**Summary**

The goal in this workshop is to implement a trading strategy with google trends information, the strategy is based from the paper Quantifying Trading Behavior in Financial Markets Using Google Trends. Our data corresponds to the Google trend information from 2005 to 2011 of the word “debt” and the Down jones index from yahoo finance.

The following libraries were used in the code.

import pandas as pd

import numpy as np

from datetime import datetime, timedelta

import matplotlib.pyplot as plt

import yfinance as yf

One of the drawbacks of the initial data is about the frequency of Google trends information, unfortunately the trend report only contains monthly information and the Down Jones information have a daily frequency, in this exercise the objective is to merge those databases in a weekly frequency table to calculate the signals of short or long investment.

1. Download Dow jones data from 2005 to 2011

The data from yahoo finance starts in 2005-01-01 and finishes in 2011-12-31, the start time is 1 week before 2005-01-01 allowing a gat for handling the weekly data frequency.

Stardate=datetime(2005, 1, 1)

Enddate=datetime(2011, 12, 31)

tck\_DJ\_data = yf.download('^DJI', Stardate-timedelta(weeks=1),Enddate)

Close\_DJ\_data=pd.DataFrame(tck\_DJ\_data['Close'])

The daily data from Yahoo finance have some missing days because they are holidays, so it necessary to add those records and perform a forward fill.

Close\_DJ\_data = (Close\_DJ\_data.set\_index('Date')

      .reindex(pd.date\_range(Stardate-timedelta(weeks=1), Enddate, freq='D'))

      .rename\_axis(['Date'])

      .fillna(method='ffill')

      .dropna()

      .reset\_index())

1. Load the monthly google trends data for term “debt”

Google\_trends=pd.read\_csv('debt\_google\_trend.csv')

Google\_trends['Mes']=Google\_trends['Mes'].astype('datetime64[ns]')

For Google trends monthly data base was transform to have all the days between 2005 and 2011 and filled with forward fill.

Google\_trends = (Google\_trends.set\_index('Mes')

      .reindex(pd.date\_range(Stardate-timedelta(weeks=1), Enddate, freq='D'))

      .rename\_axis(['Mes'])

      .fillna(method='ffill')

      .dropna()

      .reset\_index())

1. Combine the Down Jones Close price data base with the google trends database and set date frequency as weekly starting every Monday.

The function used was merge from pandas with an inner join.

Cnsl\_db\_Gtr=pd.merge(Close\_DJ\_data,Google\_trends,how='inner',left\_on='Date',right\_on='Mes')

Cnsl\_db\_Gtr=Cnsl\_db\_Gtr.loc[Cnsl\_db\_Gtr['Date']>=Stardate+timedelta(days=2)]

Once both databases with a daily frequency in dates were merged the frequency was updated to weekly.

Cnsl\_db\_Gtr2 = (Cnsl\_db\_Gtr.set\_index('Date')

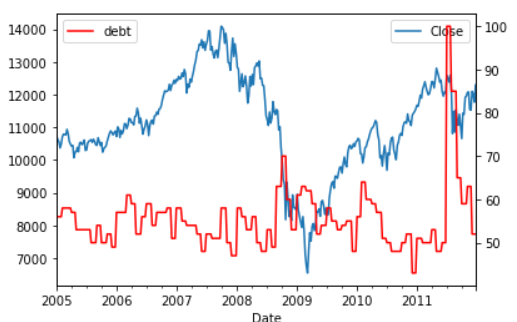
      .reindex(pd.date\_range(Stardate+timedelta(days=2), Enddate, freq='W-MON'))

      .rename\_axis(['Date'])

      .dropna()

      .reset\_index())

1. Google trends plot



1. Strategy configuration

If the current week Google trend search value is higher than moving average for previous 3 weeks select short on the other hand if the current week Google trend search value is lesser than the moving average for previous 3 weeks select long.

try:

    del StrategyDataframe

    StrategyDataframe=Cnsl\_db\_Gtr2.copy()

except:

    StrategyDataframe=Cnsl\_db\_Gtr2.copy()

StrategyDataframe['MovAvg']=StrategyDataframe['debt'].rolling(3).mean()

StrategyDataframe['MovAvg']=StrategyDataframe['MovAvg'].shift(1)

1. Create signal column in dataframe for trading strategy.

StrategyDataframe['Signal']=np.sign(StrategyDataframe['MovAvg']-StrategyDataframe['debt'])

StrategyDataframe['Signal']=StrategyDataframe['Signal'].shift(1)

StrategyDataframe['Signal']=StrategyDataframe['Signal'].replace(to\_replace=0, method='ffill')

1. Calculate weekly returns for buy and hold and the strategy.

StrategyDataframe['DJRet']=StrategyDataframe['Close']/StrategyDataframe['Close'].shift(1)-1

