

Analise de resultados em E-Sports

Felipe Moraes¹, Hugo Teixeira², João Victor Costa³

¹Departamento de Estatística e Informática – Universidade Federal Rural de Pernambuco
(UFRPE)
Recife – PE – Brasil

{felipe.oficial.bsi,hugosteixeira22,joaovictorf3}@gmail.com

Abstract. *This article has as main goal to explain the methods used in the process to analyze the data collected from a game named league of legends. With these data an analis with the correlations of in game data and pro-leagues games was made, so the win predicts could be made.*

Resumo. *Este artigo tem como principal objetivo explicar os métodos utilizados no processo para analisar os dados coletados de um jogo chamado "League of Legends". Com esses dados foi feita uma análise com as correlações dos dados do jogo e ligas profissionais, para que possa ser feita a previsão do ganhador.*

2

1. Introdução

Vários esportes como o xadrez, pôquer, entre outros "esportes de mesa", são exemplos de esportes que necessitam raciocínio logico muito além da necessidade de capacidade física. Com o avanço da tecnologia e criação de jogos eletrônicos novos desafios que necessitavam de raciocínio lógicos surgiram. Assim como o xadrez e o pôquer em algum momento deixaram de ser apenas jogos para se tornar esportes profissionais, os jogos eletrônicos também passaram por essa transição. E-esportes ou "esportes eletrônicos", surgiram após a profissionalização dos jogos eletrônicos, ganhando enorme evidencia na ultima década, com premiações de campeonatos que ultrapassam premiações de grandes campeonatos de esportes como futebol e que possuem a capacidade de lotar estádios e arenas com a mesma força de esporte tradicionais.

Os E-esportes englobam diversos tipos de modalidades de jogos eletrônicos, jogos de luta, FPS(jogos de tiro), *battle royale*, *MOBAs*("Multiplayer Online Battle Arena"), dentre outros gêneros. Cada um desses gêneros possuem características específicas e modos de jogo específico.

Dentre os E-esportes o *MOBA* é o gênero mais jogado, esse gênero de jogo consiste em dois times de jogadores adversários(geralmente 5 versus 5) onde a condição de vitoria é derrubar a base do time adversário. Entre os *MOBA* disponíveis para se jogar o "League of Legends" mais conhecido como LOL é o mais jogado do gênero e além disso é o mais jogado dentre todos os jogos eletrônicos do mundo.

O "League of Legends" é jogado por 10 jogadores simultâneos onde cada um utiliza um herói diferente, cada jogador pode escolher entre os 144 heróis disponíveis para se jogar, cada um desses heróis possui 4 habilidades para serem utilizadas dentro do jogo

e podem se equipar com até 6 itens para aumentar sua força, que podem ser escolhidos entre os por volta de 290 itens disponíveis no jogo, e essas são apenas algumas das características desse jogo, e consequentemente o torna um jogo repleto de possibilidades que podem levar um time a vitória ou a derrota.

Entender melhor os fatores que levam um time a vitória, ou a derrota, é o objetivo desse artigo, para isso será analisado diversos dados extraídos de partidas profissionais de "League of Legends" que ocorreram em diversos países do mundo.

2. Metodologia

A metodologia usada neste trabalho científico, foi focada na análise dos dados quantitativos das bases de dados, ainda que não sejam dados absolutos pode-se retirar uma boa visão de probabilidade de vitória, dado algumas características que podem ser facilmente analisadas.

Procurando por bancos de dados de e-esportes mais especificamente do jogo LoL, "League of Legends" encontramos no Kaggle, uma comunidade online de cientistas de dados de propriedade do Google, um banco de dados com mais de 7 mil registros de jogos oficiais deste e-esporte registrados desde 2015.

Além dos dados encontrados no Kaggle fazia-se necessário o conhecimento de outros aspectos do jogo, como por exemplo as classes dos campeões e suas habilidades, a Riot Games, desenvolvedora do jogo tem disponível um arquivo em formato Json, com os dados de classes de cada personagem, porém os dados de habilidades não estão disponíveis em nenhum formato de arquivo ou base de dados, o que fez necessário a extração destes por meio de mineração de dados.

Fazendo uso da linguagem Python, que contém várias bibliotecas para plotagem de gráficos como matplotlib que possui vários tipos de modelos o que facilita a visualização destes dados e pandas para processamento e manipulação dos dados em csv.

