

# CHECKPOINT 3

## Projeto ETL - Dimensão Tempo | DBurger Parte I

**Professor:** Prof. Salvio Padlipskas

**Turma:** 2TSCPR

### **Grupo Insight Hunters:**

- Diego Alves Moreira – RM552603
- Guilherme Yuiti Matsushita Nakamura – RM85355
- João Pedro de Souza Nunes – RM: 554066

**SÃO PAULO, 2025**

## Sumário

<b>1) INTRODUÇÃO .....</b>	<b>3</b>
<b>2) PREPARAÇÃO DO AMBIENTE .....</b>	<b>3</b>
<b>3) DESENVOLVIMENTO .....</b>	<b>6</b>
a) <b>Criação do Data Factory e Conexões .....</b>	<b>6</b>
b) <b>dataset do Storage Account e Containers .....</b>	<b>8</b>
c) <b>Implementação do Dataflow .....</b>	<b>9</b>

## 1) INTRODUÇÃO

Este trabalho tem como objetivo demonstrar a implementação de um *data warehouse* utilizando ferramentas da plataforma Azure, como *Storage Account*, *Azure SQL Server* e *Data Factory*. O processo inclui a criação de containers, configuração de bancos de dados, integração de dados via *dataflows* e geração de saídas em formato CSV e SQL. A estrutura segue as boas práticas de integração e transformação de dados para apoio à tomada de decisões.

## 2) PREPARAÇÃO DO AMBIENTE

Primeiro preparamos o ambiente foi necessário criar as ferramentas em cloud que utilizaríamos no Azure Data Factory

### Storage Account

Microsoft Azure

Search resources, services, and docs (G+)

Copilot

RM554066@fiap.com.br  
FIAP - FACULDADE DE INFORMÁTICA E ADMINISTRAÇÃO PAULISTA

### Create a storage account

Azure Storage is a Microsoft-managed service providing cloud storage that is highly available, secure, durable, scalable, and redundant. Azure Storage includes Azure Blobs (objects), Azure Data Lake Storage Gen2, Azure Files, Azure Queues, and Azure Tables. The cost of your storage account depends on the usage and the options you choose below. [Learn more about Azure storage accounts](#)

**Project details**

Select the subscription in which to create the new storage account. Choose a new or existing resource group to organize and manage your storage account together with other resources.

Subscription \* Azure for Students

Resource group \* (New) cp3-dw  
[Create new](#)

**Instance details**

Storage account name \* storagedbm554066

Region \* (South America) Brazil South  
[Deploy to an Azure Extended Zone](#)

Primary service \* Azure Blob Storage or Azure Data Lake Storage Gen 2

Performance \* ☒ Standard: Recommended for most scenarios (general-purpose v2 account)  
☐ Premium: Recommended for scenarios that require low latency.

Redundancy \* Locally redundant storage (LRS)

[Previous](#) [Next](#) [Review + create](#)

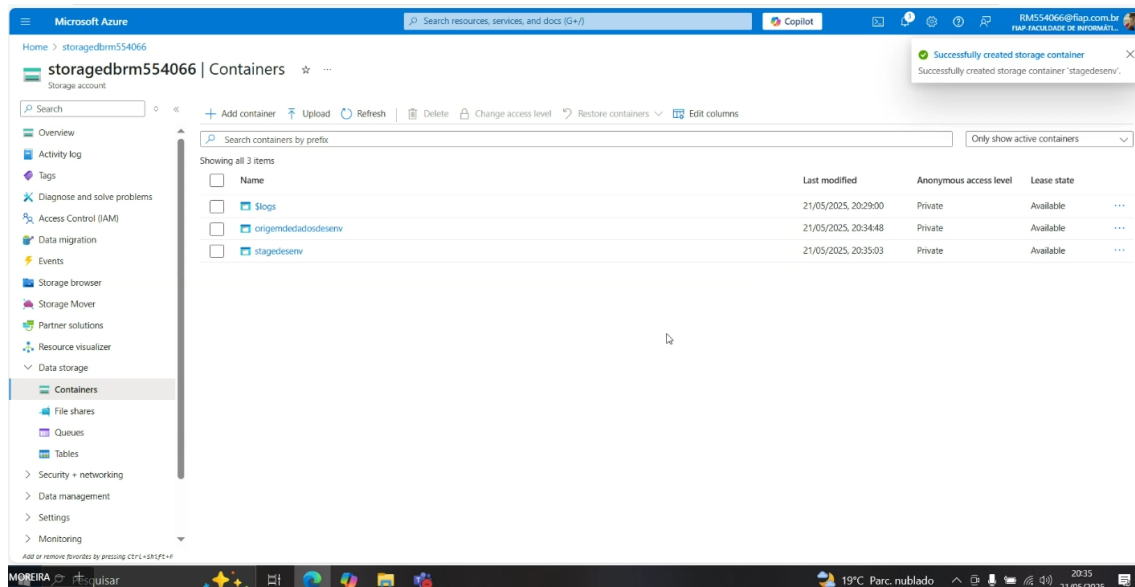
[Give feedback](#)

