

## 2ª Etapa

Eniedson Fabiano Pereira da Silva Junior - 119110512

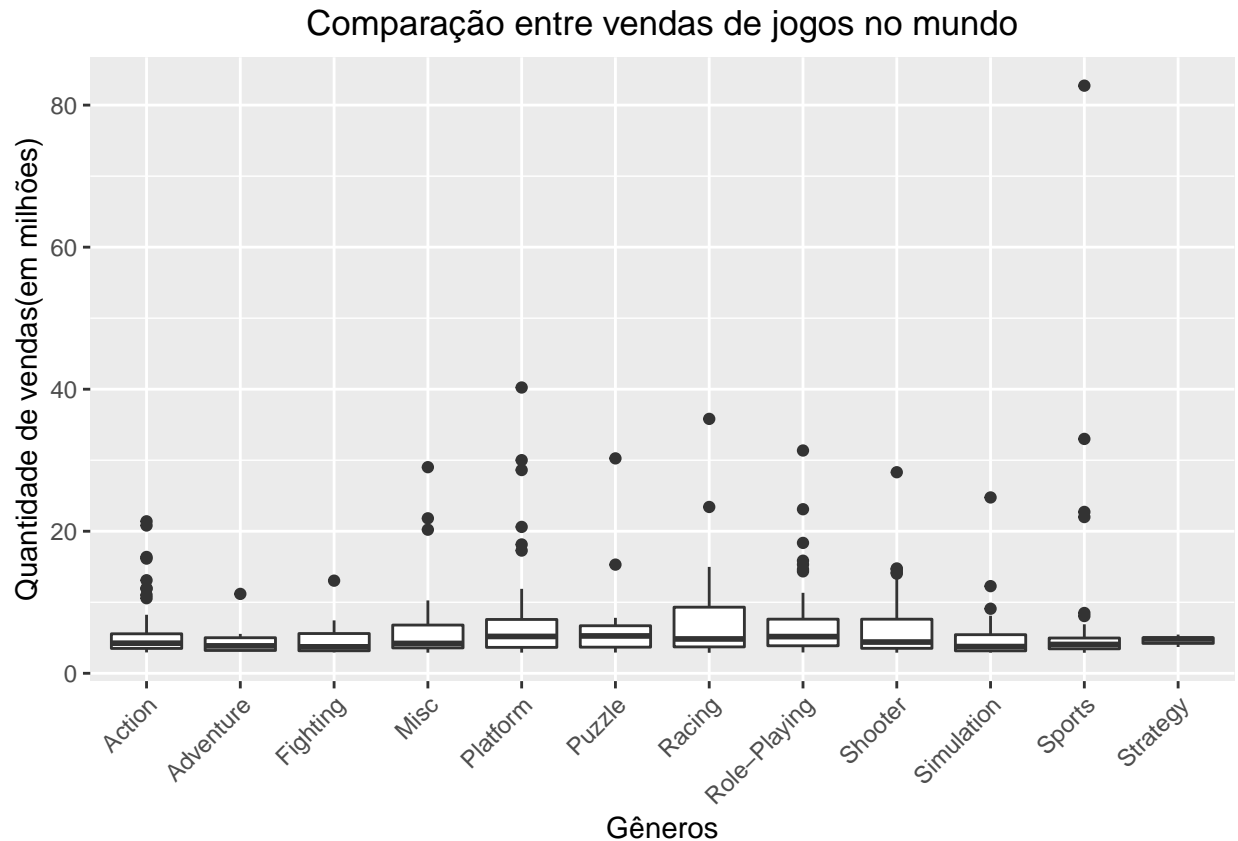
João Paulo Alves Dantas - 119210258

### Descrição do problema e variáveis

#### Análise

##### Global

Na primeira parte iremos analisar os números no cenário mundial, o primeiro passo é exibir esses dados de maneira a analisar as médias, medianas e o IQR. Para isso escolhemos uma visualização baseada em um `boxplot`, que podemos ver a seguir:



Analisando o resultado exibido no gráfico podemos chegar a suposição que de fato o gênero não influencia nas vendas totais de um jogo, já que todas as médias apresentam resultados bem-parecidos, e somente alguns casos específicos fogem dessa curva, levando a acreditar que outros fatores influenciem mais que o gênero.

