**Teste Engenheira de Dados - Esfera**

Passo a passo completo.

**1) Criação do banco de dados e das tabelas**

Realizei a criação do banco de dados e das tabelas Despesa e Receita através do código abaixo:

create database Esfera

go

use Esfera

create table Despesa (

Fonte\_de\_Recursos varchar (300),

Tipo\_Despesa varchar (300),

Liquidado varchar (50));

create table Receita (

Fonte\_de\_Recursos varchar (300),

Tipo\_Receita varchar (300),

Arrecadado varchar (50))

**2) Tratamento dos dados**

Realizei os tratamentos dos dados da seguinte forma:

* Remoção das aspas;
* Remoção de espaços duplicados;
* Remoção da coluna "TOTAL";
* Substituído separadores de coluna de vírgula para ponto e vírgula;
* Corrigido layout do arquivo, pois a separação de colunas estava por vírgula, porém existia textos utilizando vírgula, o que causava quebra do arquivo;
* Importei o arquivo no SQL com a coluna valor sendo uma string, precisei tratar a string e convertê-la para float, comado utilizado:
  + ALTER TABLE DESPESA ADD VLR\_LIQUIDADO FLOAT

GO

UPDATE DESPESA SET VLR\_LIQUIDADO = CAST(REPLACE(REPLACE(REPLACE(liquidado, '.', ''), ' ', ''), ',', '.') AS FLOAT)

Obs: -- Inclui a coluna VLR\_LIQUIDADO e passei as informações da coluna Liquidado para VLR\_LIQUIDADO substituindo '.' por '', " " por '', e ',' por '.'

Realizei a mesma operação para RECEITA.

Dropei a coluna arrecadado e liquidado e depois alterei o nome das colunas “**DESPESA.VLR\_LIQUIDADO**”, para LIQUIDADO e “**RECEITA.VLR\_ARRECADADO**” para ARRECADADO.

Segue o comado abaixo:

alter table despesa drop column liquidado

sp\_rename 'DESPESA.VLR\_LIQUIDADO', 'Liquidado', 'COLUMN'

* Extrai os selects abaixo, removi os "tabs" e substitui as vírgulas por ponto para conseguir importar o arquivo com o campo de valor com tipo de dado float:

SELECT FONTE\_DE\_RECURSOS, ';', TIPO\_DESPESA, ';', LIQUIDADO FROM DESPESA

SELECT FONTE\_DE\_RECURSOS, ';', TIPO\_receita, ';', arrecadado FROM RECEITA

* Alterei o tipo de dado da coluna liquidado e arrecadado de **varchar** para **float,** comado abaixo:

DROP TABLE DESPESA

DROP TABLE RECEITA

create table Despesa (

Fonte\_de\_Recursos varchar (300),

Tipo\_Despesa varchar (300),

Liquidado float);

create table Receita (

Fonte\_de\_Recursos varchar (300),

Tipo\_Receita varchar (300),

Arrecadado float)

* Importei novamente o arquivo no SQL.

**Script de Importação:**

BULK INSERT DESPESA FROM 'C:\Users\Usuario\PythonSQL\EsferaTeste1\gdvDespesasExcel.csv'

WITH

( FIRSTROW=2 -- Define a partir de qual linha que o sql irá importar

, FIELDTERMINATOR = ';' -- separador das colunas

, ROWTERMINATOR = '\n') -- separador das linhas, /n representa a quebra de linha (a linha acaba quando pula a linha)

**3) Criação da Procedure**

Criei uma PROC para rodar no Python, segue o comado abaixo:

CREATE OR ALTER PROC SP\_ESFERATESTE1

AS

BULK INSERT DESPESA FROM 'C:/Users/Usuario/PythonSQL/EsferaTeste1/gdvDespesasExcel.csv'

WITH

( FIRSTROW=2

, FIELDTERMINATOR = ';'

, ROWTERMINATOR = '\n')

BULK INSERT Receita FROM 'C:/Users/Usuario/PythonSQL/EsferaTeste1/gdvReceitasExcel.csv'

WITH

( FIRSTROW=2

, FIELDTERMINATOR = ';'

, ROWTERMINATOR = '\n')

**4) Entregáveis: Tabela Final**

Segue os comados abaixo para informações solicitadas:

-- Soma da Receita e Despesa agrupados pela fonte de recursos.

SELECT R.FONTE\_DE\_RECURSOS, SUM(ARRECADADO) AS RECEITA, SUM(LIQUIDADO) AS DESPESA FROM RECEITA AS R

INNER JOIN DESPESA AS D ON R.FONTE\_DE\_RECURSOS = D.FONTE\_DE\_RECURSOS

GROUP BY R.FONTE\_DE\_RECURSOS

-- Soma da despesa agrupada pelo tipo da despesa da maior para a menor

SELECT TIPO\_DESPESA, SUM(LIQUIDADO) AS DESPESA FROM DESPESA

GROUP BY TIPO\_DESPESA

ORDER BY DESPESA DESC

-- 10 maiores fontes de receita

SELECT TOP 10 FONTE\_DE\_RECURSOS, SUM(ARRECADADO) AS RECEITA FROM RECEITA

GROUP BY FONTE\_DE\_RECURSOS

ORDER BY RECEITA DESC

-- 10 maiores fontes de despesa

SELECT TOP 10 FONTE\_DE\_RECURSOS, SUM(LIQUIDADO) AS DESPESA FROM DESPESA

GROUP BY FONTE\_DE\_RECURSOS

ORDER BY DESPESA DESC

**5) Script Python**

Obs: Antes de rodar o comado, necessário excluir as tabelas e criar novamente através do comado abaixo:

DROP TABLE DESPESA

DROP TABLE RECEITA

create table Despesa (

Fonte\_de\_Recursos varchar (300),

Tipo\_Despesa varchar (300),

Liquidado float);

create table Receita (

Fonte\_de\_Recursos varchar (300),

Tipo\_Receita varchar (300),

Arrecadado float)

O script abaixo conecta o Python com SQL através da biblioteca **pyodbc**

import pyodbc  
  
dados\_conexao = (  
 "Driver={SQL Server};"  
 "Server=DESKTOP-RCO0BTU;"  
 "Database=Esfera;"  
)  
  
conexao = pyodbc.connect(dados\_conexao)  
print("Conexão Bem Sucedida")  
  
cursor = conexao.cursor()  
  
comado = """  
exec SP\_ESFERATESTE1  
"""  
  
cursor.execute(comado)  
cursor.commit()

**END.**