# COVID-19 PUBLIC FIRM RESPONSE ANALYSIS

**MSBA7012**

**Class A Group 6**


| | |
|---|---|
| CHEN Sijia | 3035809792 |
| GU Leling | 3035809091 |
| JIA Weiwei | 3035808619 |
| LI Zixin | 3035809778 |
| QU Ying | 3035811018 |
| WANG Xueru | 3035819670 |
| WANG Shu | 3035814711 |
| WU Yi | 3035808750 |
| YANG Kaixi | 3035811185 |

# CATALOG

# I. Introduction

## 1. Project Background

This project idea originally came from a competition held by S&P Global aiming to get the enterprises' responses to COVID-19, to differentiate those responses across different dimensions and to categorize the S&P 500 firms by those dimensions. Our team focused on the following two aspects of the competition idea and tried deepening its business impacts and applications.

**A reflection of public firms' responses in era of COVID-19 during late 2019 to early 2021**

With the outbreak of COVID-19 coronavirus pandemic, the heightened economic uncertainty and risk have significant financial implications on companies in difference sectors. Many firms have thus taken actions including closing out offices, laying off employees, giving public donations etc. By collecting and analyzing the S&P 500 firms' public release materials, this project aims to find out the differences between those actions by action types and action speed, further categorizing the public firms by these two dimensions.

**An interpretation of the categorization by public firms' commercial characteristics**

It's usually not a coincidence for a firm to take certain actions. With the rapid spread of the COVID-19 in both US and other parts of the whole world, public firms' response actions and reaction swiftness will be affected by various aspects including company culture, company industry, risk appetite and financial robustness. While coming up with those categorizations, the team also tried to interpret them and find out the commonalities within one classification, thinking behind those actions for specific firms, how they intend to deliver messages, maximize the return of shareholders and reflect their sense of corporate social responsibility.

## 2. Data and Methodologies

The target companies are the companies that showed up in the S&P 500 list during the year 2019Q1-2021Q1. The team started from getting raw data mainly from web scraping earnings call, financial report and downloading financial data. For the firms that do not have any records or have very few records after utilizing all three methods, the team filtered out those companies, gave higher priorities and manually searched online to get those companies responses to COVID-19.

When all data was collected, we did data cleaning first, tokenized the documents to sentences and tried to target the sentences with firms' COVID-19 responses by utilizing word2Vec model and K-means clusters along with manual-defined rules. After all the remaining sentences were mostly related to the companies' reactions, we manually tagged 10,500 sentences into one of the following 12 categories: 1). donation, 2). public support, 3). service convenience, 4). health safety, 5). operation and financial strategy adjustment, 6). general positive actions, 7). cut expenses not related with employee, 8). cut expenses related with employee, 9). delay, 10). general negative, 0). COVID-19 impact, 99). not related with COVID-19 response of impact. Then the 10,500 tagged sentences are used as training data and put into two supervised learning models -- dictionary-based supervised model and BERT (Bidirectional Encoder Representations from Transformers) to label the rest of 20K+ sentences.

With all sentences tagged, the team then did four follow-up analysis: 1) conduct overall analysis to see how many responses companies made and what are the main actions; 2) categorize the companies by their response speed into response leaders, response followers and response laggers; 3) categorize the companies by their actions types into Donation, Product Innovation, Strategy Adjustment, Health Protection, and Negative Response; 4) perform case studies for two specific companies with their detailed actions and performances.

## 3. Key Findings and Results

States with higher response volume tend to have less deaths due to COVID-19, indicating that they may develop a more thorough strategy. Except for some common actions in response to COVID-19, different industries have their own measures which show the industrial characteristics.

Consumer-oriented companies, labor-intensive firms and healthcare companies are more likely to take action in the first place. Multinational companies owning Asian sites would react faster than others in the same industries.

Companies' responses are highly relative with the companies' industry and their financial performance. Different industries have different weakness facing the COVID challenges, commonalities appeared for companies in the same industry. Company which didn't take positive and rapid responses had worse performance compared with other companies in the same industry. Those companies who have higher market value and higher net income growth ratio tend to donate more to the society and the community, while those who have lower net income growth rate tend to have negative response to COVID-19.

# II. Data

## 1. Getting Data

The data we collected mainly came from four sources and was constituted of six parts.

### 1.1 Earnings Call Script

Earnings call script is the script of conference calls during which the management of a public company announces and discusses the financial results of a company for a quarter or a year. The original data came from Motley Fool Official Website https://www.fool.com/earnings-call-transcripts. Python package BeautifulSoup is leveraged to web scrape all text scripts during 2019Q4 to 2021Q1 and save as text for S&P 500 companies. In the meanwhile, the fiscal time covering the firms' performance in the call is recorded by quarter as response time.

