Toxicity & Profitability in Gaming Communities

An Empirical Analysis of Player Behavior and Game Success

PROJECT REPORT

Submitted

In partial fulfillment of the requirements

for the awarded degree

of

MASTER OF SCIENCE IN BUSINESS ANALYTICS

To - Professor Chao Fang

By - (Zubaish, Ester, Noor, Tino, Steve)

SCHOOL OF BUSINESS

CLARK UNIVERSITY

950 Main St, Worcester, MA 01610

December, 2025

Table of Contents

Back to Table of Contents

1. INTRODUCTION

1.1 Background

The gaming industry has transformed into one of the most dynamic entertainment sectors, generating billions in annual revenue and hosting massive player communities. As online gaming grew, so did the visibility of toxic behavior, negative language, harassment, and aggressive interactions in player reviews and online discussions. Because player sentiment can influence retention, reputation, and monetization, understanding the real effect of toxicity on business outcomes has become increasingly important.

1.2 Objective

This project aims to quantify whether **community toxicity impacts engagement and profitability** across video game genres. Using scraped player reviews combined with Steam's public performance metrics, we examine relationships among:

a) **Toxicity** (negative review sentiment)

b) **Engagement** (average playtime)

c) **Profitability/Popularity** (owners)

We combine descriptive analysis, correlation, OLS regression, logistic regression, and ANOVA to provide a comprehensive understanding of whether toxicity affects game success.

Back to Table of Contents

2. RESEARCH QUESTION

2.1 Statement

Does community toxicity (measured by negativity in player reviews) influence engagement and profitability across different video game genres?

2.2 Rationale

Many studios assume toxic communities harm a game's reputation and success. However, some highly toxic games (e.g., Grand Theft Auto, Call of Duty, and other Multiplayer Online Battle Arena a.k.a. MOBA type games) remain massively popular. This contradiction motivated us to empirically evaluate whether toxicity truly predicts performance or whether engagement outweighs sentiment.

2.3 Relevance

As online communities shape purchasing decisions and long-term engagement, understanding toxicity's role can help game studios:

a) manage community health

b) optimize moderation

c) prioritize features that drive retention

d) improve business strategies

This research provides data-driven insights into how sentiment interacts with gameplay engagement and profitability.

Back to Table of Contents

3. DATA SOURCES

3.1 Identification

We used two primary datasets:

a) Player Review Data (train.csv)
   - Contains ~17,500 user reviews with sentiment labels and toxicity flags.

b) Steam Game Performance Data (steam.csv)
   - Contains ~27,000 game listings with ratings, owners, playtime, price, and genre information.

3.2 Justification

- Steam provides one of the world's largest public gaming datasets.
- Review sentiment is a meaningful measure of community toxicity.
- Owner and playtime metrics directly capture profitability and engagement.
- Both datasets are widely used in academic and industry research.

3.3 Data Preparation

Key cleaning steps:

- Converted owners from ranges ("5000000–10000000") to numeric lower bounds.
- Standardized price, playtime, and toxicity columns.
- Merged datasets on **game_title**(renamed features in both datasets to game_title and then merged them), resulting in **42 matched titles**.
- Converted genre lists into a primary genre field for ANOVA.
- Removed duplicates and missing values.

Back to Table of Contents

4. METHODOLOGY

We performed the following analyses:

a) Descriptive Statistics
b) Data Visualizations
c) Correlation Analysis
d) OLS Regression
e) Logistic Regression (Advanced)
f) ANOVA (Genre vs Toxicity) (Advanced)

All analysis was performed using Python 3.10, utilizing the pandas and NumPy libraries for data preprocessing, seaborn for visualization, statsmodels for OLS regression and ANOVA, and scikit-learn for logistic regression modeling.

Back to Table of Contents

5. Data Analysis

5.1 Descriptive Statistics

5.1.1 Descriptive Analysis

To understand the distribution and characteristics of key variables like average toxicity, game popularity (owners), price, and average playtime, we computed descriptive statistics for all 42 games in the merged dataset.
Figure 1 summarizes the results.

	avg_toxicity	owners	price	average_playtime
count	42.000000	4.200000e+01	42.0	42.000000
mean	0.382879	5.650000e+06	0.0	2153.857143
std	0.236917	1.544831e+07	0.0	3955.685957
min	0.028986	2.000000e+05	0.0	75.000000
25%	0.149212	1.000000e+06	0.0	364.000000
50%	0.382473	2.000000e+06	0.0	957.500000
75%	0.524318	5.000000e+06	0.0	1924.250000
max	0.942993	1.000000e+08	0.0	23944.000000

(*Figure 1*)

Interpretation

- Average Toxicity:
  Mean toxicity is 0.38, but values range widely from 0.03 to 0.94, indicating substantial variation in community behavior across games.
- Owners (Popularity):
  The mean ownership is 5.65 million, but the range is extremely large (200k - 100 million), reflecting a mix of niche and highly popular games.

Back to Table of Contents

- Average Playtime (Engagement):
  Engagement ranges from 75 minutes to nearly 24,000 minutes, suggesting sharp differences in retention and replayability.

- Price:
  All games in this matched dataset are free-to-play, ensuring that popularity is not driven by pricing differences.

This variability provides a strong foundation for examining whether toxicity influences success metrics like owners or playtime.

