

# MVP3 - Engenharia de Dados

Aluno: José Roberto Assis Silva

RA: 4052023000345

---

## 1. Objetivo

---

Foi criada uma campanha para premiar os vendedores de equipamentos na empresa em que trabalho. Vou tratar esta campanha neste MVP coletando dados do nosso ERP que utiliza um banco de dados relacional (SGDB ORACLE). O objetivo desta campanha é rankear os vendedores por total de equipamentos vendidos mostrando também a quantidade vendida por grupo de equipamentos. A partir destes dados será feita uma premiação para o primeiro, segundo e terceiro colocados. Esta campanha utilizará apenas dados do ano de 2023, mas vou gerar os CSVs com todas as informações de vendas que estão em nosso database para criar meu dataset. As visualizações destas informações deveram estar em um dashboard mostrando um grid rankeando os vendedores pela quantidade total de equipamentos vendidos e deve conter a quantidade vendida de cada grupo de equipamentos com os totais gerais do grupo e do vendedor. Também terá um gráfico de barras rankeando os vendedores com suas respectivas quantidades totais vendidas. Neste dashboard devemos ter filtro por loja e data. A partir do database (Oracle) vou criar algumas queries para gerar alguns arquivos CSVs e vou utilizá-los para carga na nuvem Azure para fazer ETL e posteriormente criar o dashboard. Vou mudar algumas descrições nos CSVs para manter a confidencialidade dos dados. Os CSVs foram gerados a partir de uma conexão feita com o SGDB Oracle utilizando o Dbeaver e queries SQL.

Objetivos:

1. Criar um grid com ranking dos vendedores por quantidade total vendida.
  2. Nas colunas deste grid deveram ter a quantidade por grupo de equipamento vendido, com o total geral de todos dos grupos no final das colunas.
  3. O grid terá o total por vendedor nas linhas.
  4. Fazer um gráfico de barras com ranking dos vendedores por total de equipamentos vendidos.
- 

## 2. Busca pelos Dados

---

Vou utilizar uma conexão com o SGDB Oracle utilizando o Dbeaver para gerar os CSVs que farão parte do meu dataset. Como conhecedor do database que utilizamos, criei algumas queries, com visões temporárias, joins e filtros para gerar um CSV com itens vendidos apenas de EQUIPAMENTOS (VENDAS), porque também trabalhamos com outras linhas de produtos. Vou gerar também CSVs de 2 tabelas dimensões, as quais utilizarei quando criar um dataflow no Azure Data Factory, fazendo alguns joins para gerar somente uma tabela flat VENDAS não **normalizada**. Farei alterações em algumas descrições preservando a confidencialidade dos dados, também inibirei alguns dados nas imagens. Serão gerados os seguintes CSVs:

1. VENDAS.csv (Contém os atributos das vendas)
2. LOJAS.csv (Contém os atributos das lojas)
3. GRUPOSEQP (Contém os atributos dos grupos dos equipamentos)

### 2.1. Executando query e gerando o CSV VENDAS

DBeaver 23.2.0 - <NBS\_TESTE> Script-2

File Edit Navigate Search SQL Editor Database Window Help

Database Navigator X

ACesso\_LOG <NBS\_TESTE> ... <NBS\_TESTE> ...

Enter a part of object name here

```

WITH OPERACAO AS (
    SELECT O.COD_EMPRESA
        , O.COD_OPERACAO
        , O.OPERACAO
        , O.COD_GRUPO_PC
        , O.COD_NATUREZA_RECEITA_DESPESA
        , NRD.DESCRICAO_DESCR_NRD
        , NRD.GRUPO
        , O.COD_NATUREZA
        , NAT.DESCRICAO AS DESCR_NATUREZA
        , NAT.DESCRICAO_REFERENCIA AS DESCR_NAT_REF
        , O.COD_CONTA_CORRENTE
        , CC.DESCRICAO AS DESCR_CONTA_CORRENTE
        , O.COD_TIPO_PAGAMENTO
        , TP.DESCRICAO AS DESCR_TIPO_PAGAMENTO
    FROM NBS.OPERACOES O
    LEFT JOIN NBS.NATUREZA_RECEITA_DESPESA NRD ON O.COD_GRUPO_PC = NRD.COD_GRUPO_PC
                                                AND O.COD_NATUREZA_RECEITA_DESPESA = NRD.COD_NATUREZA_RECEITA_DESPESA
    LEFT JOIN NBS.NATUREZA NAT ON O.GRUPO = NAT.GRUPO
                                                AND O.COD_NATUREZA = NAT.COD_NATUREZA
    LEFT JOIN NBS.CONTA_CORRENTE CC ON O.COD_EMPRESA = CC.COD_EMPRESA
                                                AND O.COD_CONTA_CORRENTE = CC.COD_CONTA_CORRENTE
    LEFT JOIN NBS.TIPO_PAGAMENTO TP ON O.COD_EMPRESA = TP.COD_EMPRESA
                                                AND O.COD_TIPO_PAGAMENTO = TP.COD_TIPO_PAGAMENTO
    WHERE o.cod_operacao in (-1, --Peças Balcão
                            --2, --Serviços
                            --3, --Serviços Garantias
                            4, --Máquinas
                            9, --Máquinas
                            --7, --Peças Devolução
                            --13,
                            --19, --Peças Atacado
                            22, --Diversos
                            --26
                            66, --Devolução de Máquinas fins diversos
                            28 --Máquinas Devolução
                            )
)
SELECT TRUNC(CAB.EMISSAO) AS DATA_FATURAMENTO
    , CAB.CONTROLE
    , CAB.SERIE
    , ED.DESCRICAO AS DEPARTAMENTO
    , CAB.VENDEDOR
    , CAB.COD_CLIENTE_NOTA
    , CAB.COD_CLIENTE
    , CAB.COD_OPERACAO
    , CAB.TOTAL_NOTA AS VALOR_VENDIDO
    , CAB.VALOR_ICMS*(-1) as VALOR_ICMS
    , CAB.ICMS_SUSTITUIDO
    , CAB.ANO_MODELO
    , CAB.ANO_FABRICACAO
    , CAB.VALOR_PIS*(-1) as VALOR_PIS
    , CAB.VALOR_COFINS*(-1) as VALOR_COFINS
    , CAB.STATUS
    , VS.DESCRICAO AS DESCRSTATUS
    , CAB.END_CID_DESCRICAO
    , CAB.END_CID_UF
    , CAB.END_CID_DESCRICAO||', '|||CAB.END_CID_UF AS CIDADEUF
--Dados da Fábrica

```

Results 1

WITH OPERACAO AS (SELECT O.COD\_EMPRESA, O.COD\_OPERACAO, O.OPERACAO, O.COD\_GRUPO\_PC, O.COD\_NATUREZA\_RECEITA\_DESPESA, NRD.DESCRICAO\_DESCR\_NRD, NRD.GRUPO, O.COD\_NATUREZA, NAT.DESCRICAO AS DESCR\_NATUREZA, NAT.DESCRICAO\_REFERENCIA AS DESCR\_NAT\_REF, O.COD\_CONTA\_CORRENTE, CC.DESCRICAO AS DESCR\_CONTA\_CORRENTE, O.COD\_TIPO\_PAGAMENTO, TP.DESCRICAO AS DESCR\_TIPO\_PAGAMENTO)

	DATA_FATURAMENTO	123 CONTROLE	ABC SERIE	ABC DEPARTAMENTO	ABC VENDEDOR	123 COD_CLIENTE_NOTA	123 COD_CLIENTE	123 COD_OPERACAO	123 VALOR_ICMS
1	2013-08-28 00:00:00.000	72,254	1	Administracao	MATEVEIRA	[NULL]			4
2	2013-08-28 00:00:00.000	72,250	1	Administracao	MATEVEIRA	[NULL]			4
3	2013-08-28 00:00:00.000	72,251	1	Administracao	MATEVEIRA	[NULL]			4
4	2017-12-29 00:00:00.000	59,336	1	Administracao	MMACHADO	[NULL]			4
5	2013-10-04 00:00:00.000	74,418	1	Administracao	MATEVEIRA	[NULL]			4
6	2013-10-03 00:00:00.000	74,273	1	Administracao	MATEVEIRA	[NULL]			4
7	2013-08-28 00:00:00.000	72,249	1	Administracao	MATEVEIRA	[NULL]			4
8	2013-08-28 00:00:00.000	72,244	1	Administracao	MATEVEIRA	[NULL]			4
9	2013-08-21 00:00:00.000	71,887	1	Administracao	EMOREIRA	[NULL]			4
10	2013-08-28 00:00:00.000	72,245	1	Administracao	MATEVEIRA	[NULL]			4
11	2014-03-22 00:00:00.000	26,317	1	Administracao	SMARTINS	[NULL]			4
12	2013-11-12 00:00:00.000	23,100	1	Administracao	SIOSEMAR	[NULL]			4
13	2013-11-12 00:00:00.000	23,101	1	Administracao	SIOSEMAR	[NULL]			4
14	2013-08-14 00:00:00.000	71,583	1	Administracao	MATEVEIRA	[NULL]			4

200 row(s) fetched - 6.487s (24ms fetch), on 2023-09-21 at 11:18:20

Pesquisar

Refresh Save Cancel Export data 200 200+ 200 row(s) fetched - 6.487s (24ms fetch), on 2023-09-21 at 11:18:20 POR PTB2 21/09/2023 11:18 26

DBeaver 23.2.0 - <NBS\_TESTE> Script-2

File Edit Navigate Search SQL Editor Database Window Help

Database Navigator X

ACESSO\_LOG <NBS\_TESTE> ... <NBS\_TESTE> ...