### 1.2 Financial Report DEF

A proxy statement is a document containing the information the Securities and Exchange Commission (SEC) requires companies to provide to shareholders so they can make informed decisions about matters that will be brought up at an annual or special stockholder meeting. To scrape data from the website, we use both selenium and BeautifulSoup packages as our tools. CIK is the unique index of each company and the most useful key word for us to search the information related to a specific company in SEC website. Thus, we need to firstly obtain CIK index from the company-CIK search website. After that, we use the special website design to get the latest DEF document source website, and use this source website to scrape sentences, find all pages in the document and save them to a file.

### 1.3 Financial Report 10Q

10Q document is the quarterly financial report for a company. In the 10Q, firms are required to disclose relevant information regarding their financial position. We identify the report time of the document from 2020Q1 to 2020Q4 and use the similar method as we use in DEF 14A to get the pages of 10Q document.

### 1.4 Financial Statement and Financial Market Data

Financial data is collected from Wharton Research Data Services (WRDS), including income statement, balance sheet, cash flow statement from 2017 to 2020 and stock price from 2018 to 2020 for all S&P companies.

### 1.5 Manual Search (mostly from Official Website)

Manual data is collected by manually searching from the search engine like Google. We used words like "Company name", "Response" and "COVID" as searching key words to do the relevant response information searching. During the data collecting process, we collected the important information like company response content, the information sources and the response date.

## 2. Data Constitution

| Data Source | File Count | Result | Approximate Constitute % |
|---|---|---|---|
| Company Information | 1 csv | 536 firms/records | 1% |
| Earnings Call | 2,239 txt | 957,781 sentences after tokenization | 42% |
| DEF | 1 csv | 842,116 sentences after tokenization | 35% |
| 10Q | 1 csv | 593,265 sentences after tokenization | 20% |
| Financial data | 7 csv | Used original files | 1% |
| Manual Search | 1 csv | 3,412 sentences | 1% |

*Table II-2 Data Constitution Detail*

# III. Methodologies

## 1. Data Preprocessing

Text tokenization and lemmatization are the first two preprocessing steps and most time-consuming steps. Text normalization includes converting all letters to lower cases, removing extra spaces and special characters. Stop words are kept in the sentences to increase readability.

## 2. Filter out Garbage Information

We sequenced all the sentences after cleaning and remained sentences containing 'COVID' or 'coronavirus' and following six to ten sentences. Then we further filtered out non-covid-response sentences by rules utilizing three dictionaries:

1) Basic Response Dictionary: start with words including response, react, action, effort, and so on, use glove.6B package to find its synonyms, manually review and delete irrational ones. The final dictionary contains 115 words;
2) Basic Effect Dictionary: start with words including affect, cause, impact, and so on, use glove.6B package to find its synonyms, manually review and delete irrational ones. The final dictionary contains 24 words;
3) General Covid19 Dictionary: review 5000 sentences, tag 1 for COVID-19 response and 0 for non-COVID-19, use those tagged sentences as a training dataset and find the top words that differentiate COVID-19 responses, manually review and delete irrational ones. The final dictionary contains 451 words.

For earnings call, DEF and 10Q, if a sentence includes phrases structure like response+ 'the'+ effect (such as mitigate the effect, reduce the impact), then this sentence shall be kept. If a sentence includes effect dictionary only, then the sentences shall be filtered out. If a sentence contains words in the response dictionary or COVID-19 dictionary, then the sentences shall be kept. After that, we create variables including the sentence distance to COVID-19 sentence (for instance, 5 means the current sentence is 5 sentences away from the sentence including words of 'covid' and 'coronavirus', applicable on earnings call only), utilize the output of Word2Vec model and apply K-means with K=5. After viewing the output five clusters,for def and 10q, remain sentences with cluster label 1; For earnings call script, we built two separate models, one with stop words removed(model a) and another without stop words removed(model b), we utilized those two outputs and remain sentences with cluster label 0,3,4 on model a and cluster 1 on model b; For manual search data, since it is all about COVID-19 response, we do not apply any rules and use all original 3K+ records for follow-up analysis.

| Source | Original sentences | Sentences after filtering with COVID sentence distance rules | Sentences after filtering with COVID dictionary rules | Sentences after clustering |
|---|---|---|---|---|
| Earnings Call | 957,958 | 142,085 | 74,494 | 18,576 |
| DEF | 842,116 | 67,526 | 2,242 | 763 |
| 10Q | 593,265 | 86,833 | 27,881 | 8,816 |
| Manual Search | 3,412 | 3,412 | 3,412 | 3,412 |

*Table III-2. 1 Sentences after Unsupervised Learning*

After manually reviewing around 100 sentences for each of the three outputs, the model accuracy can be summarized in Table III-2.2. What can be seen is that the false negative rate is quite low while true positive rate is not as high as 80%+, showing that sentences that the model identified as 'garbage' are most likely to be true 'garbage' but the sentences that the model identified as useful may still contain some 'garbage' information.

