

*SQL Para Data Science*

SQL Para Data Science

Definindo o Problema

Usando SQL Para Limpeza e Transformação
de Dados de Câncer de Mama



Neste capítulo aplicaremos uma série de transformações a variáveis de um conjunto de dados de pacientes que desenvolveram câncer de mama.

O objetivo é limpar e transformar os dados através de categorização, codificação e binarização e então gerar um novo dataset que poderá ser usado mais a frente por outros profissionais da equipe de Ciência de Dados.

Vamos ao trabalho.