Predicting Player Retention from Steam Metadata and Review Signals

Justin Mabry
Department of Arts & Sciences
Indiana University Bloomington
:Bloomington, U.S.
jlmabry@iu.edu

Cole Bradfield
Department of Computer Science
Indiana University Bloomington
Bloomington, U.S.
cmbradfi@iu.edu

Luke Sundquist
Department of Computer Science
Indiana University Bloomington
Bloomington, U.S.
lsundqui@iu.edu

Sam Berebitsky
Department of Computer Science
Indiana University Bloomington
Bloomington, U.S.
sambereb@iu.edu

***Abstract*—This project is about comparing Steam game information with user reviews to see what makes players enjoy and keep playing them. The metadata includes genre, price, rating, and tags, while the reviews provide sentiment, playtime, and helpfulness scores. By joining these datasets through game IDs, we aim to identify patterns between game mechanics and user responses.**

Keywords—Steam, game reviews, player satisfaction, metadata, sentiment analysis, marketplace

I. Introduction

Steam offers a large amount of structured game information and user-generated reviews. This combination makes it possible to study how game features relate to player opinions and behavior. Previous work has looked at mobile game retention and what factors induce higher results, [1], [2], but fewer studies combine metadata with review-level data. As students, we want to use this project to practice dataset normalization, statistical analysis, and visualization. The main question is: which game attributes are most connected to positive reviews and longer playtime? Reference [3] was the first study we found that was somewhat similar to what we wanted to do, however, it's study was limited to 1000 games on the marketplace. Our study on the other hand, has more than 50,000 entries before thinning. There have been studies, [4], that look at the data within individual games to see what makes them stop playing. But this doesn't explain what types of games keeps hold of players the most.

II. Methods

A. Game Metadata

The first step is preparing the game metadata. The dataset, [5], was originally in JSON format, with nested dictionaries that had to be flattened into a table where each row represents a single game. Using python and the pandas library, we converted the data into a data frame and removed entries that could not be analyzed, such as unreadable titles or reviews written in unsupported languages. We also filtered tags by applying a threshold, so that only tags with enough votes remain attached to the game. This normalization ensures that the metadata is consistent and ready for analysis.

B. Review Dataset

The review datasets are organized by game ID, with each file containing individual reviews. These reviews include whether the player's feedback was positive or negative, their playtime, and whether the review was marked helpful or not. We plan to preprocess the reviews by cleaning out non-standard symbols, removing entries in unsupported languages, and tokenizing the text for sentiment analysis. Libraries such as NLTK or spaCy could be used to text processing, while scikit-learn provides tools for sentiment classification and feature extraction.

C. Comparative Analysis

Once both datasets are cleaned, the next step is to join them together using the game ID as the key. This allows me to directly compare game-level features (such as genre, price, and tags) with review-level outcomes (such as sentiment and playtime). To explore these relationships, we will start with descriptive statistics and visualizations using matplotlib and seaborn. For example, we can calculate average playtime by genre or compare the distribution of positive reviews across price ranges. Beyond descriptive analysis, we plan to use correlation tests and regression models to see which features are most predictive of positive reviews or longer playtime. Logistic regression could be applied to model the probability of a review being positive based on game attributes, while linear regression could be used to predict playtime. Clustering methods, such as k-means, may also help identify groups of games with similar review patterns.

III. Expected Results (Hypothesis)

We're expecting that certain genres, such as role-playing or strategy games, will show higher average playtime and more positive reviews compared to casual or puzzle games. Certain First-person Shooters should fall into this category as well with the boom in the genre post-COVID. We also hypothesize that the lower-priced games may receive more mixed reviews, while mid-ranged games could show stronger satisfaction if they balance cost to content. As mentioned above, another expectation is the tags related to multiplayer or replayability will correlate with longer playtime. Overall, we predict that metadata features like genre and price will align with review sentiment, but that playtime will be the strongest indicator of user retention.

