



STEAM[®]

Data Science Project: Steam Games Case Study

By: Randy Koliha, Owen Telis,
Isaiah Sarria, and Jack Levitt

Introduction

- Our project was on the popular game hosting platform, Steam. We decided to do this data set because all of us have played a video game at least once before.
- This data set will help us dive deeper into the behavior of gamers on the platform and have a better understanding of the trends.

	game	year	number_of_month	month	avg	gain	peak	avg_peak_perc
1	Counter-Strike: Global Offensive	2021	1	January	743209.66	25405.91	1124553	66.0893%
2	Dota 2	2021	1	January	432671.65	10119.33	694613	62.2896%
3	PLAYERUNKNOWN'S BATTLEGROUNDS	2021	1	January	201247.19	12013.61	451998	44.5239%
4	Apex Legends	2021	1	January	71766.74	7296.99	129928	55.2358%
5	Rust	2021	1	January	142117.25	80945.60	244394	58.1509%
6	Team Fortress 2	2021	1	January	83148.17	724.66	111102	74.8395%
7	Grand Theft Auto V	2021	1	January	101250.59	-5663.68	184941	54.7475%
8	Tom Clancy's Rainbow Six Siege	2021	1	January	77717.42	6995.92	123637	62.8594%
9	Rocket League	2021	1	January	59449.16	-6315.92	112482	52.8522%
10	Path of Exile	2021	1	January	47685.61	34987.98	157091	30.3554%
11	Football Manager 2021	2021	1	January	55982.74	4415.95	93775	59.6999%
12	ARK: Survival Evolved	2021	1	January	63160.05	6836.00	92006	68.6478%
13	Destiny 2	2021	1	January	44864.92	-22135.76	79290	56.5833%
14	Sid Meier's Civilization VI	2021	1	January	36731.49	3914.54	65951	55.6951%
15	PAYDAY 2	2021	1	January	29201.19	-841.71	54841	53.247%
16	Wallpaper Engine	2021	1	January	44647.85	4077.52	65041	68.6457%
17	Warframe	2021	1	January	36124.14	134.09	60019	60.1878%
18	Euro Truck Simulator 2	2021	1	January	24090.05	138.39	51312	46.9482%
19	Stardew Valley	2021	1	January	57166.55	27545.84	94479	60.5071%
20	Garry's Mod	2021	1	January	28054.43	3691.88	46668	60.1149%
21	Dead by Daylight	2021	1	January	36342.49	710.31	53583	67.8247%
22	Dyson Sphere Program	2021	1	January	25524.93	NA	59339	43.0154%
23	War Thunder	2021	1	January	27477.69	441.49	39887	68.8888%
24	Cyberpunk 2077	2021	1	January	82146.66	-250248.99	225670	36.4012%
25	Total War: WARHAMMER II	2021	1	January	24622.40	-3376.01	38565	63.8465%
26	DayZ	2021	1	January	26336.03	3524.05	45236	58.2192%
27	Farming Simulator 19	2021	1	January	18998.34	981.63	37332	50.8902%

Prerequisites

- We first installed these packages:
library(tidyverse)
library(scales)
library(directlabels)
- Then we cleaned up the data by creating a data frame by converting the months in chr to a mutate of int equivalent values, and arranged the months ascending.



Data Exploration

There was a total
1,258 games in
the data frame
after sorting by
distinct names.

Some of the questions we would like to answer

Do all games lose popularity overtime?
How do multiplayer and single player titles differ?

Is the Steam platform affected by seasonality?

