Decoding Victory in League of Legends: A Comparative Study of Early Game Statistics*

Cristina Burca

April 20, 2024

Abstract

Esports, especially League of Legends, have rapidly evolved into a major professional sports phenomenon, featuring international tournaments and attracting a massive global player base. This paper delves into the statistical analysis of League of Legends Worlds Championship data from 2018-2023, focusing on early game variables gold, expereince and creep score and their impact on the win probability. Our results indicate that while generally higher early game statistics compared to the opposing team are associated with increased chances of victory, our findings suggest that paradoxically, lower statistics at the 10-minute mark may actually enhance the likelihood of winning, highlighting the complex dynamics of strategic gameplay in League of Legends.

Table of contents

1 Motivation

Esports (short for electronic gaming) as professional sporting tournaments have grown exponentially in popularity, due to the rise of multiplayer games such as Valorant, Counter-Strike: Global Offensive, Tom Clancy's Rainbow Six Siege, and the focus of this paper, League of Legends. Developed by Riot Games in 2009, League of Legends is a competitive multiplayer online battle arena (MOBA) with an active player base of over 150,000 million as of 2022. It has grown to be the largest esport in the world – with some events having higher viewership than NBA and Stanley Cup championships – and consequently promoted

---

*Code and data are available at: https://github.com/crisburca/LoLstats.git.

the popularity of the esport industry (Hardenstein 2017). International tournaments are held every year, with leagues and teams established around the world. This paper will focus on statistics from the League of Legends World Championship hosted by the developers, Riot Games. Commonly known as Worlds, it is a championship tournament hosted annually, involving 24 teams from all around the world.

There are many factors to be aware of when in a multiplayer team-based game, with some invariable factors, and others that are completely up to chance. However, the games are also greatly influenced by strategy – with decisions made before and during the first stages of the match that can impact the outcome of the game. Games typically last between 30-60 minutes but can range much shorter or longer. Arguably, the early stages of the game (referred to as "early game") can influence the game's outcome the most. As the game goes on, the increase of players' gold income, creep score, and experience levels when compared to their opponents influence the outcome. These variables and their importance will be further discussed in the next section.

This paper will examine these variables in the early stages of the game, specifically during the 10 and 15-minute marks, to understand their significance in winning a game. We also compare them with older data as well as data from 2023, to see the difference in variables. These variables of interest are tested through a logistic regression model. The findings suggest that lower 10-minute statistics might paradoxically be linked to a greater probability of winning, relative to statistics at the 15-minute mark. The accumulation of team statistics early in the game, such as securing multiple kills, can significantly boost a team's gold reserves while not necessarily reflecting a similarly high experience or creep score.

In section 1.1 of the paper, a brief introduction to the game is defined, as well as an introduction to the variables in the game and their correlations. In section 2, the data set is and the variables of interest are discussed and observed. Section 3 will establish the logistic model, and how it was created. Section 4 discusses the results.

1.1 Introduction to the Game

In the game, there are two teams of five players each, competing within a square-shaped map. They begin on opposing corners, of which each half is referred to as either the red or the blue side. Both sides have a single "nexus," and the game is ended when one of the nexuses is destroyed. The objective for each team is to reach the other end of the map, and destroy the other team's nexus in order to win. Each player controls a character, or "champion," of their choice, from an available pool of 160+ unique options with distinct abilities and play styles. During a match, players can advance towards the enemy's base through three lanes, two of which go around the perimeter of the map, and one that cuts diagonally across. These lanes are called "top," "mid(/middle)," and "bot(/bottom)," of which both teams have designated players to occupy the lanes. In each lane on both sides of the map, there are two "turrets" that must be destroyed to advance towards the enemy "base," where the nexus is located. Both red and blue sides spawn waves of minions, or "creeps," every 30 seconds - consisting of six or seven creeps that advance down each lane. The allied creeps contribute to destroying the opposing turrets, and will grant enemy players gold if killed. In each lane, players will defend their turrets by defeating creeps and engaging with enemy laners, to push towards and destroy the opposing turrets. Players earn gold and experience through killing creeps, defeating or assisting in the defeat of enemy champions, and destroying turrets. Each player will passively earn gold, but capturing objectives for greater income and experience can be used to purchase items and level up abilities faster - consequently making one's champion more powerful, and more efficient in combat. Killing creeps is the main source of gold for players within the lanes. The fifth player that is not assigned to a lane is the designated "jungler," of who passes through the "jungle" in between the lanes and kills neutral jungle monsters in order to get gold and experience. The jungler's role is to assist in defeating enemy laners, and secure neutral objectives that can offer team-wide buffs and gold. By acquiring more gold and experience than the enemy team through creeps, objectives, and kill/assist count, a team is more effective in resisting against the other and pushing towards the enemy's nexus in order to win the game (Games 2024).

