UNIVERSITY OF SOUTH AUSTRALIA

STEAM GAMES VISUALIZATION

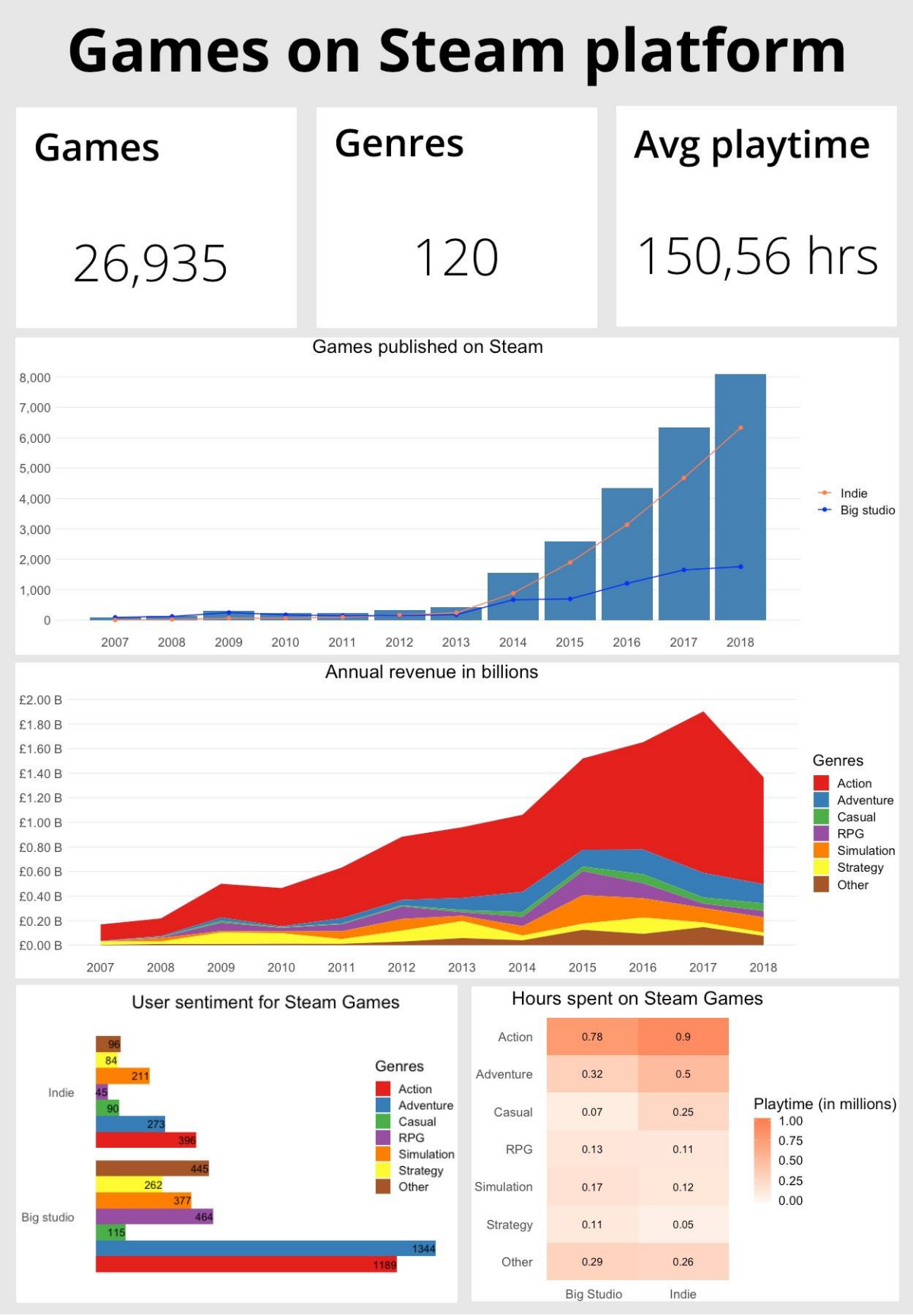
Anonymous student

DATA VISUALIZATION – FINAL REPORT

Table of Contents

Dashboard *2*
Introduction *3*
Data Sources *3*
Visualization Results *5*
Indie vs Non-indie 5
Prices 7
Game Genres 8
Gameplay mode 8
Achievement system 9
Average playtime 10
User Sentiment 11
Publisher 12
Total Revenue 13
Owners 14
Conclusion *15*
References *15*

Dashboard


Games on Steam platform
Games
26,935
Genres
120
Avg playtime
150,56 hrs
Games published on Steam
8,000
7,000
6,000
5,000
4,000
3,000
2,000
1,000
0
2007
2008
2009
2010
2011
2012
2013
2014
2015
2016
2017
2018
Indie
Big studio
Annual revenue in billions
£2.00 B
£1.80 B
£1.60 B
£1.40 B
£1.20 B
£1.00 B
£0.80 B
£0.60 B
£0.40 B
£0.20 B
£0.00 B
Genres
Action
Adventure
Casual
RPG
Simulation
Strategy
Other
User sentiment for Steam Games
Indie
Big studio
96
84
211
45
90
273
396
445
262
377
464
115
1344
1189
Hours spent on Steam Games
Action
Adventure
Casual
RPG
Simulation
Strategy
Other
0.78
0.9
0.32
0.5
0.07
0.25
0.13
0.11
0.17
0.12
0.11
0.05
0.29
0.26
Big Studio
Indie
Playtime (in millions)
1.00
0.75
0.50
0.25
0.00

Introduction

Gaming is a big industry nowadays encompasses PC games, console games, game streaming, and many other forms. In term of PC gaming, there is the most popular platform to not only gamers but to internet users, which is Steam. According to Prescott (2019), Steam is now the largest game distributor and Gough (2020) showed that the latest number of concurrent users in this platform is 20.31 million in March 2020. Moreover, game genres in Steam are diverse, not to mention multiple languages support. Some games were made by indie (independent) developers, and these games are considered a whole new kind to gamers. Ultimately, Steam is the pioneer in game distributing who has distributed free games and an achievement system to enrich user experience.

