

Investigando a base de dados dos tripulantes do Titanic

Introdução

O Titanic foi um navio de passageiros britânico, construído entre os anos de 1909 e 1911. Apesar de ter sido pensado para ser o navio mais seguro e luxuoso da época, no ano de 1912, durante uma viagem da Inglaterra para os Estados Unidos, ele naufragou após se chocar com um bloco de gelo, deixando um grande número de mortos no acidente.

Diante dos dados demográficos e informações de 891 dos 2224 passageiros e tripulantes à bordo do navio, disponível na nossa base de dados, ficamos diante de algumas indagações tais como:

- Quantas pessoas conseguiram sobreviver e quantas não sobreviveram?
- Qual foi a taxa de sobrevivência e mortalidade?
- Qual a média de idade geral da tripulação, dos sobreviventes e dos não sobreviventes?
- O navio era mais tripulado por homens ou mulheres?
- O sexo dos passageiros influenciou na taxa de sobrevivência das pessoas?
- Qual classe de embarque possuía maior número de passageiros?

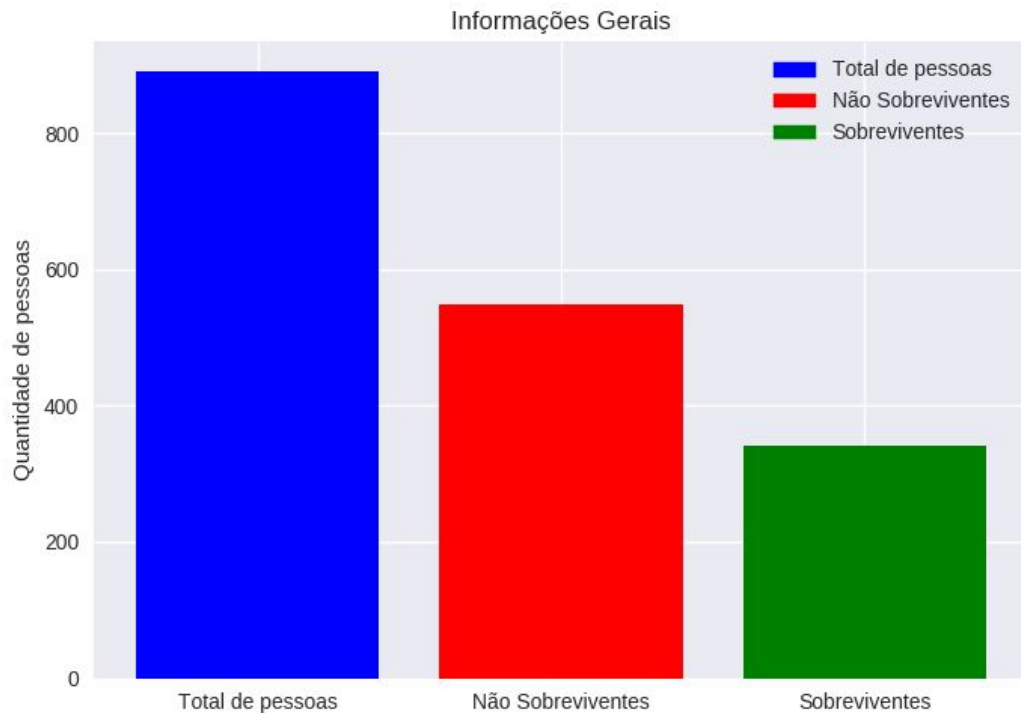
Tendo em vista essas questões, este presente trabalho visa levantar quais foram os possíveis fatores que contribuíram por deixar as pessoas que estavam dentro do navio, mais predispostas à sobrevivência.

Informações gerais

Conforme pode ser visto abaixo, dos 891 tripulantes analisados na base de dados do titanic, apenas 342 sobreviveram, totalizando uma taxa de sobrevivência de apenas 38,4%, enquanto que o número de não sobreviventes foi de 549 pessoas, o que nos leva à uma taxa de 61,6% não sobreviventes.