2.1. Pré Processamento dos Dados

Antes de começar a execução das análises sobre nossos dados foram necessários alguns ajustes para que eles se aplicassem aos nossos interesses, e facilitassem nossas análises nos passos seguintes.

Para evitar ou utilização de vários data sets por análise ajustamos o principal data sets da Kaggle que se aplicava a nosso projeto o *matchinfo*, acrescentando nele algumas colunas existentes em outros data sets estudados e gerando um único data set que se aplicasse a vários casos e facilitasse nosso estudo o *matchinfo-bans*.

Outra alteração feita foi a criação de um data set a parte, onde os nomes dos heróis foram substituídos por suas classes no jogo gerando assim o segundo data set que utilizamos em nossas análises, esta substituição foi realizada em um data set a parte pois ele passou por outros ajustes como '*pivotar*' suas colunas o que iria afetar o desempenho de outras análises.

2.2. Analisando os Dados

Após o pré-processamento dos dados, em uma primeira análise realizada buscou-se encontrar duplas consideradas mais efetivas, ou as que possuíam mais vitórias, ob-

servando a figura 1 escolheu-se analisar os *outliers* ou pontos flutuantes que foram encontrados, estes vistos na representação do gráfico. Observando primeiramente em dados absolutos não foram levados em consideração dados relativos como ano das partidas e porcentagem de vitórias. A partir daí foram definidos matematicamente qual os valores que seriam utilizados, através dos cálculos dos quartis e da faixa interquartil para assim encontrarmos os valores relativos aos *outliers*.

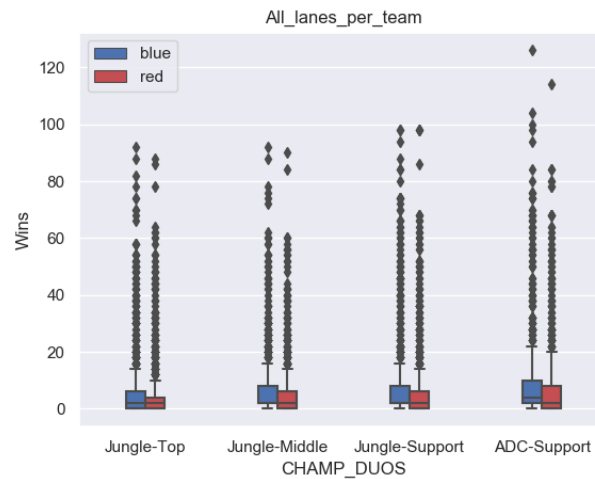


Figura 1. Vitórias absolutas de cada lane(duplas escolhidas) por time desconsiderando o ano.

Observando novamente os dados buscou-se a visualização as duplas analisando-as por ano abstraindo o time em que estava jogando assim obteriam-se todas as suas vitórias somadas. Figura 2

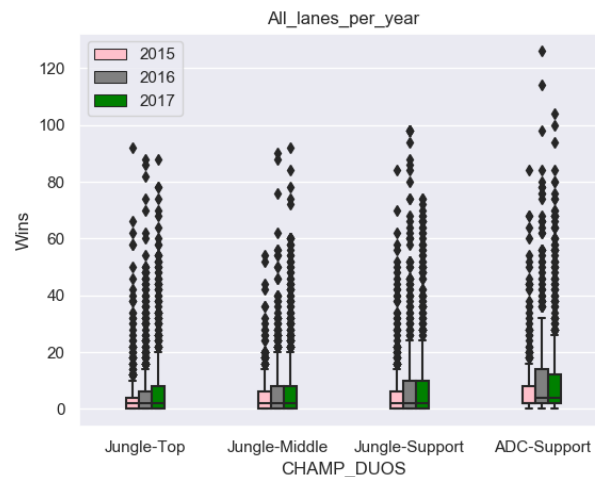


Figura 2. Vitórias absolutas de cada lane(duplas escolhidas) por ano desconsiderando o time

1º Quartil:

$$Q1 = \frac{X_n + 3}{4}$$

3º Quartil:

$$Q3 = \frac{3.X_n + 1}{4}$$

Faixa interquartis:

$$FiQ = Q3 - Q1$$

Assim foi percebido que a análise por lane levando em consideração não apenas um parâmetro como ano ou time mas o ideal seria analisar o desempenho de cada dupla em dado ano e time assim poderia-se melhorar os dados extraídos e obter os *outliers* com um maior nível de confiança visto as constantes mudanças em status e habilidades dos heróis e às diferenças de mapa referente ao time em que se estaria jogando figuras (3,4,5,6).

