

# Is League of Legends Slowly Dying?

## Github Repository Link:

[Lucaszhuokaiyu/data\\_engineer\\_project \(github.com\)](https://github.com/Lucaszhuokaiyu/data_engineer_project)



Lucas Yu

# Elevator Pitch



**Who I Helped:** Stakeholders and game developers



**Problem:** Limited visibility into new player retention and engagement



**Solution:** Built a fully automated data pipeline with SQL analysis and actionable insights using scraped and API-based player data



# Job Description

Design and implement advanced ETL process and data Integration strategies to handle diverse data sources and formats.

## Job Description

Riot's Data Engineers build products that improve the experience of our players worldwide using petabytes of data and state-of-the-art processing technology. Maximizing the potential of this data is a complex task. The League Data Central team is made up of engineers, data scientists, producers, and product managers focused on evolving the data powered capabilities of the League Studio.

As a Senior Data Engineer on the League Data Central team, you will help shape the data architecture and systems that provide amazing experiences to millions of League of Legends and Teamfight Tactics players around the world. You'll have the opportunity to work on projects covering a wide variety of problems spaces including player facing features powered by machine learning, improving game design with data focused tooling, and partnering with our Insights teams to help us make key strategic decisions and drive future planning.

### Responsibilities:

- Design and implement architectures that support advanced analytics, machine learning models, and real-time data processing, ensuring scalability and performance
- Develop complex data models that cater to both current and future needs focused on accessibility for large datasets
- Design and implement advanced ETL processes and data integration strategies to handle diverse data sources and formats
- Collaborate closely with game teams as well as central platform teams to understand their needs and provide solutions
- Analyze existing systems for bottlenecks and inefficiencies, and implement solutions to optimize data flow and processing for high-volume, complex datasets.
- Work on initiatives in data governance, data quality, and metadata management that ensure data accuracy, consistency, and security
- Explore and evaluate new technologies, tools, and data management practices to enhance the capabilities of the data platform
- Collaborate closely with data scientists and analysts, providing them with the data required for complex analytics, machine learning models, and data exploration
- Communicate complex data concepts and the value of data projects to stakeholders at all levels of the organization, influencing data-driven decision-making and strategy

