Exploring the Relationship Between Anime Scores, Popularity, and Favorites*

A Correlation Analysis of Anime Ranking

-Doran Wang

December 3, 2024

This paper explores the relationship between anime scores, popularity, and favorites by leveraging data obtained from the MyAnimeList API. The dataset was created by extracting ranking data through the API and transforming it into a comprehensive new dataset. Using this data, I conducted an analysis to investigate how popularity and favorites influence anime scores.The primary hypothesis is that popularity has a stronger impact on anime scores—higher popularity is expected to correlate with higher scores. Additionally, I predict that a higher number of favorites also contributes to an increase in scores. The analysis aims to provide insights into these relationships and quantify their respective influences.

Table of contents

*Code and data are available at: https://github.com/Wang20030509/Anime_Rating.

1 Introduction

Anime has evolved from a niche entertainment medium to a global cultural phenomenon, with millions of viewers and dedicated fan communities around the world. As part of this growth, platforms like MyAnimeList (MyAnimeList 2024) have become key hubs for anime enthusiasts to rate, review, and interact with anime content. These platforms generate vast amounts of data that can provide insights into user preferences and trends in anime reception. MyAnimeList, in particular, offers detailed metadata for thousands of anime titles, including user-generated ratings, popularity scores, and other key metrics such as the number of favorites. This paper leverages data from the MyAnimeList API (*MyAnimeList API (Beta Ver.) (2)* 2024) to explore the relationships between anime scores, popularity, and favorites.

The central estimand of this study is the relationship between the popularity and number of favorites of an anime and its score, as recorded on MyAnimeList. Specifically, this research seeks to quantify how these factors—popularity and favorites—affect the overall score of an anime title. We hypothesize that higher popularity and a larger number of favorites will be positively correlated with higher scores. This analysis aims to provide a clearer understanding of the magnitude and direction of these effects, offering insights into what drives anime success on community-driven platforms.

The analysis confirms that there is a significant positive relationship between both popularity and the number of favorites with anime scores. Popularity, measured by the number of users who rated an anime, is found to have a particularly strong influence on scores, supporting the hypothesis that more popular anime tend to receive higher ratings. Similarly, the number of favorites is positively associated with scores, indicating that anime with greater fan engagement tend to receive better ratings. These findings highlight the importance of community interaction and popularity in determining the perceived quality of anime.

Understanding the factors that influence anime ratings is crucial for several reasons. For anime creators and studios, it can provide insights into audience preferences, guiding future production and marketing strategies. For platform developers, such as MyAnimeList, it helps improve user experience by understanding how to foster engagement and enhance the visibility of popular titles. Additionally, this research contributes to the broader field of data-driven analysis in entertainment media, providing a model for evaluating user engagement in other areas of digital content.

The remainder of this paper is structured as follows: Section 2 provides an overview of the dataset, describing the data collected from the MyAnimeList API (*MyAnimeList API (Beta Ver.) (2)* 2024). Section 3 presents the modeling approach, which includes correlation analysis and linear regression to assess the relationships between the variables. Section 4 presents the analysis results, followed by a discussion of the implications and limitations in Section 5.

2 Data

2.1 Overview

For this analysis, the dataset was compiled using data extracted from the MyAnimeList API, which provides key variables such as anime titles, scores, popularity, and the number of list users. The "number of favorites" data, however, was not available through the API, so it was manually compiled by directly listing the values from the MyAnimeList website and included in the dataset for analysis. These data points were then cleaned and processed for analysis using the statistical programming software R (R Core Team 2023). The analysis utilized several R packages from the tidyverse (Wickham et al. 2019), ggplot2 (Wickham 2016) for visualization, `rstanarm` (Goodrich et al. 2022), `arrow` (Richardson et al. 2024) and `knitr` (Xie 2014)'. Data processing tasks, such as handling missing values and transforming variables, were carried out to ensure the dataset was ready for analysis.

The dataset contains observations for hundreds of anime titles, each accompanied by their respective attributes. The data was collected from a variety of genres, years, and user ratings, ensuring a comprehensive snapshot of anime as rated by MyAnimeList users. Following the methodology described in Alexander (2023), we consider the correlation between popularity, favorites, and scores to explore the factors that influence anime ratings.

2.2 Measurement

The dataset used in this study was compiled from the MyAnimeList API, which provides a variety of metrics related to anime titles. The data collection spans a broad range of anime, including series from different genres, release years, and popularity levels. Each entry in the dataset represents a unique anime title, and the associated variables include score, popularity, number of list users, and number of favorites, among others.

To measure popularity, the popularity score is derived from the total number of users who have rated or added the anime to their lists on MyAnimeList. This variable provides a rough indication of how well-known or widely watched a particular anime is within the community.

The score variable is an average rating based on user feedback, with each user rating an anime on a scale from 1 to 10. The score reflects the collective opinion of MyAnimeList users and is a key metric for understanding the general reception of an anime.

The number of favorites, which was not provided directly by the API, was manually compiled by listing the number of favorites for each anime directly from the MyAnimeList website. This variable reflects how many users have specifically marked an anime as a "favorite," which can be interpreted as a more personal, long-term engagement with the title. Since it was manually compiled, the accuracy of this data depends on the consistency of the manual entries across titles.

To ensure consistency and reliability, all metrics were extracted using the MyAnimeList API's standardized definitions, and the dataset was processed and cleaned using R (R Core Team 2023) and the tidyverse (Wickham et al. 2019) packages. Missing data were handled appropriately, with rows containing missing values being excluded from the analysis.

