
GEN BUS 760 BUSINESS DATA TECHNOLOGY
Final Project

Sentiment Analysis on Global Protests in 2019

Ziqian Deng

December 11, 2019
Wisconsin Business School

1. INTRODUCTION

Over the past year, mass protests have broken out around the world. Riots are still roiling through cities including Hong Kong, China, Catalonia, Spain and Paris, France. It is been six month since protests started in Hong Kong and on December 8, estimated 800,000 demonstrators again poured into Hong Kong's major thoroughfares, calling on the government to meet their demands; One of the biggest strikes in France in years has turned into consecutive days of protests against government's plans to overhaul the pension system. In Brussels, Belgium, thousands of people holding hands formed a human chain in central Brussels on Sunday to draw public attention to the need for urgent, joint action against climate change.

Although every demonstration has a distinct cause and goal and protestors across different countries have various methods of expressing demands, there are some common themes that connect them. While thousands of miles apart, protests have begun for similar reasons in several countries, and some have taken inspiration from each other on how to organize and advance their goals. People in Chile, Ecuador, and Bolivia are marching on the streets due to the deteriorating levels of income equality. In Lebanon and Iraq, claims of government corruption are at the heart of the protests and are closely linked to the issue of inequality. Demonstrations that were started this summer in Hong Kong are driven by protestors' opposition to the Fugitive Offenders amendment bill and demands for more political freedom. Activists from the Environmental campaign group Extinction Rebellion have been protesting in cities including London, Berlin, New York and Sydney, as they demand urgent action from governments.

Based on such unprecedented frequency and range, 2019 is believed to become the year of global protests. Therefore, this report is conducted to analyze the sentiment discrepancy across the world by scraping text data from Twitter and mainstream newspaper, aiming at extracting different attitudes against the widespread protests and detecting the potential correlation between protests and economy.

2. METHODOLOGY

- Tweets Scraping

Scraping data is the most vital part of this project because each step of analysis is based on the data collected from Twitter. R and its package *rtweet* are leveraged to scraping tweets. However, standard Twitter API only allows Twitter development account users to gather tweets posted within the recent 10 days, which causes a lack of tweets from certain search queries. In order to collect more data for analysis purposes, a premium Twitter API is required. Fortunately, Twitter provides limited free trials for a premium API which mitigates the problem of shortage of tweets.

- Sentiment Analysis

Getting the sentiment from the tweets is the essential stage. First, the noises and stop words are cleaned out by using R function *gsub*. Then *syuzhet* package is utilized for sentiment analysis, which decomposes the emotions and sentiments of each word. Also, Python and its package *TextBlob* are used for analyzing sentiments of the mainstream newspapers from the United States, the United Kingdom, and China.

- Visualization

Visualization is conducted by R and Tableau. R package *wordcloud2* which is the updated version of package *wordcloud*, performs better in creating a word cloud. Also, R package *plotly* is used to produce bar charts of the words that have a high frequency. Besides, Tableau is used for obtaining the global sentiment maps and producing time series line graph of Google search interest. Additionally, Excel helps to plot other supportive graphs.

- Acquisition of Other Data

The change in Google search interest is downloaded from Google Trends.

3. ANALYSIS

Overview of Global Protests in 2019

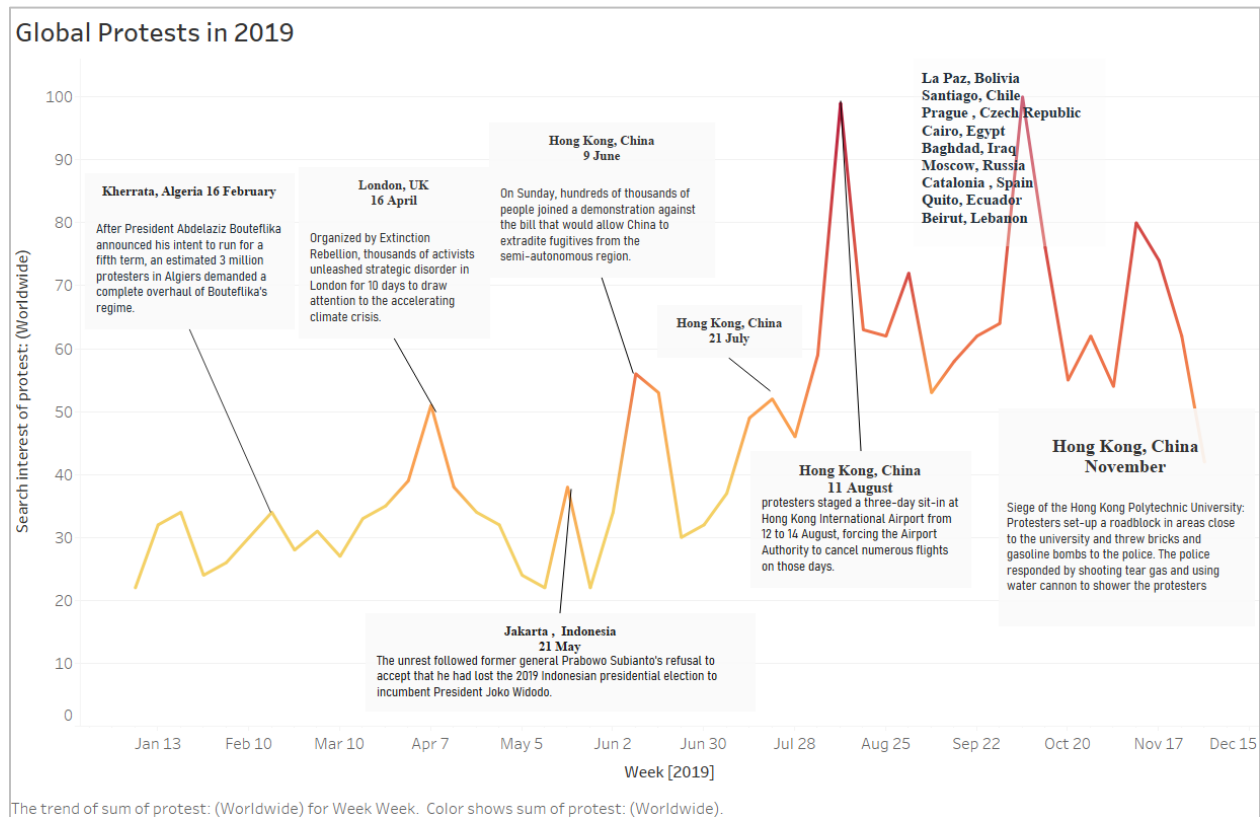


Figure 1: The trend of search interest of keyword “protest”

