

# Specific Audio Features are not Strong Predictors of K-pop Idol Groups' Success in the Western Market \*

Randall Ni

April 20, 2023

With the success of K-pop idol groups such as BTS and Twice entering the Western market, the question for many other companies became whether they can reproduce the success for their own idol groups. To explore whether specific elements of the songs correlate with increasing hits and audiences, this paper aims to investigate the relationship between various audio features of tracks and the popularity value on Spotify. In order to maintain fairness, this paper will only analyze female K-pop idol groups that were formed due to survival shows. This ensures that they have an adequate fanbase and are comparable to each other. By utilizing data pulled from Spotify API, we found that popularity value has a weak negative correlation with the acoustictness and speechiness variable. In addition, implications and possible reasons behind this correlation are discussed and explanations for outliers are provided.

## 1. Introduction

The Korean pop music industry, commonly known as K-pop, has seen a surge in popularity worldwide over the past few years. K-pop has become a global phenomenon, and its popularity has continued to grow steadily, particularly in Western markets. One factor that has contributed to the increasing popularity of K-pop in the West is the emergence of survival shows that form K-pop groups. These shows, such as “Produce 101” and “Girls Planet 999”, requires viewers to vote for their favorite idols to stay on the show; they have been instrumental in the formation of many successful K-pop groups.

The success of these survival show groups has also led to an increase in the overall popularity of K-pop in the Western market. As the popularity of K-pop continues to grow, there is a growing

---

\*Code and data are available at: <https://github.com/randalln1140/K-pop-Female-Group-Western-Success>

need for the industry to expand in the Western market. This expansion is important for several reasons. Firstly, the Western market represents a significant opportunity for growth, with a large and diverse audience that is hungry for new and exciting music. Secondly, expanding into the Western market can provide K-pop groups with access to new revenue streams, such as merchandise sales, concerts, and collaborations with Western artists. Finally, expanding into the Western market can help to establish K-pop as a mainstream genre, and solidify its position as a global cultural phenomenon. Therefore, I'm interested in whether increasing hits and audiences is correlated with any factors from the songs. If so, which attributes will affect the song's popularity the most? And is it possible to predict how popular a song can be in the future based on certain attributes?

The data used in this report is obtained from Spotify API. The datasets are specified to the five survival show K-pop idol groups that I have selected ("IZ\*ONE", "fromis\_9", "I.O.I", "Momoland", and "Kep1er"), and it provides different kinds of information about every released songs such as danceability, tempo, and valence; a brief explanation of each variables can be found in section 2. In these data sets, I'm particularly interested in the different attribute values between their most popular songs and least popular songs, and if we can establish any trends based on these statistics. The estimand for this paper is the popularity variable provided by the Spotify API. In section 2, cleaned data sets are obtained and explained to perform further analysis. Visualizations in the forms of tables and figures are presented to help to explain the possible correlation between the different attributes and the popularity variable in Spotify API. In section 3, a model is constructed to justify the relationship between acoustiness, speechiness and popularity, and to make future predictions. I found that there are weak negative correlations between the variables examined, and audio features are not good predictors for K-pop songs' success in the Western market. The interpretation of the final model and its statistical relevance is presented in section 4. A discussion is carried out in section 5 on the implications of the findings regarding the generated graphs and the Bayesian prediction model, possible application to K-pop songwriters and composers, as well as the weaknesses and future steps of this paper.

## **2. Data**

### **2.1 Data Set and Methodology**

The data set used in the analysis is a combination of 10 different data sets pulled from the Spotify API; they include audio features of the selected idol groups and their popularity on the Spotify app. The meta data provided is collected from all Spotify users from available Spotify markets (eg. North America, Europe...). The API is being updated constantly; however, I have pulled and stored my data internally on 19 April 2023. Having static data is reasonable for the overall scope and validity of the questions, since 3 out of the 5 selected groups have been disbanded (won't release anything new) and the rest did not announce new album releases in the near future. The data will be relatively stable in the future, unless a black swan event

occurs and particular songs suddenly gain popularity. Using R (R Core Team 2022), tidyverse (Wickham et al. 2019), knitr (Xie 2023), dplyr (Wickham et al. 2022), spotifyr (Thompson et al. 2022), broom.mixed (Bolker and Robinson 2022), rstanarm (Goodrich et al. 2023), and janitor (Firke 2021), I cleaned and extracted the necessary data to complete an exploratory analysis and modelling. All figures are generated using ggplot (Wickham 2016).

## 2.2 Variables Explanation

There are several variables that come with the Spotify audio features that might affect a song's popularity factor, and a brief explanation for each variable that I have graphed is shown below.

- **Danceability:** Danceability describes how suitable a track is for dancing based on a combination of musical elements including tempo, rhythm stability, beat strength, and overall regularity. A value of 0.0 is least danceable and 1.0 is most danceable.
- **Energy:** Energy is a measure from 0.0 to 1.0 and represents a perceptual measure of intensity and activity. Typically, energetic tracks feel fast, loud, and noisy. For example, death metal has high energy, while a Bach prelude scores low on the scale. Perceptual features contributing to this attribute include dynamic range, perceived loudness, timbre, onset rate, and general entropy.
- **Speechiness:** Speechiness detects the presence of spoken words in a track. The more exclusively speech-like the recording (e.g. talk show, audio book, poetry), the closer to 1.0 the attribute value. Values above 0.66 describe tracks that are probably made entirely of spoken words. Values between 0.33 and 0.66 describe tracks that may contain both music and speech, either in sections or layered, including such cases as rap music. Values below 0.33 most likely represent music and other non-speech-like tracks.
- **Acousticness:** A confidence measure from 0.0 to 1.0 of whether the track is acoustic. 1.0 represents high confidence the track is acoustic.
- **Valence:** A measure from 0.0 to 1.0 describing the musical positiveness conveyed by a track. Tracks with high valence sound more positive (e.g. happy, cheerful, euphoric), while tracks with low valence sound more negative (e.g. sad, depressed, angry).
- **Tempo:** The overall estimated tempo of a track in beats per minute (BPM). In musical terminology, tempo is the speed or pace of a given piece and derives directly from the average beat duration.
- **Popularity:** The popularity of a track is a value between 0 and 100, with 100 being the most popular. The popularity is calculated by algorithm and is based, in the most part, on the total number of plays the track has had and how recent those plays are. Generally speaking, songs that are being played a lot now will have a higher popularity than songs that were played a lot in the past.

