* 36025955
* 35984301
* 35493311
* 35362177
* 36018505

MSCI581 Coursework

**Group: 12**

[Executive summary 1](#_Toc99949764)

[Preliminary data analysis 2](#_Toc99949765)

[Section: 1 3](#_Toc99949766)

[Empirical analysis: 3](#_Toc99949767)

[Statistical analysis: 5](#_Toc99949768)

[Section: 2 8](#_Toc99949769)

[Section: 3 10](#_Toc99949770)

[Dota 2 is F2P: 11](#_Toc99949771)

[Crowdfunding tournaments: 11](#_Toc99949772)

[Dota Plus and Dotabuff Plus Subscriptions: 13](#_Toc99949773)

[Section: 4 – Promotion Strategy 14](#_Toc99949774)

[Section: 5 15](#_Toc99949775)

Table of Figures

[Figure 1: Language of channels that played Dota 2 in the past 1 year 3](#_Toc99949741)

[Figure 2: Most popular games on Twitch by all-time views 3](#_Toc99949742)

[Figure 3: Most followed games on Twitch 4](#_Toc99949743)

[Figure 4: Players and Twitch viewers plot when events happened 4](#_Toc99949744)

[Figure 5: Players and Twitch viewers plot when events have not happened 5](#_Toc99949745)

[Figure 6: Simple regression summary output 5](#_Toc99949746)

[Figure 7: Simple regression residual output 6](#_Toc99949747)

[Figure 8: Dota 2 AR models AICc output 6](#_Toc99949748)

[Figure 9: Residuals of the AR (1,0) model which has the lowest AICC 6](#_Toc99949749)

[Figure 10: Diagram of fit and residuals of Twitch viewers and Players 7](#_Toc99949750)

[Figure 11: IRF plot from Players 7](#_Toc99949751)

[Figure 12: IRF plot from Twitch viewers 8](#_Toc99949752)

[Figure 13: Plot for original data, the fitted values and the Innovators and Imitators considering a beta version 10](#_Toc99949753)

[Figure 14: Plot for original data, the fitted values and the Innovators and Imitators 10](#_Toc99949754)

[Figure 15: Team Fortress - 2 VS Dota 2 VS CS-GO 11](#_Toc99949755)

[Figure 16: Battle Pass Owners 12](#_Toc99949756)

[Figure 17: International Prize money and players base relationship 12](#_Toc99949757)

[Figure 18: The International championship prize money 13](#_Toc99949758)

[Figure 19:Concurrent players with time, (Red) lines show "The International" event happening 14](#_Toc99949759)

[Figure 21: ETSX Model Residual analysis 15](#_Toc99949760)

## Executive summary

This report answers the questions about the Dota2 game. It starts with the preliminary analysis, and it discusses the impact of Twitch viewers on the players. In addition, it also captures the pricing and promotional strategies of Dota 2 and their effects on the player base and predicts the number of players for the next 14 days.

The data used for the analysis of the report was until March 1st, 2022. Steam dB was the main source for data collection, however, some references from various articles and Dota 2 channels were considered for acknowledging the real insights of the players. The report starts with empirical and statistical analysis. Regression, AR, and VAR models are created to answer the five questions in the discussion. Firstly, Twitch viewers impact insignificantly on players concerning the data, nonetheless, it is still recommended that they should have more presence on the Twitch platform because of its role in shaping the community for the games in recent half a decade. Furthermore, it is noted that the Twitch streamers and viewers are active members of the Dota 2 community. New heroes, patches, items, and builds make the game quite complex and thus continuously some professional players are producing new strategies to play, thus providing a simultaneous increase and decrease in the innovator-imitator relationship.

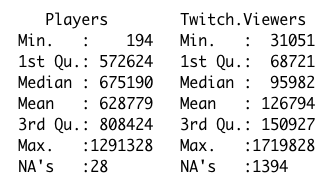
Thirdly, Dota 2 is a free-to-play game that attracts players to try the game and thus makes it the most effective pricing strategy. This strategy leads to in-game microtransactions which is an important revenue source for Dota2. It also offers subscription-based strategic assistant guides for new and existing players. Fourthly, Dota2 has an effective promotional strategy overall and it is based on sponsoring events in six regions across the globe under Dota Pro Circuit. To increase the player, count in the next 14 days outbound and contribution strategy is recommended in this report. Fifth, the next 14 days’ forecast of players is evaluated using four forecasting models and the best outcome was chosen based on the lowest value of MAPE at the valuation set, and ARIMA was chosen as our model. Later, it was compared with the actual values, and it was close to it.

## 

## Preliminary data analysis

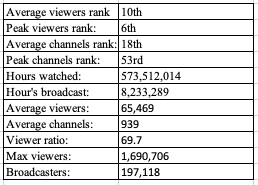
Dota 2 is a free-to-play multiplayer online battle arena video game developed and published by Valve and initially released date on 9 July 2013 where two teams consisting of five players must defend the Ancient - a protected structure located in their base.

Below is the quick summary of average active players from the date 22 Sept 2011 when it was in beta version to the 1st of Mar 2022.



Also, the twitch viewers we gathered from 17 July 2015.

Below are some key statistics, seeing it we can gauge how popular Dota 2 is on Twitch:



Dota 2 has a little over 11 million monthly active users worldwide as per the latest announcements made by the company. Between 2015 to 2016 Valve company which owns Dota 2 managed to get 3 million users, but the number decreases as soon as PUBG launches.

Language of channels that played Dota 2 in the past 365 days.