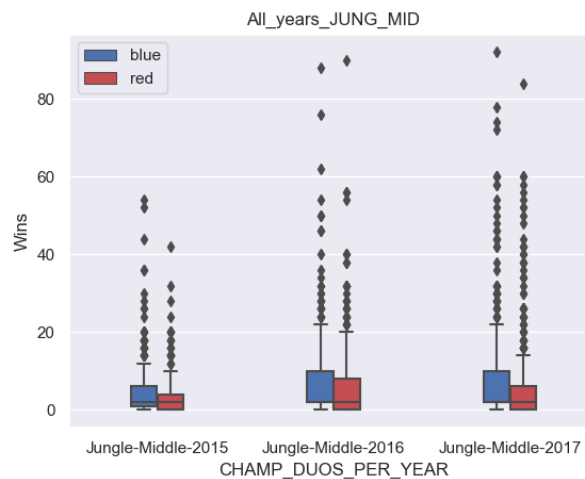


Figura 3. Vitórias absolutas da dupla Jungler-Midlaner separados por ano.

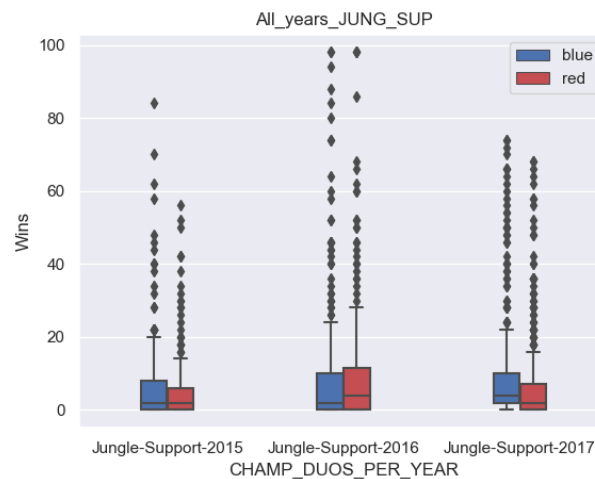


Figura 4. Vitórias absolutas da dupla Jungler-Support separados por ano.

Observando os gráficos podemos dizer que a grande maioria das duplas possuem um rendimento muito próximo, ou seja estão dentro das "caixas", sendo assim as duplas que estiverem fora destas "caixas" devem possuir uma maior possibilidade de vitória visto que não estão na distribuição comum de vitórias. Sendo assim as partidas que possuísem estes *outliers* em suas composições possuiriam uma maior probabilidade de vitória.

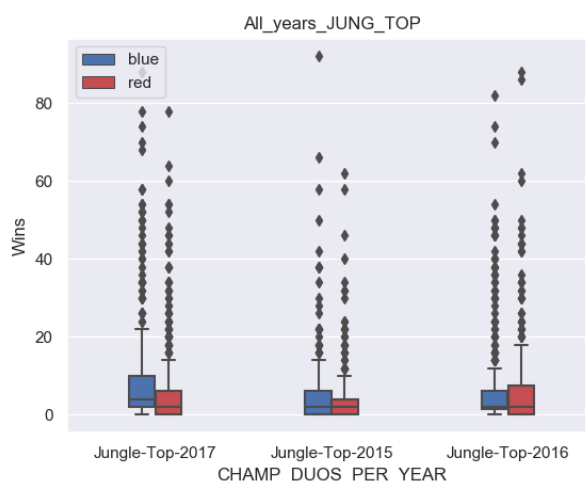


Figura 5. Vitórias absolutas da dupla Jungler-Toplaner separados por ano.

