

## DESAFIO FINAL - BD/BI - 2023.2

### FÁBRICA DE SOFTWARES

#### EXCEL

- 1 - Junte na coluna de Verdadeiro Número do Pagamento, o Número do Documento + Número do Pagamento, mas com um "\_" como delimitador entre eles.
- 2 - Junte na coluna Detalhes Bancários, o código da Coluna Número de Registro da Empresa, o código da Coluna Número da Conta, e o código da Coluna Verificação, mas com um "-" como delimitador entre cada um deles
- 3 - Na coluna Mês da Fatura, extraia o mês da coluna Fatura, ou seja, as três primeiras letras.
- 4 - Na coluna Número da Ordem de Compra, extraia o código da Coluna Ordem de compra do Cliente, ou seja, os últimos 6 dígitos.
- 5 - Extraia somente os nomes da Localização, ou seja, "Sydney" ou "Melbourne"
- 6 - Nota-se que na Coluna Valor Pago, os números presentes nela, são na verdade Strings/Textos e não podem ser transformados para valores numéricos realmente por causa de um tipo de espaçamento que não pode ser removido através da função TIRAR E ARRUMAR. Sendo assim, extraia os valores monetários da coluna Valor Pago, para a Coluna Valor e utilize a função VALOR para transformar em números de fato.
- 7 - Para não perdermos o costume, utilize o PROCV na planilha 2, mas utilizando os valores de referência presentes da Ref Planilha 2.

#### LINK:

<https://docs.google.com/spreadsheets/d/1LXF5dZWs41PHITjMQXh58Li91Erz3vcL/edit?usp=sharing&ouid=107595620328338051756&rtpof=true&sd=true>

## POWER BI

- 1 – Importe o Dataset tratado em Excel para o Power BI.
- 2 – No Power BI, verifique a tipagem das colunas do seu Dataset
- 3 – Faça no mínimo 3 gráficos de tipos diferentes (Exemplo: Barras, Colunas e Pizza) baseados em insights tirados do seu Dataset.
  - Atente-se para o VISUAL, cuidado na escolha de cores e efeitos.
- 4 – Com o Dashboard pronto, discorra sobre as conclusões da sua análise.
  - 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 da tabela e criar quantos gráficos quiser.

## SQL

- 1 - Crie 3-4 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.

## 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/1LTYnIkr9f12f-rCvo4JgAumbPbinB8Ep?usp=share\\_link](https://drive.google.com/drive/folders/1LTYnIkr9f12f-rCvo4JgAumbPbinB8Ep?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

- a. Quantas linhas e colunas temos no banco de dados?
- b. Existem dados nulos no banco? Se sim, quais colunas apresentam e quantos são os dados nulos nelas?
- c. 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:

- a. Plote um gráfico de barras (bar) que exiba a quantidade de funcionários em cada setor (coluna Department);
- b. Plote um gráfico de linha (line) que exiba a quantidade de funcionários para cada nível de escolaridade (coluna Education);
- c. 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 da pasta ou repositório com o nome DESAFIO-FS-BDBI na planilha compartilhada no grupo do whatsapp.

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 08:00h do dia 03 de setembro.**