5.1.2 Genre Distribution

To explore the diversity of game types in the dataset, we examined genre frequencies. *Figure 2* and *Figure 3* summarize the results.

genres	count
Action;Free to Play;Early Access	3
Action;Free to Play	2
Action;Free to Play;Strategy	2
Action;Free to Play;Indie;Massively Multiplayer;RPG	2
Adventure;Free to Play;Massively Multiplayer;RPG	2
Action;Adventure;Casual;Free to Play;Massively Multiplayer;RPG	2
Action;Free to Play;Massively Multiplayer	2
Action;Free to Play;Massively Multiplayer;Early Access	2
Free to Play;Massively Multiplayer;RPG	1
Action;Free to Play;Indie;Massively Multiplayer	1

(Figure 2)

Action;Free to Play;Massively Multiplayer;Simulation	1
Action;Adventure;Free to Play;Indie;Massively Multiplayer;RPG	1
Action;Adventure;Free to Play;Simulation;Sports	1
Free to Play;Indie;Simulation	1
Action;Adventure;Casual;Free to Play;Indie;Massively Multiplayer;RPG;Simulation	1
Free to Play;Simulation	1
Free to Play;Racing;Simulation;Sports	1
Action;Free to Play;Indie;Massively Multiplayer;RPG;Simulation	1
Action;Free to Play;Indie	1
Free to Play	1

(Figure 3)

Back to Table of Contents

Interpretation

- The dataset is dominated by Action, Free-to-Play, and Massively Multiplayer titles.
- These genres naturally tend toward competitive gameplay, which previous literature suggests may elevate toxic behavior.

5.1.3 Toxicity by Genre

We computed average toxicity across genres to understand which game categories experience the most negative community sentiment.

genres	avg_toxicity
Action;Free to Play;Indie;Massively Multiplayer;RPG;Simulation	0.942993
Action;Free to Play;Indie;Massively Multiplayer	0.889933
Action;Free to Play;Massively Multiplayer;Simulation	0.830556
Action;Adventure;Free to Play;Massively Multiplayer;RPG	0.787921
Action;Adventure;Free to Play;Simulation;Sports	0.555985
Action;Free to Play;Strategy	0.543374
Free to Play;Racing;Simulation;Sports	0.533654
Free to Play;Strategy	0.526316
Casual;Free to Play;Indie;Massively Multiplayer;Strategy	0.518325
Adventure;Free to Play;Massively Multiplayer;RPG	0.468933

(Figure 4)

Key Insights

- The most toxic genres are all combinations of Action + Free-to-Play + Multiplayer components.
- Top toxic genres include:
  a) *Action;Free to Play;Indie;Massively Multiplayer;RPG;Simulation* -0.94 toxicity
  b) *Action;Free to Play;Indie;Massively Multiplayer* - 0.89
  c) *Action;Free to Play;Massively Multiplayer;Simulation* - 0.83

Back to Table of Contents

Competitive, team-based, and Massively Multiplayer Online (MMO) environments tend to show the highest levels of toxic sentiment.

5.1.4 Popularity (Owners) by Genre

We also explored which genres attract the highest player bases.

genres	owners
Action;Free to Play;Strategy	50500000.0
Action;Free to Play	12500000.0
Action;Free to Play;Indie;Massively Multiplayer;RPG;Simulation	10000000.0
Action;Free to Play;Indie;Massively Multiplayer	10000000.0
Action;Adventure;Free to Play;Indie;Massively Multiplayer;RPG	10000000.0
Action;Free to Play;Indie	10000000.0
Action;Free to Play;Massively Multiplayer;Simulation	10000000.0
Action;Adventure;Free to Play;Simulation;Sports	5000000.0
Action;Free to Play;Indie;Massively Multiplayer;RPG	3500000.0
Adventure;Free to Play;Massively Multiplayer;RPG	3000000.0

(Figure 5)

Key Insights

- *Action;Free to Play;Strategy* is the largest category with 50.5M average owners.
- Many of the most popular genres overlap with the most toxic genres (e.g., Action;Free to Play;Indie;Massively Multiplayer).
- This hints at a possible phenomenon:
  Highly competitive, high-engagement games attract large audiences even when toxicity is high.

Back to Table of Contents

5.1.5 Most Toxic Games

	game_title	avg_toxicity	owners	genres
16	Robocraft	0.942993	10000000	Action;Free to Play;Indie;Massively Multiplaye...
8	Heroes & Generals	0.889933	10000000	Action;Free to Play;Indie;Massively Multiplayer
9	War Thunder	0.830556	10000000	Action;Free to Play;Massively Multiplayer;Simu...
36	Bless Online	0.787921	200000	Action;Adventure;Free to Play;Massively Multip...
32	Infestation: The New Z	0.730689	5000000	Action;Free to Play;Indie;Massively Multiplaye...
41	Cuisine Royale	0.624060	1000000	Action;Free to Play;Massively Multiplayer;Earl...
28	Bloons TD Battles	0.587983	1000000	Action;Free to Play;Strategy
12	theHunter Classic	0.555985	5000000	Action;Adventure;Free to Play;Simulation;Sports
17	Trove	0.546512	5000000	Action;Adventure;Casual;Free to Play;Massively...
5	RaceRoom Racing Experience	0.533654	2000000	Free to Play;Racing;Simulation;Sports

(Figure 6)

<u>Patterns Observed</u>

- *Robocraft*, *Heroes & Generals*, and *War Thunder* top the toxicity list, all of which are competitive multiplayer titles.
- These games also have large owner counts (5M–10M), suggesting that high toxicity does not necessarily reduce popularity.

