

**UNIVERSIDADE DE SÃO PAULO  
ESCOLA DE ARTES, CIÊNCIAS E HUMANIDADES**

**ENTREGA 1: SELEÇÃO, PREPARO E ANÁLISE DE BANCO DE DADOS**

**ADA MARIA CAYRES FERNANDEZ - 9884922  
BEATRIZ THOMPSON SATHLER FREITAS - 8586728  
LEONARDO COPPOLA BIAZUCCI - 10723728  
MICHELLY RODRIGUES DA SILVA - 11270607  
THIAGO DE OLIVEIRA DEODATO - 10258938  
RAFAEL MARTINS CAMPOS - 10367158**

**SÃO PAULO  
2021**

## Explicação do dataset

O banco de dados escolhido para a análise é o 'League of Legends Diamond Ranked Games (10 min)', que é uma base que contém informações de aproximadamente 10 mil partidas ranqueados do top 1,5% jogadores (jogadores classificados entre os ranques Diamante e Mestre) do jogo League of Legends.

O League of Legends é um jogo online onde 10 jogadores jogam simultaneamente dividido entre 2 times de 5 jogadores. Cada jogador dentro do jogo possui uma função específica e o objetivo é destruir o Nexus do time adversário, que pode ser alcançado a partir da destruição das torres de cada uma das 3 rotas do jogo. Cada rota possui 3 torres e o Nexus é defendido por mais 2 torres. Os jogadores devem defender as torres de sua rota e destruir a torre dos adversários, enquanto acumulam ouro que pode ser utilizado para comprar equipamentos para fortalecê-los e experiência que pode aumentar o nível de cada jogador.

A análise a ser realizada verificará se certos conceitos, partindo de uma tese central, se relacionam entre si ou não. Um exemplo a ser dado seria a relação entre a diferença de ouro e abates obtidos pelo time azul, em que é esperado que ambas variáveis tenham uma relação proporcional entre elas.

O dataset possui 18 atributos para cada time (sendo 36 ao total) dos dados de cada time até os 10 minutos da partida. Desse modo, resolvemos escolher os seguintes atributos:

- **blueWardsPlaced** (Sentinelas posicionadas pelo time Azul)
  - As sentinelas são itens que um jogador pode utilizar para revelar a área próximo de onde é posicionada ao seu time e desse modo é utilizada para tomadas de decisões significativas e prevenir ser emboscado pelo time inimigo em uma situação desvantajosa que irá favorecer o time inimigo.
- **blueKills** (Abates realizados pelo time Azul)
  - Os abates realizados por um time fornecem ouro adicional ao time que realizou o abate e permitindo que comprem equipamentos melhores e aumentem seu poder no jogo e o jogador que foi abatido fica fora do jogo por um período de tempo, impedindo que ele colete ouro por esse tempo, enfraquecendo esse jogador conseqüentemente.
- **blueDeaths** (Mortes do time Azul)
  - Semelhante aos abates, as mortes contém o aspecto negativo do seu jogador ficar temporariamente fora do jogo sem poder coletar ouro enquanto o jogador que o abateu recebe ouro extra de recompensa por esse abate, além de continuar coletando o ouro de outros modos no jogo.

- **blueEliteMonsters** (Monstros Épicos derrotado pelo time Azul)
  - Monstros que possuem muita vida e são mais difíceis de serem abatidos por possuírem muita vida e dano. O time que abate esses monstros ganham uma grande vantagem em ouro e experiência.
- **blueGoldDiff** (Diferença de ouro do time Azul)
  - A diferença de ouro pode ser utilizada como parâmetro para facilitar o entendimento da vantagem de um time em comparação ao outro. O time com uma vantagem de ouro em geral consegue comprar mais equipamentos e consequentemente obtém mais poder no jogo do que o time inimigo.

