2025 MAMA Awards Song of the Year, Best Male Group & Best Female Group Prediction

Stephanie (Qing) Chen *

December 22, 2025

Abstract

The MAMA Awards is an annual music award ceremony that awards for outstanding achievements in the K-pop and Asian music industry. This study predicts the Song of the Year, Best Male Group and Best Female Group of the 2025 MAMA Awards held in Hong Kong, China on November 28th and 29th. Public popularity indicators such as Spotify followers and popularity, YouTube subscribers and interaction statistics, Google trend data, Circle Chart streaming data and physical album sales were collected through APIs and chart websites. Exploratory data analysis was conducted to identify the trends and correlation among the variables. The study implements multiple models including weighted ranking model, principal components analysis and K-Means clustering to predict the 2025 winner of these three categories. The results show that APT by Rose and Bruno Mars, Stray Kids and aespa are the strongest candidates in categories Song of the Year, Best Male Group and Best Female Group, respectively. The study indicates that social media activities and chart performance are strongly correlated with awards outcomes, though limitations such as data availability and evaluation system complexity exist.

Introduction

The MAMA Awards is one of the most influential music award ceremonies in the K-pop industry. The event was first launched in 1999 as a music video awards ceremony and has attracted international interest since mid-2000s. As social media and streaming platforms gain more importance in artist exposure, predicting awards winners using analysis of these popularity data becomes increasingly feasible.

With the expansion of fan economics these days, awards winners predictions rises more interest in the authorities of music industry. This study aims to predict the winners of 2025 MAMA Song of the Year, Best Male Group and Best Female Group using publicly accessible popularity indicators. Data collected from Spotify, YouTube, Google trend and Circle Chart

*Code and data are available at: `https://github.com/stephvk130/2025-MAMA-Prediction`

(South Korea music ranking chart) are used in this study. For some indicators, the study only considers the data ranged from October 2024 to the end of September 2025 to match the evaluation criteria of the awards.

The challenges in awards predictions arise as the judging criteria involve not only popularity performance but subjective expert evaluation as well. The 2025 MAMA award for Song of the Year criteria consist of 40% judges panel evaluation and 60% digital song downloads and streaming (40% from South Korea and 20% global). The criteria for Best Male Group and Best Female Group are not explicitly published. The accuracy in public platform data and the noise from fan-driven behaviors also increase the uncertainty of the prediction.

To address these challenges, this study compiles popularity data from multiple platforms, conducts exploratory data analysis and employs different models and techniques for investigation. The predictions indicate that **APT by Rose and Bruno Mars**, **Stray Kids** and **aespa** are the strongest candidates in Song of the Year, Best Male Group and Best Female Group categories, respectively. The study successfully predicts the winners of Song of the Year and Best Female Group. The winner of Best Male Group is different from our prediction, but it is reasonable given the limitation of data and uncertainty from the judge panel evaluation.

Data

The full nominee lists for 2025 MAMA Song of the Year, Best Male Group and Best Female Group are included in the Appendix. For Song of the Year category, Spotify popularity score and YouTube music video views, likes, and comments counts collected from APIs are the main variables. In terms of the Best Male Group and Best Female Group categories, this study focuses on the group performance over the 2025 MAMA evaluation period from October 2024 to September 2025. Through Spotify API, group followers, group popularity score, sum of popularity scores of albums released during the evaluation period are collected. YouTube group official channel subscribers and views count are scraped through YouTube API. To reflect the group performance over the specific period (2024.10 - 2025.09), the study also collects the album sales and album count for main albums released between the evaluation period from Circle Chart, a Korean national music chart that tabulates the relative weekly, monthly and yearly popularity of songs and albums in South Korea. The study also includes the streaming rank of songs in those albums released during the evaluation period, but only songs hit the top 100 of the weekly streaming chart are collected, as published on Circle Chart. To see the overall group popularity from the global search engine, google trend data between October 2024 and September 2025 are collected for each group.

Exploratory Data Analysis

Song of the Year

For this category, the variables collected are Spotify popularity score, YouTube music video views, likes, and comments counts. To evaluate the fan interaction, interaction ratio is

calculated by

$$interaction\ \ ratio = \frac{MV\ likes + MV\ comments}{MV\ views}$$

From the scatter plot in Figure 1, most songs have an interaction ratio below 0.05, and those with higher interaction ratios are spread across different Spotify popularity scores.

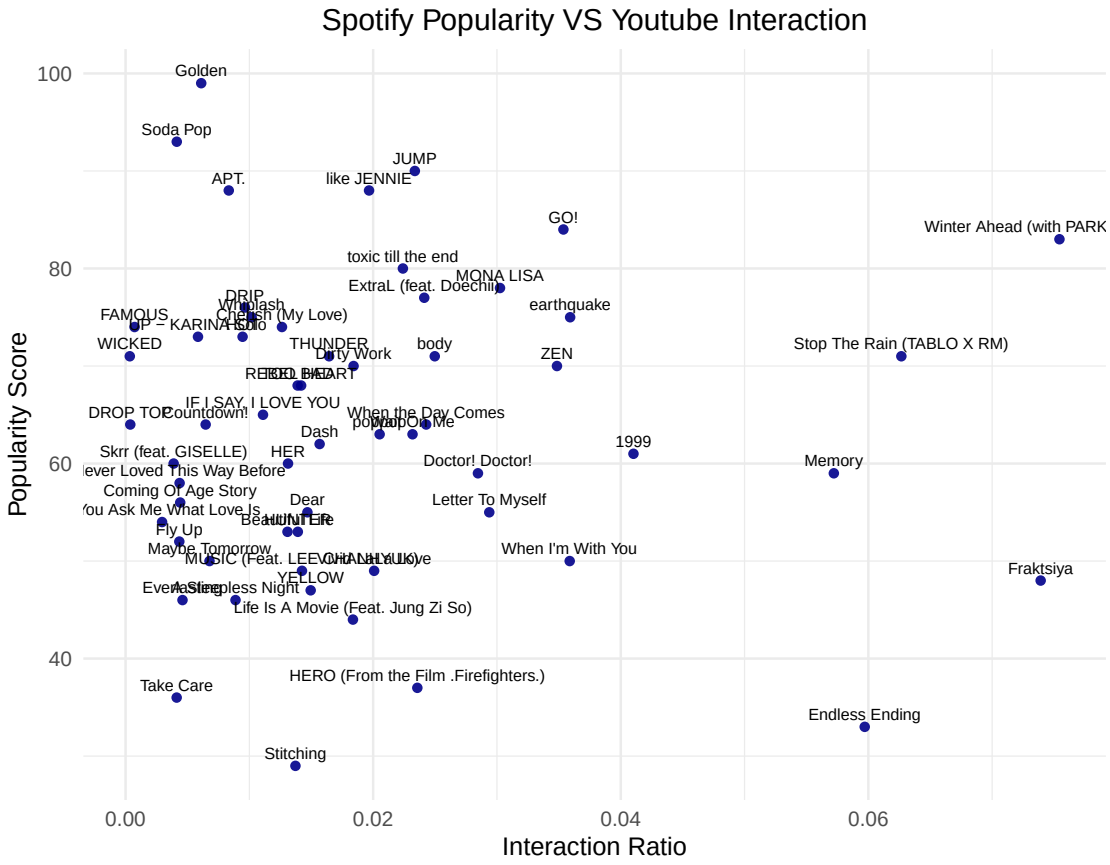

Figure 1: Spotify Popularity VS YouTube Interaction Ratio Each point corresponds to a song. While most songs have a low YouTube interaction ratio, their Spotify popularity scores show considerable variation.

