MSc Data Analysis for Business Intelligence  
MA7092 – Data Analytics for Esports

**PROJECT PLAN – League of Legends**

## **ABSTRACT**

The famous multiplayer online game League of Legends is the subject of this study’s analysis of game statistics. For providing an insight, the champions are analysed for their performance and effectiveness and are ranked based on the factors related to these. With several patches released for the game over the years, for some cases, the same analytical process is performed with the gaming data before any sort of comparative conclusion is made. The target stakeholder that the analysis will concentrate on would be the fresh players in-order to help them win more games, as well as esports commentators in-order to provide a better champions past. This will contribute to the cultural value and symbolic value subsequently.

This analysis can only be possible with a proper source of data for the game, which in this case is League of Legends 2014-2022 combined data from Oracles Elixir, exported from the google drive. In-order to produce the meaningful and desired results with visualizations we would use Python programming language, R programming language and Microsoft excel.

Table of Contents

COVER PAGE ................................................................................................................................................1

[ABSTRACT 1](#_Toc129810969)

[TOPIC 3](#_Toc129810970)

[SOURCE OF DATA 4](#_Toc129810971)

[CONTEXT – ECOSYSTEM & STAKEHOLDERS 6](#_Toc129810972)

[MINIMUM VIABLE PRODUCT – 3 QUESTIONS & DATAPOINTS EXPLAINED 7](#_Toc129810973)

[DEVELOPMENT NEEDS ANALYSIS 9](#_Toc129810974)

[REFERENCES 10](#_Toc129810975)

# **TOPIC**

League of Legends (LoL), a “multiplayer online battle arena[3] falls under the real-time tactical and strategical category.” In this game, “the players control a single figure, referred to as a champion"[3] and battle with other teams of champions on a symmetrical battleground. The goal is to demolish the base building of the opposing squad, which is typically situated on the other side of the map. Along with quick action and intense fights, the game frequently requires players to make smart decisions and work together as a team.

‘On October 27, 2009, League of Legends (LoL) was initially published for Microsoft Windows”[2]. The game was designed by “Brandon Beck and Marc Merrill”[2], who established “Riot Games”[4] in “2006”[2]. “Riot Games is a video game company with headquarters in Los Angeles, California.”[4] Since its debut, League of Legends has experienced several updates and expansions, introducing new champions, game modes, and gameplay adjustments to keep the game interesting and fun for its millions of players worldwide. Since then, it has grown to be among the most well-known and promising multiplayer games in the entire globe.

In furthermore detail about the game, the players take on the part of a summoner who commands a champion with special abilities and engages in combat with another champion squad in a variety of game modes, with the  goal of destroying the rival team's core, a defensive building within their base. Depending on the mode and the players ability level, each game usually lasts “20 to 50 minutes.”[5]

Players guide their champion using a mix of keypad and mouse commands in the top-down perception of the game. To obtain an early advantage and win, players must outsmart and destroy their opponents, work as a team, and use their skills strategically and making ultimate use of their champion’s distinctive skills set and playstyles. Players can even make use of their in-game cash or spend real money on skins to alter and modify their champion’s look and skill set.

Keeping this in mind, our research mainly focuses on analysing such champions from the gaming data source that can prove to be handier with outstanding performance in causing the damage to the opponent team player’s champion, effectiveness in attaining the creepscore and comparing the damage incurred by the champions over the last few years after receiving gaming updates or applying new patches.

With the popularity that the game achieved over the years, League of Legends professional leagues and competitions were introduced and conducted all over the globe with millions of dollars in cash prizes for the winners.

# **SOURCE OF DATA**

For performing an analysis, the first thing that we would need is collecting data for it. With proper data, insightful information on the game's balance, user behaviour, esports success, and commercial aspects can be offered.

