

## **Final Paper**

### **Introduction about Esports and related concepts**

Esports, also known as electronic sports, is a form of sport competition using video games. Esports games are as popular and profitable as traditional sports games. The attention and passion that their fans put in are unprecedented. Right now, it is still an incompletely developed market.

Personally, I am a big fan of Esports specially League of Legends (aka. LOL). I enjoy watching the World championships. In this article, I will introduce some key gaming terms to help you understand my research, then I will conduct analysis for how and why the team can win the game.

LOL is a comprehensive game, basically, its included two major factors to win the game. The first one is good basic practical skills like performed the last hit on enemy minions and obtained the gold to purchase weapons to make your champion stronger, which is named "Farming" in gaming terminology. The second one is about strategy. I think LOL is a decision-making game. As a summoner (nick name for every human player), you need to make choices consistently in the game. Even before the game, in the preparation stage, you need decide which champions you want to pick and ban in this incoming game. This process is named "Ban & Pick" and it is one of decisive factors in determining victory or defeat. Since each champion has distinctive ability, when some champions abilities match up, they will maximize their strengths, just like in NBA, the combination of Kobe and O'Neal had destroyed other teams for 3 years. What I said above is only a little part of how the decision-making influence the game.

Compare with common players, professional players are skillful and experienced with most of champions. So, it is hard to outplay your rivals just by your technical skills, advanced strategy is a necessary requirement for a top team as same as Jordan's need for triangle offense. Therefore, I will focus more on the strategy part in my study

#### **Data Source:**

I got the data from <https://oracleselixir.com/match-data/>, it's data resource for League of Legends E-sports statistics.

There are tons of columns contained in the dataset, I chose some of them as my analysis targets, the position, side, result, ban1-ban5 and KDA. In the position column, there are five different labels named "Top, Mid, Jungle, Support and ADC". They five are 5 different roles in the game. There are two sides in the game, blue and red as same as the white and black side in the chess. The geography in each side is different, so develop side-specific tactics will help you win the game. The result column is binary, 0 means lost and 1 means won, which is the most important criteria for my analysis because my final goal is to train a model can predict who will win the game. The "ban1-5" is a pre-game step, each team can ban 5 champions as they like before the game. I will go further about it in the following paragraphs. "KDA" is a parameter that used to measure a player's performance in each game as same as the shooting average for a NBA player.

For the data cleansing, since many columns in this dataset are useless for my study and the dataset looks chunky and messy, I firstly deleted unrelated columns. Afterward, I used the calculation formula (I will introduce it in the following paragraphs) to simplify the "Kill", "Death" and "Assist" to one column "KDA", which is great indicator that used to evaluate a player's performance.

## What is the “KDA”?

I want to introduce an indicator, Kill Death Assist. It is one of the most crucial criteria to evaluate a player’s performance in each game, which is as same as the shooting percentage of a basketball player. In the LOL game, players will get a large amount of golds by killing enemy champions and then use the golds to buy weapons to make their champions more powerful. We call this process “**Snowballing**”. So your snowball will be as big as your KDA.

$$\text{KDA Formula} = (\text{Kills} + \text{Assist}) / \text{Deaths}$$

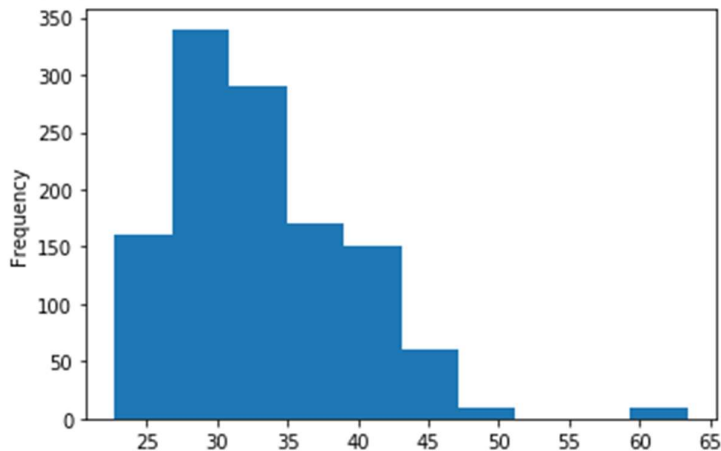
Typically, if a player’s KDA is lower than 1, his contribution to this game is negative because he had been killed more than his rivals, so he brings as many golds for the enemy as his many times deaths. Otherwise, the player influences the game positively.