## 2.3 Data Visualization

### 2.3.1 Release Date

Since Spotify developers took the recency of play into account when calculating the overall popularity of the tracks, Table 1 and Table 2 below shows the album release date of top 10 most popular tracks within the selected groups and the 11 least popular tracks within the selection respectively. The reason why this paper has included the bottom 11 instead of 10 is because there are tracks that have the same popularity score, and it is reasonable to argue that not choosing to only include one or using any of them may affect the validity of the analysis. We can see a very clear trend in the tables below. For the most popular tracks, apart from “La Vie en Rose” and “BBoom BBoom”, all other songs can be considered relatively recent; they were released during and after 2020. On the other hand, for the least popular tracks, 8 out of 11 were released before 2020, with 7 of them released between 2016 and 2017. This shows that the “album\_release\_date” variable correlates with popularity to a certain degree.

Interestingly, We can also identify the relative popularity of the groups using the tables below. “IZONE” has 4 tracks in the top 10, while “Kep1er” only has 3. Furthermore, if we are taking into account album release time, “IZONE”’s tracks were all at least one year prior to “Kep1er”’s. Since we cannot obtain the exact formula for how Spotify calculates the popularity variable, we cannot examine the importance of recent plays. However, it is reasonable to argue that “IZONE”, on average, has the most popular tracks out of the five groups selected. For the least popular 11 tracks, “I.O.I” contributed 6 of them. In addition, none of the tracks sung by “I.O.I” made it to the top 10 table. Therefore, “I.O.I” is ranked lowest in terms of track popularity in the western market.

Table 1: Table 1: Extracting Top 10 Most Popular Tracks made by Selected K-pop Idol Groups

Track Name	Popularity	Album Release Date	Artist Name
WA DA DA	69	2022-01-03	Kep1er
Panorama	69	2020-12-07	IZ*ONE
La Vie en Rose	68	2018-10-29	IZ*ONE
BBoom BBoom	66	2018-01-03	Momoland
FIESTA	66	2020-02-17	IZ*ONE
Secret Story of the Swan	66	2020-06-15	IZ*ONE
DM	65	2022-01-17	fromis_9
Up!	64	2022-06-20	Kep1er
Giddy	63	2023-04-10	Kep1er
Wrap Me In Plastic	63	2021-02-05	Momoland

Table 2: Table 2: Extracting the 11 Least Popular Tracks made by Selected K-pop Idol Groups

Track Name	Popularity	Album Release Date	Artist Name
Knock Knock Knock	30	2016-05-04	I.O.I
Hold On	38	2016-10-17	I.O.I
Dream Girls	39	2016-05-04	I.O.I
Freeze	39	2017-08-22	Momoland
Crush	40	2016-04-05	I.O.I
Okay!	43	2021-08-27	Momoland
When The Cherry Blossoms Fade	44	2016-05-04	I.O.I
We Fresh (Japanese version)	44	2023-03-15	Kep1er
BANANA CHACHA	44	2019-04-03	Momoland
Pick Me	45	2016-05-04	I.O.I
Ready Or Not	45	2020-11-17	Momoland

### 2.3.2 Danceability

For danceability and all the graphs later, the track names are ordered in a descending order(the first name being the most popular and the last being the least popular). In Figure 1 and 2 below, I did not spot any visible trends. Both the most popular 10 and the least popular 11 songs have danceability values ranging between 0.7 to 0.8 with one outlier each.

