Predicting Chance of Winning in League of Legends

Jason Lu

Abstract

League of Legends is one of the most popular online games to date with millions of potential team compositions. This project will attempt to predict whether a team will win a game based on the champions chosen by converting champions into their attack, defense, magic, and difficulty values and associate them to a team's chance of winning. Increased values of stats is associated with a reduction in the probability of winning, however it will be shown that many other factors are at play. The game of League of Legends is more than just numbers.

Introduction

League of Legends is an online game in which two teams of 5 face off against each other to destroy the other's base. Colloquially, there are two sides of the map: red side on the upper right half and blue side on the lower left half. Furthermore, there are 5 positions each with their own designated spot on the map: one person in the top lane, one person in the jungle, one person in the mid lane, and two people in the bottom lane. Players select a unique champion to play each with their own strengths and weaknesses to combat the opposition. Within the highest echelon of the game is the professional scene, in which the best of the best face off in their own region to qualify for the world stage and compete to be world champions.

The overarching question I would like to answer is whether I can predict a team will win based on the champions they have chosen. There are also two other questions I am interested in that can be answered through data analysis:

In the professional scene in League of Legends, regions fundamentally play differently from each other. For example, China has an extremely aggressive playstyle while Korea plays slow and calculated. So, my first question is as follows:

1. Do certain regions pick champions based on their playstyle, or play the same champions as any other region, just differently?

Furthermore, in each iteration of the game, there is a certain "meta" to it, meaning that some champions will be fundamentally better than others and will be picked very often, but it leads to my second question:

2. Do meta champions have a higher win rate than other champions?

Method

The dataset I have been working on deals with all professional League of Legends games in the 2020 season across all regions. It has over 5000 games played and contains the following variables:

- league – league under the region these games were played in (China, North America, Europe, etc.)
- blue/red team – team on the blue/red side of the map
- blue/red top – champion in the top lane
- blue/red jungle – champion in the jungle

- blue/red mid – champion in the mid lane
- blue/red adc – one of two champions in the bot lane
- blue/red support – one of two champions in the bot lane
- result – indicator of whether the blue team won (1 if won, 0 if lost)

I also have a dataset containing every champion in the game along side their stats such as attack, defense, magic, and difficulty values.

The data was collected from a user on Kaggle who got the match data from League of Legends esports stats website "Oracle's Elixir" as well as champion data from Riot Game's (the company that made League of Legends) API. I used arm, data.table, dplyr, and ggplot2 packages for my analysis.

For the cleaning, I first separated the teams into two sets and then merged them back together so that each team has their own row and added a variable indicating what side the team played on. I then took a subset of the data to only include major regions the games were played in (since the data came from every professional Esport game played in 2020), so we only have games from the LEC, LPL, LCK, and LCS representing Europe, China, Korea, and North America respectively. Finally, for any champion with 10 or less frequency in a region, I replaced with 'NA' since the champion would not have enough frequency to be accurately represented in predicting a win.

I then created two metrics for EDA and model–building purposes respectively. For EDA, win rate is calculated by summing all instances of a champion winning (having a '1' for result) and dividing by the total number of games the champion was in. I did this both by region, meaning each champion's win rate in their respective region, and not by region, meaning the champion in general (Figure 1).