Therefore, the purpose of this visualization report is to provide an insight into questions like how many games have been published on this platform, what top games and their genres are, and whether a user rating affect the sale or not. The target audience could be indie game developers or big studios who can find it fruitful for their development strategy since most PC games take at least 6 months to a year to develop and a significant amount of money to run advertising campaigns. Furthermore, the indie gaming trend will be unveiled through this visualization report.

Data Sources

The dataset was the second result of the searching term “Steam” on Kaggle website. It was the final product made and cleaned by Davis (2019) by gathering data from two API (application programming interface) endpoints of Steam Store and SteamSpy. Games data in the dataset were collected from the earliest day to roughly March 2019. The dataset contains 27,075 rows across 409 columns and, clearly, its dimension is significantly large for the purpose of this project. Fortunately, it was partitioned into 5 subsamples by the owner, and they are steam, steam_description_data, steam_media_data, steam_requirements_data, steam_support_info, and steamspy_tag_data. The main dataset is “steam” subsample, while other subsamples are merely supplementary. Hence, “steam” subsample is the only considered dataset for the project.

The main dataset has 27,075 observations across 18 columns under a CSV file with a size of 5.8 MB on disk. And, the data types are given as the table below.

Variable	Description	Data Type
appid	Unique identifier for each title	Nominal
name	Game title	Nominal
release_date	Game release date	Nominal
english	Main language in game is English or not	Binary
developer	Game studio who made this game	Nominal
publisher	Game publisher name	Nominal
platforms	Support Operating Systems	Nominal
required_age	Minimum age required to play	Interval
categories	Game categories	Nominal
genres	Game genres	Nominal
steamspy_tags	Game tags provided in Steamspy	Nominal

achievements	Number of achievements	Interval
positive_ratings	Number of positive ratings	Interval
negative_ratings	Number of negative ratings	Interval
average_playtime	The average playtime from Steamspy	Ratio
median_playtime	The median playtime from Steamspy	Ratio
owners	Range of estimated customers	Nominal
price	Full price in GBP	Interval

With 18 features, the main dataset suffices to give diverse perspectives of game publishing on Steam, and there are some variables that can be transformed or merged to create new variables.

Regarding the software to produce graphs, Python and R have been leveraged and, particularly, matplotlib library and ggplot2 package are used in those languages, respectively. Ultimately, to produce an amazing dashboard, an original dashboard will be produced by R then fed to Photoshop to create spacing, other short but important facts, and a big title that makes the dashboard be more compelling to the audience.

Visualization Results

Before presenting graphical results, there are some notes need to be conveyed to the audience first. The dataset has the data from 1997 to 2019; however, in some graphs, the year data starts with 2007 and ends at 2018. The reason is that the numbers of games before 2007 are significantly small; hence, they are hard to be seen in most of the charts. Data collected in 2019 is not a full year one, hence to let it in visualizations may lead to misunderstanding to the audience.

Indie vs Non-indie

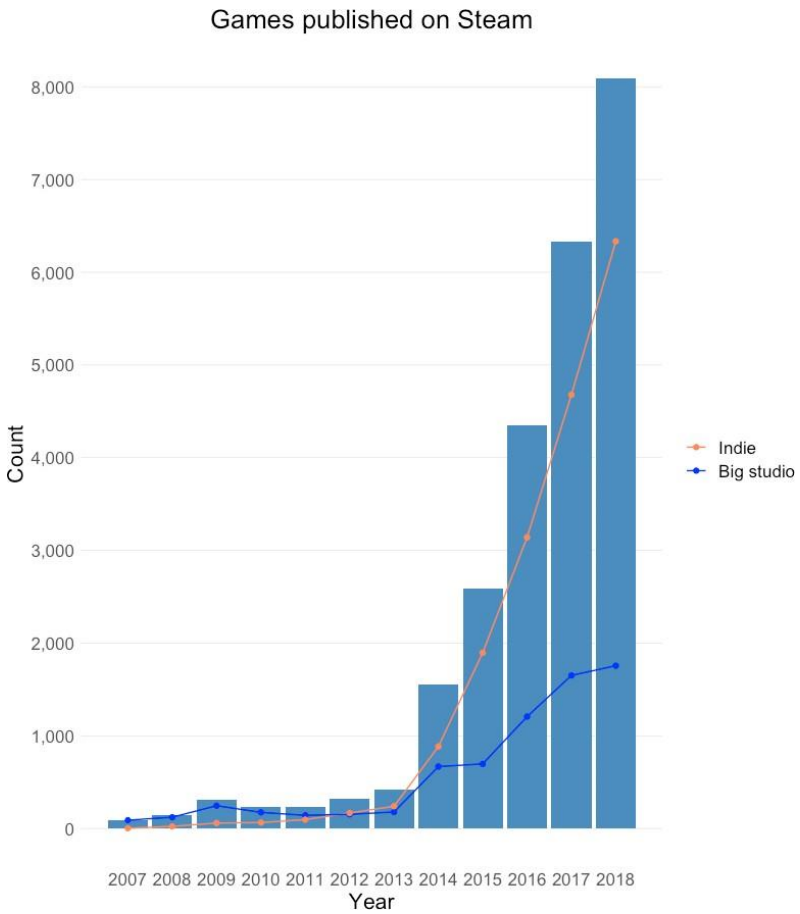


Figure 1: Total games released on Steam from 2007 to 2018 with indie game trend.

