



Steam Game Analysis

Analyzing game features, release trends, and key factors behind high-popularity games on Steam

Presented by: Nguyen Thanh Truc

Mentor: Ms Vo Tran Vy Thao and Mr Nguyen Chi Cuong



Table of contents

I

Introduction

IV

Conclusion

II

Cleaning

III

EDA

Index

- Objective
- Expected
- Sources

01

Introduction



I. Introduction

01

Objectives

Analyze release trends and seasonal patterns of games on Steam

Explore relationships between game attributes such as:

- Price
- Popularity Score
- Review Ratio
- Release Year

Identify top-performing genres, developers, and publishers

Evaluate value-for-money through metrics like Popularity per Dollar

Investigate how price correlates with perceived quality and genre competitiveness

I. Introduction

02


Expected outcome

- Discover when (seasonally) games are most likely to succeed
- Understand what features most strongly influence popularity and reviews
- Identify which genres and studios consistently deliver high performance
- Highlight games that offer the best value for money
- Clarify if premium pricing leads to higher quality or market reach
- Find a key to survive in an increasingly competitive and saturated market

I. Introduction

03

Data sources

- Source: Kaggle  ([Steam Games Dataset](#)– fronkongames)
- Description: This dataset has been created with **MIT** code and use the API provided by **Steam**, the largest gaming platform on PC. Data is also collected from **Steam Spy**.
- **Size:** 111,452 rows x 39 columns

I. Introduction

AppID	Unique ID of the game on Steam
Name	Name of the game
Release date	Release date of the game
Estimated owners	Estimated range of game owners
Peak CCU	Peak number of concurrent players
Required age	Minimum age required to play
Price	Price in USD
Discount	Discount percentage
DLC count	Number of DLCs (Downloadable Content)
About the game	Overview/about section from Steam

User score	User rating score (usually 0-100)
Positive	Number of positive reviews
Negative	Number of negative reviews
Score rank	Rank/tier of the game's score
Achievements	Number of Steam achievements
Recommendations	Number of user recommendations
Notes	Additional notes or warnings
Average playtime forever	Average total playtime (in minutes)
Average playtime two weeks	Average playtime over last 2 weeks
Median playtime forever	Median total playtime

Supported languages	Languages supported by the game
Full audio languages	Languages with full audio support
Reviews	User or critic reviews
Header image	URL of the game's header image
Website	Official website of the game
Support url	Support or contact URL
Support email	Support contact email
Windows / Mac / Linux	Supported platforms
Metacritic score	Score on Metacritic (0-100)
Metacritic url	Link to Metacritic page

Median playtime two weeks	Median playtime over last 2 weeks
Developers	Developer(s) of the game
Publishers	Publisher(s) of the game
Categories	Game categories on Steam (e.g., Single-player)
Genres	Genres (e.g., Action, Strategy, RPG)
Tags	User-submitted tags (e.g., Indie, Casual)
Screenshots	List of in-game screenshots (URLs)
Movies	List of gameplay trailers or videos

I. Introduction

Check row dataset

AppID	Name	Release date	Estimated owners	Peak CCU	Required age	Price	Discount	DLC count	About the game	Supported languages	...	Average playtime two weeks	Median playtime forever	Median playtime two weeks	Developers	Publishers	Categories	Genres
20200	Galactic Bowling	Oct 21, 2008	0 - 20000	0	0	19.99	0	0	Galactic Bowling is an exaggerated and stylize...	[English]	...	0	0	0	Perpetual FX Creative	Perpetual FX Creative	Single-player,Multi-player,Steam Achievements,...	Casual,Indie,Sports
655370	Train Bandit	Oct 12, 2017	0 - 20000	0	0	0.99	0	0	THE LAW!! Looks to be a showdown atop a train...	[English', 'French', 'Italian', 'German', 'Sp...	...	0	0	0	Rusty Moyher	Wild Rooster	Single-player,Steam Achievements,Full controll...	Action,Indie
1732930	Jolt Project	Nov 17, 2021	0 - 20000	0	0	4.99	0	0	Jolt Project: The army now has a new robotics ...	[English', 'Portuguese - Brazil]	...	0	0	0	Camplão Games	Camplão Games	Single-player	Action,Adventure,Indie,Strategy

Index

- Basic
- Features
- Outliers

02

Cleaning



II. Cleaning

01

Corrected data misalignments

02

Removed irrelevant columns

✗ Not analyze in this project:

- About the game - Long text data
- Supported languages, Full audio languages - Too detailed, rarely used
- Reviews - not used unless performing sentiment analysis
- Header image, Screenshots, Movies - Media URLs not needed
- Website, Support url, Support email - Contact information, not relevant
- Windows, Mac, Linux - Platform support
- Metacritic score, Metacritic url - Frequently missing
- Achievements - Not a key factor
- Notes - Extra content note, low analytical value
- Score rank, User score - Frequently missing, redundant with Positive and Negative
- Average playtime forever, Average playtime two weeks, Median playtime forever, Median playtime two weeks - Too many redundant metrics
- Tags, Categories - Too messy and overlaps with Genres
- Peak CCU - Popularity metric, not in modeling scope
- Discount - Many missing values and time-sensitive

II. Cleaning

03

Handle 'Estimated owners' column

04

Convert data types

05

Fill missing and check duplicate

Missing Values before Handling:

AppID	0
Name	1
Release date	122
Estimated owners	0
Required age	0
Price	0
DLC count	0
Positive	0
Negative	0
Recommendations	0
Developers	5
Publishers	344
Genres	77
dtype:	int64

```
df_games['AppID'].duplicated().sum()  
np.int64(0)
```

II. Cleaning

06

Filter out

- Non-game entries
- Games released before 2003 or in 2025

07

Feature Engineering

- Total reviews
- Review ratio
- Game age
- Price category
- Popularity score
- Value score
- Popularity quintile
- One hot encoded 'Genres'

II. Cleaning – 08. Outlier

01

IQR

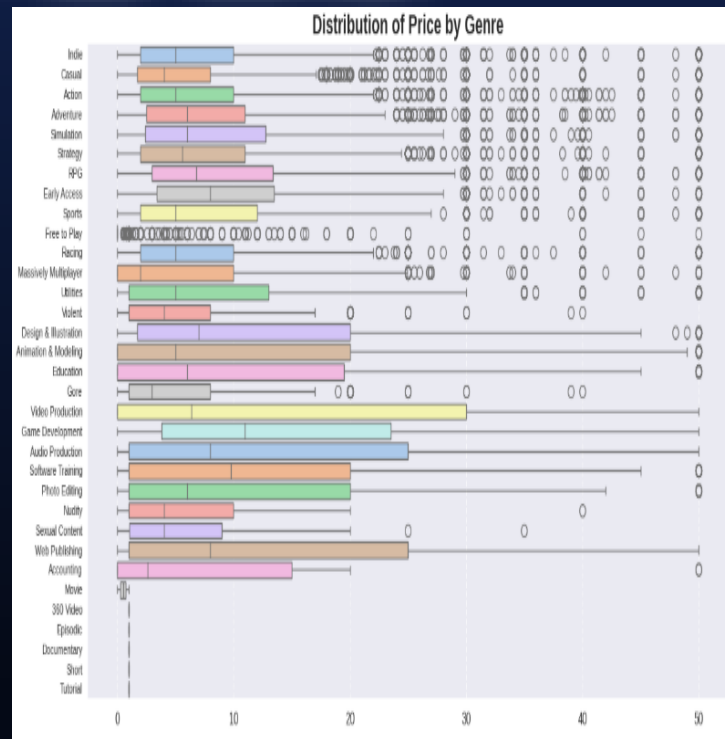
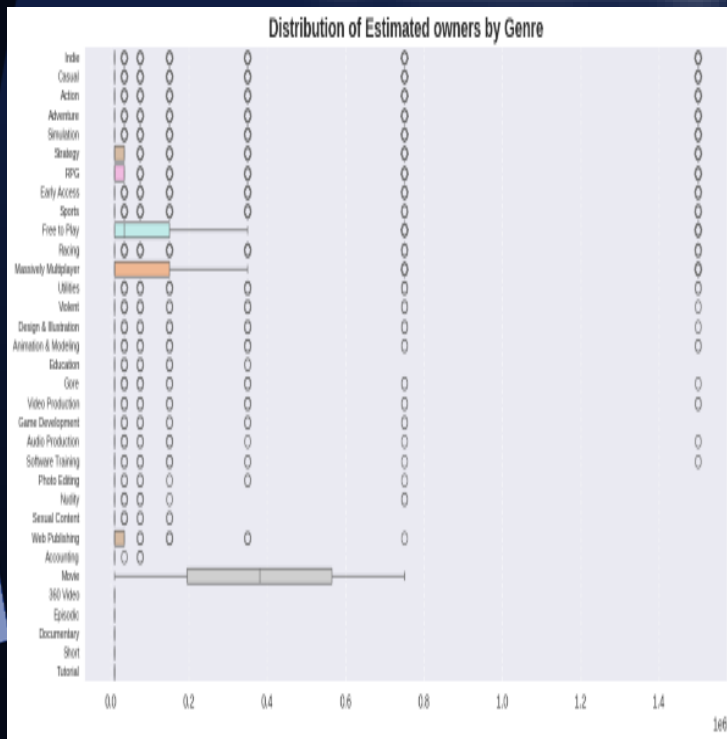
	Column	Q1	Q3	IQR	Lower Bound	Upper Bound
0	Estimated owners	10000.00	10000.00	0.0	10000.0	10000.00
1	Price	1.99	9.99	8.0	0.0	21.99
2	Positive	1.00	51.00	50.0	0.0	126.00
3	Negative	0.00	15.00	15.0	0.0	37.50
4	Recommendations	0.00	0.00	0.0	0.0	0.00
5	Required age	0.00	0.00	0.0	0.0	0.00
6	DLC count	0.00	0.00	0.0	0.0	0.00

