Video Game Recommender System

Springboard Capstone Project 2
Greg Gibson
Dec. 2020

Summary

Steam

Steam is a popular video game digital distribution service, developed by Valve Corporation, and hosts thousands of games for millions of players.

Can Steam's user vote and review data be used to develop a recommendation system to encourage players to decide which video game to purchase next?

Datasets

Video game developers increase online player involvement through acquiring achievements, treasure, rankings, in-game photos, and unlocking features. Players can link to friends and community and share these updates, as well as feedback.

The Steam API, called Steamworks, allows developers to capture and utilize this information.

The data from Steamworks consists of player votes (up/down) and reviews about individual games, the number of games owned and reviewed by the player, and how much time they have played.

PyPi.org has a tool "steamreviews" to assist with downloading game information into separate JSON files.
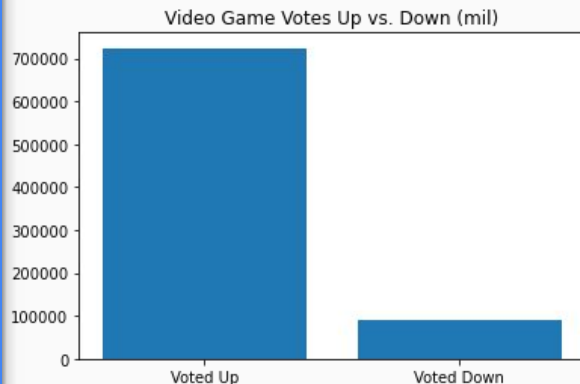
Explore the Data

Starting with:

- 748 video games
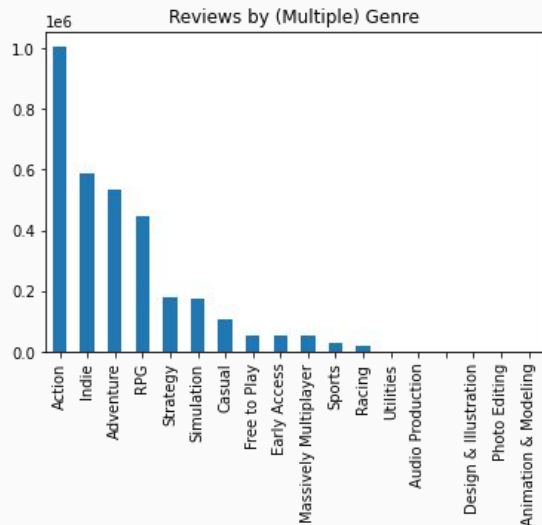- 2.3M player IDs
- 3.3M reviews

Players:

- Owns 128 video games
- Has played for 273 hours
- Plays 100 minutes per week
- 79% have logged one review
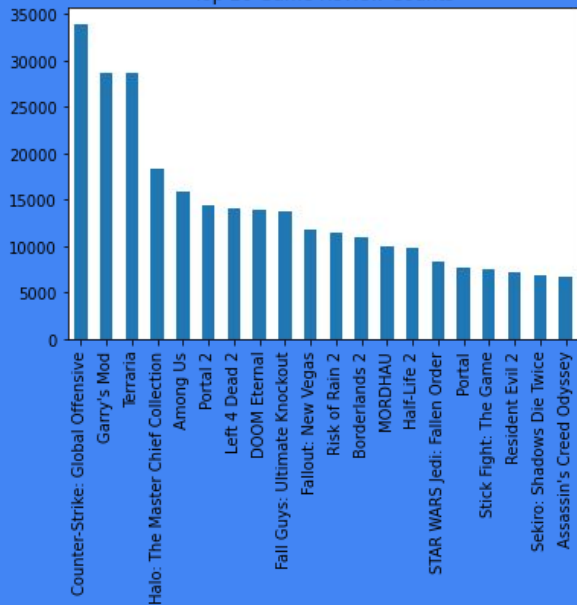- 8% have logged at least three

Players voted Up 8x more

Video Game Votes Up vs. Down (mil)

Players reviewed Action most

Reviews by (Multiple) Genre

Games

Top 20 Game Review Counts

Highest Up Vote Rate

name	reviews	up	down	up_rate
Portal 2	15268	15079	189	99%
Wallpaper Engine	318	314	4	99%
Portal	7840	7731	109	99%
Factorio	138	136	2	99%
Terraria	265148	260149	4999	98%
The Binding of Isaac:Rebirth	5719	5595	124	98%
DUSK	3230	3156	74	98%
RimWorld	389	380	9	98%
Duck Game	816	797	19	98%
DOOM	121	118	3	98%