Enter a part of object name here

NBS\_TESTE - NBS  
NIPIVIEW - 192.168.20.235:3306  
TRON-PRODUÇÃO - 192.168.20.32:3050  
TRONCG.IDB - localhost:3050

SELECT TRUNC(CAB.EMISSAO) AS DATA\_FATURAMENTO  
, CAB.CONTROLE  
, CAB.SERIE  
, ED.DESCRICAO AS DEPARTAMENTO  
, CAB.VENDEDOR  
, CAB.COD\_CLIENTE\_NOTA  
, CAB.COD\_CLIENTE  
, CAB.COD\_OPERACAO  
, CAB.TOTAL\_NOTA AS VALOR\_VENDIDO  
, CAB.VALOR\_ICMS\*( $-1$ ) AS VALOR\_ICMS  
, CAB.INDSUSTITUIDO  
, CAB.ANO\_MODELO  
, CAB.ANO\_FABRICACAO  
, CAB.VALOR\_PIS\*( $-1$ ) AS VALOR\_PIS  
, CAB.VALOR\_COFINS\*( $-1$ ) AS VALOR\_COFINS  
, CAB.STATUS  
, VS.DESCRICAO AS DESCRSTATUS  
, CAB.END\_CID\_DESCRICAO  
, CAB.END\_CID\_UF  
, CAB.END\_CID\_DESCRICAO||', ''||CAB.END\_CID\_UF AS C  
--Dados da Fabrica

Results 1 < Enter a SQL expression

	DATA_FATURAMENTO	123 CONTROLE	REC SERIE
1	2013-08-28 00:00:00.000	72,254	1
2	2013-08-28 00:00:00.000	72,250	1
3	2013-08-28 00:00:00.000	72,251	1
4	2017-12-29 00:00:00.000	59,336	1
5	2013-10-04 00:00:00.000	74,418	1
6	2013-10-03 00:00:00.000	74,273	1
7	2013-08-28 00:00:00.000	72,249	1
8	2013-08-28 00:00:00.000	72,244	1
9	2013-08-21 00:00:00.000	71,887	1
10	2013-08-28 00:00:00.000	72,245	1
11	2014-03-22 00:00:00.000	26,317	1
12	2013-11-12 00:00:00.000	23,100	1
13	2013-11-12 00:00:00.000	23,101	1
14	2013-08-14 00:00:00.000	71,583	1

Data Transfer  
Confirm  
Check results

Source Container: NBS\_TESTE  
Source: WITH OPERACAO AS ( SELECT O.COD\_EMPRESA , O.COD...  
Target Container: VENDAS.csv  
Target: C:\Users\jose.silva\Documents\VENDAS.csv

Objects

Source settings  
Table settings:  
Open new connection(s): No  
Extract type: SEGMENTS  
Select row count: No  
Selected rows only: No  
Selected columns only: No

Target settings  
Files settings:  
Write to the single file: No  
Directory: C:\Users\jose.silva\Documents  
File name pattern: VENDAS  
On object data file name conflict: Autofix name  
On blob value file name conflict: Autofix name  
Encoding: UTF-8  
Timestamp pattern: yyyy/MMddHHmm  
Insert BLOBs: No  
Compress: No  
Binaries: SKIP  
Frontends: RINARY

Export target  
Extraction settings  
Format settings  
Output  
Confirm

Results 1 < Enter a SQL expression

	123 COD_NOTA_MODELO	REC DESCRICAO_NOTA_MODELO
1	210	
2	213	
3	217	
4	208	
5	209	
6	211	
7	215	
8	216	
9	214	
10	218	

Save task

< Back Next > Proceed Cancel

200 row(s) fetched - 6.487s (24ms fetch), on 2023-09-21 at 11:18:20

Pesquisar

Refresh Save Cancel Export data 200 200+ < Back Next > Proceed Cancel

11:20 PTB2 21/09/2023

## 2.2. Executando query e gerando o CSV GRUPOSEQP

DBeaver 23.2.0 - <NBS\_TESTE> Script-2

File Edit Navigate Search SQL Editor Database Window Help

Database Navigator X

ACESSO\_LOG <NBS\_TESTE> ... <NBS\_TESTE> ...

Enter a part of object name here

NBS\_TESTE - NBS  
NIPIVIEW - 192.168.20.235:3306  
TRON-PRODUÇÃO - 192.168.20.32:3050  
TRONCG.IDB - localhost:3050

```

SELECT COD_NOTA_MODELO , DESCRICAO_NOTA_MODELO FROM NBS.NOTAS_MODELOS
SELECT * FROM EMPRESAS ORDER BY COD_EMPRESA

```

Results 1 < Enter a SQL expression to filter results (use Ctrl+Space)

	123 COD_NOTA_MODELO	REC DESCRICAO_NOTA_MODELO
1	210	
2	213	
3	217	
4	208	
5	209	
6	211	
7	215	
8	216	
9	214	
10	218	

Save task

< Back Next > Proceed Cancel

10 row(s) fetched - 3ms, on 2023-09-21 at 11:12:34

Pesquisar

Refresh Save Cancel Export data 200 10 < Back Next > Proceed Cancel

30°C Ensolarado & POR PTB2 11:12 21/09/2023

## 2.3. Executando query e gerando o CSV LOJAS

The screenshot shows the DBBeaver interface with the following details:

- Top Bar:** DBBeaver 23.2.0 - <NBS\_TESTE> Script-2. Includes File, Edit, Navigate, Search, SQL Editor, Database, Window, Help menus.
- Toolbar:** Auto, Rollback, Undo, Redo, Save, Commit, Refresh, NBS\_TESTE, NBS, Export, Import, Find, Replace, Copy, Paste, Cut, Delete, New, Open, Close, Minimize, Maximize, Close.
- Database Navigator:** Shows connections to NBS\_TESTE, NIPVIEW, TRON-PRODUÇÃO, and TRONCGI.DB.
- SQL Editor:** Contains a complex multi-table JOIN query with numerous conditions and WHERE clauses. The WHERE clause includes:

```
WHERE CAB.COD_OPERACAO = 60 AND CAB.COD_NATUREZA = 1202
```
- Results View:** Results 1. Displays the output of the query:

COD_EMPRESA	NOME	STATUS	INSCRIÇÃO ESTADUAL	LOCAL	CGC	CONTATO	INSCRIÇÃO MUNICIPAL	FACHADA
1		S		S				345
2		S		N				N.425
3		S		N				694
4		S		N				0
5		S		N				2.257
6		S		N				3525
7		S		N				N.425
- Bottom Bar:** Includes icons for Refresh, Save, Cancel, Export data, Row Number (200), Page Number (7), and a message: 7 row(s) fetched - 5ms, on 2023-09-21 at 10:57:21. Also shows system status: 30°C Ensolarado, POR PTB2, 21/09/2023, 11:04.

### **3. Coleta dos Dados**

### 3.1. Carregando os CSVs para o Azure

Foi criada uma conta de armazenamento no Azure e criado um container (mvp3container) para onde fizemos o upload do CSVs que estavam em minha máquina local conforme imagens abaixo. Esta operação consiste em colocarmos os CSVs no Datalake criado.

A screenshot of a Windows File Explorer window. The title bar reads "Dataset" and the address bar shows the path: "Este Computador > Documentos > JR > Ciéncia de Dados e Analytics > Engenharia de Dados > MVP > Dataset". The main pane displays a list of files: "GRUPOSEQP.csv" (modified 21/09/2023 09:20, 1 KB), "LOJAS.csv" (modified 21/09/2023 10:03, 7 KB), and "VENDAS.csv" (modified 21/09/2023 08:57, 6.900 KB). The left sidebar shows navigation links: "Acesso rápido", "Área de Trabalho", "Downloads", "Documentos", and "Imagens". A search bar at the top right says "Pesquisar em Dataset".

Segue abaixo a imagem mostrando os arquivos CSVs após upload para o Datalake mvp3container.

The screenshot shows the Microsoft Azure Storage Container blade for the 'mvp3storage' container. The left sidebar has sections for 'Visão geral', 'Diagnosticar e resolver problemas', 'IAM (Controle de Acesso)', and 'Configurações'. The main area displays a table of blobs:

Nome	Modificado	Camada de acesso	Status do arquivo	Tipo de blob	Tamanho	Estado de concessão
GRUPOSEQP.csv	21/09/2023, 11:57:54	Principal (Inferidos)		Blob de blocos	209 B	Disponível
LOJAS.csv	21/09/2023, 11:57:54	Principal (Inferidos)		Blob de blocos	6.97 KiB	Disponível
VENDAS.csv	21/09/2023, 11:58:10	Principal (Inferidos)		Blob de blocos	6.74 MiB	Disponível

## 3.2. Recursos criados

Foram criados alguns recursos na nuvem e um deles é o banco de dados SQL Server (mvp3sql) para onde faremos a ingestão dos dados após fazermos o ETL utilizando o recurso Data Factory (mvp3factory) que foi criado.