quarta-feira, 21 de maio de 2025

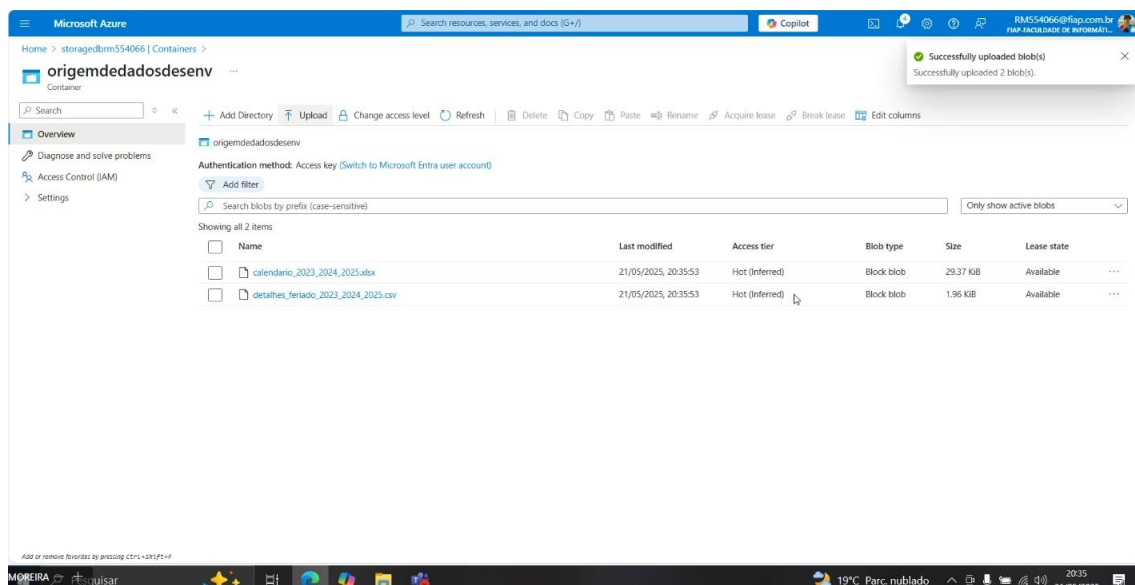
19°C Parc. nublado

2025  
21/05/2025

Foi criado os dois containers



Enviando os dados de origem



## Criando SQL Server via Script

```
--          SCRIPT PARA CRIAÇÃO DO SQL SERVER E BANCO SQL

-- Cria o SGBD Azure SQL chamado sqlserver-kidelicia
az sql server create -l brazilsouth -g cp3-dw -n kidelicia -u admsql -p
db@FIAP25 --enable-public-network true

-- Cria o banco de dados chamado dburgerdbstage0
az sql db create -g cp3-dw -s kidelicia -n dburgerdbstage0 --service-
objective Basic --backup-storage-redundancy Local --zone-redundant false

-- Libera o acesso para qualquer endereço IP da rede (somente para testes
iniciais)
az sql server firewall-rule create -g cp3-dw -s kidelicia -n AllowAll --
start-ip-address 0.0.0.0 --end-ip-address 255.255.255.255

-- Usuário e senha do banco de dados
-- admsql
-- db@FIAP25
```

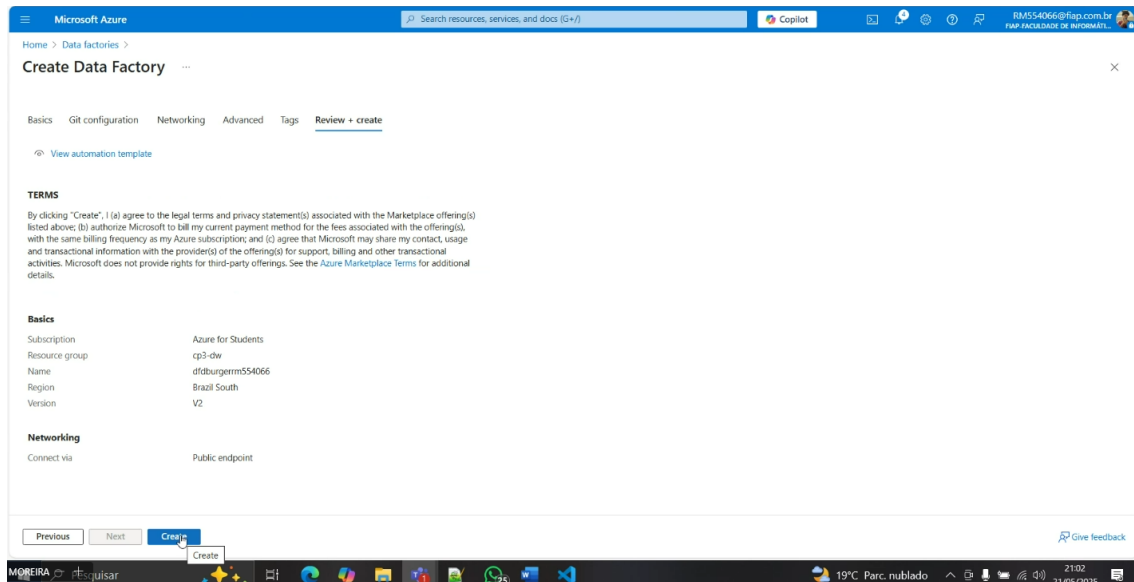
The screenshot displays the Microsoft Azure portal interface. At the top, the search bar shows 'Search resources, services, and docs (G+7)'. The main content area is for the resource group 'cp3-dw'. The 'Overview' tab is active, showing a table of resources:

