

Finance Data Analysis: A Comparison Between Yearly Economy Change and Tone Management

Norlan Prudente
University of the Pacific
3601 Pacific Ave
Stockton, CA 95211
n_prudente@u.pacific.edu

1 ABSTRACT

Companies are using tone management to gain recognition in press. How to use tone management effectively based on how the economy is doing in the US can be one of the many questions that comes up in many companies. A research to compare yearly GDP to the sentimental value of reports may benefit companies that have the same question above. Data will be downloaded two ways: GDP from an online database and forms, 10-K and 10-Q, using a web scraper. Web scraper deals with ip rotation to solve the DDOS issue by rotating through a pool of proxies. After gathering all of data required, data analysis is done through python and excel. Excel mainly does the graph. By performing this research, company may improve on using tone management based on how the economy is doing.

2 KEYWORDS

Python, Web Crawler/Scraper, Data Analysis, Economy, and Tone Management.

3 INTRODUCTION

Business, whether small or big, has their way to get their names out there. Using advertisement such as television, internet, radio, or in the street using building, bus, etc. Another way to get recognition is the use of tone management, the act of managing tone of words in order to earn press release. [1] This research paper will focus on tone management and United States yearly economy change. The goal is to determine whether the change in the economy, GDP, correlates on how tone management is utilized. The research will consist of data gathering and data analyzation.

Data gathering can be done either by manually downloading, getting data from a database, or using a web scraper, gathering them from websites directly that cannot be downloaded. Both have been used by researchers and this research will use both method due to the reasons mentioned above. As for data analysis, it will be performed through the conversion of words to sentimental scores.

Business may benefit from this research because companies will have a higher awareness of how tone management are used base on the economy change. By having this type of information in hand, business will have a better understanding on how to use tone management.

This research paper will focus on determining whether tone management and change in economy has any correlation at all. The next part of the paper is structured as follows: section 4 will cover backgrounds that will help readers understand some of the terminology used; Section 5 will cover literature review; Section 6 will cover data gathering techniques (manually downloading and using a web scraper), cover the procedures of data through sentimental analysis, and comparison of sentimental scores with yearly GDP. Section 7 will conclude the whole paper.

4 BACKGROUND

4.1 Python

Python [2] is a high-level programming language that is free. It uses an interpreter instead of a compiler. An interpreter [3] can execute a line of code without going through the whole program. A compiler [3] needs to go through the whole code, compile it to machine code, and create an executable before running it. Python follows an object-oriented [4] paradigm, which mean programming style will be based on objects, which are called classes. A class will have its own data and methods. Classes and objects are treated the same as real world objects, so no two objects should be the same. An example would be having a car with its make, model, and year as data, while turning the engine on will be one of its methods. Python uses dynamic semantics such as dynamic typing and binding. [5] Dynamic semantics makes variable assignment easier by not making the programmer initialize the type of the variable: integer, float, double, Boolean, etc. By using dynamic typing and binding, programmers can use the same variable name, and if they are done with it, assign a new type without the need to specify it. The purpose of dynamic typic and binding is to make coding faster by focusing more on the logic rather than on each variable. Python has a nice and easy structure for people who are just starting to program because it emphasizes in readability and indentation. Lastly, Python has huge libraries and packages that will assist programmers because difficult algorithms are already implemented to perform heavy calculations such as in statistics and calculus. With a wide variety of libraries and packages provided by python, solving DDOS would be easier than implementing it from scratch.

4.2 Defense Mechanism of DDOS

Distributed Denial of Service, DDOS, [6] is a defense mechanism that a website can have. DDOS can be triggered in many ways, such as when a repetitive task that a human would not likely perform occurs. Another way to trigger a defense mechanism is when a repetitive amount of request/download from a website occurs. Lastly, accessing links that are meant to catch a web crawler is called a honey pot. Honey pots are links that cannot be accessed by regular users and can only be seen if the page source is viewed, which makes it a great trap. There are good and bad reasons for having DDOS. The good part is getting rid of malicious attackers, but the bad part is not being able to automate certain tasks. Things such as automatically checking if a website is working properly through the use of web crawler and one of the biggest companies, google, uses a web crawler to make a website visible when searching through their search engine.

5.1 Quantitative VS Qualitative Data

Quantitative data [7] are data that are in forms of numbers. They can be encounter in survey such as questions that requires a number as an answer, Likert scale, bullet choice, etc. "Quantity" is the key to its name. Meaning, data that relates to measurement or measured by the quantity rather than the quality. Quantitative data is easier to analyze because it is already in the forms of number, where qualitative data are not.

Qualitative data [8] are data that comes in forms of words. Some example are questions that are answered with sentences, paragraphs, even a whole essay, reports, etc. Similar with quantitative data, key words “quality” is in its name and it looks for quality rather than quantity. Quantitative and qualitative can also be mixed up when it comes to experiment.

Mixed data [9] can be tackled in many ways, but triangulation was used in the paper called “*Mixing Quantitative and Qualitative Methods: Triangulation in Action*” written by Todd D. Jick. “They argued that more than one method should be used in the validation process to ensure that the variance reflected that of the trait and not of the method. Thus, the convergence or agreement between two methods.” [9] The paper used it for social science, but the research would be similar for data science because the result will have both quantity and quality in it. For this research paper, we used mixed

5.2 Tone Management

Sentimental analysis is widely used in research, but tone management are mainly used in business or financial analysis. This makes it a challenge to do because not a lot of study or research paper are done, therefore resources on how to perform such research is limited. The previous study was mainly concern about the frequency of abnormally positive financial sentiment. Researchers found that more positive tone are used when the company are doing good or have met their goals, but less positive when only the CEO benefits such as receiving stocks, which can result in stock price being reduced. Additional information that was mentioned in the study is that tone management can mislead investors on believing that the company is performing in a different manner than what they really are. As for this research paper, it will be studying, as mentioned in the intro, whether there is a correlation between the yearly GDP of US and the average sentimental values of each company in a yearly base.

6.1 Data Gathering

One of the many good things when it comes to data gathering is that some websites provides data that are formatted nicely and can be downloaded. A web site called The Web Bank holds GDP data for the whole world. The downloaded data is nicely formatted which are separated in countries name and yearly GDP, from 1961 up to 2007. A sample can be seen in Table 1.

