

# PROJETO DE ANÁLISE DE DADOS

# Vendas de Supermercados

## Descrição do Desafio

Um determinado supermercado está passando por um processo de modernização e contratou uma ferramenta para capturar e armazenar os dados referentes às vendas realizadas.

Pensando em utilizar os dados armazenados para impulsionar o negócio, um analista de dados foi contratado para tratar, analisar e construir um dashboard com os dados coletados de forma a gerar valor ao negócio.

Nesse *case*, você fará o papel do analista de dados e desenvolverá soluções utilizando SQL e Power Bi para anteder os objetivos necessários.

Para a avaliação do resultado produzido, serão avaliados os seguintes critérios:

- Performance e clareza do código em SQL;
- Performance e clareza das medidas do power bi;
- Assertividade no objetivo estipulado pelo negócio e visualizações escolhidas no power bi.

# Tabelas Utilizadas

## Tabelas

- Fato\_Vendas;
- Dim\_Produto;
- Dim\_Margem\_Bruta\_Perct.

## Chaves entre tabelas:

- Fato\_Vendas e Dim\_Produto: id\_linha\_produto
- Fato\_Vendas e Dim\_Margem\_Bruta\_Perct: utilizar id\_linha\_produto da Fato\_Vendas e as colunas da tabela Dim\_Margem\_Bruta\_Perct. (sugestão: transformar a estrutura dos dados da Dim\_Margem\_Bruta\_Perct)

Obs: documentação das tabelas enviada em arquivo complementar.

# Fase 1

## S Q L

Nessa fase serão realizados os tratamentos/manipulações de dados utilizando o SQL. A ideia é utilizarmos o SQL para as transformações mais complexas a fim de simplificarmos o trabalho que será feito posteriormente no Power Bi.

# Tratamento de Dados em SQL

## a) Tratamento da coluna “data” da tabela Fato\_Vendas.

A máquina que coleta de informações teve um problema e colocou todas as datas para o ano de 2019. Será necessário passar todas as datas com ano 2019 para 2022 mantendo o dia e o mês. Realize as transformações necessárias.

## b) Criação da nova coluna “vlr\_margem\_bruta” na Fato\_Vendas

Será necessário utilizar as tabelas: Fato\_Vendas, Dim\_Produto e Dim\_Margem\_Bruta\_Perct. Abaixo segue o cálculo:

- “vlr\_preco\_total” (coluna auxiliar): “vlr\_preco\_unitario” (Fato\_Vendas) acrescido do imposto (Dim\_Produto);
- “vlr\_margem\_bruta” (coluna objetivo): é o perc\_marg\_bruta (tabela Dim\_Margem\_Bruta\_Perct) do “vlr\_preco\_total”.

Obs importante: por um erro no processo, alguns registros não possuem valor na coluna “vlr\_preco\_unitario” da Fato\_Vendas, dessa forma não seria possível calcular a coluna “vlr\_margem\_bruta”.

Dessa forma, antes de se calcular a coluna “vlr\_margem\_bruta”, será necessário corrigir esse problema preenchendo os registros vazios da coluna “vlr\_preco\_unitario” com o valor médio dos últimos 20 dias (a partir da “dta\_venda” do registro) dos clientes pertencentes ao mesmo grupo.

Utilizar para grupo as colunas: nom\_sexo, id\_linha\_produto, num\_filial.

Caso, após o tratamento, ainda exista algum registro com o “vlr\_preco\_unitario” vazio por não existência de registros que atendam o critério para construção média, não é um problema.

Após o ajuste na coluna “vlr\_preco\_unitario” será possível calcular o “vlr\_margem\_bruta”.

# Tratamento de Dados em SQL

c) Criação de tabela com investimento por Filial, AnoMes e Linha de Produto.

Para cada uma das filiais existe um total de 10 mil reais para ser investido em desenvolvimentos de novos produtos por mês. (R\$ 10.000 x 3 (filiais) x 3 (meses) = R\$ 90.000 a serem investidos)

Em cada filial, esse valor será distribuído por linha de produto de acordo com a quantidade de produtos (qtd\_produtos) vendida por linha de produtos no mês.

Crie uma tabela que mostre o valor de investimento que cada linha de produto

receberá por mês em cada uma das filiais. A tabela deverá conter as seguintes

colunas:

- Filial;
- Ano mês da venda;
- Nome da linha de produtos;
- Qtd de Produtos vendidos;
- % Qtd de Produtos vendidos;
- Valor Investido.

Obs: o somatório da coluna valor investido (sem filtro) deve ser 90 mil e o somatório da coluna % Qtd de Produtos vendidos deve ser 9 (3 filiais em 3 meses).

# Fase 2

## V i s u a l i z a ç ã o + N e g ó c i o s

Para essa etapa serão elencadas algumas análises solicitadas pelo cliente e você precisará desenvolver

visualizações em Power Bi que atendam os objetivos da melhor forma.

Fique à vontade para escolher a forma para mostrar os dados e sugerir qualquer outra análise que julgue importante.

Lembrando que as medidas finais devem ser feitas utilizando linguagem DAX e uma tabela de calendário deve ser criada para apoio nas visualizações históricas.

# Análise de dados

## Sugestão

Inúmeros indicadores e medidas podem ser utilizados para as análises. Para ganharmos tempo, abaixo seguem as colunas das tabelas que acreditamos serem suficientes para as análises de todos os itens: "tpo\_cliente", "nom\_sexo\_cliente", "nom\_linha\_produto", "qtd\_produtos", "dta\_venda" (AnoMes é suficiente), hra\_venda (olhar apenas para a hora, removendo os minutos, é recomendado), num\_avaliação, nom\_filial .

- a) O Supermercado precisa entender qual linha de produtos é a melhor linha para se investir em 2023. Qual é a linha de produto que melhor atende a esse requisito? Deveríamos investir em linhas de produtos diferentes para homens e mulheres? Demostre por meio de visualizações da sua preferência.

Sugestões de análise: crescimento na quantidade de vendas (mensal), sexo etc.

Obs: como a fato\_vendas possui apenas três meses de dados, caso exista crescimento em todos os três esses meses, esse crescimento pode ser considerado como significativo.



# Análise de dados

## b) Alocação de Caixas

Nas diferentes filiais, existem funcionários fixos, que trabalham das 10h00 às 21h00, e existe a possibilidade de alocarmos funcionários terceirizados em horários de pico. Quais seriam os horários mais críticos para alocarmos os funcionários terceirizados?

## c) Avaliação dos clientes

É necessário nos preocuparmos com os atendimentos realizados (satisfação de clientes) em alguma filial? Se sim, existe alguma característica que justifique o problema?

**OBRIGADO**