Name	Type	Location
dburgerdbstage0 (kidelicia/dburgerdbstage0)	SQL database	Brazil South
kidelicia	SQL server	Brazil South
storagebrms54066	Storage account	Brazil South

Below the table, a terminal window shows the execution of the Azure CLI commands from the previous block, confirming the creation of the SQL server, database, and firewall rule.

```
ms54066 [ ~ ]$ az sql server firewall-rule create -g cp3-dw -s kidelicia -n AllowAll --start-ip-address 0.0.0.0 --end-ip-address 255.255.255.255
{
  "type": "Microsoft.Sql/servers/firewallRules",
  "id": "/subscriptions/az1be329-b0d5-4257-a490-6e2b61d82e7e/resourceGroups/cp3-dw/providers/Microsoft.Sql/servers/kidelicia/firewallRules/AllowAll",
  "name": "AllowAll",
  "resourceGroup": "cp3-dw",
  "startIpAddress": "0.0.0.0",
  "endIpAddress": "255.255.255.255",
  "type": "Microsoft.Sql/servers/firewallRules"
}
```

## Criando Data Factory

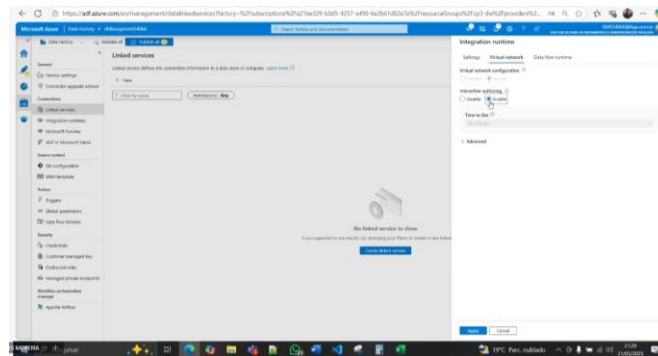


### 3) DESENVOLVIMENTO

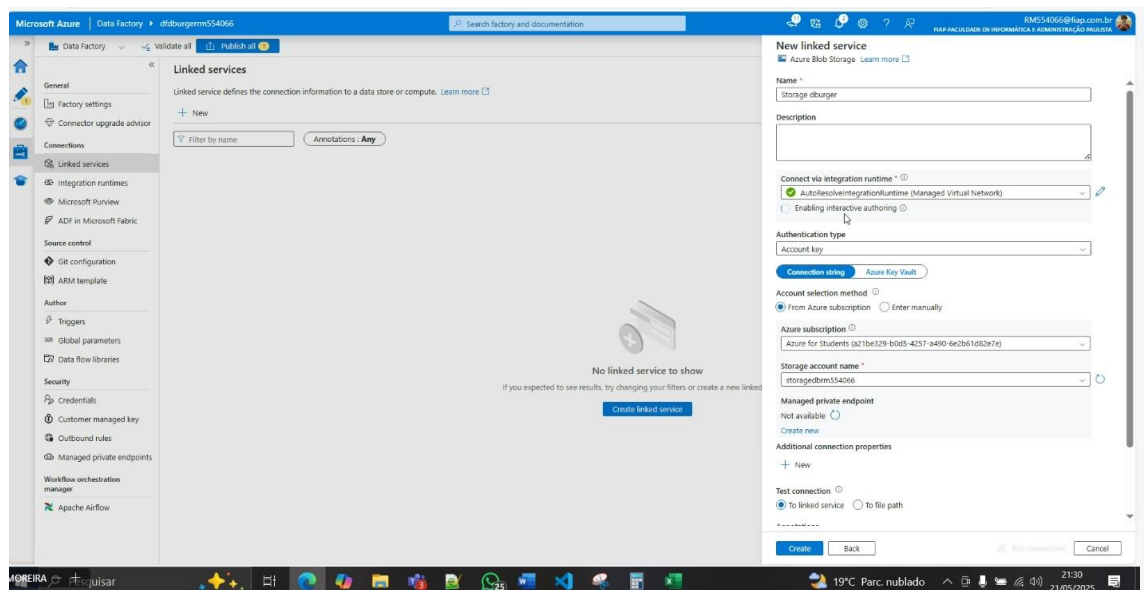
#### a) Criação do Data Factory e Conexões

O *Data Factory* foi configurado para orquestrar os pipelines. A conexão com o *Storage Account* e o SQL Server garantiu a integração entre fontes de dados.