StrategyDataframe['SRet']=StrategyDataframe['Signal']\*StrategyDataframe['DJRet']

1. Calculate cumulative returns

StrategyDataframe['DJCumRet']=(1+StrategyDataframe['DJRet'] ).cumprod()

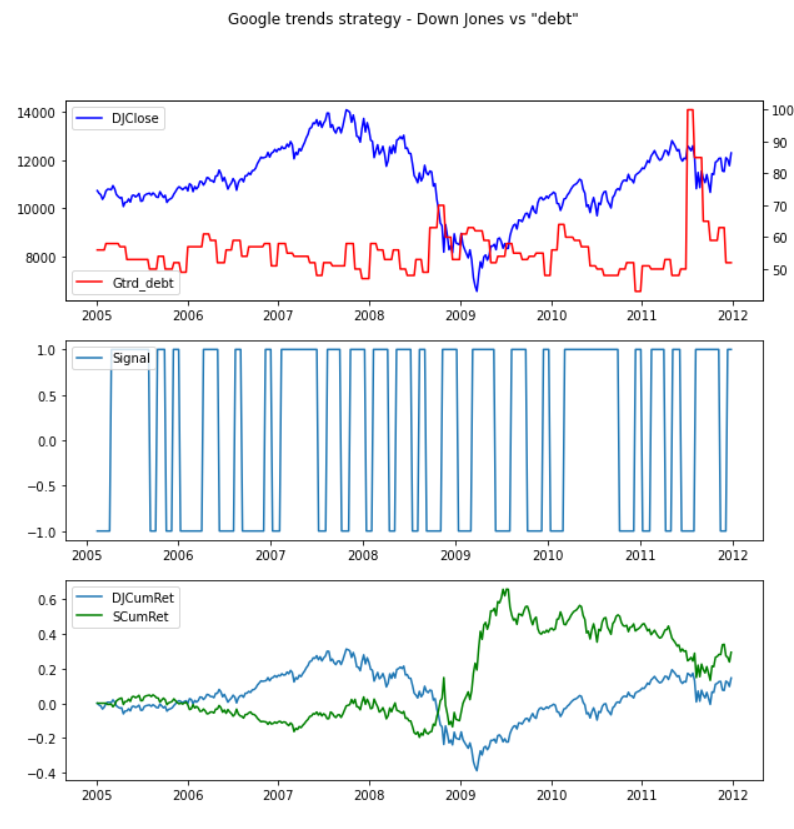
StrategyDataframe['SCumRet']=(1+StrategyDataframe['SRet'] ).cumprod()

StrategyDataframe['DJCumRet']=StrategyDataframe['DJCumRet'].fillna(1)

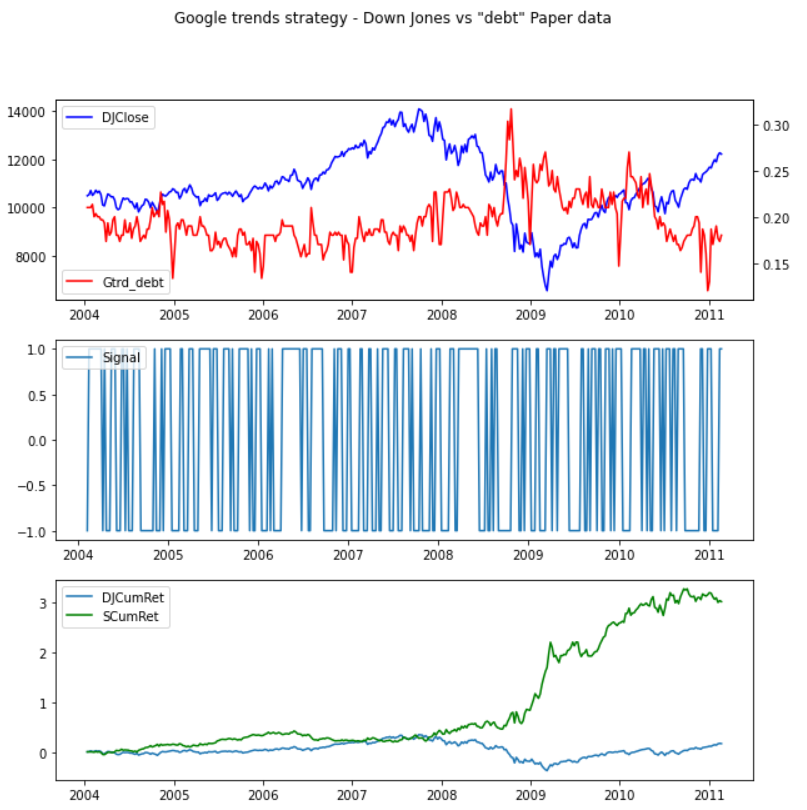
StrategyDataframe['SCumRet']=StrategyDataframe['SCumRet'].fillna(1)

