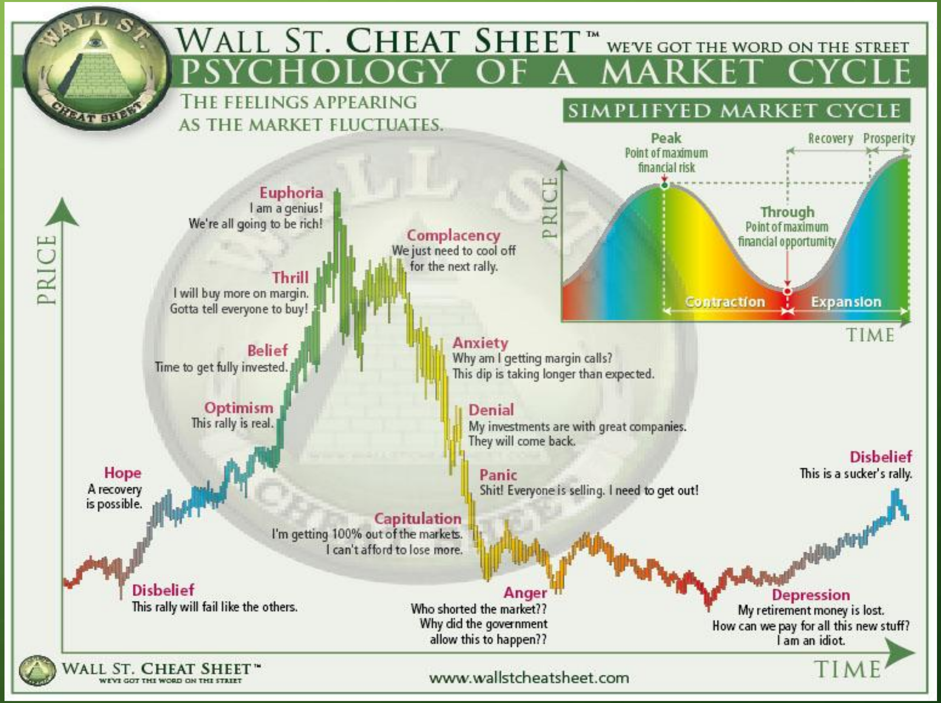
SENTIMENT ANALYSIS TEAM

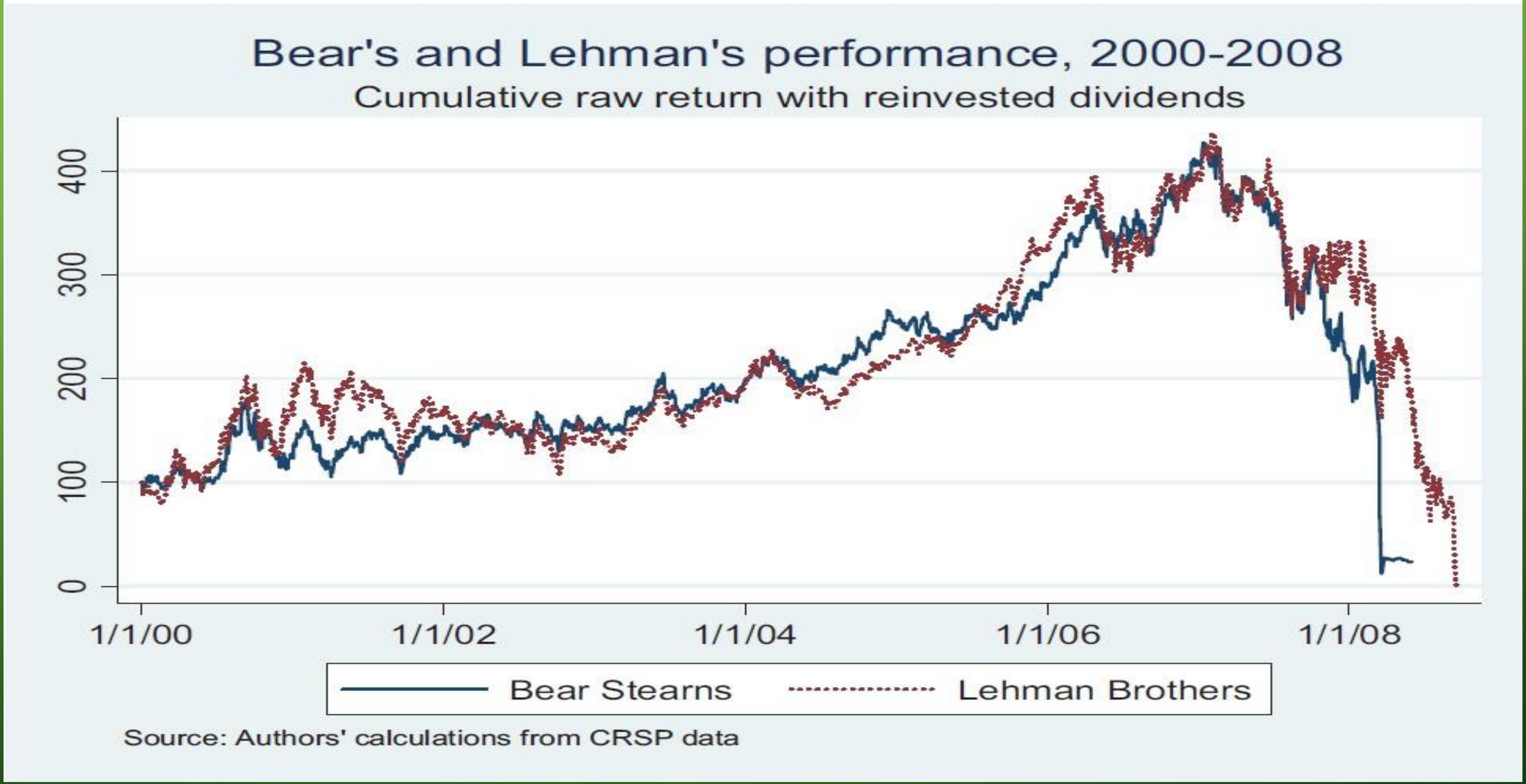
- Taehoon Kim
- Joe Gorfinkle
- Wenjie Gao
- Bharathan Mayavaram Sridharan



