

Aluno: Pedro Juan Pereira dos Santos

Turma: TADS044

Tratamento e análise exploratória de dados

1. As mulheres tiveram uma taxa de sobrevivência muito maior que os homens? Sim, a diferença é muito significativa. A análise de frequência mostrou que:

Das 233 mulheres no dataset, 197 sobreviveram (aproximadamente 84.5%).

Dos 577 homens no dataset, apenas 109 sobreviveram (aproximadamente 18.9%).

Isso confirma a política de "mulheres e crianças primeiro", que foi priorizada no resgate do navio.

2. A idade média dos sobreviventes é diferente da dos não sobreviventes? Sim, existe uma diferença. As estatísticas mostraram que a idade média dos sobreviventes é ligeiramente mais baixa que a dos não sobreviventes, embora a mediana seja muito próxima.

A idade média dos que não sobreviveram é de cerca de 28 anos.

A idade média dos que sobreviveram é de cerca de 29 anos.

O box plot da idade também revelou que os sobreviventes tinham idades mais uniformes, enquanto o grupo que não sobreviveu tinha uma distribuição de idade mais ampla. Isso pode ser influenciado pelo maior número de crianças no grupo de sobreviventes.

3. Quem pagou mais caro pela passagem tinha maior chance de sobreviver? Sim, a tarifa paga pela passagem teve uma forte influência na taxa de sobrevivência.

A tarifa média dos que sobreviveram foi de cerca de R\$ 48,2 (valores originais, antes da padronização).

A tarifa média dos que não sobreviveram foi de apenas R\$ 22,1.

