Win Condition for League of Legends

1. Data set:

The data set I found is from the  [ORACLE'S ELIXIR](https://oracleselixir.com/future) website which gathered the data from all the League of Legends pro games in 2020.

1. Project Goal:

The goal of this project is to find which variables could affect the win condition of a League of Legends game. The variables may include the gold difference, experience difference, side, game length and so on.

1. Selected Variables:

* side: Red Side and Blue side of the map
* gamelength: the length of the game in second
* result: 0 and 1, “0” represent loss, “1” represents win
* teamkills: total kills earned by the team in one game
* teamdeaths: total deaths by the team in one game
* dpm: damage per min, damage the player deals to opponent players per min
* earned.gpm: gold earned by the player per min
* golddiffat10: gold difference between the play and the opponent player of the same position at 10 mins of the game
* gold diffat15: gold difference between the play and the opponent player of the same position at 15 mins of the game

1. EDAs:

First, game length is important in this data set. I characterize the game length into 3 different level. Games under 25 mins are short games while the games longer than 38 mins are long games. Games between 25 mins and 38 mins are regular games.

##Long Regular Short

##12600 59376 5568

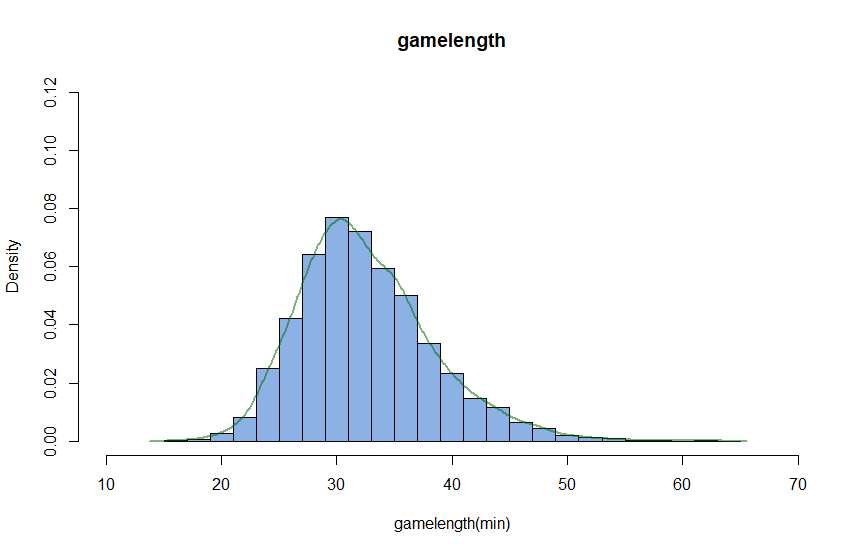
This tells that most of the games are regular games and games under 25 mins are relatively less than the other two.

Next, I summarize the win rate of each side in total.