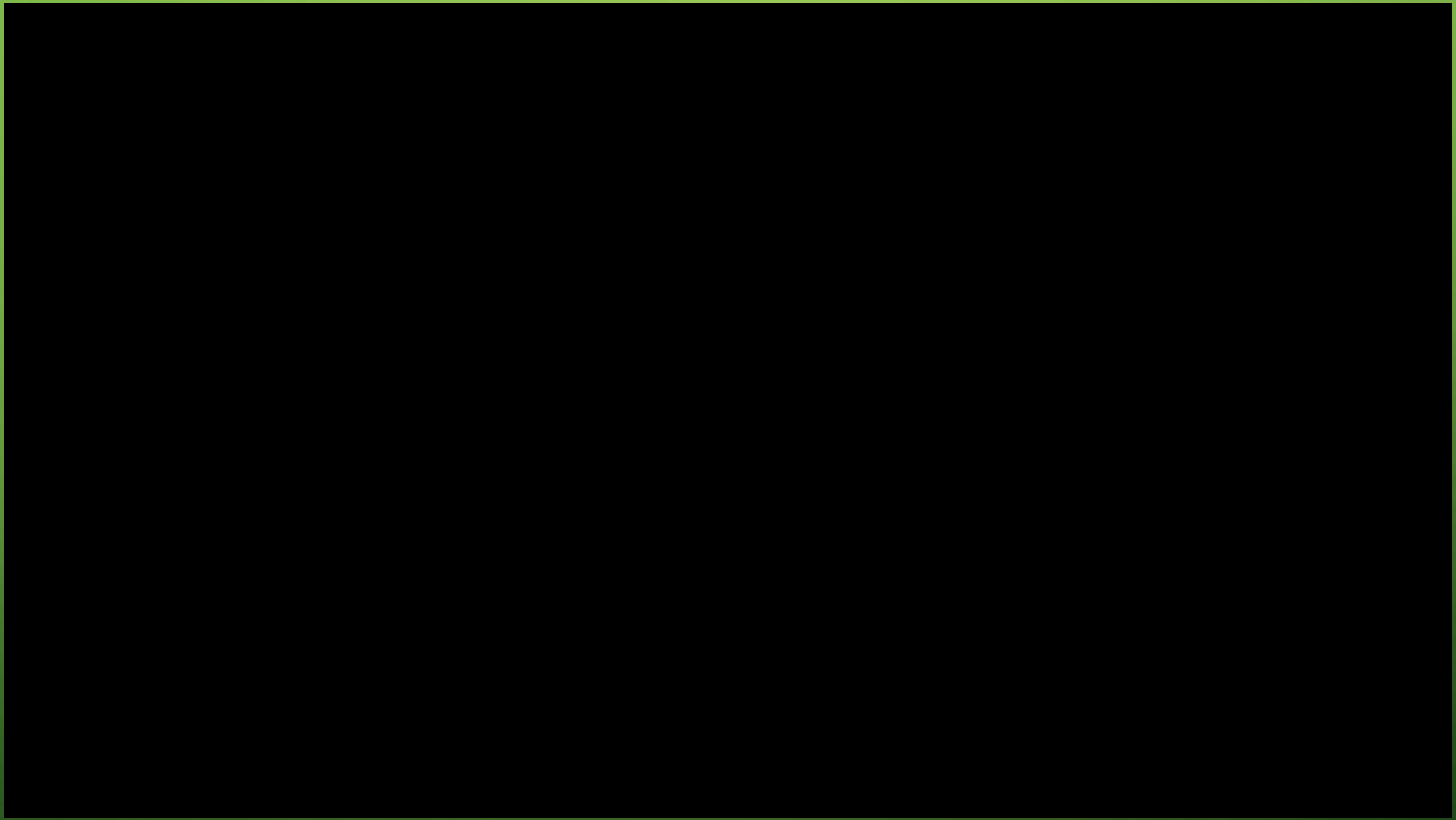
WALL ST. CHEAT SHEET™ WE'VE GOT THE WORD ON THE STREET
PSYCHOLOGY OF A MARKET CYCLE
THE FEELINGS APPEARING AS THE MARKET FLUCTUATES.
SIMPLIFYED MARKET CYCLE
Peak Point of maximum financial risk
Recovery
Prosperity
Through Point of maximum financial opportunity
Contraction
Expansion
PRICE
TIME
Euphoria I am a genius! We're all going to be rich!
Complacency We just need to cool off for the next rally.
Thrill I will buy more on margin. Gotta tell everyone to buy!
Anxiety Why am I getting margin calls? This dip is taking longer than expected.
Belief Time to get fully invested.
Optimism This rally is real.
Denial My investments are with great companies. They will come back.
Hope A recovery is possible.
Panic Shit! Everyone is selling. I need to get out!
Disbelief This is a sucker's rally.
Capitulation I'm getting 100% out of the markets. I can't afford to lose more.
Disbelief This rally will fail like the others.
Anger Who shorted the market?? Why did the government allow this to happen??
Depression My retirement money is lost. How can we pay for all this new stuff? I am an idiot.
WALL ST. CHEAT SHEET™
www.wallstcheatsheet.com

GLOBAL FINANCIAL CRISIS

FIGURE I. 2000-2008 PERFORMANCE

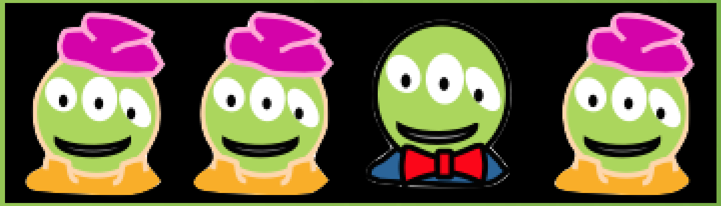


STOCK MARKET – THE ULTIMATE ILLUSION ?

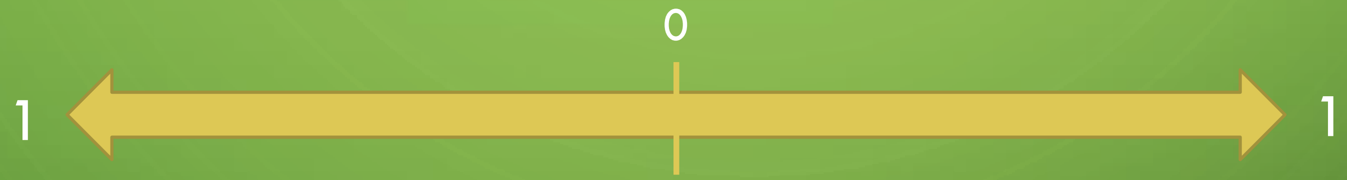


SENTIMENT DEFINITION - WIKIPEDIA

We should have played the video…



The use of natural language processing and text analysis to systematically identify, extract, quantify, and study affective states and subjective information.

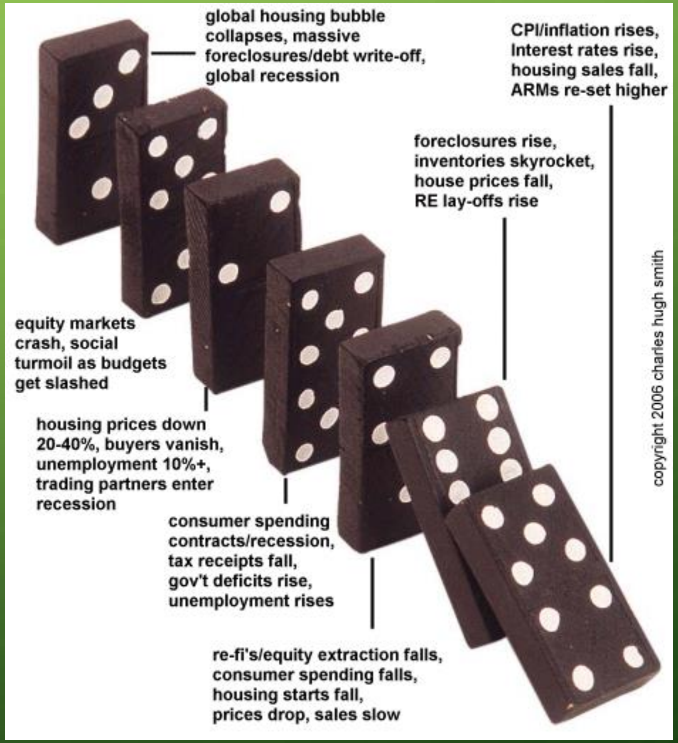


TextBlob

www.seekingalpha.com

GLOBAL FINANCIAL CRISIS

- ❖ Who pushed the 1st Domino ?
- ❖ Who does the market listen to ?


global housing bubble collapses, massive foreclosures/debt write-off, global recession
CPI/inflation rises, Interest rates rise, housing sales fall, ARMs re-set higher
foreclosures rise, inventories skyrocket, house prices fall, RE lay-offs rise
equity markets crash, social turmoil as budgets get slashed
housing prices down 20-40%, buyers vanish, unemployment 10%+, trading partners enter recession
consumer spending contracts/recession, tax receipts fall, gov't deficits rise, unemployment rises
re-fi's/equity extraction falls, consumer spending falls, housing starts fall, prices drop, sales slow
copyright 2006 charles hugh smith

WHAT DOES A STOCK PRICE REFLECT ?

