



HANDS-ON LAB

20 de agosto de 2025

Luis Assunção

Sr. Solutions Architect

luis.assuncao@databricks.com





Hands-On Labs

Democratizando dados com Databricks



Apresentação do Time Databricks



Luis
Assunção



Gabriel
Caballeria



Julandro
Figueiró



Vinicius
Fialho



Jean
Ertzogue

Sr. Solutions Architect
DATABRICKS

Solutions Architect
DATABRICKS

Solutions Architect
DATABRICKS

Solutions Engineer
DATABRICKS

Account Executive
DATABRICKS



Hands-On Labs

Democratizando dados com Databricks



AGENDA

1	Conceituação do Lakehouse e Plataforma Databricks	20 min	09:00 às 09:20
2	LAB 01 - Carregando os Dados	20 min	09:20 às 09:40
3	LAB 02 - Entendimento de Dados com Genie	30 min	09:40 às 10:10
4	LAB 03 - Análise Exploratória com Dashboards AI/BI	50 min	10:10 às 11:00
5	LAB 04.1 - Manipulação de Dados e Geo	30 min	11:00 às 11:30
6	LAB 04.2 - Aplicações Avançadas de Dados	30 min	11:30 às 12:00



Jornada de Dados

É importante ter uma "fundação" que esteja preparada para o crescimento da empresa e dos novos projetos.

Não dá para pular STEPs.



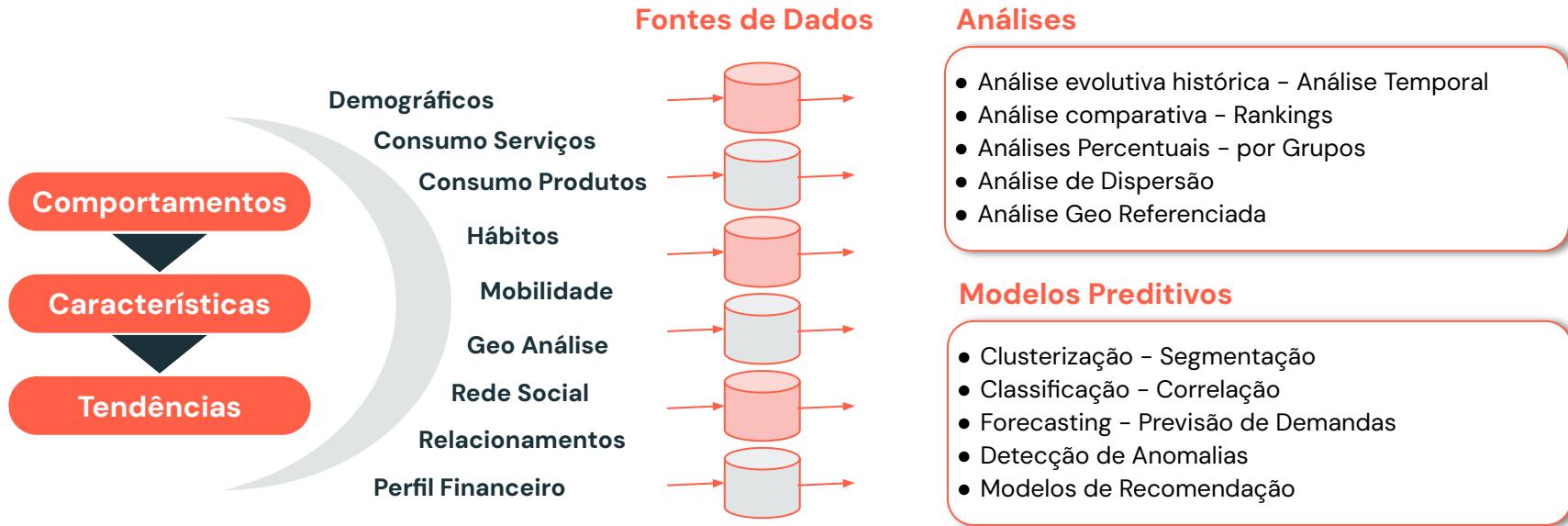
PLANO DE PROJETO – Aplicações Analíticas & IA



Objetivos Propósito	Tangibilizar Problema	Definir Hipóteses	Preparar Dados	Análises e Estudos	Modelos de IA e Agentes IA
Necessidades	Métricas	Teses	Fontes internas	Análise Exploratória	Avaliação Comportamentos
Problemas	Dimensionar	Possíveis causas	Fontes externas	Estatísticas	Modelos Preditivos
Impactos	Tangibilizar	Potenciais estudos	Enriquecimento	Ranking	Alertas / Ações
Ganhos	Quantificar	Avaliações iniciais	Tratamento	Correlações	Recomendações
Desafios	Qualificar	Correlações	Categorização	Geo Análise	Projeções
Melhorias	Fontes de Dados	Conhecimento Tácito	Geolocalização	Aplicações	Agentes IA
		Técnica 5W2H			



PLANO DE PROJETO - Definição das Fontes e das Análises





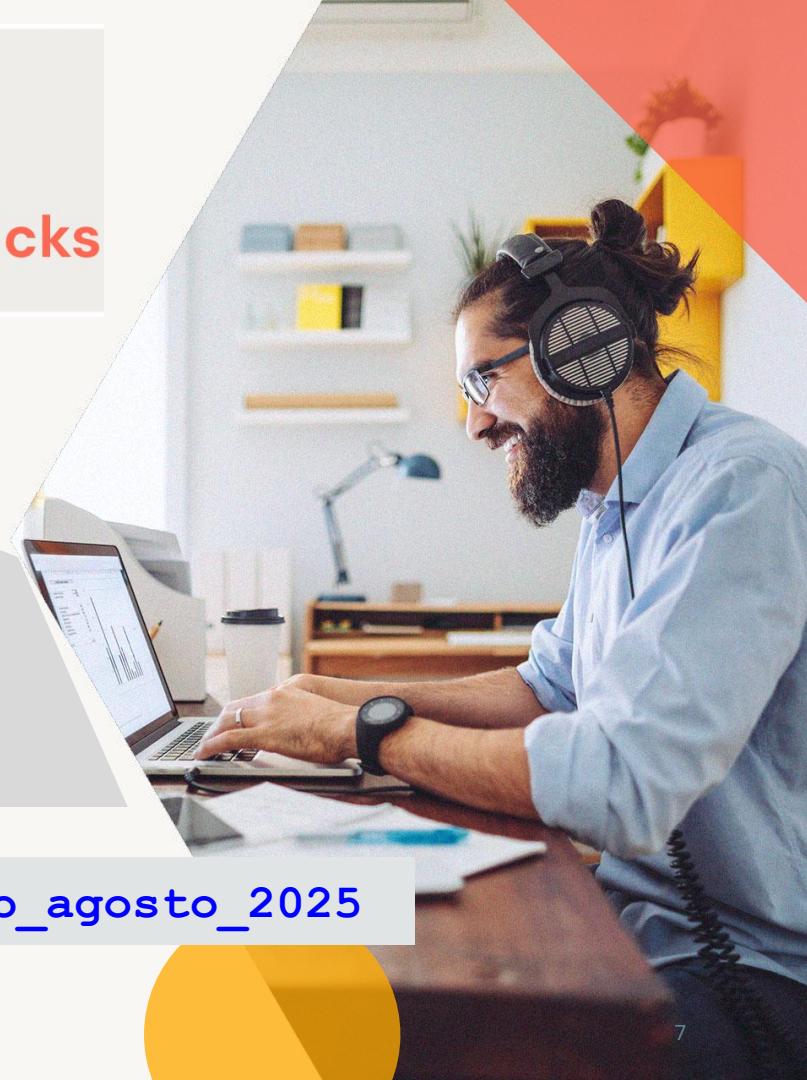
Hands-On Labs

Democratizando dados com Databricks

LABORATÓRIO 1

Carregando os dados...

https://github.com/Databricks-BR/lab_agosto_2025



Escopo dos Dados

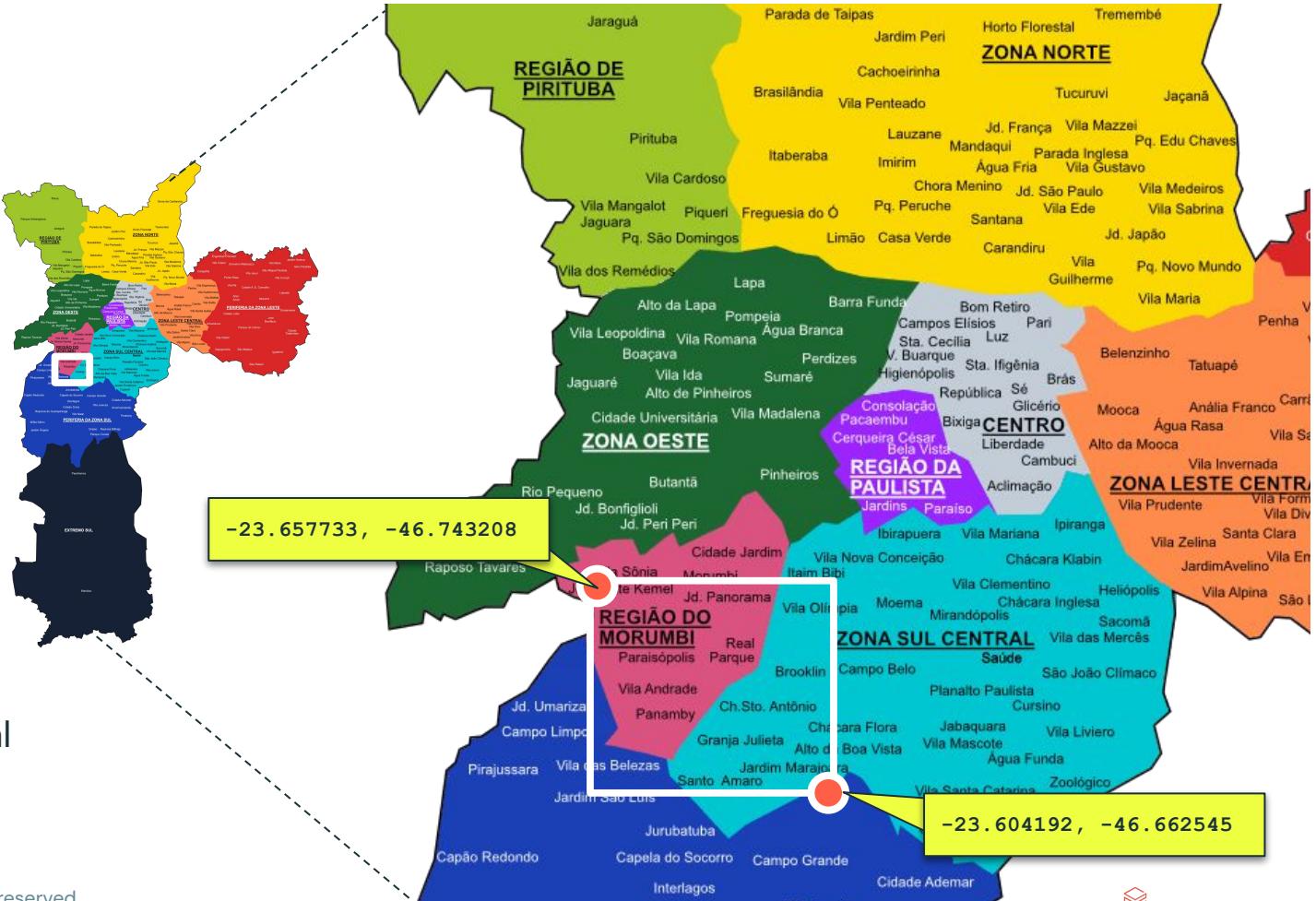
Temática: Inadimplência

Dados fictícios

+

Dados abertos

- Receita Federal
 - Geo Ref



Receita Federal

 > Repositório de Dados Abertos da RFB

Index of /cnpj/dados_abertos_cnpj/2025-08

	Name	Last modified	Size	Description
	Parent Directory		-	
	Cnaes.zip	2025-08-10 14:50	22K	
	Empresas0.zip	2025-08-10 14:51	430M	
	Empresas1.zip	2025-08-10 14:51	74M	
	Empresas2.zip	2025-08-10 14:51	75M	
	Empresas3.zip	2025-08-10 14:51	81M	
	Empresas4.zip	2025-08-10 14:51	86M	
	Empresas5.zip	2025-08-10 14:51	93M	
	Empresas6.zip	2025-08-10 14:51	90M	
	Empresas7.zip	2025-08-10 14:51	94M	
	Empresas8.zip	2025-08-10 14:52	95M	
	Empresas9.zip	2025-08-10 14:52	90M	
	Estabelecimentos0.zip	2025-08-10 14:54	1.7G	
	Estabelecimentos1.zip	2025-08-10 14:54	324M	
	Estabelecimentos2.zip	2025-08-10 14:54	320M	
	Estabelecimentos3.zip	2025-08-10 14:55	336M	
	Estabelecimentos4.zip	2025-08-10 14:55	339M	
	Estabelecimentos5.zip	2025-08-10 14:55	320M	
	Estabelecimentos6.zip	2025-08-10 14:56	320M	
	Estabelecimentos7.zip	2025-08-10 14:56	350M	



Databricks-BR / lab_agosto_2025

< > Code

Issues

Pull requests

Actions

Projects

Wiki



lab_agosto_2025

lab_agosto_2025 / dados /

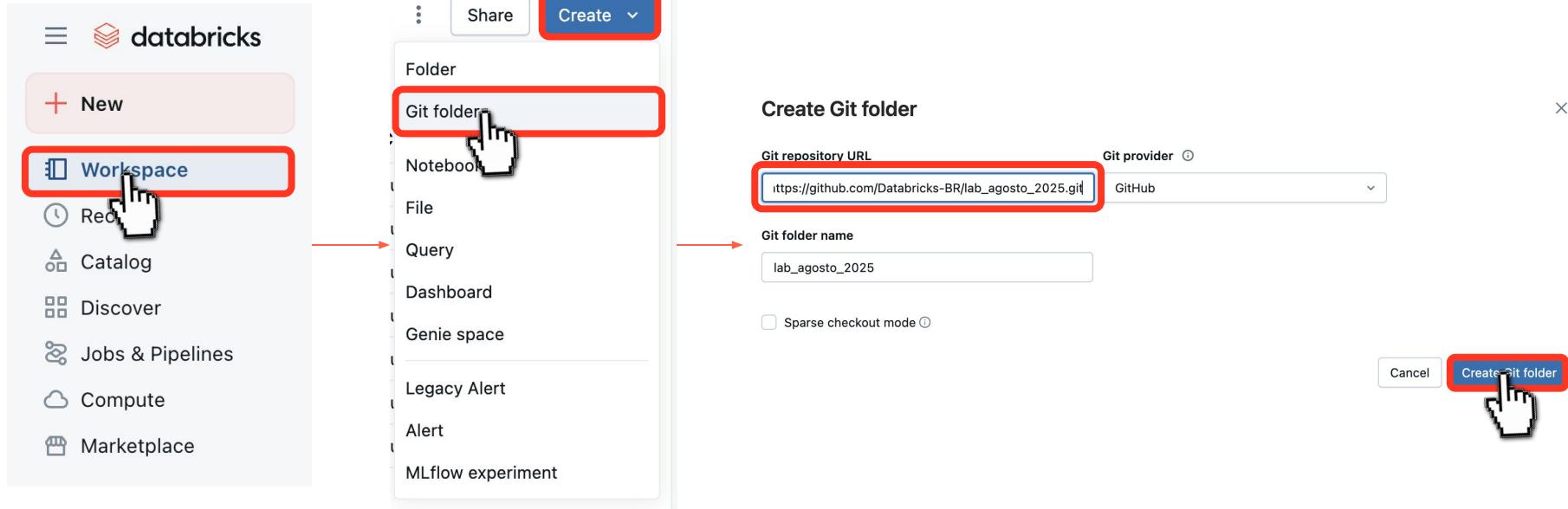


Databricks-BR dados

Name	Last commit message
..	
README.md	dados
bairros_sao_paulo.geojson	dados
cnae.csv	dados
empresas_sp.csv	dados
faturamento.csv	dados
ibge_senso.csv	dados
municipios.csv	dados
naturezas.csv	dados



Passo 1 – Importar o Notebook



https://github.com/Databricks-BR/lab_agosto_2025

Passo 2 – Alterar os parâmetros

The screenshot shows the Databricks workspace interface. On the left, there's a sidebar with various navigation options like Workspace, Recents, Catalog, Discover, Jobs & Pipelines, Compute, Marketplace, SQL, SQL Editor, Queries, Dashboards, Genie, Alerts, Query History, and SQL Warehouses. A large white hand cursor is positioned over the 'Workspace' icon in the sidebar.

The main area shows a notebook titled 'Parâmetros Iniciais'. The first cell contains Python code:

```
import pandas as pd
from pyspark.sql import SparkSession

url = f"https://raw.githubusercontent.com/Databricks-BR/lab_agosto_2025/main/dados/"
```

The second cell contains Python code with a red placeholder text:

```
catalog_name = f"workshop_08_2025"
```

A white hand cursor is positioned over the word 'catalog_name' in this cell.

The third cell contains Python code with a green placeholder text:

```
#schema_name = f"<<<-----COLOQUE SEU USER NAME AQUI ----->>>"
```

The word 'schema_name' is highlighted in green, and a white hand cursor is positioned over it.

At the top right of the workspace, there's a search bar with the placeholder 'Search data, notebooks, recents, and more...' and a '⌘ + P' keyboard shortcut.

Passo 3 – Conectar o recurso de "compute"

The screenshot shows a Databricks notebook interface. The notebook is titled "Parâmetros Iniciais". The code cell contains the following Python code:

```
import pandas as pd
from pyspark.sql import SparkSession

url = f"https://raw.githubusercontent.com/Databricks-BR/lab_agosto_2025/main/dados/"

catalog_name = f"workshop_08_2025"
```

The sidebar on the right displays the connection status. A hand cursor is pointing at the "Connected" status under "Serverless". The sidebar also shows recent resources and environment details.

Connected
Connected
Serverless

Recent resources
Field Eng Shared UC LTS Cluster 16.4 LTS + 6 workers

Environment version 1 • CPU



Passo 4 – Criar o Catálogo e o Schema

✓ 3 minutes ago (2s)

8



```
create_catalog = f"CREATE CATALOG IF NOT EXISTS {catalog_name}"  
spark.sql (create_catalog)
```

```
create_schema = f"CREATE SCHEMA IF NOT EXISTS {catalog_name}.{schema_name}"  
spark.sql (create_schema)
```

> [See performance \(2\)](#)

DataFrame []

Passo 5 – Executar a importação (notebook)

The screenshot shows the Databricks notebook interface with two code cells. A large hand cursor icon is positioned over the first code cell.

