Team AJA Project Report

Our project aims to capture the trends of online entertainment (namely, the streaming services and their usage) since the COVID-19 outbreak and lockdown. When people are stranded at their homes, and most outdoor entertainment activities are forced to shut down to protect public health, people would naturally choose to do more entertainment online, at their homes. Today, we would discover two aspects of this potential social-economic impact of COVID-19 has on the entertainment industry: do people spend more time on streaming services now, and if they do, has there been a particularly more popular genre that people enjoy during these challenging times than others?

The project mainly uses two open datasets: Google Trends and TMDb. While the former provides detailed information on the popularity of a certain topic on Google, the latter is an online movie database containing relevant information of most known movies. There are packages in R allowing us to directly access these data, namely *library(gTrendsR)* and *library(TMDb)*. Note here TMDb requires an API to access their data, which is free for all non-business users-- that is why we consider it as an open database.

Since all streaming services have their database kept private, we have to simply use Google Trends to roughly capture the hype and popularity changes they have gone through since the lockdown. As it turns out, the popularity for Netflix, Disney+ have all gone significantly up since late March when the lockdown began, and HBO suffered some external volatility due to the introduction of HBO MAX in late May.

Now that we have confirmed a significant spike in terms of popularity in online streaming services, we want to move on to explore the details behind this trend, namely: now that since people are enjoying more content online, what genres of entertainment do they prefer? For this research, we have limited the term “entertainment” to only movies (sorry about the TV shows, we just prefer movies).

We have used the *movie_popular()* function to sort out the 100 most popular movies trending worldwide as of June 13th, 2020 (unfortunately, TMDb does not provide historical trends on popular movies). As such, we have obtained 100 movies and their unique IDs, which is very convenient for us to process large amounts of movie information. We have extracted the column containing the IDs as a new vector to locate 100 most popular movies.
Using the *movie()* function, we can retrieve the genre information of a movie. Using a *for loop* here, we have successfully collected the genres for each and every individual 100 popular movies, and have plotted them onto a bar chart. Unsurprisingly, most of these movies are big genres such as action, adventure, sci-fi, drama, and family movies.

Another concern has risen: how do we know if these movies have become popular due to the pandemic, or they have been popular all along? To explore the relationship, we have plotted the popularity of every big genre from the popular movies list in Google Trends.
The results have demonstrated a significant increase in favour of genres such as action, family, drama, and comedy movies, while thriller, sci-fi, and adventure movies do not seem to have increased much in popularity. This is a proof that people tend to watch more family-friendly, easy-watching movies during hard times like these.