***League of Legends (LCS) Performance Dashboard***

**Ivan Sheng**

**Johns Hopkins University**

**EN.605.662.84 Data Visualization Fall 2020**

**Final Project**

**Abstract:**

This paper discusses the instrumentation of data acquired on professional League of Legends games to develop a tool that can be useful for the coaching staff in order to strategize during the different phases that make up a competitive match. The differences between the developed visualizations in comparison to what is currently out in the public market is discussed, and traditional sports analytics is mentioned too since this project is heavily based on box score data visualizations. The developed tool contains a team overview page and a player versus player matchup phase, so that analysts and coaches can get a high-level outlook on matchups as well as an in-depth, position by position view too.

**Introduction:**

Competitive gaming, also known as esports, has exploded in the last few years with millions of dollars in investment into teams and infrastructure. In 2019, the *North American League Championship Series (LCS)* was ranked by Nielsen as the United States’ third-most popular professional sports league among 18-to-34-year olds based on live average minute audience1.

With games being built on computers, it’s inherent by nature that these mediums of entertainment contain more data than traditional sports. However, there’s a general lack of data expertise in this industry due to both the lack of developed infrastructure and the stigma of video games - albeit both are not as relevant factors as they were during the inception of the *LCS*.

Despite the growing popularity of the broadcasts, there’s a lack of availability (or even transparency) in the development of professional sports analytical tools. The game itself, however, is littered with sites and tools that help casual players analyze their personal performance such as, *Mobalytics* and *Leaguespy*.

What’s been developed by the traditional sports analytics field for the purposes of broadcast analysis or sports betting can be applied to esports to aid in the development of analytical tools that can appeal to coaching staff, sports betters, or enthusiasts when the focus is on professional play, instead of the general population of players.

**Background:**

In *League of Legends*, there are two important phases to strategize for:

1. the pick and ban phase, where teams determine what they will play into each other, and what champions they don’t want the other team to play. Teams first ban out three champions one by one, for a total of six unique bans. Then teams begin a pick phase where three unique champions per team are selected. The final ban phase begins with two bans per team for a cumulative total of ten unique bans. Lastly, the final pick phase begins with two unique picks per team to round out the five-versus-five matchup.
2. The actual gameplay, where teams play against each other on a five versus five map with five different assigned roles, and over 150 unique champions to choose from.

