EXAMENSARBETE INOM TEKNIK,
GRUNDNIVÅ, 15 HP
STOCKHOLM, SVERIGE 2018

Sentiment analysis of tweets in comparison to a company's financial performance

ERIC LEWERENTZ

ELLINOR WESTERBERG

KTH
SKOLAN FÖR ELEKTROTEKNIK OCH DATAVETENSKAP

Sentiment analysis of tweets in comparison to a company's financial performance

ERIC LEWERENTZ, ELLINOR WESTERBERG

Bachelor in Computer Science
Date: June 6, 2018
Supervisor: Richard Glassey
Examiner: Örjan Ekeberg
Swedish title: Sentimentanalys av tweets i jämförelse med ett företags finansiella resultat
School of Electrical Engineering and Computer Science KTH

Abstract

This study analyzes peoples reactions in social media to the release of a company's quarterly report. Sentiment analysis was performed on tweets about a company both from a short-and long-term perspective. On the long-term perspective, a two year period of sentiment was studied in relation to its quarter's percental change of net income and revenues. Three companies was investigated in this process. On the short-term perspective, a linear regression was conducted based the quarter's percental change of net income and revenues and on sentiment score, ranging from release day to 2 days after the release of a company's quarterly report, in total five companies were studied. The result inferred that there was no correlation between the company's net income and the reaction on Twitter on both long and short-term. Our conclusion is that the attitude towards a company is influenced by other factors than net income and revenues. The study also indicates that sentiments found in Twitter messages regarding a company name are related to a positive or negative expectation of the company. Findings suggest future studies to focus on companies greatly dependent on one product and analyze sentiment regarding that product instead, as the product is likely to impact financial results.

Sammanfattning

Denna studie undersöker hur människors reaktioner utspelas på sociala medier efter att ett företags kvartalsrapport släppts. En sentimentanalys utfördes på tweets angående ett företag ur både ett kortsiktigt och ett långsiktigt perspektiv. I det långsiktiga perspektivet jämfördes sentimenten från en tvåårsperiod med resultatet från kvartalsrapporterna under samma tid. Tre företag undersöktes. I det kortsiktiga perspektivet utfördes en linjär regression baserat på förändringen av rörelseresultatet och intäkterna, samt sentimentet med ett tidsintervall på 2 dagar från det att bolagets kvartalsrapport publiceras, detta utfördes på fem bolag. Resultaten från både det långsiktiga och det kortsiktiga perspektivet tyder på att det inte finns någon korrelation mellan företagets rörelseresultat och intäkter, samt sentimentet på Twittermeddelanden innehållande bolagets namn. Vår slutsats är att inställningen till ett företag påverkas av andra faktorer än rörelseresultatet och intäkterna. Studien indikerar också att inställningen på Twitter till ett företagsnamn är baserat på en positiv eller negativ förväntan. Vi föreslår att framtida studier fokuserar på företag som är starkt beroende av en produkt och analyserar sentiment för den specifika produkten istället eftersom produkten i detta läge har en stor inverkan på kvartalsrapportens resultat.

Contents

Chapter 1

Introduction

The interest in sentiment analysis has grown rapidly in the past few years (Pang and Lee 2008). Sentiment analysis is the process for automatically obtaining opinions in text and see if the sentiment, i.e. the attitude, is positive, negative or neutral (Maynard 2017). Lately, new techniques have made it possible to develop refined methods (Cambria, Schuller, et al. 2013) and there are many applications for the field. In many cases, they are about comprehending how people in general react or feel towards different topics.

Another growing trend is the use of social media. One example is the micro-blogging service Twitter where people posts tweets, i.e. messages, about various topics that everyone with an account can find and comment on. Due to a large number of users, and the huge variation in data, Twitter makes a good base for retrieving data that can be used in a sentiment analysis (Pak and Paroubek 2010). There are multiple studies where trends and opinions have been detected by using sentiment analysis on Twitter.

Furthermore, sentiment analysis on tweets has proven useful in a financial context. One example is the study conducted by Bollen, Mao, and Zeng (2011) that uses it as a tool in order to predict the stock market. They performed a sentiment analysis on tweets about a company. The result was then compared to the share price. It was proven that a correlation between the share price and the sentiment could be found. However, the study also claims that the financial discussion is news driven and suggest that focus should be on analyzing reactions close to the release of the news.

A company goes through different phases and is impacted by the market it is involved in; the welfare of the company is always changing. These changes are regularly presented as the result in the quarterly reports. These reports are used by investors in order to decide if the company is worth investing in and it is quite common to share such opinions in social media, such as Twitter.

In this report, we will proceed from the assumption that the financial discussion is news driven and examine if financial news, such as a quarterly report, can influence the sentiment on Twitter.

1.1 Purpose

Ticknor (2013) highlights that the complexity of the market is related to a number of factors, including quarterly earnings reports. Therefore, the purpose of this study is to use sentiment analysis in order to see if any direct connection can be found between how individuals reason about companies and how well the company is actually performing in financial terms, such as an increase or decrease in net income. Findings of this study might contribute to the discussion regarding the usefulness of sentiment analysis in the context of risk assessment of a company. It could also conduce to the understanding of what factors affect the attitude towards different companies. Furthermore, results of this study could contribute in identifying how sentiment analysis can be applied in a financial context.

1.2 Research Questions

This paper will focus on how a company's financial results affect the sentiment of comments found on Twitter. The following question will be investigated:

How does a company's release of financial results impact the sentiment of comments found on Twitter that are related to the company?

1.3 Approach

In order to answer the research question, we will compare the results from a company's quarterly report to the sentiment of comments on

Twitter. We will first investigate how the sentiment related to a company changes over time and if any correlation can be found between the sentiment and the fundamentals (see section 2.3) of a company for a period of multiple years. Secondly, we will look closer at the time around the release of the report in order to see if the sentiment is affected by the results over a shorter period. We will use a method similar to the one described by Asur and Huberman (2010) (see section 2.2.1).

1.4 Scope

This study will focus on the application of sentiment analysis in a financial context rather than studying the actual framework and algorithms of the sentiment analysis. It also focuses on analyzing how companies are perceived in comments found on Twitter. Other social media is not included. More formal texts are not taking into account and due to sentiment analysis algorithms being language specific, we will only consider text identified as English.

Chapter 2

Background

This chapter will describe how the sentiment analysis process is done. Then, related studies will be discussed with relevance to sentiment analysis in a financial context.

2.1 The sentiment analysis process

Sentiment analysis (also known as opinion mining in a broader context) is the process of systematically determining if there is an emotion or opinion in a text and to what degree that emotion or opinion is expressed. It is a new research field that has been important since around the year of 2000. There are many applications for the field, everything from analyzing the political situation in a country to see how people react to a new product (Liu 2012).

The rest of this section describes parts the process in more detail: different methods, on what levels the sentiment analysis can be done, preprocessing of the text and what is special when performing a sentiment analysis on data from social media. Lastly, the knowledge-based model VADER that is used in this study, is described.

2.1.1 Different methods

In essence, sentiment analysis can be divided into two categories: knowledge-based techniques and statistical methods. In knowledge-based methods, a set of linguistic rules decides how to determine the sentiment. A central part is the lexicon that labels a number of words as positive or negative and is used to classify the words in the text (ibid.).

Statistical methods use machine learning methods in order to determine the sentiment. The learning based approach is done by providing human-made examples. The examples contain annotations, typically a categorization of positive, negative or neutral, and are there to clarify for the algorithm what is considered a positive sentiment and what is considered negative (Yu, Duan, and Cao 2013). There are also hybrid approaches that use both (Cambria, Das, et al. 2017).

2.1.2 Levels

Sentiment analysis can be done on different levels: document level, entity and aspect level and sentence level. The simplest form is document level where it is assumed that there is only one opinion about only one entity in the whole document. In practice, this is most of the time not the case. In most cases, a document consists of different opinions concerning different aspects of an entity. Due to this diversity, another way is to consider sentiment analysis on an entity and aspect level. In that case, the sentiment depends on what aspect of the entity are mentioned. Another, less complex, solution is to look at a sentence level where the sentiment analysis is done on a shorter part of the text, for example, on a sentence. It is then assumed that the identity of the entity in the sentence is known and that there are only one single opinion in each sentence (Feldman 2013) (Liu 2012).

2.1.3 Preprocessing

Before performing the main part of the sentiment analysis (where the sentiment of the message is calculated) the text is often preprocessed.

Extracted data normally contains noise and irrelevant parts for the sentiment analysis, so it is sensible to filter and clean the data in order to remove irrelevant information. For example a requirement of at least one certain keyword such as "I am" or "makes me". Another example is to remove URLs from the text (Bollen, Mao, and Zeng 2011). Ensuring that there are no re-tweet messages included is necessary to ensure that a few tweets are not getting its sentiment score calculated twice.

When the text is cleaned, it needs to be divided into smaller text units, such as words, numbers, symbols and punctuation; a process called tokenizing. This list of tokens is then used as input in the other algorithms in the process. In a simpler analysis, the sentiment is simply calculated for each word and combined into a total score (Feldman 2013). In more advanced methods the preprocessing often contains other steps as well.

An example of another preprocessing method is Part-Of-Speech (POS) tagging where the words are tagged with their part of speech, for example, noun, verb or adjective (Maynard 2017). Pak and Paroubek (2010) describes how POS-tagging could be used to identify what part of sentences speaks about the subject. These parts, that are describing the subject, are more likely to contribute to a positive or negative sentiment and hence, this information can help to improve the sentiment analysis. On the contrary, Kouloumpis, Wilson, and Moore (2011) tried using POS-tagging as an input for training a learning-based sentiment analyzer and considered it not being effective for improving accuracy. The study indicated that it could have something to do with the POS-tagger tool used, or, as a contradiction to Pak and Paroubek (2010) findings, that the structure of social media content is not formal enough to make use of POS-tagging.

2.1.4 Sentiment analysis in social media

Sentiment analysis in social media is often more complicated to obtain accurate scoring compared to more formal texts. The text in social media is often short, has unusual spelling and uses slang and emoticons. The emotional undertone can also be dependent on the context, for example, a link to another comment or article, or social relations. This puts high demands on the algorithm used for the analysis (Maynard 2017).

2.1.5 VADER

VADER (Valence Aware Dictionary for sEntiment Reasoning) is a lexicon based model used in sentiment analysis that is customized specifically for the English language in social media. As described above, social media has special requirements in natural language which VADER was constructed to take into account. VADER does not only classify words but also emoticons, acronyms and slang. The score for the message is based on how the words in it are classified and how they are written. Since the language in social media is very informal, people tend to use uppercase letters or multiple exclamation marks in order to emphasize their words. A sentence that is slightly negative will then be scored as more negative if there is an exclamation mark in the end than if there is a period (Gilbert 2014).

Every message gets a positive score, a negative score and a neutral score, each between 0 and 1. The message also gets a compound score between -1 and 1 where -1 is extremely negative, 1 is extremely positive and 0 is neutral (ibid.).

The process does not involve any Part-of-Speech (POS) tagging or similar processes that can detect what part of the sentence is of interest. However, that might not be necessary according to Kouloumpis, Wilson, and Moore (2011). Since tweets often are short and only concerns one aspect of a topic they can be treated on a sentence-level (ibid.) where the target and subject are assumed to be known.

2.2 Related studies

Below sentiment analysis will be put in a financial context. In addition, social media's relevance to the financial context is explained. Finally, studies discussing the collection of data from the internet will be discussed.

2.2.1 The financial context

Bollen, Mao, and Zeng (2011) mentions that financial discussions are news driven, suggesting focus should be on analyzing reactions nearby the time when news has occurred in order to draw conclusions of the outcome.

Word of mouth (WOM) is the process of sharing information from mouth to mouth from one person to another. In commercial situations, this means that consumers share their opinions, or reactions about businesses, products, or services (Jansen et al. 2009). Duan, Gu, and Whinston (2008) describes positive word of mouth as a powerful marketing medium to influence customers. Furthermore Jansen et al. (2009) mentions that the ease of sentiment analysis of a brand on micro-blogging services such as Twitter could be viewed as an intelligence source to empower a company's competitive strength. To conclude, this indicates that there is text found on social media which is relevant for sentiment analysis about a company.

One issue regarding sentiment occurring for company branding is highlighted, only 20% of the tweets mentioning a company's brand contains sentiment according to Jansen et al. (2009) study. As a result, a great deal of data is required to get a valid sentiment scoring.

Predicting box-office revenues for movies based on social media has been done by Asur and Huberman (2010). The aim is to find a correlation between the attitude towards newly released movies on Twitter and the revenues made by the movie title. In order to judge the correlation, they make a linear regression based on the tweet rate and sentiment ratio before and after the movie release. The sentiment ratio is calculated as the absolute amount of the quotient of all positive tweets divided by the negative sentiment. The study concludes that the tweet rate is a more determining factor than the sentiment for predicting the box-office revenue of a movie. Furthermore, a generalized methodology is presented in the study for predicting revenues for certain product.

2.2.2 Collection of data

To actually extract text from the internet, a scraping tool is normally used. Scraping on blogs, forums and social media can be done by using web crawlers and other scraping techniques (Yu, Duan, and Cao 2013). Something worth mentioning is that scraping may violate homepages' terms of service, as is the case for Twitter (*Twitter, Terms of Service* 2018).

2.3 Central Terminology

Consolidated Statements of Income

Usually called income statement. This is a company's reporting of revenue for a time period, normally for a year or a quarter of a year (a three month period). It also clarifies all costs associated to generate the revenue, such as payment for working hours. What is left over once all costs have been covered to generate the revenue is normally called net income. Finally, it also tends to show earnings per share (EPS), that is, how much each share would be allocated if the company decided to distribute all of the net earnings for the period to its shareholders (*SEC* 2018).

Company

A company is an organization that could be viewed as executing value-adding processes. That is, a company strives to produce a value for their customers in return of revenues. The better the executed value-adding processes are done, the more effective revenues can be generated. As a result, higher competition is achieved (Engwall 2017).

Publicly traded company

A company listed on a stock exchange (a market to buy and sell parts of the ownership of a company), where this stock exchange is available for the public, i.e. private persons are eligible to obtain ownership of a publicly traded company.

Quarterly report

Most publicly trading companies present their financial status for every each quarter a year. Typically, this quarterly report describes how much revenues the company made or whether the company made a profit or not.

SECs Form 10-Q

The federal securities laws (of the united states of America) require publicly traded companies to disclose information on an ongoing basis. According to United States Securities and Exchange Commission (SEC), a company is required to hand in financial statements on their defined form 10-Q, one example of a financial statement is the income statement. A company's financial statements for the first three fiscal quarters of the company's fiscal year must be handed in to SEC (*SEC* 2018).

Fundamentals

Basic financial data, such as the numbers described in a consolidated statement of income, for example, earnings per share (Snopek 2013).

Csv-file

Comma-separated value file. Can be used for as a simple storage of data in a row/column structure, i.e. a table.

Linear regression

To approximate a linear function y dependent on x based on a certain amount of plot points. That is, tuples on the form (x, y). The least square method can be used to do a linear regression. A linear regression can, in summary, be described as the way of approximating a line on the form $y = kx + m$ based on the provided plot points with the least possible error to reproduce the plot point values using a linear function (Anton and Rorres 2014).

PNratio

To make a set of sentiment scores expressed in a single number, a PN-ratio can be calculated as in the methodology of Asur and Huberman (2010). This is used in the linear regression. Below is the formula:

$$PNratio = \left| \frac{tweets\, with\, positive\; sentiment\; score}{tweets\, with\, negative\; sentiment\; score} \right|$$

Correlation

The correlation between two values can be measured by the Pearson correlation coefficient. It ranges from -1 to 1 where -1 denotes a total negative linear correlation and 1 denotes a total positive linear correlation. If the value is 0, there is no correlation. The formula for Pearson correlation coefficient is described below:

$$r = \frac{\sum_{i=1}^{n} (x_i - \overline{x})(y_i - \overline{y})}{\sqrt{\sum_{i=1}^{n} (x_i - \overline{x})^2}\sqrt{\sum_{i=1}^{n} (y_i - \overline{y})^2}}$$

(Lee Rodgers and Nicewander 1988)

Chapter 3

Method

This section initially describes how we executed our study in order to obtain our results. We will then motivate why we chose to proceed this way. We will finally discuss problems or weaknesses in our methodology.

3.1 Execution

We have two different approaches for analyzing sentiment and a company's fundamentals. One investigates how the sentiment is affected by the quarterly reports result over a longer period of time and one investigates how it is affected closer around the time the report was released. They are both based on the same data obtained from the quarterly reports and Twitter.