The screenshot shows the Microsoft Azure home blade. The top navigation bar includes 'Microsoft Azure', 'Atualizar', and a search bar. Below the navigation bar, there's a section for 'Serviços do Azure' with icons for 'Criar um recurso', 'Todos os recursos', 'Gerenciamento de custo +...', 'Data factories', 'Grupos de recursos', 'Assinaturas', 'Bancos de dados SQL', 'Contas de armazenamento', and 'Centro de Início Rápido'. The 'Mais serviços' link is also visible. The main area shows a list of 'Recursos' under the 'Recente' tab:

Nome	Tipo	Última visualização
mvp3factory	Data factory (V2)	16 minutos atrás
mvp3keyvault	Cofre de chaves	4 horas atrás
mvp3sql	SQL Server	4 horas atrás
MeuCatalogo	Conta do Microsoft Purview	5 horas atrás
MVP	Grupo de recursos	5 horas atrás
Azure subscription 1	Assinatura	6 horas atrás
mvp3sql	Banco de dados SQL	18 horas atrás
mvp3storage	Conta de armazenamento	5 dias atrás

At the bottom, there are links for 'Ver todos' and 'Navegar' with categories like 'Assinaturas', 'Grupos de recursos', 'Todos os recursos', and 'Painel'. The taskbar at the bottom shows various pinned icons and the system status bar indicates '30°C Ensolarado' and the date '21/09/2023'.

## 4. Carga dos Dados

O ETL foi realizado utilizando Data Factory. Foi feito primeiramente a etapa de extração dos dados criando um dataset para cada arquivo csv que tinham como origem o storage account no container mvp3container (Datalake). Na segunda etapa foi feita a transformação através da criação de um dataflow para fazer as conversões e eliminação de atributos. Na terceira etapa foi feita a ingestão dos dados no banco de dados SQL Server mvp3sql.

### 4.1. Extração

Criei um dataset para cada arquivo csv para posteriormente incluirmos os mesmos como sources de um dataflow.

The screenshot shows the Microsoft Azure Data Factory interface for creating a dataset. The left sidebar lists resources: Pipelines (pipeline1), Captura de Dados de Alterações..., Conjuntos de Dados (AzureSqlTable1, Grupos), Grupos, Fluxos de dados (dataflow1), and Power Query. The main area shows a 'Vendas' dataset of type 'DelimitedText'. The 'Conexão' tab is selected, showing 'Serviço vinculado' set to 'AzureBlobStorage1', 'Caminho do arquivo' set to 'mvp3container / Diretório / GRUPOSEQP.csv', and other settings like 'Delimitador de colunas' (Comma ()), 'Delimitador de linha' (Default (\r\n, or \r\n)), 'Codificação' (Padrão(UTF-8)), 'Caractere de aspas' (Double quote (")), and 'Caractere de escape' (Backslash (\)). The 'Esquema' and 'Parâmetros' tabs are also visible. The top navigation bar shows the URL <https://adf.azure.com/pt-BR/authoring/dataset/Grupos?factory=%2Fsubscriptions%2F6f1cde3d-c65c-4d45-8d7b-180c9a609d18%2FresourceGroups%2FMVP%2Fproviders%2FMicroso...>.

Observem que no momento da criação do dataset informamos qual é o delimitador das colunas do csv para importarmos os atributos corretamente, mas todos são reconhecidos como string e posteriormente no processo de transformação vamos colocar o tipo correto e a formatação.

Aqui foi criado um dataset para o csv Vendas e observem que na visualização dos dados o atributo DATA\_FATURAMENTO veio como string e no processo de transformação vamos colocar como tipo DATE.

	DATA_FATURAMENTO	CONTROLE	SERIE	DEPARTAMENTO	VENDEDOR	COD_CLIENTE_NOTA	COD...
1	2013-08-28T03:00:00Z	72254	1	Administracao	MTAVEIRA		1037
2	2013-08-28T03:00:00Z	72250	1	Administracao	MTAVEIRA		1037
3	2013-08-28T03:00:00Z	72251	1	Administracao	MTAVEIRA		1037
4	2017-12-29T02:00:00Z	59336	1	Administracao	MMACHADO		1696
5	2013-10-04T03:00:00Z	74418	1	Administracao	MTAVEIRA		8179

## 4.2. Transformação

Criei um dataflow e nele foram inseridos os datasets criados anteriormente. Neste passo foi definido os tipos dos atributos de cada dataset, também o formato. Nesse momento poderíamos renomear os atributos.

Foi feito nesta etapa de definição das fontes de dados do dataflow a atribuição do tipo correto do atributo, também a formatação correta do atributo. Note na imagem abaixo e relatado anteriormente o atributo DATA\_FATURAMENTO do dataset VENDAS foi definido agora como tipo DATE com formato yyyy-MM-dd. Atributos que representam moeda foram definidos como LONG e alguns somente numéricos como INTEGER.

Nome da coluna	Tipo	Formato
DATA_FATURAMENTO	date	yyyy-MM-dd
CONTROLE	integer	###,##
SERIE	integer	###,##
DEPARTAMENTO	string	Especificar o formato
VENDEDOR	string	Especificar o formato
COD_CLIENTE_NOTA	long	000,000,000
COD_CLIENTE	long	000,000,000
COD_OPERACAO	short	000
VALOR_VENDIDO	double	000,000,000,000
VALOR_ICMS	double	000,000,000,000

A Microsoft anuncia recentemente a versão prévia pública do Microsoft Fabric, uma maneira nova e interessante de criar a análise de dados em nuvem. Clique [aqui](#) para começar a usar o Fabric Data Factory!

Experiência de versão prévia Desligado

**Recursos do Data Factory**

- Pipelines: pipeline1
- Captura de Dados de Alterações...
- Conjuntos de Dados: 4
- Fluxos de dados: 1
- Power Query: 0

**dataflow1**

**Validar** Depuração de fluxo de dados Configurações de Depuração

**Vendas**

**AzureSqlTable1**

**pipeline1**

**Grupos**

**Lojas**

**Visualização de dados**

Número de linhas: 10 ATUALIZAR 0 EXCLUIR 0 EXECUTAR UPSERT 0 PESQUISAR 0 ERRO 0 TOTAL 10

Atualizar Typecast Modificar Mapa em descompasso Estatísticas Remover Exportar para CSV

COD_NOTA_MODELO	DESCRICAO_NOTA...
219	GRUPO1
213	GRUPO2
217	GRUPO3
208	GRUPO4
209	GRUPO5
211	GRUPO6

Foram renomeados alguns atributos para ficarem com a mesma identificação entre tabelas fato e tabela dimensão (Exemplo "GRUPO" no dataset Vendas e no dataset Grupos).

A Microsoft anuncia recentemente a versão prévia pública do Microsoft Fabric, uma maneira nova e interessante de criar a análise de dados em nuvem. Clique [aqui](#) para começar a usar o Fabric Data Factory!

Experiência de versão prévia Desligado

**Recursos do Data Factory**

- Pipelines: pipeline1
- Captura de Dados de Alterações...
- Conjuntos de Dados: 4
- Fluxos de dados: 1
- Power Query: 0

**dataflow1**

**Validar** Depuração de fluxo de dados Configurações de Depuração

**Vendas**

**AzureSqlTable1**

**pipeline1**

**Grupos**

**Lojas**

**Inspecionar**

Selecionar configurações Otimizar Visualização de dados

24	LINHA	abc string	LINHA
25	DESCRICAO_PRODUTO	abc string	DESCRICAO_PRODUTO
26	DESCRICAO_MODELO	abc string	DESCRICAO_MODELO
27	GRUPO	123 integer	COD_NOTA_MODELO
28	IDCN	abc string	IDCN
29	COTIPO_MODELO	abc string	COTIPO_MODELO
30	CODESPECIE_MODELO	abc string	CODESPECIE_MODELO
31	VFICILIO_ANOMODFIO	abc string	VFICILIO_ANOMODFIO

Depois dos tratamentos dos atributos incluímos no dataflow 2 JOINs. Fizemos isto para deixarmos apenas uma tabela desnormalizada na qual foi incluído o atributo nome da loja pelo primeiro join e o atributo nome do grupo no segundo join, servirão para concluirmos o nosso objetivo principal. Com isso preparamos os nossos dados para fazer a ingestão no banco de dados SQL Server (mvp3sql). Foi feito na imagem abaixo o primeiro JOIN entre o dataset VENDAS (select1) e o dataset LOJAS (select2) sendo utilizada o atributo COD\_EMPRESA dos dois datasets para incluirmos no dataset VENDAS a coluna que contém o nome da loja vindo do dataset LOJAS através do JOIN.

InPrivate https://adf.azure.com/pt-BR/authoring/dataflow1?factory=%2Fsubscriptions%2F6f1cde3d-c65c-4d45-8d7b-180c9a609d18%2FresourceGroups%2FMVP%2Fproviders%2FMic... Añadir favorito

A Microsoft anuncia recentemente a versão prévia pública do Microsoft Fabric, uma maneira nova e interessante de criar a análise de dados em nuvem. Clique [aqui](#) para começar a usar o Fabric Data Factory!

Microsoft Azure | Data Factory > mvp3factory Pesquisar alocador e documentação Experiência de versão prévia Desligado

Recursos do Data Factory <<

- Pipelines 1
  - pipeline1
- Captura de Dados de Alterações... 0
- Conjuntos de Dados 4
- Fluxos de dados 1
  - dataflow1
- Power Query 0

Filtrar os recursos por nome +

Validar Depuração de fluxo de dados Configurações de Depuração

Vendas dataflow1 AzureSqlTable1 pipeline1 Grupos Lojas

Configurações de junção Otimizar Inspecionar Visualização de dados

Nome do fluxo de saída \* join1 Saiba mais ⓘ Redefinir

Descrição Junção de Interna em 'select1' e 'select2'

Fluxo esquerdo \* select1

Fluxo direito \* select2

Tipo de junção \* Interna Externa completa Externa esquerda Externa direita Personalizado (cruzado)

37°C Pred ensolarado 16:21 POR PTB 25/09/2023

InPrivate https://adf.azure.com/pt-BR/authoring/dataflow1?factory=%2Fsubscriptions%2F6f1cde3d-c65c-4d45-8d7b-180c9a609d18%2FresourceGroups%2FMVP%2Fproviders%2FMic... Añadir favorito

A Microsoft anuncia recentemente a versão prévia pública do Microsoft Fabric, uma maneira nova e interessante de criar a análise de dados em nuvem. Clique [aqui](#) para começar a usar o Fabric Data Factory!

