Does the Covid 19 pandemic affect YouTube channels positively or negatively?

STA 260 Final project report

Furqan Arshad, Yixing Lu, Zhixin Tang

1. Introduction

1.1. Background

A novel coronavirus (CoV) named '2019-nCoV' or '2019 novel coronavirus' or 'COVID-19' by the World Health Organization (WHO) is in charge of the current outbreak of pneumonia that began at beginning of December 2019 near Wuhan City, Hubei Province, China (Artificial Intelligence for Coronavirus Outbreak 2020). Since its outbreak in 2019, the pandemic has not been fully resolved after more than 2 years. The Covid-19 pandemic does not only have impacts on global health, but also has largely influenced people's lifestyles. One hard truth of the pandemic was that, in order to someday be together safely again, we had to be apart in the meantime, and for many, this meant that social media has become one of the limtied ways to be with friends and family (Molla, 2022). Moreover, with the restriction of recreation and leisure activities, more and more people chose to relax by watching videos during quarantine. Prior to the pandemic, YouTube was a powerful online platform for retailers to drive meaningful engagement among consumers – a community of over 2 billion monthly users (Sarah Travis and Elliott NIx, 2020). YouTube saw the most significant growth of any social media app among American users during the pandemic, the Pew Research Center reported Wednesday (Rodriguez, 2022). YouTube's popularity is especially high among 18- to 29-year-olds, with 95% of them saying they use the service (Rodriguez, 2022). In order to learn more about people's entertainment after the pandemic, in this project, we mainly analyze the increase or decrease of subscribers and video views on different YouTube Channels.

1.2. Goal and objectives

Given YouTube's overall increasing popularity during the pandemic, we are interested in investigating whether the pandemic has had different effect on the performance of YouTube Channels from different categories: which types of YouTube channels have become better off during the pandemic and which kinds are hit hard by the pandemic. Then, we may be able to provide some insights on how people's taste towards YouTube videos has been changed by the pandemic, which is valuable to people who would like to engage themselves in self-media. To achieve our goals, we chose 11 most popular categories of YouTube Channels and each category had around 10 famous YouTube Channels. Then, we used the increases or decreases in subscribers and video views of Youtube Channels from 2019 to 2022 to represent the positive or negative effects of the pandemic. The hypothesis is that under the lockdown and stay-at-home policies which were global for a certain period of time during the pandemic, since people's physical mobility was restricted, they might turn to online social media for entertainment more than before the pandemic. Therefore, comparing the time series plot of time spent at home with the time series plots of YouTube Channels' subscribers and video views to figure out the association between YouTube channels and time spent at home by people is especially important for us to accomplish the goal.

2. Methodology

1.1. Data source and scraping

In order to evaluate the performance of YouTube channels since the Covid 19 pandemic, historical data on YouTube channels' subscribers and video views were needed. Socialblade.com collects such data from the YouTube API and saves all the data within 3 years on their website. Figure 1 showed an example web page from which such data was scraped. The website provided information on weekly gained subscribers and views of a YouTube channel shown in line plots as in Figure 1. The websites were scraped in May 2022, thus the oldest data available was from May 2019. A python script containing the code used to scrape the websites is available in the Appendix.

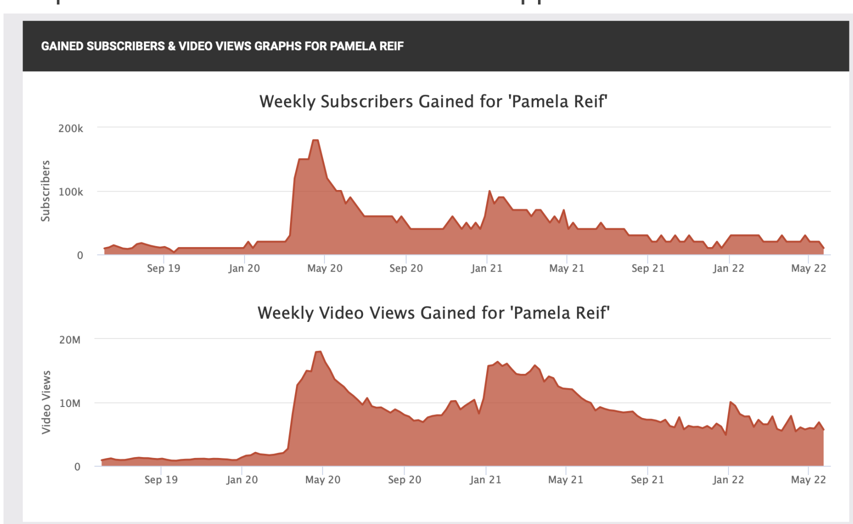


Fig. 1: Screenshot of a socialblade.com website from which data was scraped for the current project.

121 YouTube channels from 11 different categories were selected in the current project. The 11 categories covered a broad realm of available YouTube video types, including gaming, fashion & beauty, food, fitness, comedy, travel, news,

kids, product reviews, commentary/ reaction/ narrative videos, and education. 10~12 popular channels from each category were selected, each of which has at least 1M subscribers. These channels were searched on socialblade.com and the URL's of the “detailed statistics" pages were recorded. Since the SocialBlade is protected by Cloudflare, it was impossible to directly scrape from their websites. To bypass this issue, a wayback machine (web.archive.org) was used to save a historical version of each SocialBlade web page using the recorded URL's, and data was scraped from these newly generated URL's.

Another data set used in the current project was the time-spent-at-home since the Covid 19 pandemic. The data was part of the global mobility trends data provided by Google in its Covid 19 Comunity Mobility Reports and collected from https://ourworldindata.org/covid-google-mobility-trends. The specific data set we chose recorded the % change of time people spend in the residential areas after February 17th, 2020, compared to that of the baseline days (a five-week period from January 3rd to February 6th, 2020.) The data presented was already smoothed to a 7-day moving average. This data set was used to explore any potential association between the time people spent at home and the performance of YouTube channels from different categories.

1.2. Data cleaning

Figure 2 showed the data cleaning scheme used in this project on the YouTube channel weekly gained subscribers and views data scraped the website. There were five steps: date alignment, abnormal data fix, interpolation of weekly gained subscribers data, smoothing through moving averages, and finally min-max rescaling.

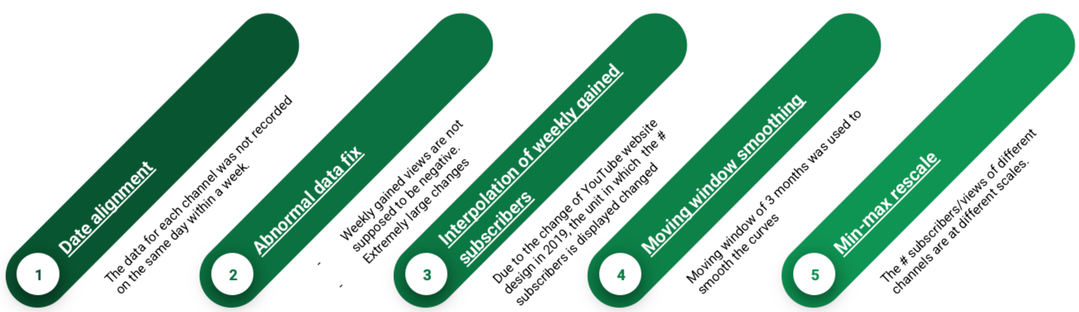


Fig. 2: Data cleaning scheme.

1.2.1. Date alignment

First, we removed channels that did not have complete data over the period from May 2019 to May 2022. One channel each from the comedy, news, and travel categories and three from the commentary/ reaction/

narrative category was removed. The data we scraped for each channel had one record per week for both the subscribers and views. However, the dates within a week on which the data was collected could be different for each channel. Therefore, if simply outer joining data from all the channels together, the data frame would contain many days with no entries for most of the channels. To clean the data, we chose to keep only a single date for each week, and if a certain channel did not have an entry for this date, the data was collected within 6 days after the designated date was assigned there. Finally, rows with NAs were omitted. After date aligning, all the dates in the data frame were 7 days apart and there were no missing values left.

1.2.2. Abnormal data fix

Two kinds of abnormality in the YouTube data set were observed: (1) negative weekly gained views; (2) extremely large change in weekly gained subscribers and views. Weekly gained subscribers can be negative since people might choose to un-subscribe. However, weekly gained views should be non-negative because a video cannot be unseen once it’s watched. Thus, a possible reason for negative weekly gained views was that some videos were deleted from the channel at that time point and all the views from those videos were subtracted from the total views. Batch uploading or deleting of videos could be the reason for an extremely large change in weekly gained views as well, but could not explain that in the number of subscribers. Since such abnormality distorted the time series trends and was not the focus of the current project, the abnormal data points were replaced with the mean value of their four neighboring points.

1.2.3. Interpolation of weekly gained subscribers data

Another observation on the weekly gained subscribers data was that after September 2019, the data became very “jumpy” and oscillated between 0 and 10,000 for many channels. The reason behind this was YouTube changed how the number of subscribers was displayed on their website. Previously, it showed the exact number to the unit digit, but after September 2019, it changed to the digit of 10,000 (e.g., 1,234,567 to 1.23 M). Therefore, the weekly gained subscribers would be either zero, or integer multiples of 10,000. Due to this issue, the current data set did not reflect the reality of how many subscribers were gained each week. To fix this, for the sequence of weekly gained subscribers of each channel, the first non-zero values were evenly distributed to the weeks with zeros after the last non-zero values.

1.2.4. Moving average smoothing and min-max rescaling

The final step of data cleaning was smoothing the curves with moving averages of a 12-week window and rescaling all the data to the range of 0-to-1 by min-max normalization. The min-max rescaling does not make the normality assumption on data, thus was selected to bring all the data to the same scale and make them comparable. After all the data cleaning steps, the final data frame contained 115 channels with weekly data collected from 2019-07-07 to 2022-03-20.

