

# STEAMCHARTS

## WHAT'S POPULAR?

Searching for trends and hidden gems  
amongst the most popular games on Steam.

Ryan McDermott | DSCI510 Final Project





# DATA SOURCES



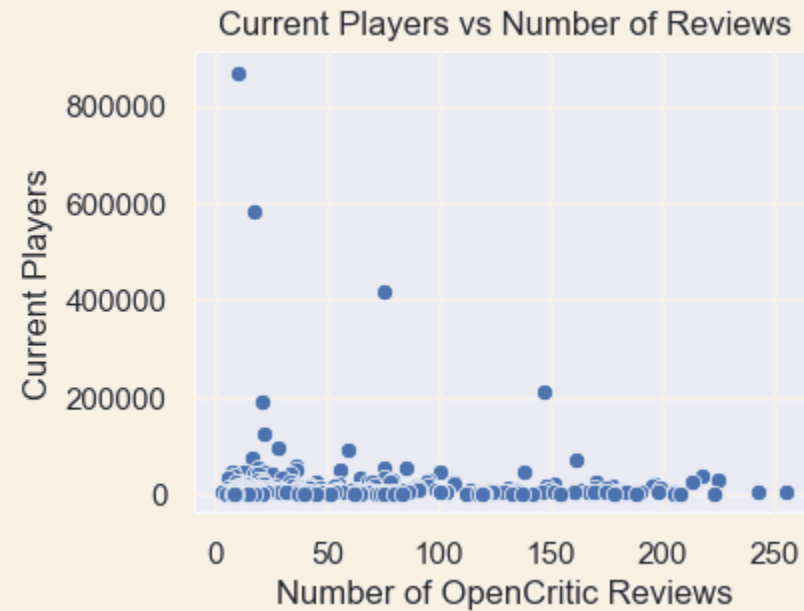
Dataset	Description/Type	Processing Approach	Data Size
<b>SteamCharts</b>	Website that presents the most popular games on Steam currently. Includes game names, appids, current players and peak players	The site is scraped via beautifulsoup to collect information one the requested number of games.	375 rows x 4 columns
<b>Steam Store Web API</b>	API that returns specific information such as price, discounts, if it's free, release dates	The Steam appids scraped from SteamCharts are used to request information from the Steam Store Web API.	375 rows x 9 columns
<b>OpenCritic API</b>	API that returns review information such as top critic review, median critic scores, % recommended, total reviews and a ranking tier.	The OpenCritic API is used to and the game name is searched and the initial results returned.	375 rows x 7 columns

# RATINGS, DISCOUNTS, & POPULARITY

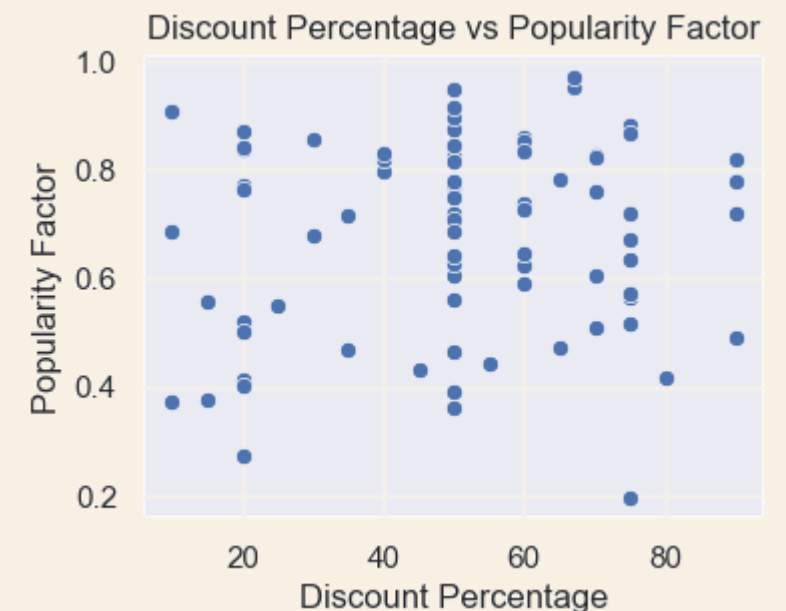
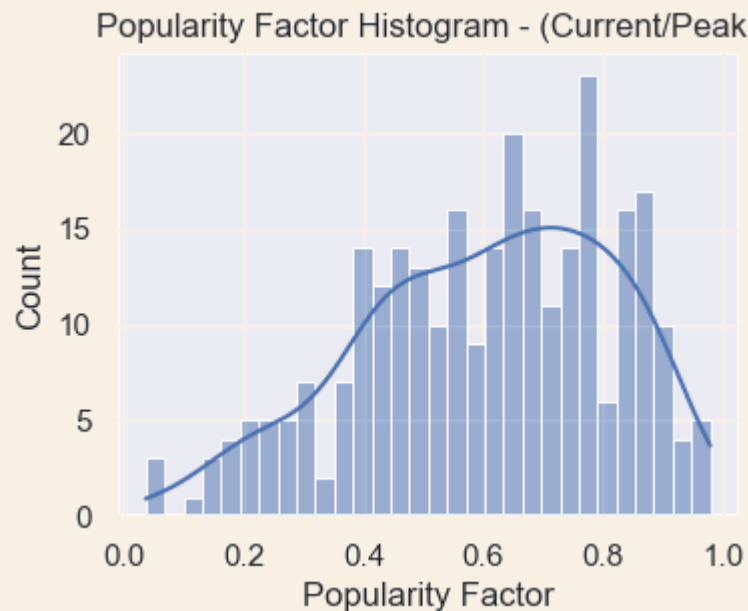
$$\text{Popularity Factor} = \frac{\text{Current Players}}{\text{Peak Players}}$$