The key things to note, and the variables that will be examined in this paper:

- Gold is primarily earned through killing creeps, and also earned through killing or assisting in killing opponents, and completing objectives. When a champion accumulates consistent kills without dying,

the champion will receive a bounty that the opposing team gets when killing the champion. Gold is then used to purchase utility items specific to champions to become stronger.

- Experience is earned the same as gold, but is counted as levels from 1-18, where each level, the champion is able to upgrade one of their four abilities. Experience is also gained through standing near minions that are being killed.
- Creep score is counted by the amount of creeps a player has killed. Comparing each player's individual creep score to that of their opposing laner can provide an estimation of how well they are performing. Although all of these variables correlate, they are not necessarily dependent on each other, and they do not grow together.

An important detail to note, which will become a relevant theme throughout the analysis, is the developers release a considerable update to the game at the beginning of each year. This update includes changes to the map, champion abilities, utility items, in-game mechanics, and other aspects of the game. Players and teams may change their play style and strategy to account for these updates. Thus, it is difficult to establish trends for variables over the years, as it is important to account for the varying fundamentals of these statistics.

2   Data

The data is sourced from OracleElixir (Sevenhuysen 2024), a popular League of Legends esports stats website used by coaches, analysts, broadcasters, and many others. The website is composed of upcoming, live, and recent statistics of matches, where statistics can be viewed by player, team, and league. The data files, categorized by year, are available through the website and are updated daily. The data contains many in-depth analytics on players and tournaments played in all championships and leagues including the World Series. We analyze data from years 2108 to 2023 with gaming statistics for each match.

The analysis of the data is done using the R programming language R Core Team (2023), and cleaned using Wickham et al. (2019). Plots, tables, and models were created using the packages 'modelsummary' Arel-Bundock (2022), 'stanarm' Goodrich et al. (2022), 'kableExtra' Zhu (2024), 'grideExtra'Auguie and Antonov (2022), Wickham (2007). Packages 'tinytex' Xie (2024), 'readr' Wickham, Hester, and Bryan (2023), and 'here' Müller (2020) were used for importing the data and formatting the document.

2.1   Variables of Interest

Each team and individual player have over 100 variables recorded for each match. We chose to focus of Worlds Championship data from 2018 to 2023. The variables considered for this paper are:

- `golddiffat10`: the total difference of gold at 10 minutes between teams
- `xpdiffat10`: the total difference of experience at 10 minutes between teams
- `csdiffat10`: the total difference of creep score at 10 minutes between teams
- `golddiffat15`: the total difference of gold at 15 minutes between teams
- `xpdiffat15`: the total difference of experience at 15 minutes between teams
- `csdiffat15`: the total difference of creep score at 15 minutes between teams

Each total difference is calculated by summing the respective statistics from all players on each team at the specified time and then subtracted.

We examine the difference in the variables between teams instead of the tangible values. Due to the many factors in the game and the many sources of income and experience, the values of each variable can differ greatly from match to match. Instead, we focus on the total difference for each variable between teams, to center our analysis on the importance of each variable relative to the likelihood of winning.

