Unlocking League of Legends Strategy: Insights from Challenger Matches in 2020

Jason Kim
Professor Steven Rashin
Fall 2023

The University of Texas at Austin

Statistics + Data Sciences

Introduction

Motivation: The reason why I chose League of Legends as my research topic is because I grew up playing the game, which brings a sense of nostalgia, and a large part of winning involves strategic elements of gameplay. League of Legends is a highly complex videogame that has a crazy number of factors that can impact game performance. Thus, investigating key elements of the game and its objectives (i.e., first Baron kill, gold accumulation, and damage dealt) would fulfill my curiosities of the game.

Research Questions:

Research Question 1 – In League of Legends, does securing the first Baron kill increase the likelihood of dealing more damage to the enemy team's champions?

Research Question 2 – In League of Legends, how does the total amount of gold earned by a team relate to the total damage dealt to the opposing team's champions?

Methods

Data Collection: The data collection comes from Kaggle, specifically a dataset that details challenger (highest rank) ranked matches in 2020. The individual sampling units for this study will be League of Legends matches, otherwise referred to as "games." The final sample size is n=26904 matches.

Measures: I will be focusing on one team (blue) for this project. In extracting the relevant variables, I plan on using R studio to subset the dataset of the relevant columns – *gameId, blueFirstBaron, blueTotalGold, blueChampionDamageDealt.*

The <u>response variable</u> will be the total damage dealt to champions (of the opposing team) by the chosen team (*blueChampionDamageDealt*), measuring the cumulative amount of damage inflicted upon the enemy team. This game stat has predominately been used as a metric to determine the contribution of each team/player, hence why I chose it as the main response variable:

- **Measurement**: typically done games themselves and displayed as a post-game statistic.
- **Units**: dmg pts ("damage points")

The <u>numeric explanatory variable</u> will be *blueTotalGold*, a measurement of the total gold earned by the team throughout the game:

- **Measurement**: typically recorded in the game's post-game statistics.
- **Units**: gold

The binary categorical explanatory variable will be "First Baron Kill." There are two possible values: *yes or no*. If First Baron Kill = yes, then it is expected that it'll be positively related to the "Damage dealt to champions," meaning it'll likely result in the team having a higher output of damage dealt to champions on the enemy team.

Analysis Method: We used a general linear model (GLM) – specifically utilizing linear regression methods – running it in R studio.
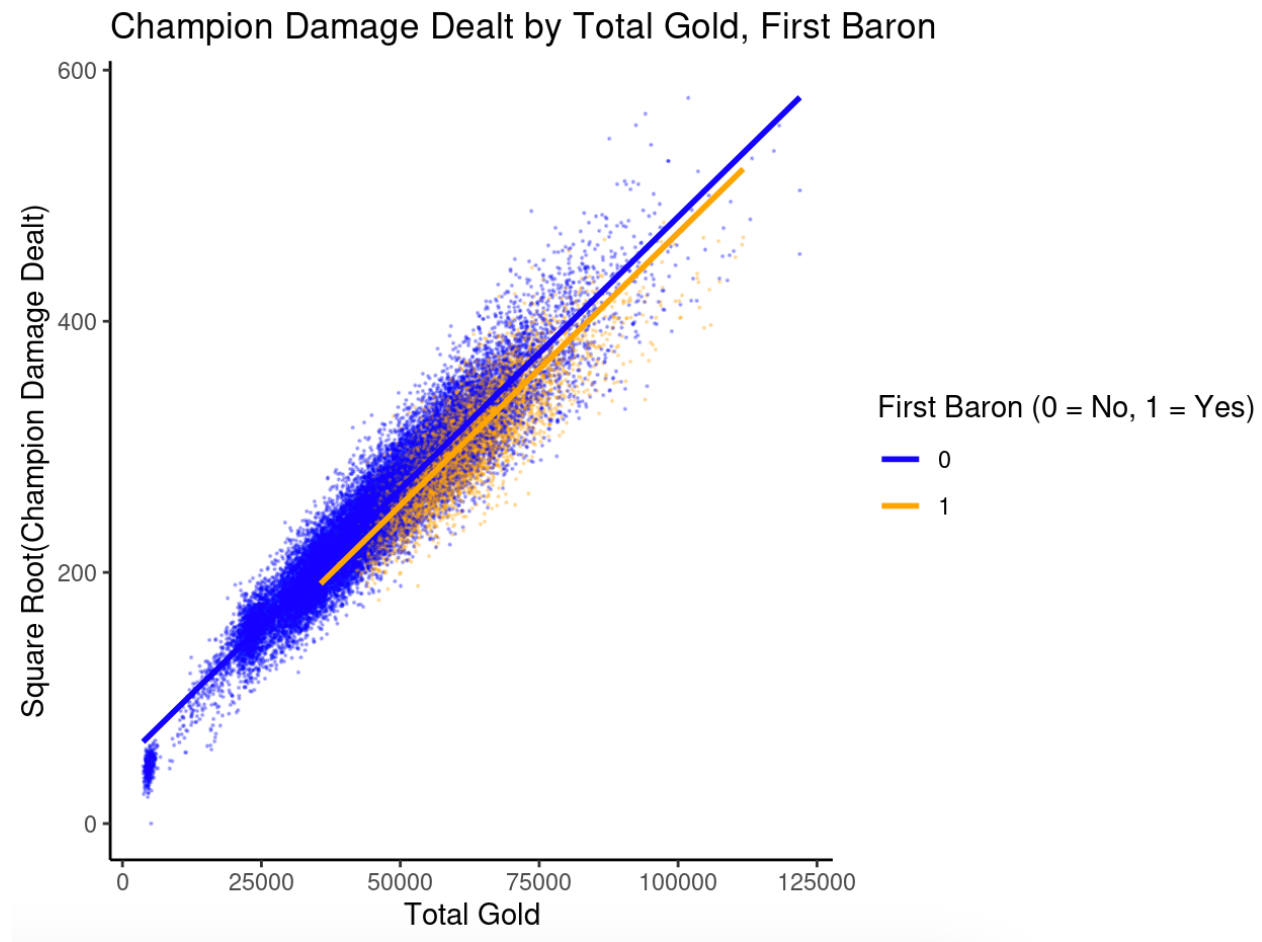
Descriptives

