettingWithCopyWarning:
A value is trying to be set on a copy of a slice from a DataFrame.
Try using .loc[row_indexer,col_indexer] = value instead
```

See the caveats in the documentation: <http://pandas.pydata.org/pandas-docs/stable/indexing.html#indexing-view-versus-copy>

```
This is separate from the ipykernel package so we can avoid doing imports until
/anaconda3/lib/python3.6/site-packages/ipykernel_launcher.py:4: SettingWithCopyWarning:
A value is trying to be set on a copy of a slice from a DataFrame.
Try using .loc[row_indexer,col_indexer] = value instead
```

See the caveats in the documentation: <http://pandas.pydata.org/pandas-docs/stable/indexing.html#indexing-view-versus-copy>

```
after removing the cwd from sys.path.
/anaconda3/lib/python3.6/site-packages/ipykernel_launcher.py:5: SettingWithCopyWarning:
A value is trying to be set on a copy of a slice from a DataFrame.
Try using .loc[row_indexer,col_indexer] = value instead
```

See the caveats in the documentation: <http://pandas.pydata.org/pandas-docs/stable/indexing.html#indexing-view-versus-copy>  
"""

In [51]:

```
# TODO: Compute transactions for each interval (minute)
ADITradea['Transcum'] = ADITradea.groupby(['Datetime'])['trans'].apply(lambda
x: x.cumsum())
FIDTradea['Transcum'] = FIDTradea.groupby(['Datetime'])['trans'].apply(lambda
x: x.cumsum())
FISTradea['Transcum'] = FISTradea.groupby(['Datetime'])['trans'].apply(lambda
x: x.cumsum())
GPNTradea['Transcum'] = GPNTradea.groupby(['Datetime'])['trans'].apply(lambda
x: x.cumsum())
JUNTradea['Transcum'] = JUNTradea.groupby(['Datetime'])['trans'].apply(lambda
x: x.cumsum())
```

/anaconda3/lib/python3.6/site-packages/ipykernel\_launcher.py:1: SettingWithCopyWarning:  
A value is trying to be set on a copy of a slice from a DataFrame.  
Try using .loc[row\_indexer,col\_indexer] = value instead

See the caveats in the documentation: <http://pandas.pydata.org/pandas-docs/stable/indexing.html#indexing-view-versus-copy>

```
"""Entry point for launching an IPython kernel.
/anaconda3/lib/python3.6/site-packages/ipykernel_launcher.py:2: SettingWithCopyWarning:
A value is trying to be set on a copy of a slice from a DataFrame.
Try using .loc[row_indexer,col_indexer] = value instead
```

See the caveats in the documentation: <http://pandas.pydata.org/pandas-docs/stable/indexing.html#indexing-view-versus-copy>

```
/anaconda3/lib/python3.6/site-packages/ipykernel_launcher.py:3: SettingWithCopyWarning:
A value is trying to be set on a copy of a slice from a DataFrame.
Try using .loc[row_indexer,col_indexer] = value instead
```

See the caveats in the documentation: <http://pandas.pydata.org/pandas-docs/stable/indexing.html#indexing-view-versus-copy>

```
This is separate from the ipykernel package so we can avoid doing imports until
/anaconda3/lib/python3.6/site-packages/ipykernel_launcher.py:4: SettingWithCopyWarning:
A value is trying to be set on a copy of a slice from a DataFrame.
Try using .loc[row_indexer,col_indexer] = value instead
```

See the caveats in the documentation: <http://pandas.pydata.org/pandas-docs/stable/indexing.html#indexing-view-versus-copy>

```
after removing the cwd from sys.path.
/anaconda3/lib/python3.6/site-packages/ipykernel_launcher.py:5: SettingWithCopyWarning:
A value is trying to be set on a copy of a slice from a DataFrame.
Try using .loc[row_indexer,col_indexer] = value instead
```

See the caveats in the documentation: <http://pandas.pydata.org/pandas-docs/stable/indexing.html#indexing-view-versus-copy>  
"""



In [52]:

```
# TODO: Keep only last.
ADITradea = ADITradea.groupby('Datetime', as_index=False).last()
FIDTradea = FIDTradea.groupby('Datetime', as_index=False).last()
FISTradea = FISTradea.groupby('Datetime', as_index=False).last()
GPNTradea = GPNTradea.groupby('Datetime', as_index=False).last()
JUNTradea = JUNTradea.groupby('Datetime', as_index=False).last()
```

In [53]:

```
# TODO: Compute size
ADITradea['size'] = ADITradea['Volumecum']/ADITradea['Transcum']
FIDTradea['size'] = FIDTradea['Volumecum']/FIDTradea['Transcum']
FISTradea['size'] = FISTradea['Volumecum']/FISTradea['Transcum']
GPNTradea['size'] = GPNTradea['Volumecum']/GPNTradea['Transcum']
JUNTradea['size'] = JUNTradea['Volumecum']/JUNTradea['Transcum']
```

In [54]:

```
# TODO: Keep Some columns.
ADITradea = ADITradea[['Datetime', 'Price', 'Volumecum', 'Transcum', 'size']]
FIDTradea = FIDTradea[['Datetime', 'Price', 'Volumecum', 'Transcum', 'size']]
FISTradea = FISTradea[['Datetime', 'Price', 'Volumecum', 'Transcum', 'size']]
GPNTradea = GPNTradea[['Datetime', 'Price', 'Volumecum', 'Transcum', 'size']]
JUNTradea = JUNTradea[['Datetime', 'Price', 'Volumecum', 'Transcum', 'size']]
```

In [55]:

```
# TODO: Merge datasets.
Task6ADI = pd.merge(Task5ADI, ADITradea, on='Datetime', how='outer')
Task6FID = pd.merge(Task5FID, FIDTradea, on='Datetime', how='outer')
Task6FIS = pd.merge(Task5FIS, FISTradea, on='Datetime', how='outer')
Task6GPN = pd.merge(Task5GPN, GPNTradea, on='Datetime', how='outer')
Task6JUN = pd.merge(Task5JUN, JUNTradea, on='Datetime', how='outer')
```

In [56]:

```
# TODO: Delete missing.
Task6ADI = Task6ADI[np.isfinite(Task6ADI['Irregular'])]
Task6FID = Task6FID[np.isfinite(Task6FID['Irregular'])]
Task6FIS = Task6FIS[np.isfinite(Task6FIS['Irregular'])]
Task6GPN = Task6GPN[np.isfinite(Task6GPN['Irregular'])]
Task6JUN = Task6JUN[np.isfinite(Task6JUN['Irregular'])]
```

In [57]:

```
# TODO: Dealing with missing.
Task6ADI['Volumecum'].fillna(0, inplace=True)
Task6FID['Volumecum'].fillna(0, inplace=True)
Task6FIS['Volumecum'].fillna(0, inplace=True)
Task6GPN['Volumecum'].fillna(0, inplace=True)
Task6JUN['Volumecum'].fillna(0, inplace=True)
Task6ADI['Transcum'].fillna(0, inplace=True)
Task6FID['Transcum'].fillna(0, inplace=True)
Task6FIS['Transcum'].fillna(0, inplace=True)
Task6GPN['Transcum'].fillna(0, inplace=True)
Task6JUN['Transcum'].fillna(0, inplace=True)
Task6ADI['size'].fillna(0, inplace=True)
Task6FID['size'].fillna(0, inplace=True)
Task6FIS['size'].fillna(0, inplace=True)
Task6GPN['size'].fillna(0, inplace=True)
Task6JUN['size'].fillna(0, inplace=True)
```

In [59]:

```
# TODO: Keep Some columns.
ADImid = ADIa[['Datetime', 'midprice', 'spread']]
FIDmid = FIDa[['Datetime', 'midprice', 'spread']]
FISmid = FISa[['Datetime', 'midprice', 'spread']]
GPNmid = GPNa[['Datetime', 'midprice', 'spread']]
JUNmid = JUNa[['Datetime', 'midprice', 'spread']]
```

In [61]:

```
# TODO: Keep only last.
ADImid = ADImid.groupby('Datetime', as_index=False).last()
FIDmid = FIDmid.groupby('Datetime', as_index=False).last()
FISmid = FISmid.groupby('Datetime', as_index=False).last()
GPNmid = GPNmid.groupby('Datetime', as_index=False).last()
JUNmid = JUNmid.groupby('Datetime', as_index=False).last()
```

In [62]:

```
# TODO: Merge datasets.
Task7ADI = pd.merge(Task6ADI, ADImid, on='Datetime', how='outer')
Task7FID = pd.merge(Task6FID, FIDmid, on='Datetime', how='outer')
Task7FIS = pd.merge(Task6FIS, FISmid, on='Datetime', how='outer')
Task7GPN = pd.merge(Task6GPN, GPNmid, on='Datetime', how='outer')
Task7JUN = pd.merge(Task6JUN, JUNmid, on='Datetime', how='outer')
```

In [63]:

```
# TODO: Delete missing.
Task7ADI = Task7ADI[np.isfinite(Task7ADI['Volumecum'])]
Task7FID = Task7FID[np.isfinite(Task7FID['Volumecum'])]
Task7FIS = Task7FIS[np.isfinite(Task7FIS['Volumecum'])]
Task7GPN = Task7GPN[np.isfinite(Task7GPN['Volumecum'])]
Task7JUN = Task7JUN[np.isfinite(Task7JUN['Volumecum'])]
```

In [64]:

```
# TODO: Compute effective spread.
Task7ADI['effective'] = 2*((Task7ADI['Price']-Task7ADI['midprice'])).abs()
Task7FID['effective'] = 2*((Task7FID['Price']-Task7FID['midprice'])).abs()
Task7FIS['effective'] = 2*((Task7FIS['Price']-Task7FIS['midprice'])).abs()
Task7GPN['effective'] = 2*((Task7GPN['Price']-Task7GPN['midprice'])).abs()
Task7JUN['effective'] = 2*((Task7JUN['Price']-Task7JUN['midprice'])).abs()
```

In [65]:

```
# TODO: Compute absolute price changes.
Task7ADI['vol'] = (Task7ADI['Price'] - Task7ADI['Price'].shift(1)).abs()
Task7FID['vol'] = (Task7FID['Price'] - Task7FID['Price'].shift(1)).abs()
Task7FIS['vol'] = (Task7FIS['Price'] - Task7FIS['Price'].shift(1)).abs()
Task7GPN['vol'] = (Task7GPN['Price'] - Task7GPN['Price'].shift(1)).abs()
Task7JUN['vol'] = (Task7JUN['Price'] - Task7JUN['Price'].shift(1)).abs()
```

In [67]:

```
# TODO: Ascengind order. For descengind write False
Task7ADI = Task7ADI.sort_values('Datetime', ascending=True)
Task7FID = Task7FID.sort_values('Datetime', ascending=True)
Task7FIS = Task7FIS.sort_values('Datetime', ascending=True)
Task7GPN = Task7GPN.sort_values('Datetime', ascending=True)
Task7JUN = Task7JUN.sort_values('Datetime', ascending=True)
```

In [68]:

```
# TODO: Lag values of dummy.
Task7ADI.dummy = Task7ADI.dummy.shift(-1)
Task7FID.dummy = Task7FID.dummy.shift(-1)
Task7FIS.dummy = Task7FIS.dummy.shift(-1)
Task7GPN.dummy = Task7GPN.dummy.shift(-1)
Task7JUN.dummy = Task7JUN.dummy.shift(-1)
```

In [69]:

```
# TODO: Add a column of stock name to datasets.
Task7ADI['RIC'] = 'ADI'
Task7FID['RIC'] = 'FID'
Task7FIS['RIC'] = 'FIS'
Task7GPN['RIC'] = 'GPN'
Task7JUN['RIC'] = 'JUN'
```

In [70]:

```
# TODO: Formatting.
ADI['Date'] = pd.to_datetime(ADI['Date-Time'])
FID['Date'] = pd.to_datetime(FID['Date-Time'])
FIS['Date'] = pd.to_datetime(FIS['Date-Time'])
GPN['Date'] = pd.to_datetime(GPN['Date-Time'])
JUN['Date'] = pd.to_datetime(JUN['Date-Time'])
```

/anaconda3/lib/python3.6/site-packages/ipykernel\_launcher.py:1: SettingWithCopyWarning:

A value is trying to be set on a copy of a slice from a DataFrame.  
Try using .loc[row\_indexer,col\_indexer] = value instead

See the caveats in the documentation: <http://pandas.pydata.org/pandas-docs/stable/indexing.html#indexing-view-versus-copy>

"""Entry point for launching an IPython kernel.

/anaconda3/lib/python3.6/site-packages/ipykernel\_launcher.py:2: SettingWithCopyWarning:

A value is trying to be set on a copy of a slice from a DataFrame.  
Try using .loc[row\_indexer,col\_indexer] = value instead

See the caveats in the documentation: <http://pandas.pydata.org/pandas-docs/stable/indexing.html#indexing-view-versus-copy>

/anaconda3/lib/python3.6/site-packages/ipykernel\_launcher.py:3: SettingWithCopyWarning:

A value is trying to be set on a copy of a slice from a DataFrame.  
Try using .loc[row\_indexer,col\_indexer] = value instead

See the caveats in the documentation: <http://pandas.pydata.org/pandas-docs/stable/indexing.html#indexing-view-versus-copy>

This is separate from the ipykernel package so we can avoid doing imports until

/anaconda3/lib/python3.6/site-packages/ipykernel\_launcher.py:4: SettingWithCopyWarning:

A value is trying to be set on a copy of a slice from a DataFrame.  
Try using .loc[row\_indexer,col\_indexer] = value instead

See the caveats in the documentation: <http://pandas.pydata.org/pandas-docs/stable/indexing.html#indexing-view-versus-copy>

after removing the cwd from sys.path.

/anaconda3/lib/python3.6/site-packages/ipykernel\_launcher.py:5: SettingWithCopyWarning:

A value is trying to be set on a copy of a slice from a DataFrame.  
Try using .loc[row\_indexer,col\_indexer] = value instead

See the caveats in the documentation: <http://pandas.pydata.org/pandas-docs/stable/indexing.html#indexing-view-versus-copy>

"""

In [71]:

