Journal of Counter Strike 2 Analysis

The Process

Alvaro Cervan

2024-09-07

Table of contents

1 Introduction

The choice for this game comes as personal preference. I have been playing Counter Strike since I was a kid and I have always been interested in the competitive scene. When Global Offensive came out, I was already a fan of the game and a few friends and I started playing it. We started not being very good, but we got better and better, and reached some of the highest ranks in the game.

The game even requires a lot of strategy, team play and study of tactics,position and decision making. That meant that watching professional games, professional players and teams could help us improve our game.

Grenades

Smoke grenades are used to block vision and by watching tutorials and professional games, we could learn how to throw them to the sky and make them land in the right spot. Having this knowledge is a great advantage in the game.
Before CS2 fixed this, there was a "bug" that allowed to achieve a "one way smoke" that allowed to see the enemy but not be seen.

2 Data

Counter Strike produces a lot of data, some of it even shown in game, such as the damage done to the player who killed you, money spent each round, bullets shot, etc.

The data is not recorded or transcript in a manual way, nor there is a specific signal or event that needs to be triggered to start recording. The data is always being recorded and can be accessed at any time. That is a big advantage for analysis in E-sports compared to traditional sports where different sensors need to be placed in the field and or players to record data.

This data is not available to the public, but there are some websites that collect data from professional games and make it available to the public.

2.1 Data Choice

The data used for this analysis comes from Kaggle which is scraped from HLTV.

Note

The was available a script to scrape the data from HLTV, but the speed of the scraping was very slow to avoid being banned by the website. Hence, the data was already scraped and available in Kaggle was used.

The data is from the professional scene as well as the normal (best) players and contains information about the players, teams, rounds, kills, deaths, etc.

Also some general data about the game is available, such as the maps, the weapons, etc.

3 Analysis

The analysis is focused on the competitive scene (casual and professional players) and the data is used to answer some questions that could help improve the game or the players.

The analysis took longer than expected, as some data wrangling was needed to get the data in the right format and some data cleaning was needed to remove some outliers and missing values.

Also I wanted to spend more time on the professional players and tournament games data, but after doing some analysis on the normal players, weapons and maps, that was already quite a big analysis.

About the tournament games and performance related to cheating scandals, I expected to find more information about it, but the data was not available or not enough to make a proper analysis. Also, as this game is very competitive and the players are very good, it is difficult to say if a player is cheating or not, so there is a lot of controversy and accusations against basically every player.

4 Learning Curve

This project was a great opportunity to learn new tools and techniques. The analysis was done in Python and Quarto, being the last one a new tool that I had never used before.

I am very happy with the results and the final document as it gives a very beautiful and professional look to the analysis. It is also very flexible and allows to produce different types of documents, such as pdf, html, word, etc. The choice for me is HTML, as it is the only one that supports the interactive plots and tables. Finally, the plethora of options and themes available make the document look very good.

The only bad thing about Quarto is that it seems a bit heavier and in my not so powerful computer, it takes a bit longer to compile and render the document. Towards the end, for making final touches, rendering the new document could take a couple of minutes. This problem had a fix in Rmarkdown with the option to cache the chunks, but I could not find a similar option in Quarto.

Finally, I tried to use Plotly for most of my plots and tables, as it is interactive and allows to zoom in, zoom out, hover, etc. This makes the document more interactive and the reader can explore the data in a more interactive way. The downside is that as an HTML element, it is not supported in the pdf or word documents.

4.1 Tools

As mentioned in the description of the project, the tools used for this analysis should be Python and Jupyter Notebook. Nevertheless, the week before the **Rbootcamp** took place and we learnt about **RMarkdown** and **Knitr**, which produced a better looking document.

The flexibility of Jupyter Notebook is great, but the final document is not as good looking as the one produced by **RMarkdown**. There is a bigger selesction of elements, themes and options that can make the document look better.

Also, producing different types of documents is easier with **RMarkdown**. For example, producing a presentation, a pdf, a word document, a html document, etc. is easier as with Jupyter Notebook.

4.1.1 Streamlit

> Streamlit is a tool that allows to create web applications with Python. It is very easy to use and the final result is very good looking. It is a great tool to show the results of the analysis in a more interactive way.

After hearing and checking Streamlit, I got really excited to learn how to use it and create a web application with the results of the analysis.

At least that was what I thought. Sadly there was some software problem that does not allow me to use Streamlit locally in my computer. I tried to solve it but I had not been able to, and after spending several hours trying to solve it, I decided to move on and use **Quarto**.

Nevertheless, I managed to host the app in the cloud and was able to access it from there, but not seeing in "real time" the changes I was making and having to compile and push it to GitHub was not very productive.

4.1.2 Quarto

This is where **Quarto** comes in. Quarto is a new tool that is being developed by the same people that developed **RMarkdown** and **RStudio**. This tool is an improvement of RMarkdown and it is more flexible and powerful. Specially for different languages. While RMarkdown is focused on R, Quarto allows multiple languages including **Python**, R, Julia, etc.

5 Discussions

Being one of the few students that took the course online made it difficult to join a group and work online while the rest of the group was in the classroom.

This lead to the decision of working alone, which was not a problem for me, but it would have been nice to work with someone else and share ideas and knowledge. On the other hand, working alone allowed me to work on this topic that I find interesting and allows me to work on it at my own pace.

Team mates were replaced then by some friends and my partner, who game insight and feedback on the work and decisions made.

> **Warning**
>
> Please note that sections 1.1 and 1.2 have a drop down section that can be opened by clicking on it. This is a feature of Quarto that allows to hide some information, by default is hidden as it is an introduction to the game of itself, and is not indispensable for the analysis.

> **Tip**
>
> On the top right of the HTML document, there is a button that allows to toggle the theme of the page from light to dark and vice versa. This is a feature of Quarto that allows to change the theme of the document to make it more comfortable for the reader.

> **Post data**
>
> While working on this project, on one of the breaks, I found about Mongolian throat singing and I got really interested in it. I started watching videos an songs and it is very fascinating. Here is my favorite:
> https://youtu.be/7mQ6z65K3KI?t=15