So back to my case, I create a column named “KDA” used the formula above to calculate each player’s performance in the games.

## How long is a normal LOL game?

As the histogram below I generated by Python shows, most of the games ended in 25-35 minutes, so 15 minutes is fulfilling my research requirement. I decided to focus on the correlation between the gold difference at 15 minutes in the game and the victory, because the “snowball size” you would get at 15 minutes usually decides the winner of the first team fight

```
In [18]: LOL_data['gamelength'].plot.hist()  
Out[18]: <matplotlib.axes._subplots.AxesSubplot at 0x2260148ca58>
```



## What is “Ban&Pick” phase?

In a professional League of Legends match, there is a pre-game phase named “Ban&Pick”, Summoners (Aka players) have chance to pick one champion they like and ban a champion they don’t want to face in the incoming match. This pre-game phase is an important part for team tactic, because it is a good chance to ban 5 champions who your rivals played very well in former games to limit the rival players performance. Furthermore, some certain champions in a same team can coordinate very well, it’s named “Ability-Combo” in LOL terminology, but each LOL champion has its “natural enemy”, we called it “counter-pick” in LOL terminology, so it is a good strategy to ban your champion’s “counter-pick” before you pick it.

The rule in LOL professional game match is that the “Ban&Pick” progress alternately, typically, the third to fifth bans is a “counter-ban”, which means it’s a tactical reaction based on the rival’s pick, in other words, to destroy rival’s “Ability-Combo”. Therefore, I decided to focus on the

first and second ban rounds because it will show which champions are most likely used as the “Playmaker”, because typically, only the overpowering champions in this patch will be designed as a team’s tactic cores.

## Part 2 Explanatory analysis:

### Champions’ Capabilities Analysis:

#### Which champions were banned a lot and Why?

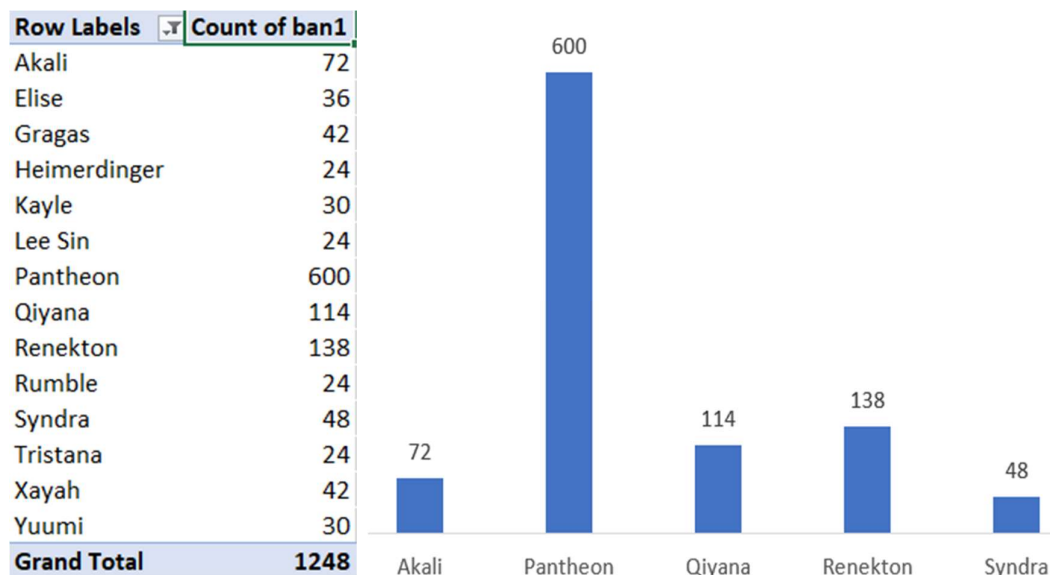
To figure out this question, I must know which champions were banned a lot and what kind of distinctive abilities these champions have.