The line graph demonstrates the changing trending of the Google search interest from January to December 2019. Most of the spikes can correspond to a protest that broke out at that time. In February, an estimated 3 million protesters in Algiers demanded a complete overhaul of Bouteflika's regime. In April, thousands of activists unleashed strategic disorder in London for 10 days to draw attention to the accelerating climate crisis. In June, the ongoing and endless Hong Kong demonstration started to fight against the Fugitive Offenders amendment bill. Between October and December, more and more protestors from different cities or countries got involved in the movements, which is the primary reason that the search interest stayed at a high level during that period.

Although different protests are triggered by various events or policies and protestors have various demands and goals, there are some common themes that connect them. Basically, those themes can be divided into 4 categories ---

national income inequality, government corruption, citizen political freedom, and global climate change. Therefore, the following sentiment analysis is conducted by selecting a certain country's protests for each category, to get an insight into sentiment distribution for different kinds of movements.

National Income Inequality --- 2019 Ecuadorian Protests

In October, A series of protests and riots erupted in Ecuador over President Lenin Moreno's austerity measures that proposed ending fuel subsidies and cutting the benefits and salaries of civil servant. The Ecuadorians feared that the measure would result in increased costs for public transport and food.

To obtain the sentiment scores, a total of 1484 tweets created between November 6 and December 5 were collected by searching the keyword “Ecuador protest.”

A bar chart and a word cloud are created to demonstrate the high-frequency words that appear a lot among the tweets. Meaningful information can be extracted by studying those words. First, it is not hard to find out that some other countries' names appear. Twitter users would also mention other countries such as Bolivia, Venezuela, Chile, Colombia, Haiti, Peru and Argentina when they were talking about Ecuador. All of these are the countries that close to Ecuador and most of them were under mass protests or unrests during the past few months. It seems that the worldwide and viral movement has roiled over South America. Second, words like "policies", "strike" and "fighting" reveals the deteriorating conflicts between policies and protestors to a certain extend. Words like "neoliberal", "truth", "corruption" and "austerity" implies the trigger of the protests is the new austerity policies and one of the goals of the demonstrators is urging the government to lower the price through neoliberal policies.

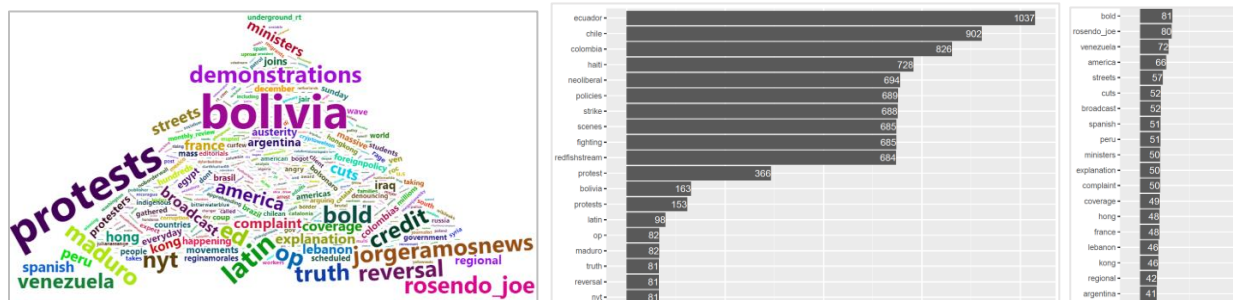
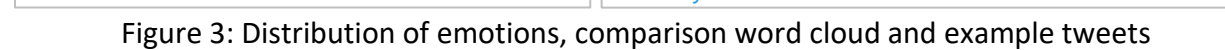


Figure 2: The high-frequency words related to Ecuador protest

Emotion Category	Count (approx.)
negative	1950
anger	1700
pride	600
trust	400
fear	250
surprise	220
disgust	200
anticipation	180
joy	100
sadness	100



In addition, a global map is drawn for showing sentiment scores on Ecuador protests by different countries. It is evident that nearly every country has a negative sentiment on Ecuador protests, especially the countries that are close to Ecuador. Peru has the lowest sentiment score (-3.75), Brazil has a relatively low score (-1.26) and Venezuela also has a low score (-1.39). This further supports the fact that the protests have gone too far and have greatly annoyed people. Another finding is that Russia and China feel negative about Ecuador protests as well (-1.80 and -1.70), probably because the mass protests that happened in Russia and China recently have people feeling the same way as South American countries.

