### Executive Summary

Video games industry has gained a lot of popularity in recent times. Games are no longer just played by kids, it’s something for everybody. Video games have moved out of peoples’ rooms into international arenas where players compete to see who is the best and carve out careers for themselves as professional players. This report focuses on exploring business opportunities in the game industry for our clients based on expert analysis of Steam data. Our goal is to come up with suggestions and recommendations that will help the company to remain competitive in the market, know the customer pulse and develop appealing brands. Thereby, increasing revenue.

While analyzing the latest video game data, our team chose to focus on playtime. We feel that if a video game is played more, then the game is more likely to get recommended to friends. It also increases the chances of in-game purchases hence increasing engagement and forming preferences of customers. We believe that this would compel people to purchase more games, driving up sales revenue.

Game playtime has increased tremendously over a period of five years from 2014 to 2018. Our analysis is to deep dive into the factors that affects playtime. We are specifically interested in how factors like ratings, price, platforms, genre and game categories could influence playtime.

As the video games industry is changing rapidly, it is essential to keep up with the trends and preferences of the customers to increase sales. The thing that surprised us the most was the number of hours that players have put into playing video games. We have 898 video games that have been released between 2014 to 2018 years and have a total of 160 thousand hours approximately.

New technology has opened up avenues to explore, allowing players to connect with each other virtually and play a video game together. But players also prefer to be the hero of their game and play single player games. The average playing time for single player games is less than games which are either multilayer or both [(1)](#30j0zll). However, the lore of saving the princess and being the hero of the story appeals to the masses even today and motivates the people to play more. The highest playing time on Steam has been observed for single player games which is almost three times the playing time for multiplayer games [(1)](#30j0zll). Steam is releasing more single player games as compared to others [(1).](#30j0zll) Steam should focus on releasing less single player games in future and increase the number of multiplayer games.

Video games are more accessible to players, now more than ever. The availability of games on Windows, Linux or Mac has made it more convenient for gamers to download games without worrying about the compatibility with the system. We observed that there was a steep increase in the number of games launched on single platform from year 2017 through 2018 [(2)](#3znysh7). The decision to launch more games on single platforms is not justified by average playtime for games released on single platform [(3)](#2et92p0). We strongly urge Steam to release games on at least two platforms and increase accessibility of games to players thereby increasing sales.

Steam offers a variety of games in different genres. With over a hundred genres to choose from, gamers can play action video games to trivia games. We explored average playtime for top 8 genres. We can observe that Action games have the highest playtime followed by Casual and RPG games. Games in the adventure genres were found to have an average playtime of 86 hours which is low when compared to others [(4).](#tyjcwt) Action genre has the highest average playtime despite the low number of games released over time [(5)](#3dy6vkm). We would recommend steam to release more games in Action genre followed by Casual and RPG.

We analyzed price vs playtime and found that average playtime for the games priced between $21 to $30 was the most and was least for games priced between $1 to $10. In order to understand why price does not affect playtime, we compared the positive ratings at the same price groups [(6)](#4d34og8). This comparison of graphs led to a discovery that positive rating influences playtime. Great games will find their players and popularity of games in dependent on word of mouth.

As discussed, customers like to play games in some genres more than others and it is important to understand how genre can affect the playing time of a game. The average playing time for the top 8 genres does not differ significantly [(7).](#3rdcrjn) But the average playing time for the genre Adventure is 86 hours which is very less as compared to Action which has around 287 hours of playing time on average [(8)](#26in1rg). Steam releases games in a lot of different genres but the average playing time for those which are not in the top eight genres is very less. So, Steam should focus on releasing games in these eight genres.

If we look at other customer preferences, some customers prefer multiplayer games while others prefer single player games. People like to play multiplayer games for the fun and the challenge of competing against over real-life players but single player games have their own fan base. The difference between the average playtime for these games is significant and Steam should consider providing more flexibility to customers by having both multiplayer and single player choices [(9).](#lnxbz9)

Customer preferences is an important factor to keep in mind before releasing new games but the availability of games to the customers also plays an important role in determining the playing time. We already observed how the availability of games on multiple platforms such as Mac, Linux and Windows, increases the playing time for those games. The average playing time for the games available on single platform is significantly lower than the games which are available on at least two platforms [(10)](#35nkun2). Steam should release games on multiple platforms and make the games more available to customers to increase the playtime.