3.1.1 Collect data from quarterly reports

The fundamentals of each company were retrieved manually from their quarterly reports, which can be found on the company's website on the form Q-10. From each report, the net income and revenues for that quarter were used as a measure of the company's results. These data points were found in the consolidated statements of income in the Q-10 report. In order to be able to compare companies of different size and assets, the relative increase or decrease in revenues and net income was calculated:

$$relative\,net\,income = \frac{net\,income\,this\,year - net\,income\,last\,year}{net\,income\,last\,year}$$

$$relative\,revenues = \frac{revenues\,this\,year - revenues\,last\,year}{revenues\,last\,year}$$

3.1.2 Collect data from twitter

Data were retrieved from Twitter from a period ranging from 2014 to 2018. The used search word was the name of the company. For more details look at table 3.1. The final datasets obtained for each company was stored in an individual csv-file.

Table 3.1: Input parameters to Twitter API

Company name	**arg to parameter: track**
Wells fargo	wells fargo
Hasbro inc	hasbro inc
Activision Blizzard	activision blizzard
EOG Resources	eog Resources
Herbalife	herbalife

3.1.3 Long-term analysis

The long-term analysis used three companies: Activision Blizzard, Eog Resources and Hasbro Inc.

The obtained Twitter messages were used as input in the lexicon based algorithm VADER which calculated the sentiment and gave each message a positive, a negative, a neutral and a compounded sentiment score as described in section 2.2.6. The compounded sentiment was used as sentiment score for each message and for each day the average compounded sentiment was calculated. See listing 3.1 for the calculation of the average compounded sentiment for each day.

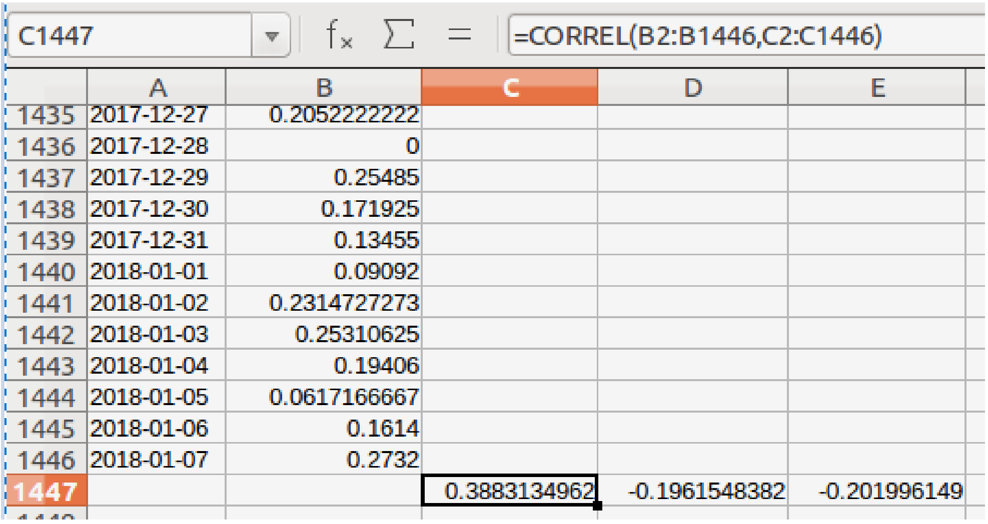
Listing 3.1: Code for VADER

```
#Calculates the average per day
score_avg_sentence_per_day = [] #date is on format
   'Y-m-d'
for key, group in groupby(dates_and_sentiments, lambda
   x: str(x[0].strftime('%Y-%m-%d'))):
   score_sentence_per_day = []
   for thing in group:
      score_sentence_per_day.append(thing[1])
```

```
    avg = sum(score_sentence_per_day) /
        len(score_sentence_per_day)
    score_avg_sentence_per_day.append((key, avg))
```

The output of the sentiment analysis was returned as a tuple on the form: (date, sentiment score for the day). The tuples were plotted into a graph together with the data from the quarterly reports, using GNUPlot.

For each company, the Pearson Product-Moment Correlation Coefficient was calculated using Excel's correlation formula, with respect to the sentiment score for each day and the financial key factors associated to a few specific days a year.



C1447 =CORREL(B2:B1446,C2:C1446)

	A	B	C	D	E
1435	2017-12-27	0.2052222222			
1436	2017-12-28	0			
1437	2017-12-29	0.25485			
1438	2017-12-30	0.171925			
1439	2017-12-31	0.13455			
1440	2018-01-01	0.09092			
1441	2018-01-02	0.2314727273			
1442	2018-01-03	0.25310625			
1443	2018-01-04	0.19406			
1444	2018-01-05	0.0617166667			
1445	2018-01-06	0.1614			
1446	2018-01-07	0.2732			
1447			0.3883134962	-0.1961548382	-0.201996149

Figure 3.1: Correlation calculus using excel, column A is the date, B its sentiment score, C, D, E are financial performance measures

3.1.4 Short-term analysis

For the short-term analysis, all companies were chosen.

The PNratio (see section 2.3) was calculated using all of the sentiment scores starting at the time of the release of a company's quarterly report and ending two days later. The sentiment score was calculated using the lexicon based algorithm VADER and gave each message a positive, a negative, a neutral and a compounded sentiment score as described in section 2.2.6.

A linear regression was made in excel using the change in the net income (as described in section 3.1.1) as a function based on PNratio.

The compound sentiment score was also plotted in a graph using GNUPlot.

3.2 Choice of method

The choice of sentiment analyzer and the method for comparing data will be motivated in this section together with problems that the choice of method might cause.

3.2.1 Data collection

Related studies, such as the ones conducted by Bollen, Mao, and Zeng (2011) and Jansen et al. (2009), have been using Twitter as a data source to conduct sentiment analysis related to finance. Therefore, Twitter should be a good data source for this study. Revenues and net income were the chosen data points from the financial statement. The reasoning behind this was that revenue is the total amount of value (measured in a currency) that a company can generate. Moving on to net income. This is the final result. As in the actual value generated for the owners (shareholders) of a company. We believe that out of two data points chosen these are the most central ones, revenues can be interpreted as what quantity of value we are talking about. Net income clarifies how efficient this value is generated for its owners.

3.2.2 Sentiment analysis process

Instead of using a lexicon based algorithm we could have used an algorithm based on machine learning. However, we could not find an algorithm that was trained on the suitable dataset. Additionally, the shortage of time made it impossible to create a dataset from scratch in order to train a learning-based algorithm.

Despite Pak and Paroubek (2010) suggestion of using POS-tagging as a way to get rid of neutral messages and obtain the messages actually containing a positive or negative sentiment value, we believe that based on Kouloumpis, Wilson, and Moore (2011) findings and the fact that VADER is designed for short social media messages, the accuracy gain would be too little to impact our findings.

3.2.3 Linear regression

Based on Asur and Huberman (2010) study, described in 2.1.1, an approach where linear regression is used to analyze the data appears reasonable. Calculating the PNratio and comparing it to a financial performance unit is exactly what this study aims to do and using linear regression makes it possible to base our study on a statistical method rather than subjective judgments.

3.2.4 Correlation

We have not found any other report that uses Pearson's correlation coefficient in order to compare sentiment in social media to other data. However, Pearson's correlation coefficient has been used in various applications where the goal is to find correlation (Lee Rodgers and Nicewander 1988) and our study mainly relies on finding a correlation by visually reviewing the graphs. The correlation coefficient is used to support the result or, if it is not in line with the graphs, indicate that something is wrong.

3.2.5 Chosen companies

In order to retrieve equivalent financial data all the chosen companies are publicly traded companies. It is also a diverse set of companies, operating in different markets which ensures a broad set of data.

Wells Fargo

Wells Fargo is one of the largest banks in the US and was founded in 1852 (Fargo 2018).

Eog Resources

Eog resources is one of the largest independent crude oil and natural gas companies in the United States and is operating in United States, Trinidad, the United Kingdom and China (EOG Resources 2018).

Activision Blizzard

Activision Blizzard is an interactive entertainment company that for example has produced World of Warcraft, Hearthstone and Candy

Crush Saga (Blizzard 2018).

Hasbro Inc

Hasbro is a play and entertainment company that mainly produces toys and games (Hasbro 2018).

Herbalife

Herbalife is a global nutrition company with a focus on creating healthy products (Herbalife 2018).

3.2.6 Critque and problems

There are limitations in the choice of companies to analyze in regards to crowd-source-based sentiment analysis. For crowd-source-based sentiment analysis, it has been found that Twitter provides the greatest amount of data. Therefore, companies not frequently mentioned on Twitter could not be included. Furthermore, a company was required to be filing its reports to SEC, that is a limitation to American companies obligated to file their income statements to SEC. In addition, due to the massive amount of tweets posted for certain companies on a two year period, some companies had to be excluded from the long-term perspective, because of limitations in obtaining a complete data set.

Chapter 4

Result

This section presents three types of graphs. The first type, the long-term graph, will show the average sentiment compound score obtained for a company for every five days. The days where the financial reports were presented have been marked on the graph with two dots. One of the dots indicates whether the revenue has increased or decreased from the previous financial recordings, the other dot represents the profit in a similar manner. There will also be a table of the correlation that was computed between the sentiment score and the changes in revenues and net income.

The second graph type, the short-term graph, provides the sentiment compound score over the time around the report. The third type will show the linear regression between PNratio and net income that was made for the short-term analysis.

4.1 Long-term graphs

These graphs show the average sentiments for a company over a longer period. The sentiment is calculated as the average sentiment over 5 days.

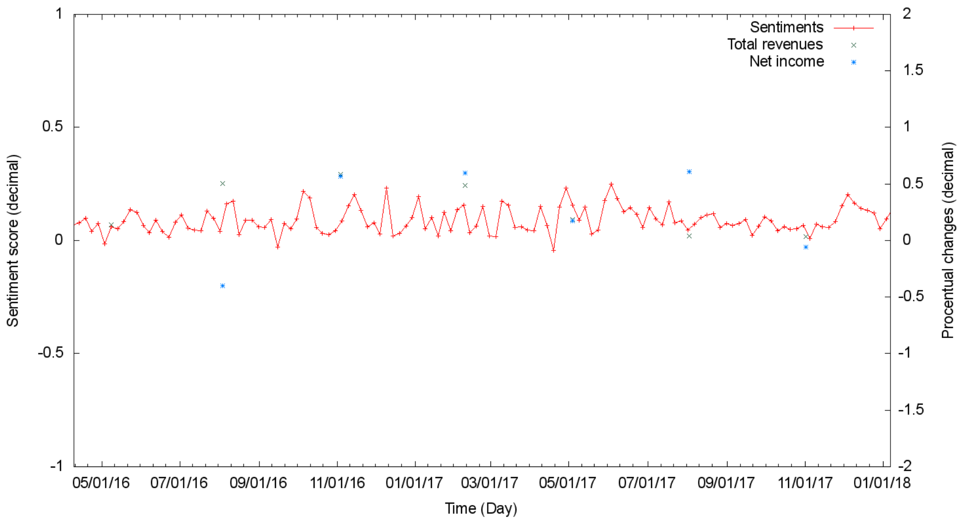


Figure 4.1: Sentiment compound score of Activision Blizzard. The left vertical axis is the sentiment compound score. The horizontal axis describes the time, as a discrete set of days. The right vertical axis describes the increase or decrease of a data point found in the financial statement

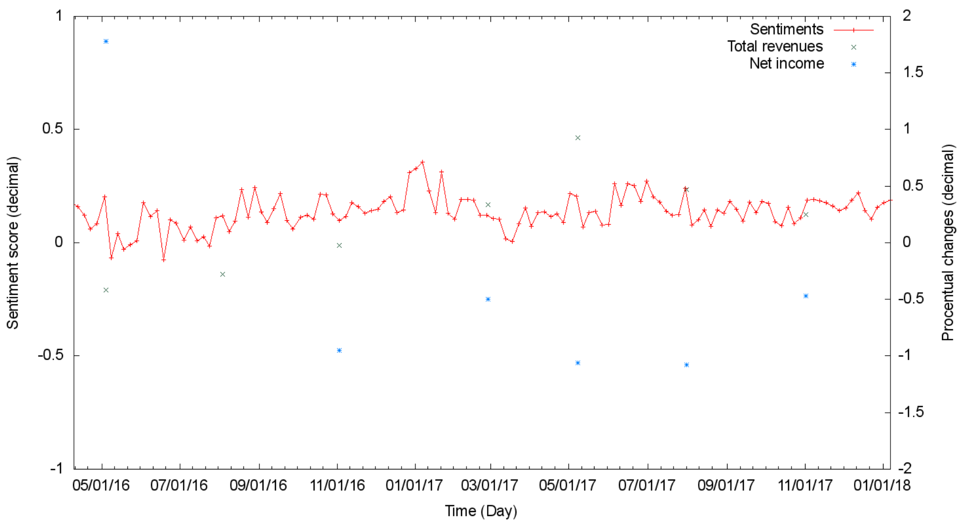


Figure 4.2: Sentiment compound score of EOG Resources. The left vertical axis is the sentiment compound score. The horizontal axis describes the time, as a discrete set of days. The right vertical axis describes the increase or decrease of a data point found in the financial statement

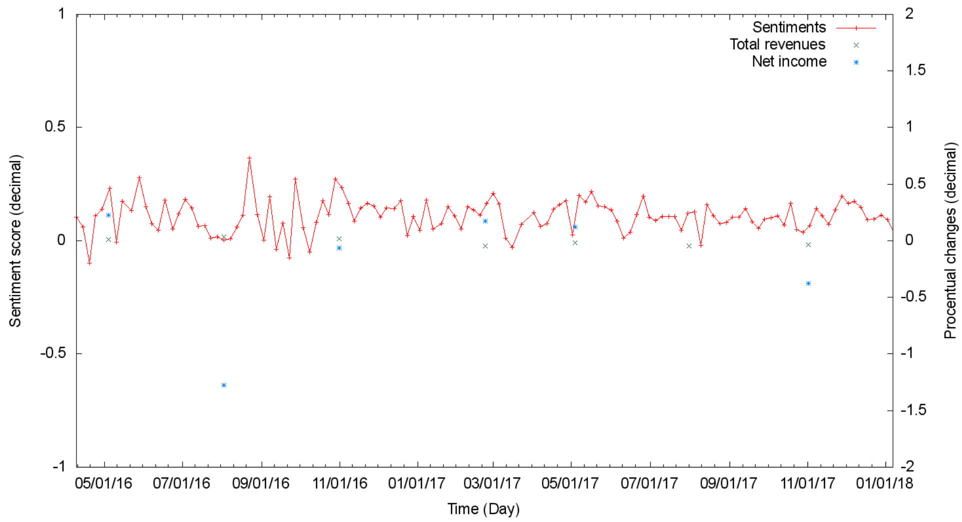


Figure 4.3: Sentiment compound score of Hasbro Inc. The left vertical axis is the sentiment compound score. The horizontal axis describes the time, as a discrete set of days. The right vertical axis describes the increase or decrease of a data point found in the financial statement

4.1.1 Correlation

This table shows the correlation between the sentiment and the net income or revenues over a period of two years.

Company name	**Sentiment and total revenues**	**Sentiment and net income**
Activision Blizzard	-0.030590567	-0.3620436334
Eog Resources	0.3114888892	-0.2114620952
Hasbro Inc	0.0322633732	0.1618314487

Table 4.1: Computed correlation between sentiment compound score and data points found in the financial statement

4.2 Short-term graphs

These graphs show some examples of the sentiments for a company two days before the release of the report and two days after. All graphs can be found in Appendix A.

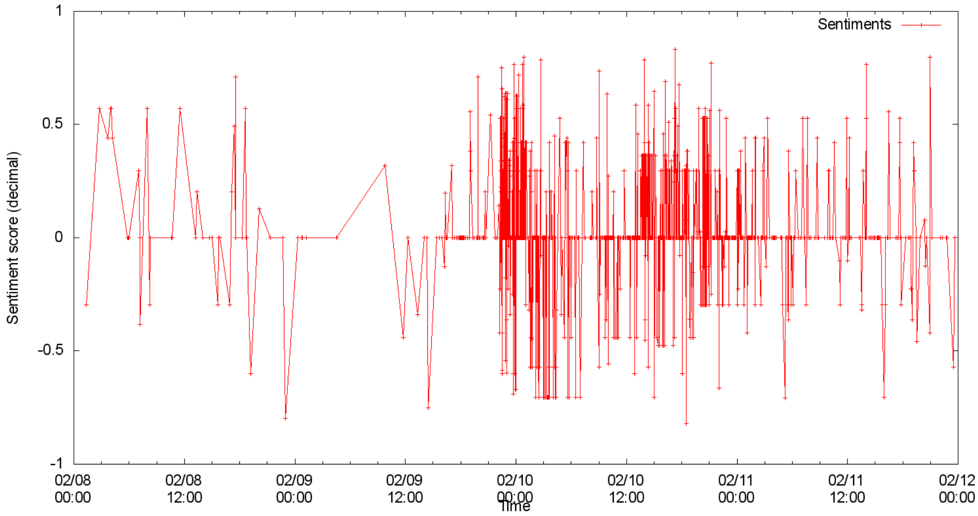


Figure 4.4: Sentiment compound score of Activision Blizzard at the release of Q4 2016. The vertical axis is the sentiment compound score. The horizontal axis describes the time as a discrete set of seconds.

