Predicting the Results of LoL Games

An Analysis of the 2018 League of Legends World Championships

Lewis Zou

12/8/2019

Abstract

League of Legends (LoL) is a multiplayer online battle arena that allows users to fight in either teams of 3 or 5. Each player, or "summoner", controls a champion with unique abilities and item builds, and sets out to destroy the enemy soul point, the "Nexus". LoL games are discrete — the progress a player makes in one game does not translate to the next. League became the most played PC game in Europe and North America in 2012 (by hour). Riot Games, the LoL developer, estimates that over 100 million users play the game each month. With its immense community, League has grown its entertainment and competitive segments. They maintain strong viewership in streaming and video services like Youtube and Twitch.tv, often ranking first in the number of hours watched. This popularity has developed its competitive scene — competitions exist in the United States, Europe, China, South Korea, and more. These seasonal long regional competitions culminate in a series of playoffs, in which top teams move on to the annual world championships. In the championships's early days, the focal points of viewership were flashy outplays and individual dominance. As the scene has become more competitive year over year, the game has become more focused on strategy and analysis. In this project, I aim to find a way to predict the results of these games with factors that impact these matches.

Background and Significance

League of Legends (LoL) is a multiplayer online battle arena with competitions that exist in the United States, Europe, China, South Korea, and more. These seasonal long regional competitions culminate in a series of playoffs, in which top teams move on to the annual world championships. The League of Legends World Championships brings 16 teams from various regions to compete for the "Summoner's Cup". In a format similar to that of traditional sports, teams are seeded and drawn into groups. The 2018 finals were watched by nearly 100 million people, a staggering number that reflects the popularity and impact of LoL. Given its worldwide influence, LA2024, the group that oversees Los Angeles' bid for the summer 2024 Olympic Games, would highly consider including eSports and LoL's virtual reality and entertainment in the games. I chose the 2018 championships because that was the year my favorite team, SK Telecom of South Korea, failed to make the tournament. As three-time winners of the entire competition (2014, 2016, 2017), I was shocked and disappointed. However, I wanted to understand the new environment. I wanted to take a deeper look at the statistics that mattered in 2018.

Methods

The 2018 World Championships dataset only contains about 1,500 rows. Although this may seem low, this set includes datapoints about every game that occured in the championships. Greater detail about collection and times were not available — this dataset presents the necessary indicators (gold, dragons, barons, etc.) that help predict the result of games. This data is extensive — it includes datapoints about nearly every part of the game. Starting from the left-most column, the data presents the game id, url, and general time statistics — these columns I will remove in the future, but will keep for now for the label. After those, we get into playerid, side, position, player, and team. These points will help keep track of each player and their historical performance, and will be a critical factor in connecting each player with their impact on the potential to win. Diving deeper, we get champions and bans — the draft can often determine the result of a game, and I will use this as part of my calculations. The data then describes the actual games that are played. We get game length, kills, deaths, assists, and more. The data includes objective and monster kills, things that help champions progress through the game. Often, there is a high correlation between objective control, gold generation, time, and performance — we will use these datapoints to see if we can in hindsight predict the outcome of the game. The data is extremely comprehensive and includes sections that are not typically relevant to the outcome of the game. They can definitely have an impact on a player's ability to move across the map, defend their towers, and generate gold, which can in turn impact the result.

Errors and Complications

There were many errors that occured and could have occured. Although we aimed to be accurate and concrete in our conclusions, we lack of data and number of possible confounding variables make it difficult to truly produce a strong conclusion. Notably, we could only use World Championship game data. We had to filter out the play in stage, in which the top teams do not play. Those games tend to boost the stats of teams that win the stage, who may end up performing poorly in the actual tournament. As such, we could only use around 1,000 data points of players and teams. After often further condensing this down to teams, we had very few datapoints to create models off of. The issue is that it was extremely difficult to get those datasets — we could not use other tournament data and other LoL game data to produce our results. The playstyle and boosts given to each champion and objective also varies every few months, making it difficult to compare the statistics from different games. The difference in preparation and drastic variances in level of competition make a side by side comparison biased and skewed. Also, the number of confounding variables was also concerning. Given that we could not find data tracking each second of the game, we could only use end-of-game and 10/15 minute time stamps. Because the game is based on strategy, shotcalling and other game statistics could greatly skewed the data. Although we see a correlation between wins and dragons and barons, the win could be a result of strong teamwork and shotcalling instead of strictly objective control. Furthermore, shotcalling is even critical to game control, and as something not quantifiable, makes it difficult to truly understand what matters in predicting the result of a game. Furthermore, given that the meta (playstyle) changes year over year, the importance of different objectives also changes year over year. In some metas, capturing the baron may boost offensive statistics much more than capturing dragons, and as such would increase the impact of baron on the overall rseult of the game.

Results

We aimed to create a model that can potential predict the results of games. Often, analysts declare that gold is a major reason for a victory or loss. They emphasize the importance of gold difference early on — however, from our analysis, we have found that gold difference fails to show a major impact. Instead, we found that the number of dragons and barons that teams captured increased their chances of winning even more. While gold remains important in thriving in the game, the actual gold difference in the early game before dragons and barons are taken is too minimal to create a major separation. We now create two models; one that predicts based off gold difference at 10 and 15 minutes, and another that predicts based on dragons and barons killed.