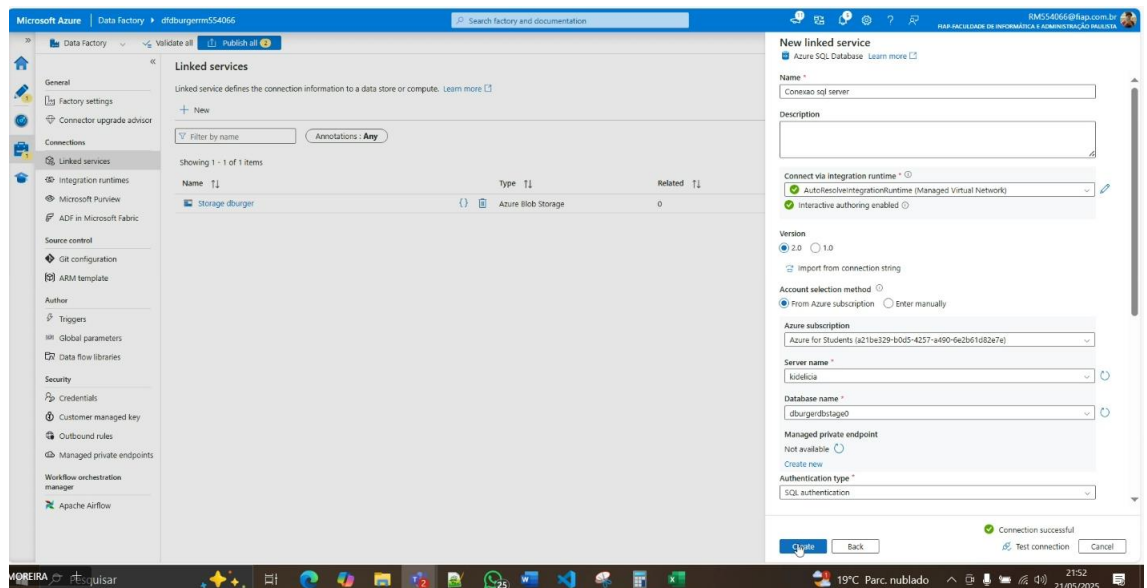
**Em todos os passos habilitamos a virtualização para facilidade de ver os dados**



Storage Account



SQL Server

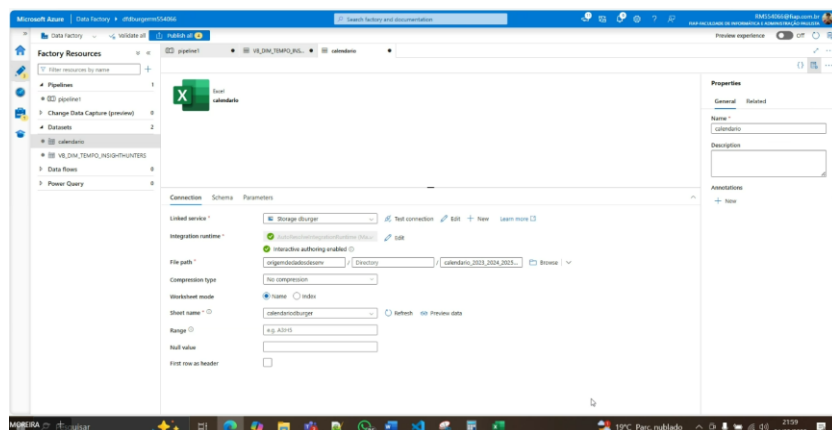


## b) dataset do Storage Account e Containers

Para armazenar os dados brutos, foi criado um Storage Account no Azure, contendo dois containers: um para dados de entrada (input) e outro para saída (output).

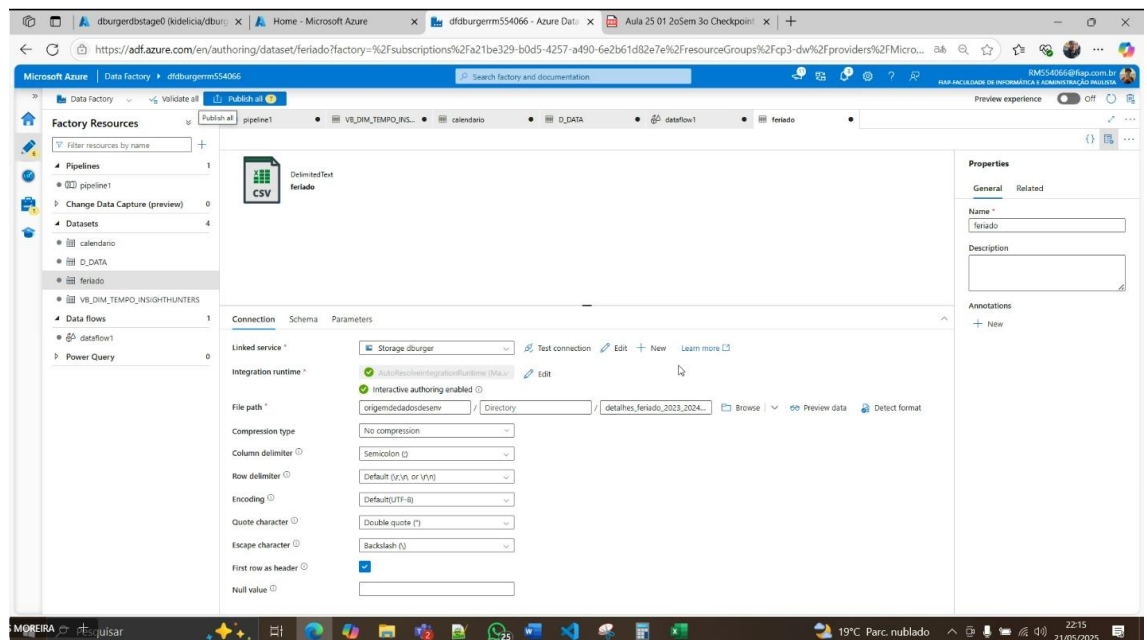
### Dados de entrada (input):