1.3. Data analysis

1.3.1. Cross correlation and clustering of time series

The cross-correlation function (CCF) is used to evaluate the similarity between the spectra of two different systems, for example, a sample spectrum and a reference spectrum (Jack L. Koenig, 1999). This technique can be used for samples where background fluctuations exceed the spectral differences caused by changes in composition (Jack L. Koenig, 1999). In this project, cross-correlation was used to estimate the correlations between different YouTube Channels. The function of cross-correlation is defined as

$$r(d) = \frac{\sum_{i=1}^{n}[(x_i-\mu_x)*(y_{(i-d)}-\mu_y)]}{\sqrt{\sum_{i=1}^{n}(x_i-\mu_x)^2}\sqrt{\sum_{i=1}^{n}(y_{(i-d)}-\mu_u)^2}}, \text{ d = 0, 1, …, N-1; i = 0, 1, …, N-1}$$

The number of i-d is the number of lags. According to Google, the lags refer to how far the series is offset, and its sign determines which series is shifted. In the following results of cross-correlation and heatmap, to get clear correlation values between different YouTube Channels, cross-correlation is calculated when lag equals 0.

To categorize similar YouTube Channels by weekly subscribers and weekly video views, we used a method of hierarchical clustering. Hierarchical clustering, also known as hierarchical cluster analysis, is an algorithm that groups similar objects into groups called clusters (Tim Bock, 2022). We grouped the total 115 YouTube Channels into 4 clusters by using the method “ward”. Then, we would visualize all YouTube Channels from the 4 clusters.

1.3.2. Rolling window time-lagged cross correlation of univariate time series

To investigate the association between time-spent-at-home and the weekly gained subscribers/ views for channels from different categories, we first calculated the category median curves for both the weekly gained subscribers and views time series. Then the world average time-spent-at-home time series was obtained. Data within the common time period covered by both data sets were chosen for this analysis. The cross-correlation between each of these median curves and the time-spent-at-home time series was calculated to compare the association between time-spent-at-home and channel performance for different categories.

We also hypothesized that (1) in different time frames within the three-year periods we looked at, the relationship between time-spent-at-home and YouTube channel performance might not be consistent; (2) there might be a time-lagged effect. To test these two hypotheses, a time-lagged cross-correlation with a rolling window of 24 weeks was calculated between the category median of weekly gained subscribers/ views and the time-spent-at-home time series. The step size of the rolling window was chosen to be 2 weeks. For a given time window t:t+24, denote the time-spent-at-home data series as X[t:t+24], and the median curve for a certain category within the same time window as Y[t:t+24]. Cross-correlation was calculated between X[t:t+24] and Y[(t+l):(t+24+l)], where l is the lagged time, with values from -12 to 12. That is, moving the category median curve 12 weeks prior to or ahead of the current time window, then calculating the cross-correlation with the time-spent-at-home time series for each of the lagged time values. If the peak cross-correlation occurred at a positive lagged time, then the time-spent-at-home time series might lead the YouTube channel performance time series, vice versa. And if the peak correlation occurred at lagged time=0, then there might be no lagged effect. Moreover, with the rolling window, we were able to see how such correlation evolved over time.

2. Results and discussion

2.1. Data cleaning visualization

Figure A1 in the Appendix provided visualization of the weekly gained subscribers and views data after each data cleaning step: (a,b) after data alignment and abnormal data fix; (c) after interpolation of weekly gained subscribers data; (d,e) after smoothing; (f,g) after min-max normalization.

2.2. Weekly gained subscribers/ views of YouTube channels in different categories

As Figure A1 was too messy for us to detect any common patterns between YouTube channels, we plotted the time series of YouTube channels by category. Figure A2 showed the visualizations of weekly gained subscribers and weekly gained video views of YouTube channels by category. In the categories of Comedy, food, and kids, we could clearly find that the trends of weekly subscribers and weekly video views reached a peak around April 2020, which coincides with the first peak of time-spend-at-home in Google mobility data. From the category of fashion & beauty, there was no clear pattern in the plot of weekly gained subscribers and the trend in the plot of weekly gained views showed an overall decreasing pattern. The trends of weekly subscribers and weekly video views in the category of fitness perfectly matched the 3 peaks of time-spend-at-home in the Google mobility data. In the plot of weekly gained subscribers and weekly gained video views of the category of product reviews, the trends showed a decreasing pattern after January 2021. In the rest of the 5 categories, which were commentary, education, gaming, news, and travel, we could not find any common patterns in the plots.

Figure 3 showed the median curves of weekly gained subscribers (a) and views (b) for each channel category. According to Figure 3, we could find that some categories showed a peak in weekly gained subscribers and weekly gained video views from late April to early May in 2020, which consisted of the first peak of time-spend-at-home in the Google mobility data. The fitness and Food categories had the most substantial increase compared to other categories. In contrast, comedy, fashion & beauty, product review & unboxing, and travel categories showed an overall decreasing trend of weekly gained subscribers and weekly gained video views since the pandemic. Therefore, the pandemic might have positive influences on fitness and food and might have negative impacts on comedy, fashion & beauty, product review & unboxing, and travel.

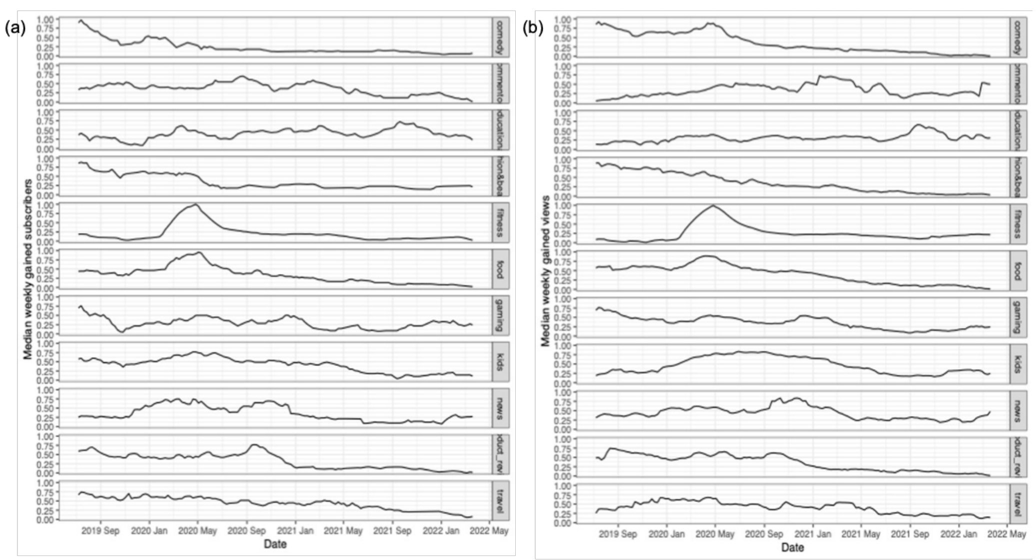


Fig. 3: Median curves of weekly gained subscribers (a) and views (b) for each channel category.

2.3. Cross correlation and clustering of different YouTube channels' weekly gained subscribers/ views time series

2.3.1 Cross correlation and clustering of different YouTube channels' weekly gained subscribers time series

To investigate the correlation of weekly gained subscribers between YouTube channels, we calculated the CCF value between 115 different YouTube channels. As weekly gained subscribers of some YouTube channels did not change over time, which means some YouTube channels do not gain new subscribers from 2019 to 2022, we could not get the CCF of these channels. Therefore, we deleted these channels, which were jennamarbles, pewdiepie.x, ryanhiga, rclbeauty, byrontalbott, gamenewsoffical, pediepie.y and tldtoday. Then, the remained number of YouTube channels was 107. We got a 107x107 correlation matrix. As the correlation matrix was really big, we could not post it in this report. According to the correlation matrix, we could find that there are a lot of YouTube channels' cross-correlations are higher than 0.9. Figure 4 showed a heatmap plotted by the correlation matrix. The square with a lighter color in the heatmap represented the higher cross-correlation. From the heatmap, we could also find that there were a lot of YouTube channels that were highly correlated.

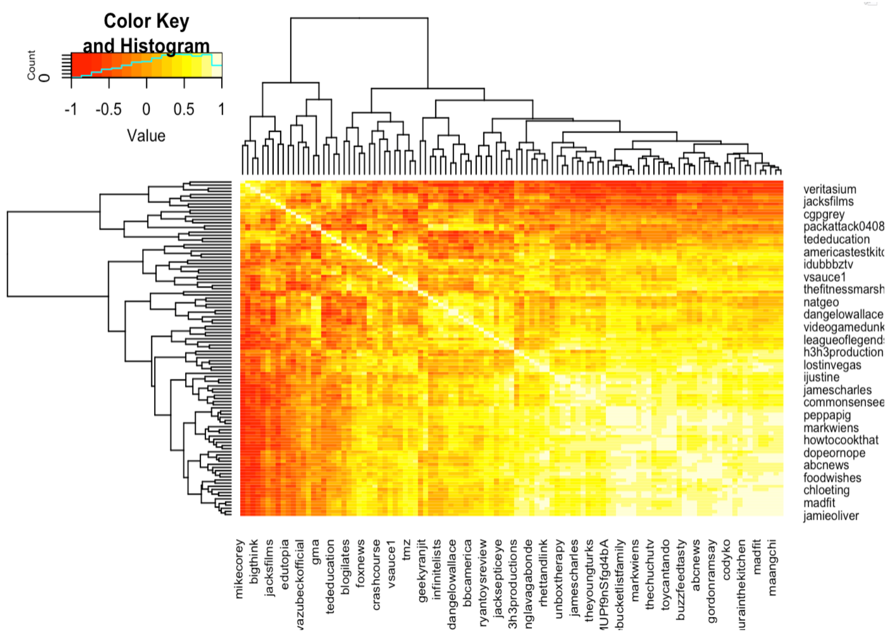


Fig. 4 Heatmap of different YouTube channels' weekly gained subscribers

Figure 5 showed the Dendrogram of hierarchical clustering on different YouTube channels' weekly gained subscribers. We grouped the total of 115 different YouTube Channels into 4 clusters. The first cluster contained 25 YouTube channels, the second cluster contained 46 YouTube channels, the third cluster contained 25 YouTube channels, and the last cluster contained 19 YouTube channels. Then, we visualized all the different YouTube channels in each cluster, and we smoothed the time-series graph of each YouTube channel. According to Figure A3, in cluster 1, it was easy to find that the trends of the total 25 YouTube channels had an increasing pattern from around January 2020 to around April 2020 and then had a decreasing pattern from around April 2020 to around January 2021. The YouTube channels in cluster 1 met the first peak of time-spend-at-home Google mobility data. In cluster 2 and cluster 3, it was hard for us to find a common pattern in the time series plots of YouTube channels. In cluster 4, we could find that all trends of the 19 YouTube channels had an overall decreasing pattern over time. Therefore, the pandemic might have positive effects on the YouTube channels in cluster 1 and might have negative effects on the YouTube channels in cluster 4.