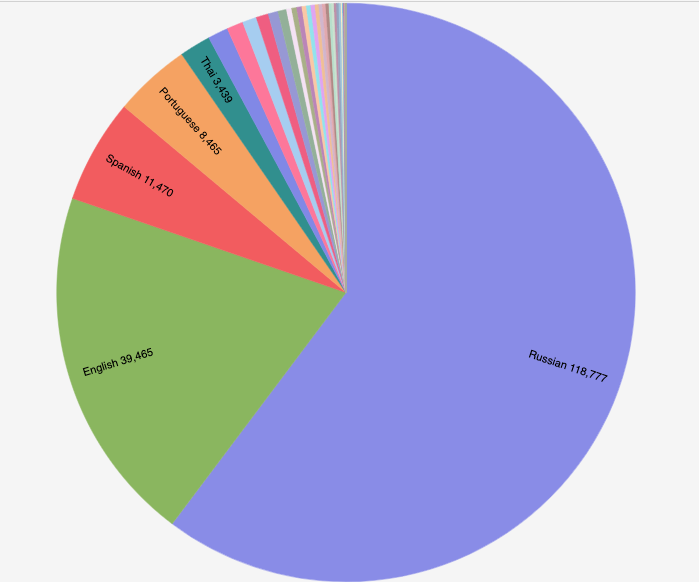


Figure 1: Language of channels that played Dota 2 in the past 1 year

## 

## Section: 1

### Empirical analysis:

Dota 2 has been the 5th most popular game on Twitch worldwide, with 19.09 billion views. Also, Dota 2 is the 8th most followed game on Twitch as shown in Fig. 2, and Fig 3 below, respectively.

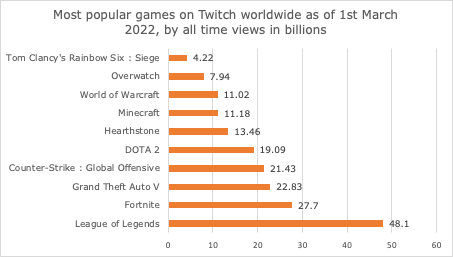


Figure 2: Most popular games on Twitch by all-time views

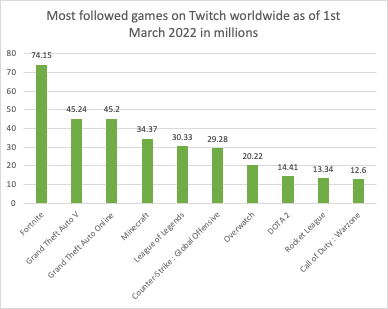
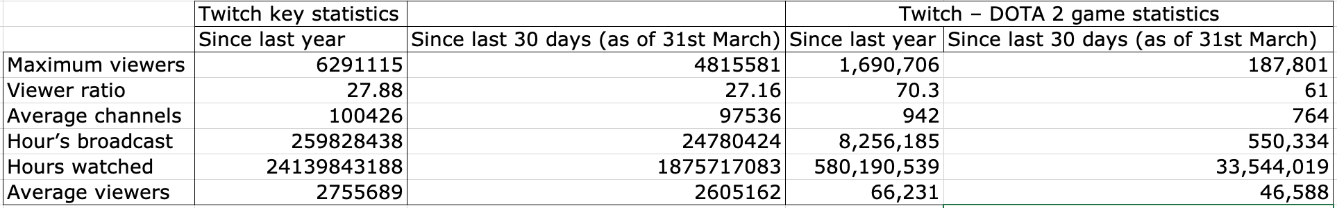


Figure 3: Most followed games on Twitch

If you refer to the table below, you will see the Twitch platform overall statistics for all games and the Dota 2 game on the Twitch platform.

Table 1:Twitch key statistics



Now we will assess whether the twitch viewers and the number of players playing Dota 2 are correlated? To answer this, we plotted 2 graphs as shown below between the twitch viewers and players plot dated from Aug 2011 till 1st Mar 2022, when there were Dota 2 championships held across the world and the second is when no championships were happening across the world.

When the twitch viewers curve is increasing, it is not increasing the number of players' curves. If you see at the end Twitch viewers are going down but it is not impacting the number of players as much. Between Aug 2018 to Aug 2019 if you see there is a spike in several players, but twitch viewers are following the same level.

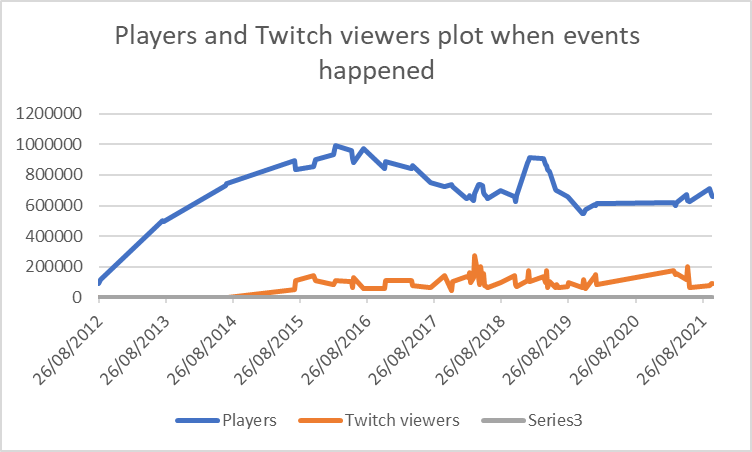


Figure 4: Players and Twitch viewers plot when events happened

In the below plot, no events were happening around the world, and we plotted the graph between the number of players and twitch viewers. There are significant spikes that could be seen in the twitch viewers, but it is not increasing the number of players at the same time. This could mean that twitch viewers are not the significant factor that impacts the number of players. There could be a correlation between the number of players with twitch viewers but empirically we could not prove whether there exists causation between them.

