Data Science for Competitive Improvement in *League of Legends*

Andrew Laurence T. Fat
College of Computing and Information Technologies
National University
Manila, Philippines
fatat@students.national-u.edu.ph

***Abstract*—Competitive performance in multiplayer online battle arena (MOBA) games is influenced by mechanical skill, strategic decision-making, and consistency across matches. This study investigates whether measurable improvement has occurred in the performance of a single top-lane player in League of Legends by comparing gameplay metrics across two consecutive time periods. Using match-history data retrieved via the Riot Games API, key performance indicators—including creep score per minute (CS/min), kill–death–assist ratio (KDA), gold earned, damage dealt, damage taken, and win rate—are analyzed using descriptive statistics, correlation analysis, and hypothesis testing.**

Descriptive trends suggest modest improvements in several performance metrics; however, inferential statistical tests did not reveal statistically significant mean differences between periods at the $\alpha = 0.05$ level. The findings highlight the challenges of detecting performance growth in high-variance, team-based competitive environments. This work demonstrates how personal gameplay analytics can be leveraged using data science techniques to support evidence-based competitive improvement.

***Index Terms*—Data science, League of Legends, performance evaluation, hypothesis testing**

I. Introduction

The rapid growth of esports has generated increasing interest in the application of data science techniques to competitive gaming performance analysis. *League of Legends* (LoL), a widely played multiplayer online battle arena (MOBA) game developed by Riot Games, provides a rich environment for quantitative analysis due to the availability of detailed match-level statistics through public application programming interfaces (APIs) [3].

While many studies focus on population-level trends or predictive modeling across thousands of players, fewer works examine longitudinal performance improvement at the individual level. For competitive players seeking self-improvement, understanding whether performance has objectively improved over time—and which metrics drive that improvement—is critical. Subjective impressions of progress may be misleading without statistical validation.

This study conducts a longitudinal analysis of a single player's top-lane performance across two consecutive time periods. By applying statistical hypothesis testing, correlation analysis, and regression-based modeling to consistently collected ranked match data, the study evaluates whether performance improvements are statistically significant and identifies the metrics most closely associated with winning matches.

A. Research Questions

This study addresses the following research questions:

- Which in-game performance metrics are most strongly correlated with match victories in *League of Legends*?
- Has the player's top-lane performance improved over time when comparing two three-month periods?
- Are improvements consistent across multiple champions or limited to a subset of frequently played champions?

B. Research Objectives

1) Primary Objective: The primary objective of this study is to determine whether measurable improvement has occurred in the player's top-lane performance during the three months following the start of the research period compared to the preceding three months. Improvement is operationalized as statistically significant increases in selected performance metrics.

2) Secondary Objectives: The secondary objectives are to:

- Identify which performance metrics exhibit the largest changes over time.
- Evaluate whether performance improvements are consistent across different champions.
- Assess whether changes in performance metrics are associated with improved match outcomes.
- Compare predictive model performance between the two time periods.

II. Literature Review

Previous research has demonstrated that in-game metrics such as gold accumulation, objective control, and damage contribution are strongly associated with match outcomes and player skill differentiation [2], [4]. Prior work has also explored player skill differentiation and expertise through fine-grained input behavior and performance indicators in *League of Legends* [1].

Lee et al. [1] characterized and quantified expert input behavior in *League of Legends*, providing insights into how mechanical skill differentiates player expertise levels. Martinez et al. [2] applied deep learning techniques to identify key performance indicators related to win prediction, map navigation, and vision control, demonstrating the importance of economy-related metrics in determining match outcomes. Schubert et al. [4] developed esports analytics frameworks through encounter detection, establishing methodological approaches for performance evaluation in competitive gaming contexts.

These studies emphasize the importance of economy-focused metrics and strategic decision-making in MOBA games, but few examine longitudinal individual performance improvement using rigorous statistical validation. This study builds on these foundations by applying similar analytical methods to a single-player case study over time.

III. Methodology

A. Data Collection and Preprocessing

1) Data Source: Match data were collected using the Riot Games API, specifically the *match-v5* and *timeline-v5* endpoints [3]. Only ranked solo queue matches were included to ensure consistency in competitive conditions. Matches played outside the top-lane role were excluded.