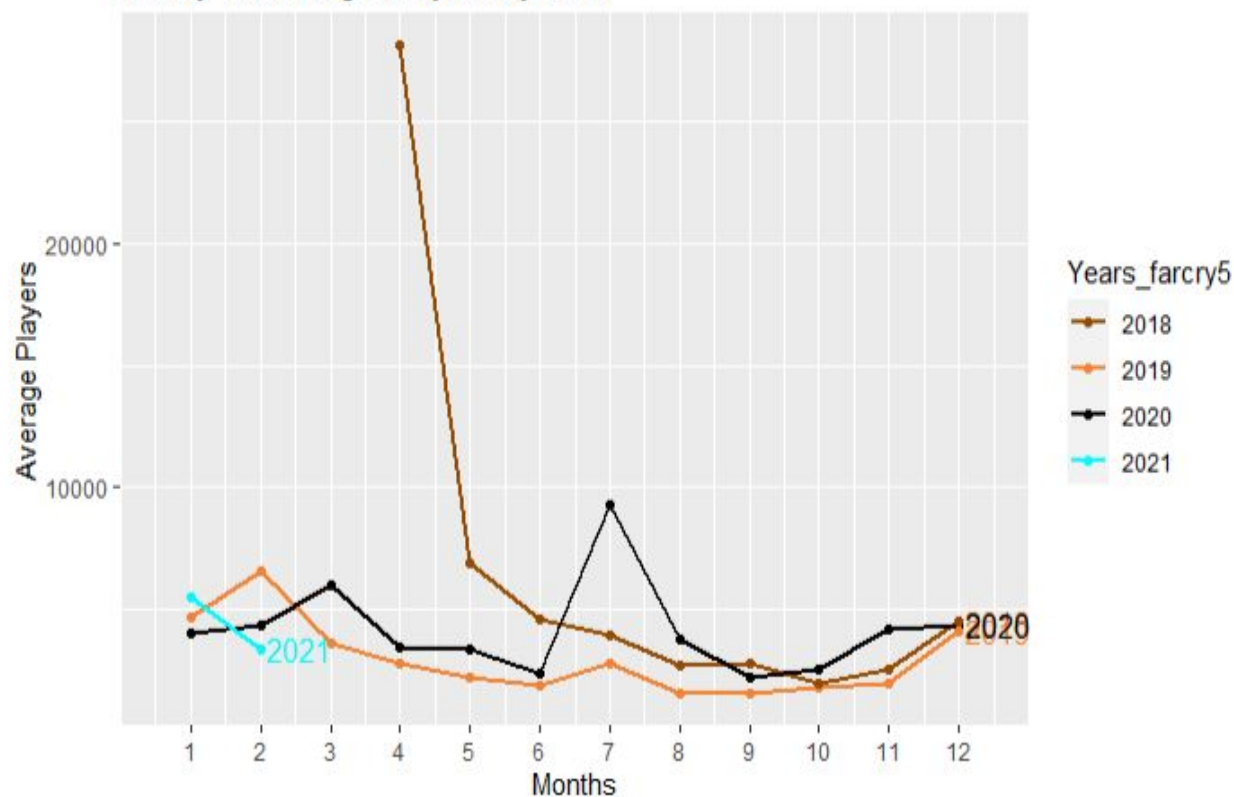
What does the growth of Steam players look like from 2012 – 2021?

How many game titles were published to the Steam platform from 2012 - 2021?