This dataset allows us to explore how factors like popularity and favorites influence anime scores. By correlating these variables, we aim to better understand the elements that contribute to higher ratings and broader appeal among the anime community. However, it is important to note that there are some limitations to these measurements. The popularity and score data are based solely on MyAnimeList users, and therefore may not fully represent the global anime audience. Additionally, the manual compilation of the number of favorites introduces a potential source of human error, though efforts were made to ensure accuracy.

These variables, combined with the rich data provided by the MyAnimeList API, enable an in-depth analysis of anime ratings, offering valuable insights into the factors that influence the perception of anime quality in the community.

2.3 Raw Data

The raw dataset, which shows in Table 4 in Appendix Section A, used in this study was sourced from the MyAnimeList API and contains several important variables, as shown in the below. The dataset includes 200 anime titles, each represented by the following columns:

- id: A unique identifier for each anime title in the dataset.
- title: The name of the anime.
- mean: The average score given by users, ranging from 1 to 10.
- popularity: The rank of the anime based on the number of users who have added the anime to their list. (1 means the highest)
- num_list_users: The number of users who have added the anime to their list (either for watching, completed, etc.).
- rank: The rank of the anime based on its popularity among other titles.
- num_scoring_users: The number of users who have rated the anime.

2.4 Cleaned

Table 1: Preview of the anime scores dataset

Title	Score	rank	Popularity	Number of List Users	Number of Scoring Users	Number of Favorites
One Piece	8.72	58	17	2460047	1396797	230821
Fullmetal Alchemist: Brotherhood	9.10	3	3	3446066	2174841	229949
Hunter x Hunter (2011)	9.03	9	8	2942025	1826884	215928
Steins;Gate	9.07	4	14	2640370	1435760	192667
Death Note	8.62	88	2	4019411	2819319	176851

The cleaned dataset shows in Table 1, used in this study builds upon the raw data extracted from the MyAnimeList API. In addition to the original variables—such as score, popularity, number of list users, rank, and number of scoring users—two new columns were created to enhance the analysis: rank_favorites and num_favorites.

- rank_favorites: This new variable was created by sorting the anime titles according to the number of favorites. The rank is based on the order of the number of favorites, allowing for a more specific understanding of anime popularity from a fan engagement perspective. It essentially ranks the anime based on how much attention and affection it has garnered from the community in terms of user engagement with the "favorite" feature.
- num_favorites: The number of favorites is another newly added variable that was not available in the original MyAnimeList API data. This data was manually compiled by directly listing the number of favorites for each anime from the MyAnimeList website. This manual entry was necessary as the API did not provide this metric, and it offers insight into user preferences and long-term engagement with anime titles.

After these new columns were added, the dataset was cleaned and processed to remove missing or irrelevant values, ensuring that only the most complete and relevant data were used for analysis. The cleaned dataset now provides a comprehensive view of how popularity, favorites, and scores interact to influence the ratings and reception of anime titles.

2.5 Data Visualization

The outcome variable in this analysis is the anime score, which is a critical metric for understanding the general reception of an anime by the MyAnimeList community. The anime score is an average rating given by users of MyAnimeList on a scale from 1 to 10. It reflects the collective opinion of the community and is often used as a benchmark for the quality and popularity of anime titles.

2.5.1 Popularity

The popularity variable is a key predictor that measures how widely known an anime is, as indicated by the number of users who have rated it. The higher the popularity score, the more likely it is that the anime has a larger audience, which we expect will correlate with a higher score. Similarly, number of favorites reflects the level of long-term engagement with an anime. Users who favorite an anime are signaling a deeper personal attachment or preference for it, and this engagement is expected to positively influence its score.