Firstly, I must prepare the dataset first. What I did first is delete all the unnecessary columns but only left “ban1 - ban5” and the match result.

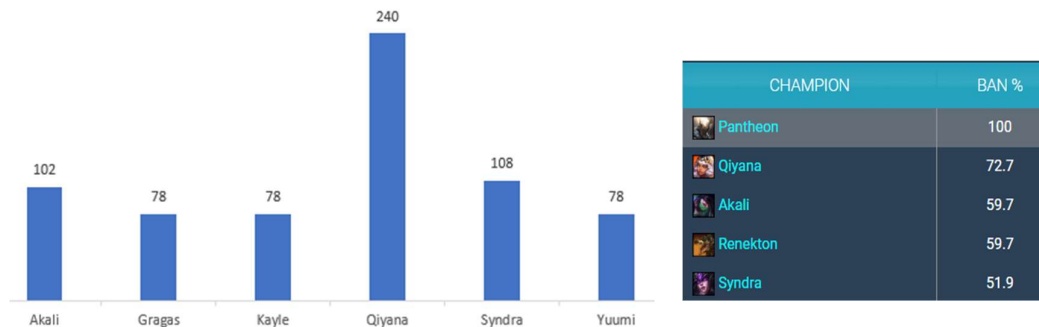
Before we go deep in this topic, let me introduce the ‘Ban&Pick’ rule in LOL game. In the LOL professional games, each team has chance to ban and pick 5 champions as they like before the game started, which is an important part for team tactic, because it is a good chance to ban 5 champions who your rivals played very well in former games to limit the rival players performance. Furthermore, some certain champions in a same team can coordinate very well, it’s named “Ability-Combo” in LOL terminology, but each LOL champion has its “natural enemy”, we called it “counter-pick” in LOL terminology, so it is a good strategy to ban your champion’s “counter-pick” before you pick it.

The specialty of LOL games is that the ban and pick progress alternately, typically, the third to fifth bans is a “counter-ban”, which means it’s a tactical reaction based on the rival’s pick, in other words, to destroy rival’s “Ability-Combo”. Therefore, I decided to focus on the first and second ban rounds because it will show which champions will most likely be used as the “Playmaker”, because typically, only the overpowering champions in this patch will be designed as a team’s tactic cores.

The column graph below shows the Top 5 of the first banned champions and the left is the times each champion was banned in first round of each game. Pantheon was banned most. So, we can



know how overpowering Pantheon was in that patch.



The second column chart above is the Top 5 (There are 6 because Yummi, Gragas and Kayle were banned same times) banned champions in the second round. The most banned champion in the second round is Qiyana, 240 times. And by combining both two graphs, Syndra, Akali and Qiyana were banned either in the first or second round, which means they are also overpowering in that patch and can play important role in team tactics.

After figuring out which champions are overpowering or important in that patch, I want to know why they be banned? There must be some characteristics they have in common. Before we dive into this topic, I want to introduce each role in LOL game. Marksman is the attack damage core in a team. Support in the game is to support and protect the marksman, its abilities are normally shield and healing. Mage is the magic damage core in the team. These three roles have small amount health but long attack/ability range, typically they need to play safe in the game because they are fragile and easy to be assassinated. Tank is just as its name implies, the job is damage-taken and engage the team fight such as charging to rival's position, so Tank has large amount of health and armor. The last one is Assassin, a distinctive role. Assassins have small amount of health and short range but some special assassinated abilities such as stealth, dash and invisible, and they are able to make nuke damage in a second. Based on its special design, assassin champions' ceiling is as high as your skill, which is the reason why they are popular in professional leagues.

By analyzing the common features of these champions, Qiyana, Akali and Pantheon are all "Assassins". Once the marksman be taken down by them, rival team would win the team fight easily. In addition, they have another characteristic in common, great ganking ability. Gank means to ambush a rival champion usually with greater numbers of our champions, so high movement speed is one of the requirements.

Therefore, the conclusion for my champion analysis is that the professional teams value the "Ganking (support)" and "assassinating" abilities most in that patch because professional players are more skillful on using champions' abilities. By picking an Assassin, the ceiling of the team will be higher as well.