1	Data Source	World Development Indicators																		
2																				
3	Last Updated Date	11/14/18																		
4																				
5	Country Name	Country Code	Indicator Name	Indicator Code	1960	1961	1962	1963	1964	1965	1966	1967								
6	Aruba	ABW	GDP growth (annual %)	NY.GDP.MKTP.KD.ZG																
7	Afghanistan	AFG	GDP growth (annual %)	NY.GDP.MKTP.KD.ZG																
8	Angola	AGO	GDP growth (annual %)	NY.GDP.MKTP.KD.ZG																
9	Albania	ALB	GDP growth (annual %)	NY.GDP.MKTP.KD.ZG																
10	Andorra	AND	GDP growth (annual %)	NY.GDP.MKTP.KD.ZG																
11	Arab World	ARB	GDP growth (annual %)	NY.GDP.MKTP.KD.ZG																
12	United Arab Emirates	ARE	GDP growth (annual %)	NY.GDP.MKTP.KD.ZG																
13	Argentina	ARG	GDP growth (annual %)	NY.GDP.MKTP.KD.ZG		5.42784288	-0.852021523	-5.308196827	10.13029766	10.56943334	-0.65972612	3.191996623								
14	Armenia	ARM	GDP growth (annual %)	NY.GDP.MKTP.KD.ZG																
15	American Samoa	ASM	GDP growth (annual %)	NY.GDP.MKTP.KD.ZG																
16	Antigua and Barbuda	ATG	GDP growth (annual %)	NY.GDP.MKTP.KD.ZG																
17	Australia	AUS	GDP growth (annual %)	NY.GDP.MKTP.KD.ZG		2.485605037	1.29647769	6.214278414	6.978723719	5.983450032	2.382491187	6.302385712								
18	Austria	AUT	GDP growth (annual %)	NY.GDP.MKTP.KD.ZG		5.537979317	2.648675113	4.138267582	6.124353738	3.480174918	5.642861411	3.008047798								
19	Azerbaijan	AZE	GDP growth (annual %)	NY.GDP.MKTP.KD.ZG																
20	Burundi	BDI	GDP growth (annual %)	NY.GDP.MKTP.KD.ZG		-13.74613505	9.063157935	4.13540742	6.273037903	3.967225609	4.612992896	13.8215186								
21	Belgium	BEL	GDP growth (annual %)	NY.GDP.MKTP.KD.ZG		4.978423063	5.212003509	4.351584282	6.956684733	3.56065984	3.155894982	3.868146947								
22	Benin	BEN	GDP growth (annual %)	NY.GDP.MKTP.KD.ZG		3.141280459	-3.426409823	4.730027975	6.650759119	5.293862834	3.575862733	1.078398666								
23	Burkina Faso	BFA	GDP growth (annual %)	NY.GDP.MKTP.KD.ZG		4.043927977	6.128798781	-1.268016373	2.282819223	3.764545381	0.537613083	8.823314975								
24	Bangladesh	BGD	GDP growth (annual %)	NY.GDP.MKTP.KD.ZG		6.058160825	5.45303097	-0.45589429	10.95278855	1.606258229	2.566812001	-1.87586392								
25	Bulgaria	BGR	GDP growth (annual %)	NY.GDP.MKTP.KD.ZG																
26	Bahrain	BHR	GDP growth (annual %)	NY.GDP.MKTP.KD.ZG																
27	Bahamas, The	BHS	GDP growth (annual %)	NY.GDP.MKTP.KD.ZG		10.66709743	10.45197189	10.5240552	10.5015672	10.42559649	9.071928711	9.598113867								
28	Bosnia and Herzegovina	BIH	GDP growth (annual %)	NY.GDP.MKTP.KD.ZG																
29	Belarus	BLR	GDP growth (annual %)	NY.GDP.MKTP.KD.ZG																
30	Belize	BLZ	GDP growth (annual %)	NY.GDP.MKTP.KD.ZG		4.892277219	4.89265218	4.956434195	4.99987991	4.930540961	4.770845113									

The data of interest, for this research, is the GDP of the US, therefore cleaning the data will be required. Cleaning up is done by writing a script that grabs only the US data and saving it to another file. The script is a simple for loop that looks at each row for a string containing "United States." Since the position of the countries name is the first row, then comparing the first index of each data to the string "United States" is all that was needed. If a match is found, then the data in that index is grabbed. Table 2 shows the new file.

Table 2. US GDP

United States	USA	GDP growth (NY.GDP.MKTP.KD.ZG)	2.3	6.1	4.4	5.8	6.4	6.5	2.5	4.8	3.1	3.20680726	3.29547673
Insert Function													

6.1.2 Web Scraper Using Python

6.1.2.1 Gathering Procedure

To start gathering the data, all the CIKs, a number representing companies, will be loaded and used by the scraper to find forms that will be gathered, 10-K and 10-Q. The scraper will then ask for the range of dates, beginning and end dates, of the forms that the user wants to gather. Leaving the two dates blank will gather all the forms available. Additionally, the user will be asked to enter the type of form they want to gather: 10-K, 10-Q, or both. Upon answering the prompt, the program will then get proxies to use from <https://free-proxy-list.net/>. The purpose of the proxies is to prevent DDOS from websites, fooling a website by letting it think that multiple computers are using it even though only one computer are accessing it. [10] The scraper will get ten https type of proxies and test it our site of interest, <https://www.sec.gov/>, to make sure it can access the website within 5 seconds. After getting ten proxies and testing them, the program will then load any data that was previously gathered.

The behavior of how the data are gathered is in a Breadth First Search Algorithm, BFS. [11] The scraper behaves in BFS algorithm because it will gather all the links at the current URL, for each CIK, before visiting all the links that was found. The first set of data to gather are links that point to the links of forms. The second set of data will be the direct links to the actual forms. Lastly, the third set are individual sections of the forms. A section called item 7 is what being gathered for 10-K forms and item 2 for 10-Q forms. Those two sections are both titled "Management's Discussion and Analysis of Financial Condition and Results of Operations." For every 100 CIKs or failure to connect, the proxy will choose another one from the ten proxies that was picked. If all ten were to fail to connect, then a new set will be gathered.