Sentiment scores on Ecuador protests

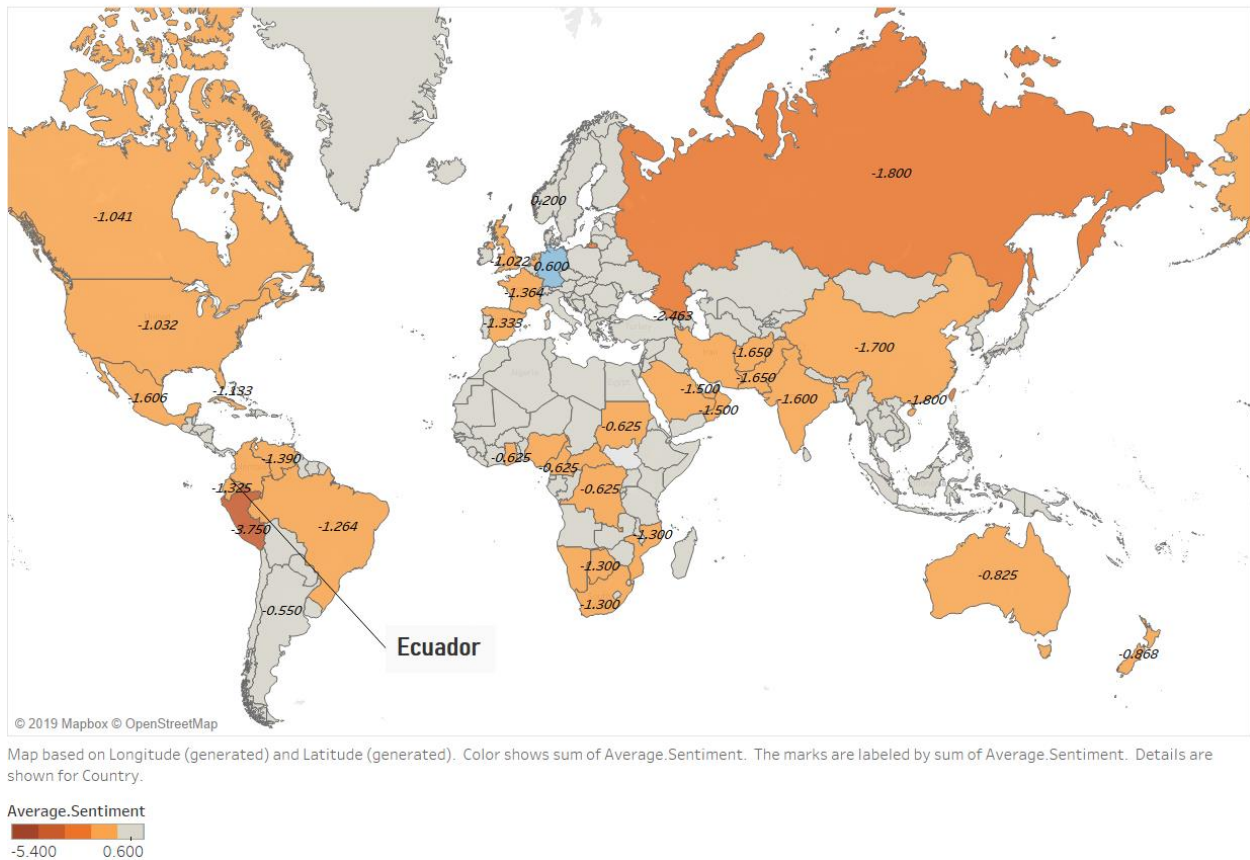


Figure 4: Sentiment scores on Ecuador protests around the world

Government Corruption --- 2019 Lebanese Protests

After dissecting tweets about Ecuador, the focus of the analysis is shifted to a Middle East country --- Lebanon. Since October, the country-wide protests in Lebanon has sparked attentions on Twitter. The protesters are demanding an end to corruption, calling for a new government made up entirely by non-politicians. They are requesting more jobs and improved services such as electricity, water and health care.

For this section of the analysis, 1310 tweets with keyword "Lebanon protest" were gathered covering from November 28 to December 7.

Same as Ecuador protests, the high-frequency words are firstly explored. Likewise, other protests from countries in the vicinity of Lebanon are noticed. Words "iranprotests" and "iraqprotests" are frequently mentioned. Second, the "economic", "corruption" and "government" suggest that Lebanese protests focus on the economy and widespread corruption. In addition, the word "women" also indicates a distinct feature of Lebanese protests ---

women have played a critical role in leading the uprising. Therefore, this finding shows that the Lebanese protests are not only against corruption but also sexual harassment and rape.