Code Cell 9: Gravando a tabela DELTA - FATURAMENTO

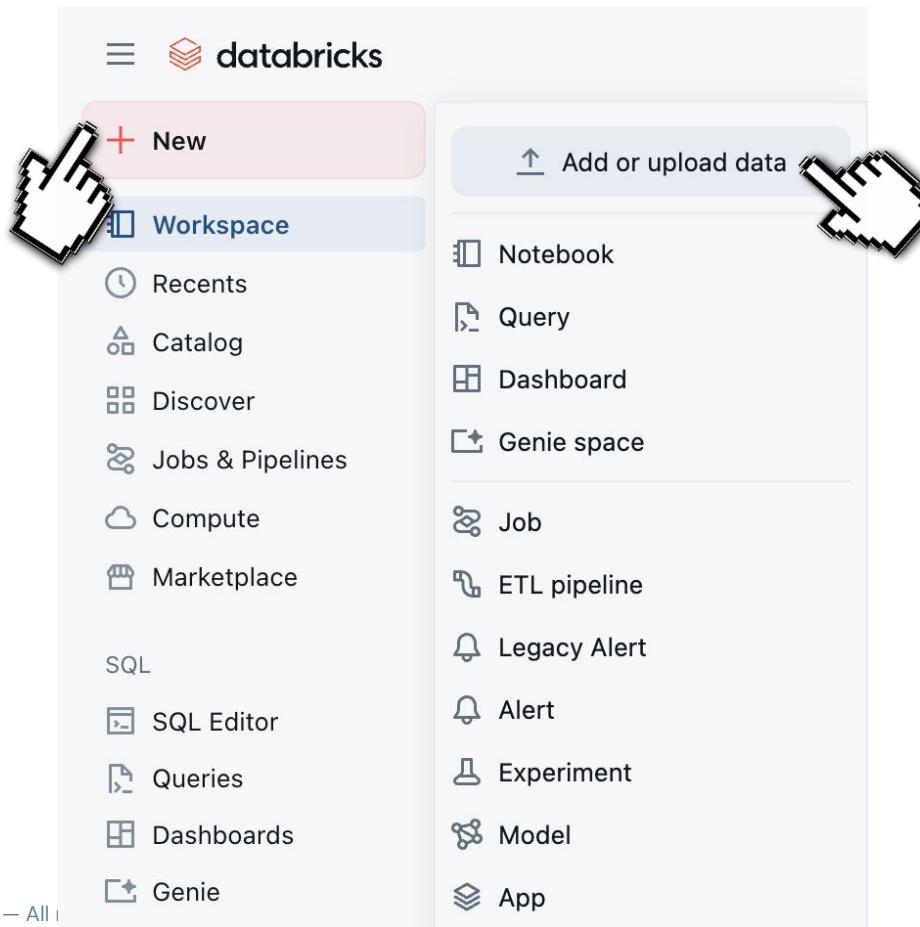
```
%python  
entity_name = f"faturamento"  
  
table_name = f"{catalog_name}.{schema_name}.{entity_name}"  
file_name = f"{url}{entity_name}.csv"  
  
df = pd.read_csv(file_name) # leitura arquivo CSV utilizando Dataframe Pandas  
s_df = spark.createDataFrame(df) # converte Dataframe Pandas em Spark Dataframe  
s_df.write.mode("overwrite").saveAsTable(table_name) # grava o DataFrame na Tabela Delta
```

Code Cell 10: Gravando a tabela DELTA - CNAE

```
%python  
entity_name = f"cnae"  
  
table_name = f"{catalog_name}.{schema_name}.{entity_name}"  
file_name = f"{url}{entity_name}.csv"  
  
df = pd.read_csv(file_name) # leitura arquivo CSV utilizando Dataframe Pandas  
s_df = spark.createDataFrame(df) # converte Dataframe Pandas em Spark Dataframe  
s_df.write.mode("overwrite").saveAsTable(table_name) # grava o DataFrame na Tabela Delta
```



Passo 6 – Importar manualmente



Passo 6 – Importar manualmente

Add data

Get started by connecting to a data source or uploading a local file.

Search data sources ▾

Databricks connectors [Provide feedback](#)



Salesforce



Workday Reports



ServiceNow



Google Analytics Raw Data



SQL Server



PostgreSQL



Dynamics 365

Files

Create or modify table

Upload tabular data files to create a new table or replace an existing one



Upload files to a volume

Add files in any format to a non-tabular dataset managed in Unity Catalog



Create table from Amazon S3

Create a table from tabular data files in S3 using a Unity Catalog external location



Passo 6 - Importar manualmente

Searched for "faturamento.csv" in 00vsdb

Add data > Create or modify table from file upload

faturamento.csv uploaded 846.65KB

Preview mode Catalog Schema Table name Advanced attributes

workshop_08_2023 inadimplencia faturamento2

Viewing 50 rows, 22 columns

1 ² ₃	num_cliente	genero_cliente	1 ² ₃	cep_num	cep_formatado	uf	1 ² ₃	cic
1		F		4707040	04707-040	SP		SAO PAL
2		M		4364010	04364-010	SP		SAO PAL
3		M		5725030	05725-030	SP		SAO PAL
4		F		5725030	05725-030	SP		SAO PAL
5		M		5725030	05725-030	SP		SAO PAL
6		M		5725030	05725-030	SP		SAO PAL
7		M		4754030	04754-030	SP		SAO PAL
8		M		4754030	04754-030	SP		SAO PAL
9		M		4754030	04754-030	SP		SAO PAL

Create table



Passo 7 – Verificar o catálogo e as importações

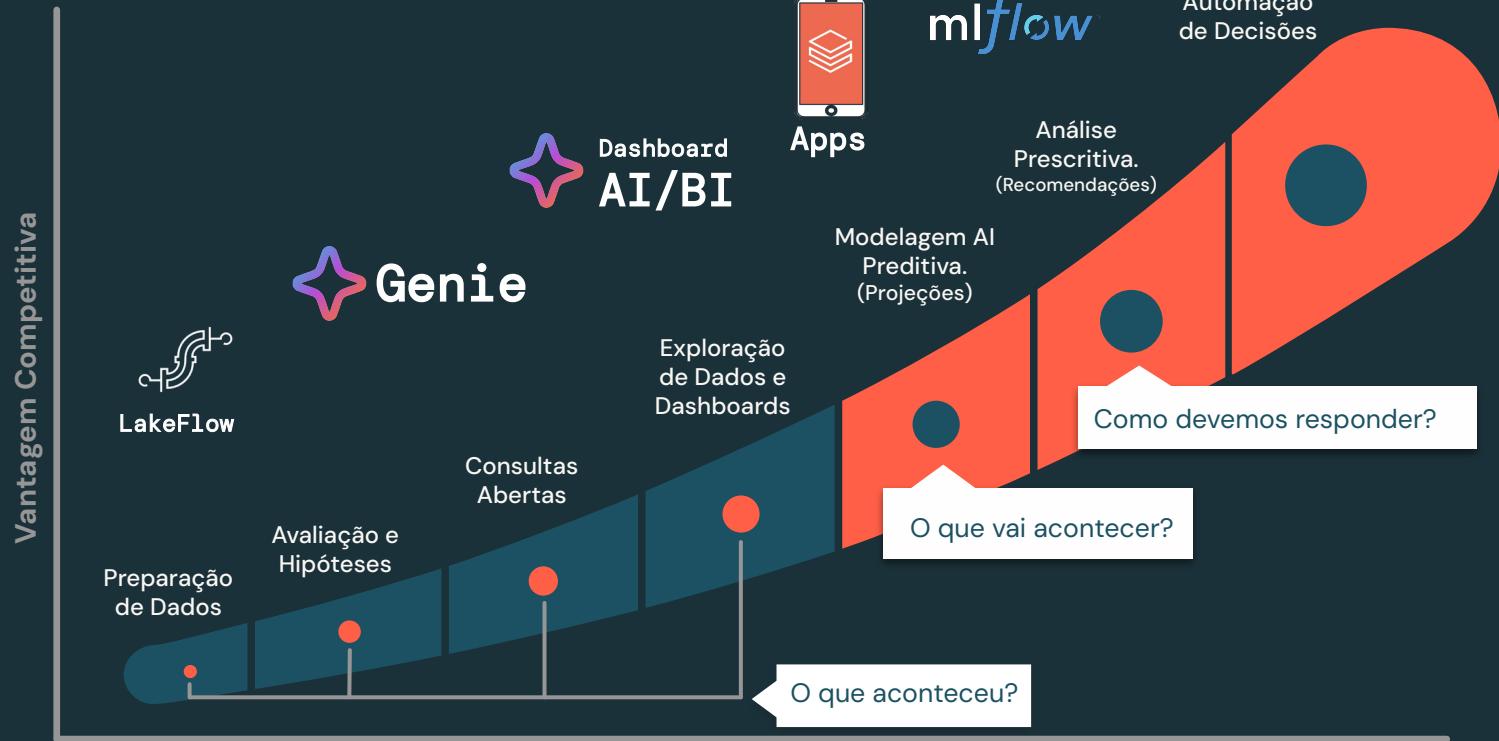
RESULTADO
Esperado
Após a carga

The screenshot shows the Databricks workspace interface. On the left, the sidebar includes options like Workspace, Recents, Catalog, Discover, Jobs & Pipelines, Compute, Marketplace, SQL, SQL Editor, Queries, Dashboards, Genie, Alerts, and Query History. The 'Catalog' section is selected. In the main area, the 'Catalog' tab is active, showing a search bar with 'workshop_'. Below it, there are two tabs: 'For you' and 'All'. Under 'My organization', the 'workshop_08_2025' schema is listed. This schema contains a 'default' database, which in turn contains the 'inadimplencia', 'cnae', 'empresas_sp', 'faturamento', 'ibge_senso', 'municipios', and 'naturezas' tables. A red box highlights this structure. The 'information_schema' database is also visible at the bottom.

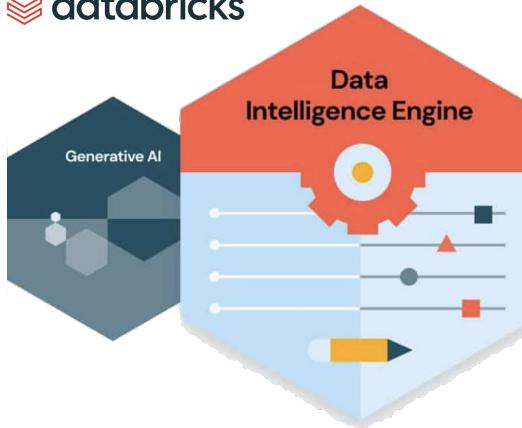


PERGUNTAS ?

Curva de Maturidade de Dados



Consumindo os Dados de forma eficiente



Visualizadores (BI)
tradicionalis de mercado



permite aos usuários de negócios criar painéis e gráficos de forma interativa utilizando assistente.



Ask your question...



permite aos usuários de negócios usar linguagem natural para fazer perguntas de negócio e obter insights valiosos e contextualizados em experiências conversacionais — como um ChatGPT.



Possibilita a criação rápida e simples de aplicações de dados em alguns cliques.



Um Portal para usuários finais das áreas de negócio, com um visual SIMPLES, CLEAN e intuitivo.

Genie



Ask your question...



Aceleradores - Go To Market

Data LAKEHOUSE



Data LAKEHOUSE + Genie Spaces





Hands-On Labs

Democratizando dados com Databricks

LABORATÓRIO 2

Entendimento de Dados com Genie

https://github.com/Databricks-BR/lab_agosto_2025

Passo 1 – Criar o Genie

The screenshot shows the Databricks sidebar menu. At the top is the 'databricks' logo. Below it is a red-highlighted 'New' button. The sidebar contains several sections with icons: Workspace (document), Recents (clock), Catalog (triangle), Discover (grid), Jobs & Pipelines (two people), Compute (cloud), and Marketplace (catalog). Under SQL, there are links for SQL Editor, Queries, Dashboards, Genie (which is highlighted with a red box and has a large white cursor pointing at it), Alerts, Query History, and SQL Warehouses.

- ≡ databricks
- + New
- Workspace
- Recents
- Catalog
- Discover
- Jobs & Pipelines
- Compute
- Marketplace
- SQL
 - SQL Editor
 - Queries
 - Dashboards
 - Genie
 - Alerts
 - Query History
 - SQL Warehouses

Vamos criar um GENIE SPACES
(Agente Conversacional)

Genie

Ask questions about your data in natural language

+ New



All

Favorites

Popular

Last modified ▾

Filter spaces

s reserved



Passo 2 – Conectar as fontes de dados

Seleccione
todos os
dados
carregados

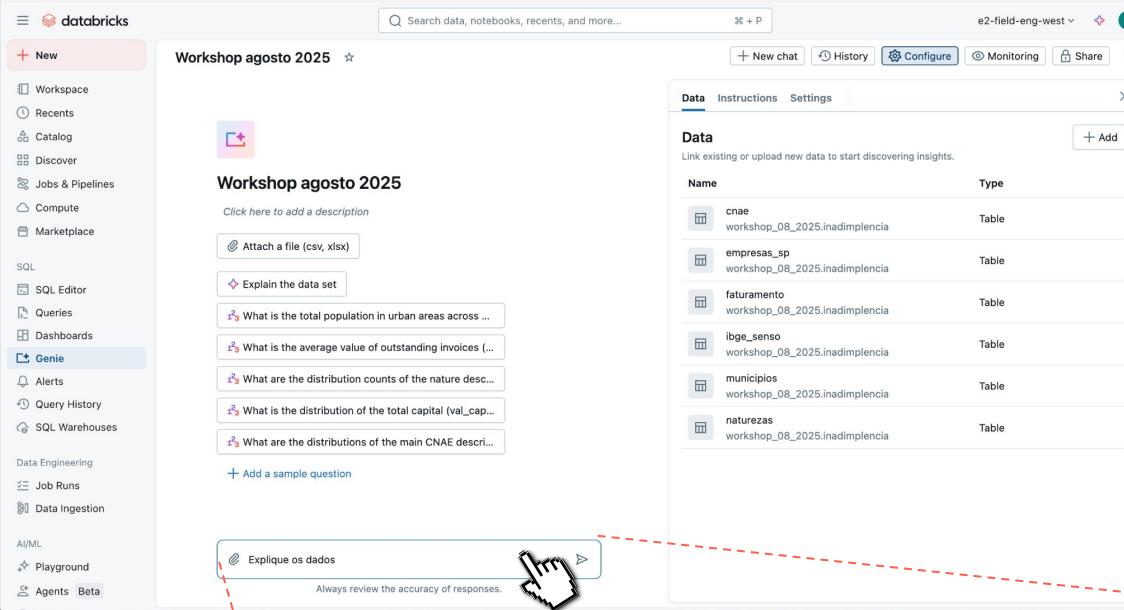
Connect your data

Genie spaces empower you to uncover meaningful insights from your data. Just upload your datasets, provide instructions, and simply ask your data questions.

The screenshot shows a user interface for connecting data. At the top, there is a search bar with the text "workshop" and a magnifying glass icon. To the right of the search bar are a close button (an 'X') and a refresh button (a circular arrow). Below the search bar are two filter buttons: "For you" and "All", with "All" being highlighted with a blue border. The main area displays a breadcrumb navigation path: "All catalogs > workshop_08_2025 > inadimplencia". Below the path is a list of datasets, each with a checkmark and a small icon. A hand cursor is shown pointing at the "naturezas" dataset. The datasets listed are: cnae, empresas_sp, faturamento, ibge_senso, municipios, and naturezas.

- ✓ cnae
- ✓ empresas_sp
- ✓ faturamento
- ✓ ibge_senso
- ✓ municipios
- ✓ naturezas

Passo 3 – Entender os dados



Workshop agosto 2025

Click here to add a description

Attach a file (csv, xlsx)

Explain the data set

What is the total population in urban areas across ...

What is the average value of outstanding invoices (...

What are the distribution counts of the nature desc...

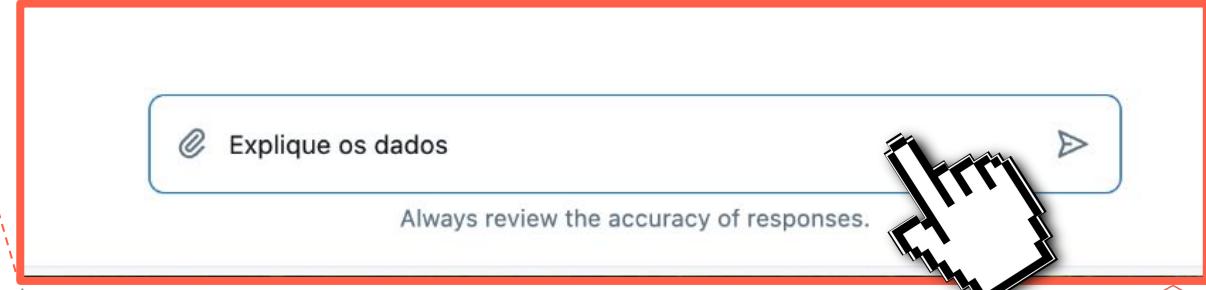
What is the distribution of the total capital (val_cap...

What are the distributions of the main CNAE descri...

+ Add a sample question

Explique os dados

Always review the accuracy of responses.



Passo 3 – Entender os dados

Workshop agosto 2025 ★

Explique os dados

Os dados consistem em seis tabelas relacionadas a empresas, inadimplência, faturamento, e informações demográficas de municípios. Aqui está uma breve descrição de cada tabela:

1. **cnae**: Contém códigos e descrições de atividades econômicas (CNAE).
2. **empresas_sp**: Inclui informações sobre empresas em São Paulo, como CNPJ, nome fantasia, razão social, CNAE principal, bairro, CEP, capital social, latitude e longitude.
3. **faturamento**: Registra dados de faturamento de clientes, incluindo gênero, CEP, cidade, bairro, logradouro, latitude, longitude, quantidade de faturas, inadimplência, valores faturados, pagos e de dívida, e faixa de dívida.
4. **ibge_senso**: Fornece dados do censo do IBGE sobre municípios, como código do município, UF, nome do município, capital, altitude, área, latitude, longitude, e populações total, urbana, rural, masculina e feminina.
5. **municípios**: Contém códigos e nomes de municípios.
6. **naturezas**: Inclui códigos e descrições de naturezas jurídicas.

Essas tabelas podem ser usadas para análises detalhadas sobre empresas, inadimplência, faturamento e características demográficas de municípios.

Is this correct?

Yes Fix it



Passo 4 – Perguntas Abertas

Façam perguntas Abertas (Converse com seus dados)

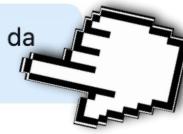
- qual a quantidade de clientes inadimplentes?
- valor da inadimplência por bairros?
- valor total das faturas em aberto?
- quantidade de faturas em aberto?
- quantidade de faturas em aberto por aging da dívida?
- quantidade de faturas em aberto por faixa de atraso?
- quais os 10 bairros com maior quantidade de faturas em aberto com aging da dívida a cima de 60 dias?



TEXT To SQL

Workshop agosto 2025 ☆

quais os 10 bairros com maior quantidade de faturas em aberto com aging da dívida a cima de 60 dias



O usuário quer saber quais são os 10 bairros com a maior quantidade de faturas em aberto que têm um aging da dívida superior a 60 dias. faturamento

1²3 aging_divida: > 60 ▾

10 rows ↓

Add as instruction

Hide code

⋮

```
1 SELECT
2   `faturamento`.`bairro`,
3   SUM(`faturamento`.`qde_faturas_abertas`) AS `total_faturas_abertas`
4 FROM
5   `workshop_08_2025`.`inadimplencia`.`faturamento`
6 WHERE
7   `faturamento`.`bairro` IS NOT NULL
8   AND `faturamento`.`qde_faturas_abertas` IS NOT NULL
9   AND `faturamento`.`aging_divida` IS NOT NULL
10  AND `faturamento`.`aging_divida` > 60
11 GROUP BY
12   `faturamento`.`bairro`
13 ORDER BY
14   `total_faturas_abertas` DESC
15 LIMIT 10
```

Passo 5 – Avaliar o Histórico de Perguntas

Workshop agosto 2025 ★ > Monitoring

History of the activities within your space. Understand the errors, feedback, and results the space is returning.