Two time windows were defined:

- Pre-Research Period: June 16, 2025 – September 15, 2025
- Post-Research Period: September 16, 2025 – January 31, 2026

2) Variables: The following variables were extracted per match:

- Kills, deaths, assists, and KDA
- Creep score and CS per minute
- Gold earned
- Damage dealt and damage per minute
- Game duration
- Champion played
- Match outcome (win/loss)

Metrics were normalized where appropriate to account for variation in game length.

3) Data Cleaning: Data preprocessing included removal of remake games, filtering for top-lane role consistency, handling missing or invalid values, and ensuring metric consistency across both periods.

B. Descriptive Analysis

Descriptive statistics, including mean and standard deviation, were computed for each metric across both time periods. Exploratory visualizations were used to examine distributional differences and performance trends.

C. Correlation Analysis

Pearson and Spearman correlation coefficients were calculated to evaluate relationships between individual performance metrics and match outcomes, identifying key performance indicators associated with winning, consistent with prior esports analytics studies [2], [4].

D. Hypothesis Testing

Independent two-sample t-tests were conducted to compare mean performance metrics between the pre-research and post-research periods.

For each metric, the following hypotheses were tested:

Null Hypothesis (H_0): There is no statistically significant difference in the mean value of the performance metric between the pre-research and post-research periods.

Alternative Hypothesis (H_1): There is a statistically significant difference in the mean value of the performance metric between the two periods.

A one-tailed independent t-test was applied when directional improvement was theoretically expected. A significance level of $\alpha = 0.05$ was used. Assumptions of independence and approximate normality were evaluated prior to hypothesis testing.

The pre-research period included $n = 22$ ranked matches, while the post-research period included $n = 163$ ranked matches.

IV. Results

This section presents the empirical findings derived from descriptive statistics, visual analysis, correlation analysis, hypothesis testing, and predictive modeling. Results are reported by comparing performance metrics between the pre-research and post-research periods, followed by an evaluation of their relationship with match outcomes.

A. Descriptive Performance Comparison

Descriptive statistics indicate consistent directional improvement across several performance metrics in the post-research period.

TABLE I
Comparison of Mean Performance Metrics Between Periods

Metric	Pre-Research	Post-Research	Change
CS per Minute	5.85	6.08	+0.23
KDA	2.99	3.30	+0.31
Damage Dealt	21,608	23,287	+1,679
Damage Taken	35,105	33,529	-1,576
Win Rate (%)	40.9%	49.7%	+8.8%
Gold Earned	11,844	11,442	-402

Table I shows improvements in CS per minute, KDA, damage dealt, and win rate, alongside reductions in damage taken. Although improvements are observable at the descriptive level, variability across matches remains substantial.

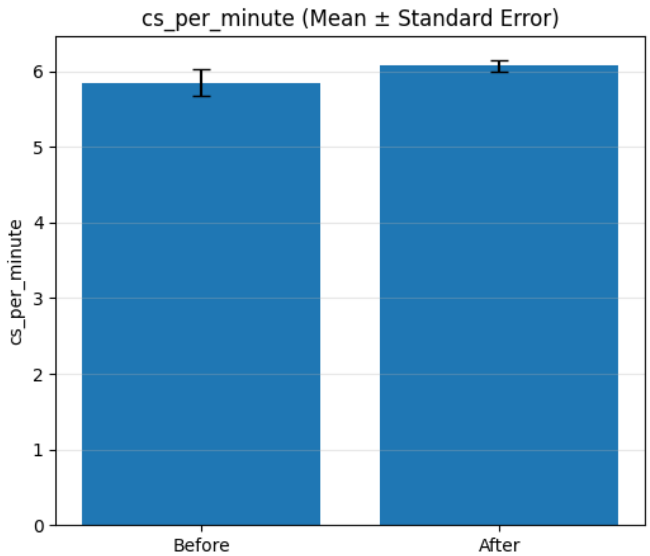

Fig. 1. Average creep score per minute before and after the project intervention. Error bars represent standard error.

