

Query Extra: Análise de Dados

Data: 31/01/2025

Candidato: João Paulo Oliveira Nolasco



Secretaria de
Saúde



Essa é uma análise sucinta dos dados do projeto desenvolvido, apenas com intuito de explorar também o papel do analista de dados numa operação de exploração de um dataset.

Acredito que o engenheiro de dados deve se colocar também no papel de quem vai usufruir dos dados tratados posteriormente, por isso me coloquei no papel de analista nesse material adicional, e realizei uma Query Extra, com intuito de atestar que as consultas anteriores realizadas no schema gold tem um propósito útil para análises.

Para isso, nessa consulta extra, o objetivo foi **filtrar as classes de produtos com baixa cobertura** de bairros, utilizando a tabela já criada no schema **Gold** (`produtos_classe_completo`). O que foi feito:

1. **Tabela de Referência:** Foi usada a tabela `produtos_classe_completo` no schema **Gold**, que já contém as informações agregadas sobre classes de produtos, bairros atendidos e registros.
2. **Filtro de Baixa Cobertura:** A consulta selecionou apenas as classes que possuem **menos de 10 bairros atendidos**. Essas classes foram identificadas como tendo "Baixa cobertura" em termos de distribuição geográfica.
3. **Seleção de Colunas:** A consulta retornou as colunas relevantes, como:
 - `classe`: A classe do produto.
 - `total_bairros_atendidos`: Quantidade de bairros onde o produto está presente.
 - `cobertura_bairro`: A classificação de cobertura como "Baixa cobertura" (para os casos com menos de 10 bairros atendidos).
4. **Ordenação:** Os resultados foram ordenados de forma decrescente pelo número de bairros atendidos.

O resultado dessa consulta fornece uma lista das classes com **baixa cobertura**, gerando o insight de que podem precisar de ações estratégicas de expansão para aumentar a presença em mais bairros.

Para consultar o código dessa consulta adicional, acesse a pasta do github "Extra".