

Using Machine Learning to

Predict the Winning Team for

League of Legends

By: Garrett Leckrone

Contents

[Introduction 3](#_Toc501131636)

[Foreword 3](#_Toc501131637)

[Data Sets 3](#_Toc501131638)

[Hypotheses 4](#_Toc501131639)

[Test Plan 4](#_Toc501131640)

[Results 5](#_Toc501131641)

[Test 1. 5](#_Toc501131642)

[Test 2. 6](#_Toc501131643)

[Test 3. 6](#_Toc501131644)

[Test 4. 7](#_Toc501131645)

[Part A. 7](#_Toc501131646)

[Part B. 7](#_Toc501131647)

[Part C. 7](#_Toc501131648)

[Test 5. 7](#_Toc501131649)

[Results 8](#_Toc501131650)

[Conclusion 8](#_Toc501131651)

[Test 1 Results. 8](#_Toc501131652)

[Test 2 Results. 8](#_Toc501131653)

[Test 3 Results. 8](#_Toc501131654)

[Test 4 Results. 8](#_Toc501131655)

# Introduction

## Foreword

League of Legends is a real-time strategy game, created by Riot Games. It is a binary game, with two teams battling against each other with only one winner. Each team has 5 players and each of those players has a role that they play. This is similar to the various positions in traditional sports, such as a Goalie. The map is divided in half and is symmetrical, where each team’s base is located at the opposite corner of the map. There are three lanes that go from base to base, Top, Mid, and Bottom. Between each lane is a winding terrain where neutral monsters and objectives spawn. To win the game, one team must defeat the opposing team’s towers and inhibitors in order to get to the Nexus: a guarded crystal located in the deepest location of each team’s base. The Nexus is only able to be damaged once the defending Turrets and 1 Inhibitors have been destroyed. The first team to destroy the opposing Nexus is the winner of the game.

League of Legends was created in America but has quickly spread in popularity to various regions around the world, including China, Japan, South Korea, Europe, Southeast Asia, and many more. It has grown to become regarded as the most popular e-sport and has big names in sports quickly joining the scene such as the Golden State Warriors, Houston Rockets, and the Cleveland Cavaliers. There is a growing need for data analysis to determine the best winning strategies, and in this report we will look at various metrics to try to achieve that goal: predicting the winning team.

## Data Sets

For this project, I found a dataset on Kaggle that has recorded stats for over 50,000 ranked games of League of Legends from the western European servers.

The dataset can be found at this URL: <https://www.kaggle.com/datasnaek/league-of-legends>

This dataset contains the following fields:

* Game ID
* Creation Time (in Epoch format)
* Game Duration (in seconds)
* Season ID
* Winner (1 = team1, 2 = team2)
* First Objective: Baron, dragon, tower, blood, inhibitor and Rift Herald (1 = team1, 2 = team2, 0 = none)
* Champions and summoner spells for each team
* Total Objectives: tower, inhibitor, Baron, dragon and Rift Herald kills each team has
* The 5 bans of each team (Again, champion IDs are used)

This analysis will focus only on the Winner, First objectives, and the Number of Objectives Taken fields.

The Game ID and Creation Time are not necessary fields for predicting the winner.   
Game Duration, Season ID, Champions Picked and Banned, and Summoner Spells Picked could also be used for predicting the winner, but they were not reliable for this analysis. Champions Picked and Champions Banned, for example, are too heavily influenced by the Season and this dataset does not contain enough information on previous seasons.

## Hypotheses

Three hypotheses will be tested for predicting the Winner:

1. The team that kills the neutral objectives, Baron Nashor and Dragons, more than the opposing team will win the game.
2. The team that kills more of the enemy’s towers and inhibitors will win the game.
3. The team that takes more first objectives will win the game.

## Test Plan

In order to test these hypothesis several tests will be run:

1. A Heatmap will be created using Seaborn in order to show any potential correlations between the number of tower kills, inhibitor kills, dragon kills, Baron kills, and Rift Herald Kills and the winner.
2. Another Heatmap will be created to show correlation between the first team to take unique objectives and the winner.
3. A Pie Chart will be used to show the percentage of times that blue team wins vs red team. This will determine if there is any bias for the winning team based on side of the map.
4. The data will be divided into training and testing datasets using SciKit Learn. One third of the data will be randomly chosen to for the training data. A Logistic Regression model will be created and fit to the training data. Lastly, we will use the predict function from the Logistic Regression library to predict the winner for the test data. Lastly, a classification report will be generated using the predictions that were made and then comparing it to the actual answer.
5. Step 4 will be repeated for each of the three hypotheses to determine how well those factors can predict the winner of the game.

