

Alistair Nunn  
Underwood  
IS 417  
12/9/2021

## Streaming During COVID-19

### Introduction

Due to Covid-19, people worldwide spend more time in their homes than ever. Quarantine mandates were issued to contain the virus, forcing many people to stay at home, whose impacts can be seen in the decrease in media production over the past year. According to a report by FX, the number of scripted TV shows in 2020 was 496, a 7% decrease from 2019.

The purpose of this project is to understand the impact of these phenomena on the popularity of TV shows released during the pandemic. According to an article published by *Blank*, the number of subscribers for streaming services jumped to 308.6 Million in 2020 - a 32% increase from the year before. From this, it would be easy to assume that the overall popularity would increase, but what does this mean for popularity over time?

It's often been noted that the public eye is always moving, people jump from topic to topic quickly and what's popular today won't necessarily be popular tomorrow. But, I wondered if this would change with people being stuck at home for a large part of the year, and many people still work and attend school from home today. My hypothesis is that because of the decrease in production and quarantine measures, tv shows would remain popular for longer because there were less things to talk about overall.

### Methods

In order to test my hypothesis, I needed to examine viewer engagement with tv shows during the pandemic. For this purpose, I set out to examine a set of TV shows produced from 2019 to 2021. To determine which shows to include in my analysis, I referenced articles about the most popular shows released from 2019 and selected those that were referenced the greatest number of times. The shows I chose to survey were: *Umbrella Academy* (2019) *The Mandalorian* (2019), and *Queen's Gambit* (2020). To build a timeline of viewer engagement, I used *snscraper* to collect tweets about each show starting from the month of its original release date. Other shows of note were *Bridgerton* (2020) and *Squid Game* (2021) of which I collected tweets, but because of gaps in the data I chose not to use them for the analysis. Overall engagement can be seen through the number of tweets each show got per day.

### Analysis: Time-Series

For analysis purposes, only the columns containing the tweet id, date created, tweet content, geographical location, and hashtags were included. From this a time-series plot was created using the content (tweet) counts for each day within the collection timeframe and the initial release dates were marked. The number of tweets collected for the surveyed shows were: *The Mandalorian* (2019): 844,956; *Queen's Gambit* (2020): 357,913; *Umbrella Academy* (2019): 575,126.

It's important to note that for Figure 1 and Figure 2, the time-series spans up to 5 months after the original release date, while the time span for Figure 2 is up to a year. As you can see, each show has a different amount of engagement, shown in the maximum tweets per day in each series. The maximum for *The Mandalorian* is over 40,000 tweets while *Umbrella Academy* and *Queen's Gambit* have 20,000 and 7,000, respectively. This means that of the three surveyed, *The Mandalorian* is by far the most popular and has the greatest amount of engagement.

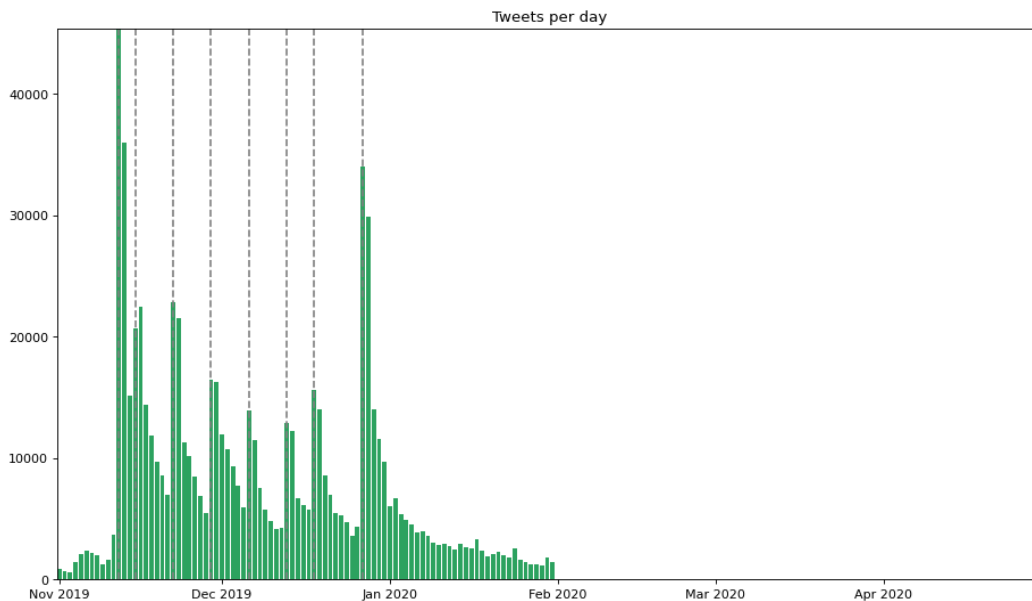


Figure 1: *The Mandalorian* (2019) tweets per day

As seen in Figure 1, the greatest number of tweets were tweeted on the initial release date, with small spikes in engagement leading up to the last episode, of which had a number of tweets more similar to the first day. However, following the last episode, the number of tweets per day dropped significantly and leveled out just after the end of January.