Data was cleaned to remove games that didn't have any OpenCritic reviews and were marked as not free but had a price of 0\$.

- This data set went from 375 games to 286 games.
- and then split between Free and Paid games
- 60 Free Games
- 226 Paid Games
  - 86 of which were discounted



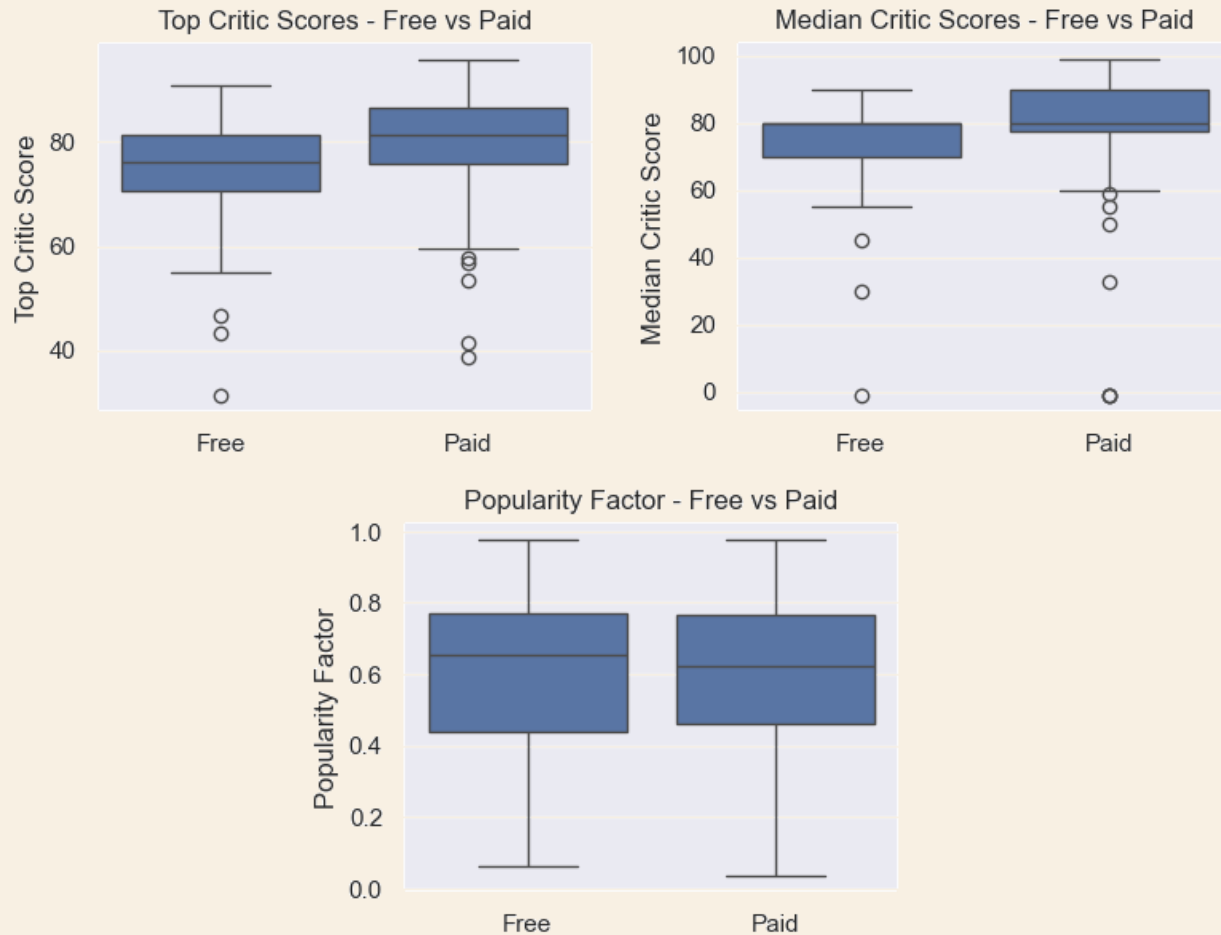
Correlation between the total OpenCritic reviews and the current number of players: 0.0206