Microsoft Azure | Data Factory > mvp3factory Pesquisar alocador e documentação Experiência de versão prévia Desligado

Recursos do Data Factory <<

- Pipelines 1
  - pipeline1
- Captura de Dados de Alterações... 0
- Conjuntos de Dados 4
- Fluxos de dados 1
  - dataflow1
- Power Query 0

Filtrar os recursos por nome +

Validar Depuração de fluxo de dados Configurações de Depuração

Vendas dataflow1 AzureSqlTable1 pipeline1 Grupos Lojas

Configurações de junção Otimizar Inspecionar Visualização de dados

Fluxo direito \* select2

Tipo de junção \* Interna Externa completa Externa esquerda Externa direita Personalizado (cruzado)

Usar correspondência difusa ⓘ

Condições de junção \*

Esquerda: Coluna do select1 123 COD\_EMPRESA

Direita: Coluna do select2 123 COD\_EMPRESA

37°C Pred ensolarado 16:22 POR PTB 25/09/2023

## 4.3. Carga dos Dados

Nesta etapa criamos no dataflow a exportação dos dados SINK1 para o dataset AzureSqlTable1 e vinculamos a ela o nosso banco de dados mvp3sql criado para receber nossos dados.

Passamos as informações de nome do domínio (mvp3sql.database.windows.net), nome do banco de dados (mvp3sql) e usuário e senha que criamos no início do projeto quando instalamos o recurso Banco de Dados SQL.

A Microsoft anuncia recentemente a versão prévia pública do Microsoft Fabric, uma maneira nova e interessante de criar a análise de dados em nuvem. Clique [aqui](#) para começar a usar o Fabric Data Factory!

**Visualizar dados**

Serviço vinculado: AzureSqlDatabase1  
Objeto: dbo.vendas

	DATA_FATURAMENTO	CONTROLE	SERIE	DEPARTAMENTO	VENDEDOR	COD_CLIENTE_NOTA	COD...
1	08/28/2013	72254	1	Administracao	MTAVEIRA		1037
2	08/28/2013	72250	1	Administracao	MTAVEIRA		1037
3	08/28/2013	72251	1	Administracao	MTAVEIRA		1037
4	12/29/2017	59336	1	Administracao	MMACHADO		1696
5	10/04/2013	74418	1	Administracao	MTAVEIRA		8179

Validamos, depuramos e publicamos o nosso Dataflow e incluímos ele no nosso Pipeline. Validamos e depuramos também o nosso Pipeline. Agora nosso banco de dados está populado e partiremos agora para a ferramenta Azure Data Studio para executarmos algumas queries e testarmos e analisarmos o nosso Data Warehouse.

**Atividades**

- > Mover e transformar
- > Synapse
- > Azure Data Explorer
- > Azure Functions
- > Serviço em Lote
- > Databricks
- > Data Lake Analytics
- > Geral
- > HDInsight
- > Iteração e condicionais
- > Machine Learning
- > Power Query

**ID de execução do pipeline:** 9db02fb9-c889-48eb-b1d2-ada9d8c3fac5

**Status do pipeline:** Bem-sucedido

Nome da atividade	Status da atividade	Tipo de atividade	Inicio da execução	Duração	Runtime de integração	Propriedades
dataflow1	Bem-sucedido	Fluxo de dados	9/25/2023, 1:51:01 PM	1m 54s	debugpool-8Cores-Ge	

#### 4.4. Utilizando o Azure Data Studio para testar o database

Criei a conexão do banco de dados com Azure Data Studio e executei algumas queries fazendo referência aos objetivos a serem atendidos. Estas queries servirão para compararmos as quantidades com os resultados obtidos no dashboard que será criado.

## SELECT 1

Azure Data Studio interface showing a query results grid. The query is:

```
select * from vendas
```

The results grid has the following columns:

	DATA_FATURAMENTO	CONTROLE	SERIE	DEPARTAMENTO	VENDEDOR	COD_CLIENTE_NOTA	COD_CLIENTE	COD_OPERACAO	VALOR_VENDA
1	2013-08-28	72254	1	Administracao	MTAVEIRA	NULL	1037124000104	4	7400
2	2013-08-28	72250	1	Administracao	MTAVEIRA	NULL	1037124000104	4	7400
3	2013-08-28	72251	1	Administracao	MTAVEIRA	NULL	1037124000104	4	7400
4	2017-12-29	59336	1	Administracao	MACHADO	NULL	1696068134	4	7576
5	2013-10-04	74418	1	Administracao	MTAVEIRA	NULL	81790368120	4	1100
6	2013-10-03	74273	1	Administracao	MTAVEIRA	NULL	81790368120	4	9897
7	2013-08-28	72249	1	Administracao	MTAVEIRA	NULL	1037124000104	4	7810
8	2013-08-28	72244	1	Administracao	MTAVEIRA	NULL	1037124000104	4	7810
9	2013-08-21	71887	1	Administracao	EMOREIRA	NULL	1409622000130	4	7700
10	2013-08-28	72245	1	Administracao	MTAVEIRA	NULL	1037124000104	4	7810
11	2014-03-22	26317	1	Administracao	SMARTINS	NULL	2947542100	4	1700
12	2013-11-12	23100	1	Administracao	SJOSEMAR	NULL	2119820163	4	9897
13	2013-11-12	23101	1	Administracao	SJOSEMAR	NULL	2119820163	4	1080
14	2013-08-14	71583	1	Administracao	MTAVEIRA	NULL	2950786000133	4	7700

## SELECT 2 - Select com agrupamento por vendedor e descrição do grupo de equipamento.

Azure Data Studio interface showing a query results grid. The query is:

```
select vendedor, descricao, sum(qtde) from vendas  
where vendedor is not null and data_faturamento >= '01/01/2023'  
group by vendedor,descricao  
order by vendedor,descricao
```

The results grid has the following columns:

vendedor	descricao	(No column name)
20	EVERTONAS	GRUPO5
21	FELIPE.V	GRUPO2
22	FELIPE.V	GRUPO4
23	FELIPE.V	GRUPO5
24	FELIPE.V	GRUPO6
25	FELIPE.V	GRUPO9
26	FGOMES	GRUPO4
27	FGOMES	GRUPO5
28	FGOMES	GRUPO6
29	FGOMES	GRUPO9
30	FMENEZES	GRUPO4
31	FMENEZES	GRUPO5
32	FMENEZES	GRUPO6
33	FMENEZES	GRUPO9
34	FSATIL	GRUPO2
35	FSATIL	GRUPO3
36	FSATIL	GRUPO4
37	FSATIL	GRUPO5

## SELECT 3 - Select com totais vendidos por descrição do grupo de equipamentos.

File Edit View Help

Welcome mvp3sql.database.windows.net:mvp3sql (mvp3admin) - Azure Data Studio

CONNECTIONS SERVERS Tables Columns

```
1 select descricao, sum(qtde) from vendas
2 where vendedor is not null and data_faturamento >= '01/01/2023'
3 group by descricao
4 order by descricao
5
```

Results Messages

descricao	(No column name)
GRUPO10	6
GRUPO2	54
GRUPO3	48
GRUPO4	678
GRUPO5	255
GRUPO6	249
GRUPO9	297

Ln 2, Col 13 Spaces: 4 UTF-8 CRLF 7 rows MSSQL 00:00:00 mvp3sql.database.windows.net : mvp3sql POR 13:36 35°C Pred ensolarado PTB 25/09/2023 [26]

SELECT 4 - Select com totais vendidos por código do grupo de equipamentos.