6.1.2.2 Data Structure

To understand how the data are gather, different data structure will be shown. The first two sets, links, are in JSON [12] form, while the last one is in list [2]. The data structure for the first set, mentioned above, is shown on Data Structure 1, Data Structure 2 for the second set, and Data Structure 3 for the last set.

Data Structure 1: Links to the links of forms

{type of form: CIK: [array of links]}

Data Structure 2: Direct links to the forms

{type of form: CIK: Year it was reported: [array of links]}

Data Structure 3: Section needed (converted to sentimental value)
[CIK, Reported Year, Sentimental Value of the section]

The last data structured showed the converted section that was gathered, but how did it get converted? The Next section will talk about how the data got converted.

6.2 Data Analysis

6.2.1 Data Conversion

As mentioned earlier, item 7 and item 2 are the data that were gathered. A partial example of this is shown in Figure 1. The section can range from 4 whole page or more. Number of pages can vary depending on the resolution of the screen and the font size that was being used by users. For clarification of how much was being gathered, more than 3000 words were used. With that in mind, the conversion started by using a predefined sentimental values using AFINN-111.txt that can be downloaded at http://www2.imm.dtu.dk/pubdb/views/publication_details.php?id=6010. The text file holds all the sentimental value, or score, of each word that were available in the file, which are both positives or negatives. The range are from -5, worst, to 5, best. The section gathered earlier were read line by line and were split into individual words. To clean up the data, numbers and symbols were removed. After cleaning up the data, each word were checked for their sentimental value and they were added together. The purpose of adding them together is to see whether they would end up on either the positive side or negative side. Additionally, by adding the sentimental score, the score shows how strongly positive or negative the data was. After getting each individual form's sentimental value, a final clean-up was done. The final clean-up was done by adding the sentimental score that goes together by CIK number and the years that they were reported. After that, a comparison is made with the yearly GDP.

Figure 1: Management's Discussion and Analysis of Financial Condition and Results of Operations

ITEM 7. MANAGEMENT'S DISCUSSION AND ANALYSIS OF FINANCIAL CONDITION AND RESULTS OF OPERATIONS

Financial Review

Abbott's revenues are derived primarily from the sale of a broad line of health care products under short-term receivable arrangements. Patent protection and licenses, technological and performance features, and inclusion of Abbott's products under a contract most impact which products are sold; price controls, competition and rebates most impact the net selling prices of products; and foreign currency translation impacts the measurement of net sales and costs. Abbott's primary products are nutritional products, branded generic pharmaceuticals, diagnostic testing products and vascular products. Sales in international markets comprise approximately 70 percent of consolidated net sales.

In October 2011, Abbott announced a plan to separate into two publicly traded companies, one in diversified medical products and the other in research-based pharmaceuticals. To accomplish the separation, Abbott created a new company, AbbVie Inc. ("AbbVie") for its research-based pharmaceuticals business which consists primarily of Abbott's historical Proprietary Pharmaceutical Products segment. On January 1, 2013, Abbott distributed all of the outstanding shares of AbbVie to Abbott's shareholders and AbbVie became an independent company trading under the symbol "ABBV". The historical operating results of the research-based proprietary pharmaceuticals business prior to separation are excluded from Earnings from Continuing Operations and are presented on the Earnings from Discontinued Operations line in Abbott's Consolidated Statement of Earnings. The assets, liabilities, and cash flows of the research-based proprietary pharmaceuticals business are included in Abbott's Consolidated Balance Sheet and its Consolidated Statements of Cash Flows for periods prior to January 1, 2013.

Sales growth and margin improvement in the nutritional and diagnostics businesses and the challenging economic and fiscal environment in many countries around the world have impacted Abbott's sales, costs and financial position over the last three years. Sales in emerging markets increased 11 percent per year in 2013 and 2012, excluding foreign exchange, despite the slowdown in several emerging economies and a weakening of key emerging market currencies in 2013. (Emerging markets include all countries except the United States, Western Europe, Japan, Canada, and Australia.)

In Abbott's worldwide nutritional products business, sales over the last three years were positively impacted by demographics such as an aging population and an increasing rate of chronic disease in developed markets and the rise of a middle class in many emerging markets, as well as by numerous new product introductions that leveraged Abbott's strong brands. At the same time, manufacturing and distribution process changes and other cost reductions drove margin improvements across the business. Operating margins for this business increased from 13.2 percent in 2011 to 18.7 percent in 2013.

In 2013 sales growth in International Pediatric Nutrition was affected by a product recall initiated in August 2013 in China and two other markets for certain pediatric nutritional products supplied to Abbott by a third-party manufacturer. While there were no health issues associated with the recalled products, and the supplier subsequently determined that the products had been safe for consumption, the recall created significant disruption in these markets. As a result, International Pediatric Nutrition sales were significantly lower than Abbott's previous expectations for this business for the second half of 2013. While Abbott initiated investments in the third quarter of 2013 in these markets to rebuild consumer confidence, Abbott expects the recall to continue to have a negative impact on sales in the first half of 2014.

In Abbott's worldwide diagnostics business, margin improvement continued to be a key focus in 2013. Operating margins increased from 19.2 percent of sales in 2011 to 22.2 percent in 2013 as the business continued to execute on efficiency initiatives in the manufacturing and supply chain functions. In addition to continued margin improvement, unit growth across geographical regions positively impacted worldwide

25

diagnostic sales. Worldwide sales for this business increased 8.3 percent in 2013 and 7.3 percent in 2012, excluding foreign exchange.

In the Established Pharmaceutical Products segment, macroeconomic and market pressures in certain emerging markets impacted this business in 2013. Nevertheless, sales in this segment's 14 key emerging markets increased 6.3 percent in 2013 excluding the effect of foreign exchange. However, the growth in emerging markets was largely offset by declines in developed markets where austerity measures have continued to impact performance.

