The Feasibility Of A Game Subscription Service

Elijah H. Harris, Antonio S. Maldonado
Colorado State University
CS455: Distributed Systems
Shrideep Pallickara

A. Introduction:

Video games are becoming one of the most popular forms of media consumption for a large portion of the world. The video game industry produced $43.8 billion in revenue a year in 2017(Shieber) with over 150 million gamers in the United States alone(Industry Facts). A new trend of subscription services that has arisen in other media consumption industries(music, TV, movies) is now making its way to the gaming industry. In this report, we will use our relevant backgrounds in business and computer science to analyze what a reasonable price would be for a game subscription service based on current game subscription models and speculate on what future services will need to do to become successful.

There are a multitude of different game subscription services already available to gamers all of which are slightly different. We will break down what makes each of the current dominant services unique and also look at pricing for each model to determine what impacts these characteristics can have on future services.

Steam is a video game market place where users can play games they purchase and connect with friends. Steam serves as a great place to pull statistical data from as they offer an open API for their platform and will serve as our primary database. Several researchers from Brigham Young University were to able to collect statistics on every user and game cataloged up until 2016(O'Neill). Various academic fields are focusing on studying the characteristics and matching behaviors of gamers, where studies have been shown up in the Journal of Computer-Mediated Communication and CyberPsychology, Behavior, and Social Networking. Many of these academic bodies use this research in order to understand the diverse gamer behavior and be able to link it to other relevant studies such as video game addiction and social networking(O'Neill).

In order to supplement the primary statistics found in Steam, we also analyzed the Video Game Sales(Smith, Gregory) dataset, which contains sales data from more than 16,500 games. The dataset contained statistics on sales based on region and included each entries' platform as long as the video game made sales greater than 100,000 copies.

After going into detail on how we obtained the data used in this report, there will be a breakdown how our values were calculated as well as assessing the accuracy of our results. This will also cover any normalization that was applied to our values as well as the reasoning behind it. There will also be sections pertaining to things we learned about the infrastructure of the software used to analyze the data and the results found in our analysis that weren't expected. Finally, we'll take a look at what the video game subscription service space will look like in the future and make conclusions based on our analysis of the data.

B. Problem characterization:

Video gaming is quickly becoming one the world's most popular pastimes with an estimated 2.5 billion people actively playing games in 2016. It should come as no surprise then that there are ludicrous amounts of money to be made in games with the market being worth $78.61 billion in 2017 with a projected growth to $90 billion by the year 2020 (wepc.com). With this much money up for grabs, game companies are being constantly pushed to innovate and create unique products and services that satisfy the ever changing needs of the industry.

Game subscription services are a relatively new subset of the industry that allows gamers to pay a monthly/yearly fee for unlimited access to a selection of games while they continue to pay for the service. This option is very attractive to gamers, especially ones that are just starting to get into gaming since the high entry prices of games and the consoles required to play them can be daunting.

These services seem streamlined and easy to use but in reality they are incredibly complex and technical programs that have to handle sorting through who has access to which games, managing licenses that provide access to the games and sometimes requiring the game to be streamed live to its users with minimal input lag and latency. The upcoming Google Stradia service touts being able to provide games running at 1080p 60fps streamed directly to a user's smart TV, google chrome tab or even their smartphone with a latency that's almost equivalent to running the game locally("Stadia").

Logistically, this sort of service is a nightmare to get right. Since most games rely on gamers being able to react as quickly as possible to ever changing surroundings, a game streaming services quality is directly affected by the lowest latency that it can achieve. Without a negligible amount of latency, fast-paced and multiplayer games quickly become unplayable which in turn discourages players from wanting to use the service.

Another technical hurdle to handle when streaming games is that games are usually a very resource intensive process on a computer disqualifying most consumer grade pc's from being able to run most modern games. While this is one of the main reasons streaming services are becoming popular, this poses a unique challenge to these services. With some games completely maxing out high-end hardware, even while running on optimised settings, the servers setup to run these games must be specialized to handle the high demand generated by millions of users wanting to stream games simultaneously.

Streaming isn't the only way to do a subscription game service though. Companies including Xbox and EA have come out with subscription services that allow users access to download and play games in their subscription libraries on the users local system. However, this method alienates those who don't already have the hardware to run these games and isn't ideal for situations where you might not have your computer/console and want to play games.