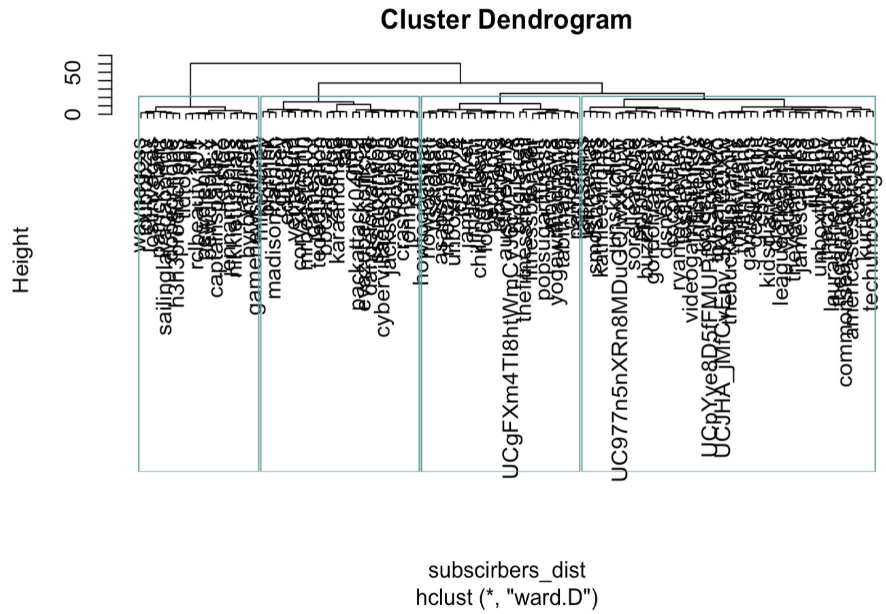


Fig. 5 Hierarchical clustering of different YouTube channels by weekly gained subscribers. The Dendrogram is divided by 4 groups

2.3.2 Cross correlation and clustering of different YouTube channels' weekly gained video views time series

As we did in 2.3.1, we also calculated the cross-correlation values of different YouTube channels' weekly gained video views. Unlike the results we had in 2.3.1, we could calculate the cross-correlation of all 115 YouTube channels. According to the cross-correlation matrix, we could find that the number of highly correlated YouTube channels based on weekly gained video views was higher than the number of highly correlated YouTube channels based on weekly gained subscribers. Figure 6 also showed that there were a lot of highly correlated YouTube channels by weekly gained subscribers. Moreover, comparing Figure 7 and Figure 5, we could also find the number of light color squares in Figure 7 was more than the number of light color squares in Figure 5. Therefore, it seemed that YouTube channels' weekly gained video views were more highly correlated to each other.

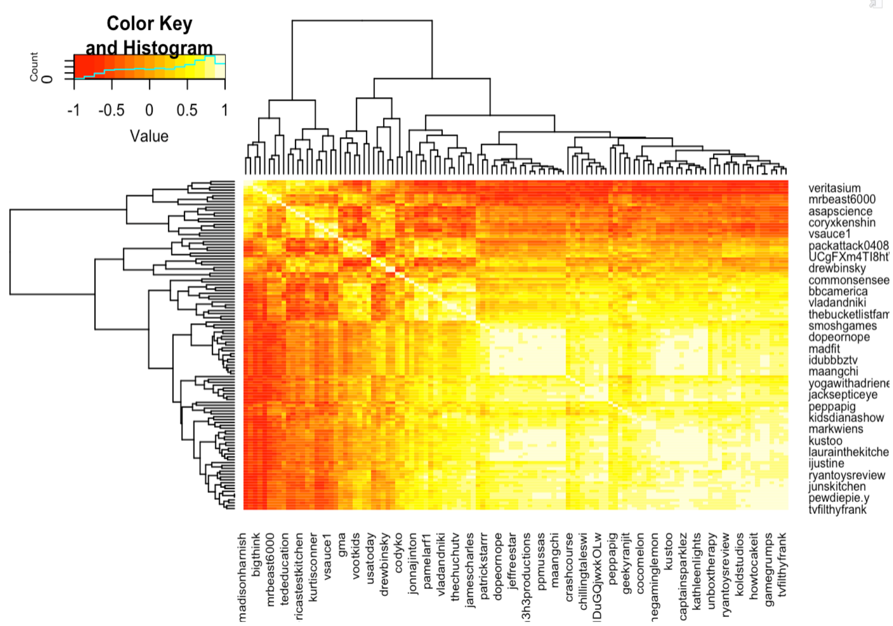


Fig. 6 Heatmap of different YouTube channels' weekly gained subscribers

Figure 7 showed the hierarchical clustering of different YouTube channels by weekly gained video views. We divided the 115 different YouTube channels into 4 groups. The first cluster contained 38 YouTube channels, the second cluster contained 21 YouTube channels, the third cluster contained 26 YouTube channels, and the last cluster contained 30 YouTube channels. To observe the differences between different clusters, we still visualized all different YouTube channels' time series in each cluster. According to Figure A4, in cluster 1, most of the trends of weekly gained video views from YouTube channels reached a peak around late April to early May 2020. Some of the trends in these plots reached a peak around January 2021. In cluster 2, most of the trends in weekly gained video views plots from YouTube channels showed an overall increasing pattern, and some of the trends in these plots reached a peak around May 2021. In cluster 3, most of the trends in weekly gained video views from YouTube channels showed an overall decreasing pattern, and some of the trends in these plots reached a peak around April 2020. In cluster 4, most of the trends in these plots showed an overall decreasing pattern, and some of the trends in these plots reached a peak around April 2020. It was hard for us to find a common pattern in each cluster. Thus, it was hard for us to make any hypothesis about how the pandemic would affect different YouTube channels by hierarchical clustering of weekly gained subscribers.

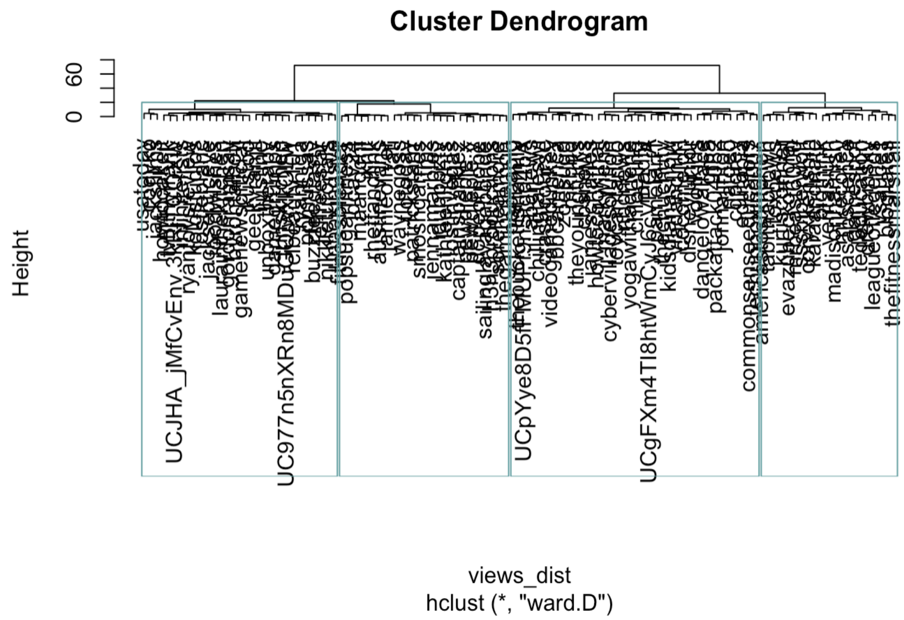


Fig. 7 Hierarchical clustering of different YouTube channels by weekly gained subscribers. The Dendrogram is divided by 4 groups

2.4. Rolling window time-lagged cross correlation between time-spent-at-home time series and by-category median weekly gained subscribers/ views time series

Figure 8 showed the world average % change in time spent at home from February 2020 to March 2022. From the figure, we could see that there were four major local maxima, in May 2020, January 2021, May 2021, and January 2022, where May 2020 had the highest increase in time-spent-at-home since that was the period where the strictest stay-at-home policies were implemented across the world. The later peaks also corresponded to substantial increases in daily Covid 19 cases worldwide.

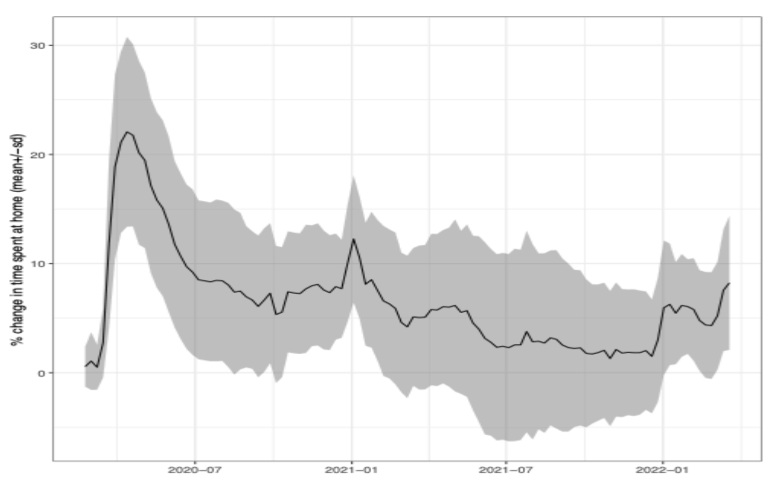


Fig. 8: World average of % change in time-spent-at-home compared to baseline days. The shaded band indicated the standard deviation.