- Total de pessoas: 891
- Sobreviventes: 342 (38,4%)

- Não sobreviventes: 549 (61,6%)



Idade dos tripulantes

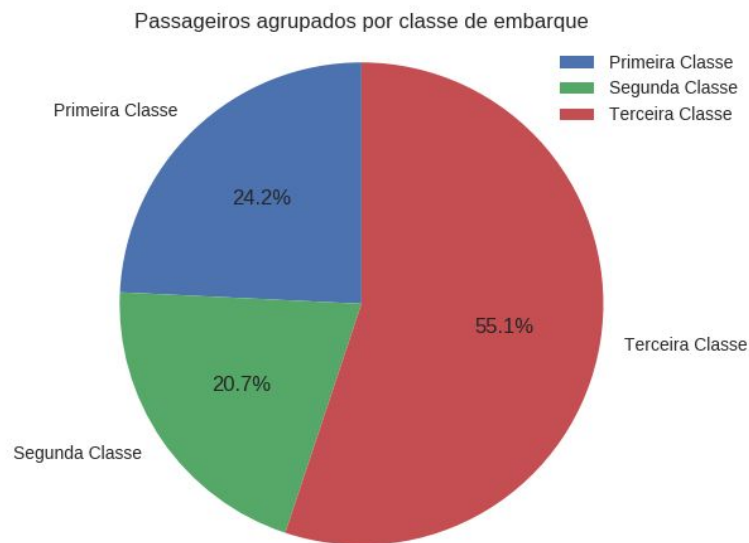
Ao verificarmos a base de dados do titanic, foi possível verificar que existe um grande número de passageiros, com idade inválida ou desconhecida. Dos passageiros sobreviventes 52 não tinham registro de idade e dos não sobreviventes 125. Para evitar erros nas informações de idade média, optamos por retirar esses valores inválidos dos nossos cálculos, por meio do uso da função *dropna*.

- Média de idade dos tripulantes: 29
- Média de idade dos sobreviventes: 28
- Média de idade dos não sobreviventes: 30
- Idades desconhecidas para os sobreviventes: 52 (17.9%)
- Idades desconhecidas para os não sobreviventes: 125(29.5%)

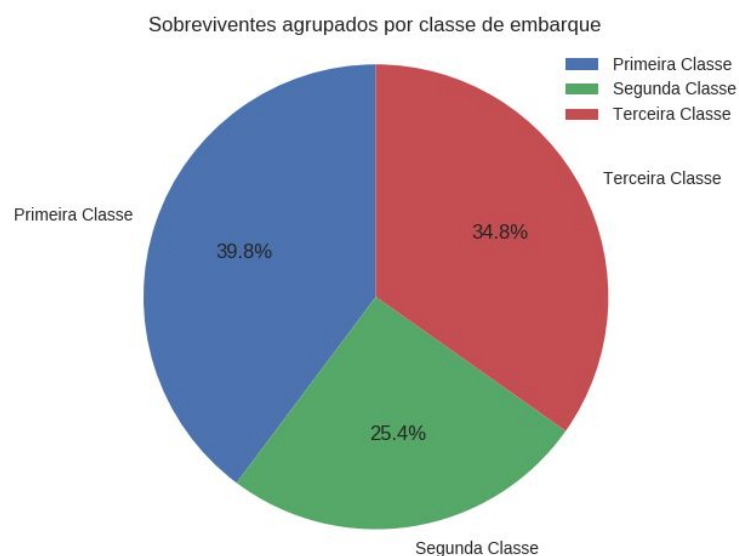
Tripulantes por classe de embarque

Podemos verificar na base de dados, que dos passageiros do titanic 491 (55,1%) eram da terceira classe, 216 (24,2%) pertenciam à primeira classe e somente 184 (20,7%) eram

pertencentes à segunda classe. Abaixo é exibido o gráfico mostrando a proporção de passageiros por classe de embarque.



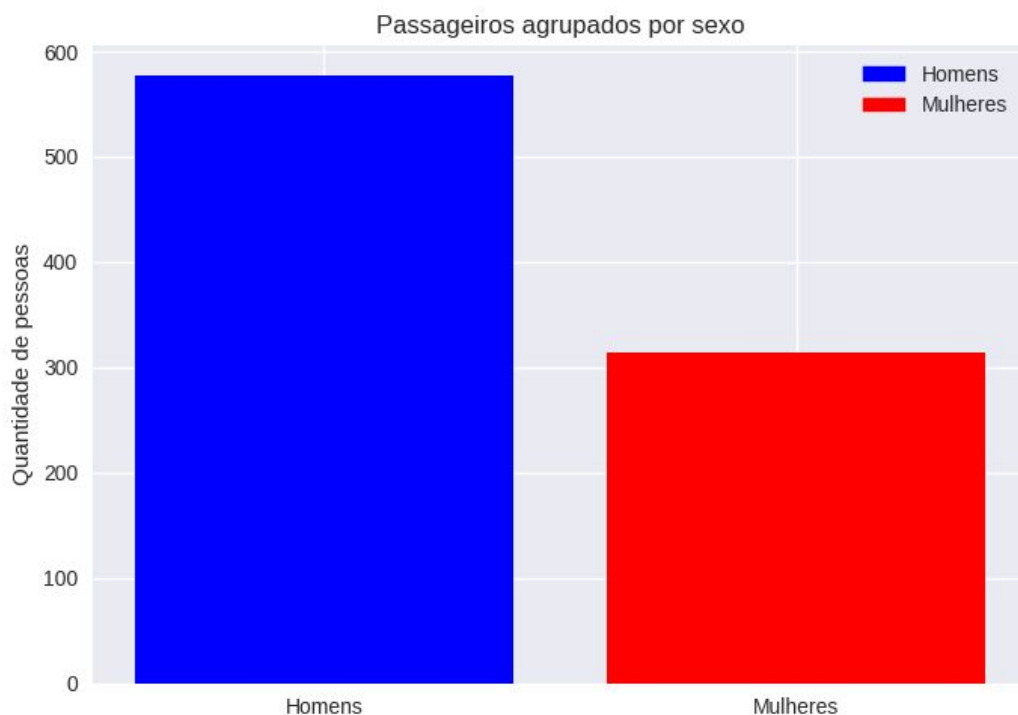
Em conformidade com os números exibidos abaixo, é possível verificar que a maioria dos sobreviventes do navio estavam concentrados na primeira classe 136 (39,8%), seguido pela terceira classe com 119 (34,8%) sobreviventes e por fim, a segunda classe com o meramente 87 (25,4%) sobreviventes.