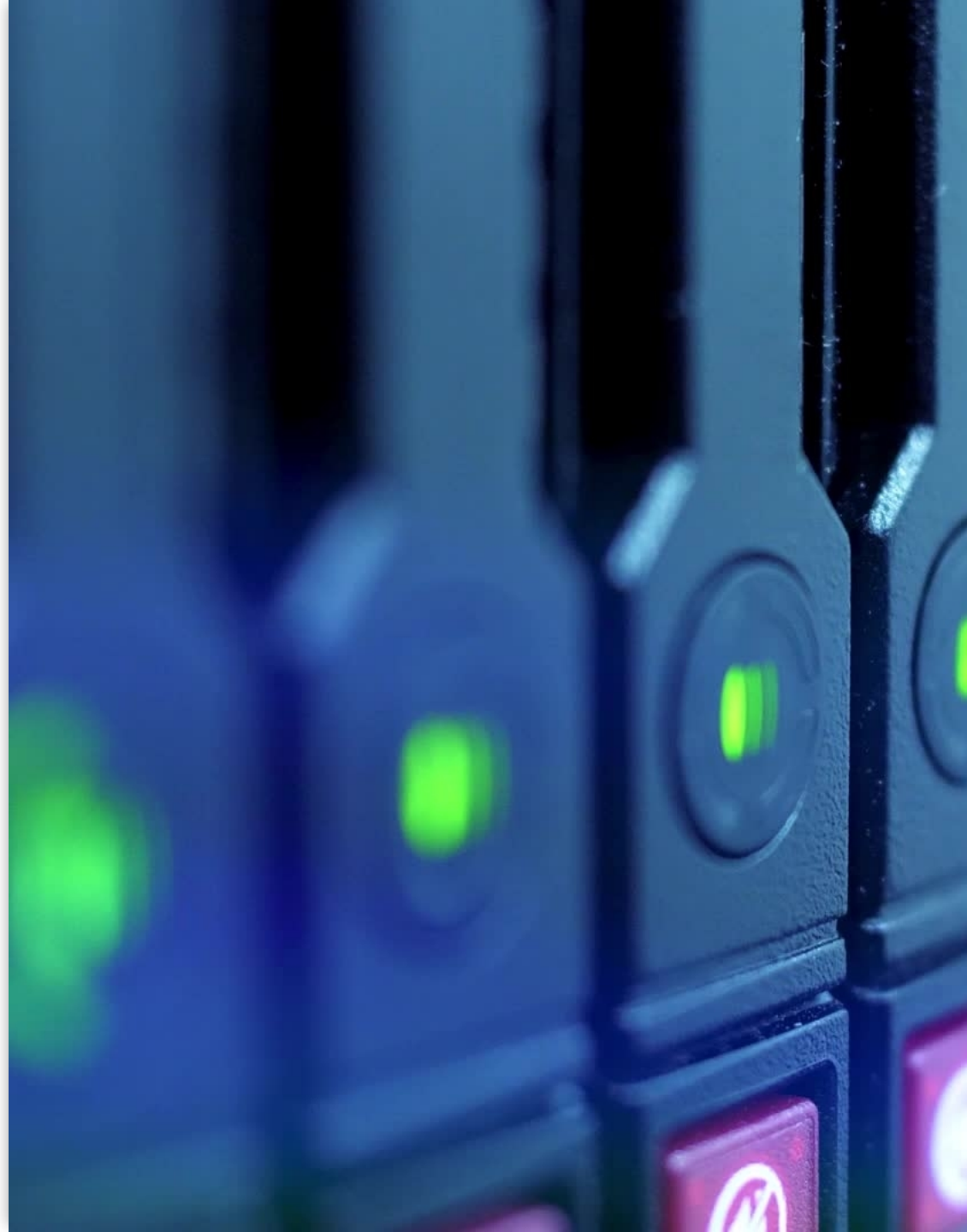
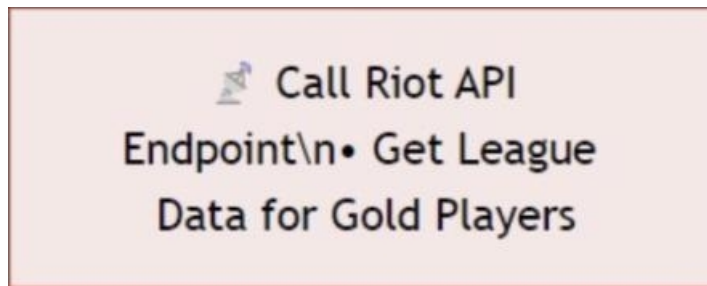
# API Data Sources

Source name: Riot Games API (Gold Ranked Players)

Data collected: Player\_type, player\_count, win\_loss\_ratio, hotStreak, player\_count, avg\_win\_loss\_ratio

Job relevance: handle diverse data sources and formats

Associated data pipeline diagram:



---

# New Players do not lose more than experienced players!

Business Question: How do win/loss performance patterns differ between new and experienced players?

Recommendation: Improve new player onboarding and matchmaking systems to reduce early losses and increase engagement.

Prediction: If new players experience more competitive matches early on, retention and conversion into experienced users will likely improve.