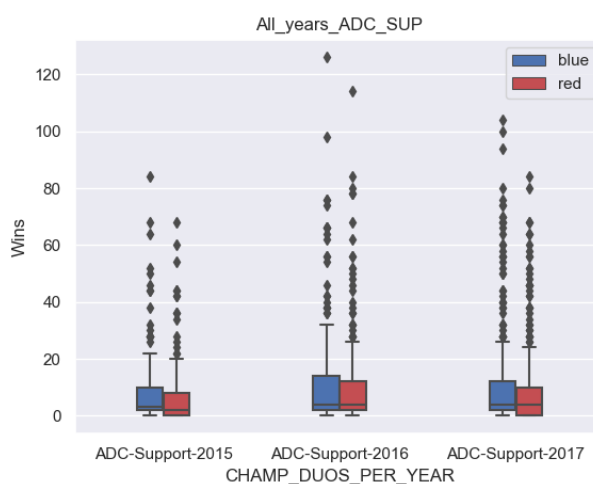


Figura 6. Vitórias absolutas da dupla ADCarry-Support separados por ano.

Confrontos individuais(*matchup*) são um grande indicio de vitoria para uma equipe, pois *champions* que ganham seus confrontos individuais conseguem distribuir essa vantagem para o time ao longo do jogo, muita vezes levando o time para a vitoria.

Para exemplificar esses confrontos foram utilizados dados do ano de 2017, para as lanes que são jogadas de forma individual pelos jogadores, logo, a *jungle*, o *top* e a *mid laner* foram utilizadas. Para cada posição apenas os *outliers* de vitorias de cada lado foram escolhidos, ou seja os confrontos onde um dos lados venceu muitas vezes.

Os confrontos individuais estão exemplificados nos gráficos abaixo:

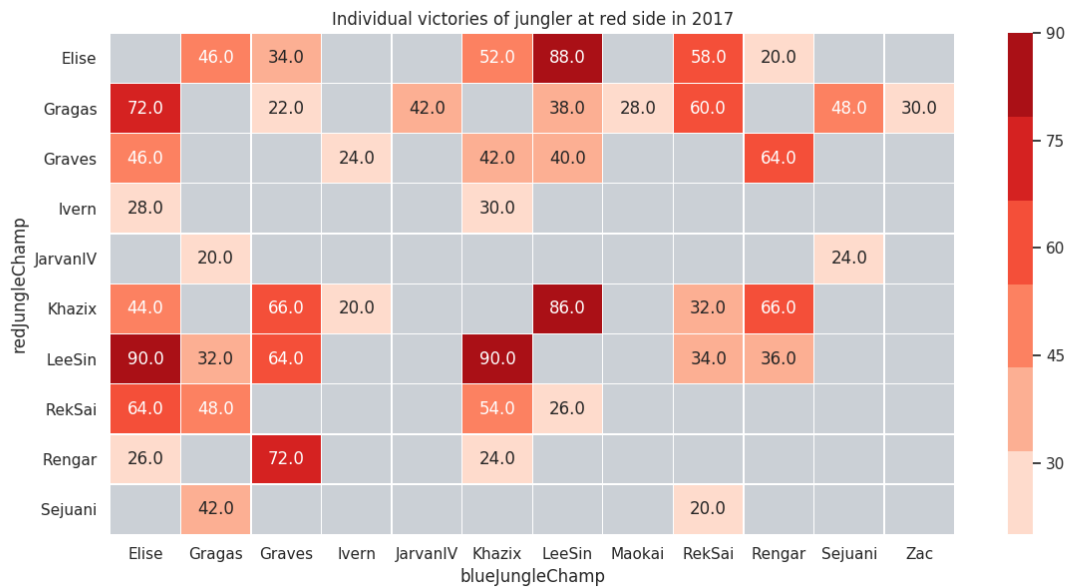


Figura 7. Vitórias absolutas dos junglers do lado vermelho em 2017

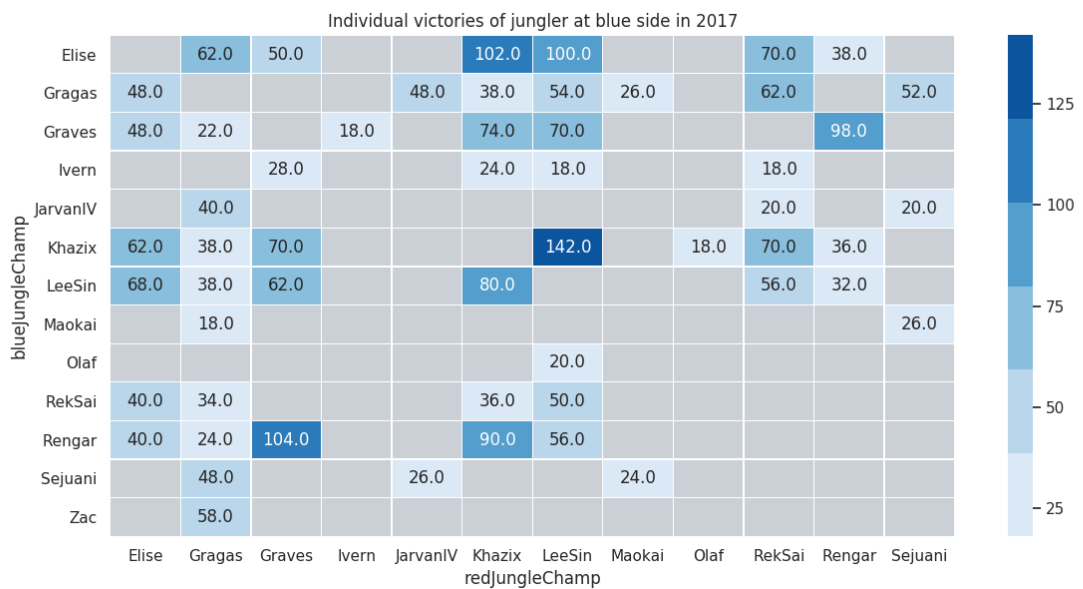


Figura 8. Vitórias absolutas dos junglers do lado azul em 2017

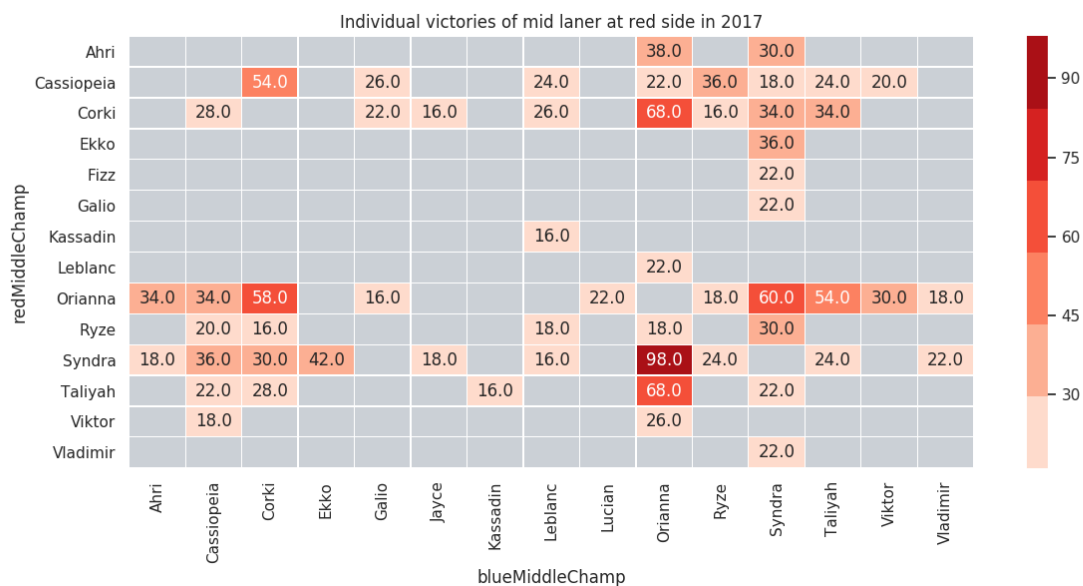


Figura 9. Vitórias absolutas dos mid laners do lado vermelho em 2017

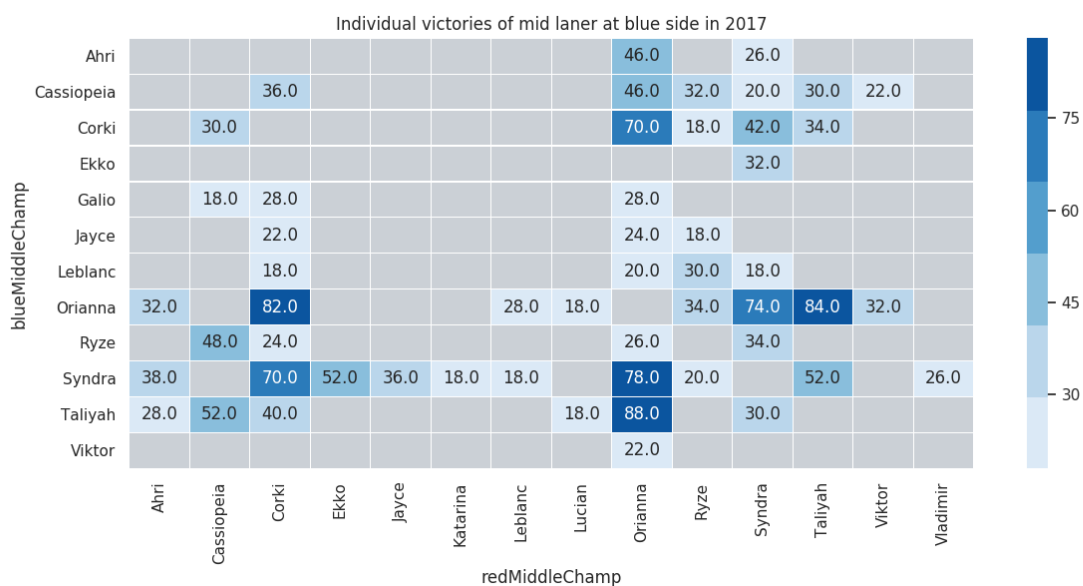


Figura 10. Vitórias absolutas dos mid laners do lado azul em 2017

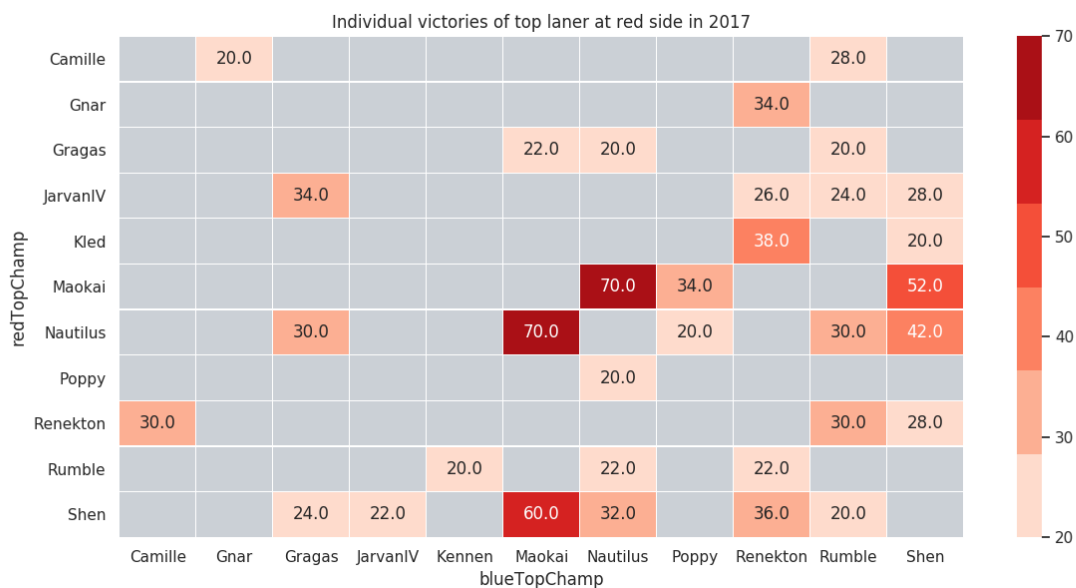


Figura 11. Vitórias absolutas dos top laners do lado vermelho em 2017

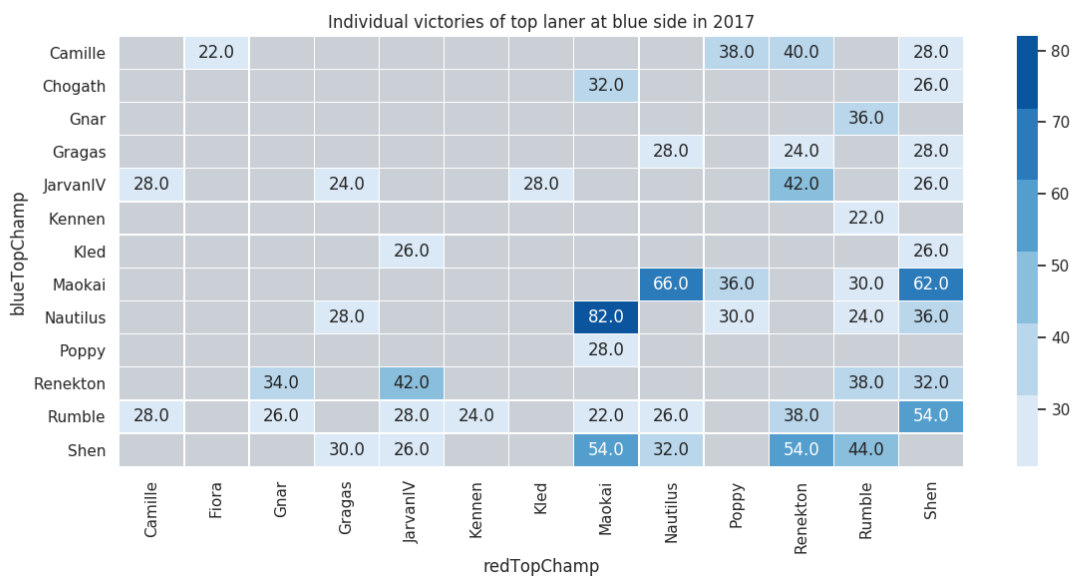


Figura 12. Vitórias absolutas dos top laners do lado azul em 2017