Last 30 days Rating Request User Status

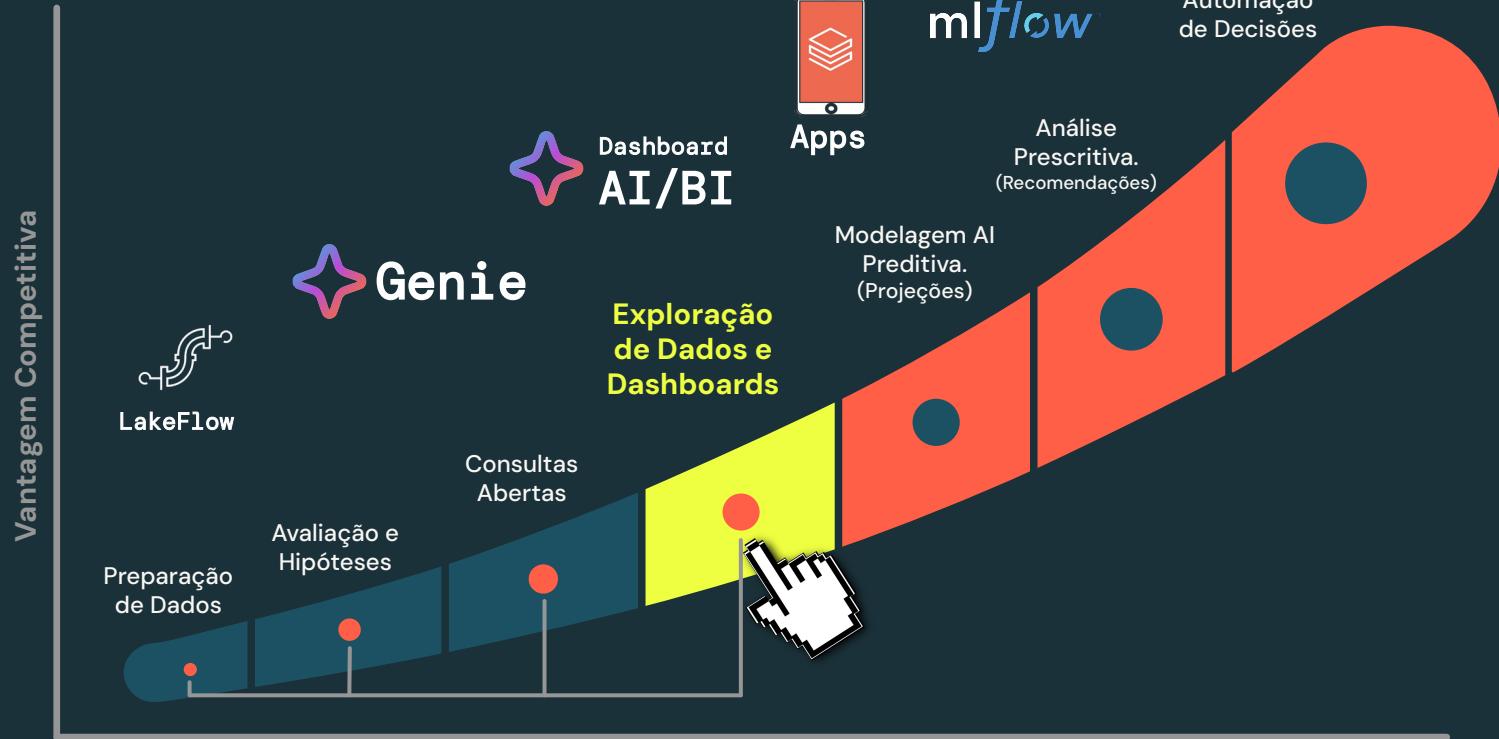
Question	Rating	Request	Comment	User	Created	Last upd...
✓ quais os 10 bairros com maior quantidade de faturas em aberto com aging ...	Good			👤 Luis ...	Aug 19, 2...	Aug 19, 2...
✓ quantidade de faturas em aberto por faixa de atraso	Good			👤 Luis ...	Aug 18, 2...	Aug 18, 2...
✓ quantidade de faturas em aberto por aging da dívida				👤 Luis ...	Aug 18, 2...	Aug 18, 2...
✓ quantidade de faturas em aberto	Good			👤 Luis ...	Aug 18, 2...	Aug 18, 2...
✓ valor total das faturas em aberto?				👤 Luis ...	Aug 18, 2...	Aug 18, 2...
✓ valor da inadimplencia por bairros	Good			👤 Luis ...	Aug 18, 2...	Aug 18, 2...
✓ qual a quantidade de clientes inadimplentes	Good			👤 Luis ...	Aug 18, 2...	Aug 18, 2...
✓ Explique os dados				👤 Luis ...	Aug 18, 2...	Aug 18, 2...

The screenshot shows a Databricks workspace interface. At the top, there is a navigation bar with several buttons: '+ New chat', 'History' (which is highlighted with a blue background and a red arrow), 'Configure', 'Monitoring', 'Share', and a three-dot menu. Below the navigation bar, the main area has tabs for 'Data', 'Instructions', and 'Settings'. The 'Data' tab is active. On the left, there's a sidebar titled 'Data' with a sub-section 'Data' containing a table with six rows. The table columns are 'Name' and 'Type'. The rows are: 'cnae' (Table), 'empresas_sp' (Table), 'faturamento' (Table), 'ibge_senso' (Table), 'municipios' (Table), and 'naturezas' (Table). The main workspace area shows a table of questions and their responses, with the 'History' tab selected in the sidebar.

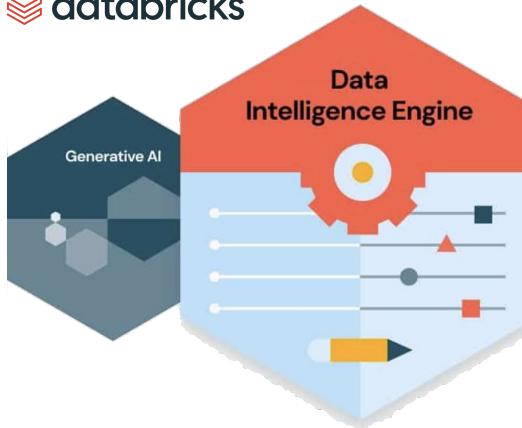


PERGUNTAS ?

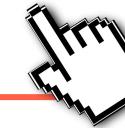
Curva de Maturidade de Dados



Consumindo os Dados de forma eficiente



Visualizadores (BI)
tradicionalis de mercado



permite aos usuários de negócios criar painéis e gráficos de forma interativa utilizando assistente.



Ask your question... ➤



Databricks
Lakehouse APPS

permite aos usuários de negócios usar linguagem natural para fazer perguntas de negócio e obter insights valiosos e contextualizados em experiências conversacionais — como um ChatGPT.



Databricks
ONE

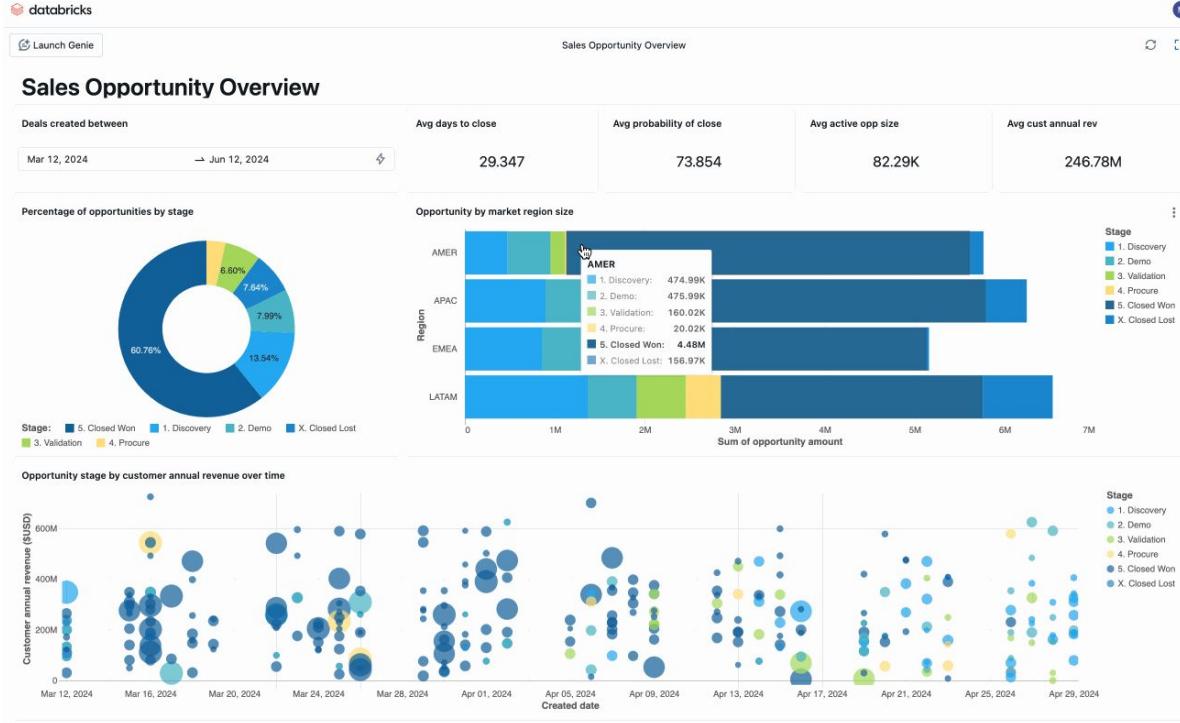
Possibilita a criação rápida e simples de aplicações de dados em alguns cliques.

Um Portal para usuários finais das áreas de negócio, com um visual SIMPLES, CLEAN e intuitivo.

Databricks AI/BI

Intelligent analytics for real-world data





Os dashboards do AI/BI facilitam a criação e a iteração de visualizações com **linguagem natural** por meio da criação assistida por IA.

Os dashboards oferecem recursos padrão de **visualização de dados**, incluindo gráficos elegantes, interações como filtragem cruzada, snapshots periódicos por e-mail, incorporação e muito mais.

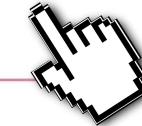
E eles convivem lado a lado com seus dados, oferecendo carregamento instantâneo e análise interativa rápida, independentemente da escala de dados ou de usuários.





databricks
Dashboard
AI/BI

crie um gráfico de



faturamento

ano

mes

regiao

faturamento

pis

cofins

icms

notas_entradas

notas_saida

to	pis	cofins			notas_saida
5578	16546	76367	458204	14301	1000
6574	3683	16997	124646	3183	354
1128	4947	22834	137003	4276	504
5220	1724	7957	45087	1490	469
0384	16578	76512	459069	14328	890
6666	2626	21750	200045	5016	200

Submit



crie um gráfico de barras do faturamento mensal



to	pis	cofins	icms	notas_entradas	notas_saida
5578	16546	76367	458204	14301	1000
6574	3683	16997	124646	3183	354
1128	4947	22834	137003	4276	504
5220	1724	7957	45087	1490	469
0384	16578	76512	459069	14328	890
6666	6666	61750	666645	5616	666



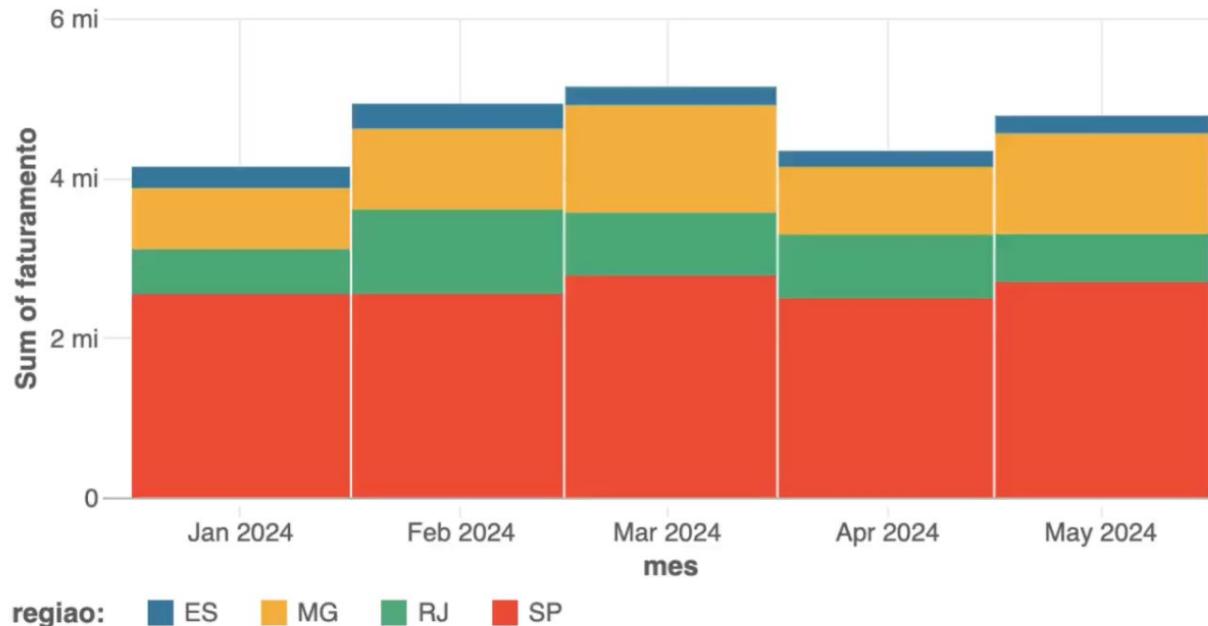
crie um gráfico de barras do faturamento mensal