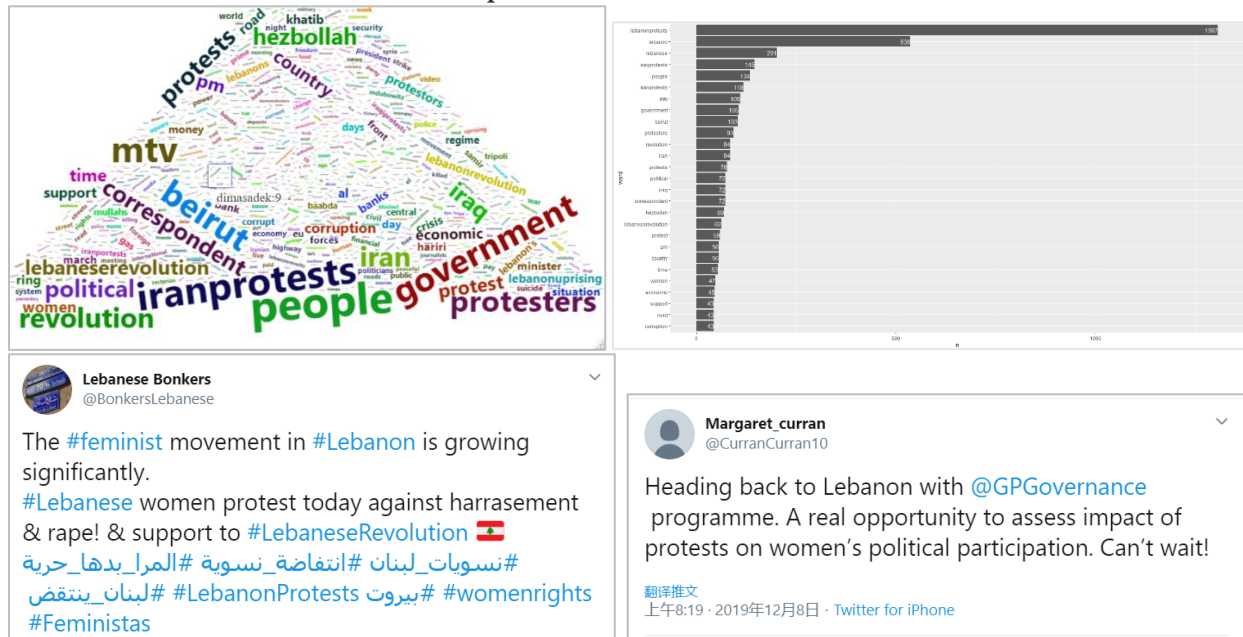


Figure 5: The high-frequency words related to Lebanon protest and example tweets

Unlike the sentiment categories of Ecuadorian protests, the positive sentiment of the Lebanese protests accounts for the biggest proportion of the tweets --- over 1600 positive words were tweeted among a total of 1310 tweets, which signifies that the Lebanese protests are supported by a considerable amount of people around the world. The conclusion can be further backed by the positive word cloud below. Words such as “support”, “trust” and “approve” show the protests have lots of supporters. Some sentiment words like “brave”, “love”, “amazing”, “happy” and “bless” illustrate the emotions of supporters.

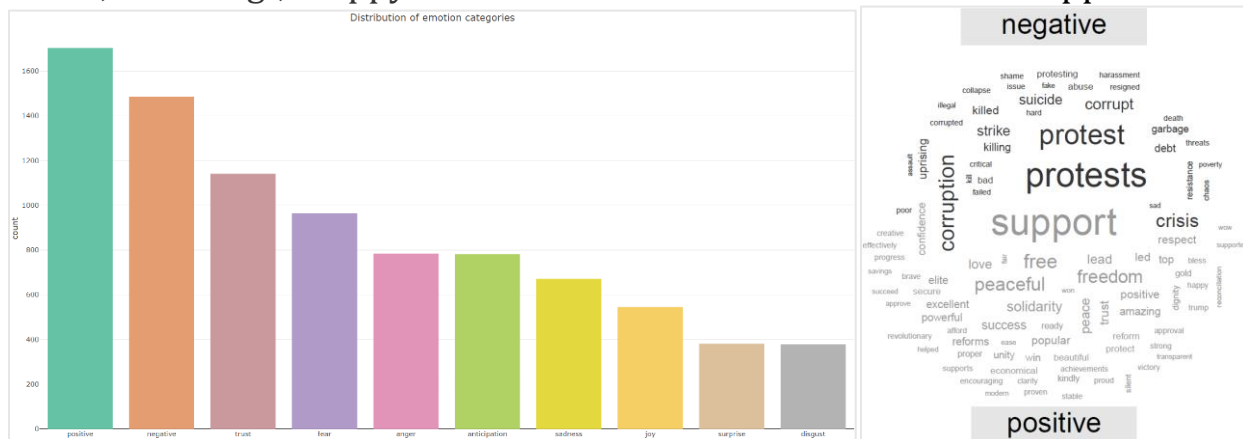


Figure 6: Distribution of emotions and comparison word cloud related to Lebanese protests

As shown in the emotion category bar chart, the negative sentiment also takes up a large percentage. Therefore, for the global sentiment distribution of Lebanese protests, while part of the South Asian countries, East Asian countries and Australia contribute to the positive sentiments, most countries around the world show the negative sentiments, especially the countries in southern Africa. This abnormal correlation between southern Africa countries and Lebanon may be brought about by the extra-marital affair between Lebanese prime minister Saad Hariri and South African model Candice van der Merwe. On September 30, the New York Times reported that Hariri gave Van der Merwe \$15.3 million in a “donation” in 2013 to buy property, which helped fuel widespread protests against corruption in Lebanon.

