An Investigation into the Factors which Influence a Video Game's Popularity

Darling Judah Hsu, Duy Nguyen, Zhanyi Qiu

Introduction

Over the years, opinions on video games have changed and become varied in quality and the player response reflects upon that. For this project, we will be analyzing different methods of visualization on the player review, to see the differences and determine which types of game perform better. We will primarily use data from Steam as a way to look at said player reviews due to there being existing classifications on review types and there being high volumes of said amount. However, we will also use Metacritic reviews too. In particular, we will be looking at AAA and indie games: AAA games being games developed by big companies and independent games being developed by independent developers. For our hypothesis, we believe that indie games will yield better results compared to that of AAA.

Limitations

It is unknown what happened to the API between 2021 to 2023, but the developer now limited the amount of pages allowed to be displayed on the data requests to 50 or less, rather than allowing the maximum value. This heavily limits the maximum data points of all queries from the standard amount to a mere 3000. Additionally, the API only allows filtering of AAA games but NOT indie games. Alternative methods of acquiring data regarding indie games will be required. Furthermore, the API can sometimes return incorrect results to that of Steam, so inspections are needed. Additionally, the reason why we use this API is because Steam no longer offers the ability to freely use their API for store lookup, as noted by the fact that the store page does not have an internal request for the API when inspecting the website. However, web scraping is still required because Cheapshark does not return publisher and developer data,

therefore a code must be implemented. Furthermore, the list of games updates quite often, which can result in our data not being up to date, but still relevant. To supplement and enhance the data acquired from the Cheapshark, we use Steam web API to obtain the review counts. One limitation is that the count was only collected for recent years, not all of it. However, these data acquired enough information to show the game's popularity and the player's thoughts on the game. This data was finalized on December 4, 2023, and is used for this report.

Methods

We will first query for data in the Cheapshark API with the AAA parameters to find games above $29.99. Then we will figure out indie games using an alternative method by calling for the games normally and then filtering only prices below $29.99. The requests will need to be looped for all possible page counts allowed by the API. The release date format is in Unix code, so that will need conversion. Then, we will use the names to parse through the Steam website itself and scrape the developer's and the publisher's data to add to both sets. Finally, the dataset's review information will be further enhanced, with additional columns detailing the counts of positive and negative reviews from another API in Steamworks. We will then use the dataset we've made for various forms of visualizations. We will not look at the counts as a way to compare at face value but in relative proportion. Due to the amount of columns, the snippet will only show a small sample of the data, but the full CSV files will be provided on the submission page.

AAA Set Sample (Not all columns Included):

steamAppID	releaseDate	normalPrice	metacriticScore	steamRatingText	steamRatingPercent
804490	2018-09-25	29.99	84	Very Positive	80
666140	2019-01-15	29.99	73	Very Positive	92
1126750	2020-06-25	29.99	76	Very Positive	82
611670	2018-04-02	59.99	81	Mostly Positive	74

developers	publishers	total_reviews	positive_reviews	negative_reviews

Survios	**Survios**	**947**	**734**	**213**
Pathea Games	**Focus Entertainment**	**11234**	**10409**	**825**
INFINITY Co., Ltd.	**BANDAI NAMCO Entertainment**	**76**	**66**	**10**
Bethesda Game Studios	**Bethesda Softworks**	**6124**	**4568**	**1556**

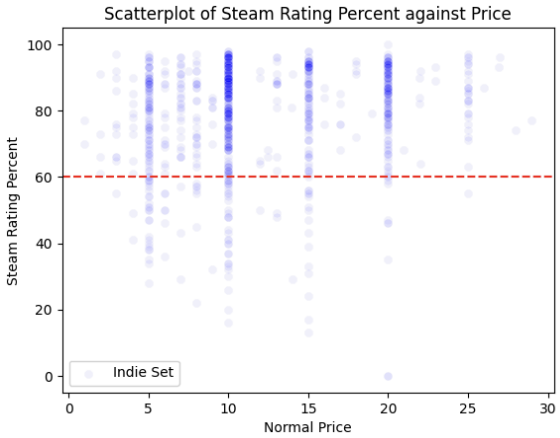
Indie Set Sample:

steamAppID	releaseDate	normalPrice	metacriticScore	steamRatingText	steamRatingPercent
294750	**2014-05-29**	**9.99**	**70**	**Mostly Positive**	**75**
1089830	**2020-02-04**	**24.99**	**72**	**Very Positive**	**88**
246760	**2014-07-21**	**19.99**	**71**	**Mostly Positive**	**76**
263860	**2014-09-17**	**9.99**	**68**	**Mixed**	**65**

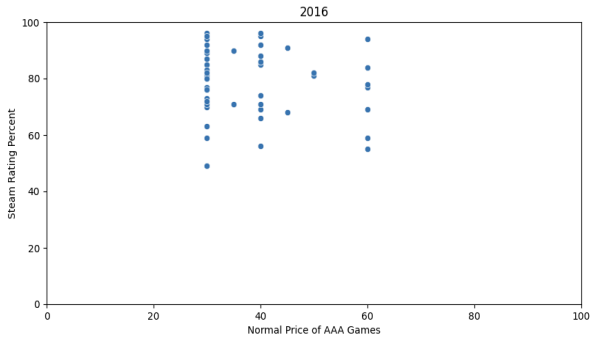
developers	publishers	total_reviews	positive_reviews	negative_reviews
11 bit studios	**11 bit studios**	**145**	**99**	**46**
Milestone S.r.l.	**Milestone S.r.l.**	**590**	**512**	**78**
Aterdux Entertainment	**Aterdux Entertainment**	**442**	**296**	**146**
Flow Combine	**11 bit studios**	**85**	**56**	**29**

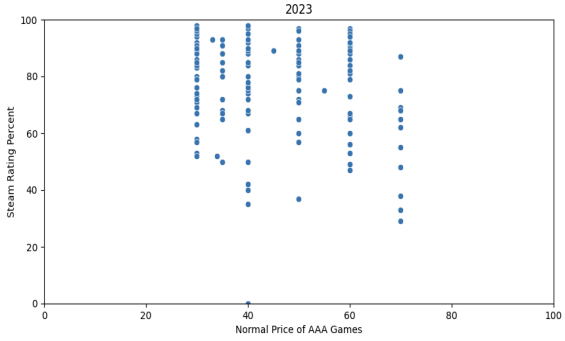
Results

In specific, the variables we'll be focusing on are the "metacriticScore", "publishers," "steamratingText," "steamRatingPercent," and "normalPrice" variables. The Metacritic score refers to scores given by critics on the Metacritic platform, steam rating text, and percentages refer to ratings given by players (with playtime on the game), normal price refers to the norm price of the games without deals, and publishers refers to the companies or groups that published the game. We'll be looking at these variables to compare the indie and AAA datasets.

