Deciphering the Strategies and Formulas Behind League of Legends Pro Gameplay

Shue Him Murphy Li

Table of Contents

Abstract

In this study, we conducted an in-depth analysis of professional League of Legends matches to elucidate the factors influencing team performance and success. Our investigation revealed that the blue side initially holds an advantage, possibly due to prevailing meta trends, map dynamics, and strategic choices during the drafting phase. However, further examination using logit regression indicated that while the models encapsulate the advantages experienced by the blue side, statistical significance points towards a slight advantage for the red side. Key objectives such as barons, heralds, and void grubs, where the blue side holds geographical superiority, are likely contributors to this advantage, highlighting the significance of neutral objectives in establishing map control and accruing gold leads. Contrary to conventional wisdom, our analysis suggests that gold leads may have less influence than previously believed, primarily serving to amplify a team's damage potential through item advantages. Additionally, our study underscores the importance of team compositions tailored to late-game dominance or early-game aggression, emphasizing the pivotal role of objectives such as dragons, towers, and inhibitors in shaping game outcomes. Furthermore, we observe that kills play a crucial role in exerting pressure and securing objectives, dispelling the notion of their minimal importance, especially at the highest levels of play. Our findings also highlight the evolving dynamics of professional gameplay, with Season 13 witnessing shorter game durations and longer kill timers, necessitating a focus on controlling engagements. Machine learning and ensemble models demonstrated a modest but notable impact of draft differentials on gameplay, emphasizing the importance of meticulous consideration during champion select. However, overarching the influence of draft compositions is the undeniable importance of player and team skill, with individual prowess often outweighing strategic choices during the draft phase. This study contributes to a deeper understanding of the multifaceted dynamics underlying professional League of Legends matches, offering insights for players, coaches, and analysts alike.

Background

Understanding League of Legends

League of Legends is an online multiplayer online battle arena (MOBA) game that features two teams, each consisting of five players. The primary objective for each team is to destroy the enemy team's Nexus, a critical structure located in the heart of their base, which determines the winner of the game. Released in October 2009 by Riot Games, League of Legends was designed to succeed and expand upon the popular MOBA game of that time, Defense of the Ancients (DotA).

The game map is divided into three main lanes: top, middle (mid), and bottom (bot) lane. The top and mid lanes are typically solo lanes where one-on-one battles (1v1s) occur between players. The bottom lane is a duo lane generally consisting of an Attack Damage Carry (ADC) and a Support, who work together in two-on-two (2v2) engagements.

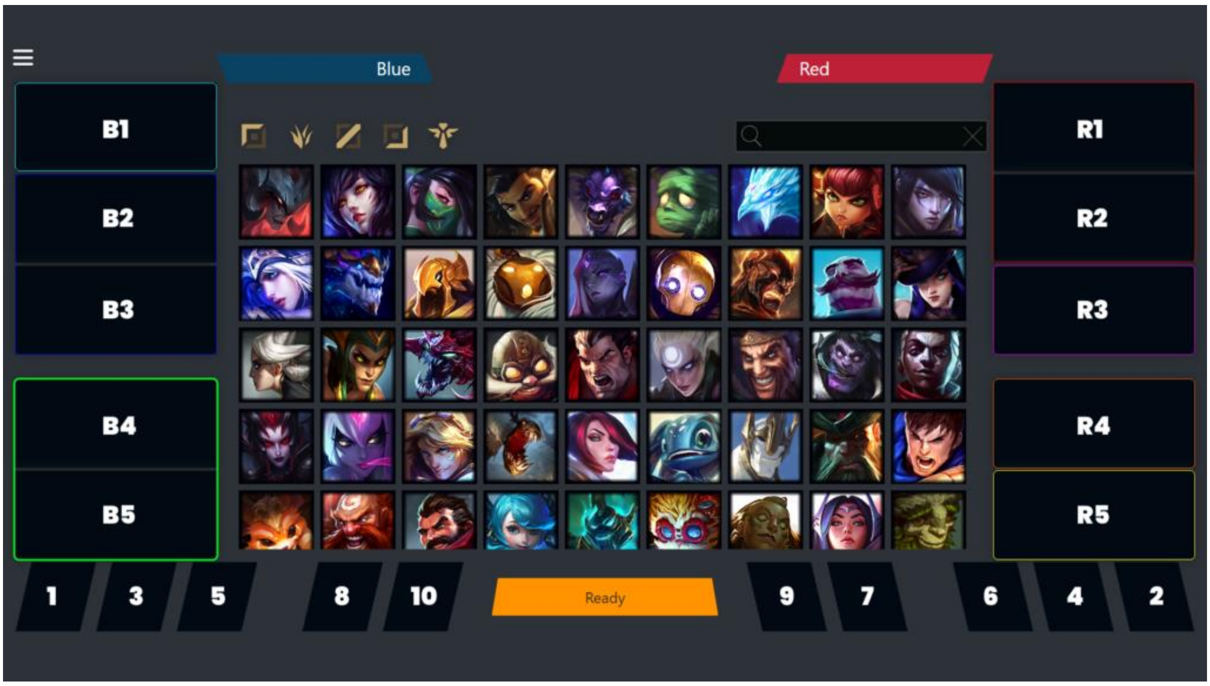
Between the lanes lies the jungle, an area filled with neutral monsters that the jungler defeats for gold and experience. The jungler's role involves clearing these jungle camps, providing assistance to laners through ganks, and securing important neutral objectives such as the Rift Herald, Dragon, and Baron Nashor using their unique summoner spell, Smite.

In the diagram above, Base 1 is referred to as the Blue Side, while Base 2 is known as the Red Side.

Each lane contains two defensive towers outside of their base: the outer tower, known as Tier 1, and the inner tower, known as Tier 2. Inside the base, each team has three inhibitors, each protected by an inhibitor tower. Destroying an inhibitor causes the attacking side to spawn super minions, which are stronger than regular minions. To open up the base and attack the Nexus, at least one inhibitor tower and its corresponding inhibitor must be destroyed.

The Nexus, the ultimate target, is also defended by two towers. These Nexus towers must be destroyed before the Nexus itself can be attacked and damaged. In the diagram above, Base 1 is referred to as the Blue Side, while Base 2 is known as the Red Side.

The above diagram shows the locations where major neutral objectives spawn. Rift Herald, Baron Nashor, and Void Grubs spawn on the top side of the map, while Dragons spawn on the bottom side of the map. Due to the unique design and geography of the map, the Red Side tends to have easier access to the bottom side of the map, whereas the Blue Side has easier access to the top side of the map.



Before each game begins, there is a draft phase where each team selects their five champions in a specific order, also banning champions to prevent the other team from picking them. The draft phase proceeds as follows:

1. Bans: 1 to 6
2. Picks: Blue Side (B1), Red Side (R1 and R2), Blue Side (B2 and B3), Red Side (R3)
3. Bans: 7 to 10
4. Picks: Red Side (R4), Blue Side (B4 and B5), Red Side (R5)

The first pick by the Blue Side (B1) is generally known as the "first blind pick," where teams select champions that are versatile and have few counters. This pick can also secure strong champions in the current meta. Conversely, the final pick by the Red Side (R5) is often a "counter pick," chosen with full knowledge of all previously selected and banned champions, allowing it to effectively counter one of the opposing team's champions.

Research Objectives and Questions

My research aims to uncover the strategies used in professional League of Legends play to secure victories. Specifically, I will investigate the following research questions:

1. How accurately can the outcome of games be predicted using early game data compared to late game data?
2. How does the presence or absence of key in-game objectives (e.g., Barons, Dragons) impact the final result of a match?
3. Does the composition of champions selected during the draft phase significantly influence the likelihood of winning a game?

To address these questions, I will explore various aspects of gameplay and strategy. For instance, I will analyze what teams should focus on regarding in-game objectives and possible play styles, such as prioritizing farming minions versus securing kills. Additionally, I will rank the importance of different in-game objectives.

Beyond in-game objectives, my research will delve into the roles of coaches and analysts, particularly in the drafting phase. I aim to determine whether the draft has a significant impact on the outcome of the game. I also intend to develop models that can effectively predict the game's outcome based on information obtained from the draft phase, including player and team information, as well as picks and bans by the teams.

Through this research, I hope to provide insights into optimal strategies and decision-making processes in professional League of Legends play, potentially offering valuable guidelines for teams and analysts in the competitive scene.

Introduction

League of Legends, a popular online multiplayer online battle arena (MOBA) game, has become a cornerstone of the global eSports industry. Its strategic depth and dynamic gameplay offer a rich subject for analysis, attracting millions of players and viewers worldwide. Understanding the factors that contribute to winning in professional play is crucial for players, coaches, analysts, and the broader eSports community. This study aims to shed light on these factors, providing valuable insights into optimal strategies and decision-making processes.

The primary goals of this research are to analyze the impact of early game versus late game data on predicting match outcomes, evaluate the significance of key in-game objectives such as Barons and Dragons on match results, and assess the influence of champion composition and the draft phase on the likelihood of winning. Specifically, this study seeks to answer the following research questions: How accurately can the outcome of games be predicted using early game data compared to late game data? How does the presence or absence of key in-game objectives impact the final result of a match? Does the composition of champions selected during the draft phase significantly influence the likelihood of winning?

In the League of Legends community, numerous unspoken rules and beliefs exist about how the game should be played. For instance, prominent coaches and figures advocate for prioritizing late game scaling compositions to secure a late game advantage, and emphasize that kills are less important than securing objectives. While some of these beliefs are more widely accepted than others, this study aims to use data to verify which statements hold true.

Coaches and analysts, relatively new roles in the League of Legends scene, provide invaluable in-game advice outside of matches. They assist players with better drafts and macro gameplay decisions. This study will explore the impact of their strategies and guidance on game outcomes. Utilizing data from Oracle's Elixir, a widely trusted source for professional League of Legends gameplay data, this research will investigate various factors determining match results. This includes the impact of early game objectives and neutral objectives. Due to limited data on champion composition, machine learning methods will be employed to assess the impact of drafts.

By analyzing how games were played across different seasons and regions, this study will examine whether play styles influence game outcomes. Comparing early game and late game variables will help identify key objectives teams should prioritize based on their composition, allowing them to fine-tune their game plans accordingly. The paper will be structured as follows: Background, detailing the game mechanics, draft phase, and in-game objectives; Methodology, explaining data collection and analytical methods; Results, presenting findings related to the research questions; Discussion, interpreting the results and their implications; and Conclusion, summarizing key insights and potential areas for future research.

This introduction sets a clear context for the study, outlines its significance, and provides a roadmap for the rest of the paper.

Literature Review

Given the novelty of my research, there have been limited in-depth studies specifically focusing on my topic. Historically, researchers have utilized similar datasets, if not the same ones, to explore basic metrics and data within this field. For instance, a study conducted by Morales Garcia et al. (2023) discovered that the team exhibiting more strategic movement across the map tends to secure victory. This finding highlights the significance of map control in determining game outcomes.

In contrast, my research delves into both pre-game and in-game data to assess the importance of drafting strategies as well as in-game objectives. A study that parallels mine is "Meta Search: A Tool for Mass Analysis of Game Strategy" by Janzen et al. This study analyzes data from Dota 2, a multiplayer online battle arena (MOBA) game similar to League of Legends. Janzen et al. developed analytical tools, including visualization techniques and correlation matrices of in-game metrics, which prove invaluable for analyzing games like League of Legends.

A noteworthy aspect shared by both Dota 2 and League of Legends is the concept of counter picks, where one champion can be countered by another, thus offering potential strategic advantages throughout the game due to the inherent strengths of the counter champion. While Janzen et al. did not specifically focus on the draft pick phase, their study provided insights and methodologies that I can apply in my research. This includes incorporating matchup analyses to simulate potential counters that are evident in professional gameplay.

Data

This study utilizes an extensive dataset covering professional League of Legends (LoL) games from Season 3 (2014) through to Season 13 (2023). The data includes all matches from Tier 1 and Tier 2 competitions across more than 10 different leagues (regions). The primary Tier 1 leagues comprise the LCK (Korea), LPL (China), LCS (North America), and LEC (Europe), among others representing various global regions.

Data Source

The data for this analysis was sourced from Oracle's Elixir, a comprehensive League of Legends esports database. Oracle's Elixir compiles match-level, player-level data and team-level data, providing a rich source of competitive gaming statistics that are widely recognized and used within the LoL esports analytics community. The reliability and accuracy of this database have been validated not only by its extensive use in industry analyses but also by its frequent citation by prominent esports commentary and news reporting sites, including TheScore Esports, ESPN Esports, and Dot Esports, which are key players in analyzing and reporting on the League of Legends scene.

Data Scope

The dataset encompasses a variety of metrics crucial for understanding the dynamics and outcomes of professional LoL matches. These metrics include, but are not limited to, game duration, player and team statistics (e.g., kills, deaths, assists, gold earned, and objectives taken like dragons, barons, and turrets). The data captures the evolving strategies and gameplay mechanics over a significant period in LoL's history, offering insights into trends and shifts in competitive play styles.

Data Collection

Data collection involved querying the Oracle's Elixir database using specific criteria to retrieve only professional-level matches from Tier 1 and Tier 2 leagues. Each record in the dataset corresponds to a single match and includes aggregated team statistics and individual player performance metrics. Data integrity checks were conducted to ensure completeness and accuracy, with missing or inconsistent records being noted and addressed through further validation steps.

Data Processing

The raw data underwent several preprocessing steps to make it suitable for analysis. These steps included data cleaning (e.g., handling of missing values), and transformation of numerical data into categorical (factor) formats suitable for statistical modelling. The final dataset was structured into a panel data format, facilitating both cross-sectional and time-series analyses to examine trends over the ten-year period.

This comprehensive data preparation enables a detailed examination of the strategic evolution in professional League of Legends, aiding in the robust statistical analysis that follows in the subsequent sections of this paper.

Variables

Draft Variables

Variable Name	Description
Champions	10 variables indicating indicated the name of the champion by side and role
Role specific matchups	5 variables indicating the role specific matchups for all 5 roles
Lane jungle dynamic	6 variables indicating the lane jungle dynamics
Champion picks	10 variables indicating champion picks by both sides,along with their pick order
Champion bans	10 variables indicating champion bans by both sides, along with their ban order

In-game Variables

Variable Name	Description
Barons	Number of barons captures by the team
Dragons	Number of dragons captures by the team
Firstbaron	Binary variable indicating whether the team captured first baron
FirstHerald	Binary variable indicating whether the team captured first herald
Firstmidtower	Binary variable indicating whether the team was first to destroy mid tower
Firsttothreetowers	Binary variable indicating whether the team was the first to destroy three towers
Gamelengthinm	Game length in minutes
Golddiffat15	Team gold difference at 15 minutes
Goldspentdiff	Difference in team gold spent at the end of the game
Heralds	Number of heralds captured
Inhibitors	Number of inhibitors destroyed by the team
Percent_total_kpm	Team kills per minute divided by total champion kills per minute (kill share)
Towers	Number of towers destroyed by the team
Void_grubs	Number of void grubs captured
Wpm	Wards placed per minute by the team
Result	Binary variable indicating whether the team won the game

Team/ Player variables

Variable Name	Description
Player name	10 variables indicating the in-game name of the player by side and position
Team name	2 variables indicating the team names of the team playing on blue and red side

Method

Exploratory Data Analysis (EDA)

Prior to in-depth analysis, an exploratory data analysis (EDA) was conducted to identify simple patterns and initial insights in the data. This phase focused on exploring potential advantages associated with playing on the blue side versus the red side, as well as analyzing which side more frequently secured key early-game objectives such as the first dragon or first turret. The EDA aimed to provide a foundational understanding of the dataset, highlighting potential predictors that influence game outcomes.