## Champion “Ban&Pick” Predictive Modeling

In my Champion Abilities analysis, I researched how some factors such as KDA, gold difference and side difference (blue or red) influence the game victory. So, I want to go deep in the correlation between each champion’s win rate and his/her/its KDA, i.e. the gaming performance in match. I want to generate a predictive model to help the coach responding the rival’s move in “Ban&Pick” phase.

In League of Legends game, if a champion’s win rate above 50%, this champion is a “positive pick”, which means he or she will influence the game positively. The one whose win rate above 52.5% is the “op pick”, which means this champion is overpowering in the game.

### Champion’s win rate and KDA

I firstly used Python to generate the Top 10 and Bottom 10 win rate champions with their average KDA and how many matches they have been picked.

In the result below, you can see some of champions have been picked way more times than the others who have a much greater win rate, for example, Ezreal and Ivern. Even Ivern has the greatest win rate, but he has been picked way less than Ezreal, 6671 vs 23480. It is not reasonable to simply use win rate to assess a champion’s performance in the game, because ‘few samples’ will let your model overfitting.

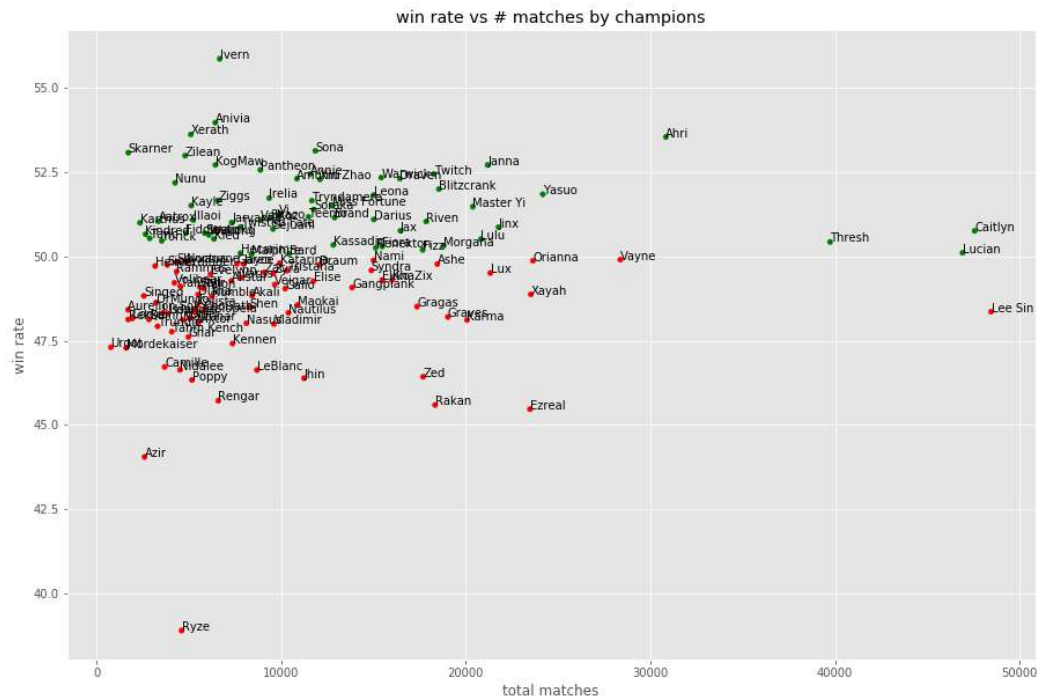
#### Top 10 win rate

	total matches	win rate	K	D	A	KDA
name						
Ivern	6671	55.9	2.6	4.2	13.3	3.8
Anivia	6433	54.0	6.4	4.8	7.5	2.9
Xerath	5108	53.6	7.3	5.4	8.6	2.9
Ahri	30841	53.5	7.3	5.6	7.7	2.7
Sona	11847	53.1	3.0	5.7	13.9	2.9
Skarner	1720	53.1	4.8	5.0	9.2	2.8
Zilean	4797	53.0	3.2	4.7	11.9	3.2
KogMaw	6440	52.7	7.9	6.6	7.1	2.3
Janna	21182	52.7	0.9	4.0	14.5	3.8
Pantheon	8872	52.6	8.1	6.5	6.8	2.3