Another factor to consider before releasing the games is its price. We discovered that the games which are free to play or are under 10 USD have less playtime as compared to games above 10 USD. As the price further increases, the playtime varies by a smaller amount. However, it is important to understand how significant this variation in the average playtime is. When we looked at the price of the games and how the playtime varies with it, we found that the difference between the average playtime is not significant as the price of games increases after 10 USD [(11).](#1ksv4uv) Price alone cannot be used to determine the playtime of the games and we need to consider other factors to better understand the variations in the playtime of a game.

We chose playtime as the topic of our interest, we want to determine which factors matters the most, which factors can be ignored and how these factors influence each other. We have already seen how different factors such as Genre, Price, Platforms, Categories can affect the playtime of games. The focus here is to analyze these factors and how they can play a role in predicting the playtime to further improve Steam’s products and services.

There is no significant linear relationship between genre and playtime. So, playtime cannot be predicted based on the genre [(12)](#44sinio). This further confirms that there is no significant difference in the average playtime for the top eight genres. The category of the games influences playtime and can be used for prediction. For multiplayer games, there is a 144 hour increase in the playtime. Games with both options have huge impact on playtime and playtime increases by 237 hours when compared to single player games [(13).](#2jxsxqh) The availability of the games on different platforms like Mac, Linux and Windows does not have a linear relationship with playtime. Hence, the availability of a game on various platforms cannot be used  in predicting the playtime of a game[(12).](#44sinio) Another factor that can have a major impact on  playtime is the price of a game. There is a significant linear relationship between price and playtime of a game and  can be used to predict the playtime [(13).](#2jxsxqh) As the price increases, playtime also increases [(14)](#z337ya). This might be due to an increase in in-game purchases and if customers like a game, they tend to buy its next versions irrespective of the price. Finally, positive ratings reflect how much a game is liked by customers. There is a significant linear relationship between positive ratings and playtime and can be used to predict the playtime [(15)](#3j2qqm3). With each additional positive rating we obtain for a game, the playtime is going to increase by 1.2 minutes [(13).](#2jxsxqh)To summarize, the category of a game if it is a multiplayer or it has both options, the price and its positive ratings will have a great impact on the game’s playtime. Hence, we need to focus our attention on these factors and improve the features for our upcoming releases to satisfy our customers.

**Conclusion:**

In conclusion, great games create their audience. Steam has a loyal customer base who have collectively logged in approximately 160 thousand hours of playtime. We would like to list few recommendations that will help Steam increase sales revenue and produce games that will tap customers’ potential and preferences.

1. More games should be developed in the top 8 genres.
2. Focus on releasing more games on two platforms as customer had more playtime on them.
3. Provide games having both the options to play as a single player and as a multi-player. This is more cost effective and addresses customer’s concern who prefer to play as single players.

We suggest that for further analysis the company analyze data on customers’ favorite features in each genre across games released. Gamers spend a lot of money on “in application purchases” and pay for features that give them a competitive advantage. This is a great way for the company to not only increase its revenue, but also capture more details about customer’s perception and preferences.

Secondly, we may also need to take a closer look at the impact of social media and YouTube influencers on popularity of games. Social media has been one of the fastest growing part of technology and has successfully engaged billions of people; these people are our potential customers if we maximize this front.