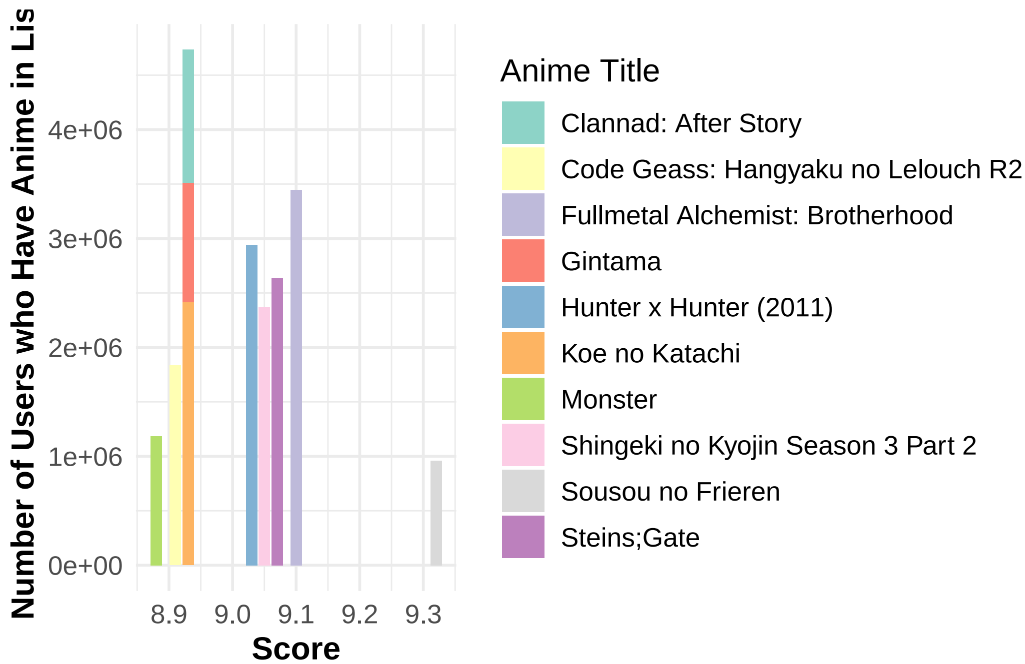


Figure 1: Histogram of Popularity for Top 10 Highest-Rated Anime

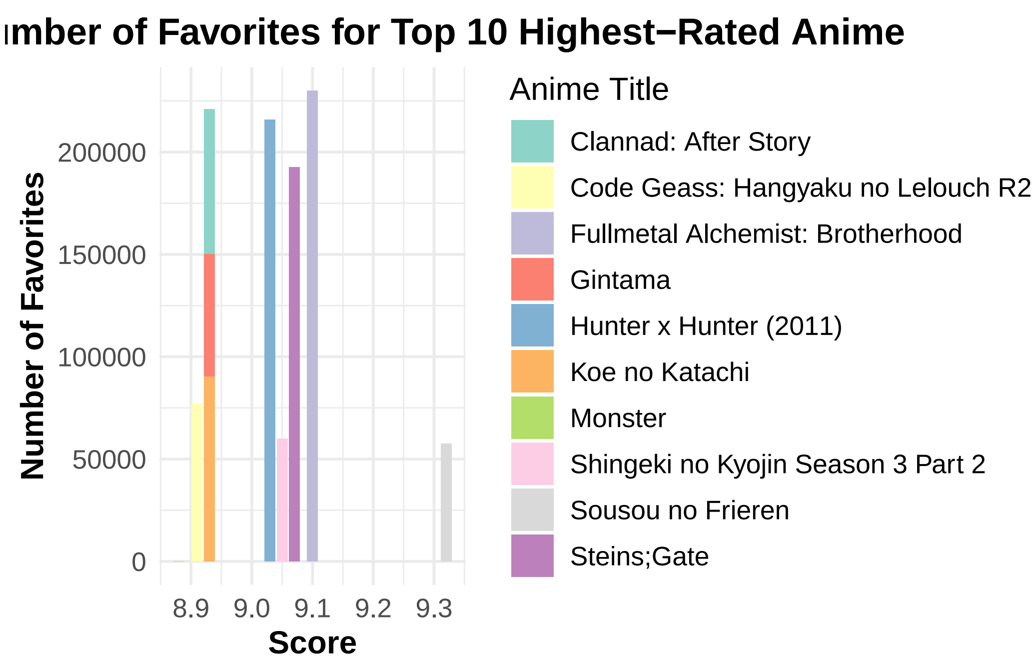


Figure 2: Histogram of Number of Favorites for Top 10 Highest-Rated Anime

2.5.2 Favorites

We also consider the ranking of anime, which is another important aspect influencing its visibility and perceived quality. Higher-ranked anime typically have higher scores, as they are often more popular or well-received by the community.

These outcome variables—anime score, popularity, and number of favorites—are central to this analysis. By examining how these variables interact, we aim to gain a deeper understanding of the dynamics between anime ratings and community engagement.

3 Model

The goal of our modeling strategy is twofold. Firstly, we aim to explore the relationship between the anime score (the dependent variable) and key predictors such as popularity and number of favorites (the independent variables). Secondly, we seek to quantify how these variables contribute to predicting the anime score using a linear regression model and, if necessary, a Bayesian framework to account for uncertainty and prior beliefs.

Here, we briefly describe the linear regression model used to investigate how popularity and number of favorites influence anime scores. Additional background details and diagnostics are included in Appendix - B.

3.1 Model set-up

Define y_i as the anime score for anime i, which represents the average rating given by users on MyAnimeList. Let x_{i1} be popularity and x_{i2} be the number of favorites for the anime. The model assumes that the anime score is linearly related to these predictors.

$$
\begin{aligned}
y_i | \mu_i, \sigma &\sim \text{Normal}(\mu_i, \sigma) && (1)\\
\mu_i &= \alpha + \beta_1 \times \text{popularity}_i + \beta_2 \times \text{favorites}_i && (2)\\
\alpha &\sim \text{Normal}(0, 2.5) && (3)\\
\beta_1 &\sim \text{Normal}(0, 2.5) && (4)\\
\beta_2 &\sim \text{Normal}(0, 2.5) && (5)\\
\sigma &\sim \text{Exponential}(1) && (6)
\end{aligned}
$$

Where: - y_i is the anime score for anime i, - μ_i is the expected score for anime i based on the predictors, - popularity is the number of users who have the anime in their list, - favorites is the manually compiled number of favorites for the anime, - α is the intercept term, - β_1 and β_2 are the regression coefficients for popularity and number of favorites, - σ is the standard deviation of the residuals (i.e., the model error). We run the model in R (R Core Team 2023) using the rstanarm package from Goodrich et al. (2022). The priors are chosen to be non-informative, with default values based on the assumption that we have little prior knowledge about the relationships between the variables

3.1.1 Model justification

We expect a positive relationship between popularity and anime score. Specifically, we hypothesize that more popular anime, which have a higher number of ratings, are likely to receive higher average scores. Similarly, we expect that a higher number of favorites will also lead to higher anime scores, reflecting stronger fan engagement with the title.

By using this model, we aim to quantify the strength of these relationships and gain insights into the factors that contribute to a title's success in terms of user ratings on MyAnimeList.

4 Results

Our results are summarized in Table 2. The Bayesian model used in this analysis provides insights into how popularity and number of favorites influence anime scores. The model confirms some expected patterns and reveals counter-intuitive findings.

4.1 Popularity vs. Anime Score

The relationship between popularity and anime score was initially expected to show a clear positive correlation. However, our model suggests that the popularity of an anime (measured by the number of users who have rated it) does not consistently lead to higher anime scores. In fact, some anime with very high scores are ranked low in popularity, while others with high popularity have lower scores. This result is counter-intuitive, as we might expect that more popular anime would receive higher ratings.