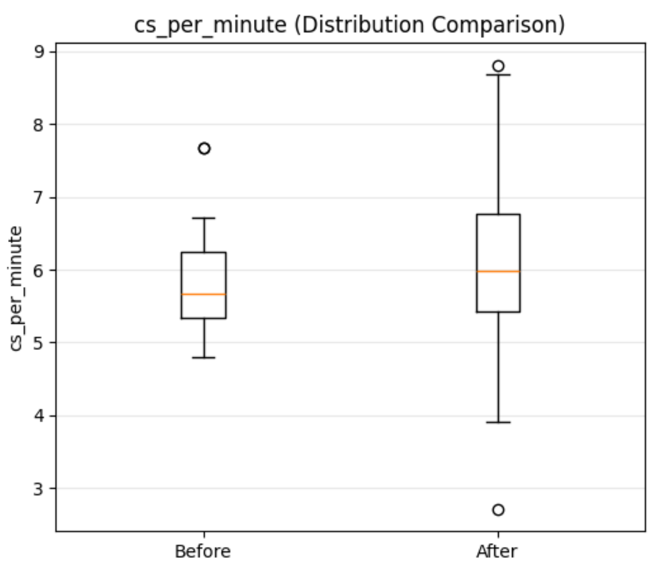

Fig. 2. Boxplot comparison of CS per minute between periods.

Figures 1 and 2 show average CS per minute. Compared to raw CS, this metric exhibits a clearer upward shift in the post-research period, suggesting improved laning efficiency independent of match duration. Nonetheless, overlapping variability indicates that performance gains were not uniformly sustained.

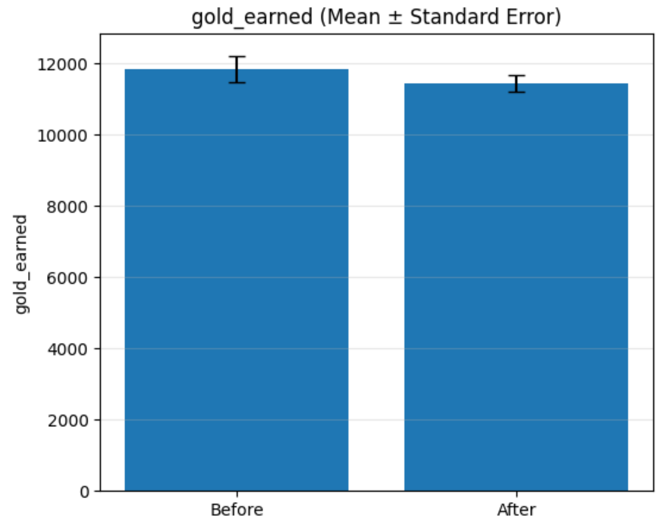

Fig. 3. Average gold earned per match before and after the project intervention. Error bars represent standard deviation.

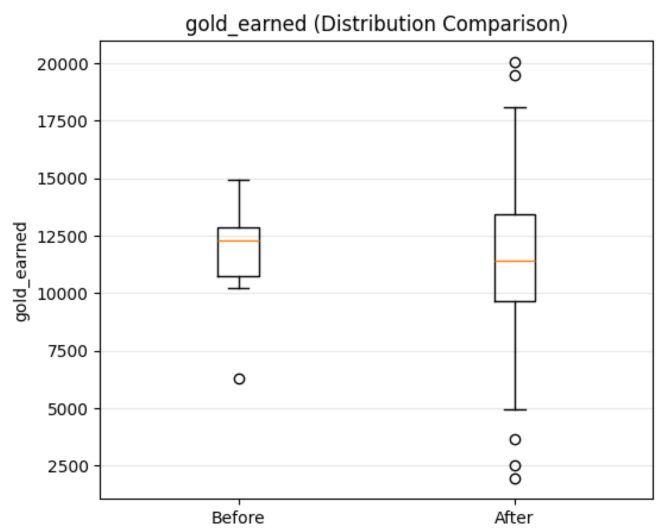

Fig. 4. Boxplot comparison of gold earned between periods.

Figures 3 and 4 present average gold earned per match. A slight decrease in mean gold is observed in the post-research period. This suggests that improvements in farming efficiency did not consistently translate into higher overall economic advantage, potentially due to shorter match durations or team-level resource distribution. The wide standard deviation highlights substantial match-to-match economic variability. Despite the slight decrease in mean gold earned, distributional spread remains similar across periods.