02

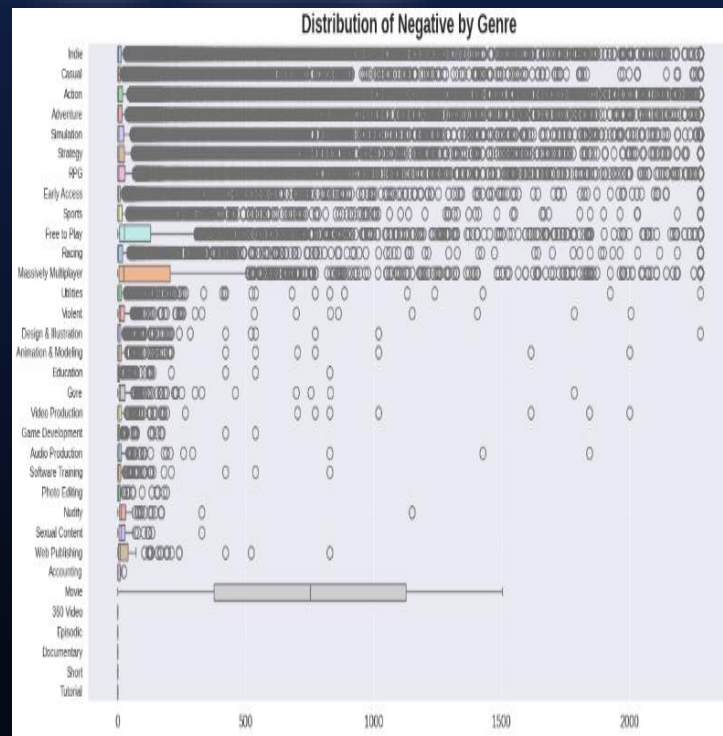
Capping

	Column	99th Percentile Cap	Rows Affected
0	Estimated owners	1500000.00	478
1	Price	49.99	646
2	Positive	12888.00	852
3	Negative	2277.33	853
4	Recommendations	10781.28	853
5	Required age	17.00	333
6	DLC count	6.00	819

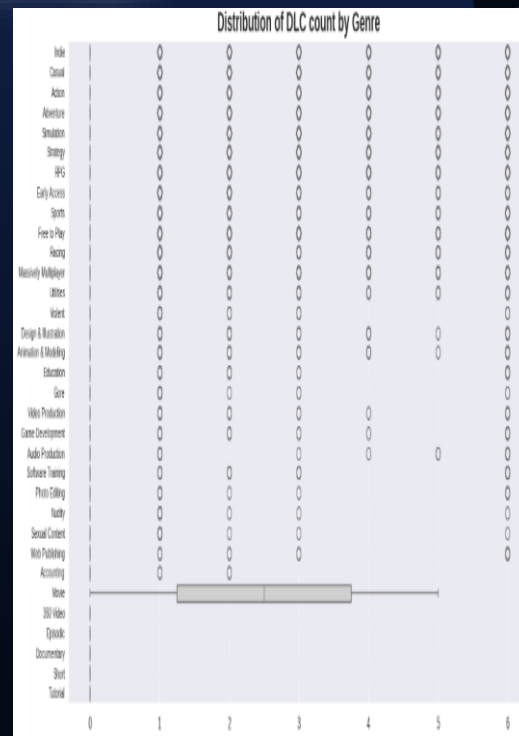
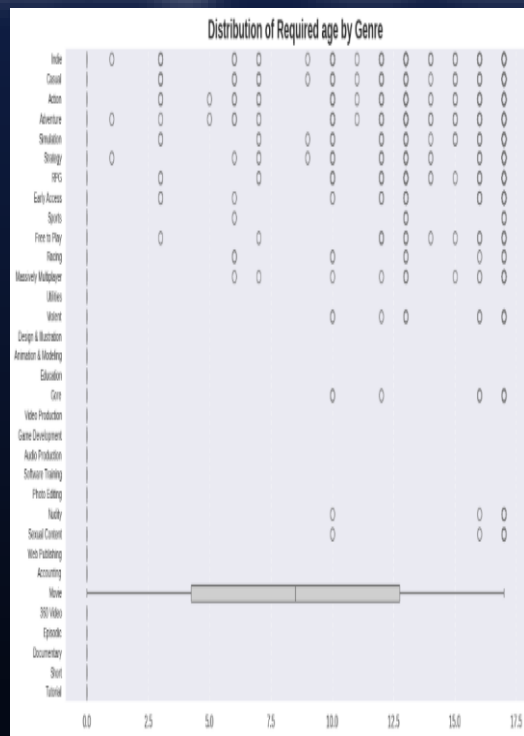
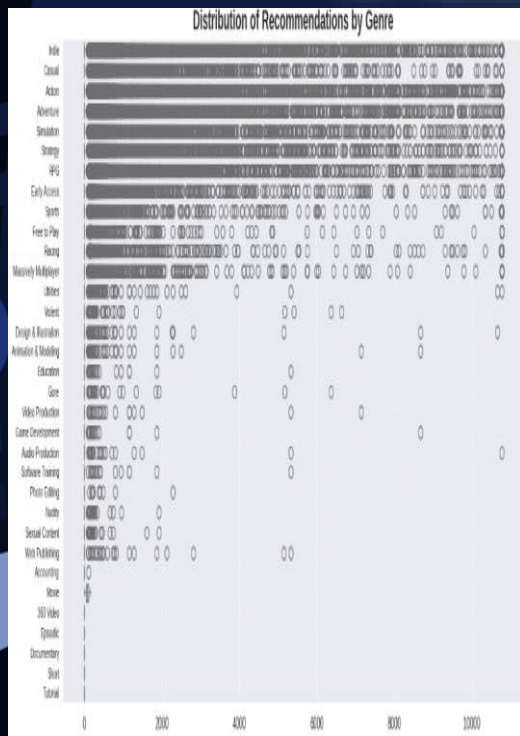
II. Cleaning – 08. Outlier



II. Cleaning – 08. Outlier



II. Cleaning – 08 Outlier



Index

- Overview
- Evolution
- Price and Value
- Genres
- Studios
- Legend Games
- Correlations

03

EDA



III. EDA

Data overview

Total Games: 85,268

Total Features: 54

	AppID	Name	Release date	Estimated owners	Required age	Price	DLC count	Positive	Negative	Recommendations	Developers	Publishers	Release year	Release month	Total reviews	Review ratio	Game age	Price category	Popularity score	Value score	Popularity quintile
0	20200	Galactic Bowling	2008-10-21	10000.0	0.0	19.99	0.0	6.0	11.0	0.0	Perpetual FX Creative	Perpetual FX Creative	2008	10	17	0.352941	17	\$10-30	0.076837	0.001292	High
1	655370	Train Bandit	2017-10-12	10000.0	0.0	0.99	0.0	53.0	5.0	0.0	Rusty Moyher	Wild Rooster	2017	10	58	0.913793	8	Under \$10	0.084167	0.038649	High
2	1732930	Jolt Project	2021-11-17	10000.0	0.0	4.99	0.0	0.0	0.0	0.0	Campião Games	Campião Games	2021	11	0	0.000000	4	Under \$10	0.000000	0.000000	Bottom 20%

