

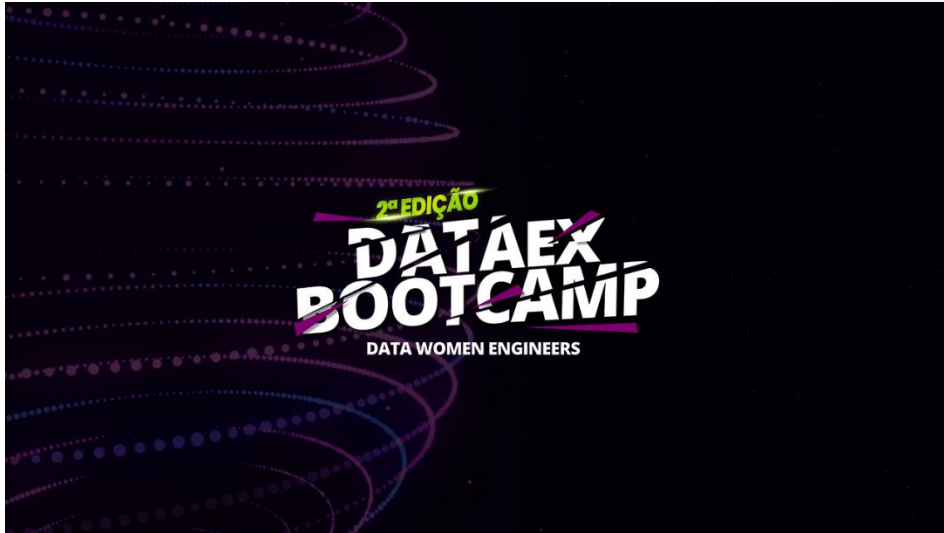
Challenge DWE 2023

Objetivo:

Avaliar conhecimentos adquiridos no DataEX BootCamp – Data Women Engineers

Público-alvo:

Alunas BootCamp DWE 2023, 2ª Edição – Data Women Engineers.



O Desafio irá testar suas habilidades nas seguintes competências:

- ✓ Construção de 1 Pipeline de dados no Notebook com Azure Data Studio (Max 700 Pts).
- ✓ Habilidades com a linguagem SQL e T-SQL Comandos DML (Max 300 Pts).

Termos e condições:

- Ser aluna inscrita e ativa no DataEX Bootcamp Data Women Engineers 2023 2ª Ed.;
- Não ser colaborador ou prestador de serviço da DataEX.

Para tornar os prêmios deste desafio elegíveis, as participantes devem:

- Entregar o Desafio até a data final (19/02/2023 as 23:59hs);
- O arquivo deverá ser enviado única e exclusivamente para o e-mail treinamentos@dataex.com.br.

Ao participar do **Challenge Data Engineer** do **DataEX Bootcamp DWE 2023 2º Ed**, realizado pela DataEX, entre os dias 30/01/2023 à 10/02/2023, eu AUTORIZO o uso de minha imagem e vídeo enviado neste formulário em todo e qualquer material impresso e/ou virtual, como sites, redes sociais, reportagens, vídeos institucionais, propaganda e outro, desenvolvidos pela DataEX, sejam essas destinadas à divulgação ao público em geral. A presente autorização é concedida a título gratuito, abrangendo o uso da imagem acima mencionada em todo território nacional e no exterior, sem data de expiração. Por esta ser a expressão da minha vontade declaro que autorizo o uso acima descrito sem que nada haja a ser reclamado a título de direitos conexos à minha imagem ou a qualquer outro, e assino a presente autorização.

Prazo Final da entrega do Desafio:

19/02/2023 as 23:59 (Domingo)

Desafios enviados após a data descrita serão desclassificados automaticamente.

Critérios de desempate

Será pela ordem de entrega, priorizando ordem de recebimento.

Forma de envio:

Entrega **Pacote único zipado e nomeado com seu usuário de banco de dados no Bootcamp:**

Ex: **aluna01.zip**

Organize todos os arquivos criados da melhor forma possível e cronológica para avaliação.

Enviar para e-mail treinamentos@dataex.com.br

Assunto E-mail "Entrega Desafio DWE 2023 AlunaXX"

Divulgação dos Resultados

01/03/2023 (Live no Sympia 15h00)

Premiação Challenge DWE 2023



CHALLENGE DWE 2022

OBS: Parte 1 e 2, possuem dependência entre elas em ordem cronológica.