Accept

Reject



Monthly Revenue



to

pis

cofins

icms

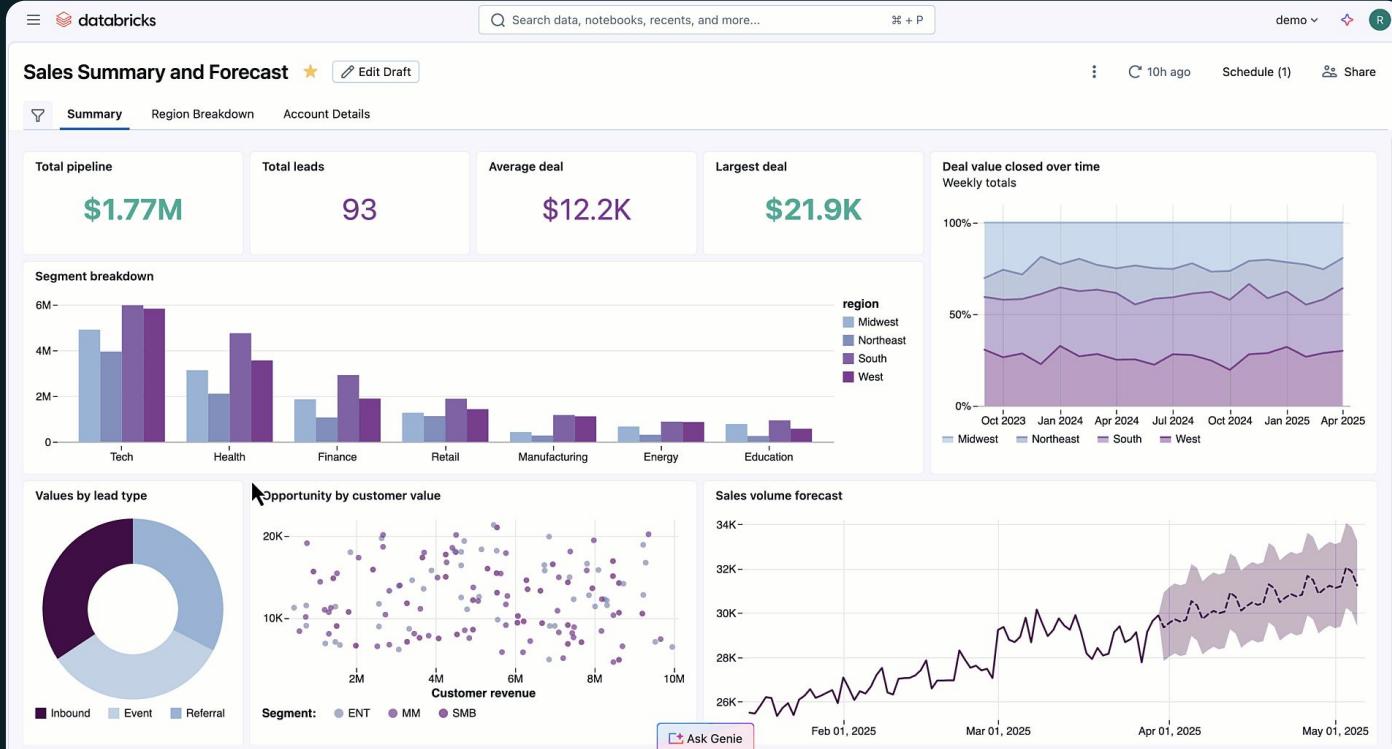
notas_entradas

notas_saida

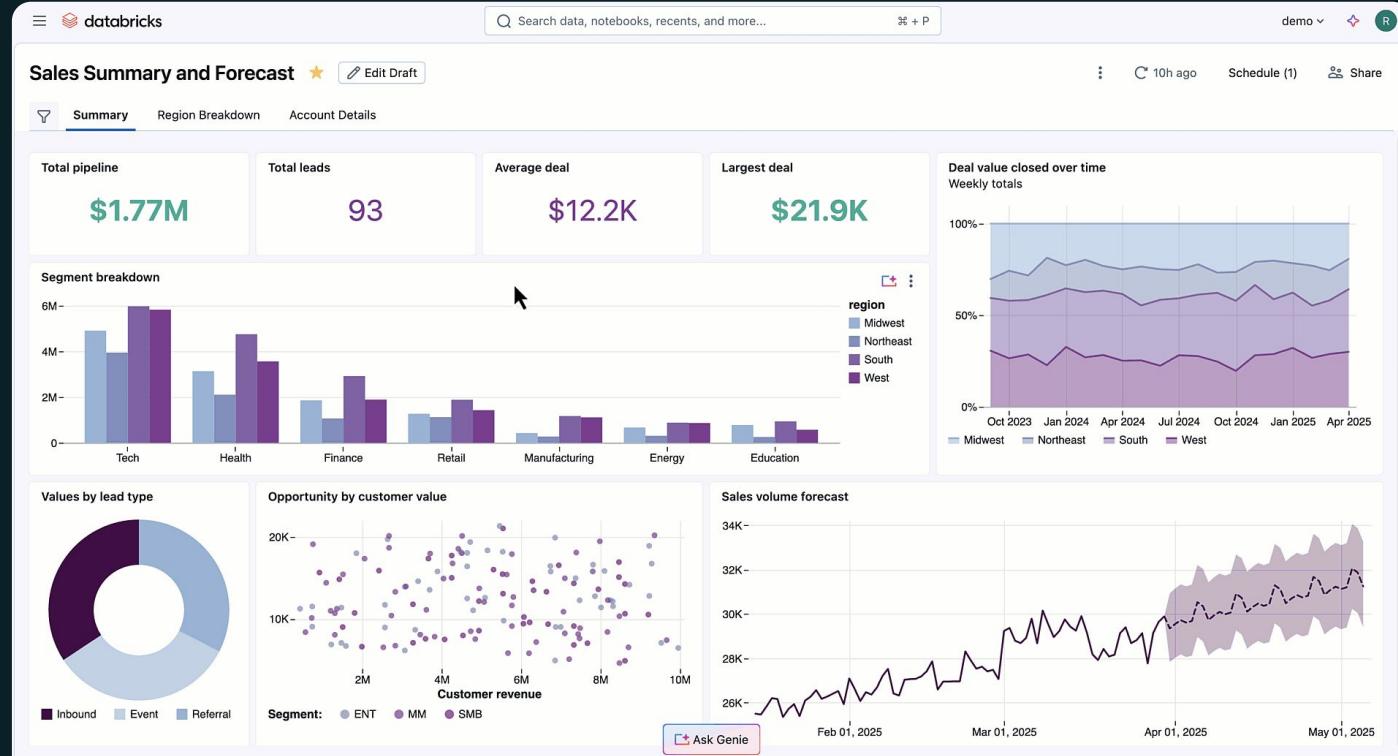


⋮

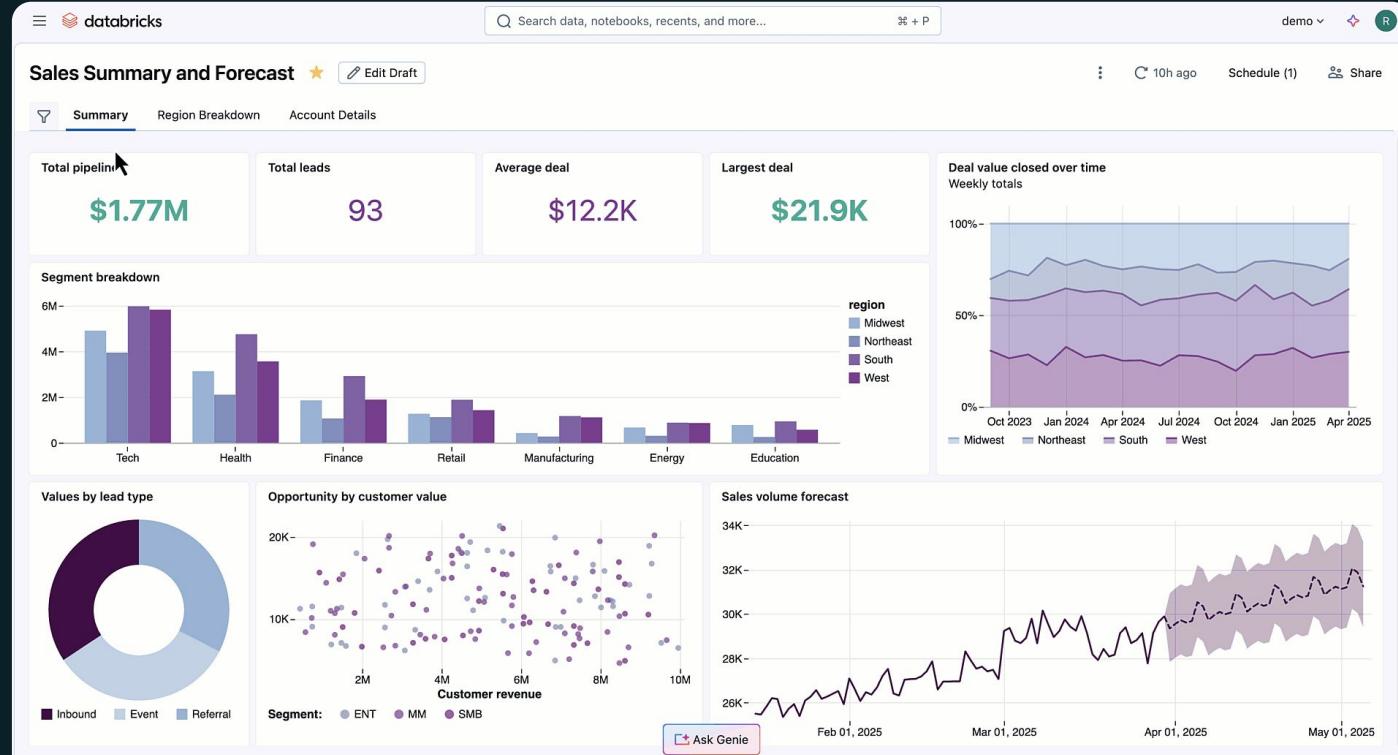
Cross-filtering



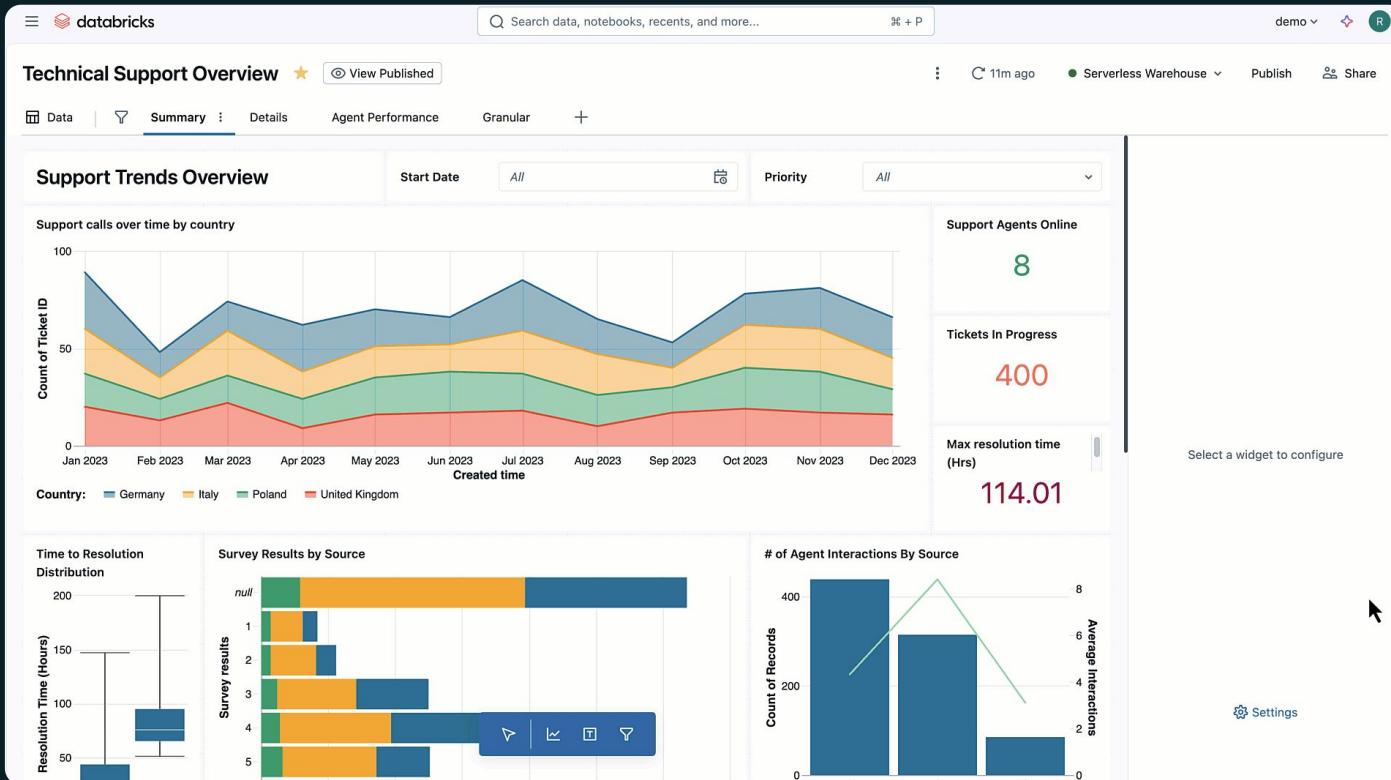
Drilling



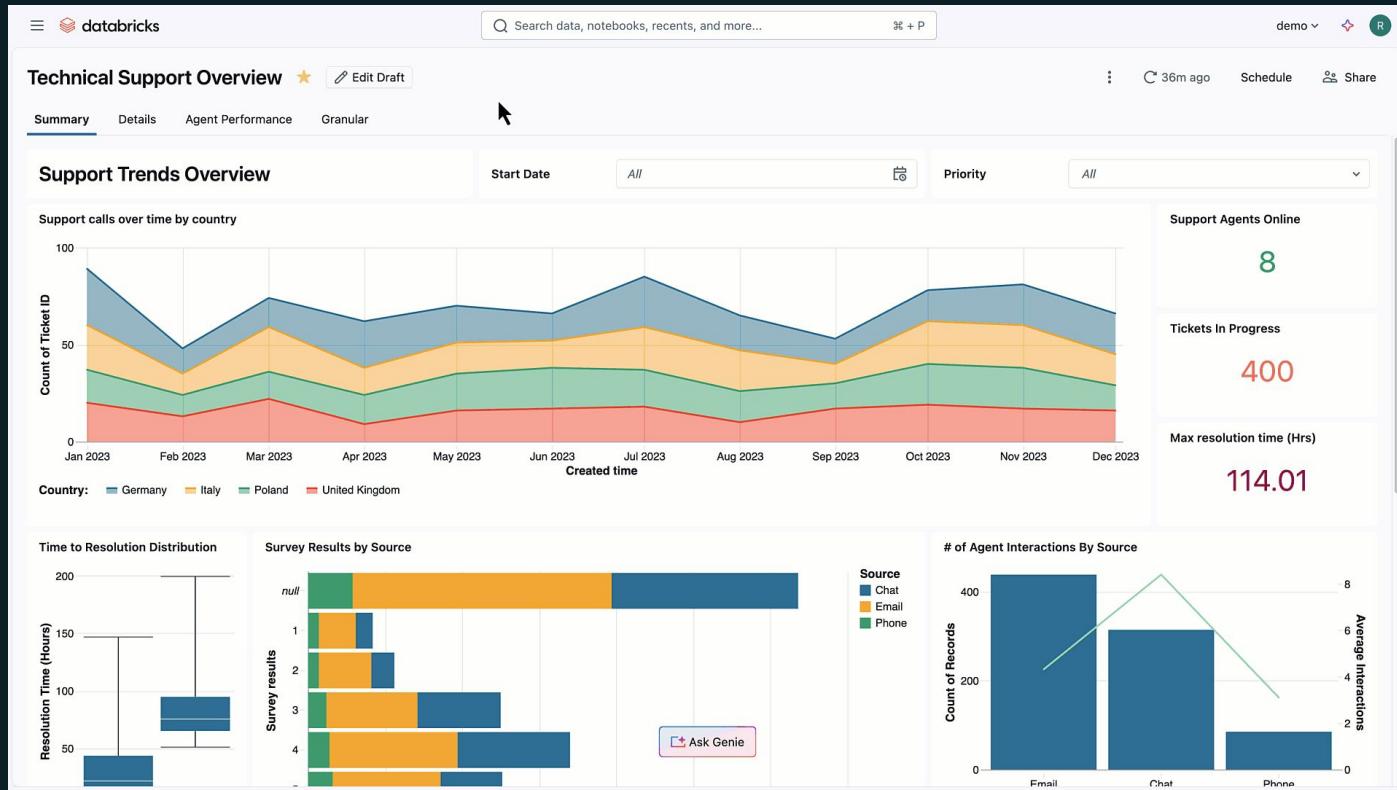
Multi-page reporting



Customização de Temas



Filtros Globais



Upload de Arquivos no Dash

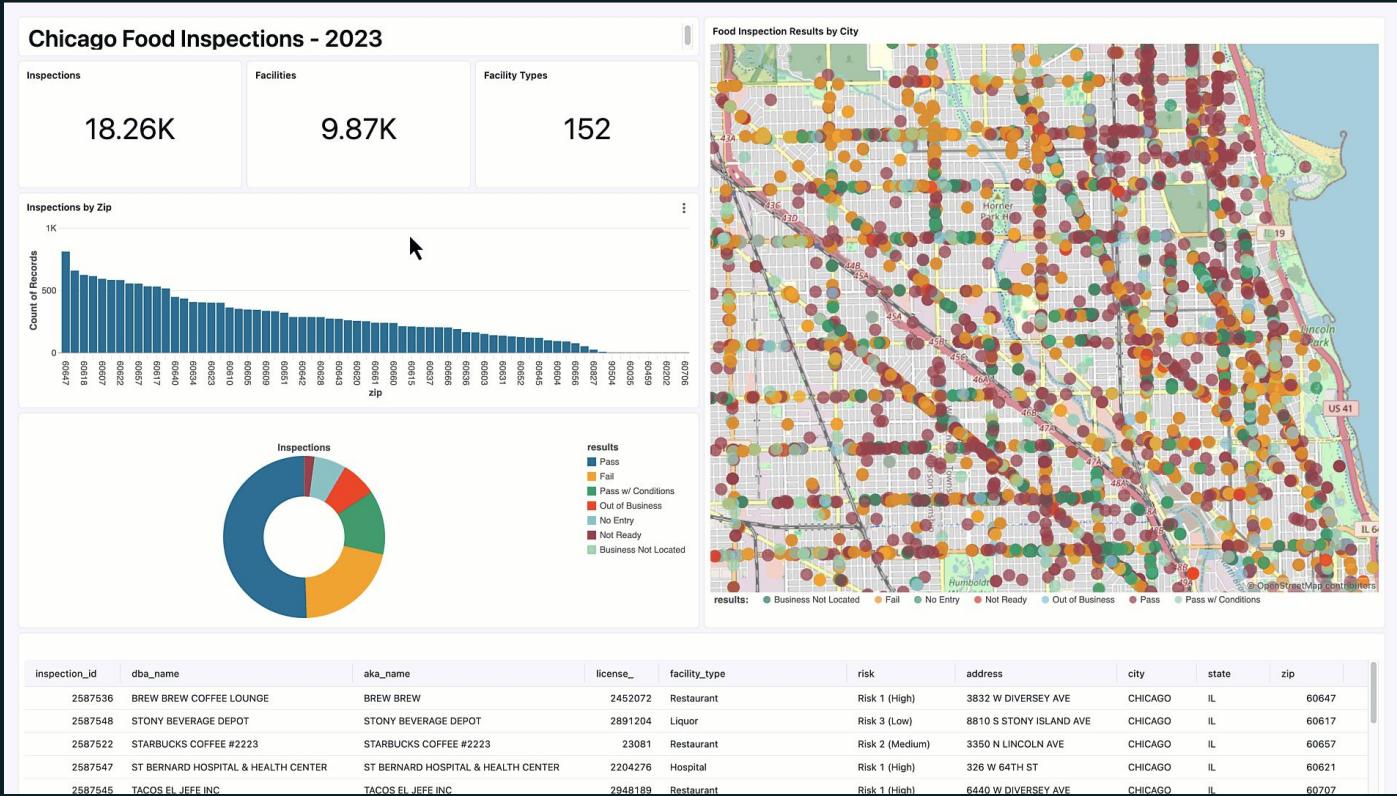
The screenshot displays the Databricks dashboard with several key features highlighted:

- Alerts are now in Git!** A section featuring a search bar and a list of alerts:
 - CSAT Star of the Week Query
 - Overdue Accounts Alert
 - Technical Support Overview Dashboard
- Git support for alerts**: A detailed description of how users can track alerts using Git.
- Welcome to Databricks**: The main header with a search bar and navigation links for Recents, Favorites, Popular, Mosaic AI, and What's new.
- Improved notebook find and replace**: A note about the enhanced search and replace functionality.
- File Editor**: A screenshot of a notebook editor showing a SQL query:

```
File Edit View Run Help SQL Tabs: ON
Find: d23b1c1e-4f2a-4a8d-9bcb-0a211e Aa ab 1 of 2 Press ⌘+F again to use browser find
SELECT
FROM
customers
WHERE
customer_id = 'd23b1c1e-4f2a-4a8d-9bcb-0a2f15413cf3'
```

- Lakeflow Jobs UI**: A screenshot of the Lakeflow Jobs UI interface.
- Lakeflow Jobs UI is now available in Public Preview**: A section describing the new UI features.

Geo Analysis



Forecasting

Sales Summary and Forecast ★ [View Published](#)

Data | **Summary** | Region Breakdown | Account Details +

Total revenue over time

Account Distribution by Industry

Sales Volume

Widget

Title Description

Dataset [Hide filters](#)

EMEA breakdown

Filter fields +

Visualization

Line

X axis Date

Y axis 1.2 SUM(Price)

Color

Tooltip

Labels

Annotation

Search data, notebooks, recents, and more... ⌂ + P

demo ▾ R

5h ago • Serverless Warehouse ▾ Publish Share

Performance by team

Team member	Values	1. Discovery	2. Demo	3. Validation	4. Procure

Dashboard embedding

Microsoft 365

Search this site

All Company

Public group 1 member

Promote Page details Preview Immersive reader Analytics Published 6/3/2025 Share Edit

Sales Home

Eason Gao

Overview

Welcome to the central resource for our Sales Team! Here you'll find everything you need to excel in your role, stay informed, and collaborate with your colleagues. Whether you're a new hire or a seasoned team member, this page is your starting point for all things sales.

databricks

15m ago

Total pipeline	Total leads	Average deal	Largest deal
\$1.77M	93	\$12.2K	\$21.9K

Deal value closed over time
Weekly totals

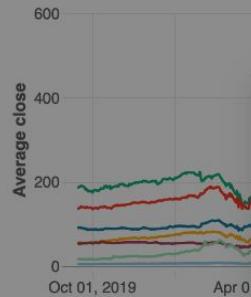
100%
50%
0%



Stock Market Dashboard

All

Gráfico de linhas do valor de fechamento por dia e por empresa



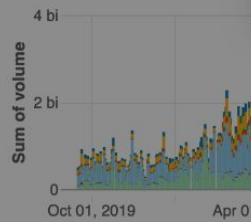
Copy embed code

You can copy the code below and paste it into your website. Note that users will still need to log in, and that your dashboard needs to be published in order for them to view it.

```
<iframe src="https://e2-demo-field-eng.cloud.databricks.com/embed/dashboardsv3/01ef756abde9176387c7a1e6e7a8245d?o=1444828305810485" width="100%" height="600" frameborder="0"></iframe>
```



Volume de ações por dia e por



Published dashboards are **allowed** to be embedded in any domain.





Hands-On Labs

Democratizando dados com Databricks

LABORATÓRIO 3

Análise Exploratória – Dashboards

https://github.com/Databricks-BR/lab_agosto_2025



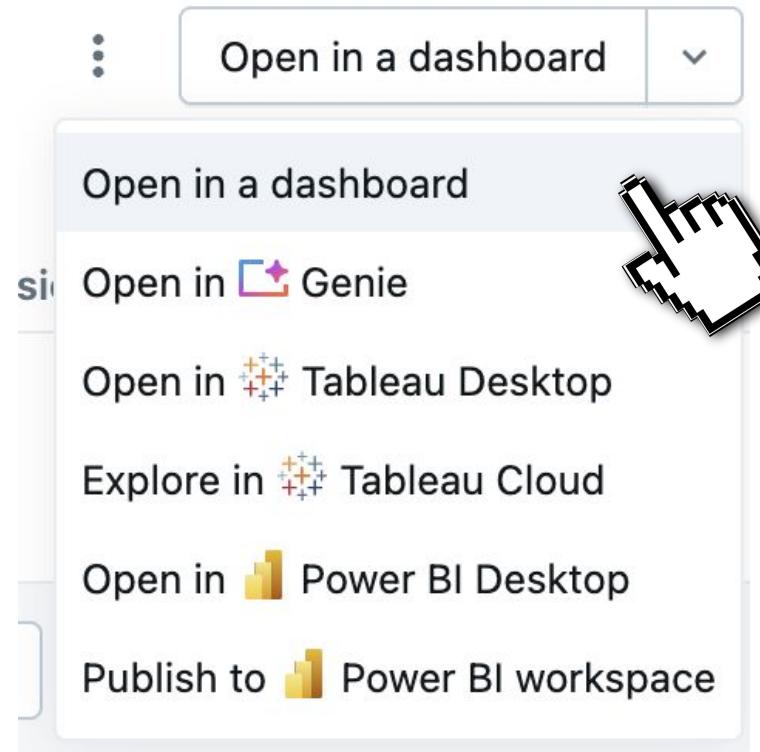
Passo 1 – Entre no menu CATALOG ...

The screenshot shows the Databricks interface with the following elements:

- Left Sidebar:** Contains links for New, Workspace, Recents, Catalog (highlighted with a red box and a cursor), Discover, Jobs & Pipelines, Compute, Marketplace, SQL, SQL Editor, Queries, Dashboards, and Genie.
- Top Bar:** Includes a search bar ("Search data, notebooks, recents, and more..."), a "P" icon, and account information ("e2-field-eng-west").
- Catalog Explorer:** Shows the "workshop_08_2025" catalog under "My organization".
 - Catalog Overview:** Displays the catalog name, a thumbnail, and a star icon. A red box highlights the "Open in a dashboard" button in the top right corner, and a cursor is clicking it.
 - Table View:** Shows the schema "empresas_sp" with columns: cnpj_basico (bigint), cnpj_txt (string), and nome_fantasia_empresa (string).



Passo 2 – Clique na Opção "Open in Dashborads" ...



Passo 3 – Crie um Indicador de "quantidades"

The screenshot illustrates the process of creating a quantity indicator in a Databricks workspace. At the top, a search bar contains the text "quantidade de empresas". Below it, two options are visible: "Company Distribution by Neighborhood" and "Capital Social Counter". A large, stylized hand cursor is positioned over the "Submit" button, which is highlighted with a blue outline. A callout bubble points from the cursor to the "Submit" button. In the bottom right corner of the workspace, a modal window is open, displaying the result "Quantidade de Empresas" followed by the value "111,53 mil". To the right of the value are three buttons: "Accept", "Reject", and a circular refresh icon.



Passo 4 – Crie um Filtro de Página

empresas_sp 2025-08-19 00:59:37 ☆ ⋮ 2m ago dbd... 

Data New Page +

Dashboard for empresas_sp table

The dashboard you created for workshop_08_2025.inadimplencia.empresas_sp is ready!

The widget to the right of this one allows you to generate a chart based on this dataset using natural language. Click the widget to start typing a description of the chart you'd like to

Widget title Select fields or parameters to filter.

cnpj_basico cnpj_txt nome fantasia empresa nome_razao
33283682 33283682/0001-46 Add a filter (field/parameter)

©2025 Databricks Inc. — All rights reserved



Widget

- Title Description

Filter

> Text entry

Multiple values

Single value

Date picker

Date range picker

> Text entry

Range slider

Case sensitive

Passo 4 – Crie um Filtro de Página

empresas_sp 2025-08-19 00:59:37 ☆ ⋮ 2m ago dbde

Data New Page +

Dashboard for empresas_sp table

The dashboard you created for workshop_08_2025.inadimplencia.empresas_sp is ready!

