

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

## A Correlation Analysis of Anime Ranking

-Doran Wang

November 26, 2024

This repository 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.

## 1 Introduction

Overview paragraph

Estimand paragraph

Results paragraph

Why it matters paragraph

Telegraphing paragraph: The remainder of this paper is structured as follows. Section 2....

---

\*Code and data are available at: [https://github.com/Wang20030509/Anime\\_Rating](https://github.com/Wang20030509/Anime_Rating).

## 2 Data

### 2.1 Overview

We use the statistical programming language R (R Core Team 2023).... Our data (Toronto Shelter & Support Services 2024).... Following Alexander (2023), we consider...

Overview text

### 2.2 Measurement

Some paragraphs about how we go from a phenomena in the world to an entry in the dataset.

### 2.3 Outcome variables

Add graphs, tables and text. Use sub-sub-headings for each outcome variable or update the subheading to be singular.

Some of our data is of penguins (?@fig-bills), from Horst, Hill, and Gorman (2020).

Table 1: Preview of the cleaned Anime Ranking dataset

Score	Popularity	Number of users who have this work in list	Ranking
8.72	17	2457763	58
9.09	3	3443448	3
9.03	8	2939130	9
9.07	14	2638313	4
8.62	2	4016610	91

Talk more about it.

And also planes (Figure 1). (You can change the height and width, but don't worry about doing that until you have finished every other aspect of the paper - Quarto will try to make it look nice and the defaults usually work well once you have enough text.)

Talk way more about it.

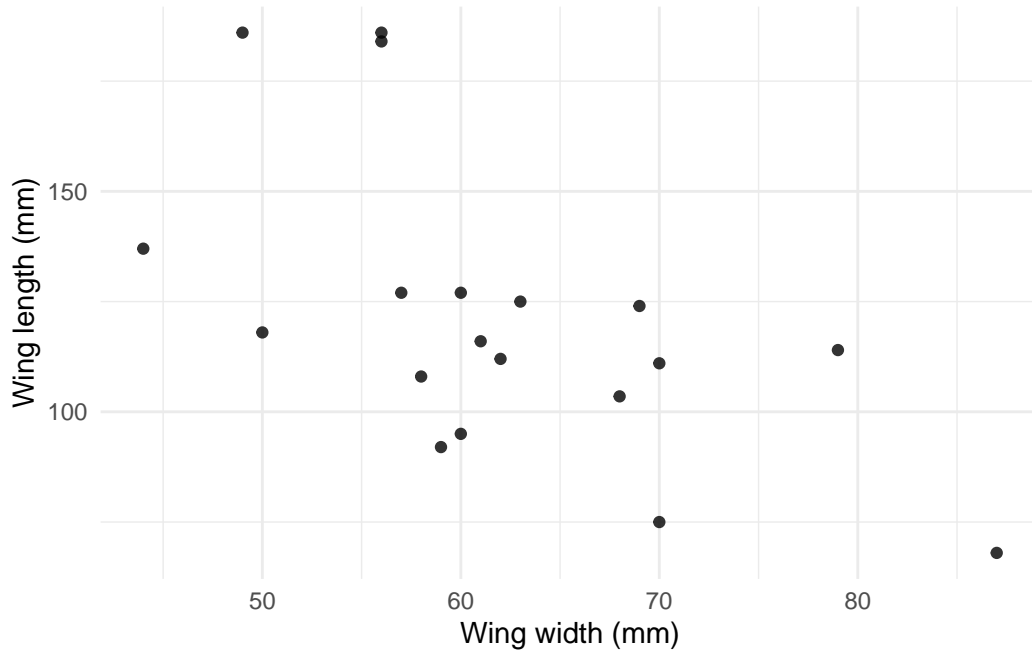


Figure 1: Relationship between wing length and width

## 2.4 Predictor variables

Add graphs, tables and text.

Use sub-sub-headings for each outcome variable and feel free to combine a few into one if they go together naturally.

## 3 Model

The goal of our modelling strategy is twofold. Firstly,...

Here we briefly describe the Bayesian analysis model used to investigate... Background details and diagnostics are included in [Appendix B](#).

### 3.1 Model set-up

Define  $y_i$  as the number of seconds that the plane remained aloft. Then  $\beta_i$  is the wing width and  $\gamma_i$  is the wing length, both measured in millimeters.

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

$$\mu_i = \alpha + \beta_i + \gamma_i \tag{2}$$

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

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

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

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

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

### 3.1.1 Model justification

We expect a positive relationship between the size of the wings and time spent aloft. In particular...

We can use maths by including latex between dollar signs, for instance  $\theta$ .

## 4 Results

Our results are summarized in Table [2](#).

## 5 Discussion

### 5.1 First discussion point

If my paper were 10 pages, then should be be at least 2.5 pages. The discussion is a chance to show off what you know and what you learnt from all this.