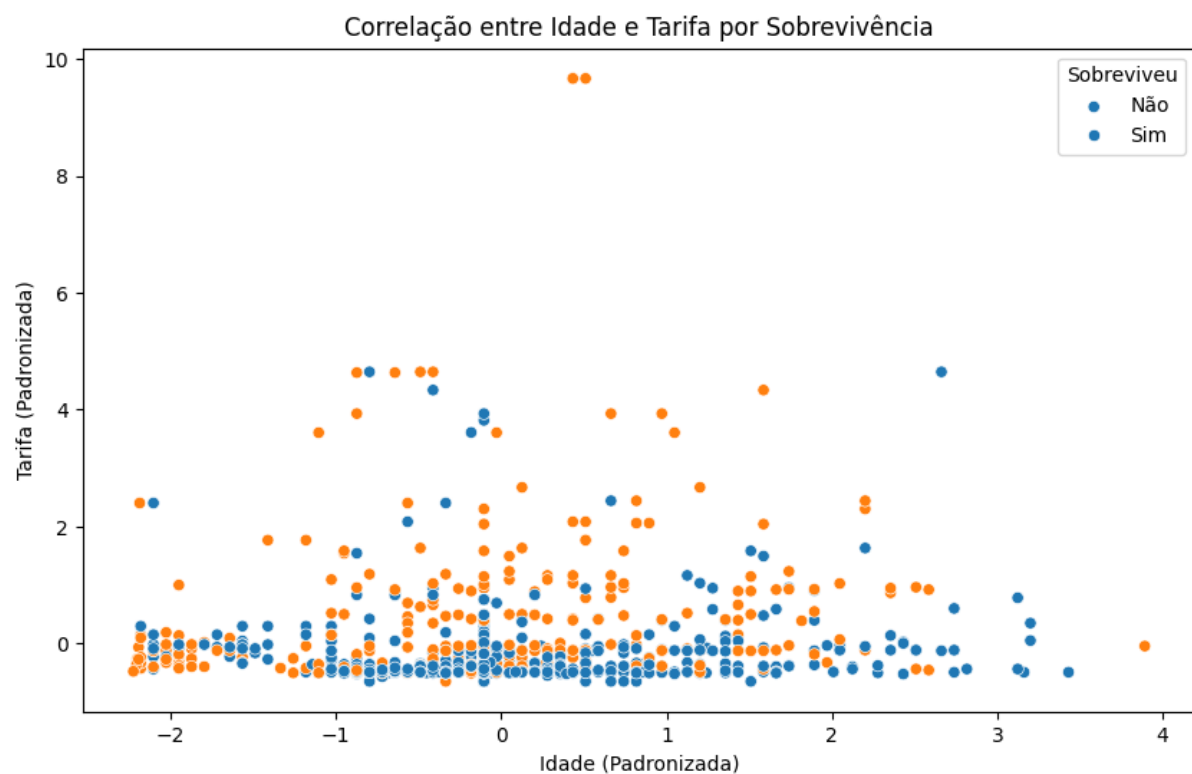
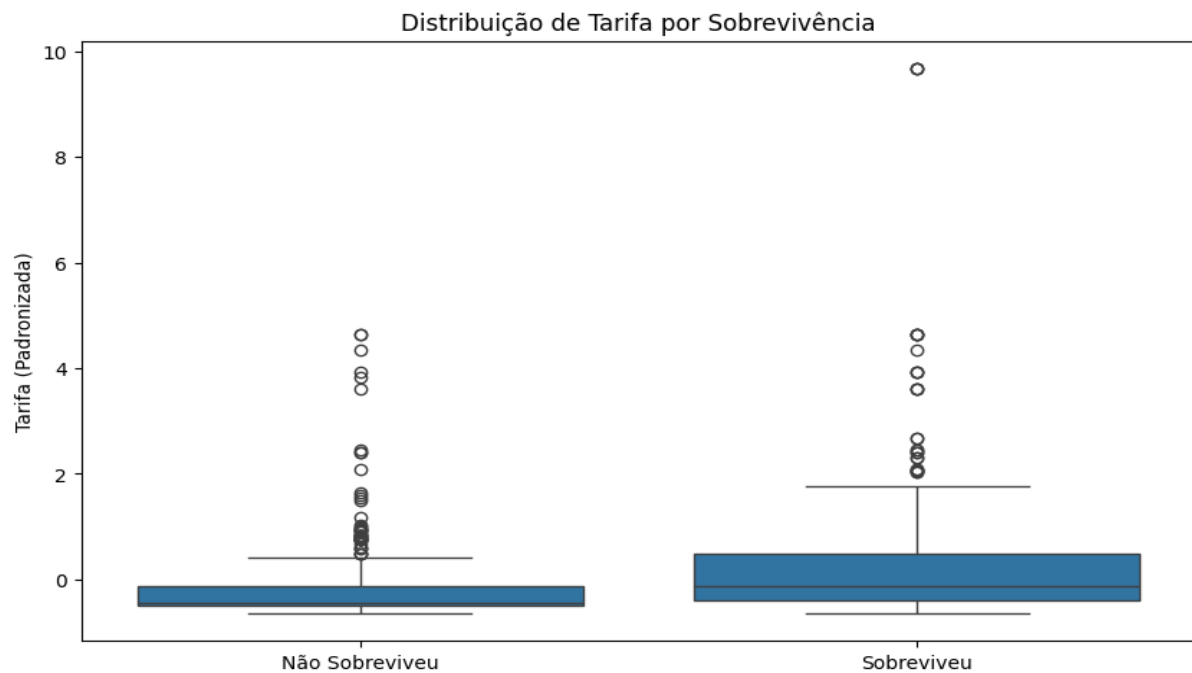
O gráfico de caixas para a tarifa mostra claramente que o grupo que sobreviveu pagou, em média, valores muito mais altos, e a presença de outliers de alta tarifa no grupo de sobreviventes reforça que os passageiros de primeira classe tiveram maior probabilidade de sobrevivência.

