The Global Decline of Music Videos on Youtube: An Analysis Using Negative Binomial Regression (2017-2024)*

Fares Alkorani

April 19, 2024

Music videos have always been a popular medium for music consumption up until recently. We look at the sum of the most viewed music videos weekly globally from Youtube Charts (n = 344) from the week of September 1 to September 7 in 2017 until now. Using a negative binomial regression model, we find that in each week, the expected number of music video views decreases by 0.7%. As Tiktok and audio-streaming platforms like Spotify and Apple continue to evolve, this information can serve useful to record labels and individuals for how they want to continue to use music videos.

1 Introduction

Music videos have been around since the early 20th century, primarily in the form of short films or promotional clips featuring musical performances. The popularity of music videos soared in the 1980s with the launch of MTV (Music Television), a cable television network dedicated to broadcasting music videos around the clock. This era marked a significant cultural shift in the music industry and popular culture. With the launch of Youtube in 2005, music videos became accessible to global audiences and were one of, if not the most popular medium for listening to music. Recently, I have generally noticed that newer music video releases do not have as many views as older music videos. In addition, there has been some discussion on forums such as Reddit on if music video views were dying, but no concrete studies have been executing to suggest as such.

As a result, in this paper, the estimand I will attempt to estimate is the impact that time has on the views of a music video. We will do this by using Youtube Charts (Youtube 2024) which

---

*Code and data are available at: https://github.com/faralk/music-video-view-analysis.

is a service hosted by Youtube that allow people to view metrics and rankings of songs on the platform. We do this by looking at the weekly top 100 music video views from the week of 2017-09-01 to 2017-09-07 to 2024-03-29 to 2024-04-04. For each week, we take a sum which leaves us with 344 weeks. From there, we use a negative binomial regression model.

We find that with time, music video views have been on a steady decline from September 2017 up until now. This is important as it reveals insights on how the music industry is evolving and can serve useful to artists or record labels on how much they want to emphasis music videos when releasing singles.

The remainder of this paper is structured as follows. Section 2 discusses the data retrieved from Youtube charts, Section 3 discusses the model I used to analyze the data, Section 4 discusses the results of the model and Section 5 provides discussion of how this impacts the music industry and further steps that can be taken in this project.

2 Data

2.1 Source

We use Youtube Charts, a service provided by Youtube which contains metrics about most listened to songs, most watched music videos, and other charts. From Youtube Charts, we use their Global Weekly Most Viewed Music Videos.

There is a similar dataset provided by "kworb" (kworb 2024) that measures the most globally viewed Youtube videos on a weekly basis as well. While it is not explicitly stated how they gather the views for their chart, it appears that their version of the chart takes the views from Youtube's view count which is publicly displayed. This view count includes paid advertisement views, where a music video may receive views if someone saw it briefly in an advertisement. However, Youtube Charts does not include paid advertisement views (Google 2024b). As a result, we choose to to proceed with Youtube Charts since we want to measure intentional actions by users and whether they are actively consuming music themselves through music videos.

2.2 Methodology

We use RSelenium (Harrison 2022) and rvest (Wickham 2024) to automatically download the 324 datasets from the Youtube Charts.

Afterward, we group by date and take the sum of all the top 100 views for that week to be left with 344 observations - corresponding to the week and total views for the top 100 music videos in the chart for that week. I do this because I want to see a general trend for music videos, rather than analyzing the music videos on a rank basis (i.e. seeing how the views decline for the first ranking videos of each week).

2.3 Measurement

2.3.1 Music Video Definition

Apart from not including paid advertisement views, we still beg the question: "What is a music video?". Youtube does not define what a music video is. In the eyes of record labels, a music video is a video that is crafted specifically. However, Youtube seems to be literal in its interpretation of the phrase "music video". To elaborate, in some datasets, Youtube allows live performance clips and lyric videos in their weekly chart ranking and may include even more videos that do not fit the standard definition of a "music video". Moreover, since they do not define their criteria of what a music video is, it is unclear if they have modified their definition overtime and how this may have affected entries in the newer weekly charts.

2.3.2 View Count

Secondly, Youtube does not explicitly define what a view is. For paid advertisement views, Youtube lists some criterions on what gets marked as a view (Google 2024a), but there is no information for how a normal attempt at engaging with a video counts for a view or not - it is unclear if the user is required to watch the full video or just a segment of it.

