



TESTE TÉCNICO

DATA SCIENTIST

Você está recebendo o teste técnico para a vaga de **cientista de dados pleno** da Data Science Brigade.

OBJETIVO

Quais são os principais bens materiais comprados pelos órgãos públicos do Rio Grande do Sul? É possível identificar algum padrão nas compras? Seja por época do ano ou por tipo de órgão comprador? O que mais é possível descobrir sobre as licitações do estado?

Neste teste, gostaríamos que você tentasse responder a essas perguntas com dados de licitações de órgãos públicos do Estado do Rio Grande do Sul que foram cadastradas no sistema LicitaCon do TCE. Cada órgão cadastra os itens licitados em campos de texto livre no sistema (colunas DS_OBJETO e DS_ITEM das tabelas fornecidas) e não há uma padronização.

Esperamos que você entregue um Jupyter Notebook em Python demonstrando o seu processo de análise exploratória e com suas conclusões. O que você descobriu de interessante? O que você não conseguiu descobrir?

Vale colocar hipóteses e análises que você testou que não deram certo também! O importante aqui não é ter uma solução perfeita. O que vamos levar em conta são: a organização das análises e do código, exposição do processo e técnicas utilizados, documentação e embasamento das conclusões.

DADOS

Datasets a serem utilizados, com variação de ano:

- [2016](#)
- [2017](#)
- [2018](#)
- [2019](#)



Os arquivos **listados acima** foram baixados diretamente do [portal de dados abertos do TCE-RS](#). Lá você também encontra um [manual completo](#) com a descrição das tabelas e dicionário de dados das colunas exportadas pelo tribunal, caso seja necessário.

Segue abaixo uma lista dos campos mais relevantes para este teste:

Arquivo `licitacao.csv`

Contém todas as licitações registradas naquele ano. Cada linha representa uma licitação (ID_LICITACAO)

Campos relevantes:

- **DS_OBJETO**: A descrição do objeto da licitação, o que realmente aquele órgão está licitando.
- **TP_OBJETO**: Tipo de objeto de licitação. Recomendado utilizar apenas TP_OBJETO="Compras"

Arquivo `item.csv`

Contém todos os itens licitados. O campo (NR_LICITACAO) é equivalente ao campo ID_LICITACAO da tabela licitacao.csv.

Campos relevantes:

- **DS_ITEM**: Descreve o item licitado, ou seja, o item que o órgão público comprou.
- **VL_UNITARIO_HOMOLOGADO**: Descreve o preço pelo qual este item foi licitado.