Pode-se observar alguns comportamentos no gráfico, como o confronto direto entre os heróis, em alguns casos existe uma dominância no confronto independentemente se o confronto é no lado azul ou lado vermelho.

Outro ponto a ser observado é a quantidade de confrontos, mesmo se utilizando de *outliers* das vitórias absolutas, em alguns casos um mesmo *champion* possui em seus

confrontos mais baixos 16 á 20 vitórias, enquanto seu confronto com mais vitoria gira em torno de 90 á 100 vitórias.

Em casos de muitas vitórias precisam existir muitos confrontos, e essa quantidade de confrontos podem ser interpretados de algumas formas, o *champion* A é um *counter* direto do *champion* B ou/e vice-versa, ou os dois *champions* fazem parte do Meta Game(conjunto de *champions* dentro do jogo que se sobressaem perante aos outros por características individuais) e por consequência jogam contra, não pela vantagem do confronto direto, mas sim pelos fato de fazerem parte do meta game atual.

Os gráficos de radar são apresentados para avaliar o comportamento de uma dupla em relação a outras. Saber qual é o *matchup* que se destaca como um *outlier*.

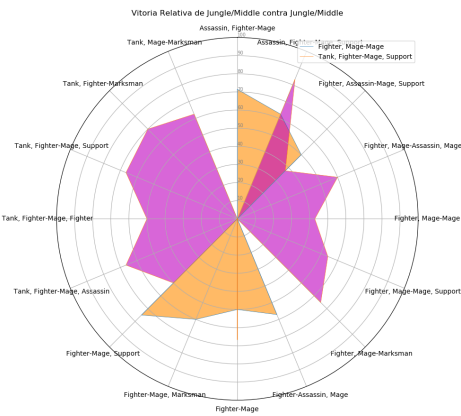


Figura 13. Vitórias relativas dupla Jungler-Mid Laner do time azul ,observando as duas duplas com mais partidas, ignorando o ano

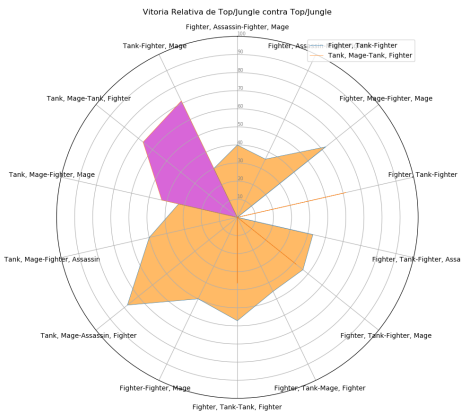


Figura 14. Vitórias relativas dupla Top Laner-Jungler do time azul ,observando as duas duplas com mais partidas, ignorando o ano.

