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 audience[1].

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.gg[2]

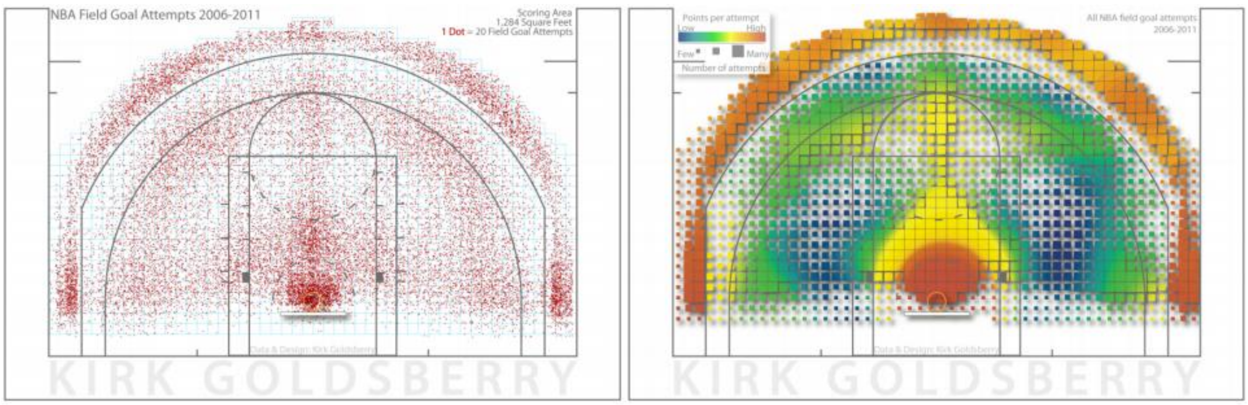
These views are fairly high-level and provide enough information for players to understand how that game went, but doesn't provide enough or irrelevant information. This is the result of attempting to provide a universal layout to accommodate all players, but there are better metrics of performance than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

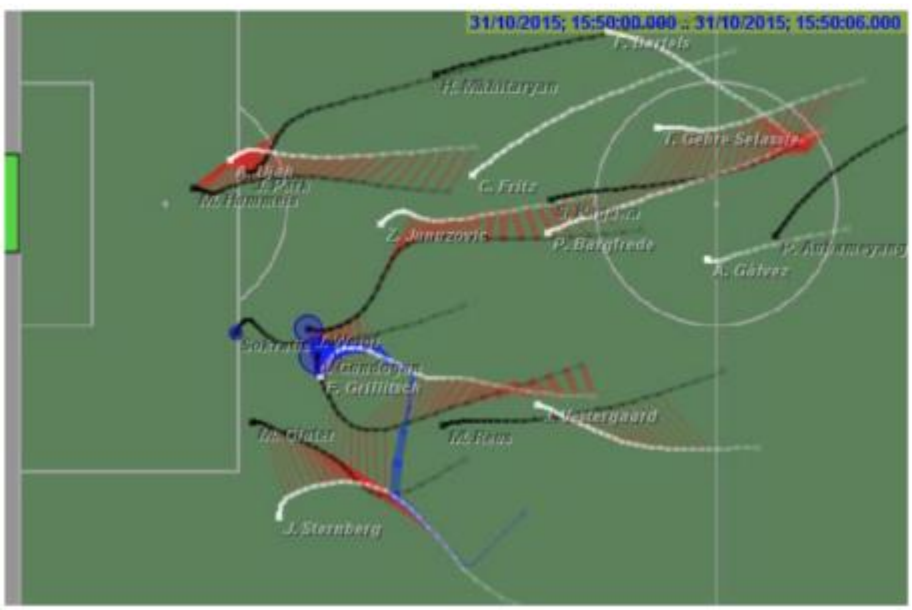
Comparison Tool developed by Mobalytics.gg[3]

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 10[4]



State of the Art of Sports Data Visualization. Computer Graphics Forum, Page 11[4]

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 121[5]

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 API[6].

Approach:

ETL and EDA were performed on the 2020 professional League of Legends dataset found on the *Oracle's Elixir* website[7]. 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.

