Are MAL Ratings Reliable*

Mohid Sharif

27 April 2022

Abstract

Many people on the internet rely on anime ratings on MyAnimeList to determine what to watch next. I obtained data from MyAnimeList, the biggest anime rating site on the web and analysed the effect of popularity, fanbase, number of episodes, and genre on an anime's rating. I found that anime is ranked rather fairly whether that be a large or a small fanbase. My study can be used to show whether ratings from MAL are reliable or if the ratings are biased depending on the popularity of the anime.

Keywords: anime, MyAnimeList, MAL, MAL ratings, manga, anime ratings, anime genre, anime popularity, anime score

1 Introduction

Watching anime has been a hobby of mine since middle school. There are thousands of anime out there and it's not easy finding one that I will enjoy. So it's very important that I have a reliable way to determine which anime I should watch next. My most reliable method is getting recommendations from friends. Most of my favorite anime have been recommended to me by my friends since they are aware of my preferences. My other method is finding anime on MyAnimeList. This method is less reliable, many times I will find an anime that interests me but I don't end up enjoying it. This could be due to false ratings or maybe the story isn't what I had expected it to be. In this paper I wanted to determine whether anime ratings on MAL are accurate and reliable.

I first found a dataset that retrieved information about 17562 different anime and 325772 different users from MAL. After this I cleaned the data to isolate variables that interested me; popularity, episodes, score, genre and members. From here I compared how the popularity, number of episodes and genre of an anime effected it's score. This was done by creating graphs which compared these variables and found any relationships between them. I noticed that anime with less episodes had more popularity and ended up having an overall higher rating. While anime with more episodes had more members in their fan base while having an overall lower rating. However anime were ranked with little bias since users only joined fanbases if they enjoyed the anime thus that anime would have a higher rating and vice versa.

I used my data as a way to determine reliability of user ratings on MAL. I have seen cases where a fan base has collectively tanked an anime's rating as a joke and thus creating false ratings. This way I could determine whether this was a common case on MAL, making the anime ratings unreliable. After my research I have concluded that MAL is in fact a reliable source for information on anime and is a great way to discover new anime.

*Code and data are available at: https://github.com/MohidSharif/Anime-Ratings-Data-Study.