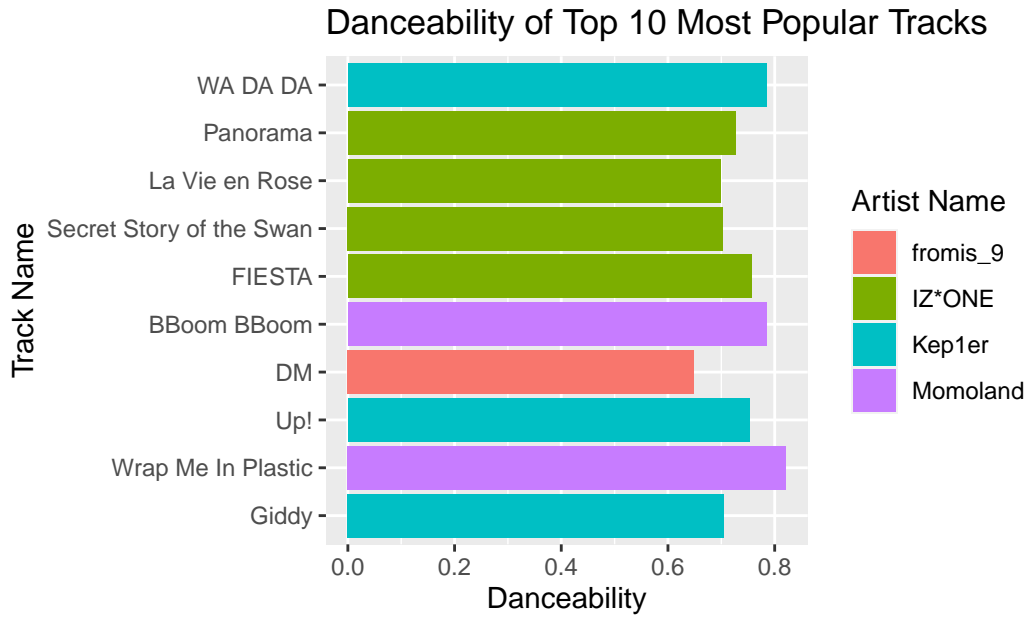


Figure 1: Danceability of each song with descending popularity order

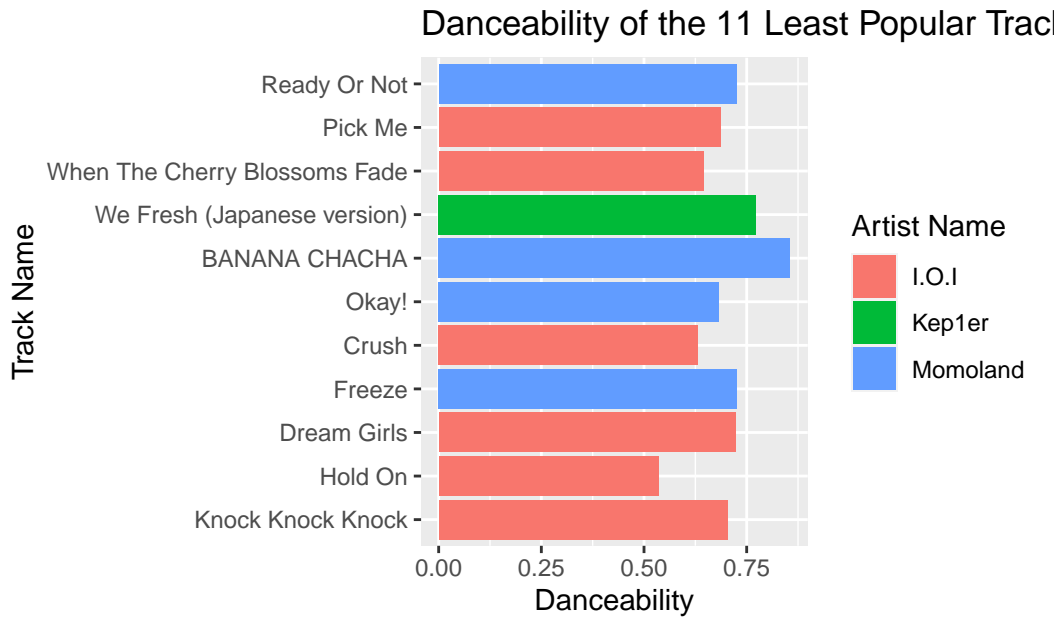


Figure 2: Danceability of each song with descending popularity order

### 2.3.3 Energy

In terms of the energy level of the tracks, there are also no visible trends. Figure 3 and Figure 4 below shows that the majority of the songs, regardless of popularity, have an energy value around 0.8 to 0.9. Interestingly, the songs that did well for “Momoland” all have low energy levels when compared to other songs that they have produced; we can see that there is a small 0.1 - 0.2 increase in energy value, but their tracks popularity value decreased around 20 points.

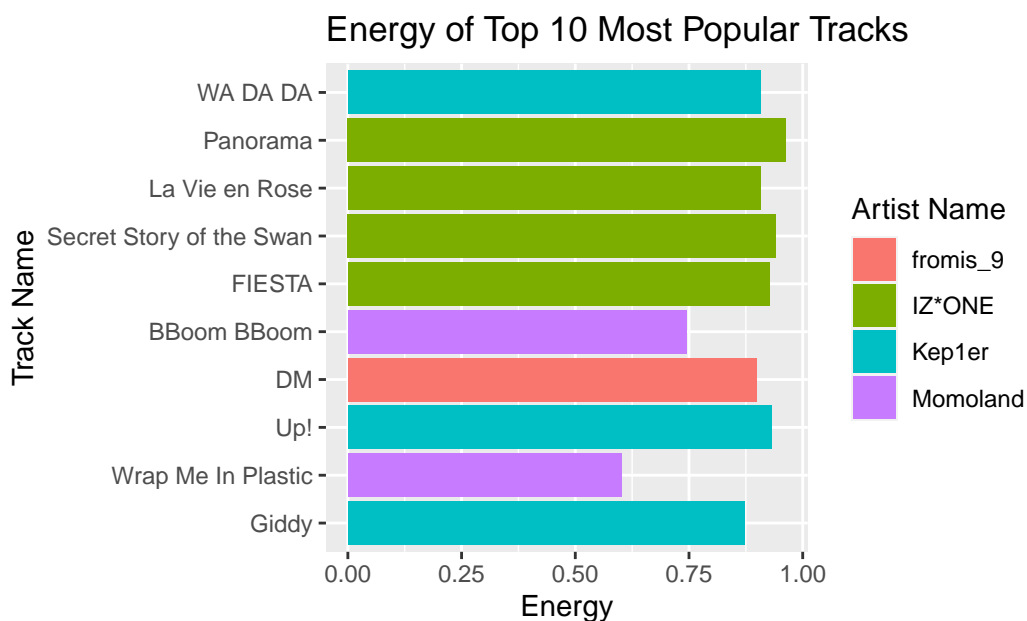


Figure 3: Energy of each song with descending popularity order

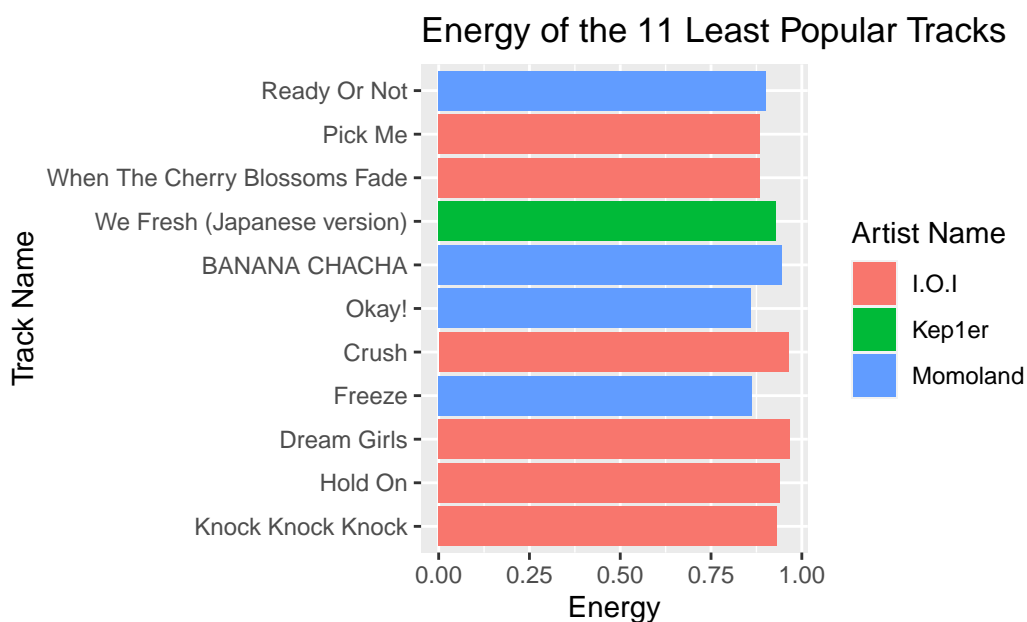


Figure 4: Energy of each song with descending popularity order

#### 2.3.4 Speechiness

The comparison of speechiness of the tracks is shown below in Figure 5 and 6. We can see that