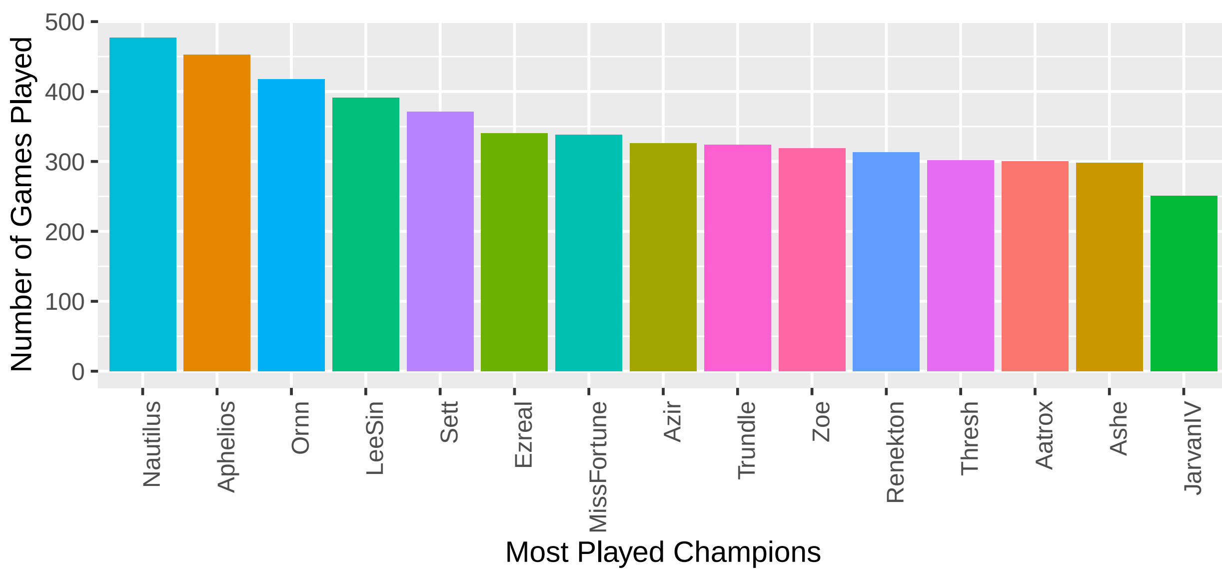

Figure 1: Frequency of Top 15 Champions Played

The meta at the time was aggression and bruisers, not too many team compositions focused on 5v5 fights but dueling power to become strong and take over the game.

As seen in Figure 2, the most played champions do not appear to have an overwhelming majority of win rates above 50%, so it seems that "meta" champions do not have a significantly higher win rate than other champions (the average win rate is approximately 50%).

In Figures 5 - 8, champions such as Aphelios, Ezreal, Sett, and Ornn were prevalent across all regions. Some regions have their own distinct picks such as Rakan in North America, Syndra over other regions' Azir in