2.3.3 Countries of Consideration

Over the years, there have been Youtube Charts made available for viewing by country. However, since we are viewing the global chart, it is made unclear what countries are included and excluded from this chart, especially since Youtube is available in over 100 countries, yet there are only Youtube Charts available for 61 countries (Google 2024b).

2.4 Features

The features given in each CSV file follow as such; Rank, previous rank, video title, artist names, periods on chart, views, growth (percentage) and YouTube URL.

However, because each file was downloaded independently of each other, Youtube does not include a date field in the table for the week of consideration. As a result, we manually create a date field such that we can reference it.

In this graph, we see a negative correlation between week of consideration and music video views. As a result, the feature we are interested in from the csv file is the date. However, since the date by itself cannot be used as a predictor, we create a variable titled "weeks_since" which is an integer value that has a value for the weeks since the week of 2017-09-01 to 2017-09-07.
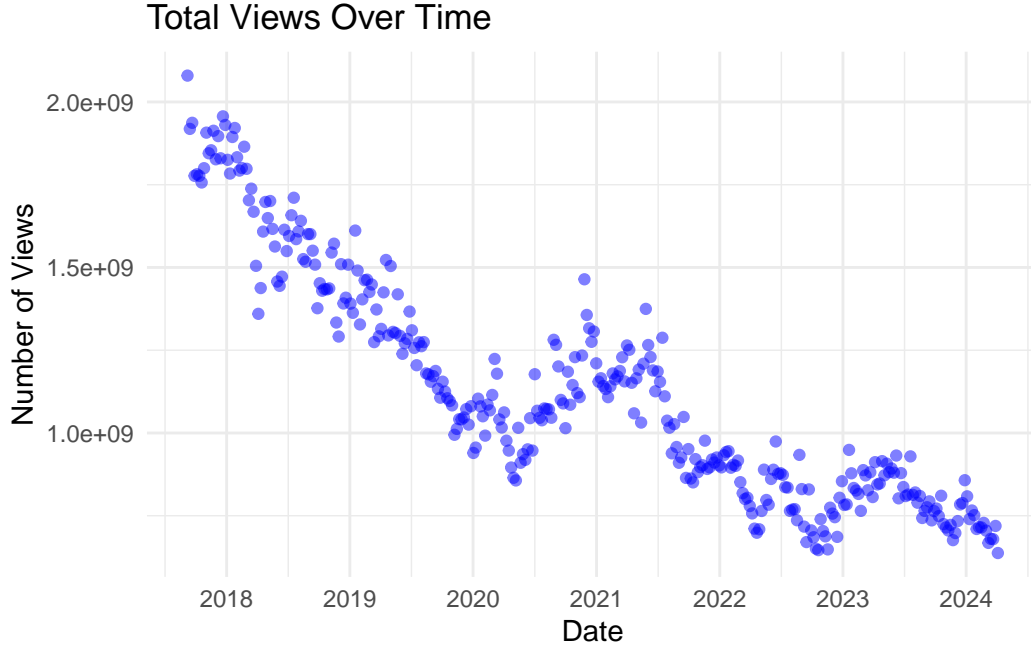
Figure 1: Relationship between week and sum of top 100 most globally viewed music video in that week

3 Model

Here we briefly describe the Bayesian analysis model used to investigate the impact that a given week has on the total number of music video views. Background details and diagnostics are included in Appendix B.

3.1 Model set-up

Define y_i as the total number of views in a given week, $Weeks since Sept 7, 2017 is the preidctor variable which is an integer that represents the number of weeks since the week of September 1, 2017 to September 7, 2017. Then α is a parameter representing the intercept with a prior distribution that is normally distributed with a mean of 21 and a variance of 2.5, meanwhile for every one week, the expected value number of views for that week will change by β - which is a parameter with a prior distribution that is normally distributed with a mean of 0 and a variance of 0.025.

$$y_i|\mu_i, r \sim \text{NegBinom}(\mu_i, r) \tag{1}$$
$$\log(\mu_i) = \alpha + \beta \times \text{Weeks since Sept 7, 2017}_i \tag{2}$$
$$\alpha \sim \text{Normal}(21, 2.5) \tag{3}$$
$$\beta \sim \text{Normal}(0, 0.025) \tag{4}$$

We run the model in R (R Core Team 2023) using the `rstanarm` package of Goodrich et al. (2022).