| Source | True Positive Rate* | False Negative Rate* |
|---|---|---|
| Earnings Call | 50.3% | 1.9% |
| DEF | 70.3% | 2.5% |
| 10Q | 73.5% | 5% |

*Table III-2. 2 TPR and FNR with Manual Review*

True Positive Rate*: # True COVID-19 related labeled as COVID-19 related / # Labeled as COVID-19 related
False Negative Rate*: # True COVID-19 related labeled as non COVID-19 related / # Labeled as non COVID-19 related

## 3. Manual Tag for 10,500 Sentences with Action Type

### 3.1 Tag Types

We have 12 tag types: 1) donation; 2) public support; 3) service convenience; 4) health safety; 5) operation and financial strategy adjustment; 6) general positive actions; 7) cut expenses not related to employees; 8) cut expenses related to employees; 9) delay; 10) general negative; 0) COVID-19 impact 99) not related with COVID-19 response or impact.

### 3.1 Distribution of Tagged Sentences

| Tag | 0 | 1 | 2 | 3 | 4 | 5 | 6 | 7 | 8 | 9 | 10 | 99 |
|---|---|---|---|---|---|---|---|---|---|---|---|---|
| Distribution | 1158 | 336 | 380 | 572 | 1194 | 1569 | 761 | 267 | 73 | 124 | 74 | 4156 |

*Table III-3-2. 1 Distribution of the Tagged Sentences*

## 4. Train Model to Give Tag for Every Sentence

To classify the companies by the actions they took in COVID-19 era, we use supervised models including both dictionary-base method and BERT method to do sentence classification.

### 4.1 Dictionary-base Method

For the dictionary-base method, we first normalize the pre-tagged sentences, remove stop words, special words and words with length shorter than two, and then tokenize and lemmatize words. After that, we use machine learning methods like maxent classifier to get a dictionary and use this dictionary to predict tags of sentences.

Although we manually classify the sentences into twelve classes, after training the model to do sentence classification, we find that twelve classes are hard to get high testing accuracy. Therefore, we regroup them into five classes: 1) Donation; 2) Product Innovation; 3) Strategy Adjustment; 4) Health Protection; 5) Negative Response.

The fraction of useful sentences to the whole sentences for earning call documents is 19.63% while that for other documents is 63.55%, which are totally different. Thus, we separate earning call dataset with others and make different models. Besides, since the accuracy of five classes is much more important, we firstly make a model to separate useful sentences from unwanted sentences, and then build another model to identify five classes.

### 4.2 BERT Method

BERT (Bidirectional Encoder Representations from Transformers) is a NLP machine learning pre-trained model with strong generalization ability of identifying noise and distinguishing semantics, provided by the Google team. After the fine-tuning process, BERT can handle a wide variety of NLP tasks, including Question Answering, Natural Language Inference and Text Categorization.

In the pre-training process, BERT uses two training strategies, Masked LM and Next Sentence Prediction. In the MLM process, before feeding word sequences into BERT, 15% of the words in each sequence are replaced with a [MASK] to token. The model then attempts to predict the original value of the masked words, based on the context provided by the other, non-masked, words in the sequence. In the BERT loss function, only the predictions of the masked values are taken into consideration. After finishing the MLM process, BERT has completed most of the semantics learning.

In the Next Sentence Prediction process, the model receives pairs of sentences as input and learns to predict if the second sentence in the pair is the subsequent sentence in the original document. During training, 50% of the inputs are a pair in which the second sentence is the subsequent sentence in the original document, while in the other 50%, a random sentence from the corpus is chosen as the second sentence. After finishing the MLM process, BERT has completed the training to recognize the noise.

In our model building process, considering our device limitation, we chose to use the BERT-Base pre-trained model, with 12-layer, 768-hidden, 12-heads and 110M parameters. We then modified the pre-trained model initial output layer and used our manually tagged dataset to do the incremental training, making slight adjustments to attention weights. We separated the manually tagged dataset into 3 datasets, the test dataset, the dev dataset and the test dataset to finish the process of fine-tuning and testing. Finally we trained models for useless information- cleaning and models for sentence classification.

### 4.3 Test Accuracy

The test accuracy of all the models is shown below. The test accuracy of BERT fine-tuning models for cleaning useless information are 80.41% and 82.32%. The test accuracy of BERT fine-tuning models for classifying the sentences are 72.72% and 69.45%. All of the BERT model accuracy performance is obviously better than the dictionary-based models. We finally choose the BERT model to finish the whole process of sentence classification.

| **Data** | **Useful Information Proportion** | **Test Accuracy** | | | |
|---|---|---|---|---|---|
| | | **BERT 1** | **Dictionary 1** | **BERT 2** | **Dictionary 2** |
| **Def 10-q Manual** | 63.55% | 80.41% | 79.59% | 72.72% | 65.28% |
| **Earning Call** | 19.63% | 82.32% | 70.68% | 69.45% | 56.17% |

*Table III-4-3. 1 Test Accuracy*

Model1*: # Model to delete useless information
Model2*: # Model to predict sentence classification

## 5. Company Classification

We clustered all the companies into five company types: the donation type, the product innovation type, the strategy adjustment type, the health care type and the negative response type by the response sentence type distribution that each company has.