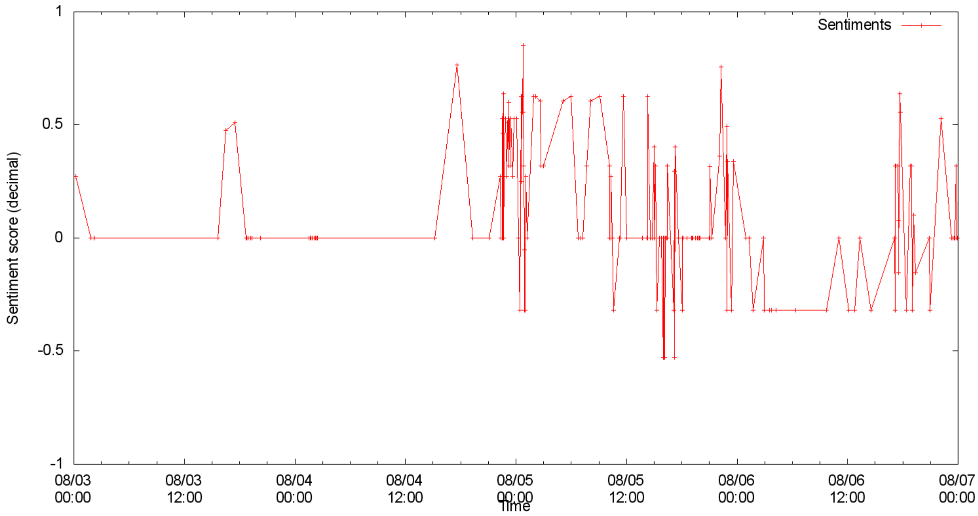


Figure 4.5: Sentiment compound score of EOG resources at the release of Q2 2016. The vertical axis is the sentiment compound score. The horizontal axis describes the time as a discrete set of seconds.

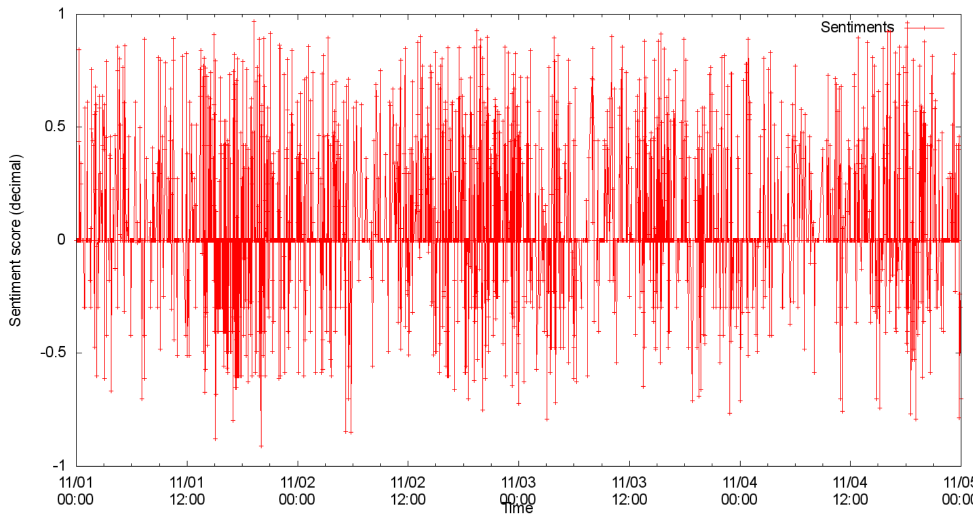


Figure 4.6: Sentiment compound score of Herbalife at the release of Q3 2017. The vertical axis is the sentiment compound score. The horizontal axis describes the time as a discrete set of seconds.

4.2.1 Linear regression

Below, graphs are presented where the change in net income is expressed as a function based on PNratio using linear regression. All graphs can be found in Appendix B.

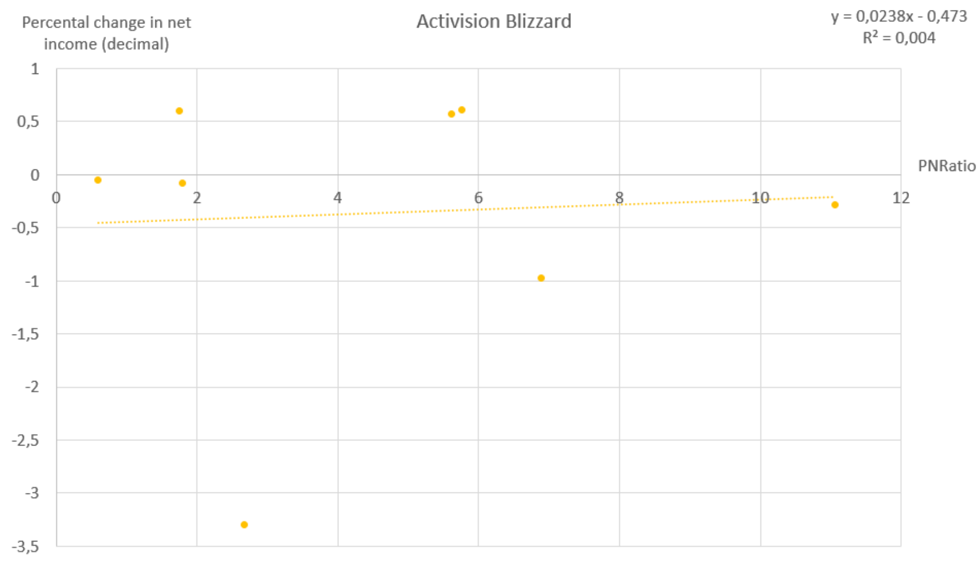


Figure 4.7: Linear regression of Activision Blizzard. The vertical axis is the change in net income for a quarterly report compared to its previous years quarterly. The horizontal axis describes the PNratio based on formula in section 3.1.4

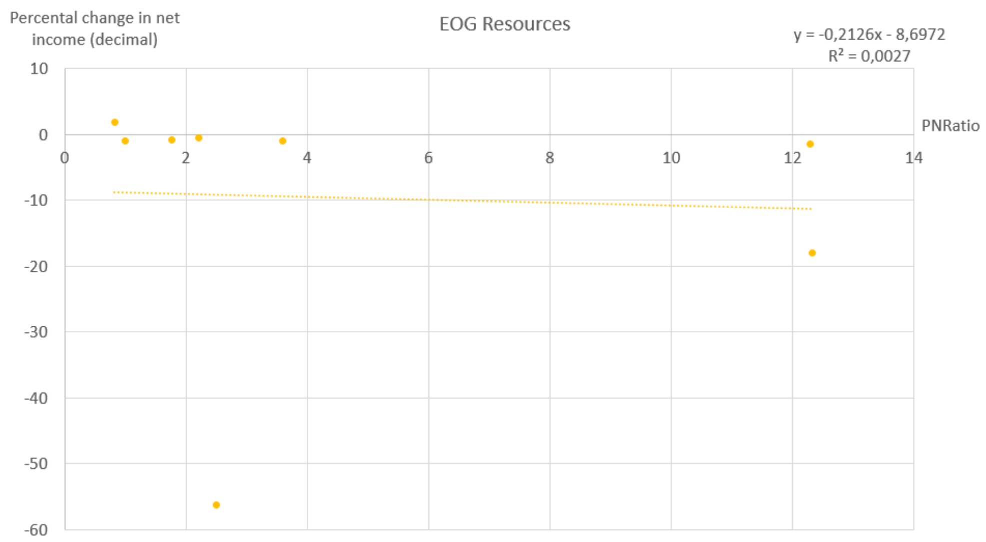


Figure 4.8: Linear regression of EOG Resources. The vertical axis is the change in net income for a quarterly report compared to its previous years quarter. The horizontal axis describes the PNratio based on formula in section 3.1.4

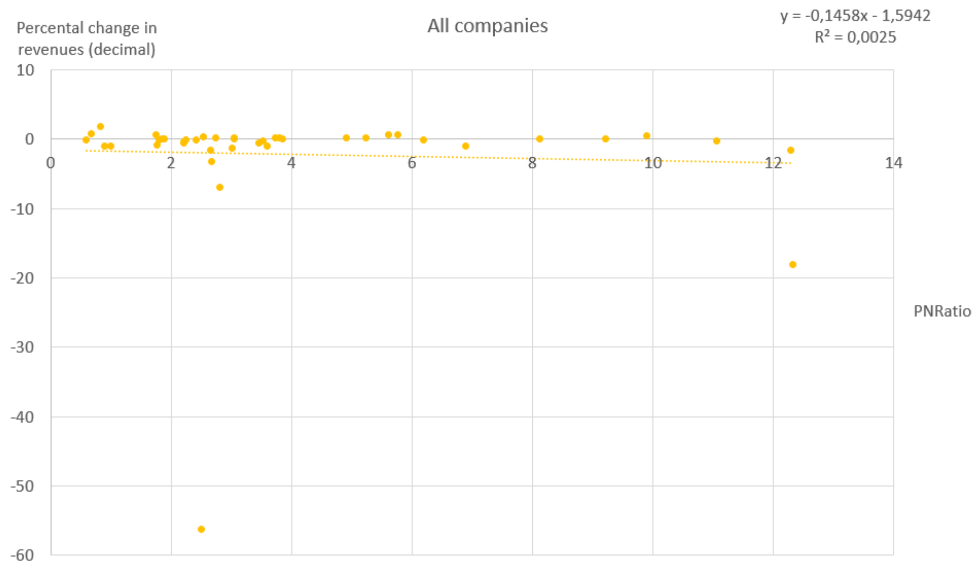


Figure 4.9: Linear regression of Activision blizzard, EOG Resources, Hasbro Inc, Herbalife and Wells fargo. The vertical axis is the change in net income for a quarterly report compared to its previous years quarter. The horizontal axis describes the PNratio based on formula in section 3.1.4

Chapter 5

Discussion

First, the results are analyzed in a short-and long-term perspective. Then, sources of errors will be discussed and we will propose what future studies could be about.

5.1 Analysis of the results

As shown above, we looked at sentiment over a time interval of both two years and just a few days. Analyzing the result from the two years interval indicated that the sentiment on Twitter is not affected by the fundamentals of a company over a longer period of time, which made us focus on the sentiment of tweets closer to a company's release of a financial report.

5.1.1 Long term

Over a longer period of time, we could not find any correlation between the sentiment and the company's fundamentals. Even though the correlation coefficient sometimes indicates that there could be a connection, it varies too much and is based on a too small amount of data to be trustworthy. This is also shown in the graphs where it is clear that the sentiment does not change to a more positive or negative score over time, it rather fluctuates over a shorter period of time.

This might indicate that there are too much noise in the data. In that case, it might be possible to find a correlation between tweets and the quarterly reports, but it would require that the tweets are selected more carefully. For example, it might be possible to find a correlation if only tweets that mentions the company in a financial context are taken into account. Further studies is needed in order to investigate this. It might also be the case that there is no correlation, regardless of how the tweets are selected, simply because the perception of a company mainly is influenced by other factors and more or less independent of the company's fundamentals.

5.1.2 Short term

When viewing some of the short-term graphs, it is clearly indicated at what time the report was released, based on the frequency of messages being more intense (figure 4.4 and 4.5). When we looked at the text of the tweets around that time we could indeed see that there were tweets related to the release of a report. This is aligned with the statement of Bollen, Mao, and Zeng (2011) that financial discussions are news driven. However, there was no visible indication of a change in the sentiment score after a report. This is clearly seen in Figure 4.6 (and the graphs in Appendix A).

The PNratio reinforces this idea. It has a large variation between the releases of the reports but does not seem to have anything to do with how well the company is doing. For example, there were times when the ratio was above one even though the net income had decreased and other times when the ratio was less than one even though the net income had increased. No relation between net income or revenues and PNratio could be found. Our linear regression supports this statement due to the fact that the inclination was close to zero for all companies.

Hence, the short-term analysis supports the idea that the sentiment in the tweets does not tell anything about the financial results. Furthermore, the results indicate that the sentiment of a company could explain expectations rather than reactions on the fundamentals of a company. To clarify, looking at Figure 4.9. Negative change in net income sometimes end up with a high PNRatio. This could be explained by news and other previous information that have made the public expect a much higher decrease than actually was the cause. Then the reaction is positive, since they expected worse. As a result the PNRatio gets positive although the percental change of net income decreased. Finally, if the sentiment of a company is based on expectations it would infer that no correlation exists between the reports and the tweets regardless of how the tweets are selected.

Compared with Asur and Huberman (2010), our study does not indicate a direct impact on revenue or net income. This infers that sentiment analysis is a useful tool for understanding one's customers rather than a key factor for boosted financial performance, and is in line with their report since their method is used rather for the release of a new product then on the release of financial news. This is also strengthened by figure 4.6 where we cannot see any indication of the release of the report. This is probably due to the fact that, for this company, people write more about the brand name and its products than the actual company. Meaning that if there are any tweets that mention the report they are not noticeable among all the other tweets. As a result, narrowing the tweets down to focusing on a brand name or a specific product would probably infer higher accuracy in tweets speaking about one particular topic.

5.1.3 Sources of error

There are a few factors that might have influenced the result of our study.

An important source of error is the small number of companies that have been conducted, particularly in the long-term analysis. The study would require more data to be sure that no correlation could be found. Hence, this study suffers from statistical insignificance. It might also have been better to concentrate on a specific business in order to limit the scope. If the sentiment scoring over a period of time is different in different businesses, there might exist a pattern that we could not see due to the fact that all our companies come from different industries.

Only looking at a few data points found in the financial statement could be another source of error. Factors other than net income and revenues affect the reaction to a company's financial performance and it is reasonable that more data points could have been used.

Moreover, the sentiment analyzer used, VADER, is very simple and might not be precise enough. VADER does for example not use any POS-tagging (section 2.2.1) and can therefore not make sure that the tweet is actually about the company. However, as described in section 2.2.6, this might not be needed since the tweets can be treated on a sentence level. We believe that VADER is precise enough to show some correlation if there are any since research has shown even such simple solution has achieved good accuracy (Feldman 2013) and the accuracy was tested when VADER was created (Gilbert 2014).

VADER does not take the financial context into account either. A sentence that for example is considered neutral in other contexts might have a positive or negative sentiment in a financial context. For example "buy company A" would probably imply a positive perception of the company. Since VADER is made for social media in general, and not social media in a financial context, this would not be noted. Regardless, we believe that the context has a too small impact on the sentiment to make a notable impression on the result in this case.

5.2 Conclusion

The quarterly report appears to not affect the sentiment on tweets over a longer period but we cannot exclude that such a correlation would exist solely based on the long-term analysis. It would require a refined study where a greater amount of data was selected from Twitter more carefully in order to only contain relevant data. However, the short-term analysis indicates that the sentiment is not at all influenced by the fundamentals and that these two are not correlated at all. Indeed, our results from the short-term analysis suggest that a refined study would come to the same conclusion as ours.

In summary, our study indicates that sentiment for a company is based on expectations rather than results. It appears that sentiment on its own, cannot be used to describe the previous financial performance of a company on a longer term, for example on one year period, due to the periodic behavior of its sentiment.

5.3 Future studies

As mentioned in section 5.1.3, the study might have been more trustworthy the research focuses on companies in a specific business. Hence, a future study might consider concentrating on a separate industry. Furthermore, it might be even better to concentrate on companies strongly dependent on one product. Since it might be a way of better understanding a company's financial performance based on sentiment. This is based on two observations. Firstly, our conclusion regarding that sentiment seems to be related to expectations rather than fundamentals. This suggests that the context of the sentiment must be about a specific revenue generator rather than a company name/brand. Secondly, the study of Asur and Huberman (2010) did carry out a similar linear regression as in this study but on a specific product. These two observations might indicate that it is sensible to conduct studies on companies strongly dependent on a specific product, since then the company's financial results are strongly dependent on that particular product.

In addition, it appears sensible to perform a similar study with a more complex linear regression, as Asur and Huberman (ibid.) showed that the sentiment analysis did not have a great deal of impact on their prediction of box-office revenues. Suggestions are to consider the stock prices before the release of a report and include more data points found in the financial statement.

A future study would also benefit from coming up with a refined method for finding the correlation between the sentiment and the fundamentals in the long-term analysis. We believe that the correlation coefficient needs more data to be reliable and that it would be useful to have a more objective solution than just viewing the graphs.

Bibliography

Anton, H. and C. Rorres (2014). *Elementary Linear Algebra, Applications Version 11E with WileyPlus Card*. John Wiley & Sons, Incorporated. ISBN: 9781118938881. URL: https://books.google.se/books?id=WEuloAEACAAJ.