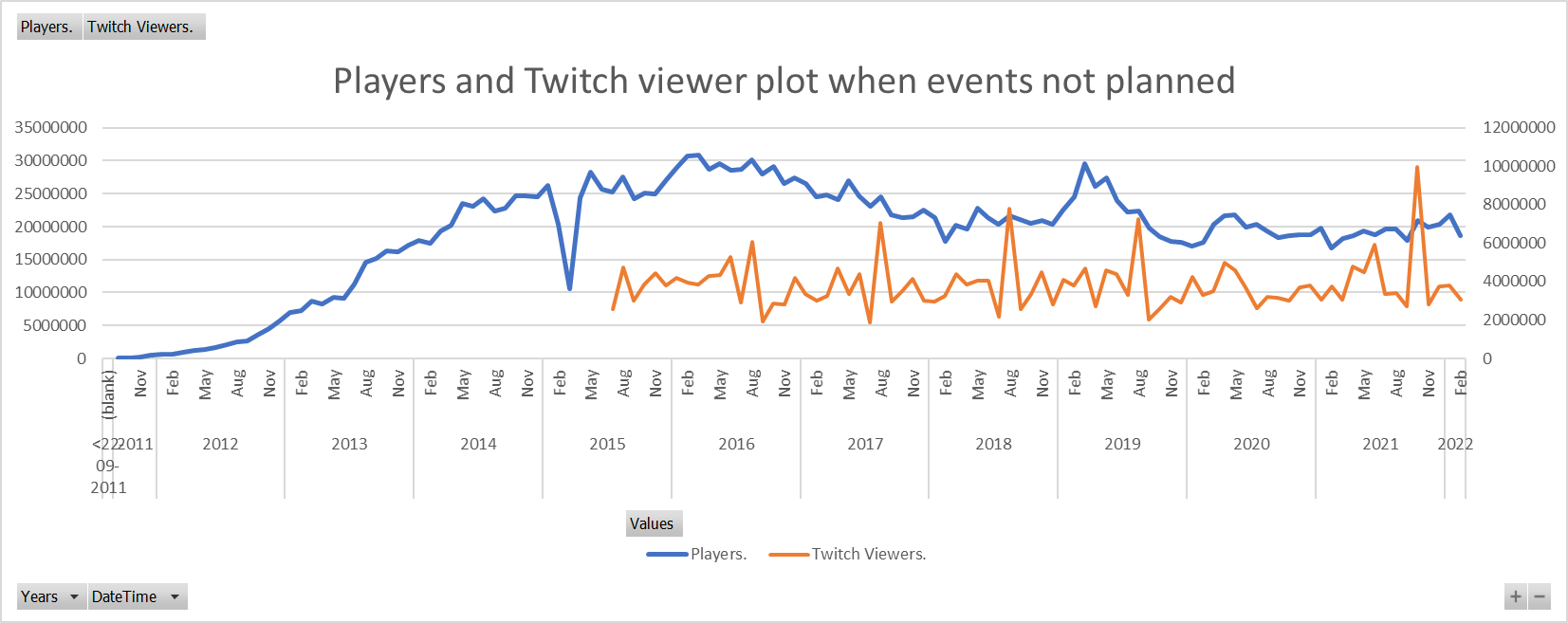
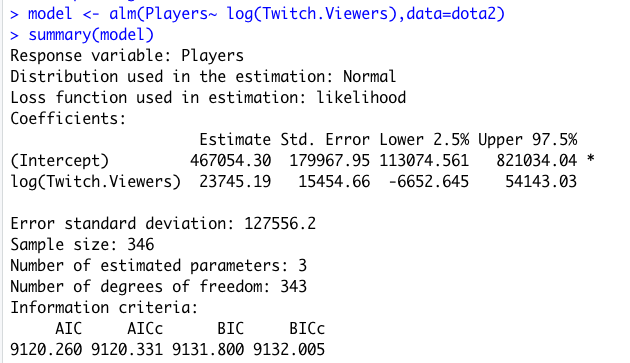


Figure 5: Players and Twitch viewers plot when events have not happened

To analyze the correlation-causation let us use statistical models, to do so we would build a simple regression – number of players as a response variable and twitch users as a predictor. We aggregated the daily active users on the weekly basis and then analyzed the models below to access the immediate and short-term impact of the Twitch viewers on the players.

### Statistical analysis:

Below is the output of the simple regression model. Estimate values of the coefficient are not that significant. The adjusted R square value is much less than – 0.003948, the p-value suggests the coefficient is not significant. The residuals have some serious issues. Standardized residuals are densely situated above the significance line.

Text

Description automatically generated

Figure 6: Simple regression summary output

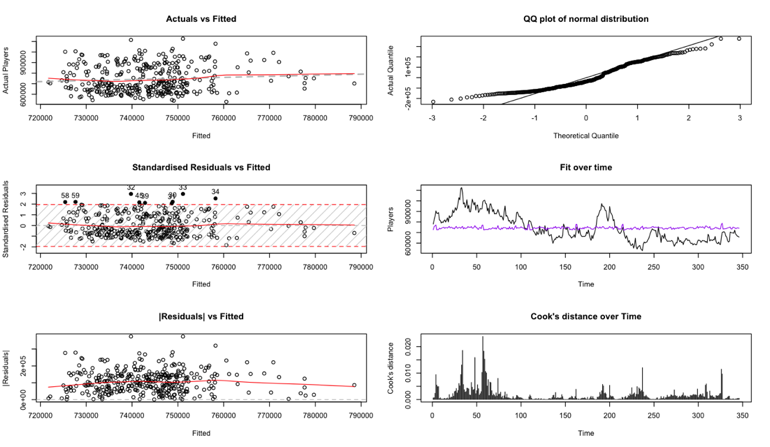


Figure 7: Simple regression residual output

Now this means alone twitch users have no direct impact on the number of players. We will now analyze the AR model incorporating the lag variable of twitch users as we supposed twitch users will not be having a real-time impact on the number of players but with some lagged timing.

Below is the AICc output of the models applied where lags up to were. The lowest AICc observed was with AR (1,0) which means an autoregressive lag of one week.

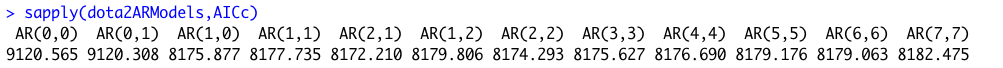


Figure 8: Dota 2 AR models AICc output

Below is the detailed output plot of the AR (1,0) model which we found has the lowest AICc among others.