File Edit View Help

Welcome mvp3sql.database.windows.net:mvp3sql (mvp3admin) - Azure Data Studio

CONNECTIONS SERVERS Tables Columns

```
1 select grupo, sum(qtde) from vendas
2 where vendedor is not null and data_faturamento >= '01/01/2023'
3 group by grupo
4 order by grupo
```

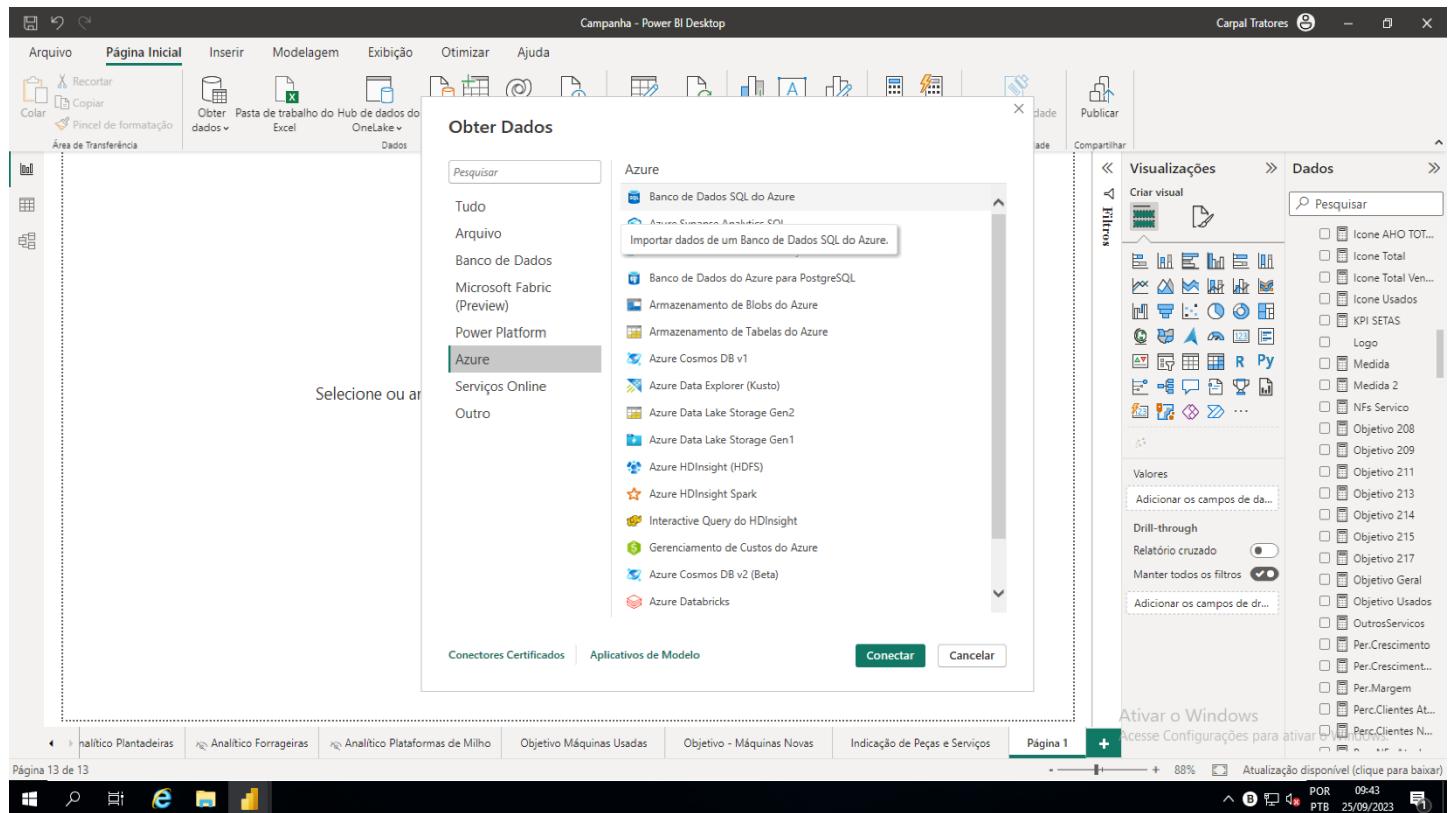
Results Messages

grupo	(No column name)
1	678
2	255
3	249
4	54
5	297
6	48
7	6

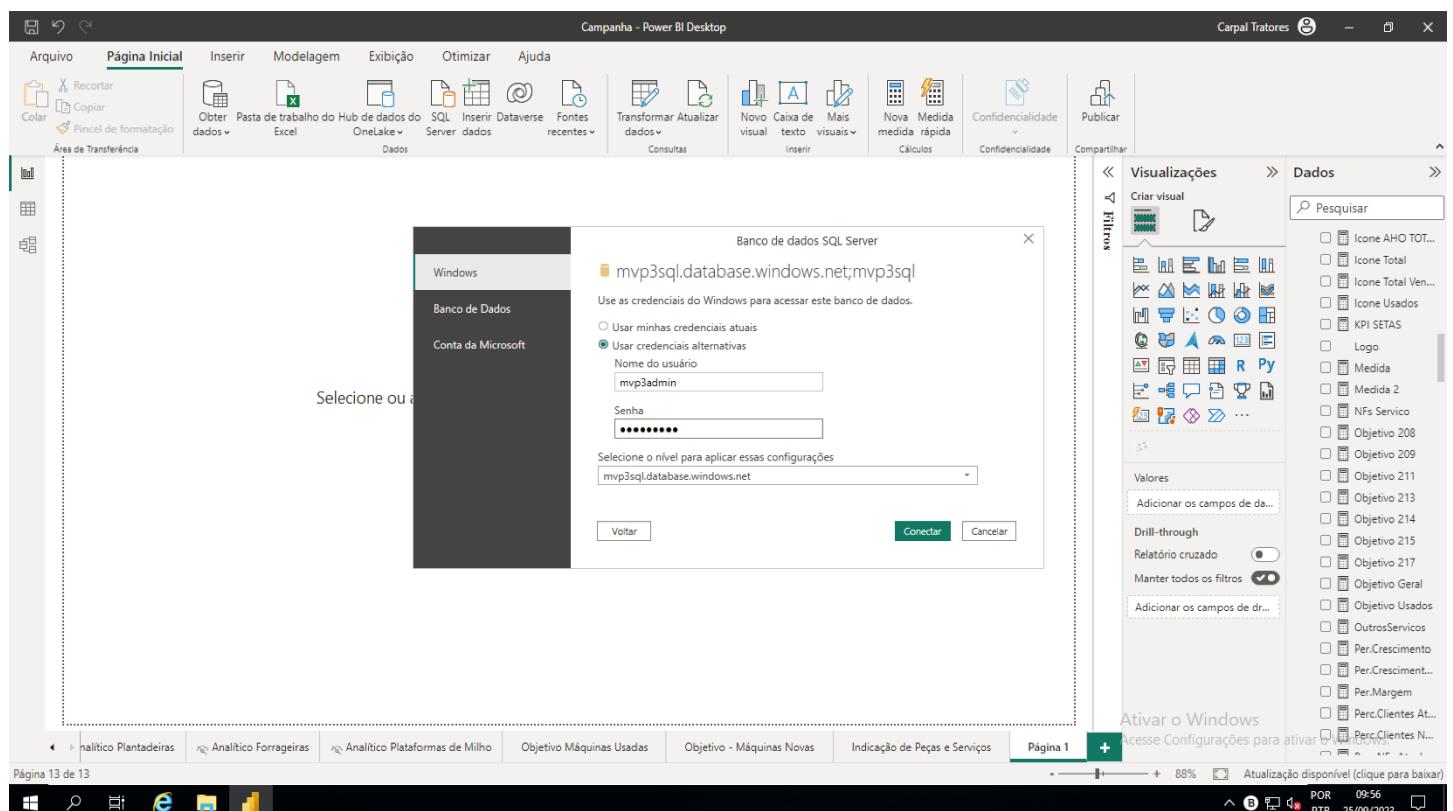
Ln 1, Col 13 Spaces: 4 UTF-8 CRLF 7 rows MSSQL 00:00:00 mvp3sql.database.windows.net : mvp3sql POR 14:08 35°C Pred ensolarado PTB 25/09/2023 [26]

## 4.5. Carga do database (mvp3sql) na Ferramenta de BI (POWER BI)

Vou utilizar o POWER BI para fazer o dashboard para atender o objetivo, desta forma vamos criar uma conexão com o banco de dados na nuvem Azure para obtermos os dados para o POWER BI. Demonstramos abaixo o processo de obter no POWER BI Desktop e selecionando a fonte BANCO DE DADOS SQL DO AZURE.



Fornecemos as credenciais cadastradas no AZURE e conectamos.



O POWER BI Desktop localiza o database e a tabela VENDAS é selecionada e carregada para desenvolvermos o dashboard.

The screenshot shows the Microsoft Power BI Desktop interface. In the top navigation bar, the 'Página Inicial' tab is selected. The main area displays a 'Navegador' (Navigator) window. Inside the Navigator, there's a tree view under 'Opções de Exibição' (Display Options) with nodes like 'mvp3sql.database.windows.net: mvp3sql [2]' and 'vendas'. A preview of the 'vendas' table is shown, containing columns: DATA\_FATURAMENTO, CONTROLE, SERIE, DEPARTAMENTO, VENDEDOR, and COD\_CI. The table data includes rows from 28/08/2013 to 21/08/2013. To the right of the Navigator, the 'Visualizações' (Visualizations) and 'Dados' (Data) panes are visible, showing various visualization icons and data-related settings.

## 5. Modelagem

Foi utilizado o Microsoft Purview para fazer a modelagem da tabela VENDAS. Fizemos a conexão do Purview com o nosso database, criamos um scan e localizamos nossa Azure Sql Table VENDAS para criarmos o catálogo dos atributos.

The screenshot shows the Microsoft Purview Data Catalog interface. The left sidebar has a navigation menu with items like 'Data sources', 'Collections', 'Monitoring', 'Metamodel', 'Asset types (preview)', 'Data sharing', 'Shares', 'Shared sources', 'Share invites', 'Source management', 'Scan rule sets', 'Pattern rules', 'Integration runtimes', 'Annotation management', 'Classifications', 'Classification rules', and 'Managed attributes'. The main area is titled 'Scope your scan' and shows a 'Data sources' section with a 'Map view' tab selected. It displays a single data source named 'Mvp3' under 'MeuCatalogo'. On the right, there's a 'Search' field and a tree view showing 'Mvp3' and 'dbo.vendas'. At the bottom, there are 'Continue', 'Back', and 'Cancel' buttons.

Microsoft Purview > MeuCatalogo

Data sources

Showing 1 collection, 1 data source

MeuCatalogo  
The root collection.