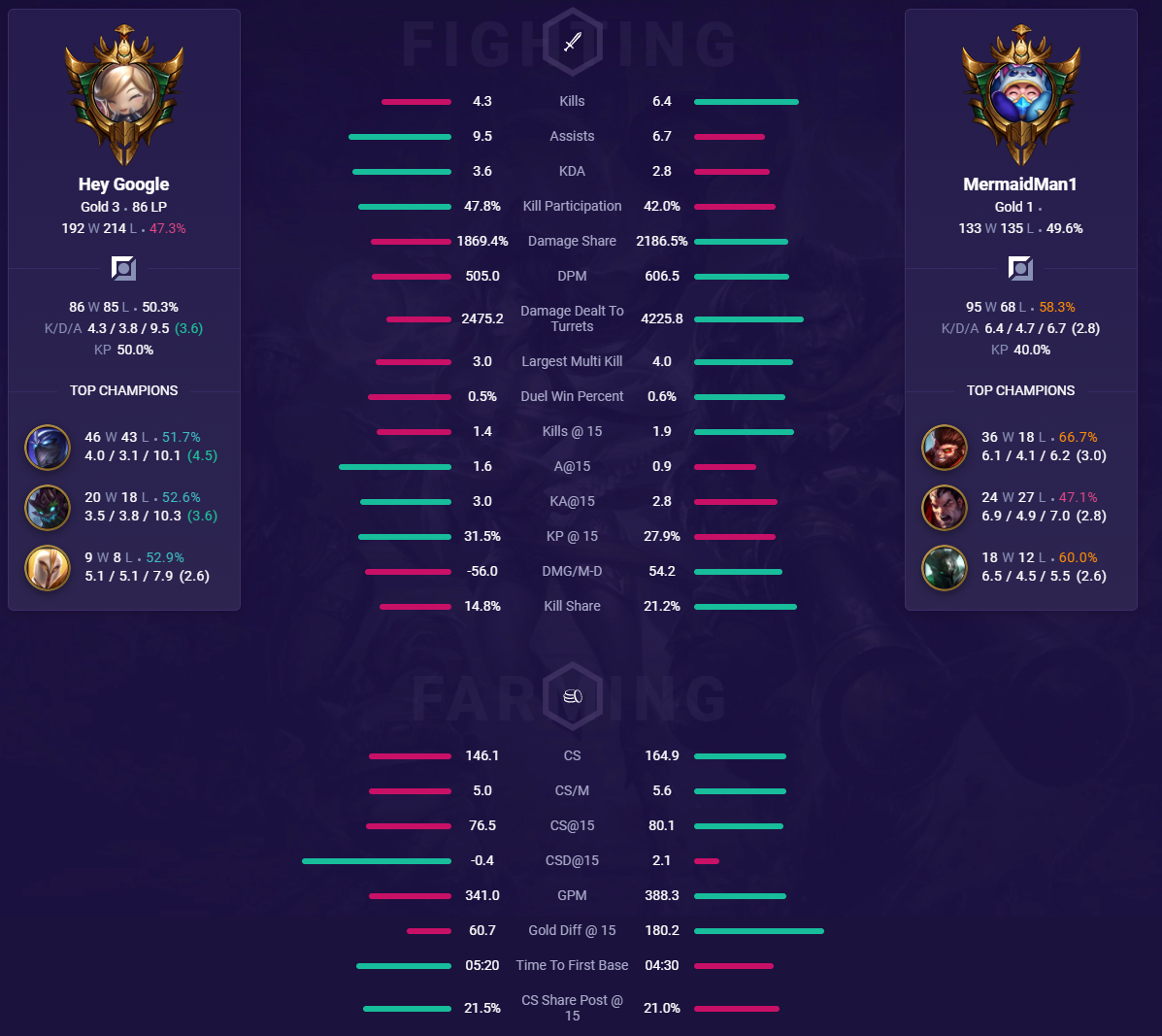
There have been plenty of attempts at producing analytical visualizations for the general public, with the over-saturated market of websites analyzing your game play. One of the most common approaches is to look at a quick overall view of how players did individually that’s mainly used for progress tracking or scouting. This can be achieved from sites such as op.gg that provide a more traditional view similar to what you would find in the in-game client, but with a more accessible search bar:



*Part of the Player Dashboard provided by na.op.gg2*

These views are fairly high-level and provide enough information for players to understand how that game went, but doesn’t provide enough or irrelevant information. This is the result of attempting to provide a universal layout to accommodate all players, but there are better metrics of performance than CS, Damage, and Kills - all of which aren’t even normalized for game time.

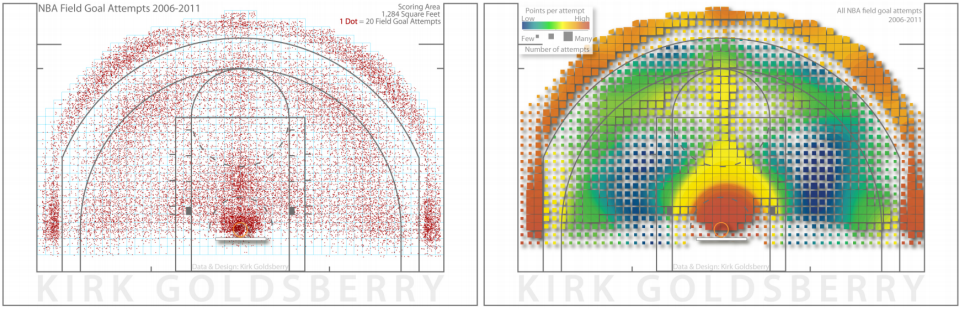
Other websites such as *Mobalytics* also provide match-up comparisons, but the quality of visualizations is poor with blanket stats that can cover all roles and play-styles, but some more than others.



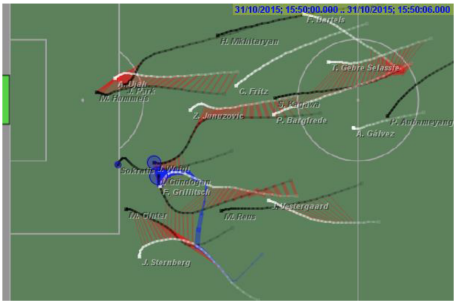
*Comparison Tool developed by Mobalytics.gg3*

Taking the two players above for example, while they play the same role, their play styles (indicated by the champions they play) are completely different. The player on the left focuses on playing champions who soak damage and act as the team’s defense, while the player on the right focuses on combat dominating champions. The approach that *Mobalytics* used to throw in all in-game performance metrics is a lazy way of comparing two players against each other.

