

Tweet Sentimental Analysis

Joi Chu-Ketterer

DSC 680

Bellevue University

April 12, 2020

lchuketterer@my365.bellevue.edu

Instructor: Catharine Williams

Introduction

Social media has no doubt made an impact on society and continues to be a social powerhouse. With over 3.5 billion users [Mohsin, 2020], or 45% of the world's population, engaging with social media in some capacity makes no question the industry is a gold mine for information. Additionally, the ability to share information instantly and for users to share their opinions allows for quick shifts in company-user dynamics. One 'wrong' move on the company's part can lead to great financial uncertainties. For this reason, sentiment analysis is an extremely useful tool to gauge the public's sentiment towards a company.

In light of recent events, many companies have adjusted their operations to meet the safety and health regulations set by the World Health Organization (WHO) and Centers for Disease Control and Prevention (CDC) regarding COVID-19 [See Appendix A]. COVID-19 is a virus that is believed to have originated at a market in Wuhan, Hubei in China [Hassanin, 2020]. The first cases were diagnosed in late 2019, and have since continued to spread across the globe. As of April 11, 2020, there are 1,777,666 confirmed cases, 108,866 deaths, and 365,631 recovered [Gutierrez, 2020]. While countless businesses have been affected by the pandemic, this project aims to understand how the public's perception of the Walt Disney Company Parks and Resorts sector has changed.

Business Understanding

The Walt Disney Company is considered a leader in many industries, but most known for their theme parks. With six resorts located around the world, the resorts alone bring in the most revenue across all sectors. Parks and Resorts continues to increase their revenue, estimating around 21 billion dollars in 2018 [Lock, 2020]. After announcing their Hong Kong and Shanghai resort closures for January 26, Tokyo resort closures for February 29, Paris resort closures for March 15, and lastly followed by their U.S. resort closures for March 15 and 16 (*Figure 1*), social media posts about the company continue to sky rocket. While the company is undoubtedly losing money during these closures, understanding the public's perception of their actions during this unprecedented time will be vital to understanding how to relate to their Guests.