### Appendix:

#### Descriptive Statistics:

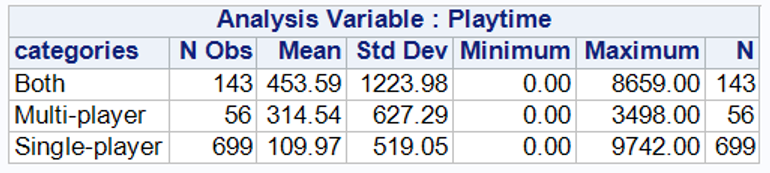


Figure 1

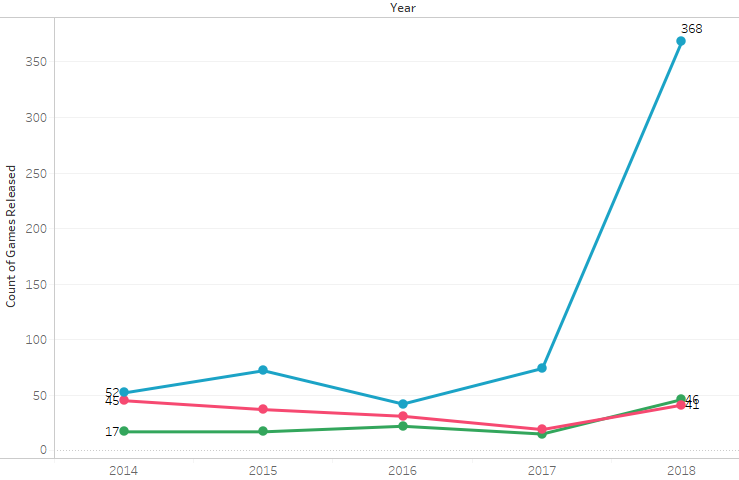
The average playing time for single player games is less when compared to multiplayer and both. But the maximum playing time for single player is almost three times than the multiplayer games. 



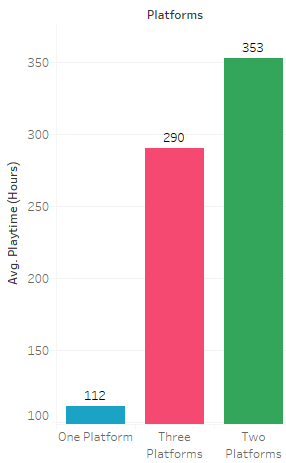
Figure 2

Figure 3

Average playtime across different number of platforms does not justify the decision to increase the number of games released. We observe a steep increase in the launch of games on a single platform from 2017 through 2018. We observed that games on 2 platforms had the most playtime and variability. However, more playtime attempts were on games available on one platform which did not really reflect on their playtime. In fact, games on one platform had the lowest mean play time.

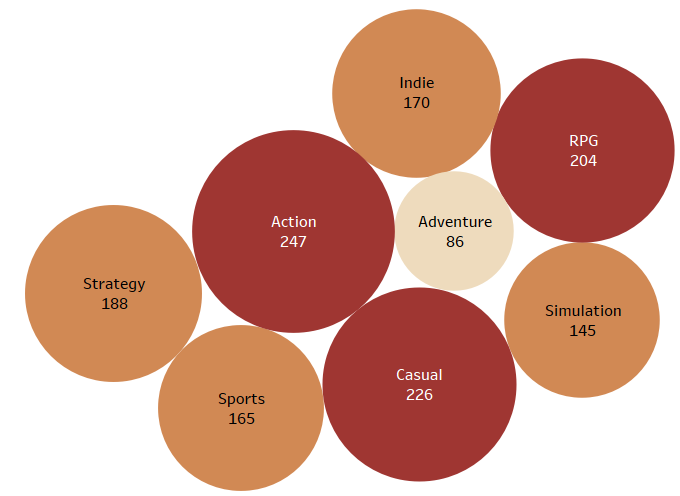


Figure 4

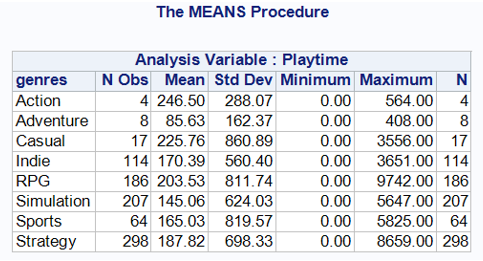
The average playtime for Action, Casual and RPG genres are the most. Therefore, video games belonging to these genres are popular.

Figure 5

Casual genre had the highest variability while Strategy had the highest number of players on it. Games with high playtime had few people who spent more time on those genres because it fits their preferences. Generally, we discovered that playtime was not just a function of how many people played but how well and long did these customers spend on the genre. We will explore this later in the report to find out if there is a significant correlation between playtime and genre using the top eight (8) genres.