1ª Parte – Desenvolvimento Pipeline/Notebook Azure Data Studio

Requisitos de Negócio

Vale até 700 pontos

O que vamos analisar?

A **Execução de despesas do Governo**, fonte de dados do Portal da transparência, referente aos dados de Jan/22 a Mar/22, 1º Trimestre

Como desejamos analisar os dados?

- Órgão Superior
- Órgão Subordinado
- Unidade Gestora
- Gestão
- Grupo de Despesa
- Tempo

Quais métricas deseja acompanhar?

- Valor Empenhado (R\$)
- Valor Liquidado (R\$)
- Valor Pago (R\$)

Requisitos de Técnicos

Desenvolva um pipeline de integração de dados usando o Notebook no Azure Data Studio e realize a modelagem Star Schema, desnormalizando os dados gerando um DW de acordo com seu entendimento do escopo do desafio e realize a carga de dados disponibilizados em seu DW_XX.

- Crie seu Notebook com o nome DESAFIO_ALUNAXX.ipynb
- As dimensões do projeto devem controlar alterações ao longo do tempo.
- Controle logs de execuções do pipeline.
- Gere scripts com possibilidade de reprocessamentos.
- Gere um diagrama dimensional com a modelagem aplicada.
- Crie um dicionário de dados dos objetos contidos em seu DW.
- Use nomes intuitivos em seus objetos.
- Use Linguagem SQL e T-SQL

Origem de dados:

Nome do Arquivo	Assunto	Fonte
202201_Despesas.csv	Despesas Jan/2022	Portal da Transparência
202202_Despesas.csv	Despesas Feb/2022	Portal da Transparência
202203_Despesas.csv	Despesas Mar/2022	Portal da Transparência

Fonte:

<https://portaldatransparencia.gov.br/download-de-dados/despesas-execucao>

Dicionário de dados:

<https://portaldatransparencia.gov.br/pagina-interna/603453-dicionario-de-dados-execucao-da-despesa>

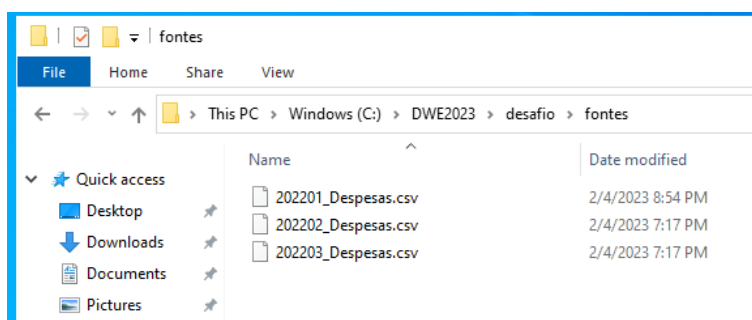
Caminho que deve ser “**fixado**” no projeto para leitura dos arquivos no Servidor.

Localização dos dados:

Servidor SQL com IP: 20.83.31.31

C:/DWE2022/desafio/fontes/

Use seus respectivos usuários e senha usados durante o curso.



Arquivos: Acesso no caminho informado acima através do servidor SQL localizado no IP 20.83.31.31

Destino dos dados:

Seu DW_XX

IP: 20.83.31.31

2ª Parte – Conhecimentos de SQL e T-SQL(DML)

DESENVOLVA SCRIPTS PARA CADA PERGUNTA DE NEGÓCIO ABAIXO:

EXERCÍCIO 1

Vale até 30 pontos

Qual o total em Jan/2022 Valor Pago (R\$) para cada Órgão Superior?

ENTREGÁVEL: SCRIPT NO NOTEBOOK DO DESAFIO

EXERCÍCIO 2

Vale até 30 pontos

Quais os top 10 Órgãos Superiores referente a Jan/2022 em Valor Pago (R\$)

ENTREGÁVEL: SCRIPT NO NOTEBOOK DO DESAFIO

EXERCÍCIO 3

Vale até 30 pontos

Qual o total no Trimestre de Valor Empenhado (R\$) para cada Gestão?