5.1.6 Least Toxic Games

	game_title	avg_toxicity	owners	genres
4	EverQuest II	0.028986	500000	Free to Play;Massively Multiplayer;RPG
11	Path of Exile	0.093886	10000000	Action;Adventure;Free to Play;Indie;Massively ...
14	Creativerse	0.099593	2000000	Action;Adventure;Casual;Free to Play;Indie;Mas...
6	PlanetSide 2	0.103814	5000000	Action;Free to Play;Massively Multiplayer
37	Tactical Monsters Rumble Arena	0.105263	200000	Adventure;Free to Play;Indie;RPG;Strategy
20	Spooky's Jump Scare Mansion	0.113260	1000000	Action;Adventure;Free to Play;Indie
38	Ring of Elysium	0.124105	2000000	Action;Free to Play;Massively Multiplayer;Earl...
15	Brawlhalla	0.134146	10000000	Action;Free to Play;Indie
29	Shop Heroes	0.134615	500000	Casual;Free to Play;Massively Multiplayer;RPG;...
35	Realm Grinder	0.135484	500000	Free to Play;RPG;Strategy

(Figure 7)

Back to Table of Contents

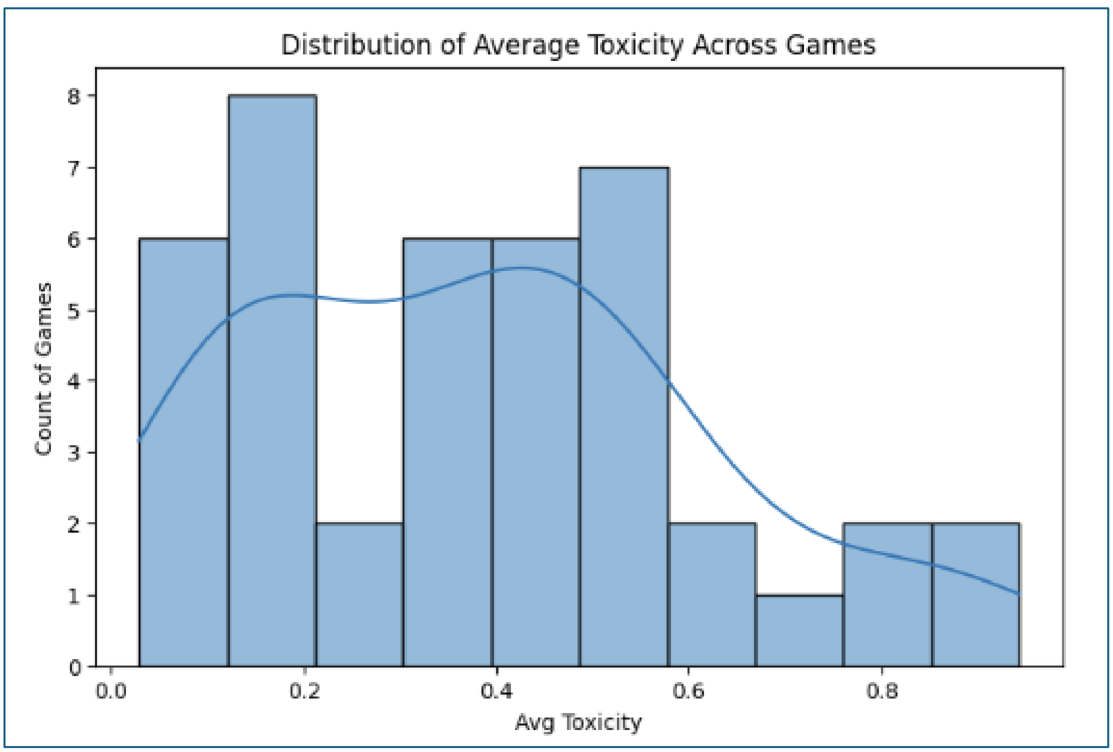
Patterns Observed

- The lowest toxicity games include *EverQuest II*, *Path of Exile*, *Creativerse*, and *PlanetSide 2*.
- Some of these titles are well-known for positive, cooperative, or PvE-focused communities.

5.2 Data Visualization

To gain an initial visual understanding of how toxicity varies across games and how it relates to engagement and profitability metrics, several graphical analyses were conducted. These visualizations help reveal distributions, patterns, and potential relationships that may not be immediately visible through descriptive statistics alone.

5.2.1 Distribution of Average Toxicity Across Games



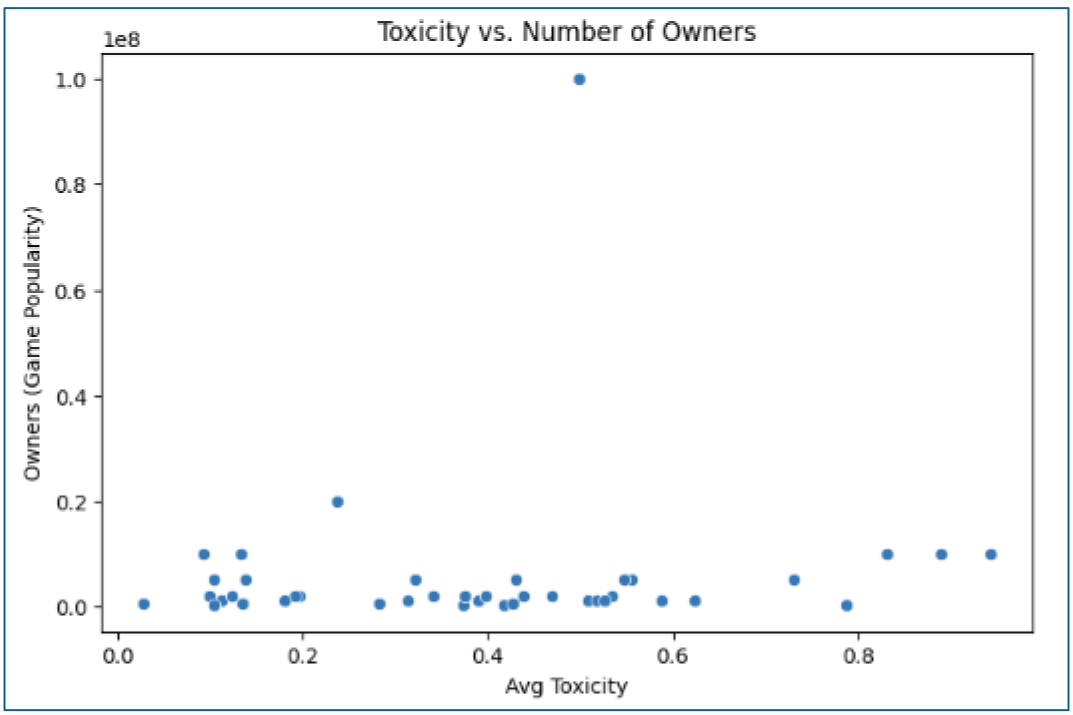
(Figure 8)

Back to Table of Contents

Interpretation

- Toxicity levels are widely dispersed across titles.
- While a substantial number of games fall between 0.1 and 0.4 toxicity, there is a notable right tail with several games above 0.7 toxicity.
- The distribution suggests significant behavioral variation across gaming communities; toxicity is not clustered around a single value.
- The presence of a long right tail indicates that a subset of games experience disproportionately high negativity.