We first gave each company a five-dimension score according to their response performance in our dataset. We set each sentence of the company response in the DEF dataset, the 10Q dataset and the Earnings Call dataset as score 1, and added them to the type that the sentence belongs to. We set the response sentences in Manual Search dataset as score 0.6 for applying an additional penalty for the reason that the manual collected data process proceeds in the open engine and unfairness of the collecting process for each company may exist.

We then calculated the overall scores of each company in each of the 5-response-type. We chose the response type with the highest score as the company type. After the company classification process finished, there were 79 companies clustered in class 1, 70 companies clustered in class 2, 167 companies clustered in class 3, 197 companies clustered in class 4, and 27 companies clustered in class 5.

### 6. Build TF-IDF Model for Speed Analysis

We first calculate the number of companies that started taking response in each quarter, according to the earnings call and Financial Report 10Q. In 2019Q4, 47 companies started to respond. In 2020Q1, 409 companies started to respond. And in 2020Q2, Q3 and Q4, the number is 49, 7 and 3, respectively. Based on these numbers, we categorized the companies into three response speed types. Companies in 2019Q4 are response leaders, companies in 2020Q1 are response followers, and companies in the rest of three quarters are response laggers.

After categorization, we used Named Entity Recognition (NER) to identify the specific location words and adjective words from S&P Business Description which is a textual description of each company's business operations. After that, in order to have an insight into these three speed types, we built a TF-IDF model to calculate their score for each type. Because one location might consist of more than 2 words, and they are not some fixed unigram or bigram words, instead of using the TfidfVectorizer or TfidfTransformer package, we built a function by ourselves to calculate scores. And the calculation of location words and adjective words are both based on following formulas.

- $tf(w, D) = f_{wD}$ , Where $f_{wD}$ is frequency for word $w$ in document $D$.
- $idf(w, D) = ln\ ln\ \left(\frac{N+1}{df_w+1}\right) + 1$, Where $N$ denotes the total number of documents in the corpus and $df(w)$ denotes the number of documents in which the word $w$ is present.
- $TF - IDF\ Score = tf(w, D) * idf(w, D)$, then we normalized the TF-IDF score between 0 and 1.

## IV. Results

### 1. Overall Analysis

In the overall analysis, we analyzed companies' overall responses to COVID-19 by state and industry, respectively.

After calculating the average number of responses in each state combined with death cases (data as of March 6, 2021), it is found that in general, states with higher responses have lower death rates, which may be caused by more complete epidemic prevention policies. It is also consistent with the manual search results: to deal with the epidemic, the state WA, which has a high response rate, released the "SafeWA" APP, while IA with lower response did not take such measures.

Additionally, the word frequency reflects the main direction of the company's response to COVID-19. Among the words with a high proportion, "customer" and "employee" often appear in "Product Innovation" and "Strategy Adjustment" contents respectively, reflecting that the company's operation and development are carried out with customers and employees as the core. In addition, the word frequencies of "health", "safety", and "work home" are also very high, reflecting that most companies attach great importance to safety during COVID-19.

The attitude of a company to COVID-19 is also closely related to its industry. Analyzing the attitude of each industry to COVID-19 (combined with Bing Liu and Lougrand and McDonald Finance Dictionary for sentiment analysis), we group six industries into three categories and conduct a detailed analysis. The first includes "Health Care Equipment & Services" and "Consumer Services", which received the direct and violent impact from COVID-19. The second group consisting of "Pharmaceuticals, Biotechnology & Life Sciences" and "Retailing" appears to have obvious industrial characteristics in response to COVID-19. The last group, which suffers not obvious but long-term effects, is "Energy" and "Banks". Despite different patterns among groups, these six industries share some common bigrams, such as "health safety", "social distancing", "work home", indicating that industries took some same actions in response to COVID-19. Regarding the characteristics of each industry, "Pharmaceuticals, Biotechnology & Life Sciences" has bigrams such as "clinical trial", "covid19 vaccine" and "gene therapy" indicate its direction, while "Banks" shows bigrams like "credit loss", "customer client", "paycheck protection" and "payment deferral".

## 2. Company Categorization by Response Speed

According to the result of location's TF-IDF Model, we found that compared with followers and laggers, leaders have the highest score of 'Asia' and 'Asia Pacific', which means that the companies in leader type might have more business that are related to Asia. And this also explains why these companies took response much earlier than other companies.