∑ All Market Sentiments:

- Macroeconomic outlook
- Industry / Sector outlook
- Fundamental Analysis
- Technical Analysis
- **<u>Analyst Reports</u>**

Goldman Sachs OPPENHEIMER RBC® MorganStanley [A B] BERNSTEIN

OUR APPROACH

BEARS

- BB
- GE
- DISH
- AMD

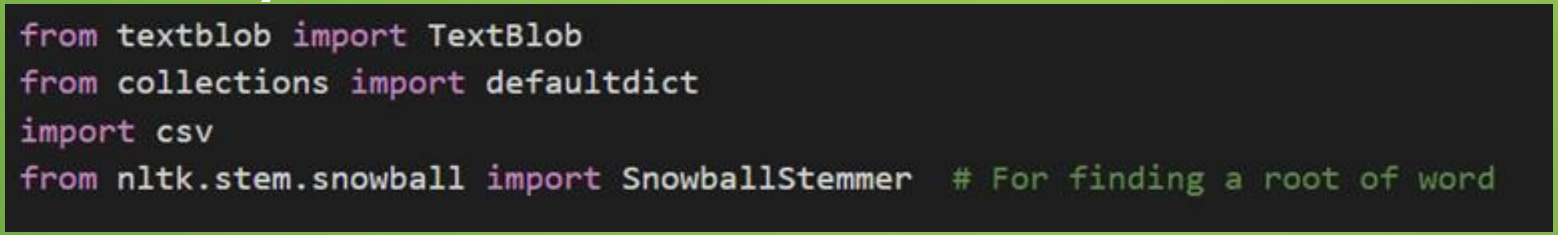
BULLS

- ✓ FB
- ✓ AMZN
- ✓ NVIDIA
- ✓ GOOG

- 150 SENTIMENT SCORES
- SA EARNINGS CALL TRANSCRIPTS
- 2013-18 W/HISTORICAL PRICES
- STATISTICAL CORRELATIONS

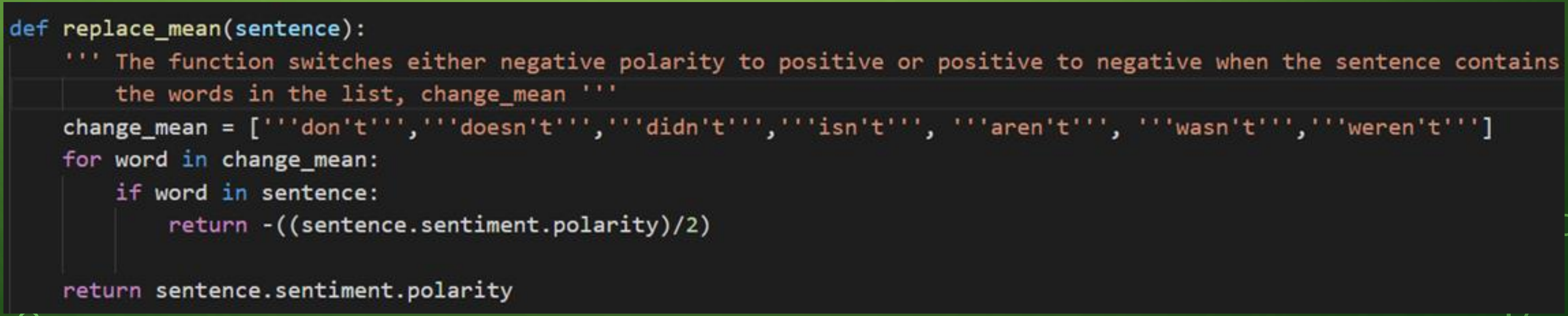
THE CODE

Library



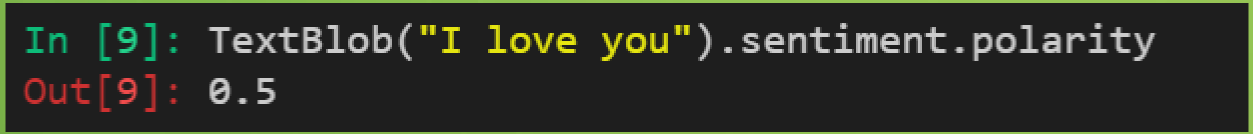
```python
from textblob import TextBlob
from collections import defaultdict
import csv
from nltk.stem.snowball import SnowballStemmer  # For finding a root of word
```

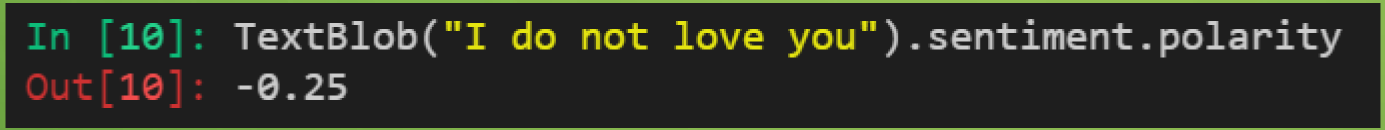
The Fucntion, Replace_mean

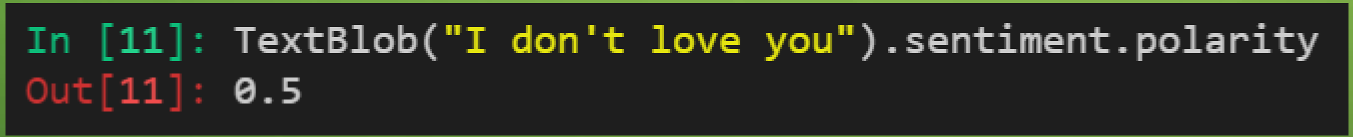


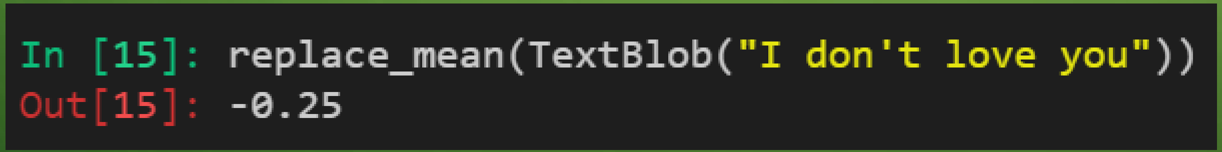
```python
def replace_mean(sentence):
    ''' The function switches either negative polarity to positive or positive to negative when the sentence contains
        the words in the list, change_mean '''
    change_mean = ['''don't''','''doesn't''','''didn't''','''isn't''', '''aren't''', '''wasn't''','''weren't''']
    for word in change_mean:
        if word in sentence:
            return -((sentence.sentiment.polarity)/2)

    return sentence.sentiment.polarity
```

EXAMPLE OF FUNCTION, Replace_mean

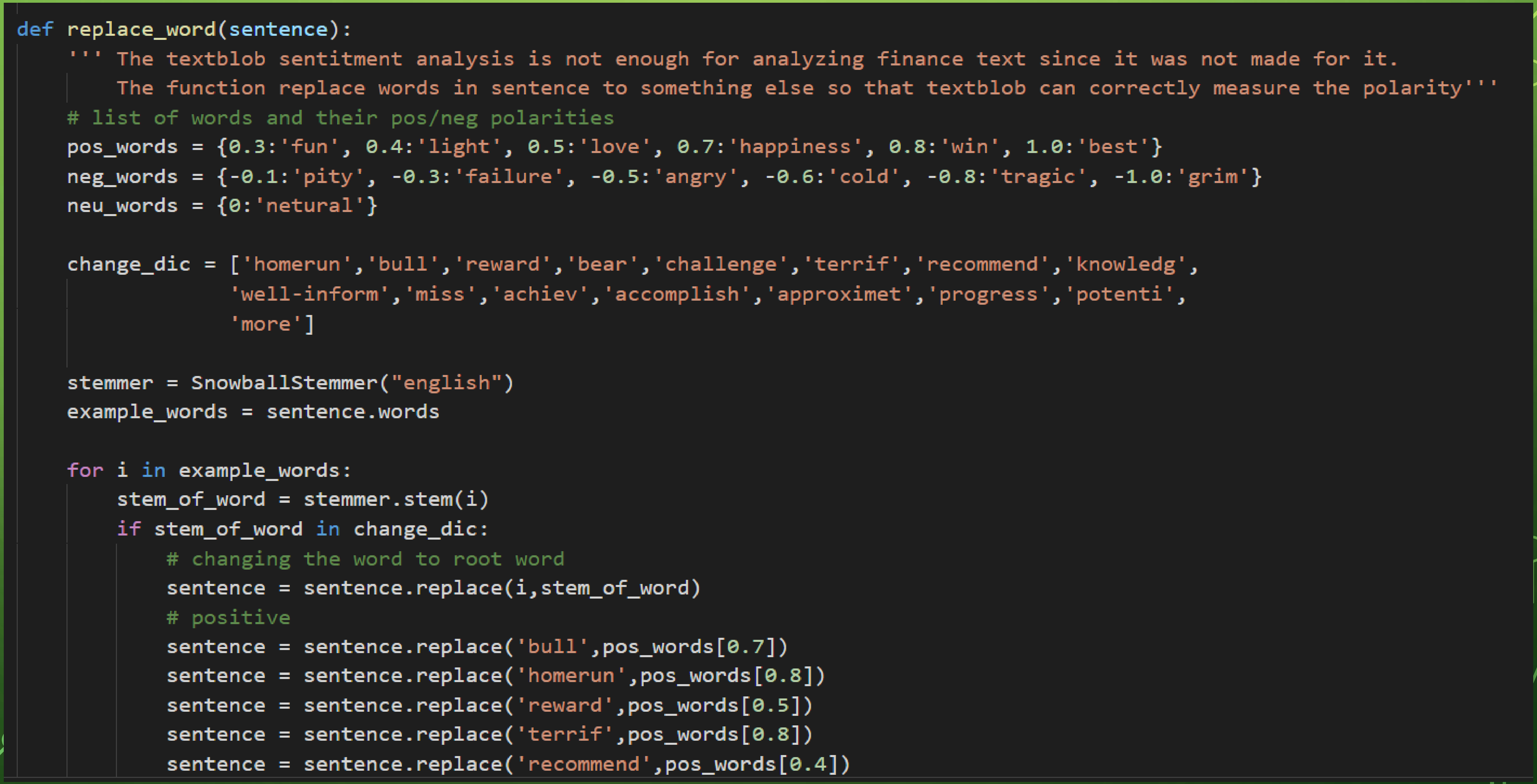

```
In [9]: TextBlob("I love you").sentiment.polarity
Out[9]: 0.5
```


```
In [10]: TextBlob("I do not love you").sentiment.polarity
Out[10]: -0.25
```


```
In [11]: TextBlob("I don't love you").sentiment.polarity
Out[11]: 0.5
```