Asur, Sitaram and Bernardo A. Huberman (2010). "Predicting the Future with Social Media". In: *Proceedings of the 2010 IEEE/WIC/ACM International Conference on Web Intelligence and Intelligent Agent Technology - Volume 01*. WI-IAT '10. Washington, DC, USA: IEEE Computer Society, pp. 492–499. ISBN: 978-0-7695-4191-4. DOI: 10.1109/WI-IAT.2010.63. URL: http://dx.doi.org/10.1109/WI-IAT.2010.63.

Blizzard, Activision (2018). *Our Company*. URL: https://www.activisionblizzard.com/about-us (visited on 05/09/2018).

Bollen, Johan, Huina Mao, and Xiaojun Zeng (2011). "Twitter mood predicts the stock market". In: *Journal of Computational Science* 2.1, pp. 1–8. ISSN: 1877-7503. DOI: https://doi.org/10.1016/j.jocs.2010.12.007. URL: http://www.sciencedirect.com/science/article/pii/S187775031100007X.

Cambria, Erik, Dipankar Das, et al. (2017). *A practical guide to sentiment analysis*. Vol. 5. Springer.

Cambria, Erik, Björn Schuller, et al. (2013). "New avenues in opinion mining and sentiment analysis". In: *IEEE Intelligent Systems* 28.2, pp. 15–21.

Duan, Wenjing, Bin Gu, and Andrew B. Whinston (2008). "Do online reviews matter? — An empirical investigation of panel data". In: *Decision Support Systems* 45.4. Information Technology and Systems in the Internet-Era, pp. 1007–1016. ISSN: 0167-9236. DOI: https://doi.org/10.1016/j.dss.2008.04.001. URL: http://www.sciencedirect.com/science/article/pii/S0167923608000754.

Engwall, Mats (2017). *Modern industriell ekonomi*. City: Studentlitteratur AB. ISBN: 9789144116914.

EOG Resources, Inc. (2018). *Overview*. URL: http://www.eogresources.com/about/overview.html (visited on 05/09/2018).

Fargo, Wells (2018). *History of Wells Fargo*. URL: https://www.wellsfargo.com/about/corporate/history/ (visited on 05/09/2018).

Feldman, Ronen (2013). "Techniques and applications for sentiment analysis". In: *Communications of the ACM* 56.4, pp. 82–89.

Gilbert, CJ Hutto Eric (2014). "Vader: A parsimonious rule-based model for sentiment analysis of social media text". In: *Eighth International Conference on Weblogs and Social Media (ICWSM-14). Available at (20/04/16) http://comp. social. gatech. edu/papers/icwsm14. vader. hutto. pdf*.

Hasbro (2018). *About Hasbro*. URL: https://hasbro.gcs-web.com/corporate/ (visited on 05/14/2018).

Herbalife (2018). *About Herbalife*. URL: http://company.herbalife.com/ (visited on 05/14/2018).

Jansen, Bernard J. et al. (2009). "Micro-blogging As Online Word of Mouth Branding". In: *CHI '09 Extended Abstracts on Human Factors in Computing Systems*. CHI EA '09. Boston, MA, USA: ACM, pp. 3859–3864. ISBN: 978-1-60558-247-4. DOI: 10.1145/1520340.1520584. URL: http://doi.acm.org/10.1145/1520340.1520584.

Kouloumpis, Efthymios, Theresa Wilson, and Johanna D Moore (2011). "Twitter sentiment analysis: The good the bad and the omg!" In: *Icwsm* 11.538-541, p. 164.

Lee Rodgers, Joseph and W Alan Nicewander (1988). "Thirteen ways to look at the correlation coefficient". In: *The American Statistician* 42.1, pp. 59–66.

Liu, Bing (2012). "Sentiment analysis and opinion mining". In: *Synthesis lectures on human language technologies* 5.1, pp. 1–167.

Maynard, Diana (2017). *Natural language processing for the semantic web*. eng. Synthesis lectures on the semantic web, theory and technology ; Number 15. ISBN: 1-62705-632-7.

Pak, Alexander and Patrick Paroubek (2010). "Twitter as a corpus for sentiment analysis and opinion mining." In: *LREc*. Vol. 10. 2010.

Pang, Bo and Lillian Lee (2008). "Opinion Mining and Sentiment Analysis". In: *Foundations and Trends® in Information Retrieval* 2.1–2, pp. 1–135. ISSN: 1554-0669. DOI: 10.1561/1500000011. URL: http://dx.doi.org/10.1561/1500000011.

SEC (2018). www.sec.gov/reportspubs/investor-publications/investorpubsbegfinstmtguidehtm.html. Accessed: 2018-04-18.

SEC (2018). https://www.sec.gov/fast-answers/answersform10qhtm.html. Accessed: 2018-04-18.

Snopek, Lukasz (2013). "Benjamin Graham's Approach". eng. In: pp. 133–136.

Ticknor, Jonathan L. (2013). "A Bayesian regularized artificial neural network for stock market forecasting". In: *Expert Systems with Applications* 40.14, pp. 5501–5506. ISSN: 0957-4174. DOI: https://doi.org/10.1016/j.eswa.2013.04.013. URL: http://www.sciencedirect.com/science/article/pii/S0957417413002509.

Twitter, Terms of Service (2018). https://twitter.com/en/tos. Accessed: 2018-02-18.

Yu, Yang, Wenjing Duan, and Qing Cao (2013). "The impact of social and conventional media on firm equity value: A sentiment analysis approach". In: *Decision Support Systems* 55.4. 1. Social Media Research and Applications 2. Theory and Applications of Social Networks, pp. 919–926. ISSN: 0167-9236. DOI: https://doi.org/10.1016/j.dss.2012.12.028. URL: http://www.sciencedirect.com/science/article/pii/S0167923612003946.

Appendix A

All short-term graphs

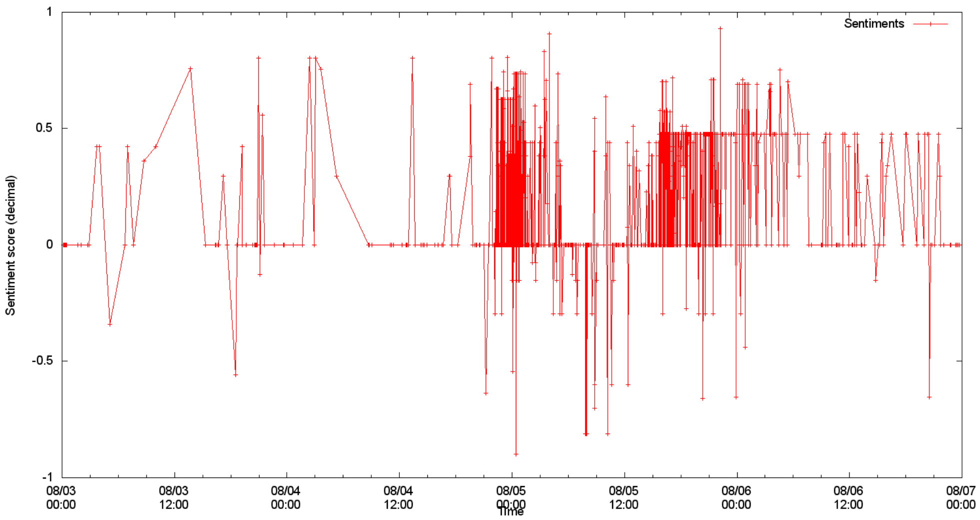


Figure A.1: Sentiment compound score of Activision Blizzard at the release of Q2 2016. The vertical axis is the sentiment compound score. The horizontal axis describes the time as a discrete set of seconds.

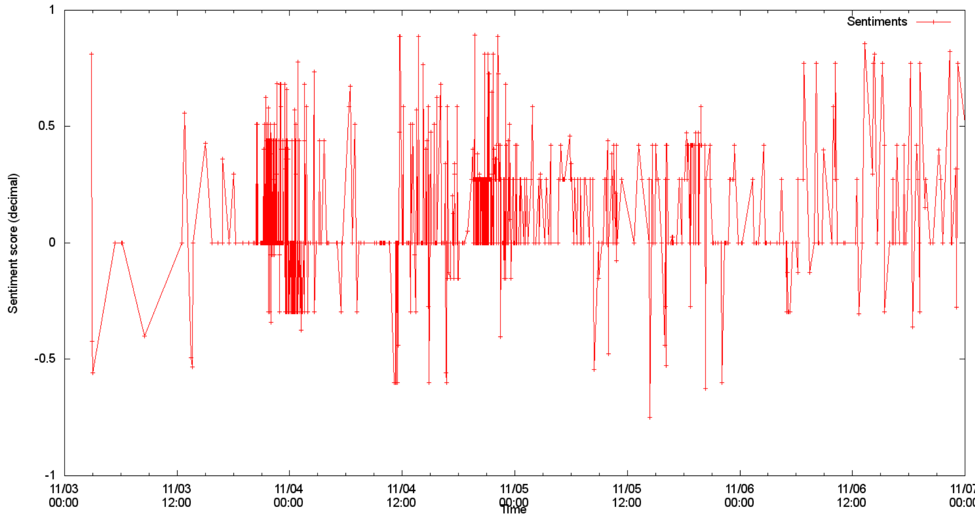


Figure A.2: Sentiment compound score of Activision Blizzard at the release of Q3 2016. The vertical axis is the sentiment compound score. The horizontal axis describes the time as a discrete set of seconds.

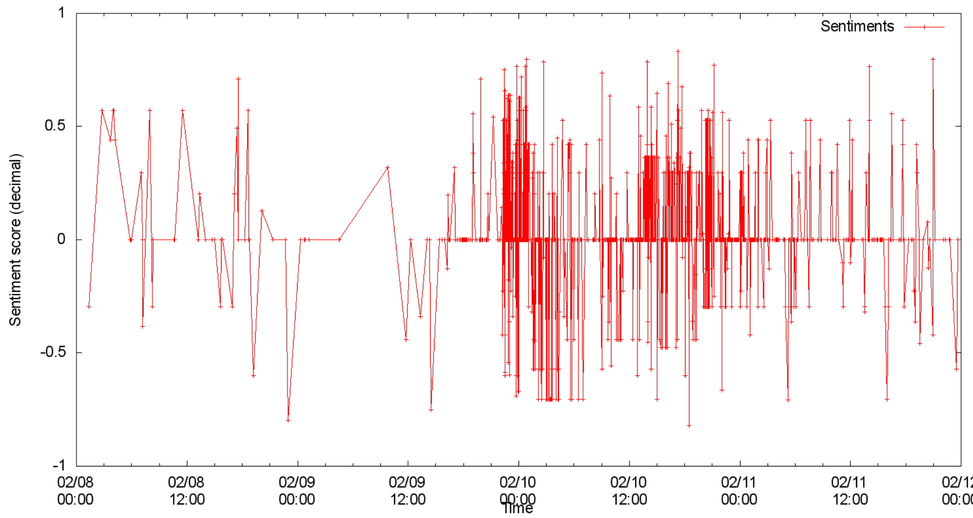


Figure A.3: Sentiment compound score of Activision Blizzard at the release of Q4 2016. The vertical axis is the sentiment compound score. The horizontal axis describes the time as a discrete set of seconds.

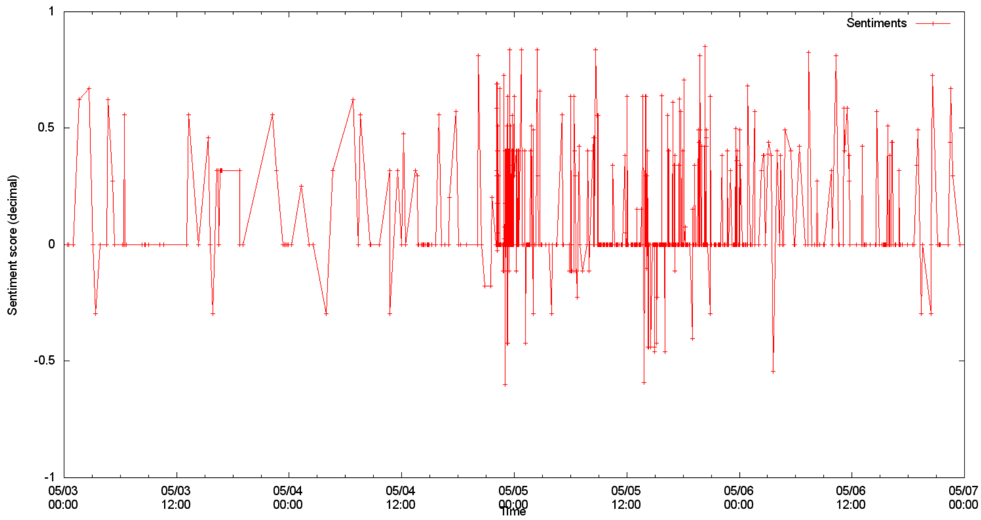


Figure A.4: Sentiment compound score of Activision Blizzard at the release of Q1 2017. The vertical axis is the sentiment compound score. The horizontal axis describes the time as a discrete set of seconds.

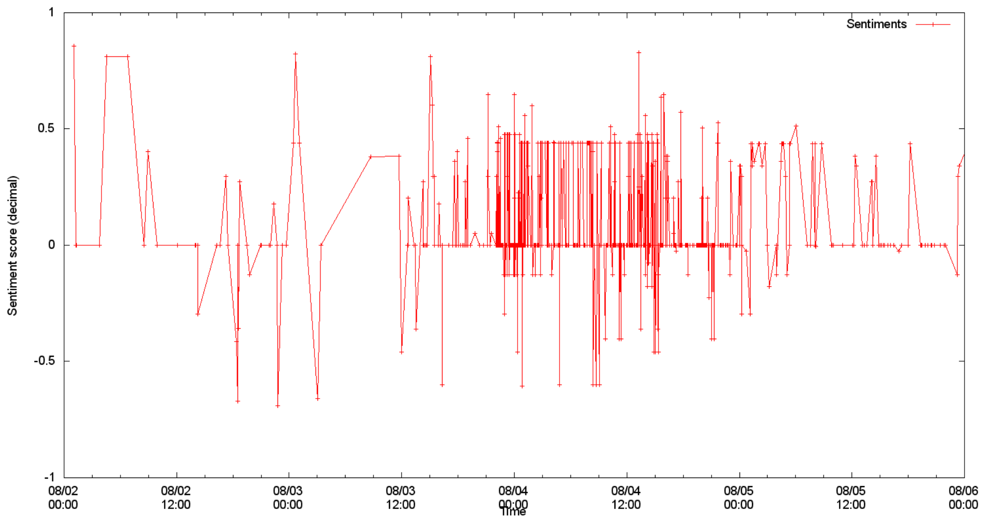


Figure A.5: Sentiment compound score of Activision Blizzard at the release of Q2 2017. The vertical axis is the sentiment compound score. The horizontal axis describes the time as a discrete set of seconds.

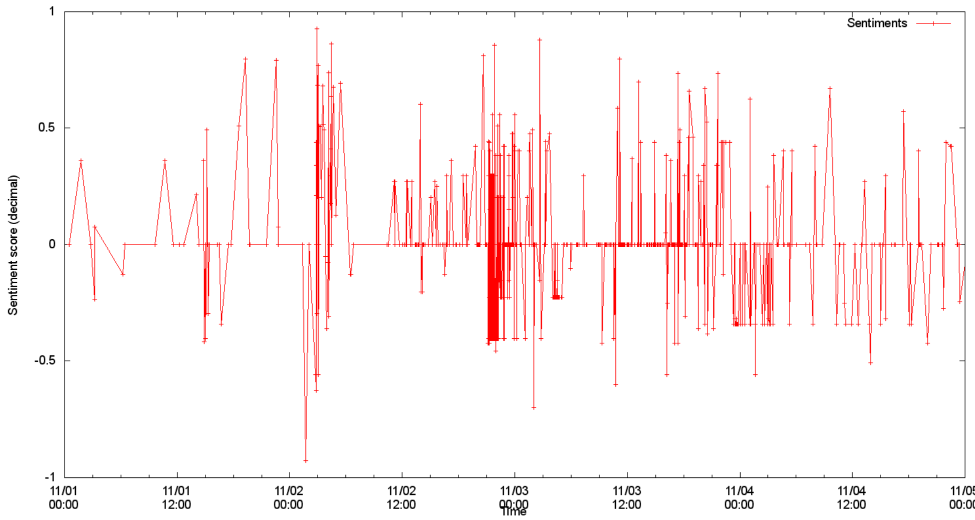


Figure A.6: Sentiment compound score of Activision Blizzard at the release of Q3 2017. The vertical axis is the sentiment compound score. The horizontal axis describes the time as a discrete set of seconds.