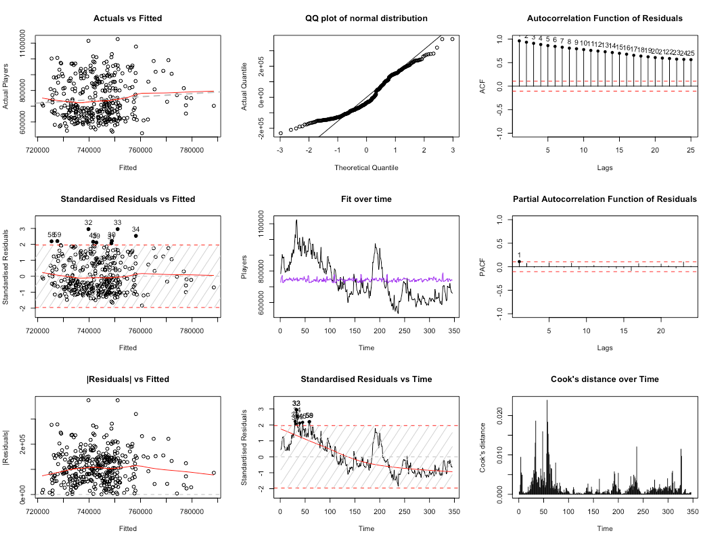


Figure 9: Residuals of the AR (1,0) model which has the lowest AICC

If analyzed the residuals for the model which has the lowest AICc, you would see QQ plot seems not to be with normal distribution, Residuals Vs fitted, residuals are widely distributed, in ACF seem a lot of lags are still correlated, In Actual Vs Fitted, lot of observations are above and below mean line and closely distributed the to the mean line. All depict this would not be a good fit model. That means that Twitch Viewers' lag variables do not explain the Players and could not be related.

On the final note, we checked the VAR model applying to the Players and Twitch viewers to find the correlation. Below are the fitted plot and residuals along with ACF and PACF functions for residuals. For the player's residuals PACF plot, some legs are still correlated but not significantly as these lags are under the significance level.

Figure 10: Diagram of fit and residuals of Twitch viewers and Players

In IRF plot for Players

Impact on Twitch viewers - The impact of Players on Twitch users is not so significant over the period. It impacted a little bit initially and then touched down to the zero level, again increases a bit and finally settling down.

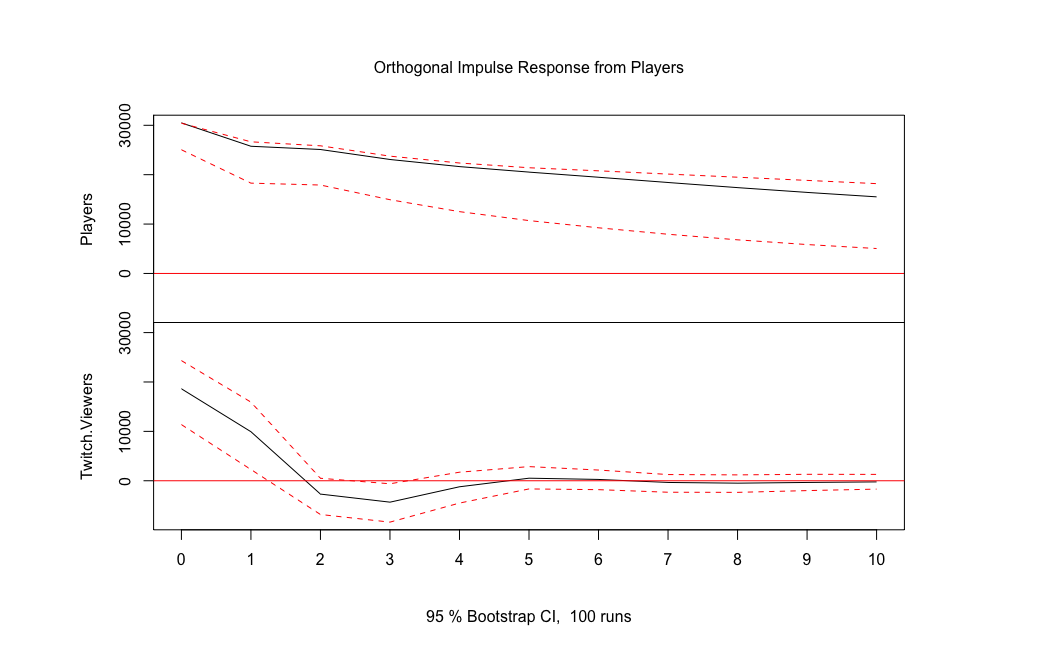


Figure 11: IRF plot from Players

In IRF plot for Twitch viewers

Impact on Players –The impact of Twitch viewers on the Players is not that significant, the line lies around the zero line this means that the impact could be very less or not at all significant. However, in the start, it started to increase gradually on a steady level but still not significant enough.