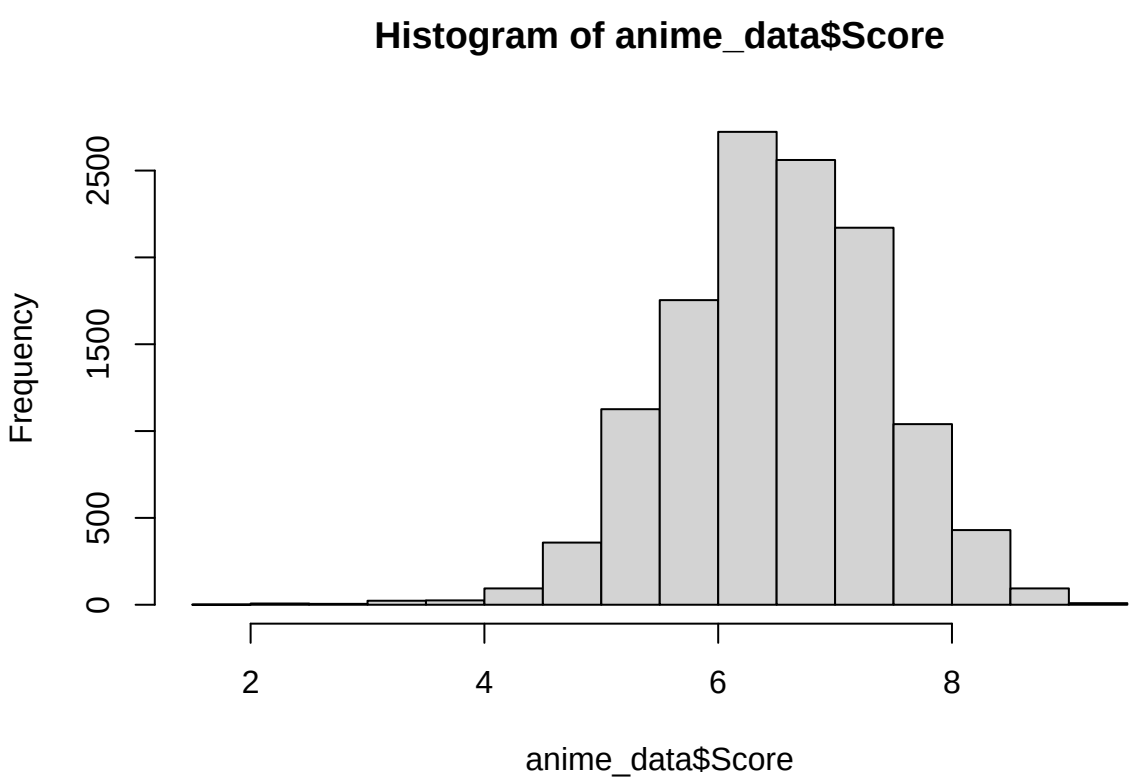
2 Data

Table 1: Sample Data From Our MAL Dataset

ID	English Name	Score	Type	Episodes	Ranked	Popularity	Members	Completed
1	Cowboy Bebop	8.8	TV	26	28	39	1251960	718161
5	Cowboy Bebop:The Movie	8.4	Movie	1	159	518	273145	208333
6	Trigun	8.2	TV	26	266	201	558913	343492
7	Witch Hunter Robin	7.3	TV	26	2481	1467	94683	46165
8	Beet the Vandel Buster	7.0	TV	52	3710	4369	13224	7314
16	Honey and Clover	8.1	TV	24	468	687	214499	81145
19	Monster	8.8	TV	74	30	169	614100	214491
20	Naruto	7.9	TV	220	660	8	1830540	1462223
22	The Prince of Tennis	7.9	TV	178	675	1039	141832	76881
24	School Rumble	7.9	TV	26	625	514	275464	157789

3 Model

```
hist(anime_data$Score)
```



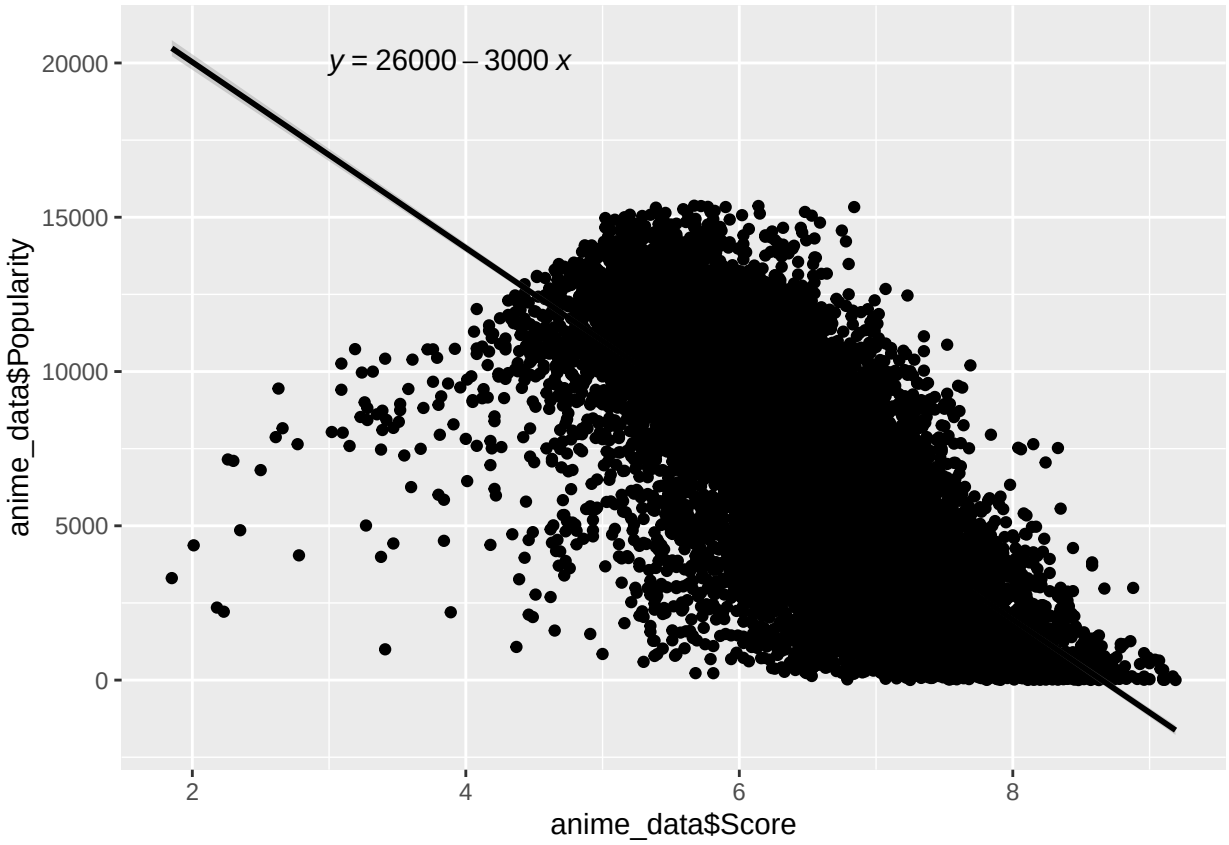

```
anime_data.lm <- lm(Score ~ Popularity, data = anime_data)

summary(anime_data.lm)
```

```
##
## Call:
## lm(formula = Score ~ Popularity, data = anime_data)
```

```
##
## Residuals:
##     Min      1Q  Median      3Q     Max
## -5.1633 -0.3494  0.0487  0.4206  1.9920
##
## Coefficients:
##               Estimate Std. Error t value Pr(>|t|)
## (Intercept)  7.543e+00  1.118e-02   674.9   <2e-16 ***
## Popularity  -1.602e-04  1.489e-06  -107.6   <2e-16 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Residual standard error: 0.638 on 12419 degrees of freedom
##   (5141 observations deleted due to missingness)
## Multiple R-squared:  0.4824, Adjusted R-squared:  0.4823
## F-statistic: 1.157e+04 on 1 and 12419 DF,  p-value: < 2.2e-16
```

```
score.graph<-ggplot(anime_data, aes(x=anime_data$Score, y=anime_data$Popularity))+
                    geom_point()+
                    geom_smooth(method="lm", col="black")+
                    stat_regline_equation(label.x = 3, label.y = 20000)
score.graph
```



