RLCS Performance Analysis
CS430 Final Project

Thomas Moyse

02.01.2023

Abstract

Analyzing statistics relating to a given sport or field can help provide novel insight with regards to performance differentials that would otherwise be difficult to identify. In this report, such an analytical study is performed over recent performances in the Rocket League esports scene. Various hypotheses relating to performance of players and teams will be investigated, specifically: regional trends, insightful statistics and optimal team structure.

1 Introduction

Finding specific areas of required focus and improvement within competitive scenarios can be a challenging and often misguided endeavour. Such focuses may be misplaced courtesy of inaccurate assumptions and preexisting biases which in turn may harm development. This issue can be seen presenting itself within the top-level competitions in the Rocket League space, with many teams across regions being unable to fix under-performances or correctly acknowledge potential reasons for losses, and this is seen through continual inter-regional performance discrepancies. Their is of course many circumstances where a major issue is internal mentalities and relationships but this report aims to focus and identify technical reasons leading to differing levels of success.

2 Background

Some level of familiarity with the underlying game Rocket League is assumed throughout this report but relevant details will be noted when relevant and a brief summary is given as follows:

Rocket League is a competitive car football game released in 2015 by Psyonix. The game can be viewed as an arcade-style physics simulation with a variety of mechanics such as the ability to jump, flip and boost which in turn lead to a great depth of mechanical possibilities. Each individual game is allocated 5 minutes of gameplay time with additional overtime added to resolve ties with a 'golden goal' style system. Much like a traditional sport the 'physics simulation' nature of the game leads to a virtually unreachable skill ceiling with the top-level players continuing to raise the current level of play.

The premier competitive scene for the game named the Rocket League Championship Series (RLCS) is an annual circuit overseen by Psyonix and Epic Games with a combined prize-pool of approximately $6 million. This involves teams of 3 (not including substitutes and coaches) competing over the year within their respective regions and the top performers potentially competing at any of the 3 international majors and/or the world championship.

3 Data

The dataset used for this analysis was obtained from Kaggle and contains the most complete possible summary of the 2021-2022 season.[1] This dataset manifests as a series of 6 CSV files: 4 files describe games and matches by player and team perspectives respectively and the remaining 2 files describe the context of games and players (these two files were however not required during this report). The largest CSV file that being the 'games by players' file and main focus contains 106795 entries, each entry coinciding with a specific players perspective of any given game. This relates to the 5297 total matches played throughout the season where each match typically consists of a best-of-5 series.

This is a sufficient amount of data to be able to provide reasonable levels of insight into playstyles and approaches. Although the data has be processed and cleaned by the author their is additional need to handle parts of the dataset with missing or incomplete fields. For the ease of handling in the report it was decided to drop any records with missing or null values.

Each data-point within the main CSV consists of 103 attributes of varying significance that can be used for the analysis. All these data-points were originally sourced and collected from the websites ballchasing.com[2] and octane.gg[3] which serve as the two main third-party game replay repositories. Notably, the game itself features a replay system which records and allows for playback of any saved game from the replay viewer. This is significant as it allows such information to be shared and if processed correctly viewed. The two aforementioned sites have implemented software that can parse these replay files and extract relevant statistical data (not otherwise available) which they present publicly in addition to hosting the files themselves.

These sites are commonly used by coaches, analysts and fans alike to compare players and teams as well as by admins in verifying match results, this means almost all of the professional RLCS games are recorded on the two respective sites. Due to the open nature of these replay repositories and the presence of their supported APIs it is both very possible and not overly complicated to expand the dataset to include additional seasons. Despite this, the analysis within this report as previously mentioned will focus on the 2021-2022 season and will leave contrasting or adding additional results as a future potential extension.

As previously mentioned the main CSV file contains approximately 107000 entries and has a size of about 80 MB. Although it is definitely possible to work directly with this volume of data, their exist both superfluous attributes and entries that can be removed to provide a clearer and more effective analysis.

Currently each datapoint has attributes that can be summarized as follows:

- Indexing data
- Core in-game stats (shots, goals, saves etc.)
- Boost statistics (Usage, Collections etc.)
- Movement statistics (Speed, powerslide etc.)
- Positioning statistics (Forward third, behind ball etc.)
- Misc. info (Camera settings, platform etc.)