First, we looked at the association between YouTube channels' performance and the time-spent-at-home with the full-range time series data, results shown in Figure 9. For subscribers, the fitness, food, and kids channel categories had the highest cross-correlation with the time-spent-at-home time series, while commentary/narrative and educational categories had the lowest correlation. The educational category was the only one showing a negative correlation with time-spent-at-home. This result implied that when people were confined at home, the two activities they likely would spend more time in were cooking and working out. Also, when kids could not go to school, they would have more time to spend on YouTube channels with kids-focused content. The negative correlation between educational channel subscribers and time-spent-at-home was not expected. We thought with limited in-person schooling, students might rely more heavily on online educational resources. The trends in weekly gained views were similar to those for subscribers, except that the gaming category surpassed kids and food categories and became the second highest in correlation. That is, the views of gaming videos had a stronger correlation with time-spent-at-home than the subscribers of the game streaming channels. This might imply that viewers of the gaming channels had a lower tendency to subscribe to the channels. Also, for most of the categories, views had a higher correlation with time-spent-at-home than subscribers.

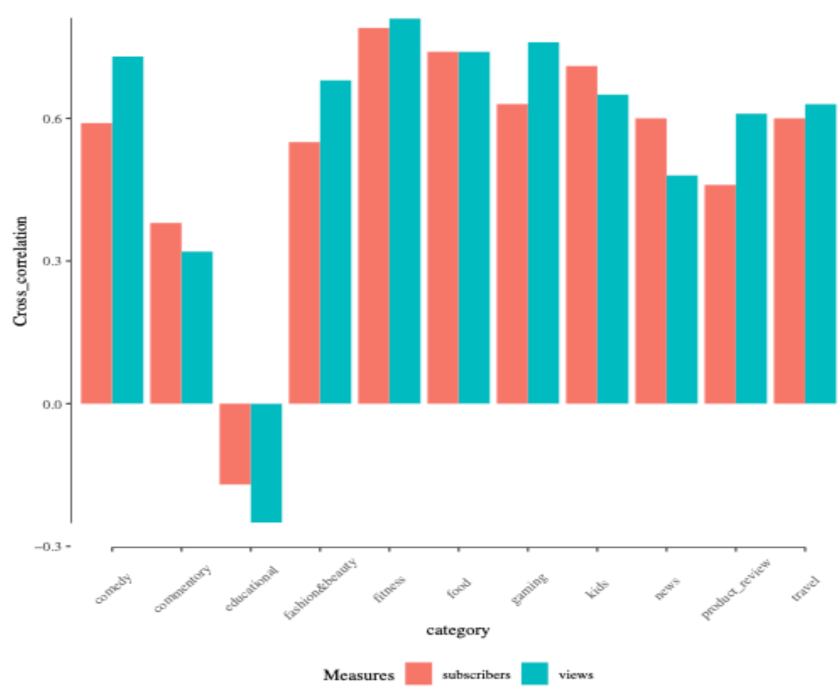


Fig. 9: Cross correlation between weekly gained subscribers (red)/ views (blue) and time-spent-at-home for different channel categories.

After this initial analysis, we then took into consideration the lagged effect and time-heterogeneity effect by calculating the rolling window time-lagged cross-correlation. Figure 10 and 11 were heat maps showing the cross-correlation between the time-spent-at-home time series and category media subscribers and views respectively. The rows of each heat map were the starting date of all the rolling windows, and the columns were the lagged time from -12 weeks to +12 weeks. Red, white, and purple colors dictated positive, zero, and negative correlation respectively. From the figures, we could tell that for most of the categories, there was no evidence for a clear pattern of lagged effect: for most of the time windows, the largest correlation (either positive or negative) occurred near-zero lagged time. Also, it was confirmed that the correlation changed over time. With the same lagged time, the cross-correlation could be highly positive within a one-time window but negative within another.

2.5. Comparison of YouTube channels' subscribers and video views before and after the Pandemic

From all the categories of YouTubers, we looked at the number of subscribers and viewers of each channel to analyze the outcome of covid on the subscribers and viewers. Figure 12 showed the comparison of subscribers of 11 categories. Kids category had the highest number of subscribers among all the categories where the number of subscribers got increased after the covid and has the highest number of subscribers increase. The remaining categories had also shown increases except for the comedy category as shown in figure 11. After covid people also started subscribing more to channels such as gaming, news, travel, educational channels, etc.

After careful analysis of the number of YouTube subscribers before and after covid, we looked at the trend of the number of views gained by each channel before and during covid. The highest number of views were gained by the Kids category during the covid, which meant that during the covid there were no schools and kids were mostly home spending time on the YouTube. The kids category had the highest number of views both before and during covid. Furthermore, increases in the number of views were seen in the Fitness category, although there was not much increase in the number of subscribers of the Fitness category/ Fitness category had the highest percentage increase. The gaming category had also increased during the covid in 2020, Comedy category had not shown any increase in the number of views as shown in Figure 13.

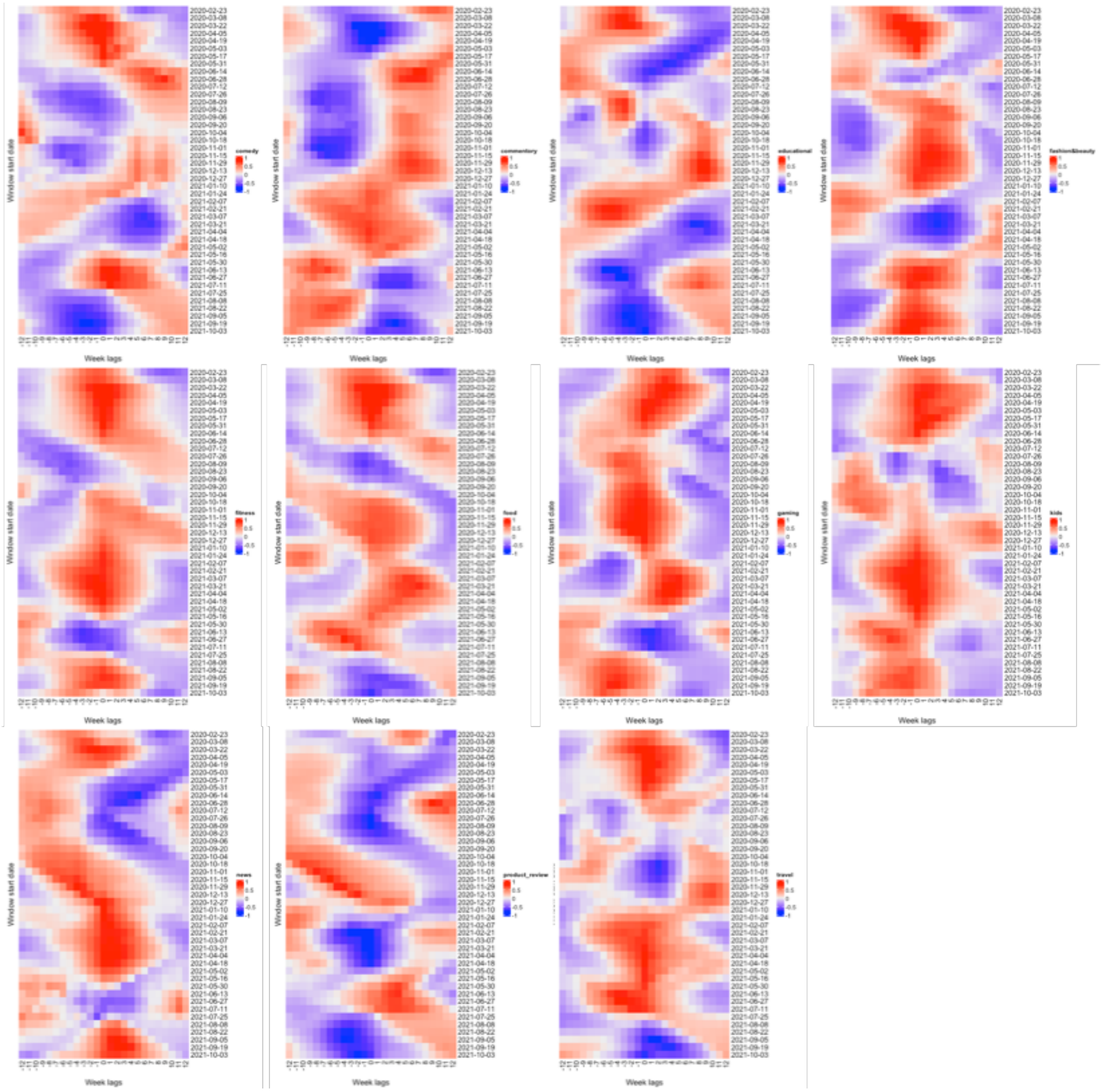


Fig. 10: Heat map of rolling window time-lagged cross correlation between time-spent-at-home and category median of weekly gained subscribers for different categories. The rows of the heat maps are the start date of each time window, and the columns are the lagged time, from -12 weeks to +12 weeks. The heat maps from left to right, top to bottom were for comedy, commentary, educational, fashion & beauty, fitness, food, gaming, kids, news, product reviews, and travel.