Logistic Regression Analysis

Following the EDA, logistic regression was employed to analyze the relationship between in-game variables (objective secures, kills, and gold) and match outcomes (win/loss). A series of logistic regression models were developed, and the optimal model was selected based on criteria such as the McFadden pseudo R^2 and the Akaike Information Criterion (AIC). This selected model was further applied to different subsets of the data, filtered by season and by league (LCK, LPL, LCS, LEC), to observe variations in the influence of these variables across different competitive environments.

Random Forest Models

To capture more complex patterns and interactions between variables, Random Forest models were utilized. Two configurations were implemented:

A Random Forest model incorporating both draft/player variables and in-game performance metrics.

A Random Forest model focusing solely on in-game performance metrics.

These models allowed for an examination of the relative importance of pre-game strategies (such as player drafts) versus actual in-game performance in determining match outcomes.

XGBoost Models

Given the complexity and size of the dataset, XGBoost models were developed for their efficiency in computational speed and performance improvement. The following configurations were tested:

An XGBoost model including both draft and player variables.

An XGBoost model exclusively using draft variables.

XGBoost was chosen for its ability to handle large-scale data and its effectiveness in improving model performance through advanced regularization techniques, which prevent overfitting. These models aimed to refine the predictions by incorporating the nuanced interactions among the predictors.

Specialized XGBoost Analysis

To specifically address potential shifts in gameplay strategy over time, an additional XGBoost model was applied to a subset of the data from Season 12. This analysis was intended to determine whether changes across the seasons affected the predictive performance of the models previously developed. This step was crucial for understanding the temporal dynamics in League of Legends competitive play and for adjusting the model parameters to enhance predictive accuracy based on recent data.

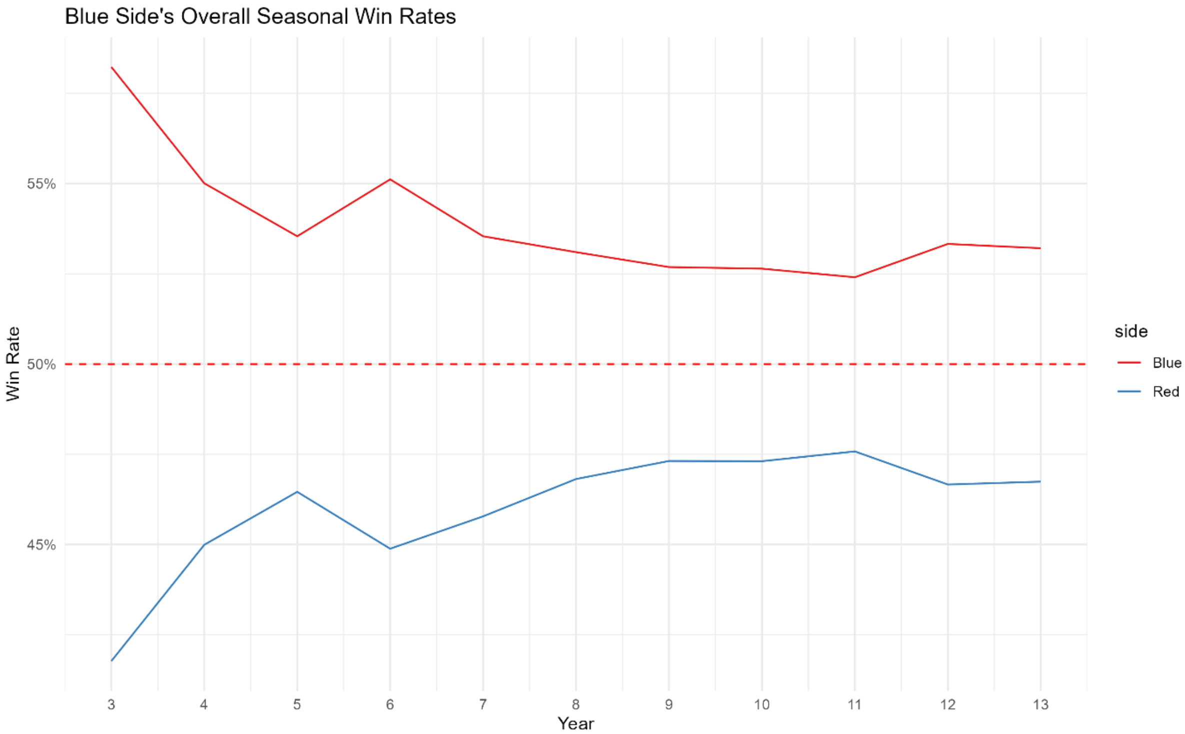
Statistical Software

All statistical analyses and model development were performed using R. All statistical analyses and model development were performed using R. The following packages were extensively utilized throughout the analysis: tidyverse for data manipulation and visualization, hms and lubridate for handling time-based data, GGally for extended visualizations, pscl and stargazer for statistical modeling and reporting, caret and car for advanced machine learning techniques, corrplot for visualizing correlations, randomForest for building Random Forest models, recipes for data preprocessing, hash for efficient data handling, RColorBrewer for enhanced color palettes in graphics, xgboost for executing XGBoost models, and pROC for evaluating model performance metrics. These tools facilitated a comprehensive approach to data analysis, model building, and evaluation, ensuring robust and reproducible results.

Exploratory Data Analysis

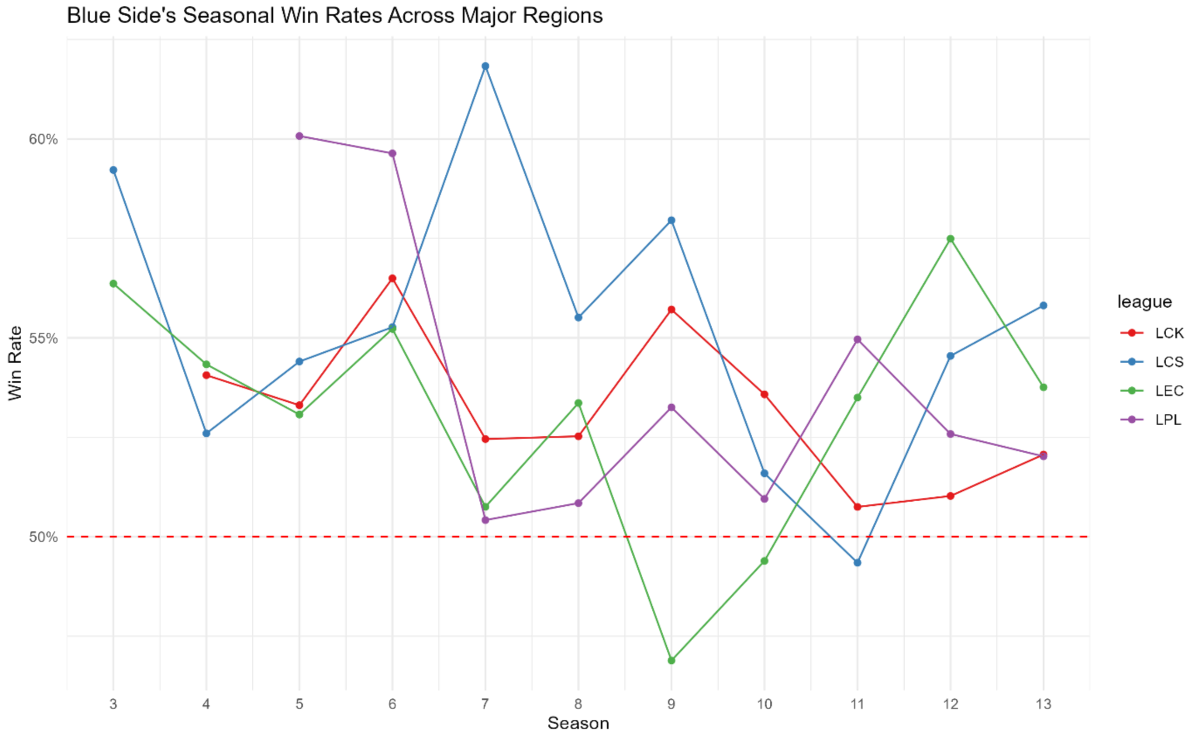
Conditional Probabilities

Blue Side Win Rate Over Time



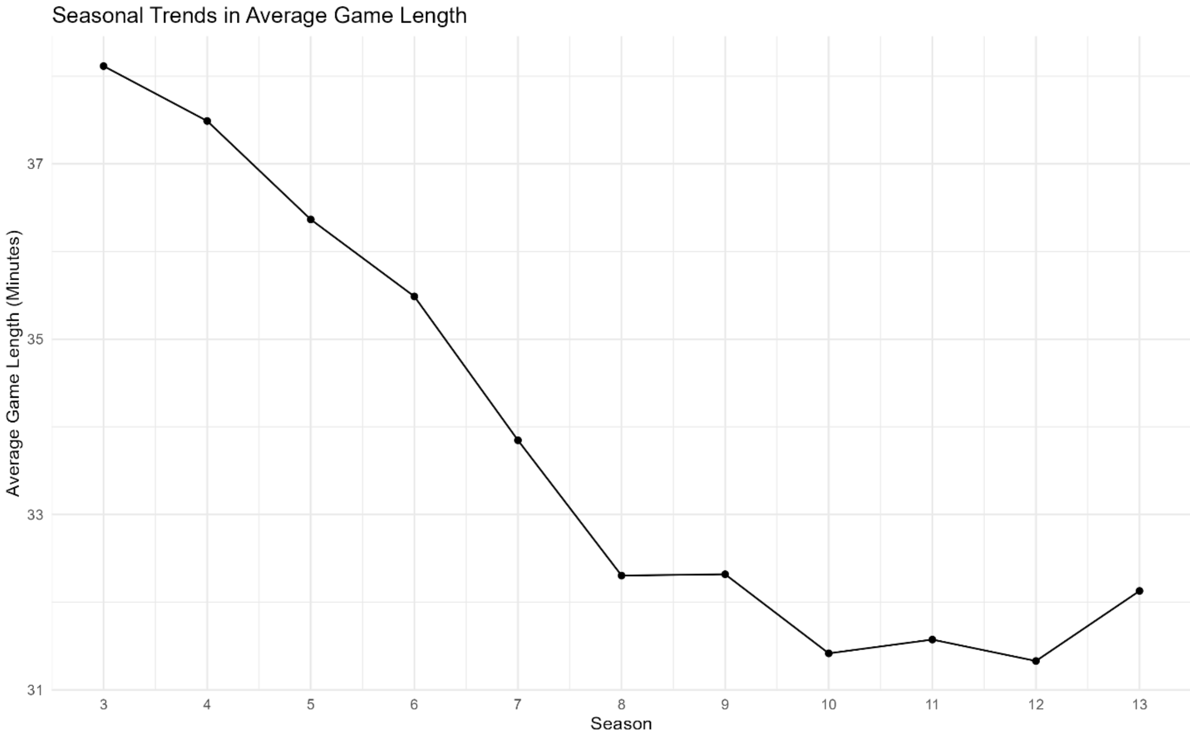
The graph above indicates that the blue side has a slight advantage in winning games, as evidenced by their consistently higher win rate compared to the red side in professional play from seasons 3 to 13.

Blue Side Win Rate by League



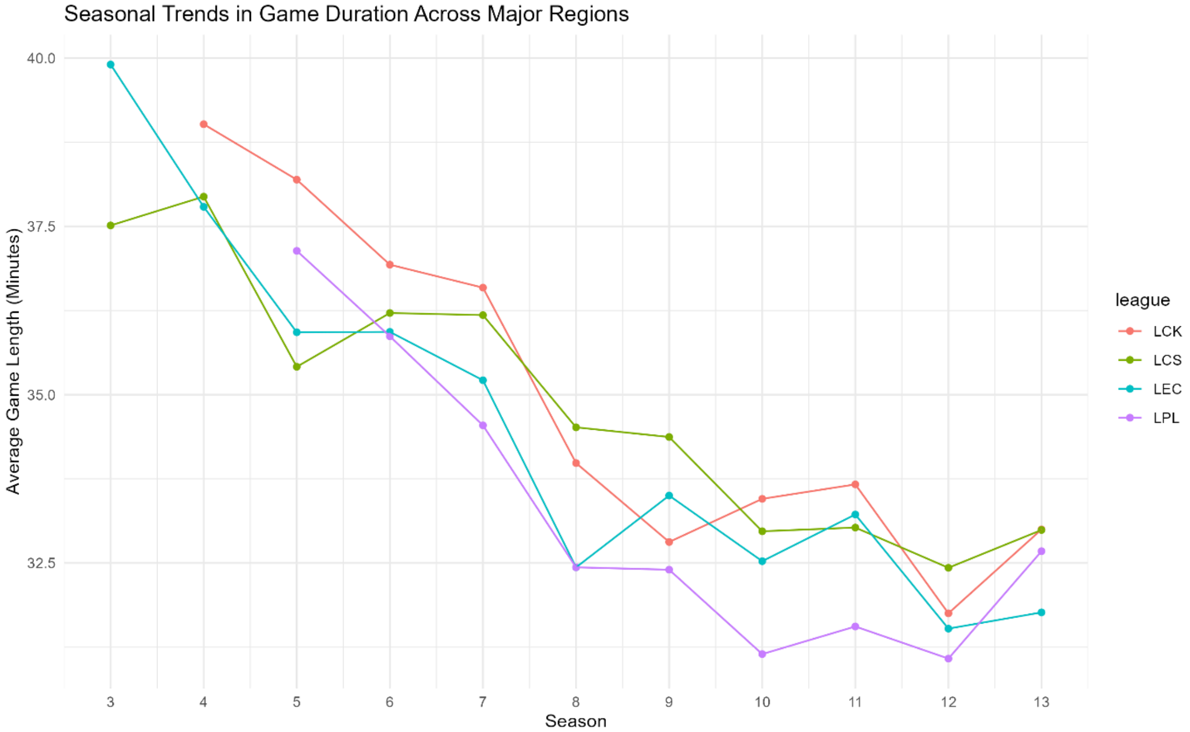
The graph above illustrates the win rates for the blue side across four major regions: LCK, LCS, LEC, and LPL. These regions represent the highest level of gameplay in professional League of Legends. In recent years, all regions have demonstrated a blue side win rate exceeding 50%, with the exception of the LEC, where the red side had a higher win rate during seasons 9 and 10. This trend suggests that the blue side may possess an inherent advantage, as teams at the highest levels of play can leverage its strengths to achieve more victories compared to the red side.