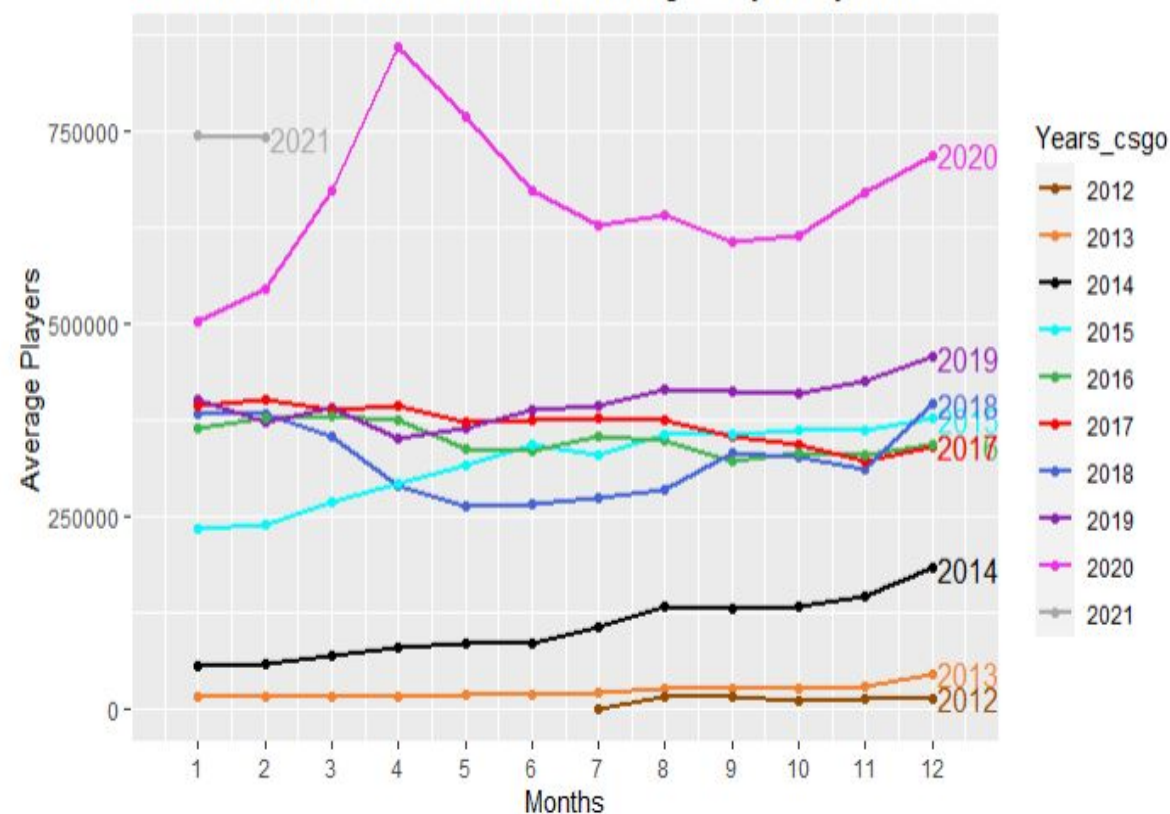
What's the change in new relevant titles being published to Steam from 2012 - 2021?

Do all games lose popularity overtime? How do multiplayer and single player titles differ?

