

Modeling Player Performance in Battlefield 4

Keywords

Videogame, Battlefield 4, Player Statistics, Player Reflex, Performance

Objective

Battlefield 4 is a first person shooter game distributed worldwide on various console platforms such as the PS3, XBOX and PC. Players compete against one another to complete different game objectives whether it be flag carrying, getting kills or controlling a designated area. This capstone collects tens of thousands of player data on the XBOX ONE platform, in an attempt to analyze which players are most successful and why they are in the game. The method used to obtain player statistics include web scraping and API calls. The data includes the kill stats of each player on the XBOX ONE as well as what weapons and vehicles they used, how accurate their weapon usage is within the game and how effective they are with certain vehicles. With the data collected, the capstone summarizes what separates top Battlefield 4 players from the rest of the player base and also attempts to find additional patterns that is reflected by the game data.

The Data

The data for this capstone is split into two parts. The first part is the corresponding player list which contains all player names on the XBOX ONE platform, which is provided [here](#). This data requires web scraping to obtain. The second part is the detailed statistics of each individual player, which can be obtained via API calls. The documentation for the API is [here](#). The features that this capstone focuses on are mainly player performance stats (bestStreak, deaths, accuracy, shotsFired, shotsHit, streak, timePlayed) and weapon usage (weapon_name, weapon_accuracy, weapon_shots, weapon_kills, weapon_kills_per_minute etc.)

Method

1. Utilize Python libraries to pull data with beautifulsoup(web scraping) and API calls for data analysis.

2. While the BF4 API site does not state how many calls can be made within a given time, it's best to include some call limits in the script to make sure the process goes smoothly. Currently the plan is to only make 100 calls per hour until it manages to download all player data.
3. Obtain the list of all active BF4 players on the XBOX ONE via beautifulsoup (31812 players 637 pages)
4. Utilizing API calls, pull each individual player stat from the list
5. Analysis the features and look for simple patterns via EDA
6. Use machine learning methods to test algorithms and build a prediction model that will help identify high performing players.
7. Based on the player score, segment the player base into 1000 parts and randomly choose from each segment $\frac{2}{3}$ of the players
8. Use $\frac{2}{3}$ of the players as training data to build a model then compare it with the remaining $\frac{1}{3}$ of the players and see if the model fits the player stats
9. Compare the results and draw a conclusion.
10. Evaluate the methods and approaches used in this capstone and summarize ways to improve the capstone.

Target Audience and Why

This capstone is inspired by two phenomena happening in the real world. One is the better collection of data in the sports industry. Two is the growth of competitive gaming along with esports betting. In the sports industry, coaches are now taking note of how data can monitor the performance and status of an athlete and are using new equipment to collect that data. While, this capstone is an analysis on competitive video games instead of sports, the competitive aspects of analyzing game data can in some ways be carried to other sports. Sports enthusiasts should be able to see the similarities between this capstone and data collected during competitive sports.

To further support my argument, in 2017, major league baseball allowed coaches to [collect player biometrics](#) via wearable devices. During the 2018 - 2019 NFL season, teams inserted [tiny chips into players shoulder pads](#) to track metrics such as running distance per hour and total distance ran within a game. In the competitive gaming industry, similar metrics are being collected and analyzed as well in order to find correlations between player success and that specific metric.

Another phenomena is the introduction of gambling sites and it's push to have competitive games well balanced. In recent years, a lot of gambling sites such as betway and bet365 are making headway in esports cash betting. Within the industry,

29% of current esports bets are made on CSGO, another popular first person shooter game. With the development of the sports industry and the competitive gaming industry, the demand for advanced data analysis will eventually increase. If the video game was not well balanced before hand, and certain players with specific loadouts were to gain an unfair advantage over the majority of the players, this information would cause players to focus on one single playstyle or loadout, which reduces the entertaining aspects of the game. In fact, the goal of the many esports competitions nowadays is to drive viewership and the more entertaining and diverse the gameplay is, the more viewers it gets. Using this capstone, we can showcase how data analytics can be used to bring insight to the game and uncover biases or imbalances the game has within it. This data can in turn be used by game developers to make adjustments to the game, making the game more fair or varied.

Summary

On a macro level, the eSports industry is projected to generate [\\$905 million](#) in revenue, reaching over \$1 billion over in 2019 and 2020. And with more younger generations exposed to video game and competitive gaming. The demand for solid advanced data science analysis for the industry will increase in the future. This capstone explores what methods can be used on competitive gaming data on the players side, in hopes of uncovering key components that lead to player success and also to explore whether certain imbalances exist within the game.