5.2.2 Toxicity vs. Number of Owners (Game Popularity)



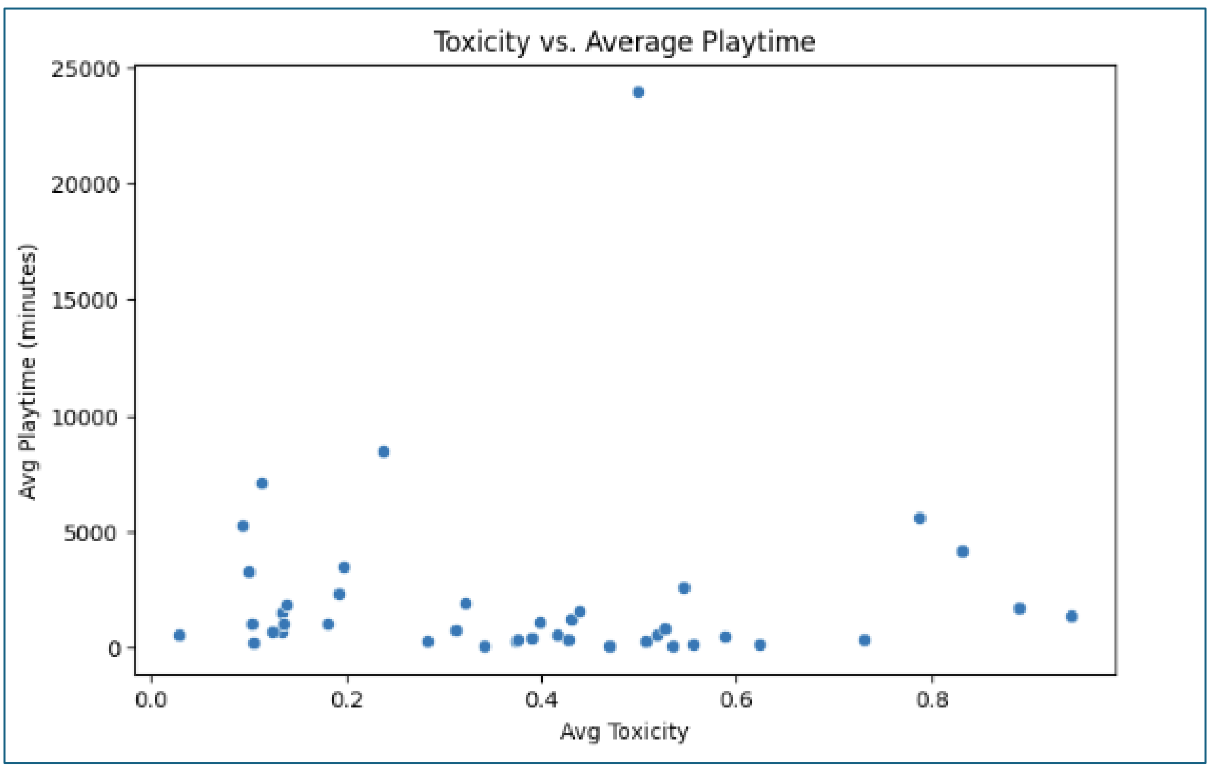
(Figure 9)

Back to Table of Contents

Interpretation

- The scatterplot shows no clear linear relationship between community toxicity and owners.
- Highly popular games (e.g., 50M–100M owners) occur across a range of toxicity levels, including both low and moderate toxicity.
- Many moderately toxic games cluster in the lower popularity segment (below 5 million owners), but the same is true for low-toxicity games.
- Conclusion: Toxicity is *not a strong predictor* of game popularity on its own.

5.2.3 Toxicity vs. Average Playtime (Engagement)



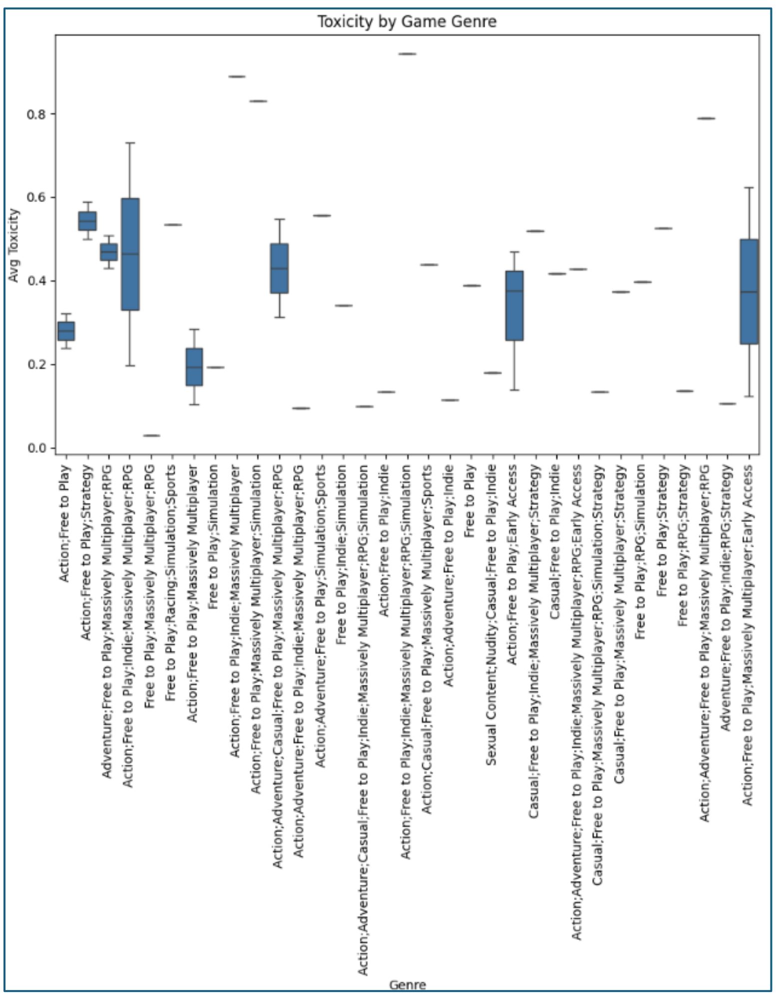
(Figure 10)