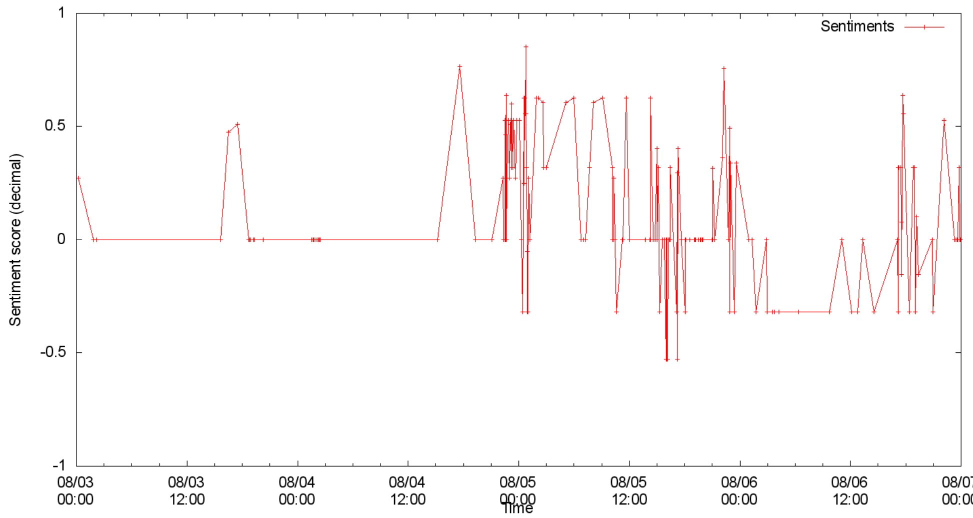


Figure A.7: Sentiment compound score of EOG resources at the release of Q2 2016. The vertical axis is the sentiment compound score. The horizontal axis describes the time as a discrete set of seconds.

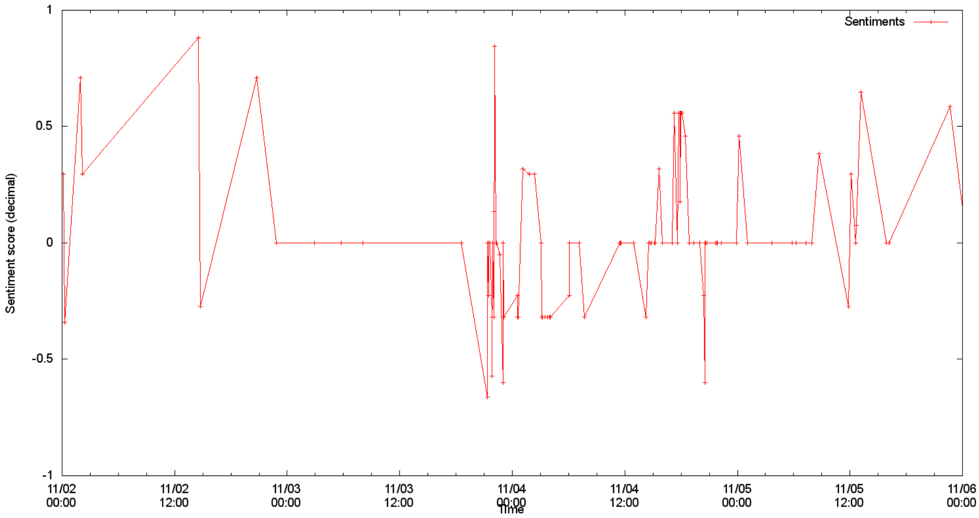


Figure A.8: Sentiment compound score of EOG resources at the release of Q3 2016. The vertical axis is the sentiment compound score. The horizontal axis describes the time as a discrete set of seconds.

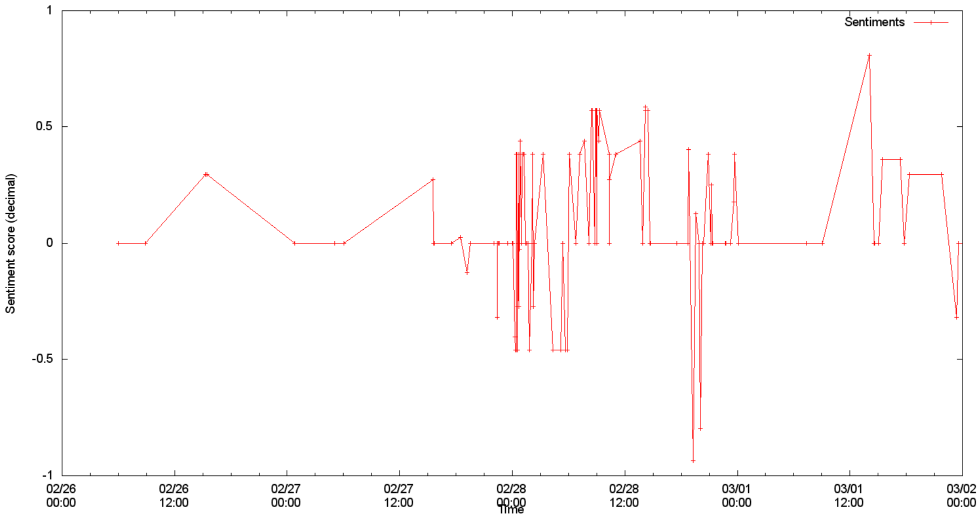


Figure A.9: Sentiment compound score of EOG resources at the release of Q4 2016. The vertical axis is the sentiment compound score. The horizontal axis describes the time as a discrete set of seconds.

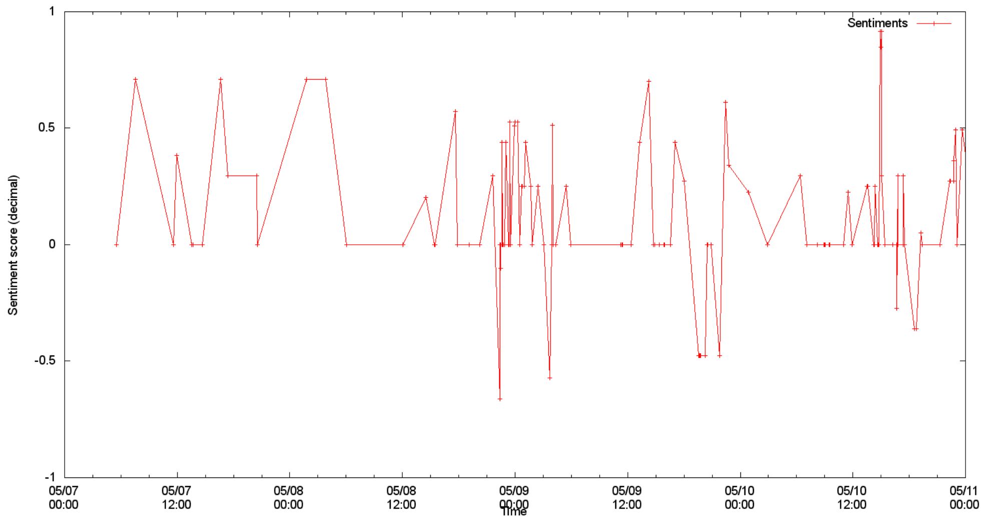


Figure A.10: Sentiment compound score of EOG resources at the release of Q1 2017. The vertical axis is the sentiment compound score. The horizontal axis describes the time as a discrete set of seconds.

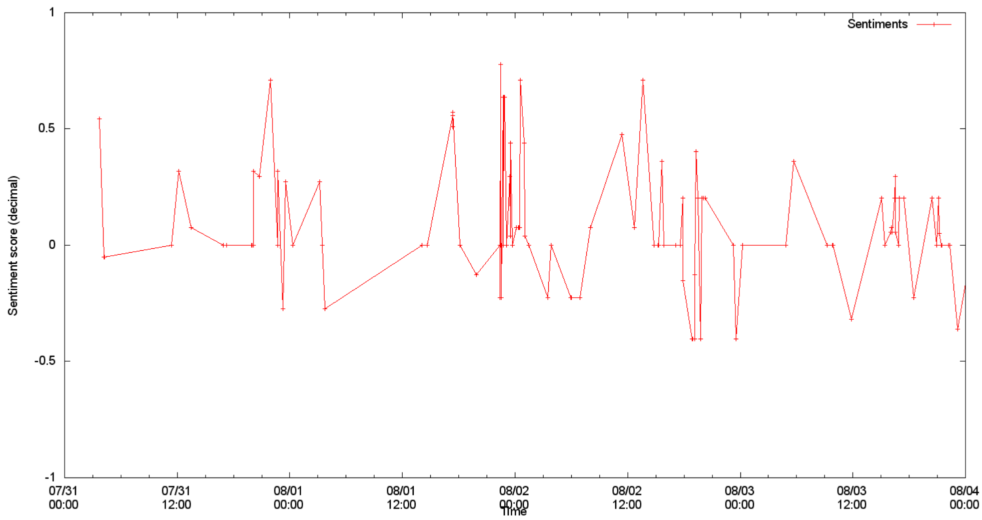


Figure A.11: Sentiment compound score of EOG resources at the release of Q2 2017. The vertical axis is the sentiment compound score. The horizontal axis describes the time as a discrete set of seconds.

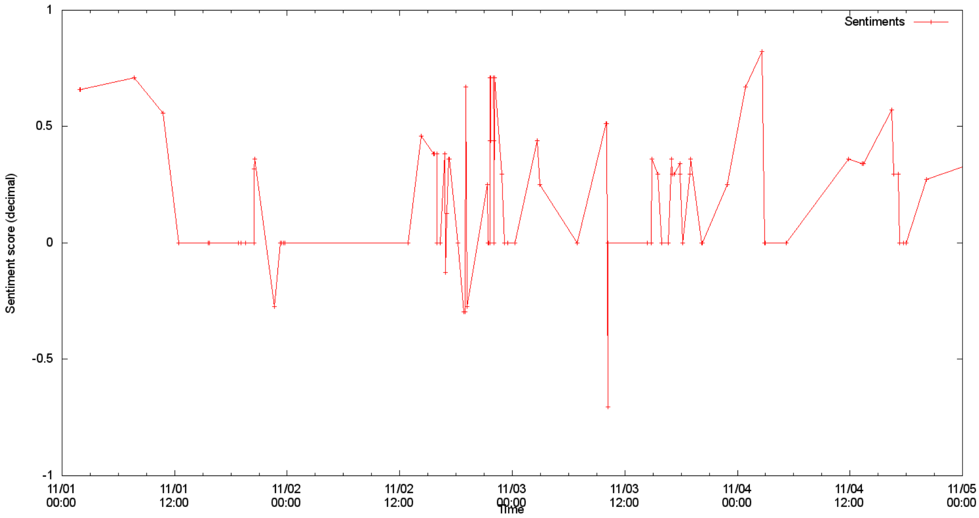


Figure A.12: Sentiment compound score of EOG resources at the release of Q3 2017. The vertical axis is the sentiment compound score. The horizontal axis describes the time as a discrete set of seconds.

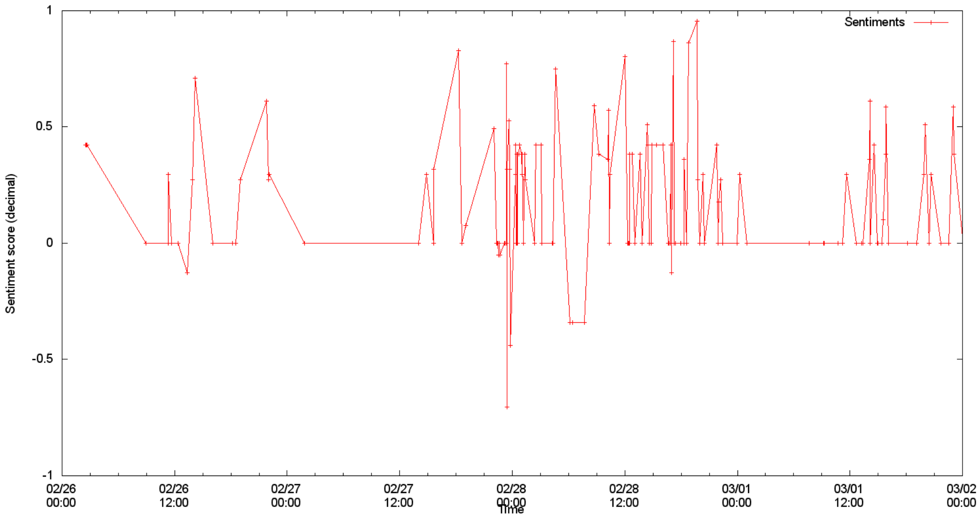


Figure A.13: Sentiment compound score of EOG resources at the release of Q4 2017. The vertical axis is the sentiment compound score. The horizontal axis describes the time as a discrete set of seconds.

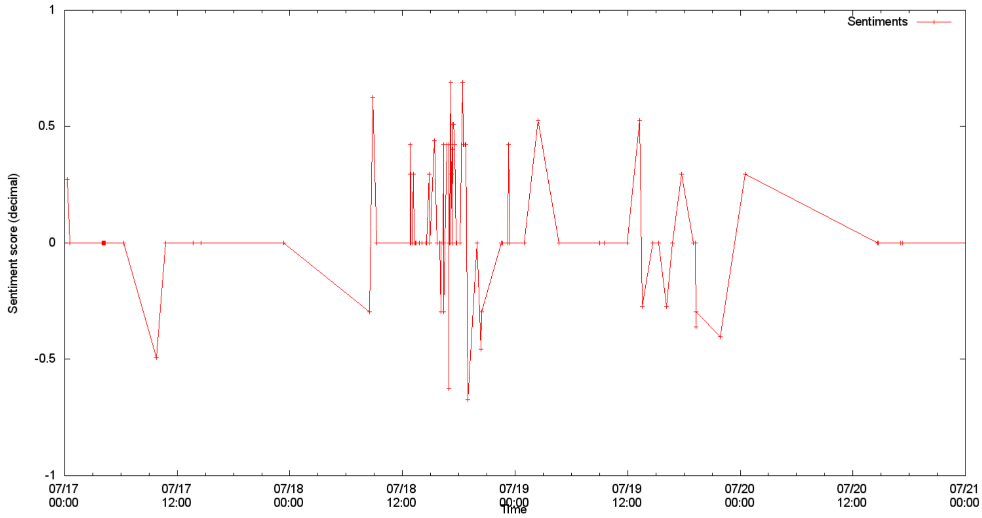


Figure A.14: Sentiment compound score of Hasbro inc at the release of Q2 2016. The vertical axis is the sentiment compound score. The horizontal axis describes the time as a discrete set of seconds.

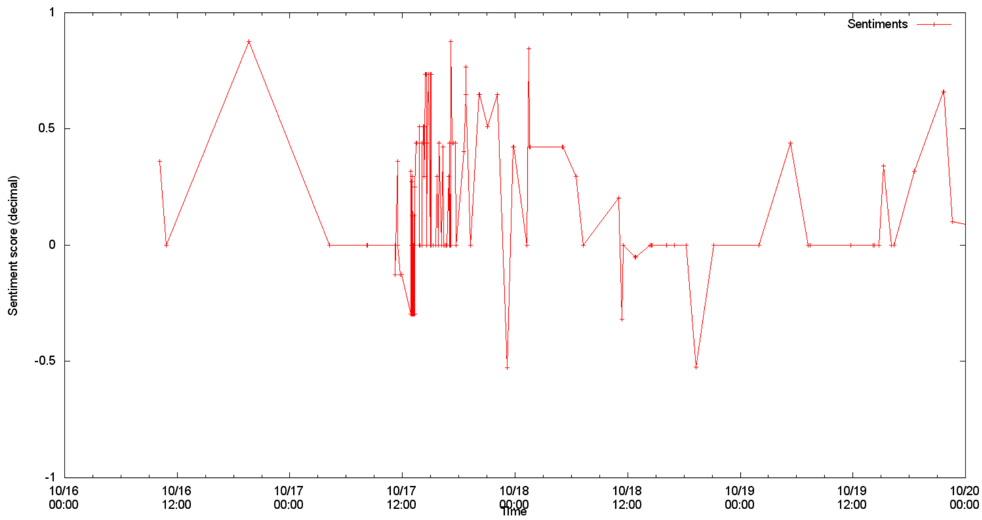


Figure A.15: Sentiment compound score of Hasbro inc at the release of Q3 2016. The vertical axis is the sentiment compound score. The horizontal axis describes the time as a discrete set of seconds.