As Figure 1 shows, the number of games which were published on Steam fluctuated under approximately 300 games from 2007 to 2013 before rocketing from around 1,500 games in 2014 to just above 8,000 games in 2018. Simultaneously, the numbers of Indie games released on Steam platform were marginal and lower than numbers of Big studio games until 2011. By 2012, Steam witnessed the equal proportions of Indie and non-indie games before Indie games had become a phenomenon and Indie developers raised the numbers rapidly.

The reason for the surge of indie games could be the emergence of game engines like Unity, cocos2d-x, and Unreal that help indie game developers easily to develop and release their game with the multiple times faster than they used to while writing games with native graphic libraries such as OpenGL and DirectX. Note that, big game studios

have spent a significant amount of money to produce a game, and it frequently takes them years to finish. Clearly, the content of games from big studios are rich, and they often are labelled as AAA games, which means excellent on every perspective.

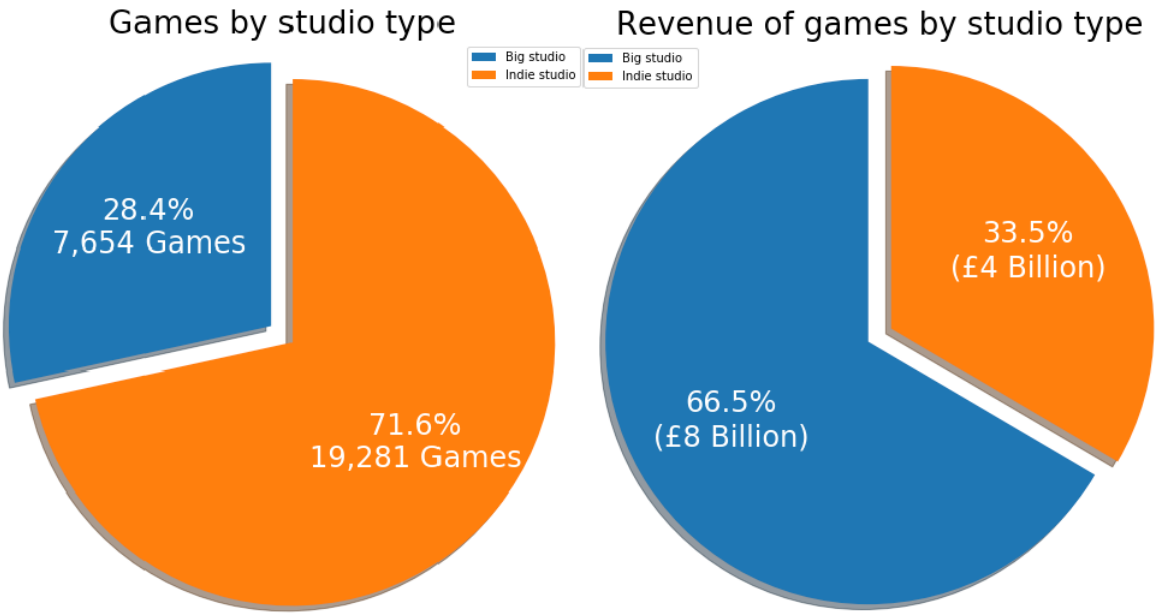


Figure 2 - 3: Piecharts of Games by studio type and Revenue of games by studio type.

Figure 2 shows that Indie games account for 71.6% of total games published on Steam, while 28.4% is from Big studios. Nevertheless, the piechart in Figure 3 depicts an opposite story that revenue from indie games is just a half compared to big studio games, which are 4 billion pounds for the former and 8 billion pounds for the matter. This can be explained by the price of games from Big studios is likely higher than the price of indie games. The graph in the Prices section below will clarify this theory.