Sentiment scores on Lebanon protests

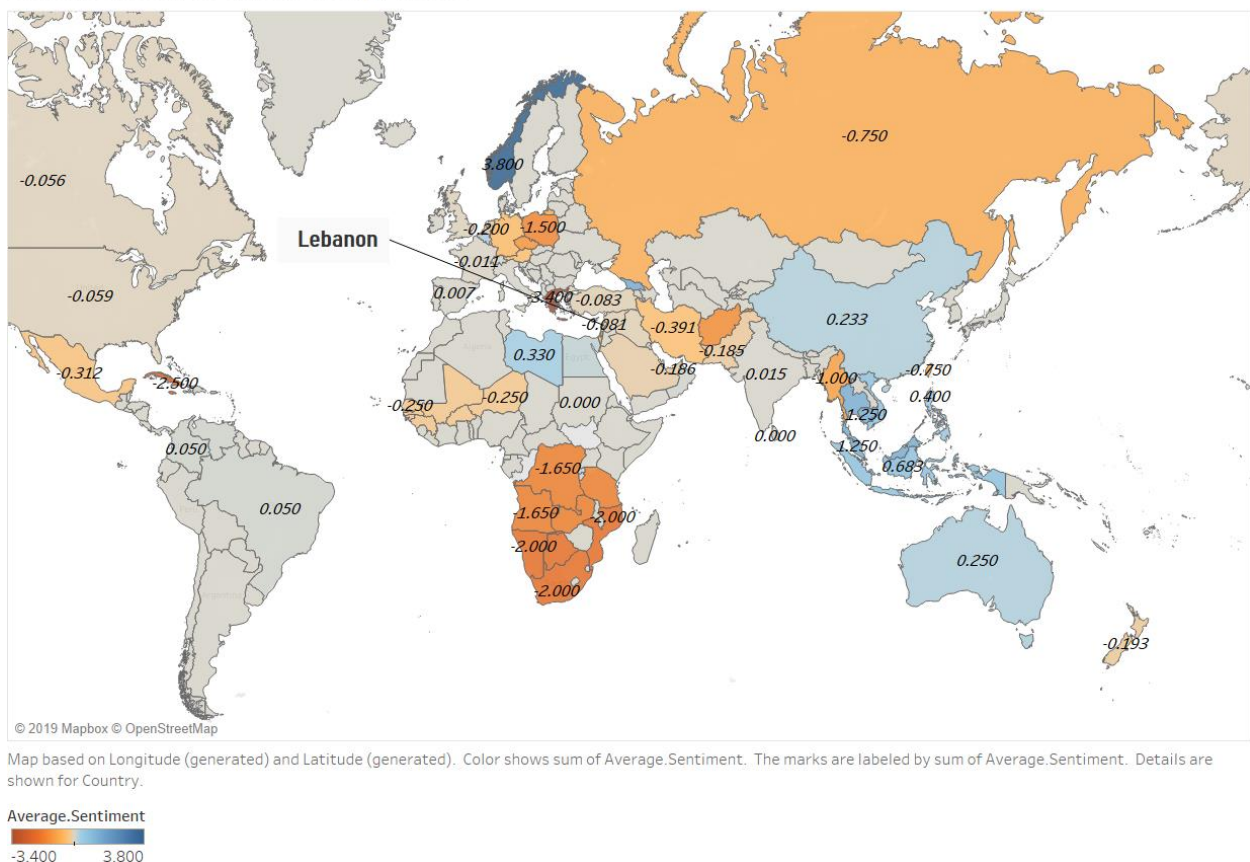


Figure 7: Sentiment scores on Lebanon protests around the world

Citizen Political Freedom --- 2019 Hong Kong protests

Since this summer, Hong Kong has witnessed one of the longest protests in history. The ongoing series of demonstrations in Hong Kong triggered by the introduction of the Fugitive Offenders amendment bill by the Hong Kong Government. Opponents said this new measure risked exposing Hongkongers

to unfair trials and violent treatment. They also argued the bill would give China greater influence over Hong Kong and could be used to target activists and journalists. After the marching and demonstrations during early summer, the protests have rapidly developed into city-wide riots and chaos, where the subways and roads across the city have been blocked with burning fires or bricks. And the opinions towards the protests are so sharply polarized that the situation shows no sign of dying down. Therefore, researching on the emotion distribution might help better understanding the public opinions.

The topic of Hong Kong protests is prevailing on Twitter, which enables numerous tweets to be available for collection. A total of 16936 tweets created between November 28 and December 8 are gathered.

Similar to the one for Lebanon protests, the emotion bar chart for Hong Kong protests shows that positive words dominate among the tweets (nearly 20,000 words). Also, negative emotion words have second place (nearly 15,000 words). For the positive side in the comparison word cloud, various derivations of verb “support” indicate more individuals choose to stand with Hong Kong protestors. Secondly, the negative and positive corpus are so contradictory that reach a conclusion that the supporters and opponents have completely different perception of the movements. While the opponents describe the protests as “unrest”, “chaos”, “killing” and “nightmare” which result in “recession” and “slump”, the supporters fancy a “peaceful”, “happy” and “harmless” “victory.”

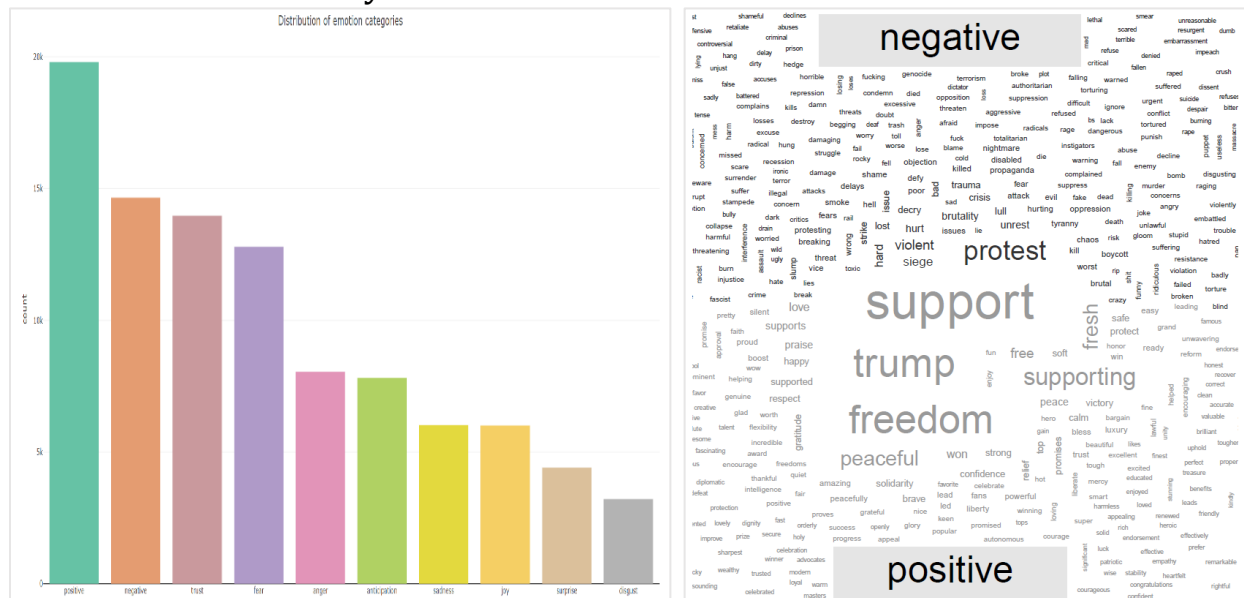
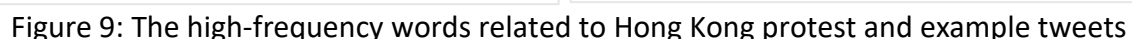


Figure 8: Distribution of emotions and comparison word cloud related to Hong Kong protests

For more detail, the high-frequency words can suggest what the tweets are most focusing on. Words such as “hongkongpoliceterrorism” and “tear gas” indicate that the Hong Kong police is under harsh criticism. And “democracy”, “5demandsnot1less”, “right” and “freedom” again articulate the underlying reasons for demonstration. “5demandsnot1less” stands for “five key demands are all needed to be satisfied and none of them can be rejected” which is a slogan of the pro-democracy protestors. The five demands include the withdrawal of the bill, investigation into alleged police brutality and misconduct, the release of arrested protesters, etc.