The source of data that I have used in the League of Legends (LoL) gaming analysis is google drive which provided me a combined dataset with the gaming statistics from the year 2014-2022. The link to this source is [https:League of Legends 2014-2022 combined data](https://drive.google.com/file/d/1EPZ5CIBJrXtloMNMVec7WoOgNbFxcH1w/view). To extract the data, one can just click on the link provided that would take you to the google drive and request you to download the excel sheet file containing the gaming statistics in a simpler way to understand and analyse. This data source suffices quite sufficient and insightful information required to answer my analytical questions, and the data appears to be taken from the League of Legends (LoL) API (Application Programming Interface) that can be viewed through the “Riot Games”[3] developer site.

The option of extracting the data from its API manually is also possible. For this, we require to register their official programme by setting up an account initially and then get an API key. Using various programming languages we can make API calls to the endpoints using the API key. Different endpoints provide distinct types of data, such as champions performance, banned champions and game events like kills, death, total golds earned etc. The League of Legends demo files and computer vision are the additional sources of data, but their use is not as widespread as the API's. Demo files need a particular software programme to view and evaluate them, whereas computer vision might need more sophisticated tools and resources. But I feel it would be a complicated process to get an accurate, sufficient, and balanced data for our analysis from the API or its demo or from computer and we might require conducting the data cleaning task as well in-order to remove the duplicate and unnecessary data which is also time consuming and require meticulous attention.

Considering the above factors I decided to proceed with the data source provided in the google drive as it provides the specific datapoints that would help me in analysing and answering the questions selected related to the game.

For the first question, I need to analyse the Damage taken per minute by the champions and how it is related to their creep score attained, so the datapoints to be used are “champion”, “damagetakenperminute” and “total cs” where the “champion” datapoint defines the player-controlled character selected to play as in a game and who has a special skill set and team position to fulfil, “damagetakenperminute” datapoint defines the harm inflicted to a player's champion over the course of one minute. It maintains track on how much harm a champion takes during a match and can offer information about how well they are surviving and performing their team duties, and “total cs” refers to the total creep score that defines the number of minions and monsters a player’s champion has killed throughout the game.

Similarly, for the second question, I need to analyse the creep score results for the champions over its last updates in-order to rank them. The datapoints to be used for this are “total cs”, “champion” both of which are already defined above, and “patch” datapoint which defines the software upgrade that the game's designers and developers issue to make significant changes to the game.

And, for the third and final question, I focus on analysing the champions that appears frequently over the yearly game updates when we compare the damage to the champions. So, here the “champion”, “patch” and “damagetochampions” datapoints are used. The “damagetochampions” datapoint defines the harm a player’s champion has inflicted on rival champions throughout a match. This number is crucial because getting wins and winning team battles frequently depend on delivering damage to the opponent champions. Likewise what does the “champion” datapoint and “patch” datapoint define is already explained above.

In additional to the above mentioned datapoints that are specific to our questions, there are many other datapoints that are also given in the data that is extracted from the drive. Some of these datapoints are game events like “kills”, “deaths”, “assists”, “double kills”, “triple kills”, “earned golds”, “game length” etc, gameplayer information like their “name”, “id”, used “champion”, “banned champions” etc, team performance information like “team kills”, “team assists” etc, and more.

# **CONTEXT – ECOSYSTEM & STAKEHOLDERS**

In the ecosystem encircling professional gaming, esports data analysis has grown in significance. “A broad variety of stakeholders, including participating players, fans, and marketers, as well as game makers and esports communities”[1], are present in this ecosystem displayed in Figure 1. Our goal in conducting the esports research surrounding this ecosystem is to offer some kind of benefit or value to the stakeholders we hope to assist, such as "cultural" value, "social" value, "economic" value, and "symbolic" value.