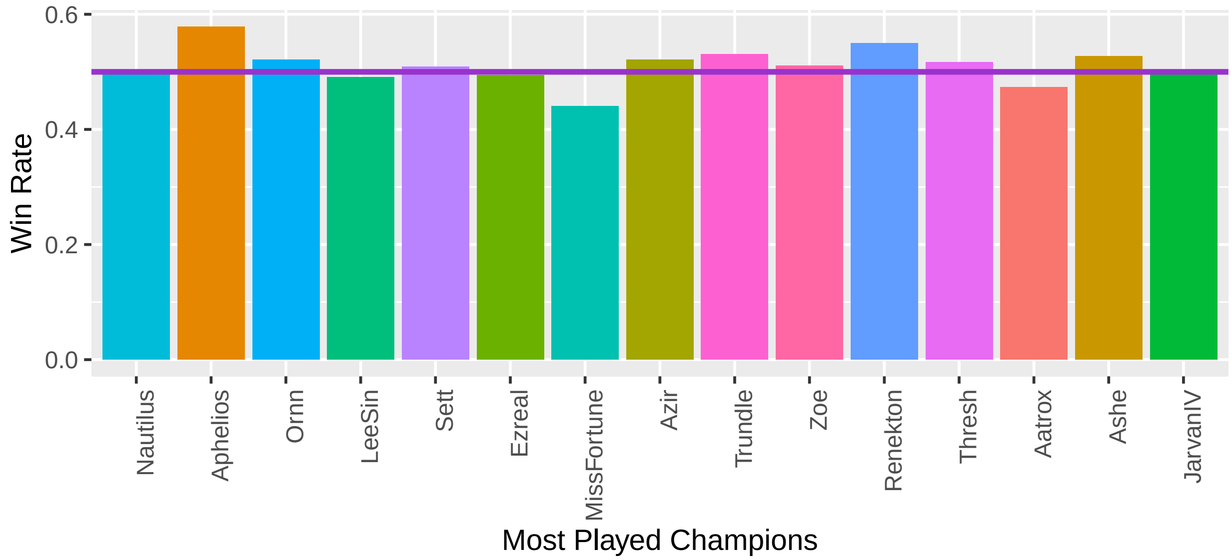

Figure 2: Win Rate of Top 15 Champions Played

China, Kalista in Korea, and Gangplank in Europe. Although not overwhelming, there does seem to be certain champions some regions favor over others. The difference in the amount of games played comes from the different formats, with best of one's played in NA and Europe and best of three's in China and Korea.

To help build the model I incorporated the attack, defense, magic, and difficulty stats of each champion for each team to acquire a more wholistic view of the data that is not just the champion name. I referenced the champions on each team with the dataset containing each champions' stats and created the average attack, defense, magic, and difficulty of each team composition to use in the model. I took the average since some teams would contain 'NA' for some positions due to the cleaning I did for severely underplayed champions, thus affecting how much the total stats would be for each team.

When creating the model, incorporating the champions as a factor was extensive and did not allow me to add champion stats as there were not enough observations to incorporate them both together. Thus I stuck with a multilevel logistic mixed effects model that predicts chances of winning with champion statistics as predictors (Figure 3) and team name as random effects (values in Appendix) since teams play the game and would perform differently from other teams.

Results

Based on the model, average attack, defense, and magic of a team composition seem to have evidence of a negative effect on the chance of winning a game, with a decrease of 0.14, 0.13, and 0.11 in the log odds of winning a game for each point increase of average attack, defense, and magic respectively. Meanwhile, average difficulty does not seem to affect the outcome of a game. The team random effect from -0.80 to 0.92 depending on the team.

The binned residuals (Figure 4) asses the overall fit of the regression model. The more the points lie within the confidence limits, the better the fit of the model. Figure X a decent fit of the model, although it tends to underpredict team compositions with less average stats and overpredict team compositions with higher average stats. As a whole, it seems to be okay as a large factor in determining a composition winning is the team that plays it.

