Google trends strategy
ALGORITHMIC TRADING
Hugo E. Ramirez
Maestria en Finanzas Cuantitativas – Universidad del Rosario

In this workshop we are testing a strategy based on signals constructed by the trends in google searches. And we are based on the paper *Quantifying Trading Behavior in Financial Markets Using Google Trends* by Tobias Preis, Helen Susannah Moat and H. Eugene Stanley.

Follow the next steps:

1. As in the document we start by downloading the data for the Dow Jones Industrial Average at range of dates (for the document from 2005 to 2011) and we work only with the close price.

2. We use the trends in searching the word "debt", which can be found in the document `debt_google_trend.csv` with the monthly data for the trend results according to some google measure. You could also download the data from `https://trends.google.com/trends/`. or even use an API or web scrapping technique

* If you want to compare our results with those in the paper you could use the file `PreisMoatStanley2013.dat`. The following code may be useful in this case.

```
1 paper = pd.read_csv('PreisMoatStanley2013.dat',delimiter=' ',parse_dates=[0,1,100,101])
2 df_gt = pd.DataFrame({'GoogleWE': paper['Google End Date'], 'debt': paper['debt'].astype(np
      .float64), 'DJIADate': paper['DJIA Date'], 'DJIAClose': paper['DJIA Closing Price'].
      astype(np.float64)})
```

3. Note that the fetched historic data has daily frequency and the google trends data is monthly frequency. We are going for middle ground and use weekly data, and we need to drop data if there is too much and keep Monday's data (DJIA historic data) and use forward filling in the need of more information (google trends).
   Merge the data in a single `DataFrame`. For this you need to convert the string monthly data of the csv file to `datetime` format. Moreover if the dates do not coincide you may need to shift one or more days.

4. Plot the Google trends.

5. The strategy consists of:
   - If there is a higher search volume at the end of a Google Trends week than in the previous three-week average of the search volume, then go short.
   - If there is a decline, we will go long the following week

   so you need to calculate the rolling mean for three weeks.

6. Create the signals in a new column of the `DataFrame`. In this case: Decrease in volume means Buy (signal 1) and increase in volume means Sell (signal −1).

7. Calculate the weekly returns for the asset and calculate the returns of the strategy by using a direct product. For example, if the return for week t_i is $r_{t_i} = 0.2324$ and the signal was sell then the return of the strategy is $sr_{t_i} = -0.2324$, on the other hand if the signal is buy the return for the strategy is $sr_{t_i} = r_{t_i} = 0.2324$.

8. Calculate the cumulative returns and compare with the cumulative returns of buy and hold.

9. Repeat this example but you choose the trend word(s), the asset(s) and the period of time, does your intuition improve the strategy?

10. Draw conclusions

** In the paper by Preis, Moat and Stanley, the prove to have cumulative return of 326% by 2011 compared with a 16% of buy and hold, but I guess our results are not as good as theirs. what could be happening there? to notice you have to repeat the exercise with their information, in `PreisMoatStanley2013.dat`.

Write a **report** where you explain what you did and how you did it (code snippets and graphs), and what conclusions you may derive. Upload a single file of no more than 4 pages with letter no smaller than $10pt$ named *HW2_yourname.pdf* to the e-aulas in a date no later than 03/10/2022.