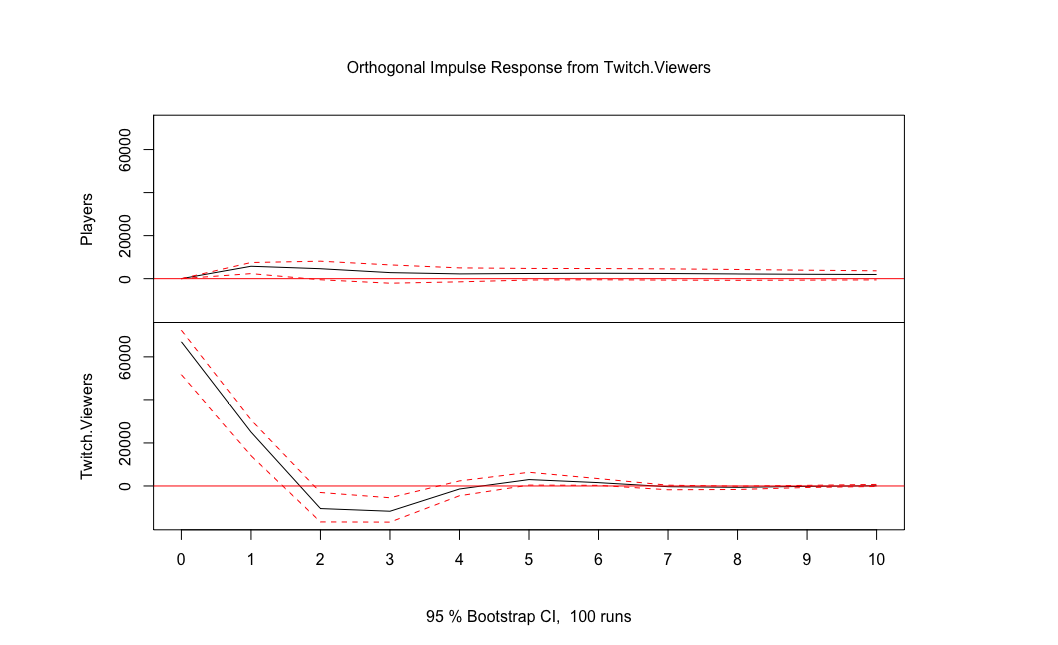


Figure 12: IRF plot from Twitch viewers

So statistically we could conclude that twitch viewers do not impact the number significantly, it could impact but in a very less manner. This could mean that if there are other games players could switch to Dota 2 when they would see an increase in Dota 2 streaming on the Twitch platform. Now should the company make efforts to increase the Dota 2 presence on Twitch? One reason could be diverting these other game players towards Dota 2. Other reasons could be engaging the community and elevating streamers that may not typically attract attention from big sponsors.

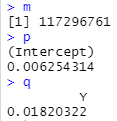
The twitch platform is helping to socialize the games industry, it is a powerful marketing tool, Twitch holds events that create awareness, also gives second chance at life to the old games and versions and most importantly Inspires the next generation of game creators. So yes, companies should encourage more viewers on Twitch even though it may not necessarily increase the number of players’ growth solely based on Twitch viewers.

## 

## Section: 2

To understand the imitators and innovators, the possible relation [coefficient of innovators] between p and q is explored in the Bass model, on this data. To explore the relation, the Bass model is constructed manually by creating a data frame with dota2 users and their cumulative values of the players by month.

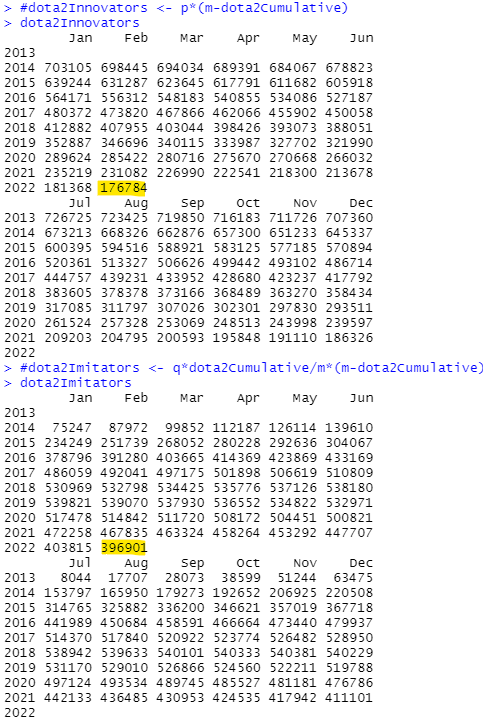
After that, the coefficients of the model are extracted by deriving the parameters of the Bass model based on the estimated parameters. p and q are identified using m (their relations with parameters of the model).



About the innovators and imitators of the product:

The innovation and imitation parameters are p = 0.006, q = 0.018. So, the coefficient of imitators is higher than the coefficient of innovators, which means that Dota2 has more imitators than innovators. In our context, more users start playing Dota2 only when innovators broadcast the game to others or even the latest updates. For the same reason, sites like Twitch and YouTube are important for imitators to get started with Dota2.

To produce fitted values of the Bass model and user volumes dues to innovators and imitators, the fitted values are extracted, and a time series object based on dota2 is created with it. Then the Dota2 user's volumes contributed by innovators and imitators are prepared.



The innovators and imitators at a time (February) are coloured in Yellow

Finally, the plot below is drafted using the original data, the fitted values, the innovators, and imitators' curves:

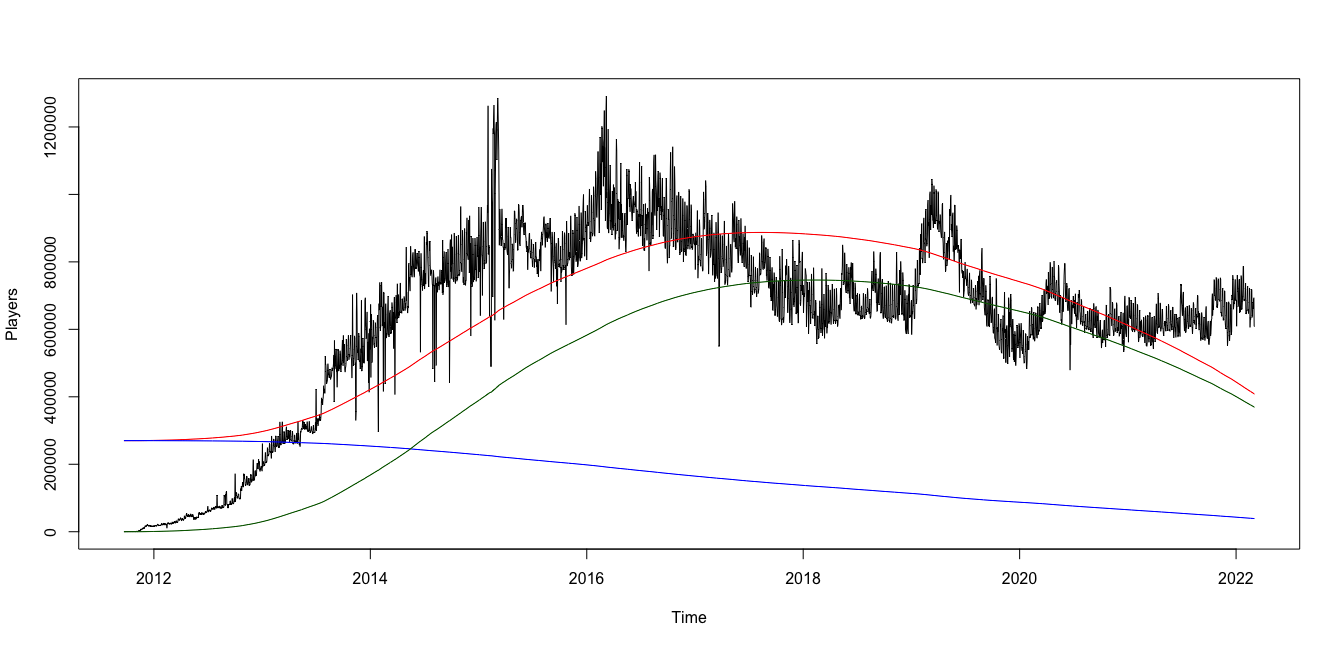


Figure 13: Plot for original data, the fitted values and the Innovators and Imitators considering a beta version