Advanced finding can be discovered through analyzing the global sentiment map. According to the sentiment distribution, all the countries can be roughly separated into two group --- western countries and the other countries including African, Australian and Asian countries.

Global scores on Hong Kong protests

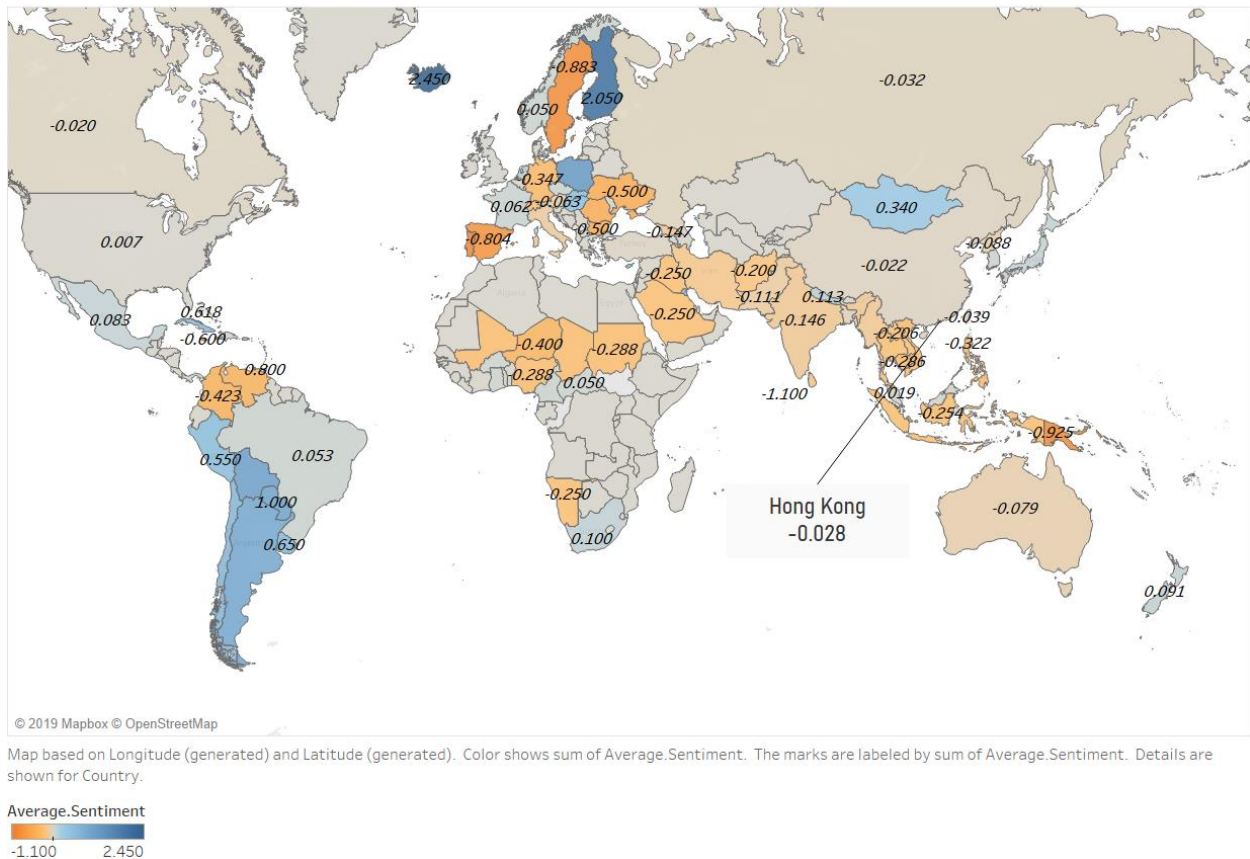


Figure 10: Sentiment scores on Hong Kong protests around the world

Western countries such as the United State, Canada, South American countries and some European countries all convey the positive sentiment towards Hong Kong protests. Such a combination indicates western countries might be the financial group that manipulates and finances the Hong Kong protestors. This assumption is in line with the Hong Kong protest history. The past protests that happened in Hong Kong such as the 2014 Umbrella Revolution, were confirmed being funded by National Endowment for Democracy which is a U.S. non-profit soft power organization that was founded in 1983 with the stated goal of promoting democracy abroad. Last month, the US enacted the so-called Hong Kong Human Rights and Democracy Act of 2019, intending to meddle in the internal affairs of China. Different facts are implying that the

Hong Kong demonstrations are not just because Hongkongers aim to combat the extradition law, more because of the instigation of western countries.

On the other hand, China is opposed to the movement with a doubt. At the same time, African countries and other Asian countries also stand by the Chinese government, showing a low sentiment score. This probably because China has been maintaining a good and solid partnership with them and the Belt and Road Initiative proposed by China in 2013 to improve connectivity and cooperation on a transcontinental scale has created fruitful results so far.

Global Climate Change --- Extinction Rebellion

While the first 3 protests are the movements that broke out in a country, the Extinction Rebellion is a global environmental movement. Launched in 2018, it describes itself as an international "non-violent civil disobedience activist movement" and its organizers say it now has groups willing to take action in dozens of countries. On April 16, thousands of activists initialized the protests in London by barricading roads and bridges at major city landmarks.

For the Extinction Rebellion, a total of 9639 tweets are scraped, with time period ranging from November 28 to December 8.