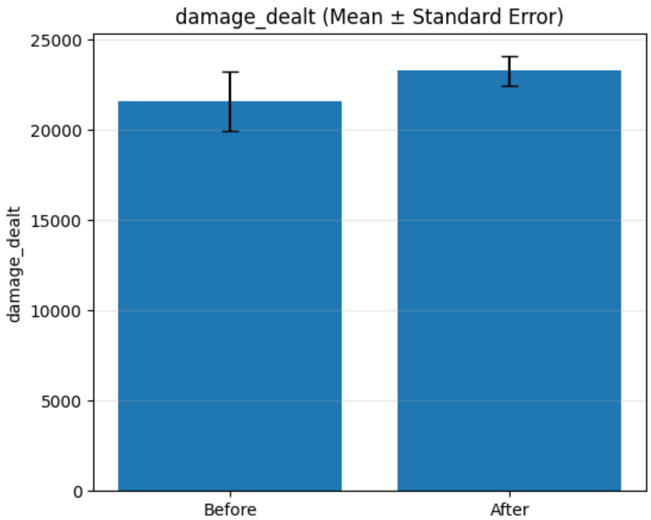

Fig. 5. Average total damage dealt before and after the project intervention. Error bars represent standard deviation.

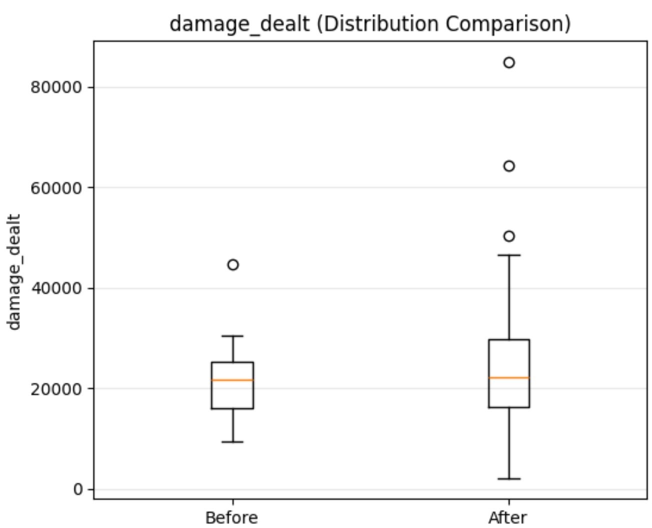

Fig. 6. Boxplot comparison of damage dealt between periods.

Figures 5 and 6 show an increase in average damage dealt during the post-research period, indicating greater engagement in combat and objective-related encounters. The large standard deviation suggests that damage output varied significantly across matches, influenced by champion selection and team composition. Damage dealt exhibits a higher median and slightly expanded upper quartile range in the post-research period, suggesting improved combat contribution in high-impact matches.

Fig. 7. Average KDA before and after the project intervention. Error bars represent standard deviation.

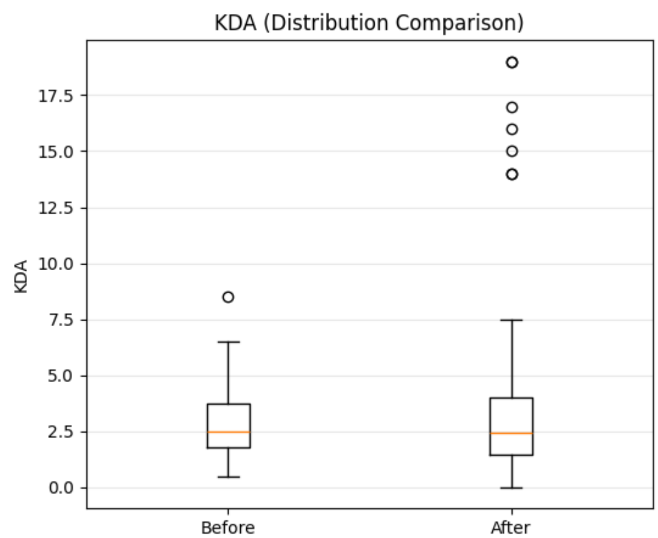

Fig. 8. Boxplot comparison of KDA between periods.

Figures 7 and 8 compare average KDA values between periods. A modest increase in mean KDA is observed post-research, indicating slightly improved combat effectiveness. However, high variability suggests that performance remained strongly dependent on match-specific factors such as team coordination. The boxplot shows a modest upward shift in median KDA during the post-research period. However, considerable overlap in interquartile ranges indicates that improvements were not uniformly sustained across matches.