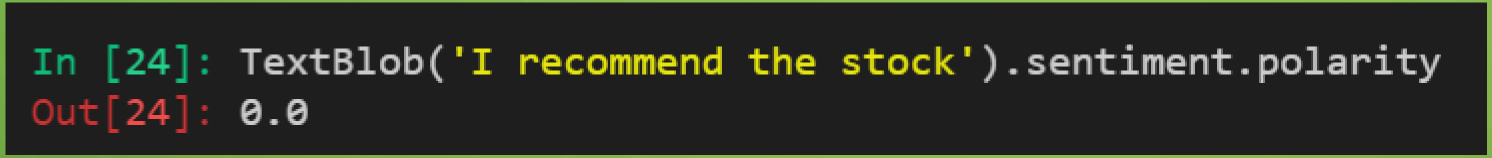

```
In [15]: replace_mean(TextBlob("I don't love you"))
Out[15]: -0.25
```

The Function, Replace_word

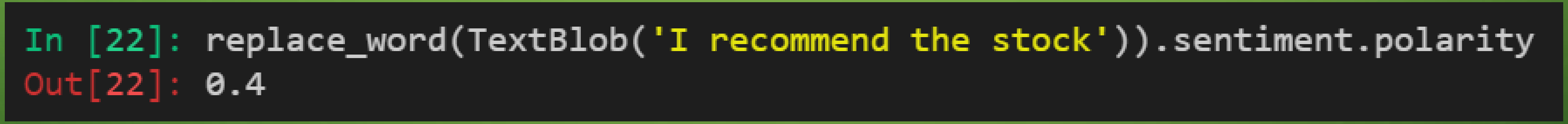


```python
def replace_word(sentence):
    ''' The textblob sentiment analysis is not enough for analyzing finance text since it was not made for it.
        The function replace words in sentence to something else so that textblob can correctly measure the polarity'''
    # list of words and their pos/neg polarities
    pos_words = {0.3:'fun', 0.4:'light', 0.5:'love', 0.7:'happiness', 0.8:'win', 1.0:'best'}
    neg_words = {-0.1:'pity', -0.3:'failure', -0.5:'angry', -0.6:'cold', -0.8:'tragic', -1.0:'grim'}
    neu_words = {0:'netural'}

    change_dic = ['homerun','bull','reward','bear','challenge','terrif','recommend','knowledg',
                  'well-inform','miss','achiev','accomplish','approximet','progress','potenti',
                  'more']

    stemmer = SnowballStemmer("english")
    example_words = sentence.words

    for i in example_words:
        stem_of_word = stemmer.stem(i)
        if stem_of_word in change_dic:
            # changing the word to root word
            sentence = sentence.replace(i,stem_of_word)
            # positive
            sentence = sentence.replace('bull',pos_words[0.7])
            sentence = sentence.replace('homerun',pos_words[0.8])
            sentence = sentence.replace('reward',pos_words[0.5])
            sentence = sentence.replace('terrif',pos_words[0.8])
            sentence = sentence.replace('recommend',pos_words[0.4])
```

EXAMPLE OF FUNCTION, Replace_word

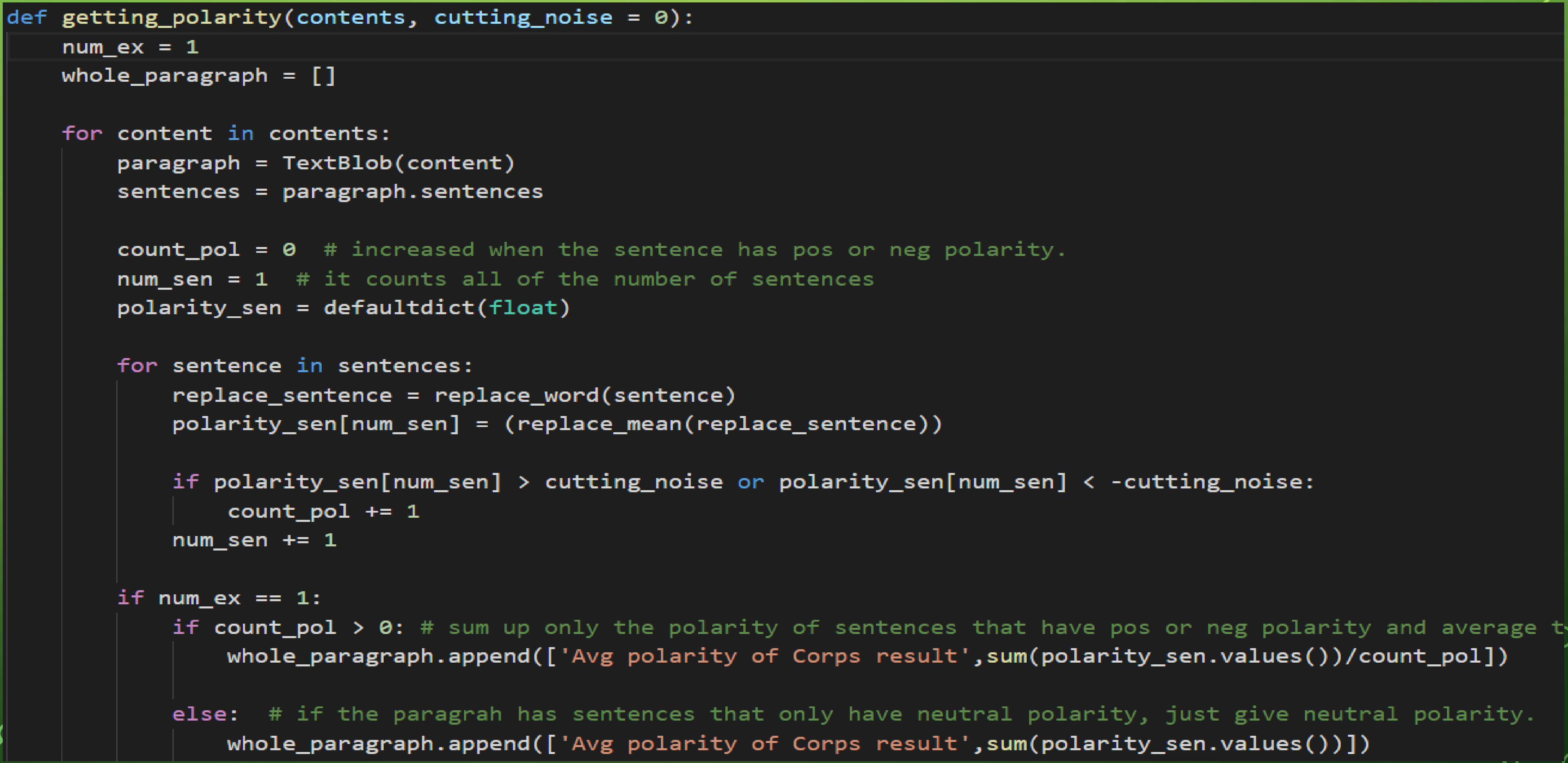


```
In [24]: TextBlob('I recommend the stock').sentiment.polarity
Out[24]: 0.0
```