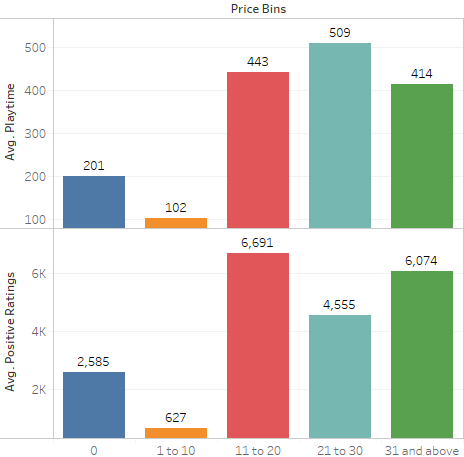
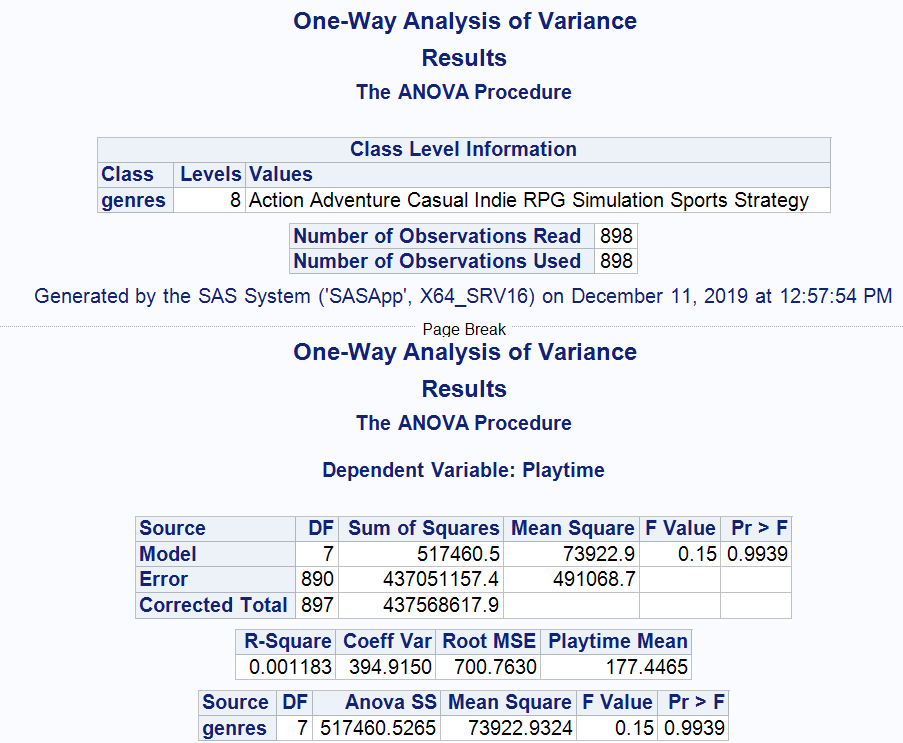


Figure 6

We can observe that average playtime and average positive ratings against price bins follow a similar pattern. In conclusion, positive ratings influence playtime.

#### Hypothesis Testing:

Genre

1. Null Hypothesis: The average playing time for the top eight genres is equal.

Alternate Hypothesis: At least one of the means is different.

1. Level of significance (α): .05
2. Test statistic: .15
3. P-value: .9939
4. As p-value (.9939) > α (.05), we fail to reject the null hypothesis
5. The average playing time for the top eight genres is equal

