

Market Sentiment & ETF Analysis

財金四 王聖揚

中文三 曹佳穎

工管一 孫利東



*We built a model that generated
36% return between 2015 – 2018*

In other words,

“發大財嘍” - 孫利東 2019

Backtest Results



Initial Portfolio Value: \$100,000
Final Portfolio Value: \$136,285



How we got these results

Research Topics

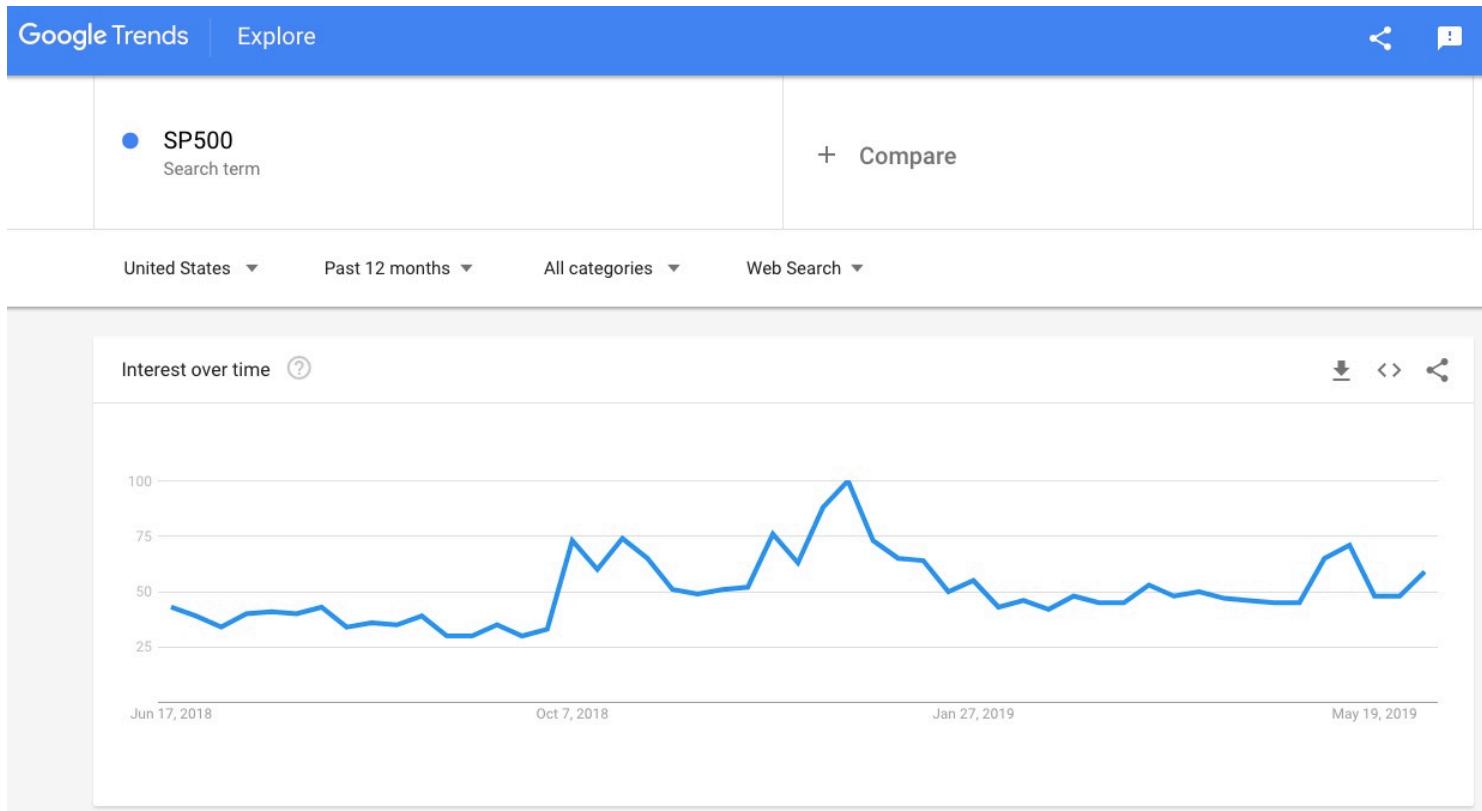
- 新聞是否為預測股價合理指標
- 網路搜尋流量與成交量是否成正比
- 市場情緒對各產業影響程度

We included these sectors in our analysis:

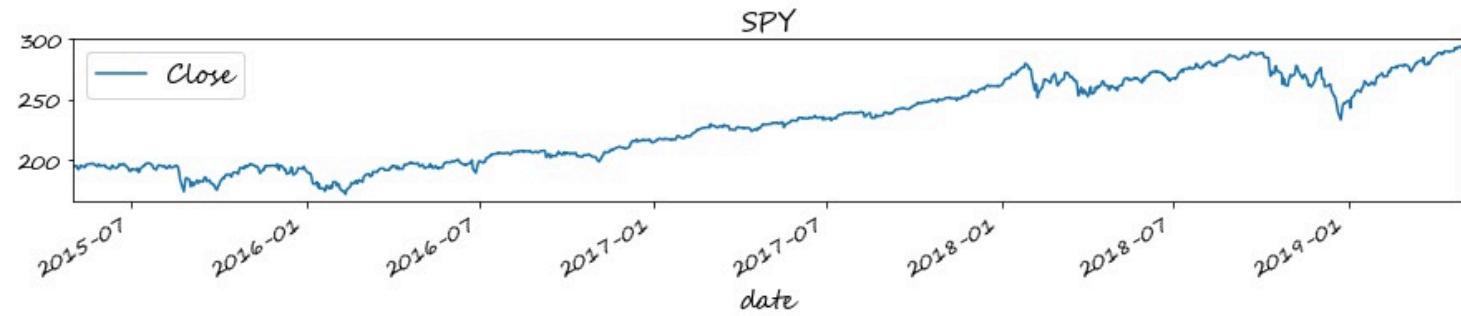
Overall Market	Financial	Technology	Industrial	Energy	Medical & Healthcare
SPY	XLF	XLK	XLI	XLE	XLV