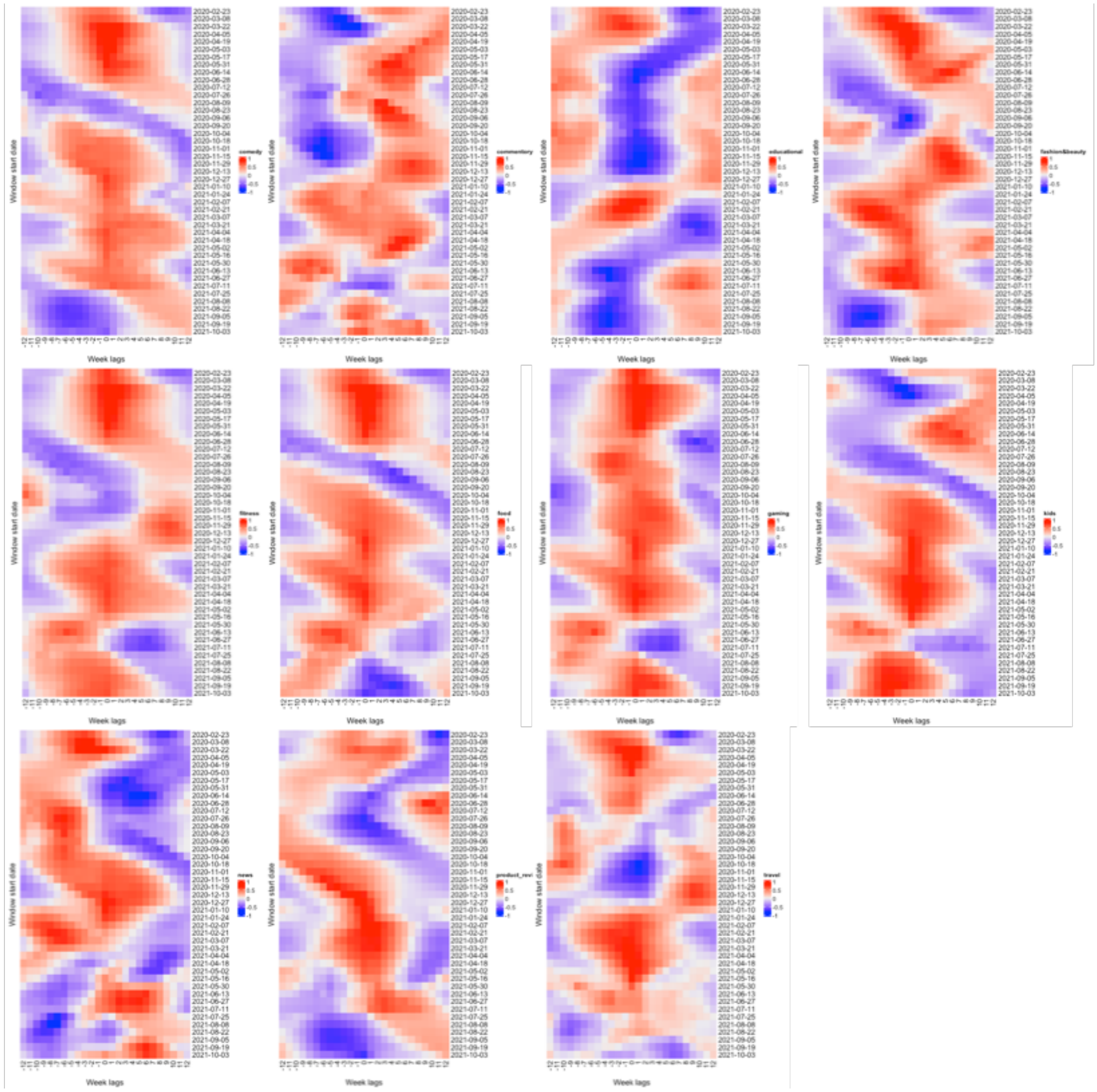


Fig. 12: Heat map of rolling window time-lagged cross correlation between time-spent-at-home and category median of weekly gained views for different categories. The rows of the heat maps are the start date of each time window, and the columns are the lagged time, from -12 weeks to +12 weeks. The heat maps from left to right, top to bottom were for comedy, commentary, educational, fashion & beauty, fitness, food, gaming, kids, news, product reviews, and travel.

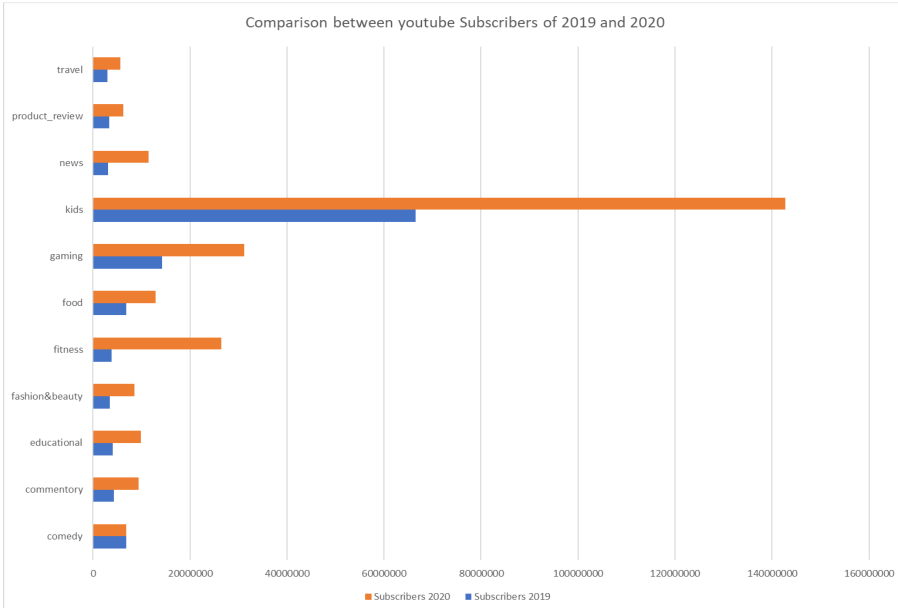


Fig. 12: Comparison of YouTube subscribers before and during covid.

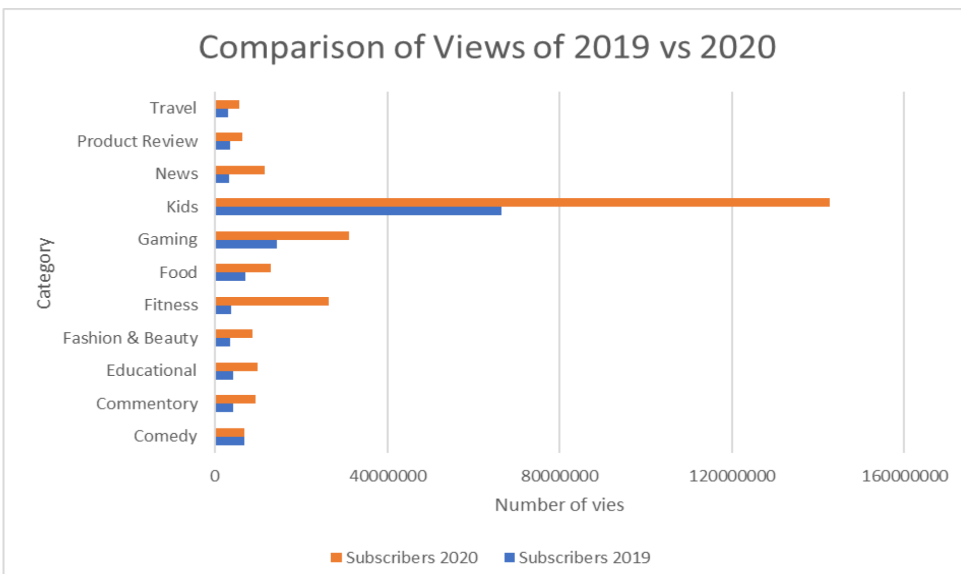


Fig. 13: Comparison of youtube views before and during covid.

We looked at different top YouTubers of the kids category to figure out why there was an increase in the number of subscribers after covid. Figure 14 showed the bar graph of different YouTubers in the kids category with subscribers during 2019 and 2020. The graph showed that different YouTubers have gained a lot of popularity during the pandemic and the number of subscribers had boosted. Only one channel out of the 13 channels had not shown much increase in the number of subscribers while other channels showed increases in the number of subscribers which showed that the increase in the number of subscribers was due to the popularity of these top channels getting promoted during the pandemic.

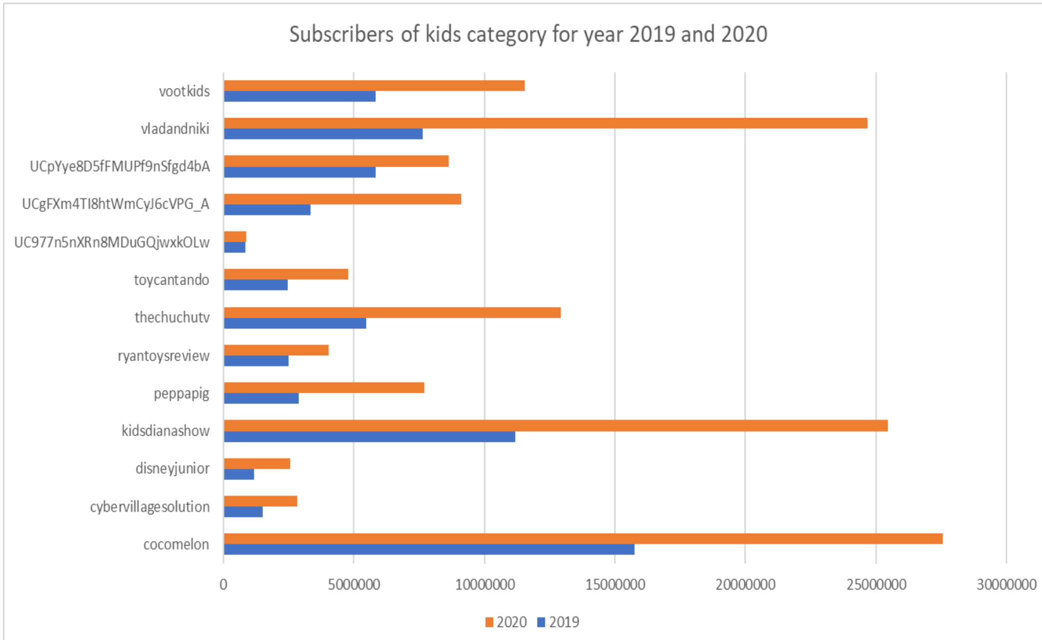


Fig. 14: Comparison of Subscribers of Kid Category before and during covid 2020

We analyzed the increase in the number of subscribers of the Gaming category before and during covid to find out the reasoning for the abrupt increase. We looked at the top 10 YouTube channels among the gaming categories which broadcast gaming videos on a weekly and daily basis. The top YouTuber with the highest number of subscribers in 2019 did not gain many subscribers in 2020 but another YouTuber “mrbeast6000” who had fewer subscribers than “pwediepie” gained more popularity during covid and became the number 1 YouTuber among the ranking of the number of subscribers gained in 2020 as shown in Figure 15. Furthermore, the increase was also seen in young newcomers who did not have many subscribers in 2019 and gained some in 2020. The only YouTuber who lost subscribers during covid captainparklez as he might not be uploading content regularly or would have been playing games that were not much popular.