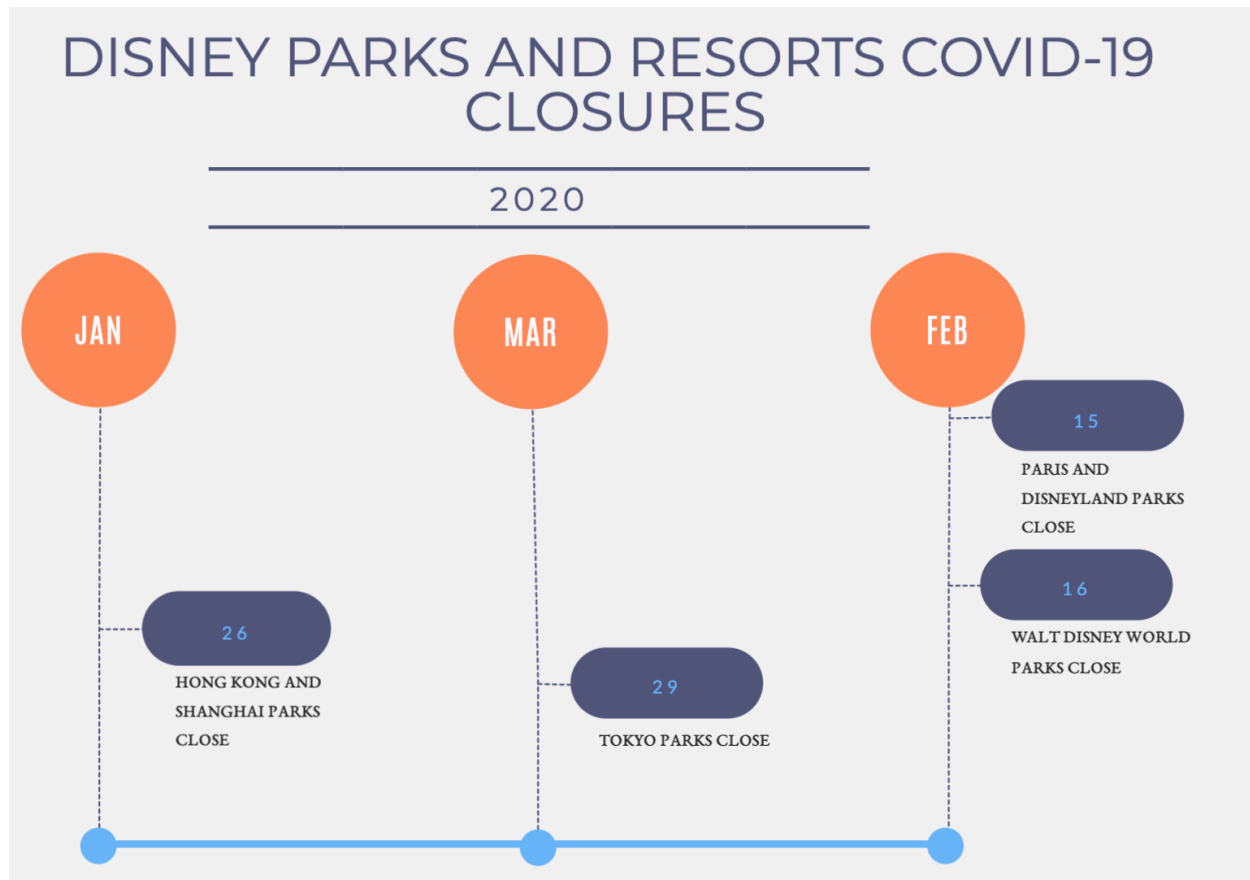


Figure 1. Timeline showing the closures of the Disney Parks and Resorts.

Sources: <https://www.oregister.com/2020/01/26/hong-kong-disneyland-closes-amid-chinas-coronavirus-outbreak/>, <https://variety.com/2020/biz/news/shanghai-disney-resort-disneyland-reopen-china-coronavirus-1203527299/>, <https://english.kyodonews.net/news/2020/02/24fdc61d5646-breaking-news-tokyo-disneyland-to-close-through-march-15-over-coronavirus-operator.html>, <https://newsdio.com/disneyland-disney-world-disneyland-paris-to-be-temporarily-closed-amid-fears-of-coronavirus-2/69188/>

Twitter is one of the more popular social media outlets in the United States with over 70 million users [Tien, 2018]. With just 280 characters, users are sharing their thoughts quickly and often. A simple Twitter search for *Disney* shows users are tweeting about Disney every 3 seconds or so. With so many opinions being shared, knowing if the general public feels positive, negative, or neutral towards the company can help them plan a more strategic re-opening.

Analysis

Data Extraction

To understand the public's sentiment towards the Walt Disney Parks and Resorts during their COVID-19 closures, text analysis was conducted on user's tweets. Tweets of interest were identified by their hashtags; including, but not limited to, *#Disney*, *#Disneyland*, *#WaltDisneyWorld*. Relevant tweets were extracted using the Twitter API in Python, and analyzed using text analysis packages. Findings were summarized and presented visualizations created in Tableau and with Matplotlib.

Method

With an ever changing social climate, extracing long periods of tweets provides a good sense into how the public's opinion is changing over time. While a constant data stream was not an option for this particular project, five hour or more bursts of tweets were extracted at a time, during the following time segments:

- March 18 – 19
- March 21 – 24
- March 27 – 30

Tweets were extracted and saved into .csv files and imported into a Jupyter notebook for analysis.

Data Preparation

Each extracted tweet had over 16 attributes, most of which were irrelevant. Many attributes related to the retweet count of a single tweet. While this may be useful data for other projects, it would not affect the results of this particular project. With sentiment analysis, if a tweet is retweeted three times, the tweet will get extracted three times and counted that way. If the analysis included the retweet counts, the tweets would be overcounted. Additional preparation methods include stripping the entries of characters like '@Text:', '@Created on:', and converting all text to lowercase.