Table 1 – Descriptive Statistics

	Center (Median)	Spread (IQR)
Total Champion Damage Dealt	65013 dmg pts	46955 dmg pts
	Center (Mean)	**Spread (IQR)**
Total Gold	48169 gold	21465 gold
	No	**Yes**
First Baron Kill	20337	6567

Results

Table 2 – Model Results: **Model df:** 26901, *Model fit statistic*: 0.8979392

	Estimate	t-stat	St. Error	P-Value
Intercept	4.900	108.53	4.515e-01	<2e-16
Total Gold	4.342e-03	454.12	9.561e-06	<2e-16
First Baron Kill	-1.277e+01	-37.08	3.445e-01	<2e-16

Multivariate Plot

Champion Damage Dealt by Total Gold, First Baron

First Baron (0 = No, 1 = Yes)

— 0
— 1

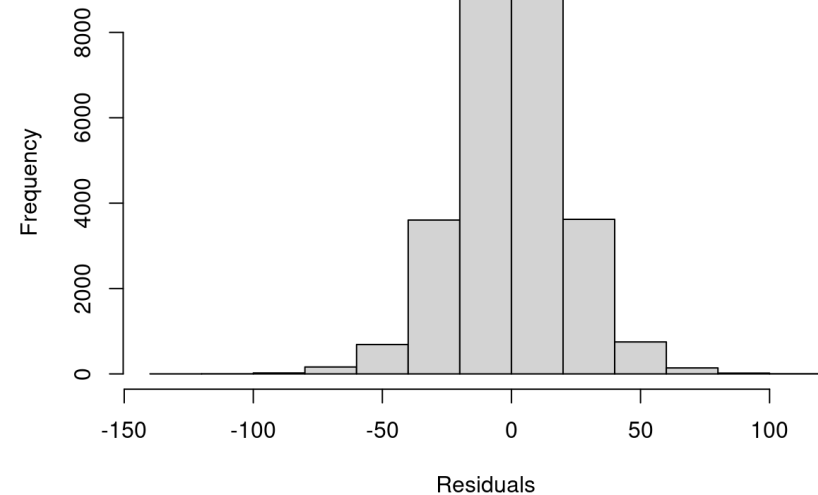Square Root(Champion Damage Dealt) vs. Total Gold