Data pertaining to miscellaneous information is not of focus and thus can be removed/ignored. During analysis each group of attributes will predominantly be considered independently.

More interesting to consider is the removal of certain data-points entirely. Due to the open nature of the RLCS circuit format there are many teams and players that do not have a sufficient number of recorded games to allow meaningful analysis and as such it was decided to remove such points from the dataset that had not competed in at least 7 matches (which would indicate reasonable performance in a split). Making this removal is somewhat involved as it not only requires identifying players who have not played sufficient games but then removing all other players references to any affected games. Due to the nature of how the data is organized this is done on a game-by-game basis as opposed to match-by-match however this means some matches may only be partially reflected within the games dataset and this does require extra consideration in some areas.

To achieve this process matches had to first be grouped by player_ids such that the number of matches associated with each player could be easily identified. In turn, these groups could be filtered to remove any players with less than 7 corresponding matches. It then suffices to read the number of records pertaining to each game, and in any case where this does not equal 6 all references to the game can be removed. This process reduced the size of the game dataset to 91452 entries, approximately a 15 % reduction. This should not only help speed up processing but should also help limit the number of potentially non-representative games.

All this processing as well as any additional processing seen throughout this report was written within python notebooks making use of common and popular libraries such as numpy, pandas, matplotlib and sklearn.

4 Hypotheses

The primary hypotheses to be investigated within the report are as follows:

1. There exist noticeable regional trends regarding player tendencies which may be a factor in performance differentials.
2. Many important statistics that provide meaningful insight exist that are not currently focused on.
3. There is evidence to suggest that it is not optimal to have a fixed first, second, third man

structure

The process for each of these hypotheses is quite distinct and as such they can be considered independently of each other.

5 Analysis

5.1 Regional Tendencies

With regards to hypothesis 1, it is very difficult to ascertain whether the regional tendencies are the causing factor of performance differentials. It is however possible to attempt to identify whether such tendencies are sufficiently reflected within the data. It is first important to denote what is meant by tendencies within the context of the data. A tendency denotes the general approach taken by a player within any given game and this spans several areas such as positioning, decision-making etc. In relation to the data it is most accurately described by the positioning statistics, certain movement statistics, certain boost statistics and demolition statistics.

Although their were eight total regions part of the RLCS 21-22 season we will focus on the 5 most prominent regions:

- Europe (EU)
- North America (NA)
- South America (SAM)
- Middle East & North Africa (MENA)
- Oceania (OCE)

These regions were selected as they are historically more competitive at the highest level especially when compared to the much more recently added Sub-Saharan Africa and Asia Pacific regions.

As some background information it is commonly believed within the community that different regions approach the game differently. For example one commonly accepted view is that EU is a more aggressive region with relation to pressuring the ball than NA. It is also commonly believed that regions that may play on more distant higher ping servers have developed a style of play where more aggressive risk taking is present, touted examples being the MENA and SAM regions. Below you can see the results of some analysis of these sentiments as well as other noticeable trends.

The first group of stats to be reasonably considered can be referred to as core stats. These are the stats that are most easily accessible and recordable, being presented on the in-game stats leaderboard and stating simple totals these stats are

- Shots
- Saves
- Assists
- Goals

By performing min-max normalization on each of these stats when averaged over each region it becomes easy to distinguish relative discrepancies even when the underlying stats were originally quite close together in a proportional sense. But even when proportionally quite close, such discrepancies can be very significant when observed a sufficient number of times across a wide range of samples such as this scenario. And the impact of such discrepancies can be seen clearly manifested within the international results where some regions are able to vastly outperform others.