The 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?

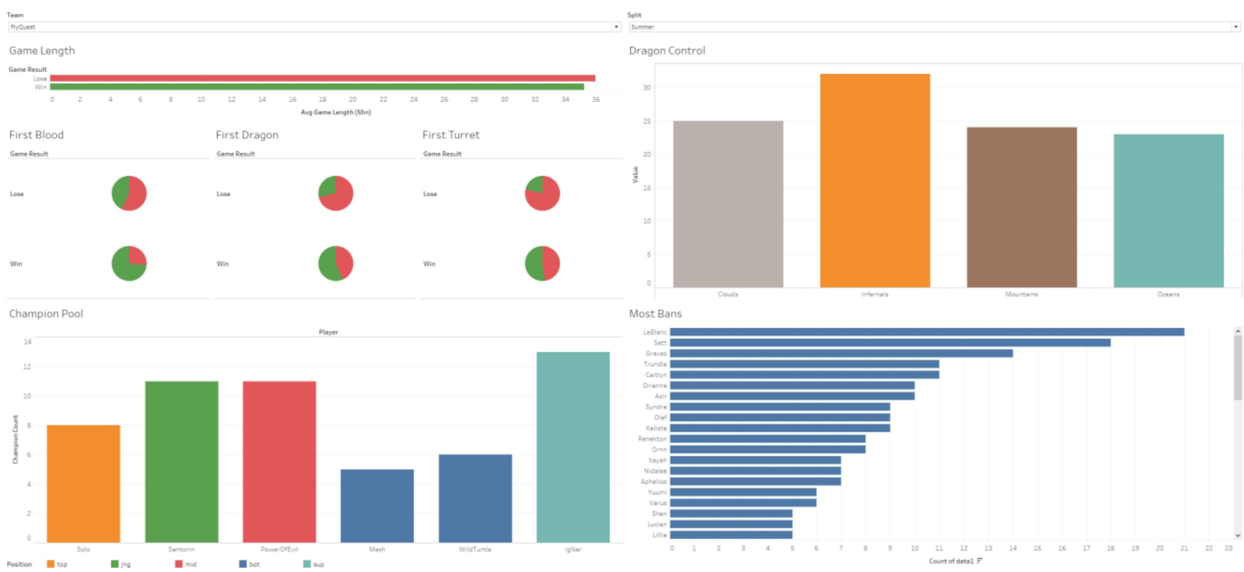
For the in-game match up, the goal was for the coaching staff to be able to compare their players against opponent players. The questions that had to be answered were:

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

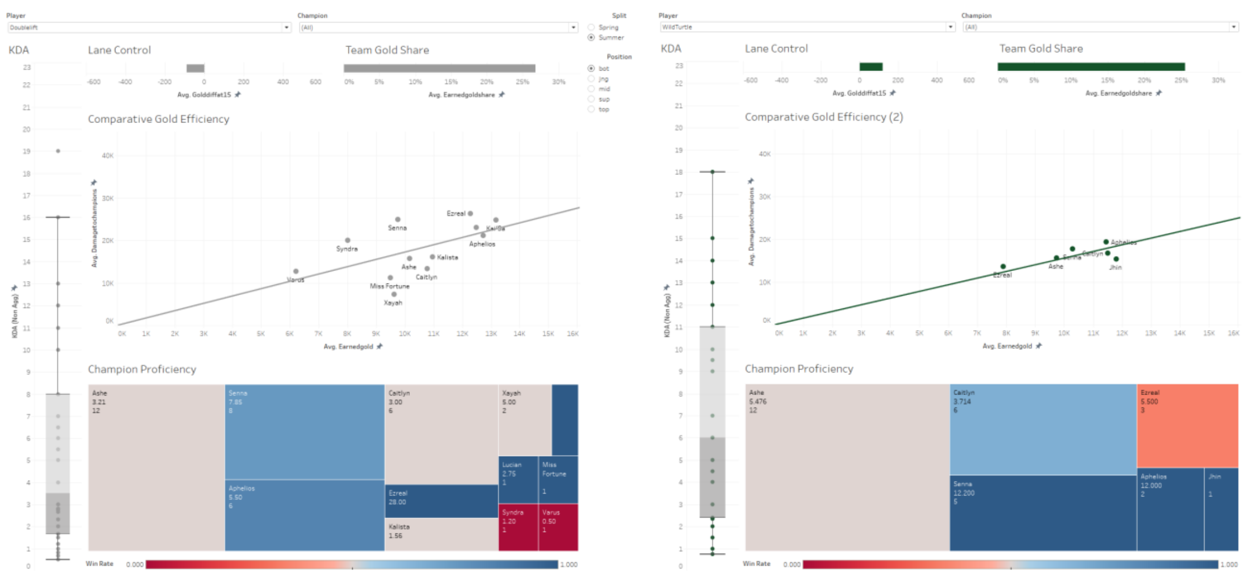
For FlyQuest, they seem to heavily value inferno dragons, which indicate that they like finding fights.