Farcry 5: Average Players by Year



Counter Strike: Global Offensive: Average Players by Year





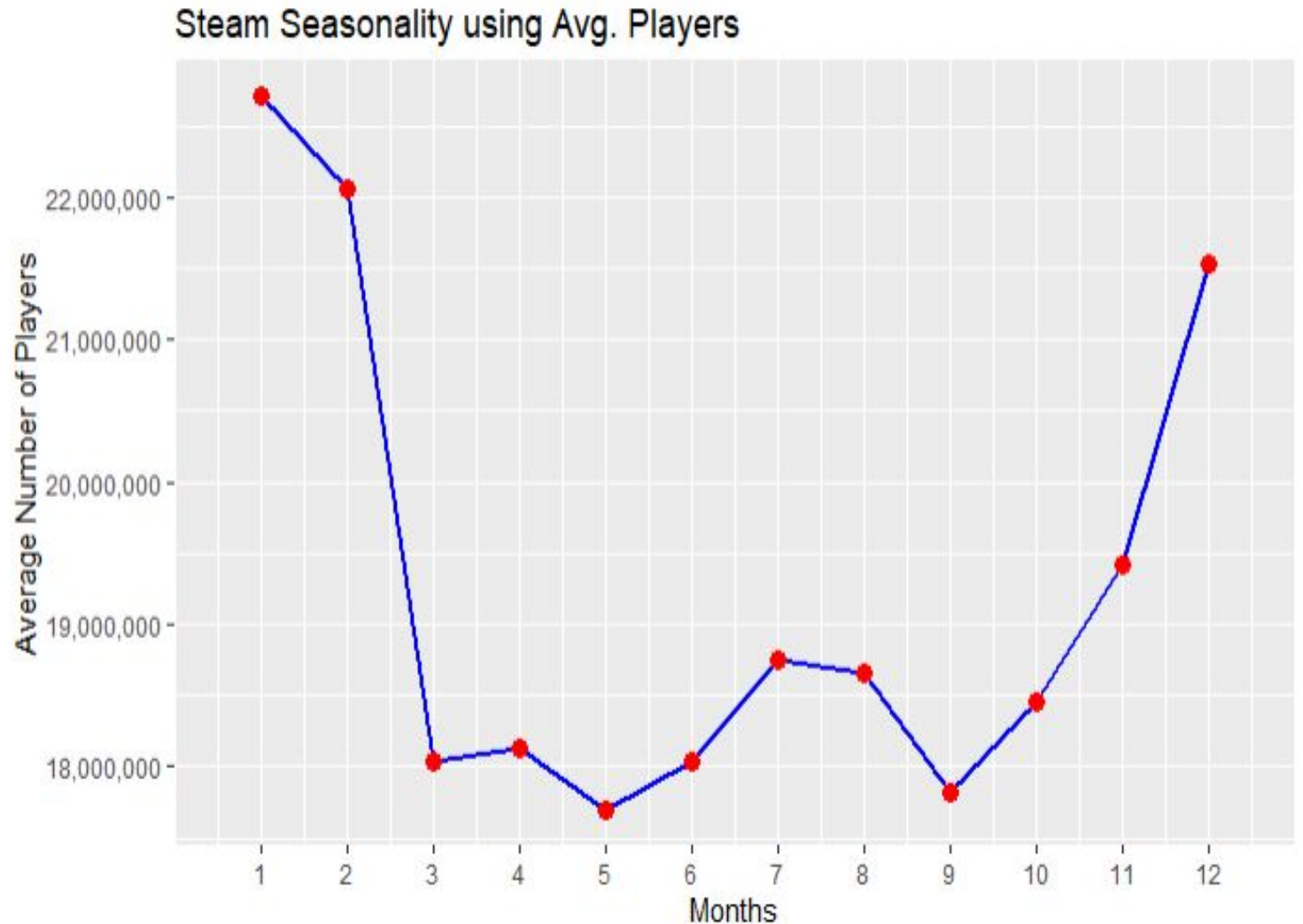
STEAM®

Do all games lose popularity overtime? How do multiplayer and single player titles differ?

- The data shows that many multiplayer games seem to gain player over time and often have their peak average players occur in a period significantly after the launch of the game.
- The data points to the conclusion that single player titles seem to have a drastic fall off in their player base after launch. We can infer that this may likely happens because once someone beats a single player game, they are less likely to return to it. This can also be attributed to the fact that multiplayer/esports games are usually continually updated with new content single player game typically are less likely to receive these updates consistently.

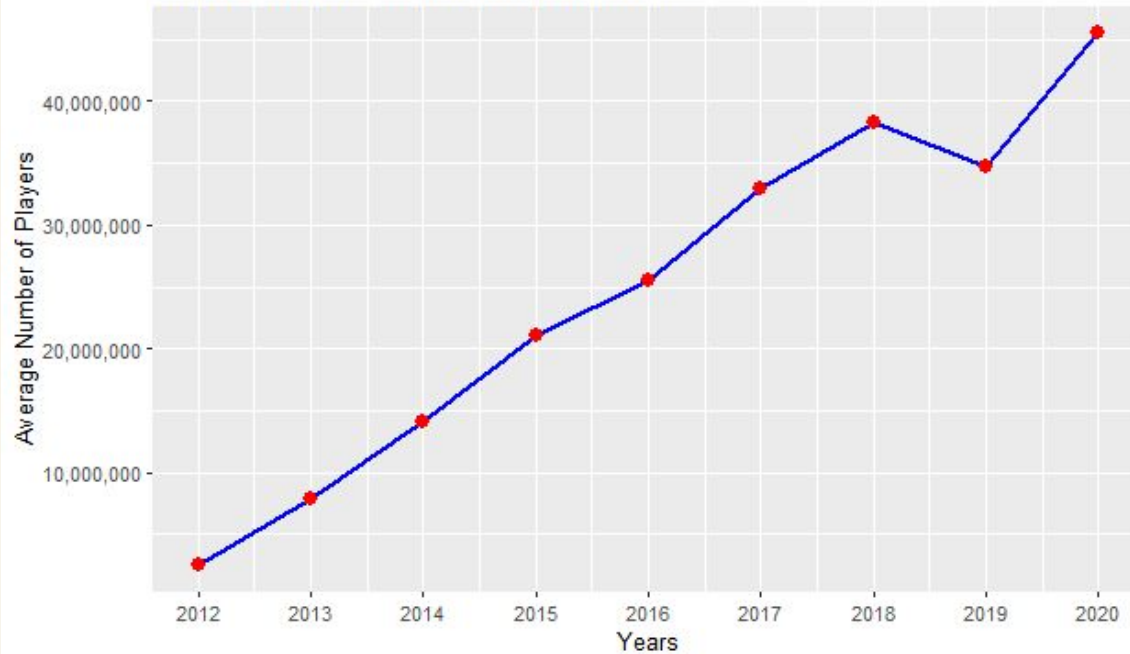
Is the Steam platform affected by seasonality?