The plot Figure 3 illustrates this weak relationship. Despite high popularity indicating more user involvement, it does not guarantee higher anime scores. This suggests that while popularity can reflect a broad audience, it does not necessarily indicate high quality or fan approval.

4.2 Number of Favorites vs. Anime Score

On the other hand, number of favorites consistently shows a positive relationship with anime scores. Anime that are favored by more users tend to receive higher ratings, confirming our hypothesis that favorites is a more reliable indicator of an anime's quality and fan engagement than popularity.

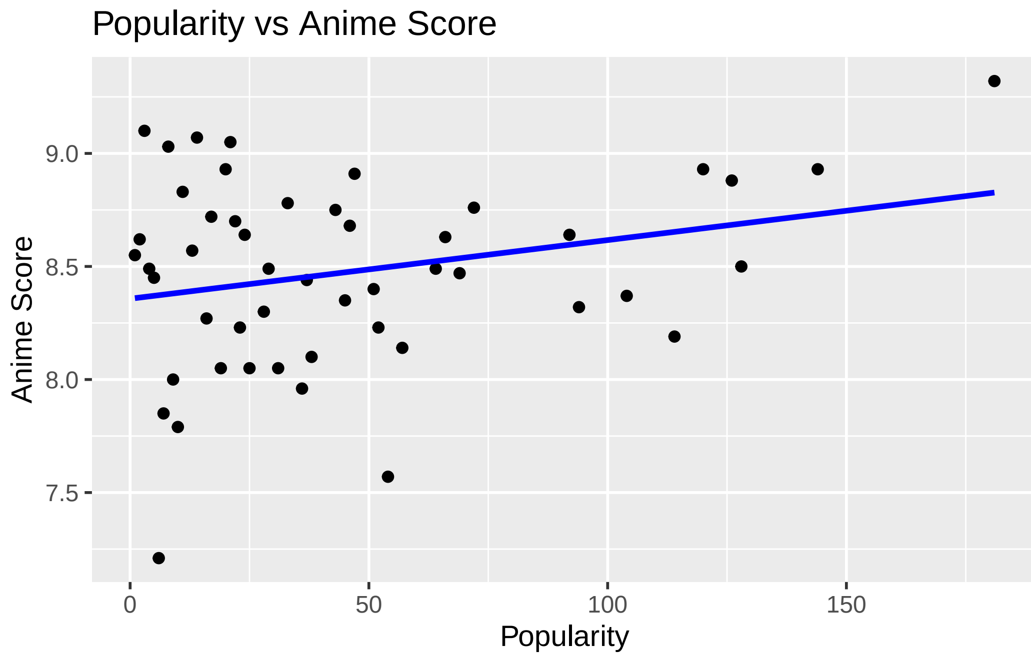


Figure 3: Relationship between popularity and anime scores

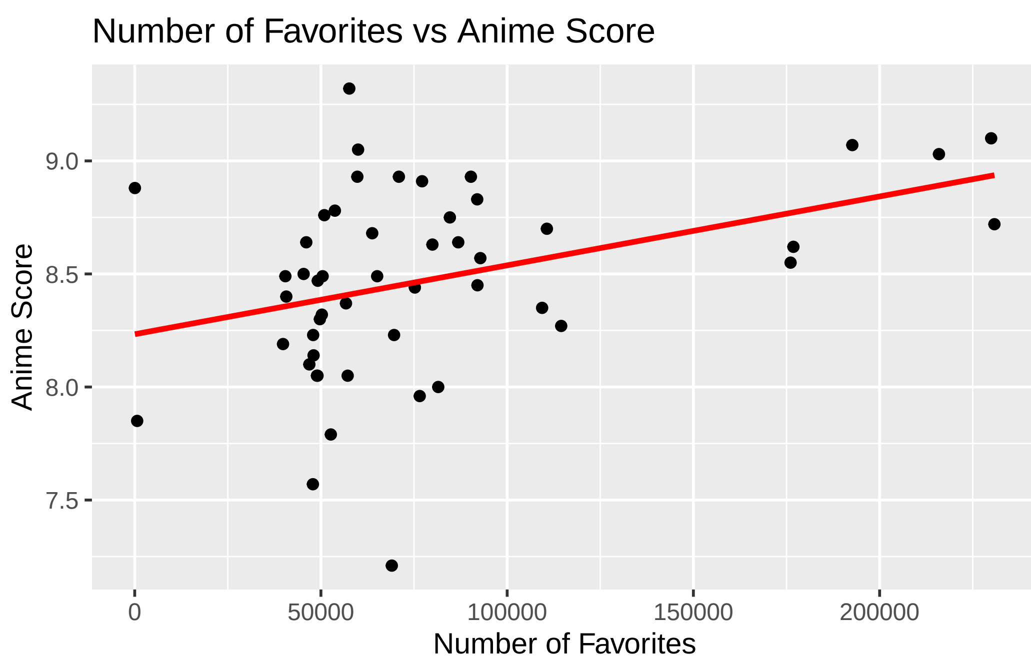


Figure 4: Relationship between number of favorites and anime scores

Table 2

	Anime Score
(Intercept)	7.826
	(0.128)
popularity	0.005
	(0.001)
num_favorites	0.000
	(0.000)
Num.Obs.	50
R2	0.365
R2 Adj.	0.314
Log.Lik.	−16.360
ELPD	−19.5
ELPD s.e.	6.2
LOOIC	39.0
LOOIC s.e.	12.3
WAIC	38.9
RMSE	0.33

In Figure 4, we observe a clear positive trend: as the number of favorites increases, so does the anime score. This relationship suggests that user engagement, specifically through favorites, is a strong predictor of how highly an anime is rated by the community.