B. Correlation Analysis with Match Outcomes

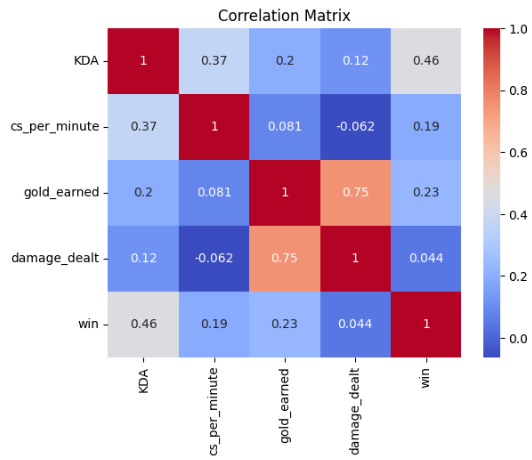

Fig. 9. Correlation heatmap illustrating relationships between performance metrics and match outcomes.

Correlation analysis examined relationships between individual performance metrics and match outcomes. Figure 9 shows that economy-related metrics, particularly CS per minute and gold earned, exhibit the strongest positive correlations with winning outcomes. Damage-related metrics demonstrate moderate correlations, while KDA shows weaker and less consistent associations.

The correlation structure remains stable across time periods, indicating that although absolute performance levels changed modestly, the relative importance of key performance indicators remained consistent. The heatmap reveals that economy-related metrics exhibit the strongest positive correlations with match victories, supporting the importance of sustained resource acquisition and lane pressure in determining match success for the top-lane role.

Damage-related metrics show moderate positive correlations, indicating that higher damage output contributes to winning outcomes, though less consistently than economic indicators. In contrast, KDA exhibits weaker and less stable correlations with win outcomes, suggesting that kill participation and survivability alone are insufficient predictors of success.

C. Champion-Specific Performance Changes

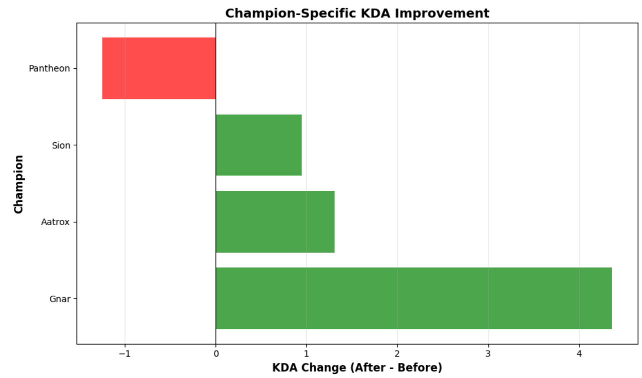

Fig. 10. Champion-specific KDA improvement between periods.

Figure 10 shows analysis of the four most frequently played champions, indicating that KDA improvement was not uniform across champion selections. Certain champions demonstrated substantial increases in average KDA, while others showed minimal change. This suggests that observed improvement may be partially champion-dependent rather than universally applied across all matchups.

D. Hypothesis Testing

TABLE II
HYPOTHESIS TESTING RESULTS (ONE-TAILED INDEPENDENT T-TEST, $\alpha = 0.05$)

Metric	Mean Before	Mean After	t-Statistic	p-Value	Decision
Win Rate	0.409091	0.496933	0.768816	0.224345	Fail to Reject H_0
CS per Minute	5.846818	6.075399	1.209553	0.118002	Fail to Reject H_0
KDA	2.992273	3.303313	0.613781	0.271529	Fail to Reject H_0
Gold Earned	11844.363636	11441.889571	-0.910178	0.815930	Fail to Reject H_0
Damage Dealt	21608.318182	23286.907975	0.917459	0.182740	Fail to Reject H_0
CS	182.772727	185.288344	0.345023	0.366108	Fail to Reject H_0
Damage Taken	35104.727273	33528.595092	-0.656015	0.741811	Fail to Reject H_0

Table II presents hypothesis testing results. None of the tested metrics showed statistically significant improvement at the $\alpha = 0.05$ level, leading to failure to reject the null hypothesis for all performance indicators.

V. DISCUSSION

A. Interpretation of Performance Trends

The boxplot visualizations reveal modest upward shifts in median values for laning-related metrics, including CS and CS per minute. Although mean gold earned decreased slightly, median values increased, suggesting possible distributional skew or the influence of extreme values in the pre-research period.