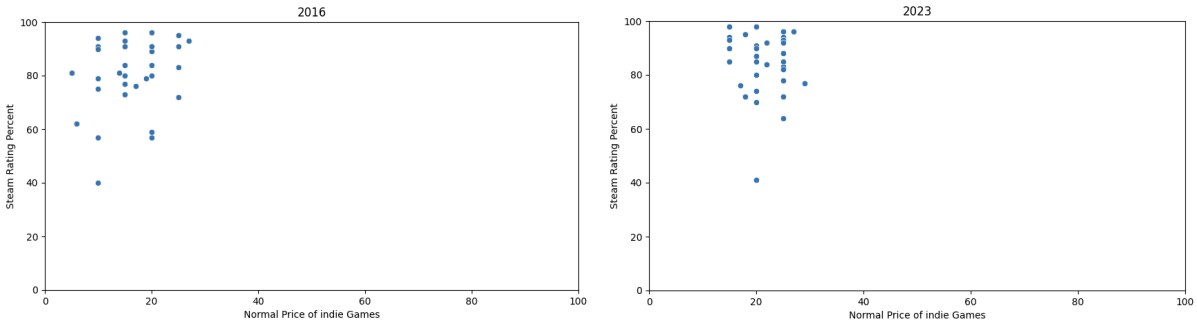


Plotting the rating percentage against their prices in this way when including all years, we see that while there are more ranges of priced games in the indie section below our designated red line, their relative proportions for indie games appear to be lesser for indie games compared to that of AAA games. Over all years, it would appear as if AAA games were doing better, but when we compare them during different years, we will see the difference in performance over time.

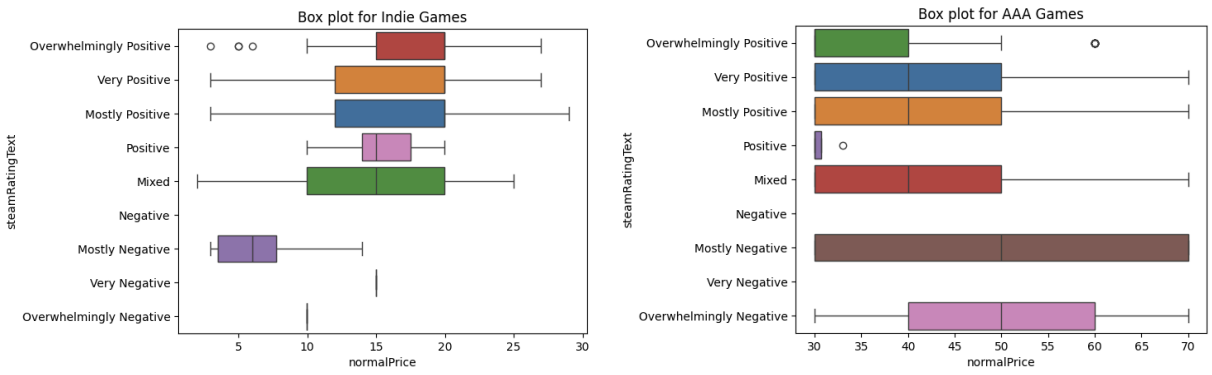




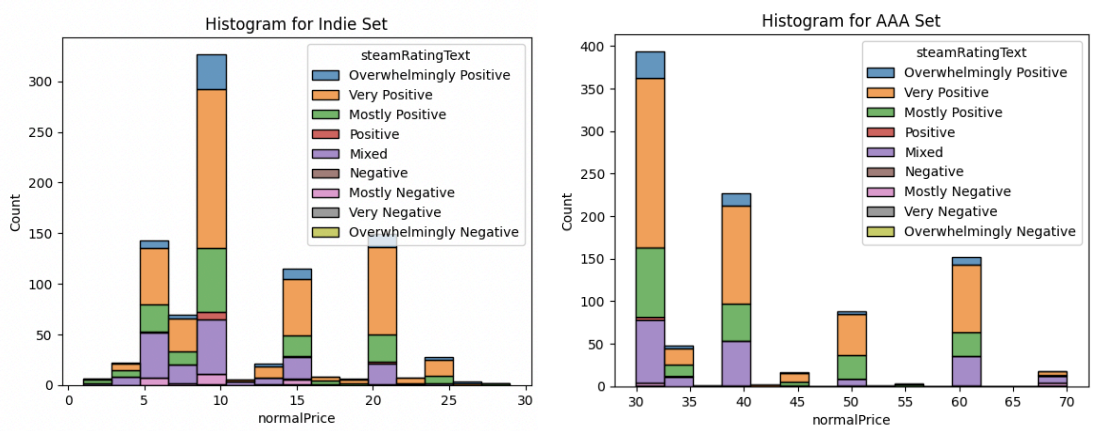
Starting with a rating percentage against the prices of video games, 2016 had a pretty small range in reviews. 2023 is when $70 games have become a part of AAA games and it appears that the ranges have increased, including lower review percentages.