Furthermore, the result of the adjective word shows that words that are closely related to medical, such as "Medical", "Bio-pharmaceutical", and "Diagnostic", tend to have higher TF-IDF score compared with others. Therefore, we can also say that companies in leader type are more likely to relate to the pharmaceutical or healthcare industry. Regarding follower and lagger, "Financial", "Commercial" and "Retail", tend to have relatively high scores among all adjective words. Therefore, the companies in follower and lagger might mainly be related to financial or retail industries.

The overall response speed varies across industries. Consumer-oriented companies or labor-intensive firms tend to react in the first place such as hotels, restaurants or manufacturing and construction firms. At the same time, many healthcare companies and their stakeholders responded quickly by investing COVID-19 related research or adjusting the production lines to meet derivative demands. Specifically, we find that most relevant R&D in 2019Q4 were initiated under collaboration with the government, which may be the reason for why those companies could be one step ahead of followers. Even in the same industries, the response speed can be different because of companies' business natures, operation locations and major clients. For instance, a lagger in consumer services group provides tax services in the American market and all laggers in healthcare sectors focus on fields not related to coronavirus. So, it is reasonable for them to be a response lagger.

Above analyses are aligned with the net income performance. As shown in *Figure IV-2. 1*, the net income performance for response leaders is significantly different from another two classes. As mentioned before, many response leaders are easily impacted by such a pandemic and some leaders invested a lot in R&D. Therefore, the decrease in net income for them lasted for a longer period. While the quicker recovery may result from the surging demand for their products and services such as breathing machines and treatments after the worldwide outbreaks.

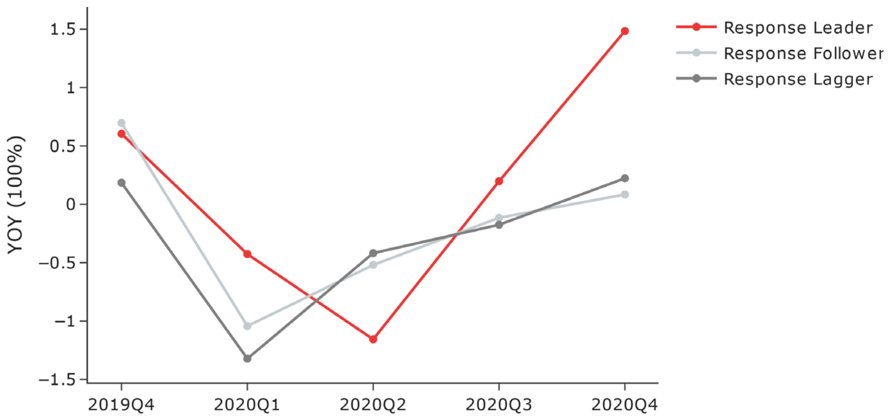



*Figure IV-2. 1 Net Income YOY Comparison for Action Takers*

## 3. Company Categorization by Action Type

### 3.1 Industry Analysis

We then did the industry analysis to see if the company response types have any relationship with its industry. We chose eleven industries from the S&P Global Industry Classification Standard. We calculated to see how industries gather in these five company types and we found the strong relationships do exist. In our calculation, if the overall company proportion of industry j in company type $i_1, i_2, i_3, i_4, i_5$ is $P_{i1j}, P_{i2j}, P_{i3j}, P_{i4j}, P_{i5j}$, the overall proportion of company type $i_1, i_2, i_3, i_4, i_5$ in all the companies is $P_{i1}, P_{i2}, P_{i3}, P_{i4}, P_{i5}$, company type that the industry j gathered is $T_j$, we have:

$$Ti = Industry\ with\ MAX\left[\frac{P_{i1j}-P_{i1}}{P_{i1}},\dots,\frac{P_{inj}-P_{in}}{P_{in}}\right](P_{i1j} > P_{i1},\dots,P_{inj} > P_{in})$$

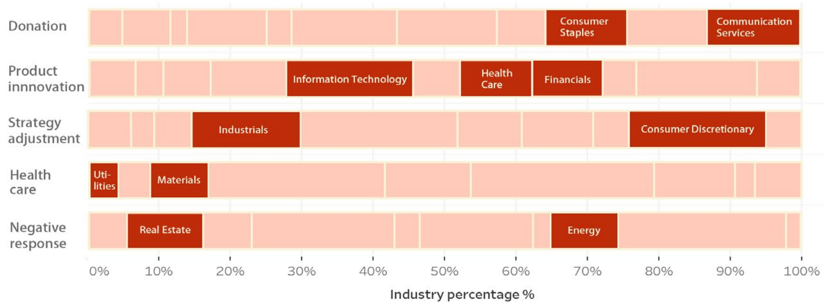
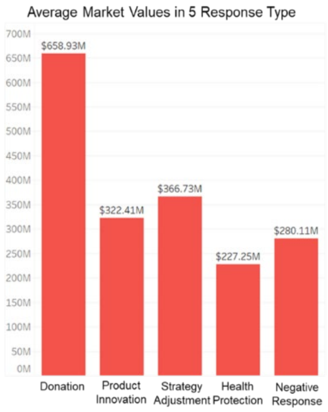