These trends indicate incremental improvement in early-game efficiency and resource acquisition, which are widely recognized as core competencies for the top-lane role. However, substantial overlap in metric distributions across periods indicates persistent variability in match outcomes and individual performance.

Damage dealt also exhibits modest improvement, potentially reflecting improved positioning or more effective participation in mid- and late-game team fights. In contrast, damage taken decreased while KDA improved, suggesting gains in combat efficiency rather than simply increased aggression.

The most substantial descriptive change occurred in win rate, increasing from 40.9% to 49.7%. Although not statistically significant, this magnitude of change may carry practical competitive relevance.

B. Champion-Specific Development

Champion-level analysis suggests that improvement was not uniformly distributed across all played champions. Performance gains appear stronger for specific champion selections, indicating that mastery progression may occur through specialization rather than global skill elevation.

C. Implications for Competitive Improvement

Despite the lack of statistically significant results, the observed trends offer practical insights for competitive improvement. The consistent association between economy-related metrics and match outcomes, as revealed by correlation analysis and predictive modeling, emphasizes the importance of laning fundamentals for top-lane success, aligning with findings from prior esports performance studies [2], [4]. Incremental gains in CS efficiency and gold generation may contribute meaningfully to long-term performance improvement, even if short-term statistical evidence remains inconclusive.

Correlation analysis consistently identified economy-focused metrics—particularly CS per minute and gold earned—as the strongest predictors of match success. These findings reinforce the strategic importance of laning fundamentals for top-lane performance. Although statistical significance was not achieved, consistent directional improvement across multiple metrics suggests incremental development. Performance growth in competitive gaming may be gradual and cumulative rather than abrupt or easily detectable within short time windows.

D. Limitations

Several limitations must be acknowledged. This study focuses on a single player, limiting the generalizability of findings to broader player populations. Additionally, the analysis does not explicitly control for confounding variables such as opponent rank, champion matchups, or patch changes, all of which may influence performance metrics.

Sample size constraints further limit the statistical power of hypothesis tests, increasing the likelihood of Type II errors. Finally, reliance on high-level match statistics may overlook lower-level behavioral factors, such as mechanical input efficiency or map awareness, that also contribute to competitive success.

E. Methodological Reflections

From a methodological perspective, this study demonstrates the feasibility of applying data science techniques to personal esports analytics. The combination of standardized data collection, longitudinal comparison, and multiple analytical approaches provides a robust framework for self-evaluation. However, future analyses may benefit from incorporating effect size measurements, time-series modeling, or mixed-effects models to better account for variance and contextual factors.

VI. Conclusion

This study applied data science techniques to evaluate longitudinal competitive performance in a top-lane League of Legends player by comparing gameplay metrics across two consecutive time periods. Using match data retrieved from the Riot Games API, the analysis examined economy-, combat-, and outcome-related metrics through descriptive statistics, correlation analysis, and hypothesis testing.

Descriptive trends suggest modest improvements in laning efficiency and combat contribution; however, inferential statistical tests did not detect statistically significant mean differences between periods. These findings highlight the challenges of measuring performance growth in high-variance, team-based competitive environments.

A. Future Work

Future work may extend this analysis by incorporating larger datasets, additional players, or longer time horizons to improve statistical power and generalizability. Advanced methods such as time-series analysis, mixed-effects modeling, or effect size estimation may also provide deeper insight into performance trends. Additionally, integrating contextual features such as champion matchups, opponent rank, and patch versions could further refine performance evaluation. Overall, this study demonstrates the viability of data-driven self-analysis as a tool for evidence-based competitive improvement in esports.

References

[1] H. Lee, J. Kim, and S. Park, "Characterizing and Quantifying Expert Input Behavior in *League of Legends*," *Elsevier Journal of Information Processing*, Yonsei University, 2024.

[2] A. Martinez, L. Chen, and R. Gupta, "Deep Learning Techniques for Identifying Key Performance Indicators in *League of Legends*: Win Prediction, Map Navigation, and Vision Control," *Procedia Computer Science*, vol. 225, pp. 112–121, 2025.

[3] Riot Games, "Riot Games Developer API," 2025. [Online]. Available: https://developer.riotgames.com/

[4] T. Schubert, J. Drachen, and M. Mahlmann, "Esports Analytics Through Encounter Detection," in *Proc. IEEE Conf. Computational Intelligence and Games*, 2016, pp. 1–8.