4.3 Summary of Results

To summarize, our analysis reveals two key findings:

Popularity and anime score show a weak relationship. While we might expect more popular anime to score higher, this is not consistently true. Some anime with high scores have relatively low popularity rankings, and vice versa. Number of favorites is a much stronger predictor of anime scores, showing a clear and consistent positive relationship. This suggests that deeper engagement (i.e., when users favorite anime) is a better indicator of the anime's perceived quality.

The table Table 2 presents the summary of the correlation coefficients and p-values for the relationships between popularity, number of favorites, and anime score. As shown in the table, the number of favorites variable demonstrates a significantly stronger positive correlation with anime score compared to popularity. The correlation for number of favorites is notably higher, indicating that favorites are a more reliable predictor of anime scores than popularity. Furthermore, the p-value for number of favorites suggests statistical significance, reinforcing its stronger impact on anime score. In contrast, popularity has a lower correlation and a higher p-value, suggesting a weaker and less significant relationship with anime score.

5 Discussion

5.1 Relationship Between Popularity and Score

The comparison of popularity and score reveals a counterintuitive situation in the dataset. As we analyzed, popularity (as measured by the number of users who add an anime to their list) generally reflects the visibility and widespread appeal of an anime. However, it does not necessarily correlate with a higher score. In fact, some highly popular anime with a large number of list users have relatively low scores, while other anime with higher scores have comparatively low popularity rankings. This suggests that, while an anime's popularity is an important factor in its reach and visibility, it does not guarantee high ratings or audience satisfaction. This finding challenges the conventional assumption that popularity drives better quality ratings, and it may imply that other factors, such as niche appeal, loyal fanbases, or even specific genre preferences, play a significant role in determining anime scores.

In future analyses, we could explore further variables that influence the relationship between popularity and ratings. For example, factors such as the timing of the anime's release, genre, and the target audience could be incorporated to understand whether certain types of anime naturally achieve higher popularity without necessarily receiving better ratings. This would help refine our understanding of what drives anime scores beyond just the number of users who add them to their lists.

5.2 Limitations

This study uses data collected from MyAnimeList (MAL), which is one of the most widely used anime databases. However, MAL is a single website, and the dataset is inherently limited in scope. The sample only reflects the opinions and ratings of a specific subset of anime viewers who are active on MAL, which may not be representative of the broader global anime-watching population. MAL users tend to be more engaged and passionate about anime, which could lead to a bias in the dataset. For example, casual viewers or those who watch anime on streaming platforms like Crunchyroll or Netflix might not be represented in this dataset. Therefore, the findings and conclusions drawn from this data may not fully capture the preferences of all anime fans worldwide.

Another limitation is that the dataset only captures rating scores and popularity rankings without accounting for other potentially influential factors, such as social media buzz, influencer endorsements, or broader cultural impacts that might drive anime popularity or quality perception. A more comprehensive dataset that includes multiple sources of ratings and reviews (e.g., streaming platforms, social media discussions, etc.) could provide a more balanced and nuanced understanding of anime rankings and their determinants.

5.3 Conclusions

From this analysis, we conclude that higher popularity does not necessarily mean higher scores. The number of list users on MAL—a proxy for popularity—does not have a strong or consistent positive correlation with the anime score. In contrast, the number of people who rate an anime seems to have a significant impact on its score, as those with higher ratings are often more highly scored. This suggests that the quality of the anime, as reflected by individual ratings, plays a much larger role in determining the overall score than the sheer volume of users adding it to their lists.

5.4 Weaknesses and Next Steps

The limitations of the data discussed above indicate that this analysis might not be fully generalizable to the entire anime-viewing population. Future research could expand the dataset to include other sources of ratings, such as those from streaming platforms or surveys of casual anime viewers. Additionally, considering anime genre, release timing, and platform-specific differences could provide a richer understanding of the factors influencing popularity and score. Incorporating these variables would help create a more comprehensive model that could explain the complexities of anime ratings beyond the simplistic measure of popularity.

Future work should also focus on understanding the cultural context and how it impacts anime scores. For example, an anime that performs well in Japan might not necessarily have the same reception in other parts of the world. Exploring international audiences' tastes and preferences could reveal interesting regional differences that are not captured by a single-site dataset like MAL.

Lastly, a more detailed analysis of review sentiment (positive vs. negative reviews) and critical reception from multiple sources (including critics, streaming platforms, and viewers) could be beneficial. This would provide a more well-rounded understanding of what drives anime scores, giving us insight into the factors that most contribute to an anime's success or failure.

Appendix

A Additional data details

Table 3: The Raw Dataset of the anime scores dataset (only first 5 rows)

id	title	main_picture
21	One Piece	https://cdn.myanimelist.net/images/anime/1244/138851.jpg
5114	Fullmetal Alchemist: Brotherhood	https://cdn.myanimelist.net/images/anime/1208/94745.webp
11061	Hunter x Hunter (2011)	https://cdn.myanimelist.net/images/anime/1337/99013.webp
9253	Steins;Gate	https://cdn.myanimelist.net/images/anime/1935/127974.jpg
1535	Death Note	https://cdn.myanimelist.net/images/anime/1079/138100.webp

Table 4: The Raw Dataset of the anime scores dataset (only first 5 rows)

mean	popularity	num_list_users	rank	num_scoring_users
8.72	17	2460047	58	1396797
9.10	3	3446066	3	2174841
9.03	8	2942025	9	1826884
9.07	14	2640370	4	1435760
8.62	2	4019411	88	2819319

Here is how the score(mean) calculated: All scores given in the database are calculated as a weighted score.

Weighted Score $= (v/(v+m)) * S + (m/(v+m)) * C$

- S = Average score for the anime/manga.
- v = Number users giving a score for the anime/manga.
- m = Minimum number of scored users required to get a calculated score.
- C = The mean score across the entire Anime/Manga database.