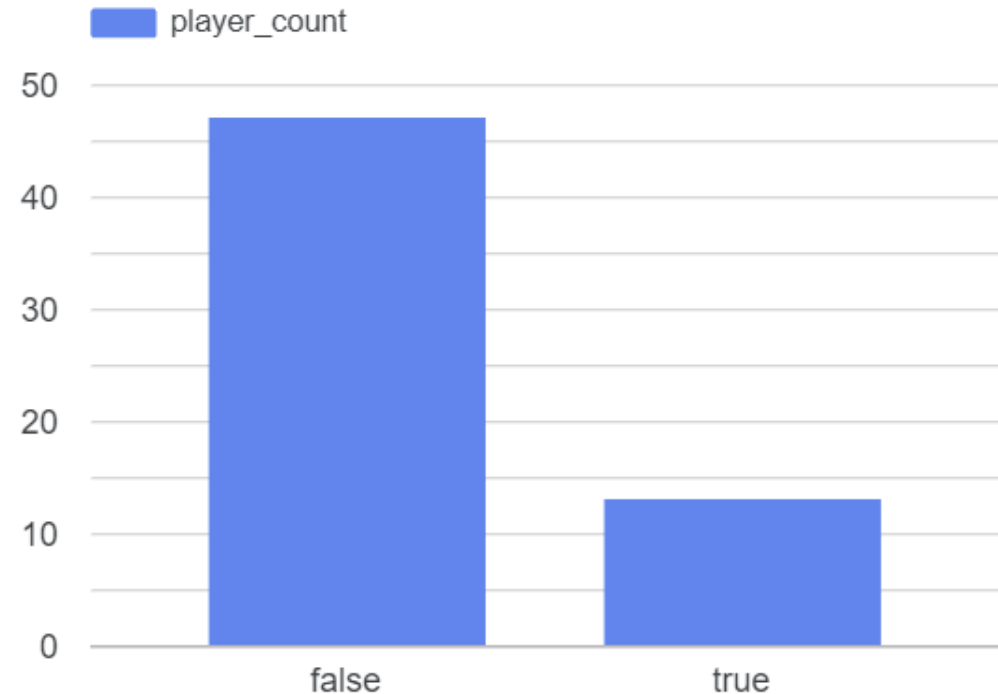


## 27.6% of new players are highly valuable!

Business Question: Does being on a hot streak correlate with better performance and possibly better retention among new players?

Recommendation: Identify and reward new players who enter hot streaks with in-game incentives to encourage continued play.

Prediction: Supporting new players during winning streaks can boost their emotional investment and prolong their gameplay lifecycle.




# Web-Scraped Data Source

Source name:steamcharts.com

Data collected: Perios, avg\_players, avg\_peak\_players, Month, avg\_to\_peak\_ratio

Job relevance: handle diverse data sources and formats

Associated data pipeline diagram:

 Scrape HTML using BeautifulSoup\n• Extract Avg/Peak Player Counts\n• Extract Monthly Gain/Loss