Figure 7

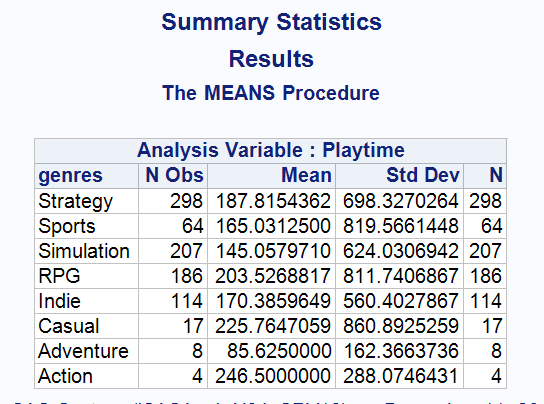
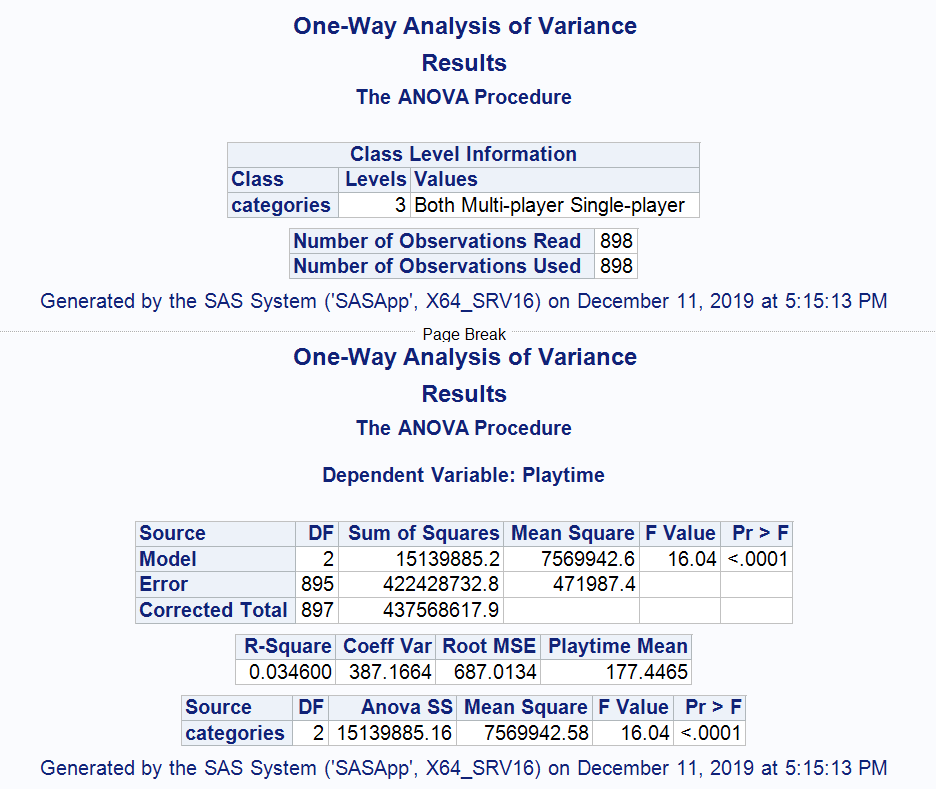
Descriptive Statistics for Genre:

Figure 8

The average playing time for Adventure is very less as compared to Action which has the highest average playing time. Adventure is in top eight genres of games played on steam. So, the other genres will have less than 96 hours of playing time.

Category:

1. Null Hypothesis: The mean playtime is same for all the different categories.

Alternate Hypothesis: At least one of the means is different.

1. Level of significance (α): .05
2. Test Statistic: 16.04
3. P-value: <.0001
4. As p-value (<.0001) < α (.05), we reject the null hypothesis.
5. The mean playtime is different for games with different categories.