3.2 Model justification

Since we expect a negative relationship between total music video views and a given week Figure 1, we decide to use the given week as a sole predictor for music video views.

Since we are dealing with count data (number of music video views), we decide between a negative binomial regression model or a poisson regression model. Since we do not have the same mean and variance, it is less likely that the poisson regression model will work. Based on findings in Appendix B, we decide to use negative binomial regression as it fits the data better and based on the fact that the ELPD is larger, too.

4 Results

Term	Estimate	Std. Error
(Intercept)	21.266	0.014
weeks_since	-0.003	0.000

Our model observes a intercept of 21.280, which is to say that on the week of September 1, 2017 to September 7, 2017, the total sum of the views should be 21.820 - which we can disregard in this case, since music videos on that week clearly had more views - and this is just to fit the model.

On the other hand, for each additional week since the start of the observation period, the log of the expected count decreases by approximately 0.003 units. This is to say that the expected video count by week decreases by a factor of $\exp(-0.003) = 0.997$ which amounts to a roughly 0.7% decrease in music video views each week.

5 Discussion

5.1 Tiktok and Spotify's Impact on Music Videos and Artistry

In this day of age, Tiktok has become a huge platform for discovering and engaging with music. In addition, streaming services like Spotify and Apple Music have taken off, and now watching music videos is not the most convenient option for consumers. However, now that Spotify and Tiktok are becoming more popular options with respect to music consumption: the budget is going towards "visualizers" - looped videos used on Tiktok or Spotify that contain some artistic elements in them (Cragg 2024). It begs the question on how record labels and independent musicians will continue to invest into their creativity - or if they will channel it in other artistic forms rather than traditional music videos.

5.2 Future Research

There is much future research that needs to be done. To begin, we analyzed the total views of the top 100 videos. However, we did not explore other factors such as the language of the music video, the Youtube Charts by country (instead of globally). In addition, the rise of Tiktok for music consumption can be explored in relation to duration of music videos - maybe a music video is too long for this current generation for it to be worth their time. In addition, it is worth exploring genres of music videos that appear high on the ranking each week - especially since K-Pop and Latin music seem to be on the forefront of highly viewed videos (Cragg 2024).

Appendix

A Additional data details

B Model details

B.1 Posterior predictive check

In Figure 2a and Figure 2b, we implement a posterior predictive check for negative binomial regression and poisson regression models respectively. This shows that the negative binomial model fits the data better.

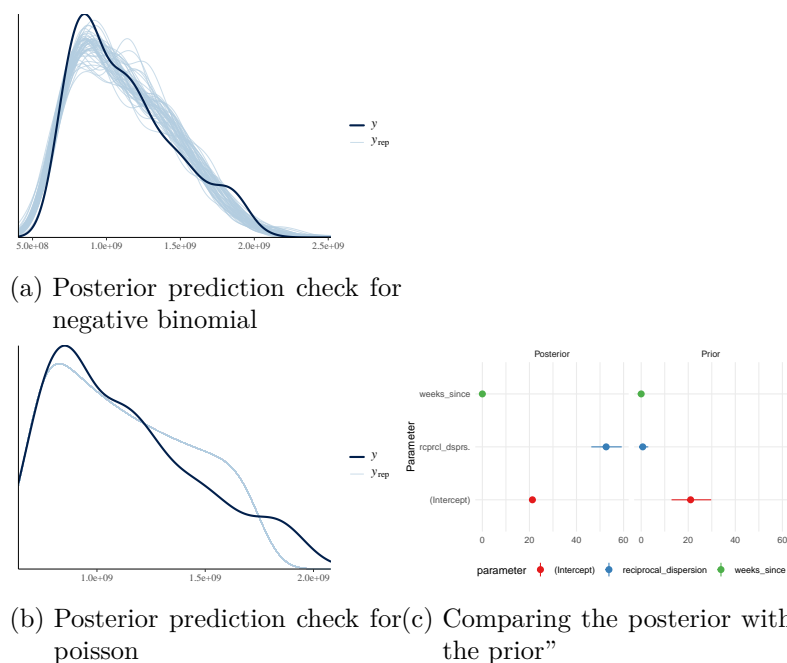In Figure 2c we compare the posterior with the prior.