0	AppID	85268 non-null	int64	28	Design & Illustration	85268 non-null	int64
1	Name	85268 non-null	string	29	Documentary	85268 non-null	int64
2	Release date	85268 non-null	datetime64[ns]	30	Early Access	85268 non-null	int64
3	Estimated owners	85268 non-null	float64	31	Education	85268 non-null	int64
4	Required age	85268 non-null	float64	32	Episodic	85268 non-null	int64
5	Price	85268 non-null	float64	33	Free to Play	85268 non-null	int64
6	DLC count	85268 non-null	float64	34	Game Development	85268 non-null	int64
7	Positive	85268 non-null	float64	35	Game	85268 non-null	int64
8	Negative	85268 non-null	float64	36	Indie	85268 non-null	int64
9	Recommendations	85268 non-null	float64	37	Massively Multiplayer	85268 non-null	int64
10	Developers	85268 non-null	string	38	Mod	85268 non-null	int64
11	Publishers	85268 non-null	string	39	Moddy	85268 non-null	int64
12	Release year	85268 non-null	int32	40	Photo Editing	85268 non-null	int64
13	Release month	85268 non-null	int32	41	RPG	85268 non-null	int64
14	Total reviews	85268 non-null	int64	42	Racing	85268 non-null	int64
15	Review ratio	85268 non-null	float64	43	Sexual Content	85268 non-null	int64
16	Game age	85268 non-null	int32	44	Short	85268 non-null	int64
17	Price category	85368 non-null	category	45	Simulation	85268 non-null	int64
18	Popularity score	85368 non-null	float64	46	Software Training	85268 non-null	int64
19	Value score	85368 non-null	float64	47	Sports	85268 non-null	int64
20	Popularity quintile	85268 non-null	category	48	Strategy	85268 non-null	int64
21	360 Video	85268 non-null	int64	49	Tutorial	85268 non-null	int64
22	Accounting	85268 non-null	int64	50	Utilities	85268 non-null	int64
23	Action	85268 non-null	int64	51	Video Production	85268 non-null	int64
24	Adventure	85268 non-null	int64	52	Violent	85268 non-null	int64
25	Animation & Modeling	85268 non-null	int64	53	Web Publishing	85268 non-null	int64
26	Audio Production	85268 non-null	int64	dtypes: category(2), datetime64[ns](1), float64(18), int32(1), int64(35), string(3)			
27	Casual	85268 non-null	int64	memory usage: 33.7 MB			

	Estimated owners	Required age	Price	DLC count	Positive	Negative	Recommendations	Total reviews	Review ratio	Game age	Popularity score	Value score
count	85,268.00	85,268.00	85,268.00	85,268.00	85,268.00	85,268.00	85,268.00	85,268.00	85,268.00	85,268.00	85,268.00	85,268.00
mean	53,948.73	0.31	7.82	0.27	350.79	67.17	276.96	1,126.56	0.63	4.93	0.11	0.02
std	183,902.46	2.22	8.60	0.87	1,594.10	283.19	1,321.38	28,099.14	0.36	3.11	0.14	0.05
min	10,000.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	1.00	0.00	0.00
25%	10,000.00	0.00	1.99	0.00	1.00	0.00	0.00	2.00	0.41	2.00	0.03	0.00
50%	10,000.00	0.00	4.99	0.00	9.00	2.00	0.00	12.00	0.76	4.00	0.06	0.01
75%	10,000.00	0.00	9.99	0.00	51.00	15.00	0.00	69.00	0.94	7.00	0.13	0.02
max	1,500,000.00	17.00	49.99	6.00	12,888.00	2,277.33	10,781.28	6,531,097.00	1.00	21.00	1.00	0.92

III. EDA – 01. Overview



Scale

- Total Games: 85,268
- Unique Genres: 33
- Average Games per Genre: 2584

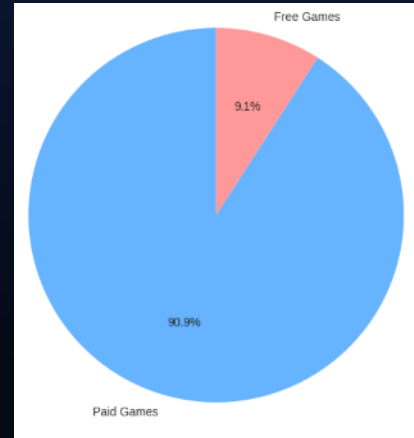


Economic

- Average Price: \$7.82
- Median Price: \$4.99
- Average Reviews: 1,127



Distribution



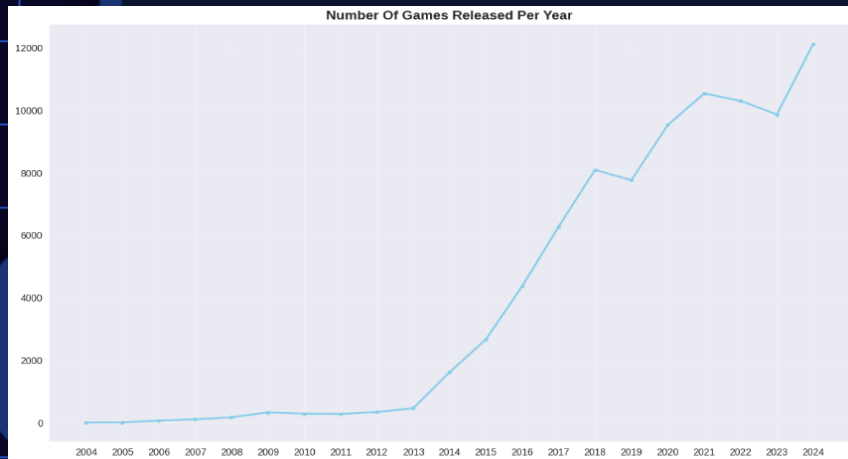
III. EDA – 02. Evolution



Release Trends

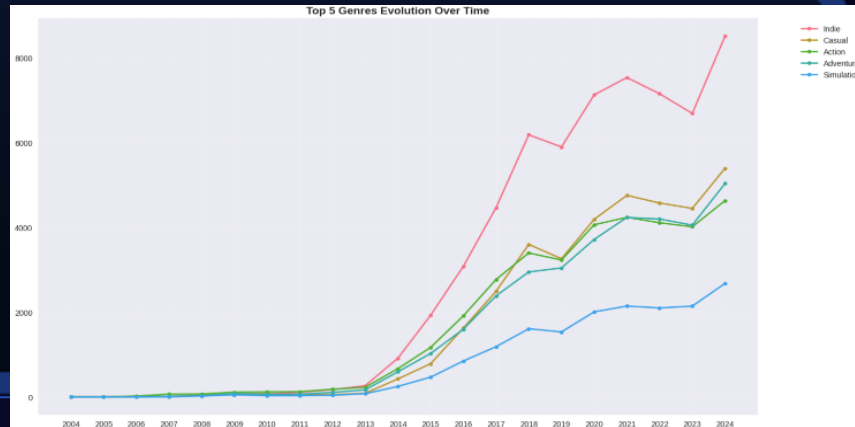
Peak Year: 2024 (12,128 games)