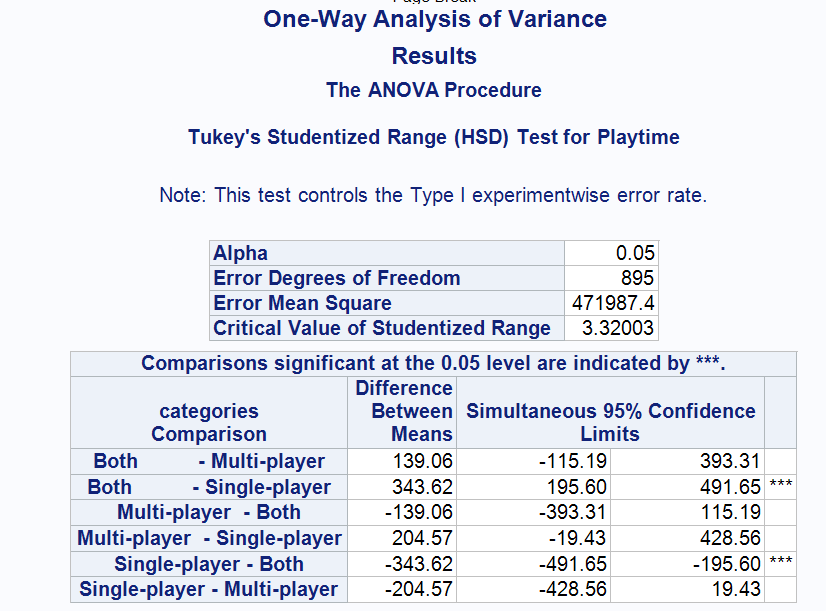
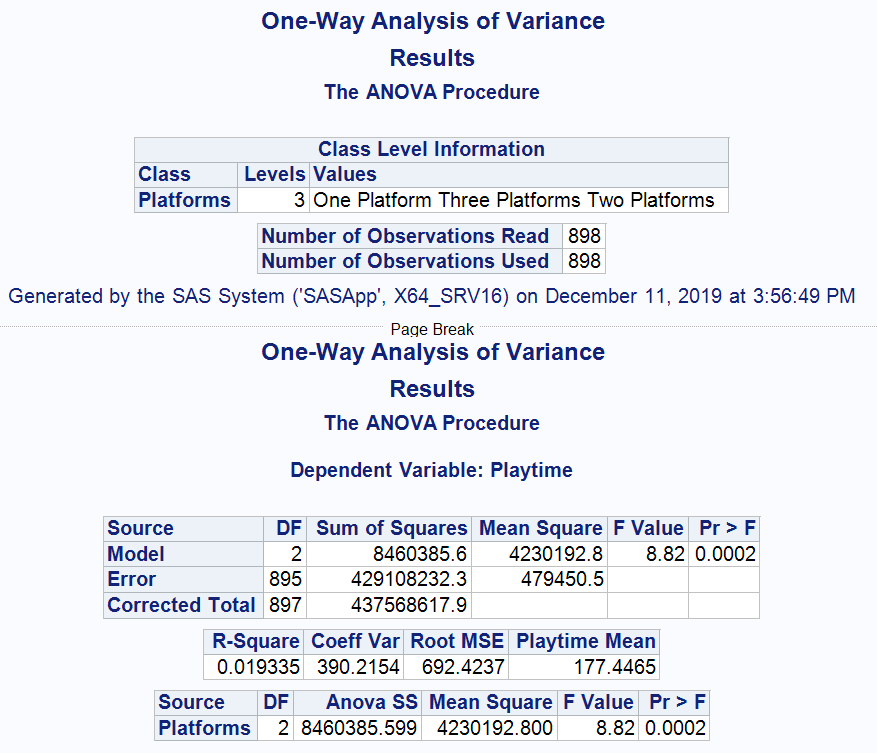


Figure 9

The games which are only multiplayer or both single layer and multiplayer have significantly higher playtime than games which are only single player.

Platform:

1. Null Hypothesis: The mean playtime for games available on different platforms is equal.

Alternate Hypothesis: At least one of the means is different.

1. Level of significance (α): .05
2. Test statistic: 8.82
3. P-value: .0002
4. As p-value (.0002) < α (.05), we reject the null hypothesis.
5. The average playing time of a game differ for different platforms.