Seasonal Trends in Average Game Length



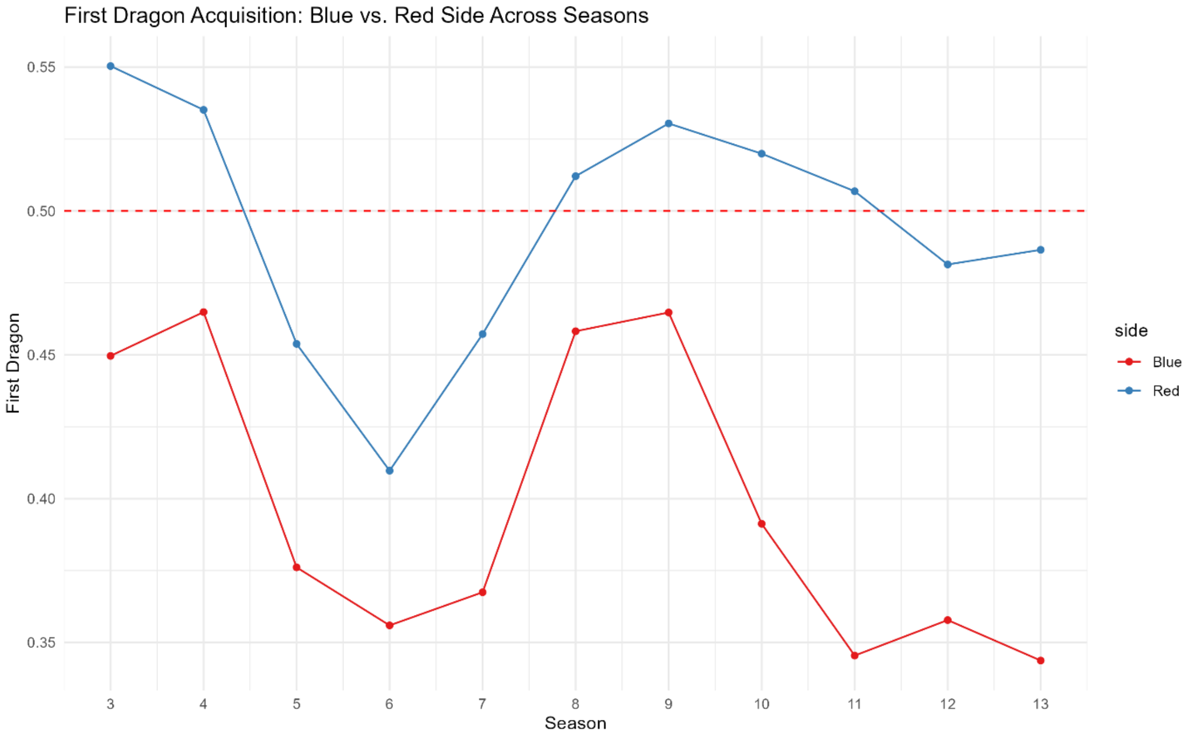
The summary statistics table indicates that the average game length from seasons 3 to 13 is approximately 32 minutes. Over this period, we observe a steady decrease in average game length from seasons 3 to 12. This trend could suggest that players have become more skilled over time, leading to quicker resolutions of matches. Alternatively, it could reflect changes in the game itself that have contributed to shorter game durations.

Seasonal Trends in Game Length by League



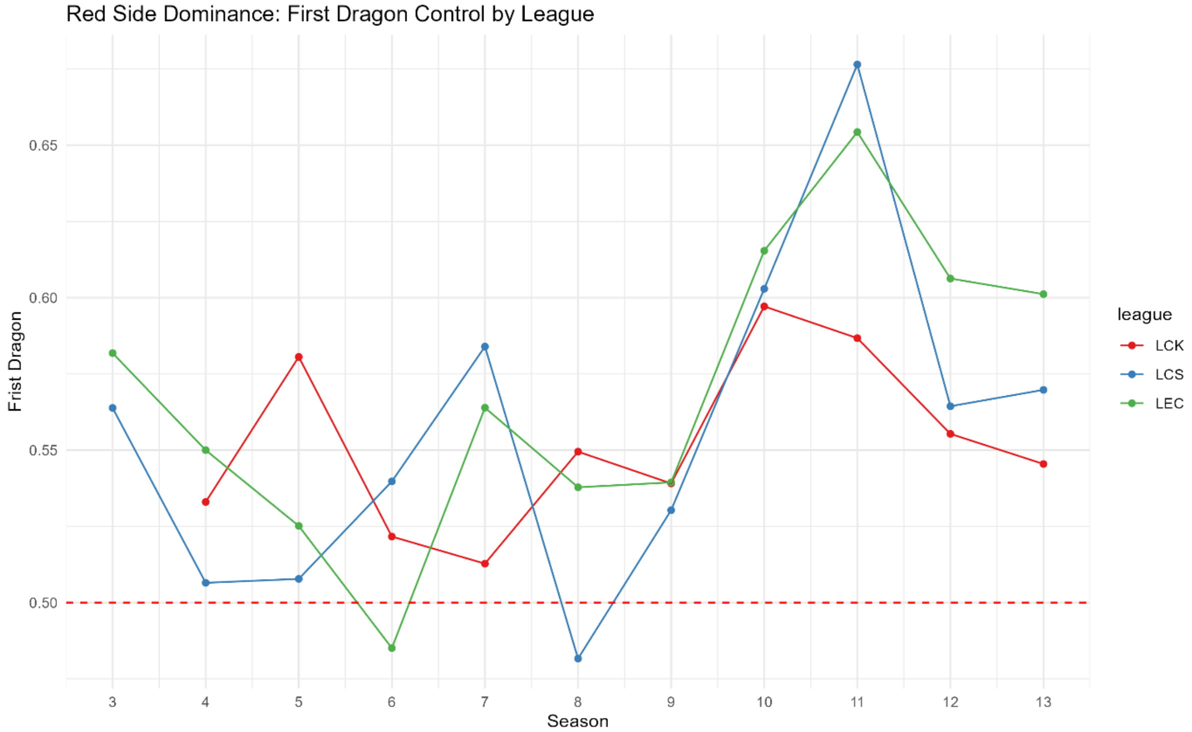
The graph above shows that the LPL has had shorter average game lengths compared to other regions, while the LCS and LCK have consistently experienced the longest average game durations over the years. This disparity may suggest a difference in play style, with the LPL favoring strategies that enable them to close out games more quickly. On the other hand, considering the general consensus that the LCS is the least strong region among the four, this could imply that LCS teams lack the knowledge or execution skills to finish games efficiently, unlike their counterparts in other regions. Meanwhile, the longer game lengths in the LCK could indicate a preference for late-game compositions, which inherently extend the duration of matches.

First Dragon Acquisition: Blue vs Red Side



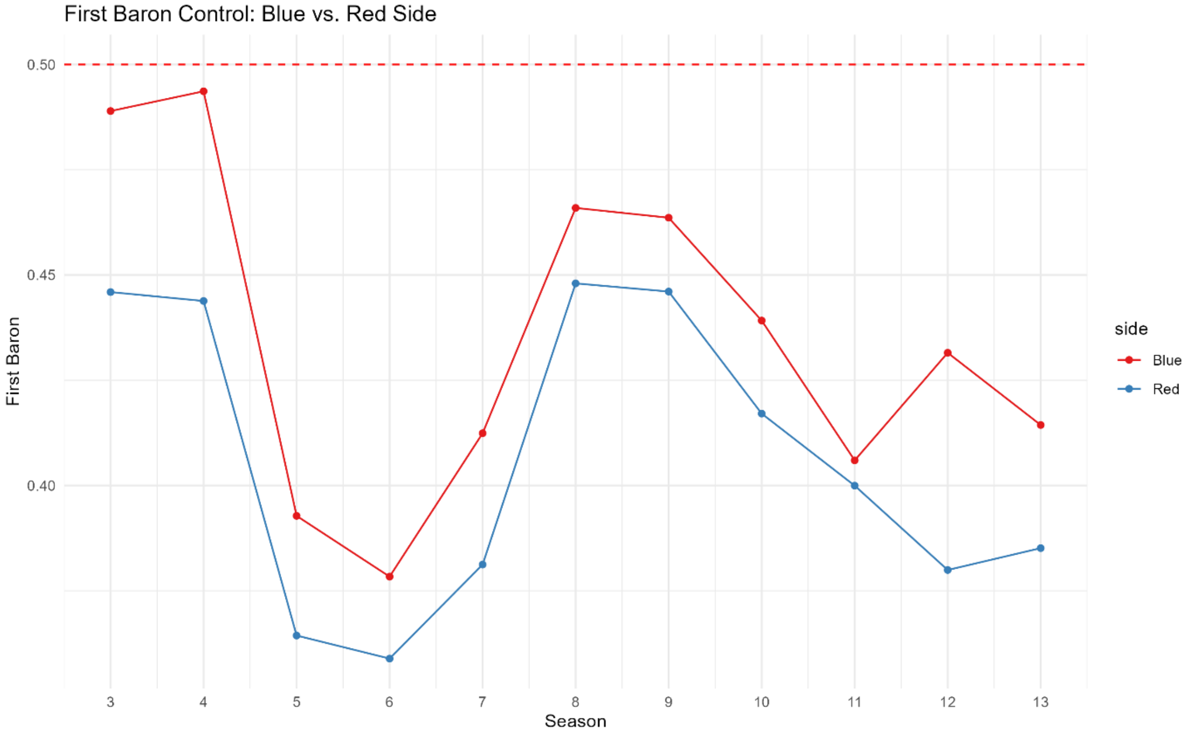
The graph above shows that the red side consistently secures the first dragon throughout the seasons. The only constant variables over these years are the geography of the map and the drafting process. Firstly, this trend could indicate that the map's layout inherently provides the red side with easier access to the bottom half of the map, enabling them to take control and secure early key objectives more effectively. Secondly, it could suggest that the red side strategically leverages their counter picks during the draft phase to select strong early-game bottom lane compositions and jungle champions, thereby gaining early advantages that facilitate securing the first dragon.

Red Side Dominance: First Dragon Control by League



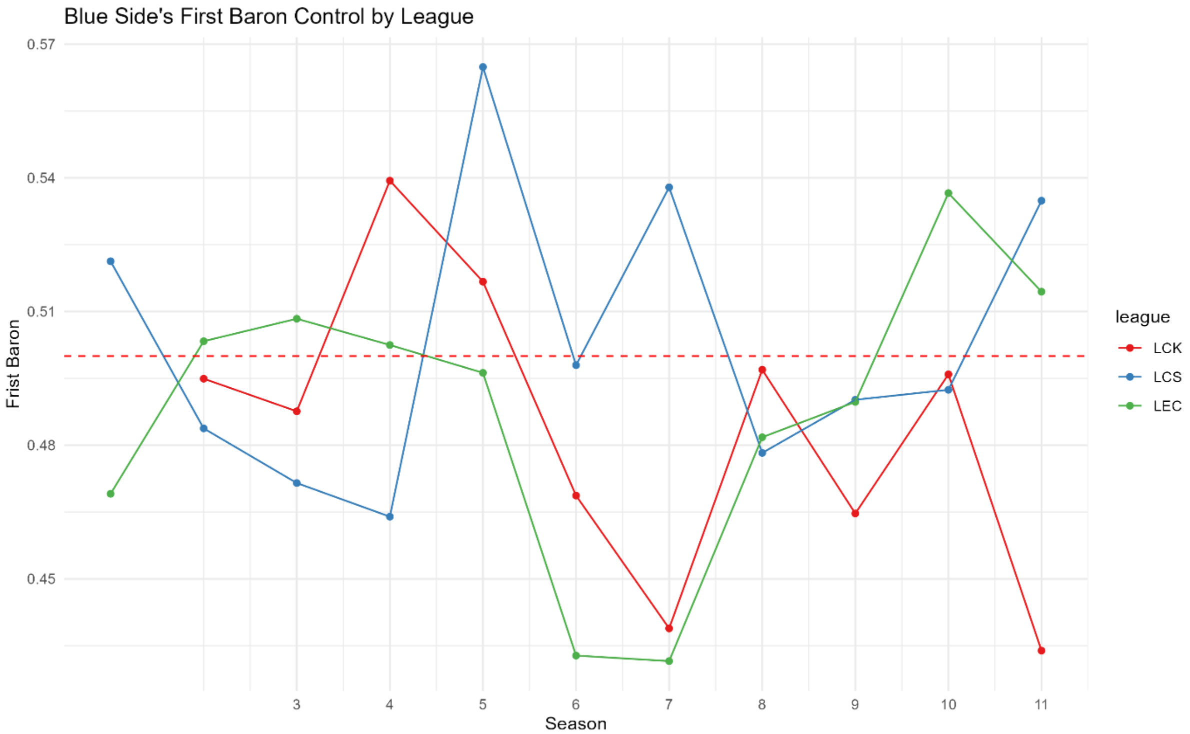
In all three major regions, the red side has demonstrated dominance in securing the first dragon, particularly after season 9, where this advantage became more pronounced. This trend could be attributed to the map's geography and the strategic drafting mentioned earlier, as well as the introduction of the Rift Herald objective by the game developers. The Rift Herald spawns on the top side of the map, where the blue side has a geographical advantage. Consequently, teams often trade objectives to minimize opportunity costs, with the blue side prioritizing the Rift Herald while the red side focuses on securing the first dragon.

First Baron Control: Blue vs Red Side



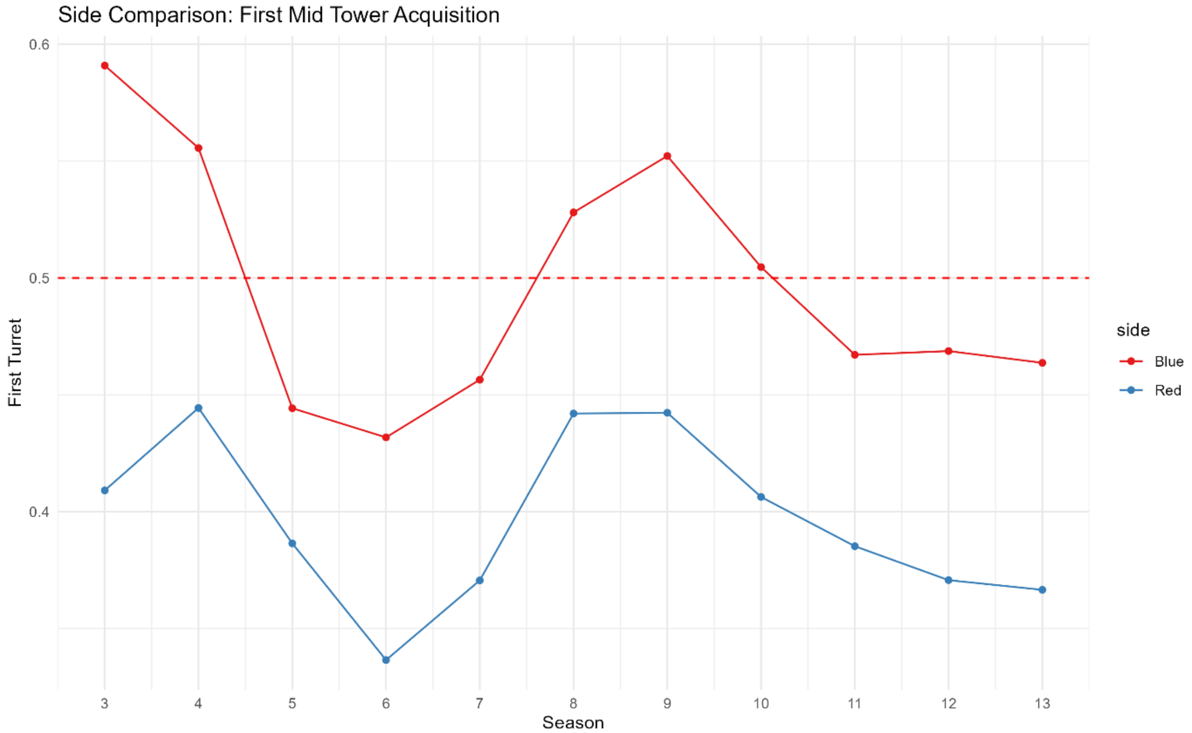
Over the seasons, the blue side has consistently been the first to secure Baron Nashor before the red side. This trend can be attributed to the geographical advantage, as Baron spawns on the top side of the map, which is more accessible to the blue side.