Google Search Counts & Trading Volume

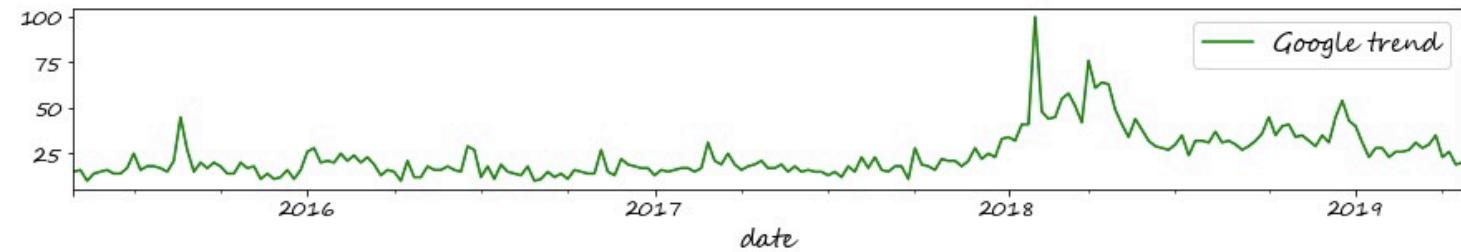
Search Frequency – Google Trends



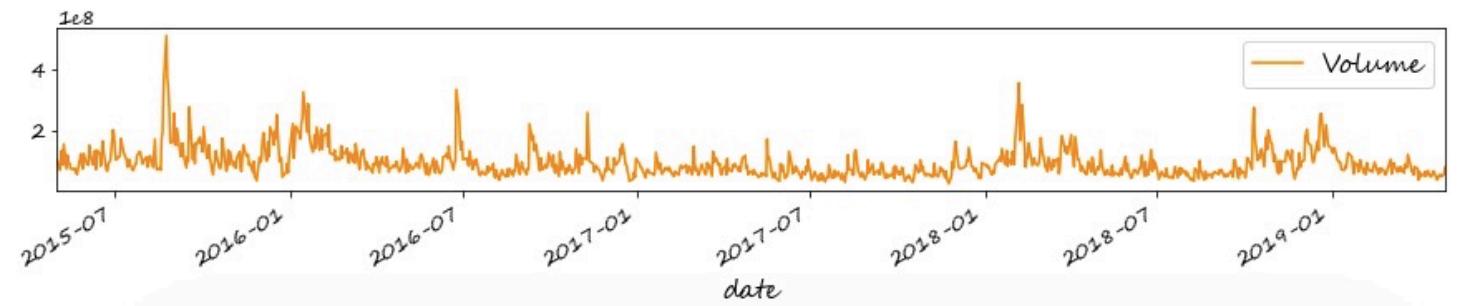
SP500 Closing Price



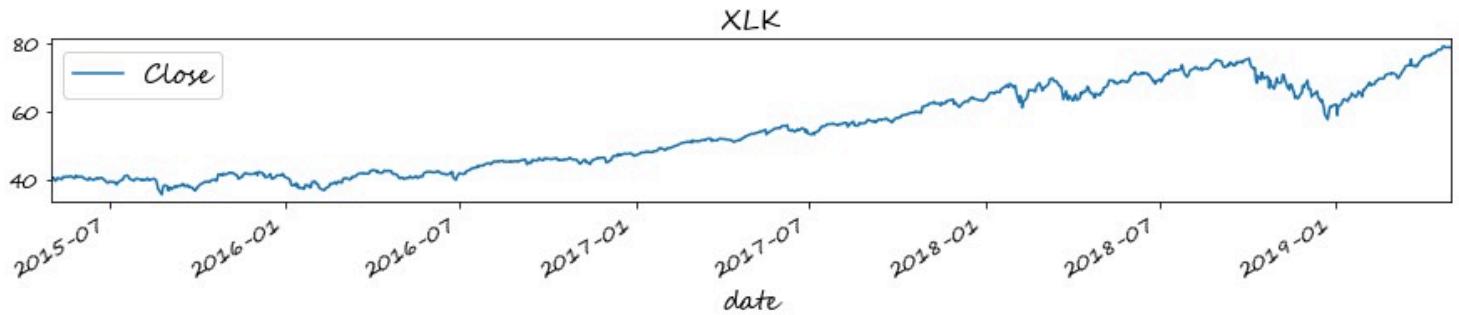
Search Counts



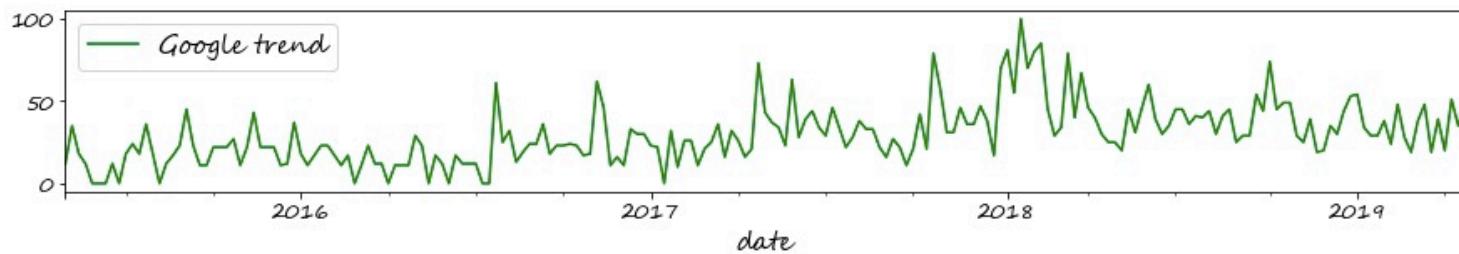
Daily Trading Volume



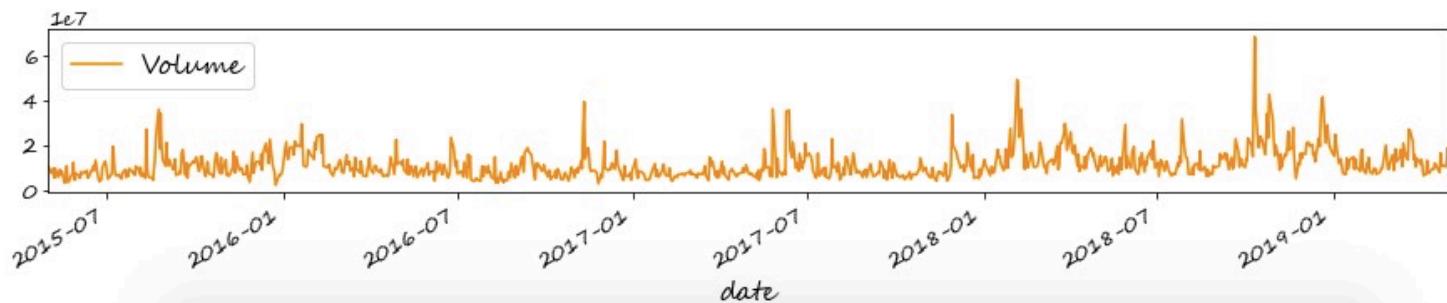
XLK (Tech) Closing Price



Search Counts



Daily Trading Volume