calendario\_2023\_2024\_2025.xlsx

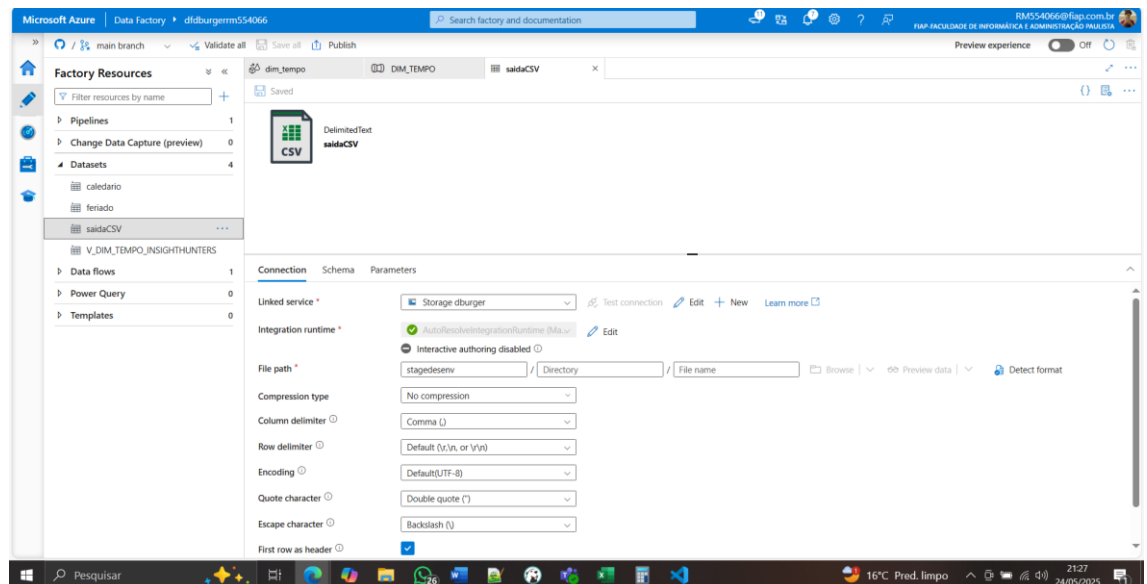




detalhes\_feriado\_2023\_2024\_2025.csv



Dados de saída (output):