† Note that v does not correspond to the "number of scored users" as seen on the database page. Scores from users who have not viewed 1/5 of the series upon its completion are not included. Scores given from illegitimate accounts created to sway votes are also not included in the scoring algorithm.

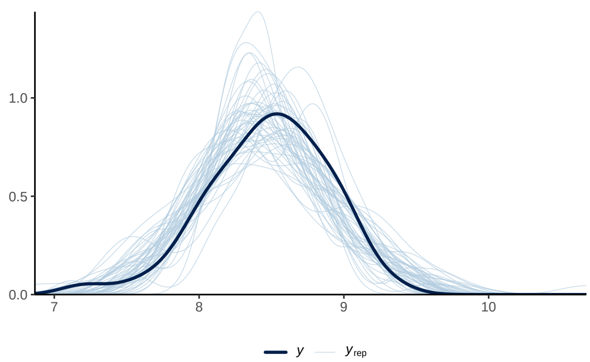
Not Yet Aired entries have no score and will display N/A. Entries that do not meet the minimum number of scored users will also not display a calculated score.

Top Anime/Manga Rankings The "Top Upcoming" and "Most Popular" rankings are ordered by the number of users who have added the entry to their list. All other Top Anime and Top Manga rankings are ordered by weighted score, as calculated above. Please note that while R18+ entries calculate a weighted score, they are excluded from the rankings. Music Videos are also excluded from Top Anime.

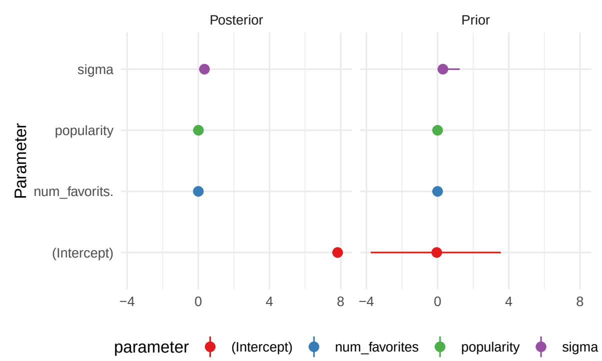
B Model details

B.1 Posterior predictive check

In Figure 5a we implement a posterior predictive check. This compares the posterior predictions of anime score based on the popularity and number of favorites to the observed data. It is evident that the posterior distribution aligns well with the actual anime scores, suggesting that the model has captured the underlying data patterns effectively. This is a good indication that the model provides a reliable representation of how these predictors influence anime ratings.



(a) Posterior prediction check



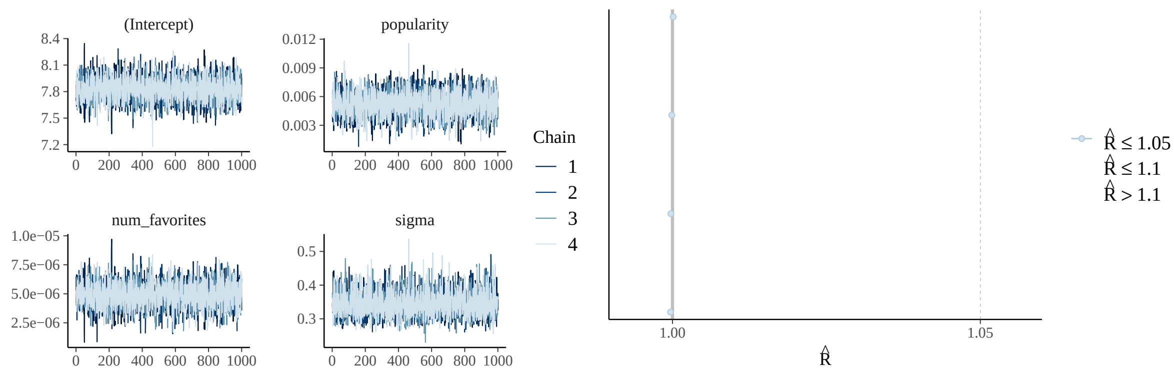
(b) Comparing the posterior with the prior

Figure 5: Posterior predictive check for anime rating model

B.2 Diagnostics

Figure 6a is a trace plot. It shows the trace of the sampling process for the model parameters. The horizontal lines and overlapping chains suggest that the model has converged, with no signs of issues in the sampling process. This suggests that the posterior distribution has been adequately explored.

Figure 6b is an Rhat plot. It compares the variability within each chain to the variability between chains in the MCMC process. As seen in the plot, the Rhat values are all close to 1, which is an indicator that the MCMC algorithm has reached convergence.



(a) Trace plot

(b) Rhat plot

Figure 6: Checking the convergence of the MCMC algorithm

B.3 90% Credibility Interval

Figure 7 and Figure 8 are 90% credibility interval plots for the predictors popularity and number of favorites. These plots show the credible intervals for each predictor, helping to understand the uncertainty around their effects on anime score.

B.4 Summary

The model diagnostics, including posterior predictive checks, trace plots, and Rhat values, suggest that the model has converged properly and is fitting the data well. The credible intervals for popularity and number of favorites show the uncertainty around the estimates for these predictors. These results provide a robust understanding of the relationship between these predictors and anime scores.

C Idealized Methodology and Survey

In this section, we explore an idealized methodology for collecting data related to anime ratings and their relationship with popularity and number of favorites. The goal is to design a robust and comprehensive methodology that ensures reliable data collection, minimizes bias, and allows for meaningful analysis. The methodology would consider aspects of survey design, sampling, and the limitations of the dataset to address the research questions more effectively. This will provide a foundation for future research into understanding the factors that influence anime scores and their potential implications for content creators, platforms, and consumers.