One of the less explored realms of eSports data visualization is often tracking data. In traditional sports, tracking data can be used to visualize events in sports such as Basketball that can highlight where on the court a player took a shot, or in Soccer/Football where player pathing can be visualized.



*State of the Art of Sports Data Visualization. Computer Graphics Forum, Page 104*



*State of the Art of Sports Data Visualization. Computer Graphics Forum, Page 114*

In a paper done by students at The University of Lisbon, the group developed a tracking tool called *VisuaLeague II*, that was intended to provide insights via tracking projected onto a top-down perspective of the *League of Legends* map.



*Comparison of Visualization Tools for Matches Analysis of a MOBA Game. 2019 23rd International Conference Information Visualisation (IV), Page 1215*

This type of visualization helps to identify decisions and events at specific parts of the map. This represents macro decisions made by the team during the game, and is more helpful in understanding how strategies were implemented during snapshots of a game, whereas the previous examples focused on pre- and post-game analysis.

Unfortunately, positional data isn’t as widely available as box-score data for this game, and is even difficult to obtain. In fact, in the API documentation, the developers of the game even recommend implementing a machine learning approach to identify map position - meaning this data isn’t even natively available via the API6.

**Approach:**

ETL and EDA were performed on the 2020 professional League of Legends dataset found on the *Oracle’s Elixir* website7. The descriptive statistics can be found in the Appendix or in the provide Excel document, and the exploratory data analysis can also be found in Appendix or in the provided Jupyter Notebook file, and was developed in Python.

**T**he priorities were to first identify metrics that would be able to properly summarize a team’s or player’s performance since my target audience is mainly for the analytics professionals and the coaching staff of teams, with a secondary audience being fans of sports analytics.

The target use-case I had in mind was for the coaching staff to use this as a tool to assist in strategizing against other teams in both pick-ban phase and in-game, with a secondary role as a roster formation tool.

For the pick-ban phase, the plan is to convey which champions and players were the most threatening on another team. This can be best summarized by answering the following questions:

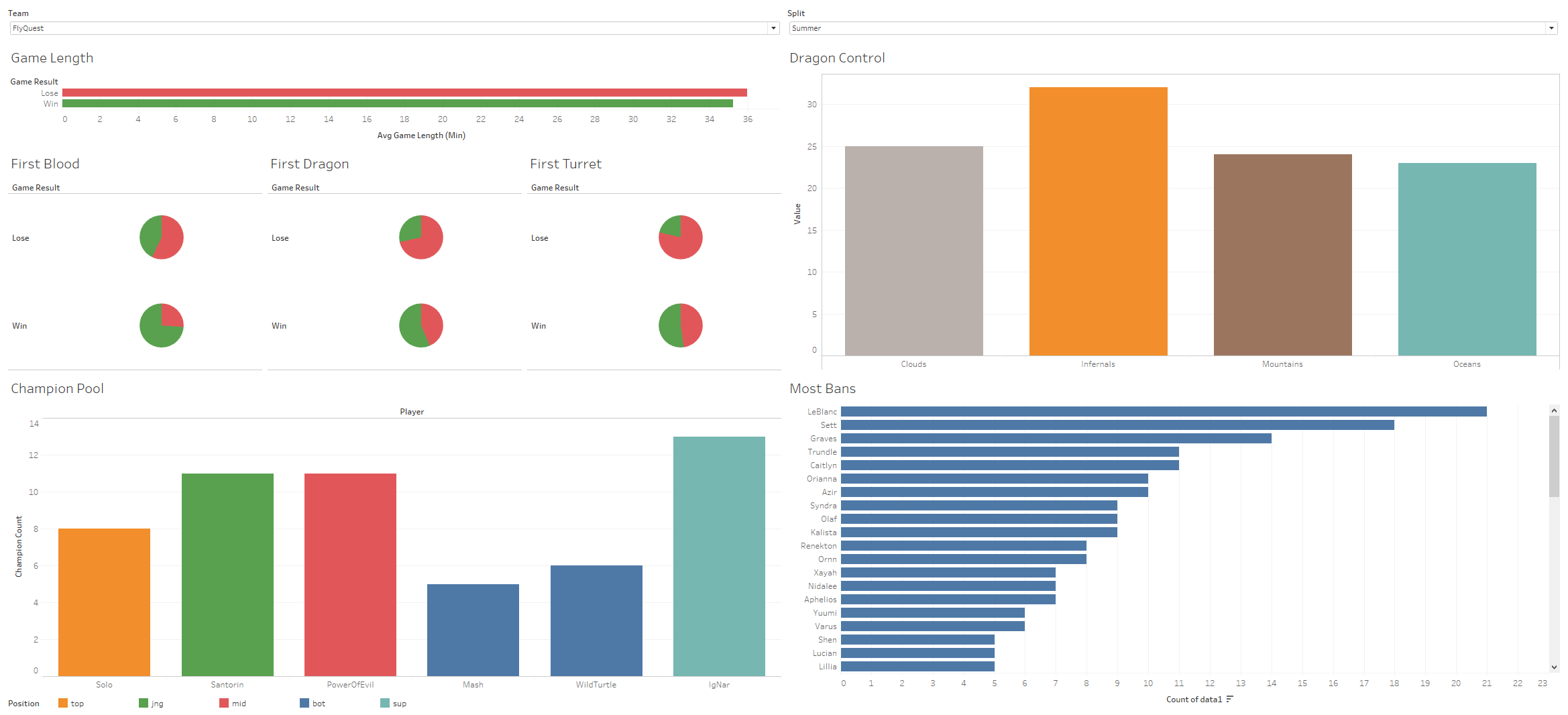
1. How do the team’s stats compare when they win versus when they lose?
2. How strong do they take advantage of their early game leads?
3. Do they value one elemental dragon over others?
4. How many champions does a player play and can we ban this player out?
5. What are they scared to play against?

For the in-game match up, the goal was for the coaching staff to be able to compare their players against opponent players. The questions that had to be answered were:

1. Which positions can they expect to win, lose or go even?
2. How effective will shutting a player down be? If the enemy laner is very effective in using their gold, but their laner isn’t as much, then putting resources and time into shutting that lane down won’t have a big impact because the disparity is too big.
3. How dominant is this player in the first ten minutes of the game? Lane phase usually lasts approximately 15 minutes (essentially when the first turret would have fallen), and one of the biggest advantages is gained during lane phase because it’s largely a one-versus-one (two-versus-two in the bottom lane) game.
4. How dependent is another team on a specific player?
5. How is this player’s overall performance?
6. Does anything about the matchup change when looking at specific champions played?

**Results:**

For the pick-ban phase, I developed the following view in Tableau to answer the questions I posed in the *Approach* section:



The horizontal bar graph covers game length split by winning and losing matches. This view is important to see what play-style teams like to play. A short average game length for winning could mean that the team translates early leads very effectively and often the game quickly.

The bottom six pie charts act as supportive charts for the game length visualization. This details out the percentage of first blood (kills), first dragon killed, and first turret destroyed by the team. These are often key objectives that are fought over in the early game because they provide boosts in gold or stats that can be heavily advantageous in the early game (sub 15 minutes). We can see here that Flyquest doesn’t really translate their early game advantages all too well since their average game length for winning and losing are practically the same, but they get almost 75% of first bloods in the majority of their winning games, as well as over 50% of first dragons and first turrets. It also appears that they lose over 70% of the time when they lose the first dragon or when their first turret drops - it’s possible that they aren’t well-versed in playing from behind.

