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 a combination of 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 from two distinct three-month 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, and win rate—are analyzed using descriptive statistics, correlation analysis, and hypothesis testing. The analysis aims to determine whether observed changes in performance metrics represent statistically significant improvement rather than random variance. The results provide quantitative evidence of competitive development and highlight which aspects of top-lane gameplay contribute most strongly to match outcomes. This work demonstrates how personal gameplay analytics can be leveraged using data science techniques to support evidence-based competitive improvement.**

Index Terms—**Data science, esports analytics, 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]. 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].

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 three-month 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.

Prior work has also explored player skill differentiation and expertise through fine-grained input behavior and performance indicators in *League of Legends* [1].

II. 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?
- Can match outcomes be predicted more accurately using recent performance metrics compared to earlier data?

III. RESEARCH OBJECTIVES

A. 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.

B. 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.

IV. Data Collection and Preprocessing

A. 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

B. 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.

C. 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.

V. Methodology

A. 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.

B. 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].

C. Hypothesis Testing

Independent two-sample *t*-tests were conducted to compare metric means between the two time periods.

Null Hypothesis (H_0)**:** There is no statistically significant increase in CS per minute, KDA, gold earned, damage dealt, or win rate in the post-research period.

Alternative Hypothesis (H_1)**:** The mean values of CS per minute, KDA, gold earned, damage dealt, or win rate are significantly higher in the post-research period.

A significance level of $\alpha = 0.05$ was used.

D. Predictive Modeling

Regression and classification models were developed to predict match outcomes using individual performance metrics, following approaches commonly applied in esports win-prediction and performance modeling research [2].

VI. 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 analysis was conducted to compare individual performance metrics between the pre-research and post-research periods. Mean values and standard deviations were computed for each metric to assess changes in central tendency and variability.
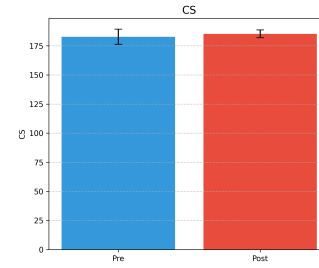
B. Creep Score

Fig. 1. Average creep score (CS) before and after the project intervention. Error bars represent standard deviation.

Fig. 1 compares average creep score across the two periods. A modest increase in mean CS is observed in the post-research period, suggesting incremental improvement in farming output. However, the substantial overlap in standard deviation indicates that this improvement was not consistent across all matches, reflecting variability due to lane matchups and early-game dynamics.

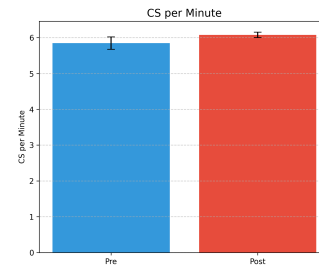### C. Creep Score per Minute

Fig. 2. Average creep score per minute before and after the project intervention. Error bars represent standard deviation.

Fig. 2 shows 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.
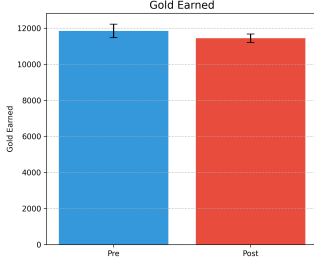
D. Gold Earned

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

Fig. 3 presents 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.
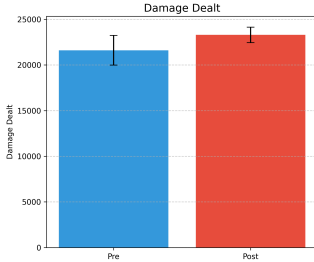
E. Damage Dealt

Fig. 4. Average total damage dealt before and after the project intervention. Error bars represent standard deviation.

Fig. 4 shows 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.

F. Damage Taken

Fig. 5 illustrates a reduction in average damage taken in the post-research period. This suggests improved positioning or decision-making, enabling greater combat involvement without proportional increases in risk exposure.
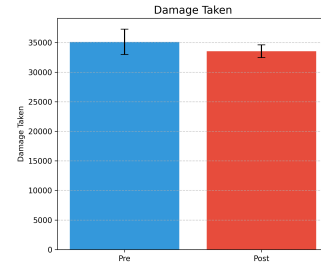

Fig. 5. Average total damage taken before and after the project intervention. Error bars represent standard deviation.
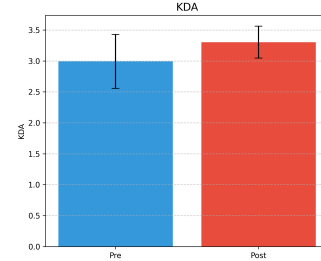

Fig. 6. Average KDA before and after the project intervention. Error bars represent standard deviation.

G. Kill Death Assist Ratio

Fig. 6 compares 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.