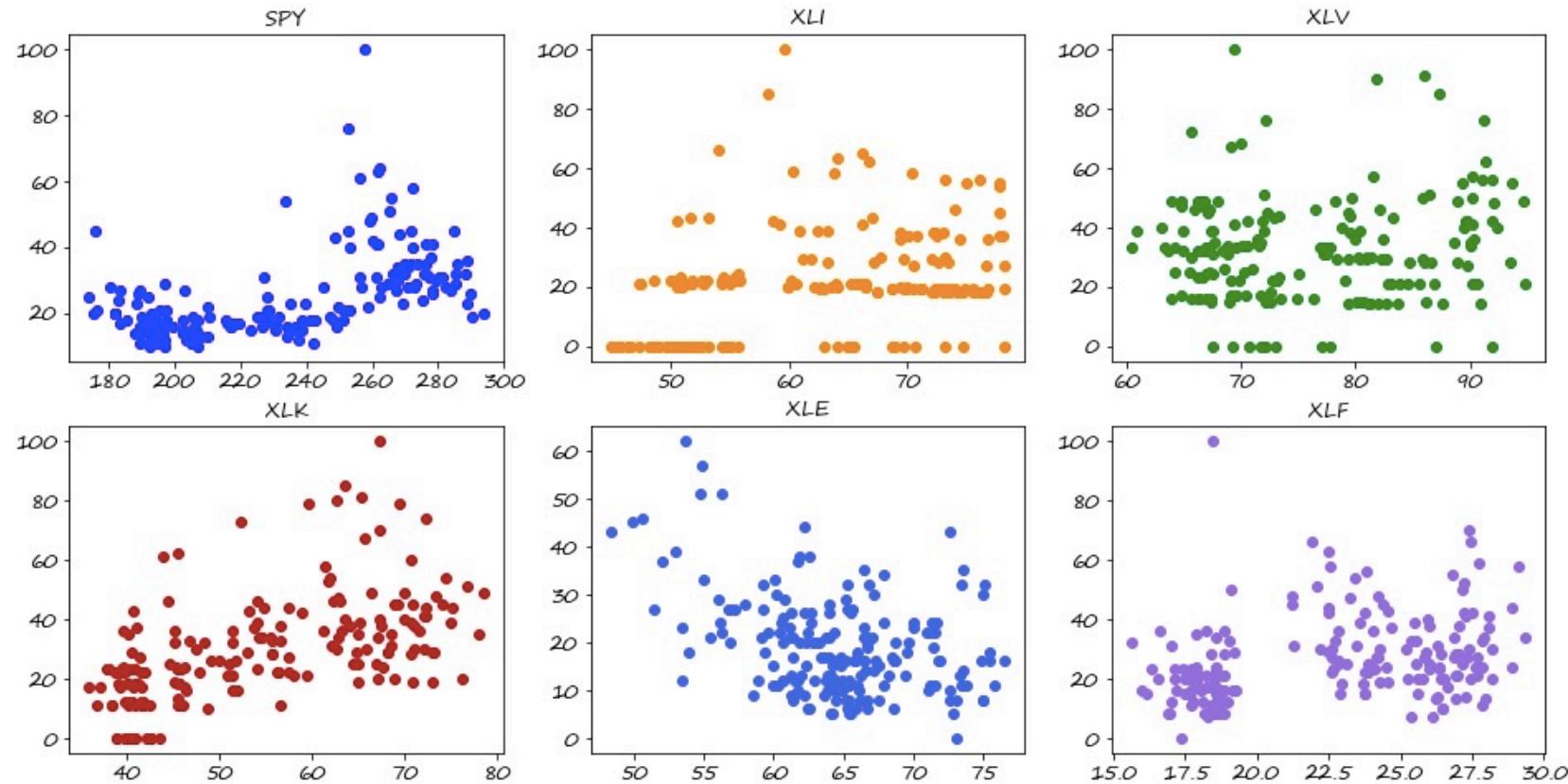


Tech seems to be a sector that has an obvious relationship between its price and search data

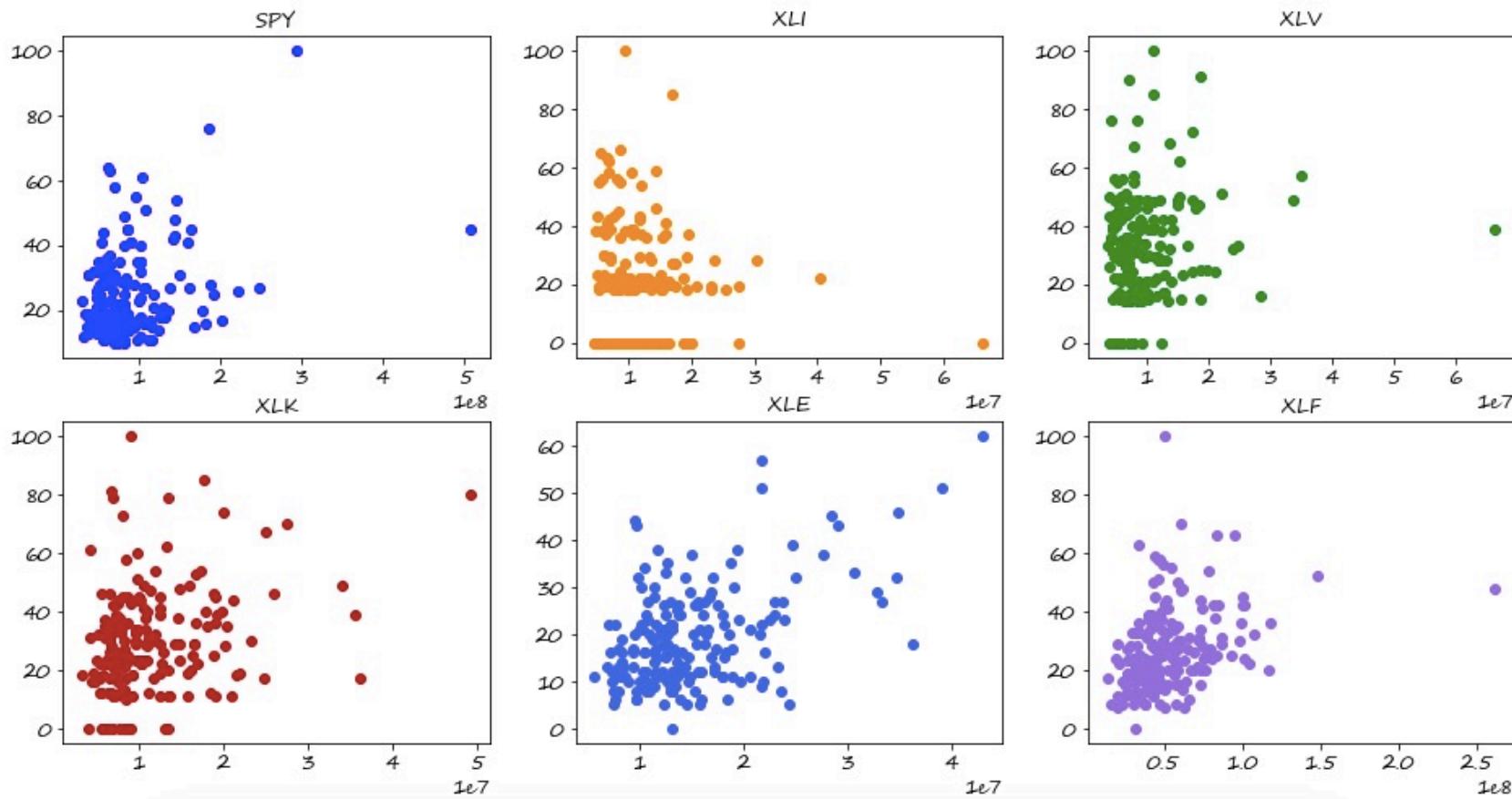
Observations

- Whenever google trend has a drastic change, close prices either surge or plunge
- Google trend data and volume data show similar pattern

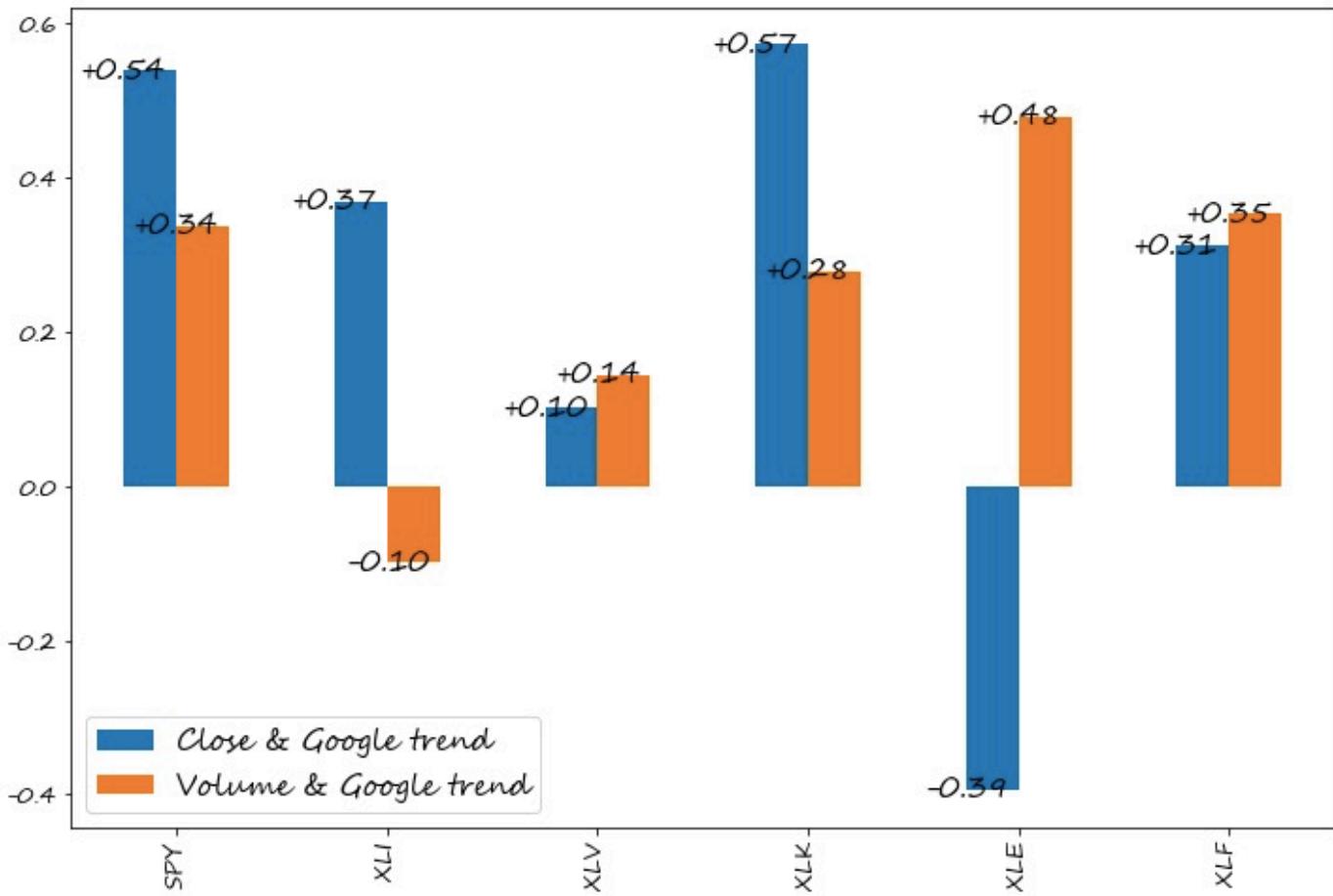
Search volume & price shows no correlation