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:



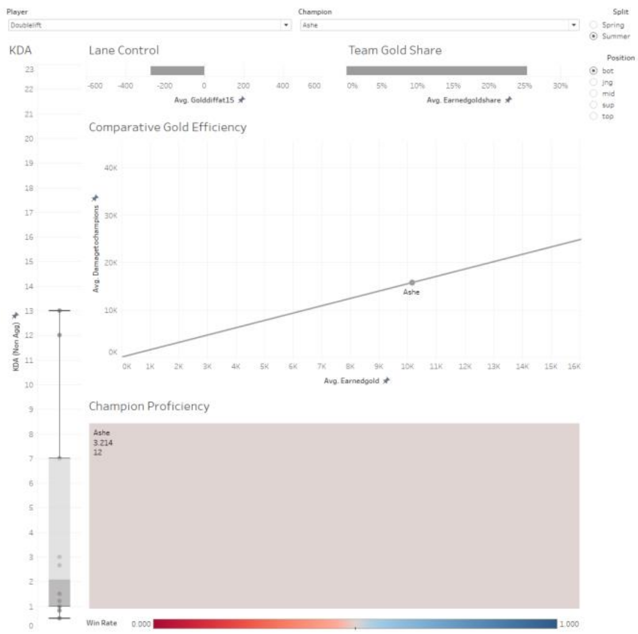
The first graph is the KDA boxplot. It shows the expected performance for two players based on KDA. This visualization contributes to answering the first two questions because it shows where the team’s expected performance is for that player versus the opponent.

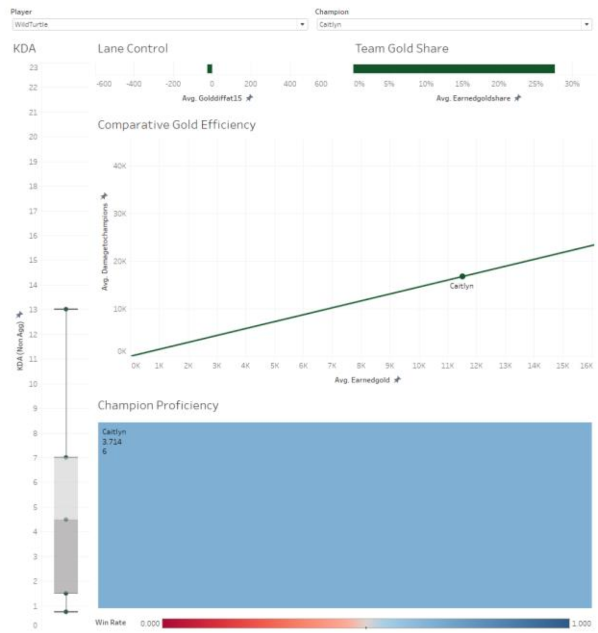
The next graph is the Lane Control bar chart which shows how much more gold this player has by the time turrets would expect to fall.
The Team Gold Share bar graphs answers the fourth question by showing the percentage of a team’s total gold that they have. It looks like both teams depend on their respective 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:

	mean	std	min	25%	50%	75%	max	data type	data type subcategory	description
gamelength	2062.306818	388.4404228	1337	1793.25	1988.5	2292.25	3680	quantitative	ratio	match time in seconds
kills	3.799873737	4.7019509	0	1	2	5	30	quantitative	ratio	number of kills
deaths	3.803030303	4.513988744	0	1	2.5	4	30	quantitative	ratio	number of deaths
assists	9.107323232	11.10522496	0	3	6	10	86	quantitative	ratio	number of assists
teamkills	11.39962121	6.139008588	0	6	11	16	30	quantitative	ratio	number of kills aggregated by team
teamdeaths	11.40909091	6.136995595	0	6.75	11	16	30	quantitative	ratio	number of deaths aggregated by team
doublekills	0.38510101	0.800650345	0	0	0	1	9	quantitative	ratio	number of double kills (2 in succession within a set time limit)
triplekills	0.053661616	0.245501057	0	0	0	0	2	quantitative	ratio	number of triple kills (3 in succession within a set time limit)
quadrakills	0.005681818	0.075175255	0	0	0	0	1	quantitative	ratio	number of quadra kills (4 in succession within a set time limit)
pentakills	0.001262626	0.035516619	0	0	0	0	1	quantitative	ratio	number of penta kills (5 in succession within a set time limit)
team kpm	0.335271212	0.187798682	0	0.192725	0.3141	0.460025	0.9662	quantitative	ratio	kills per minute by team
ckpm	0.670544697	0.204346886	0.139	0.51125	0.6482	0.7982	1.4493	quantitative	ratio	Average combined kills per minute (team kills + opponent kills)
dragons	2.520833333	1.528831599	0	1	3	4	7	quantitative	ratio	number of dragons secured
opp_dragons	2.520833333	1.528831599	0	1	3	4	7	quantitative	ratio	number of dragons secured by the opponent
elementaldrakes	0.393939394	1.058497967	0	0	0	0	4	quantitative	ratio	number of elemental dragons secured
opp_elementaldrakes	0.393939394	1.058497967	0	0	0	0	4	quantitative	ratio	number of elemental dragons secured by the opponent
infernals	0.101010101	0.407413448	0	0	0	0	4	quantitative	ratio	number of internal dragons secured
mountains	0.089015152	0.36542112	0	0	0	0	3	quantitative	ratio	number of mountain dragons secured
clouds	0.096590909	0.383773818	0	0	0	0	4	quantitative	ratio	number of cloud dragons secured
oceans	0.107323232	0.430708693	0	0	0	0	4	quantitative	ratio	number of ocean dragons secured
dragons (type unknown)	2.4	1.173787791	1	2	2	3.5	4	quantitative	ratio	dragon type unknown - probably caused by an error in data collection from tournament realm
elders	0.018623737	0.158837345	0	0	0	0	3	quantitative	ratio	number of elder dragons secured
opp_elders	0.018623737	0.158837345	0	0	0	0	3	quantitative	ratio	number of elder dragons secured by the opponent
heralds	0.965909091	0.804855242	0	0	1	2	2	quantitative	ratio	number of heralds secured
opp_heralds	0.965909091	0.804855242	0	0	1	2	2	quantitative	ratio	number of heralds secured by the opponent
barons	0.6875	0.77855203	0	0	1	1	3	quantitative	ratio	number of barons secured
opp_barons	0.6875	0.77855203	0	0	1	1	3	quantitative	ratio	number of barons secured by the opponent
towers	6.130681818	3.804517178	0	2	7	10	11	quantitative	ratio	number of towers destroyed
opp_towers	6.130681818	3.804517178	0	2	7	10	11	quantitative	ratio	number of towers destroyed by the opponent
inhibitors	0.34469697	0.760899313	0	0	0	0	6	quantitative	ratio	number of inhibitors destroyed
opp_inhibitors	0.34469697	0.760899313	0	0	0	0	6	quantitative	ratio	number of inhibitors destroyed by the opponent
damagetochampions	21215.41982	22785.61582	970	7359	12947	23843.25	172471	quantitative	ratio	damage dealt to champions
dpm	605.7755118	604.4298853	35.2013	225.105675	395.11355	654.037475	3445.974	quantitative	ratio	damaged per minute