```
# TODO: Formatting.
ADI['Date1'] = ADI['Date'].dt.date
FID['Date1'] = FID['Date'].dt.date
FIS['Date1'] = FIS['Date'].dt.date
GPN['Date1'] = GPN['Date'].dt.date
JUN['Date1'] = JUN['Date'].dt.date
```

```
/anaconda3/lib/python3.6/site-packages/ipykernel_launcher.py:1: Se
ttingWithCopyWarning:
A value is trying to be set on a copy of a slice from a DataFrame.
Try using .loc[row_indexer,col_indexer] = value instead
```

See the caveats in the documentation: <http://pandas.pydata.org/pandas-docs/stable/indexing.html#indexing-view-versus-copy>

```
"""Entry point for launching an IPython kernel.
/anaconda3/lib/python3.6/site-packages/ipykernel_launcher.py:2: Se
ttingWithCopyWarning:
A value is trying to be set on a copy of a slice from a DataFrame.
Try using .loc[row_indexer,col_indexer] = value instead
```

See the caveats in the documentation: <http://pandas.pydata.org/pandas-docs/stable/indexing.html#indexing-view-versus-copy>

```
/anaconda3/lib/python3.6/site-packages/ipykernel_launcher.py:3: Se
ttingWithCopyWarning:
A value is trying to be set on a copy of a slice from a DataFrame.
Try using .loc[row_indexer,col_indexer] = value instead
```

See the caveats in the documentation: <http://pandas.pydata.org/pandas-docs/stable/indexing.html#indexing-view-versus-copy>

This is separate from the ipykernel package so we can avoid doing imports until

```
/anaconda3/lib/python3.6/site-packages/ipykernel_launcher.py:4: Se
ttingWithCopyWarning:
A value is trying to be set on a copy of a slice from a DataFrame.
Try using .loc[row_indexer,col_indexer] = value instead
```

See the caveats in the documentation: <http://pandas.pydata.org/pandas-docs/stable/indexing.html#indexing-view-versus-copy>

```
after removing the cwd from sys.path.
/anaconda3/lib/python3.6/site-packages/ipykernel_launcher.py:5: Se
ttingWithCopyWarning:
A value is trying to be set on a copy of a slice from a DataFrame.
Try using .loc[row_indexer,col_indexer] = value instead
```

See the caveats in the documentation: <http://pandas.pydata.org/pandas-docs/stable/indexing.html#indexing-view-versus-copy>

```
"""
```

In [72]:

```
# TODO: Formatting.
ADI['Time1'] = ADI['Date'].dt.time
FID['Time1'] = FID['Date'].dt.time
FIS['Time1'] = FIS['Date'].dt.time
GPN['Time1'] = GPN['Date'].dt.time
JUN['Time1'] = JUN['Date'].dt.time
```

/anaconda3/lib/python3.6/site-packages/ipykernel\_launcher.py:1: SettingWithCopyWarning:

A value is trying to be set on a copy of a slice from a DataFrame.  
Try using .loc[row\_indexer,col\_indexer] = value instead

See the caveats in the documentation: <http://pandas.pydata.org/pandas-docs/stable/indexing.html#indexing-view-versus-copy>

"""Entry point for launching an IPython kernel.

/anaconda3/lib/python3.6/site-packages/ipykernel\_launcher.py:2: SettingWithCopyWarning:

A value is trying to be set on a copy of a slice from a DataFrame.  
Try using .loc[row\_indexer,col\_indexer] = value instead

See the caveats in the documentation: <http://pandas.pydata.org/pandas-docs/stable/indexing.html#indexing-view-versus-copy>

/anaconda3/lib/python3.6/site-packages/ipykernel\_launcher.py:3: SettingWithCopyWarning:

A value is trying to be set on a copy of a slice from a DataFrame.  
Try using .loc[row\_indexer,col\_indexer] = value instead

See the caveats in the documentation: <http://pandas.pydata.org/pandas-docs/stable/indexing.html#indexing-view-versus-copy>

This is separate from the ipykernel package so we can avoid doing imports until

/anaconda3/lib/python3.6/site-packages/ipykernel\_launcher.py:4: SettingWithCopyWarning:

A value is trying to be set on a copy of a slice from a DataFrame.  
Try using .loc[row\_indexer,col\_indexer] = value instead

See the caveats in the documentation: <http://pandas.pydata.org/pandas-docs/stable/indexing.html#indexing-view-versus-copy>

after removing the cwd from sys.path.

/anaconda3/lib/python3.6/site-packages/ipykernel\_launcher.py:5: SettingWithCopyWarning:

A value is trying to be set on a copy of a slice from a DataFrame.  
Try using .loc[row\_indexer,col\_indexer] = value instead

See the caveats in the documentation: <http://pandas.pydata.org/pandas-docs/stable/indexing.html#indexing-view-versus-copy>

"""

In [73]:

```
# TODO: Clean millisecond.
ADI['Time1'] = ADI['Time1'].apply(lambda x: x.replace(microsecond=0))
FID['Time1'] = FID['Time1'].apply(lambda x: x.replace(microsecond=0))
FIS['Time1'] = FIS['Time1'].apply(lambda x: x.replace(microsecond=0))
GPN['Time1'] = GPN['Time1'].apply(lambda x: x.replace(microsecond=0))
JUN['Time1'] = JUN['Time1'].apply(lambda x: x.replace(microsecond=0))
```

/anaconda3/lib/python3.6/site-packages/ipykernel\_launcher.py:1: SettingWithCopyWarning:  
A value is trying to be set on a copy of a slice from a DataFrame.  
Try using .loc[row\_indexer,col\_indexer] = value instead

See the caveats in the documentation: <http://pandas.pydata.org/pandas-docs/stable/indexing.html#indexing-view-versus-copy>

```
"""Entry point for launching an IPython kernel.
/anaconda3/lib/python3.6/site-packages/ipykernel_launcher.py:2: SettingWithCopyWarning:
A value is trying to be set on a copy of a slice from a DataFrame.
Try using .loc[row_indexer,col_indexer] = value instead
```

See the caveats in the documentation: <http://pandas.pydata.org/pandas-docs/stable/indexing.html#indexing-view-versus-copy>

```
/anaconda3/lib/python3.6/site-packages/ipykernel_launcher.py:3: SettingWithCopyWarning:
A value is trying to be set on a copy of a slice from a DataFrame.
Try using .loc[row_indexer,col_indexer] = value instead
```

See the caveats in the documentation: <http://pandas.pydata.org/pandas-docs/stable/indexing.html#indexing-view-versus-copy>

This is separate from the ipykernel package so we can avoid doing imports until

```
/anaconda3/lib/python3.6/site-packages/ipykernel_launcher.py:4: SettingWithCopyWarning:
A value is trying to be set on a copy of a slice from a DataFrame.
Try using .loc[row_indexer,col_indexer] = value instead
```

See the caveats in the documentation: <http://pandas.pydata.org/pandas-docs/stable/indexing.html#indexing-view-versus-copy>

```
after removing the cwd from sys.path.
/anaconda3/lib/python3.6/site-packages/ipykernel_launcher.py:5: SettingWithCopyWarning:
A value is trying to be set on a copy of a slice from a DataFrame.
Try using .loc[row_indexer,col_indexer] = value instead
```

See the caveats in the documentation: <http://pandas.pydata.org/pandas-docs/stable/indexing.html#indexing-view-versus-copy>

```
"""
```

In [74]:

```
# TODO: Clean second.
```

```
ADI['Time1'] = ADI['Time1'].apply(lambda x: x.replace(second=0))
FID['Time1'] = FID['Time1'].apply(lambda x: x.replace(second=0))
FIS['Time1'] = FIS['Time1'].apply(lambda x: x.replace(second=0))
GPN['Time1'] = GPN['Time1'].apply(lambda x: x.replace(second=0))
JUN['Time1'] = JUN['Time1'].apply(lambda x: x.replace(second=0))
```

/anaconda3/lib/python3.6/site-packages/ipykernel\_launcher.py:1: SettingWithCopyWarning:

A value is trying to be set on a copy of a slice from a DataFrame.  
Try using .loc[row\_indexer,col\_indexer] = value instead

See the caveats in the documentation: <http://pandas.pydata.org/pandas-docs/stable/indexing.html#indexing-view-versus-copy>

"""Entry point for launching an IPython kernel.

/anaconda3/lib/python3.6/site-packages/ipykernel\_launcher.py:2: SettingWithCopyWarning:

A value is trying to be set on a copy of a slice from a DataFrame.  
Try using .loc[row\_indexer,col\_indexer] = value instead

See the caveats in the documentation: <http://pandas.pydata.org/pandas-docs/stable/indexing.html#indexing-view-versus-copy>

/anaconda3/lib/python3.6/site-packages/ipykernel\_launcher.py:3: SettingWithCopyWarning:

A value is trying to be set on a copy of a slice from a DataFrame.  
Try using .loc[row\_indexer,col\_indexer] = value instead

See the caveats in the documentation: <http://pandas.pydata.org/pandas-docs/stable/indexing.html#indexing-view-versus-copy>

This is separate from the ipykernel package so we can avoid doing imports until

/anaconda3/lib/python3.6/site-packages/ipykernel\_launcher.py:4: SettingWithCopyWarning:

A value is trying to be set on a copy of a slice from a DataFrame.  
Try using .loc[row\_indexer,col\_indexer] = value instead

See the caveats in the documentation: <http://pandas.pydata.org/pandas-docs/stable/indexing.html#indexing-view-versus-copy>

after removing the cwd from sys.path.

/anaconda3/lib/python3.6/site-packages/ipykernel\_launcher.py:5: SettingWithCopyWarning:

A value is trying to be set on a copy of a slice from a DataFrame.  
Try using .loc[row\_indexer,col\_indexer] = value instead

See the caveats in the documentation: <http://pandas.pydata.org/pandas-docs/stable/indexing.html#indexing-view-versus-copy>

"""



In [75]:

```
# TODO: Create datetime.
ADI.loc[:, 'Datetime'] = pd.to_datetime(ADI.Date1.astype(str)+' '+ADI.Time1.
astype(str))
FID.loc[:, 'Datetime'] = pd.to_datetime(FID.Date1.astype(str)+' '+FID.Time1.
astype(str))
FIS.loc[:, 'Datetime'] = pd.to_datetime(FIS.Date1.astype(str)+' '+FIS.Time1.
astype(str))
GPN.loc[:, 'Datetime'] = pd.to_datetime(GPN.Date1.astype(str)+' '+GPN.Time1.
astype(str))
JUN.loc[:, 'Datetime'] = pd.to_datetime(JUN.Date1.astype(str)+' '+JUN.Time1.
astype(str))
```

/anaconda3/lib/python3.6/site-packages/pandas/core/indexing.py:357

: SettingWithCopyWarning:

A value is trying to be set on a copy of a slice from a DataFrame.

Try using `.loc[row_indexer,col_indexer] = value` instead

See the caveats in the documentation: <http://pandas.pydata.org/pandas-docs/stable/indexing.html#indexing-view-versus-copy>

```
self.obj[key] = _infer_fill_value(value)
```

/anaconda3/lib/python3.6/site-packages/pandas/core/indexing.py:537

: SettingWithCopyWarning:

A value is trying to be set on a copy of a slice from a DataFrame.

Try using `.loc[row_indexer,col_indexer] = value` instead

See the caveats in the documentation: <http://pandas.pydata.org/pandas-docs/stable/indexing.html#indexing-view-versus-copy>

```
self.obj[item] = s
```

In [76]:

```
# TODO: Add a column of value 1 to compute minutely messages.
ADI['ACT'] = 1
FID['ACT'] = 1
FIS['ACT'] = 1
GPN['ACT'] = 1
JUN['ACT'] = 1
```

```

/anaconda3/lib/python3.6/site-packages/ipykernel_launcher.py:1: Se
ttingWithCopyWarning:
A value is trying to be set on a copy of a slice from a DataFrame.
Try using .loc[row_indexer,col_indexer] = value instead

See the caveats in the documentation: http://pandas.pydata.org/pan
das-docs/stable/indexing.html#indexing-view-versus-copy
    """Entry point for launching an IPython kernel.
/anaconda3/lib/python3.6/site-packages/ipykernel_launcher.py:2: Se
ttingWithCopyWarning:
A value is trying to be set on a copy of a slice from a DataFrame.
Try using .loc[row_indexer,col_indexer] = value instead

See the caveats in the documentation: http://pandas.pydata.org/pan
das-docs/stable/indexing.html#indexing-view-versus-copy

/anaconda3/lib/python3.6/site-packages/ipykernel_launcher.py:3: Se
ttingWithCopyWarning:
A value is trying to be set on a copy of a slice from a DataFrame.
Try using .loc[row_indexer,col_indexer] = value instead

See the caveats in the documentation: http://pandas.pydata.org/pan
das-docs/stable/indexing.html#indexing-view-versus-copy
    This is separate from the ipykernel package so we can avoid doin
g imports until
/anaconda3/lib/python3.6/site-packages/ipykernel_launcher.py:4: Se
ttingWithCopyWarning:
A value is trying to be set on a copy of a slice from a DataFrame.
Try using .loc[row_indexer,col_indexer] = value instead

See the caveats in the documentation: http://pandas.pydata.org/pan
das-docs/stable/indexing.html#indexing-view-versus-copy
    after removing the cwd from sys.path.
/anaconda3/lib/python3.6/site-packages/ipykernel_launcher.py:5: Se
ttingWithCopyWarning:
A value is trying to be set on a copy of a slice from a DataFrame.
Try using .loc[row_indexer,col_indexer] = value instead

See the caveats in the documentation: http://pandas.pydata.org/pan
das-docs/stable/indexing.html#indexing-view-versus-copy
    """

```

In [77]:

```
# TODO: Minutely messages.
ADI['ACTIVITY'] = ADI.groupby(['Datetime'])['ACT'].apply(lambda x: x.cumsum())
FID['ACTIVITY'] = FID.groupby(['Datetime'])['ACT'].apply(lambda x: x.cumsum())
FIS['ACTIVITY'] = FIS.groupby(['Datetime'])['ACT'].apply(lambda x: x.cumsum())
GPN['ACTIVITY'] = GPN.groupby(['Datetime'])['ACT'].apply(lambda x: x.cumsum())
JUN['ACTIVITY'] = JUN.groupby(['Datetime'])['ACT'].apply(lambda x: x.cumsum())
```

/anaconda3/lib/python3.6/site-packages/ipykernel\_launcher.py:1: SettingWithCopyWarning:  
A value is trying to be set on a copy of a slice from a DataFrame.  
Try using .loc[row\_indexer,col\_indexer] = value instead

See the caveats in the documentation: <http://pandas.pydata.org/pandas-docs/stable/indexing.html#indexing-view-versus-copy>

```
"""Entry point for launching an IPython kernel.
/anaconda3/lib/python3.6/site-packages/ipykernel_launcher.py:2: SettingWithCopyWarning:
A value is trying to be set on a copy of a slice from a DataFrame.
Try using .loc[row_indexer,col_indexer] = value instead
```

See the caveats in the documentation: <http://pandas.pydata.org/pandas-docs/stable/indexing.html#indexing-view-versus-copy>

```
/anaconda3/lib/python3.6/site-packages/ipykernel_launcher.py:3: SettingWithCopyWarning:
A value is trying to be set on a copy of a slice from a DataFrame.
Try using .loc[row_indexer,col_indexer] = value instead
```

See the caveats in the documentation: <http://pandas.pydata.org/pandas-docs/stable/indexing.html#indexing-view-versus-copy>

This is separate from the ipykernel package so we can avoid doing imports until

```
/anaconda3/lib/python3.6/site-packages/ipykernel_launcher.py:4: SettingWithCopyWarning:
A value is trying to be set on a copy of a slice from a DataFrame.
Try using .loc[row_indexer,col_indexer] = value instead
```

See the caveats in the documentation: <http://pandas.pydata.org/pandas-docs/stable/indexing.html#indexing-view-versus-copy>

```
after removing the cwd from sys.path.
/anaconda3/lib/python3.6/site-packages/ipykernel_launcher.py:5: SettingWithCopyWarning:
A value is trying to be set on a copy of a slice from a DataFrame.
Try using .loc[row_indexer,col_indexer] = value instead
```

See the caveats in the documentation: <http://pandas.pydata.org/pandas-docs/stable/indexing.html#indexing-view-versus-copy>

```
"""
```

In [78]:

```
# TODO: Keep columns.
ADI = ADI[['Datetime', 'ACTIVITY']]
FID = FID[['Datetime', 'ACTIVITY']]
FIS = FIS[['Datetime', 'ACTIVITY']]
GPN = GPN[['Datetime', 'ACTIVITY']]
JUN = JUN[['Datetime', 'ACTIVITY']]
```

In [79]:

```
# TODO: Keep only last.
ADI = ADI.groupby('Datetime', as_index=False).last()
FID = FID.groupby('Datetime', as_index=False).last()
FIS = FIS.groupby('Datetime', as_index=False).last()
GPN = GPN.groupby('Datetime', as_index=False).last()
JUN = JUN.groupby('Datetime', as_index=False).last()
```

In [80]:

```
# TODO: Merge datasets.
Task8ADI = pd.merge(Task7ADI, ADI, on='Datetime', how='outer')
Task8FID = pd.merge(Task7FID, FID, on='Datetime', how='outer')
Task8FIS = pd.merge(Task7FIS, FIS, on='Datetime', how='outer')
Task8GPN = pd.merge(Task7GPN, GPN, on='Datetime', how='outer')
Task8JUN = pd.merge(Task7JUN, JUN, on='Datetime', how='outer')
```

In [81]:

```
# TODO: Delete missing.
Task8ADI = Task8ADI[np.isfinite(Task8ADI['Volumecum'])]
Task8FID = Task8FID[np.isfinite(Task8FID['Volumecum'])]
Task8FIS = Task8FIS[np.isfinite(Task8FIS['Volumecum'])]
Task8GPN = Task8GPN[np.isfinite(Task8GPN['Volumecum'])]
Task8JUN = Task8JUN[np.isfinite(Task8JUN['Volumecum'])]
```

In [83]:

```
# TODO: Compute HFT proxy.
Task8ADI['HFT'] = Task8ADI['ACTIVITY']/Task8ADI['Transcum']
Task8FID['HFT'] = Task8FID['ACTIVITY']/Task8FID['Transcum']
Task8FIS['HFT'] = Task8FIS['ACTIVITY']/Task8FIS['Transcum']
Task8GPN['HFT'] = Task8GPN['ACTIVITY']/Task8GPN['Transcum']
Task8JUN['HFT'] = Task8JUN['ACTIVITY']/Task8JUN['Transcum']
```

In [84]:

```
# TODO: SAVE HFT DATE FOR TASK4.
Task8ADI.to_csv('/Users/khaladdin/Desktop/Twitter Project/HFTADI.csv')
Task8FID.to_csv('/Users/khaladdin/Desktop/Twitter Project/HFTFID.csv')
Task8FIS.to_csv('/Users/khaladdin/Desktop/Twitter Project/HFTFIS.csv')
Task8GPN.to_csv('/Users/khaladdin/Desktop/Twitter Project/HFTGPN.csv')
Task8JUN.to_csv('/Users/khaladdin/Desktop/Twitter Project/HFTJUN.csv')
```

In [198]:

```
# TODO: Combine data.
TotalTASK5 = pd.concat([Task8ADI, Task8FID, Task8FIS, Task8GPN, Task8JUN])
```

In [200]:

```
# TODO: Set index for FIXED REGRESSION
TotalTASK5 = TotalTASK5.set_index(['RIC', 'Date'])
```

In [123]:

```
# TODO: Compute LAG values
TotalTASK5['LOGVOLUME'] = np.log(TotalTASK5['Volumecum'])
TotalTASK5['LOGsize'] = np.log(TotalTASK5['size'])
TotalTASK5['LOGHFT'] = np.log(TotalTASK5['HFT'])
```

```
/anaconda3/lib/python3.6/site-packages/ipykernel_launcher.py:1: Ru
ntimeWarning: divide by zero encountered in log
    """Entry point for launching an IPython kernel.
/anaconda3/lib/python3.6/site-packages/ipykernel_launcher.py:2: Ru
ntimeWarning: divide by zero encountered in log
```

In [136]:

```
# TODO: Cleaning
TotalTask6 = TotalTASK5.replace([np.inf, -np.inf], np.nan)
```

In [137]:

```
# TODO: Add intercept
TotalTask6['intercept'] = 1
```

In [139]:

```
# TODO: Cleaning
TotalTask6 = TotalTask6[np.isfinite(TotalTask6['LOGHFT'])]
```

In [ ]:

```
# TODO: FIXED REGRESSION
from linearmodels import PanelOLS
mod = PanelOLS(TotalTask6.Level, TotalTask6[['intercept', 'dummy', 'effective', '
LOGVOLUME', 'LOGsize', 'vol', 'LOGHFT']], entity_effects=True)
res = mod.fit(cov_type='clustered', cluster_entity=True)
res
```

In [169]:

```
# TODO: Winsorised
from scipy.stats import mstats
def WinsorizeStats(TotalTask6):
    out = mstats.winsorize(TotalTask6, limits=[0.05, 0.05])
    return out
```

In [280]:

```
# TODO: Winsorised
TotalTask7 = TotalTask6[['RIC', 'Level', 'effective', 'LOGVOLUME', 'LOGsize', 'vol',
'LOGHFT']].apply(WinsorizeStats, axis=0)
```

In [283]:

```
# TODO: Dealing with Missing
TotalTask7['vol'].fillna(0, inplace=True)
```

In [284]:

```
# TODO: Keep some columns.
dummy = TotalTask6[['intercept', 'dummy']]
```

In [285]:

```
# TODO: Create new column. I need it since, I should merge dummy and TotalTask
5 datasets based on any column.
dummy['C']=dummy.reset_index().index
TotalTask7['C']=TotalTask7.reset_index().index
```

/anaconda3/lib/python3.6/site-packages/ipykernel\_launcher.py:1: SettingWithCopyWarning:

A value is trying to be set on a copy of a slice from a DataFrame.  
Try using .loc[row\_indexer,col\_indexer] = value instead

See the caveats in the documentation: <http://pandas.pydata.org/pandas-docs/stable/indexing.html#indexing-view-versus-copy>

"""Entry point for launching an IPython kernel.

In [286]:

```
# TODO: Merge two datasets and set RIC (stock) and DATE as index for FIXED eff
ect regression
TotalTask8 = TotalTask7.reset_index().merge(dummy, on='C', how='outer').set_in
dex(['RIC', 'Date'])
```

In [ ]:

```
# TODO: FIXED effect regression
from linearmodels import PanelOLS
mod = PanelOLS(TotalTask8.Level, TotalTask8[['intercept', 'dummy', 'effective', '
LOGVOLUME', 'LOGsize', 'vol', 'LOGHFT']], entity_effects=True)
res = mod.fit(cov_type='clustered', cluster_entity=True)
res
```