The widget to the right of this one allows you to generate a chart based on this dataset using natural language. Click the widget to start typing a description of the chart you'd like to

Widget title Select fields or parameters to filter.

cnpj_basico	cnpj_txt	nome_fantasia_empresa	nome_razao_social
33283682	33283682/0001-46	EBR ELETRONICOS	EBR ELETRO

Quantidade de Empresas
111,53 mil

©2025 Databricks Inc. — All rights reserved

Widget Title Description

Filter > Text entry

Fields

Parameters Search fields

Matched 

Condition	Value
<input type="checkbox"/> C	bairro
<input type="checkbox"/> C	cnae_descricao
<input type="checkbox"/> C	cnpj_txt
<input type="checkbox"/> C	nome_fantasia_empresa
<input type="checkbox"/> C	nome_razao_social

Passo 5 – Faça uma pesquisa TEXTUAL

empresas_sp 2025-08-19 00:59:37 ★

⋮ ⏪ 40s ago ● dbde

Data | New Page : +

Dashboard for empresas_sp table

The dashboard you created for
workshop_08_2025.inadimplencia.empresas_sp is ready!

The widget to the right of this one allows you to generate a
chart based on this dataset using natural language. Click the
widget to start typing a description of the chart you'd like to

Quantidade de Empresas

634

cnae_descricao

farma



Passo 6 – Crie um MAPA (Point Map)

Data | New Page :: +

Dashboard for empresas_sp

Quantidade de Empresas
111,53 mil

cnae_descricao

nome_razao_social

cnae_principal

cnpj_txt

crie um mapa com a localização das empresas ➔

Add a visualization

Filter icon



©2025 Databricks Inc. — All rights reserved



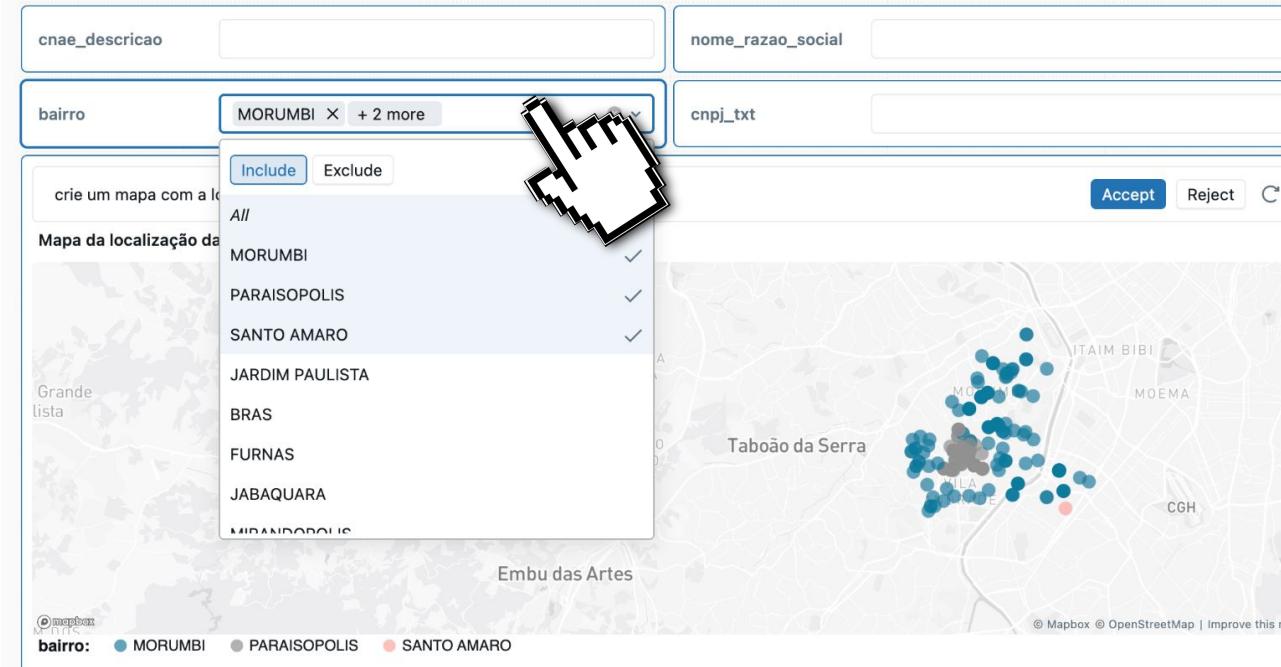
Passo 7 – Faça investigações usando os filtros

empresas_sp 2025-08-19 00:59:37 ☆

2m ago · dbdemos-shared-en... · Publish · Share

Data | New Page | +

Dashboard for empresas_sp



Widget

Title Description

Filter

Multiple values

Fields

empresas_sp.bairro

Parameters

Default Value

All

Passo 8 – Explore as configurações do Mapa

Catalog

Discover

Jobs & Pipelines

Compute

Marketplace

SQL

SQL Editor

Queries

Dashboards

Genie

Alerts

Query History

SQL Warehouses

Data Engineering

Job Runs

Data Ingestion

AI/ML

Playground

Agents Beta

Experiments

Features

Models

Serving

Quantidade de Empresas

Dashboard for empresas_sp

111,53 mil

cnae_descricao

nome_razao_social

bairro

All

cnpj_txt

Mapa da localização das empresas

The dashboard displays a map titled "Mapa da localização das empresas" showing the distribution of businesses across the São Paulo metropolitan area. The map is filled with numerous colored dots of varying sizes, representing individual business locations. The colors appear to correspond to different industry codes (CNAE). The map includes labels for several cities and towns such as Guarulhos, Itaquaquecetuba, Mogi das Cruzes, Suzano, Vila Barros, and many others. The overall density of dots is highest in the central urban core and decreases towards the outskirts.

Passo 9 – Outro Dashboard – Inadimplência

The screenshot shows the Databricks Catalog Explorer interface. On the left, the sidebar has 'Catalog' selected, highlighted with a red box and a mouse cursor. The main area shows a catalog tree under 'workshop_08'. A table named 'faturamento' is selected, also highlighted with a red box and a mouse cursor. To the right, the table details page is displayed for 'faturamento' in the 'inadimplencia' schema. A context menu is open on the table name, with the 'Open in a dashboard' option highlighted with a blue box and a mouse cursor. Other options in the menu include 'Share', 'Create', 'Open in Genie', 'Open in Tableau Desktop', 'Explore in Tableau Cloud', 'Open in Power BI Desktop', and 'Publish to Power BI workspace'. The table schema is listed below:

Column	Type	Comment	Tags	Column masking rule
num_cliente	bigint			
genero_cliente	string			
cep_num	bigint			
cep_formatado	string			

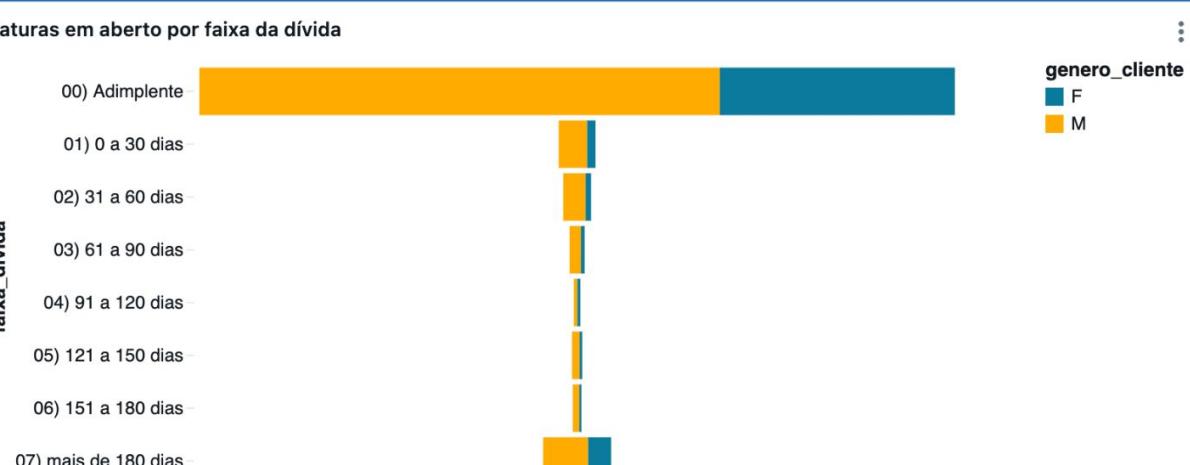


Passo 10 – Configuração manual de componente

Faturamento e Inadimplência

Quantidade d... 5629

Faturas em aberto por faixa da dívida



Count of Records

faixa_divida	Count of Records
00) Adimplente	5629
01) 0 a 30 dias	0
02) 31 a 60 dias	0
03) 61 a 90 dias	0
04) 91 a 120 dias	0
05) 121 a 150 dias	0
06) 151 a 180 dias	0
07) mais de 180 dias	0

Widget

Title Description

Dataset faturamento Show filters

Visualization Funnel

Stage faixa_divida

Value COUNT(*)

Color genero_cliente

F

M

Passo 11 – Mapa da Inadimplência

Passo 12 – Configurando a imagem do cabeçalho

The screenshot shows the Databricks interface with a sidebar on the left containing various navigation options like Workspace, Recents, Catalog, Discover, Jobs & Pipelines, Compute, Marketplace, SQL, and Dashboards. The main area displays a map titled "HandsOn LAB 20 de agosto" with a subtitle "Democratizando dados com Databricks". The map shows several blue bubbles of varying sizes scattered across a city map, representing data points. To the right of the map is a configuration panel with sections for Visualization, Coordinates, Color, Size, and Tooltip. A red box highlights the "Color" section, which is set to "qde_faturas_abertas". A large callout box with a hand cursor icon is overlaid on the top right, pointing to the "Color" section. The URL for the image being pointed to is: https://raw.githubusercontent.com/Databricks-BR/lab_agosto_2025/refs/heads/main/images/head_lab.png.

Passo 13 – Criando Filtros GLOBAIS

HandsOn LAB 20 de agosto ★ View Published 34h ago dbdemos-shared-en... Publish Share

A hand cursor is hovering over the "Global filters" section.

Global filters

- aging_divida: All
- bairro: All
- faixa_divida: All
- genero_cliente: All
- latitude: Min → Max
- longitude: Min → Max

Widget Title Description

Filter Range slider

Fields 1.2 faturamento.longitude...

Parameters

Default Value Min → Max

Mapa com qde_faturas_abertas

The map interface includes navigation controls at the bottom: a magnifying glass icon, a north arrow, and arrows for zooming in and out.

Passo 14 – Publicando o Dashboard

HandsOn LAB 20 de agosto  

⋮ ⏴ 5m ago • dbdemos-shared-en... ⋮ Publish

Data |  **Mapa** ⋮ +

Publish HandsOn LAB 20 de agosto 

Embed credentials (default)
 Shared data permissions
All viewers of the published dashboard use your data permissions to run queries. This enables a shared cache which can improve performance.

Don't embed credentials
 Individual data permissions
Each viewer of a published dashboard uses their own data permission to run queries. This can lead to more frequent refresh operations.

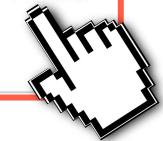
People with access
Users with at least "Can View" permissions can see the published dashboard. Click [here](#) to change user permissions.
luis.assuncao@databricks.com, Admins, All workspace users, kholoud.ibrahim@databricks.com

Notify viewers(emails individuals and small groups)

Enable Genie Beta 
Genie is a no-code tool that helps answer questions about your dashboard. The published dashboard's context is used to build this Genie space.

Auto-generate Genie space
Genie in dashboards is an experimental feature. Users may see a higher than expected number of errors and inaccuracies. You can [open the draft Genie Space here](#).

Link existing Genie space

©2025 Databricks Inc 

Passo 15 – Testando INSIGHTS com Ask Genie

HandsOn LAB 20 de agosto ★ [Edit Draft](#)

Mapa

databricks
Hands-On Labs
Democratizando dados com Databricks

Faturas 5,63 mil

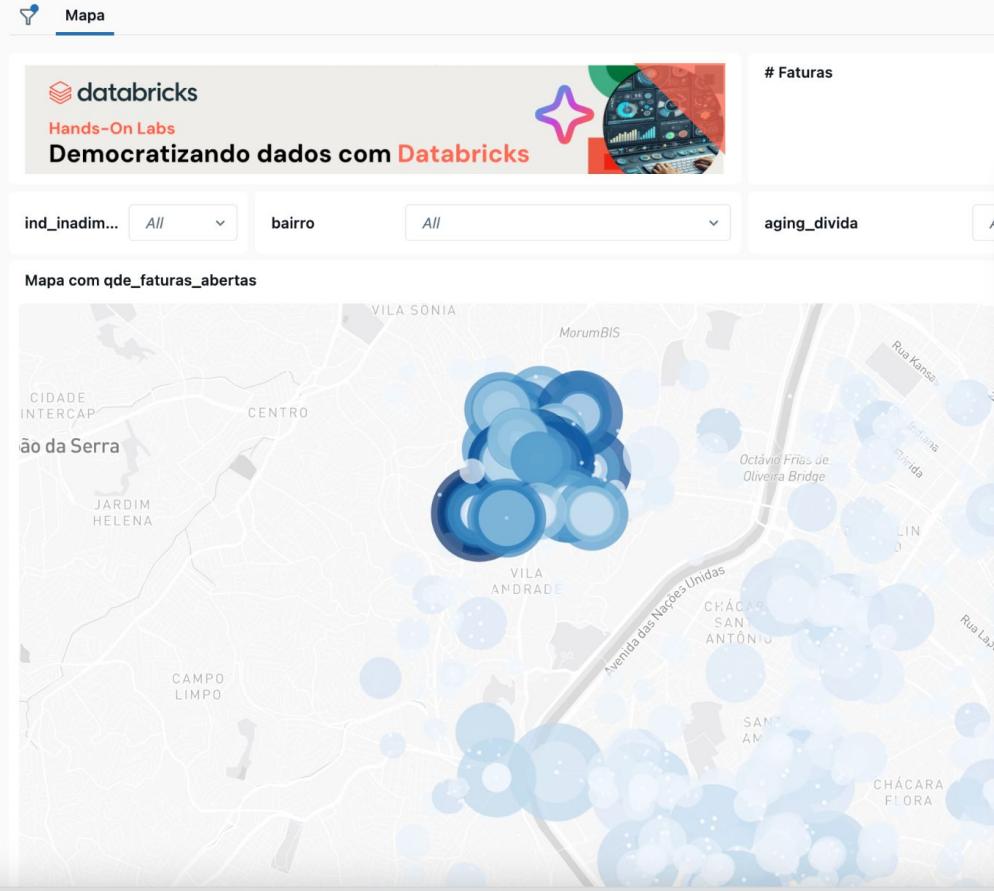
ind_inadim... All bairro All aging_divida All

Mapa com qde_faturas_abertas

VILA SÔNIA
Morumbi
CIDADE INTERCAP
CENTRO
Vila da Serra
JARDIM HELENA
VILA ANDRADE
CHÁCARAS SAN ANTONIO
SANTO ANDRÉ
VILA SANTA CATARINA
CAMPOLIMPO
Cidade VARGAS
Lixo
Ask Genie

Passo 15 – Testando INSIGHTS

HandsOn LAB 20 de agosto ★ [Edit Draft](#)



Genie Beta

Ask Genie about the data in this dashboard

porque essa concentração no bairro de Paraisópolis

O usuário quer entender a situação financeira no bairro de Paraisópolis, especificamente o número de faturas em aberto e o valor da dívida associado a esse bairro.

936 rows [Show code](#)

qde_faturas_abertas	val_divida
8	20
9	0
10	0
11	9
12	0
13	0
14	0
15	0
16	11
17	0

↳ Qual é o valor médio das faturas em aberto no bairro de →

Ask your question... [Ask](#)

Always review the accuracy of responses.

Passo 15 – Testando INSIGHTS

HandsOn LAB 20 de agosto ★ [Edit Draft](#)

Mapa

databricks
Hands-On Labs
Democratizando dados com Databricks

Faturas

ind_inadim... All bairro All aging_divida

Mapa com qde_faturas_abertas

VILA SÔNIA
Centro
Morumbi
Cidade InterCap
Jardim Helena
Vila Andrade
Centro
São da Serra
VILA ANDRADE
Avenida das Nações Unidas
CHÁCARA SÃO ANTONÍO
SANT'AM
CHÁCARA FLORA
CAMPO LIMPO
RUA KANSAI
IRACEMA
FLÓRIDAS
AVENIDA DAS NAÇÕES UNIDAS
RUA LARANJEIRAS
VILA ANDRADE
CENTRO
SÃO PAULO

Genie Beta

18 | 0 | 0

Was this correct?

Qual é a quantidade total de faturas em aberto no bairro de Paraisópolis?

O usuário quer saber o total de faturas em aberto no bairro de Paraisópolis.

1 row Show code

	1 ² 3 total_faturas_abertas
1	6224

Was this correct?

↳ Qual é a quantidade total de faturas pagas no bairro de

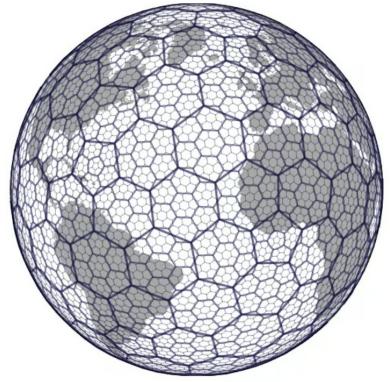
Ask your question...

Always review the accuracy of responses.

PERGUNTAS ?

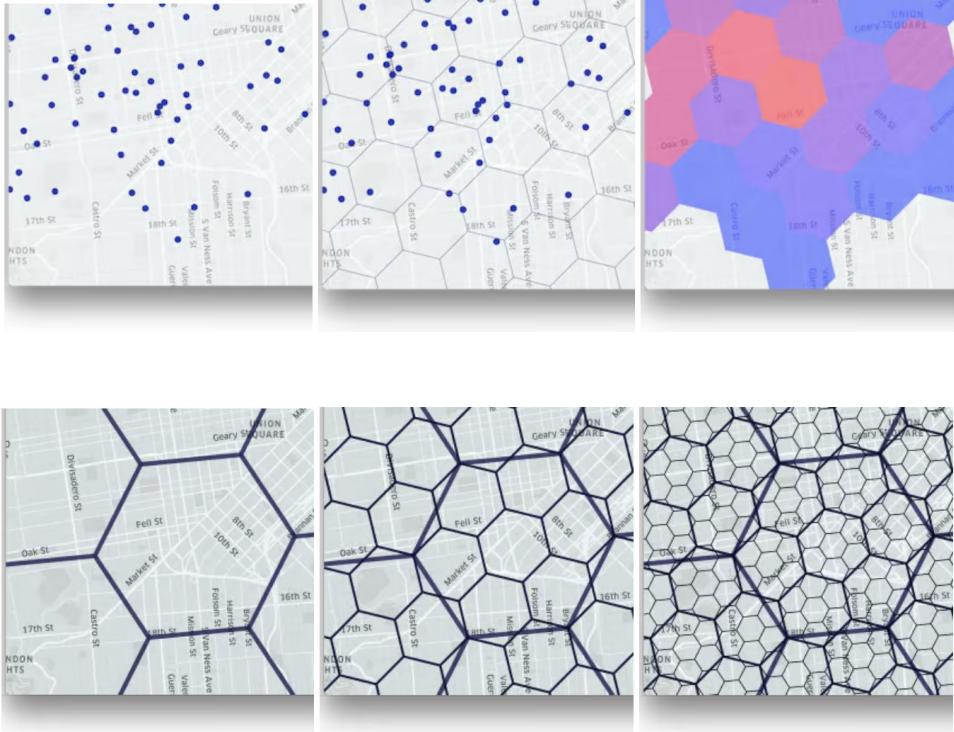
Manipulando Dados. SQL Editor. Funções de GEO (H3).