C. Dominant approaches to the problem:

There are currently three different types of game subscription/streaming services: PlayStation Now, Xbox Game Pass, and Humble Bundle Monthly.

PlayStation Now is a game streaming service from Sony that allows the user to play a wide variety of PlayStation exclusive games on a PC or PlayStation 4(Gurwin). This service is provided on a subscription or rental basis. The subscription options range from one month at $19.99 to one year at $99.99(playstation.com). The advantages of this service are that the user can both stream the games or download the games directly and also allows you to access over 750 PlayStation 3 and 4 exclusives, far more than any other service on this list. The disadvantage of this service is that on the games can only be streamed on PC, so if there is a poor internet connection, the visual quality and performance will drop.

Xbox Game Pass is a service different from other game streaming service in that it only allows the user to download the game directly to an Xbox One console instead of streaming the game. This service is provided on a monthly basis at only $10(Gurwin). The advantages of this service is that users can download the games directly and not be limited by a poor internet connections . The disadvantages of this service are that if the storage space on a computer is low, the downloaded games will not be able to be installed.

Humble Monthly is a game subscription service that gives subscribers a variety of PC games to the user each month. The subscription costs $12 per month and although the user does not know what each month's package will hold, the user does get to keep the game after the month is over(humblebundle.com). The advantages of this service are that the user keeps the games from every bundle. The disadvantages of this are that the user cannot choose which games will show up in each bundle.

D. Methodology:

Answering the question of what is a reasonable price for a game service subscription requires an analysis of key statistics like game price and user data. The most important question that needs to be answered in this analysis is how much the average user spends on games specifically on a per month/year basis. To do this we estimated the average cost of each person's Steam library and divided it by how old the average account is to get the average amount spent per month/year on games.

We also considered that some users may have games in their library that have never been or rarely played. This number helps calculate how much money a user can save if they didn't buy these games since the user would still have access to them through a subscription service. We decided the optimal number of hours for a game to be considered rarely played would be two

hours since that is the amount of time you can play a game on Steam and still be able to refund it. The cost per month/year gets multiplied by the percentage of games the user has actually played to create an adjusted cost per month/year.

The average cost of a user's library was normalized by not counting games that are free to own. The reason for this is that a free game would not be included in a subscription so it shouldn't factor into what gamers would need to pay for a subscription.

There are also some values included in the final results that aren't necessary to calculate the final answer but were included because they show what the numbers look like if they hadn't been normalized. We think these values would be relevant to anyone that's considering a game subscription service.

Steam is one of the most widely used video game launchers with over 90 million monthly active users from around the world as of October 2018 (Boxer). Steam, a platform which only sells PC games without making any of their own, pulls in about $4.3 billion revenue or 10% of the video game industry's total revenue every year (statistica.com). Steam's large market share makes it a good platform to conduct an in-depth analysis given its diverse userbase and abundant user and game data.

The "Steam Dataset" created by researchers at Brigham Young University serves as our primary dataset to analyze(O'Neill). The dataset was originally created in 2016 for the purpose of analyzing gamer behavior and includes key statistics that are needed for this analysis. Data used in this analysis that came from this dataset includes the price of every video game for sale on steam, how long each user has played each game, how many games each user has in their Steam library and much more. This dataset was chosen over others that were similar due to the completeness of the data and that it was created using SQL making it easier to parse with Apache Spark.

The primary database consists of 10+ sql tables totaling over 160gb of data that provides various statistics about Steam's users and games. For this analysis we only need data from the App_ID_Info(pricing), Games_2(play time) and player_summaries(account age) tables. This makes it would be unnecessary and time consuming to load the entire dataset into our mySQL database so only these three tables were loaded.

Apache Spark parses data in a similar way to how SQL databases parse their data. This combined with Spark's robust JDBC functionality makes it the perfect environment to analyze a large amount of SQL data like the primary database. Inside Spark, all that's needed to read from an external mySQL database is an SQL connector. It's then possible to capture tables or even the results of SQL statements in Spark dataframes which can be manipulated into relevant data easily.

The "Video Game Sales Dataset" is the secondary dataset analyzed to supplement the data in the Steam dataset(Smith, Gregory). The secondary dataset contains data on video games with at least 100,000 copies sold. This dataset also contains data on videos games released between 1980 and October 26, 2016. The fields in the dataset included the names of the video

games, the platforms that the games were released on, and the number of sales in separated by area such as North America, Europe, Japan, and other countries.