*Figure IV-3-1. 1 Relationship between Response Type and Industry*

We can see that companies in Communication Service industries are likely to gather in the "Donation" company type for the reason that they have limited business fields related to COVID. Most of them choose to give more charitable society support as COVID responses. Companies in Healthcare industries are likely to gather in the "Product Innovation" company type for the reason that they faced lots of opportunities and needs in the pandemic period so that their most frequent response is to develop new products and medical facilities. Companies in Consumer Discretionary are likely to gather in the "Strategy Adjustment" company type for the reason that they faced a lot of management challenges in 2020 because of the COVID, like the supply chain adjustment need and the increasing need for the digital market. Companies in the Utilities industry are likely to gather in the "Health Care" company type for the reason that most of the companies are government relevant and had large on-sight employees need. Companies in Energy are likely to gather in the "Negative Response" company type for the reason that most of the energy relevant market like oil market faced lots of uncertainties during the 2020 and there were few things they can do to get rid of the situation.

### 3.2 Financial Analysis

As the *Figure IV-3-2.* 1 shows, the "Donation" class has the largest average market value, which means that companies which are richer tend to donate more to the community and society.



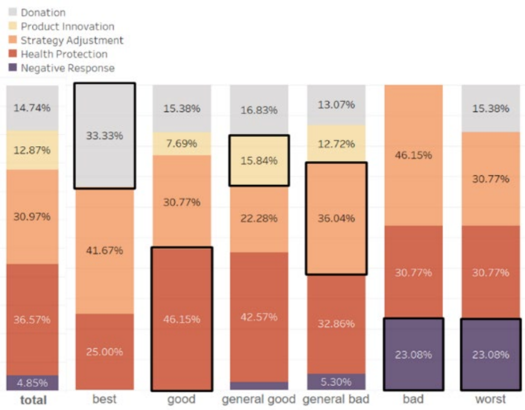




*Figure IV-3-2. 1 Average Market Values in 5 Response Type* *Figure IV-3-2. 2 Relationship between Response Type and Financial Class*

We use the net income growth rate as our financial benchmark to assess the financial performance of each company. After removing the largest growth rate and the lowest growth rate, we classify companies by the method as follows, and donate the largest value and lowest value of the remaining data as "MAX" and "MIN".

| Upper bound | Lower bound | Class |
|---|---|---|
| **MAX** | 0.1MAX | Best |
| **0.1MAX** | 0.05MAX | Good |
| **0.05MAX** | 0 | General good |
| **0** | 0.05MIN | General bad |
| **0.05MIN** | 0.1MIN | Bad |
| **0.1MIN** | MIN | Worst |

*Table IV-3-2. 1 Method to Classify Net Income Growth Rate*

We use the similar method we used before to analyze the relationship between the five action classes and six financial classes. After calculation, the result is shown in *Figure IV-3-2*. Comparing the proportion of five action classes in the total dataset, we conclude that the "best" financial class tend to gather in the "Donation" action type, since companies that earn more money could donate more. Besides, the "good" financial class tend to gather in the "Health Protection" action type, it is because most of the companies in the "good" class are from the "Utilities" industry and they have stable income even in the COVID-19 period. Thus, they tend to focus on health protection but not make other adjustments on their product and strategy, while those whose net income does not perform well need to adjust their strategy or do product innovation. Additionally, companies who present negative responses on COVID-19 also perform bad in the financial index, which is reasonable.

## 4. Individual Case Study

To apply our previous findings, we select two companies from S&P 500 companies to do the case study. One has negative performance in response to COVID-19 and the other has positive performance. We would evaluate companies' responses speed and type.

## 4.1 Negative Performance

We define that a company underperforms the market (S&P 500 index) in a quarter if, in most trading days, its daily percentage is lower than the market's one. L Brands, an American fashion retailer, is selected to assess whether there is a connection between responses to COVID-19 and unsatisfied market performance.

The company owns flagship brands including Victoria's Secret and Bath & Body Works. Its products include bras, panties, athletic attire, body care accessories and more. As of May 20, 2020, the company operated more than 2500 stores all over the world. However, affected by the epidemic, L Brands closed 250 stores in 2020 and more than 50 stores in 2021. Meanwhile, its net sales decreased by more than 40% during 2020.

As shown in *Figure IV -4-1-1. 1*, L Brands started to take actions in 2020Q2 and took more in 2020Q4. At the same time, more than half of retailers responded early in 2020Q1. Regarding the response type, L Brands invested more in health protection, such as wearing masks and taking temperature tests in retail stores. However, many retailing companies also adjusted their strategy in response to the restrictions of retail store shopping. It is concluded that L Brands does not follow the main patterns with the industry from both response speed and type aspects.