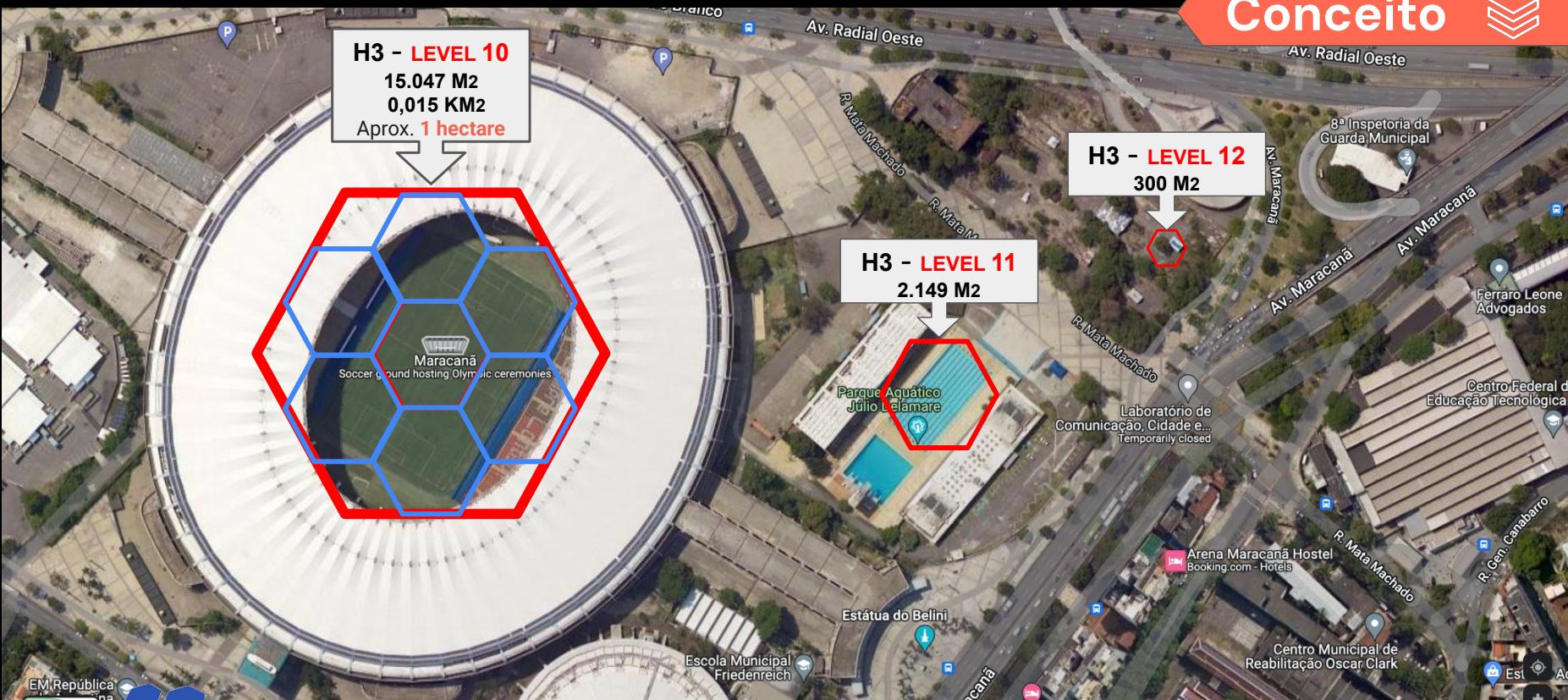


H3 é um sistema de indexação espacial hexagonal hierárquico desenvolvido pela Uber.

Ele divide o globo em células hexagonais de diferentes tamanhos, permitindo a indexação de dados geoespaciais de forma eficiente.



Conceito



H3

Databricks + H3

Funcionalidades Nativas

30+ funções nativas no Databricks
SQL, Python, Scala

The screenshot shows a Databricks notebook interface. In the top-left cell, the user has typed:

```
1 SELECT FROM mp_routes
```

A tooltip is displayed over the word "FROM", listing the columns and their types for the table "mp_routes". The tooltip content is as follows:

Location	string
URL	string
Avg Stars	double
Route Type	string
Rating	string
Pitches	bigint
Length	double
Area Latitude	double
Area Longitude	double
desc	string
protection	string

In the bottom-left cell, the user has typed:

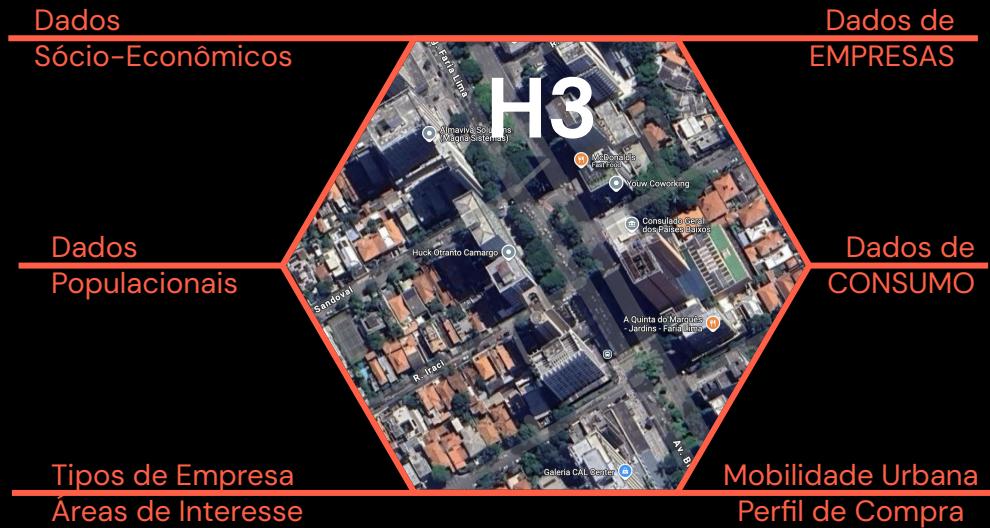
```
1 SELECT Rating
2 FROM mp_route
3 LIMIT 100 ;
```

Below the cells, there is a summary section titled "(2) Spark Jobs". At the bottom left, there are buttons for "Table" and "+".

At the very bottom, there is a preview table with two columns: "Rating" and "Pitches". The data rows are:

	Rating	Pitches
1	5.10b/c	4
2	5.10b/c	1
3	5.10b/c	1
4	5.10b/c	1

Visão 360° da microrregião



Visão 360° da microrregião



- PIB região
- Atividade Econômicas
- Assentamentos
- Área Urbana
- Área extrativista
- Madeireiras

Dados
Sócio-Econômicos

Dados da
Região

- Estado
- Município
- Clima
- Umidade

- População
- Densidade populacional
- Área indígena

Dados
Populacionais

Dados da
Vegetação

- Bioma
- Vegetação nativa
- Reforestada
- Tipo vegetação

- Proteção Ambiental
- Reserva Florestal
- Parque Estadual
- Parque Nacional
- Reserva Biológica

Áreas de
Proteção e Conservação

RFV
Rescência, freq., valor

- Qde de incêncios
- Recorrências
- % área nativa
- % área desmatada
- Última queimada

github.com/Databricks-BR/amazonia_geoui



Hands-On Labs

Democratizando dados com Databricks

LABORATÓRIO 4.1

Manipulando Dados e GEO.

https://github.com/Databricks-BR/lab_agosto_2025

Passo 1 – Entrando no Editor de Consultas (Query)

The screenshot shows the Databricks interface with the SQL Editor selected. A red box highlights the 'SQL Editor' button in the sidebar. The main area displays a catalog named 'workshop_08' and a query editor window titled 'Query_WorkShop'. The query window has a red border and contains the following SQL code:

```
1 create table faturamento_h3
2 as
3 select fat.*,
4     h3_longlatash3(longitude,latitude , 10) as h3_10,
5     h3_longlatash3(longitude,latitude , 9) as h3_09
6 from faturamento fat
```

The code uses the `h3_longlatash3` function to generate hierarchical geolocation data for the `faturamento` table.

Passo 2 – Criação de uma Tabela com Geo H3

Convertendo Coordenadas Geo em H3



```
create table faturamento_h3
as
select fat.*,
       h3_longlatash3(longitude,latitude , 10)   as h3_10,
       h3_longlatash3(longitude,latitude , 9)    as h3_09
from faturamento fat
```

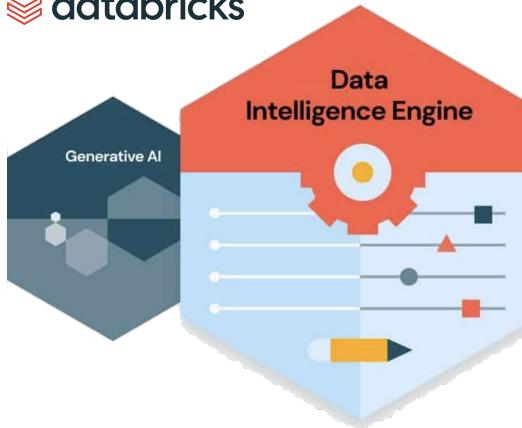
Passo 3 – Criação de uma Tabela para o APPs

```
CREATE OR REPLACE TABLE gold_faturamento_h3
AS
SELECT
    h3_09 AS h3_09_id,
    genero_cliente,
    bairro,
    faixa_divida,
    COUNT(num_cliente) AS contagem_clientes,
    SUM(val_divida)     AS valor_inadimplencia
FROM   faturamento_h3
WHERE  ind_inadimplente = 'S'
GROUP BY 1, 2, 3, 4
```



PERGUNTAS ?

Consumindo os Dados de forma eficiente



Visualizadores (BI)
tradicionalis de mercado



Databricks
ONE

permite aos usuários de negócios criar painéis e gráficos de forma interativa utilizando assistente.

permite aos usuários de negócios usar linguagem natural para fazer perguntas de negócio e obter insights valiosos e contextualizados em experiências conversacionais — como um ChatGPT.

Possibilita a criação rápida e simples de aplicações de dados em alguns cliques.

Um Portal para usuários finais das áreas de negócio, com um visual SIMPLES, CLEAN e intuitivo.



O que é um **Data Apps ?**

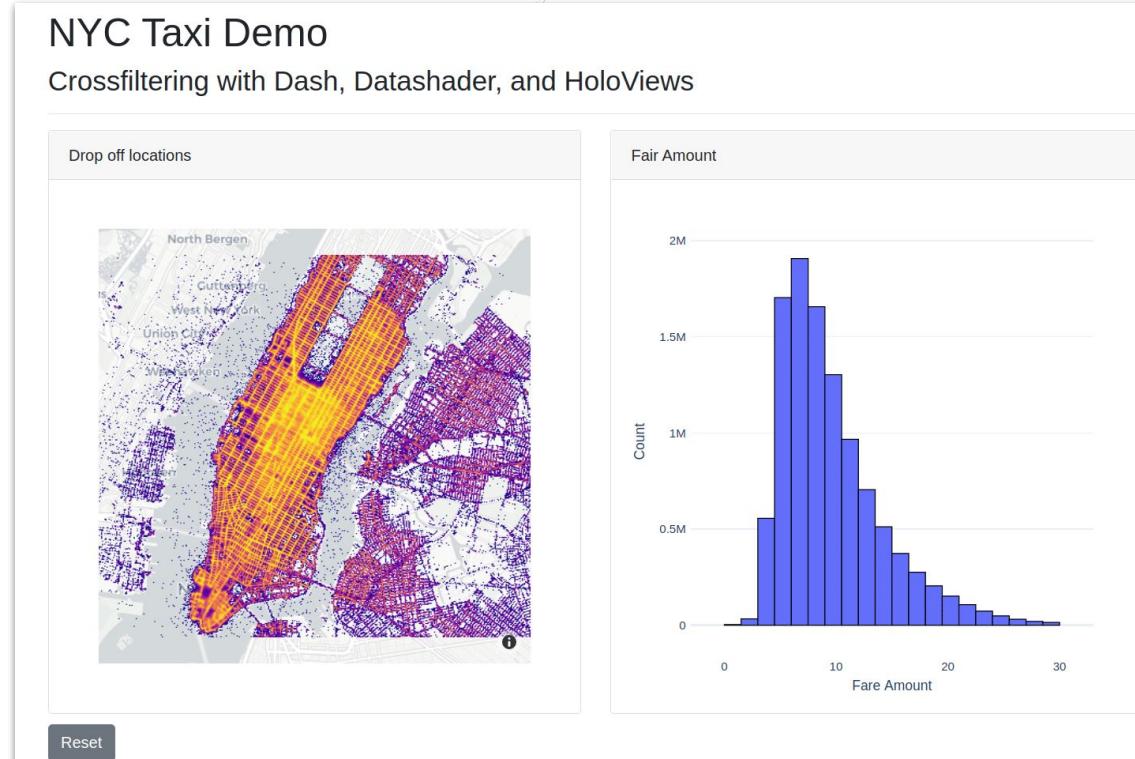
Um "Data Apps" é uma **Aplicação** (APP) de **dados e IA**, que foca na exploração de grandes conjuntos de dados para **propósitos analíticos**.





O que é um **Data Apps ?**

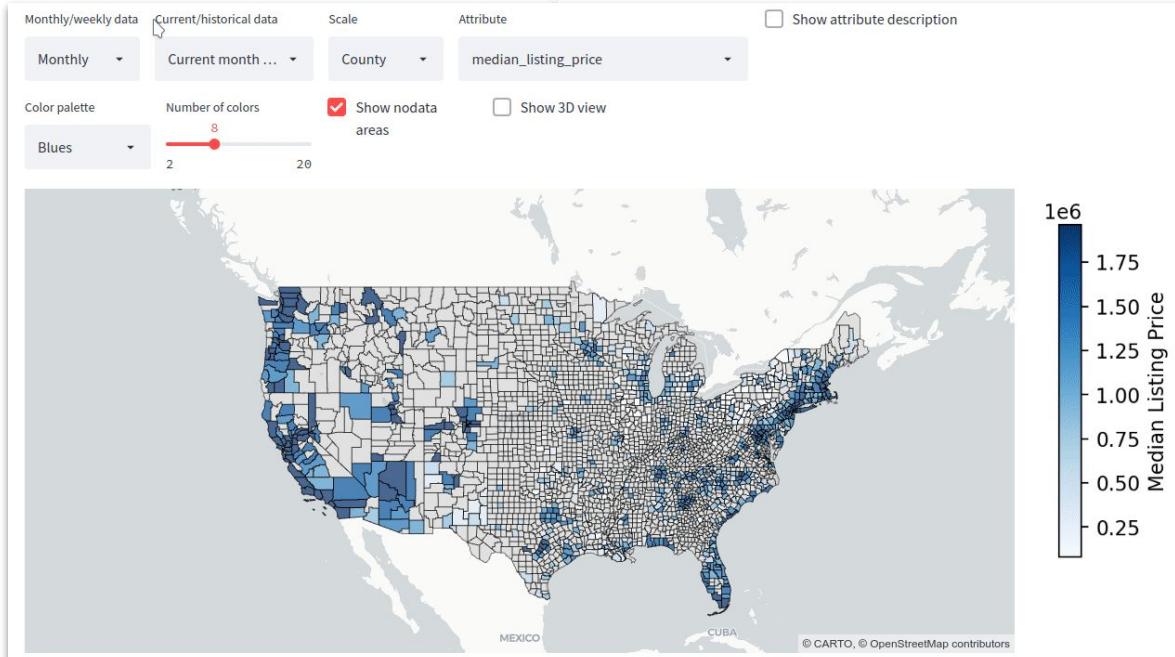
É uma aplicação
direcionada para um
determinado perfil de
consumo de dados, com o
propósito de:
simplificar e agilizar a
gestão.





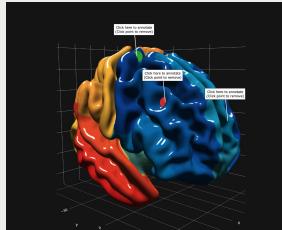
O que é um Data Apps ?

Um programa interativo
que coloca dados e
ferramentas
personalizadas nas mãos
de **usuários de negócio**
(principalmente)
não técnicos.

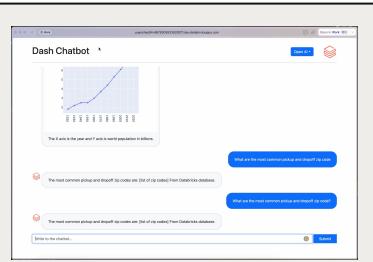




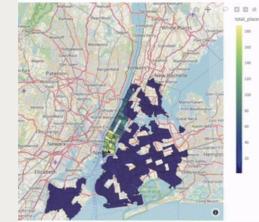
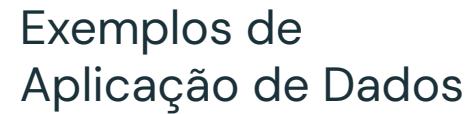
Casos de Uso de **Data Apps**



Visualizações Complexas



Assistente Gen AI



Gráficos de Geo Análise

A screenshot of a conversational AI application. At the top center is a purple star icon with a yellow outline. To its right, a green rounded rectangle contains the text "I am your Gen AI assistant". Below this is a large input field with a black border, containing the placeholder text "Ask me anything...". To the right of the input field is a large black arrow pointing to the right. At the bottom left is a circular icon with a microphone symbol, and at the bottom right is a circular icon with a speaker symbol.

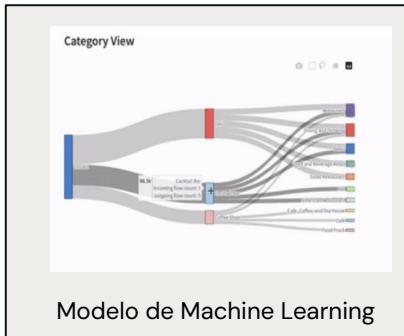
©2025 Databricks Inc. – All rights reserved

PRODUCTS ↓

-  ADIDAS | CLASSIC BACKPACK | LEGEND INK MULTICOLOUR
OS / blue
SKU: AD-001-blue-OS
-  NIKE | CRACKLE PRINT TB TEE
s / blue
SKU: NK-01-blue-s
-  PUMA | SUDE CLASSIC REG
S / red
SKU: PM-01-red-S



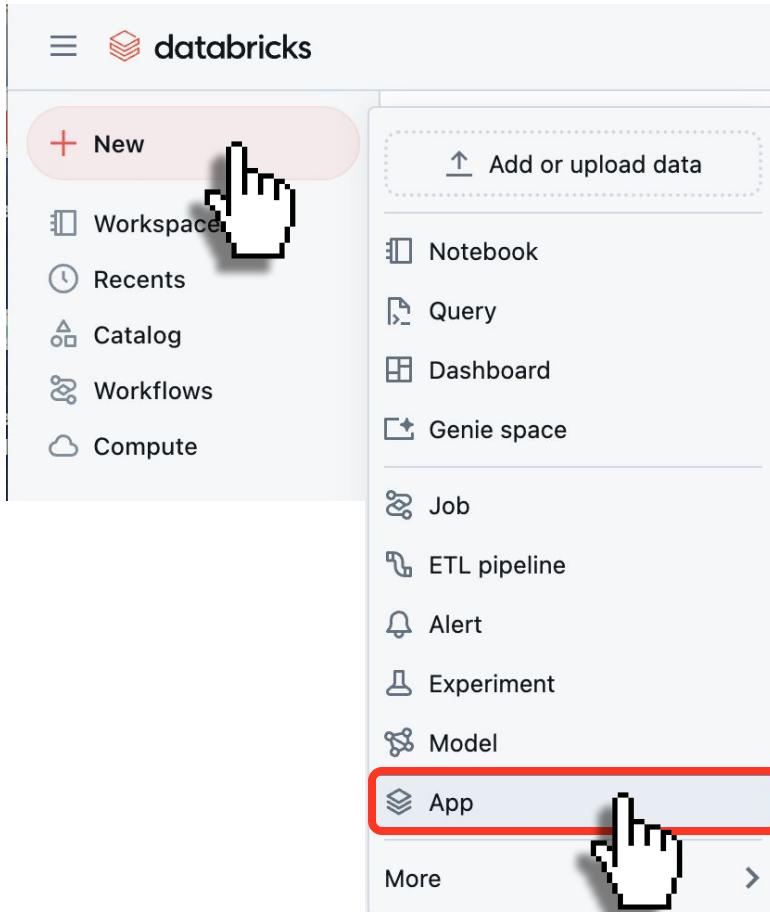
Integração de periféricos



Modelo de Machine Learning

```
POST /api/2.0/harvests  
{  
  "site": "Morris",  
  "yield": "35",  
  "variety": "Manchuria"  
}
```

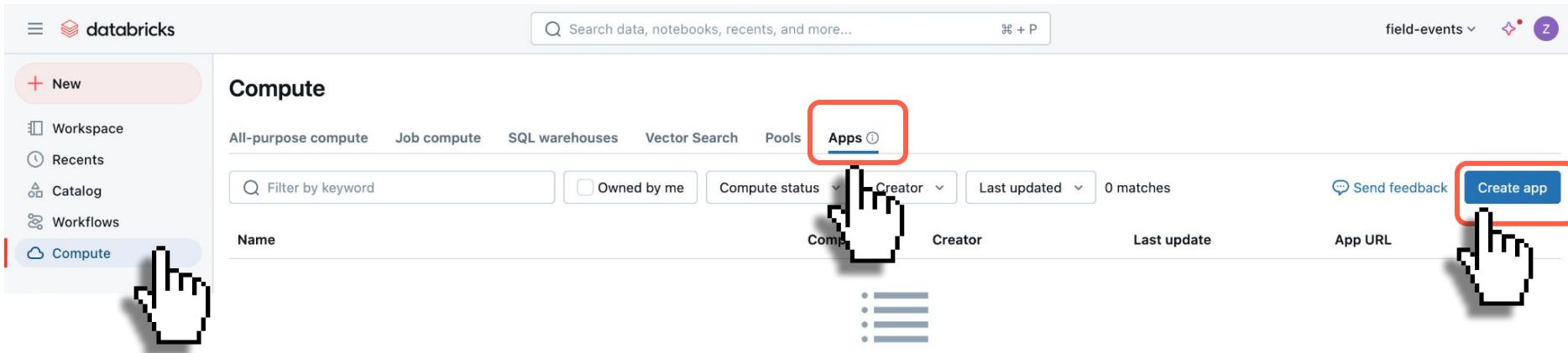
APIs



databricks
Lakehouse
APPS

Entrando na funcionalidade,
para criar um APP

Outra opção de navegação na funcionalidade de Lakehouse Apps



The screenshot shows the Databricks interface. On the left, there's a sidebar with a 'Compute' tab highlighted by a large white mouse cursor. At the top, there's a search bar and some navigation icons. In the center, there's a 'Compute' section with tabs for 'All-purpose compute', 'Job compute', 'SQL warehouses', 'Vector Search', 'Pools', and 'Apps'. The 'Apps' tab is currently selected, indicated by a red box and a smaller mouse cursor. Below the tabs, there are filters for 'Owned by me', 'Compute status', 'Creator', 'Last updated', and a search bar for 'Filter by keyword'. A message says '0 matches'. To the right, there are buttons for 'Send feedback' and a prominent blue 'Create app' button with a red box around it, also targeted by a mouse cursor. The main area shows a table with columns for 'Name', 'Creator', 'Last update', and 'App URL'. At the bottom, it says 'No apps found'.