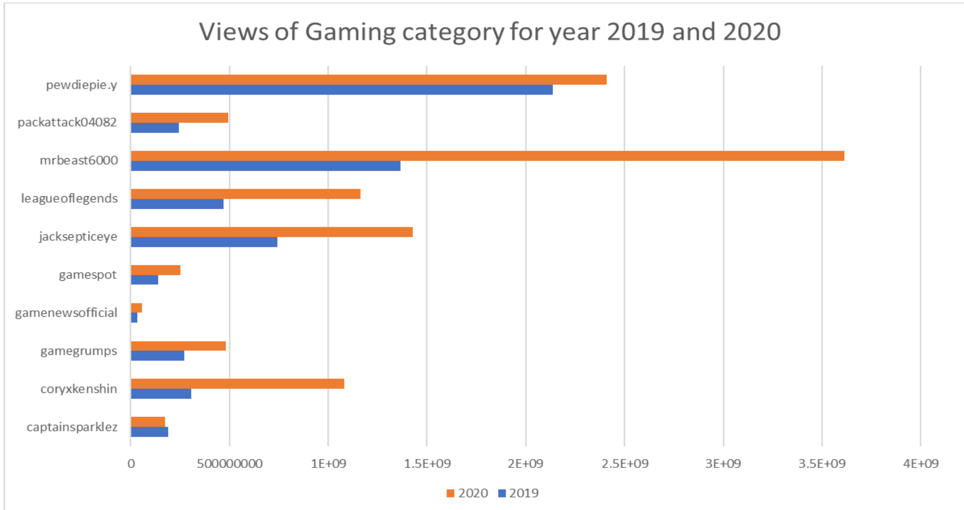


Fig. 15: Comparison of Subscribers of Gaming Category before and during covid 2020

3. Conclusion

In this project, the impact of the Covid 19 pandemic on the performance of YouTube channels in different categories was investigated. We collected data on the weekly gained subscribers and video views of more than 100 popular YouTube channels from 11 categories, from 2019 May to 2022 May. After data cleaning and transformation, we were able to observe some trends in the change of subscribers and views overtime. Channels in the food and fitness categories had substantially rapid increase in the number of subscribers and views after May 2020, when a global scale lockdown took place. On the contrary, channels in the fashion & beauty, comedy, product reviews, and travel categories showed slower growth or even decay after the pandemic. The cross correlation between each pair of time series was calculated for all the channels. Then the channels were clustered using the correlation matrix as a distance measure. The results showed that many of the channels were highly correlated in terms of both weekly gained subscribers and views. The channels within the same clusters showed similar trends over time. Cross correlation between the channel data and the time-spent-at-home time series revealed that the performance of channels from the following four categories were most highly correlated with how much time people spent at home: food, fitness, comedy, and gaming. We also observed that the kids category gained the most number of subscribers after the pandemic. Besides its large base to start with, it is believed that when kids were confined at home, there is a huge need of kids-friendly entertainment and the promotion of these channels during the lockdown might also contribute to their prosperity.

Appendix

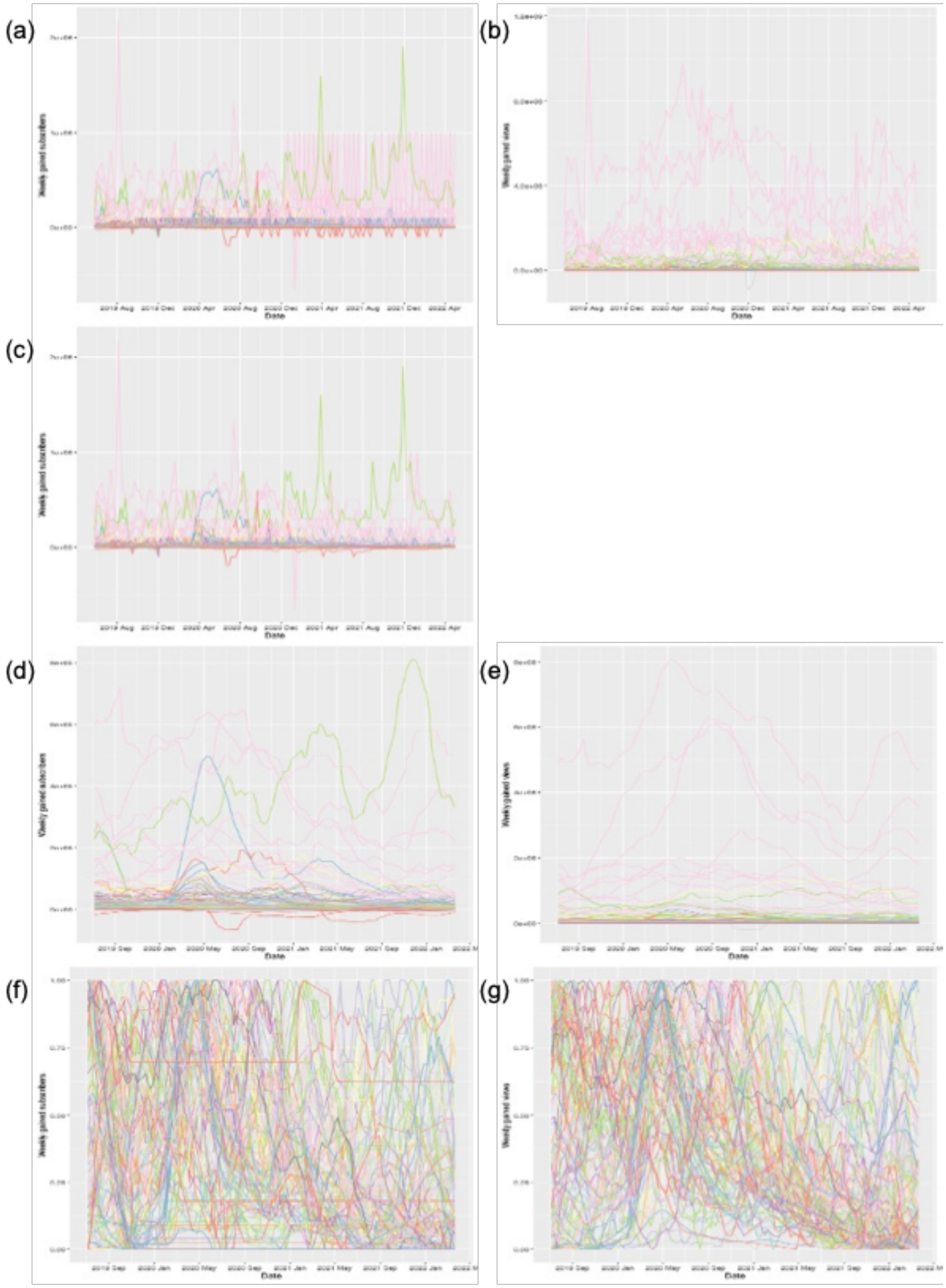
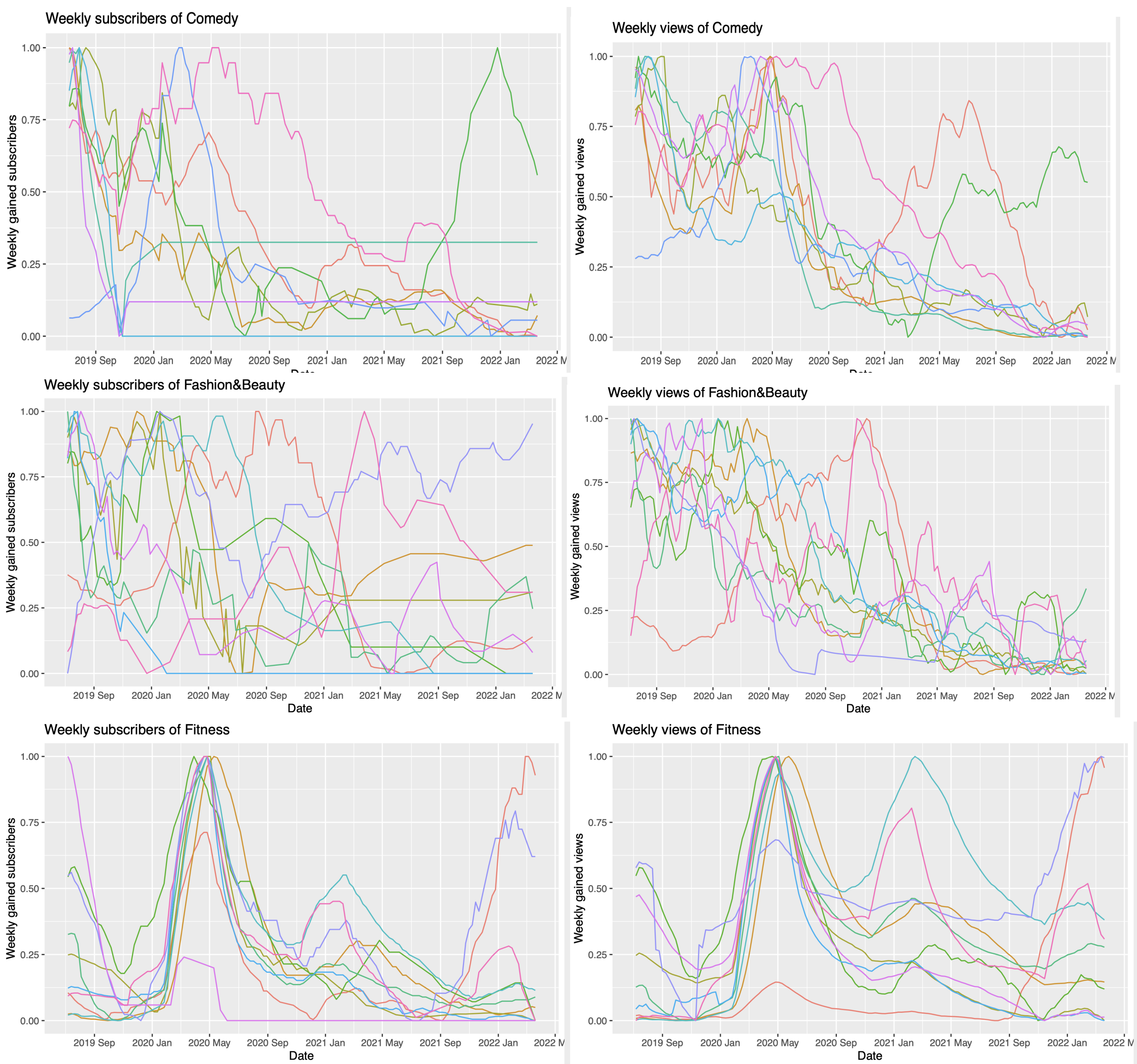
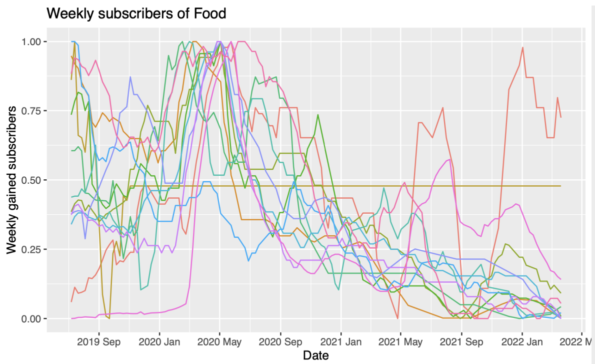
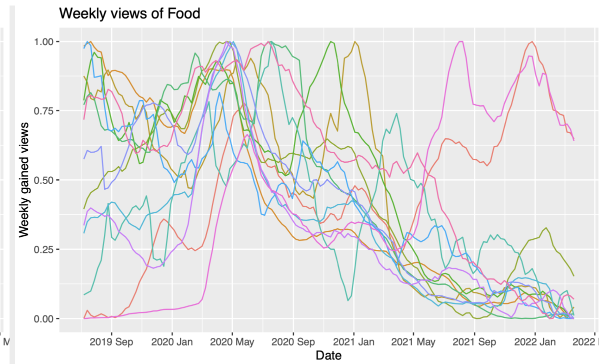
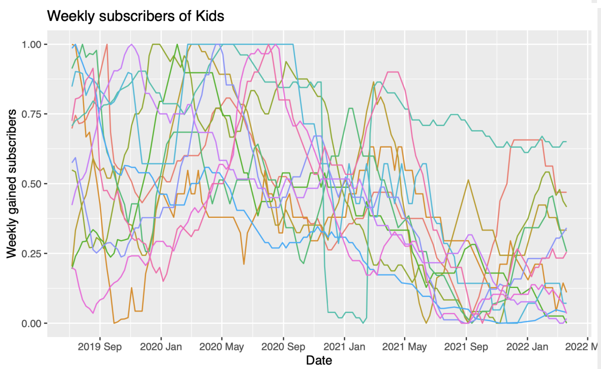