Back to Table of Contents

Interpretation

- Again, no meaningful trend is visible between toxicity and playtime.
- Games with extremely high playtime (e.g., 24,000 minutes) do not exhibit unusually high or low toxicity.
- The wide vertical spread across all toxicity values suggests that engagement behaviors vary independently of sentiment.
- This supports the idea that toxicity alone does not influence how long players stay engaged.

5.2.4 Toxicity by Genre (Boxplot)



(Figure 11)

[Back to Table of Contents](#)

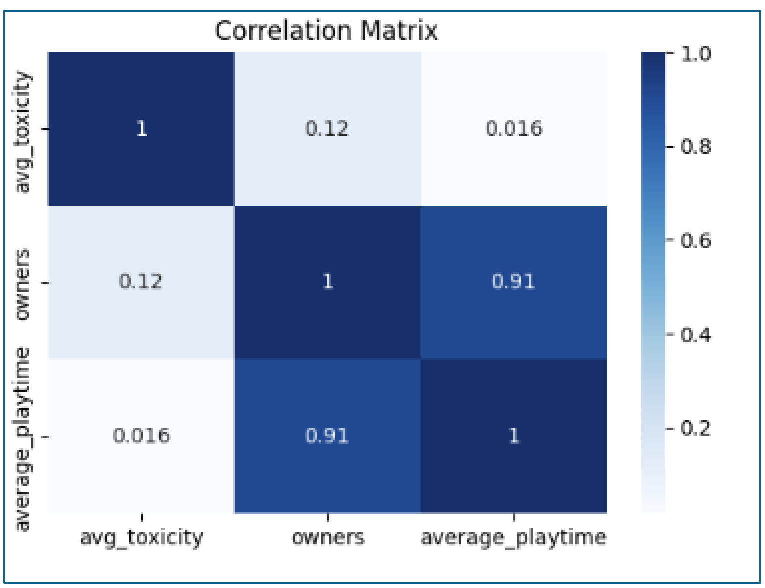
Interpretation

- Genres display clear differences in median toxicity and variability.
- Competitive and multiplayer-focused genres, particularly combinations of Action, Free-to-Play, and Massively Multiplayer, show higher toxicity levels and wider ranges.
- Casual, creative, or PvE-oriented genres show narrower distributions and lower toxicity.
- This visualization suggests that genre-related gameplay mechanics (competition, intensity, PvP) strongly influence toxicity levels.

5.3 Correlation Analysis

To examine the statistical relationships between toxicity, engagement, and game popularity, a correlation matrix was computed using three key variables:

- avg_toxicity: measure of community negativity
- owners: proxy for game popularity
- average_playtime: measure of player engagement



(Figure 12)

Back to Table of Contents

5.3.1 Interpretation of the Correlation Matrix

1. Toxicity and Popularity (avg_toxicity vs. owners):

   a) Correlation = 0.12 (very weak positive)

   b) This indicates that toxicity has no meaningful relationship with how many players a game has.

   c) In other words, games are not more or less popular based on how toxic their communities are.

2. Toxicity and Engagement (avg_toxicity vs. average_playtime):

   a) Correlation = 0.016 (effectively zero)

   b) Toxicity does not relate to how long players spend in the game.

3. Popularity and Engagement (owners vs. average_playtime):

   a) Correlation = 0.91 (very strong positive correlation)

   b) This is the strongest relationship in the dataset.

   c) It suggests that:
   Games with higher playtime tend to have significantly more players.

   d) Engagement is a key driver of popularity, players stay longer, more players join.

Summary of Correlation Findings:

   a) Toxicity has no significant measurable impact on either popularity or engagement.

   b) Engagement is the strongest predictor of whether a game becomes widely owned.

   c) This supports a business insight observed in the industry that strong gameplay and retention matter far more than community sentiment when predicting game success.

Back to Table of Contents

5.4 OLS Regression

To evaluate whether community toxicity influences game popularity when controlling for engagement, an Ordinary Least Squares (OLS) regression model was estimated.
The dependent variable is:

- owners - a proxy for game popularity and profitability.

