个股与指数的回归分析（python） - CSDN博客 https://blog.csdn.net/csqazwsxedc/article/details/51336322

# 一 个股与指数的回归分析 ## 1.1 数据加载 加载分析所需的Python库

import statsmodels.api as sm

import statsmodels.formula.api as smf

import statsmodels.graphics.api as smg

import patsy

import matplotlib.pyplot as plt

import numpy as np

import pandas as pd

from pandas import Series,DataFrame

from scipy import stats

import seaborn as sns

import datetime

start = datetime.datetime(2015,1,1)

end = datetime.datetime(2015,12,31)

#获取”推特”2018年股价数据，记为twtr； 获取”谷歌”公司2018年股价数据，记为google。

**from pandas\_datareader import data, web**

**import fix\_yahoo\_finance as yf**

**yf.pdr\_override()**

**twtr =web. DataReader("TWTR","yahoo",start,end)**

**google =web.DataReader("GOOG","yahoo",start,end)**

twtr.head()

google.head()

## 1.2数据探索性分析

close\_twtr = twtr["Close"]

close\_goggle = google["Close"]

#得到推特2018年各交易日收盘价的简单统计结果，如下所示。共有129个推特的股价数据，指数的平均值为32.6725，最小值为22.16，最大值为46.75。

close\_twtr.describe()

# count 129.000000

mean 32.672558

std 6.415918

min 22.160000

25% 28.450001

50% 31.910000

75% 34.849998

max 46.759998

Name: Close, dtype: float64

#得到谷歌公司2018年各交易日收盘价的简单统计结果，如下所示。共有129个谷歌的股价数据，股价的平均值1093，最小值为1001，最大值为1175。

close\_google.describe()

count 129.000000

mean 1093.039144

std 47.666273

min 1001.520020

25% 1054.790039

50% 1097.569946

75% 1129.790039

max 1175.839966

Name: Close, dtype: float64

>>> 观察推特和谷歌公司的股价波动图，如下所示，可以看到，推特股价波动整体上升趋势，谷歌公司股价波动较推特大。

fig,ax = plt.subplots(nrows=1,ncols=2,figsize=(15,6))

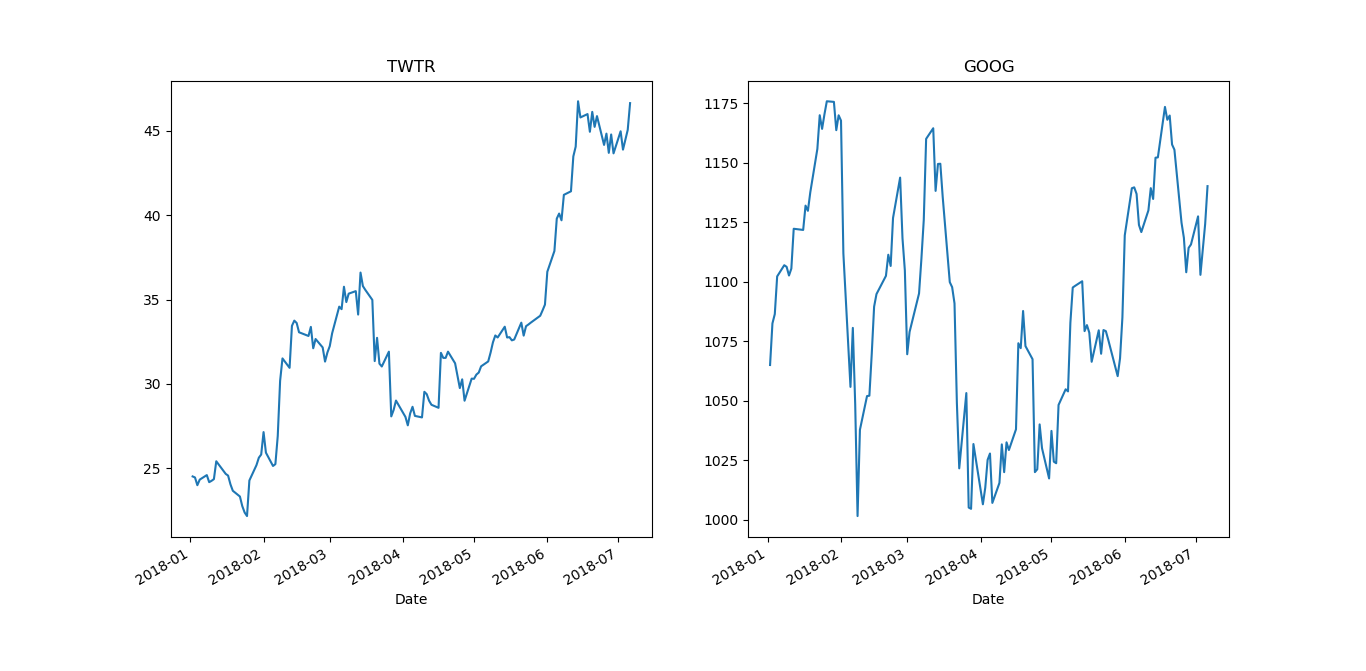
close\_twtr.plot(ax=ax[0])

ax[0].set\_title("TWTR")

close\_google.plot(ax=ax[1])

ax[1].set\_title("GOOG")

plt.show()