### 5.2 Second discussion point

Please don't use these as sub-heading labels - change them to be what your point actually is.

Table 2: Explanatory models of flight time based on wing width and wing length

	First model
(Intercept)	1.12 (1.70)
length	0.01 (0.01)
width	−0.01 (0.02)
Num.Obs.	19
R2	0.320
R2 Adj.	0.019
Log.Lik.	−18.128
ELPD	−21.6
ELPD s.e.	2.1
LOOIC	43.2
LOOIC s.e.	4.3
WAIC	42.7
RMSE	0.60

### **5.3 Third discussion point**

### **5.4 Weaknesses and next steps**

Weaknesses and next steps should also be included.

# Appendix

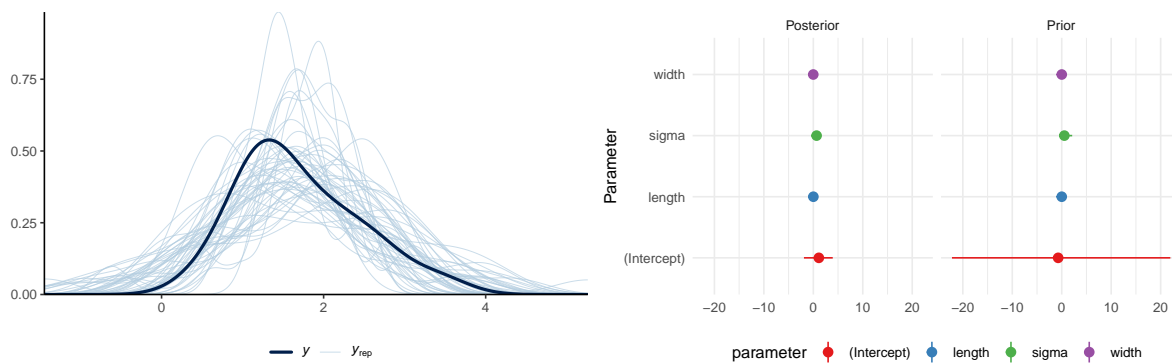
## A Additional data details

## B Model details

### B.1 Posterior predictive check

In Figure 2a we implement a posterior predictive check. This shows...

In Figure 2b we compare the posterior with the prior. This shows...



(a) Posterior prediction check

(b) Comparing the posterior with the prior

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

### B.2 Diagnostics

Figure 3a is a trace plot. It shows... This suggests...

Figure 3b is a Rhat plot. It shows... This suggests...

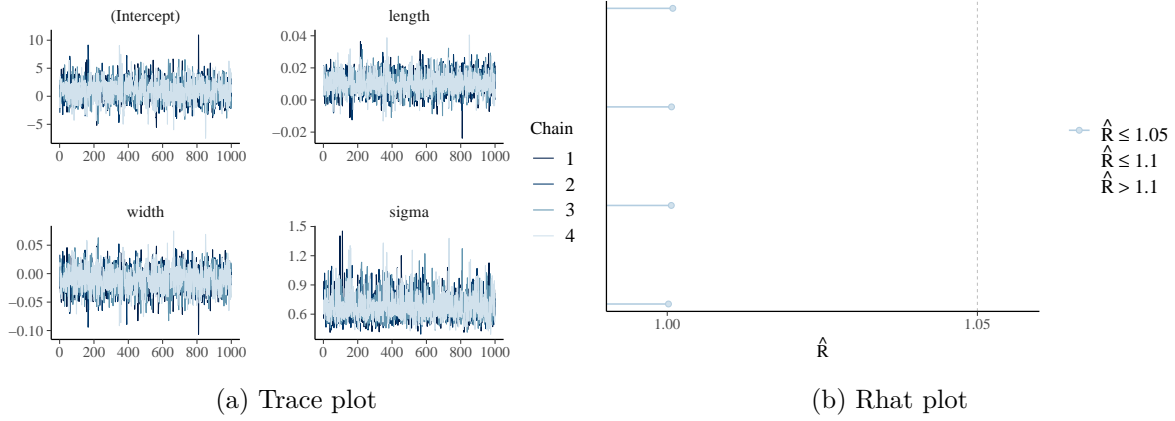


Figure 3: Checking the convergence of the MCMC algorithm

## 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/>.
- Horst, Allison Marie, Alison Presmanes Hill, and Kristen B Gorman. 2020. *palmerpenguins: Palmer Archipelago (Antarctica) penguin data*. <https://doi.org/10.5281/zenodo.3960218>.
- R Core Team. 2023. *R: A Language and Environment for Statistical Computing*. Vienna, Austria: R Foundation for Statistical Computing. <https://www.R-project.org/>.
- Toronto Shelter & Support Services. 2024. *Deaths of Shelter Residents*. <https://open.toronto.ca/dataset/deaths-of-shelter-residents/>.