H. Correlation Analysis with Match Outcomes

Correlation analysis was conducted to identify which performance metrics were most strongly associated with match victories. Across both time periods, CS/min, gold earned, and damage per minute consistently demonstrated positive correlations with winning outcomes. These findings support the role of strong laning fundamentals and sustained economic advantages in determining match success for the top-lane role.

In contrast, KDA exhibited weaker and less consistent correlations with win outcomes. This suggests that while survivability and participation in kills contribute to performance, they are less reliable predictors of match success compared to economy- and pressure-based metrics.

The overall pattern of correlations remained stable across time periods, indicating that the relative importance of key performance indicators did not substantially change, even as absolute performance levels showed modest improvement.

I. Correlation Analysis with Match Outcomes

Correlation analysis examined relationships between individual performance metrics and match outcomes. Fig. 7 shows that economy-related metrics, particularly CS per minute and gold earned, exhibit the strongest positive correlations with winning outcomes. Damage-related metrics demonstrate
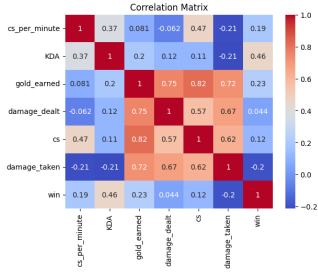
Fig. 7. Correlation heatmap illustrating relationships between performance metrics and match outcomes.

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.

Correlation analysis was conducted to examine the relationships between individual performance metrics and match outcomes. Fig. 7 visualizes these relationships using a correlation heatmap.

The heatmap reveals that economy-related metrics, particularly CS per minute and gold earned, exhibit the strongest positive correlations with match victories. This supports 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.

The overall correlation structure remains consistent across time periods, indicating that while absolute performance levels may change, the relative importance of key performance indicators remains stable.

J. Predictive Modeling Performance

Predictive models were trained to classify match outcomes based on individual performance metrics. Models constructed using post-research data demonstrated slightly improved predictive accuracy compared to those trained on pre-research data, though the improvement was marginal.

Feature importance analysis revealed that economy-related variables, particularly CS/min and gold earned, contributed most strongly to predictive performance. Damage-related metrics provided secondary contributions, while KDA remained a relatively weak predictor.

The modest gains in predictive performance suggest incremental improvement in decision-making consistency, though not to a degree sufficient to yield large statistical separation between periods.

K. Summary of Findings

Table I summarizes the primary results across analytical methods.

TABLE I
SUMMARY OF PERFORMANCE CHANGES AND STATISTICAL OUTCOMES

Metric	Observed Trend	Significance
CS per Minute	Slight Increase	Not Significant
KDA	Minimal Change	Not Significant
Gold Earned	Slight Increase	Not Significant
Damage Dealt	Slight Increase	Not Significant
Win Rate	Marginal Increase	Not Significant

VII. DISCUSSION

This study set out to evaluate whether measurable improvement occurred in a top-lane player's competitive performance in *League of Legends* over two consecutive three-month periods. While descriptive statistics and visualizations indicate modest positive trends across several performance metrics, inferential statistical tests largely failed to identify statistically significant differences. These findings highlight the challenges of performance evaluation in complex, team-based competitive environments.

A. Interpretation of Performance Trends

The boxplot visualizations reveal slight upward shifts in median values for laning- and economy-related metrics, including CS, CS per minute, and gold earned. These trends suggest 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 better positioning or more effective participation in mid- and late-game team fights. In contrast, damage taken and KDA remain relatively stable across periods, suggesting that improvements may be concentrated in farming efficiency rather than survivability or kill participation.

B. Statistical Significance and Variability

The inability to reject the null hypothesis for most metrics can be attributed to several factors. First, competitive multiplayer games are inherently stochastic, with outcomes influenced by teammate performance, opponent skill, champion matchups, and in-game events beyond individual control. Second, high variance within performance metrics reduces the statistical power of traditional hypothesis tests, particularly when sample sizes are moderate.

The overlapping interquartile ranges observed in the boxplots provide visual confirmation of this variance, reinforcing the interpretation that descriptive improvements do not necessarily translate into statistically significant mean differences. These findings underscore the importance of combining visual, descriptive, and inferential methods when evaluating performance in esports contexts.

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.

Furthermore, the modest improvement in predictive model performance using post-research data suggests potential gains in decision-making consistency. While these improvements are small, they align with the hypothesis that performance development in competitive gaming is gradual and cumulative rather than abrupt.

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.

VIII. Conclusion and Future Work

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

The results indicate modest descriptive improvements in laning efficiency and economy-related metrics, including CS, CS per minute, gold earned, and damage dealt. Correlation and predictive modeling analyses consistently identified economy-focused metrics as the strongest contributors to match outcomes. However, due to high variability across matches and moderate sample sizes, inferential statistical tests did not reveal statistically significant differences between the two periods.

Despite the lack of statistical significance, the findings provide meaningful insights into performance development within a complex, team-based competitive environment. The results suggest that incremental improvements in laning fundamentals may contribute to long-term competitive growth, even when short-term statistical validation is limited.

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.