In Table 1, we introduce the data by displaying averages of differences in gold, experience, and creep score at 10 and 15 minutes between teams by year. Note that all of these variables can highly vary between games. The average differences in gold between teams usually vary between 500 to 2,000 gold, average differences in experience vary between 300 to 1,000, and average differences in creep score vary between 5 and 20 creeps.

Table 1: Average difference in Gold, Experience and Creep Score at 10 and 15 minutes by Year

Year	Gold at 10	XP at 10	CS at 10	Gold at 15	XP at 15	CS at 15
2018	580.07	460.38	10.95	1411.52	1138.50	20.06
2019	520.18	340.15	7.12	1582.01	874.52	14.28
2020	750.52	549.97	11.12	1732.56	1161.57	21.51
2021	909.69	493.14	9.73	2000.14	1027.89	17.79
2022	749.04	343.09	5.72	1696.67	875.41	14.26
2023	564.02	394.23	8.61	1292.53	899.05	14.56

In Figure 1, we plot the values displayed Table 1 along with the overall averages for easier observation.
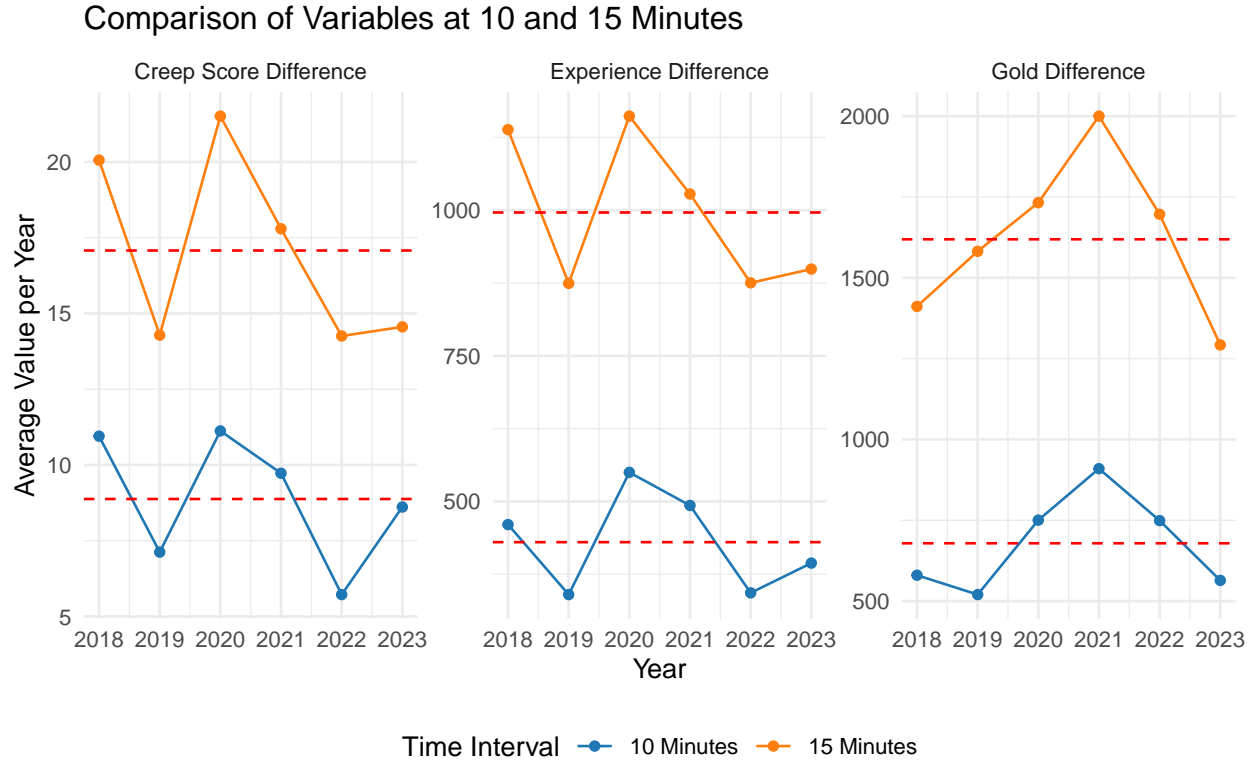