Predicting Results using Gold

Gold Difference at 10 and 15 Minutes

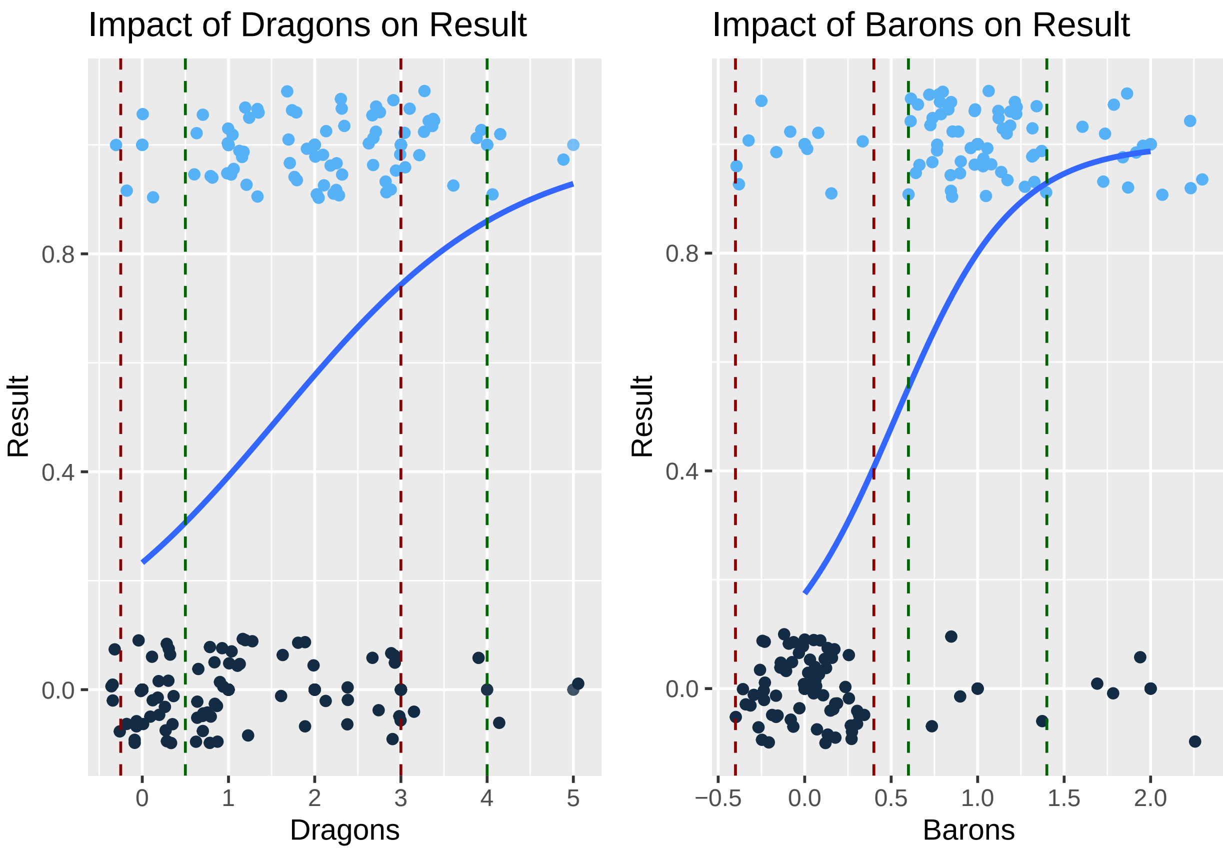
	Type	Percentage
accuracy	binary	71.43%
kap	binary	42.86%

Predicting Results using Objectives

Dragon and Baron kills

	Type	Percentage
accuracy	binary	88.31%
kap	binary	76.62%

When comparing the impact of dragons and barons, we found that a single baron was more impactful.



Conclusions

League of Legends is a game of strategy and objective control. Everything is tied — in essence, gold generated helps gain items to take dragons and barons, which ultimately is a major difference maker in the result of games. When looking at the 2018 data we started with, we see that the model based on gold difference had an overall accuracy of .71, doing better than random by .43. Even though the linear models previously found showed only slight increases to winning, we find that this predictive model is relatively accurate. However, we find that the model based on dragons and barons has an overall accuracy of .88, nearly 25% more. Notably, it performs better than random by .77, a near 80% increase over the model of gold difference. This extends further when using these models on other years. Although gold difference and gold generated do have a critical impact on the result of games, its important to note that gold difference, while important, is not the most defining factor in the game. While gold difference does give an early indicator on the potential result of the game, the barons and dragons captured also present an important factor. In conclusion, we are in agreement that gold difference is extremely important. From our findings however, the impact of dragons and barons should not be minimized. They hold the key as important indicators in the result of League world championship games. This can be applied further — even when considering normal League of Legends games, players can focus more on objectives than actual kills. This is a huge change in mindset — given that the ultimate goal is to win teamfights, players forget the importance of objectives in improving their chances to win teamfights. This project was a great way to better understand the strategy and mechanics of the game, and helps better explain the strategies that LoL teams adopt.