Prices

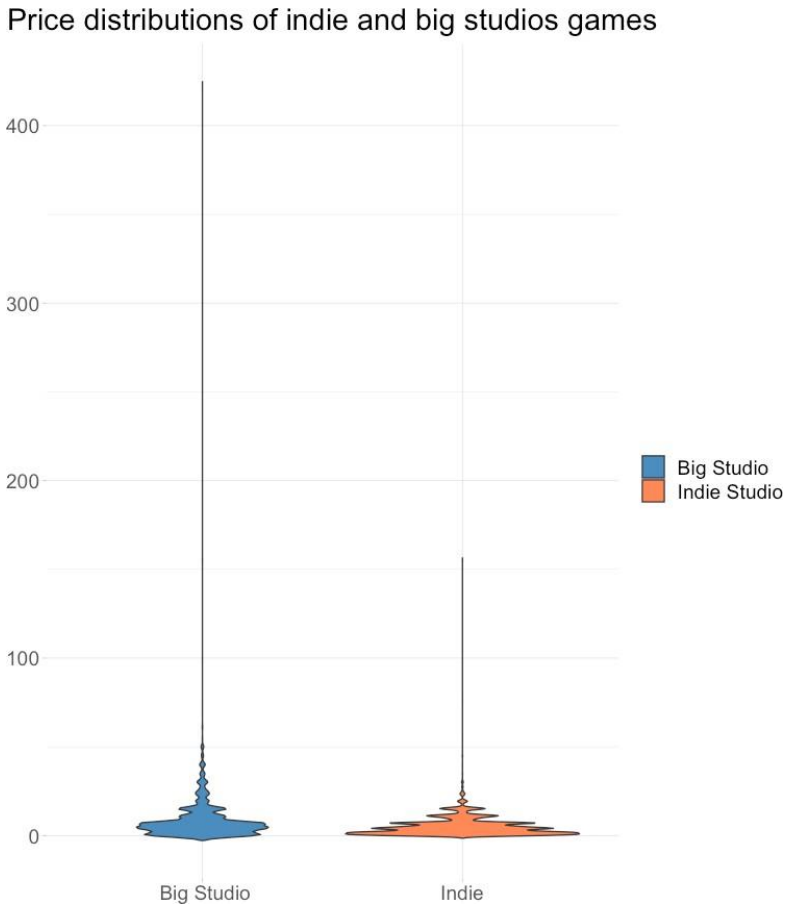


Figure 4: Violin chart of price distributions of indie and big studios games

Looking at Figure 4 now, the violin chart demonstrates the price distributions of indie and big studios games and indicates that the price of indie games is likely to be lower than triple-A games from big studios. It is fair because the fund spent on AAA games is frequently significant high, and not to mention the marketing fee for them. And, gamers can accept that the quality of indie games would not be as good as AAA games, but fun and have funny ideas. Most of the indie games have prices under 20 pounds, and the highest price was around 150 pounds, while it was over 400 pounds in Big studio type.

Game Genres

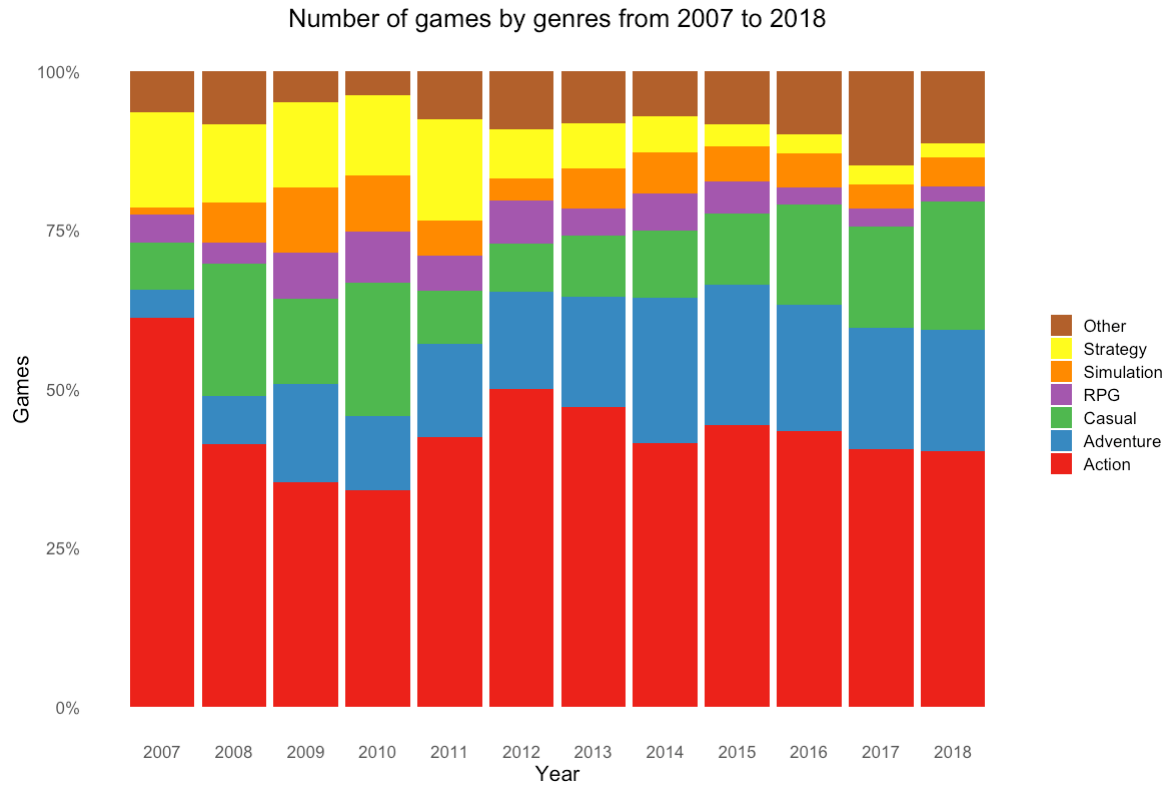


Figure 5: Proportion of game genres

The dashboard indicates that there are 120 game genres on Steam. However, only 6 game genres have a powerful impact on the number of released games and other statistical indicators which are Action, Adventure, Casual, RPG (Role Playing Game), Simulation, and Strategy. The proportion of Action games illustrated in Figure 5 dominates other game genres that began with around 60% then fluctuated and sat at roughly 40% in 2018. Adventure, Casual, and Simulation had seen increments during the period, while RPG and Strategy suffered the opposite trend. Other genres had growth, but it was just a slight increase which may be explained by the appearance of new genres by time.