Figure 1: Plot of average difference in Gold, Experience and Creep Score at 10 and 15 minutes by Year

We notice that these differences greatly fluctuate from year to year. Apart from variation in games, this fluctuation can also likely be due to the updates released by the developers every year, which affect the game entirely. Although the values themselves fluctuate, the difference between the averages stays approximately uniform. From this, we deduce that between the 10 and 15 mark in the game, on average, the increase in all variables is uniform, but over the years, the actual values of the difference change.

With these ideas in mind, it is hard to base our current predictions on games played in tournaments on older releases of the game. However, for overall analysis, we examine the variables over the years for overall predictions, as well as in the current year for current predictions.

We introduce the averages of all variables for winning and losing teams, as well as their difference in Table 2.

Table 2: Difference in Gold, Experience and Creep Score at 10 and 15 minutes for years 2018-2023

Result	Gold at 10	XP at 10	CS at 10	Gold at 15	XP at 15	CS at 15
Lost	15323.12	18134.15	313.64	23892.57	29026.46	500.87
Won	16003.27	18560.71	322.41	25512.12	30017.17	517.81
Difference	680.15	426.56	8.77	1619.55	990.71	16.94

Notice that the average difference between the winning and losing teams of each variable is much greater at 15 minutes compared to 10 minutes. This suggests that an advance in the first 10 minutes of the game can exponentially raise the winning team's statistics in the subsequent 5 minutes.

It is evident that the winning team has higher statics overall, but we are interested in the importance of these early-game variables and their influence on winning the game. To evaluate this, we build a logistic regression model.

One small thing to note is the small bias of teams winning that begins on the blue side of the map relative to the red side throughout the game's history. In the past, it has been known in the community that the blue side is 'advantageous', most likely due to the positioning and the map layout. We show this is true in Worlds Championship matches as well, in Table 3, where the average win rate for the blue side is greater than the red side, with win rates of 53% and 47% respectively. In Figure 2, we notice this bias stays true throughout the years.

Table 3: Win percentages by side for all Worlds data from 2018-2023

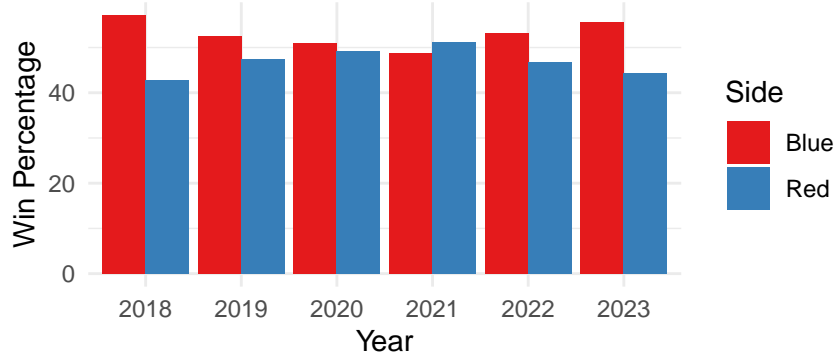| Side	Avg. Win rate
Blue	53.04
Red	46.96

Figure 2: Win percentages by side for all Worlds data from 2018-2023

3   Model

The goal of this model is simple. We analyze the relation of gold, experience, and creep score difference at 10 and 15 minutes to understand their correlation to the likelihood of winning a game through a logistic regression model. We create different models for data from 2018 to 2023, as well as specifically 2023, to understand the overall and current correlation.

3.1   Model Set-up

Since each of these differences is calculated for both teams in each game, there are duplicate entries for each variable, one for the team with the positive difference and the other with a negative difference. We consider the statistics of teams that played on the blue side of the map to account for this duplication.