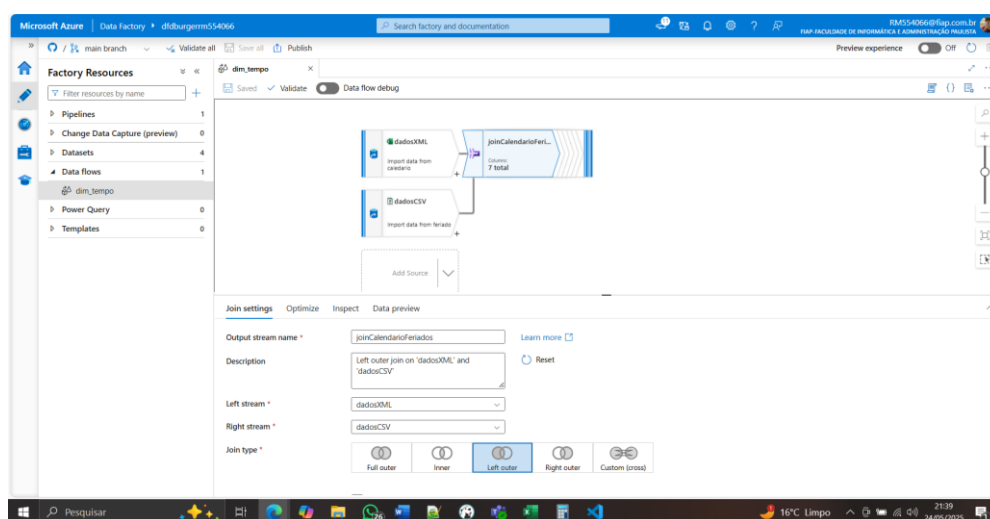


### C) Implementação do Dataflow

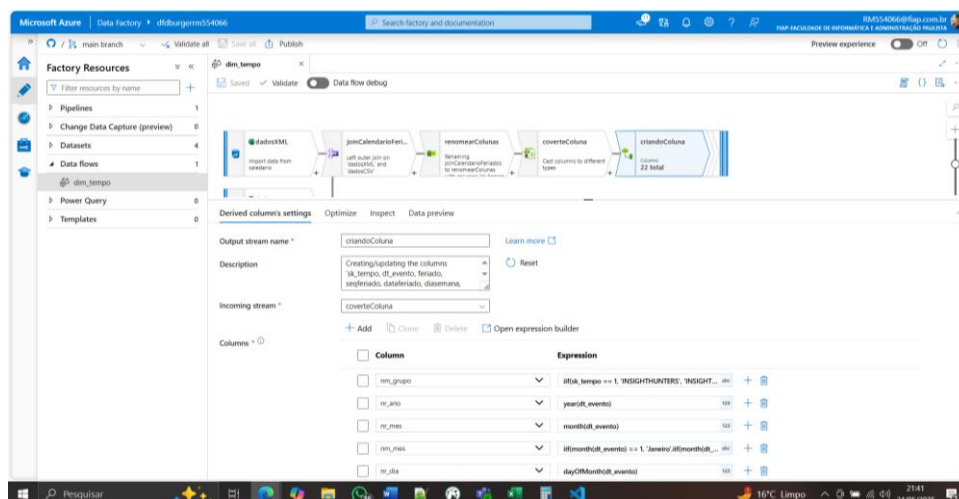
O *dataflow* foi construído em etapas:

FIAP - FACULDADE DE INFORMÁTICA E ADMINISTRAÇÃO PAULISTA

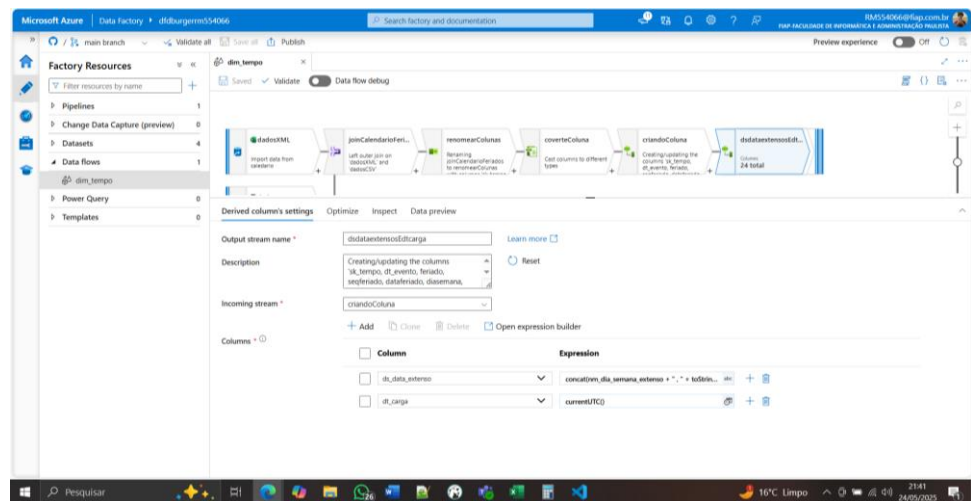
1. **Junção (Left Join)** das tabelas "calendario" e "feriados".



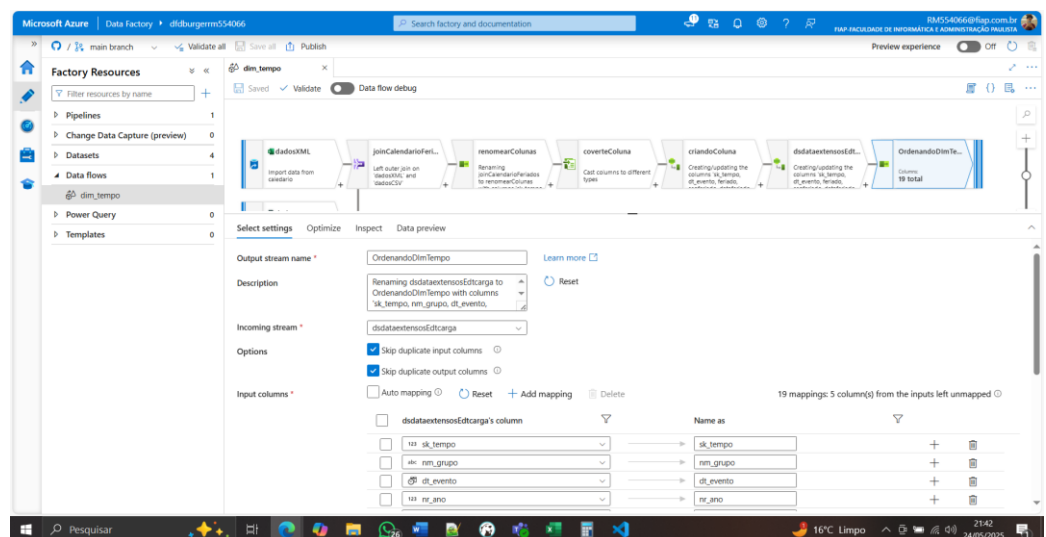
2. **Seleção e Transformação** de colunas, incluindo conversão de *dt\_evento* para *timestamp*.