Looking closer at the data, it was found that the average number of tweets per day from February 2020 on was just around 36. Though, given that the second season of the

show was released in late 2020, similar results to those shown leading up to the first season are expected.

Looking at the graph alone, engagement for The Mandalorian seems inline with what would be expected in non-Covid times, with most of the engagement occurring during the show run and dropping during the off-seasons.

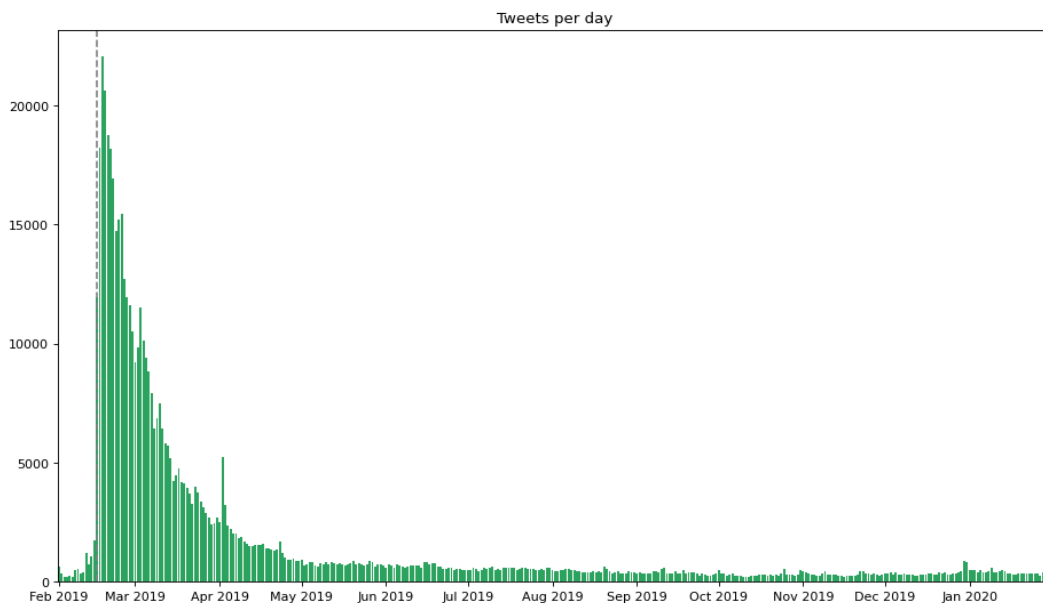


Figure 2: *Umbrella Academy* (2019) tweets per day

Similarly to Figure 1, *Umbrella Academy* had the highest engagement closer to its original release date, but engagement drops exponentially over the next 3 months. This can be explained by the fact that both *Umbrella Academy* and *The Queen's Gambit* are produced by Netflix, which releases all of the episodes at once, while *The Mandalorian* is distributed by Disney and episodes are released on a regular week schedule.

Since this show was released pre-covid it makes sense that it shows less engagement as time goes on, but there is a spike in April of 2019 that would be interesting to examine later. Overall, I think that the results from this show would be a good benchmark for normal engagement, and it would be interesting to compare with engagement for the second season which was released in summer of 2020.