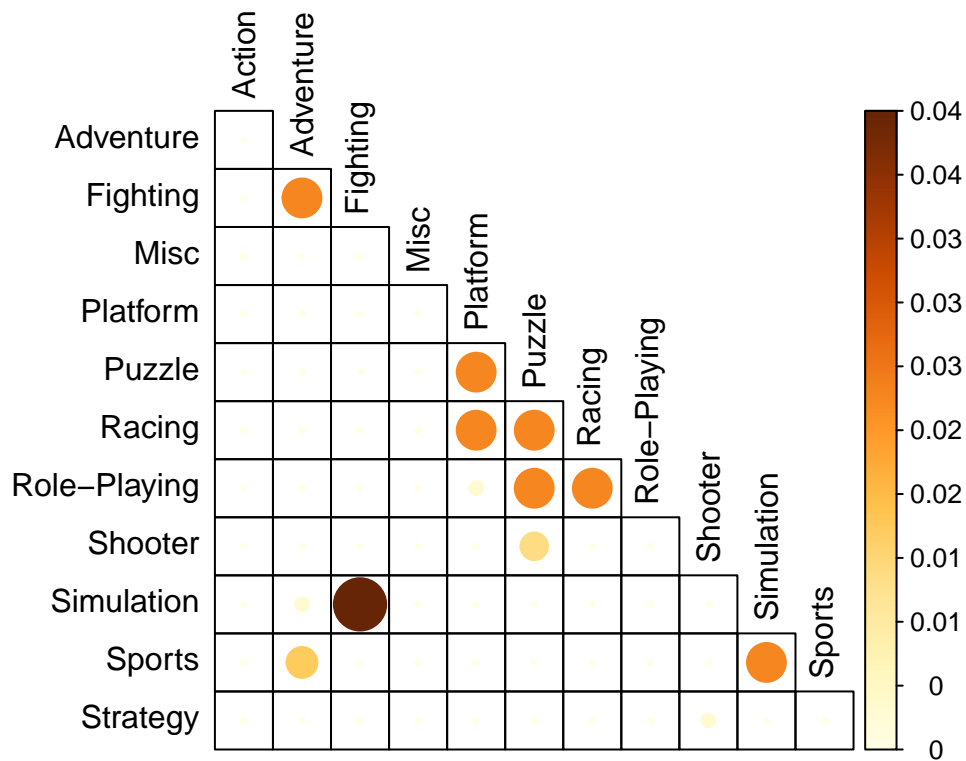
Para dar continuidade a nossos testes é importante testar a normalidade dos dados, e para isso utilizaremos o teste de Shapiro-Wilk nos grupos.

Genre	W	P.value
Sports	0.3127973	4.011810e-15
Platform	0.6247107	8.162158e-11
Racing	0.6565865	1.131400e-07
Role-Playing	0.6849663	7.141094e-10
Puzzle	0.6397710	8.002547e-05
Misc	0.5892975	9.334666e-10
Shooter	0.7477099	4.163680e-10
Simulation	0.5926912	3.749323e-06
Action	0.6253278	4.746311e-15
Fighting	0.7249293	2.190581e-05
Adventure	0.6919917	1.827851e-03
Strategy	0.9692752	8.705817e-01