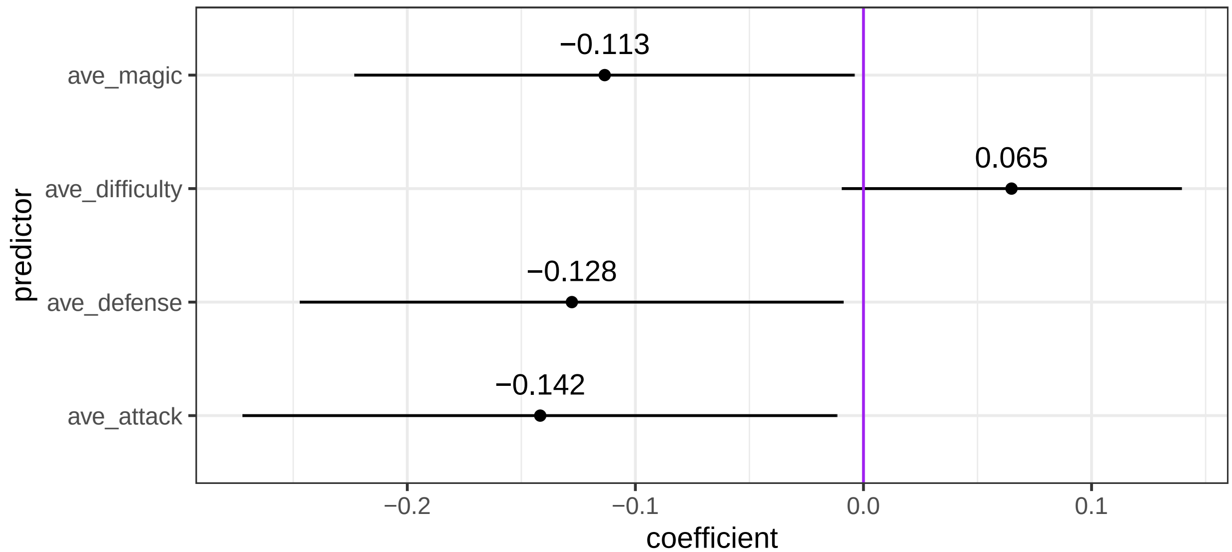

Figure 3: Multilevel Logistic Model Coefficient Values

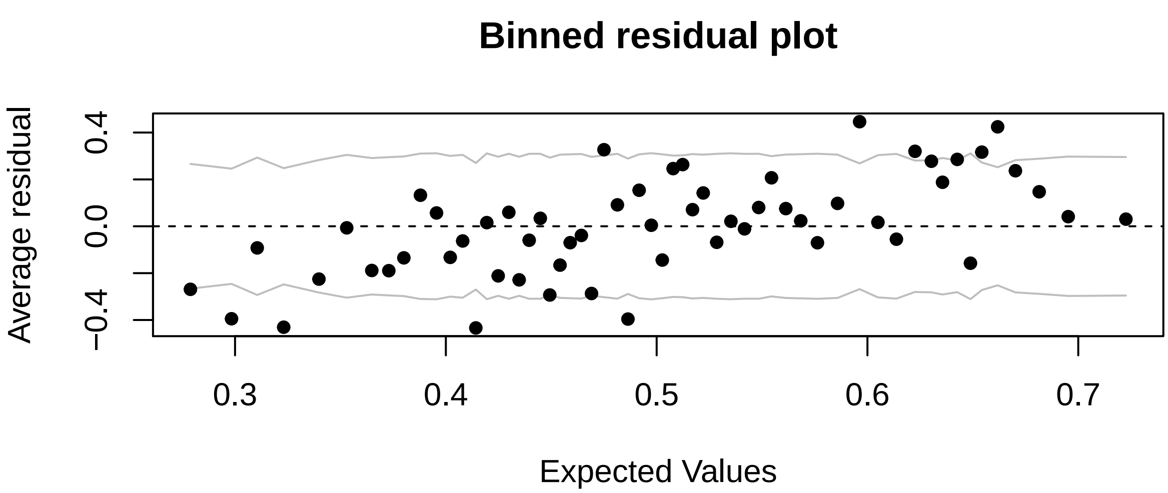

Figure 4: Binned Residuals for Model

Discussion

Having increased average attack, defense, and magic stat of a team lowering our chances of winning is unexpected, since normally higher stats mean you would be stronger and thus be able to win more, though too much of one stat would be detrimental to your chances. A large part of how a game is won is due to teamwork and coordination of players, not just the champions that are chosen, which is why the random effects of the team have a significant influence on the outcome of a match, causing the model to have negative coefficients for the predictors. For example, teams like Cloud9 have a great roster that works well with each other and do not make many mistakes in game, leading to a high effect value compared to other teams.

The game of League of Legends stems far from just numbers on how strong a champion is, limiting the generalizability of the model heavily. It encompasses an infinitesimal among of factors not captured in the model such as how you play your champion, how good your decision-making is, and even how well any individual is playing that day.

Future implementations of the model would be metrics to incorporate champion interactions with each other (some champions work extremely well with each other), what team of champions you are facing (similar to how a champion works well with others, some champions work very well against others), which side you are playing on (it determines the draft order, which can decide a game before it even starts), and individual champions as factors in the model alongside their stats (sometimes having this one champion will win or lose the game).