damageshare	0.199999997	0.09954119	0.0178724	0.1125795	0.2009085	0.27199175	0.512909	quantitative	ratio	damage percent share compared to team total damage
damagetakenperminute	881.5339629	838.7780866	34.1207	368.003225	556.14045	868.126	4304.4777	quantitative	ratio	damage taken per minute
damagemitigatedperminute	888.4430374	900.0534998	18.4467	306.042275	526.7762	1013.037	5331.1574	quantitative	ratio	damage mitigated per minute (shields, debuffs, etc)
wardsplaced	37.9854798	39.99008139	1	13	18	50	266	quantitative	ratio	wards placed
wpm	1.094628157	1.090561185	0.0299	0.3851	0.5078	1.48975	4.8031	quantitative	ratio	wards placed per minute
wardskilled	16.8219697	18.00310108	0	6	10	19	109	quantitative	ratio	wards killed
wcpm	0.478308586	0.477804019	0	0.182825	0.29915	0.520025	2.4988	quantitative	ratio	wards killed per minute
controlwardsbought	14.71338384	14.75799679	0	5	9	16	87	quantitative	ratio	number of control wards bought
visionscore	85.54671717	86.61673158	7	34	49	91	578	quantitative	ratio	vision score
vspm	2.454996338	2.337343743	0.2911	1.0206	1.38975	2.59315	10.5921	quantitative	ratio	vision score per minute
totalgold	19795.36174	18698.31621	4462	9607	12599.5	16682.25	96558	quantitative	ratio	total gold
earnedgold	12324.19634	11977.85766	1075	5434.5	8193.5	11728.25	59866	quantitative	ratio	earned gold
earned gpm	358.5096236	337.3447907	37.348	171.68575	241.86355	322.142825	1534.8269	quantitative	ratio	earned gold per minute
earnedgoldshare	0.200000003	0.070852481	0.0484213	0.15348475	0.213014	0.25401175	0.36265	quantitative	ratio	earned gold percent share compared to team total
goldspent	18356.3125	17344.46865	3900	8943.75	11625	15612.5	88493	quantitative	ratio	gold used on items
gspd	0	0.129199643	-0.355302	-0.0960151	0	0.0960151	0.355302	quantitative	ratio	Average gold spent percentage difference
total cs	220.1522727	116.2972547	0	152	238	301	558	quantitative	ratio	total cs (minion score)
minionkills	295.0492424	294.9390802	0	41	244	328	1401	quantitative	ratio	minions killed
monsterkills	71.87121212	85.60406551	0	7	26	141	383	quantitative	ratio	neutral monsters killed (not minions)
monsterkillsownjungle	47.22032828	56.55701542	0	4	16	96	249	quantitative	ratio	neutral monsters killed in jungle
monsterkillsenemyjungle	6.820707071	11.24720309	0	0	1	9	105	quantitative	ratio	neutral monsters killed in enemy jungle
cspm	10.70577986	10.03965438	0	5.164775	8.01985	9.78845	38.5259	quantitative	ratio	cs per minute
goldat10	5115.973485	4615.674975	1793	2935	3289.5	3651.25	18463	quantitative	ratio	gold at 10 minutes
xpat10	6070.258838	5501.251951	1543	2994.5	3866.5	4861	20305	quantitative	ratio	experience points at 10 min
csat10	106.614899	99.35426209	0	58	79	93	375	quantitative	ratio	cs at 10 min
opp_goldat10	5115.973485	4615.674975	1793	2935	3289.5	3651.25	18463	quantitative	ratio	opponent gold at 10 minutes
opp_xpat10	6070.258838	5501.251951	1543	2994.5	3866.5	4861	20305	quantitative	ratio	opponent experience points at 10 min
opp_csat10	106.614899	99.35426209	0	58	79	93	375	quantitative	ratio	opponent cs at 10 min
golddiffat10	0	638.6319874	-3718	-285	0	285	3718	quantitative	ratio	gold difference versus opponent at 10
xpdiffat10	0	572.6002887	-2568	-294.25	0	294.25	2568	quantitative	ratio	xp difference versus opponent at 10
csdiffat10	0	14.689202	-78	-7	0	7	78	quantitative	ratio	cs difference versus opponent at 10
goldat15	8060.518939	7289.122428	2647	4537.75	5171.5	5916	29792	quantitative	ratio	gold at 15 minutes
xpat15	9687.400253	8786.1861	2681	4889	6321.5	7720.25	33530	quantitative	ratio	experience points at 15 min
csat15	169.6616162	158.0455635	0	90	126	149	594	quantitative	ratio	cs at 15 min