there is a trend within the graphs. Only 3 out of the 10 most popular tracks have speechiness over 0.1, while over half of the least popular tracks are closely approaching or surpassed 0.1. Interestingly, tracks that “Momoland” sang that did well generally have lower speechiness value when compared to the songs that they did not do well in.

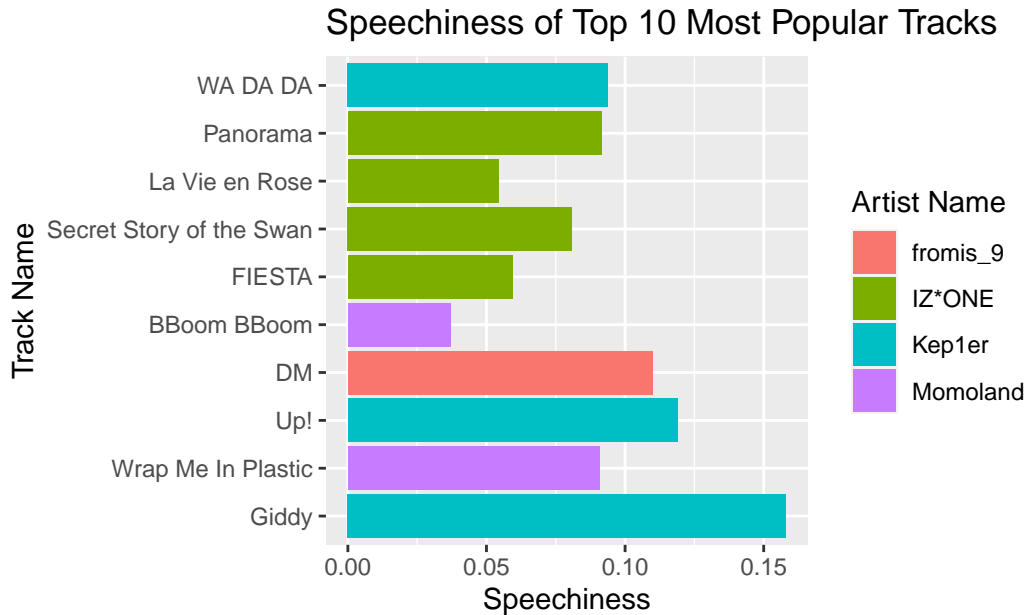


Figure 5: Speechiness of each song with descending popularity order

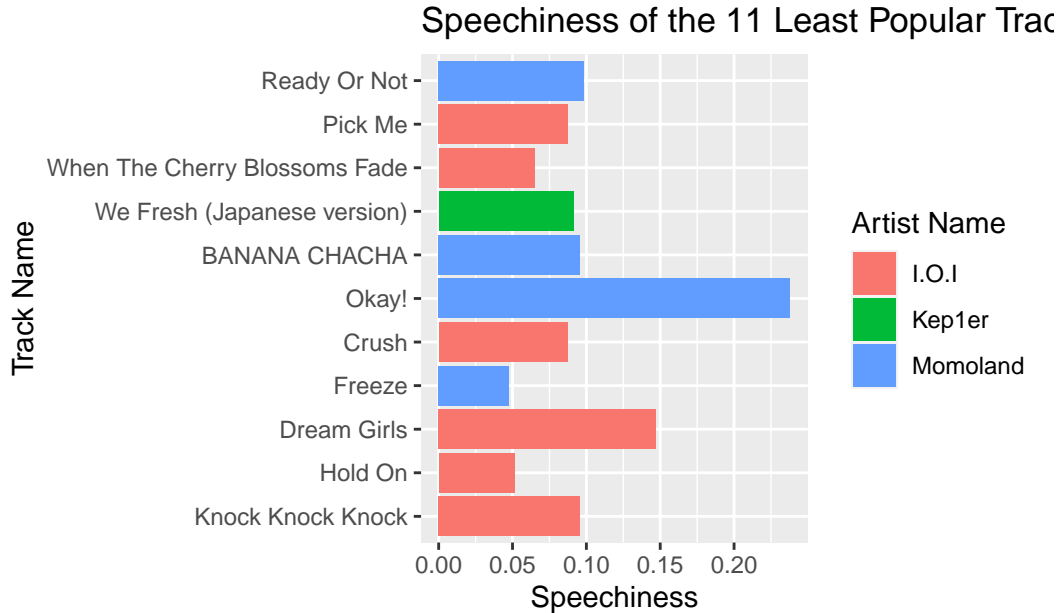


Figure 6: Speechiness of each song with descending popularity order



### 2.3.5 Acousticness

Figure 7 and 8 below showcased the difference in acousticness for the selected tracks. We can see that there is a negative correlation with acousticness and popularity. Only 2 tracks in the most popular category have an acousticness value over 0.1, while around half of the least popular tracks surpassed the acousticness value of 0.1.