For indie games, it would appear that earlier years had a pretty high range of reviews extending into the lower numbers. Over recent years, it would appear that the performance has improved because the ranges are smaller and more compact. Compared to the earlier plot, it seems like indie games are indeed performing better in reviews, supporting our hypothesis. The pdfs do not support gifs, so we had to only include 2 frames of it in this report but the code for the file will be included in the Jupyter notebook.

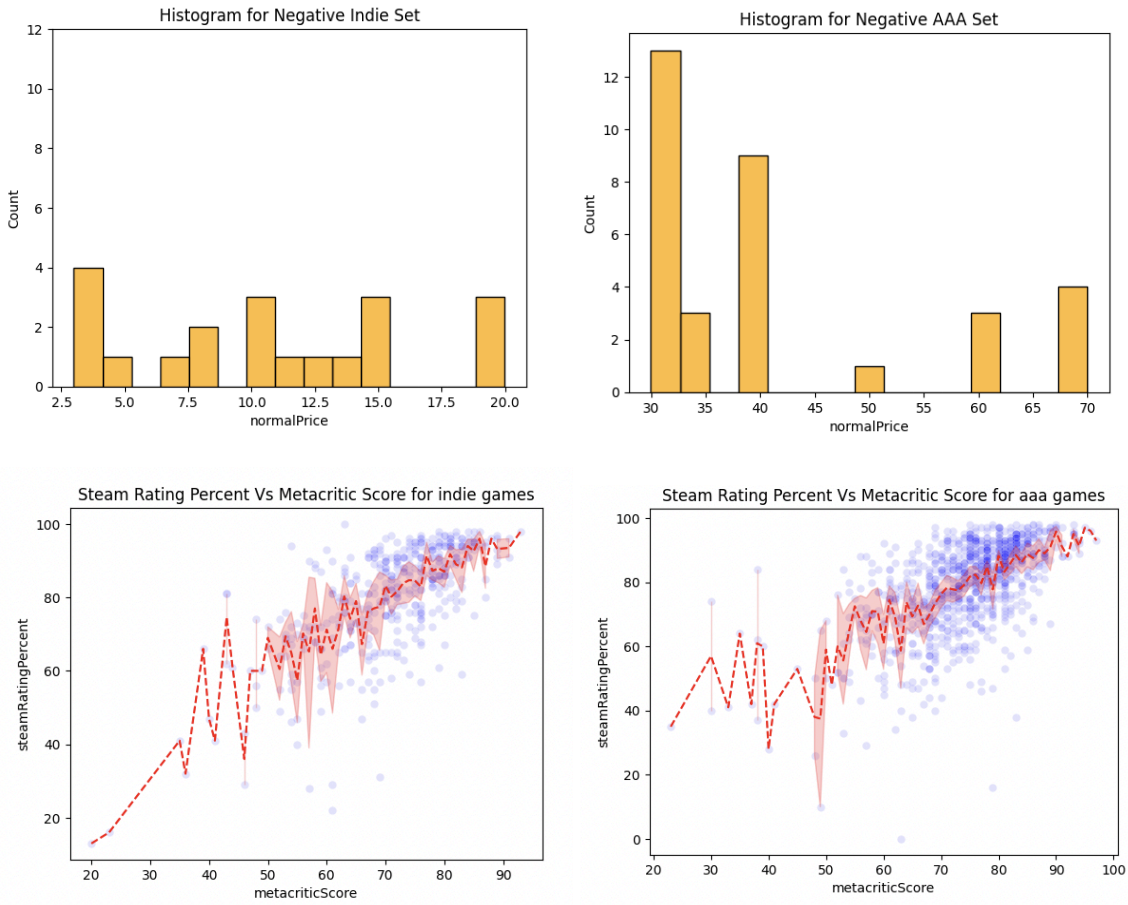


It would seem like for AAA games there are a noticeable amount of games that are in the negative reviews, with lower price ranges instead being in the positive area. Indie games also feature a similar aspect where they have better performance at lower prices. It is however worth noting that the mostly negative reviews cover lower price ranges, which makes sense because these prices reflect games with not too much effort in development. Regardless It's evidence to support that not only do indie games have better performance, but that it seems reviews are higher for lower-priced games.