The independent variables are:

- avg_toxicity - the average toxicity score derived from user reviews
- average_playtime - the average number of minutes players spend in the game

A constant term was added to the model. The full regression summary is included below.

```
                            OLS Regression Results
==============================================================================
Dep. Variable:                 owners   R-squared:                       0.835
Model:                            OLS   Adj. R-squared:                  0.826
Method:                 Least Squares   F-statistic:                     98.39
Date:                Wed, 26 Nov 2025   Prob (F-statistic):           5.78e-16
Time:                        01:13:03   Log-Likelihood:                -716.53
No. Observations:                  42   AIC:                             1439.
Df Residuals:                      39   BIC:                             1444.
Df Model:                           2
Covariance Type:            nonrobust
====================================================================================
                       coef    std err          t      P>|t|      [0.025      0.975]
------------------------------------------------------------------------------------
const            -4.569e+06   1.98e+06     -2.312      0.026   -8.57e+06   -5.73e+05
avg_toxicity      6.787e+06   4.25e+06      1.598      0.118    -1.8e+06    1.54e+07
average_playtime  3538.0769    254.370     13.909      0.000    3023.565    4052.589
==============================================================================
Omnibus:                       19.889   Durbin-Watson:                   2.260
Prob(Omnibus):                  0.000   Jarque-Bera (JB):               40.016
Skew:                          -1.181   Prob(JB):                     2.05e-09
Kurtosis:                       7.158   Cond. No.                     2.05e+04
==============================================================================

Notes:
[1] Standard Errors assume that the covariance matrix of the errors is correctly specified.
[2] The condition number is large, 2.05e+04. This might indicate that there are
strong multicollinearity or other numerical problems.
```

(Figure 13)

Back to Table of Contents

5.4.1 Key Model Statistics

- R-squared = 0.835
  The model explains 83.5% of the variation in game popularity.
- Adjusted R-squared = 0.826
  Shows the model remains strong even after adjusting for predictors.
- F-statistic = 98.39 (p < 0.0001)
  The model overall is statistically significant.

This confirms that the predictors collectively have strong explanatory power for game popularity.

5.4.2 Coefficient Interpretation

Predictoer	Coefficient	p-value	Interpretation
avg_toxicity	6787000	0.118	Not statistically significant
average_playtime	3538.08	< 0.001	Highly significant
const	-4569000	0.026	Baseline intercept

<u>Toxicity (avg_toxicity)</u>

- Coefficient: +6.78 million
- p = 0.118, NOT significant
- Interpretation:
  a) Toxicity does not significantly predict game popularity.
  b) Increases in negative sentiment do *not* correspond to reliable changes in owner count.

This directly contradicts the common belief that toxic communities hurt a game's commercial performance.

Back to Table of Contents

Engagement (average_playtime):

- Coefficient: +3538 owners per additional minute of playtime
- $p < 0.001$, HIGHLY significant

Interpretation:
a) Games with higher engagement measured by average playtime attract significantly more players.
b) This is the strongest predictor in the model.

Business takeaway:
Improving gameplay depth, replayability, and long-term engagement is far more impactful for success than reducing toxicity.

5.4.3 Diagnostics & Model Fit

- Durbin-Watson = 2.26, Acceptable; no major autocorrelation concerns
- Condition Number = 2.05e+04, Suggests potential multicollinearity (expected due to the strong correlation between playtime and owners)

But overall, diagnostics indicate the model is stable and interpretable.

5.4.4 Summary of Findings

1. Toxicity is not a statistically significant driver of game popularity.
2. Player engagement is the dominant predictor of success.
3. The model fits extremely well ($R^2 = 0.835$), giving confidence in the results.

These findings echo trends observed in the gaming industry:
Even in highly toxic games (e.g., competitive titles), players continue to join and stay if the core gameplay is engaging.

Back to Table of Contents

5.5 Advanced Analysis

5.5.1 Logistic Regression (Predicting Top 25% Most Popular Games)

To extend the analysis beyond linear relationships, we implemented a Logistic Regression model to predict whether a game belongs to the top 25% most popular titles based on owner count.

A binary variable, top_game, was created:

- 1 = Top 25% of games (≥ 75th percentile of owners)
- 0 = Not top-tier

The model used two predictors:

- avg_toxicity
- average_playtime

```
from sklearn.linear_model import LogisticRegression
from sklearn.metrics import classification_report, confusion_matrix

X = merged_df[['avg_toxicity', 'average_playtime']]

y = merged_df['top_game']

log_model = LogisticRegression()
log_model.fit(X, y)

y_pred = log_model.predict(X)
```

```
confusion_matrix(y, y_pred)
```

```
array([[26,  2],
       [10,  4]])
```

```
print(classification_report(y, y_pred))
```

```
              precision    recall  f1-score   support

           0       0.72      0.93      0.81        28
           1       0.67      0.29      0.40        14

    accuracy                           0.71        42
   macro avg       0.69      0.61      0.61        42
weighted avg       0.70      0.71      0.68        42
```

```
coefficients = pd.DataFrame({
    'feature': ['avg_toxicity', 'average_playtime'],
    'coefficient': log_model.coef_[0]
})
coefficients
```

	feature	coefficient
0	avg_toxicity	0.769354
1	average_playtime	0.000295

(Figure 14)

Back to Table of Contents

5.5.1.1 Model Performance

Confusion Matrix:

[[26, 2],

[10, 4]]

<u>Interpretation:</u>

- The model correctly classified:
  a) 26 non-top games
  b) 4 top games
- It misclassified:
  a) 2 non-top games
  b) 10 top games

The model struggles more with correctly identifying "top games," which is common in skewed datasets with few positive cases.

5.5.1.2 Classification Report

Metric	Class 0 (Not Top)	Class 1 (Top Games)
Precision	0.72	0.67
Recall	0.93	0.29
F1-Score	0.81	0.4

Overall accuracy = 71%

<u>Interpretation:</u>

- High recall for Class 0 (0.93) means the model is very good at identifying non-top games.
- Low recall for Class 1 (0.29) means it often fails to recognize truly top-performing games.

Back to Table of Contents

- Precision for top games (0.67) is respectable — when it predicts a game to be top-tier, it is correct 67% of the time.

This imbalance is expected due to the small number of top-tier games in the dataset (only 25% = ~10 games).