#### Bottom 10 win rate

	total matches	win rate	K	D	A	KDA
name						
Nidalee	4520	46.6	7.0	6.1	7.3	2.3
LeBlanc	8690	46.6	8.0	5.9	6.1	2.4
Zed	17687	46.4	8.5	6.5	5.7	2.2
Jhin	11240	46.4	7.1	5.9	8.6	2.6
Poppy	5165	46.3	4.4	5.3	8.2	2.4
Rengar	6590	45.7	7.0	6.6	6.4	2.0
Rakan	18339	45.6	1.5	5.6	14.2	2.8
Ezreal	23480	45.5	6.7	5.5	8.0	2.7
Azir	2597	44.1	5.9	6.6	6.6	1.9
Ryze	4607	38.9	5.4	6.3	6.2	1.9

So, I created a **scatter plot** for each champion's two variables, win rate and matches (i.e. picked times) in order to filter those champions out. Green for win rate smaller than 50%, red for larger



Lee Sin has been picked most but the win rate is not good enough, lower than 50%. Ivern has the greatest win rate but less picked times, smaller than 10000. Based on this scatter plot, Ahri got my attention. She has a great win rate (above 52.5%) as well as more picked times (more than 30,000 times). So Ahri, Janna and Yasuo are good examples for the champion who have both good matches and win rate data to avoid overfitting in future machine learning model.

In order to predict the winner of the game, we need to bring a new factor in, the “**Role**”

### Champions Role and win rate Analysis

Since every champion in League of Legends should play a role in each game (Top, bottom, mid, jungle and support), I am curious about the correlation between the champion at a certain role and its win rate. Sometimes, coaches like to place a certain champion at an unusual role, it is part of the gaming strategy. I created the table about Top 10 win rate with role. In this table, Pantheon, the overpowering champion I mentioned in preliminary, normally plays at Top, but when place him in Mid, his win rate reaches even 60%. However, not all champions played at unusual role had good performance. Lee Sin is normally at Jungle but when he played at Mid, his win rate is only about 33%. So by my this analysis, team coaches will know which champion(s) can perform even better at an unusual role. In this Top 10 champions, Pantheon, Yasuo, Draven (for two), Sona, Velkoz, and Leblanc were played at an unusual role. It is an interesting finding that most of champions who were played at an unusual role had great win rates



```

Top 10 win rate with role (occur > 0.01% of total # matches)
              total matches  win rate  K  D  A  KDA
name  adjposition
Pantheon MID              417      60.0 9.2 6.5  7.0  2.5
Ivern  JUNGLE             6515      56.3 2.7 4.1 13.3  3.9
Yasuo  DUO_CARRY          203      56.2 8.4 7.3  6.9  2.1
Zilean MID             1380      55.8 4.5 4.1  9.8  3.5
Draven TOP              185      55.7 7.2 7.0  4.8  1.7
Sona   MID              254      55.1 6.5 5.7 10.2  2.9
VelKoz TOP              178      55.1 7.5 6.7  7.1  2.2
Draven MID              151      55.0 7.9 7.1  5.2  1.8
LeBlanc TOP             241      54.8 7.8 5.9  5.5  2.2
Anivia MID             5923      54.2 6.5 4.7  7.5  3.0
Bottom 10 win rate with role (occur > 0.01% of total # matches)
              total matches  win rate  K  D  A  KDA
name  adjposition
Teemo  JUNGLE            415      39.8 6.2 7.2  6.8  1.8
Ashe   DUO_SUPPORT       200      39.0 2.2 6.8 11.0  1.9
Lee Sin DUO_SUPPORT      186      38.7 3.9 7.3 10.7  2.0
        TOP              584      38.0 6.6 7.0  6.3  1.8
Talon  JUNGLE            185      37.8 6.7 7.8  6.6  1.7
Varus  MID              297      37.7 6.9 7.2  7.6  2.0
Ryze   MID             3197      36.8 5.6 6.2  6.1  1.9
Udyr   TOP              273      36.6 4.4 6.6  4.6  1.4
Jhin   MID              156      36.5 7.3 7.0  7.5  2.1
Lee Sin MID             163      33.7 7.9 7.2  6.3  2.0

```

Afterwards, I conducted an analysis for some specific rivals at each role, it will help team coaches to make “counter pick”, the term I interpreted in the preliminary analysis. The larger the counter score, the more the first outplayed the second. The counter score larger than 0 means the first outplayed the second, vice versa.