Gameplay mode

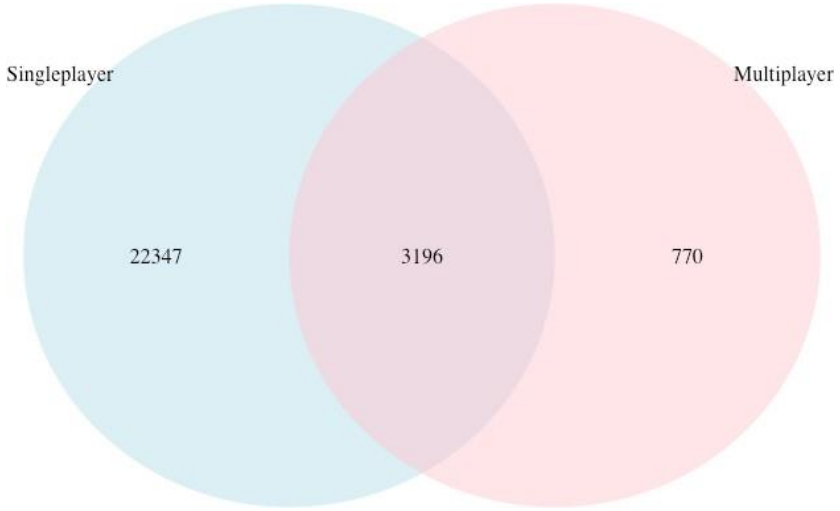


Figure 6: Venn diagram of the number of games in 2 game modes.

Venn diagram in Figure 6 indicates that there are 770 games on Steam having only Multiplayer mode, so-called fully online games. Moreover, there are 25,543 games support Singleplay mode, and just an eighth of them have Multiplayer mode. Hence, the typical approach for game developers is to begin with Singleplayer mode then Multiplayer.

Achievement system

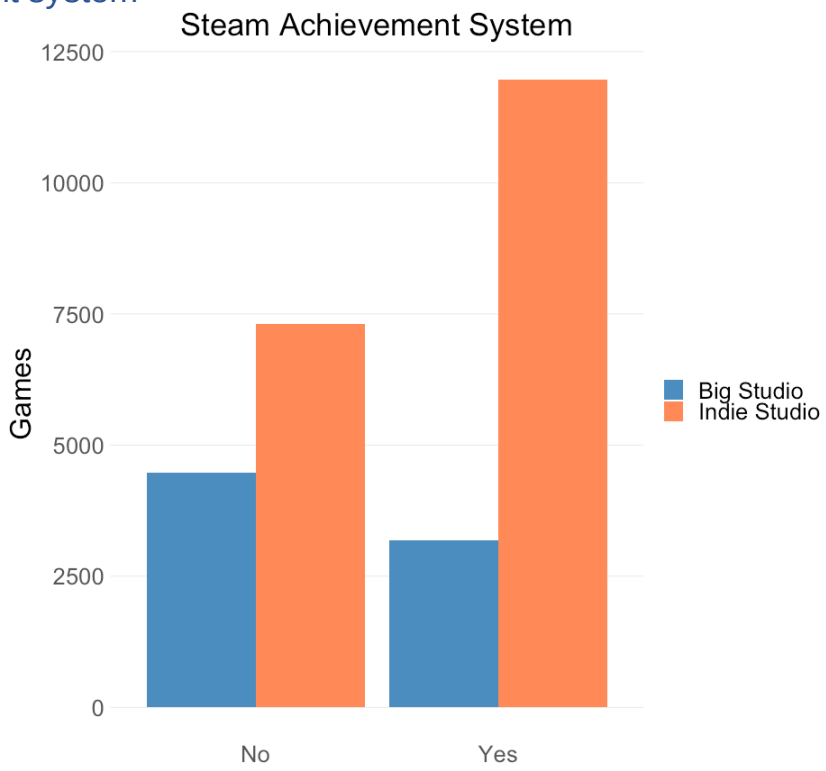


Figure 7: Grouped Bar chart of the numbers of games adopted Steam Achievement System.

The grouped bar chart in Figure 7 implies that around two thirds of total games on Steam have adopted the achievement system of Steam. The achievement in Steam enriches user experience, provides trophies in games, and especially, achievements and trophies unlocked will be shown in user social account on Steam.