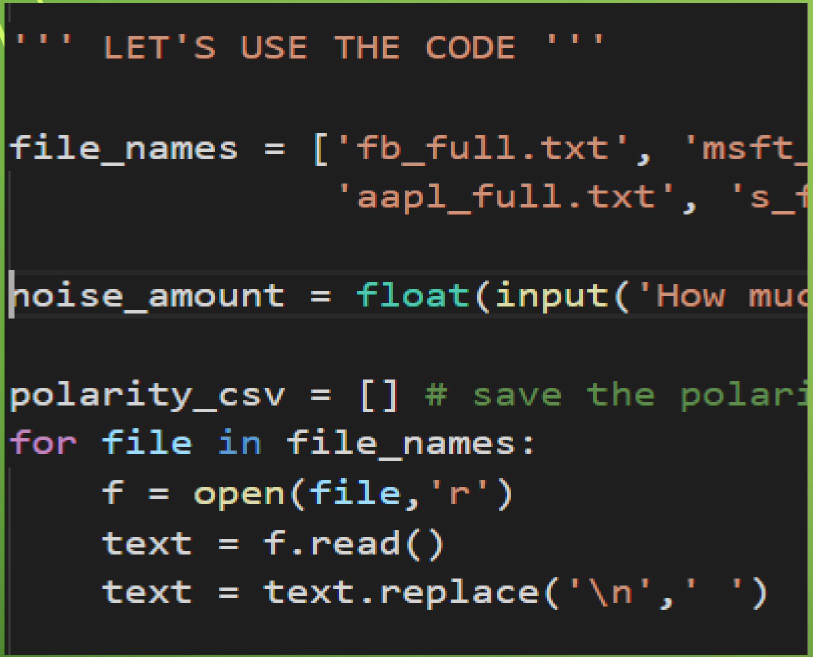
```
In [22]: replace_word(TextBlob('I recommend the stock')).sentiment.polarity
Out[22]: 0.4
```

The Function, getting_polarity

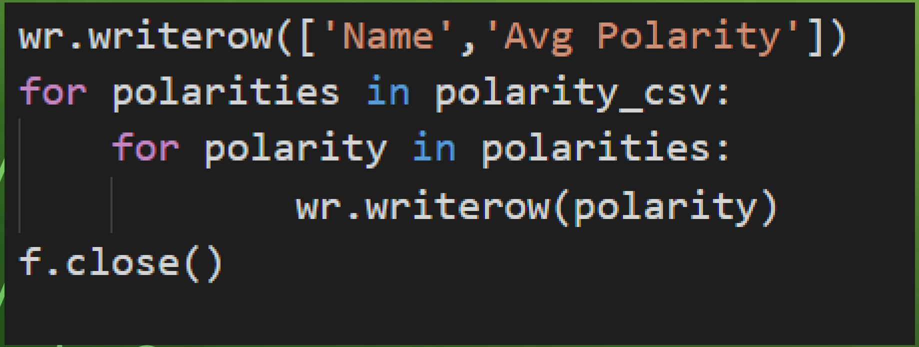


```
def getting_polarity(contents, cutting_noise = 0):
    num_ex = 1
    whole_paragraph = []

    for content in contents:
        paragraph = TextBlob(content)
        sentences = paragraph.sentences

        count_pol = 0  # increased when the sentence has pos or neg polarity.
        num_sen = 1  # it counts all of the number of sentences
        polarity_sen = defaultdict(float)

        for sentence in sentences:
            replace_sentence = replace_word(sentence)
            polarity_sen[num_sen] = (replace_mean(replace_sentence))

            if polarity_sen[num_sen] > cutting_noise or polarity_sen[num_sen] < -cutting_noise:
                count_pol += 1
            num_sen += 1

        if num_ex == 1:
            if count_pol > 0: # sum up only the polarity of sentences that have pos or neg polarity and average t
                whole_paragraph.append(['Avg polarity of Corps result',sum(polarity_sen.values())/count_pol])

            else:  # if the paragrah has sentences that only have neutral polarity, just give neutral polarity.
                whole_paragraph.append(['Avg polarity of Corps result',sum(polarity_sen.values())])
```

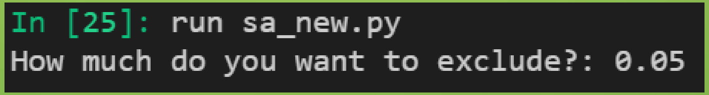
Result of the Code



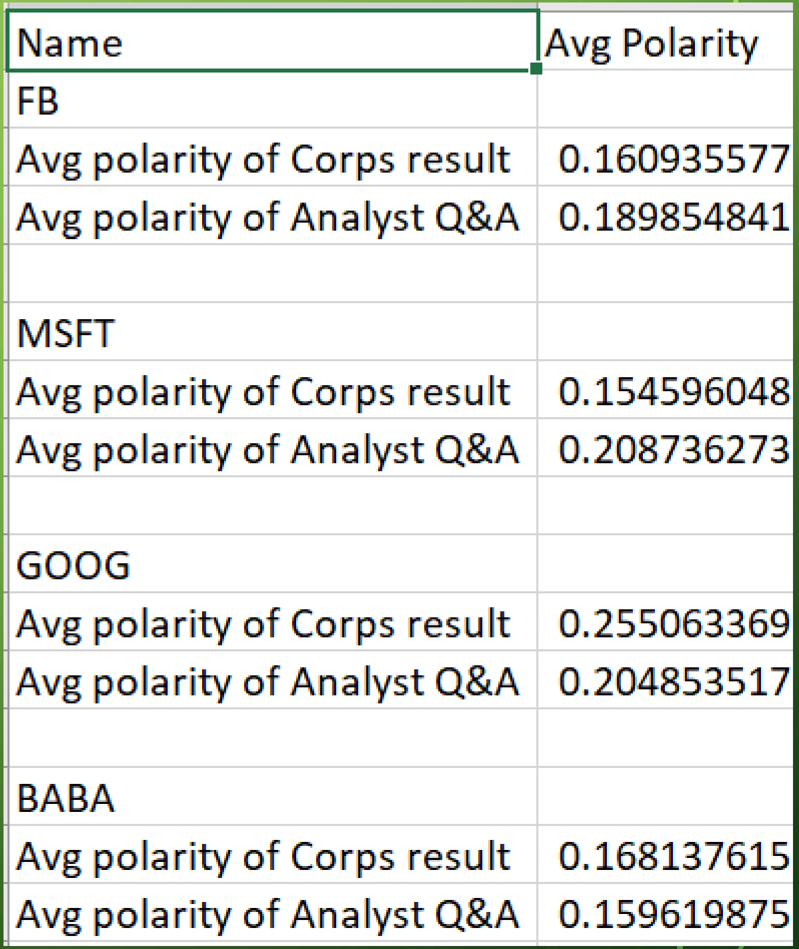
```
''' LET'S USE THE CODE '''

file_names = ['fb_full.txt', 'msft_
              'aapl_full.txt', 's_f

noise_amount = float(input('How muc

polarity_csv = [] # save the polari
for file in file_names:
    f = open(file,'r')
    text = f.read()
    text = text.replace('\n',' ')
```



```
wr.writerow(['Name','Avg Polarity'])
for polarities in polarity_csv:
    for polarity in polarities:
            wr.writerow(polarity)
f.close()
```