Over the last three years in the vascular business, Abbott continued to build its *Xience* drug-eluting stent franchise with the receipt of approval to market *Xience Xpedition* in various countries, including Japanese approval in the third quarter of 2013 and U.S. approval in the fourth quarter of 2012. *Xience Pro* received CE Mark approval in the second quarter of 2012. Abbott's market share also benefited from the U.S. launches of *Xience nano* and *Xience PRIME* in 2011, and the Japanese launches of *Xience PRIME* small vessel DES in 2013 and *Xience PRIME* in April 2012. *Xience*, which includes *Xience V*, *PRIME*, *nano*, *Pro*, and *Xpedition*, ended 2013 as the market-leading drug eluting stent globally. In 2013, *ABSORB* and *MitraClip* also contributed to sales growth. In 2011, the third party distributor of the *Promus* product began transitioning away from the product and that supply agreement ended in 2012. The effect of the winding down of the agreement continued into the first quarter of 2013.

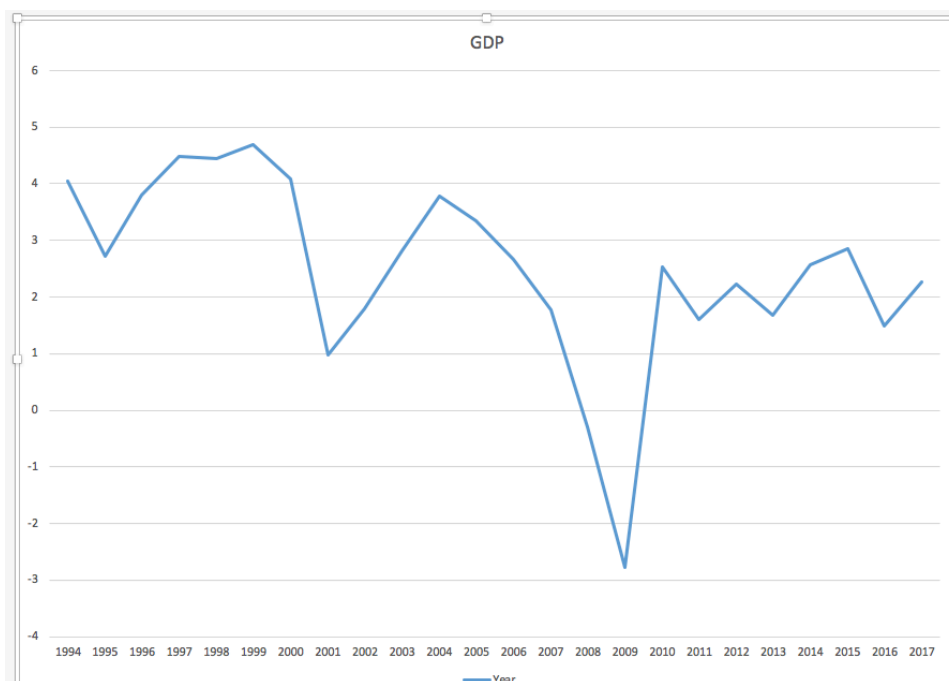
Abbott's short- and long-term debt totaled \$6.6 billion at December 31, 2013. At December 31, 2013, Abbott's long-term debt rating was A+ by Standard and Poor's Corporation and A1 by Moody's Investors Service. In the fourth quarter of 2012, Abbott extinguished \$7.7 billion of long-term debt and incurred a charge of \$1.35 billion related to the early repayment, net of gains from the unwinding of interest rate swaps related to the debt. In October 2013 Abbott announced a 57 percent increase in Abbott's quarterly dividend to \$0.22 per

6.2.2 GDP and Sentimental Comparison

For the purpose of graphing the data, the data is fixed in the format that excel works on, horizontally with each column as year.

Figure 2 shows the graph of GDP change in the United States from 1994 up to 2017. Since the 10-K and 10-Q data, that we currently have, only covers 1994 to 2019, the graph focused on those years. Which is why a zoomed in of the GDP graph is shown.

Figure 2: GDP Yearly Changes



6.2.2.1 10-K Analysis

Looking at the yearly changes in GDP graph, 2009 is the biggest downfall, worst economy, and this year would be the most interesting to compare the data into. To be clearer, looking at the data from the worst economy year and how tone management are used can give the researchers, and the readers, a view on how companies used tone management. Table 3 shows the final modification on 10-K data and Figure 3 shows the graph of the first ten data. Figure 3 looks messy, but as the graph get zoomed in, 2007

to 2011 that is shown in Figure 4, the graph shows that an increase in sentimental value is used in tone management on majority of the companies. Another interesting finding is that looking at 1999 data, the economy is rising, but the sentimental values are falling. A reason for this is that companies probably think that if the economy is good then tone management is not as important as it is in the worst time. Additional reason might be that tone management was not practiced as much as it is in the later years. Figure 5 shows 1999 zoomed in data.