## 2 - Pré-processamento dos dados

Primeiramente nós procuramos por dados faltantes das variáveis escolhidas, e esse foi o resultado obtido:

```
blueWardsPlaced    0
blueKills           0
blueDeaths          0
blueEliteMonsters   0
blueGoldDiff        0
dtype: int64
```

Após não termos encontrado nenhum dado faltante, nós prosseguimos para a normalização desses dados, e obtivemos o seguinte resultado:

	blueWardsPlaced	blueKills	blueDeaths	blueEliteMonsters	blueGoldDiff
0	0.316980	0.935254	-0.046924	-0.879186	0.256215
1	-0.570963	-0.393196	-0.387777	-0.879186	-1.191194
2	-0.404474	0.271029	1.657340	0.719467	-0.483590
3	1.149426	-0.725309	-0.387777	0.719467	-0.544323
4	2.925312	-0.061084	-0.046924	-0.879186	-0.415112
5	-0.237985	-0.393196	-1.069482	0.719467	0.278634
6	-0.237985	0.271029	-0.046924	0.719467	0.976863
7	-0.348978	-0.393196	2.339046	-0.879186	-1.071765
8	-0.348978	0.271029	0.293929	-0.879186	-0.812528
9	-0.515467	-0.725309	-0.387777	0.719467	-0.636849

## 3 - Análise Descritiva

### 3.1 - Medidas de resumo

- **Média:**

Variável	Valor da Média
blueWardsPlaced	22.288288
blueKills	6.183925
blueDeaths	6.137666
blueEliteMonsters	0.549954
blueGoldDiff	14.414111

- **Mediana:**

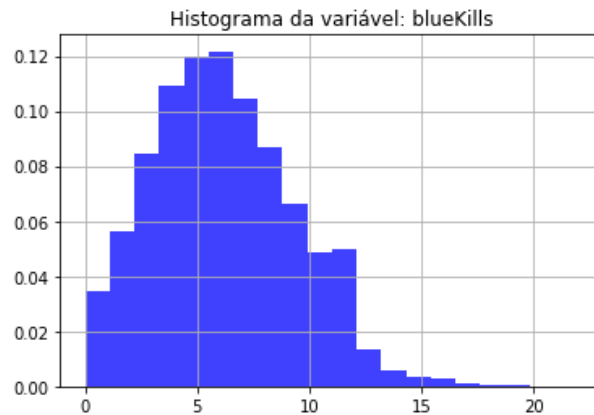
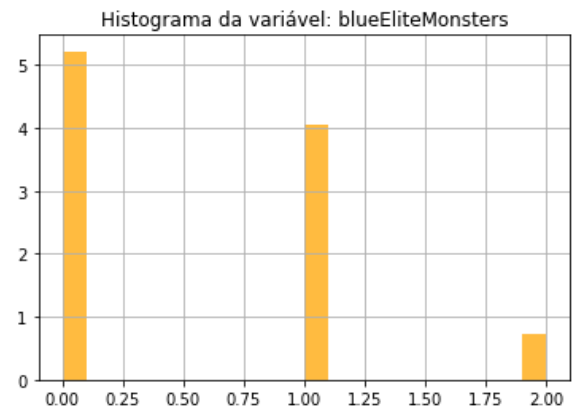
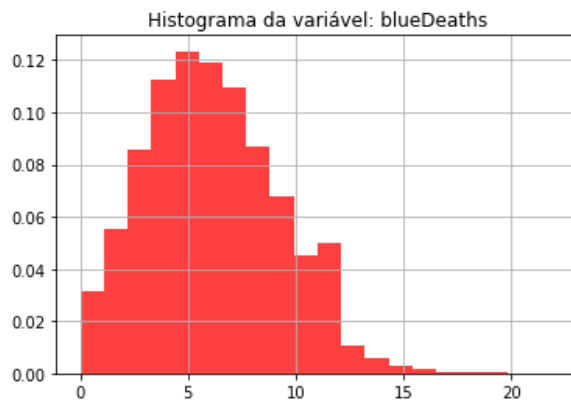
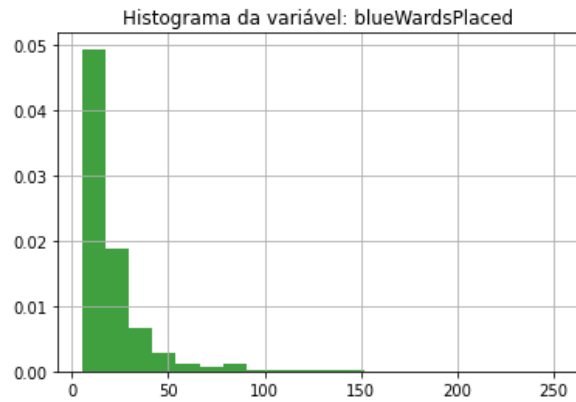
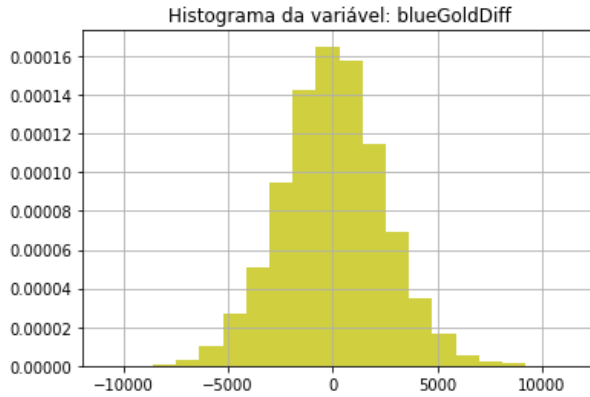
Variável	Valor da Mediana
blueWardsPlaced	16.000000
blueKills	6.000000
blueDeaths	6.000000
blueEliteMonsters	6.000000
blueGoldDiff	14.000000