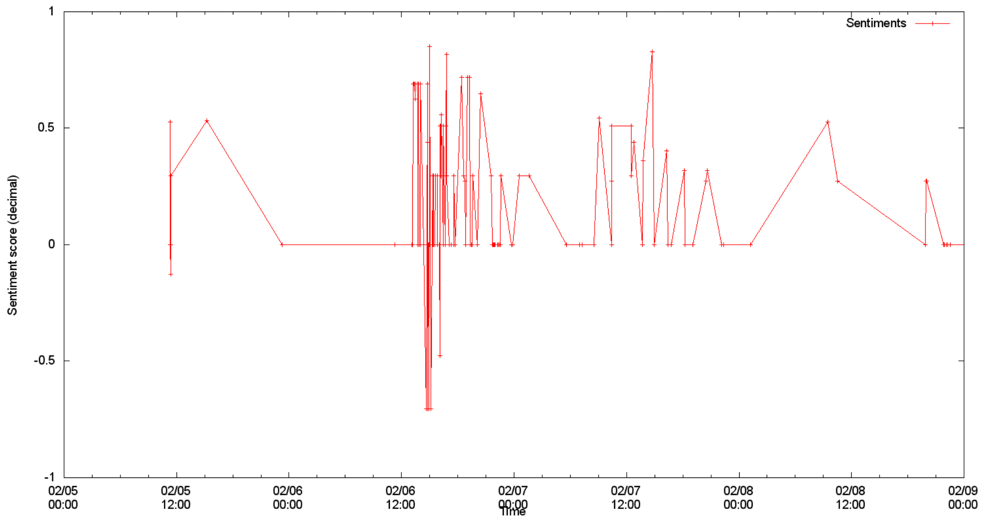


Figure A.16: Sentiment compound score of Hasbro inc at the release of Q4 2016. The vertical axis is the sentiment compound score. The horizontal axis describes the time as a discrete set of seconds.

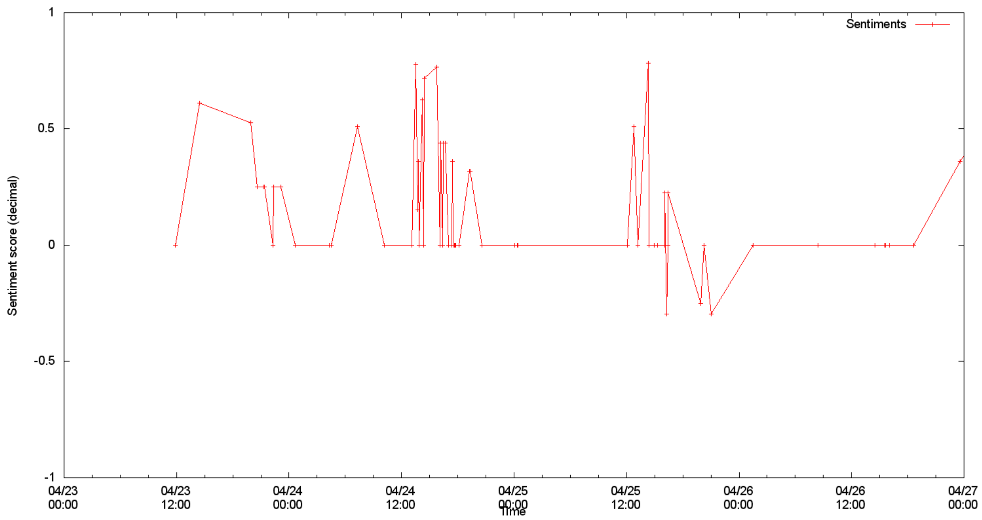


Figure A.17: Sentiment compound score of Hasbro inc at the release of Q1 2017. The vertical axis is the sentiment compound score. The horizontal axis describes the time as a discrete set of seconds.

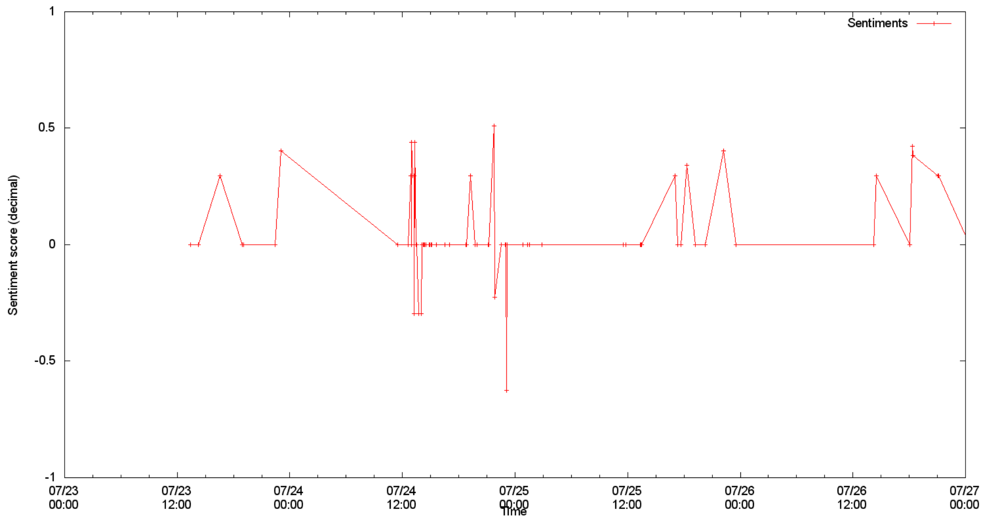


Figure A.18: Sentiment compound score of Hasbro inc at the release of Q2 2017. The vertical axis is the sentiment compound score. The horizontal axis describes the time as a discrete set of seconds.

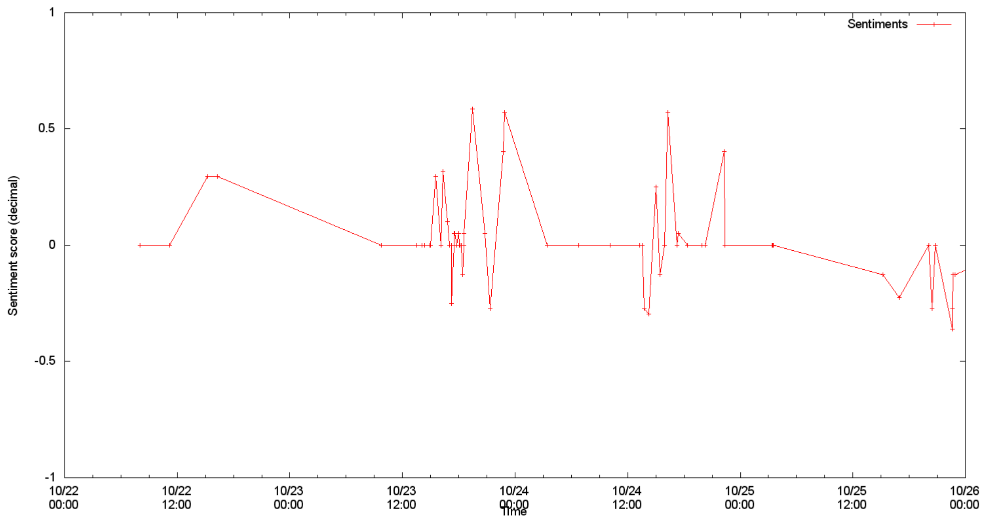


Figure A.19: Sentiment compound score of Hasbro inc at the release of Q3 2017. The vertical axis is the sentiment compound score. The horizontal axis describes the time as a discrete set of seconds.

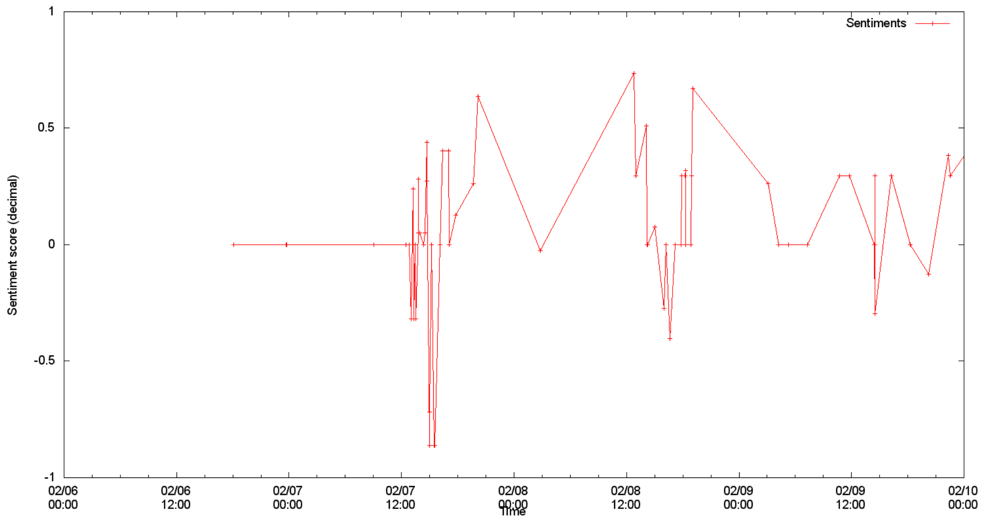


Figure A.20: Sentiment compound score of Hasbro inc at the release of Q4 2017. The vertical axis is the sentiment compound score. The horizontal axis describes the time as a discrete set of seconds.

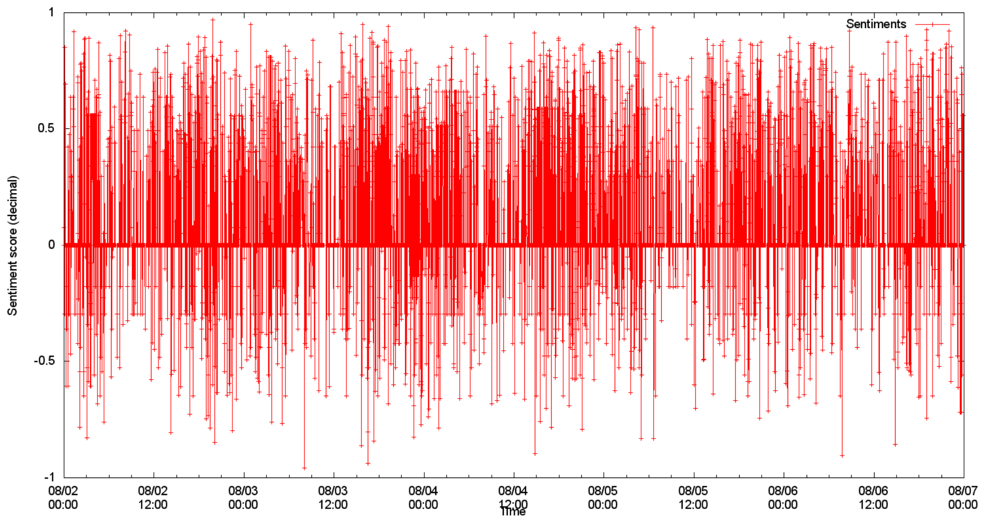


Figure A.21: Sentiment compound score of Herbalife at the release of Q2 2016. The vertical axis is the sentiment compound score. The horizontal axis describes the time as a discrete set of seconds.

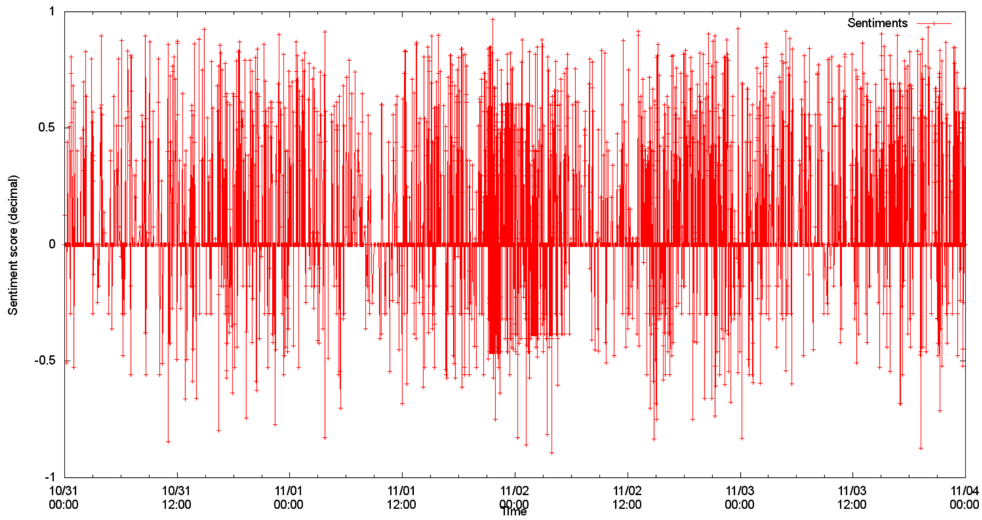


Figure A.22: Sentiment compound score of Herbalife at the release of Q3 2016. The vertical axis is the sentiment compound score. The horizontal axis describes the time as a discrete set of seconds.

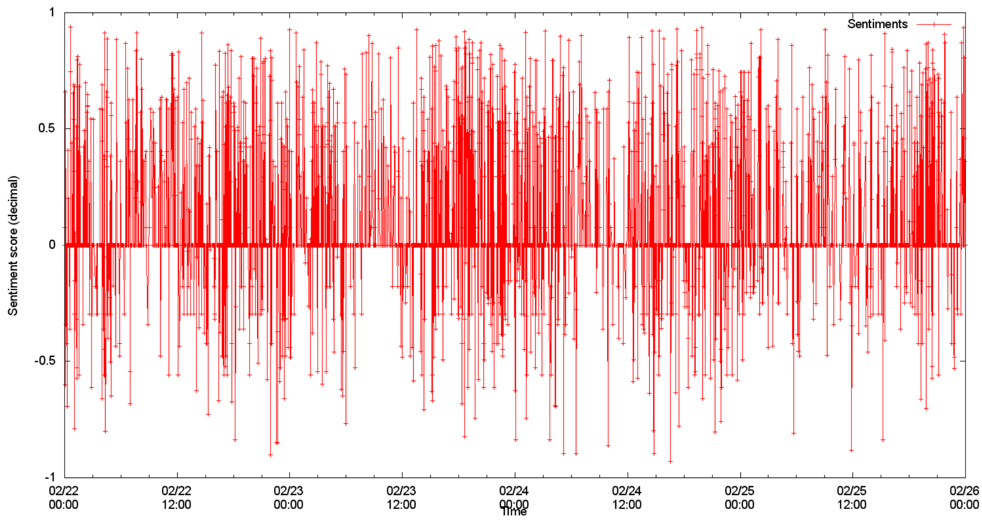


Figure A.23: Sentiment compound score of Herbalife at the release of Q4 2016. The vertical axis is the sentiment compound score. The horizontal axis describes the time as a discrete set of seconds.

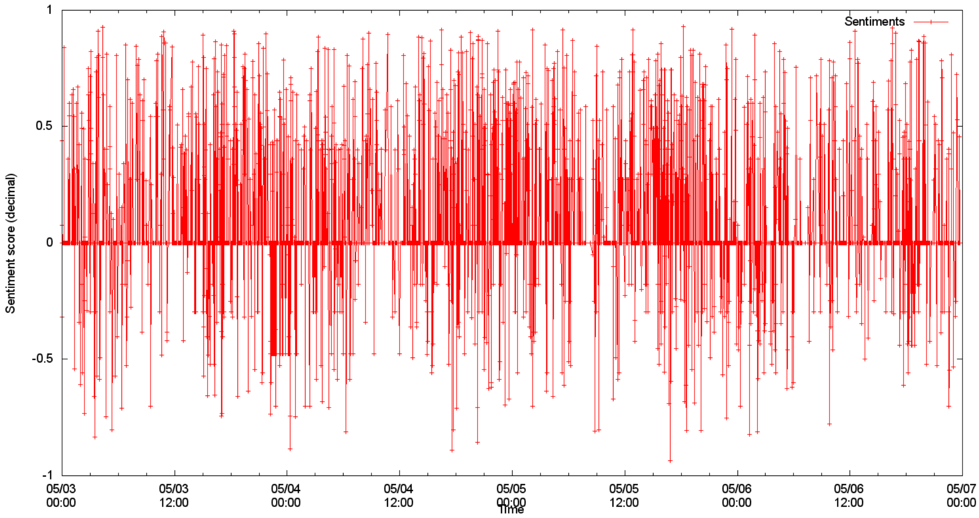


Figure A.24: Sentiment compound score of Herbalife at the release of Q1 2017. The vertical axis is the sentiment compound score. The horizontal axis describes the time as a discrete set of seconds.