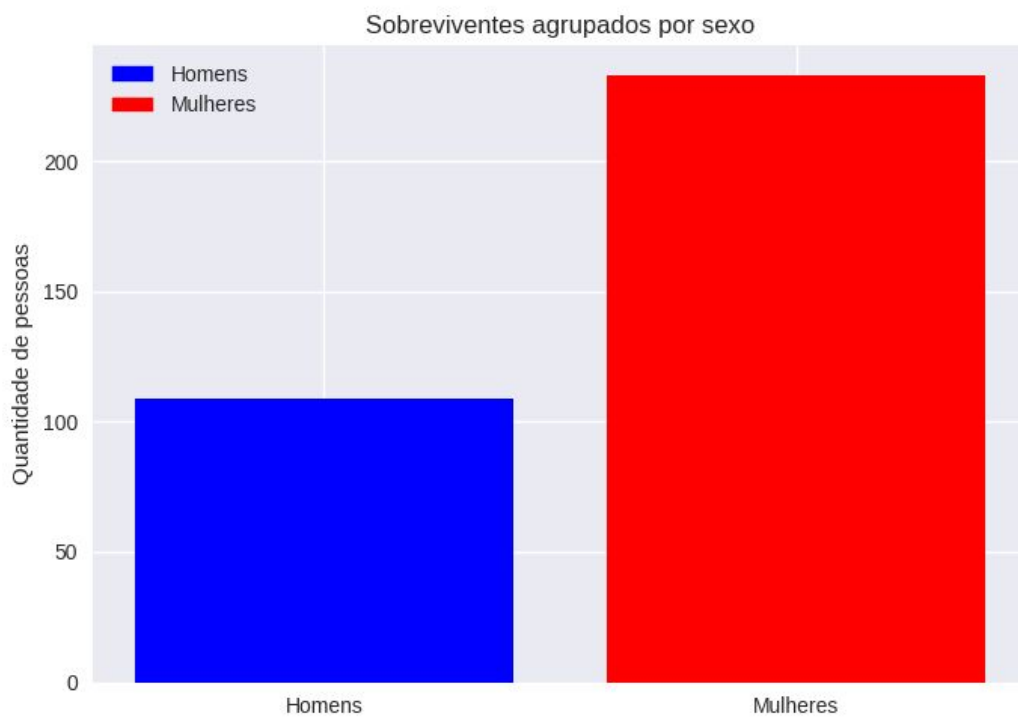
Se analisarmos o número de sobreviventes por classe de embarque em relação ao número de pessoas contidas naquela classe, podemos verificar, que a proporção de sobreviventes da primeira classe foi de 63%, a da segunda classe foi de 47% e a da terceira classe somente 24%.

Tripulantes por Sexo:

De acordo com a base de dados, a maioria dos tripulantes presentes, eram do sexo masculino. Os homens totalizavam 577 pessoas, com isso, eles representavam aproximadamente 64,8% da população do navio. As mulheres, contabilizavam 314 pessoas, o que denota que elas eram responsáveis por 35,2% da população do titanic, conforme pode ser visto abaixo.