Table 3: Final Data Modification, 10-K

AI	AJ	AK	AL	AM	AO	AP	AQ	AR	AS	AT	AU	AV	AW	AX	AY	AZ	BA	BB	BC	BD	
1994	1995	1996	1997	1998	1999	2000	2001	2002	2003	2004	2005	2006	2007	2008	2009	2010	2011	2012	2013	2014	2015
4.037643	2.718976	3.795881	4.487026	4.449911	4.6852	4.092176	0.975982	1.786128	2.806776	3.785743	3.345216	2.666626	1.77857	-0.291621	-2.77553	2.531921	1.601455	2.224031	1.677332	2.569194	2.861587
		23472	6525	487	574	897	924	742	1394	2203	3814	4156	3511	3978	9368						
27153	10180	13629	26060	0	0	0	1151	787	1788	3014	4592	4920	5317	4230	4711	5427	7701	9340	14420	7653	8205
	16157	5629	6556	24937	19529	1198	868	2235	3012	3886	2621	1725	3156	0	4143	4303	5217	6470	6242	5713	0
	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
		5063	5041	8073	5203	205	6808	1120	988	1330	1814	1680	1944	2814	2869	0	2448	2924	3058	2782	3353
		23336	36983	14773	19233	3866	5782	9620	12186	8832	14735		116281	14410	15331						
		11538	6327	6502	1290	1617	1457	1747	3347	4400	3928	0	0	0	0	3314	3955	4205	3835	3695	3176
		0	17529	16822	358	727	1268	1902	3181	4661	4850	5714	5796	5140	3165	4050	4004	2637	2989	10236	10573
73348	45491	27990	35678	6358	0	0	0	0	7036	10377	14348	12963	12481	15307	18581	13190	13887	14144	7371	7846	5901
	17995	22296	149162	8827	9491	7577	7276	7858	6922	8487	0	7878	27381	8283	7329	6947	8001	6759	6689	6808	
		4499	3245	3794	0	0	3637	0	0	0	0	0	0	0	0	0	5306				
	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
22760	31335	41704	0	0	0	0	0	0	0	0	0	0	0	0	6059	11782	12795	11562	37516	13086	14029
		42292	10871	6954	1434	3518	3127	13990	15913	7455	7630	7669	8691	7511	22218	10665	10537	12719	9440	8655	6648
78112	101956	45706	0	0	0	0	0	1830	2633	3089	6388	2942	5132	6034	6140	29183	5701	6886	7125	7008	8923
52153	5955	7108	5907	515	996	1042	1343	1859	2521	3313	2518	3671	3427	3386	3183	2342	2773	3475	2594	2530	
	15346	40922	1633	1579	2748	8335	8112	3893	2446	4062	4625	5887	7432	5199	4917	21118	19695	29932	27222	30405	
	0	0	35056	45635	37713	7069	39625	116780	43033	57121	15733	19616	53510	12093	14772	2508	4953				
	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
0	3944	4460	4262	6305	6407	5617	5282	5878	7032	7730	10570										
0	0	0	22130	12377	2706	11126															
		10081	12869	12634	2534	1334	2204	1984	0	16970	9196	4110	5755	3525	2169	4461	2563	3139	3275		
	0	5251	4595	849	1218	2571	3026	3234	1672	1330	57325	10723	10211	54757	10924	7215	8669				
	163491	64827	34477	15262	2022	5948	2982	10775	5998	6722	24900	27594									
		10197	894	682	1425	0	0	0	0	2275	3233	2340	0	0	17420	11756	7186	6257	22324	23392	176197
7340	10571	7873	16121	0	0	0	0	0	0	0	0	0	0	0	18275	15252	16804	21740	39070	21117	24274
		0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
49493	10748	18849	11580	0	9391	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
												3215									
5588	5523	4908	5282	2300	3603	2705	1884	2385	2670	2710											
12037	12502	12811	13279	0	0	0	0	0	0	0	9286	7322	7915	8303	9099	10496	10190	15509	10269	6007	10649
		7863	21235	9447	1272	1122	512	2022	3095	5543	6247	9180	6997	25647	23236	20303	15989	0	0	0	0
	27894	27457	34006	41406	0	23933	20	20	-337	-888	0	0	0	0	0	0	0	0	0	0	11171
43952	31077	59235	46954	46812	40901	42559	569	21745	1101	728	700	0	0	0	0	0	0	0	0	0	0
		7269	8525	18789	8094	13396	6840	7830	10108	7639	11514	0	4038	4374	0	3170	2301	2114	1985	1550	1669
26951	0	61998	0	0	0	0	181854	137470	34011	31076	33065										
	0	35636	8888	8047	7726	0	0	3174	0	29059	0	16354	0	0	0	0	0	0	0	0	0
	0	0	0	0	0	0	0	0	0	0	11675	2329	2759	3712	38375	4347	3184	3820	3120	3048	3125
14900	31347	23757	33318	119432	121040	44932	20676	16887	19471	25108	24869	31989	91077	0	30779	32249	31954	66448	47373	93047	13463
6800	7824	6107	8436	11108	6227	6146	7125	5109	7121	7585	18260	8237	5189	4125	2933	3709	9773	12232	5544	14762	18102
11168		39629	10584	580	997		4893	6046	17087	4364											
	20921	28023	26383	0	0	0	0	2061	9476	17364	4511	6310	8441	7705	6026	6016	7155	26910	27890	32908	31074
7914	8003	8147	7780	8795	10136	12443	9553	2037	6581		4822	3527	5069	6329	6412	5945	5889	6832	6190	5133	4447

