

Report for Steam Platform

The entire report is divided into these three parts:

1. Data Extraction
2. Data Definition
3. Insights

1. Data Extraction

Steam has assigned a unique 'appid' for each game. This fact is used for extracting all the data for games.

Data Extraction is further divided into 3 stages:

- 1.1. In the first stage, appid is obtained for all game from steam. This is done by sending requests to url: <https://steamspy.com/api.php?request=all&page=0>. Then iterating though all pages
- 1.2. In the second stage, details of all games are obtained from appid. This is done by sending requests to url: <https://steamspy.com/api.php?request=appdetails&appid=540>. Then iterating through all appid values
- 1.3. In the third stage, dates of release is obtained for each game. This is done though web-scraping using BeautifulSoup python package. Url used for getting date of release is <https://store.steampowered.com/app/540/> .Then iterating through all appid values

2. Data Definition

This section is divided into 3 parts:

- 2.1. Extracted Columns
- 2.2. Created Columns
- 2.3. Glossary

2.1. Extracted Columns

Sr. No.	Column Name	Definition
1.	appid	Steam Application ID
2.	name	game's name
3.	developer	developers of the game
4.	publisher	publishers of the game
5.	positive	number of people who have rated this game as positive
6.	negative	number of people who have rated this game as negative
7.	score_rank	score rank of the game based on user reviews
8.	owners	owners of this application on Steam as a range

9.	average_forever	average playtime since March 2009. In minutes
10.	average_2weeks	average playtime in the last two weeks. In minutes
11.	median_forever	median playtime since March 2009. In minutes
12.	median_2weeks	median playtime in the last two weeks. In minutes
13.	ccu	peak CCU yesterday
14.	price	current US price in cents
15.	initialprice	original US price in cents
16.	discount	current discount in percents
17.	tags	game's tags with votes
18.	languages	list of supported languages
19.	genre	list of genres
20.	date_of_release	Date on which game was released on steam. Given in format: DD MMM, YYYY

2.2 Created Columns

Sr. No.	Column Name	Definition
1.	date_of_release	Date on which game was released on steam. Given in format: DD
2.	month_of_release	Month in which game was released on steam. Given in format: MMM
3.	year_of_release	Year in which game was released on steam. Given in format: YYYY
4.	positive_percent	Calculated as: $\text{number of positive} / (\text{number of positive} + \text{number of negative})$. Value lies between 0 and 1
5.	genre_list	List of all the genres assigned to game by steam
6.	Indie	If it exists in genre_list, then its value is 1, otherwise 0
7.	Action	If it exists in genre_list, then its value is 1, otherwise 0
8.	Casual	If it exists in genre_list, then its value is 1, otherwise 0
9.	Adventure	If it exists in genre_list, then its value is 1, otherwise 0
10.	Strategy	If it exists in genre_list, then its value is 1, otherwise 0
11.	Simulation	If it exists in genre_list, then its value is 1, otherwise 0
12.	RPG	If it exists in genre_list, then its value is 1, otherwise 0
13.	Early Access	If it exists in genre_list, then its value is 1, otherwise 0
14.	Sports	If it exists in genre_list, then its value is 1, otherwise 0
15.	Racing	If it exists in genre_list, then its value is 1, otherwise 0

16.	Massively Multiplayer	If it exists in genre_list, then its value is 1, otherwise 0
17.	free_to_play	If a game has Free to Play in genre_list, then its value is 1, otherwise 0

2.3 Glossary

2.3.1 Active games: Games which have 'median_forever' value greater than 0

2.3.2 Non-active games: Games which have 'median_forever' value equal to 0

2.3.3 Paid games: Games for which a player has to buy the game first before he/she can start playing

2.3.4 Free-to-play games: Games for which a player does not need to spend any money for playing the game

3. Insights

All the insights are divided into 7 parts:

1. Overall
2. Ownership
3. Engagement
4. Publisher and Developer
5. Game Genre
6. Free-to-play games vs Paid games
7. Game release date

1. Overall

- Total number of games on Steam are 46,101
- Median price of games on Steam is USD 5
- Median Positive Percent for all games on Steam is 0.78
- There are specific price points for games on Steam which are USD 1, 5, 10, 15, 20

2. Ownership

- 69.18% of games are owned by less than 20 thousand players
- 1% of games are owned by 1 mil to 2 mil players

3. Engagement

- 75.36% of all games are non-active games
- 23.57% of active games have less than 1 hour of median playtime

Inference: 1. Median Positive Percent for all games is 0.78 This means that roughly 3 out of 4 people have rated games as positive. But rating a game as positive does not necessarily mean that games will have more playtime. As 75.36% of all games are non-active games. So we need to look at some other metric to know why a player is not actually playing a game.

4. Publisher and Developer

- 'Choice of Games' is the biggest developer on steam
- Median price for 'Choice of Games' (biggest developer) is USD 4.99
- 'Big Fish Games' is the biggest publisher on steam
- Median price for 'Big Fish Games' (biggest publisher) is USD 9.99

5. Game genre

- Top 5 genre for all games are:
 1. 'Indie'
 2. 'Action'
 3. 'Casual'
 4. 'Adventure'
 5. 'Strategy'
- All game genre categories have increased between 2014 and 2020
- Number of 'Casual' games released has increased 10.34 times between 2014 and 2020

6. Free-to-play games vs Paid games

- 15.47% of all games on steam are free to play games
- When player base is less than 20 thousand, more players are playing paid games
- But as player base crosses 1 mil, more players are playing free-to-play-games
- Top 5 genre for both free-to-play games and paid games are same
- 19% of Free-to-play games have Massively Multiplayer genre
- Only 2% of Paid games have Massively Multiplayer genre

Inference: 2. Between 2014 and 2020 all game genre have increased. Most significant increase has been in 'Casual' game genre. This is also in top 5 genre for paid games. Hence, 'Casual' paid game represents an opportunity which can be explored for monetization.

Inference: 3. In Massively Multiplayer genre there is significant difference in number of games for paid vs free-to-play. One possible reason could be this: 'Massively Multiplayer genre are played by group of people who are friends in real life. Not everyone in the group is ready to pay for such game'. The available data is not sufficient to test this hypothesis.

7. Game release date

- Number of games released have increased at a CAGR of 43.47% between 2014 and 2018
- Maximum number of games are released in month of August
- The difference between games released in August and July (2nd highest) is not significantly large
- Release of games are spread evenly across all months
- 44.57% of all games are released in both Week 1 and Week 2 combined
- 46.26% of all games are released in Week 3