Search volume & trading volume shows positive correlation in most sectors



Correlation



Overall Market Industrial Medical & Healthcare Technology Energy Financial

SPY	XLI	XLV	XLK	XLE	XLF
-----	-----	-----	-----	-----	-----

Market News

Sentiment Model

News Articles

CNBC MARKETS BUSINESS INVESTING TECH POLITICS CNBC TV

MARKETS | WORLD MARKETS

All ETF Daily News Articles

Will growth stocks continue to outpace value stocks?



From Todd Shriber: Growth stocks and the related exchange traded funds took their lumps in May, but year to date, the **iShares Russell 1000 Growth ETF** (NYSE: **IWF**) is higher by 18.40 percent, an advantage of 530 basis points over the **iShares Russell 1000 Value ETF** (NYSE: **IWD**).

NYSE:VUG © June 11, 2019 11:45am

Could there be a new marijuana ETF on the way?



From Todd Shriber: Global X, one of the largest purveyors of thematic exchange traded funds, is looking to add that lineup with a cannabis ETF.

REUTERS Business Markets World Politics Tech Breakingviews Wealth Life

REGIONAL MARKETS NEWS

EUROPE & MIDDLE EAST ASIA PACIFIC

Consumer inflation boosted by services

Inflation rose solidly for a second straight month in May, boosted by a surge in the cost of hotel and restaurant service fees.

U.S. bond yields stalls for now

U.S. bond yields added to a recent rebound on Tuesday as tech shares stalled from earlier in the day.

Treasury yields continue to rise as global trade concerns

U.S. Treasury yields added to a recent rebound on Tuesday as fears of a trade war between the U.S. and China paused and investors looked toward避險 assets.

European stocks close higher despite renewed trade tariff threats

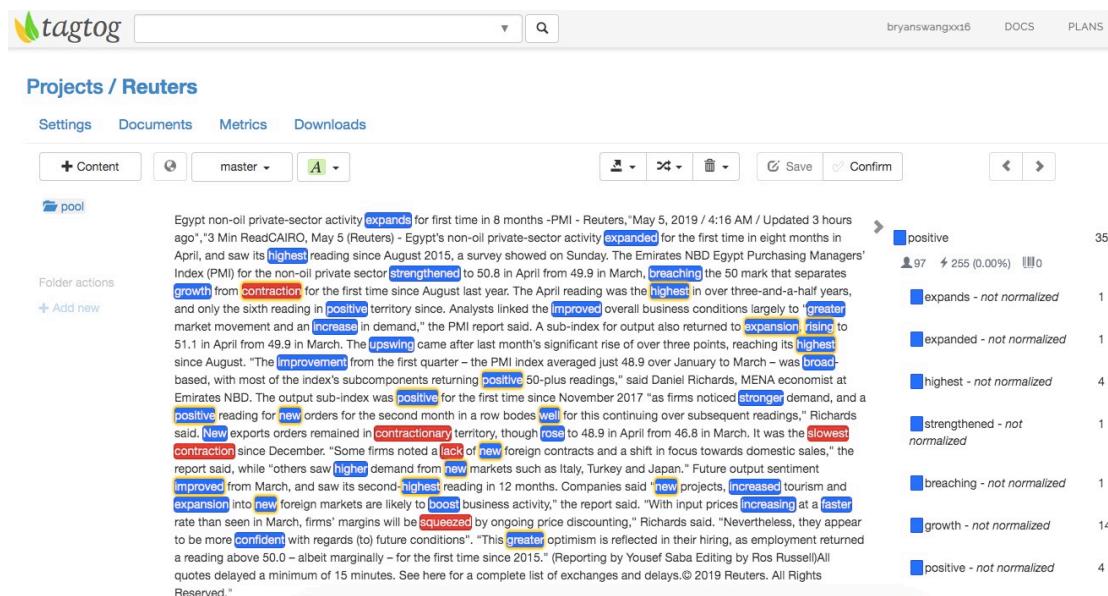
European stocks traded higher Tuesday after U.S. President Donald Trump told CNBC that immediate tariffs would be imposed on up to \$300 billion of Chinese goods if President Xi does not agree to a deal at this month's G-20 meeting.

Asian markets continue rally on trade, Fed hopes

Asian markets continued to rally on Tuesday, with the Nikkei 225 index rising 1.1% to a record closing level.

Wordlist

- Loughran McDonald word list
- Tagtog (from 100+ Reuters/CNBC news articles)



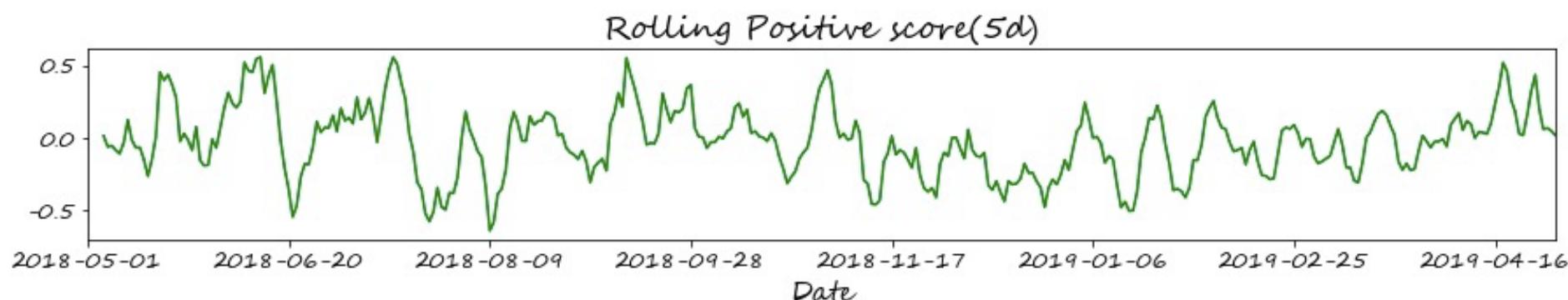
Naive Bayes Classifier

```
for i in range(len(word)):
    if is_number(word[i]) or word[i]=="$" or word[i]=="&":
        word="the"
        break
    #if word.isdigit():
    #print(word)
    #continue
#print(type(word))
#print(word)

if word not in stop_words:
    #print(word)
    classResult = classifier.classify( word_feats(word))
    #print(word)
    if classResult == 'neg':
        #print("neg",word)
        neg = neg + 1
    if classResult == 'pos':
        #print("pos",word)
        pos = pos + 1
    #if classResult== "neu":
    #print("neu",word)
print('Positive: ' + str(float(pos)/len(words)))
print('Negative: ' + str(float(neg)/len(words)))
return float(pos)/len(words), float(neg)/len(words)
#print(nltk.classify.accuracy(classifier,sentence))
```