The first group is Mid, the most important and well-farmed role in the game, and the “damage-maker” in a team. In the mid group, you can see Ryze is the most defeated champion, which is also matched my last KDA analysis result, Ryze is the lowest win rate champion (38.9%). And we also can know how easy Annie outplay Ryze, 25 counter score. So, when your rival pick Ryze, you may first want to pick Annie.

```

MID:
              match up  total matches  counter score
0      Annie vs Ryze           80.0           25.0
1      Ahri vs Ryze          319.0           21.2
8      Azir vs Lux           150.0          -14.7
12     Ryze vs Twisted Fate    78.0          -14.1
13     Ekko vs Ryze          130.0           13.8

```

```

JUNGLE:
              match up  total matches  counter score
2      Rengar vs Xin Zhao     200.0          -17.5
4      Evelyn vs Rengar       118.0           15.3
5      Hecarim vs Twitch       89.0           15.2
10     Graves vs Ivern        410.0          -14.4
11     Nunu vs Vi              95.0           14.2

```

TOP:

	match up	total matches	counter	score
3	Gangplank vs Singed	102.0		15.7
6	Yasuo vs Yorick	109.0		15.1
9	Shen vs Teemo	189.0		-14.6
14	Maokai vs Yasuo	113.0		-13.7
15	Irelia vs Kennen	186.0		13.4

DUO\_SUPPORT:

	match up	total matches	counter	score
7	Janna vs Tahm Kench	168.0		14.9
23	Blitzcrank vs Galio	103.0		13.1
37	Leona vs VelKoz	124.0		12.1
61	Tahm Kench vs Zyra	84.0		-10.7
74	Brand vs Lux	149.0		10.4

DUO\_CARRY:

	match up	total matches	counter	score
29	Kalista vs KogMaw	115.0		-12.6
72	Ezreal vs Graves	81.0		-10.5
95	Graves vs Vayne	149.0		-9.7
110	Sivir vs Twitch	201.0		-9.2
176	KogMaw vs Twitch	320.0		-8.1

### Ban & Pick Advising Model:

All above are just Top 5 counter scores at each role in the game (either positive or negative), but in the real world (i.e. the professional matches), the team coach will want a model that can return quickly that which group of champions counter the rival's pick.

**Important:** Since some champions can play at different role, I associated the role with the champion we want to query.

Let me show you some examples:

Scenario 1: Imagine your rival picked Yasuo and you want to counter him....

```
In [52]: champion = 'Yasuo'
...: role = 'MID'
...: get_best_counter(champion, role)
Best counter for MID - Yasuo:
champion  total matches  advantage
55   Karthus           77.0         11.0
156  Kayle             84.0          8.3
314  Ziggs            181.0          6.4
343  Annie            344.0          6.1
450  Xerath            181.0          5.2
990  Anivia            211.0          2.6
1006 Viktor            215.0          2.6
1300 Ahri           1,147.0          1.5
1391 Azir             125.0          1.2
1402 Kassadin          428.0          1.2
1444 Twisted Fate       241.0          1.0
1484 Lissandra          114.0          0.9
1590 Veigar            289.0          0.5
1622 Ekko              448.0          0.4
1655 Diana             161.0          0.3

In [53]: champion = 'Yasuo'
...: role = 'TOP'
...: get_best_counter(champion, role)
Best counter for TOP - Yasuo:
champion  total matches  advantage
177   Aatrox          117.0          8.1
333   Wukong           137.0          6.2
530   Pantheon          252.0          4.8
645   Jax              399.0          4.1
857   Renekton          670.0          3.1
936   Fizz              229.0          2.8
1070 Tryndamere        417.0          2.3
1099 Darius             596.0          2.2
1186 Riven              779.0          1.9
1637 Irelia             403.0          0.4
1652 Swain             153.0          0.3
```

You can see the results of mid lane Yasuo and Top lane Yasuo are different. So you might want to figure out which lane the Yasuo will go first. This kind of champions are named “Swingman” in the gaming/sports term which means a champion/player who can play two or more roles. It is usually a good choice at the first or second pick (5 picks totally), which can give the team more flexibility. So, as a counter, the best strategy is to find a champion who can outplay him



at both roles.