Blue Side First Baron Control by League



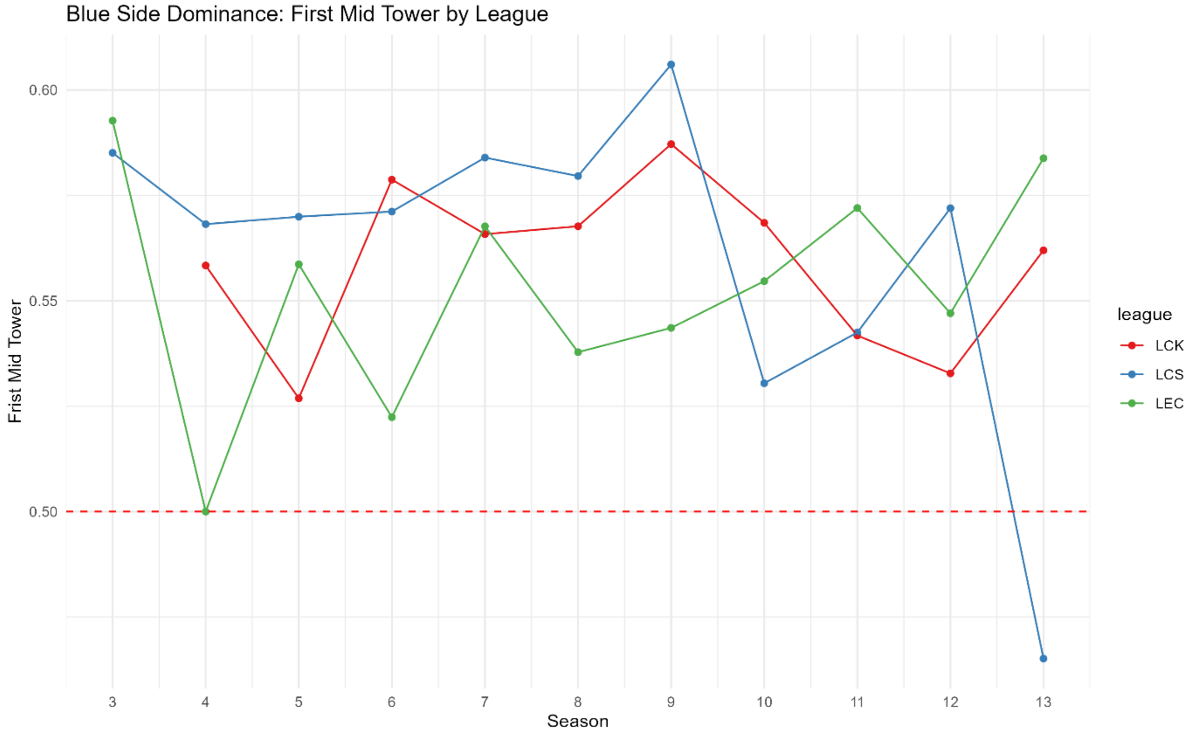
Over the seasons, the blue side has consistently been the first to secure Baron Nashor before the red side, likely due to the geographical advantage, as Baron spawns on the top side of the map, which is more accessible to the blue side. However, at the highest level of play, this trend appears to be relatively balanced, with no significant imbalances observed in either side's ability to secure the first Baron.

Side Comparison: First Mid Tower Acquisition



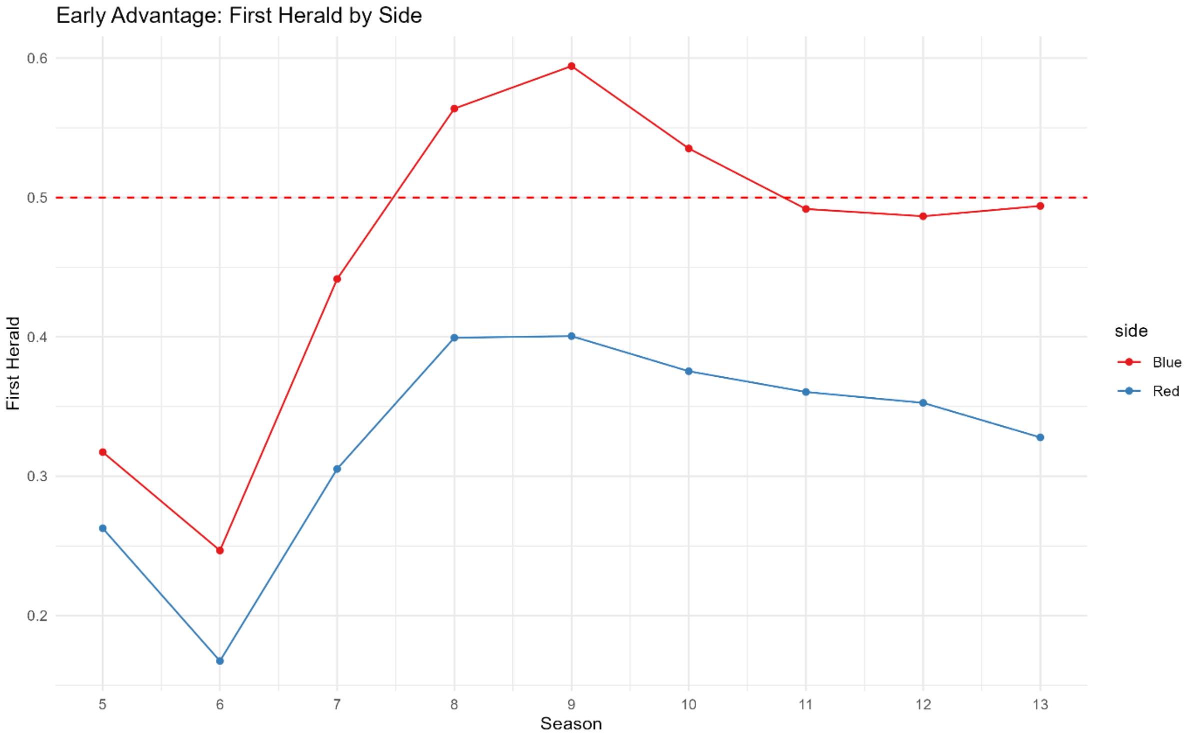
Blue side has consistently secured the first mid tower over red side, which could be attributed to the strategic advantage of securing safe blind picks for their mid player during the draft phase. However, there is limited evidence to conclusively support this claim, making it unclear exactly how the blue side maintains this advantage in securing the first mid tower.

Blue Side Dominance: First Mid Tower by League



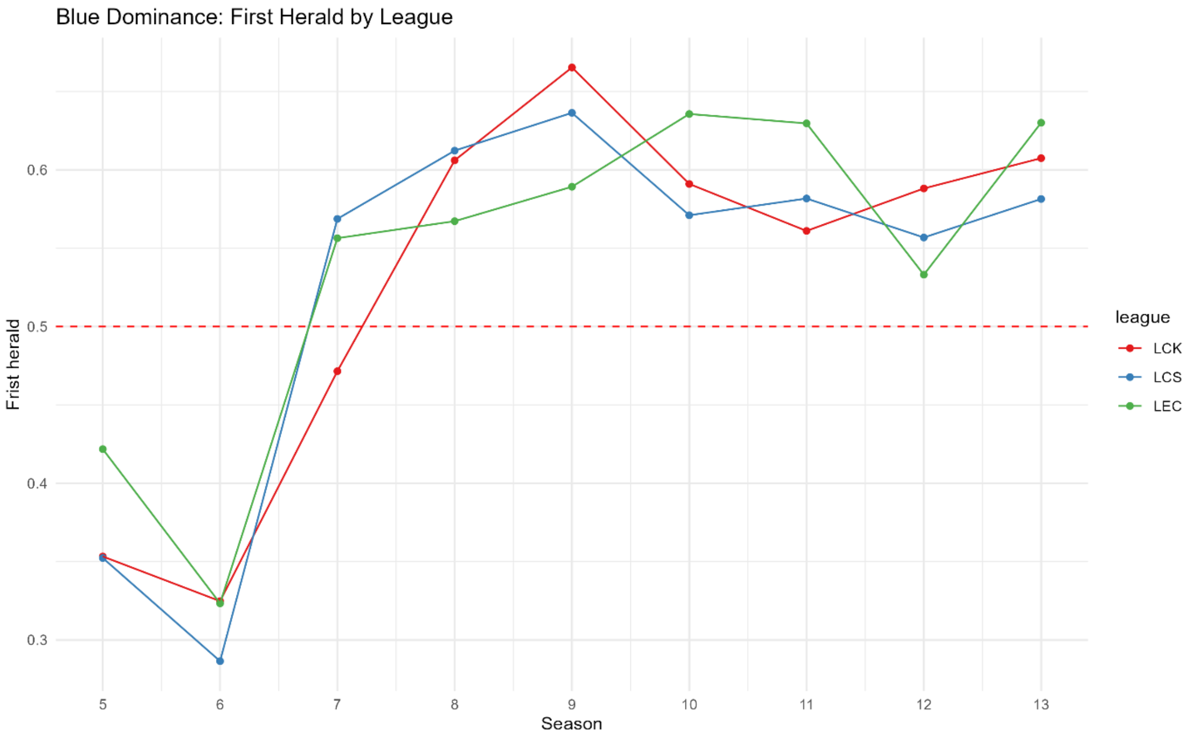
In the three major regions, the LCS has strategically focused on controlling the mid lane when playing from the blue side, from seasons 3 to 9. They achieve this by adopting draft strategies that facilitate a strong push and command over the mid lane, consistently allowing them to secure the first mid tower. Additionally, most other regions have also demonstrated a clear advantage for the blue side in securing the first mid tower. This trend underscores a widespread strategic pattern across the professional League of Legends landscape.

Early Advantage: First Herald by Side



Due to geographic advantages inherent to the map layout, the blue side has consistently secured the first Rift Herald more frequently than the red side over the years.

Blue Dominance: First Herald by League



This phenomenon is also observable in professional play, where the blue side often prioritizes the first Rift Herald over the first dragon. Since season 7, they have consistently taken advantage of their geographical positioning to secure this objective more frequently than the red side.

Correlation Matrix

	result	percent_total_kpm	dragons	barons	wpm	firstmidtower	firsttothreetowers	towers
result	1	0.858	0.528	0.621	0.076	0.416	0.502	
percent_total_kpm	0.858	1	0.539	0.561	0.087	0.461	0.536	0.843
dragons	0.528	0.539	1	0.452	0.114	0.305	0.348	0.565
barons	0.621	0.561	0.452	1	0.141	0.267	0.327	0.698
wpm	0.076	0.087	0.114	0.141	1	0.029	0.037	0.125
firstmidtower	0.416	0.461	0.305	0.267	0.029	1	0.709	0.477
firsttothreetowers	0.502	0.536	0.348	0.327	0.037	0.709	1	0.579
towers	0.884	0.843	0.565	0.698	0.125	0.477	0.579	1
inhibitors	0.707	0.632	0.438	0.621	0.072	0.399	0.476	0.794
golddiffat15	0.512	0.613	0.377	0.307	0.064	0.538	0.609	0.591
firstbaron	0.595	0.585	0.363	0.693	0.059	0.436	0.510	0.632
firstdragon	0.197	0.209	0.395	0.123	0.008	0.252	0.271	0.203
firstherald	0.168	0.189	0.164	0.100	-0.005	0.387	0.375	0.211
heralds	0.213	0.245	0.224	0.127	-0.060	0.458	0.437	0.260
void_grubs	0.010	0.009	0.022	0.014	0.034	0.042	0.047	0.020
goldspentdiff	0.746	0.814	0.473	0.590	0.082	0.477	0.576	0.812

	inhibitors	golddiffat15	firstbaron	firstdragon	firstherald	heralds	void_grubs	goldspentdiff
result	0.707	0.512	0.595	0.197	0.168	0.213	0.010	0.746
percent_total_kpm	0.632	0.613	0.585	0.209	0.189	0.245	0.009	0.814
dragons	0.438	0.377	0.363	0.395	0.164	0.224	0.022	0.473
barons	0.621	0.307	0.693	0.123	0.100	0.127	0.014	0.590
wpm	0.072	0.064	0.059	0.008	-0.005	-0.060	0.034	0.082
firstmidtower	0.399	0.538	0.436	0.252	0.387	0.458	0.042	0.477
firsttothreetowers	0.476	0.609	0.510	0.271	0.375	0.437	0.047	0.576
towers	0.794	0.591	0.632	0.203	0.211	0.260	0.020	0.812
inhibitors	1	0.435	0.562	0.193	0.182	0.215	0.015	0.642
golddiffat15	0.435	1	0.391	0.201	0.348	0.391	0.029	0.678
firstbaron	0.562	0.391	1	0.253	0.224	0.263	0.028	0.625
firstdragon	0.193	0.201	0.253	1	0.094	0.124	-0.016	0.199
firstherald	0.182	0.348	0.224	0.094	1	0.786	0.050	0.234
heralds	0.215	0.391	0.263	0.124	0.786	1	-0.015	0.276
void_grubs	0.015	0.029	0.028	-0.016	0.050	-0.015	1	0.016
goldspentdiff	0.642	0.678	0.625	0.199	0.234	0.276	0.016	1

The correlation matrix reveals that the primary drivers for victories in the game include kill share (which is the percentage of total kills per minute), barons, towers, inhibitors, and the difference in total gold spent. However, upon examining the specific values, it becomes clear that no variables show significant correlation except for gold difference at 15 minutes and experience difference at 15 minutes. These two variables, which both indicate early game advantages, could potentially induce multicollinearity in subsequent models. This suggests that careful consideration must be given to how these variables are included and analyzed to avoid skewing the results due to overlapping influences.

Variance Inflation Factor (VIF)

Variable	VIF Value	Variable	VIF Value
Side	1.070	Inhibitors	1.837
Percent Total KPM	1.358	Gold Diff at 15	1.715
Dragons	1.188	First Baron	2.365
Barons	1.746	First Dragon	2.304
WPM	1.040	First Herald	2.640
First Mid Tower	1.725	Heralds	2.825
First to Three Towers	1.978	Void Grubs	1.022
Towers	2.207	Gold Spent Diff	1.588
First Blood	1.120	First Baron:First Dragon	3.145

Upon examining the Variance Inflation Factor (VIF) of the variables, it is evident that due to the relatively low VIF values, no multicollinearity is observed among the selected set of variables. This indicates a low likelihood of interdependencies that could skew analytical outcomes in the model.

Summary of EDA

In summary, the analysis reveals that the blue side holds a slight advantage in winning games, possibly attributed to factors such as draft advantages and map geography. The blue side's geographic positioning allows for greater control over objectives like the first Rift Herald and first Baron, facilitating game closure. Additionally, blue side teams tend to secure the first mid tower, granting them map control. Conversely, the red side capitalizes on its own advantages, particularly in securing objectives on the bottom side of the map, such as the first dragon.

Building on the patterns observed in our exploratory data analysis, we will now transition to modeling to better understand the impact of various factors on game outcomes. By incorporating variables such as neutral objectives, we aim to demonstrate that the perceived advantage of being on the blue side diminishes when these key factors are taken into account. This approach will allow us to study the game evenly across both sides, highlighting the importance and significance of each objective regardless of the starting side.

