

DESAFIO FINAL - BD/BI - 2024.2 FÁBRICA DE SOFTWARES

POWER BI

- 1 – Baixe e importe um dataset de alguns desses sites: **kaggle**, **dados.gov.br**, **UCI Machine Learning Repository**. Se preferir, use qualquer outro dataset (menos o que foi visto em sala e os que foram passados para o desafio diário) que tenha pelo menos 800 linhas;
- 2 – No Power BI, **atenção** à tipagem das colunas do seu dataset, e aos **erros** (valores nulos, duplicatas, etc.) que o dataset pode apresentar;
- 3 – Faça no **mínimo** 3 gráficos de tipos diferentes (Exemplo: Barras, Colunas, Pizza, Treemap, etc.) a partir do seu dataset. Atente-se no visual, cuidado na escolha de cores e efeitos;
- 4 – Com o Dashboard pronto, discorra sobre as conclusões da sua análise em um **arquivo .txt**. Por exemplo: “A cidade x teve uma média de vendas maior que a cidade y, isso pode ocorrer devido a...”;
- 5 – Sinta-se à vontade para analisar qualquer elemento do dataset e criar quantos gráficos e páginas quiser. Importante mostrar sua **capacidade de análise**, e **extrair insights** do dashboard;

OBS: Envie os 3 arquivos: power BI (.pbix), dataset (.xlsx ou .csv), e o texto (.txt).

Modelagem – SQL

- 0- Os modelos conceitual, lógico e físico do seguinte banco:
- 1 - Crie 5-6 tabelas e utilize 3 tipos de variáveis no total(INT, VARCHAR e uma a sua escolha), não esqueça da chave primária.
 - 2 - Adicione uma coluna em ao menos 1 tabela.
 - 3 - Insira dados em todas colunas das tabelas (3-4 linhas).
 - 4 - Mostre o conteúdo de cada tabela.
 - 5 - Salvar o arquivo em .sql

Obs: é necessário ter sentido lógico entre as tabelas, atributos e dados inseridos.

Obs2: as tabelas precisam ser sobre temas diferentes do desafio diário.

Obs3: descrever as regras de negócio que fizeram escolher as cardinalidades do Diagrama.

Obs4: modelo conceitual e lógico, subir em prints colocados em um único PDF.

PYTHON

1 - Banco de Dados

Para execução do desafio, você deve criar um notebook (.ipynb), importar as bibliotecas Pandas e Matplotlib, e o banco de dados do link abaixo:

Link:

https://drive.google.com/drive/folders/1LTYnIkr9f12frCvo4JgAumbPbinB8Ep?usp=share_link

2 - Tasks Finais

A ideia é que você faça uma Análise Completa dos Dados, desde visualizar suas características, a qualidade de seus dados e extrair valor dele através da geração de gráficos. Com isso, as questões propostas são:

1. Características do Banco de Dados

- Quantas linhas e colunas temos no banco de dados?
- Existem dados nulos no banco? Se sim, quais colunas apresentam e quantos são os dados nulos nelas?
- Quais tipos de dados estão presentes no banco de dados?
(Exemplo: object, int, float)

2. Agrupamento de Dados e Plotagem de Gráficos

Utilizando as técnicas de agrupamento de dados vistas em sala, faça as seguintes questões:

- Plote um gráfico de barras (bar) que exiba a quantidade de funcionários em cada setor (coluna Department);
- Plote um gráfico de linha (line) que exiba a quantidade de funcionários para cada nível de escolaridade (coluna Education);
- Plote um gráfico de torta/pizza (pie) que exiba a quantidade de funcionários que tem ou não algum atrito na empresa (coluna Attrition).

ENTREGA

Compartilhe o link do repositório com o nome DESAFIO-FS-BDBI neste formulário:

<https://forms.gle/xN3UBsyWWpsDFBLZ7>

A pontuação do imerscionista será computada de acordo com a conformidade das respostas e a postagem da atividade na data prevista no Git e na planilha.

Horário limite para entrega do material: até as 14:00h do dia 24/08.