Average playtime

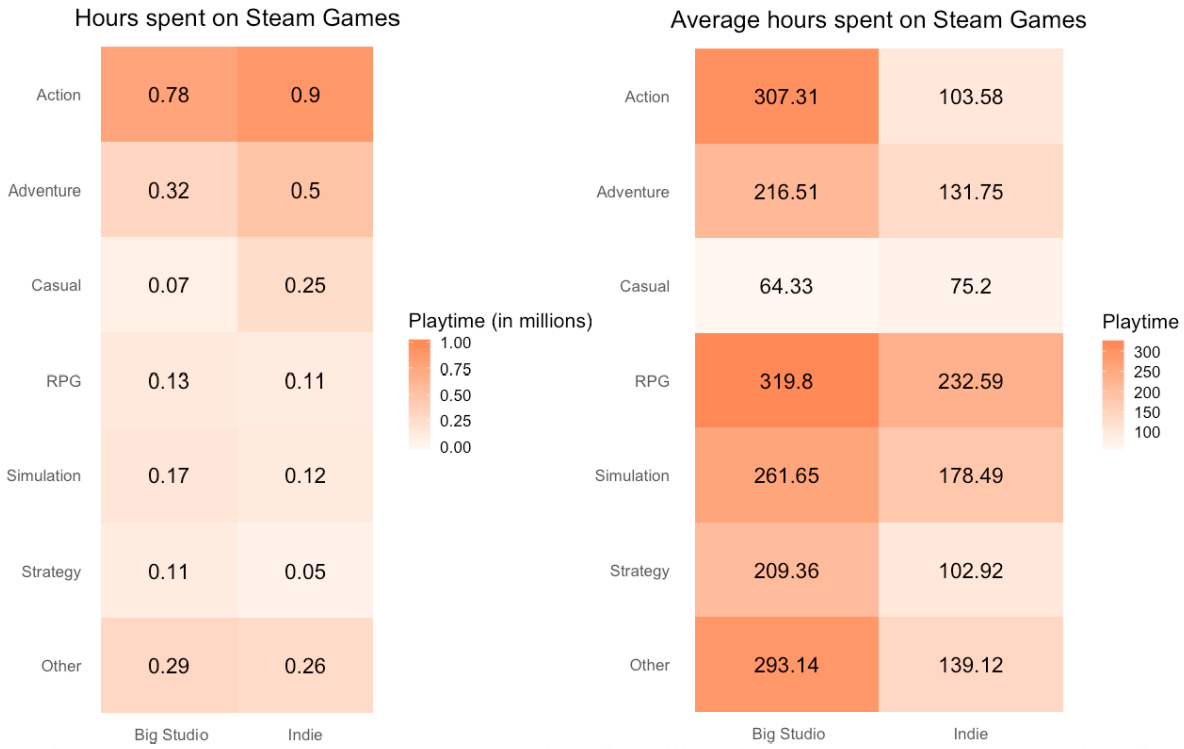


Figure 8 - 9: Heat map of user playtime in million hours and average playtime by game genres and studio type.

Indie games and Big studio games share a similar pattern in user playtime shown in Figure 8. Users spent most of their time on Action games in both types, with 0.78 million hours on Big studio games, and 0.9 million hours on Indie games. Noticeably, users spent the lowest amount of time on Casual games made by Big studios, while the lowest number in Indie games belongs to Strategy genre, with 0.07 for the former and 0.05 for the latter.

However, the average hours spent on Steam games in Figure 9 tells a definitely different story. Users were likely to spend time around two times higher on playing games made by Big studios than by the indie developers, and only Casual witnessed the opposite trend, but the difference is minor.

User Sentiment

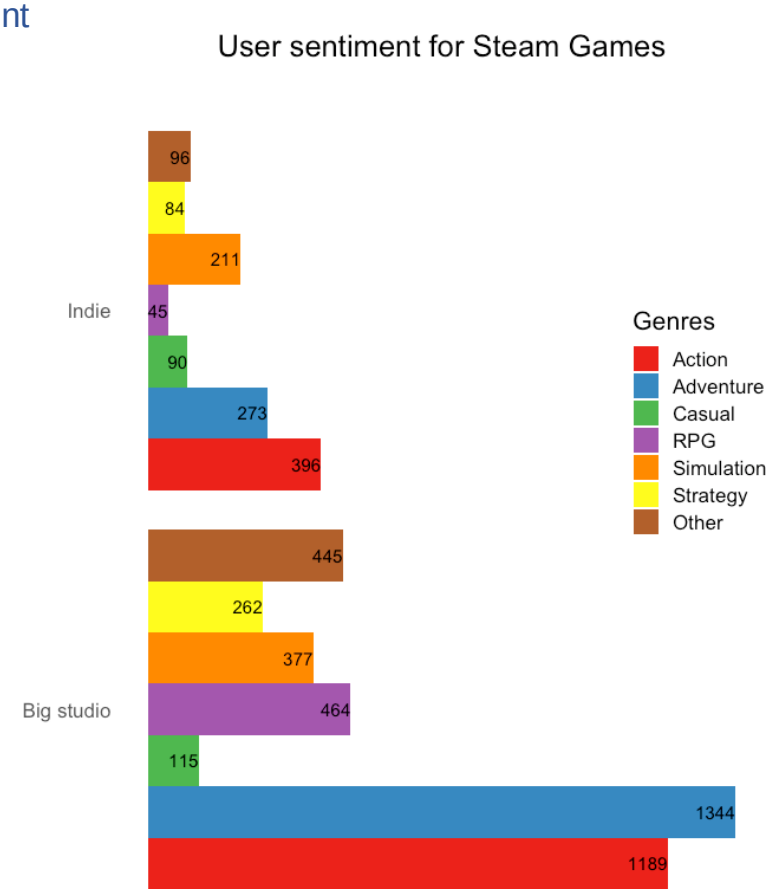


Figure 10: Grouped bar chart of user sentiment

Since games made by Big game studio were frequently carefully made, so it is clear from Figure 10 that the compliment of gamers for those games is significantly higher than indie games. Furthermore, some indie games were negligently made, so it leads to overwhelming numbers of negative comments on them. What stands out from the graph is that Action and Adventure are the 2 genres which have the highest sentiment in both types of game makers. Notably, Simulation has the third-highest number of user sentiment in Indie games.

Publisher

Figure: Word cloud of top publishers on Steam.

Figure 11 - 12: Word cloud of top Action game publishers on Steam.