Assigning positive and negative score to each day

	Unnamed: 0	Date	Content	positive	negative	positive_std	negative_std
2040	2040	2019-05-28	1 Min ReadLONDON, May 28 (Reuters) - British f...	0.029436	0.019865	0.724439	-0.785940
2041	2041	2019-05-29	3 Min ReadTUCSON, Ariz. (Reuters) - An Arizona...	0.027447	0.022797	0.283517	-0.197503
2042	2042	2019-05-30	2 Min Read(Reuters) - Juul Labs Inc is consider...	0.028865	0.021533	0.597983	-0.451094
2043	2043	2019-05-31	1 Min ReadBOGOTA, May 31 (Reuters) - Colombia'...	0.026845	0.021770	0.150212	-0.403583
2044	2044	2019-06-01	3 Min ReadMERCUREA-CIUC, Romania (Reuters) - ...	0.028560	0.022213	0.530316	-0.314619



Prices

Financial	XLF
Technology	XLK
Industrial	XLI
Energy	XLE
Medical & Healthcare	XLV
Retail	RTH



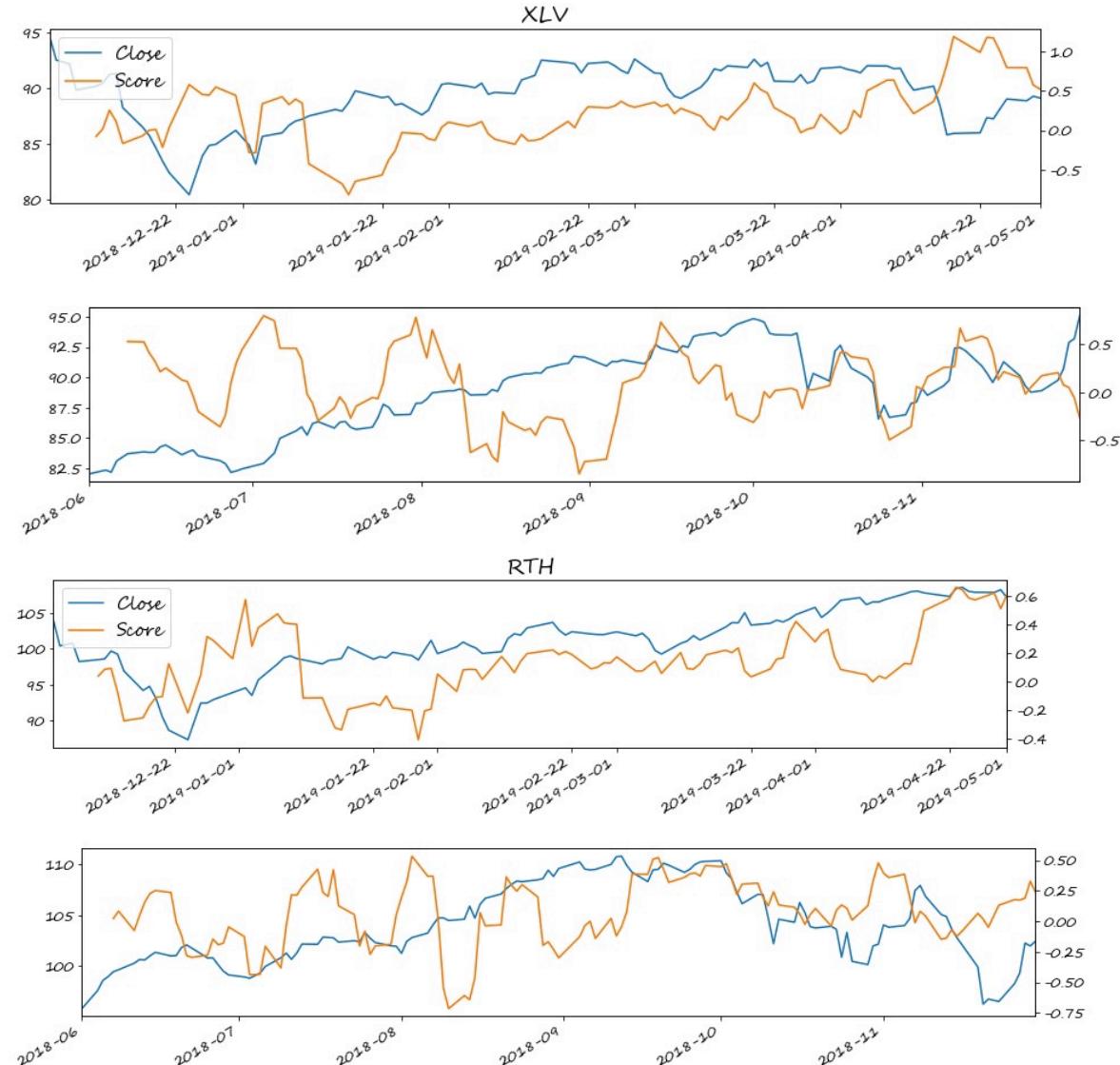
Prices

Financial	XLF
Technology	XLK
Industrial	XLI
Energy	XLE
Medical & Healthcare	XLV
Retail	RTH