Appendix

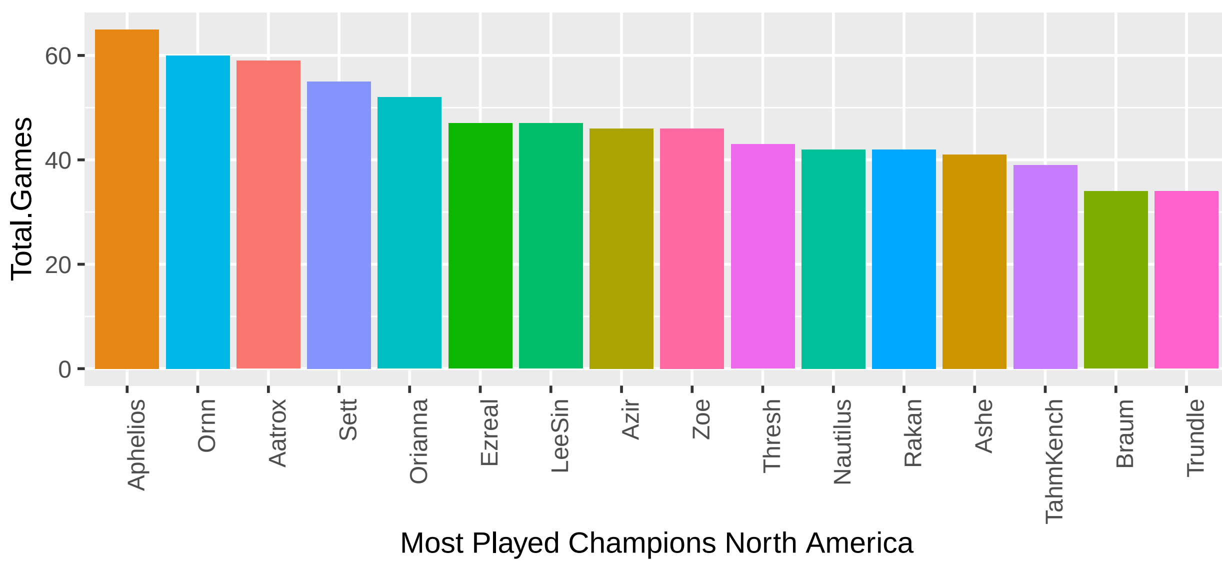

Figure 5: Frequency of Top 15 Champions Played in LCS

```
## $team
##                        (Intercept)
## 100 Thieves            -0.23701064
## Afreeca Freecs         -0.14370538
## Bilibili Gaming        -0.11548533
## Cloud9                  0.92146845
## Counter Logic Gaming   -0.79430576
## DAMWON Gaming           0.65935435
## Dignitas               -0.38651508
```