Figure 3: 10-K First Ten Data

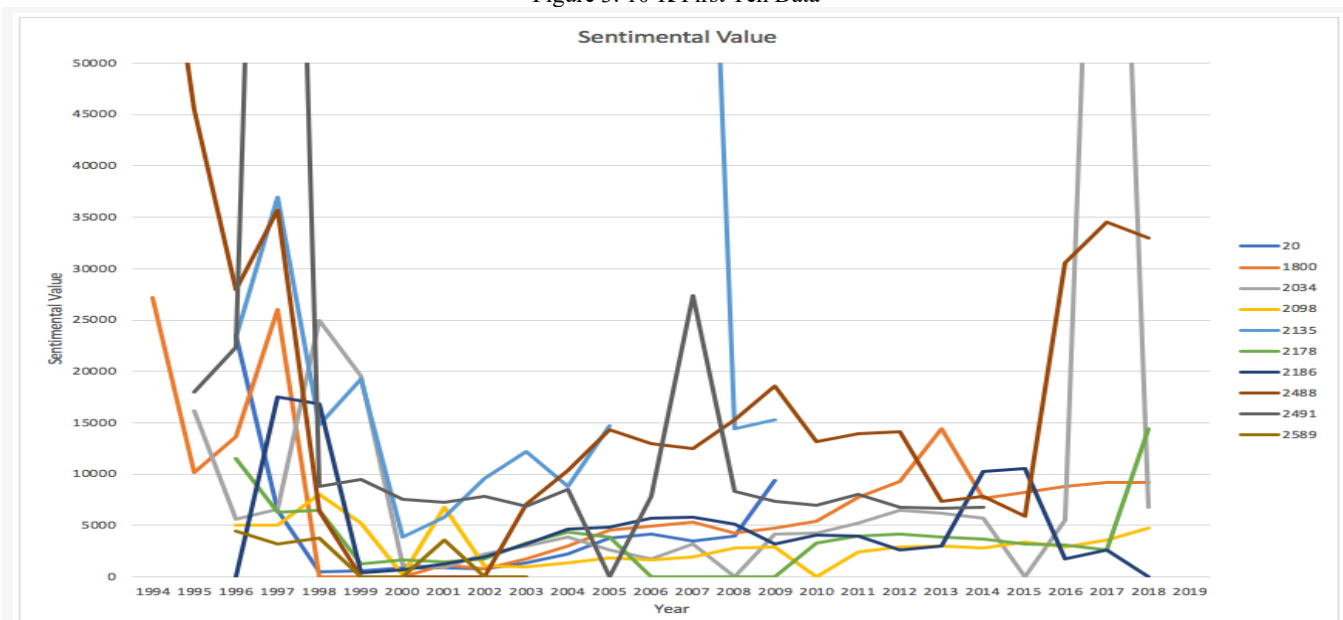


Figure 4: 10-K Zoomed-In Data, 2007-2011

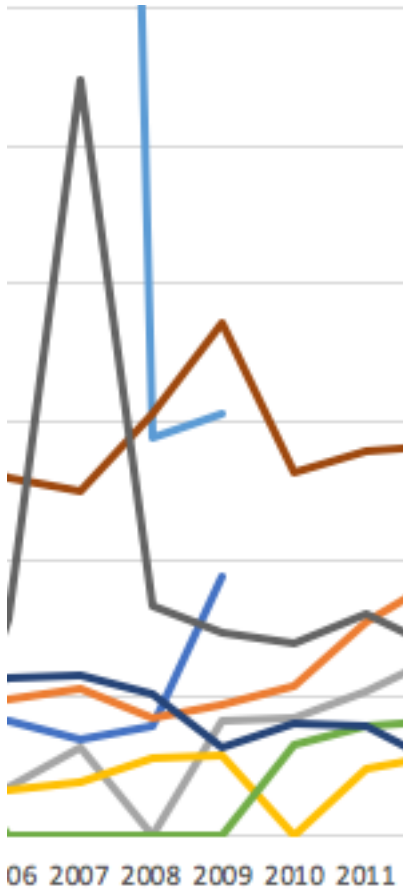
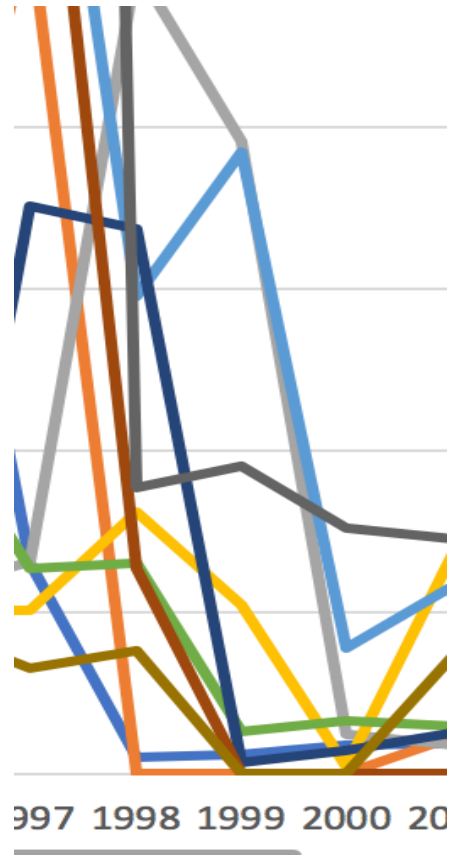


Figure5: 10-K Zoomed-In Data, 1999



6.2.2.2 10-Q Analysis

Similarly, Table 4 shows the final 10-Q data and Figure 6 shows the Graph of 10-Q. Looking at 10-Q data and comparing it to the yearly GDP changes, Figure 7 shows that 6 out of ten has a

decreased sentimental value on the worst economy, 2009. All the while, looking at the similar year as 10-K, 1999, Figure 8 shows that shows that half of the data increased 4 decreased and one neutral. Compare to 10-K, 10-Q does not seem to have a pattern when it comes to using tone management because the graph shows mixed data.

Table 4: Final Data Modification, 10-Q

1994	1995	1996	1997	1998	1999	2000	2001	2002	2003	2004	2005	2006	2007	2008	2009	2010
4.03764342	2.71897579	3.79588123	4.48702649	4.44991096	4.68519961	4.09217645	0.97598183	1.78612769	2.80677596	3.78574285	3.34521606	2.66662583	1.77857024	-0.2916215	-2.7755296	2.53192062
874	525	5208	5039	3178	6722	4635	4959	3467	3467	1736	1416	1902	1736	12	1485	1015
0	0	0	0	3225	0	0	0	0	0	917	984	626	660	524	1482	202
0	0	578	550	3520	2635	790	2170	2670	8752	5398	8696	932	0	1485	1015	1015
0	0	0	0	1756	0	0	0	0	0	0	0	0	0	0	0	0
635	-758	315	50	60	60	60	60	60	0	383	2466	2938	2731	4222	1325	1160
0	13585	10307	12913	10870	2483	8040	22936	18068	25211	50751	82137	71007	27958	27655	4032	4032
0	0	0	0	0	0	0	0	0	0	2054	2434	0	0	0	1752	3132
1071	881	815	362	1237	2844	2635	4326	1898	7169	9014	9834	5999	4258	4961	3132	3132
29892	71576	37602	112508	14495	18132	11835	9570	15859	22815	41605	45117	25236	27927	28046	31528	26539
9401	16395	49100	24066	20378	11635	9278	9330	5021	6382	2320	4039	13315	19323	28965	26478	26478
1880	2884	2535	5270	3861	3514	3309	4860	3087	0	0	0	0	20	0	0	0
0	0	0	0	305	100	0	0	0	0	0	0	0	0	0	0	0
0	0	0	0	0	0	0	0	0	6375	15313	21874	23483	32348	31235	9226	13140
1531	5986	5465	5000	6079	3939	80	40	31599	15557	14142	3967	6963	5284	14690	14832	14832
0	0	0	0	0	0	1045	2055	0	0	0	9884	2854	2481	1919	18218	18218
26270	8184	38975	7482	5668	8744	0	26944	0	40099	50947	11104	7749	8257	0	3331	3331
12337	2732	3954	8163	364	2581	2328	3328	2205	1938	1552	2102	0	0	0	3393	3016
0	0	26877	14350	21719	13391	7890	7699	14561	6748	12276	17109	20935	17551	11993	20339	20339
0	0	0	0	0	0	0	0	16224	36493	41084	24841	34638	22671	24203	20917	20917
0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
1959	3802	4306	5982	5291	11403	4869	1337	-221	567	4359	0	0	472	0	0	3331
3224	6865	5095	15875	10405	32952	4917	0	0	0	0	0	0	0	0	0	3331
5657	5930	5668	6064	7198	3609	5376	0	0	0	0	0	0	0	0	0	3331
461	1325	1639	2543	18108	7042	9813	2886	3974	9662	92470	27236	21008	19953	9752	13959	13959
0	0	0	0	0	0	0	0	0	15251	10025	13182	8951	8425	15816	7858	7858
5639	2774	5870	16979	10280	4804	8957	8158	12062	8448	6944	0	0	0	0	0	0
0	0	0	0	0	0	0	0	0	0	0	0	334	58	444	465	285
0	0	0	0	0	0	5084	2490	5792	7117	7723	12423	15028	19938	14651	9072	9072
18370	10758	29242	27839	4077	15715	4069	917	6743	9498	14535	3114	2848	0	0	0	0
0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
7668	4818	4265	4431	7179	7625	6025	6399	8841	4991	4226	6169	4171	0	0	0	0
3593	2837	4727	39163	4019	32096	6179	4459	6422	13897	10537	12436	12399	12599	10860	9363	8045
0	0	0	0	0	0	0	63	1413	11946	12676	21761	20323	16280	3444	16066	16066
0	0	0	0	0	0	69	60	60	20	108949	0	0	0	0	48760	105023
8026	9747	16317	16705	18650	13008	26810	17810	12772	8865	16658	17973	33102	28324	27837	20513	20513