- According to our analysis of the data, It suggest that Steam's player data is affected by seasonality. More users will be online using the platform between the months December, January and February. We suspect this is due to an increase of games downloaded and holidays allowing more playable hours.

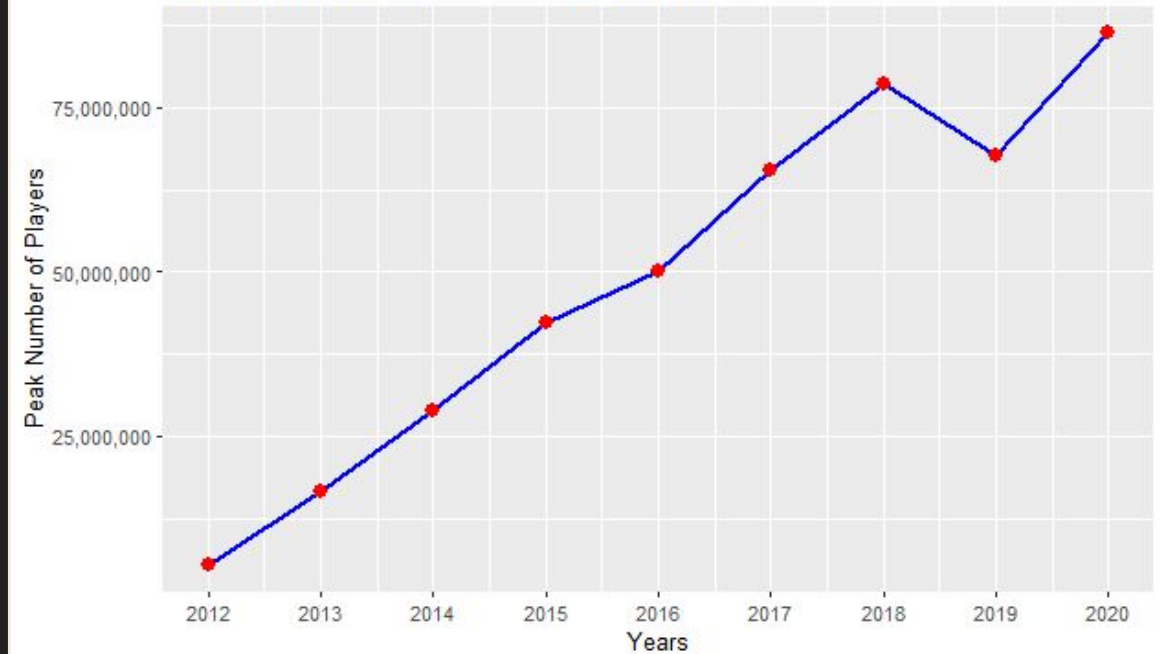


What does the growth of Steam players look like from 2012 -2021?