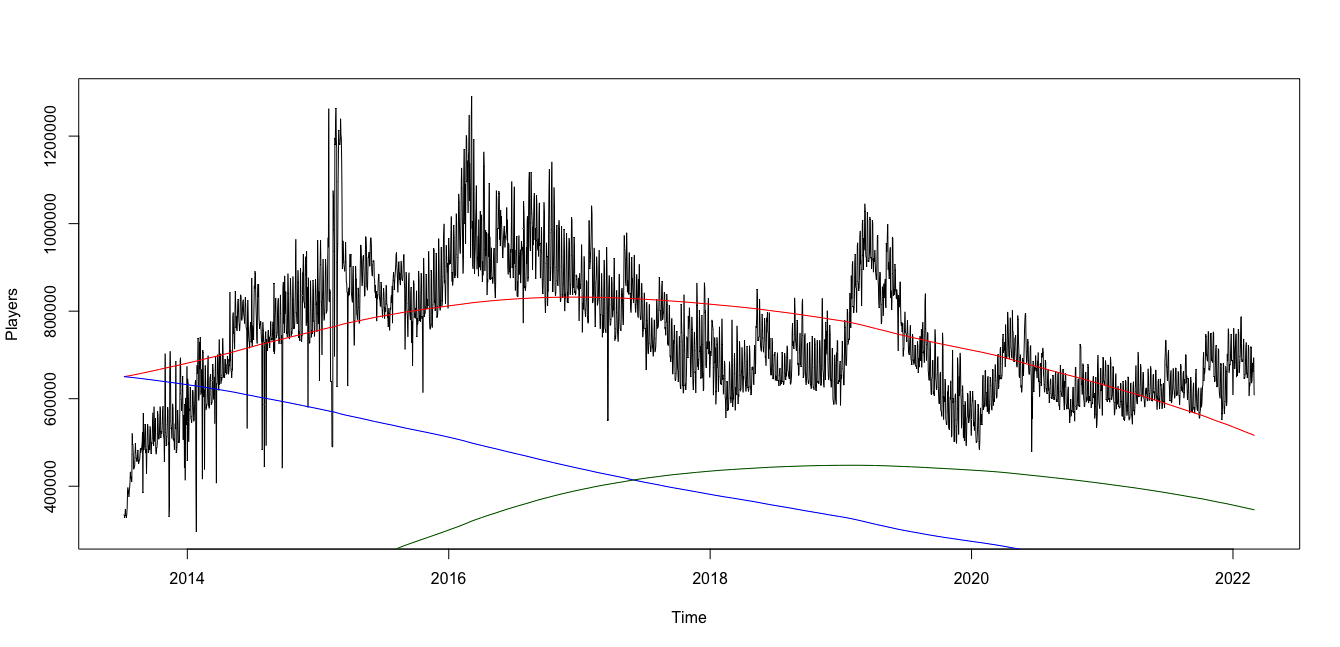


Figure 14: Plot for original data, the fitted values and the Innovators and Imitators

Alternatively, a diffusion graph can also be used

## 

## Section: 3

Dota 2 pricing strategy is divided into four components listed below:

### Dota 2 is F2P:

Dota 2 is a free-to-play (F2P) game this denotes that the entire game is free for anyone to play on Steam, and this acts as a primary pricing strategy for the game. For instance on Steam two games (CS-Go and Team Fortress) had changed their strategy from paid to free-to-play which demonstrates the applicability of the strategy.

Chart, histogram

Description automatically generated

Figure 15: Team Fortress - 2 VS Dota 2 VS CS-GO

### Crowdfunding tournaments:

Crowdfunding is the practice of funding a project by raising money from a large pool of people. Dota 2’s International is the game’s World Cup and crowdfunding plays a vital role in getting its financials. Dota2 hosts many regional majors/ tournaments like China Major, Europe Major etc. with prize money and in-game offers to attract and encourage the players. During the Dota 2 International, the “Battle Pass” is released for the users to participate in the tournament.

These cost around 4.99$ to 44.99$ and 25% of the battle pass purchase or any in-game purchase of items within the battle pass goes to the prize pool of the tournament. The battle pass be

Graphical user interface, text, application

Description automatically generated

Figure 16: Battle Pass Owners

Chart

Description automatically generated

Figure 17: International Prize money and players base relationship

Chart, surface chart

Description automatically generated

Figure 18: The International championship prize money

### Dota Plus and Dotabuff Plus Subscriptions:

For the new players, Dota2 has subscription levels “DotaPlus” and “Dotabuff Plus” available at 3.99$/month and 6$/month, respectively.

Thus, we can conclude that the pricing strategy is efficient to attract players to Dota2 as it is free for the initial play and the game is supportive of gamers who have money but less time by offering them their DotaPlus subscriptions.

## Section: 4 – Promotion Strategy

Dota 2 promotion strategy is based on sponsoring and co-sponsoring the competitions that keep the players charged. The huge cash prizes and the recognition at the highest level are a motivation for this strategy. Unlike the competitor game “League of Legends” there are no outbound or contribution marketing strategies applied by Dota 2 (Morag, 2021).

“The International” is the biggest annual event and every regional tournament around the year within the defined six regions across the globe revolves around it, as the prize money for the event increases annually and the last event, was about $40 million which is a massive amount as a motivation for the dota2 community members to upskill themselves. Figure 1 below demonstrates the happening of the “The International” event in red.