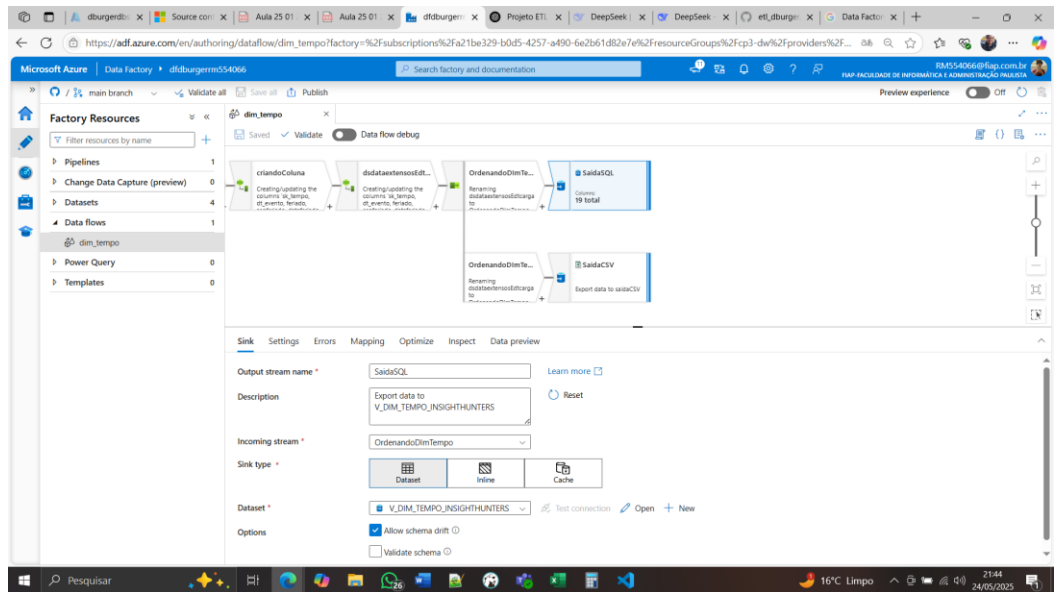
3. **Criação de Colunas Derivadas** para atender à estrutura da dimensão tempo.



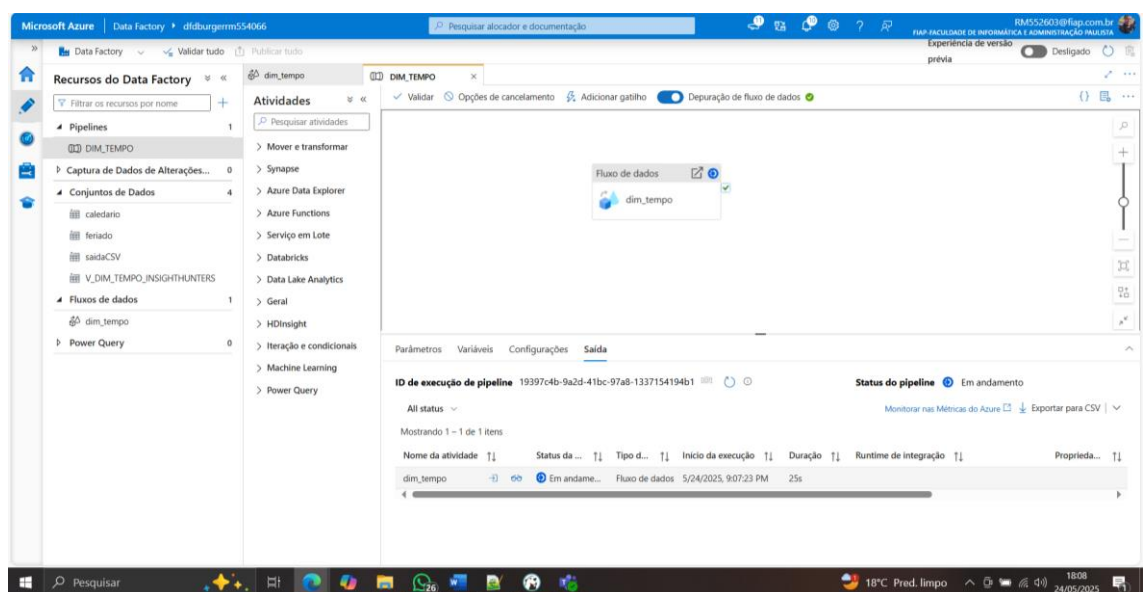
#### 4. **Ordenação** das colunas conforme padrão *DIM TEMPO*.