Mvp3  
Azure SQL Database

Apply Cancel

Abaixo mostramos a linhagem da tabela vendas que originou do nosso pipeline.

Microsoft Purview > MeuCatalogo

Browse assets

By collection By source type

View collection tree

MeuCatalogo

Sub collection(s)

Related

Narrow results by:

- Data
  - Dashboard
  - Data pipeline
  - Data share
  - Database
  - File
  - Folder
  - Report
  - Stored procedure
  - Table
- Business
  - Application service
  - Business process
  - Data domain
  - Glossary terms
  - Product

Data source type : All Add filter Clear all filters

Showing 1-4 out of 4 results Sort by: Relevance

- vendas Azure SQL Table
  - No description available
  - Fully qualified name: mssql://mvp3sql.database.windows.net/mvp3sql/dbo/vendas
  - Updated 22 minutes ago
- dbo Azure SQL Schema
  - No description available
  - Fully qualified name: mssql://mvp3sql.database.windows.net/mvp3sql/dbo
  - Updated 22 minutes ago
- mvp3sql Azure SQL Database
  - No description available
  - Fully qualified name: mssql://mvp3sql.database.windows.net/mvp3sql
  - Updated an hour ago
- mvp3sql.database.windows.net Azure SQL Server
  - No description available

Página 1 de 1

Catalogando os atributos da tabela VENDAS.

The screenshot shows the Microsoft Purview Catalog interface. In the top navigation bar, there are tabs for 'InPrivate', 'mvp3keyvault - Microsoft Azure', 'Edit vendas - MeuCatalogo - Mi...', and a search bar labeled 'Search catalog'. On the right side of the header, there are icons for user profile, notifications, and settings, along with the text 'joserobertofox@hotmail... DIRETÓRIO PADRÃO'.

The main content area is titled 'Edit "vendas"' and shows the 'Schema' tab selected. It displays a table with 13 columns. The columns are: DATA\_FATURAMENTO, CONTROLE, SERIE, DEPARTAMENTO, VENDEDOR, COD\_CLIENTE, COD\_OPERACAO, VALOR\_VENDIDO, VALOR\_ICMS, and ICMS\_SUBSTITUIDO. Each column has a 'Select...' dropdown menu next to it, indicating they are categorized. The 'Data type' column shows the data types for each column: date, int, int, nvarchar, nvarchar, bigint, smallint, float, float, float, and float. The 'Asset description' column provides a detailed description for each attribute.

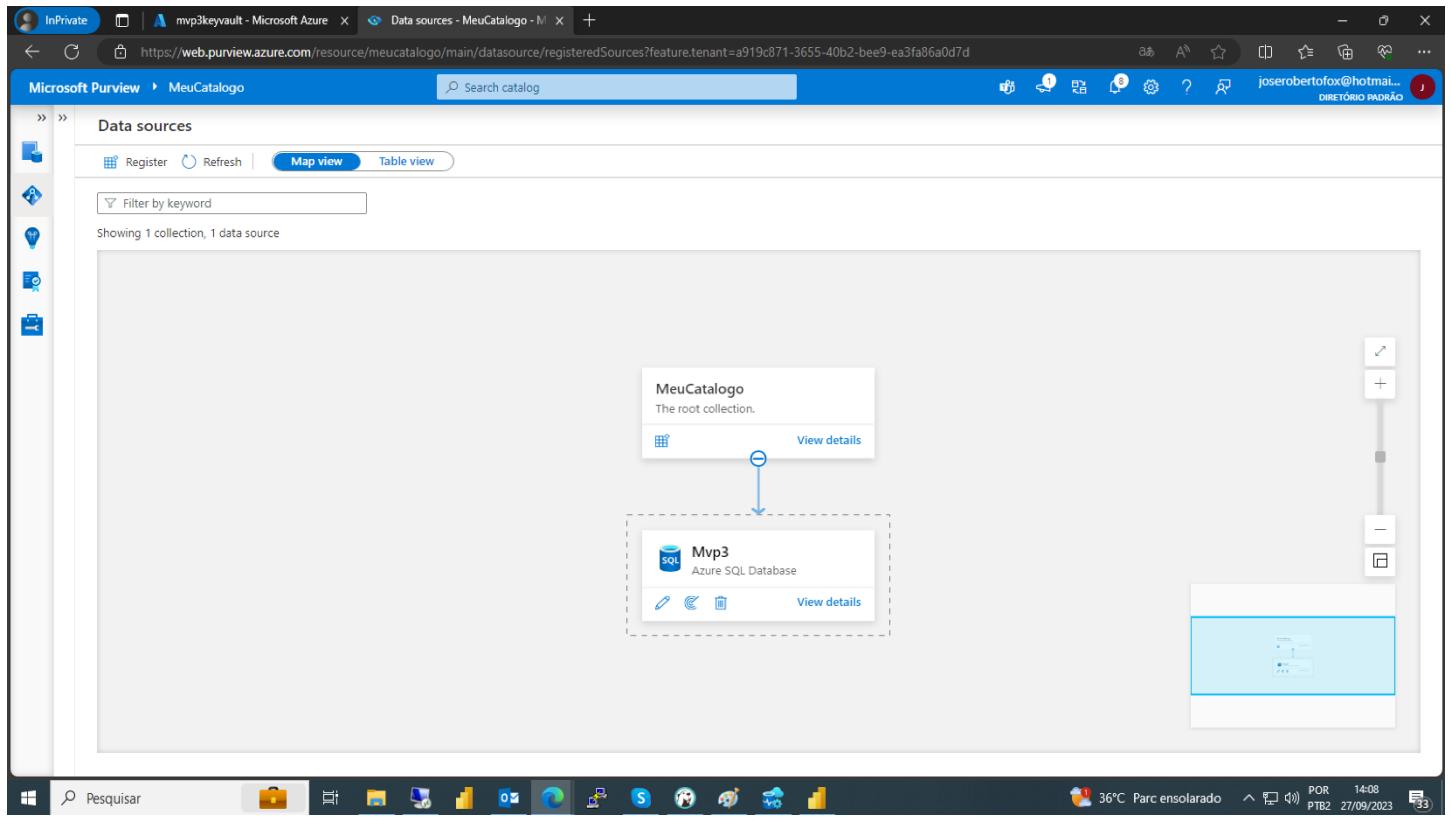
At the bottom of the table, there is a note: 'Valor do icms de substituição tributária' followed by a red link 'icms'. Below the table are 'Save' and 'Cancel' buttons. The Windows taskbar at the bottom shows various pinned icons and the system tray with the date and time.

Catalogando os atributos da tabela VENDAS. Como exemplo cito o atributo categórico STATUS com seus possíveis valores.

This screenshot is similar to the previous one, showing the 'Edit "vendas"' schema in Microsoft Purview Catalog. However, the 'STATUS' column is now highlighted, and its description is expanded.

The expanded description for the 'STATUS' column states: 'Identifica a situação da nota fiscal.' Below this, a list of possible values is provided: '0 - Ativa', '1 - Cancelada', '2 - Devolução Parcial', and '3 - Devolução Total'. The rest of the table and interface elements are identical to the first screenshot.

Mostramos na imagem abaixo que o banco de dados que contém a tabela VENDAS faz parte do catálogo de dados criado no Microsoft Purview.