Growth Rate 2014-2024: 650%



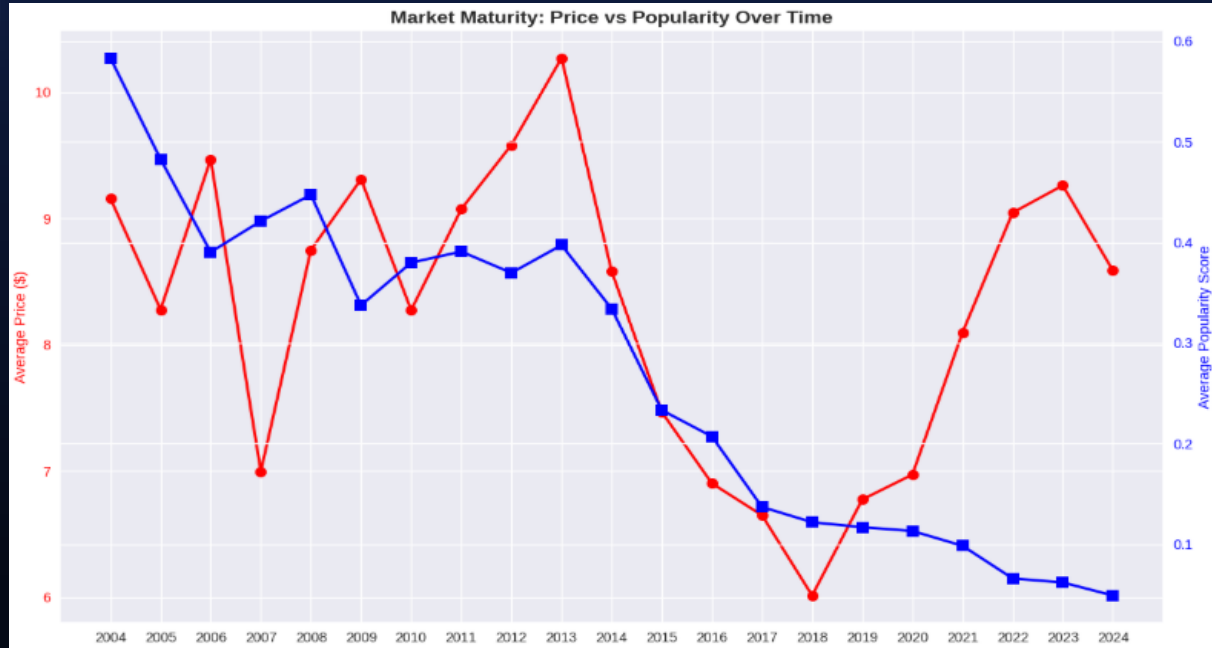
Top genres (2022 – 2024)

- Indie: 60,139
- Casual: 35,912
- Action: 35,088
- Adventure: 33,307
- Simulation: 17,190



III. EDA – 02. Evolution

Market Maturity

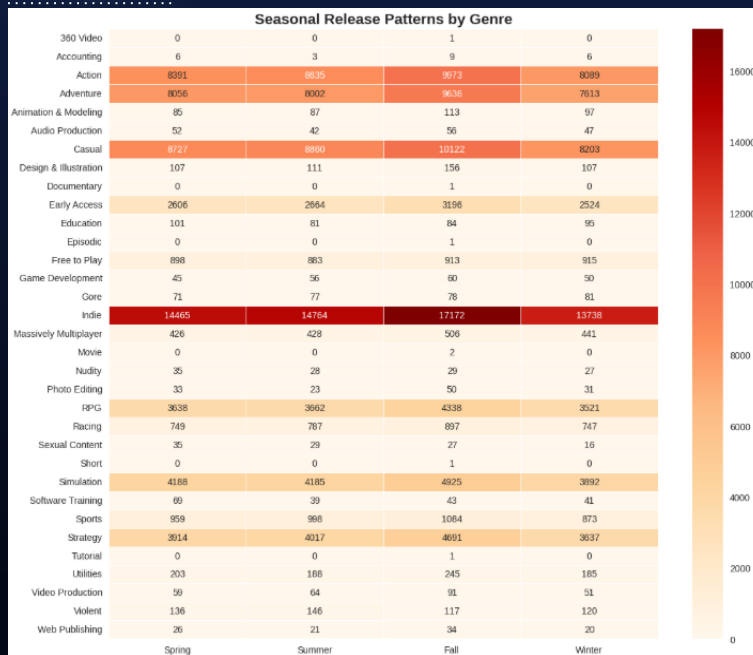


III. EDA – 02. Evolution

Seasonal



Seasonal by Genres



III. EDA – 03. Price and Value



	Price		Popularity score	Total reviews	Review ratio
	mean	count	mean	mean	mean
Price category					
Free	0.00	7764	0.17	3,527.75	0.68
Under \$10	4.76	58295	0.88	312.22	0.61
\$10-30	17.56	17247	0.17	2,060.44	0.68
\$30-60	43.15	1693	0.27	8,523.97	0.66
Over \$60	49.99	269	0.11	1,865.53	0.52



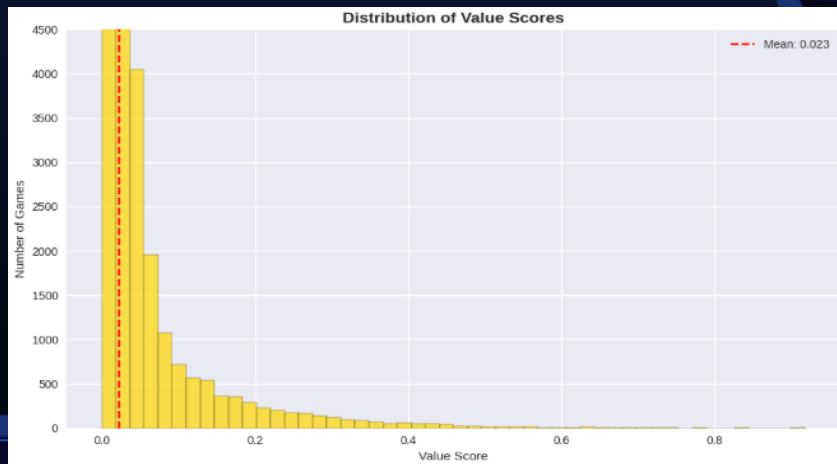
Top Value-for-Money

	Name	Popularity score	Review ratio	Value score
	Grand Theft Auto: San Andreas	1.00	0.92	0.92
	Team Fortress 2	0.97	0.94	0.91
	Counter-Strike: Global Offensive	1.00	0.88	0.88
	Half-Life 2: Lost Coast	1.00	0.88	0.88
	Grand Theft Auto V	1.00	0.85	0.85
	Grand Theft Auto: Vice City	0.89	0.93	0.83
	Destiny 2	0.99	0.84	0.83
	Rocket League®	0.91	0.89	0.81
	Life is Strange - Episode 1	0.83	0.96	0.79
	Dota 2	0.95	0.83	0.79

Average Popularity by Price Category



Distribution of Value Scores



III. EDA – 03. Price and Value

Price vs Quality

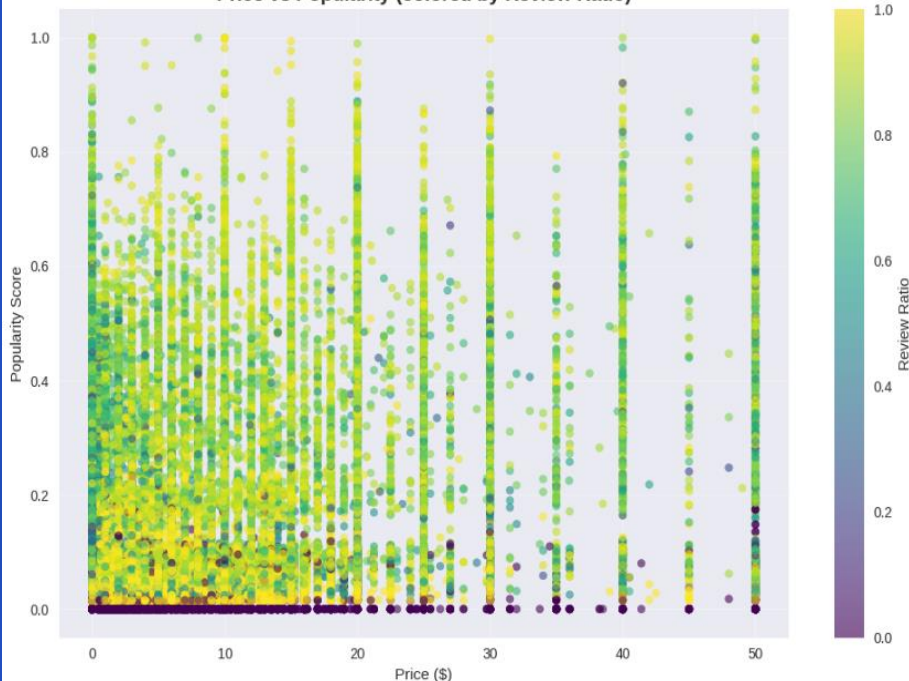


Price - Quality Correlations:
Popularity score 0.25
Review ratio 0.06
Total reviews 0.04



Premium vs Budget

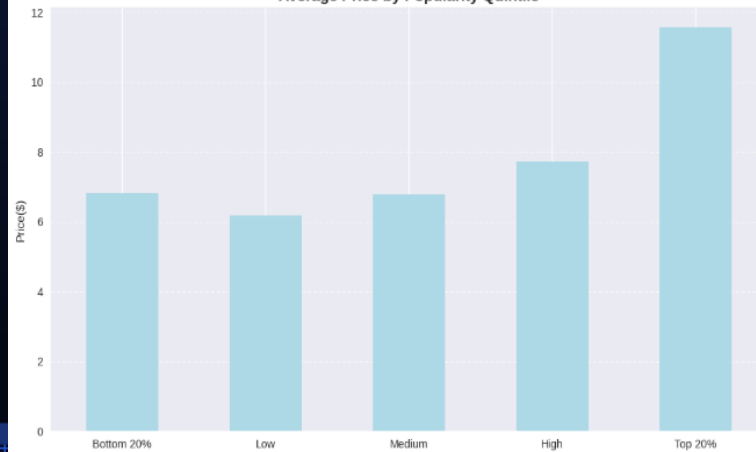
Price vs Popularity (colored by Review Ratio)