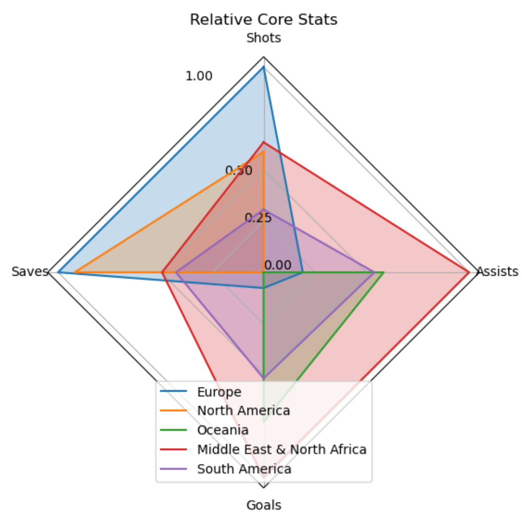

Figure 1: Core Stats Min-Max normalized

The relative dispositions of core stats alone already help to provide a significant insight into the relative game approaches within regions. It is worth noting that as the majority of data is from intra-regional competition, having a relatively low proportion of goals such as seen in Europe and North America, does not necessarily translate to inter-regional competitions and this is a fact supported when exclusively viewing international performances. The graph in figure 1 can be seen highlighting the strong links between shots and saves as well as between assists and goals which makes sense when considering the context that the majority of shots in a game typically result in a save with an average shooting percentage of between 21% to 25% depending on the region under consideration.

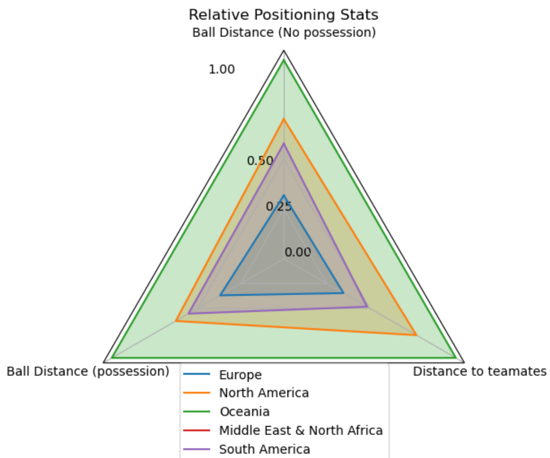

Figure 2: Positioning Stats Min-Max normalized

The plot seen in figure 2 shows an interesting result. Firstly, the absence of any significant skew within each of the plotted regions highlights the close link between these three positioning stats which would not otherwise be viewed as a certainty. Contexually, the commonality between these stats is able to display a clear ranking of how much space is afforded on the field between each of the respective regions. Specifically it is seen that at least in the 2021-2022 season MENA positioned most aggressively and OCE positioned most passively.

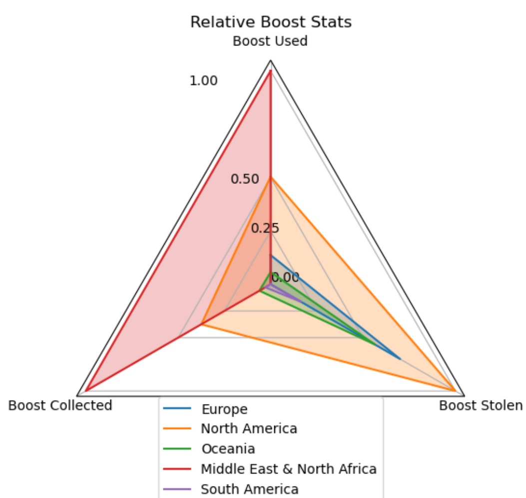

Figure 3: Boost Stats Min-Max normalized

Arguably, the most insightful remaining plot to view a clear difference in approach across regions is the boost radar seen in figure 3. Of immediate notice is the boost usage and collection within MENA, which is high enough such that the differences between it and NA (the next highest) are as large as the difference between NA and the minimum consuming regions. Although it may be expected that stolen boost is inversely proportional to collection and thus usage, this can be seen from the plot to not be the case.

One point of insight that is also derivable from the plot relates to boost utilization and availability across regions. The area occupied for each of the plotted regions signifies the utilization of boost resources/the relative lack of availability within the respective regions. A parallel perspective is that these areas relate to the relative focus on boost resources in games within any of the regions.

Numerous other relative statistics were investigated for this hypothesis but have not been presented due to either lack of novel insight being provided or an absence of any significant differences between regions. One example of this relates to several of the positioning statistics, particularly those relative to field locations e.g. back-third, which did not contain any meaningful distinctions between regions. This was in itself surprising, as it could easily be assumed that any aggressive or defensive tendencies of a region would manifest in being positioned further forwards or backwards on the field (like traditional football) but this was seen to not be the case.