rank	popularity_score	mv_view_count	mv_like_count	mv_comment_count	interaction_ratio
1	Golden	APT.	APT.	APT.	Winter Ahead
2	Soda Pop	Golden	JUMP	JUMP	Fraktsiya
3	JUMP	Soda Pop	Golden	YELLOW	Stop The Rain
4	like JENNIE	DRIP	like JENNIE	like JENNIE	Endless Ending
5	APT.	JUMP	earthquake	earthquake	Memory

Table 1: Top 5 Songs For each variable, rank the songs in descending order of their values and display the top 5.

From Table 1, *APT.* has the highest YouTube music video views, likes, and comments while *Golden* has the highest Spotify popularity score. Songs with high popularity scores and

strong YouTube music video statistics do not rank highly in interaction ratio, consistent with the pattern observed in the previous scatter plot.

Among all songs, *APT.* and *JUMP* rank top 5 in all variables except interaction ratio, following by *Golden* and *like JENNIE* ranked top 5 in 3 variables.

Best Male Group

For the seven male groups nominated (see Appendix), Spotify followers and popularity score, YouTube subscribers and total views are collected. To reflect the group performance over the evaluation period (2024.10 - 2025.09), sum of Spotify popularity scores of albums released during this period and main albums' physical sales are also included.

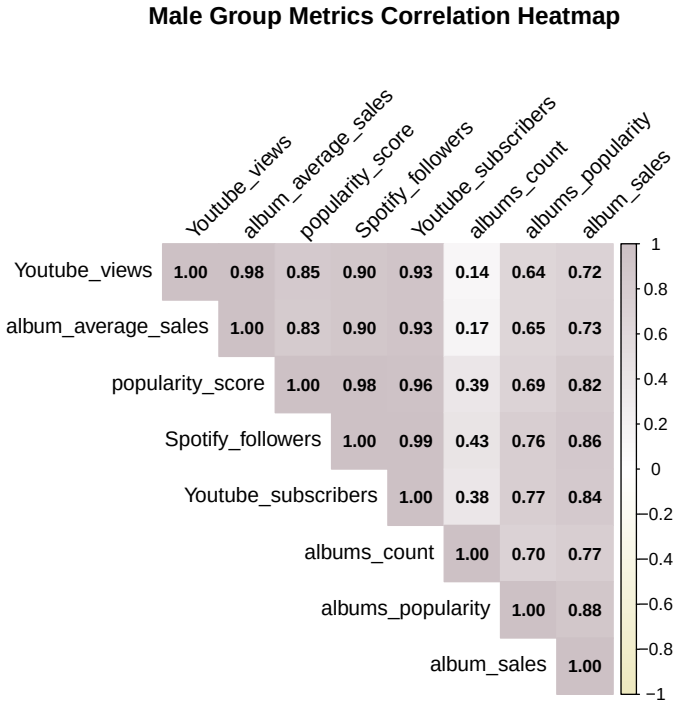

Figure 2: Male Group Correlation Heat Map The numbers represent the correlations between each pair of variables. High correlations are observed.

From the heat map in Figure 2, the correlations between these variables are high, except for albums count, the number of main albums released during the evaluation period. It is reasonable as these variables all reflect the popularity of the group.

Spotify followers, YouTube subscribers and total views are cumulative over time, while Spotify popularity score, album popularity and album sales (for albums released during the evaluation period) are based on recent data.

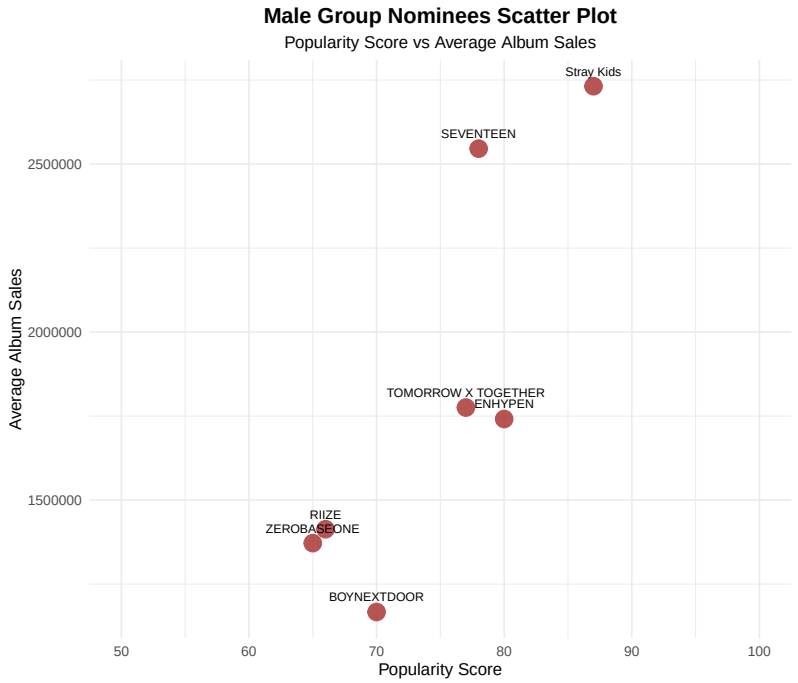

Figure 3: Male Group Spotify Popularity Score VS Average Album Sales Each point represents a male group. Three clusters and potential linear trend are observed.

Figure 3 suggests a positive linear trend between Spotify popularity and average album sales. Groups with higher Spotify popularity generally have higher sales, with *Stray Kids* leading both metrics, following by *SEVENTEEN*, which ranks second in sales and third in popularity.