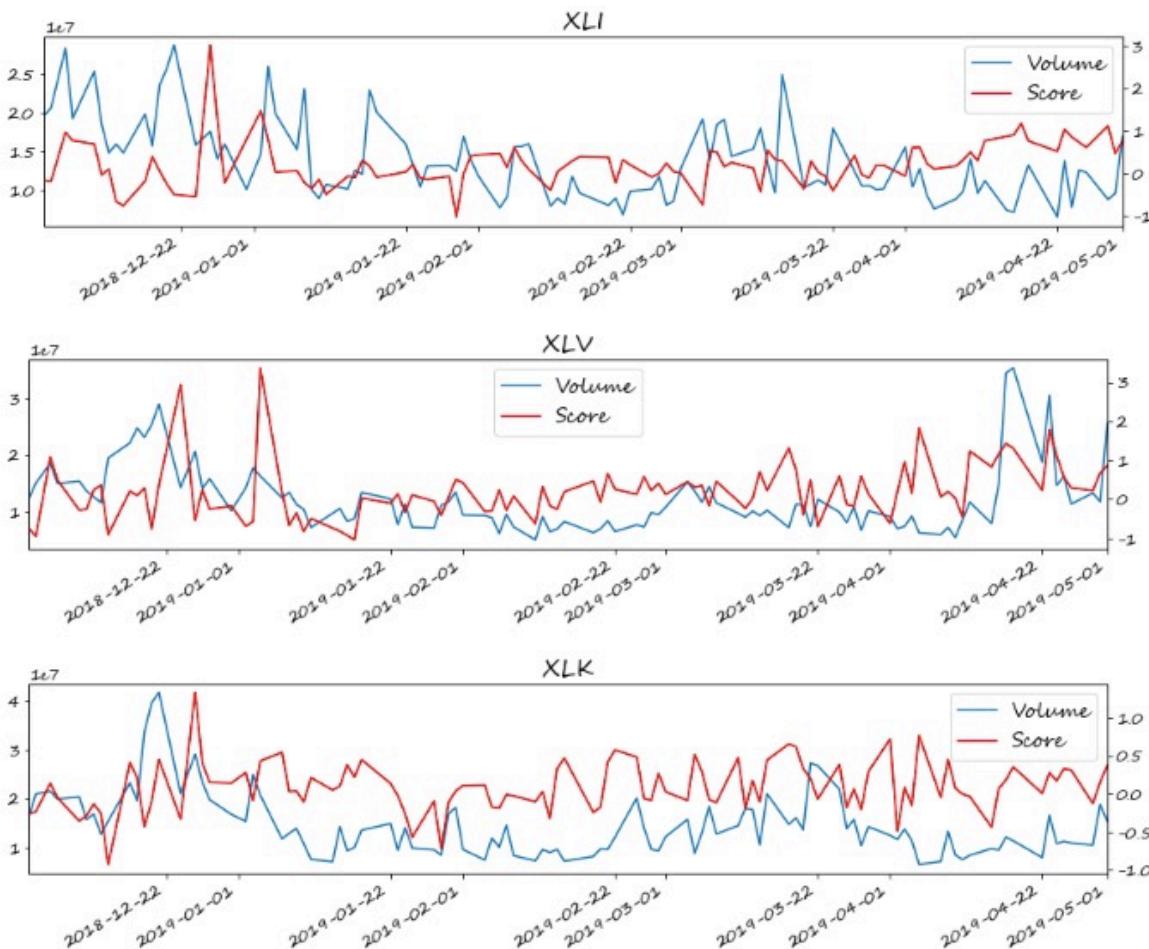
Prices

Financial	XLF
Technology	XLK
Industrial	XLI
Energy	XLE
Medical & Healthcare	XLV
Retail	RTH



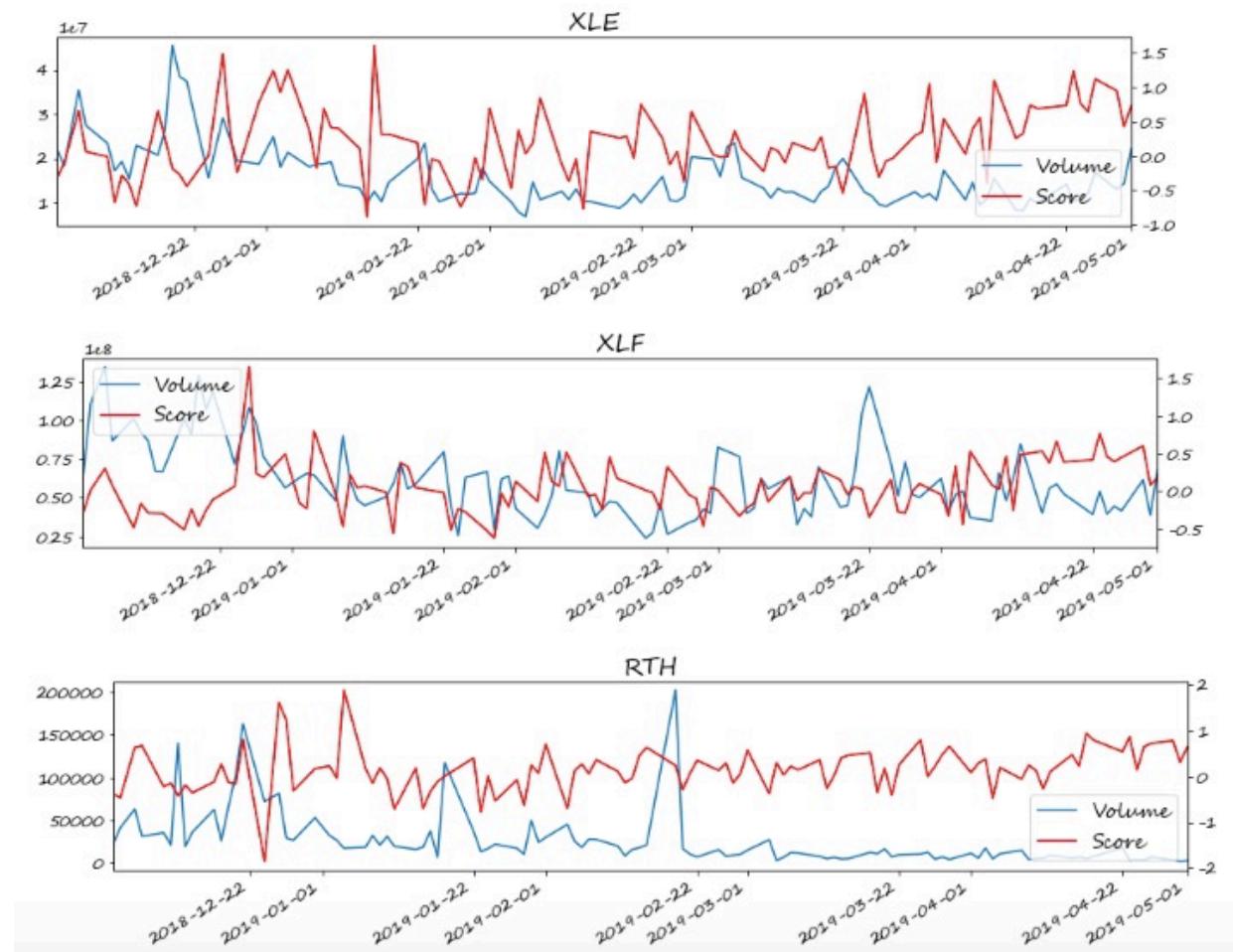
Volume

Financial	XLF
Technology	XLK
Industrial	XLI
Energy	XLE
Medical & Healthcare	XLV
Retail	RTH