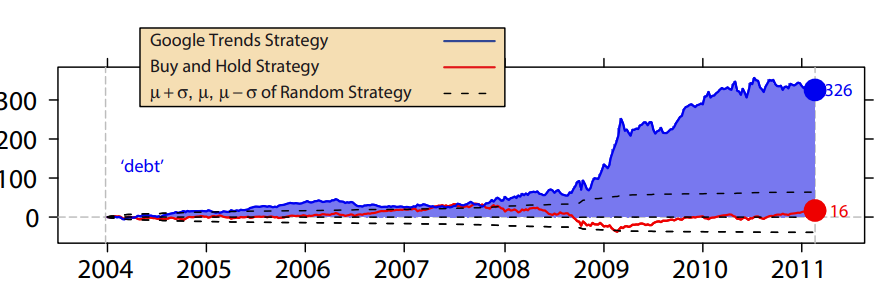
StrategyDataframe['DJCumRet']=StrategyDataframe['DJCumRet']-1

StrategyDataframe['SCumRet']=StrategyDataframe['SCumRet']-1



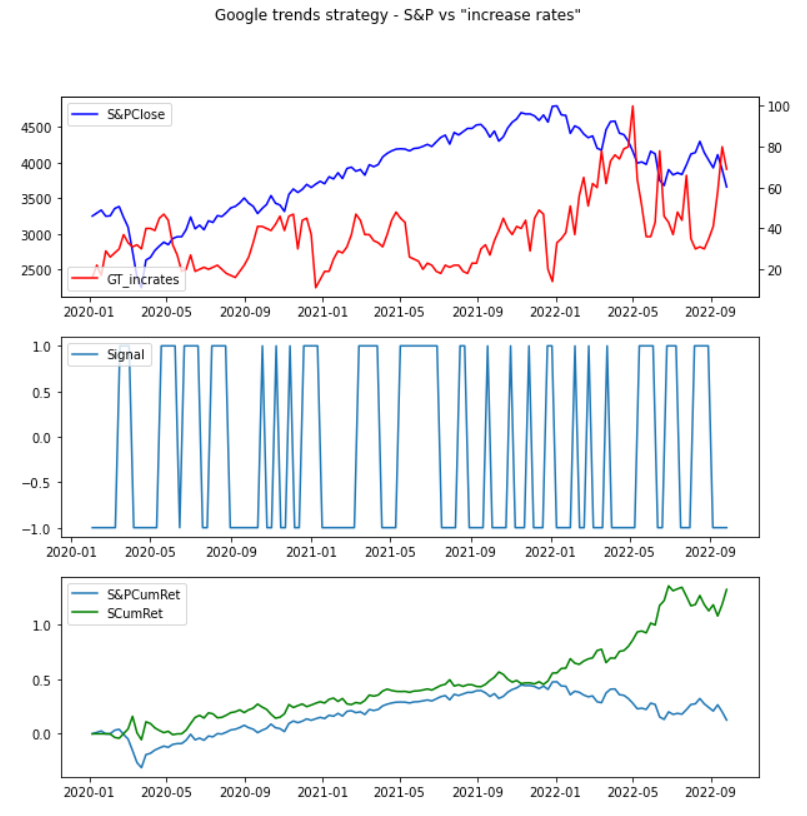
By executing the same trading algorithm with the paper data, the cumulative return obtained in 2011-02-01 is 3, this figure is similar to the exercise performed in the paper by by Tobias Preis, Helen Susannah Moat and H. Eugene Stanley.





1. Repeating the strategy for S&P500 using the term “increase rates”

The period of time selected stars in 2020-01-01 and ends in 2022-10-01, the data is evaluated in a weekly strategy with same rules for signals, when there is an increase in search versus the average of 3 previous weeks go short if there is a decrease versus the average of 3 previous weeks go long.



1. Conclusions

Using the Google trends data provided the trading strategy performed better than buy and hold right after the 2008 crisis.

The success for the paper strategy comes from the input data, they used the Google trends information but with some transformation that allow the model get a better perception of the market which translates in a higher return.

The trading strategy was used for S&P500 following the Google trend term “increase rates” did a great performance even though the period of time selected starts before the COVID pandemic. It very interesting to find that the strategy was always getting a hinger return.

This exercise is very interesting for understanding how the people trends are so related with the financial markets.

To implement this model in a real-life investment strategy is necessary to normalize the historical data with new entries. The Google trends figures are in a range of 0 to 100, that scale vary depending on the period of time consulted. Also it’s very important to consider the impact of the transactions costs in the strategy performance.