Figure 6: 10-Q First ten Data

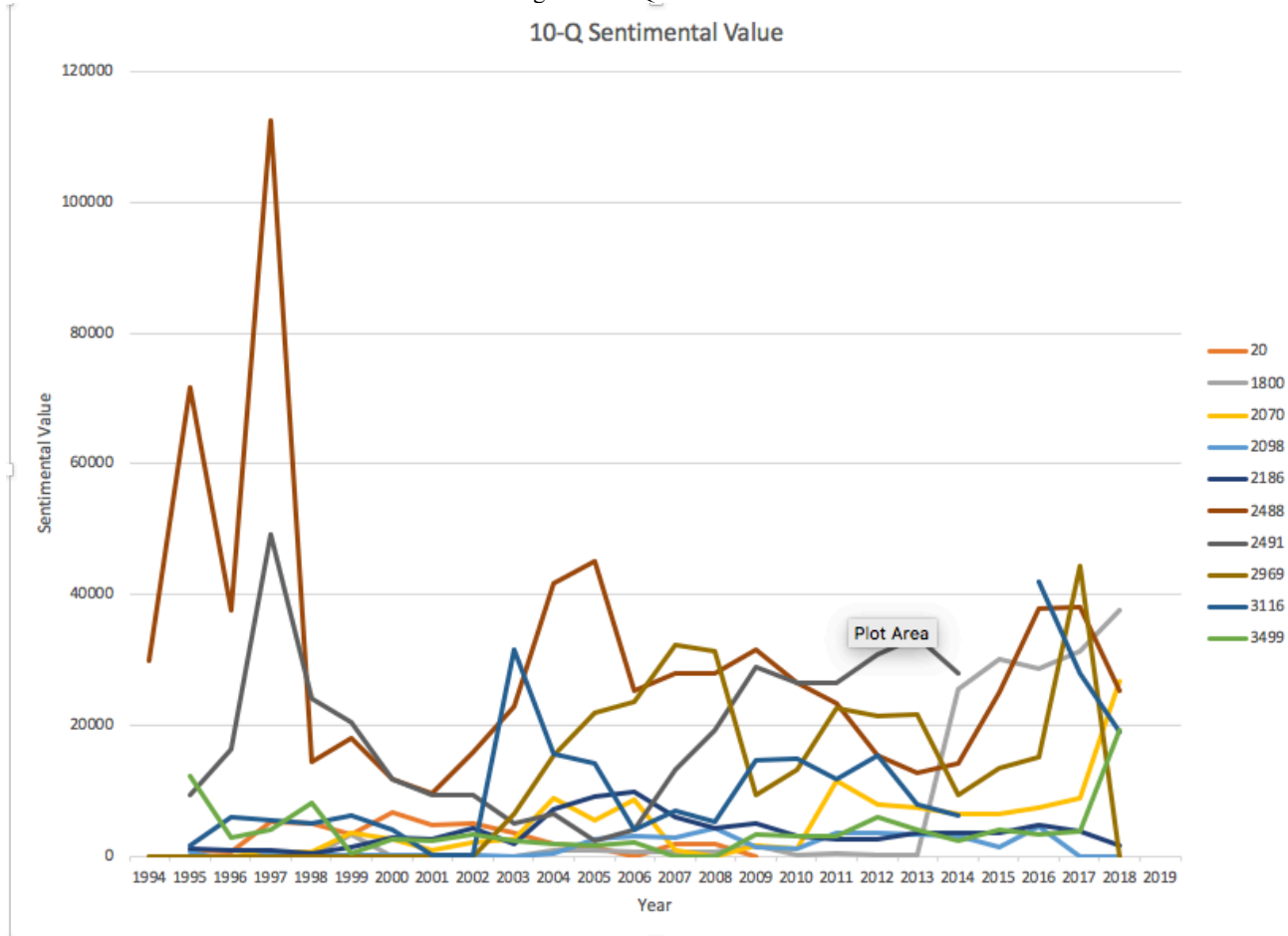


Figure 7: 10-Q Zoomed-In Data, 2006-2011

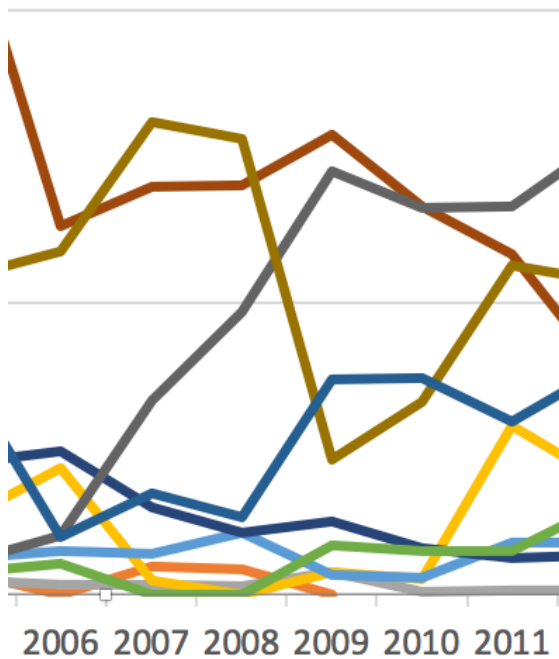
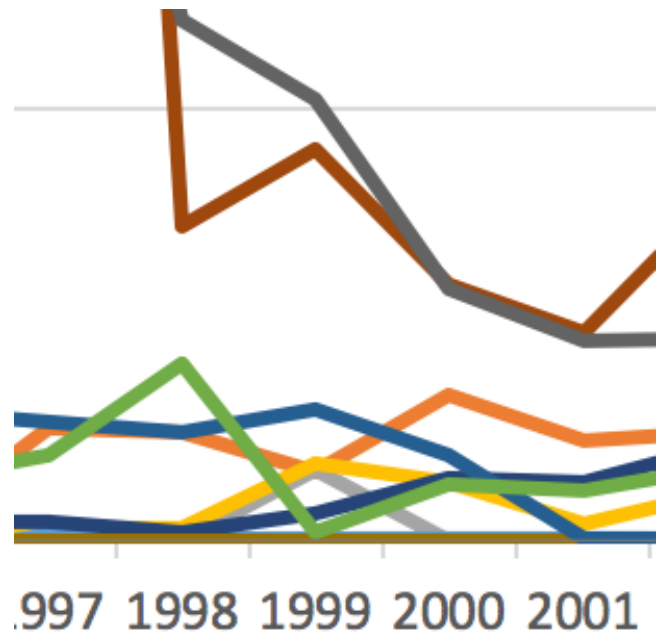


Figure 8: 10-Q Zoomed-In Data, 1999



7 CONCLUSION

7.1 Summary

Companies has been using tone management to earn recognition to the press. Researching towards how tone management are used based on the economy standing may benefit companies on how to utilize tone management in their reports. Data gathering and data analyzation was performed using python and excel. Gathering data was done by downloading the yearly GPD changes from 1961 to 2017 and was easier than using a data scraper. Getting data using web scraper was harder due to data inconsistency. Initially, data that was supposed to be collected are the reports, item 7 and item 2 sections, and the total net or total gross. The forms sometimes have tables or certain pattern, but as CIK change so as the form format. Each CIK is different and even within the same CIK, it changes the report format. Therefore, comparison went to yearly GDP instead. Analyzing the data, it seems that 10-K forms use tone management better on bad economy and relaxes during better years. As for 10-Q, there seem to be no pattern when using tone management. Of course, a lot of factors can also add up to how tone management are used such as how company are doing, total sales made, and employee performance.

7.2 Future Work

7.2.1 Web scraper

Web scraper still needs improvement on speed. It takes a long time to gather everything: links that contain the links to forms, links that holds the actual forms, and scraping html content to be converted into sentimental values. To give the actual speed for 1000+ CIK, the web scraper takes a week to gather each one that was mentioned above, 3 weeks total. Another problem is that when it comes to collecting html content, some links break the program. The solution was, for now, to delete "broken" links.

7.2.2 Data Analysis