We define the variables in the model Equation 1. y_i is the probability that the game is won, β_0 is the coefficient of the intercept, α_i is the average gold difference at 10 minutes, δ_i is the average experience difference at 10 minutes, γ_i is the average creep score difference at 10 minutes, θ_i is the average gold difference at 15 minutes, ϕ_i is the average experience difference at 15 minutes, and κ_i is the average creep score difference at 15 minutes. We look at differences in variables as opposed to individual team statistics as we are not interested in the actual value of each variable, but rather the importance of each individual variable.

$$
\begin{aligned}
y_i|\pi_i &\sim \text{Bern}(\pi_i) \\
\text{logit}(\pi_i) &= \beta_0 + \alpha_i + \delta_i + \gamma_i + \theta_i + \phi_i + \kappa_i \\
\alpha_i &\sim \text{Normal}(0, 2.5) \\
\delta_i &\sim \text{Normal}(0, 2.5) \\
\gamma_i &\sim \text{Normal}(0, 2.5) \\
\theta_i &\sim \text{Normal}(0, 2.5) \\
\phi_i &\sim \text{Normal}(0, 2.5) \\
\kappa_i &\sim \text{Normal}(0, 2.5)
\end{aligned}
\tag{1}
$$

We run this model in R Core Team (2023) using the 'rstanarm' package Goodrich et al. (2022). We scale this model since gold, experience and creep score are different 'units' and are not recorded on the same scale. We model and plot our findings based on this scale, then we scale back to the original units for insight on the specific values.

We expect for all higher values to imply a winning team, but we are interested in which variables are more significant. One might assume that all variables are equally as important, and we will use our model to either prove or disprove this.

4   Results

4.1   Model 1: From years 2018-2023

Model 1 illustrated in Figure 3 displays the logistic regression model of Worlds matches from 2018- 2023 for average team difference in gold, experience, and creep score at 10 and 15 minutes relative to win rate. Since we have scaled the model, we can assume all variables are counted in the same units. Notice that the intercept is not exactly at 0, meaning that there is not a 50% chance base win rate dependent on both sides. This is due to the analysis in Section 2.1, Table 3, where the bias was discussed. Although something to note, it does not skew the results. The model indicates that a negative difference in experience and creep score at 10 minutes is associated with an increased likelihood of winning. The difference in gold at 10 minutes and creep score at 15 minutes suggests a small increase in the likelihood of winning. The difference in experience especially gold at 15 minutes increases the likelihood of winning. Based on the first 15 minutes of a game, we can utilize this model to predict the outcome of the winning team.
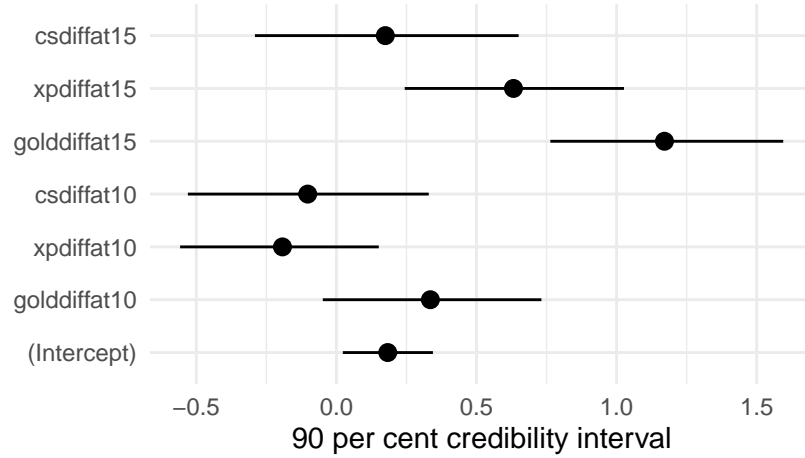
Figure 3: Model 1: Logistic regression model of Worlds matches from 2018 to 2023 for difference gold, xp and cs at 10 and 15 minutes relative to win rate

Table 4: Model summary of Model 1

	Probability of winning
(Intercept)	0.183
golddiffat10	0.335
xpdiffat10	−0.192
csdiffat10	−0.102
golddiffat15	1.171
xpdiffat15	0.632
csdiffat15	0.175