Scenario 2: Imagine your rivals picked Jayce first...

After generating the results by the model, you may want to pick Yasuo, because he can outplay Jayce at either Top or Mid, so you got a “Counter Swingman”, Hooray!



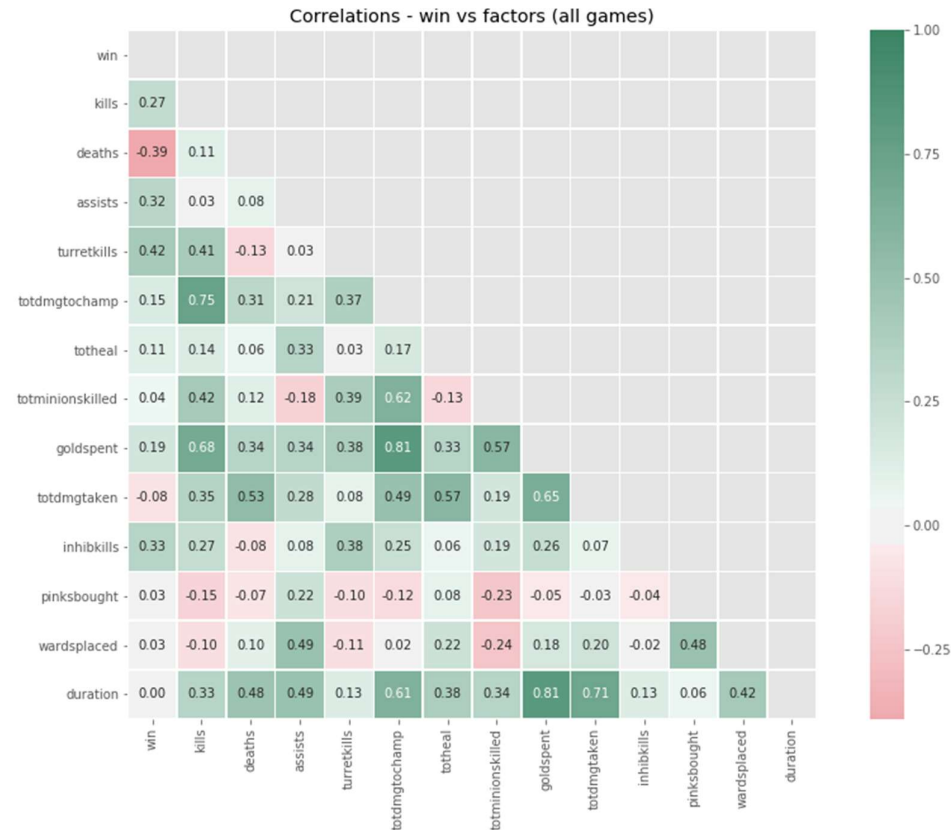
### Conclusion for the Advising Model

League of Legends is not only a operating level game but also a strategic game, especially in professional leagues when each player performance are close. By using this predictive model, professional team coaches can easily find out “Counter Champions” and “Anti Counter Champions” i.e. “Swingman”, which is quite helpful for strategy making. Normally, a team prefer to give the star player a counter pick to “snowball” on the lane. So, the counter strategy is always to pick a Swingman first for other players.

## Exploratory Analysis for other Influential Factors

After generated the predictive analysis, I wanted to conduct exploratory analysis for other influential factors in LOL game to explore which factors are most influential for win rate.