Results

Sentiment Analysis

After parsing the polarity of each tweet between *positive*, *negative*, and *neutral*, Figure 2 shows a line plot indicating that a majority of tweets about Disney were either neutral or positive. Negative tweets remain the minority of the three, although there are some spikes on March 22 and March 23. This was only a few days after the last resort closure announcements and may have impacted the public's opinion. Conversely, the count of positive tweets increased starting mid-day March 23 and stayed relatively consistent through March 25. This could be due to their choice to continue paying Cast Members [Alexander, 2020]. Additional announcements that correlate with the increase in positive sentiment include Disney+ launching in certain areas in Europe [Callaham, 2020], Megan Markle returning to Hollywood in some Disney projects [Hallemann, 2020], and Disney Senior Executives announcing large pay cuts ranging from 20% to 50% [Parker, 2020].

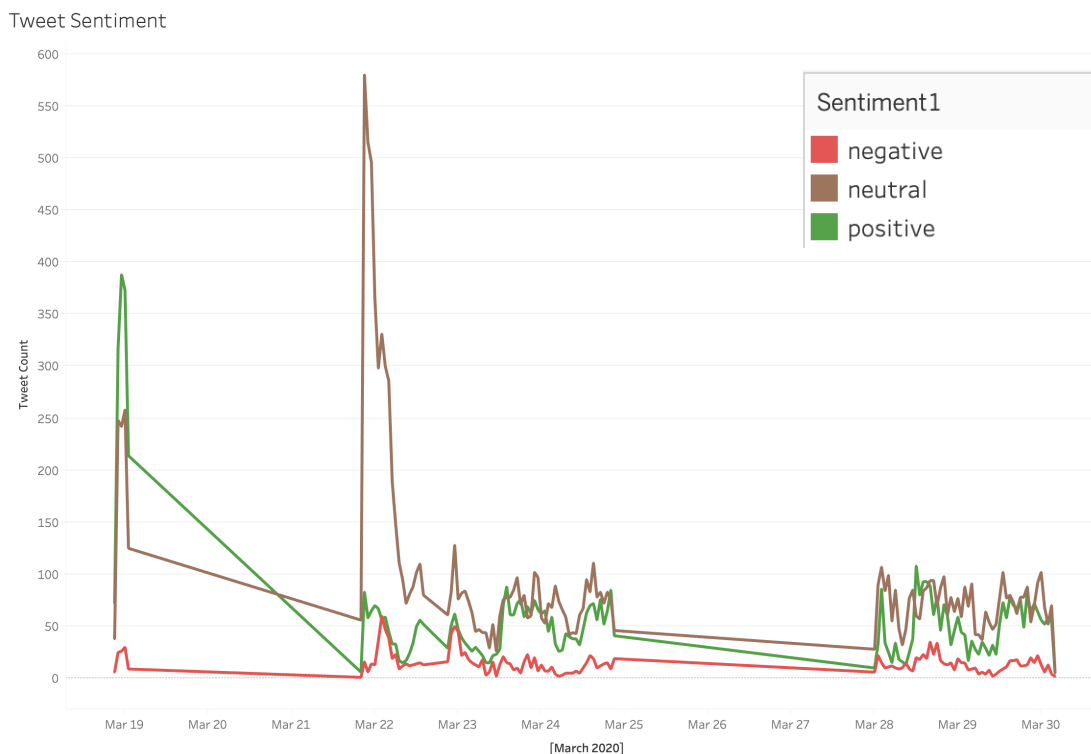


Figure 2. Twitter tweet sentiment towards the Walt Disney Company between March 19 - 30. No data was collected between March 20 -21, and March 25 – 27, leaving those periods inconclusive.

Aditonal Insights

In addition to a sentiment analysis, other exploratory analysis was conducted as a means to gain better insight. For instance, the hashtags used in each tweet were added to a dictionary and counted. *Figure 3* shows the top ten hashtags used in the tweets, with *#disneyland* holding the highest count. Converting all text to lowercase helped with the hashtag count by removing the difference between *#WaltDsineyWorld* and *#waltdisneyworld*, since they hold the same value within the scope of this project.