#两者数据进行对比

stock = pd.merge(twtr,google,left\_index = True, right\_index = True)

stock = stock[["Close\_x","Close\_y"]]

stock.columns = ["TWTR","GOOG"]

stock.head()

#根据股价得到推特和谷歌公司的日收益率序列，如下所示。

daily\_return = (stock.diff()/stock.shift(periods = 1)).dropna()

daily\_return.head()

#观察日收益率序列的简单统计值，如下所示。推特日收益率平均值为0.005608，最小值为-0.120338，最大值为0.121516.谷歌公司股价的平均值为0.000681，最小值为-0.050454，最大值数据为0.036205值。

daily\_return.describe()

#观察是否异常值数据daily\_return[daily\_return["GOOG"] > 某个数]经分析，该日股价数据异常的原因主要是

#画出推特和谷歌公司日收益率波动图

fig,ax = plt.subplots(nrows=1,ncols=2,figsize=(15,6))

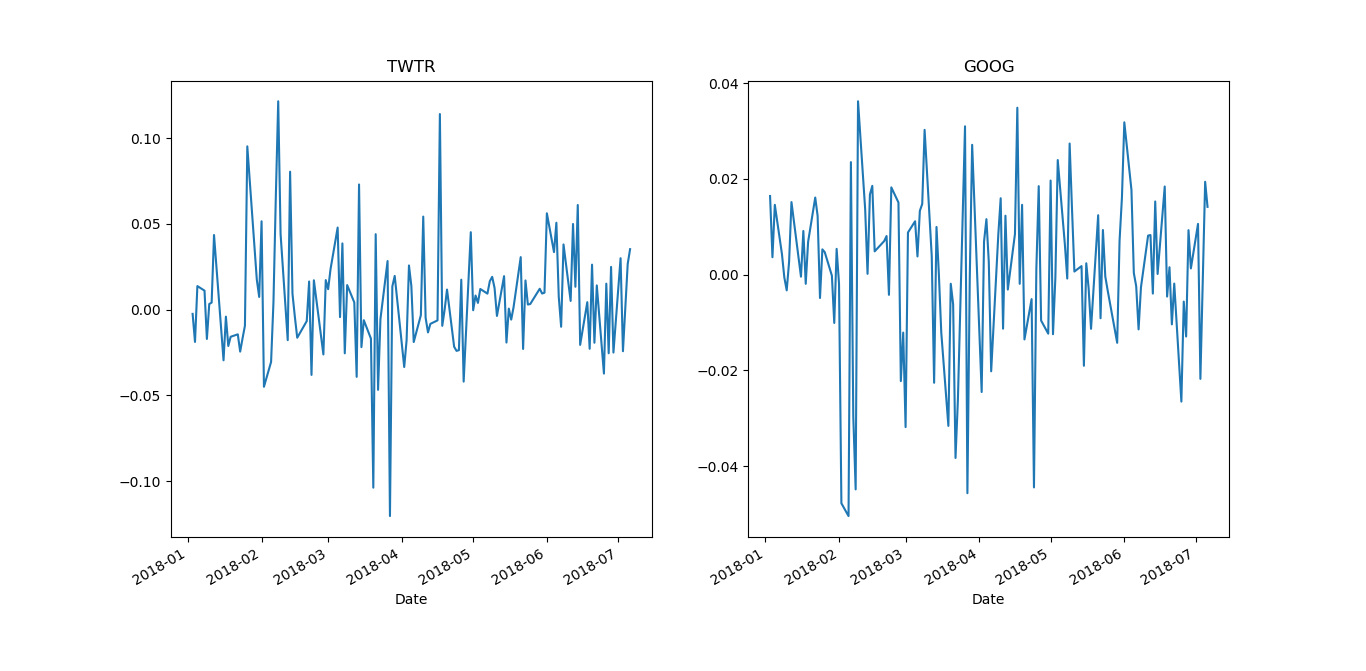
daily\_return["TWTR"].plot(ax=ax[0])

ax[0].set\_title("TWTR")

daily\_return["GOOG"].plot(ax=ax[1])

ax[1].set\_title("GOOG")

plt.show()