Nessa opção podemos listar os APPs já criados, podendo alterá-los ou criar novas aplicações.

Create new app

Create custom app from scratch



Create a custom app

Bring your code and resources to build an app from scratch



Install from a template



Dash



Flask



Gradio



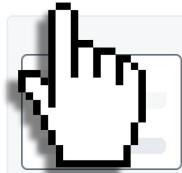
Shiny



Streamlit



Node.js



Chatbot

Create a chat UI using a large language model and Model Serving.

Dependencies: Serving endpoint



Data app

Create an app that reads from a SQL warehouse and visualizes the data using the app user's permissions.

Dependencies: SQL warehouse

User authorized API scopes: sql



<https://dash.plotly.com>

Dash by Plotly é um framework de código aberto que permite a criação de painéis interativos e aplicativos web com base em **Python**.

- O Dash é construído sobre modelos de machine learning e pode ser usado para visualização de dados;
- Abstração pura do Python em torno de HTML, CSS e JavaScript;
- Composição do layout usando Python com o módulo Dash HTML Components;
- Suporte a cerca de **50 tipos de gráficos**, incluindo mapas ;
- Código de aplicativo declarativo e reativo, o que facilita a construção de aplicativos complexos;
- Controle total sobre o visual e a sensação dos aplicativos;



Create new app

Create custom app from scratch

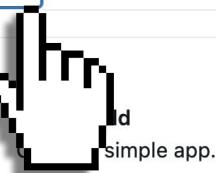


Create a custom app

Bring your code and resources to build an app from scratch



Install from a template



Flask



Gradio



Shiny

Streamlit



Node.js



simple app.



OLTP database app

Create a todo app that stores tasks in a postgres database hosted on Databricks.

Dependencies: Database



<https://flask.palletsprojects.com>

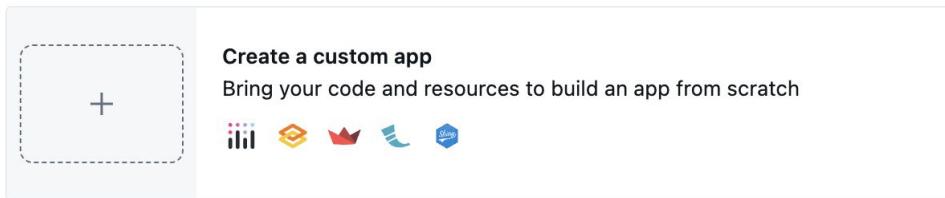
Flask é um micro-framework web escrito em **Python** que permite o desenvolvimento de aplicações web de forma rápida e eficiente:

- É conhecido pela sua simplicidade, facilidade de uso e extensibilidade
- Permite criar aplicativos robustos com menos código e menos complexidade
- É ideal para projetos que precisam de uma solução personalizada ou que devem escalar ao longo do tempo
- Oferece grande flexibilidade e controle sobre o projeto



Create new app

Create custom app from scratch



Create a custom app
Bring your code and resources to build an app from scratch

+

Dashboard 📈 API 🚜 Chatbot 🤖 Streamlit 🎨 Node.js 🏛️

Install from a template

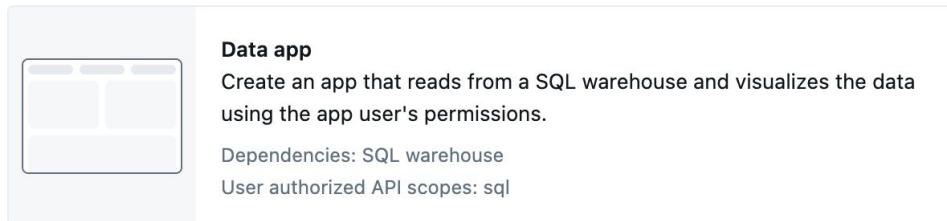
Dashboard 📈 Flask 🚜 Gradio 🎖️ Shiny 🎩 Streamlit 🎨 Node.js 🏛️



Chatbot 🤖

Create a chatbot using a large language model and Model Serving.

Dependencies: Serving endpoint



Data app 📊

Create an app that reads from a SQL warehouse and visualizes the data using the app user's permissions.

Dependencies: SQL warehouse
User authorized API scopes: sql



www.gradio.app

Gradio é uma biblioteca **Python** de código aberto projetada para tornar o desenvolvimento de **modelos de machine learning** (ML) mais acessível e amigável ao usuário.

Com o Gradio, os desenvolvedores podem criar interfaces intuitivas baseadas na web para seus modelos de ML, facilitando a interação de não especialistas com os modelos, o fornecimento de entrada e a visualização da saída.

Não importa se você está criando um modelo de regressão simples ou um sistema complexo de deep learning, o Gradio fornece uma solução versátil e direta para implantação e teste.



Create new app

Create custom app from scratch

Create a custom app
Bring your code and resources to build an app from scratch

Icons shown: Dash, Flask, Gradio, Shiny, Streamlit, Node.js

Install from a template

Shiny (selected)

Chatbot
Create a chat UI using a pre-trained language model and Model Serving.
Dependencies: Serving endpoint

Data app
Create an app that reads from a SQL warehouse and visualizes the data using the app user's permissions.
Dependencies: SQL warehouse
User authorized API scopes: sql



<https://shiny.posit.co>

Shiny é um framework de **linguagem R** para o desenvolvimento de aplicações web interativas:

- Sem necessidade de conhecimento prévio de HTML, CSS e JavaScript.
- Facilidade de desenvolvimento: permite desenvolver aplicações web completas, com o lado cliente e lado servidor diretamente em R.
- Reactive programming: possui um modelo de programação reativa que permite que as saídas reajam automaticamente a novos inputs do usuário.
- Widgets pré-construídos, personalizáveis e fáceis de usar, como plots, tables, sliders, dropdowns, e date pickers.
- Integração com R Markdown: permitindo a incorporação de várias aplicações nativamente dentro de um documento dinâmico maior.



Create new app

Create custom app from scratch

Create a custom app
Bring your code and resources to build an app from scratch

Install from a template

Dash Flask Gradio Shiny **Streamlit** Node.js

Chatbot
Create a chat UI using a large language model.
Dependencies: Serving endpoint

Data app
Create an app that reads from a SQL warehouse and visualizes the data using the app user's permissions.
Dependencies: SQL warehouse
User authorized API scopes: sql



Streamlit

<https://streamlit.io>

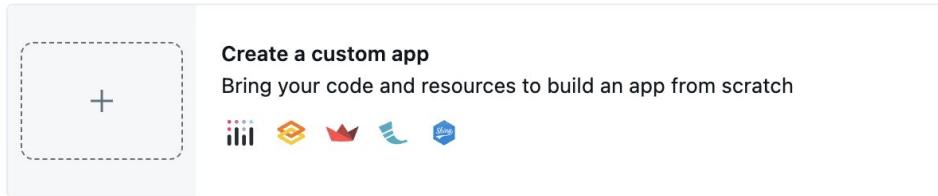
Streamlit é uma biblioteca de código aberto para **Python** que permite criar e compartilhar aplicativos da web para visualização e interação com dados:

- Popular entre cientistas de dados e engenheiros de machine learning
- Mais fácil de aprender e usar do que outras estruturas para criar aplicativos da web
- Permite criar aplicativos interativos com poucas linhas de código
- Abstrai muitos aspectos do desenvolvimento web, como a necessidade de escrever HTML e CSS
- É altamente **compatível com outras ferramentas** e bibliotecas de análise de dados, como: Pandas, Matplotlib, Plotly, TensorFlow, Keras, PyTorch, Numpy, Seaborn, Altair, Bokeh, Vega-Lite e muito mais.

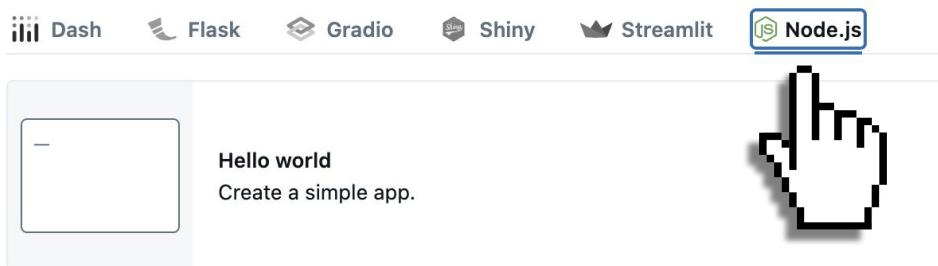


Create new app

Create custom app from scratch



Install from a template



Default templates are subject of the [Databricks license](#).



<https://nodejs.org/en>

Node.js é um ambiente de execução (runtime) de código aberto para **JavaScript** que permite construir aplicações de rede rápidas e escaláveis do lado do servidor:

- Popular entre desenvolvedores back-end, full-stack e engenheiros de DevOps.
- Permite o uso de JavaScript tanto no front-end quanto no back-end, unificando a stack de desenvolvimento.
- Sua arquitetura orientada a eventos e I/O (entrada/saída) não bloqueante é ideal para aplicações em tempo real, APIs e microserviços de alta performance.
- É a base para uma vasta gama de frameworks, bibliotecas e ferramentas modernas, como: Express.js, React, Angular, Vue.js, Next.js, NestJS, Socket.IO, e se integra com bancos de dados como MongoDB, PostgreSQL e MySQL.



Análise de Investimentos

Usando IA para avaliar uma carteira e identificar oportunidades

Open AI Assistant databricks

 Ron Gabrisko

Invested Market Assets
\$105,231

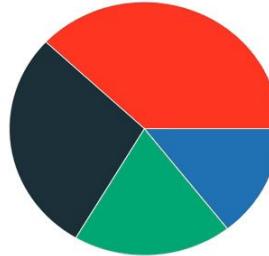
[AE26781-Investment](#)

[Have AI Assistant Analyze Portfolio](#)

Profits / Losses

Metric	Value
Value	+\$123
Percentage	+0.12%
Total Trades	10
All-time Gains / Losses	+\$523!
All-time Gains / Losses (Percentage)	+5.23%

Stocks by Industry



■ Tech ■ Finance ■ Healthcare ■ Energy

- Monte seu app como preferir
- Insights da carteira automatizados
- Análise dos documentos de resultados financeiros



Deploying Streamlit Apps to Dash Enterprise with Databricks

Authors: [Ivan Trusov](#) (Databricks), [Cal Reynolds](#) (Plotly)
Contributors: [Dave Gibbon](#) (Plotly)



<https://medium.com/plotly/deploying-streamlit-apps-to-dash-enterprise-with-databricks-1042cbb39327>



Databricks Apps Cookbook



Introduction

Tables

- Read a table
- Edit a table

Volumes

- Upload a file
- Download a file

AI / ML

- Invoke a model
- Run vector search

Business Intelligence

- AI/BI Dashboard

Workflows

- Trigger a job
- Retrieve job results

Compute

- Connect

Authentication

- Get current user



Databricks Apps Cookbook

Welcome to the Databricks Apps Cookbook!

Are you ready to serve some tasty apps to your users? You're in the right place!

Explore the recipes via the sidebar to quickly build flexible and engaging data apps directly on Databricks.

Have a great recipe to share? Raise a pull request on the [GitHub repository](#)!

Recipes

Tables	Volumes	AI / ML	Business Intelligence
Read a table	Upload a file	Invoke a model	AI/BI Dashboard
Edit a table	Download a file	Run vector search	
Workflows	Compute	Authentication	
Trigger a job	Connect	Get current user	
Retrieve job results		Retrieve a secret	

O Databricks Apps Cookbook contém **trechos de código prontos** para uso para criar dados interativos e aplicativos de IA usando o Databricks Apps.



<https://apps-cookbook.dev/docs/intro>



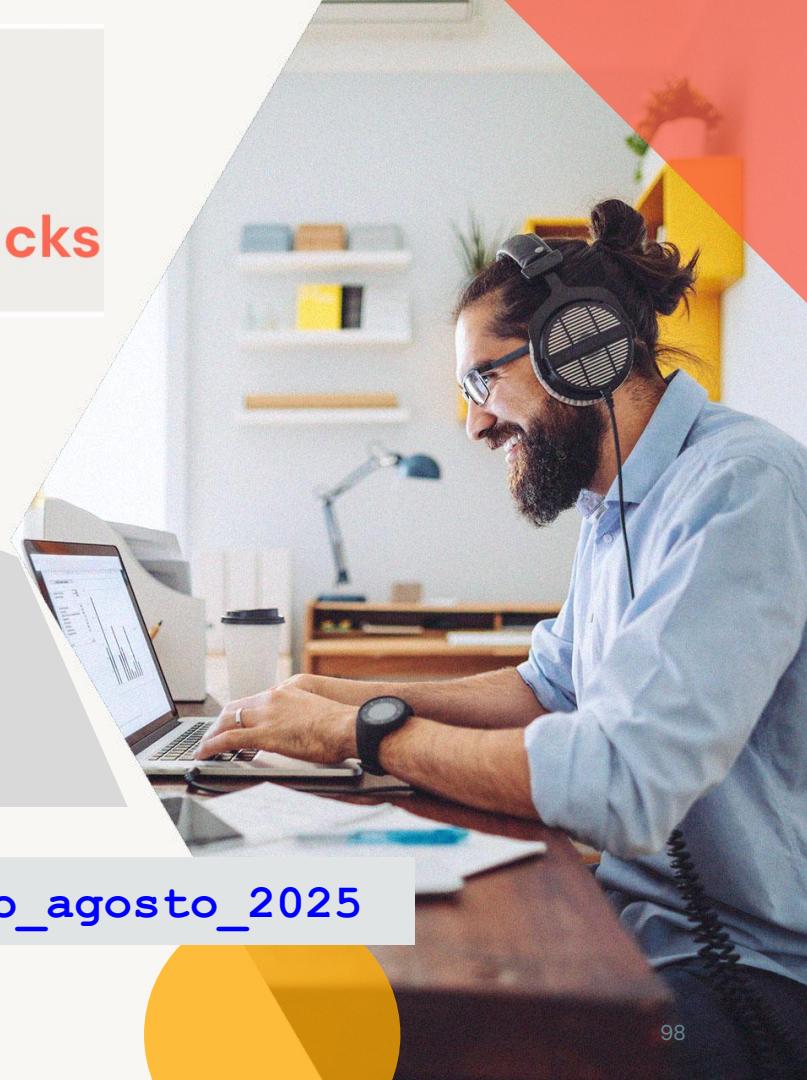
Hands-On Labs

Democratizando dados com Databricks

LABORATÓRIO 4.2

Aplicações Avançadas de Dados

https://github.com/Databricks-BR/lab_agosto_2025



Passo 1 – Criando um APPs

The screenshot shows the Databricks interface for creating a new application. On the left is a sidebar with various navigation links such as New, Workspace, Recents, Catalog, Workflows, Compute, SQL, SQL Editor, Queries, Dashboards, Genie, Alerts, Query History, and SQL Warehouses. Under Data Engineering, there are links for Job Runs, Data Ingestion, and Pipelines. Under Machine Learning, there are links for Playground, Experiments, Features, Models, and Serving. A search bar at the top right says 'Search data, notebooks, recent'. The main area is titled 'Create new app' and has a sub-section 'Choose how to start'. It offers two options: 'Template' (selected) and 'Custom'. Below these are several application templates: 'Dash', 'Flask', 'Gradio', 'Shiny', and 'Streamlit'. The 'Streamlit' option is highlighted with a blue underline. Below these are three more templates: 'Chatbot', 'Data app', and 'Hello world'. A large hand cursor icon is positioned over the 'Data app' template. At the bottom of the main area, it says 'Templates are subject of the Databricks license.' with 'Cancel' and 'Next' buttons.

Tipos de Aplicação (templates):

- Assistente Conversacional
- Aplicação de Dados (Dash)
- Aplicação genérica



Passo 2 – Nome e Descrição do APPs

Screenshot of the Databricks interface showing the "Create new app" form.

The sidebar menu includes:

- + New
- Workspace
- Recents
- Catalog
- Workflows
- Compute
- SQL
- SQL Editor
- Queries
- Dashboards
- Genie
- Alerts
- Query History
- SQL Warehouses
- Data Engineering
- Job Runs
- Data Ingestion
- Pipelines
- Machine Learning

The main "Create new app" form contains the following fields:

- App name**: test-lakehouse-apps
- Description**: (empty)
- Advanced settings**: (button to expand)
- App resources**:
 - Resource type**: SQL warehouse
 - Resource mapping**: dbdemos-shared-endpoint
 - Permission**: Can use
 - Resource key**: sql-warehouse

At the bottom right, there is a large hand cursor icon pointing towards the "Create and deploy app" button.

Bottom navigation: Back, Create and deploy app, and a small Databricks logo.