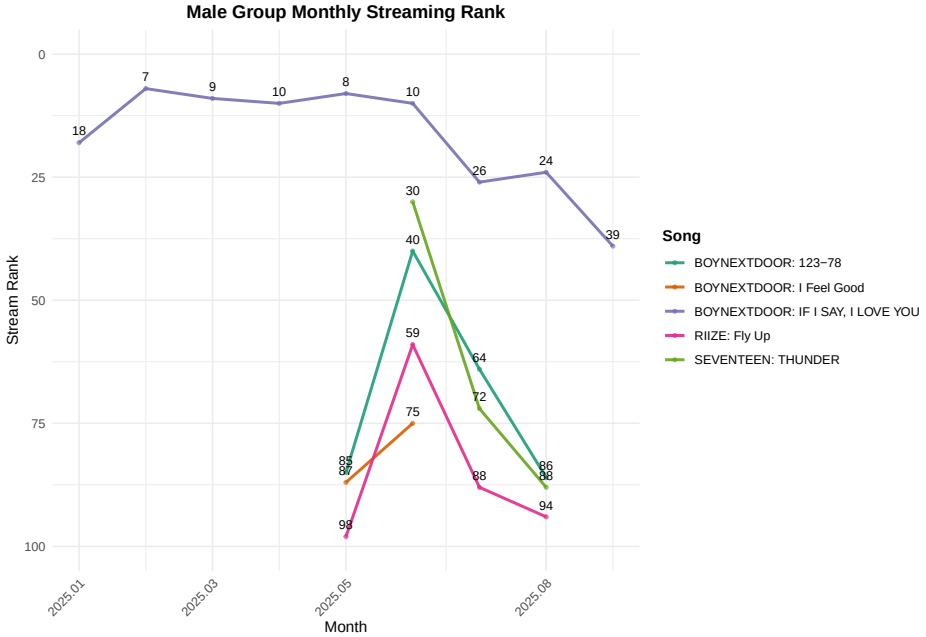

Figure 4: Male Group Stream Rank Songs by nominated male groups that entered the Circle Chart monthly stream Top 100 are recorded, with changes in monthly rankings.

In Figure 4, only songs from albums released during the evaluation period and ranked top 100 in monthly stream on Circle Chart are included. Streaming performance is completely different from the previous results. *BOYNEXTDOOR* has three songs hit the top 100 on Circle Chart, following by *SEVENTEEN* (1 song) and *RIIZE* (1 song).

Among these 5 songs, *BOYNEXTDOOR*'s *IF I SAY, I LOVE YOU* ranked top 40 for 9 months while the other songs only showed up at most 4 months with ranks lower than or equal to 30 on Chart. *BOYNEXTDOOR* has the best streaming performance over this evaluation period, based on Circle Chart.

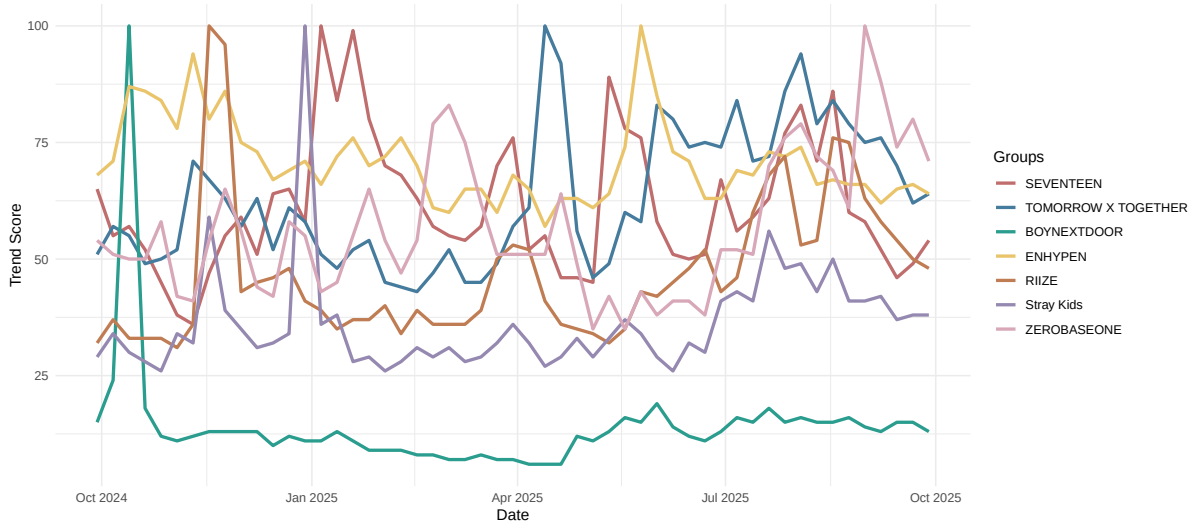

Figure 5: Male Group Google Trend Changes in weekly Google trend scores for each male group over the evaluation period are shown.

Surprisingly, from Figure 5, *BOYNEXTDOOR*, the group with the best streaming performance on Circle Chart, hit 100 in Google trend score at the start of evaluation period and then dropped to the lowest. This could be caused by the popularity difference in domestic Korea and global area. Trend score fluctuates over the evaluation period for each group. Overall, no group dominates the trend score over the entire evaluation period.

rank	Spotify_followers	popularity_score	albums_popularity	YouTube_subscribers	YouTube_views	album_average_sales
1	Stray Kids	Stray Kids	Stray Kids	Stray Kids	Stray Kids	Stray Kids
2	SEVENTEEN	ENHYPEN	TXT	SEVENTEEN	SEVENTEEN	SEVENTEEN

Table 2: Top 2 Male Group For each variable, rank the male groups in descending order of their values and display the top 2.

Table 2 shows *Stray Kids* ranks the 1st in all variables above, dominating all popularity metrics collected. *SEVENTEEN* ranks the 2nd in four variables, following by *ENHYPEN* and *TOMORROW X TOGETHER* ranked the 2nd in one variable.

Best Female Group

Similar to the Best Male Group category, Spotify followers and popularity score, YouTube subscribers and total views are collected for the six female groups nominated (see Appendix). To reflect the group performance over the evaluation period (2024.10 - 2025.09), sum of Spotify popularity scores of albums released during this period and main albums' physical sales are also included.

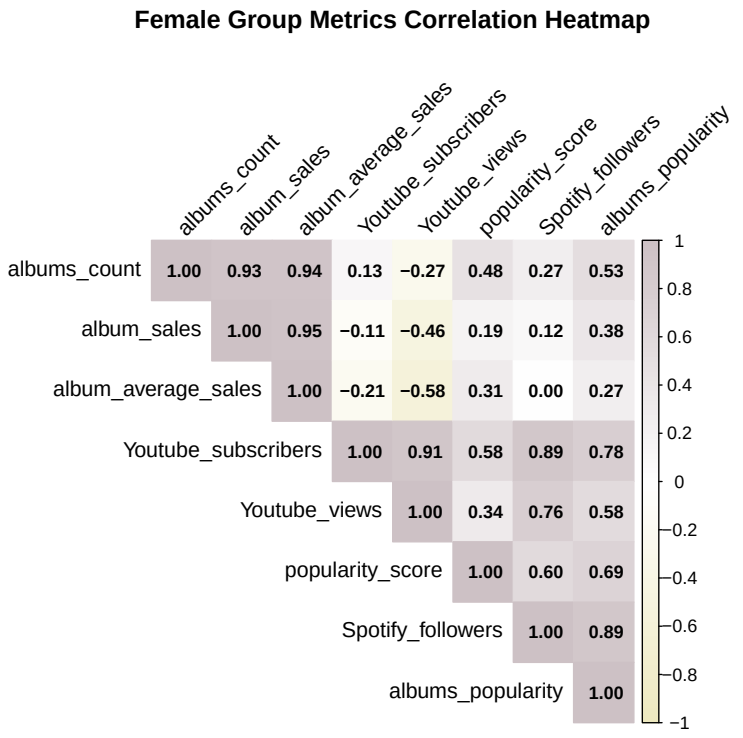

Figure 6: Female Group Correlation Heat Map The numbers represent the correlations between each pair of variables.