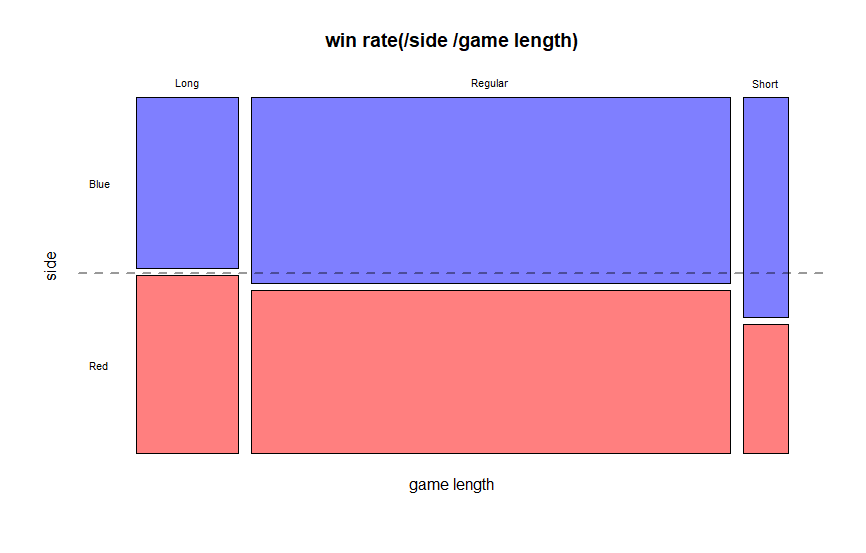
##Blue Red

##20658 18114

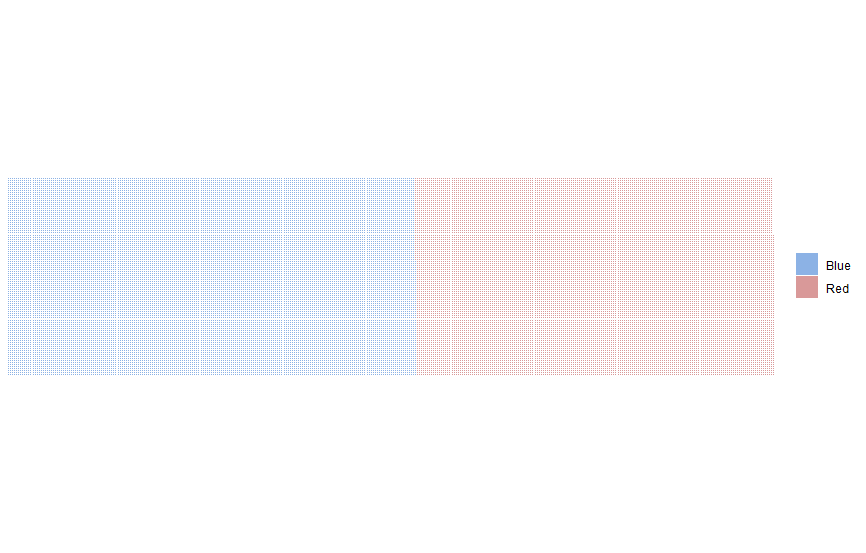
This would tell us that blue side has greater chance to win the game than the red side.

1. Visualization

This is a histogram overlaid with density plot that shows the gamelengh of games through out the year. From this graph, we can tell that most of the games varies from 25 mins to 40 mins while the longest game can take up to 65 mins which is uncommon.



This is a mosaic plot showing the win rate for each side at different game length. From this graph we can say that for short game, Blue side has more chance to win. For all the game length, blude side generally has better chance to win than red side.



This is a waffle plot that plot the total win games of different sides. This plot could tell that the win rate of blue side is higher than 50% which indicates that blue side has greater chance than red side to win in this year at League of legends.

1. Trained algorithm prediction

The project is to find out the win condition of a game and which variables that could predict the result of the game. Due to the results of the game could only be “win” or “loss”, I convert the character variable to numeric variable of “1” and “0”. And use the glm function to train the training set. Please see the summary below of the algorithm:

##Coefficients:

## Estimate Std. Error z value Pr(>|z|)

##(Intercept) -1.043e+00 2.168e-01 -4.811 1.5e-06 \*\*\*

## sideRed -7.997e-02 6.022e-02 -1.328 0.1841

##gamelength 4.205e-03 5.375e-03 0.782 0.4340

##teamkills 5.768e-01 8.628e-03 66.857 < 2e-16 \*\*\*

## teamdeaths -5.784e-01 8.687e-03 -66.582 < 2e-16 \*\*\*

## dpm -1.029e-03 2.186e-04 -4.708 2.5e-06 \*\*\*

## earned.gpm 6.235e-03 6.112e-04 10.201 < 2e-16 \*\*\*

## golddiffat10 -1.980e-04 9.746e-05 -2.032 0.0422 \*

## golddiffat15 5.704e-05 5.584e-05 1.022 0.3070

##---

## Signif. codes: 0 ‘\*\*\*’ 0.001 ‘\*\*’ 0.01 ‘\*’ 0.05 ‘.’ 0.1 ‘ ’ 1

The test set using the trained algorithm came out to have 0.92 correlation with the original result which means that this model is a really good model that could predict most of the win condition.

The model is telling us that the most relevant variables are teamkills, teamdeaths, damage per minute and gold earned per minute.

From the result we can tell that if a team get more kills and die less, also remaining high damage to opponents and earning more money could mostly win the game.