#### 5. **Saída** para CSV e SQL Server.



A execução do pipeline validou o fluxo, confirmando a integridade dos dados.



Microsoft Azure | Data Factory | dlburgerm554066 | Pesquisar alocador e documentação

Todas as execuções de pipeline -> DIM\_TEMPO - A atividade é executada -> dim\_tempo

dim\_tempo  
Tempo de Inicialização do cluster: 931ms Número de transformações: 10 Status do fluxo de dados: Success

Atualizar Atualização automática Ligado Editar o fluxo de dados

Coletores Todos os fluxos

Coletor	Status	Tempo de processamento	Maior tempo de proces...	Linhas gravadas	Fases	Linhagem
SaídaCSV	Bem-sucedido	2s 201ms	1s 141ms	1096		
SaídaSQL	Bem-sucedido	559ms	1s 141ms	1096		

Pesquisar 18°C Pred. limpo 1808 24/05/2025

Microsoft Azure | Search resources, services, and docs (5+)

Home > cp3-dw > dburgerdbstage0 (kidelicia/dburgerdbstage0) | Query editor (preview)

SQL database

Query 1 x Query 2 x

Run Cancel query Save query Export data as Show only Editor

```
1 SELECT * FROM [dbo].[VB_DIM_TEMPO_INSIGHTHUNTERS]
2
```

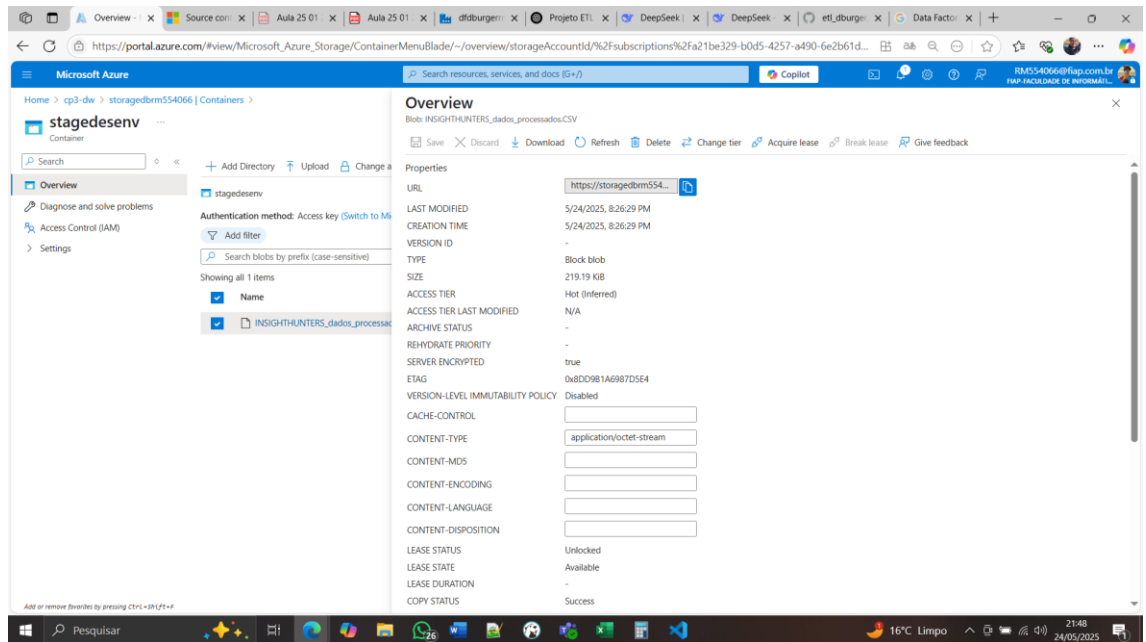
Results Messages

Search to filter items...

sk_tempo	nm_grupo	dt_evento	nr_ano	nr_mes	nm_mes	nr_dia
20230101	INSIGHTHUNTERS	2023-01-01T00:00:00.0000000	2023	1	Janeiro	1
20230102	INSIGHTHUNTERS	2023-01-02T00:00:00.0000000	2023	1	Janeiro	2
20230103	INSIGHTHUNTERS	2023-01-03T00:00:00.0000000	2023	1	Janeiro	3
20230104	INSIGHTHUNTERS	2023-01-04T00:00:00.0000000	2023	1	Janeiro	4
20230105	INSIGHTHUNTERS	2023-01-05T00:00:00.0000000	2023	1	Janeiro	5
20230106	INSIGHTHUNTERS	2023-01-06T00:00:00.0000000	2023	1	Janeiro	6

Query succeeded | 0s

Pesquisar 16°C Limpo 21:47 24/05/2025



---

### 3. CONCLUSÃO

O projeto demonstrou a viabilidade de integrar ferramentas Azure para construção de um *data warehouse*. A automatização via *Data Factory* e a estruturação de *dataflows* otimizaram a transformação e carga de dados, atendendo aos requisitos de negócio. Para trabalhos futuros, recomenda-se a adição de camadas de segurança e monitoramento contínuo.