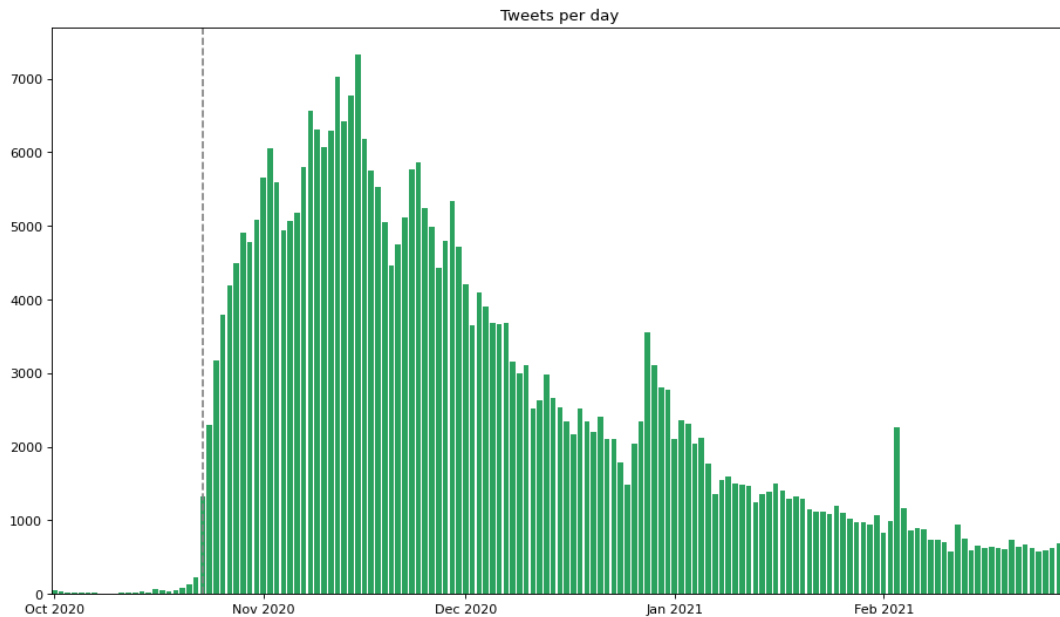


Figure 3: *The Queen's Gambit* (2020)

Unlike the other shows surveyed, *The Queen's Gambit* (2020) which was released in late October of 2020, had the highest engagement a month after its initial release. It shows similar results to *Umbrella Academy* in that there is continuous engagement for the first couple of months after its release, with numbers dipping as time goes on.

Two interesting spikes I found were during January and February of 2021. From a closer look at the data it was found that during February many people tweeted using the hashtag #GoldenGlobes for which the show was nominated and won two awards. The spike from January of the same year was found to be attributed to normal viewership.

### Analysis: Topic Modeling

For the second phase of this project I set out to do topic modeling on the tweets to discover what the top topics of discussion were for each show. For the purpose of this analysis, I attempted to parse and tokenize the tweets using spacy, and equip pyLDAvis for topic modeling. Along with this, I imported a list of stopwords for over 50 languages but because of an issue with my data I didn't solve in time, I was only able to get as far as tokenizing tweets for *The Mandalorian* and was not able to get much useful information from the tokens beyond the most frequent words, which also proved not to be useful since some special characters weren't removed.

### Conclusion

From my results, I was unable to determine whether or not Covid impacted tv show popularity over time. All shows surveyed showed high engagement around the initial release date, and lower engagement in the following months. Future research would

benefit from including subsequent seasons and comparing engagement within and across shows during the height of Covid (spring to summer 2020).

Further analysis in topic modeling would lend more information as to what people are discussing about each show and, and using the geolocation of each tweet could give insight into worldwide popularity.

## Sources

- Agate, B. (2021, April 13). *The Impact COVID-19 Had On The Entertainment Industry In 2020*. Forbes. Retrieved from <https://www.forbes.com/sites/bradadgate/2021/04/13/the-impact-covid-19-had-on-the-entertainment-industry-in-2020/>
- Beck, M. (2021, February). *How to Scrape Tweets With snsrape*. Medium. Retrieved from <https://betterprogramming.pub/how-to-scrape-tweets-with-snsrape-90124ed006afn>
- Diaz, G. (2019). *Stopwords ISO*. Github Repository. Retrieved from <https://github.com/stopwords-iso/stopwords-iso>
- Feldman, Dana. (2020, December 8). *Here Are The 10 Most-Watched Netflix Original And Limited Series Of 2020*. Forbes. Retrieved from <https://www.forbes.com/sites/danafeldman/2020/12/08/here-are-the-10-most-watched-netflix-original-and-limited-series-of-2020/>
- “V. Kindratenko, D. Mu, Y. Zhan, J. Maloney, S. Hashemi, B. Rabe, K. Xu, R. Campbell, J. Peng, and W. Gropp. (2020, July). *HAL: Computer System for Scalable Deep Learning. In Practice and Experience in Advanced Research Computing (PEARC '20)*, July 26–30, 2020. <https://doi.org/10.1145/3311790.3396649>”.
- Richwine, L., Coster, H. (2020, October 13). *Analysis: Fewer movies in theaters? Big Media turns focus to streaming video*. Reuters. Retrieved from <https://www.reuters.com/article/walt-disney-restructuring-streaming/analysis-fewer-movies-in-theaters-big-media-turns-focus-to-streaming-video-idUSKBN26Z09E>
- Zhu, Z. (2021). *Text Analysis in Foreign Languages*. Medium. Retrieved from <https://towardsdatascience.com/text-analysis-in-foreign-languages-b952d7de3f41>