Chart, line chart, histogram

Description automatically generated

Figure 19:Concurrent players with time, (Red) lines show "The International" event happening

For this coursework, “The International” and other regional events are considered as the promotional strategy for Dota2. To answer the question about the efficiency of the promotional strategy dummy variables were created for the categorized 23 events. Several lags and lead numbers were tried to capture their impact, finally, two lags and one lead were selected with an understanding that one day before the event will be great excitement, and as the tournament is on the resulting news keeps the spark which is discussed within the community.

Using the multiplicative sink model except for the intercept all the variables along with the lag, and lead effects were insignificant as there were zero in their confidence intervals, this might be due to the wrong choice of model. So, different other models were reviewed in this process. A stepwise approach was used for the selection. The parameters for the selected model are shown in Figure 14 of appendix section A. Regression diagnostic for the selected model is also attached within Figure 15 of the same section.

For a game such as Dota 2 where the dynamics are complex to understand for a new player, the current strategy seems fit as a long-term strategy, as the players which are gained with the word of mouth are the ones which will stay, and the stats complement this understanding. However, if the number of players is to be increased only in the next 14 days and the goal is not to be concerned about their association with the game in the longer run, then definitely new outbound marketing and contribution strategy can be applied to increase the number of players.

## Section: 5

We have built the regression model as discussed in question 1 where a simple regression model with Players as a response variable and a log of Twitch viewers as a predictor but the model has some serious issues and residuals were serially correlated and the variable was not significant on the p-value. We also built the ARDL model as discussed under the same question and issues related to that model. In the ARDL model, we found that the QQ plot was not following the normal line, residuals are correlated, spikes on Cook’s distance and standardized residuals are above the signature line. To assess the dynamic effect of parameters, we built the ETSX model with twitch viewers lag to 7. Below is the residual output and QQ plot align with normal distribution with a tail off to the normal distribution, standardized residuals are still widely distributed above the significance level. Actual Vs fitted seems good, but some observations are still far from the mean actual line at the end.

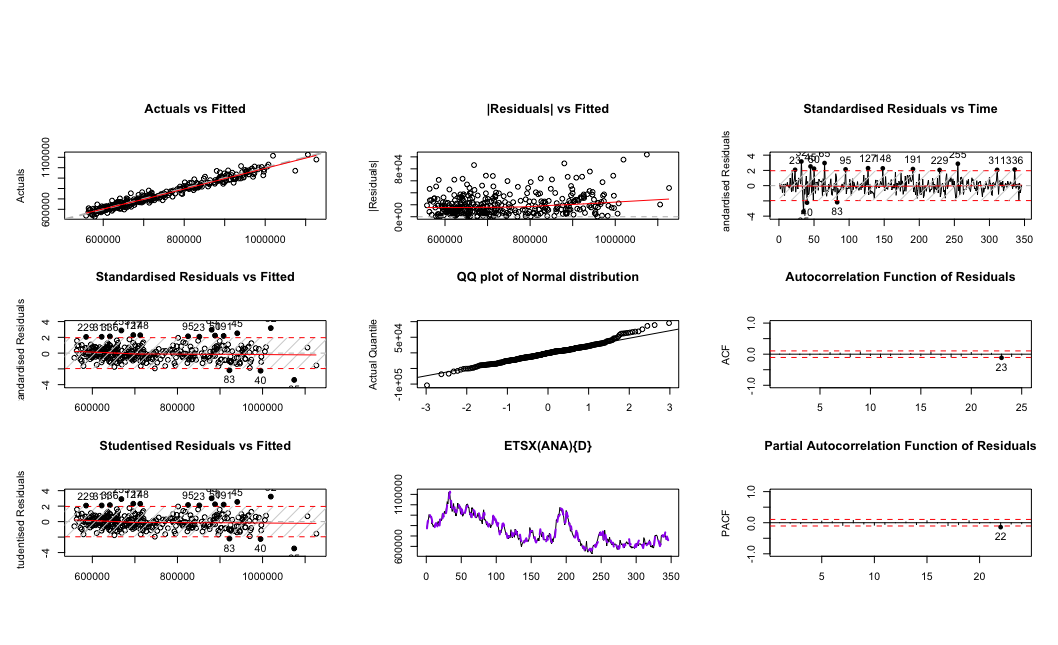
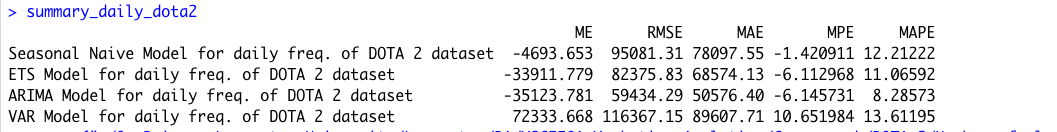


Figure 21: ETSX Model Residual analysis

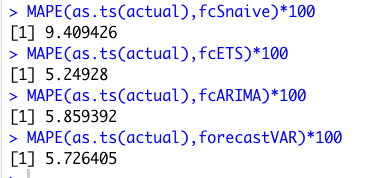
Additionally, we have built Univariate models such as seasonal naïve, ETS (ANA) and ARIMA and multivariate model as the VAR model where variables we considered as Twitch viewers Below are the error measured on the test dataset and found out that the ARIMA model is best suited for the player’s forecasting model as the MAPE value is only coming to 8.28 which is the lowest if compared with other models.

Table 2: Summary of forecasting models error measures



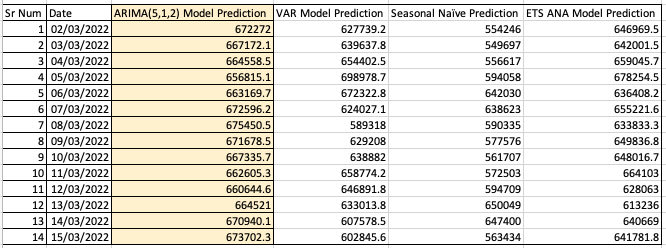
We compared the actual data available for players from 2nd Mar to 15th Mar and below is the result of the MAPE error on the actual. ETS has the lowest MAPE, but our chosen ARIMA model has not to have much difference from the ETS model.