Results

Logistic Regression Models

Logit Models 1,2,3

	Dependent variable:		
	result		
	(1)	(2)	(3)
sideRed	0.111**	0.131***	−0.062*
	(0.049)	(0.048)	(0.032)
percent_total_kpm	23.633***	23.348***	28.920***
	(0.333)	(0.328)	(0.236)
dragons	1.310***	1.386***	
	(0.073)	(0.072)	
barons	1.548***	1.199***	
	(0.142)	(0.132)	
wpm	0.014	0.020	
	(0.034)	(0.034)	
firstmidtower	−0.007		−0.158***
	(0.062)		(0.043)
firsttothreetowers	−0.212***		0.565***
	(0.066)		(0.045)
towers	1.364***	1.351***	
	(0.021)	(0.021)	
inhibitors	−0.152***	−0.154***	
	(0.029)	(0.029)	
golddiffat15	−0.0001***	−0.0001***	−0.00002**
	(0.00001)	(0.00001)	(0.00001)
firstbaron	−0.279***		0.805***
	(0.074)		(0.045)
firstdragon	0.495***		0.203***
	(0.071)		(0.044)
firstherald	0.339***		
	(0.078)		
heralds	−0.195***	−0.071**	−0.048*
	(0.054)	(0.035)	(0.025)
void_grubs	−0.059*	−0.063*	−0.002
	(0.034)	(0.034)	(0.024)
goldspentdiff	−0.0001***	−0.0001***	
	(0.00001)	(0.00001)	
firstblood	−0.342***		−0.484***
	(0.050)		(0.034)
dragons:I(gamelengthins/60)	−0.031***	−0.032***	
	(0.002)	(0.002)	
I(gamelengthins/60):barons	−0.031***	−0.026***	
	(0.003)	(0.003)	
firstbaron:firstdragon	−0.208**		0.302***
	(0.097)		(0.066)
Constant	−21.724***	−21.665***	−14.845***
	(0.281)	(0.275)	(0.124)
Observations	122,015	122,015	122,015
Log Likelihood	−6,396.885	−6,478.038	−13,607.550
Akaike Inf. Crit.	12,835.770	12,984.080	27,239.090
Pseudo R^2	92.4	92.3	83.9
Note:	*p<0.1; **p<0.05; ***p<0.01		

The regression results for three logit models are presented. Model 1 incorporates all in-game variables, while models 2 and 3 use only late-game and early-game variables, respectively. Upon evaluating the model fit, it becomes evident that relying solely on early-game variables does not provide sufficient explanatory power. Comparing models 1 and 2, model 1 exhibits a slight edge, indicated by its higher pseudo R-squared and lower AIC. Therefore, it is recommended to prioritize the utilization of model 1 for subsequent analysis.

The interaction term between game length and objectives (dragons and barons) aims to capture the diminishing utility offered by these objectives in longer games. As both teams reach the late-game phase, there is increased variance, and the potential benefits from securing these objectives decrease. Early leads still play a significant role, but their impact diminishes as the game progresses, making it important to account for this dynamic relationship between game length and objective importance in the analysis.

Logit Model on Games Shorter than 35 Minutes

	Dependent variable:		
	result		
	(1)	(2)	(3)
sideRed	0.206*	0.208**	−0.088
	(0.107)	(0.105)	(0.064)
percent_total_kpm	25.127***	24.965***	34.157***
	(0.690)	(0.683)	(0.474)
dragons	2.517***	2.590***	
	(0.367)	(0.362)	
barons	4.563***	4.271***	
	(1.060)	(1.017)	
wpm	0.025	0.037	
	(0.076)	(0.076)	
firstmidtower	−0.260**		−0.225***
	(0.132)		(0.085)
firsttothreetowers	−0.069		0.734***
	(0.143)		(0.089)
towers	1.267***	1.262***	
	(0.045)	(0.044)	
inhibitors	0.917***	0.915***	
	(0.117)	(0.115)	
golddiffat15	−0.0001**	−0.0001***	−0.00001
	(0.00003)	(0.00003)	(0.00002)
firstbaron	0.014		1.499***
	(0.180)		(0.090)
firstdragon	0.198		−0.004
	(0.153)		(0.085)
firstherald	0.413***		
	(0.158)		
heralds	−0.271**	−0.136*	−0.090*
	(0.107)	(0.072)	(0.050)
void_grubs	0.083	0.087	0.001
	(0.078)	(0.078)	(0.044)
goldspentdiff	−0.0002***	−0.0002***	
	(0.00002)	(0.00002)	
firstblood	−0.294***		−0.591***
	(0.106)		(0.065)
dragons:I(gamelengthins/60)	−0.077***	−0.080***	
	(0.011)	(0.011)	
I(gamelengthins/60):barons	−0.121***	−0.113***	
	(0.032)	(0.031)	
firstbaron:firstdragon	−0.161		0.400***
	(0.210)		(0.132)
Constant	−21.797***	−21.863***	−17.605***
	(0.544)	(0.535)	(0.250)
Observations	84,145	84,145	84,145
Log Likelihood	−1,498.254	−1,508.381	−3,763.108
Akaike Inf. Crit.	3,038.508	3,044.762	7,550.215
Pseudo R^2	97.4	97.4	93.5
Note:	*p<0.1; **p<0.05; ***p<0.01		

The regression results above compare the three models on games that were shorter than 35 minutes, categorized as shorter games based on early-game and late-game definitions. Model 3 exhibits a significant improvement, accompanied by slight improvements in models 1 and 2. This further suggests that models 1 and 2 excel in capturing effects not accounted for by model 3.

Logit Models on Games Longer than 35 minutes

	Dependent variable:		
	result		
	(1)	(2)	(3)
sideRed	0.099*	0.125**	$-$0.030
	(0.056)	(0.055)	(0.038)
percent_total_kpm	22.191***	21.805***	25.422***
	(0.400)	(0.392)	(0.278)
dragons	1.391***	1.463***	
	(0.101)	(0.100)	
barons	0.882***	0.464***	
	(0.176)	(0.168)	
wpm	0.130***	0.130***	
	(0.039)	(0.039)	
firstmidtower	0.086		$-$0.135***
	(0.071)		(0.049)
firsttothreetowers	$-$0.250***		0.468***
	(0.076)		(0.052)
towers	1.391***	1.375***	
	(0.026)	(0.025)	
inhibitors	$-$0.250***	$-$0.253***	
	(0.030)	(0.030)	
golddiffat15	$-$0.0001***	$-$0.0002***	$-$0.00005**
	(0.00002)	(0.00002)	(0.00001)
firstbaron	$-$0.430***		0.479***
	(0.084)		(0.052)
firstdragon	0.595***		0.276***
	(0.081)		(0.051)
firstherald	0.292***		
	(0.094)		
heralds	$-$0.215***	$-$0.110***	$-$0.027
	(0.065)	(0.041)	(0.029)
void_grubs	$-$0.126***	$-$0.126***	$-$0.005
	(0.041)	(0.041)	(0.029)
goldspentdiff	$-$0.0001***	$-$0.0001***	
	(0.00001)	(0.00001)	
firstblood	$-$0.319***		$-$0.410***
	(0.058)		(0.039)
dragons:I(gamelengthins/60)	$-$0.030***	$-$0.031***	
	(0.002)	(0.002)	
I(gamelengthins/60):barons	$-$0.015***	$-$0.009**	
	(0.004)	(0.004)	
firstbaron:firstdragon	$-$0.256**		0.231***
	(0.112)		(0.076)
Constant	$-$21.976***	$-$21.786***	$-$13.014***
	(0.349)	(0.340)	(0.146)
Observations	37,870	37,870	37,870
Log Likelihood	$-$4,609.399	$-$4,694.500	$-$9,549.481
Akaike Inf. Crit.	9,260.798	9,417.000	19,122.960
Pseudo R^2	82.4	82.11	63.6
Note:	*p<0.1; **p<0.05; ***p<0.01		

Comparing the three models on games longer than 35 minutes, it is evident that all three models perform worse, as indicated by their lower R-squared values. This suggests that shorter games are easier to predict and explain, whereas longer games involve more complex and diverse factors, making them inherently more challenging to model accurately.

Model Evaluation and Performance of Logit Model 1

	Reference		Classification Error
Prediction	0	1	
0	14,836	261	1.7%
1	317	15,089	2.1%

Accuracy: 0.9811

P-Value: <2.2 e-16

Kappa: 0.9621

Sensitivity: 0.9791

Specificity: 0.9830

Model Evaluation and Performance of Logit Model 2

	Reference		Classification Error
Prediction	0	1	
0	14,831	251	1.7%
1	322	15,099	2.1%

Accuracy: 0.9812

P-Value: <2.2 e-16

Kappa: 0.9624

Sensitivity: 0.9788

Specificity: 0.9836

Model Evaluation and Performance of Logit Model 3

	Reference		Classification Error
Prediction	0	1	
0	14,502	658	4.3%
1	651	14,692	4.2%

Accuracy: 0.9571

P-Value: <2.2 e-16

Kappa: 0.9142

Sensitivity: 0.9570

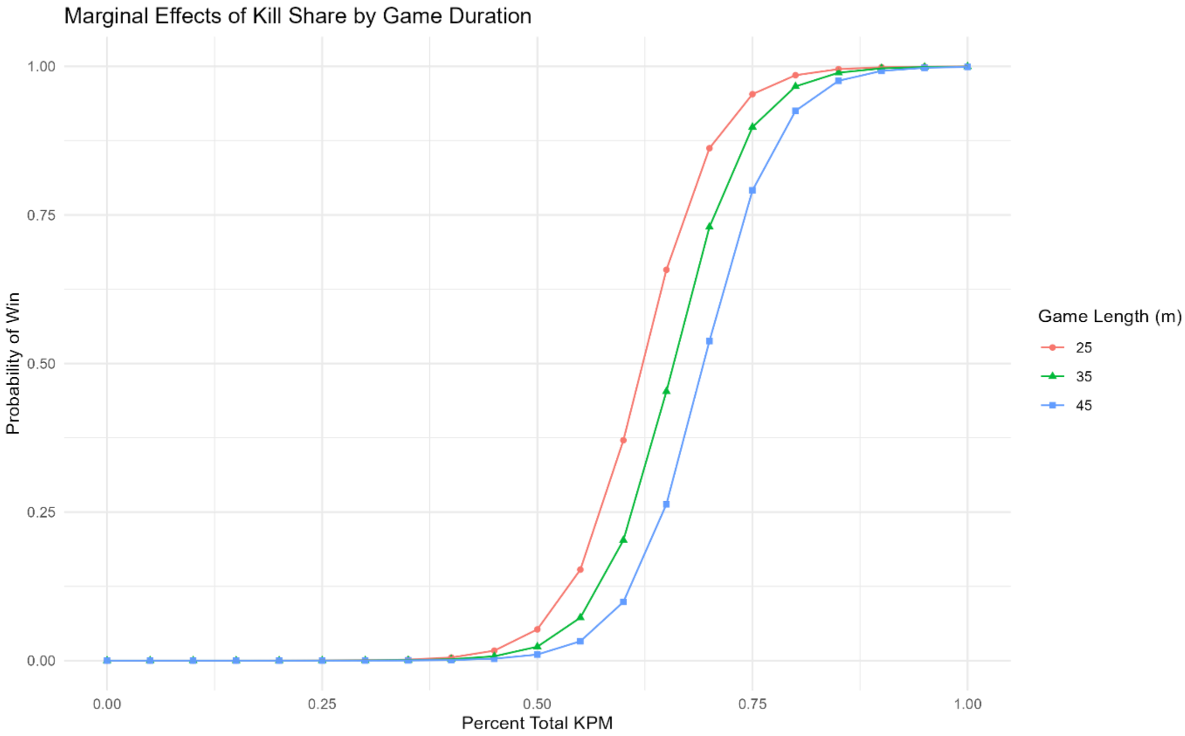
Specificity: 0.9571

Upon reviewing the confusion matrices above, it becomes apparent that model 2 outperforms the others in terms of accuracy, and specificity. This result is surprising, considering the performance of the models from the regression results. Model 1 seems to exhibit some degree of overfitting, likely due to its inclusion of more in-game variables compared to model 2. However, when considering classification errors, both models perform similarly, with model 2 demonstrating fewer type 1 and type 2 errors.

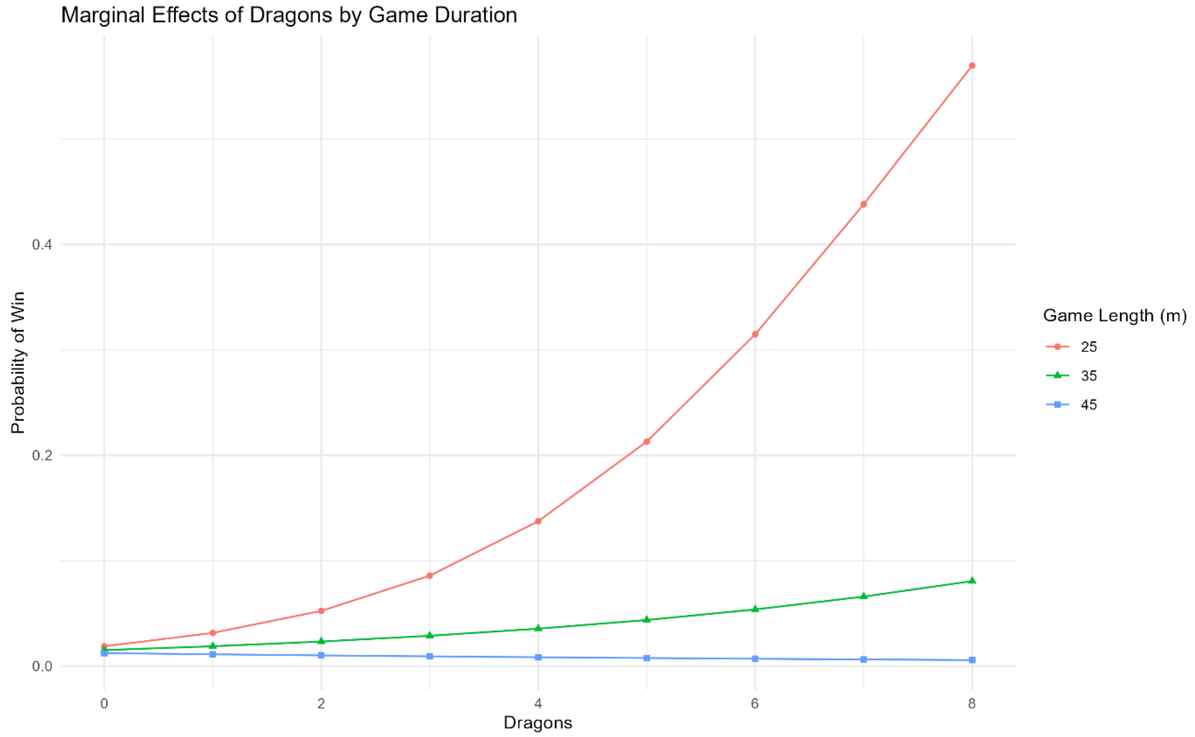
Considering the explanatory power of the models, it is evident that model 1, which includes a broader range of variables, offers more explanatory power despite its relatively lower predictive accuracy. Therefore, we will proceed with an in-depth analysis using model 1 to fully understand and interpret the importance and significance of each variable.