From Figure 11, Big Fish Games had published the highest number of games on Steam along with Strategy First, THQ Nordic, and Ubisoft. Big Fish Games is a famous casual gaming studio based in Seattle, but not all of their games are labelled as Casual on Steam but other genres. Regarding the top hot genre, which is Action, Ubisoft and THQ Nordic appear again in top publishers shown in Figure 12. Japan is one among the biggest game markets and a giant in the gaming industry, and they have relentlessly raised the gaming industry. Therefore, many Japanese studios appear in the top publisher list.

Total Revenue

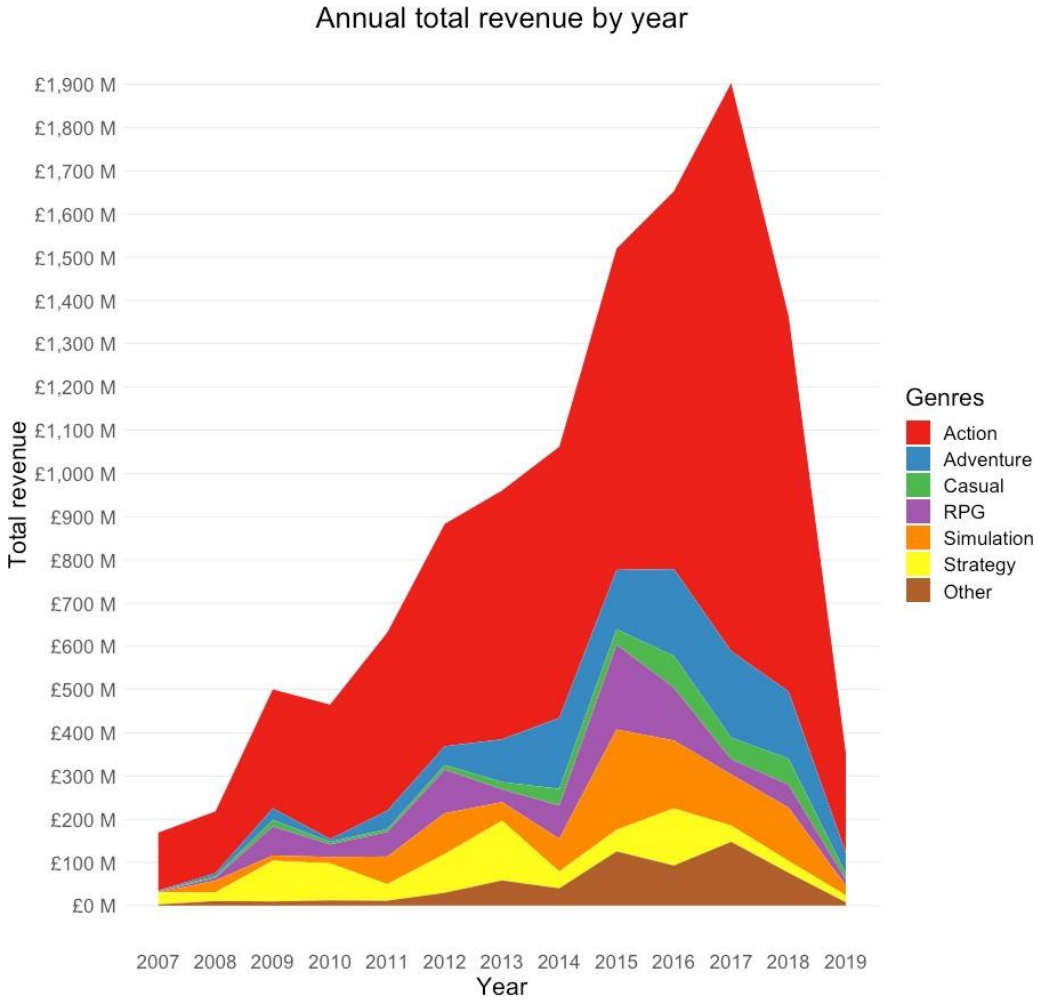


Figure 13: Annual revenue in billion pounds of Steam platform

As the stacked area chart in Figure 13 demonstrates, the revenue Steam platform has earned during the period from 2007 to 2019 soar from below 200 million pounds to around 1.9 billion pounds in 2017 before falling to around 1.4 billion pounds in 2018. Action is the genre that brought the most earnings for Steam because it had contributed at least 50% of the total revenue each year.