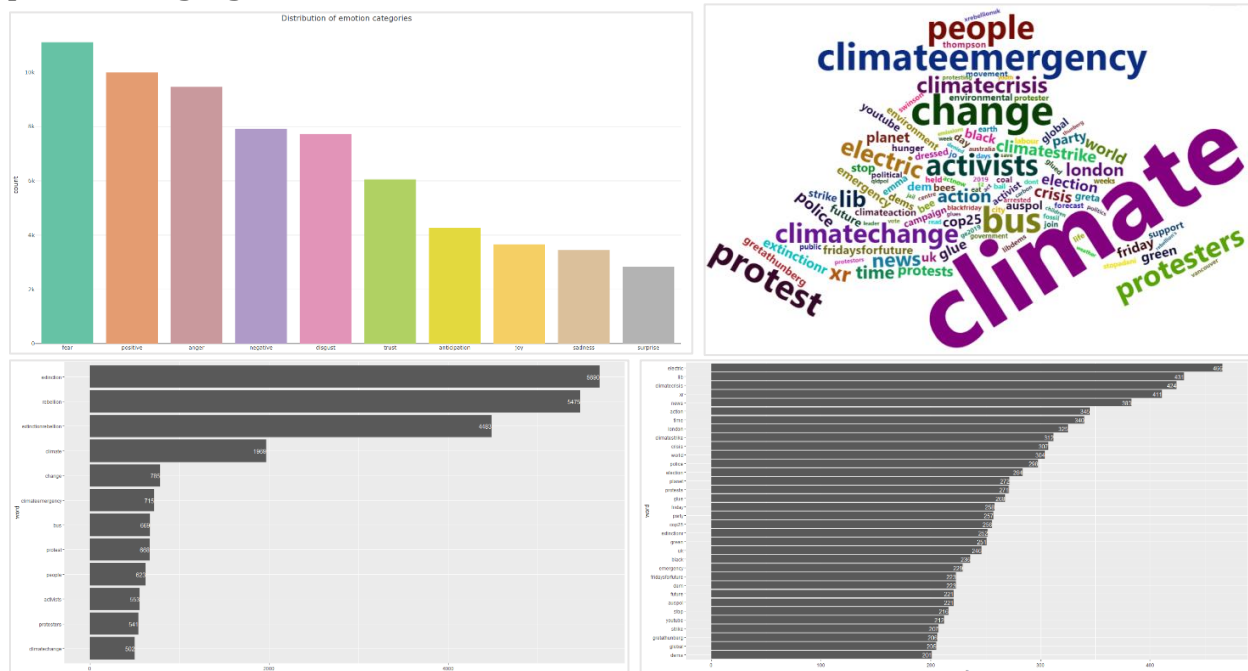
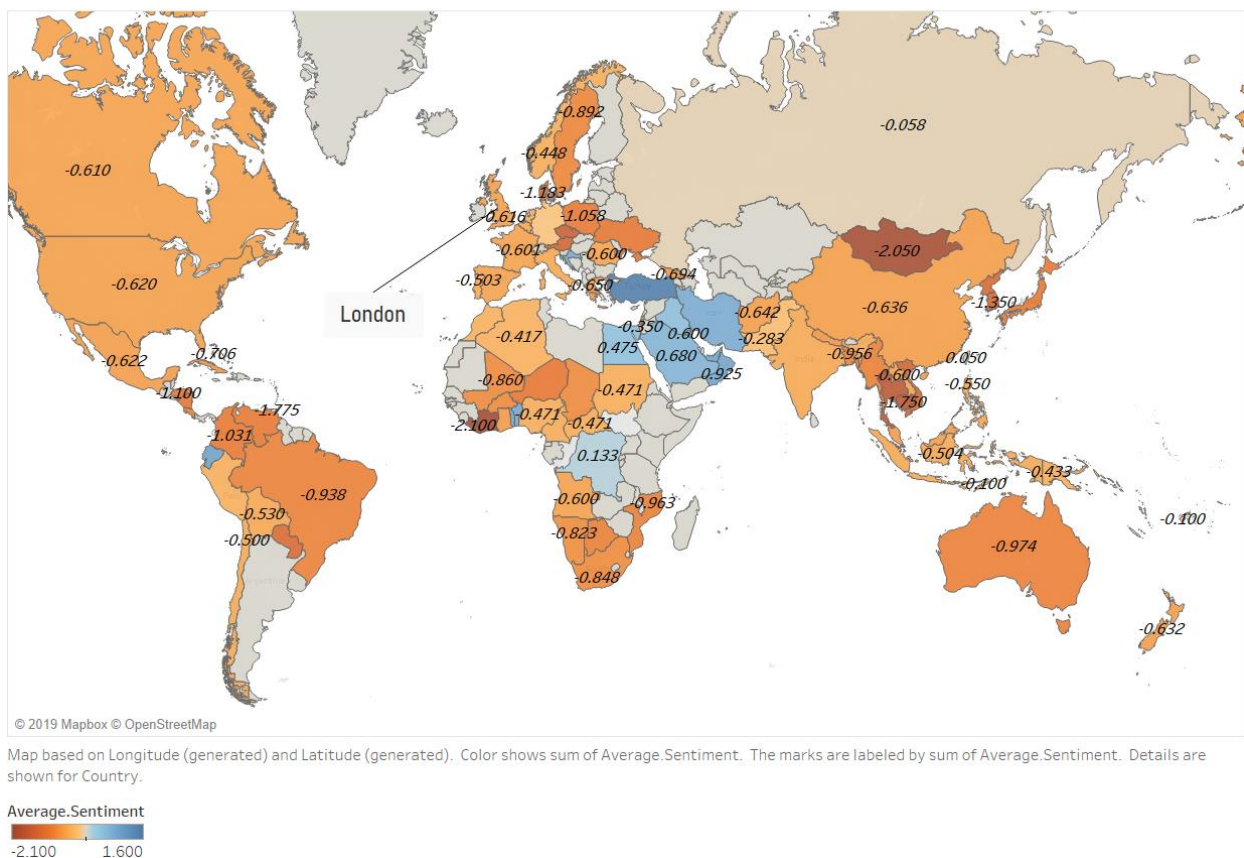


Figure 11: Distribution of emotions and high-frequency words related to Extinction Rebellion

Surprisingly, the highest emotion turns out to be “fear”, followed by “positive” and “anger”. However, it is hard to find out the cause of fear by looking at the high-frequency words. No words seem to have a connection with fear. One potential reason that people show fear and anger is that the ongoing protests have already extremely disrupted national normal running. On October 17, activists from Extinction Rebellion group disrupted London’s public transport network during rush hour and commuters dragged a protester from the roof of a train and set upon him.