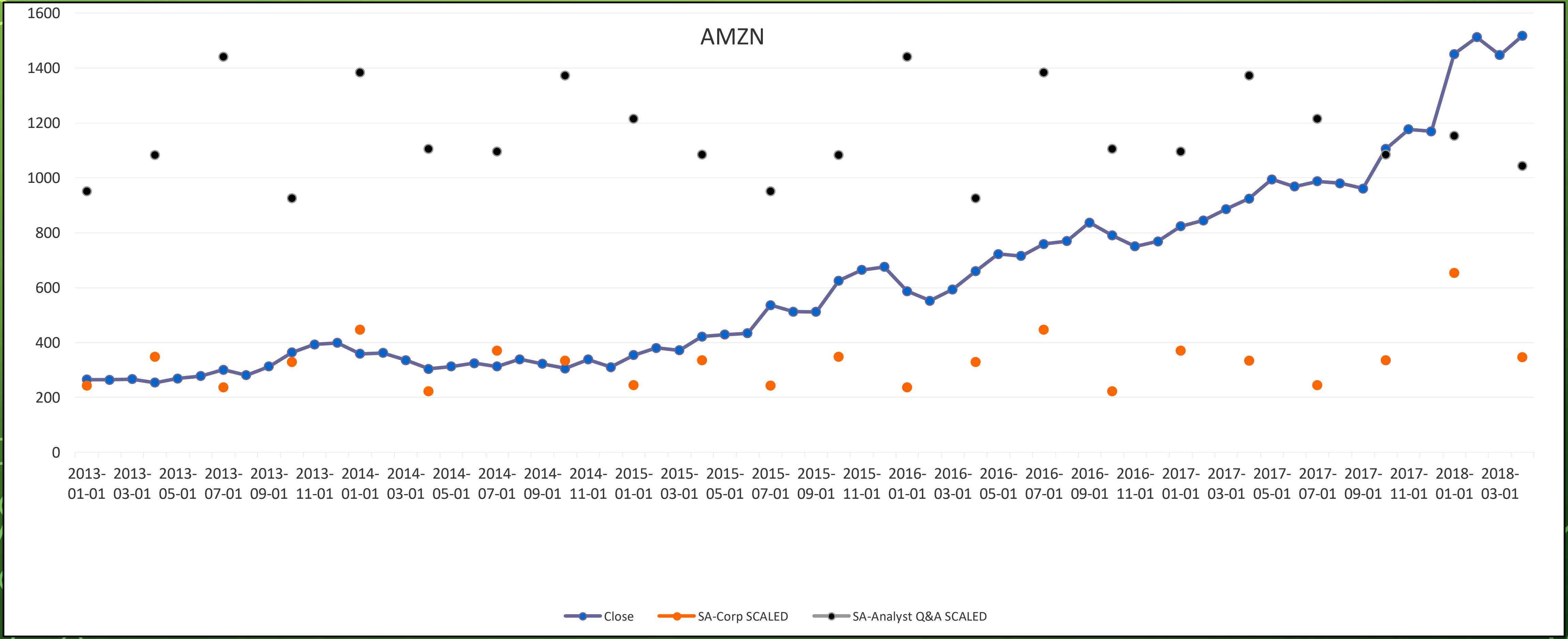


```
In [25]: run sa_new.py
How much do you want to exclude?: 0.05
```



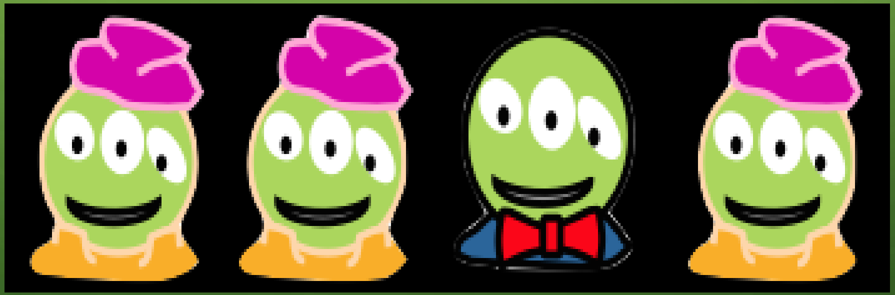
Name	Avg Polarity
FB	
Avg polarity of Corps result	0.160935577
Avg polarity of Analyst Q&A	0.189854841
MSFT	
Avg polarity of Corps result	0.154596048
Avg polarity of Analyst Q&A	0.208736273
GOOG	
Avg polarity of Corps result	0.255063369
Avg polarity of Analyst Q&A	0.204853517
BABA	
Avg polarity of Corps result	0.168137615
Avg polarity of Analyst Q&A	0.159619875

RESULTS - AMZN



INSIGHTS – PROJECT

- ETL – Huge Pain point w/o API's
- Lack of Data to discern anything useful (our Project limitation)
- Finance data becoming premium products…
- I have respect for PM now, basically how to you scale this !!!
- SA Model / Segment Interesting Idea (Google)
- Weighting Subjectiveness, ART or SCIENCE, Can ML Help?



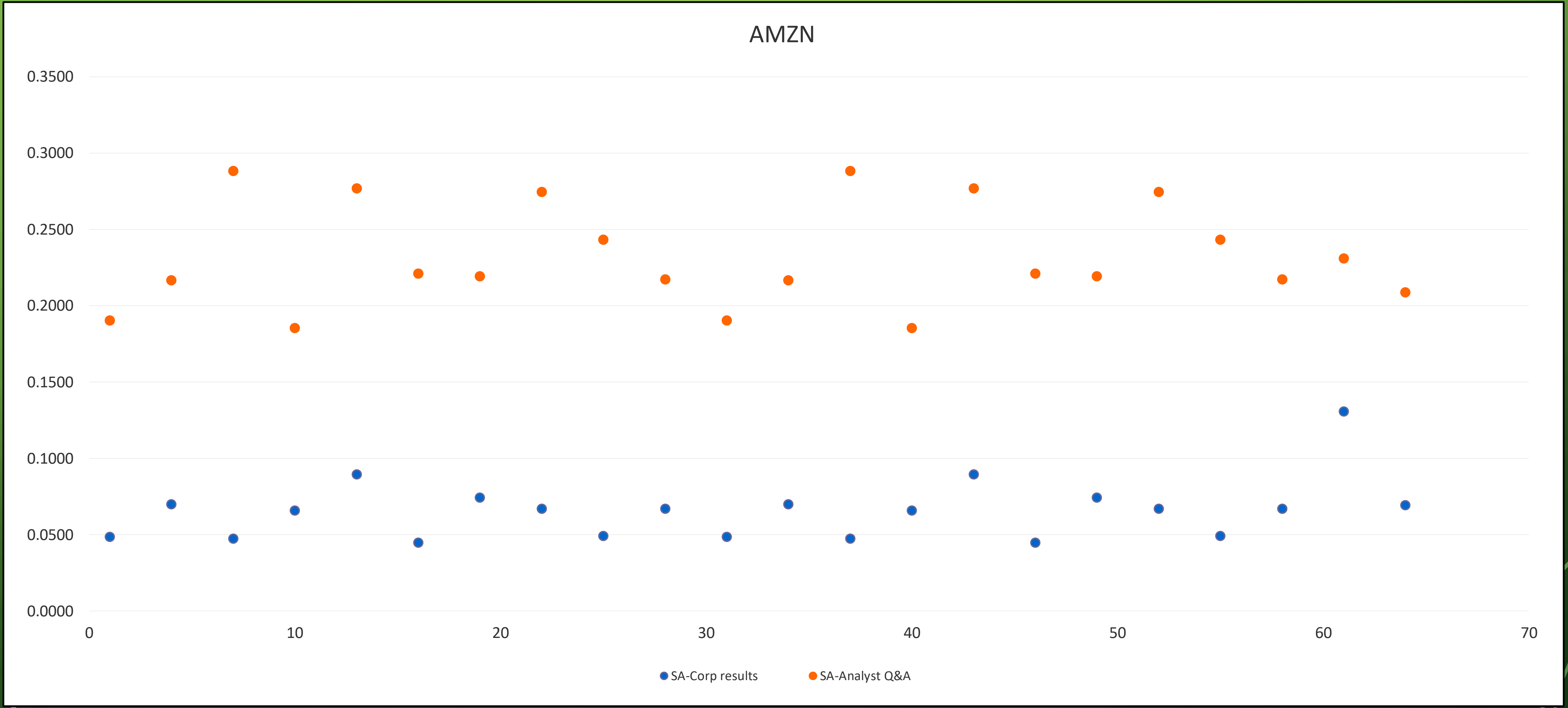
QUESTIONS ?