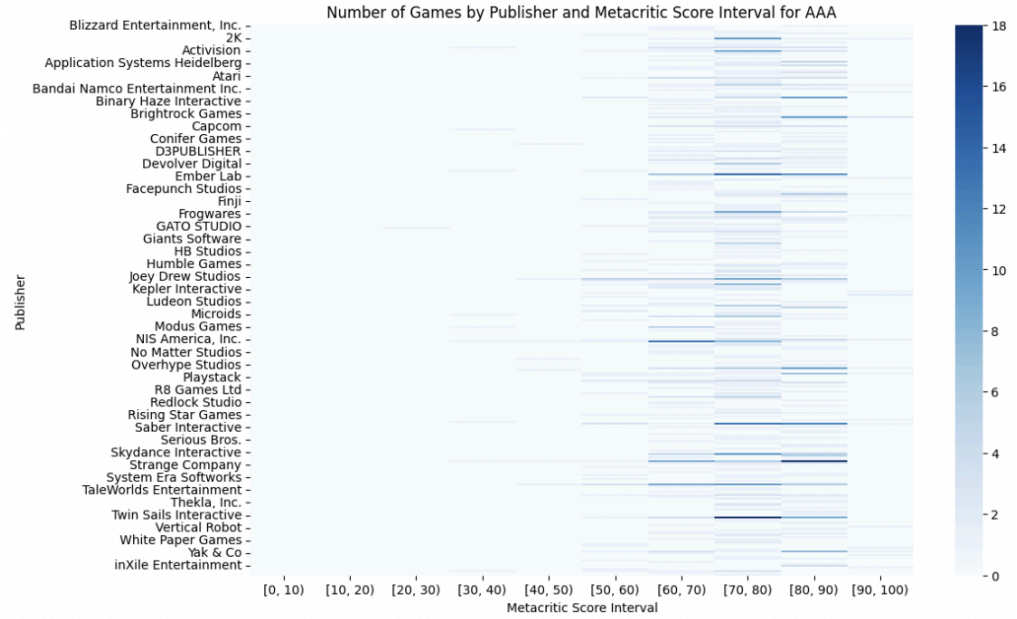


This histogram looks at the counts of games and is filtered by their respective classifications of reviews. Because there are not enough negative reviews to be seen on the histogram, 2 more histograms only for negative reviews are necessary for comparison.

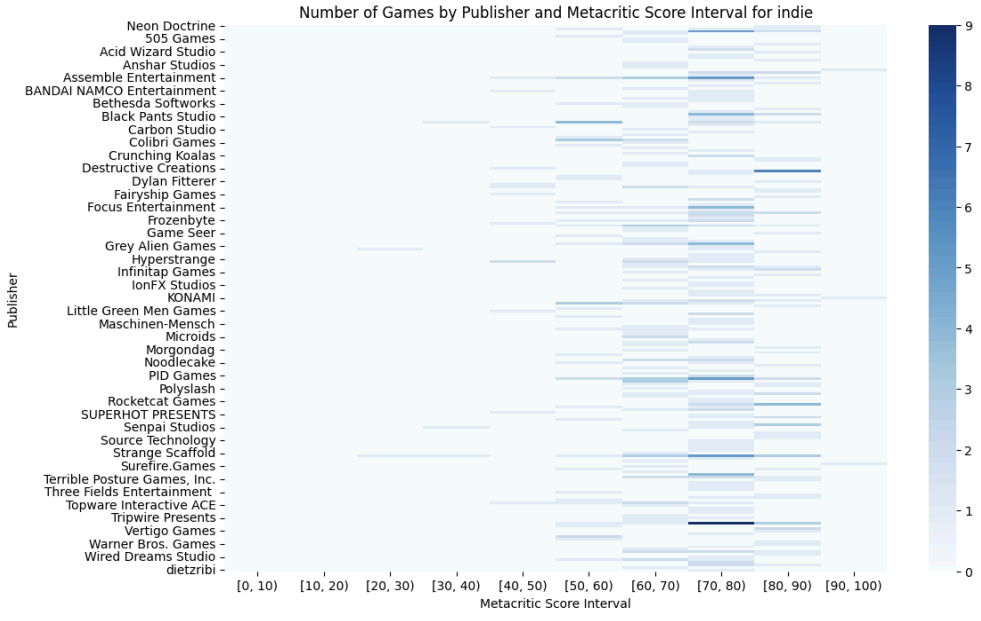
The scaling between the 2 histograms below is different on the Y axis, but AAA games do have a higher relative proportion of negative reviews in their price ranges compared to that of indie games.