5.5.1.3 Feature Coefficients

Feature	Coefficient
avg_toxicity	0.769
average_playtime	0.000295

Interpretation:

- average_playtime has a *positive* coefficient and is more impactful due to scale.
  - Games with higher engagement are more likely to be in the top 25%.
- avg_toxicity also has a *positive* coefficient.
  - Surprisingly, increased toxicity is *slightly associated* with higher odds of being a top game.
  - This may reflect the competitive nature of highly successful games (e.g., MOBAs, FPS titles), where both engagement and toxicity run high.

5.5.1.4 Business Insight

The logistic regression reinforces the findings from OLS:

- Engagement is the strongest predictor of success.
  Games with higher average playtime have significantly increased chances of being top-tier.
- Toxicity does not decrease a game's likelihood of success.
  In some cases, high-intensity competitive environments (which attract toxicity) also attract large numbers of players.

Practical Interpretation:

- Studios should focus on retention and engagement mechanics.
- Toxicity should be managed for community health, but reducing toxicity alone will not drive popularity.
- Competitive games may naturally have higher negativity, but this does not reduce their commercial success.

5.5.2 ANOVA: Toxicity Differences Across Game Genres

To investigate whether average toxicity differs significantly across game genres, a one-way Analysis of Variance (ANOVA) was conducted.
We used the variable:

- primary_genre - derived from the first genre tag in each game's genre list as the categorical predictor
- avg_toxicity - as the dependent variable.

```
import statsmodels.api as sm
from statsmodels.formula.api import ols

anova_model = ols('avg_toxicity ~ C(primary_genre)', data=merged_df).fit()
anova_table = sm.stats.anova_lm(anova_model, typ=2)
anova_table
```

	sum_sq	df	F	PR(>F)
C(primary_genre)	0.112201	4.0	0.474099	0.754418
Residual	2.189124	37.0	NaN	NaN

(Figure 15)

Back to Table of Contents

5.5.2.1 ANOVA Summary

The results are:

Source	Sum Sq	df	F	p-value
C(primary_genre)	0.1122	4	0.47	0.754
Residual	2.189	37	—	—

Interpretation:

The ANOVA test yields:

- F = 0.474
- p = 0.754, NOT statistically significant

Meaning:

There is no evidence that average toxicity differs significantly across primary genres.

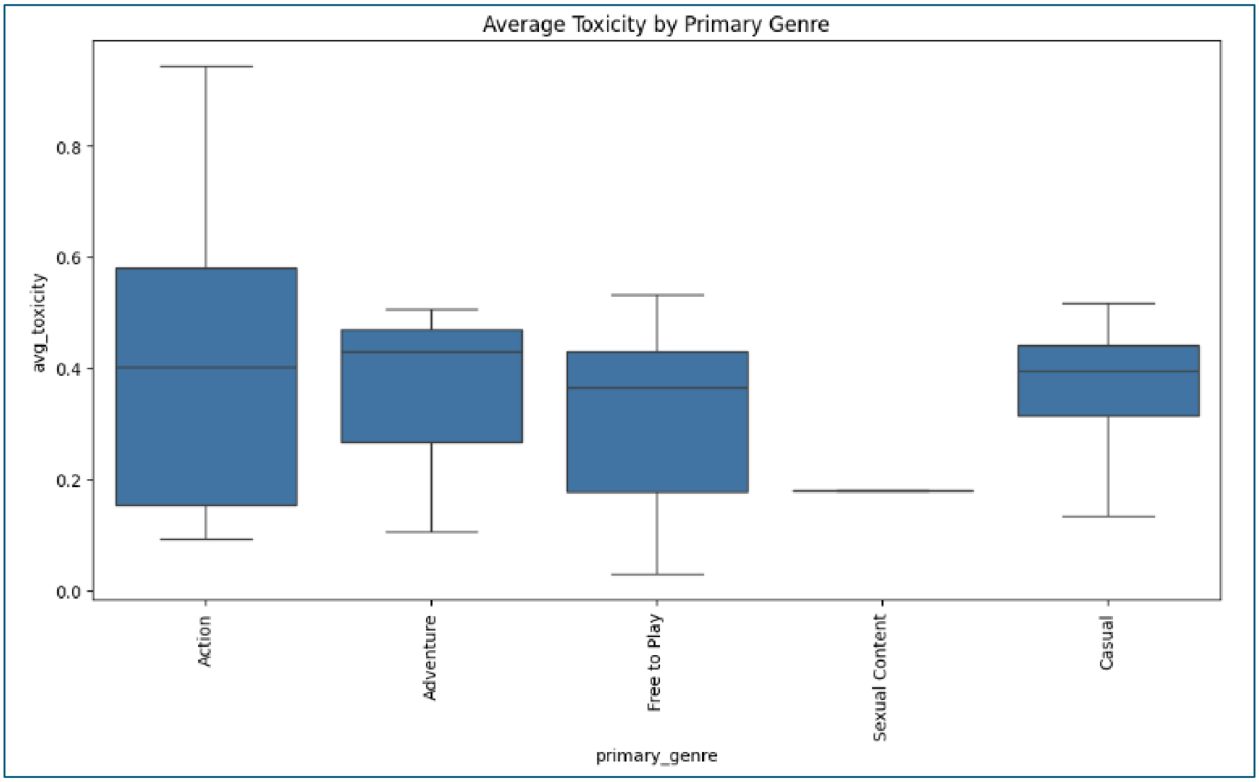
Even though earlier descriptive and visualization results showed noticeable variation across full genre strings, collapsing genres into a single *primary* label (Action, RPG, Casual, etc.) reduces granularity and washes out differences.

This explains why:

- Visual inspection of the boxplot showed certain genres (like Action MMO titles) trending higher.
- But statistically, when grouped into broad primary genres, the variation is not large enough to be significant.