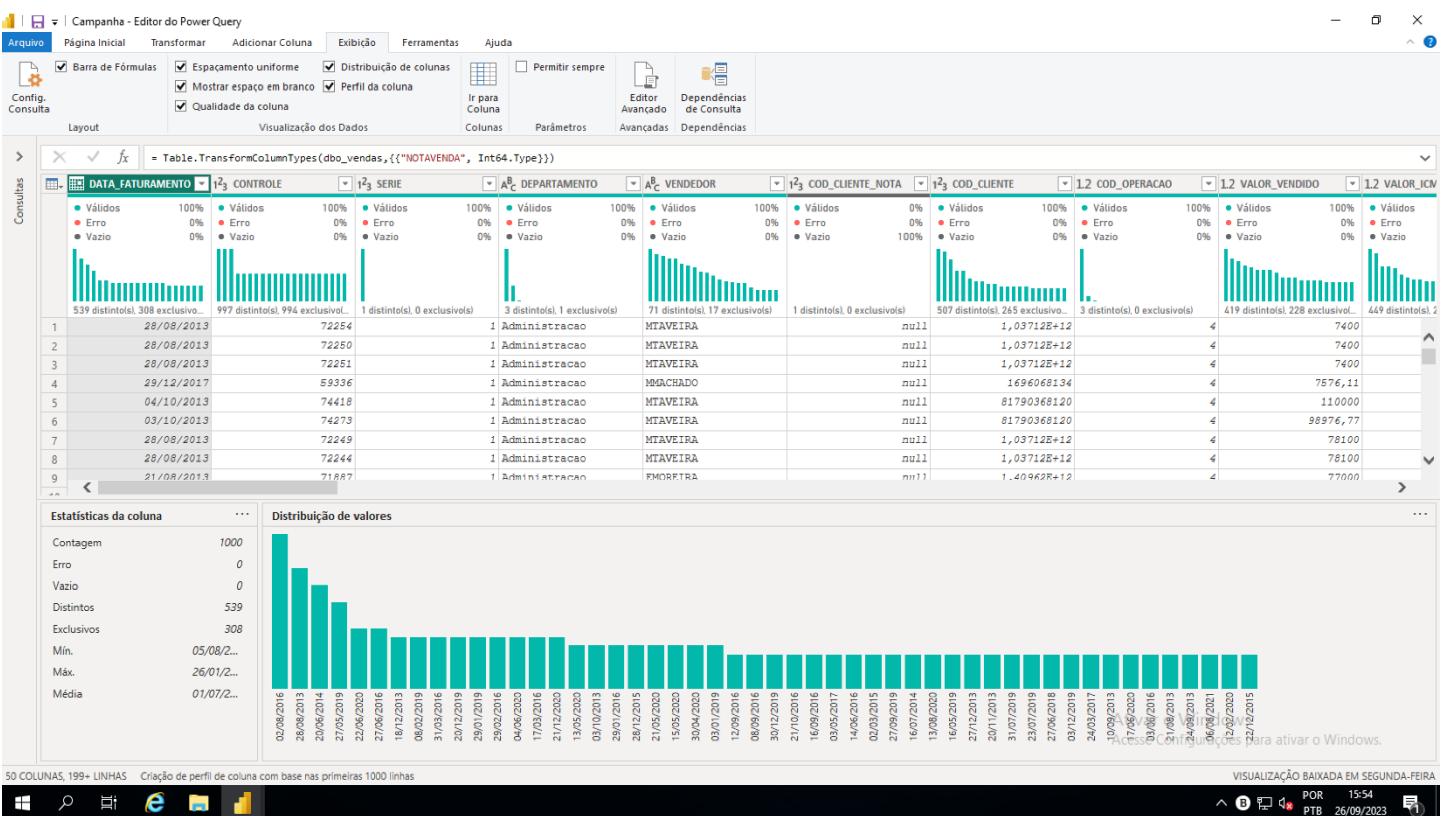


## 6. Análise dos Dados

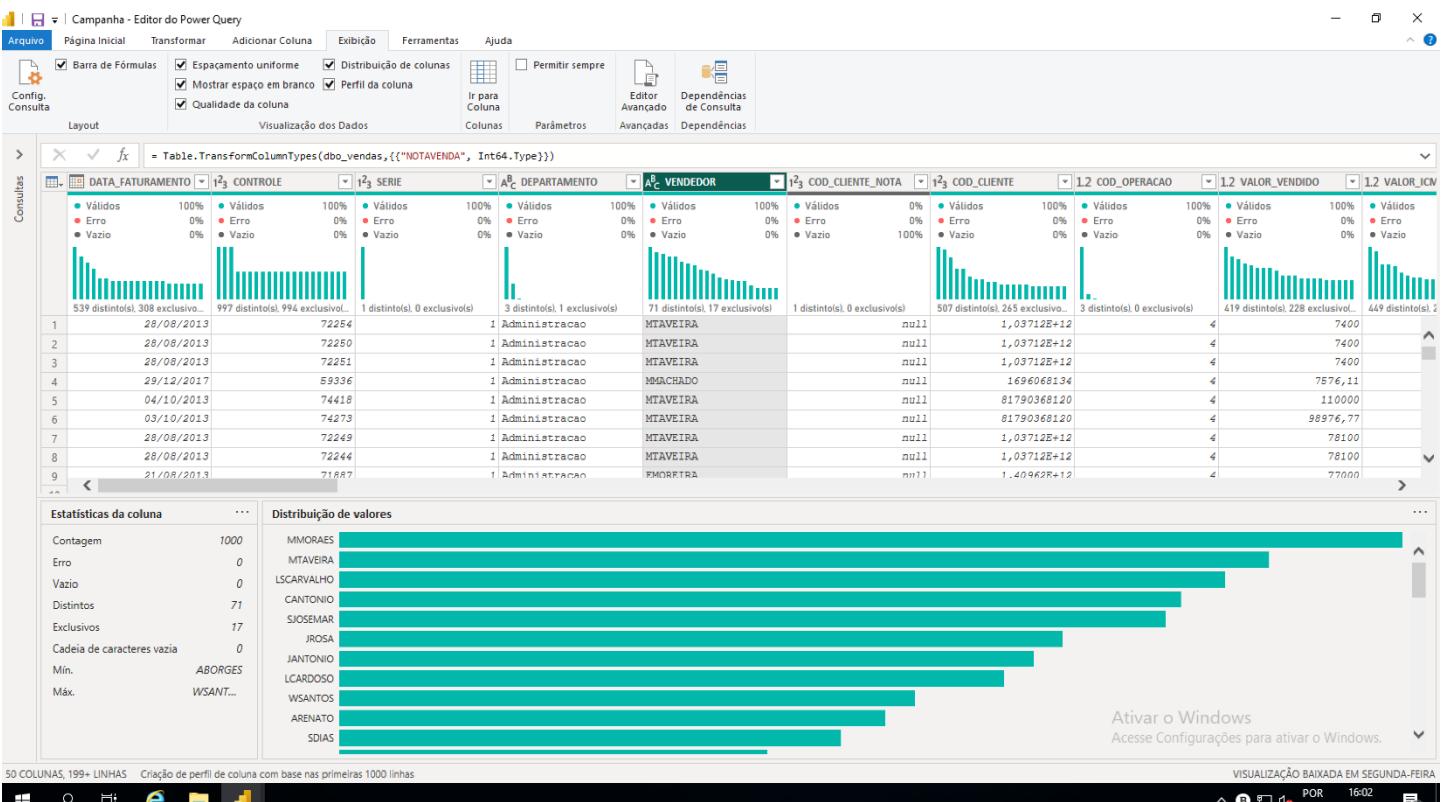
### 6.1. Qualidade dos dados

Analisei os atributos do meu conjunto de dados que será utilizado para resolver o objetivo e verifiquei que não há problemas pois ele já veio com um tratamento. Não há erros nos atributos e os valores faltantes aparecem apenas em atributos sem restrição de integridade. Como minha tabela VENDAS contém 50 atributos, mostrei abaixo algumas análises nos atributos que são base para construir meu dashboard para atender o objetivo.

Análise do atributo DATA\_FATURAMENTO que utilizaremos para fazer a segmentação por data de faturamento. Verificamos que os valores estão 100% válidos, nenhum erro e nenhum registro com valor deste atributo está vazio e nem nulo.



Análise do atributo VENDEDOR que vamos utilizar para fazer o ranking e o grid, sendo muito importante para atendermos o nosso objetivo. Verificamos que os valores estão 100% válidos, nenhum erro e nenhum registro com valor deste atributo está vazio e nem nulo.



O Atributo QTDE abaixo contém somente o valor "1" porque existe uma particularidade em nossas vendas de equipamentos que permitem somente 1 equipamento por NF, devido a algumas exigências das regras de negócio (business rules).

Campagna - Editor do Power Query

Arquivo Página Inicial Transformar Adicionar Coluna Exibição Ferramentas Ajuda

Barra de Fórmulas Espaçamento uniforme Distribuição de colunas Permitir sempre Ir para Coluna Editor Avançado Dependências de Consulta Config. Consulta Qualidade da coluna Visualização dos Dados Colunas Parâmetros Avançadas Dependências Layout

Consultas

NOTADEVOLUCAO QTD SOMAQUANT TICKET DRILL REFAT INDICOU NOME DESCRIAO

	A <sub>C</sub>	B <sub>C</sub>	C <sub>C</sub>	D <sub>C</sub>	E <sub>C</sub>	F <sub>C</sub>	G <sub>C</sub>	H <sub>C</sub>	I <sub>C</sub>	J <sub>C</sub>	K <sub>C</sub>	L <sub>C</sub>	M <sub>C</sub>	N <sub>C</sub>	O <sub>C</sub>	P <sub>C</sub>	Q <sub>C</sub>	R <sub>C</sub>	S <sub>C</sub>	T <sub>C</sub>	U <sub>C</sub>	V <sub>C</sub>	W <sub>C</sub>	X <sub>C</sub>	Y <sub>C</sub>	Z <sub>C</sub>								
100%	● Válidos	6%	● Válidos	100%	● Válidos	100%	● Válidos	100%	● Válidos	100%	● Válidos	100%	● Válidos	100%	● Válidos	100%	● Válidos	100%	● Válidos	100%	● Válidos	100%	● Válidos	100%	● Válidos	100%	● Válidos	100%						
0%	● Erro	0%	● Erro	0%	● Erro	0%	● Erro	0%	● Erro	0%	● Erro	0%	● Erro	0%	● Erro	0%	● Erro	0%	● Erro	0%	● Erro	0%	● Erro	0%	● Erro	0%	● Erro	0%	● Erro	0%				
0%	● Vazio	94%	● Vazio	0%	● Vazio	0%	● Vazio	0%	● Vazio	0%	● Vazio	0%	● Vazio	0%	● Vazio	0%	● Vazio	0%	● Vazio	0%	● Vazio	0%	● Vazio	0%	● Vazio	0%	● Vazio	0%	● Vazio	0%				
exclusivo(s)	57 distinto(s), 56 exclusivo(s)	12 QTD	1 distinto(s), 0 exclusivo(s)	1 S	1 S	1 S	1 S	1 S	1 S	1 S	1 S	1 S	1 S	1 S	1 S	1 S	1 S	1 S	1 S	1 S	1 S	1 S	1 S	1 S	1 S	1 S	1 S	1 S	1 S	1 S				
11	2631																																	
12	23100	null	1 S	S	S	S	S	S	S	S	S	S	S	S	S	S	S	S	S	S	S	S	S	S	S	S	S	S	S	S				
13	23101	null	1 S	S	S	S	S	S	S	S	S	S	S	S	S	S	S	S	S	S	S	S	S	S	S	S	S	S	S	S				
14	71583	null	1 S	S	S	S	S	S	S	S	S	S	S	S	S	S	S	S	S	S	S	S	S	S	S	S	S	S	S	S				
15	104743	null	1 S	S	S	S	S	S	S	S	S	S	S	S	S	S	S	S	S	S	S	S	S	S	S	S	S	S	S	S				
16	21523	null	1 S	S	S	S	S	S	S	S	S	S	S	S	S	S	S	S	S	S	S	S	S	S	S	S	S	S	S	S				
17	21529	null	1 S	S	S	S	S	S	S	S	S	S	S	S	S	S	S	S	S	S	S	S	S	S	S	S	S	S	S	S				
18	21530	null	1 S	S	S	S	S	S	S	S	S	S	S	S	S	S	S	S	S	S	S	S	S	S	S	S	S	S	S	S				
19	22403	null	1 S	S	S	S	S	S	S	S	S	S	S	S	S	S	S	S	S	S	S	S	S	S	S	S	S	S	S	S				
20																																		

Estatísticas da coluna

Contagem	1000
Erro	0
Vazio	0
Distintos	1
Exclusivos	0
NaN	0
Zero	0
Mín.	1
Máx.	1
Média	1

Distribuição de valores

Ativar o Windows  
Acesse Configurações para ativar o Windows.

