

Video Game Recommender System

Springboard Capstone Project 2

Greg Gibson

Nov. 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:

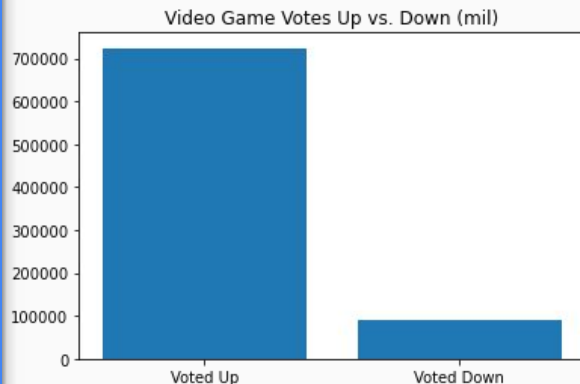
- 748 video games
- 2.3M player IDs
- 3.3M reviews

Players:

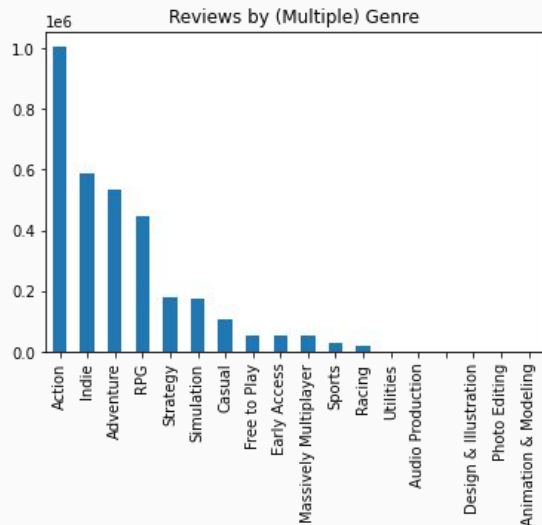
- Owns 128 video games
- Has played for 273 hours
- Plays 100 minutes per week
- 79% have logged one review
- 8% have logged at least three



Players voted Up 8x more

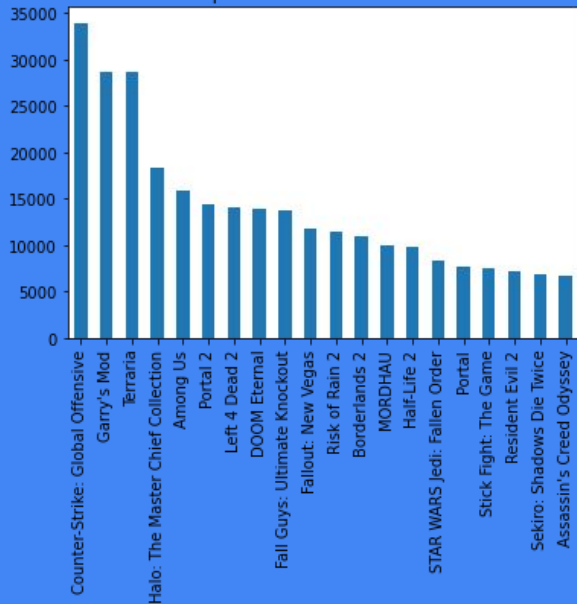


Players reviewed Action most



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

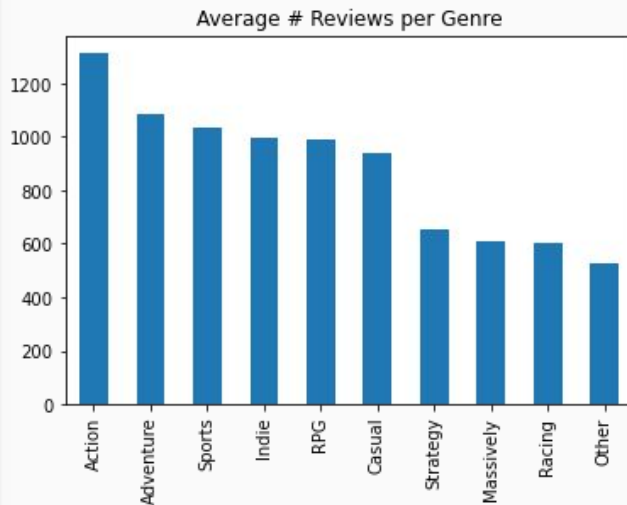
Genres

Steam groups games by:

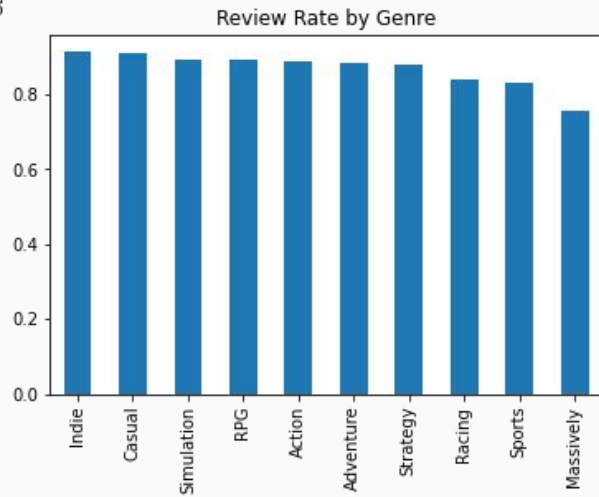
- Action
- Adventure
- Role Playing (RPG)
- Massively Multiplayer
- Strategy
- Racing
- Independent Developer

Games can be in multiple genres

Action games get the most reviews



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





VADER

Sentiment Analysis Rating Tool

778K Reviews

Average 29 Words

0.37 Dataset Score

Preprocessing:

- Removed special characters
- Removed 1-2 character reviews except 'no' and 'ok'
- Removed long reviews with no spaces
- Corrected repetitions, i.e. "yesyesyesyes"
- Corrected stretching, i.e. "yessssssss"
- 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, Terrorist -0.69, etc.