From Figure 6, the correlations between some variables are higher than others, such as albums count, album sales and album average sales, as well as YouTube subscribers and YouTube views. This is expected as variables regarding albums are directly correlated, same for variables from YouTube data.

Spotify followers, YouTube subscribers and total views are cumulative data. For 2025 awards, the study focuses on recent popularity metrics such as album popularity and average album sales.

From Figure 7, Spotify popularity scores for female groups nominated are concentrated between 70 and 85, while the average album sales vary substantially. There is no clear linear relationship between these two metrics. *IVE* has the highest average album sales while *TWICE* gets the highest popularity score.

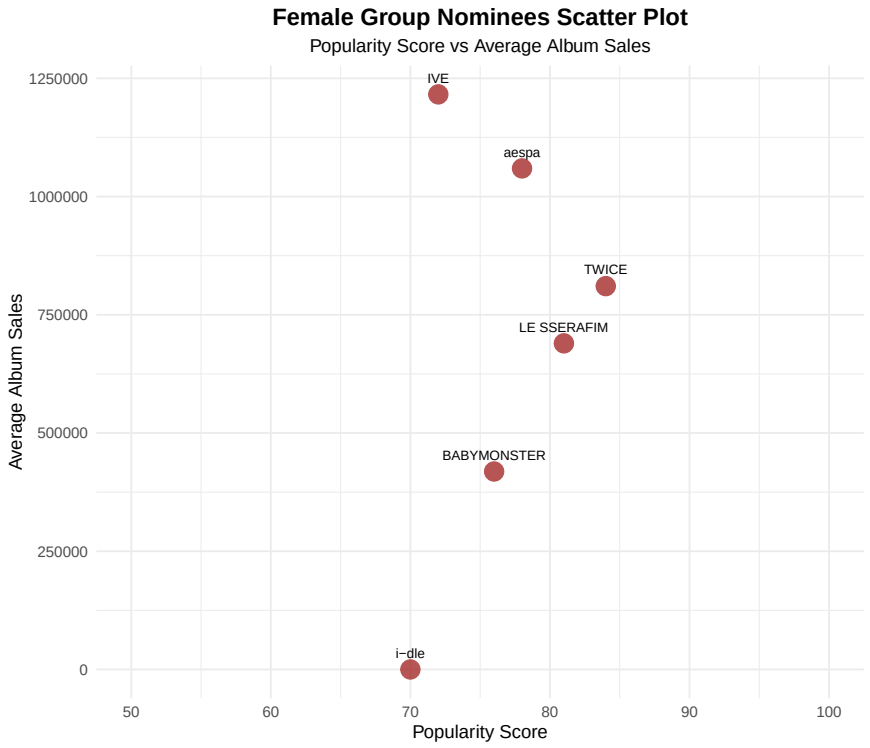

Figure 7: Female Group Spotify Popularity Score VS Average Album Sales Each point represents a female group, with popularity scores tightly clustered and average album sales widely dispersed.

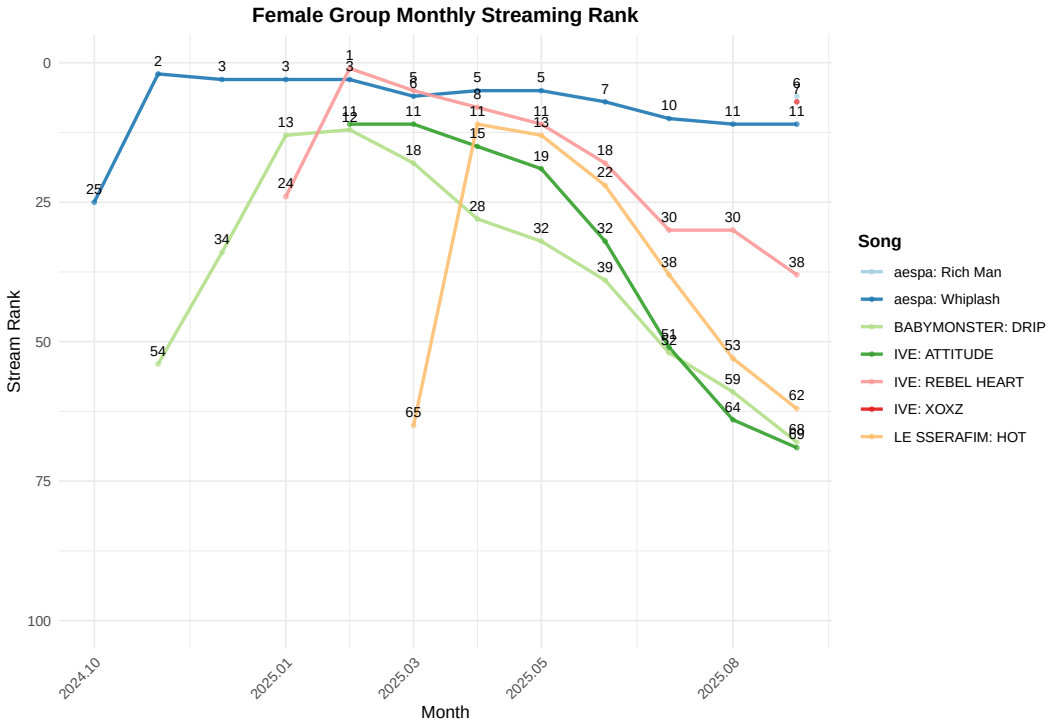

Figure 8: Female Group Stream Rank Songs by nominated female groups that entered the Circle Chart monthly stream Top 100 are recorded, with changes in monthly rankings

The streaming rank plot for female groups in Figure 8 only includes songs from albums released during the evaluation period and ranked top 100 in monthly stream on Circle Chart. Female groups have much better streaming performance comparing to that of male groups. *"Whiplash"* by *aespa* has the dominated rank for almost the entire evaluation period, following by *IVE*'s *"REBEL HEART"*. Overall, *aespa* has the best streaming performance, following by *IVE*, *LE SSERAFIM* and *BABYMONSTER*, respectively.