The secondary dataset was analyzed using an Apache Spark application that mainly processed the data using RDDs. The dataset, which was in a CSV format, was placed in the local HDFS. The dataset in HDFS was used to define a base RDD that contained every field from the CSV. Since only the video games that were released on PC were relevant in this case, the base RDD was filtered and transformed into another RDD by checking to see if the "platform" field contained "PC".

From here, it was necessary to grab the sales from each region category, so the cache method was called on the RDD to create a checkpoint for when the lineage of the RDD branches out. Every value in each column (North American sales, Europe sales, Japan sales, other sales, total global sales) was mapped out into their own RDD using a parseDouble method we created to attempt to convert the string values into doubles. If the values could not be converted into a double due to formatting issues, it would set the current value to 0.0 when mapping to the RDD.

For each region RDD, a reduce method was then called to grab those values and find the sum for each specified region. Since the reduce methods only returned a single double value instead of an RDD, we were unable to simply use the saveAsTextFile method, so we wrote to the local HDFS after specifying a Path and using spark.sparkContext.hadoopConfiguration to grab the filesystem.

E. Experimental Benchmarks:

After running the spark application for the Steam dataset, we were able to collect the following data. The average price of a game when the dataset was last updated that included free games was $10.61, whereas the average price without including free games was $11.92; however, the average price may be skewed due to Steam having multiple sales throughout the year. The average playtime of any game on Steam was 1390 minutes or 23.2 hours. The average time users have spent playing games was 359.9 hours. The average number of games in a user's library is around 15. The average percentage of games played for less than 30 minutes in a single user's library was 62.55%, whereas those played for less than or for two hours was 71.23%. The average cost of a user's Steam library is $185.12. The average account age (with more than 7 days since creation) for each user is 1001.84 days.

From this data, we can also calculate several other statistics. If we take the average time users spend playing a game and average cost of an entire Steam library, we can calculate that the average cost per hour of playtime is $0.46/hour. From here, we can also calculate the average cost per day/month/year for a user from the average cost of a Steam library and average account age: $0.16/day, $4.93/month, $60.03/year when including free games in the library; $0.18/day, $5.54/month, $67.44/year when free games not included.

However, given that some users have games in their library they have never played, we can calculate the average cost over time for the games the users have played using the average percentage of games played for less than two hours. Including free games in the library, this would be $0.05/day, $1.42/month, and $17.27/year, whereas without free games, this is $0.05/day, $1.59/month, and $19.40/year.

After running the spark application for the Video Game Sales dataset, we were able to obtain the total PC game sales for each of the different regions for games released between 1980 and 2016. In the spark application, we ran into errors when attempting to parse the dataset sale values from string to double. So, we made a method that tried to convert the string into a double, and if it could not, it would set that value to 0.0 when mapping the RDD. This would mean that there were false negatives when computing the total sales by region, but the bad records were very few in number, so the data was still reliable.

The total PC game sales for each region are as follow: North America has 93.07 million sales; Europe has 138.98 sales; Japan has 0.20 million sales; Other countries not listed before have 24.76 million sales. We were able to add each sale value together to obtain the total worldwide sales at 257.85 million sales.

F. Insights Gleaned:

Regarding the Video Game Sales dataset, one of the things that we were not overly familiar with in the beginning was the order in which the base RDD that was defined by the external dataset. Since we only wanted the video game entries that had “PC” in the platform field and every sales value in the different region columns, we needed to transform the base RDD at least twice. In the beginning, this was not as apparent, and we spent a while trying to map all of the region specific sales and then tried filtering. The simple solution was for us to first filter the base RDD for “PC” in the “platform” field of the dataset and then persist this filtered RDD. The filtered RDD served as a checkpoint for mapping the sales data for each region category. From there, getting the total PC game sales for each region was simple.

We hypothesized from the beginning that there would be many Steam users who did not completely play all of the games in their library from personal experiences with other users. It turns out that on average over half (53.18%) of all games in a user’s library are never played and that 71.23% of games are played for under two hours. Even by our greatest estimates we didn’t expect for such a large portion of games bought to go unplayed. This made for unexpected results when it came to calculating how much a gamer should spend on their games and gave necessity to creating an adjusted amount of how much gamers spend on average.