INSIGHTS – TRANSCRIPTS

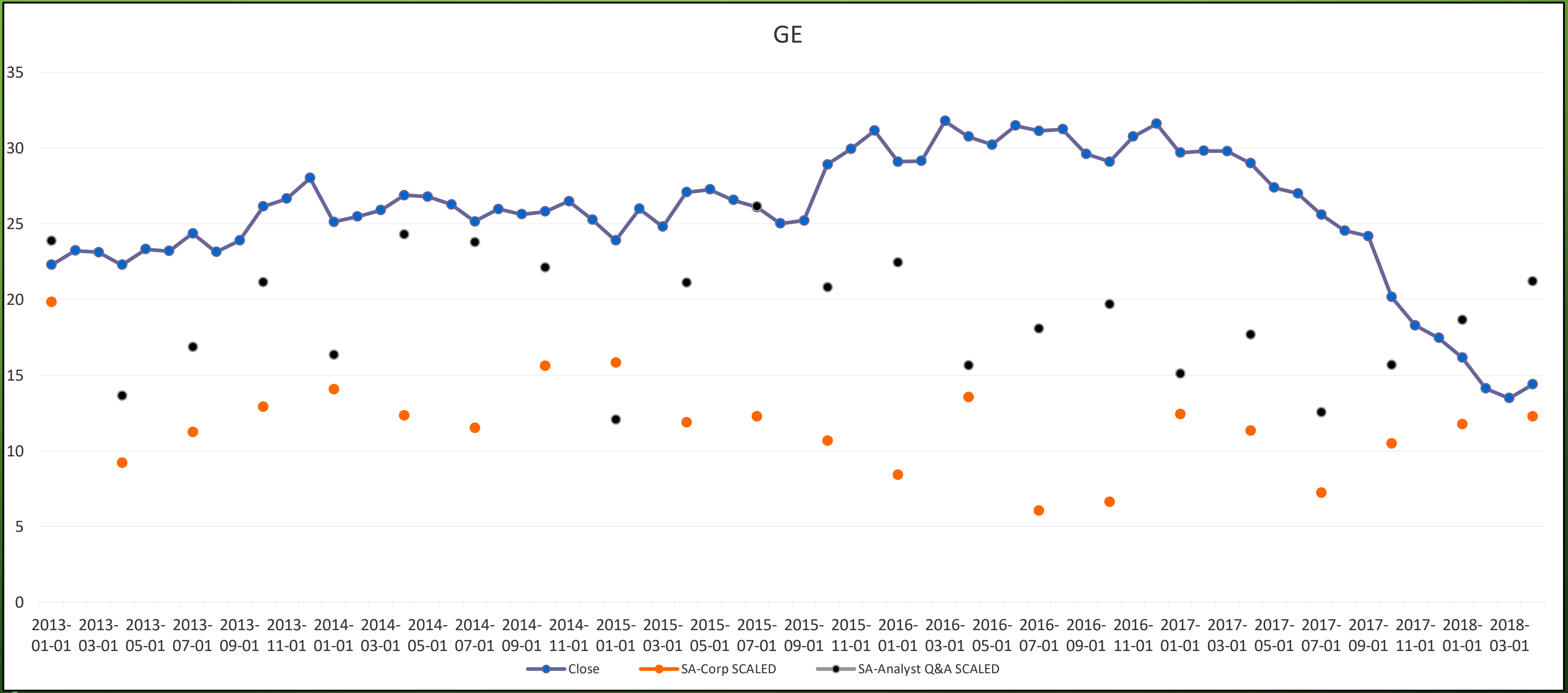
- WE NEVER FOUND NEGATIVE SENTIMENT (EVEN IN 2008)
- AMZN VERY NEUTRAL WITH NARROW RANGE
- ANALYSTS MORE OPTIMISTIC THAN CORPORATE (DISH exception)
- COMMENTS FROM LISTENERS HAVE STRONG SENTIMENT…
- THE SEARCH CONTINUES FOR THE DOMINO PUSHERS…

Michael J. Burry was one of the first investors to recognize and profit from the impending subprime mortgage crisis.