Marginal Effects Using Model 1

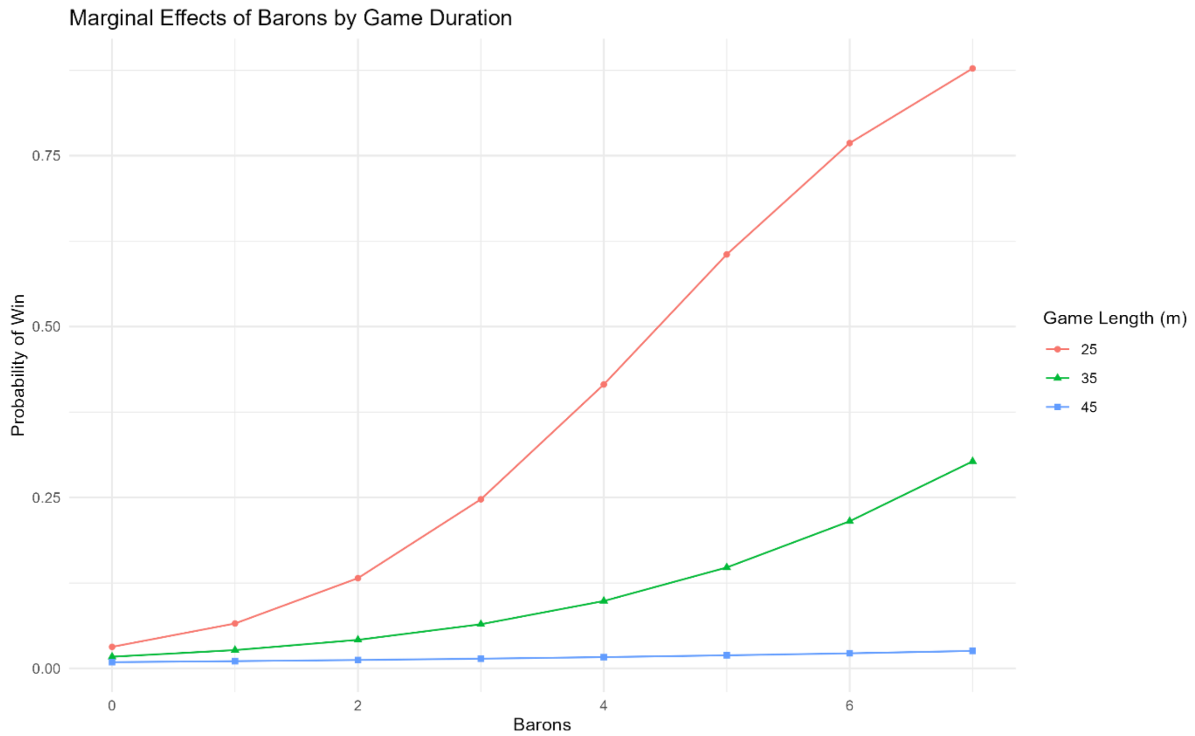
Kill share by Game Duration



Dragons by Game Duration



Barons by Game Duration



Analyzing the marginal effects of objectives using logit model 1, it becomes evident that kill share and all other metrics and objectives become less important as the game duration surpasses the 45-minute mark. Conversely, their significance is amplified in games lasting less than 25 minutes. This highlights the crucial importance of objectives, particularly in the early stages of the game, with dragons showing slightly greater importance than barons. However, for games around the average length, barons emerge as significantly more influential, as indicated by their steeper slope.

From these findings, it can be inferred that red side teams may benefit from prioritizing early-game compositions to secure as many dragons as possible and consequently close out games swiftly. Conversely, blue side teams could capitalize on their geographical advantage to secure barons, utilizing late-game compositions to stall out games and ultimately secure victories with baron buffs.

However, considering the lesser advantage of the blue side in securing the first Baron at the highest level of play, it appears that the red side holds a notable advantage. The only discrepancy lies in the blue side's ability to secure the first mid tower. At this stage, it's challenging to draw definitive conclusions. Nevertheless, given the significance of objectives such as dragons, if the red side strategically opts for a late-game composition and successfully secures dragons in the early stages of the game, leveraging the advantages offered by dragon buffs to secure the first Baron, it could be argued that the red side holds a stronger position in the highest level of professional play.

Logit Model 1 on Varying Game Lengths

	Game Length:		
	(< 30)	(> 30, < 40)	(> 40)
sideRed	−0.197	0.253***	0.005
	(0.237)	(0.074)	(0.071)
percent_total_kpm	18.898***	25.305***	21.121***
	(1.160)	(0.521)	(0.524)
dragons	5.049***	1.224***	1.311***
	(1.040)	(0.252)	(0.147)
barons	−2.600	5.181***	0.398*
	(3.218)	(0.644)	(0.236)
wpm	0.232	0.072	0.167***
	(0.197)	(0.051)	(0.049)
firstmidtower	0.132	−0.145	0.172*
	(0.310)	(0.094)	(0.090)
firsttothreetowers	−0.169	−0.225**	−0.203**
	(0.338)	(0.101)	(0.096)
towers	0.983***	1.527***	1.325***
	(0.079)	(0.036)	(0.033)
inhibitors	2.380***	0.020	−0.272***
	(0.261)	(0.062)	(0.034)
golddiffat15	0.00005	−0.0001***	−0.0001***
	(0.0001)	(0.00002)	(0.00002)
firstbaron	0.613	−0.268**	−0.468***
	(0.456)	(0.115)	(0.107)
firstdragon	0.419	0.358***	0.705***
	(0.322)	(0.109)	(0.103)
firstherald	0.041	0.512***	0.182
	(0.345)	(0.115)	(0.125)
heralds	0.074	−0.321***	−0.222**
	(0.229)	(0.078)	(0.089)
void_grubs	0.085	−0.055	−0.162***
	(0.183)	(0.050)	(0.058)
goldspentdiff	−0.0001**	−0.0002***	−0.0001***
	(0.00004)	(0.00001)	(0.00001)
firstblood	−0.022	−0.263***	−0.411***
	(0.230)	(0.076)	(0.075)
dragons:I(gamelengthins/60)	−0.173***	−0.032***	−0.027***
	(0.037)	(0.007)	(0.003)
I(gamelengthins/60):barons	0.118	−0.131***	−0.003
	(0.114)	(0.017)	(0.005)
firstbaron:firstdragon	0.208	−0.235	−0.297**
	(0.512)	(0.146)	(0.142)
Constant	−18.393***	−23.677***	−21.423***
	(1.033)	(0.436)	(0.462)
Observations	45,412	61,435	15,168
Log Likelihood	−347.325	−2,855.281	−2,765.742
Akaike Inf. Crit.	736.650	5,752.563	5,573.485
Pseudo R^2	98.9	93.3	73.7
Note:	*p<0.1; **p<0.05; ***p<0.01		

Analyzing logit model 1 across varying game lengths reveals some contradictions with previous marginal effects analyses. Notably, kill share appears to be less important in the early stages of the game compared to longer games. Dragons play a significantly larger role in shorter games, with a coefficient of 5 compared to 1, indicating their crucial importance in early game strategies. Surprisingly, towers seem to play a smaller role in shorter games, possibly due to teams having less time to capture them. Furthermore, first dragon, Baron, and Herald do not appear to influence shorter games significantly, with kill share emerging as the main driving force for victory. This suggests that, in general, securing more kills than opponents offers greater control over the game, irrespective of team composition or objective control.

A surprising finding is the negative coefficient of early game gold leads. Contrary to common belief and a key metric frequently highlighted in live professional game broadcasts, it appears that not only does an early gold lead have little to no effect on the game's outcome, but it might even negatively impact the team's chances of winning.

A possible explanation for this phenomenon is that, at the high level of professional games we are analyzing, the disparity in gold is generally low. This results in no clear item advantage for the team with an early gold lead. Consequently, the strategies adopted by teams in macro play and team fights likely play a more significant role in securing key objectives. Since these strategies are not easily measurable metrics, a deeper analysis into this aspect is beyond our current scope.

Logit Model 1 on Seasons 10-13

	Season			
	(10)	(11)	(12)	(13)
sideRed	−0.061	0.274**	0.023	0.402
	(0.135)	(0.130)	(0.150)	(0.254)
percent_total_kpm	29.180***	25.264***	26.293***	27.090***
	(1.089)	(0.906)	(1.128)	(1.885)
dragons	0.889***	1.339***	1.541***	1.346***
	(0.238)	(0.233)	(0.270)	(0.444)
barons	2.390***	2.796***	2.386***	3.499***
	(0.554)	(0.526)	(0.553)	(0.992)
wpm	−0.128	0.040	0.067	−0.051
	(0.099)	(0.113)	(0.072)	(0.203)
firstmidtower	−0.087	−0.110	−0.181	−0.125
	(0.174)	(0.170)	(0.191)	(0.316)
firsttothreetowers	−0.107	−0.193	−0.290	−0.401
	(0.188)	(0.174)	(0.194)	(0.344)
towers	1.458***	1.519***	1.665***	1.789***
	(0.062)	(0.062)	(0.076)	(0.125)
inhibitors	−0.295***	−0.365***	−0.385***	0.096
	(0.092)	(0.087)	(0.095)	(0.157)
golddiffat15	−0.0001***	−0.0001***	−0.0001	−0.0003***
	(0.00004)	(0.00004)	(0.00005)	(0.0001)
firstbaron	−0.436**	−0.364*	−0.940***	−1.049***
	(0.207)	(0.194)	(0.231)	(0.379)
firstdragon	0.572***	0.366**	0.444**	−0.146
	(0.202)	(0.186)	(0.224)	(0.371)
firstherald	0.333*	0.990***	0.970***	1.230*
	(0.193)	(0.210)	(0.249)	(0.660)

	Season			
	(10)	(11)	(12)	(13)
heralds	−0.247*	−0.425***	−0.334*	−0.572
	(0.129)	(0.146)	(0.172)	(0.568)
void_grubs				−0.012
				(0.062)
goldspentdiff	−0.0002***	−0.0002***	−0.0002***	−0.0002***
	(0.00002)	(0.00002)	(0.00002)	(0.00003)
firstblood	−0.344**	−0.351***	−0.262*	−0.0004
	(0.137)	(0.130)	(0.155)	(0.247)
dragons:I(gamelengthins/60)	−0.024***	−0.032***	−0.037***	−0.035***
	(0.006)	(0.006)	(0.007)	(0.011)
I(gamelengthins/60):barons	−0.051***	−0.061***	−0.048***	−0.079***
	(0.013)	(0.013)	(0.013)	(0.023)
firstbaron:firstdragon	−0.044	−0.343	0.128	0.052
	(0.267)	(0.255)	(0.301)	(0.479)
Constant	−23.948***	−23.643***	−25.303***	−26.406***
	(0.825)	(0.797)	(0.924)	(1.654)
Observations	19,331	19,656	16,803	6,468
Log Likelihood	−841.438	−953.728	−687.656	−269.249
Akaike Inf. Crit.	1,722.875	1,947.457	1,415.311	580.497
Pseudo R^2	93.7	93	94.1	94
Note:				*p<0.1; **p<0.05; ***p<0.01

Analyzing logit model 1 across seasons 10 to 13 reveals that the red side only holds an advantage in season 11. Kill share remains consistently important across all seasons, offering control over the pace and objectives of the games. Dragons show increasing importance over the seasons, possibly due to changes implemented by game developers, while Barons exhibit a significant spike in significance in season 13, possibly due to alterations in their buffs or changes in gameplay dynamics. Surprisingly, first dragons play a less significant role in season 13, while first herald plays a more significant and crucial role. Additionally, first blood appears to have a negative impact on game outcomes, indicating that early-game events may not heavily influence the overall result of the game.

The regression model allows us to capture the evolving effects of key objectives as they change throughout the seasons. It is important to highlight that despite the insignificant coefficients for objectives such as Void Grubs, this could be due to multicollinearity on a theoretical level, even though the data and VIF results conducted prior did not show significant multicollinearity.

For instance, both Herald and Void Grubs provide teams with greater tower-pushing power. In future research, we could attempt to capture this effect by individually modeling how these objectives contribute to a team's ability to secure other objectives. This approach could provide a clearer understanding of their impact on game outcomes.

Logit Model 1 by Major Region

	League			
	(LCS)	(LCK)	(LEC)	(LPL)
sideRed	0.304	0.120	0.360	0.351**
	(0.236)	(0.234)	(0.276)	(0.176)
percent_total_kpm	22.674***	28.211***	30.033***	21.714***
	(1.553)	(1.809)	(2.145)	(1.142)
dragons	1.462***	1.365***	1.013**	1.049***
	(0.345)	(0.331)	(0.420)	(0.240)
barons	2.313***	2.527***	1.468**	1.447***
	(0.653)	(0.658)	(0.716)	(0.478)
wpm	−0.182	−0.034	−0.318	0.094
	(0.217)	(0.190)	(0.229)	(0.116)
firstmidtower	−0.299	0.193	0.643**	−0.647**
	(0.279)	(0.308)	(0.321)	(0.321)
firsttothreetowers	−0.002	−0.390	−0.300	0.538
	(0.301)	(0.331)	(0.347)	(0.347)
towers	1.371***	1.526***	1.679***	1.224***
	(0.111)	(0.116)	(0.143)	(0.072)
inhibitors	0.053	−0.077	−0.535***	0.092
	(0.154)	(0.146)	(0.157)	(0.112)
golddiffat15	−0.0002**	−0.00001	−0.0001	−0.0001**
	(0.0001)	(0.0001)	(0.0001)	(0.0001)
firstbaron	−0.843**	−0.878**	−0.947**	0.257
	(0.395)	(0.398)	(0.440)	(0.330)
firstdragon	0.655*	0.474	−0.163	0.064
	(0.342)	(0.336)	(0.409)	(0.318)
firstherald	0.371	0.144	0.021	0.164
	(0.446)	(0.432)	(0.518)	(0.321)

	League			
	(LCS)	(LCK)	(LEC)	(LPL)
heralds	−0.192	−0.045	0.028	−0.148
	(0.309)	(0.310)	(0.378)	(0.235)
void_grubs	−0.240	−0.068	−0.028	
	(0.185)	(0.163)	(0.227)	
goldspentdiff	−0.0001***	−0.0002***	−0.0002***	−0.0002***
	(0.00002)	(0.00003)	(0.00003)	(0.00003)
firstblood	−0.625**	−0.536**	−0.455	−0.279
	(0.254)	(0.257)	(0.296)	(0.174)
dragons:I(gamelengthins/60)	−0.032***	−0.033***	−0.022**	−0.027***
	(0.007)	(0.007)	(0.009)	(0.006)
I(gamelengthins/60):barons	−0.041***	−0.045***	−0.025*	−0.028***
	(0.012)	(0.013)	(0.014)	(0.010)
firstbaron:firstdragon	−0.362	0.219	0.264	−0.585
	(0.450)	(0.456)	(0.521)	(0.454)
Constant	−21.050***	−24.916***	−25.993***	−20.200***
	(1.425)	(1.630)	(1.899)	(0.963)
Observations	5,286	7,294	4,850	8,925
Log Likelihood	−284.651	−275.315	−218.496	−499.111
Akaike Inf. Crit.	611.302	592.631	478.991	1,038.223
Pseudo R^2	92.2	94.6	93.5	91.9
Note:				*p<0.1; **p<0.05; ***p<0.01

Examining logit model 1 across the four major regions reveals that only the LPL experiences an advantage on the red side. Interestingly, the LEC appears to prioritize kill share more than other regions, while the LPL exhibits the lowest coefficient for kill share. This discrepancy could stem from differences in game length and playstyles, with kills potentially mattering less in the LPL due to frequent skirmishes throughout the game. Moreover, Barons appear to play a significant role in the LCS and LCK, but less so in the LEC and LPL. This variation could be attributed to longer game durations in certain regions and differing priorities on key game objectives.

Additionally, inhibitors play little to no positive role in all regions, with the LEC displaying a significant negative coefficient. This suggests that the advantages gained from super minions generated by destroying inhibitors may not be significant enough at the highest level of play to secure victories through the pressure exerted in lanes.