5.2 Insightful statistics

Currently most coaches and analysts pay little to no attention to many of the potential available statistics when analyzing their previous results, this is because it is not clear which if any of the non-core statistics provide useful insight. There do exist other studies relating to such statistics but these focus on areas such as lower level gameplay, or alternative modes like 1v1s [4]

In order to evaluate which statistics contribute most to game results we can analyze both correlation with game result and attempt to make predictions on game result via regression. For this section instead of the previous player-perspective dataset, we will instead focus on the team perspective dataset in order to consider the team as a whole. This is important as considering players individually when looking for insight over a single game may not be reflective of the issues the team as a whole is facing - examples being a player trying to fulfill a specific role thus skewing their individual stats but not necessarily skewing stats from a wider viewpoint. Of note regarding the team-level dataset is a reduction of features/attributes to consider, with now only 49 total attributes. These attributes are in the majority, cumulative totals of relevant individual player attributes e.g. at player level the attribute 'closest to ball percentage (on team)' would now have no meaning at cumulative team level (being 100%), and is thus not present.

We start by investigating the Pearson correlation coefficient between each of the features and the target 'winner' attribute, to establish which if any features provide meaningful insight into game result. After filtering to require a non-negligible absolute coefficient of greater than 0.1 the number of meaningful features is reduced to 17. One interesting note is the saves attribute being the only core attribute that appears largely uncorrelated with game result having a coefficient of only -0.07.

To validate if the coefficient results are meaningful an alternative feature selection method was selected as a point of comparison. Specifically, ANOVA F-values, which when selecting 17-best produced an identical list of features to that of the filtered coefficient.

Supposing that a set of features could be used to predict game outcome would imply that such features are meaningful. As such, a logistic regression classifier was established, which when passed all 17 features achieved a 90.7% accuracy. This is clearly a significant result that proves the features themselves are meaningful to analyze and consider, however it should be noted that both this 17-feature model and the reduced optimal 12-feature model (achieving comparable accuracy) contain many of the core-stats which could be heavily influencing classification success. This would be problematic, as the end goal is to identify new non-core statistics. As such, a classifier which may rely on core statistics does not prove the significance of the remaining statistics.

For this reason a new optimal classifier was built that did not contain any of the core-statistics. The new classifier contained only 6-features (identified via ANOVA F-values) to achieve an optimal accuracy of 78.8%. Although performance can be seen to have degraded without access to core stats, this lesser result still highlights that meaningful insight into the game result is possible when exclusively using said features.

A Naive Bayes classifier was also tested in addition to the logistic regression classifier and accuracies were somewhat weaker across the board which is to be expected given the nature of the classifier. For example, the 6-feature classifier achieved 71.8% accuracy, but importantly insight remained present. It was also validated that relative performances were consistent with the regression classifier across feature sets.

The optimal identified features and their correlations in order of insight are:

Rank	Feature	r
1	Time in front of ball	-0.371
2	Time spent with ball in side	-0.193
3	Time spent behind ball	0.165
4	Time in offensive third	0.156
5	Average boost	0.153
6	Time in offensive half	0.145

Table 1: Feature Correlation with Game result

It is important to note that correlation does not necessarily imply a causal relationship, and thus attempting to optimise these values in a team will not necessarily increase winning percentages, however this analysis does display that these statistics are insightful and should be paid attention to. For example, if your team loses a game and can be seen to have spent significant time in front of the ball this may in turn highlight overconfidence as an issue which could then be resolved.

5.3 Optimal team structure

This is the most difficult hypothesis to analyze and test courtesy of the data not clearly describing to what extent a first, second, third man structure is present. It should however be possible to make a few key assumptions in order to estimate a factor describing this.

The concept of a first, second and third man is very simple but does have some underlying nuance that is not strictly defined. Put most simply, if all three players were in an equally good position to attack or challenge the ball, the first man would be the player that attacks the ball, the second man would be the one that positions near to make a potential follow up play and the third man positions to cover any potential mistakes by their team, or clears and outplays by the opposing team.

The first assumption relates to assuming a player being most back is third man, most forward is first man and remaining player is second man. This is not strictly always the case but should be reasonably accurate. Experimenting with closest and furthest from ball may prove more accurate.