Table 5: Model summary of Model 2

	Probability of winning
(Intercept)	0.256
golddiffat10	0.341
xpdiffat10	-0.179
csdiffat10	0.307
golddiffat15	0.925
xpdiffat15	0.669
csdiffat15	0.375

Table 4 shows the coefficients of the predictor variables in Model 1. The table provides the values of the coefficient representing the intercept, and the other coefficients represent the average difference of the winning team for each independent variable.
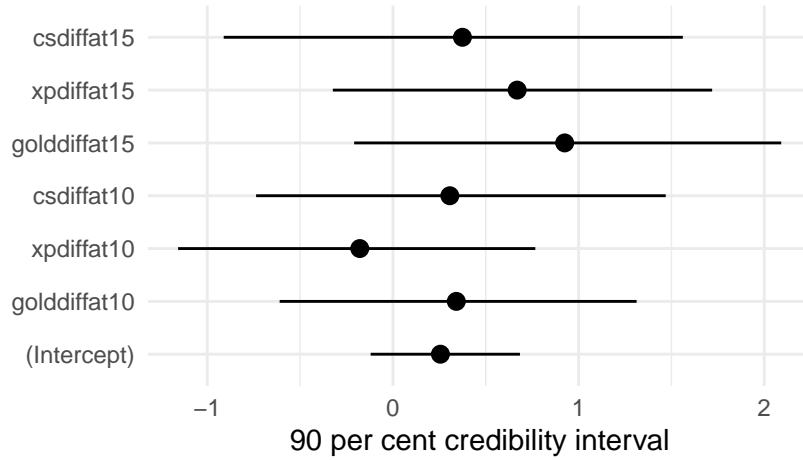
4.2 Model 2: Year 2023

Figure 4: Model 2: Logistic regression model of Worlds matches from 2023 for difference gold, xp and cs at 10 and 15 minutes relative to win rate

Figure 4 displays the logistic regression model of Worlds matches from 2023 for average team difference in gold, experience, and creep score at 10 and 15 minutes relative to win rate. Quite similar to Model 1 in Figure 3, the same variables are correlated with a higher probability of winning, expect for 'csdiffat10', which is instead positively correlated with a higher win probability. Observe that the values vary more for all variables, as we are considering a smaller data set compared to Model 1.

Table 5 shows the coefficients of the predictor variables in Model 2. The table provides the values of the same coefficients as Table 4, but considering data only from 2023.

5 Discussion

5.1 Model Analysis

As early game statistics can determine the rest of the game, it is important to consider these early game variables. Taking into consideration team statistics, if multiple players of one team secure kills in the early game, they will have a higher amount of gold relative to experience and creep score. According to Figure 3, if a team has a higher gold count at 10 minutes but a lower experience and creep score, we can assume that

they have more kills at that stage. This suggests that team kills before 10 minutes is correlated to higher difference statistics for the variables at the 15-minute mark, and thus a higher likelihood of winning.

Considering the correlation of a negative difference in experience and creep score to a higher likelihood of winning, this could suggest that game mechanics are more forgiving towards mistakes during the first 10 minutes of the game. This could also suggest that teams with "scaling" champions, ones that are weaker in the early game due to lack of items and experience, encounter "spikes" in power after the first 10 minutes. These spikes are attributed to becoming stronger due to unlocked abilities, or utility items that can amplify a champion's power. Another possibility for these negative differences can be from the losing team acquiring "bounties." If a team is ahead early in the game, individual champions will likely be assigned a bounty from securing kills - and it will be rewarded to whichever enemy "shuts down" the champion with the bounty. If this is the case, most teams will target the player with the bounty, and attempt to "feed" it to a specific player on their team in order to assist them with gaining gold and experience. This specific player is normally the "carry" of the team, which can benefit the most from a higher income by dealing more damage and/or being harder to kill. If a team's carry is fed a bounty, they can begin to "snowball" through their power spike, in which they collect more kills and secure more objectives. For this reason, a team having a weaker start to a game suggests they have champions that grow much stronger as the game progresses, and thus benefit in the later stages. Additionally, a losing team can capitalize on the overconfidence or blunders of the winning team, and be awarded bounties that a carry can utilize to reverse the disadvantage.