IV. Author Contribution Statement

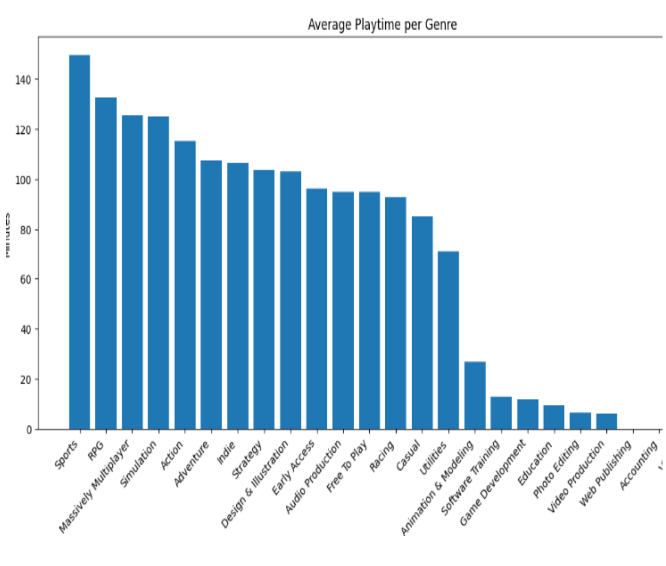
We designed the project, cleaned the game metadata, and wrote this proposal. No AI has been used yet in the creation of this project or proposal. As students, while still learning more advanced methods, we're confident in applying the data analysis and visualization techniques required.

V. Results

After cleaning our data, we identified several different trends in terms of game recommendations, genre, tags, and playtime that we found were important.

A. *Genre vs Play time*

First, we graphed each genre to their average playtime to see which genre keeps players' attention the longest. We can see that genres like RPG, action, adventure are high in playtime, while casual, puzzle, and video/photo editing games are lower down. This matches what we hypothesized: that the RPG and action games are designed to keep you engaged for long periods of time. They are challenging enough to keep you wanting to improve, but not too difficult to where you are frustrated and feel hopeless. More casual games are generally created for people to play in their free time, and to without too much stress.

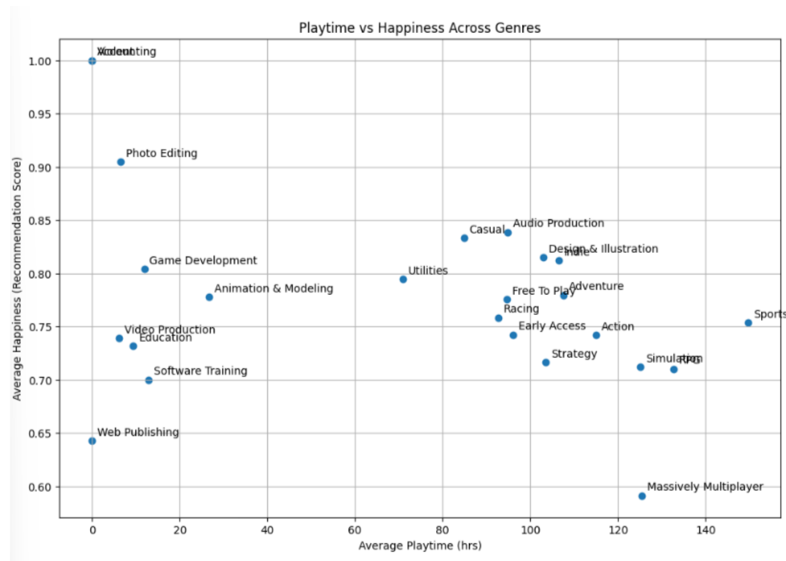


Average Playtime per Genre (Figure A)

B. *Playtime vs Happiness*

Next, we took the Average Playtime Per Genre one step further by adding another attribute, Player satisfaction. To calculate each genre's satisfaction score, we converted each review into either a 1 ("Recommended") or a 0 ("Not Recommended"). Then we added the scores for each genre together and divided by the number of reviews to get the average satisfaction, or recommendation rate.

As shown in Figure B, there is very little difference in satisfaction between genres with shorter average playtime, and almost no difference in which genre's have a high satisfaction score. The wide range of successful games across different genres and playtimes shows that players have diverse preferences. This causes us to conclude that the key to how well received a game is, is less about maybe being a popular genre and more about the creating the best playing experience .