Steam Platform Growth using Avg. Players



Steam Platform Growth using Peak Players



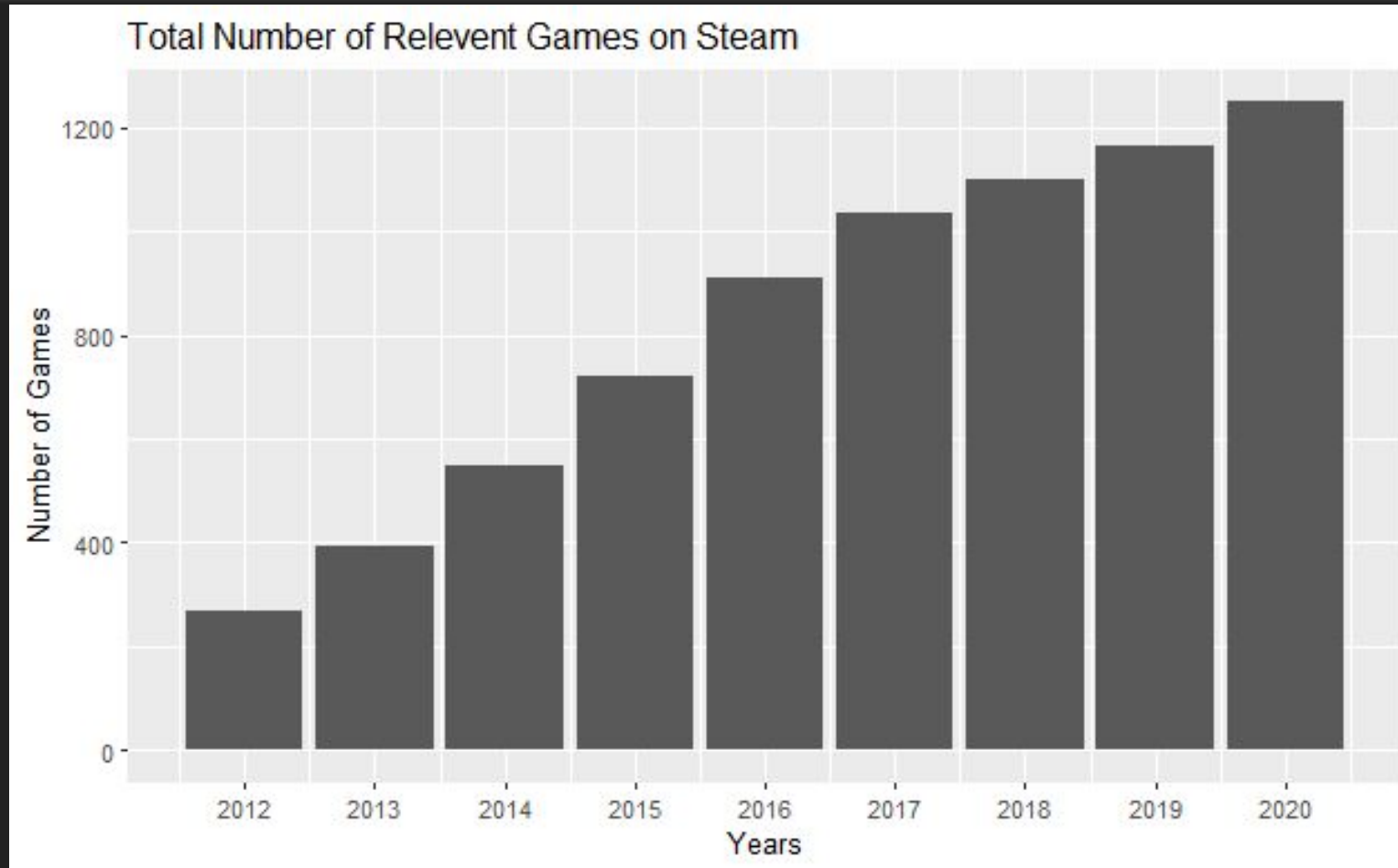
The growth of the steam platform has a stable rising slope until the years 2018-2019.