Owners

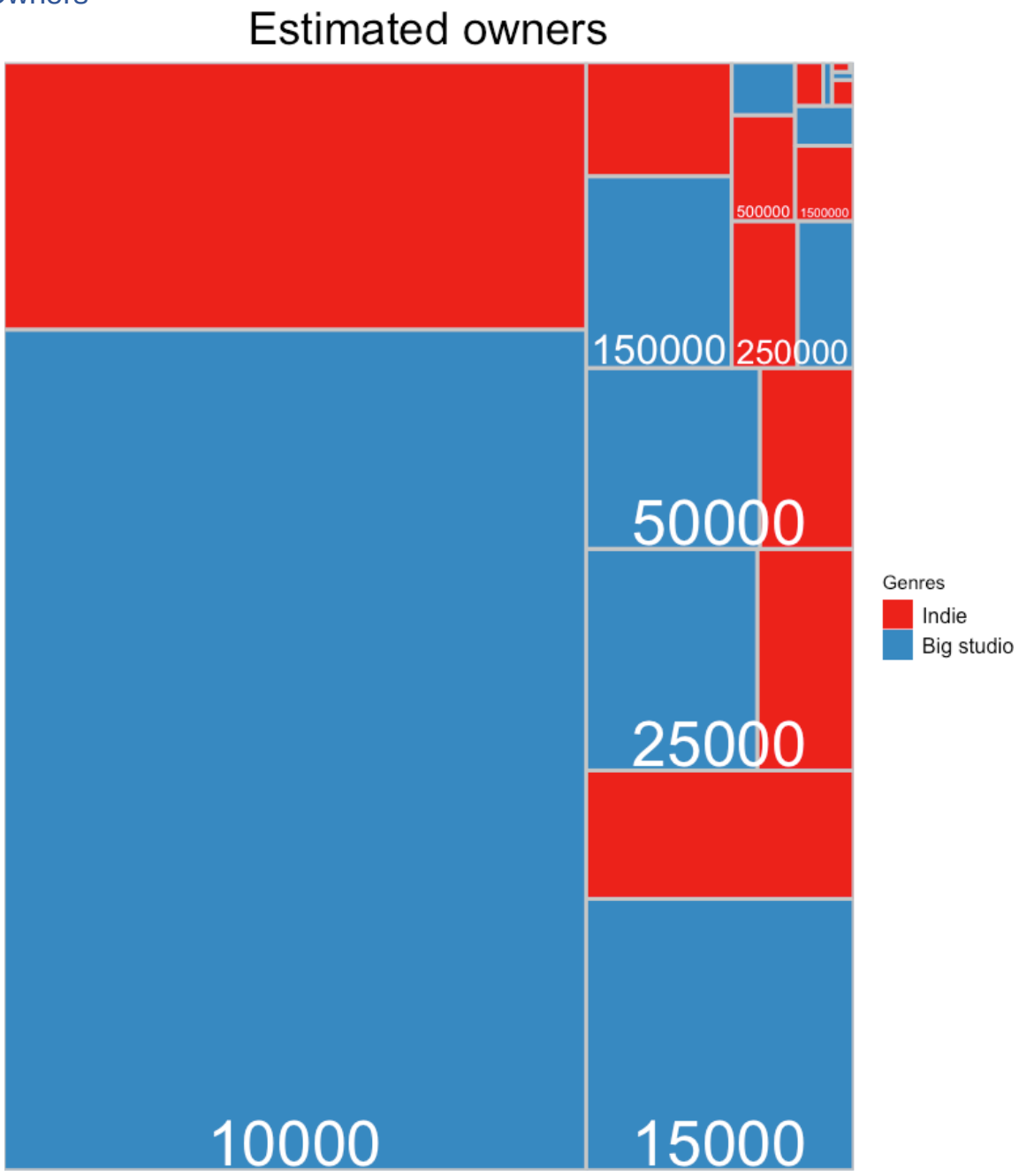


Figure 14: TreeMap of estimated owners by game genre.

The TreeMap shown in Figure 14 implies that about two thirds of total games on Steam had sold under 10,000 copies, and approximately a third of it was Action games. Action, Adventure, and Casual are the three top hot genres while they account for the most sale in every category. The proportion of games had the best sale, which has over 1 million owners is modest. It can give a hint about Steam platform that not all games are so good that gamers would like to have a copy of it.

Conclusion

Steam platform is indeed a compelling publishing platform to not Indie game developers only, but Big studios also. It is confirmed by the amazing numbers of games growing every year on Steam. Thanks for the emergence of game engines and the revolution of PC hardware that make game development be easier than it used to be in the past that required a significant afford in human resource, funding, and, importantly, production time. By 2012, the indie games had risen and became a contemporary gaming trend on Steam, and it accounts for around two thirds of total games published on Steam.

However, user sentiment indicator reflects the actual situation that not all indie games are good enough. Therefore, a small portion of them has a high user sentiment, while some of them even have zero sentiment. Once again, the evidence in the number of owners and average playtime confirmed that games from Big studios tended to have good sales and user sentiment than indie games did. Hence, indie game developers should try harder in game content to earn more user sentiment because there is a significant number of similar games on Steam and Steam users may be doubt about indie game quality than ever before.

Action is truly the top hot genre on Steam platform due to its substantial contributions to Steam revenue as well as developer earnings. A remarkable number of Action games had been owned by 1 million users and over, and gamers did spend billion hours on playing them. Besides Action, Adventure and Simulation are potential by reason of their growth not only in quantity but in other criteria.

Most games on Steam adopted the achievement system and had Singleplay mode, and just a small proportion of them have Multiplayer mode. It is unclear about the impact of the achievement system on the sale, but it is worth to adopt it to enrich the user experience.

In summary, if an indie game developer or a Big studio would like publish a game on Steam, 3 genres Action, Adventure, and Simulation will be the best genres to start with. Choosing a proper genre is the most important, but supporting Steam achievement and having a Singleplay mode can lead to a high user sentiment which means a good sale awaits. Pricing for games is analytical clear that under 20 pounds for an indie game and under 50 pounds for games from Big studios are often acceptable.

References

Davis, N 2019, *Steam Store Games (Clean dataset)*, Kaggle, viewed 18 Apr 2020, <https://www.kaggle.com/nikdavis/steam-store-game>.

Gough, C 2020, *Number of peak concurrent Steam users from November 2012 to March 2020 (in millions)*, Steam, Statista, Apr 19 2020.

Prescott, S 2019, 'The most popular desktop gaming clients, ranked', <https://www.pcgamer.com/au/the-most-popular-desktop-gaming-clients-ranked/>.