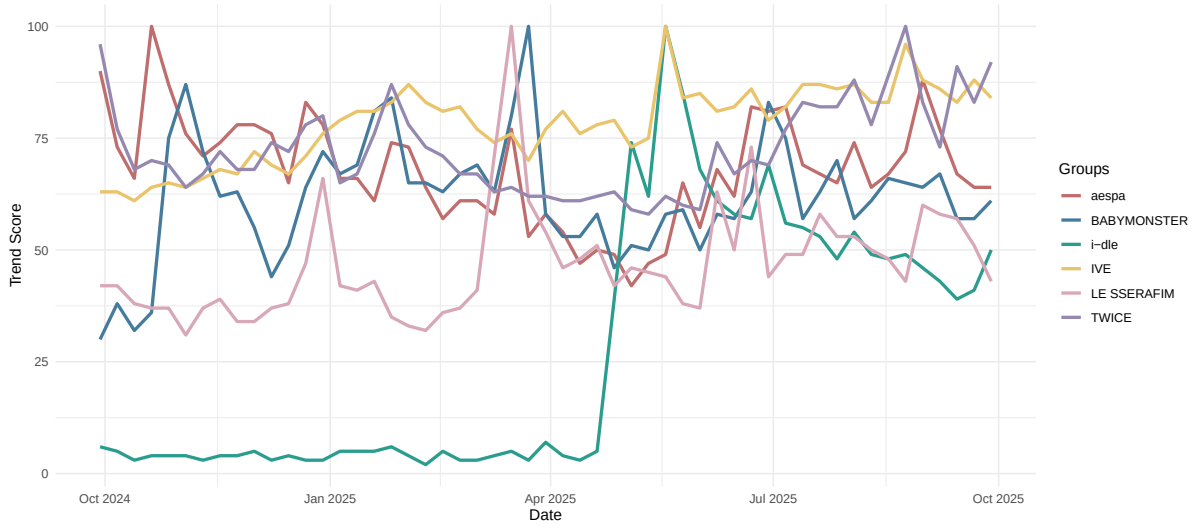

Figure 9: Female Group Google Trend Changes in weekly Google trend scores for each female group over the evaluation period are shown.

From Figure 9, the trend scores of all nominated female groups fluctuated to certain extent. However, *IVE* has the most stable high trend scores, especially from January 2025. *aespa* has the highest average trend scores at the start of evaluation period, from October 2024 to January 2025.

rank	Spotify_followers	popularity_score	albums_popularity	YouTube_subscribers	YouTube_views	album_average_sales
1	TWICE	TWICE	TWICE	TWICE	TWICE	IVE
2	i-dle	LE SSERAFIM	aespa	BABYMONSTER	i-dle	aespa

Table 3: Top 2 Female Group For each variable, rank the female groups in descending order of their values and display the top 2.

From Table 3, *TWICE* ranks the top in all Spotify/YouTube related variables, while *IVE* ranks the top in average album sales. This result is different from the previous observations, where *aespa* and *IVE* have the best performance. But this is expected, as Spotify and YouTube metrics are cumulative, at least to some extent. *TWICE* has been active for 10 years, much longer comparing to other female groups. Therefore, *TWICE* does have the generation advantage in these metrics.

Model

To increase the prediction accuracy, this study implements three models: weighted rank model, principal component analysis, and K-Means clustering. All collected data are normalized to enable evaluation on the same scale.

A weighted rank model is a way of combining multiple rankings or scores into a single ranking, assigning different weights to each component. It's often used in competitions, recommendation systems, or scoring multiple features. This study uses it as the award evaluation system similarly applies it to assign weights.

Principal component analysis is a dimensionality reduction technique which finds new uncorrelated axes (principal components) that capture most of the variance in the data. The first principal component (PC1) captures the largest variance, the second (PC2) captures the next largest variance, and so on. It is applied in this study to address multicollinearity, as the collected data are highly correlated, shown in the correlation heat maps (Figures 2 and 6).

K-means clustering is an unsupervised machine learning algorithm used to partition a dataset into a predetermined number of clusters based on feature similarity. This study employs k-means clustering to uncover patterns in unlabeled data, enabling the categorization of nominees according to their varying winning probabilities.

This study uses normalization, the process of rescaling numerical data to compare different features on the same scale. This is crucial as many algorithms (eg. PCA, k-means) are sensitive to the scale of the features. Both min-max scaling and standard scaling are used in this study.

Min-Max scaling rescales data to a fixed range, usually [0, 1], where

$$x' = \frac{x - x_{min}}{x_{max} - x_{min}}.$$

Standard scaling standardizes features to have mean 0 and standard deviation 1. In other words,

$$x' = \frac{x - \mu}{\sigma}$$

where μ and σ represents the mean and standard deviation, respectively.

Song of the Year

As shown in the Appendix, there are 58 songs nominated for this category. The criteria for 2025 MAMA Song of the Year consist of 40% judges panel evaluation and 60% digital song downloads and streaming (40% from South Korea and 20% global). To predict the winner of this award, this study uses three different methods: weighted rank model, principal component analysis (PCA), and K-Means clustering.

Normalization is applied to all variables (popularity score, mv view count, mv like count, mv comment count) prior to modeling, as data are collected in different scales.

For the weighted rank model, as popularity score is the only variable biased towards recent streaming, we put more weight on it. Moreover, we put more weights on mv like count and mv comment count comparing to mv view count, as YouTube likes and comments directly reflect fans' interaction. In contrast, repeating streaming YouTube music video is widely used by K-pop fans to boost chart rankings, leading to higher YouTube mv views count. Therefore, least weight is applied to mv view count variable. We implement the following weights:

popularity score	mv view count	mv like count	mv comment count
0.45	0.10	0.25	0.20

Then the weighted rank model is

$$\begin{aligned}\text{weighted score} = {} & 0.45 \cdot \text{popularity score} + 0.10 \cdot \text{mv view count} + \\ & 0.25 \cdot \text{mv like count} + 0.20 \cdot \text{mv comment count}\end{aligned}$$

where all variables are normalized by MinMaxScaler.

Top 5 songs by this weighted model are shown in Table 4.

rank	**song_name**	**artist_name**	**weighted_score**
1	APT.	ROSE	0.9293
2	JUMP	BLACKPINK	0.6056
3	Golden	HUNTR/X	0.5907
4	like JENNIE	JENNIE	0.4929
5	Soda Pop	Saja Boys	0.4789

Table 4: Top 5 Songs - Weighted Rank Model Rank the songs in descending order of their weighted scores and display the top 5.

From the weighted scores, *APT.* by ROSE and Bruno Mars ranks the 1st, corresponding to the result in EDA.

In terms of principal component analysis, we only used the first principal component for prediction, as it captures the largest amount of variance.

rank	**song_name**	**artist_name**	**PC1**
1	APT.	ROSE	1.4889
2	JUMP	BLACKPINK	0.6230
3	Golden	HUNTR/X	0.5857
4	like JENNIE	JENNIE	0.3832
5	Soda Pop	Saja Boys	0.3540

Table 5: Top 5 Songs - PCA Rank the songs in descending order of their PC1 values and display the top 5.

From Table 5, with normalized variables, *APT.* by ROSE and Bruno Mars ranks the 1st and has a dominate PC1 value.