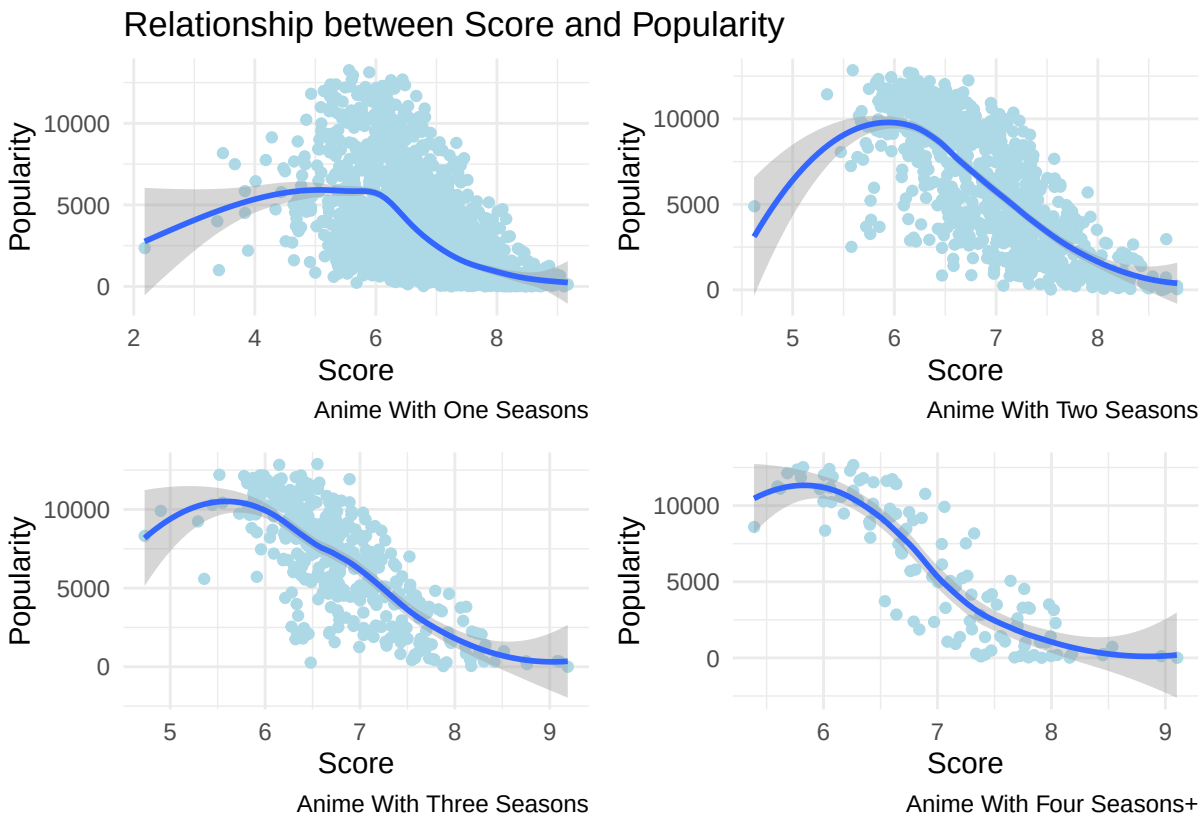


Figure 1: Relationship between Score and Popularity Categorized By Length of Anime