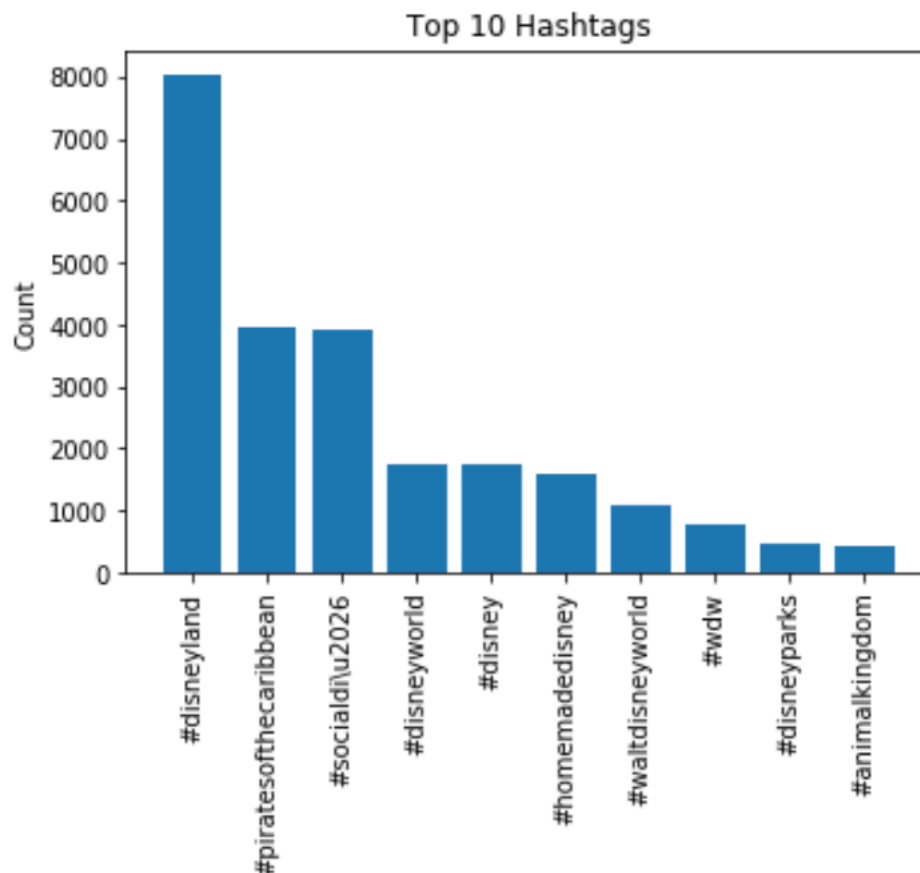


Figure 3. Total hashtag counts for the top ten hashtags used in analyzed tweets between March 18 – March 30.

While the tweet extraction process flagged all tweets relating to Disney’s domestic and international resorts, analysis shows that Twitter users are talking about domestic parks much more. This could be due to the timing in tweet extraction since most international parks were closed by February and this project started in March. Additionally, social media platforms tend to be region specific. For instance,

while Twitter is most popular in the United States, WeChat corners the market in China. For this reason, extracting social media comments from a single platform inherently introduces bias into the project. As such, no conclusions regarding the interest between domestic and international properties can be made at this time. However, data suggests that more people are interested in the Disneyland Resorts over the Walt Disney World Resorts. It is interesting to see that the Pirates of Caribbean attraction is mentioned more times than Disney World itself.

Within the context of tweet sentiment, this project did not distinguish between retweets and original tweets. However, distinguishing between the two holds more significance during hashtag count analysis. Understanding what people directly choose to talk about is just as important as understanding how users feel. *Table 1* shows the counts for original tweets for the top three hashtags used, in addition to providing sentiment counts.