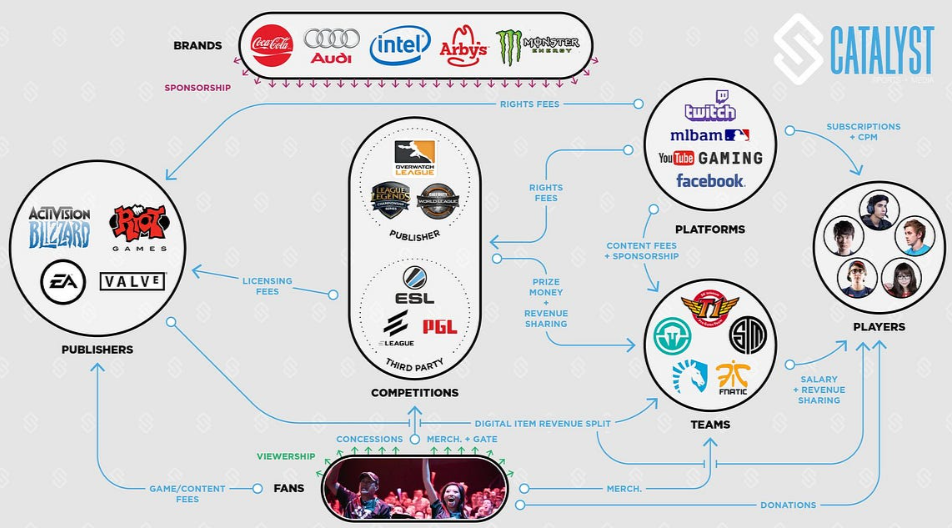


Figure 1. Esports Ecosystem

In my esports analytical research, I'm focusing on the League of Legends game. The questions I need to analyse and respond to are set up in a way that will help new or inexperienced players win more games and that will also help esports commentators provide a better champion's legacy so that players can choose a champion quickly and efficiently before playing and winning with them. This study primarily focuses on the champions in-order to correctly evaluate or grade them based on their performance ability and efficacy by taking into account the gaming security patches being applied over the years.

With this, the objective will be to boost or contribute to the cultural value by enhancing gameplayer's (new and old) expert knowledge, and to the symbolic value by enhancing the champion's rankings and outstanding quality so that more players are inclined to use it when playing.

# **MINIMUM VIABLE PRODUCT – 3 QUESTIONS & DATAPOINTS EXPLAINED**

The questions analysed in this League of Legends project concentrates mainly on the champions in order to accurately assess or grade them based on their performance ability and effectiveness while taking into consideration the security patches that have been applied to video games over time. With the chosen source of data, the following questions are framed for analysis:

1. How does Damage Mitigation per Minute relate to Creepscore, and which Champions appear most often?
2. Which Champions appear most in the top 10% highest Creepscore, and how has this changed over the last five patches?
3. Which Champions appear most in the highest and lowest 10% Damage to Champions, and how has this changed over the last five patches?

By providing answers to the aforementioned three questions, the main goal is to assist new and inexperienced players in winning more games while also strengthening the cultural value of gaming by enhancing their understanding of the use of champions in light of their performance, skill set, and effectiveness. To assist commentators or analysts in providing a more accurate account of the champion's history and triumphant accomplishments, and hence make a significant contribution to the symbolic value by raising the champion's image and exceptional quality, so that more players are compelled to use it when playing.

The datapoints to be used for the first question are "champion," "damagetakenperminute," and "total cs" as I require to analyse the Damage taken per minute by the champions and how it relates to the creep score they have achieved. The "damagetakenperminute" datapoint describes the damage a player's champion suffered in the duration of one minute, and it keeps track of how much damage a champion takes during a match as well as providing information about how well they are surviving and carrying out their team duties. The "champion" datapoint describes the player-controlled character chosen to play as in a game and who has a unique skill set and team role to perform. The term "total cs" refers to the total creep score, which refers to the total number of minions and monsters a player's champion has managed to kill throughout the game.

For the second query, I intend to rank the champions based on the outcomes of the champions creep scores over the most recent and major developments done to the game. The datapoints that will be used for this are “champion”, the player-controlled figure that is selected to play as in a game and who has a special skill set and a team duty to perform, "total cs" that defines the total creep score, which is the sum of all the monsters and minions a player's champion has killed during the course of the game, and "patch" that describes the software update that the game's creators and developers release in-order to make major changes to the game.