Something else we discovered when trying to parse through the primary dataset is that an SQL server can only deal with 151 connections at a time. This became an issue while querying the tables that had over 700 million entries in it since it turns out that the database available to us in the CS department is queried very often. The jobs we ran continually errored out at random

points due to this, no matter what the settings we chose were. After much trial and error, we found that our job runs optimally with only 10 total executors and a slightly lower partition count for the larger databases but still has the chance of erroring out if traffic on the database gets too high. Basically, we learned that if you're going to be querying a database repeatedly from multiple machines, you should be the only one accessing the database or you're gonna have a bad time.

G. How the problem space will look like in the future:

There are numerous technological advancements to be made for subscription services, specifically streaming services, that are currently under development. Google Stadia is a game streaming service that is currently under development and underwent a closed beta between 2018 and 2019(Smith, Matthew). Users are able to stream games nearly as smoothly as on a console as long as the user has a device with a web browser and a 25 Mbps download speed. However, when the game Assassin's Creed Odyssey was tested using a browser and a 25 Mbps download speed, the game did not look as visually appealing as it would if it were run on a console, with the game appearing to run at 720p resolution at most(Smith, Matthew).

From Google Stadia we can see how companies are focusing on allowing games to be accessible to a larger population, not just those with consoles or gaming PCs. This also means these types of technological advancements will allow for a move away from direct downloads of games and instead encourage a streaming/subscription service model.

Despite having the technology to stream games to the masses, it will still come down to the consumers to decide whether or not these subscription services become the new normal in the future. How these services are implemented, marketed and priced will determine whether consumers will be willing to spend the money on them.

That being said, no matter how game subscription services evolve over time they will never fully take over the gaming market as there will always be a demand to full-on purchase games rather than subscribe to them. When you subscribe to a game subscription service you will only have access for as long as you pay the subscription or until the company decides to cancel the service/stop supporting the game. For gamers that put long hours into games (we found that gamers put 23.2 hrs on average into a game), losing their progress after losing access to the game would be immensely frustrating and unacceptable.

Another problem game services may face is an oversaturation of the marketplace. Many gamers like having all their games in one place, it's one of the many reasons Steam was so successful when it first came out(Yu). Given that subscription services are trying to target gamers on a budget, if there are too many different services gamers won't be able to afford all of them and may just end up choosing to buy games the traditional way.

H. Conclusions:

In our research we found that a reasonable price for a game subscription service would be anywhere from $1.59/month ($19.40/year) to $5.54/month ($67.44/year). We come to this conclusion based on what the average gamer would spend on games if they only bought games that they actually played and what the average gamer actually spends on games respectively. This is a reasonable estimate because gamers shouldn't have to pay more than they're already paying to play the games that they want to play.

Our data also showed that gamers don't play most of the games that they buy for very long. On average 73% of a user's games get played for less than two hours total and that 53% don't get played at all. These numbers would only increase for a subscription service with a vast library like Steam's and should factor into the cost that the user ends up paying. Using those percentages is how we calculated our lower bound for a reasonable price at $1.59/month.

This data also leads us to conclude that for most gamers, the game subscription services that are currently offered today are not worth it. The cheapest game subscription service that we surveyed was Playstation Now coming in at $99.99 for a full year. While that seems like a small price to pay for access to over 750 games, in reality we found most gamers only spend around $65/year on games. When you add the fact that PS Now is notoriously laggy and buggy which limits the availability of the games, it becomes apparent how bad of a deal the subscription actually is.

In this report we speculate that given the announcement of multiple game streaming services like Google Stadia, we believe the industry is heading more towards streaming games as a way to increase availability to consumers. Given our results, we believe that for a service like this to be an attractive option to gamers they will have to keep their prices affordable in order to survive. This will be harder for streaming services in particular due to the large startup and upkeep costs from buying and maintaining all the necessary hardware.

In conclusion, game subscription services in their current form are not worth the money spent on them. We found that gamers already waste a good chunk of money on games that they never end up playing, and that subscription services currently cost more than the average user spends on video games per year. With this large price differential to overcome, subscription services will need to lower their prices before more gamers will buy their products.