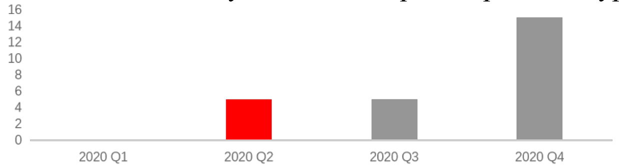



*Figure IV -4-1-1. 1 L Brands response frequency*

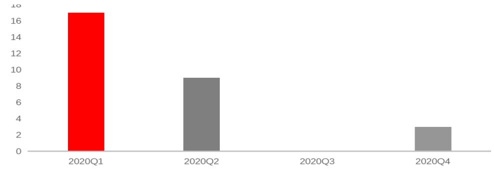



*Figure IV -4-1-1. 2 Retail industry first response*

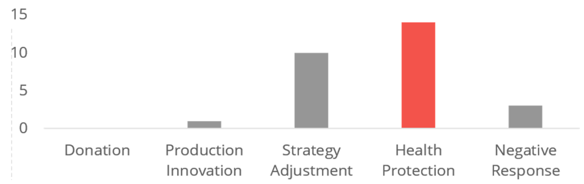



*Figure IV -4-1-1. 3 L Brands response response type*

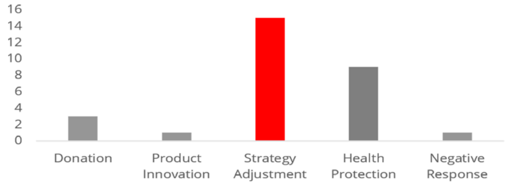



*Figure IV -4-1-1. 4 Retail industry*

Based on these analyses, some suggestions are provided to L Brand to help them handle the pandemic in the future. The first is to closely follow the government policy or change of COVID-19 cases and update the information in time. Then, it is recommended to adopt a more flexible strategy, which may enable it to respond to emergencies quickly and adapt to the new circumstance continuously, since COVID-19 is not a one-off event, it will continue for some time. A practical suggestion is to improve its digital shopping mode and increase consumers' online shopping experience.

## 4.2 Positive performance

On the contract, Merck, one of the largest pharmaceutical companies in the world, showed a positive attitude with timely responses and a comprehensive responding plan, as shown in *Figure IV -4-1-2. 1*. In the early stage, Merck took active actions to help China by donating money and medical supplies. Meanwhile, it set up emergency response teams to make adequate preparations for future uncertainty. Merck continued to focus on collaboration and research throughout the whole year. Moreover, after 2020Q2, Merck invested more in the development of vaccines and drugs.

Through the financial data analysis, we found that there is a sharp jump in the investment in research and development from the third to the fourth quarter in 2020, and correspondingly the net income declined rapidly in this period. This significant change was because Merck signed $356 million to deal with the US government for experimental COVID-19 therapy in 2020Q4.

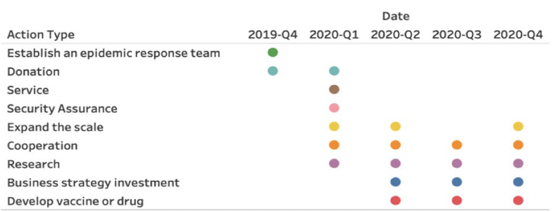



*Figure IV -4-2. 1 Action Type Distribution by Data*


*Figure IV -4-2. 2 Change of Net Income*

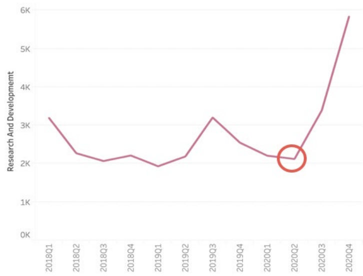



*Figure IV -4-2. 3 Change of Investment in Research*

# V. Limitations and Future Work

## 1. Work Limitations

**2020Q4 financial reports have not been completely released yet**
When the team started to do the analysis, some of the public firms have not released their financial reports yet, resulting in some missing records.

**The data scope is not enough**
The original data source is huge, however after data cleaning, some of the firms remain only one or two sentences. Moreover, some companies took active actions to respond to COVID-19 but do not release that in the financial reports or mentioned a lot in the earnings call. This could make the analysis biased.

**Manual tag for training sample led to the categorization inconsistent**
In total seven people devoted to the sentence tag process with ten categories. Each team member has their own benchmark, judgment and understanding of the sentences, since the sentences are tokenized from the whole paragraph, making the manual tagging process even more difficult. All above reasons could lead to the inconsistency of the sentence tag and thus have an impact on the training data and supervised learning modeling process.