Not limited to London, nearly every country in the world has a negative sentiment about the Extinction Rebellion movement as shown below. However, a few of the Middle East countries such as Turkey, Iraq, Iran and Saudi Arabia are delivering a positive sentiment, which leads to a hypothesis that these countries might be organizing and manipulating the movements all over the world.

Global scores on Extinction Rebellion



4. CONCLUSION

Protests have served as an outlet for marginalized groups to articulate their frustrations and make their demands known. This year, protests like Global Extinction Rebellion movements and Hong Kong's ongoing mass demonstrations are calling for governors to change their policies. The action of marching on streets or staging a silent sit-in is justified and considered as a brave move, since every individual has the right to freedom of opinion and expression. However, many of those protests have gone too far and disobeyed the initial demands of the movements. Moreover, the protests have progressively been turned into countrywide and even global turmoil and uprising, which throws entire world into disorder. Blocked roads and burning campuses have already disrupted the daily routine of people ranging from students to senior citizens. For instance, Hong Kong has been suffering from endless demonstrations for 6 months, accompanied by the mass strike and economic recession.

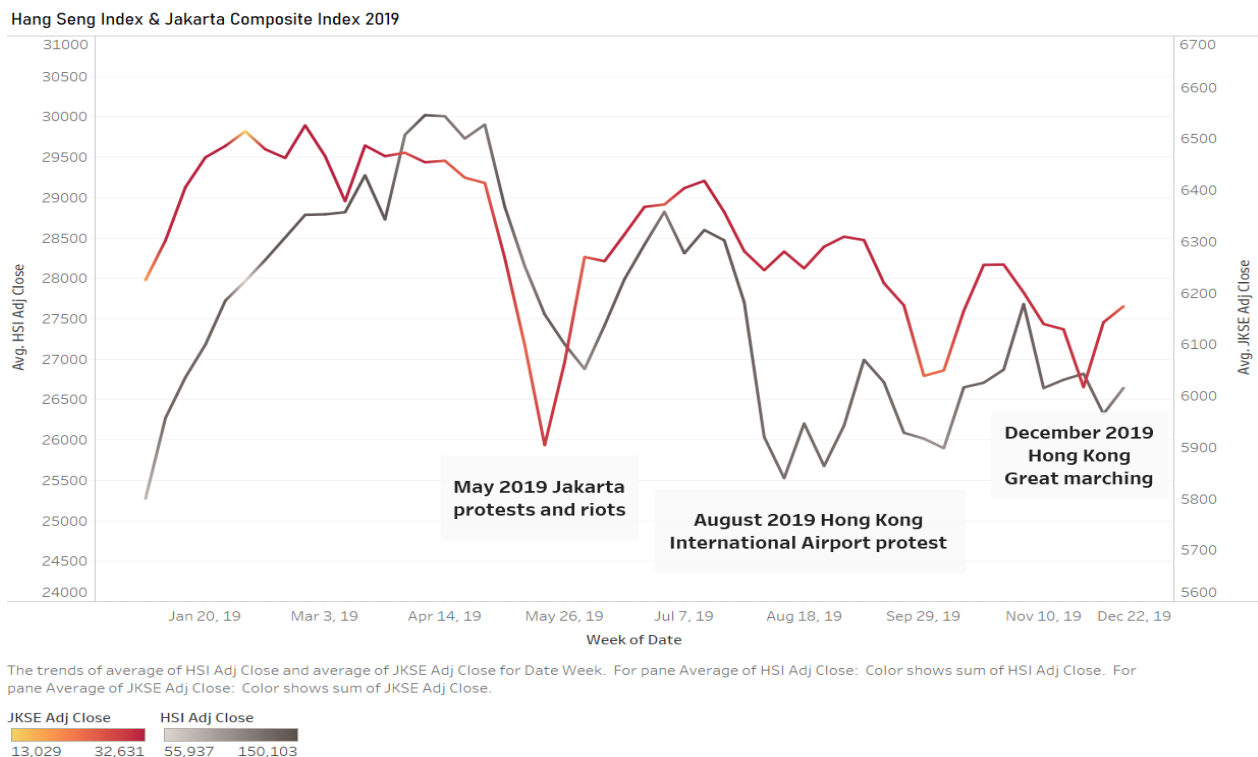


Figure 13: 2019 Hang Seng Index and Jakarta Composite Index by weeks

By comparing the stock market index with the timing of the protests, the highly positive correlation between them can be easily detected. As shown in the above dual line graph, when mass protests and riots were roiling over

Jakarta, Indonesia, the local stock index --- Jakarta Composite Index (JKSE) has witnessed a huge slump during May. From May 2 to May 21, JKSE has dropped by nearly 10%. Same as the stock market in International financial harbor Hong Kong, its stock index --- Hang Seng Index has undergoing multiple plunges since this summer. On August 12, protestors staged a three-day sit-in at Hong Kong International Airport, forcing the Airport Authority to cancel numerous flights on those days. Meanwhile, in the stock market, the Hang Seng Index has dropped to the lowest point (25,281) since February.

Based on the above discussions, while the outcomes of calling for government's acting on global climate deterioration or demanding for democracy and freedom are unknown, such global protests and the following unrests indeed have given rise to the inconvenience of citizens and recession on economic. Since most of the ongoing movements haven't shown any sign of slowing down, the widespread unrests are believed to keep disrupting every aspect of human beings.

5. REFLECTION

Data Technology for Business Analytics is my favorite course that I have enjoyed taking this semester. After exposing to various data-related tools, I had my interest in exploring data analysis sparked. I am aware that the skills I have obtained from this course are entry-level. Therefore, further independent study on leveraging R, Python and SQL are necessary, so that I can equip myself with sufficient knowledge and experience when I step into the industries in the future.

I am also grateful for having the chance to take the course of Professor Terence Ow. He is the most enthusiastic and dedicated professor that I have met so far. I believe that the experience of this semester will definitely pave my way to my ideal position and industry.