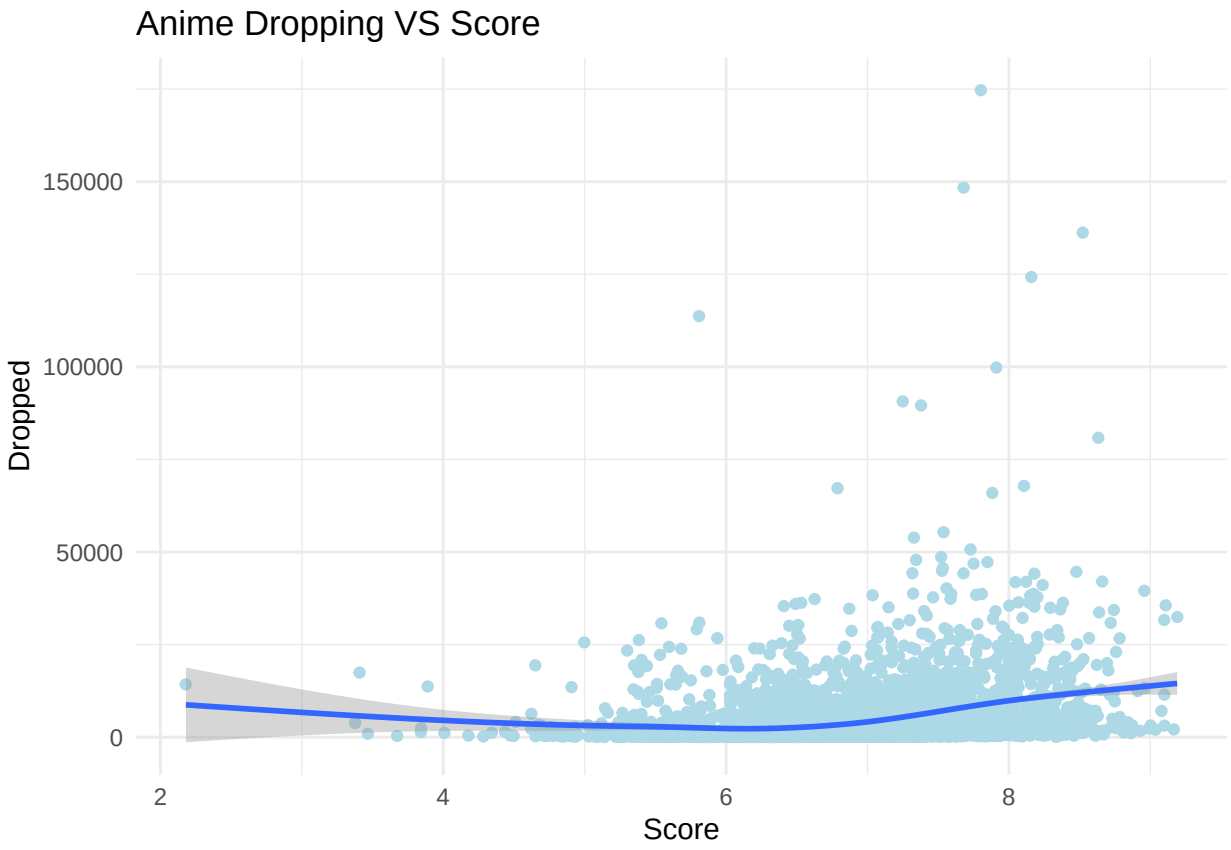


Figure 2: Effect Of Dropping An Anime On Score

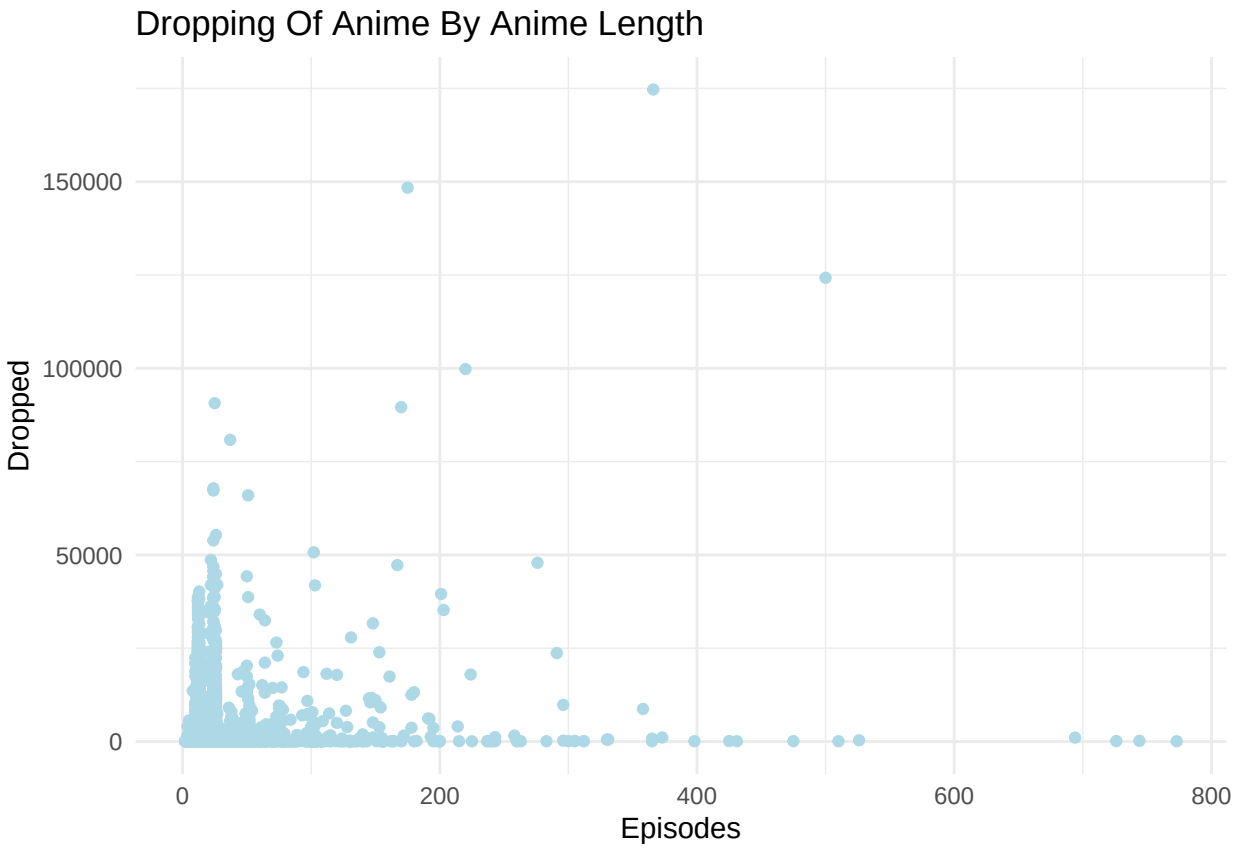


Figure 3: Anime Relationship between Dropped And Episodes

4 Results

5 Discussion

5.1 First discussion point

5.2 Second discussion point

5.3 Third discussion point

5.4 Weaknesses and next steps

Appendix

A Additional details

B References