# Results

## Test 1.

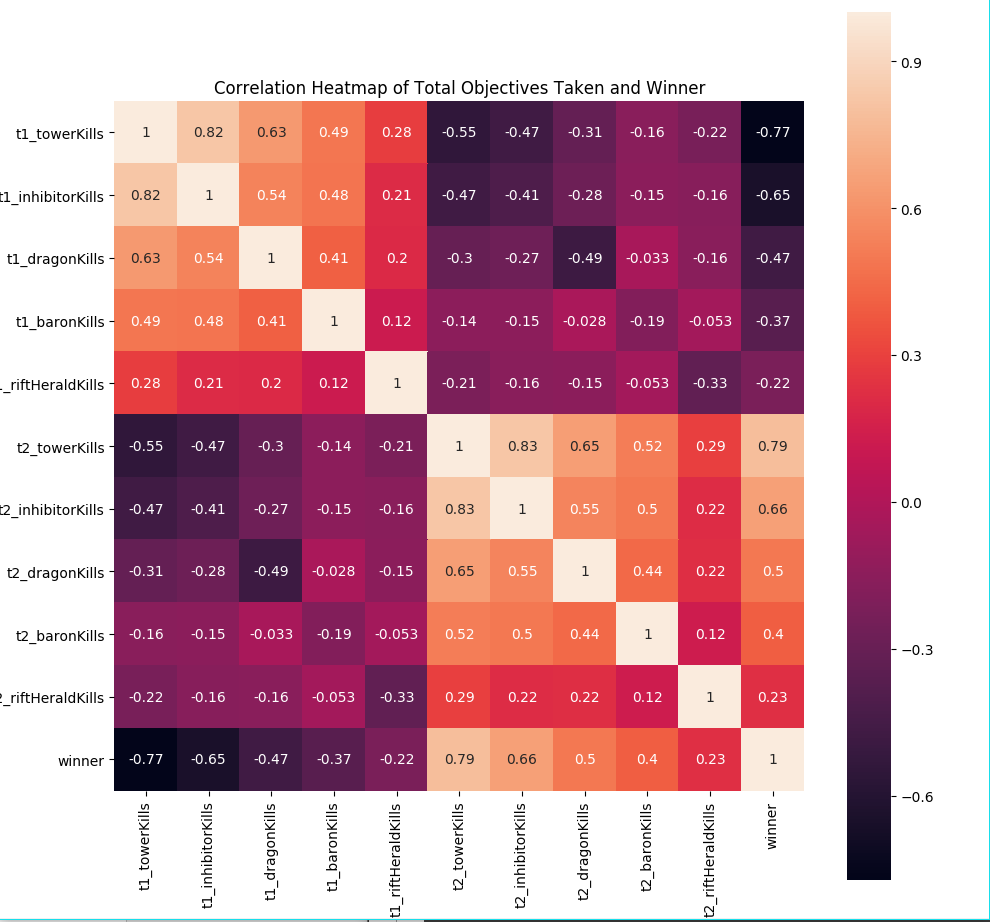


Figure 11.A Heatmap showing potential correlations between the number of tower kills, inhibitor kills, dragon kills, Baron kills, and Rift Herald Kills and the winner of the game.

## Test 2.

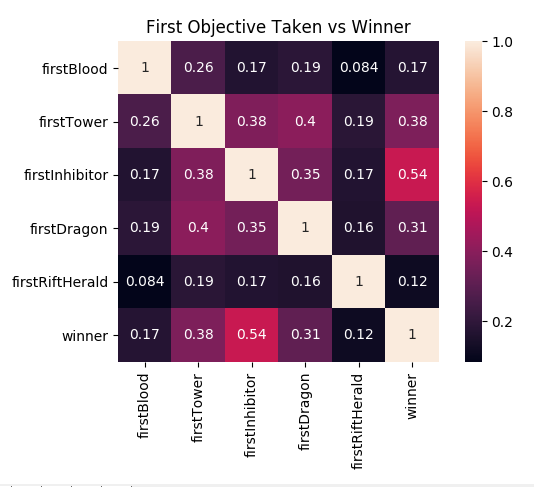


Figure 2. Heatmap showing correlation between the first team to take unique objectives and the winner of the game.