Table 3: Summary of forecasting models error measures on actual data



Below are the predicted player’s numbers over the 14 days starting from 2nd Mar to 15th Mar 2022 highlighted in the column in the colour.

Table 4: Predicted values for the number of players in the next 14 days



References

Dota 2 Prize Pool Tracker. 2022. *The International 10*. [online] Available at: <https://dota2.prizetrac.kr/international10> [Accessed 4 April 2022].

Dotabuff.com. 2022. *DOTABUFF - Dota 2 Statistics*. [online] Available at: <https://www.dotabuff.com> [Accessed 4 April 2022].

Jacobacci, K., 2022. *Dota 2 and Microtransaction Madness | Esports Edition*. [online] Esports Edition. Available at: <https://esportsedition.com/dota-2/dota-2-microtransactions/> [Accessed 4 April 2022].

LEVEL. 2022. *Dota 2 statistics and facts 2022 | LEVVVEL*. [online] Available at: <https://levvvel.com/dota-2-statistics-and-facts/> [Accessed 4 April 2022].

Medium. 2022. *Game Economics, Part 3: Free-to-Play Games*. [online] Available at: <https://medium.com/building-the-metaverse/game-economics-part-3-free-to-play-games-78aa790d55ae> [Accessed 4 April 2022].

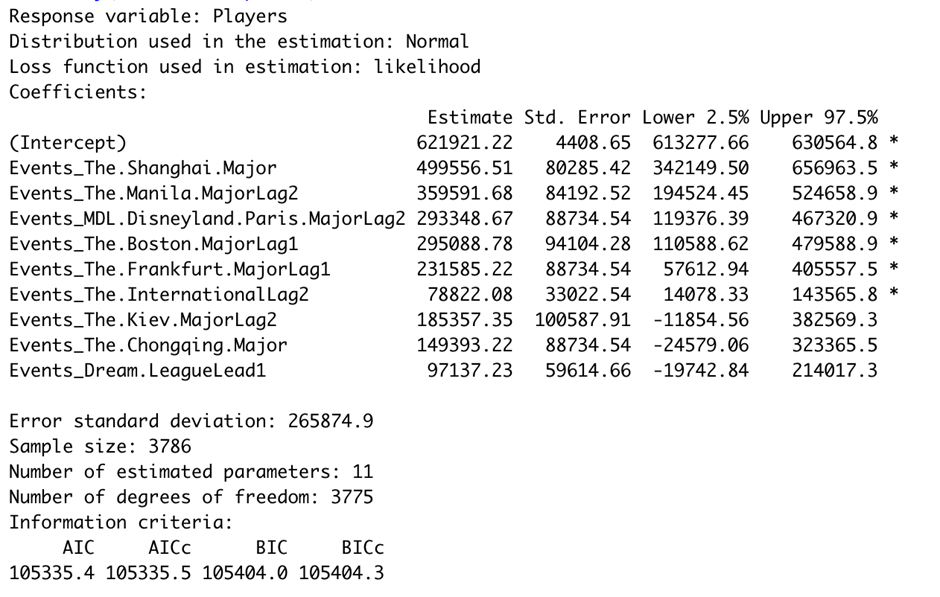
Raghuram, V., 2022. *Over 1 million Players Have Purchased the International 10 Battle Pass*. [online] AFK Gaming. Available at: <https://afkgaming.com/dota2/news/4116-over-1-million-players-have-purchased-the-international-10-battle-pass> [Accessed 4 April 2022].

Steamdb.info. 2022. *Steamdb*. [online] Available at: <https://steamdb.info/app/570/graphs/> [Accessed 4 April 2022].

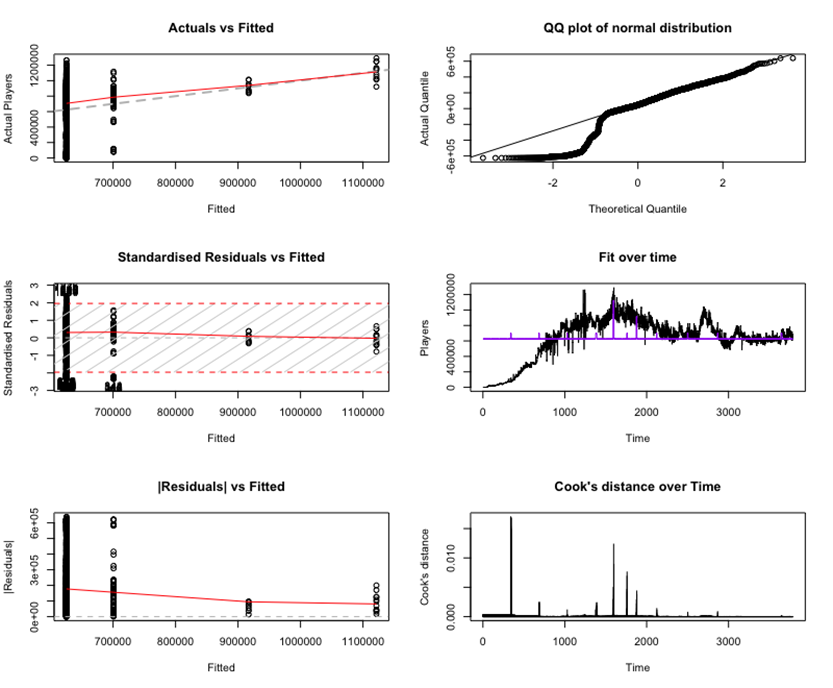
Wwwx.cs.unc.edu. 2022. *Wwwx.cs.unc.edu*. [online] Available at: <https://wwwx.cs.unc.edu/Courses/comp585-s15/Research/PrisingStrategiesofGames.pdf> [Accessed 4 April 2022].

Appendix

Section B



*Figure 14 Parameters for Selected model in section 3*



*Figure 15 Regression Diagnostic for the selected mode in Section 3*