4. Existe alguma diferença clara na distribuição de idade ou tarifa entre os grupos que sobreviveram e os que não sobreviveram? Sim, as distribuições são bem diferentes:

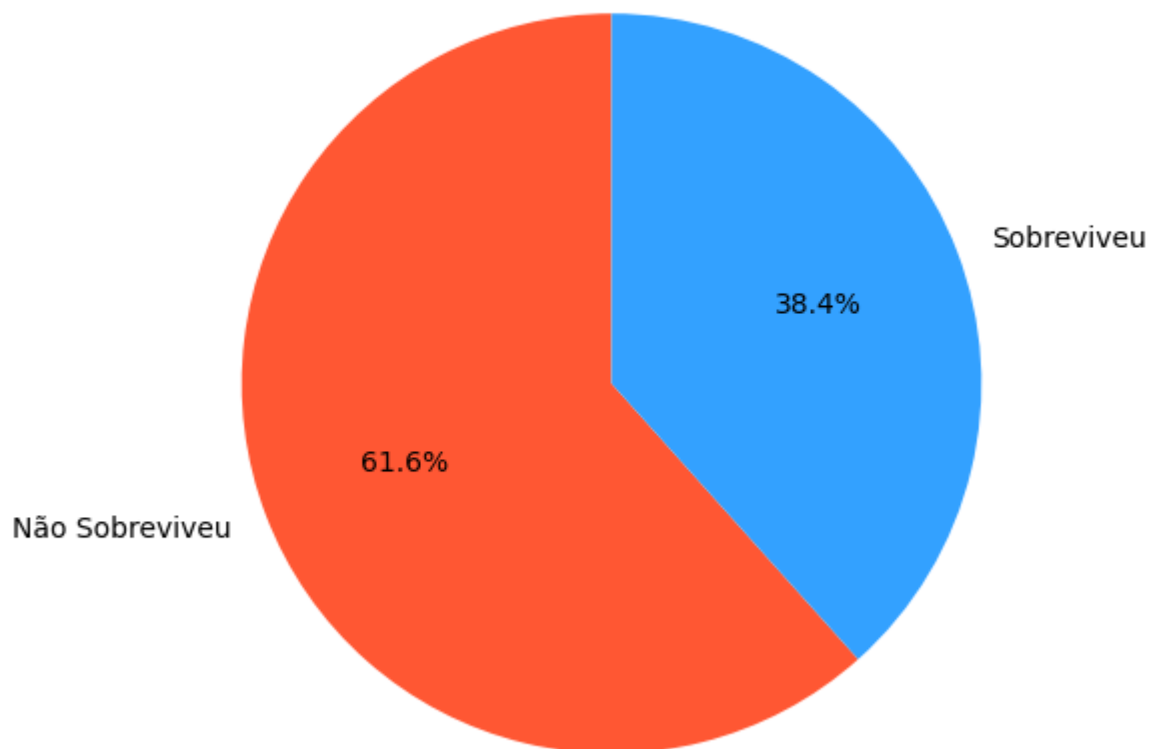
Distribuição de Idade: O histograma mostrou que a distribuição de idades entre os sobreviventes e não sobreviventes é similar, mas com um pico maior de crianças (faixa de idade mais baixa) no grupo de sobreviventes, indicando que as crianças tiveram maior chance de sobrevivência.

Distribuição de Tarifa: O box plot da tarifa deixou claro que a faixa de valores pagos por sobreviventes é, em geral, muito mais alta do que a dos não sobreviventes. A maioria das pessoas que pagaram tarifas mais baixas não sobreviveu, enquanto a maioria das que pagaram tarifas mais altas sim.

Conclusão Final: A análise mostra que a sobrevivência no Titanic não foi aleatória. Os fatores que mais influenciaram a chance de uma pessoa sobreviver foram o sexo (ser mulher), a classe social (refletida na tarifa da passagem) e, em menor grau, a idade (ser criança).



Proporção de Sobreviventes no Titanic



Distribuição de Idade por Sobrevivência