#画出推特和谷歌公司日收益率直方图和密度图，如下所示，可以发现，总体上，推特和谷歌公司日收益率服从正态分布。相对而言，谷歌公司的日收益率较推特偏低。

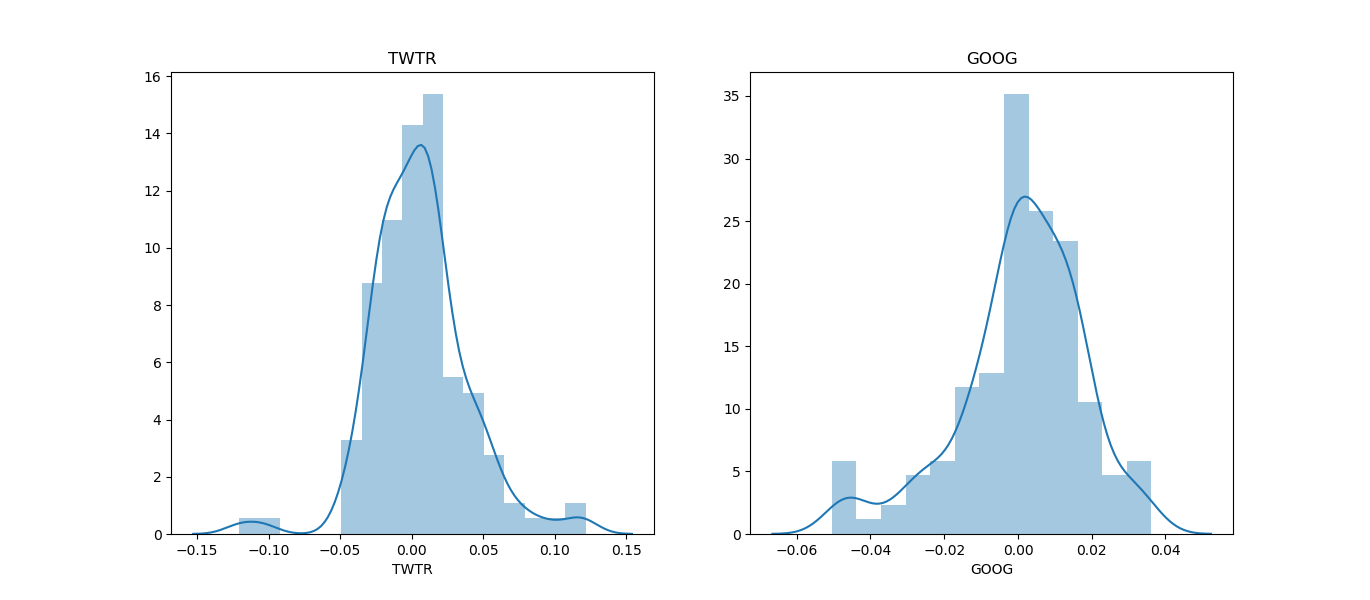
fig,ax = plt.subplots(nrows=1,ncols=2,figsize=(15,6))

sns.distplot(daily\_return["TWTR"],ax=ax[0])

ax[0].set\_title("TWTR")

sns.distplot(daily\_return["GOOG"],ax=ax[1])

ax[1].set\_title("GOOG")

plt.show()

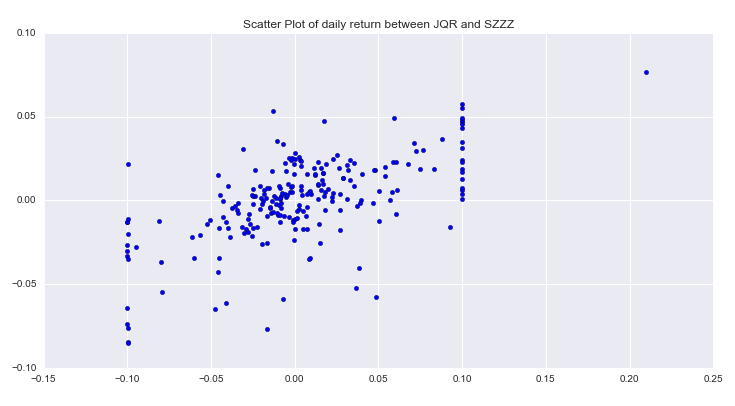
#画出推特和谷歌公司股价日收益率散点图，如下所示。

fig,ax = plt.subplots(nrows=1,ncols=1,figsize=(12,6))

plt.scatter(daily\_return["GOOG"],daily\_return["TWTR"])

plt.title("Scatter Plot of daily return between GOOG and TWTR")

plt.show()



#散点图表明，推特和谷歌公司的股价可能存在线性的正相关关系。

1.3回归分析

import statsmodels.api as sm

#加入截距项。

daily\_return["intercept"]=1.0

#以谷歌为自变量，推特为因变量，对谷歌和推特进行回归分析。得到回归结果，如下所示。

model = sm.OLS(daily\_return["GOOG"],daily\_return[["TWTR","intercept"]])

results = model.fit()

results.summary()

#一元最小二乘法回归结果表明，谷歌公司的股票日收益率与推特日收益率之间存在显著的正相关关系。其中，可决系数为 ，表明推特日收益率变量对谷歌日收益率变量有较 （强or弱）的解释力，模型拟合结果较（好or坏），F统计量和Omnibus统计量的P值都接近于（ ），自变量的作用（显著or微弱）。t统计量的P值接近于（ ），表明推特变量（显著or 微弱）。自变量系数为（ ），表明谷歌公司股票的日收益率波动比推特（大or 小），该个股的风险更（大or 小），可能获得的收益和损失也更大or 小）。平均推特日收益率波动（ ），个股日收益率波动（ ）。Durbin-Waston检验的值为（ ），表明收益率数据（or不）存在序列相关性。Jarque-Bera的P值接近于（），表明日收益率数据服从（or 不）正态分布。