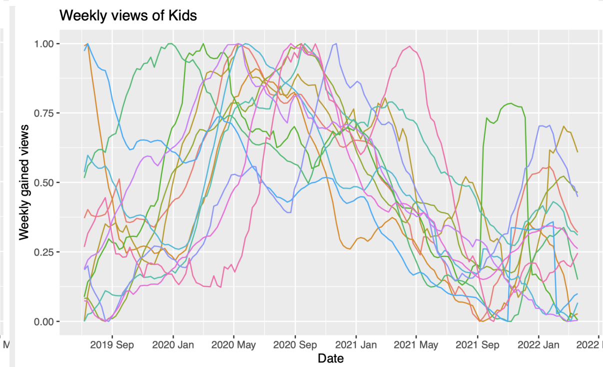
Fig. A1: Visualization of YouTube channel weekly gained subscribers (left column) and views (right columns) after each step of data cleaning. (a,b) after data alignment and abnormal data fix; (c) after interpolation of weekly gained subscribers data; (d,e) after smoothing; (f,g) after min-max normalization.

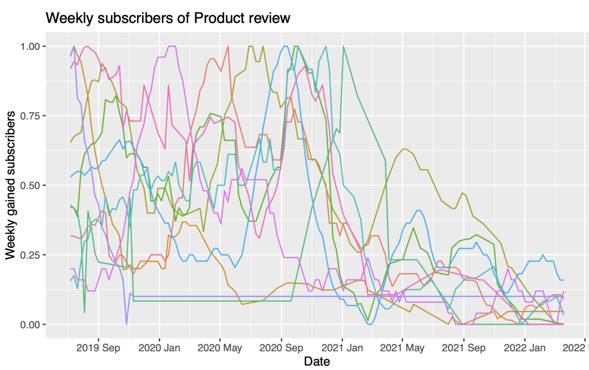


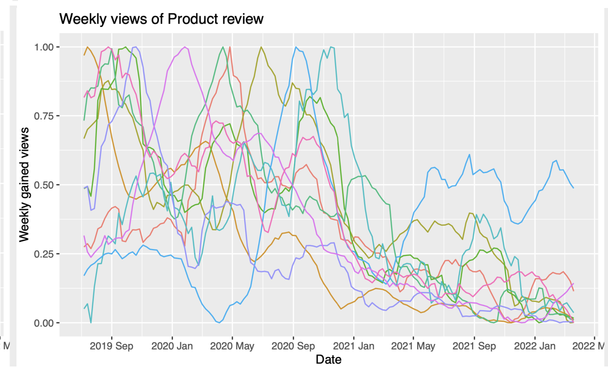

Weekly subscribers of Food
Weekly gained subscribers
1.00
0.75
0.50
0.25
0.00
2019 Sep
2020 Jan
2020 May
2020 Sep
2021 Jan
2021 May
2021 Sep
2022 Jan
Date

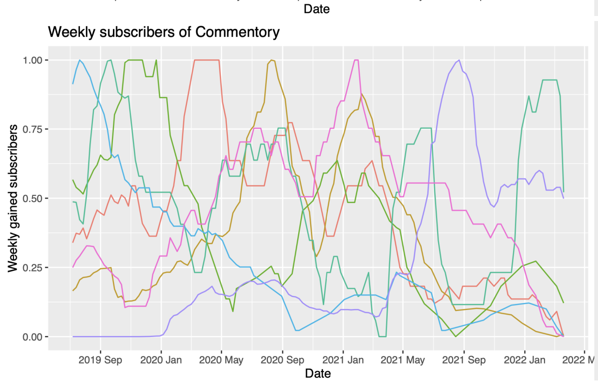

Weekly views of Food
Weekly gained views
1.00
0.75
0.50
0.25
0.00
2019 Sep
2020 Jan
2020 May
2020 Sep
2021 Jan
2021 May
2021 Sep
2022 Jan
Date

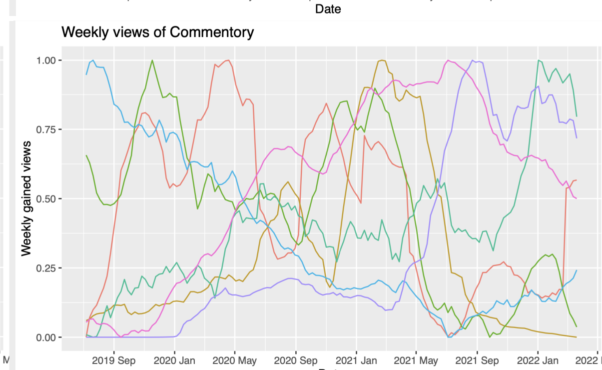

Weekly subscribers of Kids
Weekly gained subscribers
1.00
0.75
0.50
0.25
0.00
2019 Sep
2020 Jan
2020 May
2020 Sep
2021 Jan
2021 May
2021 Sep
2022 Jan
Date


Weekly views of Kids
Weekly gained views
1.00
0.75
0.50
0.25
0.00
2019 Sep
2020 Jan
2020 May
2020 Sep
2021 Jan
2021 May
2021 Sep
2022 Jan
Date


Weekly subscribers of Product review
Weekly gained subscribers
1.00
0.75
0.50
0.25
0.00
2019 Sep
2020 Jan
2020 May
2020 Sep
2021 Jan
2021 May
2021 Sep
2022 Jan
Date


Weekly views of Product review
Weekly gained views
1.00
0.75
0.50
0.25
0.00
2019 Sep
2020 Jan
2020 May
2020 Sep
2021 Jan
2021 May
2021 Sep
2022 Jan
Date


Weekly subscribers of Commentory
Weekly gained subscribers
1.00
0.75
0.50
0.25
0.00
2019 Sep
2020 Jan
2020 May
2020 Sep
2021 Jan
2021 May
2021 Sep
2022 Jan
Date


Weekly views of Commentory
Weekly gained views
1.00
0.75
0.50
0.25
0.00
2019 Sep
2020 Jan
2020 May
2020 Sep
2021 Jan
2021 May
2021 Sep
2022 Jan
Date