In order to validate the hypothesis their is a requirement to first process the data in a way that allows for required analysis to be conducted. Iterating over the games by players dataset grouped by both team and game id allows each game and team within the game to be considered independently. Over these team groupings and making use of the aforementioned assumption we can generate a new data table which contains players and the counts for the positional roles they have fulfilled across all their games. Win and loss counts were also recorded for ease of use in further analysis.

	player_id	player_tag	first	second	third	wins	losses
0	5f3d8fdd95f40596eae2412e	Amphis	118	93	71	176	94
1	5f3d8fdd95f40596eae23e01	Torsos	102	130	44	176	94
2	5f3d8fdd95f40596eae23e53	Express	50	73	155	176	94
3	604e562901d675f81a96b270	mel kin	42	41	22	48	57
4	5f7ca648ea8a0f0714fb9a20	Laxin	34	13	44	39	44

Figure 4: Sample data from new table

Using this newly formatted data table it is now much clearer how to approach and perform the positional analysis. We first, define two additional metrics win percentage and position uniformity which will allow us to clearly test the hypotheses. Uniformity is a custom numeric statistic that aims to quantify the spread across roles for a given player, it is defined in the following manner:

$$\text{Uniformity} = 1 - max\begin{cases} \left|\frac{\text{f}}{\text{f+s+t}} - \frac{\text{s}}{\text{f+s+t}}\right|, \\ \left|\frac{\text{f}}{\text{f+s+t}} - \frac{\text{t}}{\text{f+s+t}}\right|, \\ \left|\frac{\text{s}}{\text{f+s+t}} - \frac{\text{t}}{\text{f+s+t}}\right| \end{cases}$$

Figure 5: First: f, Second: s, Third: t

This definition of uniformity has a range of $(0, 1)$, where 1 is a maximally uniform player (playing all positions an equal number of times) and 0 is a player who exclusively plays a singular role. We can then perform an analysis of our new statistics.

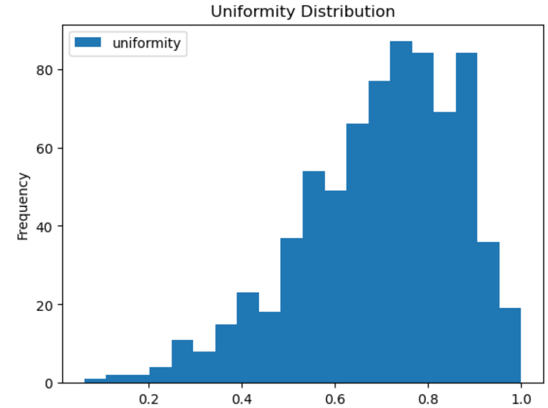

Figure 6: Distribution of derived attribute uniformity

This distribution already provides insight to support that most players although switching between positions tend to have a preferred position. The mean uniformity is approximately 0.7 which implies the gap in proportions between the filled roles is 0.3, as role proportions must add to 1 this gap is already significant as an example distribution could have the player with the split 0.2, 0.3, 0.5 which implies they play one role as much as the other two combined. As individual games are volatile and results span across a year, in which time they may have switched preference in positions, this is significant evidence that most players trend towards set roles on a team.

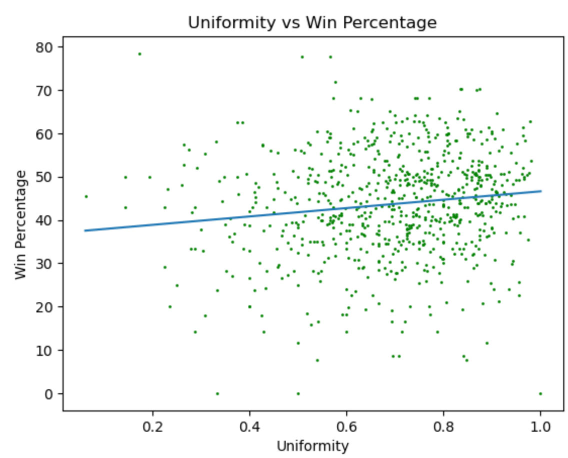

Figure 7: Plot of uniformity against wins

When comparing uniformity with win percentage a small positive correlation was observed with a coefficient of approximately 0.135 however this data included teams at slightly lower performance levels and given the shape of the distribution it was important to investigate if the pattern remained consistent even at the highest performance levels. Repeating a plot like figure 7 with worse-performing teams removed highlighted that this slight correlation was in fact simply a factor of the lower performing teams who typically play less uniformly, skewing the data.Specifically, filtering the data to only include players with 60 or more game wins reducing the player count from approximately 700 to 200 displayed a correlation coefficient -0.05 and a subsequent coefficient of determination of 0.003 and thus can be deemed insignificant. The mean uniformity also slightly increased by around 0.04 which supports that lower ranked teams are more likely to play less uniformly.

