

## Teste Cientista de Dados





## Sumário

1. Perguntas Perfil .....	3
2. Proposta de Desenvolvimento .....	4



## 1. Perguntas Perfil

**1 – Informe teu Nome Completo:**

**2 – Você se considera Desenvolvedor Nível:**

☐ Júnior      ☐ Pleno      ☐ Sênior

**3 – Informe sua Pretensão Salarial PJ:**

Valor/Hora :

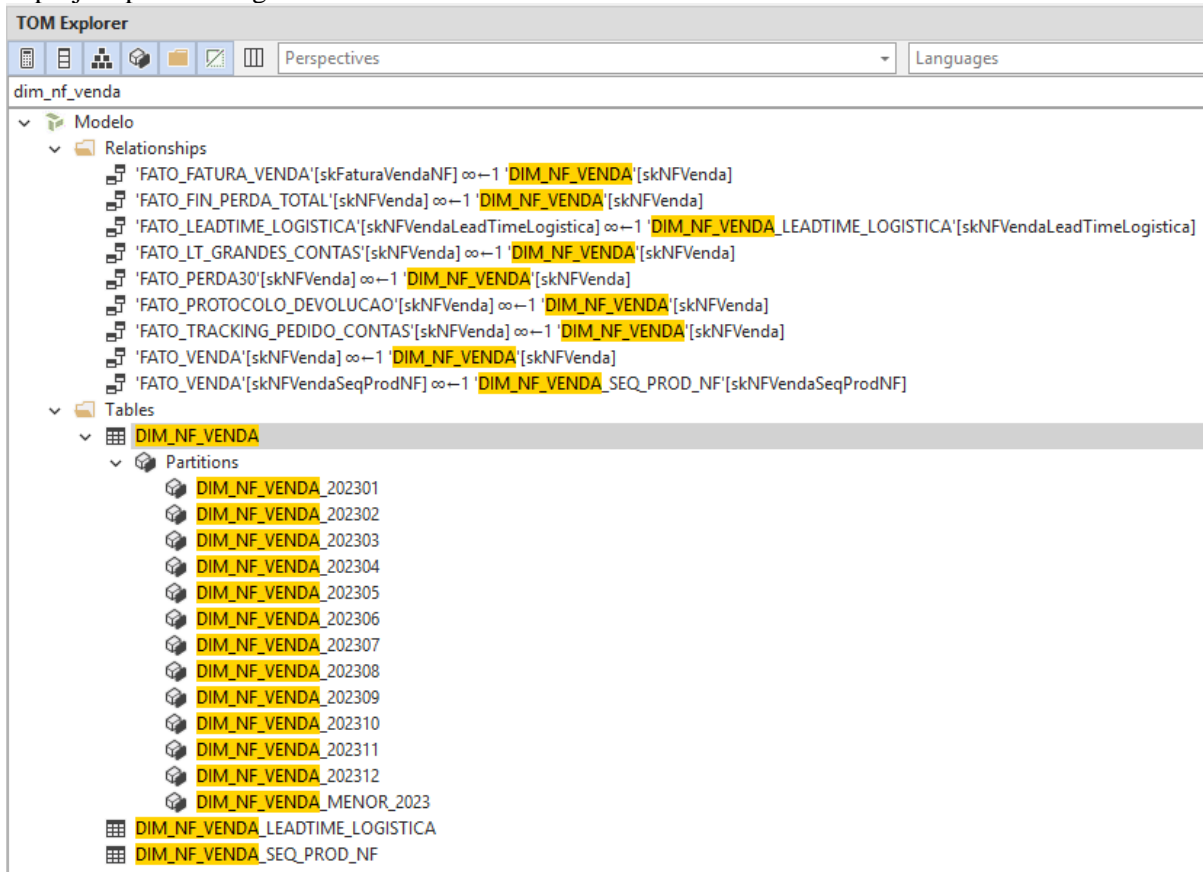
**4 – Possui Empresa Aberta ?**

☐ Sim    ☐ Não



## Decifre o problema abaixo:

O projeto possui a seguinte Dimensão:



Ao executar o comando abaixo (carga full)

```
{"refresh": {"type": "full", "objects": [{"database": "DW PROFARMA DISTRIBUICAO", "table": "DIM_NF_VENDA", "partition": "DIM_NF_VENDA_202307"}]}}
```

Recebemos de retorno a seguinte mensagem de erro:

Executing the query ...

A coluna 'skNFVenda' na Tabela 'DIM\_NF\_VENDA' contém um valor '56018333' duplicado e isso não é permitido para colunas de um lado da relação muitas-para-uma ou para colunas que estão sendo usadas como chave primária de uma tabela.

Run complete

- A primeira investigação fazemos um select na tabela e retornamos o seguinte resultado:



SQLQuery3.sql - sr..ICAO (GER)dw (88) - Mensagens - 16/07...5:17 - sv98wbd001

```
1 SELECT * FROM DIM_NF_VENDA
2 WHERE skNFVenda = 56018333
3
```

110 %

Results Messages

	skNFVenda	Num NF Venda	Data NF Venda	Num Fatura	Hr NF Venda	Hr NF Venda (Atual)	Num Pedido Venda	Num Pedido Venda (Atual)	Num Categoria Venda	Num Categoria Venda (Atual)	Cód Canal	Cód Canal (Atual)	inRegistroAtivo
1	56018333	552986	2023-07-11	1906008140	01:19:09.0000000	01:19:09.0000000	1776869076	1776869076	0	0	4	4	1

Você pode perceber que o registro não está duplicado na tabela de origem.

Responda as perguntas abaixo:

1. Você já passou por um problema parecido?
2. Qual linha de raciocínio você usaria para tentar resolver o problema? Descreva com detalhes.
3. Na sua opinião por qual motivo esse problema está acontecendo?