Correlation between the discount percentage and the popularity factor: 0.0734

# FREE VS PAID

Interestingly, paid games tend to carry higher critic scores – both top and median, but free games edge out on popularity.

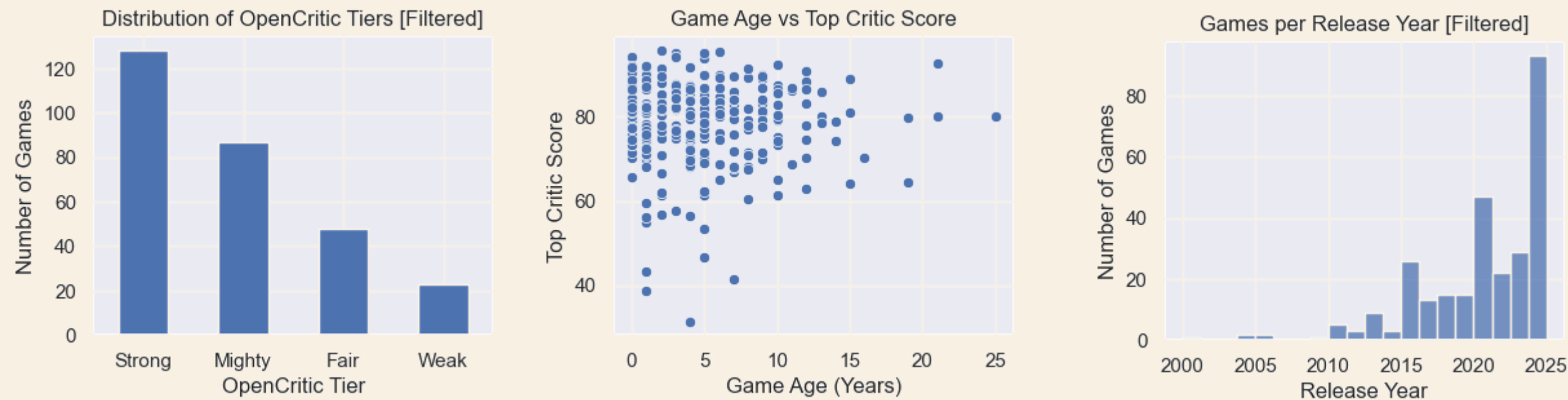


# RELEASE DATES AND AGING

Unsurprisingly, the best rated games were the most popular per the OpenCritic tiers.

The older a game is in the list, the more likely it was reviewed highly.

The most popular games are currently dominated by newer games.