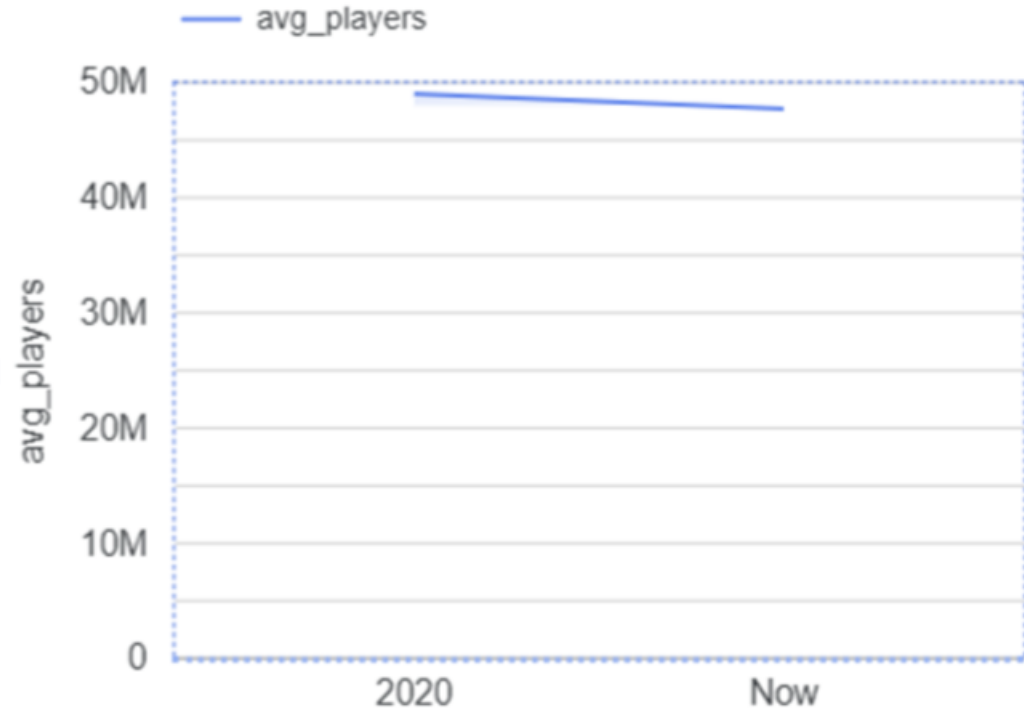
---

# Average Player Decreased by about 13% !

Business Question: How has player activity changed over time between the high-engagement period of 2020 and the current period?

Recommendation: If the current numbers are lower, consider releasing major updates, reactivation campaigns, or seasonal events to restore player interest.

Prediction: Without strategic updates or marketing interventions, the game may continue to see reduced average monthly players despite retaining peaks during specific events.





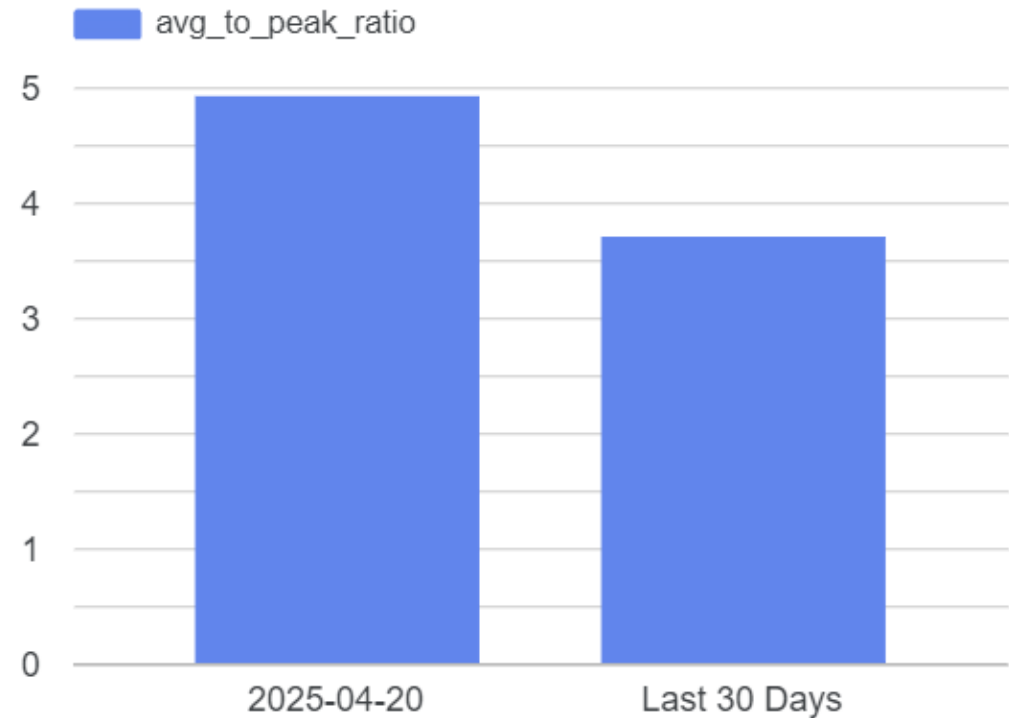
---

## Shift from day-to-day engagement to event-driven spikes!

Business Question: Is the game sustaining consistent player activity, or is it becoming more dependent on short-term peak events?

Recommendation: If the ratio is falling, invest in features that drive habitual engagement to stabilize the average base.

Prediction: A continued drop in the average-to-peak ratio may lead to greater volatility in game population and higher risk of churn during non-event periods.



# Conclusion

As a candidate for a data engineer role in the gaming industry, I focused on what matters most—new player behavior, engagement trends, and retention drivers. This aligns directly with the job's emphasis on using data to improve player experiences and guide product decisions.

What the Data Showed:

1. New players experience in this game is relatively good
2. Player base has slightly decreased throughout the past 5 years
3. The game is driven more by seasonal events rather than consistent engagement.