opp_goldat15	8060.518939	7289.122428	2647	4537.75	5171.5	5916	29792	quantitative	ratio	opponent gold at 15 minutes
opp_xpat15	9687.400253	8786.1861	2681	4889	6321.5	7720.25	33530	quantitative	ratio	opponent experience points at 15 min
opp_csat15	169.6616162	158.0455635	0	90	126	149	594	quantitative	ratio	opponent cs at 15 min
golddiffat15	0	1278.069747	-7782	-560	0	560	7782	quantitative	ratio	gold difference versus opponent at 15
xpdiffat15	0	974.875182	-5909	-497.5	0	497.5	5909	quantitative	ratio	xp difference versus opponent at 15
csdiffat15	0	21.70153777	-114	-11	0	11	114	quantitative	ratio	cs difference versus opponent at 15
gameid	N/A	N/A	N/A	N/A	N/A	N/A	N/A	nominal	categorical	unique ID
datacompleteness	N/A	N/A	N/A	N/A	N/A	N/A	N/A	nominal	categorical	complete/partial row data
url	N/A	N/A	N/A	N/A	N/A	N/A	N/A	nominal	arbitrary	url containing match data
league	N/A	N/A	N/A	N/A	N/A	N/A	N/A	nominal	categorical	league/tournament that the match took place in
year	N/A	N/A	N/A	N/A	N/A	N/A	N/A	quantitative	interval	year of match
split	N/A	N/A	N/A	N/A	N/A	N/A	N/A	nominal	categorical	split (seasion) of match
playoffs	N/A	N/A	N/A	N/A	N/A	N/A	N/A	ordinal	binary	boolean of whether or not the game was a playoff game
date	N/A	N/A	N/A	N/A	N/A	N/A	N/A	quantitative	interval	date of match
game	N/A	N/A	N/A	N/A	N/A	N/A	N/A	ordinal	numeric	game number in sequence (e.g. game 2 in a best-of-3)
patch	N/A	N/A	N/A	N/A	N/A	N/A	N/A	ordinal	non-numeric	patch number that the match was played on
playerid	N/A	N/A	N/A	N/A	N/A	N/A	N/A	ordinal	numeric	playerid for the match
side	N/A	N/A	N/A	N/A	N/A	N/A	N/A	nominal	categorical	red/blue side
position	N/A	N/A	N/A	N/A	N/A	N/A	N/A	nominal	categorical	position that player plays
player	N/A	N/A	N/A	N/A	N/A	N/A	N/A	nominal	arbitrary	player name (ign)
team	N/A	N/A	N/A	N/A	N/A	N/A	N/A	nominal	arbitrary	team the player plays for
champion	N/A	N/A	N/A	N/A	N/A	N/A	N/A	nominal	arbitrary	champion/character selected by the player
ban1	N/A	N/A	N/A	N/A	N/A	N/A	N/A	nominal	arbitrary	one of the 5 bans selected by the team
ban2	N/A	N/A	N/A	N/A	N/A	N/A	N/A	nominal	arbitrary	one of the 5 bans selected by the team
ban3	N/A	N/A	N/A	N/A	N/A	N/A	N/A	nominal	arbitrary	one of the 5 bans selected by the team
ban4	N/A	N/A	N/A	N/A	N/A	N/A	N/A	nominal	arbitrary	one of the 5 bans selected by the team
ban5	N/A	N/A	N/A	N/A	N/A	N/A	N/A	nominal	arbitrary	one of the 5 bans selected by the team
result	N/A	N/A	N/A	N/A	N/A	N/A	N/A	ordinal	binary	boolean for victory or defeat
firstblood	N/A	N/A	N/A	N/A	N/A	N/A	N/A	ordinal	binary	if player participated in the first kill (boolean)
firstbloodkill	N/A	N/A	N/A	N/A	N/A	N/A	N/A	ordinal	binary	if player got the first kill (boolean)
firstbloodassist	N/A	N/A	N/A	N/A	N/A	N/A	N/A	ordinal	binary	if player assisted in the first kill (boolean)
firstbloodvictim	N/A	N/A	N/A	N/A	N/A	N/A	N/A	ordinal	binary	if player was the first death (boolean)
firstdragon	N/A	N/A	N/A	N/A	N/A	N/A	N/A	ordinal	binary	if the first dragon was secured by this team (boolean)
firstherald	N/A	N/A	N/A	N/A	N/A	N/A	N/A	ordinal	binary	if the first herald was secured by this team (boolean)
firstbaron	N/A	N/A	N/A	N/A	N/A	N/A	N/A	ordinal	binary	if the first baron was secured by this team (boolean)
firsttower	N/A	N/A	N/A	N/A	N/A	N/A	N/A	ordinal	binary	if the first tower was destroyed by this team (boolean)
firstmidtower	N/A	N/A	N/A	N/A	N/A	N/A	N/A	ordinal	binary	if the first middle tower was destroyed by this team (boolean)
firsttothreetowers	N/A	N/A	N/A	N/A	N/A	N/A	N/A	ordinal	binary	if team was the first to destroy three towers (boolean)