The *Dragon Control* graph displays which elemental dragons the team tends to go after. Each elemental dragon provides different boosts in stats and physical changes to the map, which are valued more or less depending on the team. The buffs are as follows:

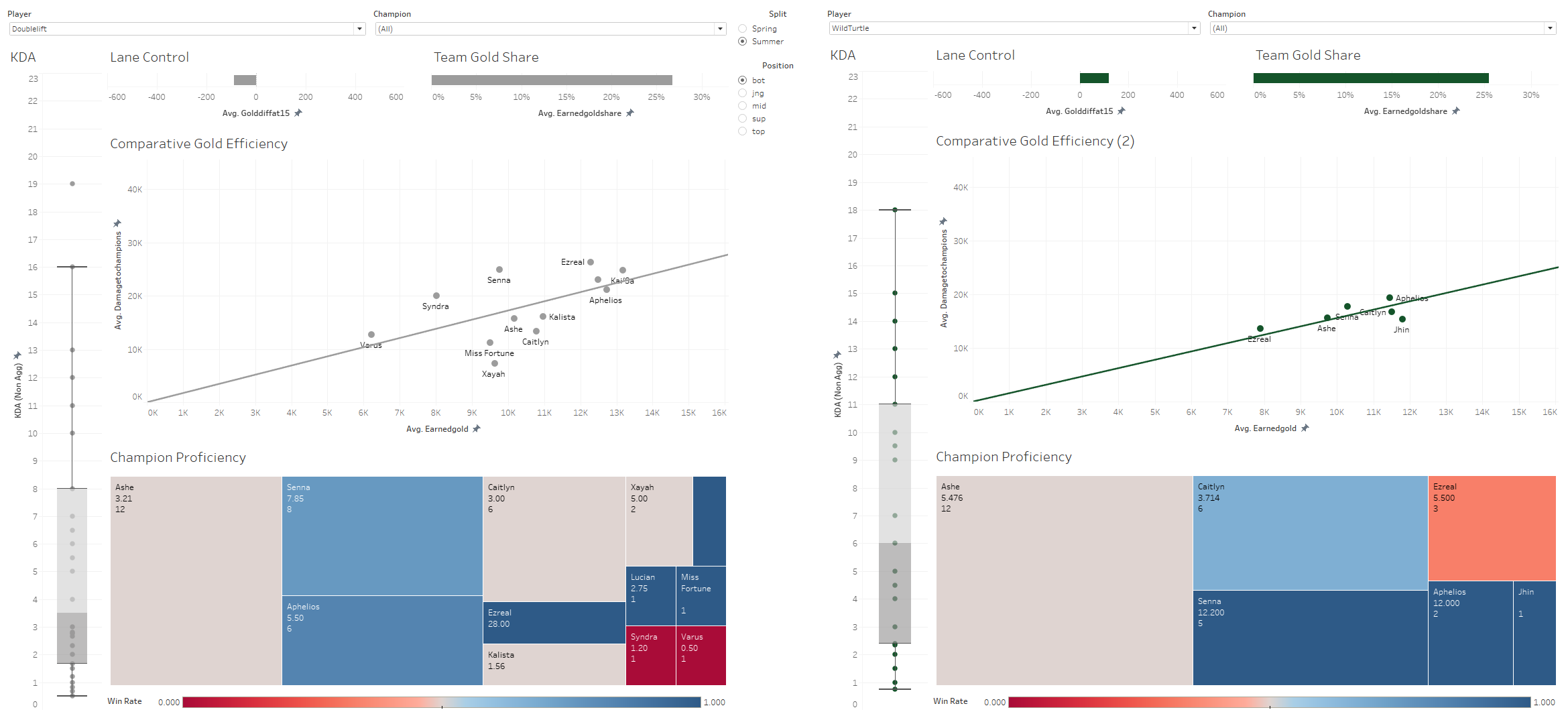
|  |  |  |
| --- | --- | --- |
| Element | Stat Boost | Map Change |
| Cloud | Ultimate ability cooldown reduction (ultimate spell reset faster) | Wind currents in specified areas that boost movement speed when walking in it |
| Inferno | Damage boosts | Certain walls and bushes are burned down - opening up the map. |
| Mountain | Defense boosts | Additional walls form across the map |
| Ocean | Health regeneration speed boosts | More Honeyfruit (healing) plants and bushes in specified points on the map |

For FlyQuest, they seem to heavily value inferno dragons, which indicate that they like finding fights.

The champion pool bar graph displays out the number of unique champions that each player has played. This answers the fourth question, and in the case for this team, it looks like Either of their bot laners, Mash and WildTurtle, are ban-able - with Mash having played four champions and WildTurtle having played six. With five total bans per team, they can be target-banned and forced onto a subpar or uncomfortable pick, however, it’s not often advised to focus so heavily on one player since that opens up other threats.