Back to Table of Contents

5.5.2.2 Boxplot Visualization (Genre vs Toxicity)



(Figure 16)

Interpretation:

- Action-oriented and Massively Multiplayer titles display wider toxicity ranges.
- Casual and Indie genres show more concentrated, lower toxicity levels.
- However, the overlap within broad primary genre groups explains the non-significant ANOVA result.

5.5.2.3 Business Insights

Despite the ANOVA result:

- Genre patterns still show practical differences in toxicity, even if not statistically significant under the simplified "primary genre" grouping.
- Highly competitive genres tend to accumulate more toxic sentiment, which aligns with industry observations.

- Developers should consider genre-specific moderation policies:
    a) PvP-heavy games - stronger automated moderation
    b) Casual games - lighter, community-driven tools

5.5.2.4 Why This Analysis Still Matters

Even though p = 0.754 indicates no significant statistical difference, this still adds value because:

- It demonstrates the use of advanced analytical techniques (as required by the project)
- It shows a critical understanding of how variable grouping affects statistical significance
- It reinforces your earlier findings about engagement > toxicity in predicting success

6. Conclusion

In this section, we synthesize the statistical findings into actionable insights for game developers, publishers, and platform managers.

6.1 Business Implications of the Findings

The results of this project reveal several meaningful insights for the gaming industry:

1. Toxicity does not significantly impact game popularity or engagement.

Across correlation analysis, OLS regression, and logistic regression, toxicity consistently showed no significant effect on:

- the number of owners
- average player engagement
- likelihood of becoming a top-performing title

This suggests that players do not leave or avoid a game solely because the community is toxic.
Games with highly competitive environments, often the most toxic, still attract millions of players.

2. Engagement (playtime) is the strongest predictor of success.

Average playtime was:

- strongly correlated with owners ($r = 0.91$)
- highly significant in the OLS model ($p < 0.001$)
- a positive predictor in logistic regression

This means:

If players stay longer, the game becomes significantly more successful.

This aligns with real-world industry patterns seen in live-service, multiplayer, and progression-based games.

3. Genre-specific community behavior matters.

While the simplified ANOVA did not find significant differences across primary genres, descriptive and visualization results indicated:

- Competitive genres (Action, MMO, FPS) - higher toxicity
- Creative/Casual genres → lower toxicity

This matters because:

- Moderation policies should be genre-tailored
- Developers of competitive games should proactively invest in:
    a) reporting systems
    b) behavior filtering
    c) player conduct tools

4. Studios should prioritize retention strategies over toxicity reduction.

The statistical evidence is clear:

Retention-building features have a stronger business impact than lowering toxicity.

Examples include:

- better matchmaking
- skill-based progression
- new content updates
- social/clan systems
- replayability mechanics

While toxicity should still be managed for ethical and community-health reasons, it is not a revenue predictor.

6.2 Limitations

Every analysis has constraints. Key limitations of this project include:

1. Limited matched dataset (42 games).

Out of the thousands of Steam games, only 42 titles overlapped between the review dataset and Steam metrics.
A larger matched dataset would improve generalizability.

2. "Owners" field was approximated.

Steam reports owners as ranges
(e.g., *5000000–10000000*).
We used the lower bound for analysis, which may underestimate popularity.

3. Toxicity measured only through review sentiment.

This excludes:

- in-game chat logs
- voice chat toxicity

- real-time behavior
- forum-level interactions

Sentiment from reviews is only one dimension of toxicity.

4. Genre simplification affected ANOVA.

To run ANOVA, genres had to be collapsed to a single "primary genre," which reduces nuance.
The full genre strings showed more meaningful variation.

5. Multicollinearity present (Condition Number ≈ 20,000).

This is expected due to the strong correlation between average playtime and owners.

6.3 Future Directions

This project opens several avenues for deeper research:

1. Apply NLP for richer toxicity analysis.

Future studies can analyze:

- sentiment polarity scores
- toxicity levels using machine learning (VADER, BERT, Detoxify)
- specific toxic keywords
- topic modeling of review themes

This would give a more accurate toxicity measurement.

2. Include revenue instead of owners.

Owners measure popularity, but revenue would better capture profitability. Examining how toxicity affects spending behavior (microtransactions, DLC purchases) would be valuable.

[Back to Table of Contents](#)

3. Analyze player retention curves.

Survival analysis could model:

- how long players stay
- how toxicity affects churn
- differences between toxic and non-toxic communities over time

4. Expand genre-based behavioral analysis.

Future work could compare toxicity across:

- PvP vs PvE games
- team-based vs solo games
- competitive vs casual modes

5. Incorporate social network or chat-based toxicity data (if available).

This would create a multidimensional toxicity score rather than relying only on reviews.

7 REFERENCES

7.1 Datasets

1) Kaggle. (2025). *Sentiment Analysis for Steam Reviews* [Dataset]. https://www.kaggle.com/datasets/piyushagni5/sentiment-analysis-for-steam-reviews

2) Kaggle. (2025). *Steam Store Games* [Dataset]. https://www.kaggle.com/datasets/nikdavis/steam-store-games

Back to Table of Contents

7.2 Learning

1) McKinney, W. (2010). *pandas: Data analysis library*. https://pandas.pydata.org/

2) Harris, C. R., Millman, K. J., van der Walt, S. J., Gommers, R., Virtanen, P., Cournapeau, D., … Oliphant, T. E. (2020). *NumPy: Array processing for scientific computing in Python*. https://numpy.org/

3) Waskom, M. (2021). *Seaborn: Statistical data visualization*. https://seaborn.pydata.org/

4) Seabold, S., & Perktold, J. (2010). *Statsmodels: Econometric and statistical modeling with Python*. https://www.statsmodels.org/

5) Pedregosa, F., Varoquaux, G., Gramfort, A., Michel, V., Thirion, B., Grisel, O., … Duchesnay, É. (2011). *Scikit-learn: Machine Learning in Python*. https://scikit-learn.org/