Table 1. Sentiment count for top three hashtags

	Positive	Neutral	Negative
#disneyland	1439	1139	61
#piratesofthecaribbean	2	4	0
#disneyworld	285	408	54

Table 1 is interesting to consider as it puts the total hashtag counts from *Figure 3* into perspective. While there were almost 4,000 total tweets with *#piratesofthecaribbean*, there were only 6 original tweets with the same hashtag. While this does not impact the goal of understanding the public's sentiment towards the company as described earlier in this report, it brings light to the power of retweeting. Additionally, more Disneyland tweets were positive while those about Disney World were neutral. Since all the public announcements by the company pertained to both parks, it's interesting to see the sentiment for both are different.

Benefits

Understanding how the public views your company is an incredible advantage for businesses working to increase their public presence. In the case of the Walt Disney Company, most of the tweets were either neutral or positive, indicating that they do not have a reputation problem at the moment. This could be an indication that they do not need to allocate as many resources to their media sector, but rather put more efforts into finance projects. In a time where your workforce is limited and revenue is tight, understanding what departments to prioritize is incredibly valuable to a company.

Additionally, with so many intellectual properties, Disney constantly highlights different ones depending on its current popularity. The hashtag count analysis indicates that the Pirates of Caribbean attraction is currently being talked about the most, and so incorporating it into future media posts will help them stay relevant and relatable. Regardless of the industry a company is in, understanding their users and audience should be a top priority.

Future Work

With a start to understanding the public's sentiment towards the Disney Parks and Resorts, there is much more analysis that can be achieved. If this project were to continue, setting up a constant data stream would help eliminate the gaps in data and provide a more cohesive look into the trends at play. It would also provide real time reactions following recent announcements. When the U.S. resorts announced their extended closures on March 27, 2020 around 1pm, additional tweet extractions did not continue until about five hours later. While this certainly still helps understand the trends starting from March 18, it does lose the impact the announcement might have made.

In addition to having a continuous stream of data, analysis into exploring the effects of retweeting and original tweets would be interesting to observe. One could argue that a retweet does not hold the same sentiment value as an original tweet given the lack of freedom to use their own words. While interesting, continuing this topic of analysis would be entering a new business domain of behavioral and psychological studies. To aid the success of such a study would require more robust sentiment polarity

outcomes. A broad sense of sentiment can easily be detected with *positive*, *neutral*, and *negative*. However, with so many emotions humans are capable of experiencing, it would certainly be more informative to break *positive* further down to *excited*, *hopeful*, or *grateful* and *negative* further down to *aggravated*, *guilty*, or *disappointed*. As the English language continues to evolve, so does sentiment analysis.

Conclusion

Despite being in an unprecedented time with communities filled with fear, the overall tone of Disney's fans continue to remain positive. Disney's commitment to keeping the *magic* alive for their audience through other segments like *Disney+* and *ShopDisney*, and continual pay for Cast Members seem to be working in their favor of keeping Disney fans' spirits high. Moreover, staying relevant during such times provides them with a great platform to bounce back on when the threats of COVID-19 lessen.

References

- Mar 15, 2020. "Interim Guidance for Coronavirus Disease 2019 (COVID-19)", *Centers for Disease Control and Prevention*. Retrieved from <https://www.cdc.gov/coronavirus/2019-ncov/community/large-events/mass-gatherings-ready-for-covid-19.html>
- 28 Feb, 2020. "Tokyo Disneyland, USJ to close through March 15 over coronavirus", *Kyodo News*. Retrieved from <https://english.kyodonews.net/news/2020/02/24fdc61d5646-breaking-news-tokyo-disneyland-to-close-through-march-15-over-coronavirus-operator.html>
- 13 Mar, 2020. "Disneyland, Disney World, Disneyland Paris to be temporarily closed amid fears of coronavirus", *NewsDio*. Retrieved from <https://newsdio.com/disneyland-disney-world-disneyland-paris-to-be-temporarily-closed-amid-fears-of-coronavirus-2/69188/>