Metric Premium Games (>\$30) Budget Games (<=\$30)

Metric	Premium Games (>\$30)	Budget Games (<=\$30)
Game Count	1,962 games	83,306 games
Avg Popularity	0.246	0.110
Success Rate (>0 popularity)	87.2%	83.8%

Average Price by Popularity Quintile

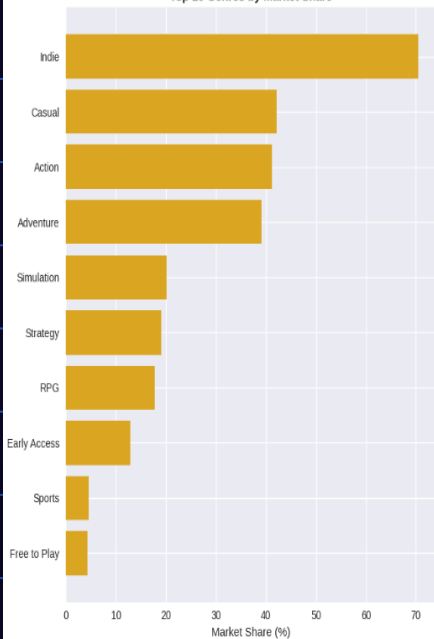


III. EDA – 04. Genres

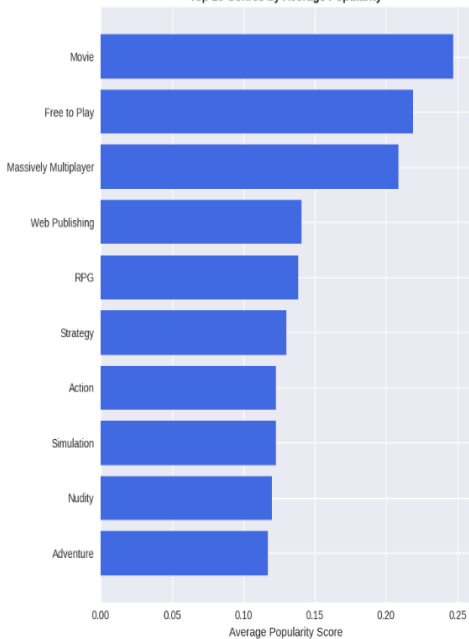


Top genres

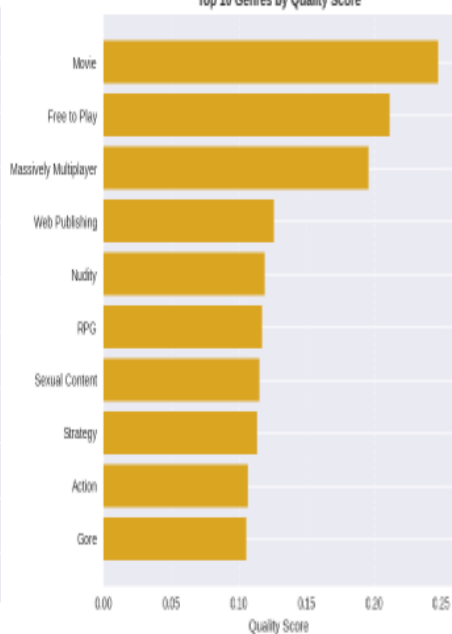
Top 10 Genres by Market Share



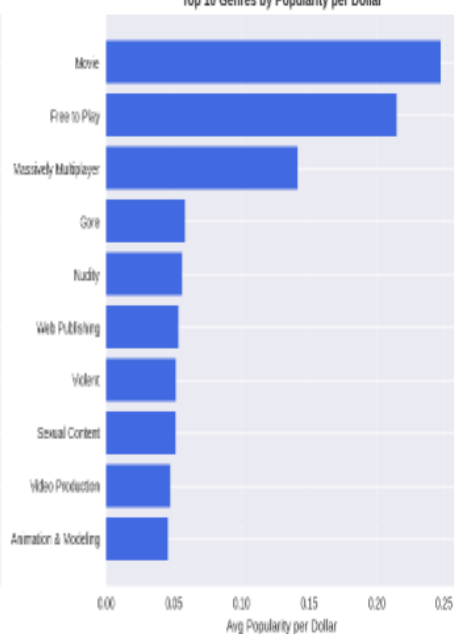
Top 10 Genres by Average Popularity



Top 10 Genres by Quality Score



Top 10 Genres by Popularity per Dollar



III. EDA – 04. Genres

Quantity vs Quality

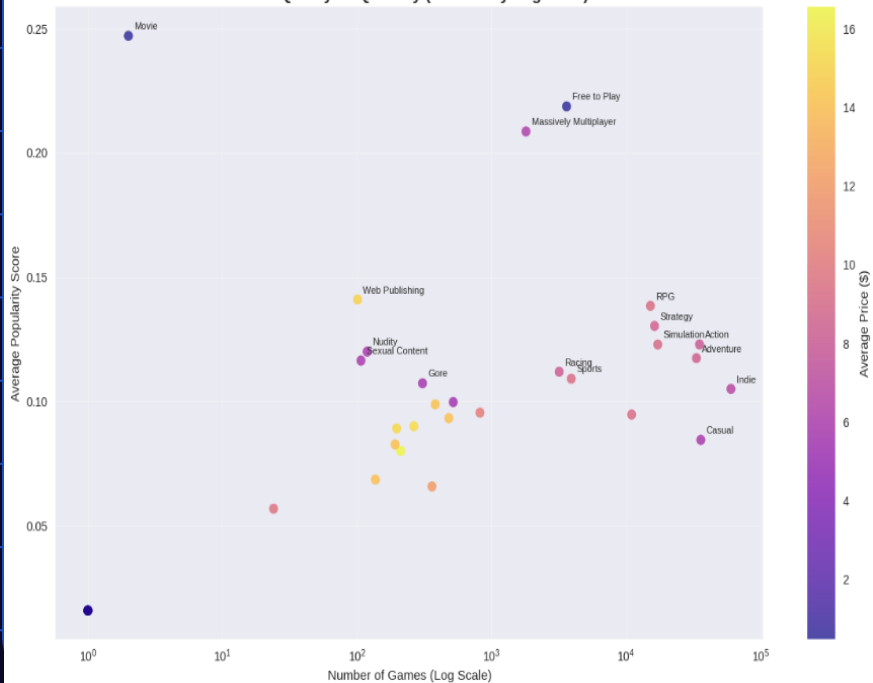


🔥 Most Competitive Genres (High Volume, Low Avg Popularity - over 10% of total games):

	Game count	Avg popularity	Competitiveness
Indie	60,139.00	0.10	572,913.76
Casual	35,912.00	0.08	425,186.77
Action	35,088.00	0.12	285,537.99
Adventure	33,307.00	0.12	283,779.63
Simulation	17,190.00	0.12	139,936.50



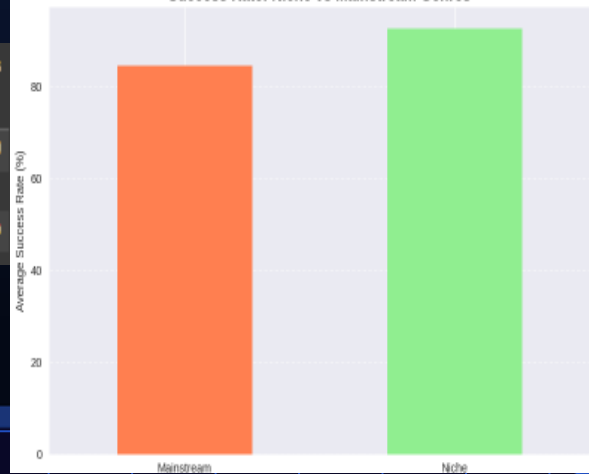
Genre Quality vs Quantity (colored by Avg Price)



Niche Genres Mainstream Genres

Metric	Niche Genres	Mainstream Genres
Genre Count	25.00	8.00
Avg Popularity	0.10	0.11
Avg Success Rate	92.50	84.49