Random Forest Models

Random Forest Model 1: Draft, Player and In-game Variables

	Reference		Classification Error
Prediction	0	1	
0	14,638	78	0.5%
1	626	15,161	4.0%

Accuracy: 0.9771

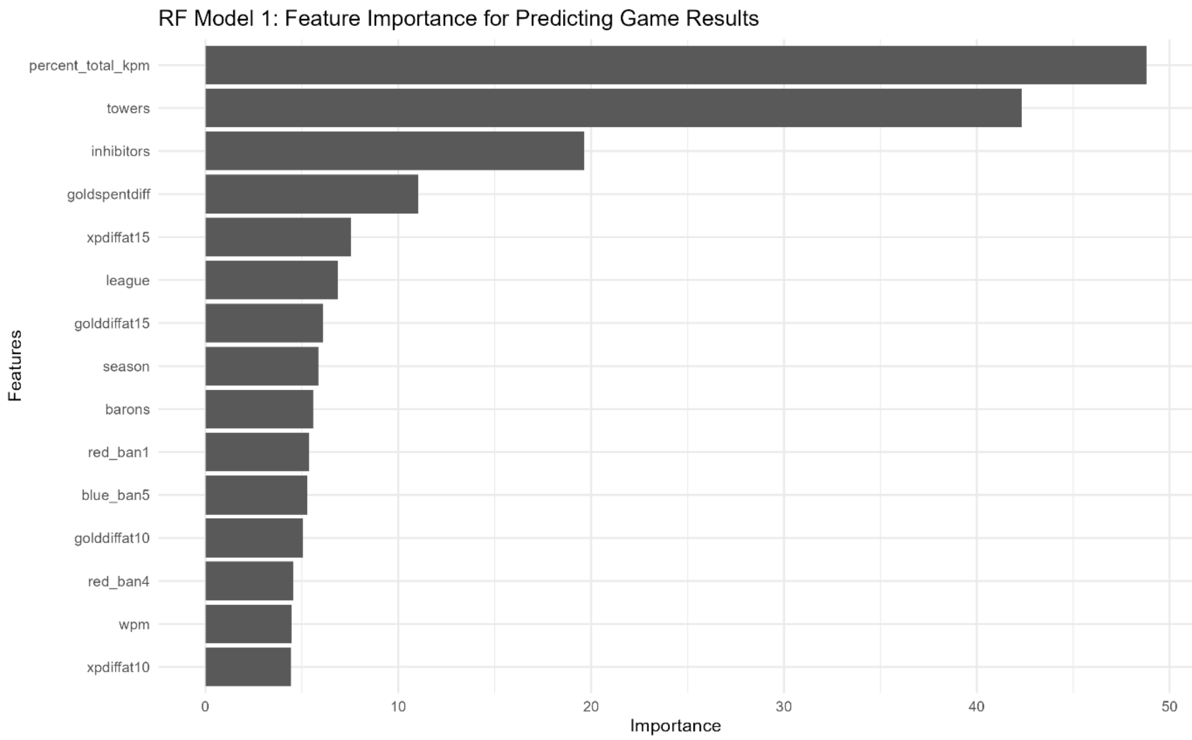
P-Value: <2.2 e-16

Kappa: 0.9541

Sensitivity: 0.9594

Specificity: 0.9947

Feature Importance (RF1)



Random forest model 1, which utilizes all available data on draft, player, and in-game metrics, demonstrates strong predictive performance in determining game outcomes. The variables of kill share, towers, and inhibitors emerge as the most influential factors in predicting game results. However, the model does suffer from relatively high false positives, indicating room for improvement in minimizing incorrect predictions of winning outcomes. On the positive side, the model excels in predicting losing games, as evidenced by its low false negative rate.

Random Forest Model 2: In-game Variables

	Reference		Classification Error
Prediction	0	1	
0	14,914	218	1.4%
1	350	15,021	2.3%

Accuracy: 0.9814

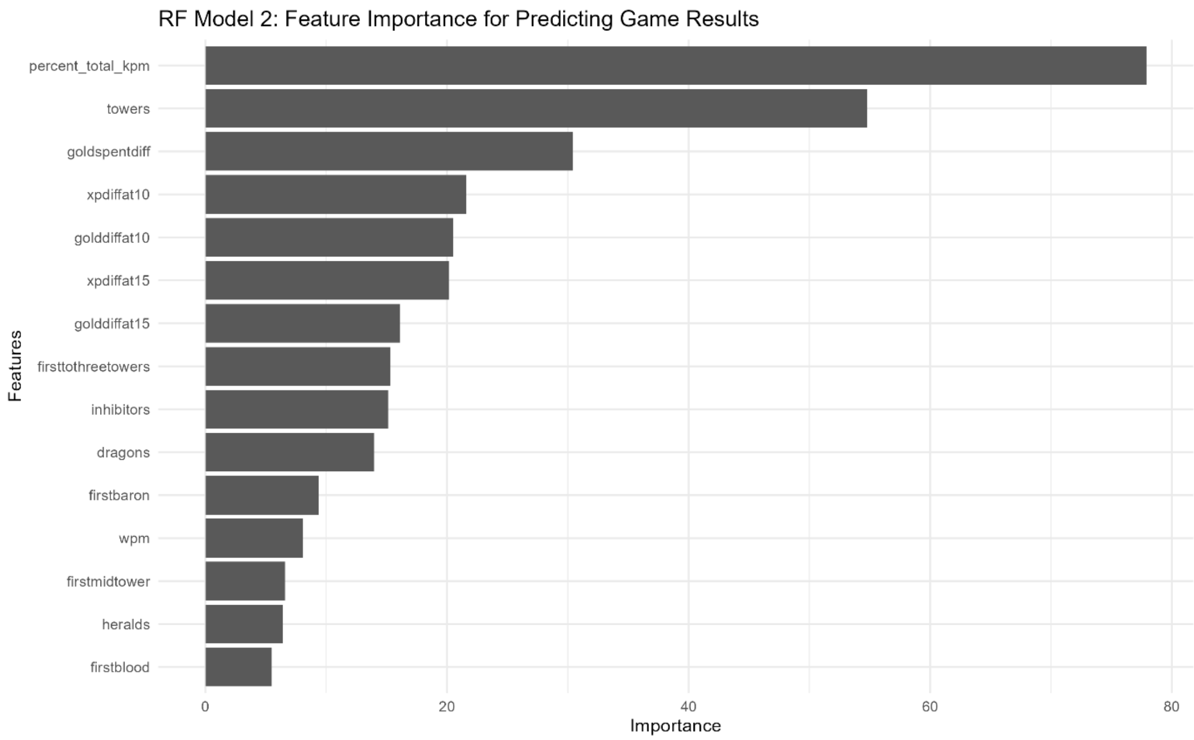
P-Value: <2.2 e-16

Kappa: 0.9628

Sensitivity: 0.9771

Specificity: 0.9857

Feature Importance (RF2)



Random forest model 2, which only utilizes in-game variables, demonstrates superior performance compared to model 1, which includes all variables. This improvement suggests that overfitting may have been an issue in model 1. However, model 2 performs less effectively in predicting losses, indicated by its higher false negative rate.

Interestingly, upon examining the feature importance graph, it is evident that inhibitors are no longer as important as they used to be in model 2. Instead, the difference in total gold spent has taken over its place as a more influential factor in predicting game outcomes.

Random Forest Model 3: Draft and Player Variables

	Reference		Classification Error
Prediction	Blue	Red	
Blue	5,639	3,652	39.3%
Red	2,477	3,483	41.6%

Accuracy: 0.5981

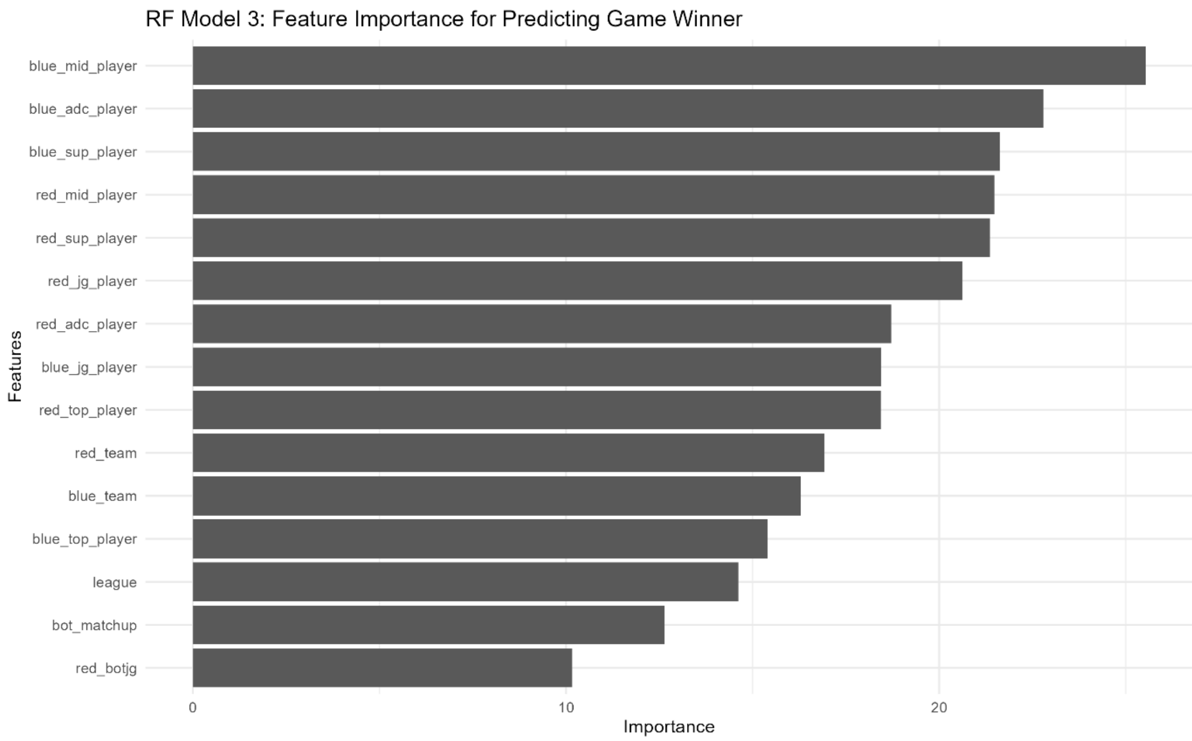
P-Value: <2.2 e-16

Kappa: 0.1848

Sensitivity: 0.6948

Specificity: 0.4882

Feature importance (RF3)



Random forest model 3, utilizing only pre-game variables such as draft and player data, exhibits lower accuracy compared to models incorporating in-game metrics. However, it still suggests that outcomes can be predicted with reasonable accuracy prior to the start of a game. Surprisingly, player skill appears to play a greater role than the actual champion picks during drafts, indicating the significance of player skill differences in determining game outcomes.

Additionally, the model highlights the mid player as the most significant role, suggesting that the mid lane holds the highest importance in professional play.

Random Forest 4: Draft variables

	Reference		Classification Error
Prediction	Blue	Red	
Blue	5,627	4,768	45.9%
Red	2,489	2,367	51.3%

Accuracy: 0.5242

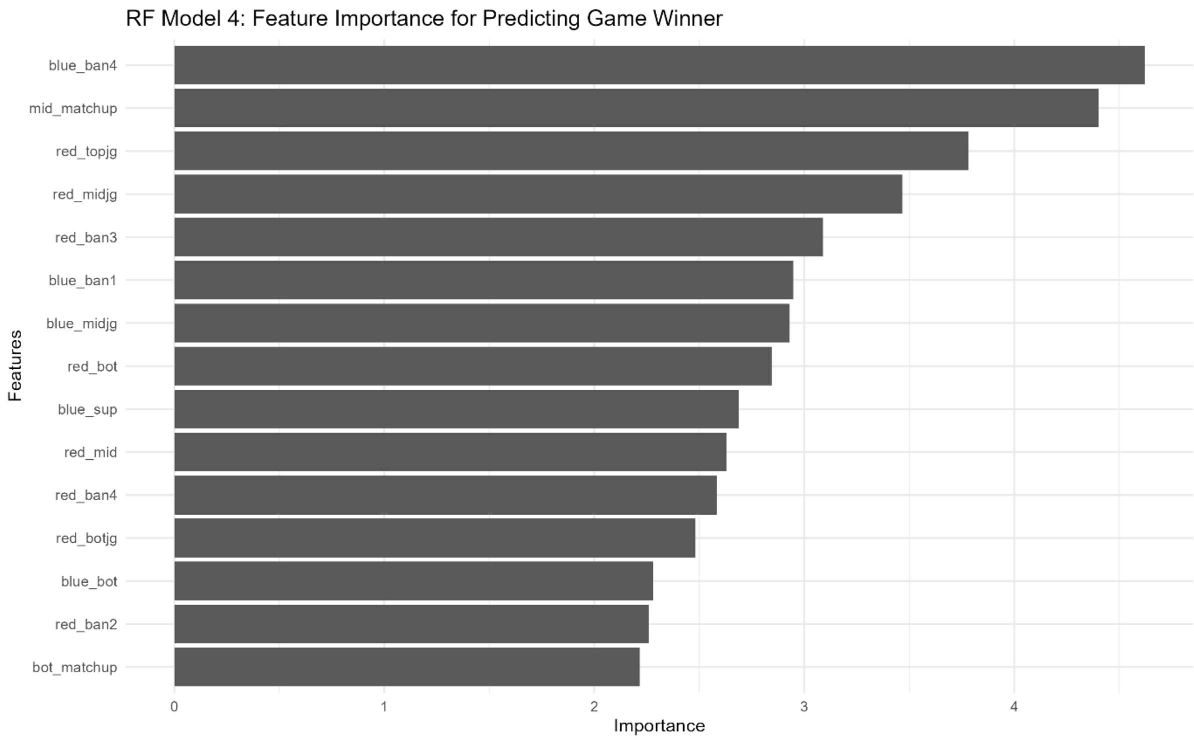
P-Value: <0.9766

Kappa: 0.0256

Sensitivity: 0.6933

Specificity: 0.3317

Feature Importance (RF4)



Random forest model 4, focusing solely on draft variables, exhibits reduced accuracy compared to model 3. This suggests that teams may already be drafting at similar levels, with no significant "draft differences" observable.

Examining feature importance, it is notable that aside from the fourth ban from the blue side, the mid lane matchup and the top-jungle dynamic play the most significant roles. This indicates that the bot lane may serve a less significant purpose in drafting, as counters are more prominent in other roles.