(a) Posterior prediction check for negative binomial

(b) Posterior prediction check for poisson

(c) Comparing the posterior with the prior"

Figure 2: Examining how the model fits, and is affected by, the data

```
                                         elpd_diff     se_diff
youtube_views_nbinom_improved_rstanarm         0.0         0.0
youtube_views_poisson_rstanarm        -2475633180.1   175718907.8
```

B.2 Diagnostics

Figure 3a is a trace plot. It shows all the lines appear to bounce around, are horizontal, and have a nice overlap between the chains, meaning that it is suitable for a negative binomial regression model.

Figure 3b is a Rhat plot. It shows that all the values are close to 1 and not more than 1.1, which shows that the negative binomial regression model has no issues.
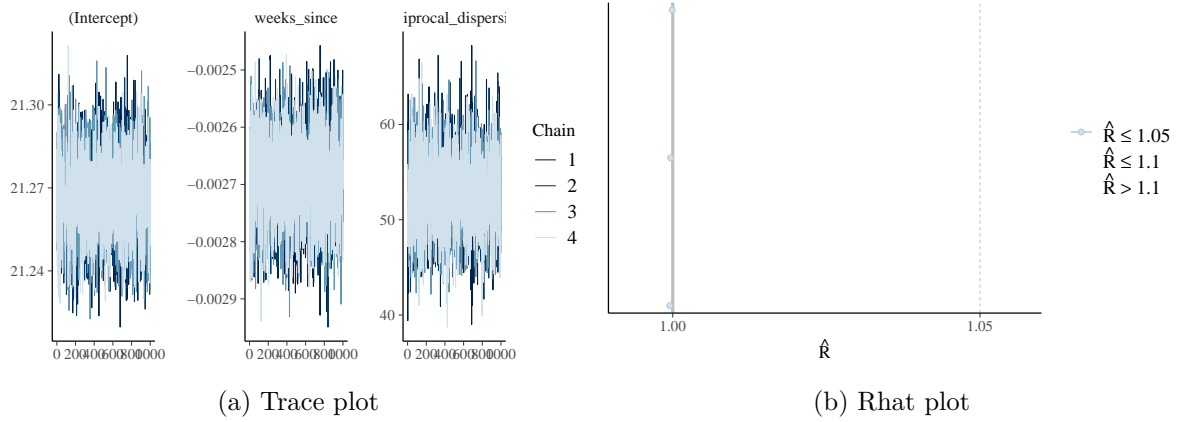

(a) Trace plot   (b) Rhat plot

Figure 3: Checking the convergence of the MCMC algorithm

References

Cragg, Michael. 2024. "'This Is an Art Form – and We're Losing It': Is the Music Video Dying?" https://www.theguardian.com/music/2024/apr/05/has-tiktok-killed-music-videos.

Goodrich, Ben, Jonah Gabry, Imad Ali, and Sam Brilleman. 2022. "Rstanarm: Bayesian Applied Regression Modeling via Stan." https://mc-stan.org/rstanarm/.

Google. 2024a. "How Engagement Metrics Are Counted." https://support.google.com/youtube/answer/2991785?hl=en.

———. 2024b. "Music Charts and Insights." https://support.google.com/youtube/answer/9014376?hl=en.

Harrison, John. 2022. *RSelenium: R Bindings for 'Selenium WebDriver'.* https://CRAN.R-project.org/package=RSelenium.

kworb. 2024. "How Engagement Metrics Are Counted." https://kworb.net/youtube/weekly/.

R Core Team. 2023. *R: A Language and Environment for Statistical Computing.* Vienna, Austria: R Foundation for Statistical Computing. https://www.R-project.org/.

Wickham, Hadley. 2024. *Rvest: Easily Harvest (Scrape) Web Pages.* https://CRAN.R-project.org/package=rvest.

Youtube. 2024. "Weekly Top Music Videos." https://charts.youtube.com/charts/TopVideos/global/weekly.