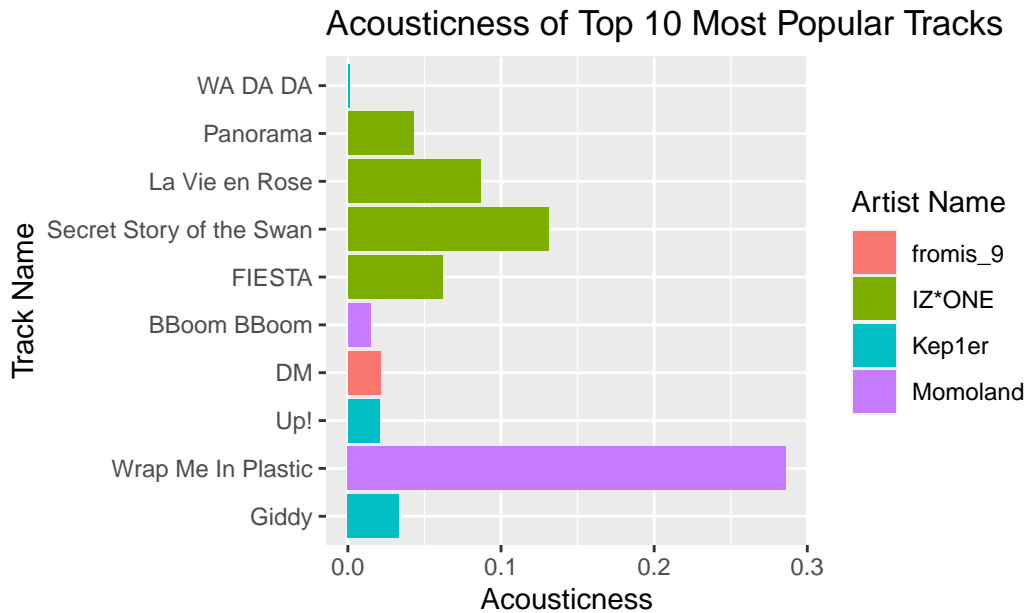


Figure 7: Acousticness of each song with descending popularity order

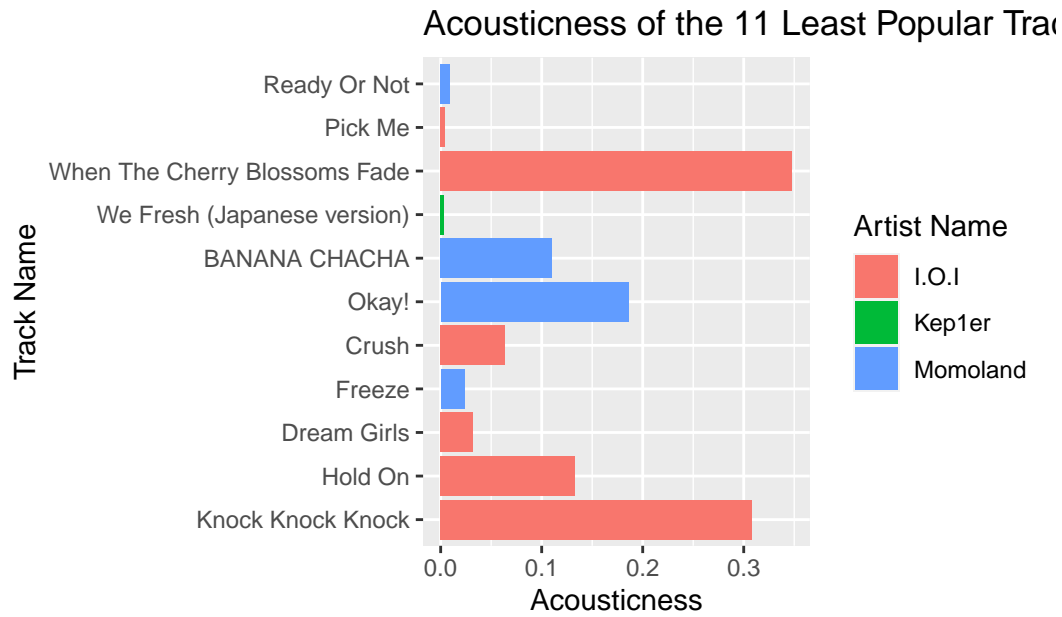


Figure 8: Acousticness of each song with descending popularity order

### 2.3.6 Valence

In terms of valence, figure 10 and 11 below does not provide us with a clear trend. The majority of the tracks fall between the value of 0.5 to 0.75, with the same amount of tracks surpassing 0.75 in both categories.