## Test 3.

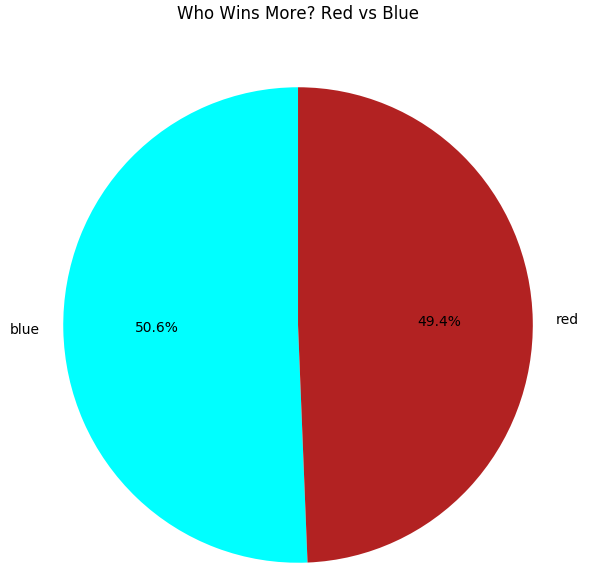
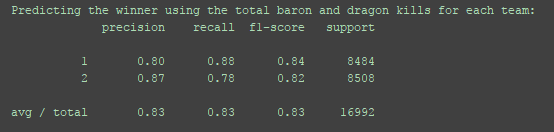


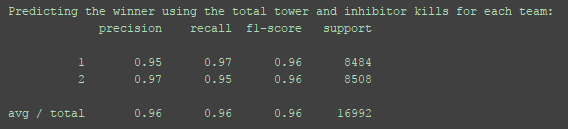
Figure 3. A Pie Chart showing the percentage of games that blue team wins vs red team.

## Test 4.

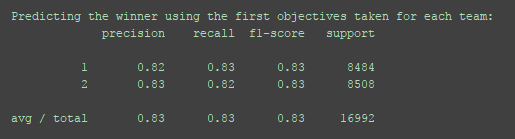
### Part A.



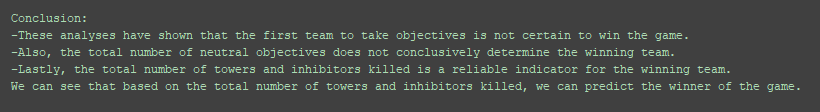
### Part B.



Part C.



## Test 5.



# Results

## Conclusion

### Test 1 Results.

From the first Heatmap, a few correlations stand out. The most notable correlation for winning the game is the number of Tower Kills for each team. With almost .80, there is a strong correlation. Inhibitor kills are the next notable correlation that stands out for winning the game. At .65 and .66, there is a moderate correlation. Less intuitively, the number of Baron Nashor kills and Dragon kills have moderate/weak correlation to winning the game. As a player of the game, I am surprised that Baron Nashor kills do not have a stronger correlation for winning the game.

### Test 2 Results.

The second Heatmap shows the correlations between the team to take unique objectives and the winner of the game. These results are fairly intuitive as a player, but I’m a little surprised that there is not a stronger correlation between winning and taking the first inhibitor. From experience, it has always felt like the team that takes the opposing team’s inhibitor first is very likely to win the game. Also, I am relieved that first blood does not correlate to winning the game, considering all of my teammates are usually monkeys and die first.

### Test 3 Results.

The pie chart showing the percentage of winning teams is also nice to see. With both teams winning about 50% of the time, it shows that the game is balanced for each team and that there are no inherent biases towards either side.

### Test 4 Results.

Using machine learning to make predictions for this project was the most fascinating part. The SciKit Learn library really makes it easy to divide a large dataset into test and training data and then run various types of machine learning algorithms to predict.

The results of the Logistic Regression analyses run for each of the hypotheses:

1. The model was able to correctly predict the winner of the game using the total Baron Nashor kills and dragon kills for each team for 83% of the test sets. This moderately supports the hypothesis but does not conclusively confirm the claim.
2. The model also correctly predicted the winner for 83% of the test sets using the first objectives taken for each team. This is very interesting to me that it was equally as likely to predict the winner of the game based on the amount of first objectives they take as it is to predict the winner based on the total number of neutral objectives. I was very surprised by this, I expected the predictability to deviate from one another.
3. Lastly, our model was very accurate as predicting the winner of the game with a precision of 96% based on the total number of towers and inhibitors that a team takes. This is highly intuitive to me, because in order to win the game you must destroy many towers. And it is very rare that the losing team has less standing turrets and inhibitors than the winner.

Overall, I was very happy with these results and the different conclusions that were able to be formed from them and it was a pleasure to be able to work with a data set that was personally interesting to me.