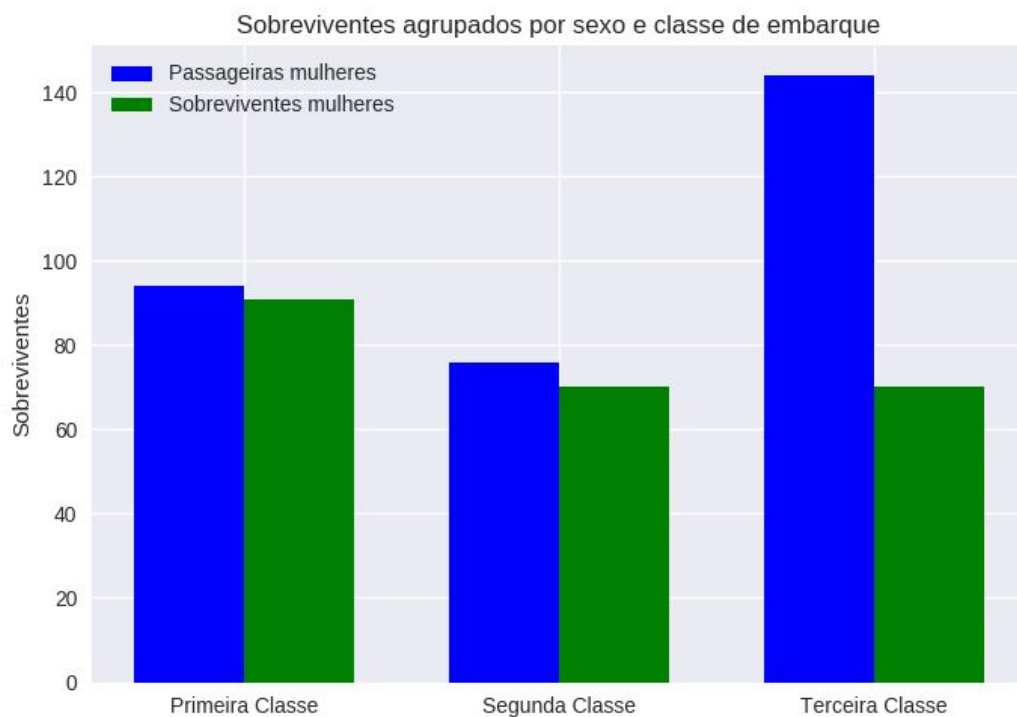


Ao analisarmos os sobreviventes dos sexos masculinos e femininos, foi possível notar que os números inverteram-se em relação à visualização anterior. A quantidade de sobreviventes do sexo feminino supera em muito o número de sobreviventes do sexo masculino. Sobreviveram 233 mulheres, ou seja, 68,1% das mulheres presentes, conseguiram sobreviver. Quanto aos homens, somente 109 sobreviveram, com isso, sobreviveram apenas 31,9% dos homens do navio.



Considerações finais

Os números indicam que os passageiros do sexo feminino que embarcaram nas classes superiores do navio, primeira e segunda classe, foram os passageiros que tiveram a maior taxa de sobrevivência.



Sobreviveram 91 passageiras da primeira classe, ou seja, taxa de sobrevivência de 97% para as mulheres que embarcaram na primeira classe, já na segunda classe, foram 70 sobreviventes do sexo feminino, o que nos dá uma taxa de sobrevivência 92% para as mulheres que embarcaram nessa classe. Conforme pode ser visto no gráfico acima.

Esses números, nos ajudam a levantar uma possível relação entre o gênero feminino e as classes de embarque superiores, com o aumento das possibilidades de sobrevivência. No entanto, nossa análise não é suficiente para provar que esses fatores foram a causa para aumento das possibilidades de sobrevivência no navio.

Dessa forma, nosso estudo atual é limitado pelo pressuposto de grande parte das mulheres da primeira e segunda classe conseguiram sobreviver. Para um estudo mais aprofundado, seria adequado a utilização de um *t-test* para determinar com maior precisão se a diferença na taxa de sobrevivência entre mulheres que embarcaram nas classes superiores e demais passageiros é estatisticamente significativa.

Referências:

matplotlib bar chart. Disponível em: http://matplotlib.org/examples/api/barchart_demo.html

matplotlib pie char . Disponível em:
https://matplotlib.org/examples/pie_and_polar_charts/pie_demo_features.html

stackoverflow. Disponível em:
<https://stackoverflow.com/questions/14734533/how-to-access-pandas-groupby-dataframe-by-key>

Unicode. Disponível em:
https://matplotlib.org/examples/text_labels_and_annotations/unicode_demo.html

Matthiaseisen. Disponível em: <http://matthiaseisen.com/pp/patterns/p0178/>

Legenda. Disponível em: https://matplotlib.org/users/legend_guide.html

Wikipedia. Disponível em: https://pt.wikipedia.org/wiki/RMS_Titanic