In order to answer the third and concluding question, I analyse the champions that have shown up most frequently over the course of the annual game updates. In this case, the datapoints "champion", "patch", and "damagetochampions" are used. The "damagetochampions" datapoint describes the injury a player's champion has caused to competing champions over the course of a battle. This figure is important because dealing harm to the rival champions is frequently necessary to secure victories and triumph in team fights. The definitions or description of the "champion" and "patch" datapoints have also already been covered in the previous section that would be sufficient to understand what they relate to.

Now, to present the analytical outcomes of each question in an efficient and effective way to our stakeholders i.e., new (fresh or inexperienced) players and commentators or analyst, I bend towards using the presentation format by providing insights, plotting some meaningful and interesting visualizations (graphs) with which important and precise information can be extracted or conclusions can be made. Also, a tabular form of the result can be shown if it expresses the outcome in the desired meaningful way.

# **DEVELOPMENT NEEDS ANALYSIS**

|  |
| --- |
| To complete this project, I need to learn Python programming language. |
| Specific features I need to know are Data preparation, Python numpy, pandas and data visualization library – matplotlib. |
| On a scale of 1 – 10, my confidence with it is 7 |
| The specific tutorials and resources I will use are…  1. <https://www.youtube.com/watch?v=a9UrKTVEeZA>  2. <https://pythonprogramming.net/data-analysis-tutorials>  3. <https://realpython.com/tutorials/data-science>  4. <https://www.youtube.com/watch?v=xQJxGaBCG5w>  5. <https://towardsdatascience.com/data-science-as-support-of-esports-performance-and-strategies-i-d511843b00a1> |
| Learning more about this feature will help me lay a solid basis in Python data analysis process like cleaning, transformation, and data visualisation, which will be helpful for a variety of uses as research project, statistical analysis, business intelligence, and much more. |

|  |
| --- |
| To complete this project, I need to learn R programming. |
| Specific features I need to know are R programming data visualization library – ggplot. |
| On a scale of 1 – 10, my confidence with it is 4 |
| The specific tutorials and resources I will use are…  1. <https://www.analyticsvidhya.com/blog/2016/02/complete-tutorial-learn-data-science-scratch>  2. <https://r4ds.had.co.nz/exploratory-data-analysis.html>  3. <https://www.youtube.com/watch?v=IYISD3dX-c0>  4. <https://www.youtube.com/watch?v=rfR9Nrpfnyg>  5. <https://www.youtube.com/watch?v=sV5lwAJ7vnQ> |
| Learning more about this feature will help me in producing high-quality and useful data visualisations that can convey your ideas and data insights clearly and accurately, as well as understanding data research, collaboration, data analysis, and improved data storytelling. |

# **REFERENCES**

[1] Alford, H. (2019). The 2017 eSports Ecosystem Explained In One Chart. [online] humble words. Available at: <https://medium.com/humble-ventures/the-2017-esports-ecosystem-explained-in-one-chart-ef6fc48b779e>

[2] Crecente, B. (2019). League of Legends Is Now 10 Years old. This Is the Story of Its Birth. [online] Washington Post. Available at: <https://www.washingtonpost.com/video-games/2019/10/27/league-legends-is-now-years-old-this-is-story-its-birth>

[3] League of Legends Wiki. (n.d.). *League of Legends*. [online] Available at: <https://leagueoflegends.fandom.com/wiki/League_of_Legends#cite_note-5> [Accessed 15 Mar. 2023].

[4] League of Legends Wiki. (n.d.). Riot Games. [online] Available at: <https://leagueoflegends.fandom.com/wiki/Riot_Games>

[5] www.leaguemath.com. (n.d.). LeagueMath --- Match duration analysis. [online] Available at: <https://www.leaguemath.com/match-duration-analysis> [Accessed 15 Mar. 2023].