Data analyzation comes as personal interpretation from the current researcher. Other researchers can have different finding when they look at the data gathered and use different ways to analyze them. This research only did a simple addition to get the total sentimental value of each year. There are other statistic techniques that can be applied to this research to make it have better findings.

8 REFERENCES

- [1] X.Huang, S. H. Teoh, and Y. Zhang, "Tone Management," *The Accounting Review*, vol. 89, no. 3, pp. 1083–1113, Dec. 2013.
- [2] Python Software Foundation Python Language Reference, version python 3.6, Available at <https://www.python.org>.
- [3] Programiz. *Interpreter Vs Compiler : Difference Between Interpreter and Compiler*. Available: <https://www.programiz.com/article/difference-compiler-interpreter>.
- [4] Tutorialspoint. OOAD Object Oriented Paradigm. Available: https://www.tutorialspoint.com/object_oriented_analysis_design/ooa_d_object_oriented_paradigm.htm.
- [5] S. Ferg, "Static vs. dynamic typing of programming languages," 2009. Available: <https://pythonconquerstheuniverse.wordpress.com/2009/10/03/static-vs-dynamic-typing-of-programming-languages/>.
- [6] M. Li, "An approach to reliably identifying signs of DDOS flood attacks based on LRD traffic pattern recognition," *Computers & Security*, vol. 23, (7), pp. 549-558, 2004. Available: <https://www.sciencedirect.com/science/article/pii/S0167404804001245>. DOI: 10.1016/j.cose.2004.04.005.
- [7] A. Bryman and D.Cramer, *Quantitative data analysis for social scientists*. Florence, KY, US: Taylor and Frances/Routledge, 1990.
- [8] A. Lacey and D. Luff, "Qualitative Data Analysis," National Institute for Health Research, Available: http://healthindisasters.com/images/Books/9_Qualitative_Data_Analysis_Revision_2009.pdf
- [9] T. D. Jick, "Mixing Quantitative and Qualitative Methods: Triangulation in Action," *Administrative Science Quarterly*, Vol. 24, No. 4, pp. 602-611, 1979. Available: <https://www.jstor.org/stable/pdf/2392366.pdf>
- [10] ScrapeHero, "How to prevent getting blacklisted while scraping," 2014. Available: <https://www.scrapehero.com/how-to-prevent-getting-blacklisted-while-scraping/>.
- [11] GeeksforGeeks. Breadth First Search or BFS for a Graph. Available: <https://www.geeksforgeeks.org/breadth-first-search-or-bfs-for-a-graph/>
- [12] JSON, Inroducing JSON. Available: <https://www.json.org/>