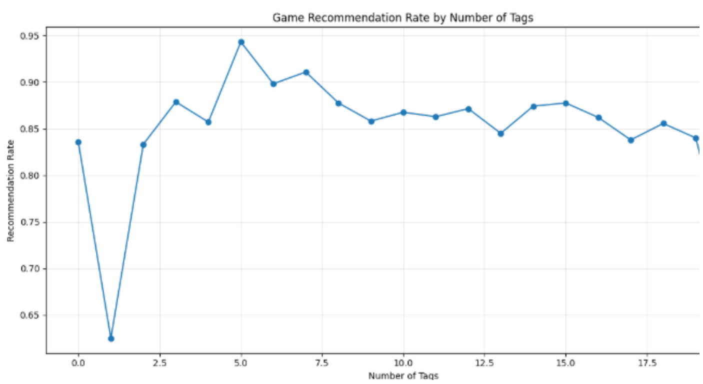
Playtime vs Happiness Across Genres(Figure B)

C. *Tag Correlation with Game Recommendations*

Another trend that may help predict a game's success is how certain tags relate to recommendation rates. We visualized this by measuring how strongly each tag correlates with being recommended (Figure C).The better the recommendations of games using that specific genre tag, the closer the correlation gets to 0. Tags such as shooter, RPG, action, and adventure, have correlations closest to 0, meaning games with these tags tend to be recommended more often. This reaffirms our prior prediction that games with these tags are better received as they are more skill-based, intense, and keep you engaged.

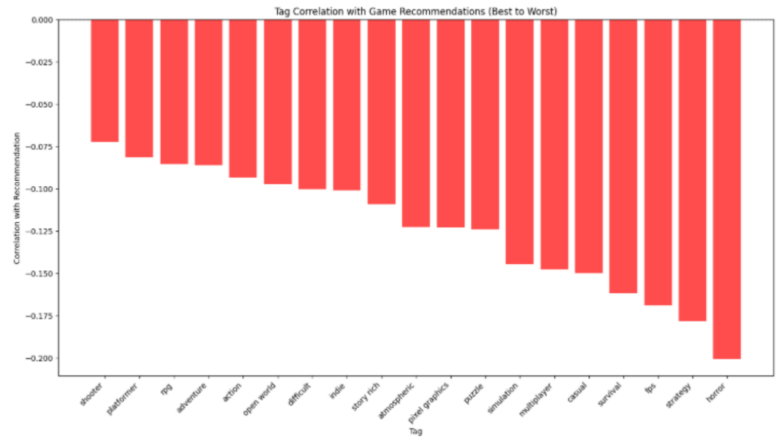
D. Game Recommendation Rate vs. Number of Tags

Continuing with trends that have to do with tags, We can calculate the recommendation rate for each number of tags that a game may have (Figure D). From the graph, we can tell that the recommendation rate peaks at 0.95 at 5 tags and stays above .9 until 7.5 tags. The recommendation rate then dips below 0.9, and plateaus. There may be extensive reasons for this, but we deduced that this may be caused by gamers not wanting an oversaturated game. When the number of tags approaches 10, players may feel the game is trying to do too many different things. The results may also be due to the games not seeming authentic, as if the game studio is adding tags in for no other reason than getting attention. Overall, this graph displays the fact that gamers want clear and concise tags that get to the core of the game.



RECOMMENDATION RATE VS. NUMBER OF TAGS

(FIGURE D)



TAG CORRRELATION WITH GAME RECOMMENDATIONS

(FIGURE C)

References

[1] S. Gupta, "Mobile Game Retention Benchmarks: An Industry Breakdown," *Nudgenow.com*, 2024. https://www.nudgenow.com/blogs/mobile-game-retention-benchmarks-industry

[2] D. Curry, "Mobile Game Retention Rates (2024)," *Business of Apps*, Jul. 23, 2024. https://www.businessofapps.com/data/mobile-game-retention-rates/

[3] Oleksandra-Shevchenko, "GitHub - Oleksandra-Shevchenko/steam-player-retention-analysis: Data analysis project exploring Steam game player retention, focusing on game lifecycle, genre, content updates, and market competition.," *GitHub*, 2025. https://github.com/Oleksandra-Shevchenko/steam-player-retention-analysis (accessed Nov. 14, 2025).

[4] T. Allart, G. Levieux, M. Pierfitte, A. Guilloux, and S. Natkin, "Design influence on player retention: A method based on time varying survival analysis," *2016 IEEE Conference on Computational Intelligence and Games (CIG)*, pp. 1–8, Sep. 2016, doi: https://doi.org/10.1109/cig.2016.7860421

[5] H. Abdelqader, "Steam Games Metadata and Player Reviews (2020–2024)," *Mendeley Data*, Jun. 2025, doi: https://doi.org/10.17632/jxy85cr3th.2.