Success Rate: Niche vs Mainstream Genres



III. EDA – 04. Genres

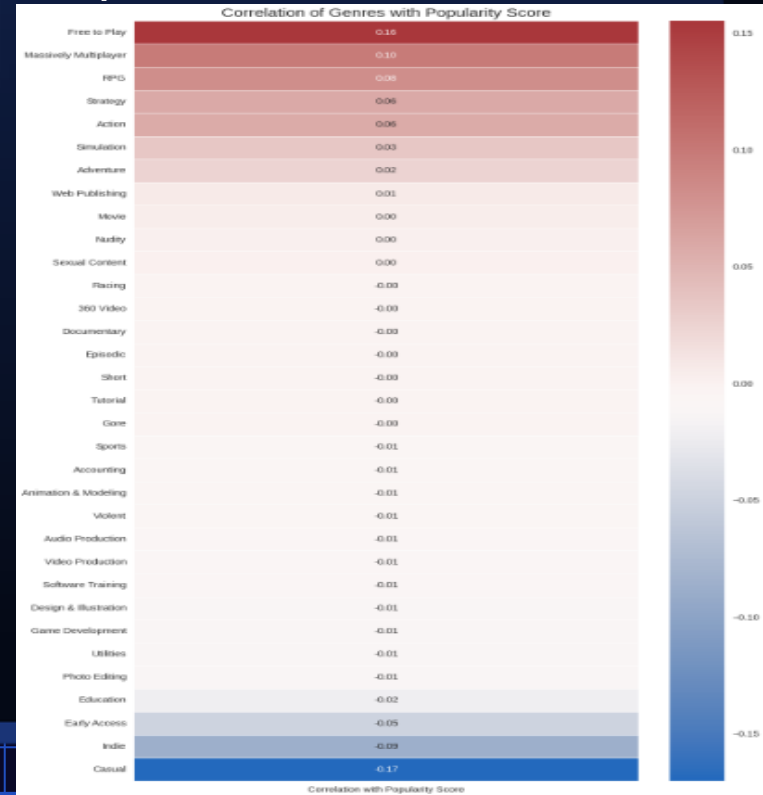
Combinations

Popularity vs Genres



Top 5 Common Genre Combinations:

('Casual', 'Indie'): 27418 games
('Action', 'Indie'): 26578 games
('Adventure', 'Indie'): 25303 games
('Action', 'Adventure'): 15555 games
('Adventure', 'Casual'): 12721 games

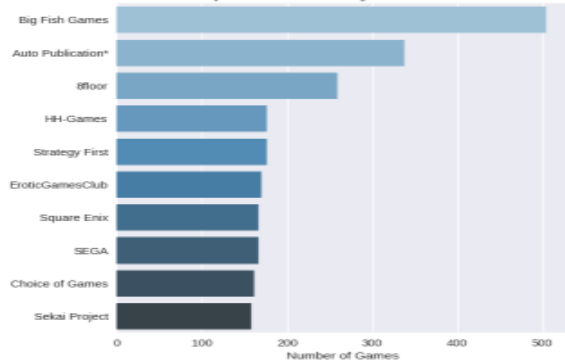


III. EDA – 05. Studios

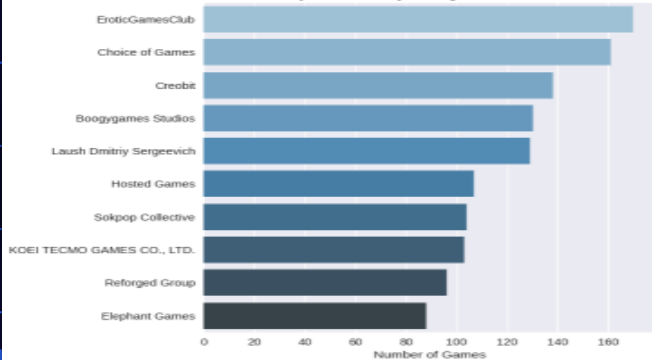
Top studios



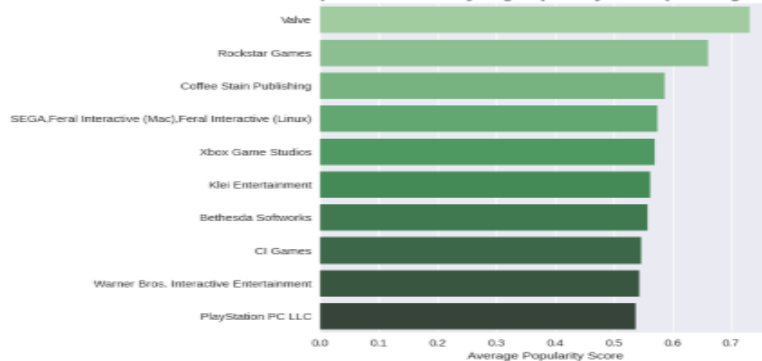
Top 10 Publishers by Game Volume



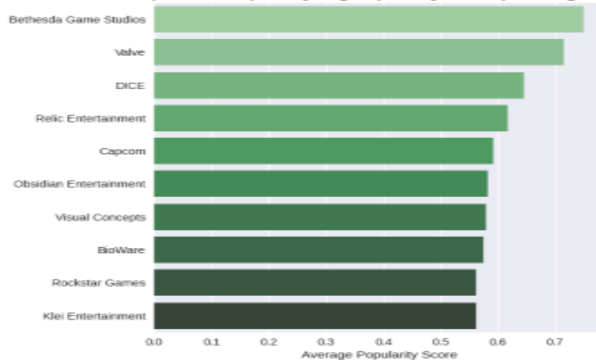
Top 10 Developers by Game Volume



Top 10 Publishers by Avg Popularity Score (min 10 games)



Top 10 Developers by Avg Popularity Score (min 10 games)



- Top 10 Publishers control: 2.7% of market
- Top 10 Developers control: 1.4% of market
- Market Structure: Fragmented

III. EDA – 05. Studios



Indie vs AAA

	Price	Popularity score	Total reviews	Review ratio
Publisher tier				
Indie	7.17	0.10	750.63	0.64
Major	9.73	0.16	2,224.50	0.62

	Price	Popularity score	Total reviews	Review ratio
Developer tier				
Indie	7.82	0.11	1,029.79	0.64
Major	7.82	0.12	1,711.87	0.57

Volume vs Quality Correlation:

- Publishers: 0.071
- Developers: 0.047
- Strategy Insight: Balanced

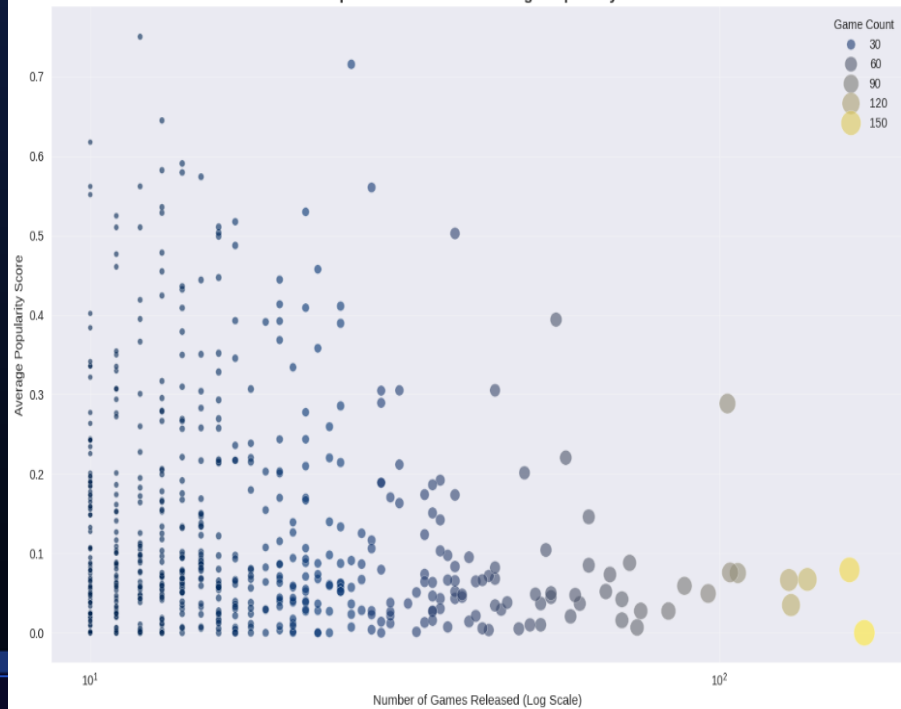
III. EDA – 05. Studios



Publisher Game Count vs. Average Popularity Score



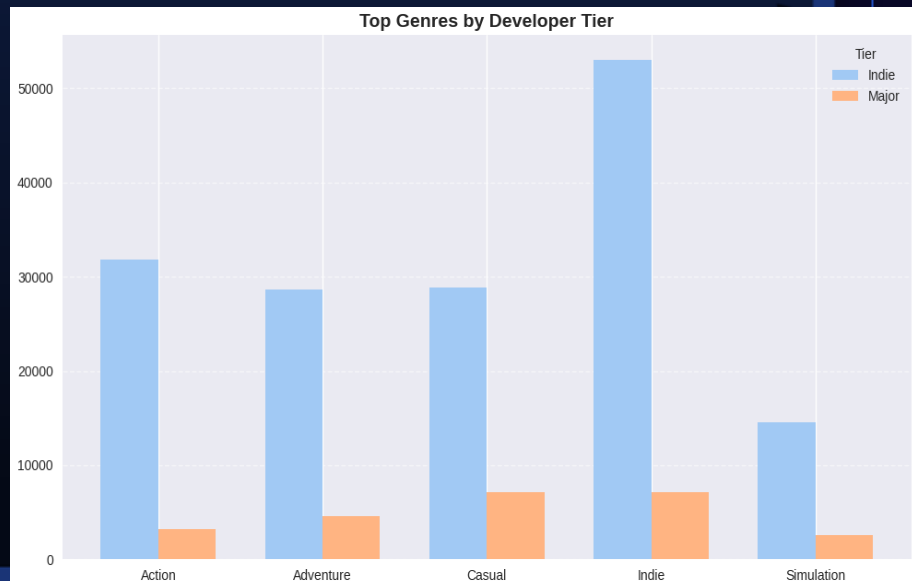
Developer Game Count vs. Average Popularity Score



III. EDA – 05. Studios

↑
TOP

Studios top Genres



III. EDA – 06. Legend Games



Criteria for Legendary Games

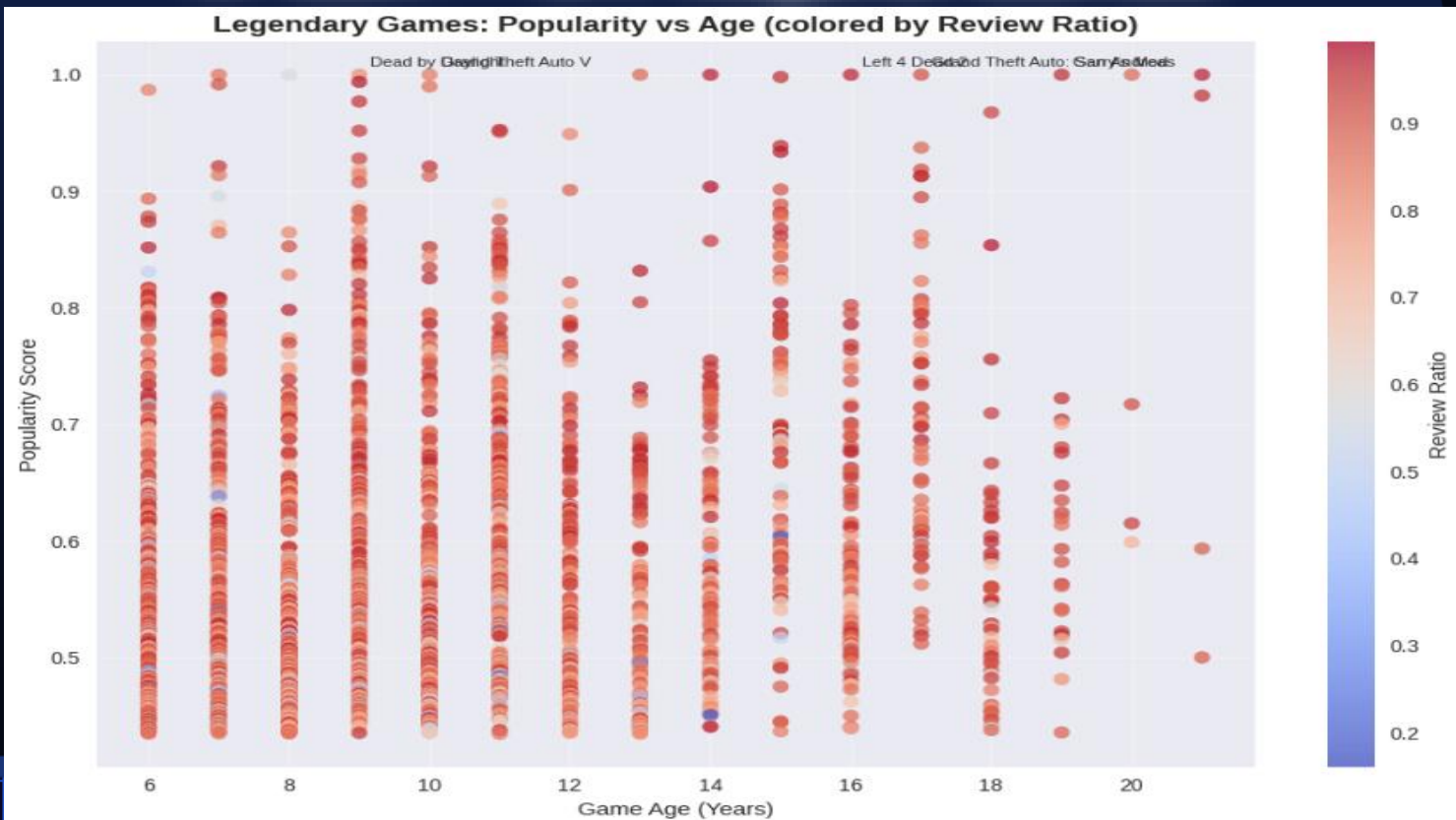
- Age: > 5 years
- Popularity: Top 5% (Score > 0.434)
- Reviews: > 1,000



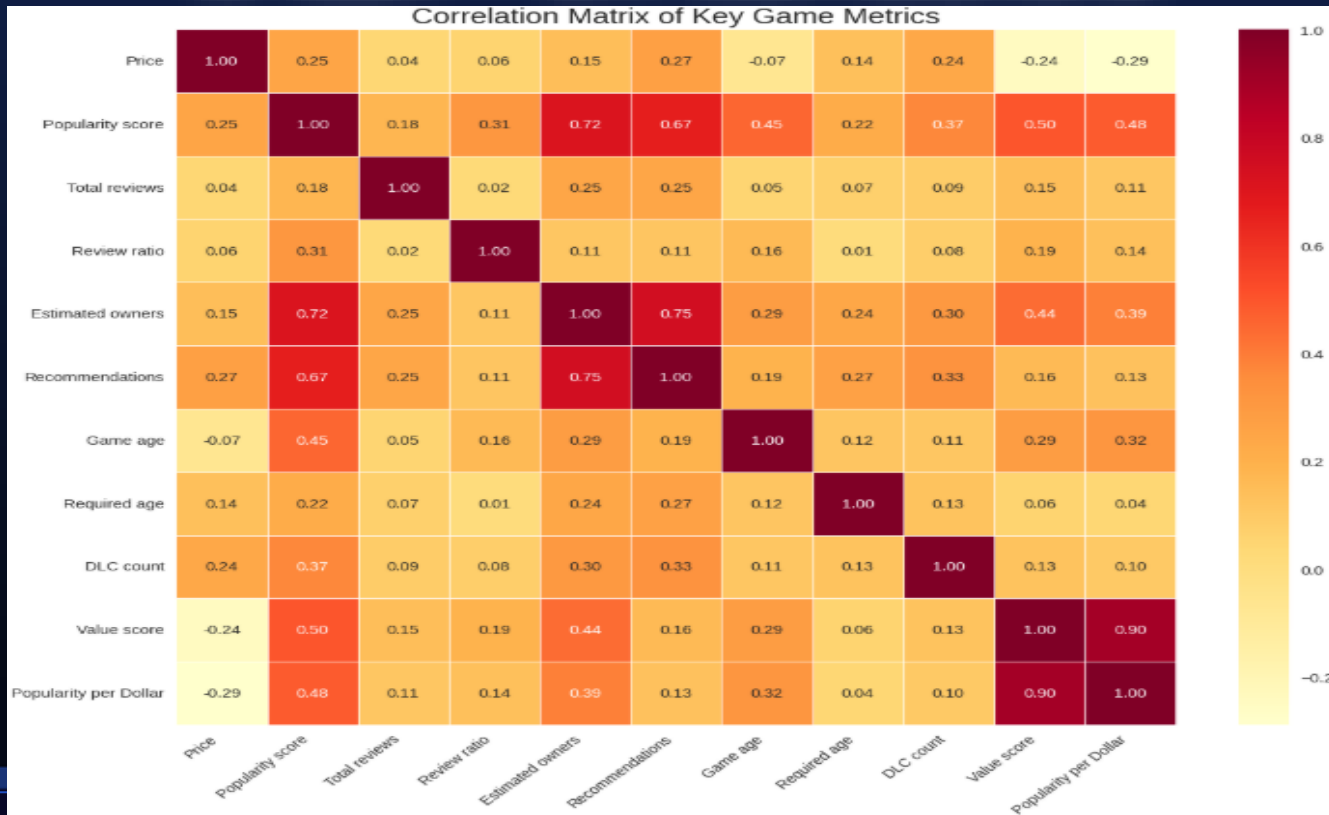
Top 10 Legendary Games:

Name	Release year	Game age	Popularity score	Review ratio	Estimated owners
Garry's Mod	2006	19	1.00	0.97	1,500,000.00
Left 4 Dead 2	2009	16	1.00	0.97	1,500,000.00
Dead by Daylight	2016	9	1.00	0.81	1,500,000.00
Grand Theft Auto V	2015	10	1.00	0.85	1,500,000.00
Grand Theft Auto: San Andreas	2008	17	1.00	0.92	1,500,000.00
Counter-Strike: Global Offensive	2012	13	1.00	0.88	1,500,000.00
Terraria	2011	14	1.00	0.98	1,500,000.00
PUBG: BATTLEGROUNDS	2017	8	1.00	0.56	1,500,000.00
Half-Life 2: Lost Coast	2005	20	1.00	0.88	1,500,000.00
Rust	2018	7	1.00	0.87	1,500,000.00

III. EDA – 07. Legend Games



III. EDA – 07. Correlations



III. EDA – 07. Correlations

1. THE "SUCCESS TRINITY" - NEARLY PERFECT CORRELATIONS

PAIR	CORR.	DESCRIPTION	INSIGHT	STRATEGIC	ACTION
Value Score ↔ Popularity per Dollar	0.90	Value drives popularity relative to cost	Value optimization multiplies success	Maximize perceived value	Build clear value propositions
Estimated Owners ↔ Recommendations	0.75	More owners → more recommendations	Size predicts advocacy	Scale user base first	Focus on acquisition over pricing
Popularity Score ↔ Estimated Owners	0.72	Popularity scales ownership	Popularity fuels exponential growth	Build and fuel hype	Leverage viral & influencer marketing

2. THE "ENGAGEMENT CLUSTER" - STRONG POSITIVE CORRELATIONS

PAIR	CORR.	DESCRIPTION	INSIGHT	STRATEGIC	ACTION
Popularity Score ↔ Recommendations	0.67	Popularity creates recommendation loops	Social proof engine	Early popularity matters	Use influencers and community
Popularity Score ↔ Value Score	0.50	Popular = valuable	Quality sustains fame	Invest in both product & promotion	Match quality to marketing
Pop. Score ↔ Pop./\$	0.48	Success breeds efficiency	Fame → better cost ratios	Upfront marketing pays off	Front-load campaigns
Pop. Score ↔ Game Age	0.45	Older games more popular	Longevity builds success	Think long-term	Support evergreen games
Est. Owners ↔ Value Score	0.44	High value → wide adoption	Value drives scale	Communicate value well	Highlight key value in marketing
Est. Owners ↔ Pop./\$	0.39	Big base = cost-effective fame	Scale boosts efficiency	Focus on volume	Retention & expansion > price

3. THE "CONTENT ENHANCEMENT" GROUP - MODERATE TO WEAK CORRELATIONS

PAIR	CORR.	DESCRIPTION	INSIGHT	STRATEGIC	ACTION
Pop. Score ↔ DLC Count	0.37	Popular = more content	Monetization via expansions	Delay content until fame	Build post-launch DLC plans
Recommendations ↔ DLC Count	0.33	More DLC = more advocacy	Content extends engagement	Fuel advocacy with expansions	Strategize DLC for recomm. spikes
Game Age ↔ Pop./\$	0.32	Older games = better efficiency	Time improves cost-efficiency	Plan for long-term	Build sustainable brands
Pop. Score ↔ Review Ratio	0.31	Pop games have better reviews	Popularity ↔ quality	Quality supports fame	Continuous product polish
Est. Owners ↔ DLC Count	0.30	Big base = more expansions	User size justifies content	Expand after scaling	Sync DLC with user growth
Est. Owners ↔ Game Age	0.29	Time → more owners	Penetration takes time	Grow steadily	Long-term acquisition strategies
Game Age ↔ Value Score	0.29	Older = more valuable	Updates refine value	Improve over time	Continuous enhancements
Price ↔ Recommendations	0.27	Higher price → more advocacy	Premium implies quality	Use price to frame perception	Position as premium product
Recomm. ↔ Required Age	0.27	Mature games → more recomm.	Older content = better advocacy	Don't shy from mature markets	Target adult segments
Total Reviews ↔ Est. Owners	0.25	More users = more reviews	Scale → engagement	Use for feedback, not success	Focus on meaningful reviews
Total Reviews ↔ Recomm.	0.25	Reviews = advocacy indicator	Reviews show engagement	Encourage reviews	Incentivize quality reviews
Price ↔ Pop. Score	0.25	Premium can be popular	Price ≠ block success	Premium must earn it	Deliver exceptional value
Est. Owners ↔ Required Age	0.24	Mature games scale	Adult markets viable	Mature ≠ niche	Develop for adult segments
Price ↔ DLC Count	0.24	High price = more DLC	Premium funds expansion	Use price smartly	Reinvest in content
Pop. Score ↔ Required Age	0.22	Mature = popular	No age cap on popularity	Build for all ages	Age-appropriate engagement

Index

- Strategies
- Key factors

04

Conclusion



IV. Conclusion – 01. Strategies

✓ Quality over Quantity

Focus on a few high-quality games.
Indies have better review ratios.
Fragmented market = big opportunity.

✓ Value-Based Approach

Best price: \$5–8.
Value > price ($r = 0.90$).
Avoid low pricing without visibility.

✓ Positioning

Action, Adventure, Casual = high reach.
Niche focus = high success.
F2P shows strong results.

✓ Timing and Marketing

Fall = max exposure.
Early marketing boosts momentum.
Avoid winter fatigue.

✓ Long-Term Value

Longer lifespan = more popular.
Major updates every 3–6 months.
DLC works better after success.

✓ Prepare for future

Steam Deck leads handheld PC.
Portable PC gaming is rising.
Ensure cross-platform support.

IV. Conclusion – 01. Strategies



New studios

- Leverage indie credibility
- Focus on niche opportunities
- Build early community engagement
- Price based on delivered value



Old Studios

- Use pricing power strategically
- Invest in strong marketing
- Avoid quantity-over-quality traps
- Build lasting IP and brand value



Publishers & Investors

- Balance niche and mainstream titles
- Support long-term content updates
- Evaluate teams by quality KPIs
- Develop in-house F2P capabilities

IV. Conclusion – 02. Key factors

Value Wins

Value is the strongest success driver
(Focus on what players get, not just what they pay)

Market Reality

Few hit games dominate the charts
(Success is concentrated — quality over volume)

Momentum Matters

Scale and player advocacy fuel growth
(Early traction builds long-term success)

Player-Centric Design

Satisfaction drives reviews
(Focus on delighting players, not chasing ratings)



★ Critical Success Formula

Niche Focus (92.5% success rate)

+

F2P Model (0.21 quality score)

+

Quality-First Approach

+

Deep Engagement Systems

+

Community Building

=

Sustainable Success



Thanks!

Does anyone have any questions?

thanhtruc034@freepik.com

github.com/Truc034



CREDITS: This presentation template was created by [Slidesgo](#), and includes icons by [Flaticon](#) and infographics & images by [Freepik](#)

Fonts & colors used

This presentation has been made using the following fonts:

Roboto

(<https://fonts.google.com/specimen/Roboto>)

Turret Road

(<https://fonts.google.com/specimen/Turret+Road>)

#ffffff

#090827

#090a67

#214fc0

#73b3e3

#e4d0f5

#8c4bc4