Predicting Online Game Player Count through SARIMAX

Ryan McArthur, Flatiron School, Data Science

Introduction

- There are 2.7 billion gamers across the globe
- Businesses must be able to handle changes in this market, moreso now with the impact of COVID-19
- By using time-series data on number of concurrent players, we will be able to make predictions on trends in player counts and suggest how production companies can prepare for the future.

Data

- Our data comes from SteamDB, or Steam Database.
- Data provides daily statistics on the number of concurrent players
- We used 4 games:
  - CS:GO
  - DOTA 2
  - Rocket League
  - Team Fortress 2

All our games share the same Trends

- We found that when visualizing trends on a monthly level, it was plain to see that all of our games follow the same trends in changes in player counts.
- This suggests that our data is capable of being modelled in the same way

Events

- We found that while events do have a positive impact on number of players, this effect is not significant.
- For viewers, however, events dramatically increase counts.

Volatility trends

- We found that volatility tends to increase as number of players increases
- This suggests that a games maximum number of players increases at a rate greater than average number of players

Modelling Process

- Two types of models:
    - First, we constructed basic models for each game, training on its own data.
    - Second, we constructed an averaged model, using average percent change in number of players across all games.

Modelling Results

Each model has strengths and weaknesses, which present in Long Run VS Short Run forecasts

- Our basic models perform best for short run predictions
- Our averaged model performs best for long run predictions

Recommendations to Production Companies

1. Capital Investment must outpace player growth due to increases in volatility
2. Events are effective at increasing viewership, not player counts.
3. Long-term predictions should be made on market-wide analysis, while short-term predictions are best made by a games own historical data

Further Work

- Collecting hourly statistics instead of daily statistics

- Event Investigations

- Company investigations

Thank You!

Ryan McArthur, Flatiron School
rmcar522@gmail.com