RESULTS - AMZN

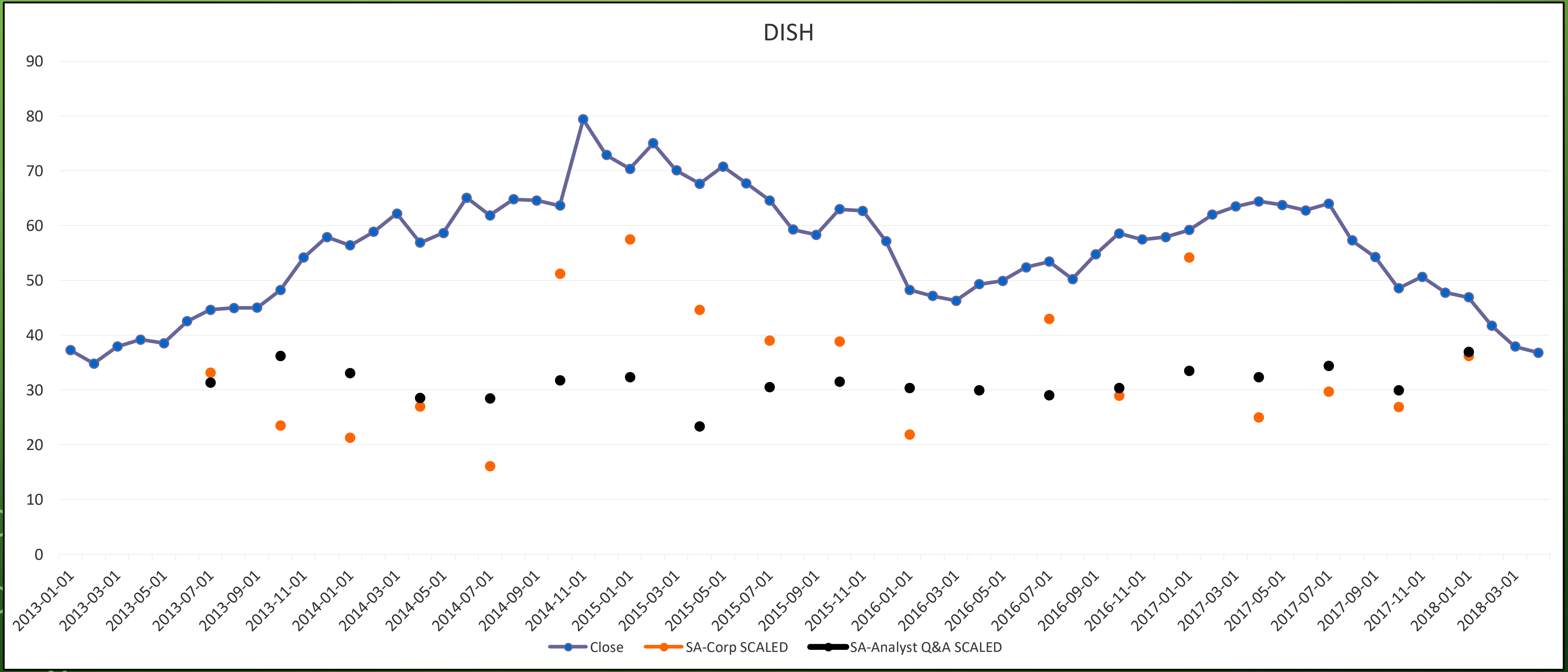


RESULTS - GE

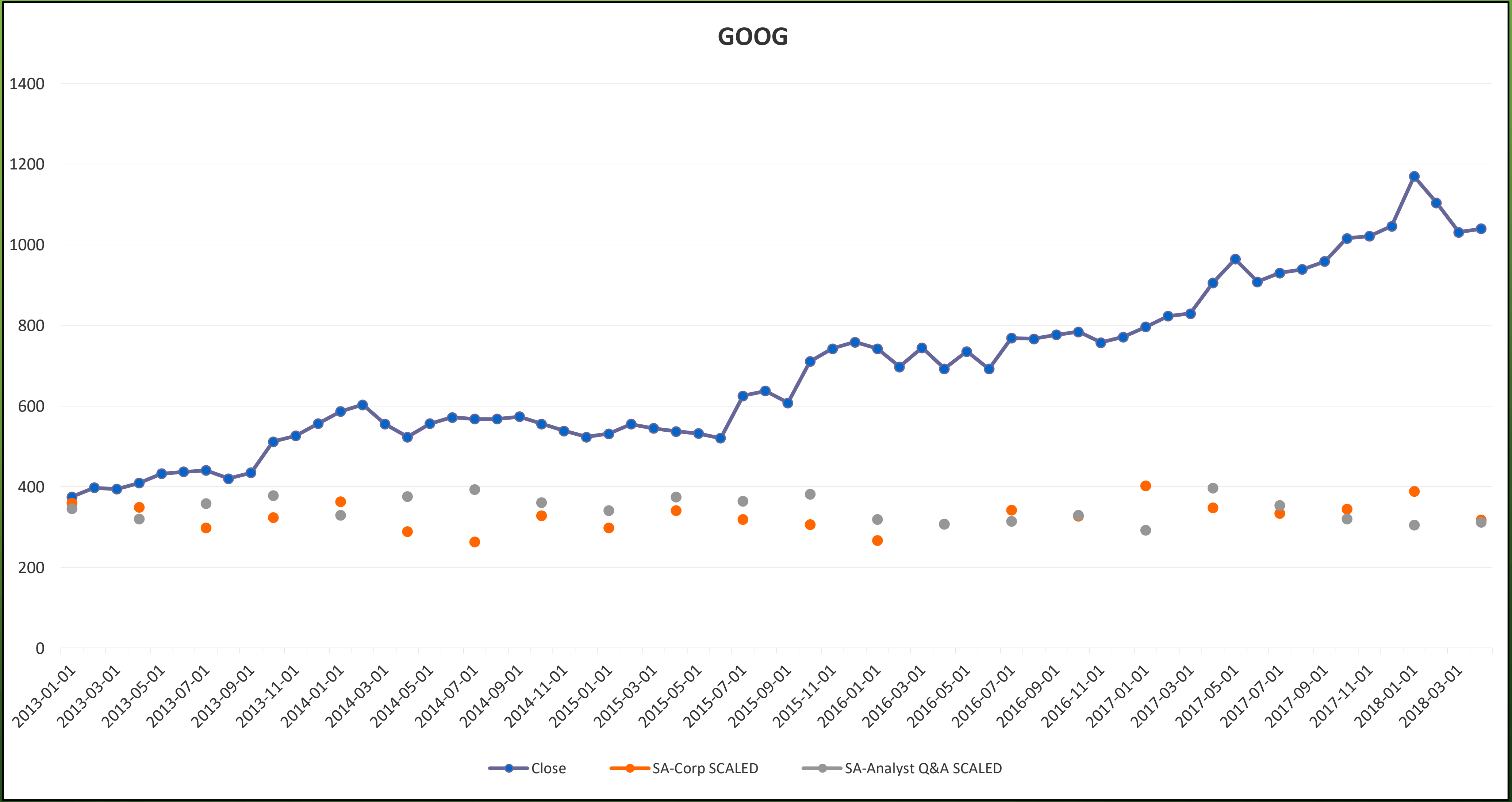


GE
35
30
25
20
15
10
5
0
2013-01-01
2013-03-01
2013-05-01
2013-07-01
2013-09-01
2013-11-01
2014-01-01
2014-03-01
2014-05-01
2014-07-01
2014-09-01
2014-11-01
2015-01-01
2015-03-01
2015-05-01
2015-07-01
2015-09-01
2015-11-01
2016-01-01
2016-03-01
2016-05-01
2016-07-01
2016-09-01
2016-11-01
2017-01-01
2017-03-01
2017-05-01
2017-07-01
2017-09-01
2017-11-01
2018-01-01
2018-03-01
Close
SA-Corp SCALED
SA-Analyst Q&A SCALED

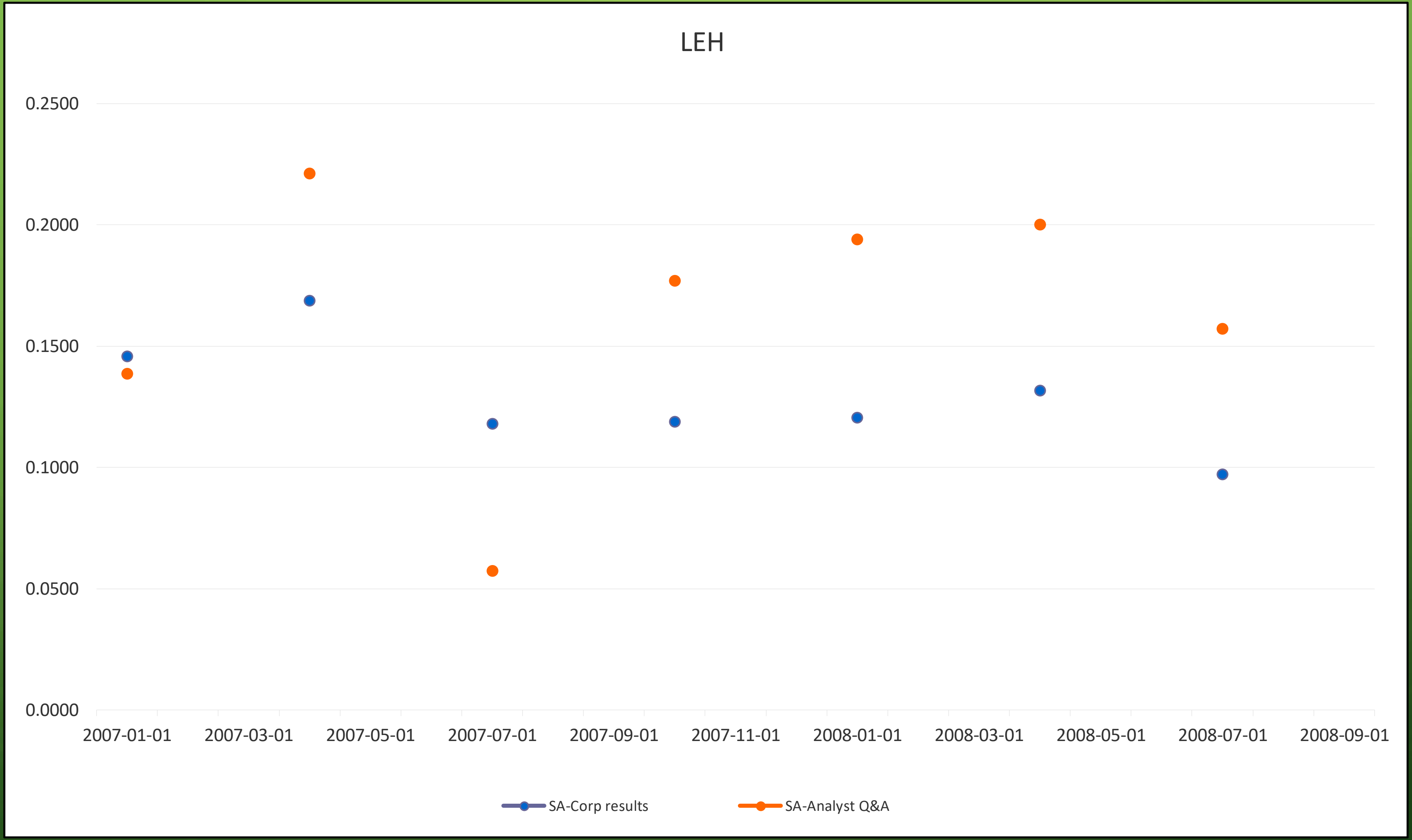
RESULTS - DISH



RESULTS - GOOG



RESULTS – LEHMAN BROTHERS


LEH
SA-Corp results
SA-Analyst Q&A