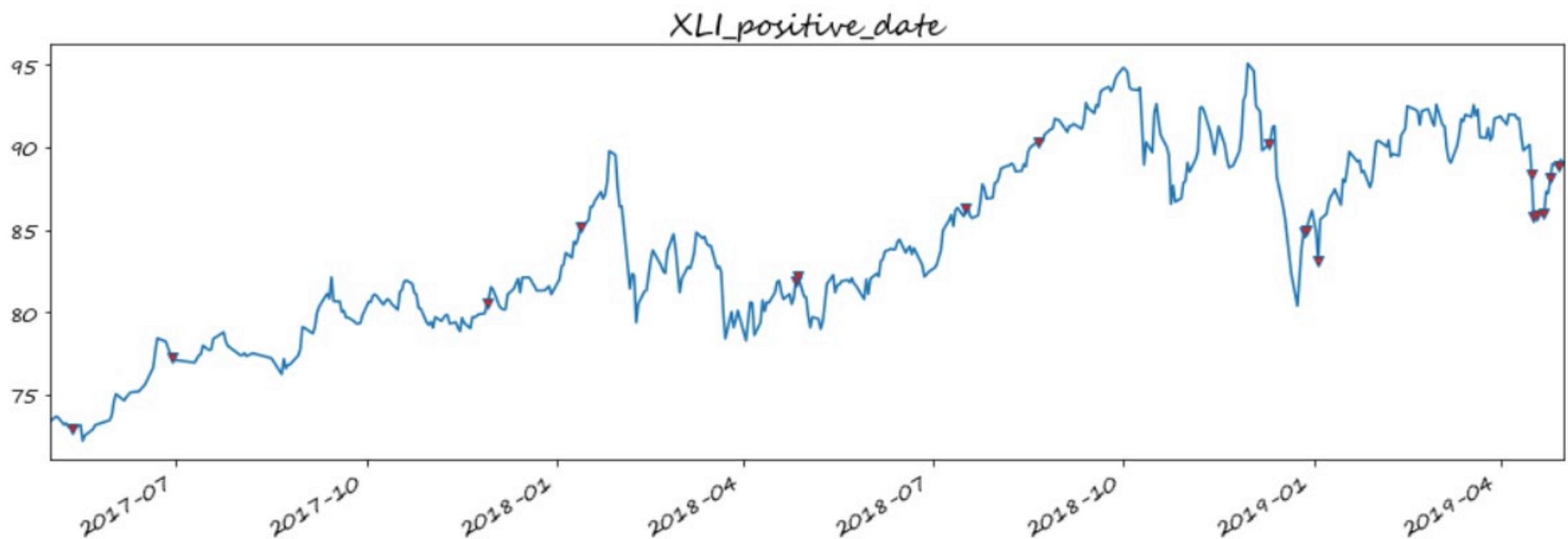


Volume

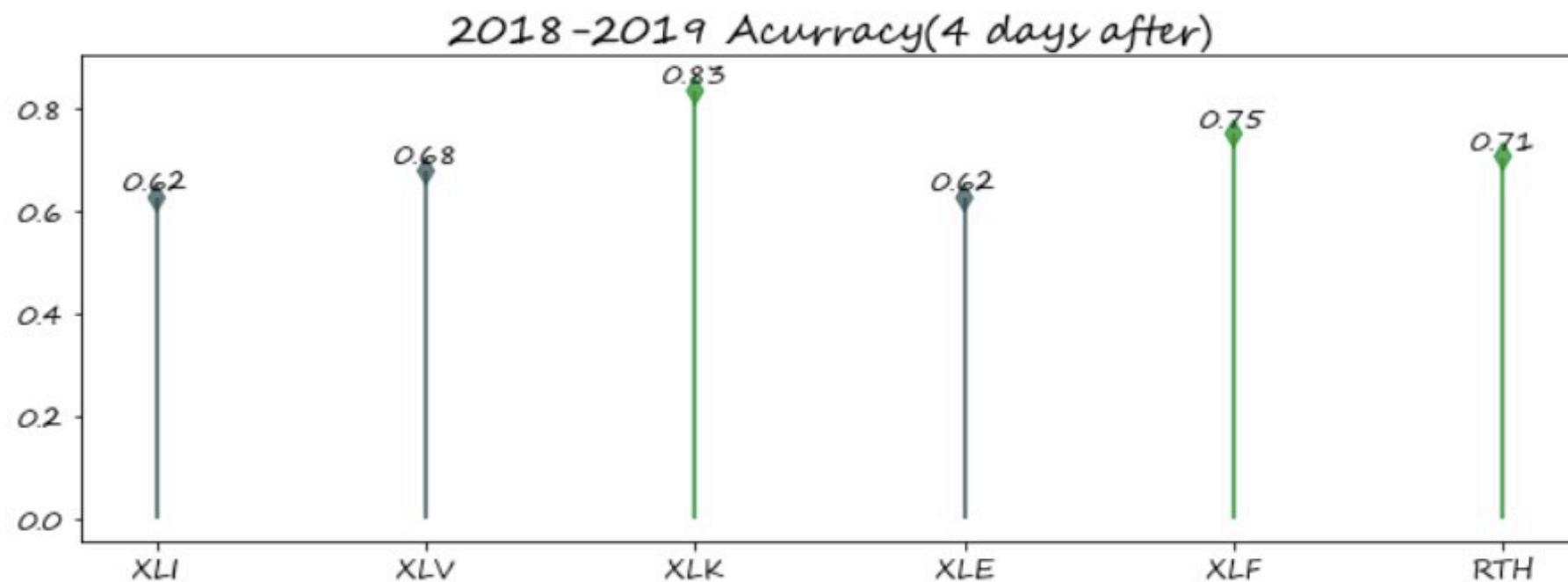
Financial	XLF
Technology	XLK
Industrial	XLI
Energy	XLE
Medical & Healthcare	XLV
Retail	RTH

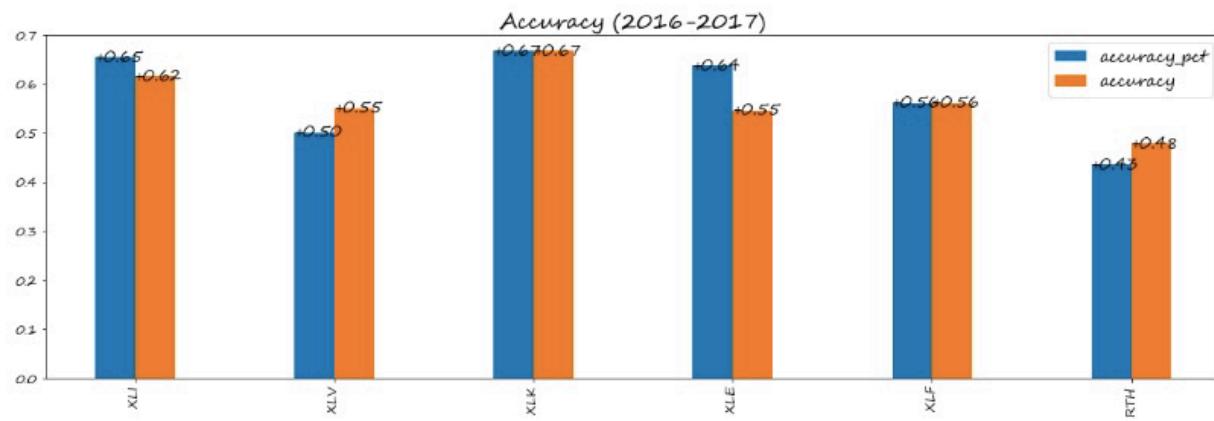
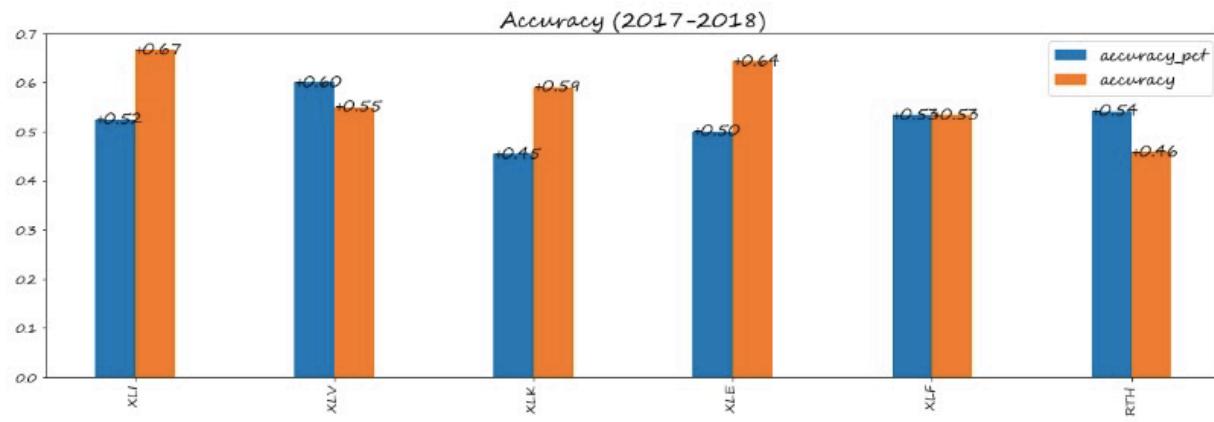