# HIDDEN GEMS

To locate the “best” and “worst” of the most popular games, a Value equation is calculated below:

$$Value = \frac{TopCriticScore * Popularity\_Factor}{Discounted\_Price + 200}$$

This value calculation attempts to prioritize popularity and the highest ratings over the cost – otherwise free games tend to dominate the value rankings.

The worst value tend to be the most expensive games that have become abandoned while also being rated lower than the majority of the games queried.

Top 5 best value games:

	game name	Current Price (USD)	TopCriticScore	Popularity_Factor	Value
119	Magic: The Gathering Arena	0.00	82.448276	0.978110	0.403218
31	Risk of Rain 2	8.24	86.052632	0.952936	0.393789
167	Fallout: New Vegas	9.99	81.166667	0.977123	0.377684
97	Lost Ark	0.00	81.691176	0.909796	0.371612
361	Disco Elysium - The Final Cut	9.99	89.235294	0.865516	0.367801

Bottom 5 value games:

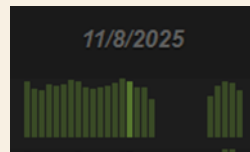
	game name	Current Price (USD)	TopCriticScore	Popularity_Factor	Value
122	Escape From Duckov	17.99	77.500000	0.036804	0.024174
314	New World: Aeternum	59.99	72.000000	0.058605	0.026374
152	NARAKA: BLADEPOINT	0.00	72.166667	0.061585	0.044444
280	Black Myth: Wukong	59.99	81.530303	0.139661	0.071171
271	The Outlast Trials	39.99	72.487179	0.164542	0.085200

Top 5 best paid value games:

	game name	Current Price (USD)	TopCriticScore	Popularity_Factor	Value
31	Risk of Rain 2	8.24	86.052632	0.952936	0.393789
167	Fallout: New Vegas	9.99	81.166667	0.977123	0.377684
361	Disco Elysium - The Final Cut	9.99	89.235294	0.865516	0.367801
128	Balatro	14.99	91.638889	0.860538	0.366802
191	God of War	19.99	94.241379	0.853585	0.365667

# CHALLENGES / ISSUES

- Initial SteamCharts scraping limited to 25 games due to the charts only displaying 25 at a time. Had to scrape through multiple pages.
- Steam Store Web API request limitations. API calls continued to fail after some point. Discovered I was limited to 200 requests every 5 minutes and found I could use `time.sleep` to slow my requests to >1.5 seconds and prevent the request limit from being hit.
- SteamCharts presents dynamic tables that include player counts per day, unsuccessfully tried to scrape data for the previous 7 days.



Dynamic Selection Example

Top Games By Current Players					Next page
Name	Current Players	Last 30 Days	Peak Players	Hours Played	
1. Counter-Strike 2	865,464		1,597,285	558,898,802	
2. Dota 2	506,260		888,896	320,797,821	
3. ARC Raiders	332,535		459,483	56,758,650	
4. Battlefield™ 6	162,344		563,989	148,225,453	
5. Where Winds Meet	161,469		249,851	10,452,578	
6. PUBG: BATTLEGROUNDS	112,077		750,442	156,228,608	
7. Rust	94,174		184,595	56,630,355	
8. Marvel Rivals	76,502		127,557	35,843,219	
9. Bongo Cat	69,417		120,791	49,300,880	
10. Banana	66,027		114,228	47,415,982	
11. Call of Duty®	62,975		91,752	21,589,343	
12. Football Manager 26	58,741		82,750	13,060,940	
13. Grand Theft Auto V Legacy	58,026		108,952	35,634,884	
14. Geometry Dash	53,104		72,511	17,861,259	
15. Baldur's Gate 3	52,812		81,071	25,537,572	
16. Team Fortress 2	50,833		71,887	23,994,121	
17. Path of Exile	48,504		184,633	32,794,789	
18. EA SPORTS FC™ 26	47,845		75,347	21,103,886	
19. Wallpaper Engine	46,715		148,230	47,682,685	
20. Apex Legends	46,186		229,804	48,197,370	
21. Warframe	44,650		86,182	31,390,597	
22. Crosshair X	44,308		61,818	22,976,005	
23. Dead by Daylight	43,468		82,802	28,769,548	
24. War Thunder	42,771		97,370	31,094,782	
25. Stardew Valley	40,983		108,550	31,008,963	Next page