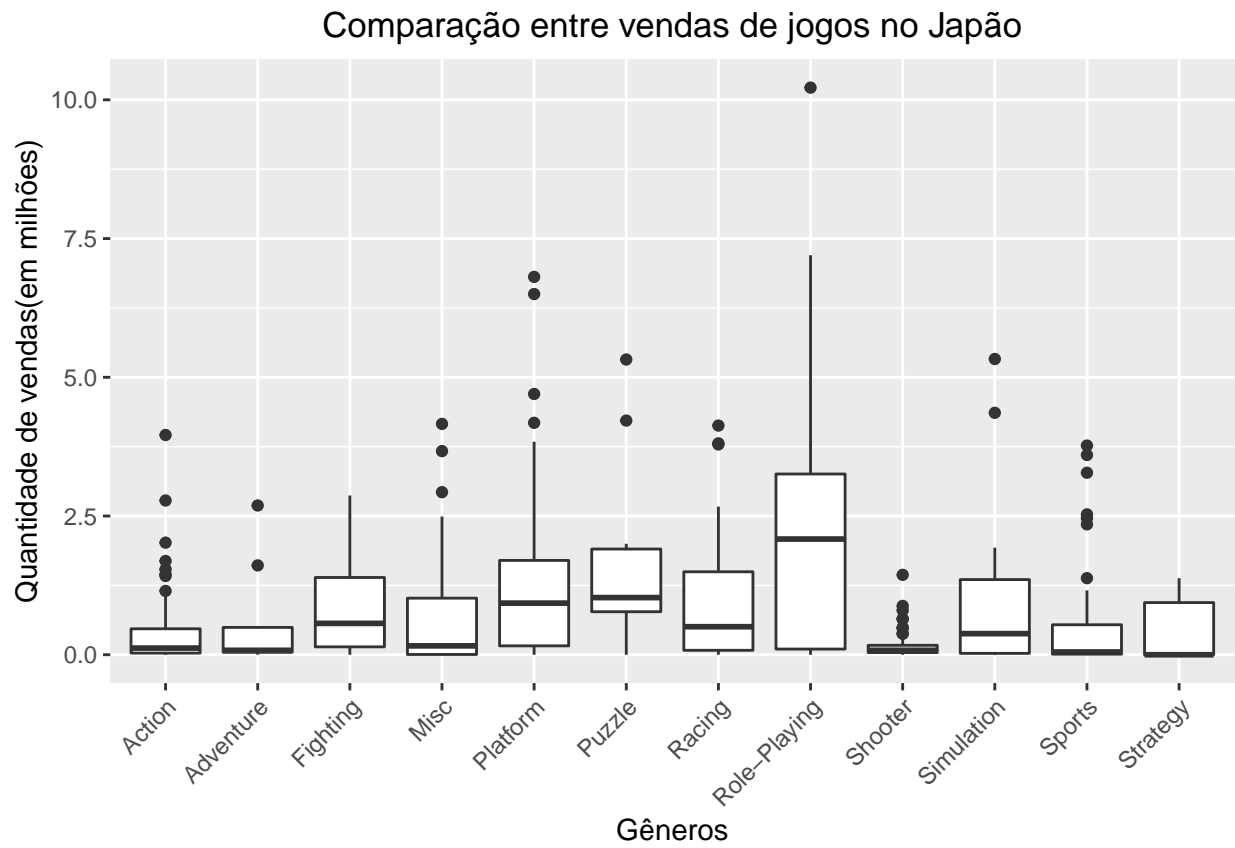
Como o valor de p foi menor que 0,05 em pelo menos um dos grupos, e seguindo a orientação do Canal Pesquisa, escolhemos o teste de Kruskal-Wallis para determinar de fato se essas variáveis são independentes.

Ao aplicar esse teste obtemos um valor de p igual a 0.0545696, ou seja, maior que o valor definido de 0,05, e por esse motivo podemos bater o martelo e definir que no cenário global gênero não é um fator determinante para as vendas de um jogo.

Porem o valor ainda é bem próximo ao nosso limiar, então para fazer uma análise mais precisa podemos ainda utilizar o teste de Pairwise Wilcoxon, que vai comparar as vendas entre cada um dos gêneros. Aplicando esse teste nos nossos dados obtemos o seguinte resultado:



Japão



```
##
## Kruskal-Wallis rank sum test
##
## data: JP_Sales by Genre
## Kruskal-Wallis chi-squared = 80.988, df = 11, p-value = 9.5e-13
```