I. Total Results:

Primary Dataset Results:

1. Game avg Price,
    a. w/ free games: $10.61
    b. w/o free games: $11.92
2. Avg playtime of a game = 1390 min / 23.2 hrs
3. Avg total time users spend playing games = 359.9 hrs
4. Avg # of games in a users library = 15.53
5. Avg % of games never played = 53.18%
   Avg % of games played for <30 min = 62.55%
   Avg % of games played for <= 2 hrs = 71.23%
6. Avg cost of library = 1. * 4. = w/ 1.a.: $164.77, w/1.b/: $185.12
7. Avg cost/hr of playtime 6. / 3. = $.46/hr
8. Avg acct. age (player_summaries.dateretrived - timecreated) (acct's age > 7 days) = 1001.84 Days
9. Avg cost per month/day/year = 6. / 8.
   w/1.a.= $.16/day, $4.93/month, $60.03/yr
   w/1.b. = $.18/day, $5.54/month, $67.44/yr
10. Adjusted avg cost per month/day/year = (6. * (1 -. 5)) / 8.

    Never played:

        w/1.a. = $.08/day, $2.31/month, $28.11/yr

        w/1.b. = $.09/day, $2.60/month, $31.58/yr

    <30 min played:

        w/1.a. = $.06/day, $1.85/month, $22.48/yr

        w/1.b. = $.07/day, $2.08/month, $25.26/yr

    <=2 hrs played:

        w/1.a. = $.05/day, $1.42/month, $17.27/yr

        w/1.b. = $.05/day, $1.59/month, $19.40/yr

Secondary Dataset Results:
Total PC Game Sales between 1980 and 2016:
NA Sales: 93.07 million
EU Sales: 138.98 million
JP Sales: 0.20 million
Other Country Sales: 24.76 million
Worldwide Global Sales: 257.85 million

J. Bibliography:

1. Shieber, Jonathan. "Video Game Revenue Tops $43 Billion in 2018, an 18% Jump from 2017 – TechCrunch." *TechCrunch*, TechCrunch, 22 Jan. 2019, techcrunch.com/2019/01/22/video-game-revenue-tops-43-billion-in-2018-an-18-jump-from-2017/.
2. "Industry Facts." *The Entertainment Software Association*, 2018, www.theesa.com/about-esa/industry-facts/.
3. O'Neill, Mark, et al. "CONDENSING STEAM: DISTILLING THE DIVERSITY OF GAMER BEHAVIOR." *Steam Dataset*, 2016, steam.internet.byu.edu/.
4. Smith, Gregory. "Video Game Sales." *Kaggle*, 26 Oct. 2016, www.kaggle.com/gregorut/videogamesales.
5. "2018 Video Game Industry Statistics, Trends & Data - The Ultimate List." *WePC.com*, 27 Apr. 2019, www.wepc.com/news/video-game-statistics/.
6. "Stadia." *Google Store*, Google, 2019, store.google.com/us/magazine/stadia.
7. Gurwin, Gabe. "Here Are All the Game Streaming Services You Should Know About." *Digital Trends*, 5 Apr. 2019, www.digitaltrends.com/gaming/best-game-streaming-services/.
8. "PlayStation™ Now | Games." *PlayStation*, 2019, www.playstation.com/en-us/explore/playstation-now/games/.
9. "Humble Monthly - FAQ." *Humble Bundle*, 2019, support.humblebundle.com/hc/en-us/articles/210560328-Humble-Monthly-FAQ.
10. Boxer, Steve. "Steam Passes 90 Million Monthly Active Users." *Green Man Gaming Newsroom*, 25 Oct. 2018, www.greenmangaming.com/newsroom/2018/10/25/steam-passes-90-million-monthly-active-users/.
11. "Steam Sales Revenue 2017 | Statistic." *Statista*, 2018, www.statista.com/statistics/547025/steam-game-sales-revenue/.
12. Smith, Matthew S. "Google Project Stream Hands-On: Can Chrome Replace Your PlayStation?" *Digital Trends*, Digital Trends, 9 Oct. 2018, www.digitaltrends.com/gaming/google-project-stream-impessions/#/2.
13. Yu, Yun. "Steam – the Gaming Platform before There Were Platforms." *Digital Innovation and Transformation*, 23 Feb. 2017, digit.hbs.org/submission/steam-the-gaming-platform-before-there-were-platforms/.