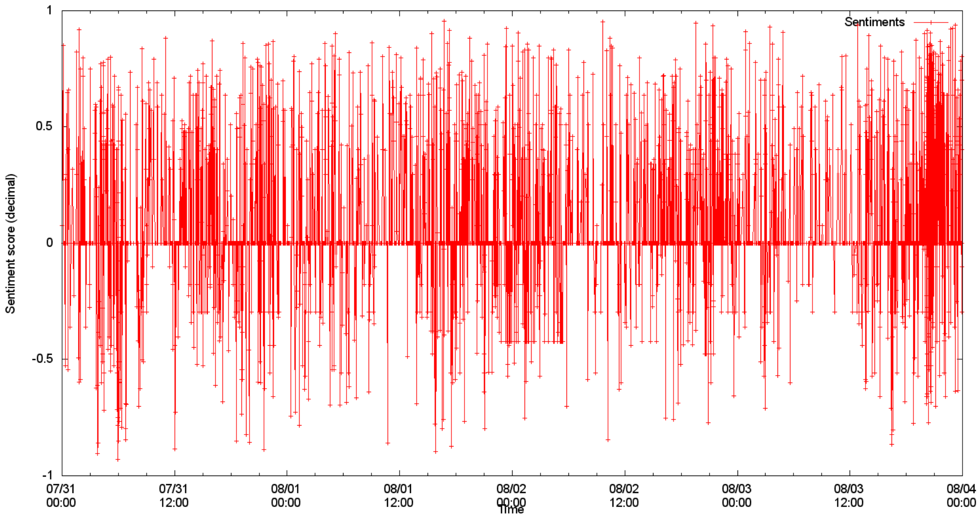


Figure A.25: Sentiment compound score of Herbalife at the release of Q2 2017. The vertical axis is the sentiment compound score. The horizontal axis describes the time as a discrete set of seconds.

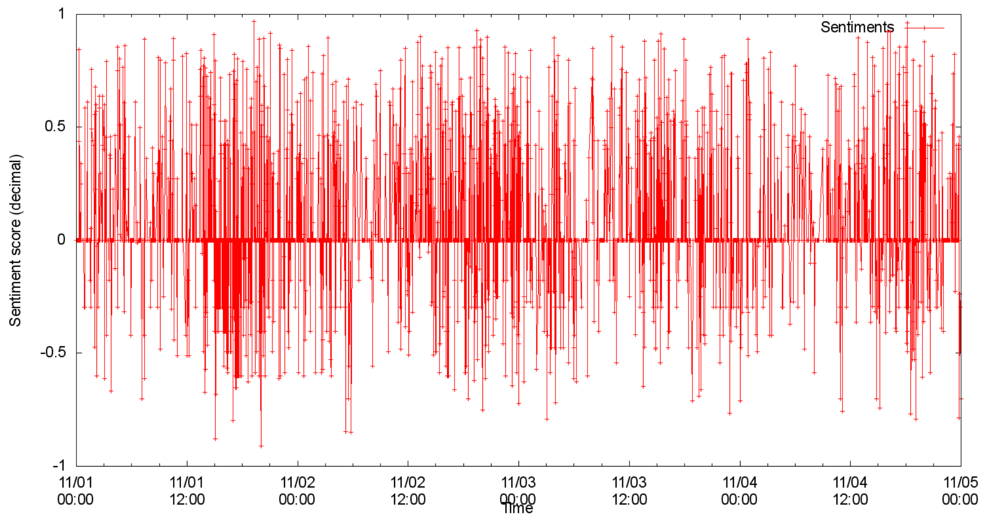


Figure A.26: Sentiment compound score of Herbalife at the release of Q3 2017. The vertical axis is the sentiment compound score. The horizontal axis describes the time as a discrete set of seconds.

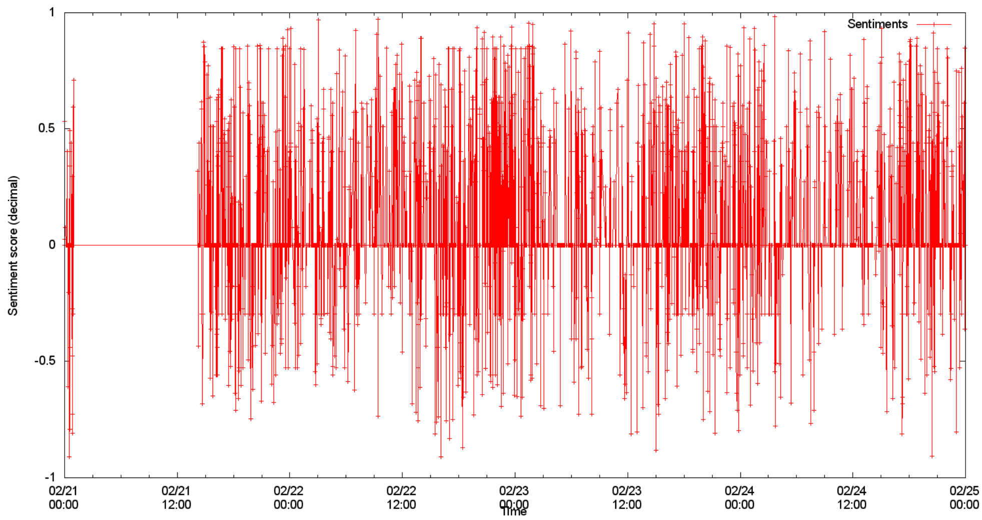


Figure A.27: Sentiment compound score of Herbalife at the release of Q4 2017. The vertical axis is the sentiment compound score. The horizontal axis describes the time as a discrete set of seconds.

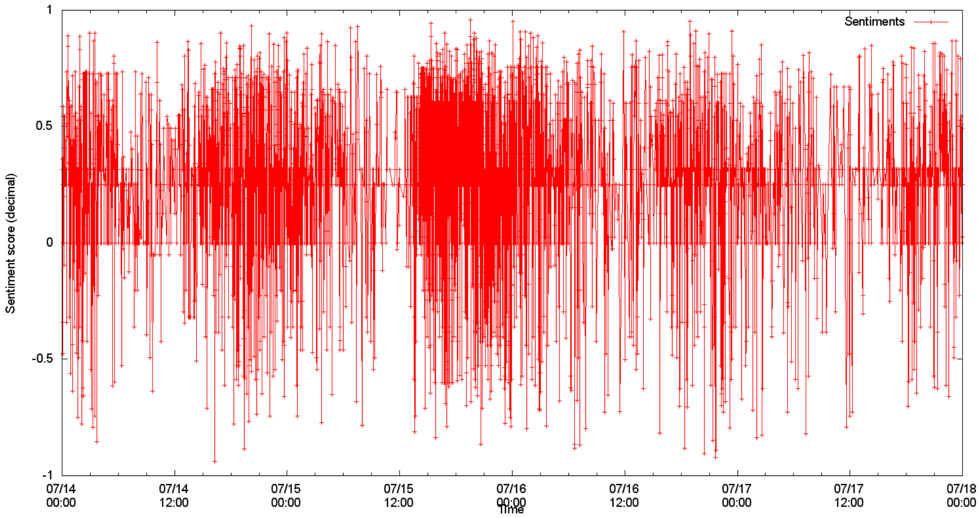


Figure A.28: Sentiment compound score of Wells Fargo at the release of Q2 2016. The vertical axis is the sentiment compound score. The horizontal axis describes the time as a discrete set of seconds.

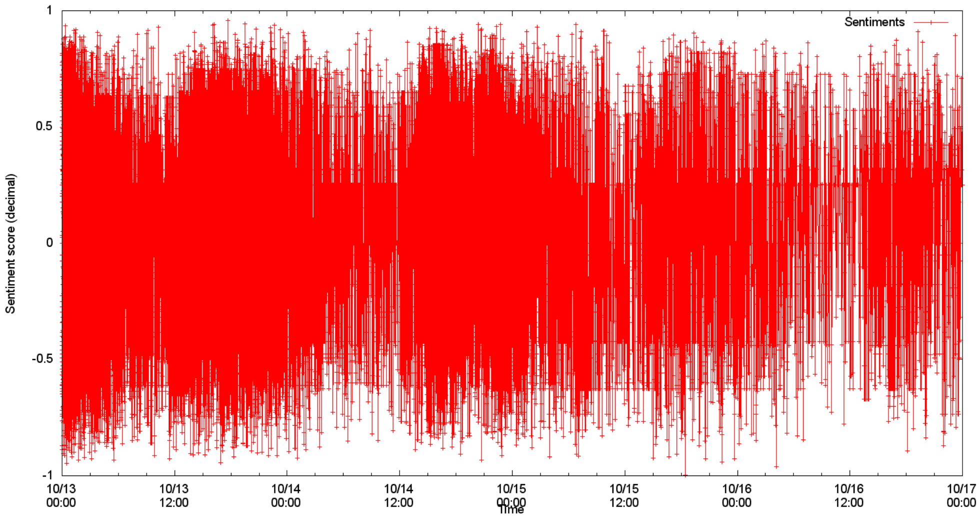


Figure A.29: Sentiment compound score of Wells Fargo at the release of Q3 2016. The vertical axis is the sentiment compound score. The horizontal axis describes the time as a discrete set of seconds.

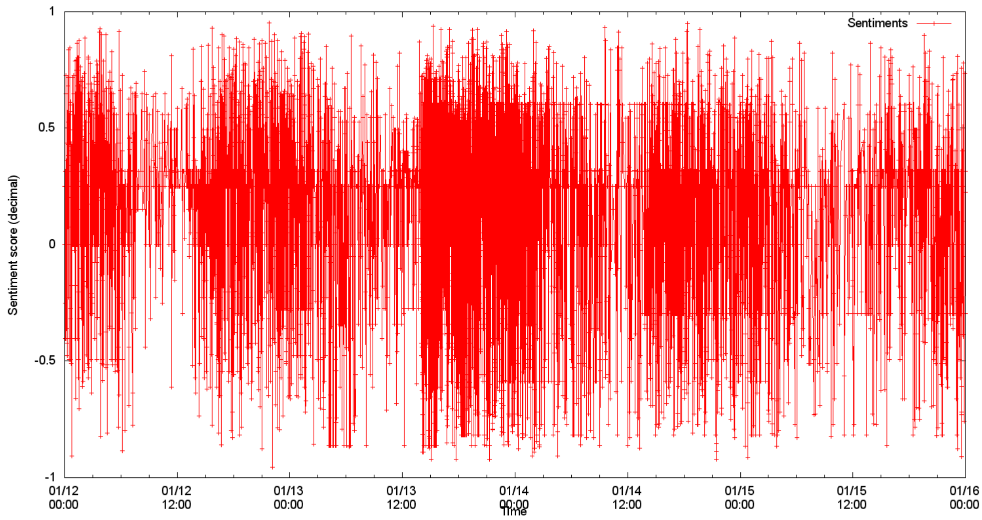


Figure A.30: Sentiment compound score of Wells Fargo at the release of Q4 2016. The vertical axis is the sentiment compound score. The horizontal axis describes the time as a discrete set of seconds.

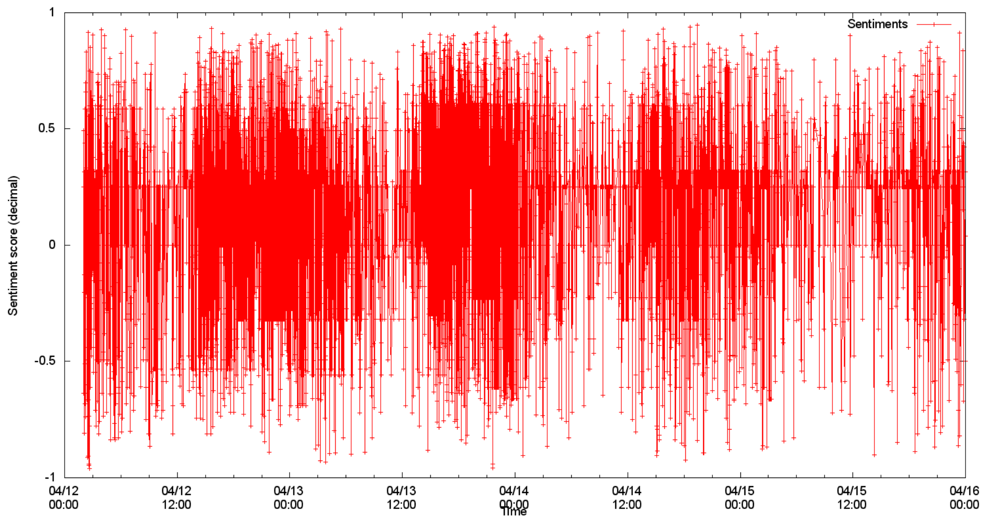


Figure A.31: Sentiment compound score of Wells Fargo at the release of Q1 2017. The vertical axis is the sentiment compound score. The horizontal axis describes the time as a discrete set of seconds.

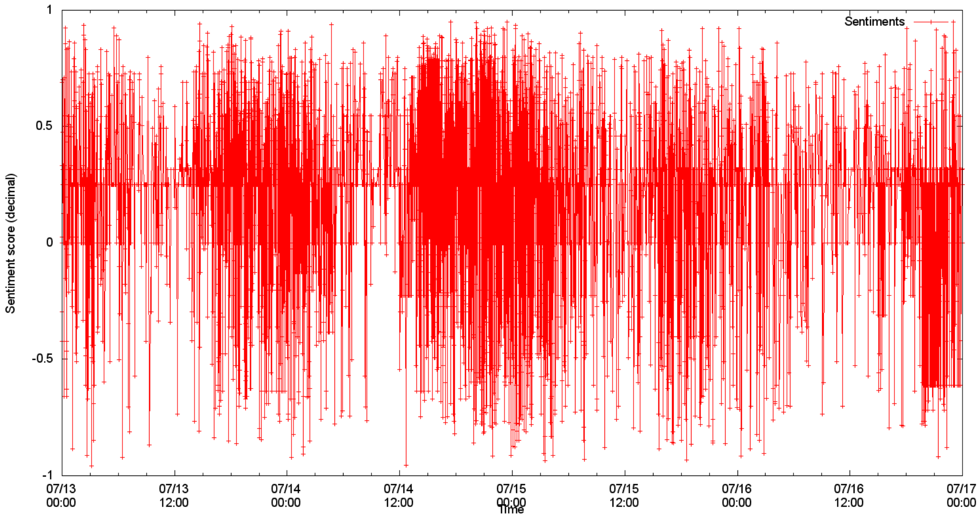


Figure A.32: Sentiment compound score of Wells Fargo at the release of Q2 2017. The vertical axis is the sentiment compound score. The horizontal axis describes the time as a discrete set of seconds.

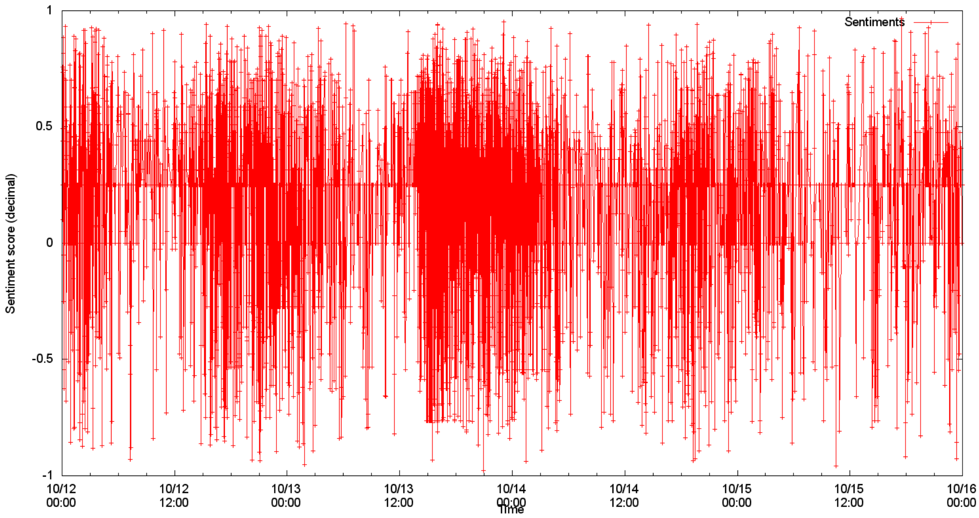


Figure A.33: Sentiment compound score of Wells Fargo at the release of Q3 2017. The vertical axis is the sentiment compound score. The horizontal axis describes the time as a discrete set of seconds.

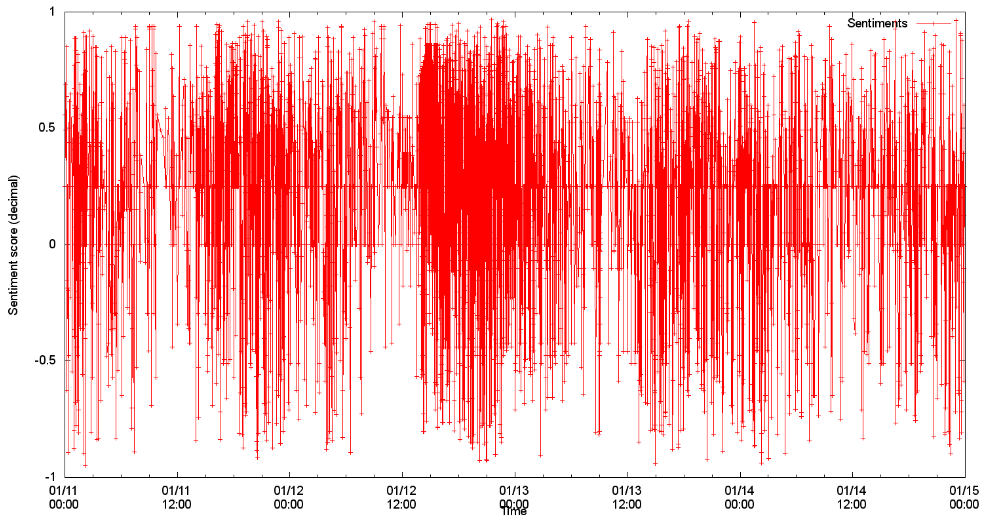


Figure A.34: Sentiment compound score of Wells Fargo at the release of Q4 2017. The vertical axis is the sentiment compound score. The horizontal axis describes the time as a discrete set of seconds.