Assumptions

Total Gold and Champion Damage Dealt
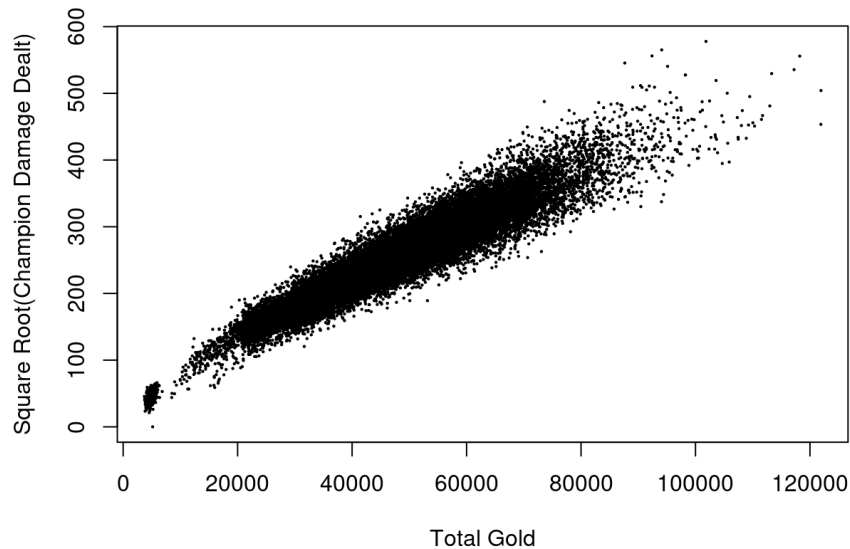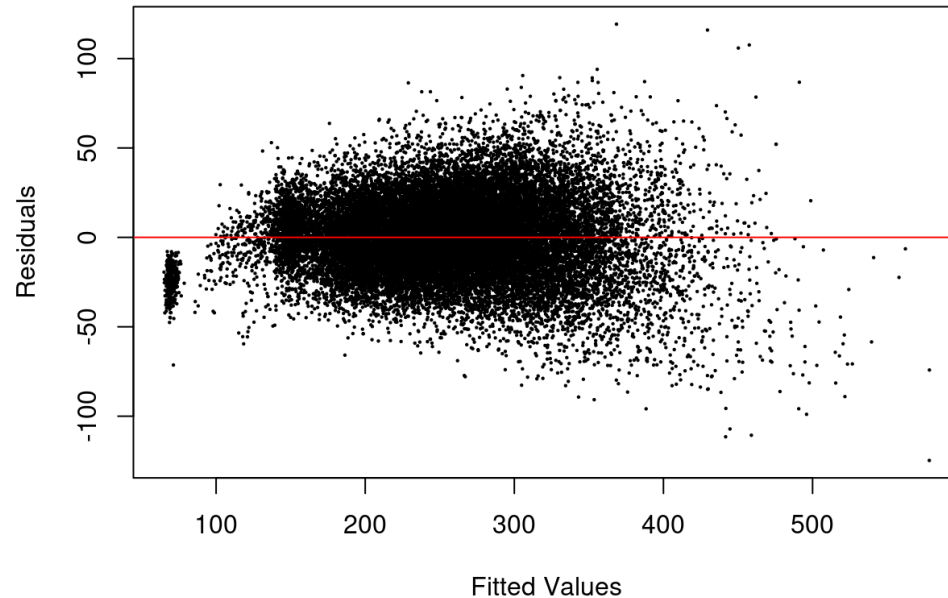

Model Residuals

There is not a violation of the linearity assumption in the scatterplot shown above, as we see a clear linear relationship between the total gold and champion damage dealt variables.

The histogram of the residuals passes the normality assumption, as seen from the normal distribution above.

Residual Plot

Though there are minor violations of the equal variance assumption (minimal cluttering and outliers), the residuals nonetheless exhibit a relatively consistent spread across the range of fitted values. Given the minor nature of these deviations, we consider the equal variance assumption to be met for the purposes of our analysis. However, we acknowledge this limitation and have conducted sensitivity analyses to assess the robustness of our results.

Discussion

Interpretation: Controlling for first Baron kill, the total gold earned is a significant predictor of total champion damage dealt to the enemy team (t=454.12, df=26901, p < 0.001). For each additional unit of *totalGold,* the *championDamageDealt* (square root transformed) increases by an average of 0.004342 units (significant estimate) dmg pts. The effect size of this would be an R-squared value of 0.8979, with our model explaining the 89.79% variation in champion damage dealt to enemy team.

Controlling for total gold earned, the first Baron kill is a significant predictor of total champion damage dealt to the enemy team (t=-12.77, df=26901, p < 0.001). Securing the first Baron kill is linked to a decrease of around 12.77 dmg pts (square root transformed) in the team's champion damage dealt to the enemy team. The effect size of this would be an R-squared value of 0.8979, with our model explaining the 89.79% variation in champion damage dealt to enemy team. This suggests that if a team secures the first Baron kill, they will tend to have lower champion damage dealt on average.

Limitations: We did not collect the data, and the dataset isn't considered a "random sample" because it is a culmination of all the challenger-ranked matches in the year 2020. This can pose the limitation of generalizability, as the conclusions and results may be specific to the conditions of in Challenger level games in 2020. Furthermore, there was a notable imbalance in the distribution of categories for our explanatory variable, *firstBaronKill,* which could have impacted the reliability and generalizability of our results. Lastly, as mentioned earlier, the equal variance assumption wasn't quite *fully* met due to very minor deviations, which might have impacted the robustness of our results and findings.

Implications: In performing statistical analysis on key factors and its effect on certain outcomes in League of Legends, we learned that in the highest level of ranked matches, teams tend to strategically shift their gameplay after securing certain objectives, which may contribute to avoiding team fights, hence less total damage dealt to champions, and that obtaining more gold is a crucial aspect of gameplay.

Future Research: Possible areas for expanding this research could be determining how relationships between game factors evolve over time within a game, specifically early, mid, and late-game as these phases significantly differ in how teams play. If I could change how I conducted this study again, I would expanded the dataset to include matches from all ranks – beyond just Challenger-level games like we did for this project – because it would allow for a more comprehensive representation of the League of Legends player base.

References:

lolmjlaver. (2021, May 11). *How to maximize your damage in League of Legends*. Dignitas.
    https://dignitas.gg/articles/how-to-maximize-your-damage

Kaggle. (2020). League of Legends Challenger Ranked Games (2020).
    https://www.kaggle.com/datasets/gyejr95/league-of-legends-challenger-ranked-games2020