Their is still interest to be gained from looking at the uniformity stat with regard to how different top-level teams were able to approach the game with differing uniformity. For example the most upper-left point in figure 7 represents oKhalid a player from the MENA region who played on the extremely successful team Falcons. He could be seen playing the role of third man in 86% of games and managed to come in second place at one of the three international majors (a historic best result for the region). Comparatively, the world championship winning team BDS played with Extra who achieved a uniformity score of 0.98 across the season. Not only do different teams feature players with different approaches to role uniformity, but within teams it is not uncommon to see players with differing uniformity scores - BDS are a clear example with teammate Seikoo having a uniformity score of 0.58.

6 Conclusions

Overall the analysis of the data, and by extension project as a whole, has been successful with all of the questioned hypotheses being sufficiently explored to make conclusions on their accuracy.

Specifically, with regards to the first hypothesis of the existence of differing regional tendencies, which in turn may being a contributing factor to performance differentials. It can be seen that this is clearly supported by the results of the analysis. It was found that their exist notable discrepancies with regards to various stats across regions and these discrepancies branch across many aspects of the game, including boost management, field spread, and ball pressure as well as notable differences in core stats like goals and assists.

Regarding the second primary hypothesis of the existence of meaningful but not often considered statistics, this was largely shown to be true. The only area of uncertainty regarding this hypothesis is whether top-level coaches already consider these areas but do not share such considerations with the wider community, however this is deemed unlikely. At least 6 new attributes have been presented as meaningful with regards to predicting the outcome of the game, all of which are thus able to provide meaningful insight into potentially required team adjustments.

The third and final primary hypothesis was unfortunately very much inconclusive, with no meaningful relationship being uncovered. It is not entirely clear whether this inconclusiveness is a derivative of the hypothesis itself or the approach taken to determine it. A reasonable estimate would be that the lack of substantial findings is indicative of the lack of a relationship between fixed roles and team success, however this is of course not certain and as mentioned may be a result of the specific process taken within the report.

7 Potential Extensions

A very simple extension is to consider more potential base statistics that are not by default provided by replay repositiories but this may involve custom parsing of replay files or using existing pattern mining extractions [5]. Being a constantly evolving game with regards to the approaches taken at the highest level, one very obvious extension is to carry out similar experiments over a wider range of seasons. This would hopefully enable the analysis of changing trends over time.

Due to the relative lack of international games in any given season when compared to the quantity of domestic games some analysis may benefit from incorporating measures of regional strength this could simply involve ranking based upon international match results or could make use of a more complex external ranking dataset such as the AI power rankings presented by Rocket Science[6], which not only aims to rate teams and players but does so fairly with respect to their regions apparent strength.

Another adjacent area that could be considered is trends across regions across all skill levels. This would involve construction of a new dataset but this could be achieved in a similar manner to the dataset that was used within this report by pulling replays statistics from sites like ballchasing.com and octane.gg

References

[1] D. Monfret, *Rocket league championship series 2021-2022*, Aug. 2022. [Online]. Available: https://www.kaggle.com/datasets/dylanmonfret/rlcs-202122/data.

[2] Can'tFly, *Rocket league replays database*. [Online]. Available: https://ballchasing.com/.

[3] Octane. [Online]. Available: https://octane.gg/stats/players?mode=3&minGames=50&group=rlcs2122fall.

[4] T. D. Smithies, M. J. Campbell, N. Ramsbottom, and A. J. Toth, "A random forest approach to identify metrics that best predict match outcome and player ranking in the esport rocket league," *Scientific reports*, vol. 11, no. 1, p. 19 285, 2021.

[5] R. Mathonat, J.-F. Boulicaut, and M. Kaytoue, "A behavioral pattern mining approach to model player skills in rocket league," in *2020 IEEE Conference on Games (CoG)*, 2020, pp. 267–274. DOI: 10.1109/CoG47356.2020.9231739.

[6] HalfwayDead, *Rl ai power rankings*, Jan. 2024. [Online]. Available: https://rocketscience.fyi/esports/ranking.