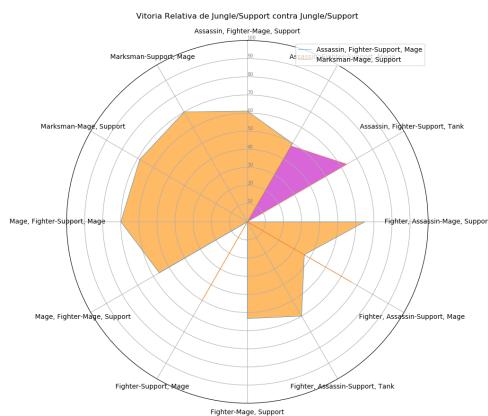


Figura 15. Vitórias relativas dupla Jungler-Support do time azul ,observando as duas duplas com mais partidas, ignorando o ano

Com os gráficos de radar é possível analisar a efetividade de uma dupla contra outras, algumas duplas com a maior quantidade de partidas foram selecionadas para serem comparadas com outras duplas. Uma dupla ter mais de 60% de taxa de vitórias é algo incomum, logo ela pode ser considerada superior nesta situação.

Um ponto de possível análise é que as duplas são em alguns pontos opostos, quando as áreas dos radares não coincidem tanto quer dizer que as duplas tem focos de adversários diferentes, mostrando a necessidade de adaptação e variação durante um campeonato para a seleção da melhor dupla.

No gráfico além da parte preenchida, a parte branca também nos informa dados, a vitória dos oponentes, se a área branca é maior que a colorida o time vermelho teve maior taxa de Vitória contra as duplas selecionadas. Com isso pode se afirmar que quanto mais perto do eixo central mais fraca a dupla selecionada é contra a do vértice.

Observando o gráfico da figura 15 vemos que a área laranja abrange uma área bem maior que a magenta, logo com isso podemos dizer que a dupla laranja é mais versátil que a dupla magenta, pois cada vértice representa uma dupla inimiga, se um conjunto ganha de várias duplas diferentes, pode ser concluído que taxa de versatilidade dele é maior baseado na área que o seu gráfico abrange.Figura 15

3. Teste de hipótese.

No LoL existe uma diferença entre a seleção de campeões do time vermelho e do time azul. Há uma especulação popular de que a equipe azul possui mais chances de ganhar que a equipe vermelha, senão assim foi analisado a veracidade desta informação. Foi assumido como Hipótese que: H-O lado azul possui mais chances de ganhar que o lado vermelho. H0-Os dois lados possuem a mesma possibilidade de vitória. Sendo assim todos os registros foram substituídos ao invés do nome do campeão foi colocado o tag deste campeão e analisado a hipótese nula deste caso, se um tipo os tipos de campeões possuíam uma maior possibilidade de vencer dependendo do seu time, estes dados foram analisados por cada lane. Este teste foi rodado para todos os registros estes que por sua maioria geraram um p-value superior a 0,10 e em pouquíssimos casos um resultado entre

0,01 e 0,05 o que nos deixa com pouca ou nenhuma evidência estatística contra nossa H0. Estes dados estão salvos em data sets para maiores validações.

BLUECHAMP	REDCHAMP	TAG_TO_COMPARE	Pvalue	Statistic
blueMiddleChamp	redADCChamp	Mage/Marksman	0.68	-0.42
blueMiddleChamp	redADCChamp	Mage	0.84	0.21
blueMiddleChamp	redADCChamp	Tank/Mage	0.37	0.91
blueMiddleChamp	redADCChamp	Mage/Assassin	0.69	0.41
blueMiddleChamp	redADCChamp	Marksman/Assassin		
blueMiddleChamp	redADCChamp	Fighter		
blueMiddleChamp	redADCChamp	Marksman	0.8	0.26
blueMiddleChamp	redADCChamp	Tank/Fighter		
blueMiddleChamp	redADCChamp	Fighter/Tank		
blueMiddleChamp	redADCChamp	Assassin/Fighter	0.63	0.48
blueMiddleChamp	redADCChamp	Fighter/Marksman	0.12	1.61
blueMiddleChamp	redADCChamp	Fighter/Assassin	0.23	1.27
blueMiddleChamp	redADCChamp	Mage/Support	0.24	1.18
blueMiddleChamp	redADCChamp	Fighter/Support		
blueMiddleChamp	redADCChamp	Support/Mage	0.39	0.88
blueMiddleChamp	redADCChamp	Assassin/Mage	0.26	1.14
blueMiddleChamp	redADCChamp	Mage/Fighter	0.41	0.84
blueMiddleChamp	redADCChamp	Mage/Tank	0.67	-0.43
blueMiddleChamp	redADCChamp	Marksman/Mage	0.81	0.25

Figura 16. Parte dos p-values gerados