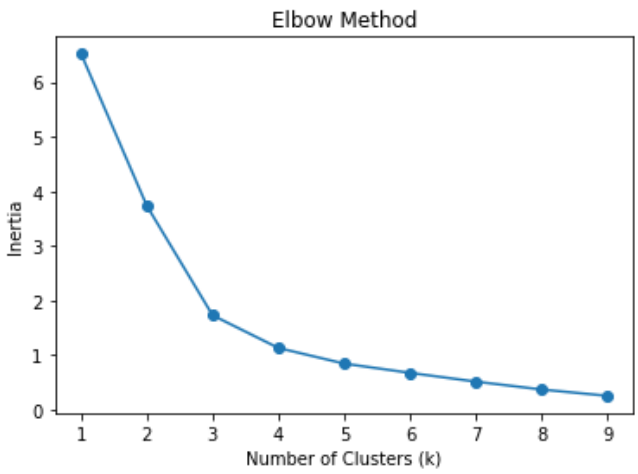

Figure 10: Elbow Plot - Song of the Year For each candidate value of k, run k-means and compute the within-cluster sum of squares (also called inertia) which measures how tightly points are clustered around their cluster centroids. Optimal k occurs where the rate of decrease sharply changes — like a bent elbow.

Based on the Elbow plot in Figure 10, use the optimal $k = 3$ for K-means clustering and obtain three clusters shown in Table 6.

cluster	popularity_score	mv_view_count	mv_like_count	mv_comment_count
0	0.6881	0.0659	0.1046	0.1292
1	0.8428	1.0000	1.0000	1.0000
2	0.3407	0.0066	0.0094	0.0235

Table 6: K-Means Clusters - Song of the Year The shown variable values represent the cluster centroids.

Cluster 1 has the highest value in all popularity metrics. Surprisingly, cluster 1 has only one song, as shown in Table 7.

song_name	artist_name	popularity_score	mv_view_count	mv_like_count	mv_comment_count	cluster
APT.	ROSE	0.8428	1	1	1	1

Table 7: K-Means Cluster 1 - Song of the Year Only one song in cluster 1.

This indicates that *APT.* by ROSE and Bruno Mars ranks the 1st by K-means clustering. This result corresponds to previous findings by weighted rank model and principal component analysis.

Best Male Group

As shown in the Appendix, there are 7 male groups nominated for this category. The criteria for 2025 MAMA Best Male Group are not explicitly published. However, we do expect judge panel evaluation, digital album downloads and streaming, and physical album sales in the evaluation criteria. To predict the winner of this award, this study implements three different methods/models: weighted rank model, principal component analysis (PCA), and K-Means clustering.

Based on the streaming rank plot in Figure 4 of EDA, we added streaming weights to groups with good streaming performance:

Group	BOYNEXTDOOR	RIIZE, SEVENTEEN	Others
Weight	2	1	0

Normalization is applied to all variables including streaming weights prior to modeling, as Spotify data, YouTube data and album sales are all in different scales.

For the weighted rank model, we only select 6 variables (popularity score, albums popularity, YouTube views, albums count, album average sales, stream score) due to high correlations. As stream score and album average sales are most biased towards evaluation period group performance, more weights are assigned. Moreover, we put higher weights on Spotify popularity score and album popularity than YouTube views and album count as Spotify popularity score are biased towards recent streaming, comparing to purely cumulative count in YouTube views. The weights implemented are:

popularity score, albums popularity	YouTube views, albums count	album average sales	stream score
0.15	0.10	0.20	0.30

The weighted rank model is

$$\begin{aligned}\text{weighted score} = {} & 0.15\cdot\text{popularity score} + 0.15\cdot\text{albums popularity} + 0.10\cdot\text{YouTube views} + \\ & 0.10\cdot\text{albums count} + 0.20\cdot\text{album average sales} + 0.30\cdot\text{stream score}\end{aligned}$$

where all variables are normalized by StandardScaler.

Male group ranks by this model are shown in Table 8. Based on the weighted scores, *Stray Kids* ranks the 1st with a dominate leading weighted score 0.8261, following by *SEVENTEEN* in the second.

In terms of principal component analysis, use the same variables as weighted rank model but only choose the first principal component for prediction. With normalized variables, male group ranks by PCA are shown in Table 9.

rank	artist_name	weighted_score
1	Stray Kids	0.8261
2	SEVENTEEN	0.4886
3	TOMORROW X TOGETHER	0.0847
4	BOYNEXTDOOR	-0.1300
5	ENHYPEN	-0.1445
6	RIIZE	-0.5025
7	ZEROBASEONE	-0.6225

Table 8: Male Group Ranks - Weighted Rank Model Rank the male groups in descending order of their weighted scores.

rank	artist_name	PC1
1	Stray Kids	3.4206
2	TOMORROW X TOGETHER	1.2736
3	SEVENTEEN	0.6103
4	ENHYPEN	0.5580
5	ZEROBASEONE	-0.8177
6	RIIZE	-2.2536
7	BOYNEXTDOOR	-2.7911

Table 9: Male Group Ranks - PCA Rank the male groups in descending order of their PC1 values.

From Table 9, *Stray Kids* ranks the 1st again with a leading PC1 value 3.42, following by *TOMORROW X TOGETHER* in the second.

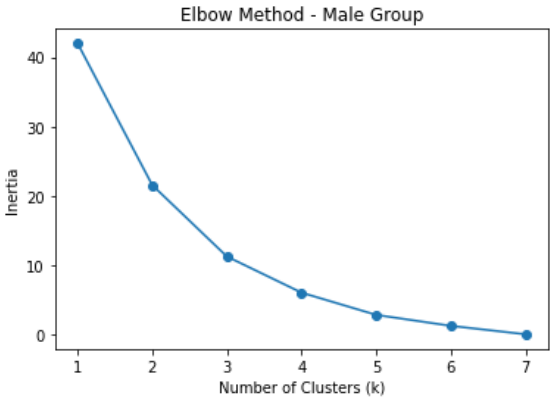

Figure 11: Elbow Plot - Best Male Group For each value of k, run k-means and compute the within-cluster sum of squares (inertia) which measures how tightly points are clustered around their cluster centroids. Optimal k occurs where the rate of decrease sharply changes — like a bent elbow.

Based on the Elbow plot in Figure 11, choose optimal $k = 2$ for K-means clustering. With the same variables as weighted rank model and PCA, obtain the resulted two clusters in

Table 10. Cluster 0 has higher average values across most variables, except for stream_score, where Cluster 1 has an advantage. Choose cluster 0 as it has the highest average values overall.

cluster	popularity_score	albums_popularity	YouTube_views	albums_count	album_average_sales	stream_score
0	0.7777	0.6185	0.6460	0.3608	0.6800	-0.4412
1	-1.0369	-0.8247	-0.8613	-0.4811	-0.9067	0.5883

Table 10: K-Means Clusters - Best Male Group Obtain two clusters and cluster centroids are shown by the variable values.