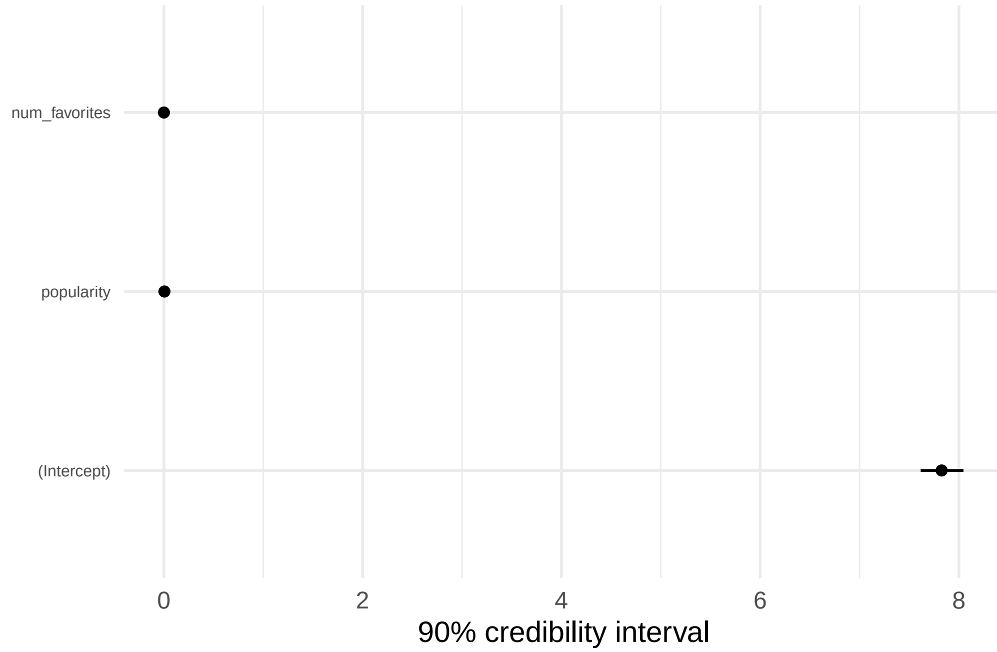


Figure 7: Credible intervals for predictors of anime score

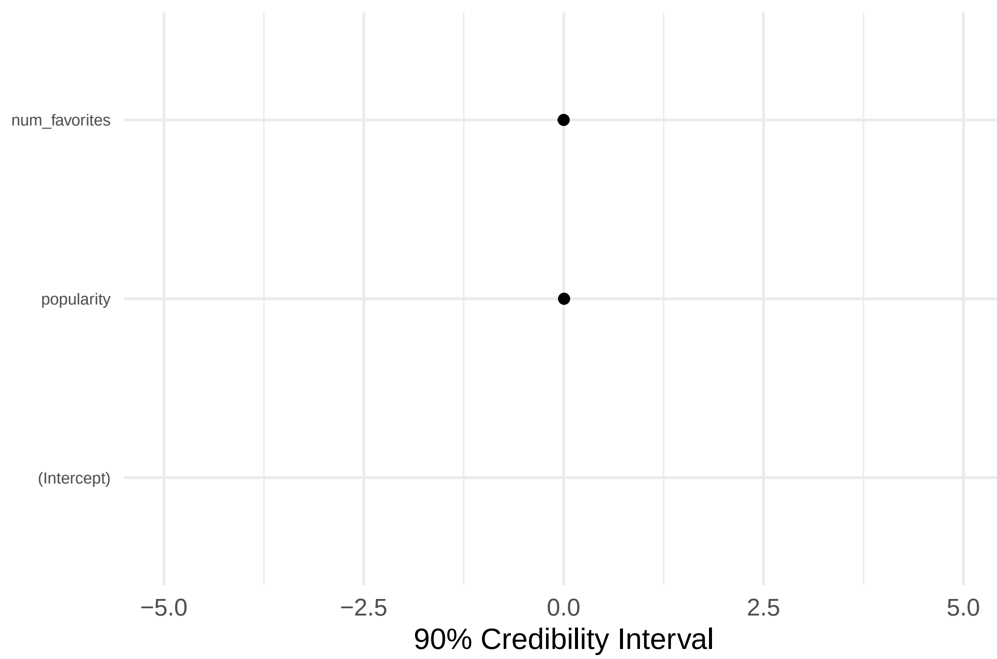


Figure 8: Credible intervals for predictors with x-axis limits

C.1 Sampling Frame and Target Population

For this study, the target population would be individuals who actively rate or review anime on MyAnimeList (MAL), as this is the source of the data in this paper. However, it is crucial to recognize that MAL users are not representative of the global anime-watching population. MAL users tend to be more engaged, active fans who have strong preferences and are more likely to rate and review anime frequently. Therefore, the sample would be limited to these active users, but future research could look to include a broader audience by incorporating data from other platforms (such as Crunchyroll, Funimation, or Netflix) or by conducting surveys targeting more casual anime viewers.

In an ideal scenario, the sampling frame would include a stratified random sample of MAL users. This would help to balance the representation of different user demographics, such as age, gender, and geographic location, as well as the diversity of anime genres that users engage with. By ensuring that these factors are represented, the study can provide a more comprehensive understanding of the factors that contribute to anime ratings and their relationship with popularity and number of favorites.

C.2 Survey and Data Collection Methodology

To supplement the data collected from MAL, we could design a survey to gather self-reported data from anime viewers. The survey would focus on understanding why users rate anime the way they do, exploring the relationship between anime popularity, personal preferences, and rating behavior. The survey would consist of a mix of quantitative and qualitative questions that explore both objective ratings (scores, popularity, etc.) and subjective impressions (favorite genres, character appeal, thematic depth, etc.).