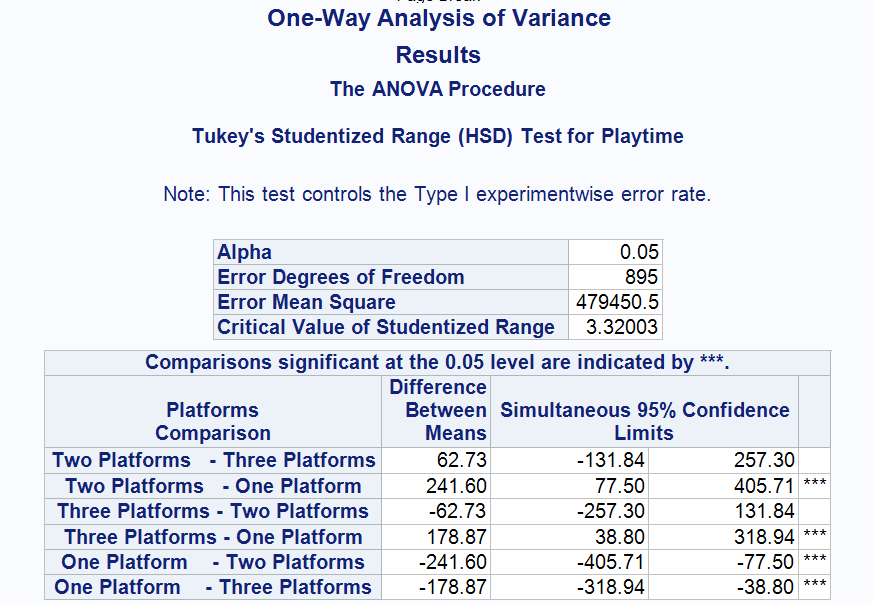
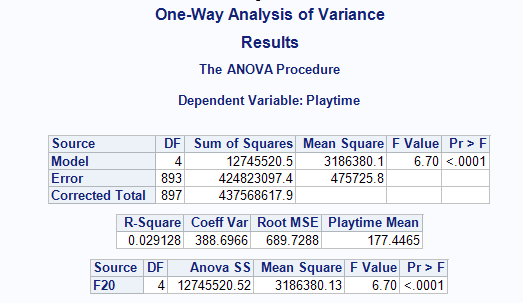


Figure 10

The games which are released only on a single platform has significant lower average play time than the games available on at least two platforms.

Price:

1. Null Hypothesis: The average playtime for games in different price range is same.

Alternate Hypothesis: At least one of the means is different.

1. Level of significance (α): .05
2. Test Statistic: 6.70
3. P-value: <.0001
4. As p-value (<.0001) < α (.05), we reject the null hypothesis.
5. The average playtime is different for games in different price range.

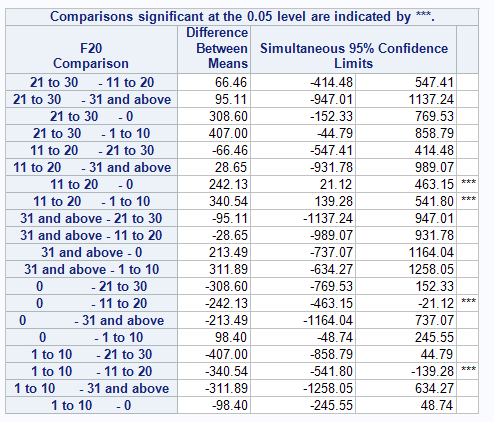


Figure 11

#### Regression Analysis:

**Full Model:** The full model includes all the variables of interest including independent variables genres, price, platforms, positive ratings, categories and dependent variable playtime.

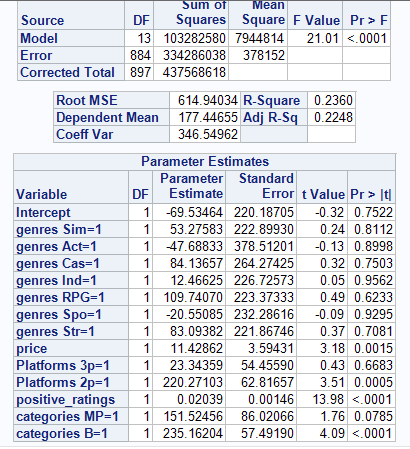


Figure 12

**Removal of Variables:**

* The variable platform does not explain a significant amount of variability in the playtime at .05 level of significance, so we decided to remove it from the final model.
* The variable Genre does not explain a significant amount of variability in the playtime at .05 level of significance, so we decided to remove it from the final model.
* For the variable category, categories B=1 is significant but categories MP=1 is not but we decided to keep the variable as it explains 3 % of variability in playtime. So, there is a significant difference in playtime for the games that are both single player and multiplayer than the games which are either single player or multiplayer.

**Best Model**

The final model includes independent variables price, positive ratings, category and dependent variable playtime.