Lowest Up Vote Rate

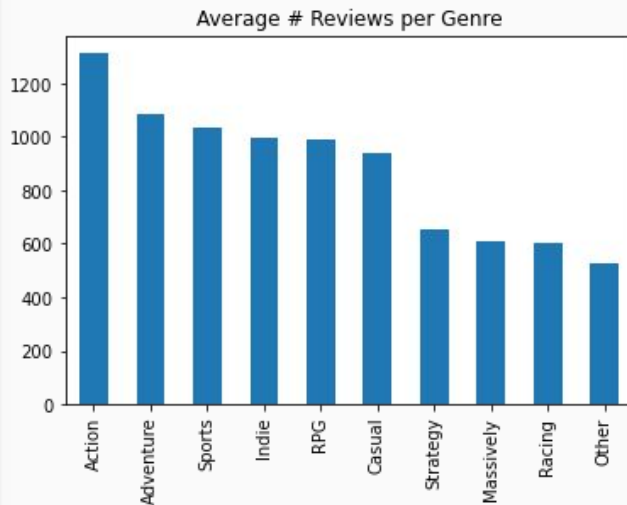| name	reviews	up	down	up_rate
Infestation: Survivor Stories 2020	115	20	95	17%
Just Survive	611	149	462	24%
Secret World Legends	173	63	110	36%
Ace of Spades: Battle Builder	356	135	221	38%
Wolcen: Lords of Mayhem	267	108	159	40%
WWE 2K20	187	78	109	42%
The Culling	201	87	114	43%
ATLAS	3003	1303	1700	43%
Dark and Light	174	83	91	48%

85%
Average
Vote Rate

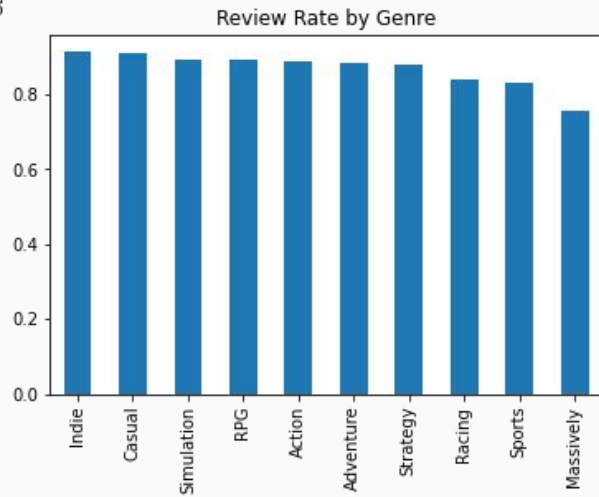# Genres

Steam groups games by:

- Action
- Adventure
- Role Playing (RPG)
- Massively Multiplayer
- Strategy
- Racing
- Independent Developer

Games can be in multiple genres

Action games get the most reviews

Average # Reviews per Genre

Most genres had similar rates of up votes, near 90%, but Massively Multiplayer scored 76% on average

Review Rate by Genre

VADER

Sentiment Analysis Rating Tool

778K Reviews
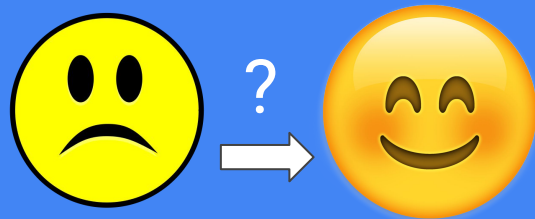
Average 29 Words

0.36 Dataset Score

Preprocessing:

- Removed special characters

- Removed 1-2 character reviews except 'no' and 'ok'

- Removed very long texts with no spaces or very few words

- Removed repetition, yes yes yes yes yes

- Corrected stretching, i.e. "yesssssss"

- Removed reviews without a single word in the English dictionary

Why is the VADER score low? Words used in gaming can have a very negative connotation outside gaming: Shoot -0.34, Battle -0.38, Destroy -0.54, Die -0.60, Violence -0.63, Villains & Apocalyptic -0.66, Kill & Terrorist -0.69, etc.

Rescore Sentiment

- 177K unique text words in reviews
- 2,700 were marked negative by VADER
- Self-identified 300 gamer words and variants to exclude from negative scoring:
  - Kill, evil, villain, attack, enemy, obstacle, etc.
- Replaced these word tokens with a zero-score nonsense word
- Recombined tokens into review
- Rescored with VADER

This effort only slightly changed the sentiment:

0.36 -> 0.41

13% increase

VADER Review

VADER (Valence Aware Dictionary and sEntiment Reasoner) is designed to read emotions from tweets. It takes into account phrasing, repetition, ALL-CAPS, and number of exclamation points: "GREAT GREAT GREAT!!!", on the -1 to 1 scale, will score much higher than "great".

- Highest single word score given is 0.66, "magnificently", for example, used 22 times
- Most negative word was "rapist" at -0.71, used 14 times
- The sentiment score was multiplied by the word's frequency to determine a weight
- Two words each weighted -24,000: "no" is used 83,000 times and "bad" 45,000 times
- The third worst weighting is "shit", 14,700 uses at -0.56 and totaling -8,200

Let's continue with recommendations

What is Collaborative Filtering?

There exists an overwhelming amount of product choices, how to narrow it down?

Recommendations!

- Items with similar characteristics - ingredients in meals, themes of movies, authors of books

  or

- Items enjoyed by other people who have similar tastes as yourself - this group loves apples and pears, just like you, but also love dragon fruit, try it!

  or

- Mix both!