Appendix B

Linear regression graphs

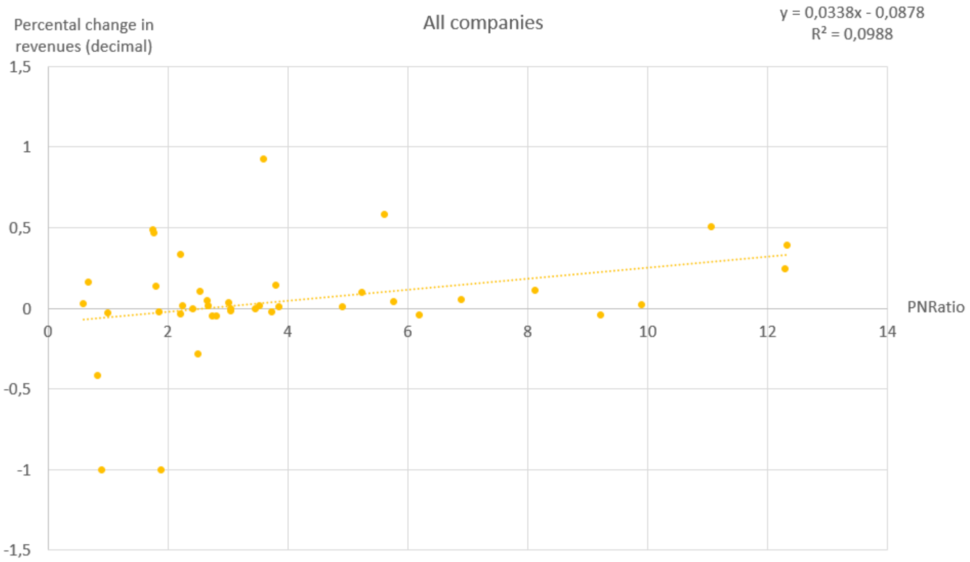


Figure B.1: Linear regression of Activision blizzard, EOG Resources, Hasbro Inc, Herbalife and Wells fargo. The vertical axis is the change in revenue for a quarterly report compared to its previous years quarter. The horizontal axis describes the PNratio based on formula in section 3.1.4

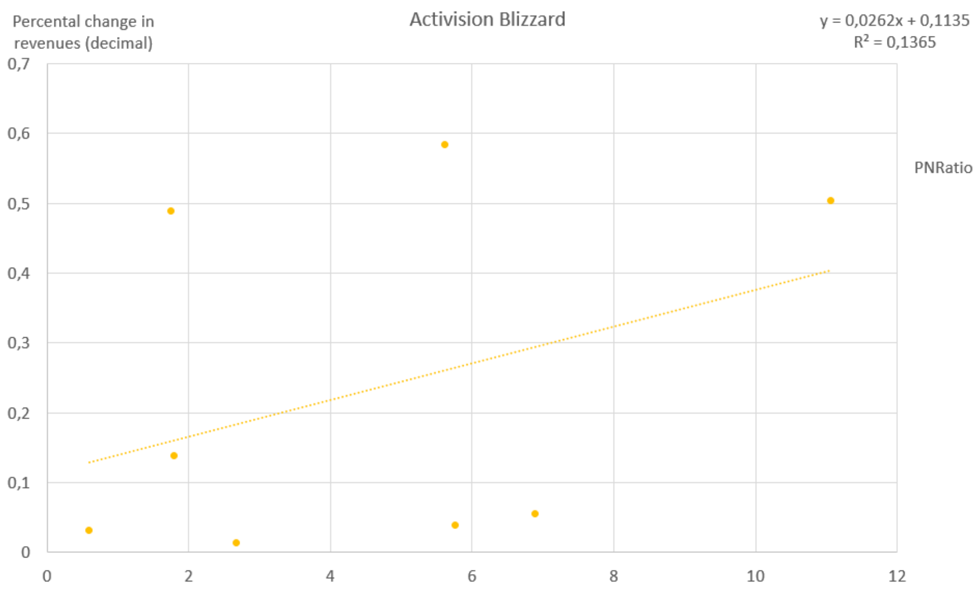


Figure B.2: Linear regression of Activision Blizzard. The vertical axis is the change in revenue for a quarterly report compared to its previous years quarter. The horizontal axis describes the PNratio based on formula in section 3.1.4

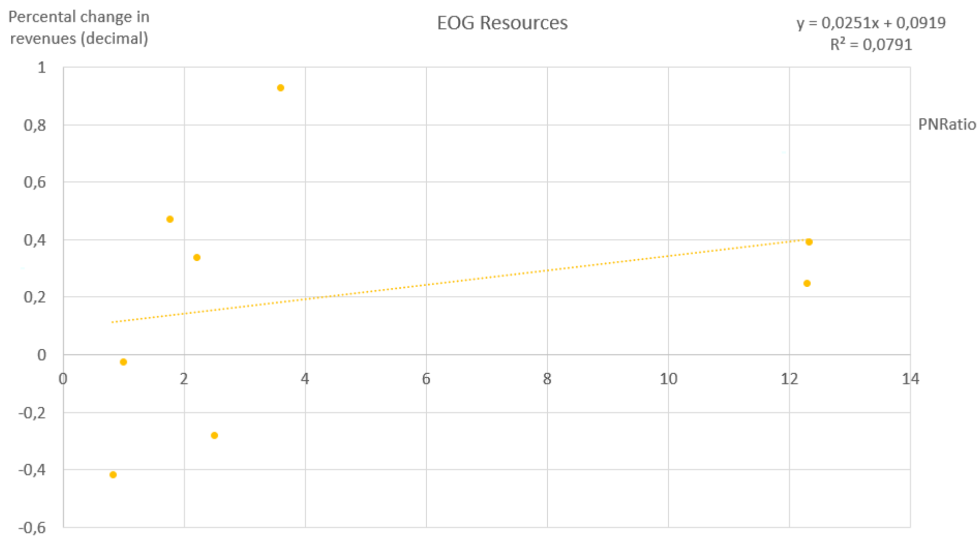


Figure B.3: Linear regression of EOG Resources. The vertical axis is the change in revenue for a quarterly report compared to its previous years quarter. The horizontal axis describes the PNratio based on formula in section 3.1.4

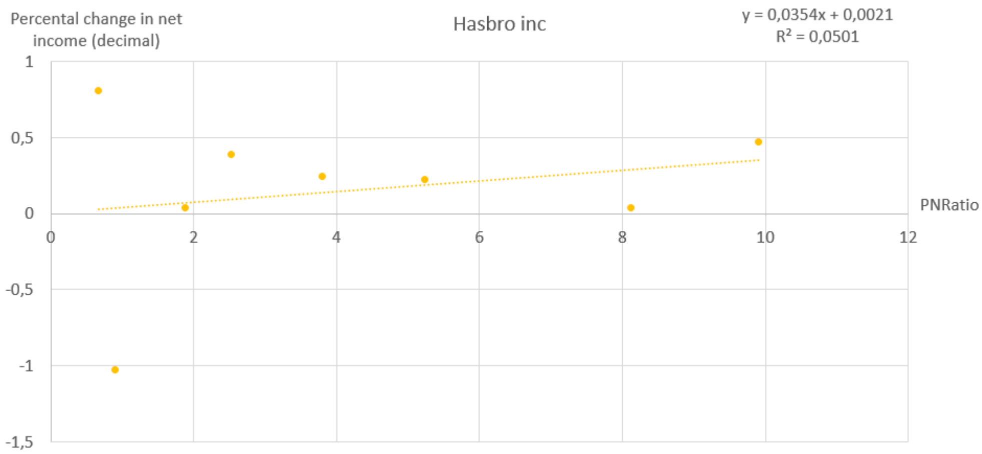


Figure B.4: Linear regression of Hasbro Inc. The vertical axis is the change in net income for a quarterly report compared to its previous years quarter. The horizontal axis describes the PNratio based on formula in section 3.1.4

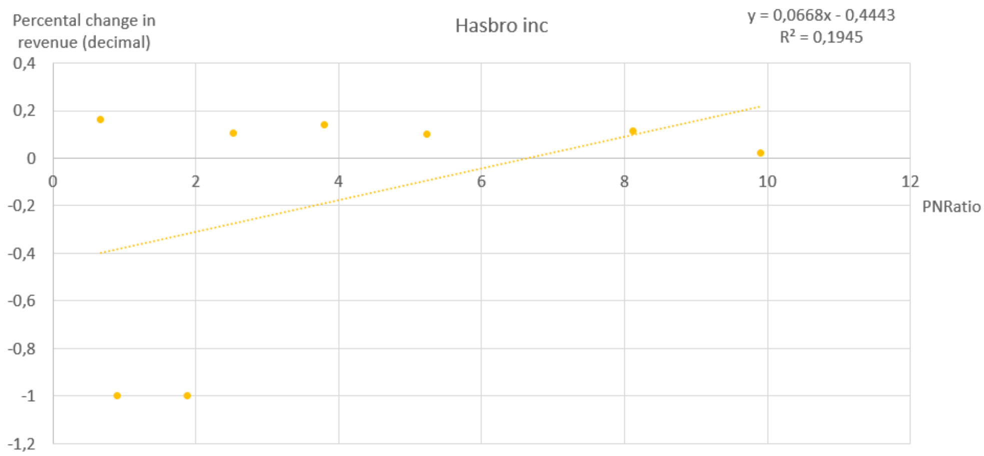


Figure B.5: Linear regression of Hasbro Inc. The vertical axis is the change in revenue for a quarterly report compared to its previous years quarter. The horizontal axis describes the PNratio based on formula in section 3.1.4

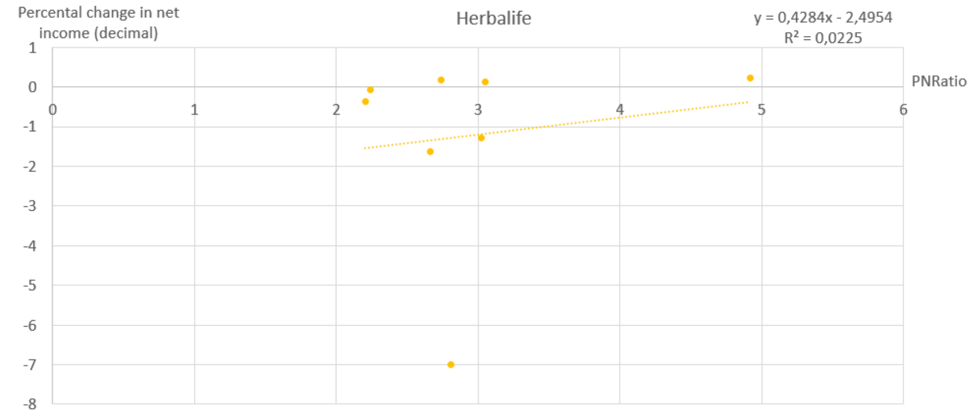


Figure B.6: Linear regression of Herbalife. The vertical axis is the change in net income for a quarterly report compared to its previous years quarter. The horizontal axis describes the PNratio based on formula in section 3.1.4

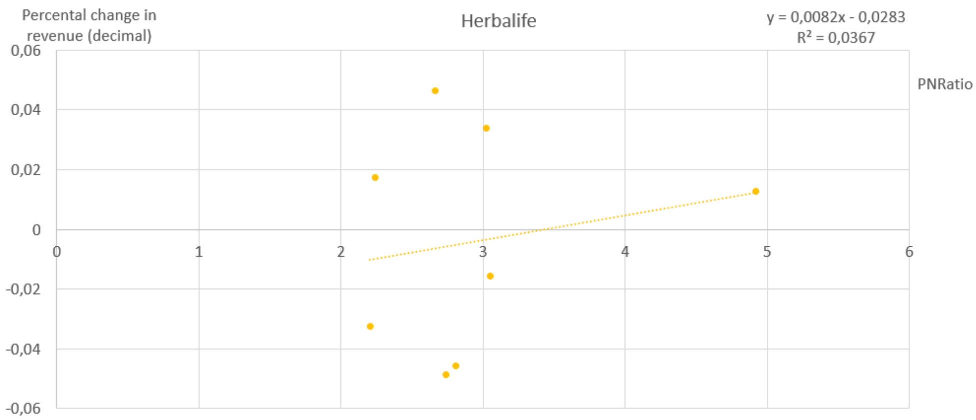


Figure B.7: Linear regression of Herbalife. The vertical axis is the change in revenue for a quarterly report compared to its previous years quarter. The horizontal axis describes the PNratio based on formula in section 3.1.4

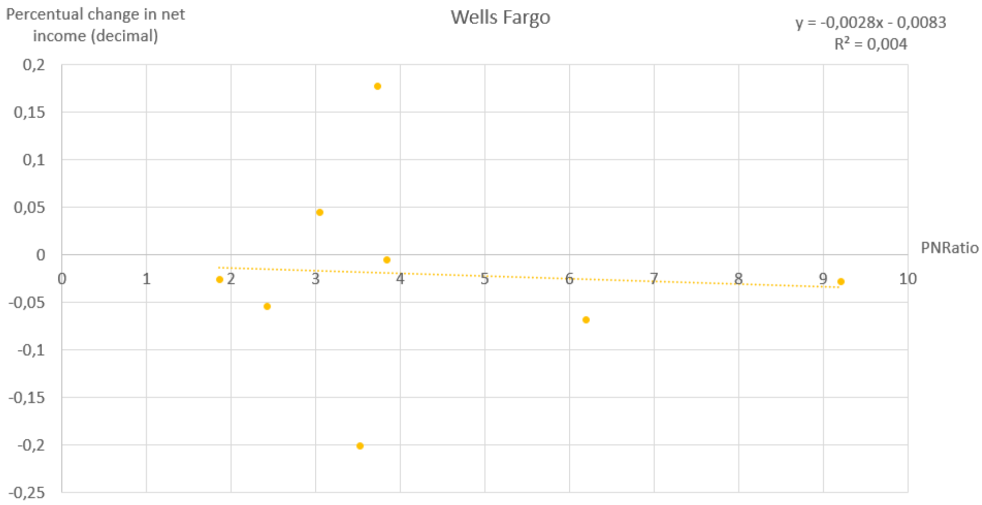


Figure B.8: Linear regression of Wells Fargo. The vertical axis is the change in net income for a quarterly report compared to its previous years quarter. The horizontal axis describes the PNratio based on formula in section 3.1.4

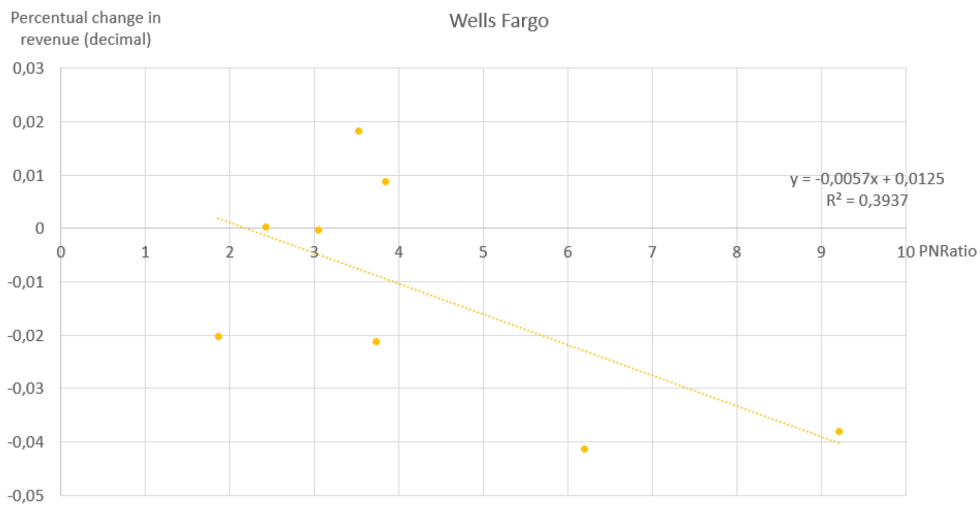


Figure B.9: Linear regression of Wells Fargo. The vertical axis is the change in revenue for a quarterly report compared to its previous years quarter. The horizontal axis describes the PNratio based on formula in section 3.1.4