XGboost models

XGboost Model 1: Draft and Player Variables

	Reference		Classification Error
Prediction	Blue	Red	
Blue	5,463	3,250	37.3%
Red	2,760	3,778	42.2%

Accuracy: 0.6059

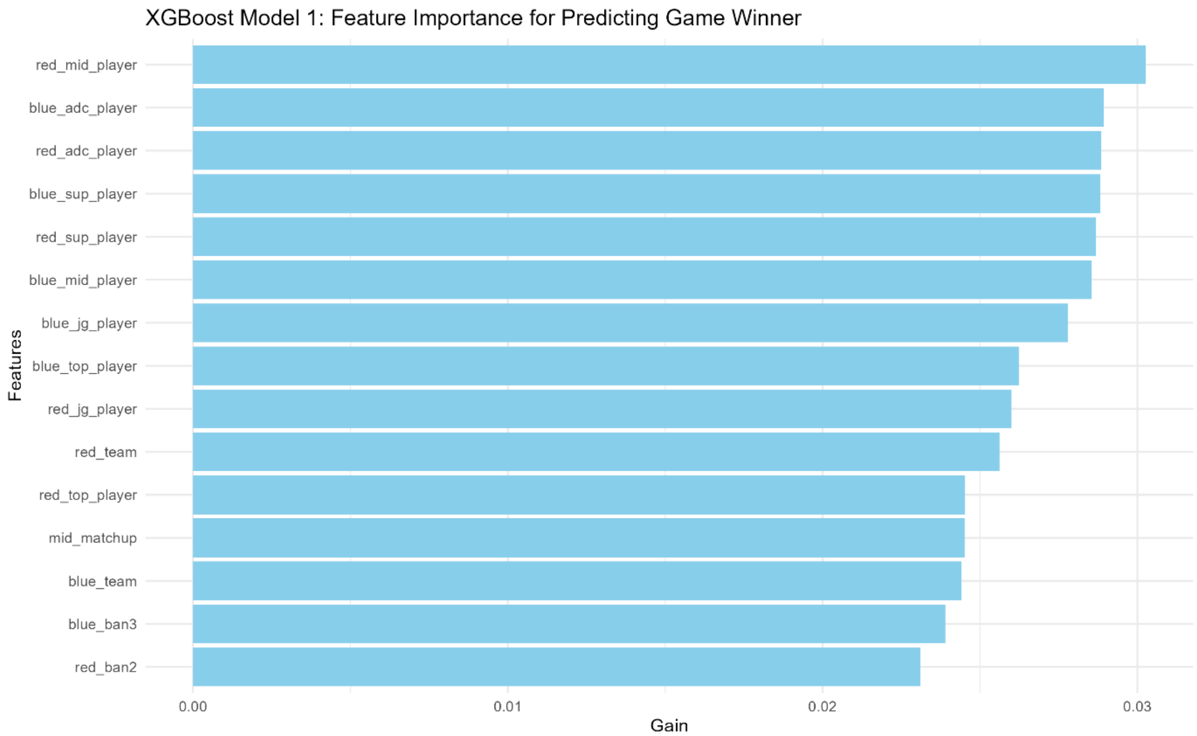
P-Value: <202e-16

Kappa: 0.2029

Sensitivity: 6644

Specificity: 0.5376

Feature Importance (XGboost 1)



The XGBoost model 1, utilizing draft and player variables, achieves a decent accuracy of 0.6. Similar to previous findings, it suggests that players play a significantly greater role than the actual draft in predicting game outcomes. This further reinforces our conclusion regarding the existence of player skill differences and their impact on game results.

XGboost Model 2: Draft Variables

	Reference		Classification Error
Prediction	Blue	Red	
Blue	5,050	4,041	44.5%
Red	3,173	2,987	51.5%

Accuracy: 0.527

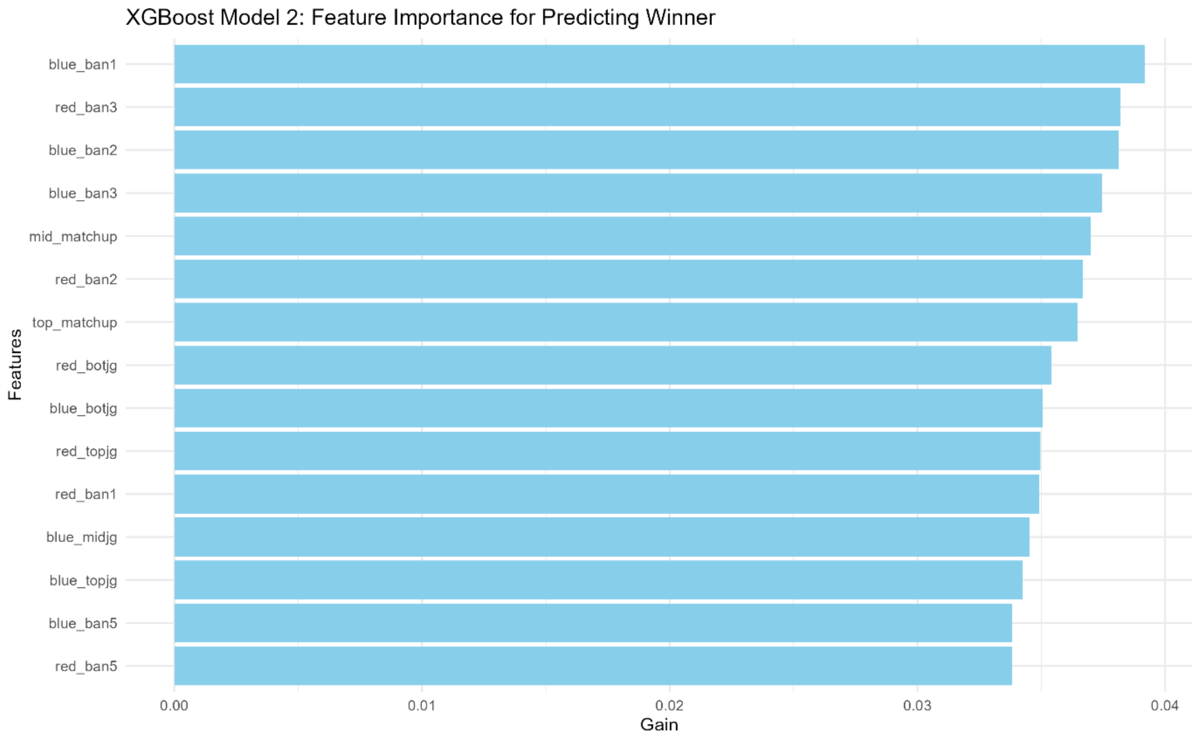
P-Value: <0.9988

Kappa: 0.0395

Sensitivity: 0.6141

Specificity: 0.4250

Feature Importance (XGboost 2)



XGBoost model 2, which only utilizes draft variables, demonstrates lower performance compared to models incorporating player variables. This suggests that drafting plays a relatively minor role in determining the outcome of the game.

Examining the feature importance, bans emerge as the most important variable, followed by solo lane matchups. This indicates that differences in performance in these solo lanes play a larger role than the bot lane and jungle. Given that solo lanes involve counter matchups that can significantly influence the game's outcome, this finding is not surprising. In contrast, the bot lane may have a relatively simpler objective in the game, contributing to its lesser importance in the draft.

XGboost Model 3: Draft Variables on Season 12

	Reference		Classification Error
Prediction	Blue	Red	
Blue	709	563	44.3%
Red	395	430	52.1%

Accuracy: 0.5432

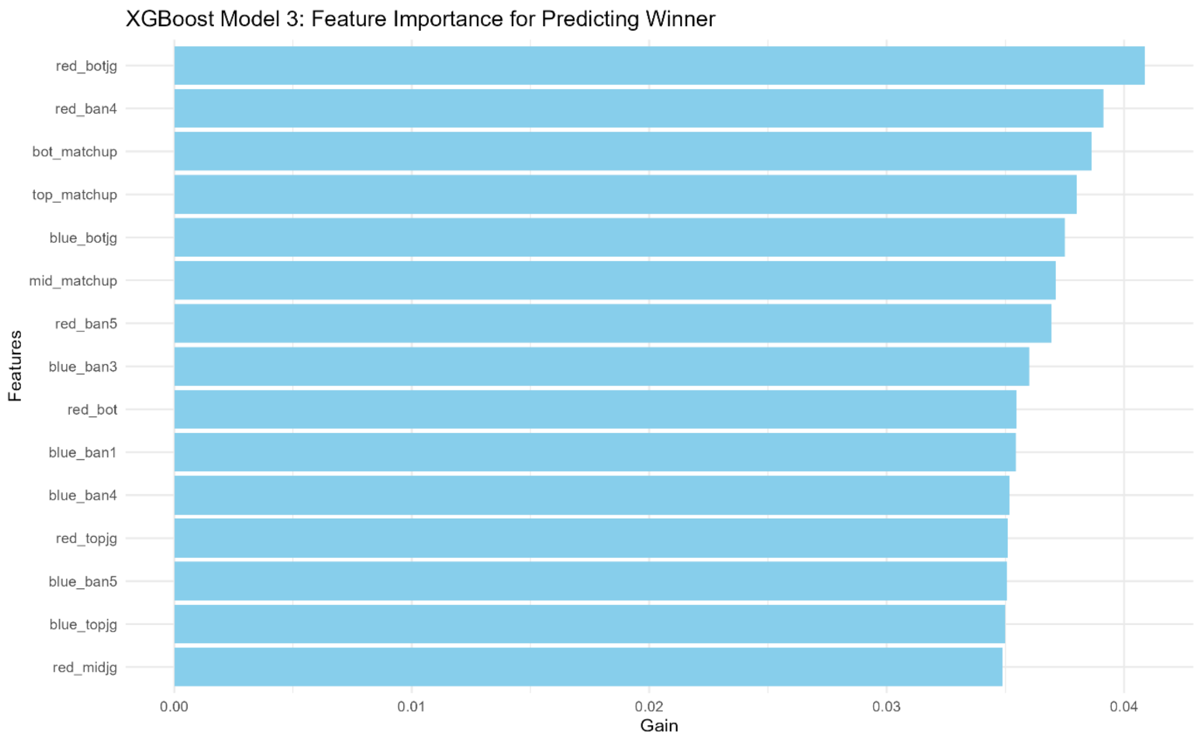
P-Value: <0.06559

Kappa: 0.0759

Sensitivity: 0.6422

Specificity: 0.433

Feature Importance (XGboost 3)



XGBoost model 3 focuses solely on draft variables from season 12, recognizing that changes in the meta and champion picks and bans could impact the analysis. The improved performance and significance of p-values suggest that when analyzing drafts, a cross-sectional approach may not be suitable.

Interestingly, in season 12, the bot jungle dynamic and bot lane matchup emerge as the most important variables. This shift in importance could reflect a meta change placing greater emphasis on the bottom lane rather than the solo lanes. Surprisingly, the mid matchup is no longer as crucial as suggested by previous models. This shift may indicate a broader evolution of the game from a kill- and lane-driven playstyle to a more team-oriented approach, where objectives take priority over individual lane dominance.

Discussion

In analyzing the results, it becomes evident that the blue side initially appears to hold an advantage, potentially attributed to various factors such as prevailing meta trends, map dynamics, and strategic choices during the drafting phase. The exact source of this advantage remains ambiguous; however, upon closer examination through logit regression, it becomes clear that the models encapsulate the advantages experienced by the blue side, with statistical significance pointing towards a slight advantage for the red side.

Key objectives such as barons, heralds, and void grubs, where the blue side holds geographical superiority, are likely contributors to this advantage. Notably, our findings suggest that objectives in the mid to late game hold greater significance compared to early game objectives, emphasizing the importance of securing neutral objectives in the jungle, establishing map control, and accruing gold leads from early tower takedowns and lane dominance. Interestingly, while gold leads are traditionally considered pivotal, our analysis suggests their influence may be overstated, serving primarily as a means to amplify a team's damage potential through item advantages.

In the League of Legends community, there's a prevalent strategy among high-level teams to draft compositions geared towards late-game dominance, prioritizing strategic advantages in later stages. However, our models do not explicitly capture this variable due to its relative and nuanced nature. Nevertheless, our research successfully quantifies and underscores the significance of every key objective within the game.

For teams opting for early-game compositions aimed at swift victories, focusing on objectives like dragons, towers and inhibitors proves pivotal, as they mitigate the risk of being outpaced by opponents in later stages. Regardless of game length, a consistent pattern emerges wherein kills wield considerable influence, enabling teams to assert pressure and capitalize on incremental gold and experience advantages. This dispels the notion that kills hold minimal importance, particularly at the highest echelons of play, where they serve as catalysts for securing objectives and exerting control over opponent movements.

Notably, with game durations shorter in Season 13 and longer kill timers compared to earlier seasons, controlling the flow of engagements becomes paramount. Our machine learning and ensemble models shed light on the significance of draft differentials, showcasing a modest but notable impact on professional gameplay. While treating all compositions equally should yield a 50% prediction accuracy, our findings demonstrate a slightly higher accuracy of 54%, underscoring the need for meticulous consideration of team compositions during champion select.

However, overarching the influence of draft compositions is the undeniable importance of player and team skill. Our analysis reveals a persistent skill gap among professional players, with individual prowess often outweighing the impact of champion selecti ons. Thus, the roles of coaches and analysts, while significant, are somewhat constrained by the dominant influence of player skill over strategic choices during the draft phase.

Conclusion

In conclusion, our research sheds light on the multifaceted dynamics of League of Legends gameplay, particularly regarding the impact of team compositions, objectives, and player skill on match outcomes. While the initial advantage of the blue side is evident, stemming from various strategic elements such as map dynamics and drafting strategies, our analysis underscores the nuanced interplay of factors contributing to this advantage.

Key objectives like barons, heralds, and void grubs play pivotal roles, with the blue side often holding geographical superiority in their control. Moreover, our findings highlight the evolving importance of mid to late game objectives over early game ones, challenging conventional wisdom about the significance of early leads.

The League of Legends community's preference for late-game compositions is apparent, though our models may not explicitly capture this due to its complexity. Nonetheless, our research provides valuable insights into the quantifiable impact of every key objective within the game, informing strategic decision-making for teams aiming for victory.

For teams pursuing early-game dominance, strategic focus on objectives like dragons and inhibitors is paramount to maintaining momentum and avoiding being outpaced in later stages. Additionally, our analysis underscores the enduring significance of kills in asserting pressure and securing advantages, debunking the notion that kills hold minimal importance.

In the evolving landscape of League of Legends gameplay, marked by shorter game durations and refined strategies, the careful consideration of draft compositions becomes increasingly crucial. While our models demonstrate a modest but noteworthy impact of draft differences on match outcomes, the overarching influence of player and team skill remains undeniable.

Ultimately, while coaches and analysts play vital roles in guiding strategic decisions during the draft phase, the dominance of player skill underscores the need for continuous improvement and refinement in gameplay. As League of Legends continues to evolve, our research provides a foundation for understanding the intricate dynamics shaping professional gameplay and offers insights into optimizing strategies for success on the Rift.

Future Works

In the future, there are several avenues of research that can expand upon our current findings and deepen our understanding of the factors influencing gameplay dynamics in League of Legends. Firstly, it would be valuable to investigate the specific variables contributing to the securing of crucial objectives that often dictate game outcomes, such as barons, dragons, heralds, and towers. Considering the impact of champion abilities, team compositions, and strategic decisions on objective control could provide insights into optimal gameplay strategies tailored to different scenarios.

Furthermore, delving into the intricate dynamics of team compositions and their interactions could offer valuable insights into the effectiveness of various strategies. Exploring how certain compositions counter or synergize with others, such as poke compositions versus hard engage compositions or split push compositions versus teams with strong pushing power, could inform strategic decision-making and draft phase prioritization.

Additionally, it would be beneficial to identify the most effective team compositions for each season and anticipate how the meta might evolve in response to patch notes, champion reworks, and item changes. By incorporating these factors into predictive models, we can anticipate shifts in gameplay trends and adapt strategies accordingly.

Moreover, understanding the effectiveness of different play styles and strategic approaches could provide valuable insights for teams seeking to optimize their performance. By analyzing the relative strengths and weaknesses of various play styles, we can identify the most effective strategies for achieving victory in different situations.

Finally, investigating the role of coaching staff and their impact on gameplay strategies and performance outcomes could revolutionize the way the game is played and approached in the future. By uncovering significant insights into coaching methodologies and their effects on team performance, we can work towards reducing the inherent skill gap between regions and promoting a more equitable competitive landscape.

Overall, by exploring these future research directions, we can continue to advance our understanding of League of Legends gameplay dynamics and contribute to the ongoing evolution of strategic decision-making in professional esports.

References

Janzen, B., & Viviani, G. . "Meta Search: A Tool for Mass Analysis of Game Strategy." Journal of Game Studies.

Morales-García, J., Llanes-Castro, A., Curado, M., & Arcas-Túnez, F. (2023). An Exploratory Data Analysis for League of Legends Professional Match Data. DOI: 10.3233/AISE230013.