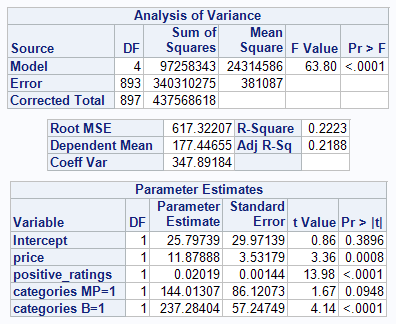


Figure 13

The model is significant at .05 level of significance as the overall p-value (<.0001) is less than level of significance (.05).

**Estimated Linear Regression Equation:**

bo=25.80, b1=11.88, b2=.02, b3=144.01, b5=237.28

Playtime = 26.77 + 11.88 X price + .02 X positive\_ratings + 144.01 X categories MP=1 + 237.28 X categories B=1

For a unit increase in price, there is a 11.88 increase in playtime.

For a unit increase in positive ratings, there is a .02 increase in playtime.

For category multiplayer, the playtime increases by 144.01 than single player.

For category Both, the playtime increases by 237.28 than single player.

**Explanation:**

Adj R-Sq: .2188

21.88 % of variability in playtime can be explained by the linear relationship between playtime and at least one of the independent variables.

Correlation between price and playtime:

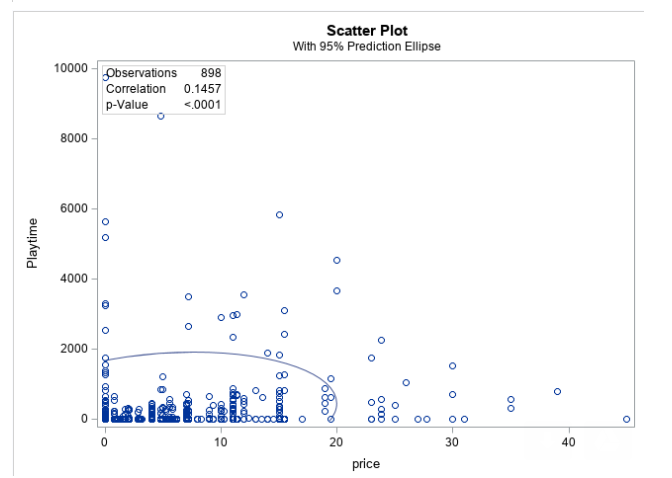


Figure 14

Correlation between positive ratings and playtime:

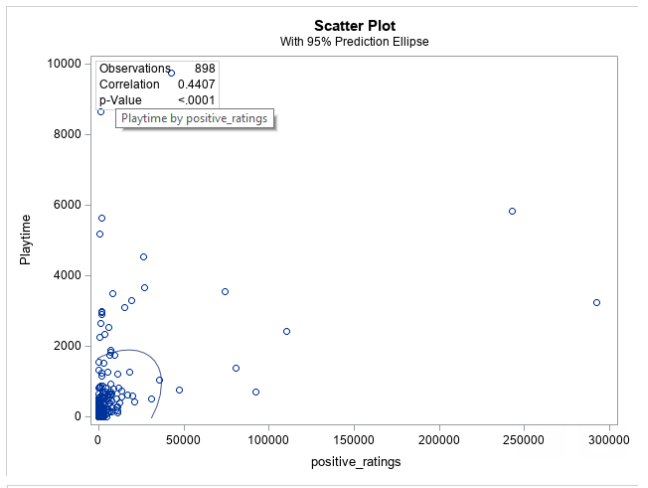
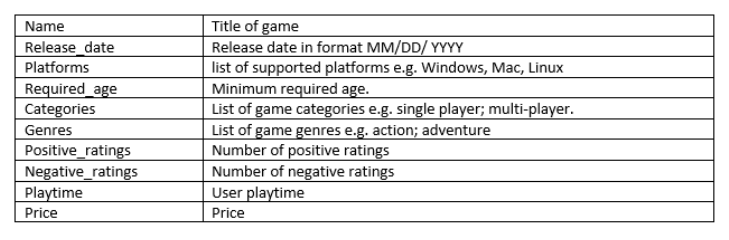


Figure 15

#### Data Dictionary:



#### Dataset Location: <https://drive.google.com/open?id=1f5sIpBXd5MlM4so49SJ5dGEF39dDken0>