## 2. Future Improvements

**Increase Data Volume and Get More Timely Data**
Instead of focusing on the official financial websites and financial database, try utilizing some other social media platform like Facebook, Twitter and LinkedIn or public news platforms. Also try the information related to COVID-19 responses posted by the company's official website to get more timely data.

**Increase Web Scraping and Text Normalization Speed**
Our Web Scraping and text normalization process took more than half a week and would take more if later we want to incorporate more public firms in the analysis. Each time when we want to change some of the text content we have to re-do the process because the dataset is too large to allow us to store all of the original sentences. To solve these issues, we could utilize the cloud data platform to either improve the data processing speed or the storage issue.

**Increase Unsupervised Learning Modeling Accuracy**
Current rule-based and unsupervised learning to position COVID-19 response related sentences has worse performance on the earnings call script than on other sources. This might be because the different sentence and paragraph structures between earnings call and financial reports. To improve the model output accuracy, we could try strengthening the rules for earnings call scripts and try other models like GloVe model or topic modeling.

**Increase Supervised Learning Model Accuracy**
In the manually tagging process, the final characteristics in each type may not be that idealistically obvious as we expected and this may finally influence the model predicting performances as we mentioned before. A more clear standard and measures like allocating the tagging workload to one person can be taken to make sure the final tagging standard is more accurate. Instead of using the based-BERT pre-trained model, the BERT pre-trained model with more layers can be used with better devices provided. Also, we can try model fusion with BERT and LightGBM, for example, to help to develop the final classification performance.

# VI. Work Allocation

**CHEN Sijia**

- Coding: 1) Web scrapped for DEF 14A and 10-Q from SEC website; 2) Used dictionary-base method to do sentence classification; 3) Summarized and prepared tags which are manually tagged
- PPT: Dictionary-base method introduction and financial analysis part in response type part
- Report: Data, Methodologies, Result
- Completed readme.docx

**GU Leling**

- Searched useful data from WRDS
- Manually tagged responses for model training for twice
- Calculated the bigram frequency of response contents generally and by group
- Carried out the overall analysis part with WANG Xueru
- Responsible for the presentation of overall analysis section, slide 4 in this part, as well as corresponding report writing

**JIA Weiwei**

- Coding: 1) Analyzed companies' financial performance by YOY comparison and consecutive quarter comparison; 2) Analyzed stock performance compared with the S&P index; 3) Applied NER to extract locations and some part of speech words from S&P business description; 4) Conducted net income YOY comparison for speed analysis
- Presentation: Speed analysis
- PPT: Speed analysis
- Reports: 1) Speed analysis; 2) Revised overall analysis and case study in Results
- Others: 1) Collected S&P 500 company list and their basic information; 2) Collected S&P index data; 3) Manually tagged sentences for model training
-

**LI Zixin**

- Coding: 1) Built key words dictionary for targeting COVID-19 response sentences process; 2)Supervised learning with BERT to build sentence classification prediction model; 3)used PCA and 3D plot to visualize the company cluster result 4)Analyzed the company industry and net income performances to see the relationship between them and companies' responses
- PPT: Collected alternative template, prepare content of BERT method introduction and industry analysis part in response type part
- Manually tagged sentences for model training
- Report: Integrated content and finished the final formatting, prepared content of Data, Methodologies, Result, Limitations and Future work
- Completed readme.docx

**QU Ying**

- Constructed ideas and analysis structures, supervise and set up time for each meeting
- Coding: 1) Web scrapped for earnings call script; 2) All data cleaning, aggregation and normalization; 3) Supervised learning with rules to target COVID-19 response sentences and filter out garbage information
- PPT: Background, Data and Methodologies, Conclusions and suggestions
- Report: Introduction, Data, Methodologies, Limitations and Future work
- Completed readme.docx

**WU Yi**

- Organized Meeting Minutes
- Collected all Financial Data for S&P 500 companies
- Manually tagged sentences for model training
- Coding: 1) Used Word2Vec Model and Document Clustering via PCA; 2) TF-IDF Model for Speed Analysis
- Presentation: Speed Analysis
- PPT: Speed Analysis, Final Formatting
- Report: Data, Methodologies and Result
- Completed readme.docx

**WANG Xueru**

- Summarized and write Initial Proposal
- Manually tagged responses for model training
- Calculated the tf-idf score of the words classified according to tag1-5
- Carried out the overall analysis part with GU Leling

- Responsible for slides 1-3 in the overall analysis section, as well as corresponding report writing.

**WANG Shu**
- Data preprocessing for unsupervised method
- Manually collected 100 companies' response sentences
- Manually tagged sentences for model training
- Case study and in charge of both presentation and report part

**YANG Kaixi**
- Manually collected 100 companies' response sentences
- Manually tagged 3500 pieces of sentences for model training
- Found positive case and do corresponding case analysis
- In charge of positive case study presentation and report