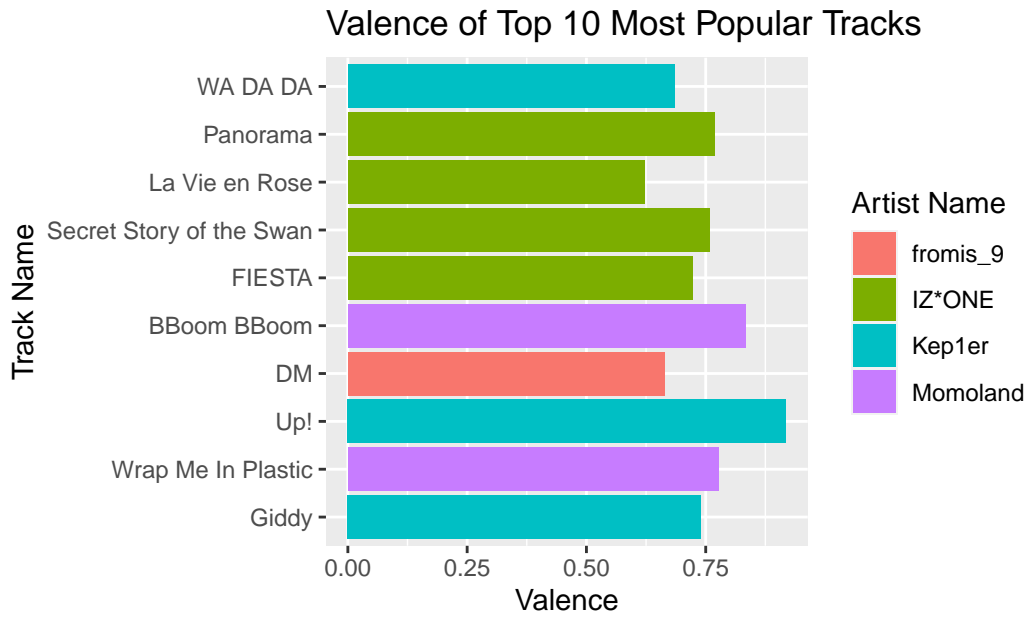


Figure 9: Valence of each song with descending popularity order

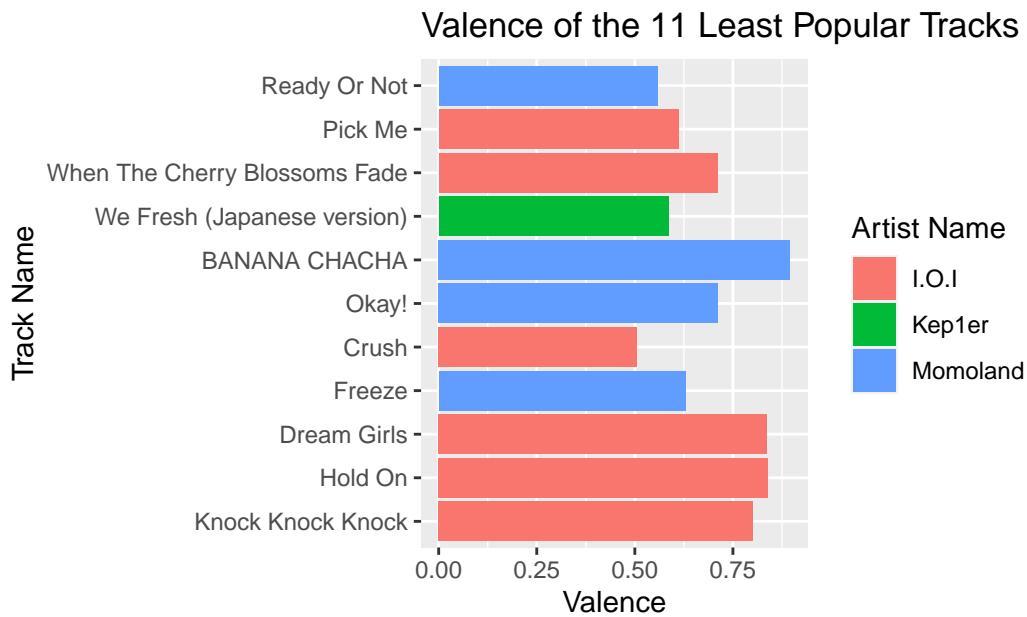


Figure 10: Valence of each song with descending popularity order

### 2.3.7 Tempo

Lastly, figure 11 and 12 below highlight the tempo of each track in the given categories. Tracks generally teeter around a tempo value of 125 for both categories with only a few outliers. This does not show a clear trend as well.