Similar Games

Based on how frequently players liked two games together creates a similarity score

The same game always likes perfectly with itself (1) and games never liked together are zero (0)

Calculated via cosine similarity

	12 is Better Than 6	7 Days to Die	A Plague Tale: Innocence	ARMA: Cold War Assault	ASTRONEER	ATLAS	Ace of Spades: Battle Builder
12 is Better Than 6	1	0.0015	0.0061	0.0055	0	0.0007	0
7 Days to Die	0.0015	1	0.0079	0.0073	0.0011	0.0453	0.0036
A Plague Tale: Innocence	0.0061	0.0079	1	0	0	0.0048	0.0007
ARMA: Cold War Assault	0.0055	0.0073	0	1	0	0	0
ASTRONEER	0	0.0011	0	0	1	0.0049	0
ATLAS	0.0007	0.0453	0.0048	0	0.0049	1	0
Ace of Spades: Battle Builder	0	0.0036	0.0007	0	0	0	1

Neighboring Games

The most similar games can be identified based on those scores

	Most Similar Pattern of Like-Votes									
	1	2	3	4	5	6	7	8	9	10
12 Is Better Than 6	12 Is Better Than 6	Reus	Prey	Beholder	Guns of Icarus Online	Call of Duty: Ghosts	Mark of the Ninja	The Darkness II	LIMBO	Dark Messiah of Might & Magic
7 Days to Die	7 Days to Die	Terraria	Raft	Space Engineers	ATLAS	Garry's Mod	Counter Strike: Global Offensive	DayZ	They Are Billions	Left 4 Dead 2
A Plague Tale: Innocence	A Plague Tale: Innocence	Shadow of the Tomb Raider: Definitive	GRIS	STAR WARS Jedi: Fallen Order	Assassin's Creed Odyssey	Resident Evil 2	Ori and the Will of the Wisps	Life is Strange: Before the Storm	Sekiro: Shadows Die Twice	Quantum Break
ARMA: Cold War Assault	ARMA: Cold War Assault	Arma 2: Operation Arrowhead	Day of Defeat	Half-Life 2: Deathmatch	Counter-Strike	Mount & Blade: Warband	POSTAL 2	Counter-Strike: Source	Zombie Army Trilogy	Half-Life
ASTRONEER	ASTRONEER	Terraria	Among Us	The Forest	Oxygen Not Included	The Incredible Adventures of Van Helsing	SUPERHOT	Risk of Rain 2	Counter-Strike: Global Offensive	The Bureau: XCOM Declassified

Recommended Games

The most similar games, to games each player likes, can be recommended

UserID	0	1	2	3	4	5	6	7	8	9
76561197960 2679	Aliens vs. Predator	Resident Evil 2	Tomb Raider	Amnesia: The Dark Descent	BioShock Infinite	A Plague Tale: Innocence	DOOM Eternal	Halo: The Master Chief Collection	Portal 2	Batman: Arkham Knight
76561197960 2687	DOOM Eternal	Resident Evil 2	STAR WARS Jedi: Fallen Order	CODE VEIN	FINAL FANTASY XV WINDOWS EDITION	Risk of Rain 2	Destiny 2	DRAGON BALL FighterZ	MORDHAU	Remnant: From the Ashes
76561197960 2691	Total War: WARHAMMER II	Assassin's Creed III Remastered	STAR WARS Jedi: Fallen Order	Total War: ROME II – Emperor Edition	Shadow of the Tomb Raider: Definitive Edition	Far Cry New Dawn	Sekiro: Shadows Die Twice	Planet Zoo	MORDHAU	A Plague Tale: Innocence
76561197960 2692	Terraria	Slay the Spire	DRAGON BALL XENOVERSE 2	Risk of Rain	Halo: The Master Chief Collection	Remnant: From the Ashes	Skullgirls	DOOM Eternal	Sekiro: Shadows Die Twice	Devil May Cry 5
76561197960 2696	Among Us	Counter-Strike: Global Offensive	Portal 2	Left 4 Dead 2	Bloons TD 6	Ravenfield	The Binding of Isaac: Rebirth	Fallout: New Vegas	Half-Life 2	Fall Guys: Ultimate Knockout

How Accurate

?

- Added average sentiment score per game to update ranking
- Identified player's favorite genre
- Removed each player's recent votes for games within favorite genre
- Re-recommended games to play

How often did the recommender identify a game we know the player does like?

28%

The recommender can recall about a quarter of a gamer's preferences

Please enjoy this player's "art" discovered in the text reviews

Before I played:

```
\ \ _
  \(ﮬـﮬ)
   <  ⌒
   /  ^\
  /  /  \\
 レ ノ    ﹨_つ
   / /
  / /|
 (  (ヽ
 | |ヽ \
 |丿 \ ⌒)
 | |   )/
`丿)  レ/
```

After I played:

```
/�            ム `ヽ
/ノ) )  ﹨
/ |    (ﮬـﮬ) ノ⌒ (ﹳ_,ノ
/  ﾉ⌒7⌒ヽー< ﹨  /
ﾍ_ノ。    ﾉ、 。|/
  `�`ー'_人`ー丿
    、 ￣_人'彡)
```

Now i will never play again. it ruined my life.

Game Over