Análise do atributo NOME que identifica a LOJA, vamos utilizar para fazer a segmentação de dados por loja. Verificamos que os valores estão 100% válidos, nenhum erro e nenhum registro com valor deste atributo está vazio e nem nulo.

Campanha - Editor do Power Query

Arquivo Página Inicial Transformar Adicionar Coluna Exibição Ferramentas Ajuda

Barra de Fórmulas Espaçamento uniforme Distribuição de colunas Permitir sempre

Mostrar espaço em branco Perfil da coluna Ir para Coluna Editor Avançado Dependências de Consulta

Qualidade da coluna Layout Visualização dos Dados Colunas Parâmetros Avançadas Dependências

Consultas

NOTADEVOLUCAO 1.2 QTDE SomaQuant TICKET DRILL REFAT INDICOU NOME DESCRIÇÃO

	Válidos	Erro	Vazio	Válidos	Erro	Vazio	Válidos	Erro	Vazio	Válidos	Erro	Vazio	Válidos	Erro	Vazio	Válidos	Erro	Vazio	Válidos	Erro	Vazio
100%	6%	0%	94%	100%	0%	0%	100%	0%	0%	100%	0%	0%	100%	0%	0%	100%	0%	0%	100%	0%	0%
0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%
0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%
exclusivo(s)	57	1	56	1	0	0	1	1	0	1	0	0	1	0	0	5	0	0	9	0	0
distintos(s), exclusivo(s)	56	0	56	0	0	0	1	0	0	1	0	0	1	0	0	0	0	0	0	0	0
11 26317	null	1	S	S	S	N	LOJA4	LOJA4	GRUPOS	11 26317	null	1	S	S	N	LOJA4	LOJA4	GRUPOS			
12 23100	null	1	S	S	S	N	LOJA4	LOJA4	GRUPOS	12 23100	null	1	S	S	N	LOJA4	LOJA4	GRUPOS			
13 23101	null	1	S	S	S	N	LOJA4	LOJA4	GRUPOS	13 23101	null	1	S	S	N	LOJA4	LOJA4	GRUPOS			
14 71583	null	1	S	S	S	N	LOJA2	LOJA2	GRUPOS	14 71583	null	1	S	S	N	LOJA2	LOJA2	GRUPOS			
15 104743	null	1	S	S	S	N	LOJA2	LOJA2	GRUPOS	15 104743	null	1	S	S	N	LOJA2	LOJA2	GRUPOS			
16 21523	null	1	S	S	S	N	LOJA4	LOJA4	GRUPOS	16 21523	null	1	S	S	N	LOJA4	LOJA4	GRUPOS			
17 21529	null	1	S	S	S	N	LOJA4	LOJA4	GRUPOS	17 21529	null	1	S	S	N	LOJA4	LOJA4	GRUPOS			
18 21530	null	1	S	S	S	N	LOJA4	LOJA4	GRUPOS	18 21530	null	1	S	S	N	LOJA4	LOJA4	GRUPOS			
19 22403	null	1	S	S	S	N	LOJA4	LOJA4	GRUPOS	19 22403	null	1	S	S	N	LOJA4	LOJA4	GRUPOS			
20																					

Estatísticas da coluna ... Distribuição de valores

Contagem	1000
Contagem	1000
Erro	0
Vazio	0
Distintos	5
Exclusivos	0
Cadeia de caracteres vazia	0
Mín.	LOJA2
Máx.	LOJA6

Ativar o Windows  
Acesse Configurações para ativar o Windows.

50 COLUNAS, 199+ LINHAS Criação de perfil de coluna com base nas primeiras 1000 linhas

VISUALIZAÇÃO BAIXADA EM SEGUNDA-FEIRA

Análise do atributo DESCRIÇÃO que identifica a descrição do grupo de equipamentos. Vamos utilizar este atributo para fazer o grid solicitado no objetivo. Verificamos que os valores estão 100% válidos, nenhum erro e nenhum registro com valor deste atributo está vazio e nem nulo.

## 6.2. Solução do problema

Depois da carga do database para a ferramenta Power BI, desenvovi um dashboard para atender os objetivos. Demonstro abaixo que o painel foi desenvolvido com conexão no banco de dados no Azure Sql Server (Data Warehouse).

Se compararmos os totais das colunas dos grupos de equipamentos vai coincidir com o select(s) número 3 e número 4 feitos no item 4.4 que faz o agrupamento das vendas por grupo. Comparando também o rankeado de número 12 (VENDEDOR: FELIPE V) com o select número 2, veremos que os

valores são equivalentes configurando assim a qualidade e acurácia dos dados. Segue abaixo o dashboard criado no POWER BI Desktop.

**CAMPANHA GANHA GANHA**

Indique, acumule pontos e ganhe uma viagem!

**TOP VENDEDORES**

Rank	Vendedor	208-GRUPO4	209-GRUPOS5	211-GRUPO6	213-GRUPO2	214-GRUPO9	217-GRUPO3	218-GRUPO10	Total
1	PAULO	60	9	15	6	21			111
2	JANTONIO	33	36	15	6	15	3		108
2	JAPARECIDA	60	3	3		30	12		108
3	LSCARVALHO	21	12	27	3	39			102
4	LROSA	72	6	6		3			87
5	FSATIL	33	18	9	6	12	6		84
6	VVIEIRA	36		18		12			69
7	RICHARD	12	30	12		12			66
8	GDIAS	39	6	6		6	6		63
8	LCARDOSO	18	12	9	3	18	3		63
9	MMORAES	24	3	9	3	9	12		60
10	WSANTOS	30		9	6	9			54
11	MCAMILORV	18	21	9		3			51
11	SDIAS	27	12	3	3	6			51
12	FELIPE.V	9	12	9	9	9			48
		678	255	249	54	297	48	6	1587

**1º LUGAR**  
Máquinas e Pós-Venda  
Consultor e Gerente: Viagem para o Mercosul com acompanhante

**2º LUGAR**  
Máquinas e Pós-Venda  
Consultor e Gerente: Viagem no Brasil com acompanhante

**Condições**

- Período: 21/08/23 a 14/12/23
- Participantes: Consultores de Máquinas/Pegas e Técnicos Mecânicos
- Os vencedores (1º e 2º lugar) serão os que acumularem mais pontos.

Objetivo - Máquinas Novas

Quanto aos objetivos criamos o grid conforme solicitado nos objetivos, satisfazendo aos valores totais por vendedor nas linhas e totais dos grupos de vendas por vendedor nas colunas, também o total geral dos grupos como totalizador das colunas. Os totais gerais batem com os resultados obtidos no Azure Data Studio. O gráfico de barras retrata o total de equipamentos vendidos por vendedor rankeando os mesmos. Chegamos à conclusão que os objetivos foram atendidos.

Dashboard publicado sendo executado.

**CAMPANHA GANHA GANHA**

Indique, acumule pontos e ganhe uma viagem!

**TOP VENDEDORES**

Rank	Vendedor	208-GRUPO4	209-GRUPOS5	211-GRUPO6	213-GRUPO2	214-GRUPO9	217-GRUPO3	218-GRUPO10	Total
1	PAULO	60	9	15	6	21			111
2	JANTONIO	33	36	15	6	15	3		108
2	JAPARECIDA	60	3	3		30	12		108
3	LSCARVALHO	21	12	27	3	39			102
4	LROSA	72	6	6		3			87
5	FSATIL	33	18	9	6	12	6		84
6	VVIEIRA	36		18		12			69
7	RICHARD	12	30	12		12			66
8	GDIAS	39	6	6		6	6		63
8	LCARDOSO	18	12	9	3	18	3		63
9	MMORAES	24	3	9	3	9	12		60
10	WSANTOS	30		9	6	9			54
11	MCAMILORV	18	21	9		3			51
11	SDIAS	27	12	3	3	6			51
12	FELIPE.V	9	12	9	9	9			48
		678	255	249	54	297	48	6	1587

---

## 7. Autoavaliação

---

Após a realização de todas as etapas do MVP conseguimos atingir todos os itens requeridos no objetivo e retratamos a visualização das informações em um dashboard conforme solicitado. A maior dificuldade foram as configurações dos recursos no Azure e não familiaridade com a ferramenta Purview usada para catalogar os dados.

---