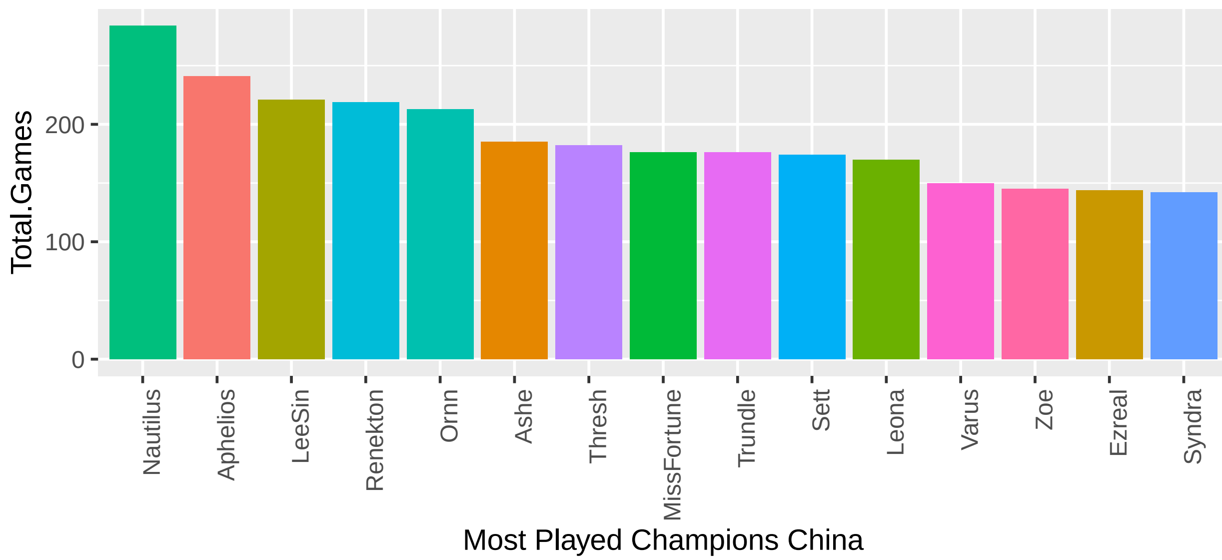

Figure 6: Frequency of Top 15 Champions Played in LPL

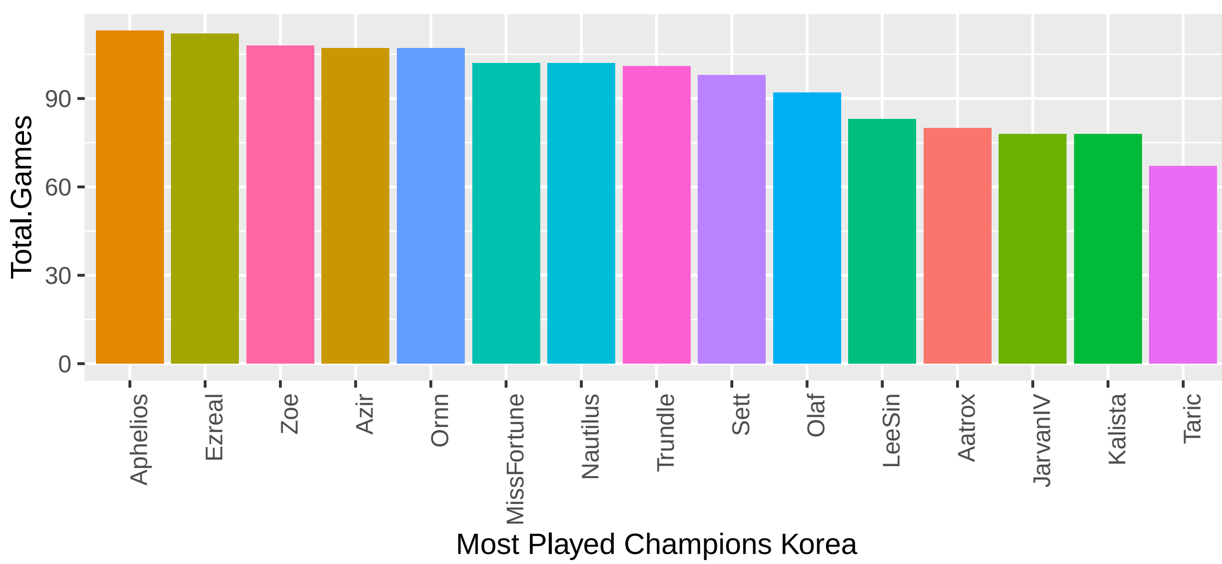

Figure 7: Frequency of Top 15 Champions Played in LCK

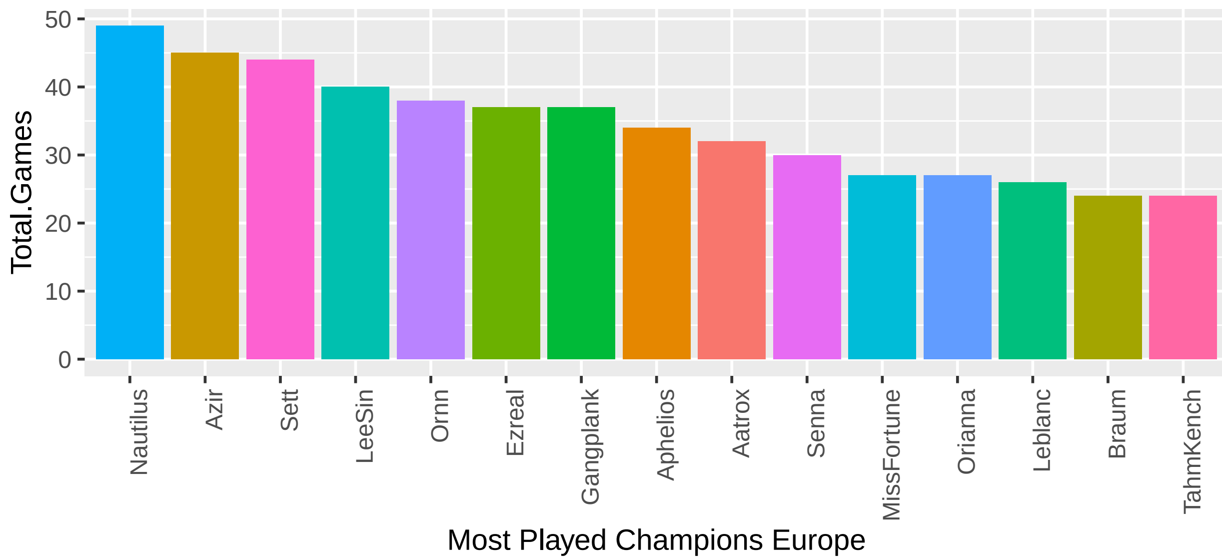

Figure 8: Frequency of Top 15 Champions Played in LEC

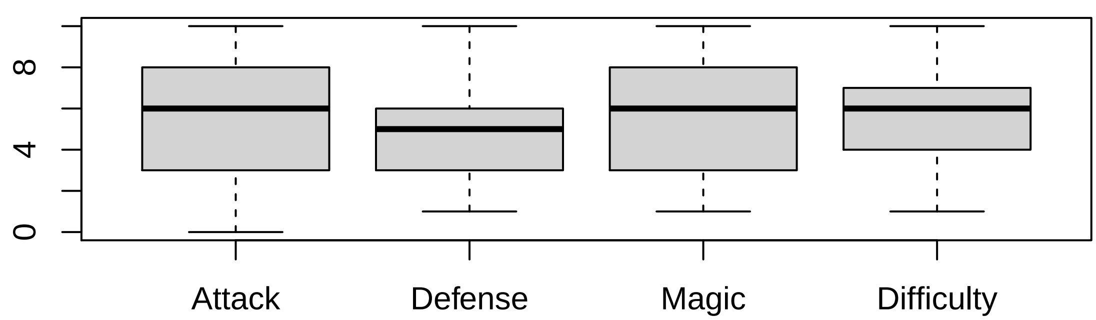

Figure 9: Distribution of Champion Stats

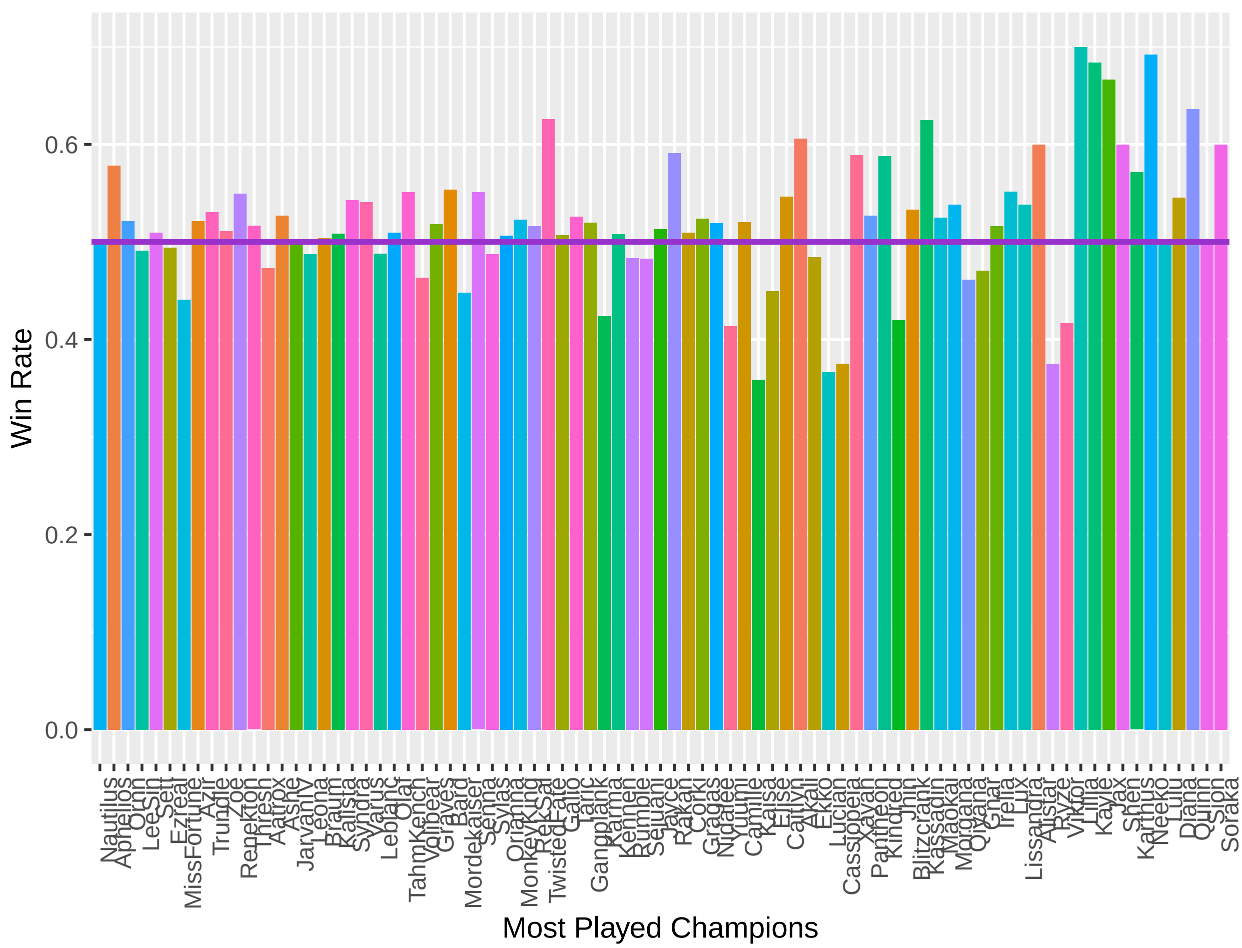

Figure 10: Win Rate of Top 80 Champions Played

```
## Dominus Esports       -0.67412482
## DRX                    0.56709965
## EDward Gaming          0.22412398
## eStar                 -0.08658329
## Evil Geniuses         -0.01724694
## Excel Esports         -0.19622532