There are four groups in cluster 0: *SEVENTEEN*, *TOMORROW X TOGETHER*, *ENHYPEN*, and *Stray Kids*. This result matches the previous conclusions from weighted rank model and PCA.

cluster	artist_name
0	SEVENTEEN
0	TOMORROW X TOGETHER
0	ENHYPEN
0	Stray Kids

Table 11: K-Means Cluster 0 - Best Male Group Four male groups in cluster 0.

Best Female Group

As shown in the Appendix, there are 6 female groups nominated for this category. The criteria for 2025 MAMA Best Female Group are not explicitly published, but we do expect weights on judge panel evaluation, digital album downloads and streaming, and physical album sales. To predict the winner of this award, this study uses three different methods/models: weighted rank model, principal component analysis (PCA), and K-Means clustering.

Based on the streaming rank plot in Figure 8 of EDA, we added streaming weights to groups with good streaming performance:

Group	aespa	IVE	BABYMONSTER, LE SSERAFIM	Others
Weight	3	2	1	0

Normalization is applied to all variables including streaming weights prior to modeling, since Spotify data, YouTube data and album sales are all in different scales.

For the weighted rank model, we select 7 variables (Spotify followers, popularity score, albums popularity, YouTube views, albums count, album average sales, stream score) based on correlation heat map in Figure 6 of EDA. As stream score and album average sales reflect the evaluation period group performance the most, more weights are assigned. The weights implemented are:

Spotify followers, popularity score, albums popularity	YouTube views, albums count	album average sales	stream score
0.10	0.10	0.20	0.30

The weighted rank model is

$$\begin{aligned}\text{weighted score} = {} & 0.10\cdot\text{Spotify followers} + 0.10\cdot\text{popularity score} + 0.10\cdot\text{albums popularity} + \\ & 0.10\cdot\text{YouTube views} + 0.10\cdot\text{albums count} + 0.20\cdot\text{album average sales} + 0.30\cdot\text{stream score}\end{aligned}$$

where all variables above are normalized by StandardScaler.

rank	**artist_name**	**weighted_score**
1	aespa	0.8287
2	TWICE	0.5179
3	IVE	0.1697
4	LE SSERAFIM	-0.1766
5	BABYMONSTER	-0.3735
6	i-dle	-0.9661

Table 12: Female Group Ranks - Weighted Rank Model Rank the female groups in descending order of their weighted scores.

Female group ranks by weighted rank model are shown in Table 12. *aespa* ranks the 1st with a dominate leading weighted score 0.8286, following by *TWICE* in the second.

In terms of principal component analysis, use the same variables as weighted rank model but only consider the first principal component for prediction. With normalized variables, female group ranks by PCA are shown in Table 13. *TWICE* ranks the 1st with the only positive PC1 value.

rank	**artist_name**	**PC1**
1	TWICE	3.7456
2	i-dle	-0.1504
3	aespa	-0.3027
4	LE SSERAFIM	-0.4820
5	BABYMONSTER	-0.6716
6	IVE	-2.1388

Table 13: Female Group Ranks - PCA Rank the female groups in descending order of their PC1 values.

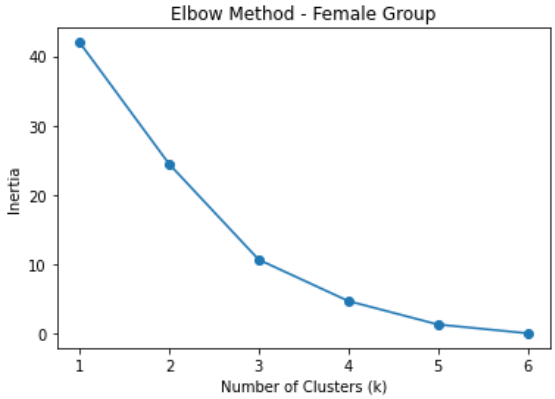

Figure 12: Elbow Plot - Best Female Group For each value of k, run k-means and compute the within-cluster sum of squares (inertia) which measures how tightly points are clustered around their cluster centroids. Optimal k occurs where the rate of decrease sharply changes — like a bent elbow.

The Elbow plot in Figure 12 suggests the optimal $k = 3$ for K-means clustering. With the same variables as weighted rank model and PCA, obtain three K-means clusters in Table 14.

cluster	Spotify_followers	popularity_score	albums_popularity	YouTube_views	albums_count	album_average_sales	stream_score
0	2.1342	1.4792	1.8581	1.5395	0.8944	0.2763	-1.0932
1	-0.4569	-0.3784	-0.0055	-0.9951	0.8944	1.0872	1.2494
2	-0.4068	-0.2408	-0.6230	0.1502	-0.8944	-0.8169	-0.4685

Table 14: K-Means Clusters - Best Female Group Obtain three clusters and cluster centroids are shown by the variable values.

Cluster 0 has the highest overall value, but is the lowest in stream score. Surprisingly, there is only one group in cluster 0.

cluster	artist_name
0	TWICE

Table 15: K-Means Cluster 0 - Best Female Group Only one group in cluster 0.

Cluster 1 has the best values in average album sales and stream score. Since these two variables are the most relevant metrics given the evaluation period time constraint, we prefer the groups in cluster 1 for prediction.

cluster	artist_name
1	aespa
1	IVE

Table 16: K-Means Cluster 1 - Best Female Group Two groups in cluster 1.

In cluster 1, *aespa* and *IVE* show up as expected from previous analysis in EDA.

Results

Based on the models/methods used, we have the following prediction results for the three awards.

Song of the Year

APT. by ROSE and Bruno Mars ranks the 1st by all methods - weighted rank model, principal component analysis, and K-Means clustering. Therefore, our predicted winner of 2025 MAMA Song of the Year is ***APT.* by ROSE and Bruno Mars**.

Best Male Group

By weighted rank model, *Stray Kids* and *SEVENTEEN* rank the 1st and 2nd, respectively. By principal component analysis, *Stray Kids* and *TOMORROW X TOGETHER* rank the 1st and 2nd, respectively. All three groups show up in the best cluster by K-Means clustering. Therefore, our predicted winner of 2025 MAMA Best Male Group is ***Stray Kids***.

Best Female Group

By weighted rank model, *aespa* and *TWICE* rank the 1st and 2nd, respectively. By principal component analysis, *TWICE* ranks the 1st with the only positive PC1 value. By K-Means clustering, *aespa* and *IVE* show up in the cluster favoring average album sales and stream score - the two most relevant metrics given the time constraint in the evaluation period.

Since MAMA has the preference of awarding current generation idols and *TWICE* is a group from the third generation of K-pop (mid-2010s), the probability of *TWICE* receiving this award is strongly affected. Therefore, our predicted winner of 2025 MAMA Best Female Group is ***aespa***.

Discussion

For the predicted results, *APT.* by ROSE and Bruno Mars and *Stray Kids* consistently rank the 1st in their categories respectively by all models and methods used. In comparison, *aespa* only ranks the 1st in weighted rank model and shows up in the preferred K-means cluster, failing to rank high in principal component analysis. However, with relevant knowledge, *aespa* has a strong advantage in group generation that boosts its winning probability.