ENTREGÁVEL: SCRIPT NO NOTEBOOK DO DESAFIO

EXERCÍCIO 4

Vale até 30 pontos

Quais as Top 10 Gestão no Trimestre referente a valor empenhado?

ENTREGÁVEL: SCRIPT NO NOTEBOOK DO DESAFIO

EXERCÍCIO 5

Vale até 30 pontos

Qual o total mês a mês de Valor Liquidado (R\$) para cada Grupo de Despesa?

ENTREGÁVEL: SCRIPT NO NOTEBOOK DO DESAFIO

EXERCÍCIO 6

Vale até 30 pontos

Quais os TOP 10 Grupo de Despesas referente a valor Liquidado (R\$) em Fev/2022?

ENTREGÁVEL: SCRIPT NO NOTEBOOK DO DESAFIO

EXERCÍCIO 7

Vale até 30 pontos

Qual a variação %(MoM) de Valor Empenhado (R\$) de Mar/2022 para Fev/2022 por Unidade Gestora

ENTREGÁVEL: SCRIPT NO NOTEBOOK DO DESAFIO

EXERCÍCIO 8

Vale até 30 pontos

Retorne o maior Valor Empenhado (R\$) de Grupo de despesas em Fev/2022 para cada Órgão Superior.

ENTREGÁVEL: SCRIPT NO NOTEBOOK DO DESAFIO

EXERCÍCIO 9

Vale até 30 pontos

Retorne o % de Valor Pago (R\$) sobre o total, por Gestão Fev/2022.

ENTREGÁVEL: SCRIPT NO NOTEBOOK DO DESAFIO

EXERCÍCIO 10

Vale até 30 pontos

Retorne o valor média de Valor Pago (R\$) por despesa agrupado por Gestão.

ENTREGÁVEL: SCRIPT NO NOTEBOOK DO DESAFIO.

ORIENTAÇÕES FINAIS

- Entregar notebook único com parte 1 e 2
- Organize e evidencie a construção do seu pipeline da melhor forma possível para entrega e avaliação.
- Documente em seu Notebook todos os passos realizados na execução do pipeline.
- Onde se lê XX substituir por seu número de aluna no Bootcamp.

ENTREGUE O MÁXIMO POSSÍVEL DE EXERCÍCIOS.

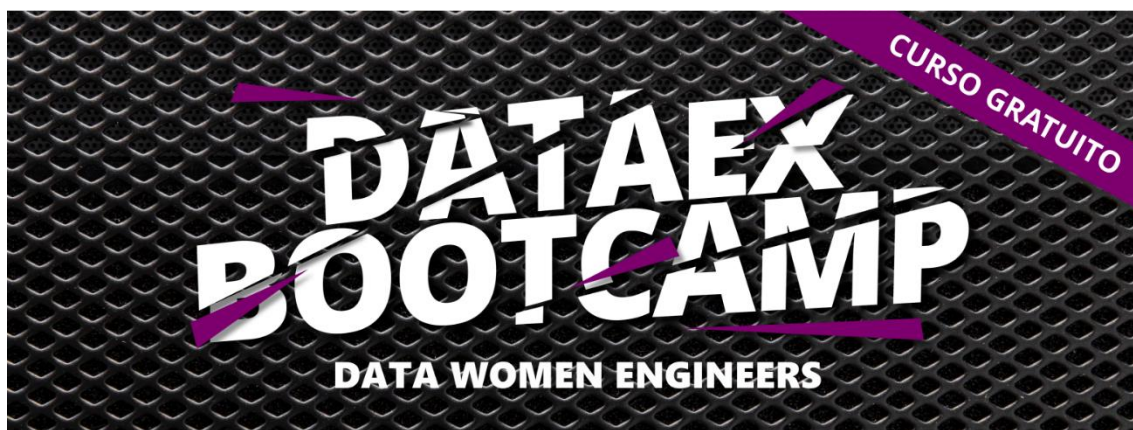
SERÃO ACEITAS ENTREGAS PARCIAIS.

Boa sorte!

Deseja ir além do solicitado? Vá além!

Agora é com você! Bora codar Bootcamper

Baixe os arquivos no Sympyla e Hands-on!



**TORNE-SE UMA
ENGENHEIRA DE DADOS**

Realização
 **dataEX**

Apoio
 **Microsoft**