## FC Schalke 04 Esports -0.18109136
## FlyQuest               0.27773606
## Fnatic                 0.35654673
## FunPlus Phoenix        0.37610652
## G2 Esports             0.67152544
## Gen.G                  0.78927877
## Golden Guardians      -0.11266147
## Griffin               -0.39946883
## Hanwha Life Esports   -0.68982353
## Immortals             -0.44270307
## Invictus Gaming        0.35992547
## JD Gaming              0.70175404
## KT Rolster            -0.03835760
## LGD Gaming            -0.03356868
## LNG Esports           -0.43219313
## MAD Lions              0.12907837
## Misfits Gaming        -0.09962153
## Oh My God             -0.28226072
## Origen                 0.01734260
## Rogue                  0.24556387
## Rogue Warriors        -0.33554428
## Royal Never Give Up    0.12826796
## SANDBOX Gaming        -0.25213558
## SeolHaeOne Prince     -0.75936463
## Seorabeol Gaming      -0.28391897
## SK Gaming             -0.43365691
## Suning                 0.19256561
## T1                     0.70378166
## Team Dynamics         -0.22773049
## Team Liquid            0.32859225
## Team SoloMid           0.27635796
## Team Vitality         -0.68421852
## Team WE                0.11710065
## Top Esports            0.71890208
## Vici Gaming           -0.01245951
## Victory Five          -0.36562133
##
## with conditional variances for "team"
```