- **Moda:**

Variável	Valor da Moda
blueWardsPlaced	16.000000
blueKills	6.000000
blueDeaths	5.000000
blueEliteMonsters	0.000000
blueGoldDiff	428.000000

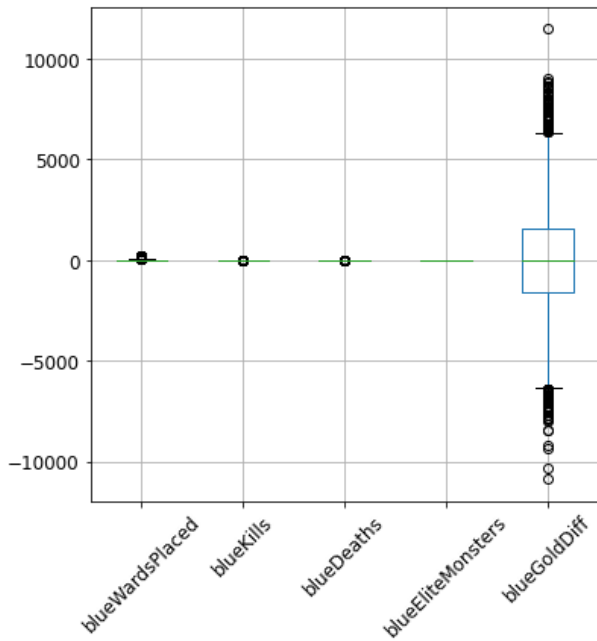
## 3.2 - Histogramas e Medidas Separatrizes

Nesse dataset, nós obtivemos os seguintes histogramas para as 5 variáveis:

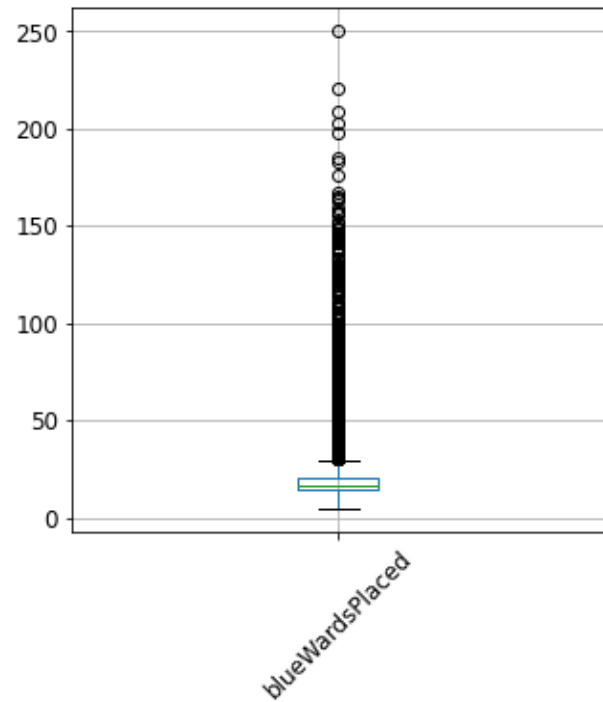


Nesse dataset também obtivemos os seguintes boxplots:

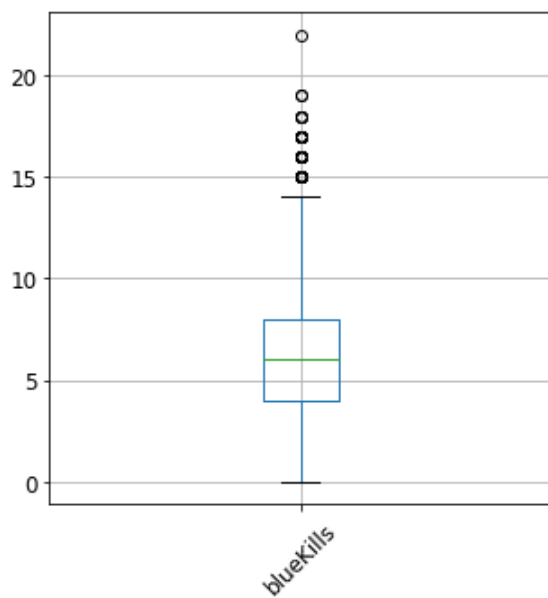
Boxplot do dataframe:



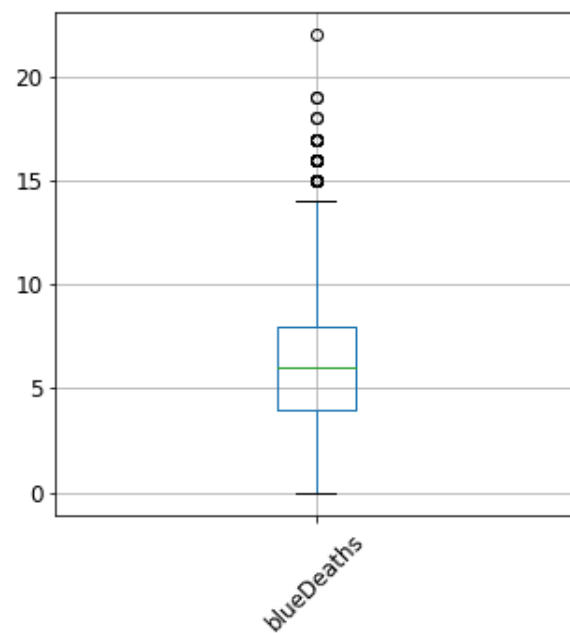
Boxplot da variável blueWardsPlaced:



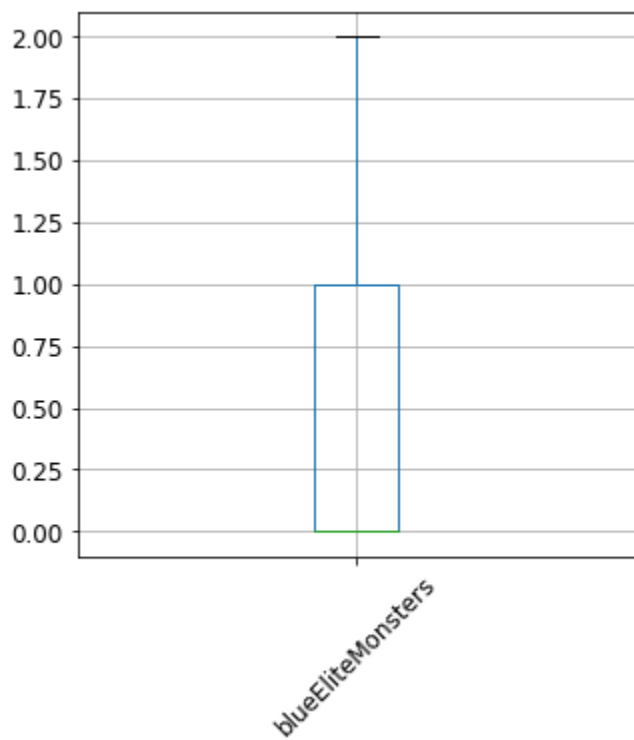
Boxplot da variável blueKills:



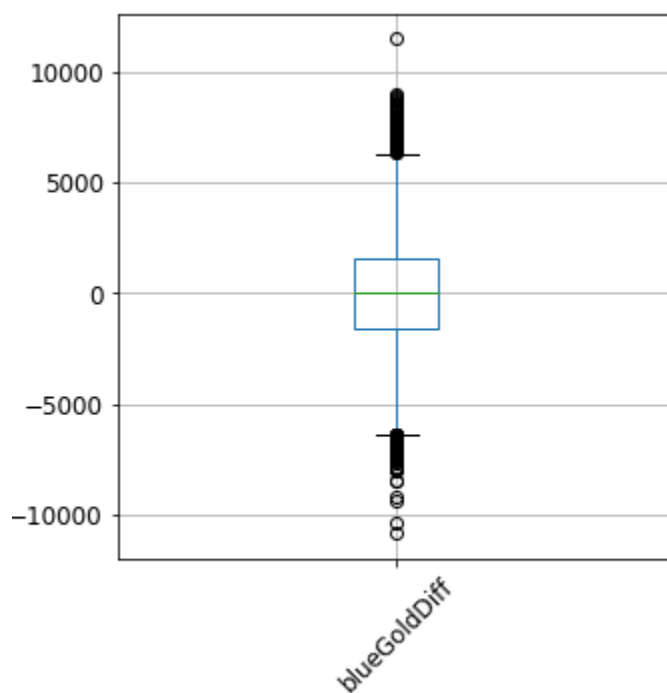
Boxplot da variável blueDeaths:



Boxplot da variável blueEliteMonsters:



Boxplot da variável blueGoldDiff:



### 3.3 - Medidas de variabilidade:

Foi calculado a variância, desvio padrão e coeficiente de variação das 5 variáveis escolhidas pelo grupo neste dataset. Os resultados obtidos foram:

- **Variância:**

Variável	Valor da Variância
blueWardsPlaced	324.690722
blueKills	9.066289
blueDeaths	8.607286
blueEliteMonsters	0.391283
blueGoldDiff	6018922.196079

- **Desvio Padrão:**

Variável	Valor do Desvio Padrão
blueWardsPlaced	18.019177
blueKills	3.011028
blueDeaths	2.933818
blueEliteMonsters	0.625527
blueGoldDiff	2453.349179

- **Coeficiente de Variação:**

Variável	Valor de Coeficiente de Variação
blueWardsPlaced	80.845942
blueKills	48.691207
blueDeaths	47.800220
blueEliteMonsters	113.741511
blueGoldDiff	17020.468509

Link para **gráficos extras** de moda, média, mediana, variância, desvio padrão e coeficiente de variação:

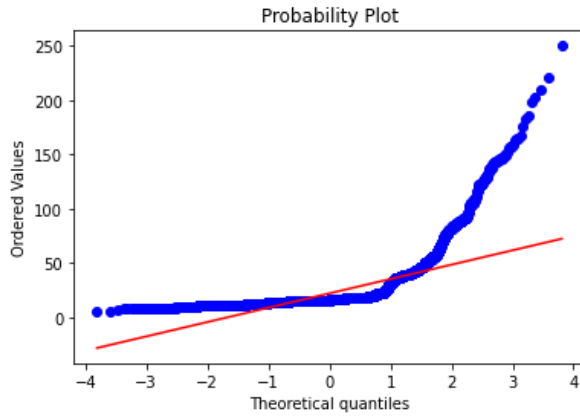
<https://drive.google.com/file/d/1uNKkfaYSgFP6KE7bk6R7cOx7Sas-2i2E/view?usp=sharing>



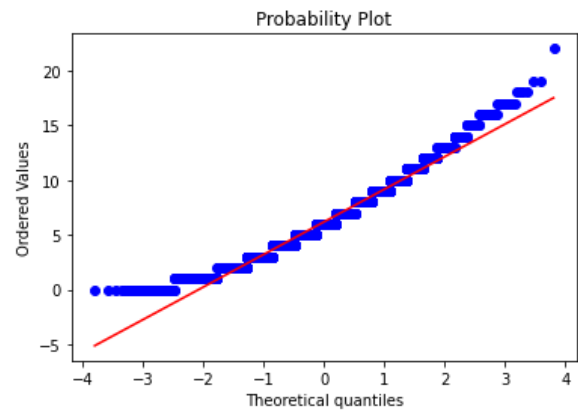
## 4 - Avaliação de Normalidade

A distribuição normal e faixa de normalidade obtidas pelo grupo foram:

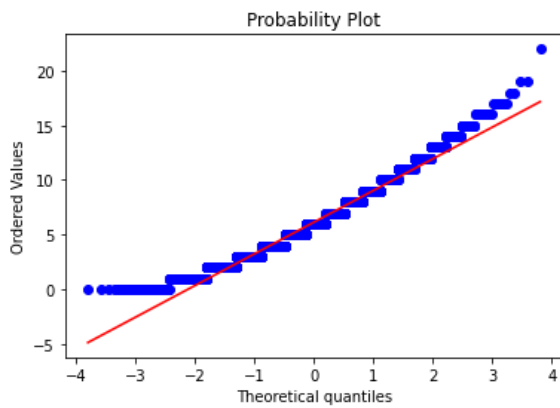
blueWardsPlaced



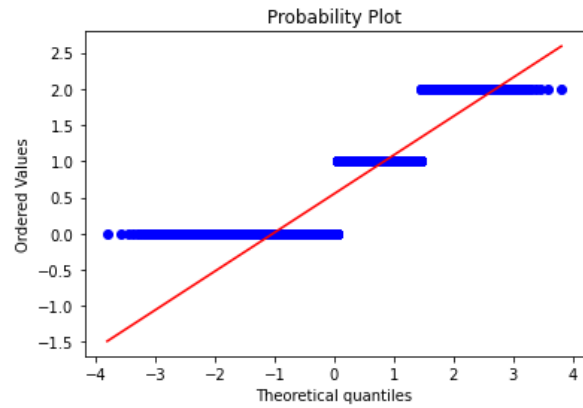
BlueKills



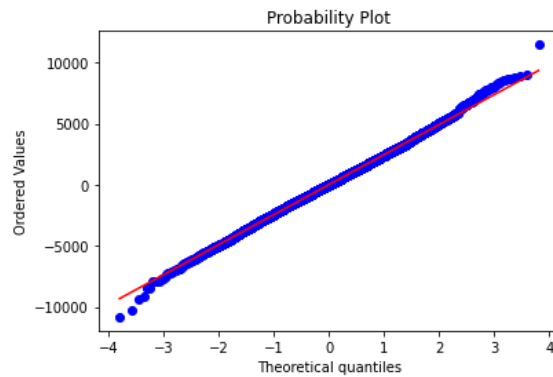
blueDeaths



blueEliteMonsters



blueGoldDiff



## 5 - Análise dos resultados

A partir dos dados obtidos pelo dataset, podemos tirar algumas conclusões sobre o andamento das partidas até os 10 minutos de jogo.

A primeira conclusão, é a de que ambos os lados do jogo são balanceados. Da mesma forma que o time azul morreu em média 6 vezes, ele também matou 6 vezes, o que significa que tanto o time vermelho quanto o time azul, estão balanceados no jogo, visto que se a pessoa de um time morre, o outro ganha um abate. Outra prova desse balanceamento é a diferença de ouro aos 10 minutos, que está muito próxima de zero (uma vantagem grande seria algo em torno dos 2000 de ouro de diferença).

A segunda conclusão é a de que os jogadores não fazem os objetivos assim que eles aparecem no mapa, demorando alguns minutos para tal, pois embora seja possível abater dois monstros elite no período de 10 minutos, a média do abate dessas criaturas pelo time azul é de 0,5 abates. Considerando que ambos os lados são balanceados, nós temos em média 1 abate de objetivo durante essas partidas, ao invés de 2.

Por fim, a terceira conclusão que podemos obter é a de que os jogadores no elo diamante utilizam quase todas as suas wards (sentinelas) disponíveis. Até os 10 minutos, cada jogador obtém 4 wards gratuitas, e podem comprar outras na loja do jogo. A média obtida pelo dataset nos diz que os jogadores do time azul colocaram 22 sentinelas no jogo durante esse período, o que nos diz que eles estão utilizando praticamente todas as wards disponíveis e ainda comprando outras na loja.