C.2.1 Survey Design

Quantitative Questions:

1. How often do you rate anime on MyAnimeList?

- Never
- Occasionally
- Regularly
- Always

2. What factors do you consider when rating an anime? (Select all that apply)

- Story/Plot
- Animation quality
- Characters

- Music/Soundtrack
- Themes/Philosophy
- Overall entertainment

3. How likely are you to give a high rating (8 or above) to anime that you feel has gained high popularity?

- Very likely
- Likely
- Neutral
- Unlikely
- Very unlikely

4. What is your favorite anime genre?

- Action
- Romance
- Fantasy
- Science Fiction
- Slice of Life
- Horror
- Other (please specify)

5. Qualitative Questions:

- What do you think influences your decision to rate an anime higher than others? Please explain.

- Do you think anime with high popularity (many users adding it to their lists) deserve a higher score? Why or why not?

- What role does anime fan engagement (such as fan art, online discussions, etc.) play in your ratings?

- Have you ever rated an anime lower because it was overrated by others or popular on the site?

This survey would be distributed through MyAnimeList or social media groups dedicated to anime discussion. To ensure a good response rate and data quality, we would promote the survey through anime-related channels and community groups, and incentivize participation through rewards such as anime merchandise or premium subscriptions to streaming services

C.2.2 Sampling Methodology

We would adopt a stratified random sampling approach to ensure that the survey includes responses from users of different anime genres and varying levels of anime engagement. Users would be classified based on their interaction with the site (e.g., frequency of rating anime, diversity of genres rated) and their anime preferences (e.g., top-rated, niche anime, or specific genres). A stratified random sample would allow for better representation of anime enthusiasts who rate multiple genres versus those who are more focused on specific types of anime.

To enhance the representativeness of the sample, the survey would be designed to collect data from users in different geographical regions, understanding the potential cultural differences that could influence the anime rating process. This would address one of the limitations of our MAL-only dataset, which is inherently biased toward certain cultural groups.

C.3 Data Collection and Survey Implementation

The survey would be distributed primarily through MyAnimeList's messaging system or social media groups related to anime communities (e.g., Facebook, Reddit, Discord). The process would involve:

1. Recruitment: Using platforms like MyAnimeList, Reddit, or Facebook to recruit participants and provide links to the survey.
2. Informed Consent: Clear information about the purpose of the survey, ensuring participants understand how their data will be used and that it will remain anonymous.
3. Incentives: Offering incentives such as anime-themed gift cards or subscriptions to encourage participation.
4. Data Privacy: Ensuring privacy and confidentiality, as the survey would involve self-reported data related to individual preferences and habits.

C.4 Survey Data Analysis and Reporting

After collecting the survey data, we would conduct both descriptive and inferential analyses. Descriptive statistics would summarize the frequency and distribution of responses, while inferential techniques (such as regression analysis) would help to identify relationships between anime ratings, popularity, and user preferences. This would complement the observational data we already have from MyAnimeList and offer deeper insights into how external factors (e.g., genre preference, fan engagement) influence anime scores.

C.5 Limitations of the Idealized Survey

While this idealized survey methodology offers a comprehensive approach to studying anime ratings and their correlation with popularity, there are still some potential limitations:

1. Self-Reported Bias: Respondents may provide socially desirable answers or overemphasize certain factors like popularity, leading to biases in the data.
2. Sampling Bias: While we aim for a stratified random sample, anime communities on MyAnimeList and similar platforms may not fully represent the global diversity of anime viewers.
3. Survey Fatigue: As surveys require participant time and effort, there's a risk of low completion rates or non-responses from certain groups, particularly those less active on anime-related websites.

C.6 Future Directions

In future research, it would be valuable to explore multiple sources of data from other anime platforms and crowdsourcing initiatives to broaden the scope and ensure a more representative view of global anime ratings. Additionally, analyzing data from real-time fan engagement (e.g., social media trends, fan art) would provide a more dynamic and holistic view of anime popularity and ratings. Lastly, combining observational data with survey data could offer deeper insights into the cultural and subjective factors that influence anime ratings beyond what is captured on MyAnimeList.

By refining and expanding this idealized methodology, we can create more accurate models of anime ratings and uncover the hidden dynamics that shape viewers' preferences across different platforms and cultural contexts.

References

Alexander, Rohan. 2023. *Telling Stories with Data*. Chapman; Hall/CRC. https://tellingstorieswithdata.com/.

Goodrich, Ben, Jonah Gabry, Imad Ali, and Sam Brilleman. 2022. "rstanarm: Bayesian applied regression modeling via Stan." https://mc-stan.org/rstanarm/.

MyAnimeList. 2024. *MyAnimeList: Anime and Manga Database and Community*. MyAnimeList Co., Ltd. https://myanimelist.net/.

MyAnimeList API (Beta Ver.) (2). 2024. https://myanimelist.net/apiconfig/references/api/v2.

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

Richardson, Neal, Ian Cook, Nic Crane, Dewey Dunnington, Romain François, Jonathan Keane, Dragoș Moldovan-Grünfeld, Jeroen Ooms, Jacob Wujciak-Jens, and Apache Arrow. 2024. *Arrow: Integration to 'Apache' 'Arrow'*. https://CRAN.R-project.org/package=arrow.

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 Grolemund, et al. 2019. "Welcome to the tidyverse." *Journal of Open Source Software* 4 (43): 1686. https://doi.org/10.21105/joss.01686.

Xie, Yihui. 2014. "Knitr: A Comprehensive Tool for Reproducible Research in R." In *Implementing Reproducible Computational Research*, edited by Roger D. Peng Victoria Stodden Friedrich Leisch. Chapman; Hall/CRC.