Marking days where positive score is higher than 0.8



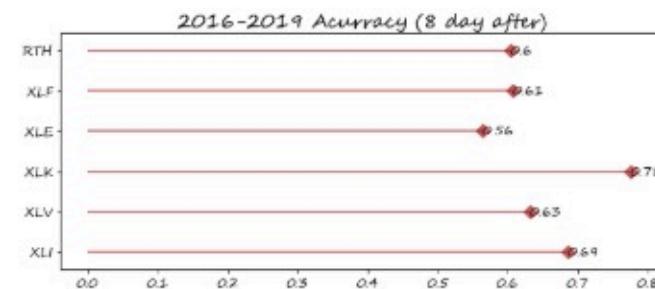
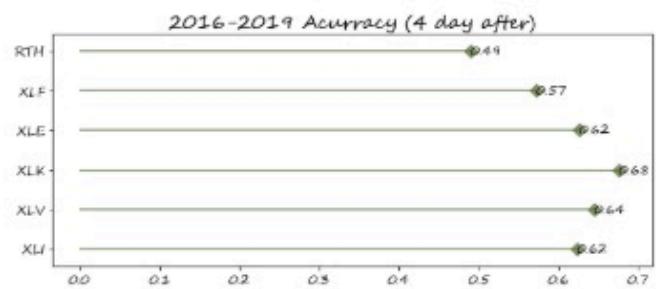
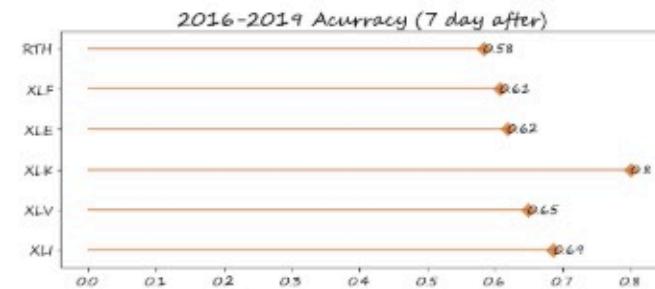
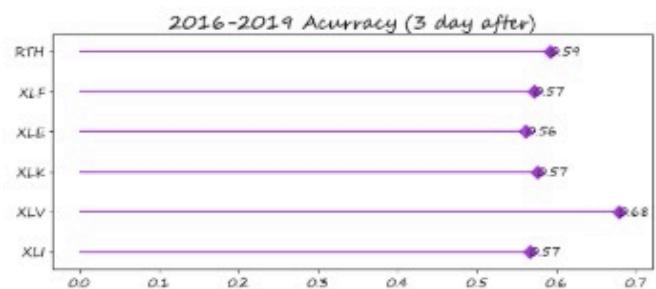
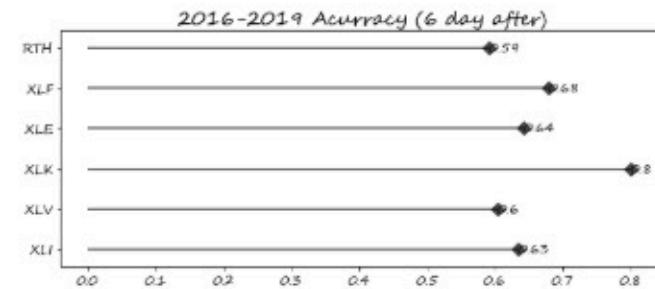
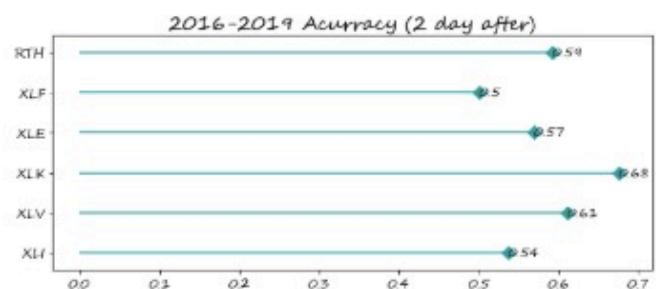
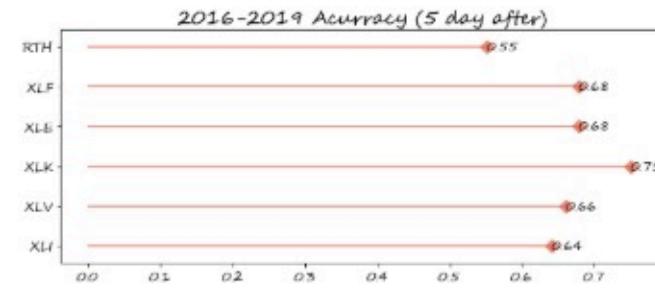
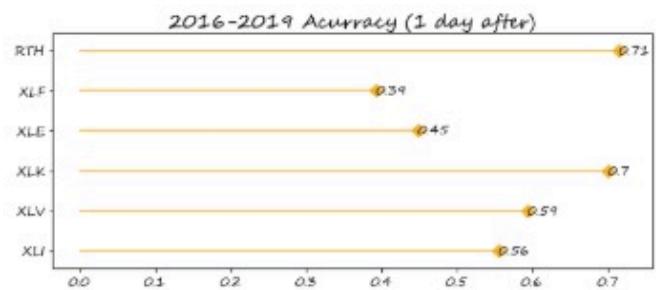
Probability that prices will be higher 4 days later after receiving a positive score > 0.8



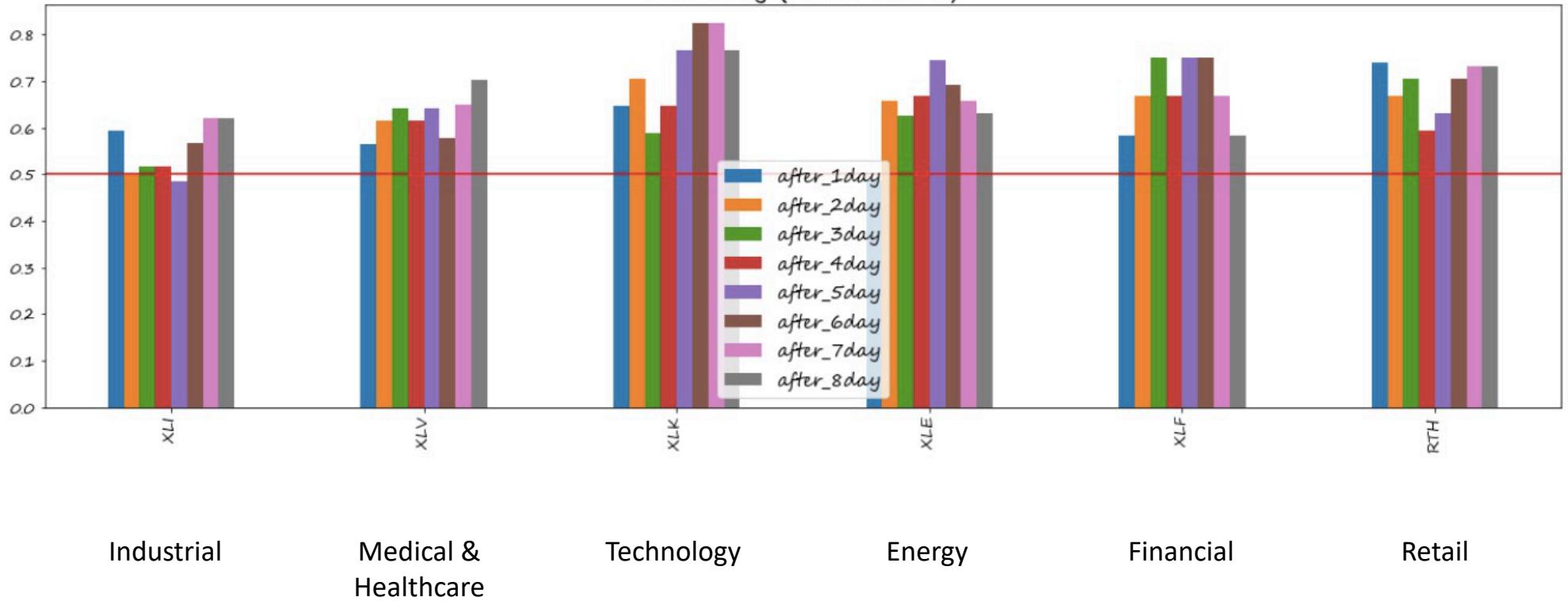


Blue line - probability that average returns over the next 4 days are positive

Orange line – probability that prices will be higher 4 days later



Accuracy (2018-2019)



Backtesting Strategy

利東教你如何發大財

Strategy



Thank You For Your Attention!