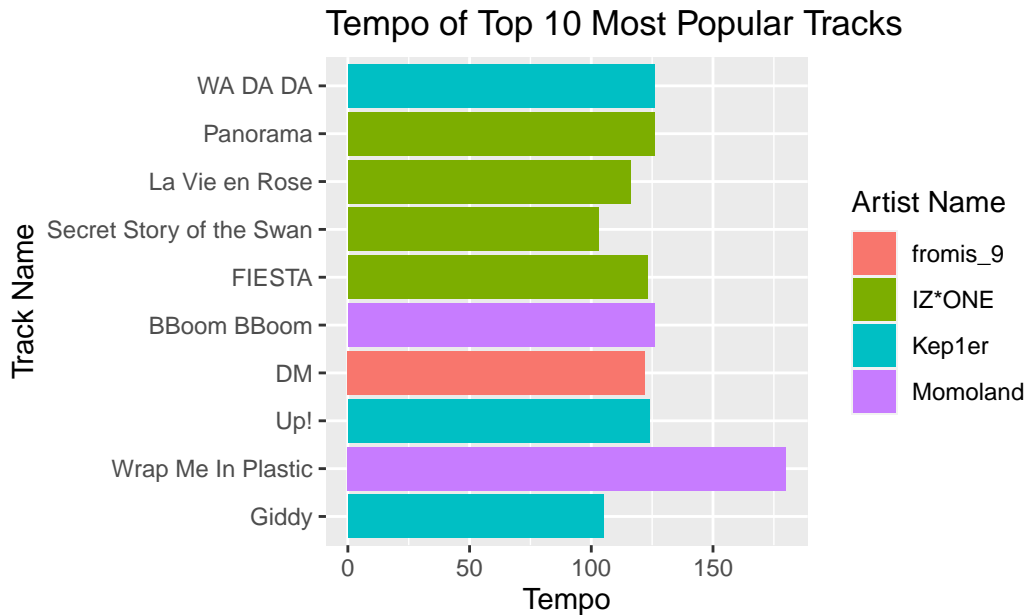


Figure 11: Tempo of each song with descending popularity order

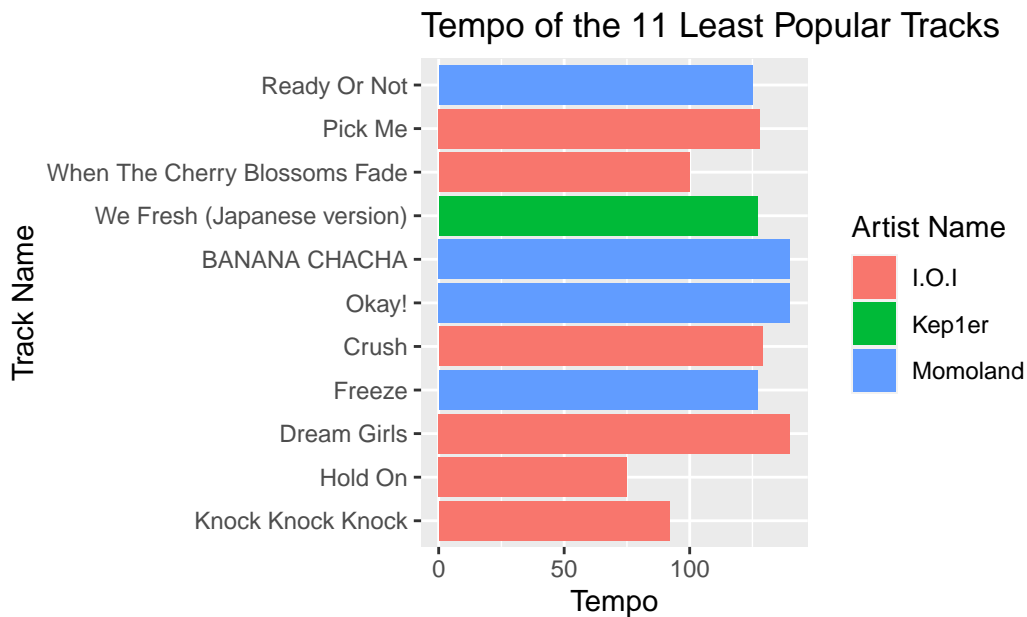


Figure 12: Tempo of each song with descending popularity order

### 3. Model

The goal of creating this prediction model is to answer my questions stated at the beginning of this paper. Will there be certain elements in the songs that contributed to their success in the Western market? If no, can we explain the popularity of tracks with other means? If yes, can we replicate the success for future releases?

Bayesian analysis model is a perfect tool for creating a prediction model that aligns with my goals. Due to a low amount of observation, the Bayesian analysis model is perfect to make predictions based on scarce and incomplete data through thousands of iterations of sampling tests within seconds.

#### 3.1 Model Set Up

Define  $y_i$  as the popularity value in Spotify. Then  $\beta_i$  is the “speechiness” variable and  $\gamma_i$  is the “acousticness” variable.

$$y_i | \mu_i, \sigma \sim \text{Normal}(\mu_i, \sigma) \quad (1)$$

$$\mu_i = \alpha + \beta_i + \gamma_i \quad (2)$$

$$\alpha \sim \text{Normal}(0, 2.5) \quad (3)$$

$$\beta \sim \text{Normal}(0, 2.5) \quad (4)$$

$$\gamma \sim \text{Normal}(0, 2.5) \quad (5)$$

$$\sigma \sim \text{Exponential}(1) \quad (6)$$

We run the model in R (R Core Team 2022) using the “rstanarm” package (Goodrich et al. 2023). We use the default priors from “rstanarm”.

#### 3.2 Model Justification

I have included 2 variables “speechiness” and “acousticness” into my model, and I expected to see a negative relationship between both variables and the popularity value. Those two variables show their negative correlations clearly in the graphs from the previous sections, and they should carry it over to the model created.

Table 3: Predicted Model	
(Intercept)	58.07 (3.58)
speechiness	-40.01 (33.74)
acousticness	-10.54 (10.86)
Num.Obs.	50
R2	0.059
R2 Adj.	-0.080
Log.Lik.	-180.628
ELPD	-183.9
ELPD s.e.	4.3
LOOIC	367.9
LOOIC s.e.	8.6
WAIC	367.6
RMSE	8.81

#### 4. Results

Table 3 above is a Bayesian predictive model comparing the “speechiness” and “acousticness” variables with popularity scores. There is a very weak negative correlation between the variables, but overall, the result shows a large amount of volatility and potentially answered my question.

The intercept term of 58.07 represents the expected value of the outcome variable (in this case the popularity value) when all predictor variables are zero. In other words, if both “speechiness” and “acousticness” are zero, we would expect the outcome variable to have an average value of 58.07. The coefficient for “speechiness” is -40.01, which means that for every one-unit increase in speechiness, we expect the outcome variable to decrease by 40.01 units (assuming all other variables are constant). However, the standard error is 33.74, which means that the estimated speechiness coefficient values can vary by about 33.74 units around the true speechiness coefficient value if there are more data or the sampling tests are repeated more times. The coefficient for “acousticness” is -10.54, which means that for every one-unit increase in acousticness, we expect the outcome variable to decrease by 10.54 units (holding all other variables constant). In terms of standard error, acousticness also has a very large value of 10.86, which means that the estimated acousticness coefficient values can vary by about 10.86 units if more ttests are carried out.

The  $R^2$  value of this predictive model shows extreme volatility as well. An R-squared value of 0.059 indicates that only 5.9% of the variability in the popularity variable can be explained by the predictor variables included in the model. In addition,  $R^2$  value of 0.059 is generally

considered to be quite low, suggesting that the predictor variables in the model are not very strongly related to the outcome variable.

## 5. Discussion

### 5.1 Further Discussion with Model Results

The Bayesian model's result concluded that it is unlikely that "acousticness" and/or "speechiness" contributed towards a song's success in the Western market. Even if they do, the current data model result suggests that there is high volatility and a very weak negative correlation. The  $R^2$  value is also not statistically significant. For more general problems regarding the prediction model and the overall data set, please refer to "5.3 Weaknesses and Next Steps", where I will delve deeper into the potential problems that might have happened in this analysis.

### 5.2 Audio Attributes Most Likely not Contributing to the Popularity of the Songs

As the graphs and the Bayesian prediction model shown in the previous sections, audio features most likely do not contribute to the popularity of the selected K-pop songs in the Western Market. Instead, the songs popularity might be mainly explained by the group's popularity or how relevant the song is in the current moment.

For example, frequent music show wins are a good predictor for the success of a K-pop song within the Asian market. Some of the examples that broadcast weekly are MBC Music's "Show Champion" and Mnet's "M!Countdown". Their metrics include "digital", "physical album sales", "Youtube music video views", "expert score", and "audience live votes". Winning those shows highlight the song's and group's popularity relatively fairly. We have predicted that "IZONE" is the most popular K-pop group out of the selected 5 through generated tables and graphs. This is also supported by their total of 26 various music show wins throughout their career as a team. They are ranked 30th overall on the K-pop idol group that won the most music show list. For other groups in the data set, "Momoland" has won 11 times (placed 44th overall), "I.O.I" has won 9 times (placed 46 overall), "fromis\_9" has won 8 times (placed 47 overall), and "Kep1er" has won 4 times (placed 51st overall).

Interestingly, apart from successfully predicting "IZONE"'s dominant popularity, the Spotify API failed to reflect other group's success accordingly; it is reasonable to argue that it is due to temporal factors, since the popularity value also correlates with how recent users played those specific songs.