Missing Data Explanation and Solution:

Explanation of Missing Data and Solutions:

Explanation: 0.38% of ban5 and 0.19% of ban3 are missing because of penalties or timeouts: Teams only have 2 rounds of 30 seconds each to ban out certain champions. Ban3 occurs in the first round of bans and ban5 occurs in the second round of bans. Teams can be penalized with a loss of ban for violations such as not following code of conduct.

Solution: Fill in missing data with "None"

```
lcs.fillna({'ban3': 'None', 'ban5': 'None'}, inplace=True)
```

Explanation: 99.68% of Dragons (type unknown) is missing. Looking at the dataset, this variable is only filled in when the dragon-type columns Infernals, Mountains, Clouds, Oceans, Elders are empty. This could be an error in data collection from the tournament realm since all dragon-types are visible by nature of the game. The occurance is also only about 0.32%.

Solution: Unfortunately, because the data wasn't properly collected, there is no way to split "Dragons (Type Unknown)" between "elementaldrakes" and "opp_elementaldrakes" (aka: team dragons versus opponent dragons). "dragons" and "Opp_dragons" are suitable replacements. In order to keep the datatype consistent throughout the columns, the data will be filled in as 0, but the rows will be filtered out in the visualization when looking at individual elemental dragon objectives since the rest of the data is still required.

```
lcs.fillna({'elementaldrakes': 0,
            'opp_elementaldrakes': 0,
            'opp_elementaldrakes': 0,
            'infernals': 0,
            'mountains': 0,
            'clouds': 0,
            'oceans': 0,
            'elders': 0,
            'opp_elders': 0}, inplace=True)
```

Explanation: 16.67% of the data is missing Player name - this is probably because the data includes aggregated team statistics as well, which are null in the Player column. The other stats that are also missing 16.67% of the data are only available on a team-level granularity, not on a smaller player-level granularity. Similarly, variables such as dragons, heralds, towers, and other map objects are missing 83.33% of the the data. This data is only available on a Player-level granularity, not on a higher team-level granularity. As was previously noted, approximately 0.32% of the missing data is contributed by the Dragons (Type Unknown) error.

Solution: Because of expected interchangable use-cases, the player and team data will remain as one dataset since *Team* data can be filtered out on *position != 'team'*.