Scatterplot between steam rating percentages and Metacritic scores for AAA games, with a moving average line. Here we see that generally Metacritic score and Steam rating percent have a high correlation, especially closer to the points. In comparison, there does not to be a big difference in shape when it comes to looking at the difference between Steam and Metacritic reviews. We deem that between Metacritic and Steam, there is not enough difference to have a large effect.



On AAA games, the respective publishers appear to have a narrow range on the heatmap for the reviews. The lowest appears to start from 50. It is worth noting that these are only the top publishers with a high number of games released.



In indie games, the spread of Metacritic scores is noticeably larger than that of AAA games (noticeable patterns can be seen at intervals [50, 90] in the indie set compared to [60, 90] in the AAA set). This is likely because they're less consistent as independent developers rather than formalized, big companies. Nonetheless, indie publishers appear to be less consistent in their game favorability. It suggests that perhaps it's a little more unpredictable as to how games might perform.

Discussion

Overall, our research supports the idea that indie games tend to have better performance than AAA games in recent years. This could be the result of the often-uniqueness of small companies, non-standardized games, or cheaper prices. Additionally, recent news has seen AAA games performing incredibly poorly such as Activision Blizzard's *Call of Duty: Modern Warfare III*. Nevertheless, it appears that indie and AAA games have similar results in terms of game favorability: both sides will have games people like and don't like. However, our evidence also suggests that indie games tend to be more inconsistent with their results. The data for the heatmap was on for all years, but our best evidence of review change comes from the plots for different years, showing that the range of performance is no longer narrow for AAA games and is improved for indie games. Still, there are a few points that must be considered for future reference.

Another point is that for the Metacritic reviews, we took reviews from critics. Though it's their job, critic reviews must always be taken with a grain of salt since a critic's intention when trying food or playing a game is different from the average enjoyer's intention. However, we chose critic reviews over user reviews partly because the Metacritic website doesn't require reviews to be by users who've played the game; on Steam, a user can only review a game if they've recorded some playtime. Regardless, this point shouldn't be noted as a huge issue since a high correlation was found between the Metacritic score and the Steam percent rating. In a future experiment, we should instead try to query for user reviews, but we can experience large differences due to anyone being able to simply make an account and potentially review bombing. As of this project, however, we believe we were successful in gathering data on AAA games and indie games to show their performance and prove our hypothesis.

Acknowledgments

Professor Kramlinger assisted in the guidance and review of decisions to ensure that the project went in the correct direction.

Cheapshark:
https://apidocs.cheapshark.com

Metacritic reviews were on official critics instead of the player, but was still reliable because the reviews were on different platforms instead of exclusively PC.
https://www.metacritic.com

Steam store was used for web scraping
https://store.steampowered.com

Call of Duty Modern Warfare III reference:
https://store.steampowered.com/app/2519060/Call_of_Duty_Modern_Warfare_III/

ChatGPT was used to assist with refining the dataset and understanding coding methods.