The true winners of 2025 MAMA awards are shown in Table 17.

Our modeling results successfully predict the winners of 2025 MAMA Song of the Year and Best Female Group. However, *SEVENTEEN* won the Best Male Group instead of *Stray Kids*.

Award	Winner
Song of the Year	APT. by ROSE & Bruno Mars
Best Male Group	SEVENTEEN
Best Female Group	aespa

Table 17: 2025 MAMA Winners True winners of 2025 MAMA Song of the Year, Best Male Group, and Best Female Group.

In our prediction, *SEVENTEEN* ranks the 2nd in the weighted rank model and the 3rd in the principal component analysis, indicating that it is a group with high winning probability. Given the limited data and vague evaluation criteria, it is reasonable that our models and methods failed to predict correctly for the Best Male Group category.

Conclusively, the study successfully predicted 2 out of 3 award winners, suggesting that our models and methods are effective, given the uncertainty in evaluation criteria and limited data. The correct prediction for Best Female Group implies the importance of focusing on most relevant metrics in prediction, as well as the benefits of prior knowledge from Bayesian perspective. The study shows the feasibility of music awards prediction through popularity metrics. However, uncertainty from subjective experts evaluations and limited data still greatly affect the prediction results.

References

1. Python Software Foundation (2022). Python 3.9.13. https://www.python.org/

2. McKinney, W. (2010). pandas: Powerful data analysis toolkit for Python. https://pandas.pydata.org/

3. Raphael, M. (2018). Spotipy: A lightweight Python library for the Spotify Web API. https://spotipy.readthedocs.io/

4. Google LLC (2023). google-api-python-client: Client library for Google's discovery based APIs. https://github.com/googleapis/google-api-python-client

5. Hurst, A. (2023). pytrends: Unofficial Google Trends API. https://pypi.org/project/pytrends/

6. Hunter, J. D. (2007). matplotlib: A 2D graphics environment. https://matplotlib.org/

7. Pedregosa, F., Varoquaux, G., Gramfort, A., Michel, V., Thirion, B., Grisel, O., Blondel, M., Prettenhofer, P., Weiss, R., Dubourg, V., Vanderplas, J., Passos, A., Cournapeau, D., Brucher, M., Perrot, M., Duchesnay, E. (2011). scikit-learn: Machine learning in Python. https://scikit-learn.org/

8. R Core Team (2024). R 4.5.1: A Language and Environment for Statistical Computing. Vienna, Austria: R Foundation for Statistical Computing. https://www.R-project.org/

9. Wickham, H., François, R., Henry, L., Müller, K. (2023). dplyr: A grammar of data manipulation. https://CRAN.R-project.org/package=dplyr

10. Wickham, H. (2016). ggplot2: Elegant graphics for data analysis. https://CRAN.R-project.org/package=ggplot2

11. Wickham, H., Henry, L. (2023). tidyr: Tidy messy data. https://CRAN.R-project.org/package=tidyr

12. Wei, T. (2017). corrplot: Visualization of a correlation matrix. https://CRAN.R-project.org/package=corrplot

13. Zeileis, A., Grothendieck, G. (2005). zoo: S3 infrastructure for regular and irregular time series. https://CRAN.R-project.org/package=zoo

14. Wikipedia contributors. (2025, December 1). MAMA Awards. In Wikipedia https://en.wikipedia.org/wiki/MAMA_Awards

15. Wikipedia contributors. (2025, November 30). MAMA Award for Song of the Year. In Wikipedia https://en.wikipedia.org/wiki/MAMA_Award_for_Song_of_the_Year

16. Wikipedia contributors. (2025, November 28). Circle Digital Chart. In Wikipedia https://en.wikipedia.org/wiki/Circle_Digital_Chart

17. Wikipedia contributors. (2025, December 3). 2025 MAMA Awards. In Wikipedia https://en.wikipedia.org/wiki/2025_MAMA_Awards

18. Wikipedia contributors. (2025, December 3). Circle Chart. In Wikipedia https://en.wikipedia.org/wiki/Circle_Chart

Appendix

Song of the Year Nominees

1. aespa — Dirty Work
2. aespa — Whiplash
3. ALLDAY PROJECT — FAMOUS
4. ALLDAY PROJECT — WICKED
5. BABYMONSTER — DRIP
6. BIG Naughty — MUSIC (feat. Lee Chanhyuk)
7. BLACKPINK — JUMP
8. BOYNEXTDOOR — IF I SAY, I LOVE YOU
9. BOYNEXTDOOR — Never Loved This Way Before
10. CNBLUE — A Sleepless Night
11. CORTIS — GO!
12. Davichi — Stitching
13. DAY6 — Maybe Tomorrow
14. Dayoung — body
15. Doyoung — Memory
16. Dynamic Duo and Gummy — Take Care
17. G-Dragon — TOO BAD (feat. Anderson .Paak)
18. HAON — Skrr (feat. Giselle)
19. HIGHLIGHT — Endless Ending
20. HUNTR/X — Golden
21. ILLIT — Cherish (My Love)
22. IVE — REBEL HEART
23. J-hope — MONA LISA
24. Jennie — like JENNIE
25. Jennie — ZEN
26. Jennie and Doechii — ExtraL
27. Jisoo — earthquake
28. Kai — Wait On Me
29. Karina — Up
30. Key — HUNTER
31. LE SSERAFIM — HOT
32. Lee Chanhyuk — Vivid LaLa Love
33. Lee Mu Jin — Coming of Age Story
34. Mark — 1999
35. Mark — Fraktsiya (feat. Lee Young Ji)
36. MEOVV — DROP TOP
37. Minnie — HER
38. N.Flying — Everlasting
39. NCT DREAM — When I'm With You
40. NCT WISH — poppop
41. Park Hyo Shin — HERO

42. pH-1 — Life Is A Movie (feat. Jung Zi So)
43. PLAVE — Dash
44. QWER — Dear
45. RIIZE — Fly Up
46. Rose — toxic till the end
47. Rose and Bruno Mars — APT
48. Roy Kim — If You Ask Me What Love Is
49. Saja Boys — Soda Pop
50. SEVENTEEN — THUNDER
51. Tablo and RM — Stop The Rain
52. Taeyeon — Letter To Myself
53. TREASURE — YELLOW
54. TWS — Countdown!
55. TXT — When the Day Comes
56. V — Winter Ahead (with Park Hyo Shin)
57. Xdinary Heroes — Beautiful Life
58. ZEROBASEONE — Doctor! Doctor!

Best Male Group Nominees

1. SEVENTEEN
2. TOMORROW X TOGETHER
3. BOYNEXTDOOR
4. ENHYPEN
5. RIIZE
6. Stray Kids
7. ZEROBASEONE

Best Female Group Nominees

1. aespa
2. BABYMONSTER
3. i-dle
4. IVE
5. LE SSERAFIM
6. TWICE