The *Most Bans* graph shows the champions that this team bans often, which indicate who they consider to be a threat or are uncomfortable playing into. It’s displayed as a horizontal bar graph instead of a vertical one because this orientation takes up less space so more information can be displayed before even scrolling. This visualization answers the last question.

The next view in the Tableau dashboard would help strategize for the actual game. The following visualizations would hopefully answer the questions posed for this stage of the match:



The first graph is the KDA boxplot. It shows the expected performance for two players based on KDA. This visualization contributes to answering the first two questions because it shows where the team’s expected performance is for that player versus the opponent.

The next graph is the Lane Control bar chart which shows how much more gold this player has by the time turrets would expect to fall.

The Team Gold Share bar graphs answers the fourth question by showing the percentage of a team’s total gold that they have. It looks like both teams depend on their respective bot laners the same amount since they each take up about a quarter of their team’s total gold.

The Gold Efficiency scatter plot answers the second question by displaying the average damage dealt to champions per average earned gold. With the trendline graphed out too, users can tell which champions players are more effective on, especially when deciding strategies on who to prioritize in a game without resulting in a serious deficit in another lane: who can do more with less?

The champion proficiency chart uses the number of games as the dimension for the area of the boxes, and the win rate is used as shading; the color palette is diverging with 50% as the white, and red as sub-50% win rate (more losses) and blue as post-50% win rate (more wins). Also displayed as labels within the boxes are the champion name, the average KDA, and the number of games played. Hovering over the boxes will reveal the actual win rate. I believe this chart answers the fifth question fairly well. In this specific scenario, the takeaway from that chart should be that WildTurtle’s comfort pick is Ashe, however his Senna (and arguably, Caitlyn) is more threatening. Since some of Doublelift’s champion proficiency overlaps, it could be a good idea to either ban out or take any of those choices from WildTurtle.

There’s an additional filter to view player performance based on specific champions too:



This allows for a more granular match up to ascertain expectations for certain matchups. In this Ashe versus Caitlyn matchup, we can expect WildTurtle to perform better in combat, since he has better lane control, better historical results, and a better distribution according to the boxplot, but Doublelift’s gold efficiency is a bit higher on Ashe meaning that in a slight disadvantage, he can still be expected perform well.

**Conclusion:**

Ultimately, *League of Legends* is a game with a large number of variables and even extensive post-game visualizations won’t be enough to paint a full picture of the game. If tracking data were to ever be made more accessible, it would greatly supplement post-game box score data since tracking data will bring another dimension to how teams play:

* How is the jungler pathing in the map, and which lane does he tend to stay nearby?
* Does the mid laner prefer to gank top lane or bot lane more often?
* Where does the support like to ward, and where does he like to roam?

However, I believe the visualization I developed to look at teams and players at a performance level can help start conversations and draft strategies on the possible prioritization teams need to consider when jumping into a match. This can be especially helpful in a best-of series where games will be played back-to-back and the coaches need to be adaptive based on the results of the previous games.

**References:**

1. Erzberger, T. (2019, December 15). League of Legends quickly gaining on traditional sports in American popularity. Retrieved December 11, 2020, from https://www.espn.com/esports/story/\_/id/28319463/league-legends-quickly-gaining-traditional-sports-american-popularity
2. <https://na.op.gg/summoner/userName=Hey%20Google>
3. <https://app.mobalytics.gg/lol/comparison?lr=NA&ls=Hey%20Google&queue=RANKED&role=top&rr=na&rs=mermaidman1>
4. Perin, C., Vuillemot, R., Stolper, C. D., Stasko, J. T., Wood, J., & Carpendale, S. (2018). State of the Art of Sports Data Visualization. Computer Graphics Forum, 37(3), 663-686. doi:10.1111/cgf.13447
5. Afonso, A. P., Carmo, M. B., & Moucho, T. (2019). Comparison of Visualization Tools for Matches Analysis of a MOBA Game. 2019 23rd International Conference Information Visualisation (IV). doi:10.1109/iv.2019.00029
6. <https://riot-api-libraries.readthedocs.io/en/latest/roleid.html>
7. Sevenhuysen, T. (2020). Oracle's Elixir - LoL Esports Stats. Retrieved December 11, 2020, from https://oracleselixir.com/tools/downloads

**Appendix:**

*Descriptive Statistics:*



*Missing Data Explanation and Solution:*