- Alexander, Julia. 27 Mar, 2020. "In an unprecedented move, Disney parks will remain closed until further notice", *The Verge*. Retrieved from <https://www.theverge.com/2020/3/27/21197449/disney-land-disneyworld-park-closure-coronavirus-film-tv-streaming-espn-cruises>
- Bizjak, Tony. 15 Mar, 2020. "Gov. Newsome asks California bars to close, tells older residents to isolate due to coronavirus", *The Sacramento Bee*. Retrieved from <https://www.sacbee.com/news/coronavirus/article241212146.html>
- Callaham, John. 24 Mar, 2020. "Disney Plus is now available in the UK, Ireland, and other European markets", *Android Authority*. Retrieved from <https://www.androidauthority.com/disney-plus-europe-launch-1097183/>
- Davis, Rebecca. 8 Mar, 2020. "Shanghai Disney Resprt, Closed in Janurary due to Coronavirus, Set to particually Reopen Monday", *Variety*. Retrieved from <https://variety.com/2020/biz/news/shanghai-disney-resort-disneyland-reopen-china-coronavirus-1203527299/>
- Gutierrez, Pablo. 11 Apr, 2020. "Coronavirus world map: which countries have the most cases and deaths?", *The Guardian*. Retrieved from <https://www.theguardian.com/world/2020/apr/11/coronavirus-world-map-which-countries-have-the-most-cases-and-deaths>
- Halleman, Caroline. 26 Mar, 2020. "Meghan Marckle Signs a Voiceover Deal With Disney", *Town & Country*. Retrieved from <https://www.townandcountrymag.com/society/tradition/a30486892/meghan-markle-disney-voiceover-deal/>

Hassanin, Alexandre. 20 Mar, 2020. "Coronavirus origins: genome analysis suggests two viruses may have combined", *We Forum*. Retrieved from

<https://www.weforum.org/agenda/2020/03/coronavirus-origins-genome-analysis-covid19-data-science-bats-pangolins/>

Lock, S. 16 Jan, 2020. "Revenue of Walt Disney's parks and resorts segment worldwide 2008-2018", *Statista*. Retrieved from <https://www.statista.com/statistics/193221/revenue-of-walt-disneys-parks-and-resorts-since-2008/>

Macdonald, Brady. 26 Jan, 2020. "Hong Kong Disneyland closes amid China's coronavirus outbreak", *Orange County Register*. Retrieved from <https://www.ocregister.com/2020/01/26/hong-kong-disneyland-closes-amid-chinas-coronavirus-outbreak/>

Mohsin, Maryam. 7 Feb, 2020. "10 Social Media Statistics You Need to Know in 2020", *Oberlo*. Retrieved from <https://www.oberlo.com/blog/social-media-marketing-statistics>

Parker, Ryan. 30 Mar, 2020. "Bob Iger to Forgo Disney Salary, Top Execs to Take Pay Cuts Amid Virus Crisis", *Hollywood Reporter*. Retrieved from <https://www.hollywoodreporter.com/news/bob-iger-forgo-disney-salary-top-execs-take-pay-cuts-1287418>

Tien, Shannon. 26 Jun, 2018. "Top Twitter Demographics That Matter to Social Media Marketers", *Hoot Suite*. Retrieved from <https://blog.hootsuite.com/twitter-demographics/>

Appendix

Health and Safety Regulations Pertaining to COVID-19

Given the unpredictable nature of COVID-19, health and government organizations constantly update their own suggestions. At first organizations and political figures were using headcounts as a metric safety in conjunction with physical distance. Soon after, specific activities were advised against and establishments were characterized as non-essential and ordered to close. On March 11, 2020 the CDC accounced that

- Any events of 250 or more people should be cancelled
- Older persons and those with severe pre-existing health conditions are at highest risk

Around the same time, the idea of social distancing was introduced to describe keeping a distance of six feet between people. At this time, it was anticipated that closures would last a two week period until the end of March. Later on March 15, the social gathering head count was reduced from 250 to 50, and extended social distancing to eight weeks [CDC, 2020]. Since then, additional regulations have been issued by state government and local officials. Also on March 15, California Govenor Newsom called for the closure of all bars and nightclubs, and reduced restuarants to take-out and delivery only [Bizjak, 2020]. Numerous states have adopted this idea since then as well as reducing people's time outside to groceries and medical appointments.