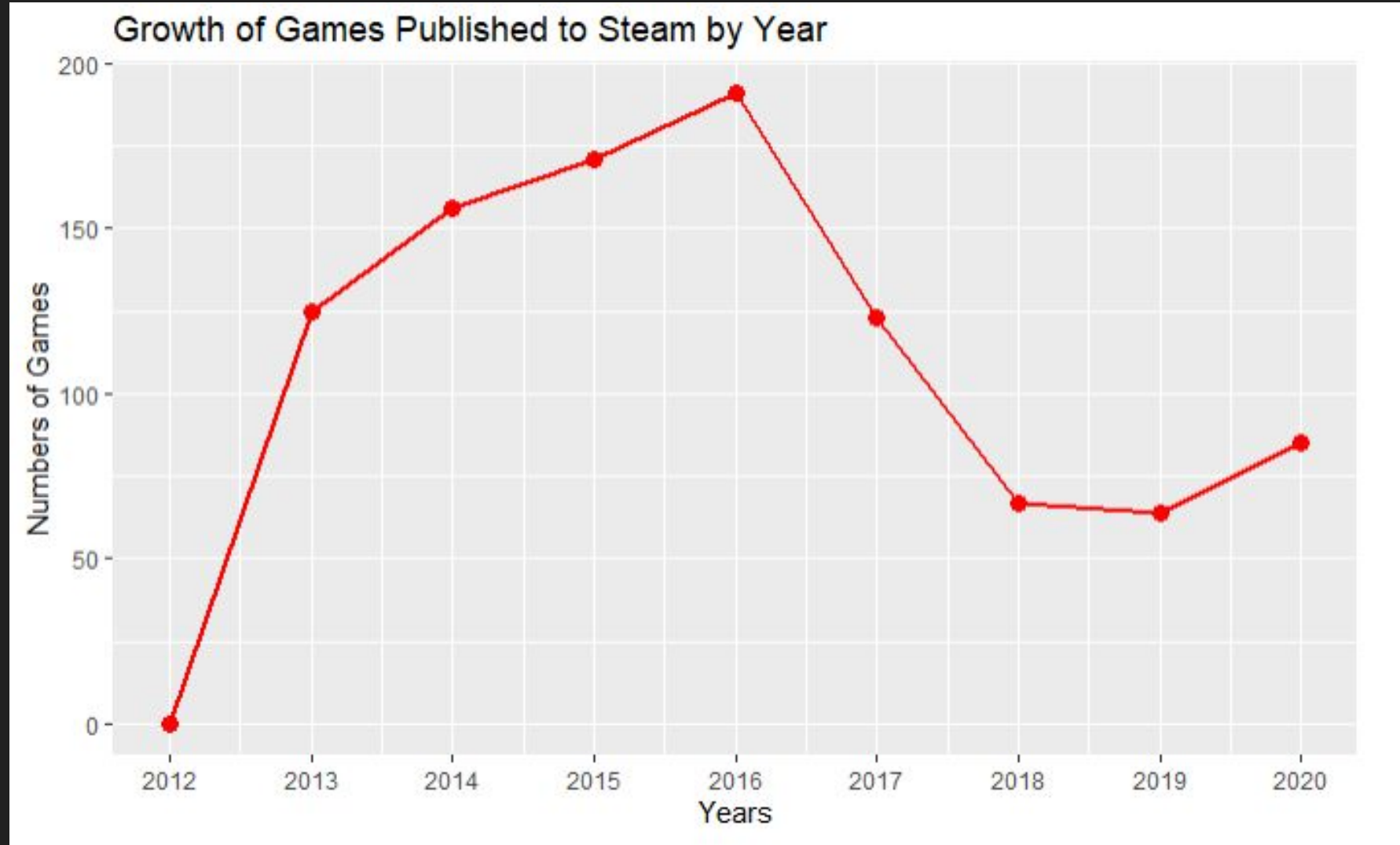
We can speculate this was due to Fortnite becoming popular which resulted in playtime dropping on the Steam platform.

This was caused because Fortnite is an Epic Games product, and they released their own launcher to host the game.

What is the total number of relevant titles on the Steam platform from 2012 - 2021?



What's the change in new relevant titles being published to Steam from 2012 - 2021?





In Conclusion

- Our data suggest that multiplayer games can hit their peak after release after a new content updates release. This is not always the case though. Single player games follow the trend of peaking on release due to players finishing the game.
- We were able to use our data to see the seasonality of Steam Games and conclude that the months December, January and February are the most actively played months.
- Lastly, our analysis concludes that Steam, at a steady rate, is a consistently growing platform that is used by millions daily.