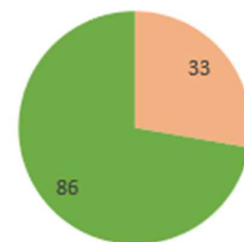
### The “Gold-diff”



Since League of Legends is a multifactor game, there are many influential factors affecting the match outcome. I generated this seaborn heatmap<sup>1</sup> above. In this plot, you can see the “goldspent” has the most “deep green” squares, which means more positive impact on victory. It affects game positively with duration,

What is “Gold-Diff”? How gold difference affect a game?

Negative gold difference at 15 minutes but Win
33
Positive Gold Difference at 15 and Win
86
Positive Gold Difference Win Rate
72.27%



“Gold-diff” aka. gold difference, it means the economics gap between both teams. In a LOL game, economic gap typically means fighting capacity gap, because the economics leader can

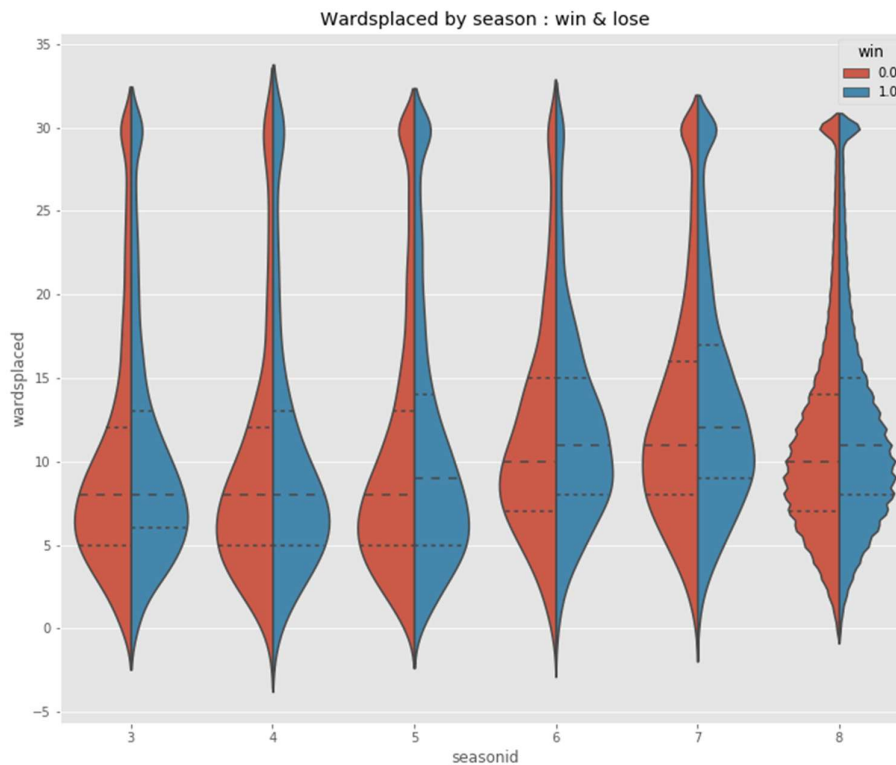
<sup>1</sup>More information about seaborn heatmap:  
<https://seaborn.pydata.org/generated/seaborn.heatmap.html>

purchase higher performance items(weapons) in the game, which could help the team sustain their advantages and keep snowballing.

### “Ward placed” Win Rate Analysis:

What is ward? A **ward** is a deployable unit that removes the fog of war in a certain area of the map in order to gain “**Vision Advantage**”. Typically, the more vision advantages the team has, the more useful information they get. The team has more information typically can take the lead on the battlefield by faster tactic adjusting, ambushing and wiser engagement.

This violin plot I generated below shows that the average amounts of ward have been placed in each game from Season 3 to Season 8 of League of Legends World Championship.



We can notice that the amounts of placed ward are increasing season by season, which means the **Vision Advantages** are getting more important and popular in LOL professional games.

### Conclusion

In this article, I conducted an explanatory analysis for explaining why those champions have always been banned; I created a model that can advise the best group of “counter-picks” based on the “KDA”, an important performance criteria; I conducted an exploratory analysis that explored the important correlation between the two influential factors (“Ward Placed” “Gold Diff”) and the game victory.