These trends are also ongoing in 2023 shown in Figure 4. The variables are consistent in 2023, with the only difference being that a slightly positive creep score difference is advantageous compared to the overall trend.

Overall, there are many reasons to consider the negative difference in the two variables relative to the rest, and this could suggest that a weaker starter to a game can be more beneficial and increase the likelihood of winning.

5.2   Next Steps

Considering the multitude of factors that contribute to victory, including individual gameplay and team collaboration, predicting the precise aspects that lead to a win, particularly in the game's early phases, is challenging. However, it is important to analyze what variables suggest an increase in the likelihood of winning.

Due to the frequent updates that significantly impact various aspects of the game, predicting outcomes for the 2024 Worlds Championship remains difficult For better analysis, it would be prudent to utilize data from other recent tournaments that have been conducted under the latest game update and involve teams qualified for the 2024 Worlds. This approach ensures that our model is grounded in the most current and relevant data. Additionally, a valuable subsequent analysis could involve examining the origins of gold accumulation for each team or player and exploring the relationship between these sources and their success rates in matches.

Another detailed analysis could involve scrutinizing in-game statistics in the context of game updates. It is possible that developers are aiming to promote a specific playstyle through these updates, influencing how strategies evolve over time.

References

Arel-Bundock, Vincent. 2022. "modelsummary: Data and Model Summaries in R." *Journal of Statistical Software* 103 (1): 1–23. https://doi.org/10.18637/jss.v103.i01.

Auguie, Baptiste, and Anton Antonov. 2022. "gridExtra: Miscellaneous Functions for 'Grid' Graphics." https://haozhu233.github.io/kableExtra.

Games, Riot. 2024. "How to Play." https://www.leagueoflegends.com/en-us/how-to-play/.

Goodrich, Ben, Jonah Gabry, Imad Ali, and Sam Brilleman. 2022. "Rstanarm: Bayesian Applied Regression Modeling via Stan." https://mc-stan.org/rstanarm.

Hardenstein, Taylor Stanton. 2017. ""Skins" in the Game: Counter-Strike, Esports, and the Shady World of Online Gambling." https://scholars.law.unlv.edu/cgi/viewcontent.cgi?params=/context/glj/article/1118/&path_info=Hardenstein_Formatted.pdf.

Müller, Kirill. 2020. "Here: A Simpler Way to Find Your Files." https://cran.r-project.org/package=here.

R Core Team. 2023. *R: A Language and Environment for Statistical Computing.* Vienna, Austria: R Foundation for Statistical Computing. https://www.r-project.org.

Sevenhuysen, Tim. 2024. "Match Data Downloads." https://oracleselixir.com/tools/downloads.

Wickham, Hadley. 2007. "Reshaping Data with the Reshape Package." *Journal of Statistical Software* 21 (12). https://www.jstatsoft.org/v21/i12.

Wickham, Hadley, Mara Averick, Jennifer Bryan, Winston Chang, Lucy D'Agostino McGowan, Romain François, Garrett Grolemund, et al. 2019. "Welcome to the tidyverse." *Journal of Open Source Software* 4 (43): 1686. https://doi.org/10.21105/joss.01686.

Wickham, Hadley, Jim Hester, and Jennifer Bryan. 2023. *Readr: Read Rectangular Text Data.* https://CRAN.R-project.org/package=readr.

Xie, Yihui. 2024. *Tinytex: Helper Functions to Install and Maintain TeX Live, and Compile LaTeX Documents.* https://github.com/rstudio/tinytex.

Zhu, Hao. 2024. "kableExtra: Construct Complex Table with 'Kable' and Pipe Syntax." https://haozhu233.github.io/kableExtra.