"I.O.I"'s songs did win a moderate amount of music shows, but their popularity is not highlighted by the Spotify API. Out of the 10 songs included for each group, over half of "I.O.I"'s songs are placed into the least popular category. Their "Knock Knock Knock" has less than half of the popularity score for "Kep1er"'s "WADADA". Since we did not find any noticeable predictor in the audio features, it is very likely that the low popularity score is due to the early release

date; a vast majority of their songs were released in 2016, which was a long time ago and might cause low present day listener count. On the other hand, “Kep1er” is a relatively new idol group compared to other selected groups. All of their songs that made into the top 10 most popular tracks were released after 2022, which gave them a significant advantage over older disbanded groups.

Another very interesting element that I found after digging deeper into individual song is the possible success factor in “Momoland”’s song “Wrap Me in Plastic”. “Momoland” is considered as one hit wonders in the Asian market, while “Bboom Bboom” was very catchy and has addictive rhythm, the group’s other songs were criticized due to being very similar. I was very surprised to see a second song made by them to make it into the top 10. After hearing the full song, I noticed that it is very different from traditional K-pop songs. Although English is vastly used by song writers to fill in words or be onomatopoeia (majority of song lyric is still Korean, unless the song is specifically made for English speaking markets), the majority of the lyrics for “Wrap Me in Plastic” is English. This might be a very important contributing factor to its surprisingly high popularity value. High inclusion of English lyrics might help K-pop songs to be more acceptable in English speaking markets.

### **5.3 Practical Implications and Conclusion**

Practical implications can be gained from this paper’s findings as well. Due to the importance of timing, K-pop agencies should focus more on marketing and making their group stay as relevant as possible. This will greatly help with the temporal factor in determining a track’s popularity score. With the result of the Bayesian model, although the negative correlation is very weak, songwriters should still take it into consideration. K-pop songs should have a relatively lower amount of speech like recordings and acoustics if they wish to have a high popularity score determined by the Spotify API. In addition, another factor explained above is the strength of including English lyrics. Although I do not consider full English songs K-pop (even when the performing artist is from a K-pop group), it seems beneficial for K-pop groups that wish to have higher popularity within the Western market to include more English words when writing their songs.

In conclusion, while audio features provided by the Spotify API can provide developers with valuable information about the acoustic characteristics of songs, they might not be reliable predictors of success in the Western music market. Other factors, such as marketing, promotion, and general cultural acceptance, might play a more significant role in determining the success of K-pop songs in the Western market.

### **5.4 Weaknesses and Next Steps**

Several weaknesses of the data sets used can be spotted and they might negatively impact my analysis. To begin with, the small set of observations really cannot speak for every K-pop group.



The data set that I have created by pulling various individual track data from the Spotify API cannot really translate well to the entire K-pop industry. The specificity of only using female K-pop groups that were formed due to survival shows makes the amount of data very limited. They are also generally more short-lived than regular K-pop groups, which resulted in less tracks available. In addition, I have only analysed female groups, which might not translate well when it comes to male groups. This has also been reflected in the Bayesian model with very large standard errors and a possible over fitting model. Next, without the exact formula that the Spotify developers used to calculate popularity value, we can never really know the extent of the recency effect for songs. The only thing that we can be certain about is to always make the group relevant and keep producing more songs; this will garner more recent hits which will positively affect the popularity value. Lastly, bias is also present in terms of the chosen data. Since I wanted to keep the groups fairly similar in terms of song types, this resulted in the overwhelming amount of songs that are generally cheery and uplifting. This will cover the majority of K-pop songs, but won't cover the entire K-pop market. For example, there are extremely popular ballad singers such as IU and Younha, whose songs won't fit the ones represented in this paper. Other genres like rap also won't fit the genres examined in this paper.

The possible next step for this paper is to include more data into the data set. There are three ways that I can advance this research further. The first being including more popular/diverse K-pop idol groups such as "Twice" and "BTS". This can possibly strengthen my claim about the popularity being more influenced by the group's popularity instead of specific song elements. Another future step that I can take is to include other genres in K-pop, and I can try to highlight any differences between the genres' success factors in the Western market. Another step that I can take to answer my question stated in the previous sub-section (about including more English words into the lyrics) is to include songs that are purely English sung by Korean singers and check their audio features and popularity values.

## 6. References

- Bolker, Ben, and David Robinson. 2022. “Broom.mixed: Tidying Methods for Mixed Models.” <https://github.com/bbolker/broom.mixed>.
- Firke, Sam. 2021. *Janitor: Simple Tools for Examining and Cleaning Dirty Data*. <https://CRAN.R-project.org/package=janitor>.
- Goodrich, Ben, Jonah Gabry, Imad Ali, and Sam Brilleman. 2023. “Rstanarm: Bayesian Applied Regression Modeling via Stan.” <https://mc-stan.org/rstanarm/>.
- R Core Team. 2022. *R: A Language and Environment for Statistical Computing*. Vienna, Austria: R Foundation for Statistical Computing. <https://www.R-project.org/>.
- Thompson, Charlie, Daniel Antal, Josiah Parry, Donal Phipps, Tom Wolff, and Stephen Holsenbeck. 2022. “R Wrapper for the Spotify Web API.” <https://github.com/charlie86/spotifyr>.
- Wickham, Hadley. 2016. *Ggplot2: Elegant Graphics for Data Analysis*. Springer-Verlag New York. <https://ggplot2.tidyverse.org>.
- Wickham, Hadley, Mara Averick, Jennifer Bryan, Winston Chang, Lucy D’Agostino McGowan, Romain François, Garrett Grolemond, et al. 2019. “Welcome to the tidyverse.” *Journal of Open Source Software* 4 (43): 1686. <https://doi.org/10.21105/joss.01686>.
- Wickham, Hadley, Romain François, Lionel Henry, and Kirill Müller. 2022. *Dplyr: A Grammar of Data Manipulation*. <https://CRAN.R-project.org/package=dplyr>.
- Xie, Yihui. 2023. *Knitr: A General-Purpose Package for Dynamic Report Generation in r*. <https://yihui.org/knitr/>.