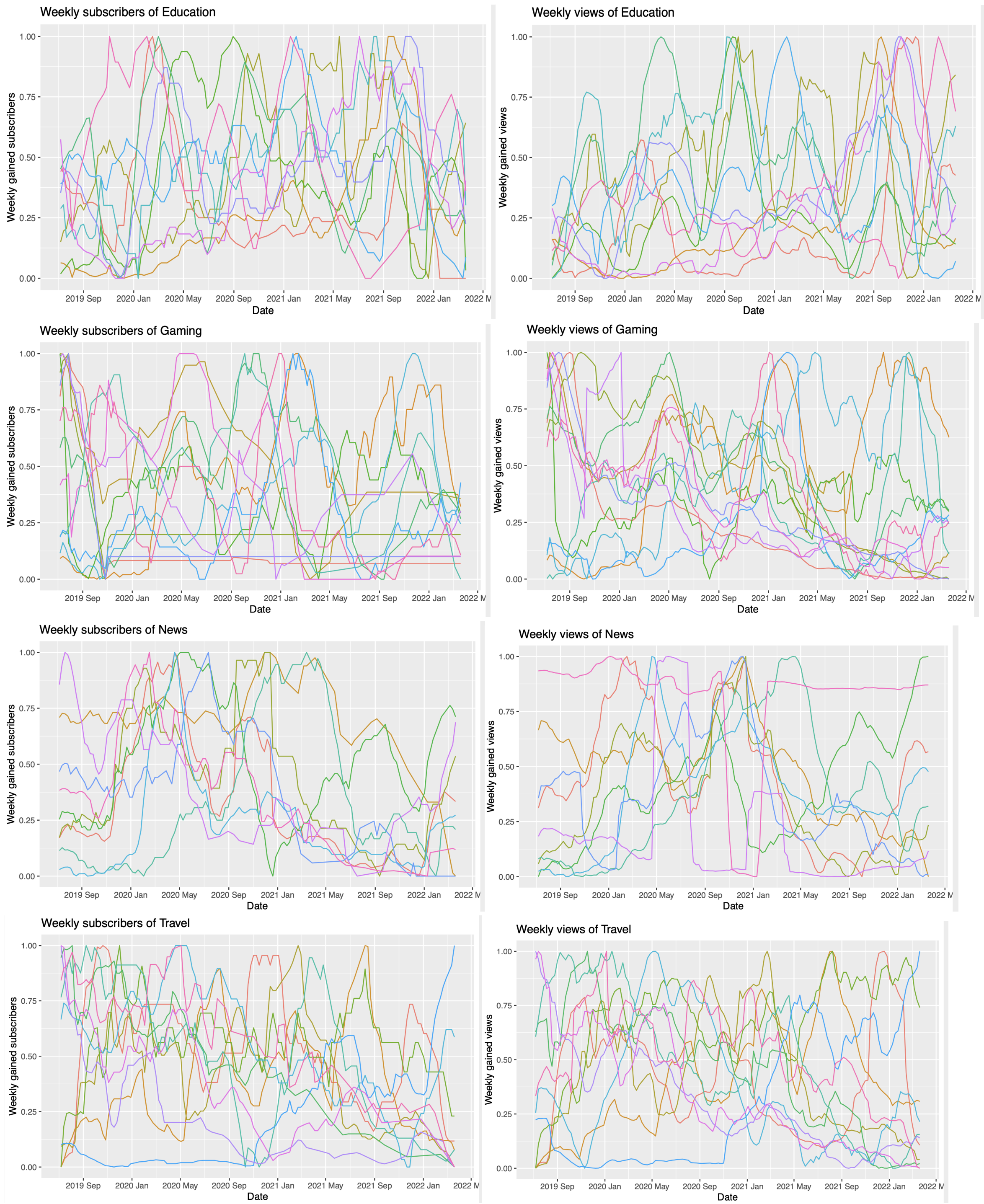


Fig. A2. Visualization of weekly subscribers and weekly video views by categories. The plots in left side is weekly gained subscribers, and the plots in right side is weekly gained video views. Each different row represents a different category of YouTube Channels.

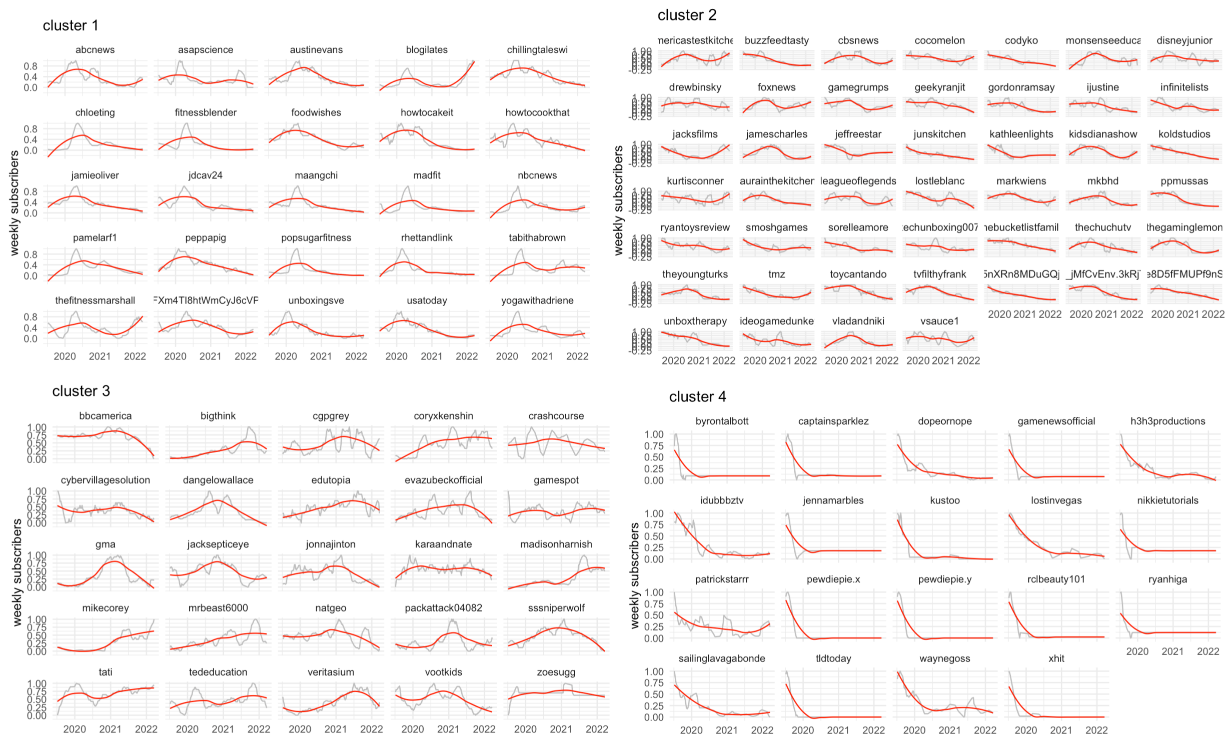


Fig. A3 Visualizations of all YouTube Channels in different cluster by using hierarchical clustering in weekly gained subscribers

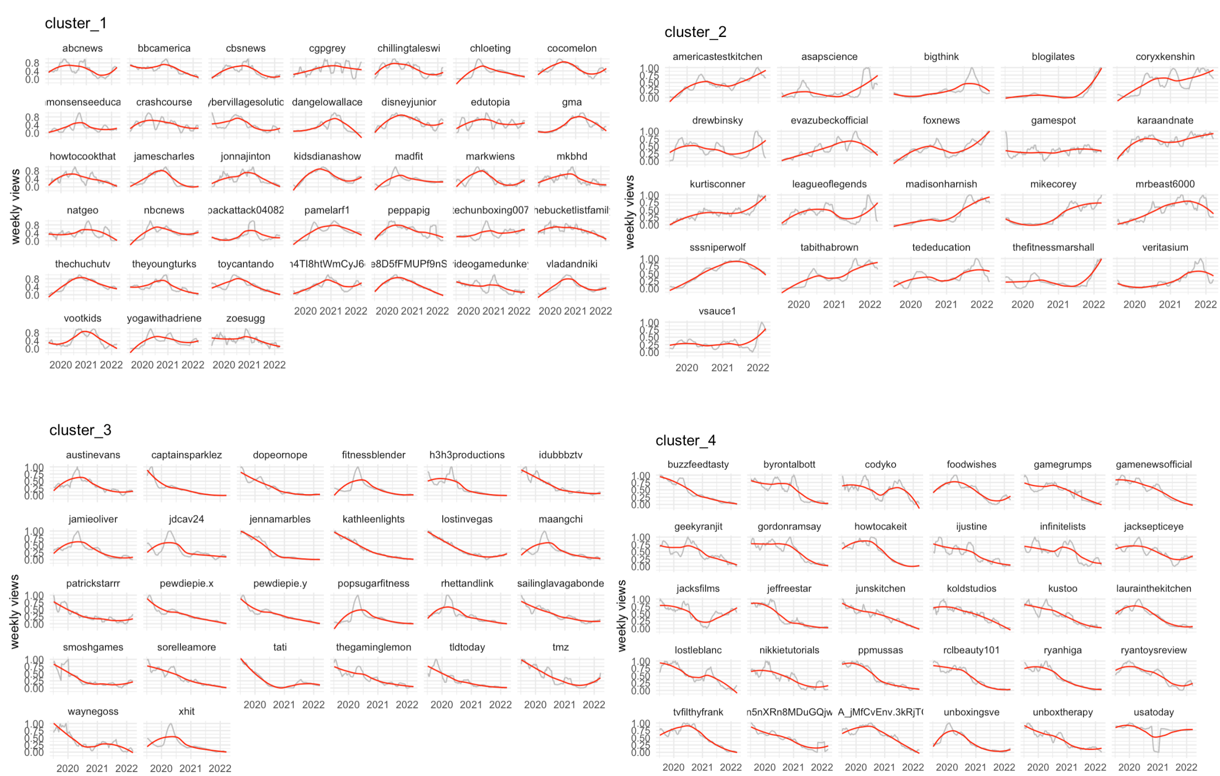


Fig. A4 Visualizations of all YouTube Channels in different cluster by using hierarchical clustering in weekly gained video views

Reference

Jamshidi, M. B., Lalbakhsh, A., Talla, J., Peroutka, Z., Hadjilooei, F., Lalbakhsh, P., Jamshidi, M., Spada, L. L., Mirmozafari, M., Dehghani, M., Sabet, A., Roshani, S., Roshani, S., Bayat-Makou, N., Mohamadzade, B., Malek, Z., Jamshidi, A., Kiani, S., Hashemi-Dezaki, H., & Mohyuddin, W. (2020, June 12). *Artificial Intelligence and covid-19: Deep learning approaches for diagnosis and treatment*. IEEE access : practical innovations, open solutions. Retrieved June 5, 2022, from https://www.ncbi.nlm.nih.gov/pmc/articles/PMC8043506/

Molla, R. (n.d.). *Posting less, posting more, and tired of it all: How the pandemic has changed social media*. Ethics at Work. Retrieved June 5, 2022, from https://ethicsatwork.nd.edu/resources/posting-less-posting-more-and-tired-of-it-all-how-the-pandemic-has-changed-social-media/

Rodriguez, S. (2021, April 7). YouTube is social media's big winner during the pandemic. CNBC. Retrieved June 5, 2022, from https://www.cnbc.com/2021/04/07/youtube-is-social-medias-big-winner-during-the-pandemic.html

Sarah, T and Elliott, N. The power of YouTube in the era of COVID-19. Retrieved Nov. 16, 2020, from https://www.retaildive.com/spons/the-power-of-youtube-in-the-era-of-covid-19/588999/

Jack L, K. (n.d.). *Cross correlation*. Cross Correlation - an overview | ScienceDirect Topics. Retrieved June 5, 2022, from https://www.sciencedirect.com/topics/earth-and-planetary-sciences/cross-correlation

Bock, T. (2022, January 17). *What is hierarchical clustering?* Displayr. Retrieved June 5, 2022, from https://www.displayr.com/what-is-hierarchical-clustering/