Passo 3 – Acessando o código do APPs

The screenshot shows the Databricks UI with the sidebar open. The 'Compute' section is selected under 'Compute > Apps'. A deployment for the app 'test-lakehouse-apps' is listed as 'Running' with the URL <https://test-lakehouse-apps-1444828305810485.aws.databricksapps.com>. A large hand cursor icon is positioned over the URL link. Below the URL, a deployment log shows a successful deployment completed 1 hour ago. The log entries are:

- Jan 02, 2025, 08:25:14 PM GMT-3 ✓ Stopped app successfully
- Jan 02, 2025, 08:25:15 PM GMT-3 ✓ Source code downloaded successfully
- Jan 02, 2025, 08:25:15 PM GMT-3 ✓ App spec validated successfully
- Jan 02, 2025, 08:25:16 PM GMT-3 ✓ Requirements installed successfully
- Jan 02, 2025, 08:25:18 PM GMT-3 ✓ App started successfully

The 'App resources' section shows a service principal named 'app-40zbx9 test-lakehouse-apps'. The 'Key' column lists 'sql-warehouse', 'Type' is 'SQL warehouse', 'Details' is 'dbdemos-shared-endpoint', and 'Permissions' is 'Can use'. The 'Edit in your IDE' section contains prerequisites and a command-line export template:

```
$ databricks workspace export-dir /Workspace/Users/luis.assuncao@databricks.com/databricks_apps/test-lakehouse-apps_2025_01_02-23_22/streamlit-data
```

The 'Sync the files' section provides instructions to copy the template to a computer and includes a sync command:

```
$ databricks sync --watch . /Workspace/Users/luis.assuncao@databricks.com/databricks_apps/test-lakehouse-apps_2025_01_02-23_22/streamlit-data
```

link da aplicação

código fonte do APP

status do deploy

recurso (Cluster)

Passo 4 – Acessando o código do APPs

Códigos fonte da aplicação

Workspace > Users > luis.assuncao@databricks.com > databricks_apps > test-lakehouse-apps_2025_01_02-23_22  main >

streamlit-data-app 

 Share 

Name 	Type	Owner	Created at	
 app.py	File	Luis Assunção	01/02/2025, 08:22:59 PM	
 app.yaml	File	Luis Assunção	01/02/2025, 08:22:59 PM	
 requirements.txt	File	Luis Assunção	01/02/2025, 08:22:59 PM	



Workspace

lakehouse_app

app.py

File Edit View Run Help Tabs: ON main Last edit was 10 hours ago Connect Schedule Share Edit

app.py

app.yaml

requirements.txt

app.py

```
1 # -----
2 # MAIN APP WITH MULTI-PAGE NAVIGATION
3 #
4 import streamlit as st
5 import pandas as pd
6 import json
7 import keplergl
8 import os
9 from databricks import sql
10 from databricks.sdk.core import Config
11 from streamlit_keplergl import keplergl_static
12 import h3
13 import pydeck as pdk
14
15 #
16 # PAGE CONFIGURATION
17 #
18
19 st.set_page_config(
20     page_title="Mapa de Inadimplência",
21     page_icon="⚠️",
22     layout="wide",
23     initial_sidebar_state="expanded"
24 )
25
26 #
27 # SIDEBAR NAVIGATION
28 #
29
30 # Logo
31 st.sidebar.image(
32     "https://raw.githubusercontent.com/Databricks-BR/lab_agosto_2025/main/images/app_logo.png",
33     width=200
34 )
35
36 st.sidebar.header("Mapa de Inadimplência")
37
38 # PAGE SELECTION
39 page = st.sidebar.selectbox(
40     "Escolha uma página"
```

Lakehouse Apps – arquivo **app.py**

```
4 import streamlit as st
5 import pandas as pd
6 import json
7 import keplergl
8 import os
9 from databricks import sql
10 from databricks.sdk.core import Config
11 from streamlit_keplergl import keplergl_static
12 import h3
13 import pydeck as pdk
```

biblioteca do **Streamlit**

biblioteca do **Kepler.gl**

biblioteca do **H3**



```
14  
15 # -----  
16 # PAGE CONFIGURATION  
17 # -----  
18  
19 st.set_page_config(  
20     page_title="Mapa de Inadimplência",  
21     page_icon="⚠️",  
22     layout="wide",  
23     initial_sidebar_state="expanded"  
24 )  
25  
26 # -----  
27 # SIDEBAR NAVIGATION  
28 # -----  
29  
30 # Logo  
31 st.sidebar.image(  
32     "https://raw.githubusercontent.com/Databricks-BR/lab_agosto_2025/refs/heads/main/images/app_logo.png",  
33     width=200  
34 )  
35  
36 st.sidebar.header("Mapa de Inadimplência")  
37  
38 # PAGE SELECTION  
39 page = st.sidebar.selectbox(  
40     "Escolha uma página!"
```

Lakehouse Apps – arquivo **app.py**

```
59 # -----
60 # DATABASE CONNECTION (SHARED)
61 #
62
63 assert os.getenv('DATABRICKS_WAREHOUSE_ID'), "DATABRICKS_WAREHOUSE_ID must be set in app.yaml."
64
65 def sqlQuery(query: str) -> pd.DataFrame:
66     cfg = Config()
67     with sql.connect(
68         server_hostname=cfg.host,
69         http_path=f"/sql/1.0/warehouses/{os.getenv('DATABRICKS_WAREHOUSE_ID')}",
70         credentials_provider=lambda: cfg.authenticate
71     ) as connection:
72         with connection.cursor() as cursor:
73             cursor.execute(query)
74             return cursor.fetchall_arrow().to_pandas()
75
76 @st.cache_data(ttl=30)
77 def getData():
78     # Insira o nome da tabela criada
79     return sqlQuery("select * from academy.genie_aibi.gold_faturamento_h3")
80
```

Temos uma função de **consulta de uma tabela** (*query*) para colocar os dados em um DataFrame

Lakehouse Apps - arquivo `app.py`

```
181     # --- Display Map ---
182     try:
183         map_1 = keplergl.KeplerGl(height=600, config=kepler_config)
184         map_1.add_data(data=final_df, name="inadimplencia_data")
185
186         map_html = map_1._repr_html_()
187         st.components.v1.html(map_html, height=600)
188
189     except Exception as e:
190         st.error(f"Erro exibindo o mapa: {e}")
191
```

Estamos utilizando o **Kepler.gl** para exibirmos o mapa de inadimplência no nosso App

Lakehouse Apps - arquivo **app.py**

```
309     # --- 4. Render the map in Streamlit ---
310     try:
311         st.pydeck_chart(pdk.Deck(
312             layers=[h3_layer],
313             initial_view_state=view_state,
314             map_style=pdk.map_styles.LIGHT,
315             tooltip=tooltip
316         ))
317     except Exception as e:
318         st.error(f"Ocorreu um erro ao renderizar o mapa com Pydeck: {e}")
319
```

Estamos utilizando o **PyDeck** para exibirmos o mapa 3D de inadimplência no nosso App

Lakehouse Apps - arquivo `app.py`

```
322 # -----
323 # GENIE CHAT PAGE
324 # -----
325 import json
326 from databricks.sdk import WorkspaceClient
327 from databricks.sdk.service.dashboards import GenieAPI
328
329 # Insira o ID do Genie Space criado anteriormente
330 genie_space_id = "<INSIRA-AQUI-O-ID-DO-GENIE-SPACE>"
331
332 workspace_client = WorkspaceClient(
333     host=os.environ.get("DATABRICKS_HOST"),
334     client_id=os.environ.get("DATABRICKS_CLIENT_ID"),
335     client_secret=os.environ.get("DATABRICKS_CLIENT_SECRET"),
336 )
337 genie_api = GenieAPI(workspace_client.api_client)
338 conversation_id = st.session_state.get("genie_conversation_id", None)
```

Estamos utilizando a API do **Genie** para realizarmos a sua integração com o nosso App

Lakehouse Apps - arquivo **app.py**

```
425      # -----
426      # ANALYTICS PAGE
427      #
428
429      st.markdown("---")
430      st.subheader("⚡ Dashboard Interativo: Databricks AI/BI")
431
432      # Insira o link do Embed Dashboard
433      src_dashboard="<INSIRA-AQUI-0-LINK-DO-EMBED-DASHBOARD>"
434
435      st.components.v1iframe(
436          src=src_dashboard,
437          scrolling=True,
438          height=600
439      )
```

Estamos embeddando o **AI/BI Dashboard** utilizando o iFrame

Passo 5 – Vamos subir um App via Notebooks

Vamos acessar a pasta **04_LAB_criando_aplicativo/01_deploy_app**;

The screenshot shows a Databricks repository interface. The top navigation bar includes the Databricks logo, the project name "Databricks-BR / lab_agosto_2025", a search bar, and navigation links for Code, Issues, Pull requests, Actions, Projects, Wiki, Security, Insights, and Settings.

The left sidebar displays the "Files" section with a dropdown menu set to "main". It lists several items:

- A dropdown menu icon with "main" and a dropdown arrow.
- A "+" button for creating new files.
- A search bar with the placeholder "Go to file".
- A folder named "01_LAB_importando_dados".
- A folder named "04_LAB_criando_aplicativo" which is expanded, showing its contents.
- A folder named "lakehouse_app".
- A file named "01_deploy_app.ipynb".
- A file named "RFADMF.md".

The right main area shows a list of commits under the path "lab_agosto_2025 / 04_LAB_criando_aplicativo /". The commits are:

Name	Last commit message
Databricks-BR LAB 04	
..	
lakehouse_app	Add LAB_04 code
01_deploy_app.ipynb	Add LAB_04 code

A red rectangular box highlights the "01_deploy_app.ipynb" file in the commit list. A large hand cursor icon is positioned over this highlighted file, indicating it is selected or about to be interacted with.



Passo 6 – Altere os parâmetros

Vá na aba do Databricks SQL Warehouse, selecione o SQL Warehouse desejado, copie o ID e cole na variável.

The screenshot shows the Databricks interface for managing SQL Warehouses. On the left, a sidebar menu lists various options: SQL Editor, Queries, Dashboards, Genie, Alerts, Query History, and SQL Warehouses. The 'SQL Warehouses' option is highlighted with a red box and a cursor icon pointing at it. The main content area is titled 'SQL Warehouses > Shared Unity Catalog Serverless'. Below the title, there are tabs for 'Overview' (which is underlined), 'Connection details', and 'Monitoring'. Under the 'Overview' tab, the 'Status' section shows a green dot next to 'Running'. The 'Name' section displays 'Shared Unity Catalog Serverless (ID: 4b9b953939869199)'. A red box highlights the warehouse name, and a cursor icon points at the ID. At the bottom left, the footer reads '©2023 Databricks Inc. — All rights reserved'. At the bottom right, there is a small logo consisting of three squares.

Passo 6 – Altere os parâmetros

Altere a **célula 6** com o ID do seu SQL Warehouse;

▶ ✓ Yesterday (4s)

6: Configura o Cliente do Workspace no Databricks

```
from databricks.sdk import WorkspaceClient
```

```
app_name = "geospatial-viewer"
```

```
# Vá na aba do Databricks SQL Warehouse, selecione o SQL Warehouse desejado, copie o ID e cole na variável
```

```
sql_warehouse_id = "<INSIRA-O-SQL-WAREHOUSE-ID-AQUI>"
```

```
w = WorkspaceClient()
```



Passo 7 – Copie o ID do Genie Spaces

COPIE o ID da Sala Genie que se encontra na URL quando entramos na sala, como no exemplo à esquerda.

Quando você abre um Genie Space na interface web do Databricks, o ID do espaço está presente na própria URL. O formato típico da URL é:



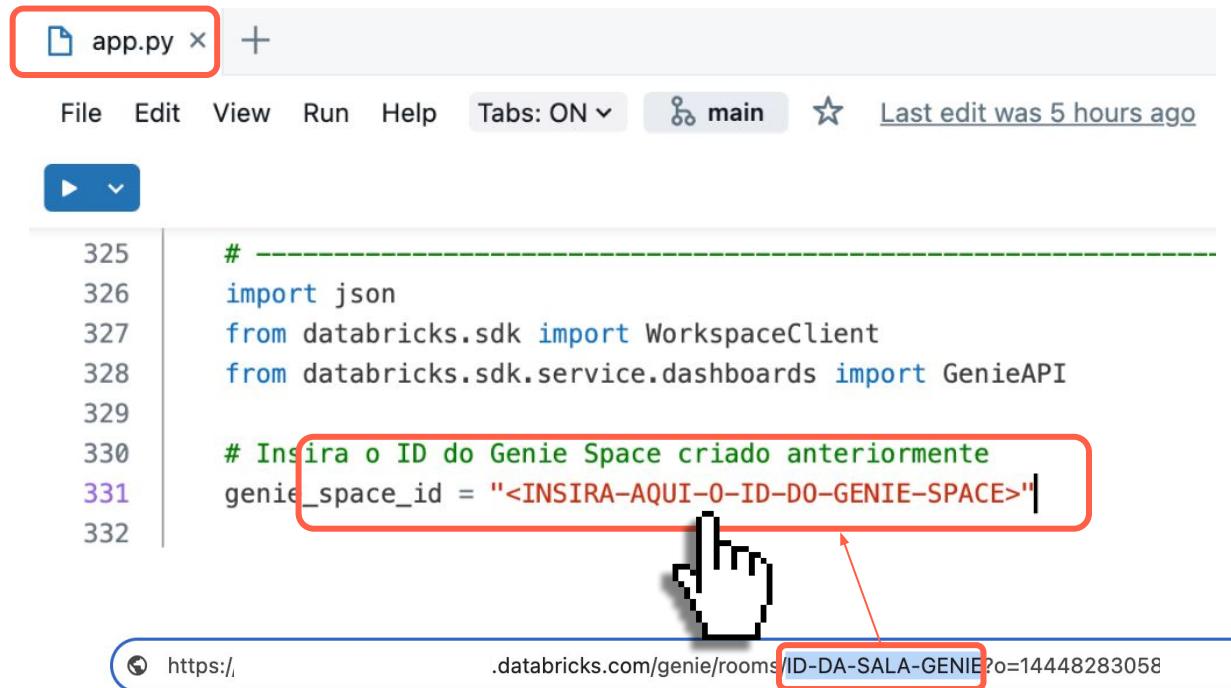
```
https://<databricks-instance>/genie/rooms/<space-id>
```

```
https://cloud.databricks.com/genie/rooms/01f07ca798c01f90b56d7638e2d00984?o=1444828...
```



Passo 8 – Altere o ID do Genie Spaces

Vá até a linha 331 e cole o **ID da sua Sala Genie** na variável;



```
app.py +  
File Edit View Run Help Tabs: ON main Last edit was 5 hours ago  
▶  
325 # -----  
326 import json  
327 from databricks.sdk import WorkspaceClient  
328 from databricks.sdk.service.dashboards import GenieAPI  
329  
330 # Insira o ID do Genie Space criado anteriormente  
331 genie_space_id = "<INSIRA-AQUI-O-ID-DO-GENIE-SPACE>"  
332  
https://databricks.com/genie/rooms/<ID-DA-SALA-GENIE>o=14448283058
```



Passo 9 – Copie o ID do Dashboard

O Link do Dashboard se encontra na página do dash,

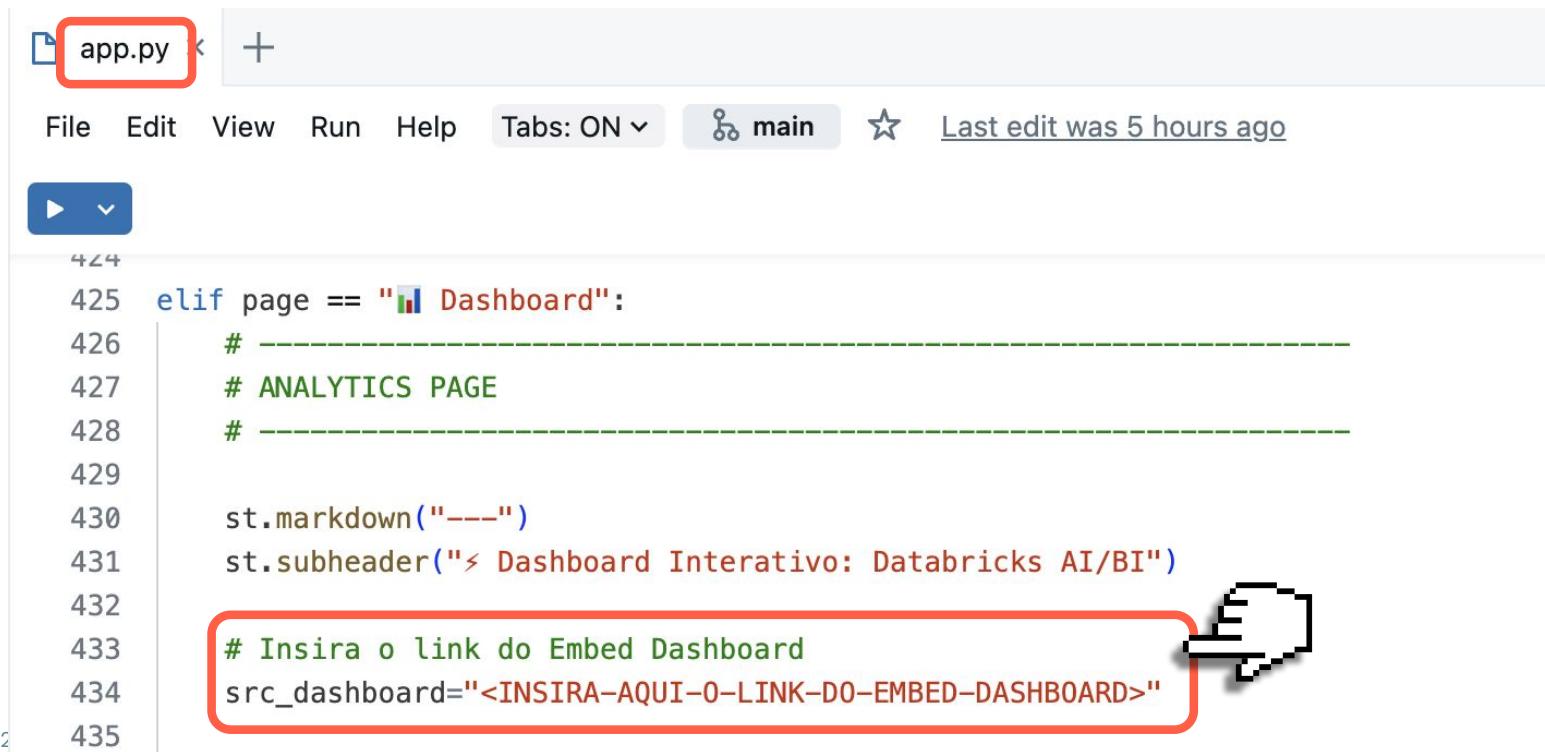
depois de publicarmos > Share > **Embed Dashboard** > **iFrame** > Copiar o Link do "src".

Dentro da página do AI/BI Dashboard



Passo 10 – Altere o parâmetro do ID do Dashboard

Vá até a linha 434 e cole o **Link do seu AI/BI Dashboard** na variável;



```
File Edit View Run Help Tabs: ON main Last edit was 5 hours ago  
app.py +  
424  
425 elif page == "Dashboard":  
426     # -----  
427     # ANALYTICS PAGE  
428     # -----  
429  
430     st.markdown("---")  
431     st.subheader("⚡ Dashboard Interativo: Databricks AI/BI")  
432  
433     # Insira o link do Embed Dashboard  
434     src_dashboard="<INSIRA-AQUI-O-LINK-DO-EMBED-DASHBOARD>"  
435
```

Passo 11 – Deploy do APPs (Run All)

The screenshot shows the Databricks workspace interface. On the left, the sidebar includes sections for Workspace, Recents, Catalog, Discover, Jobs & Pipelines, Compute, Marketplace, SQL, Data Engineering, AI/ML, and Experiments. The main area displays a notebook titled '01_deploy_app' under the workspace '04_LAB_criando_aplicativo'. The notebook contains several cells, with the first one visible showing Python code:

```
%pip install --upgrade databricks-sdk==0.36.0 -q  
dbutils.library.restartPython()
```

. A large red box highlights the 'Run all' button in the top right corner of the notebook header. Another red box highlights the 'Run all' button in the top right corner of the workspace sidebar.

Run all cells in this notebook.

File Edit View Run Help Python Tabs: ON main Connected Schedule Share

01_deploy_app

04_LAB_criando_aplicativo

lakehouse_app

01_deploy_app

README.md

Run all

Run all

Descrição e Objetivos

Este projeto tem como objetivo fornecer uma **visualização interativa da inadimplência** no estado de São Paulo, utilizando dados geoespaciais e inteligência artificial para análises avançadas.

Descrição do Cluster

Toda a demo pode ser executada utilizando Serverless ou um tipo de instância de sua preferência.

5: Instala e atualiza as bibliotecas necessárias

```
%pip install --upgrade databricks-sdk==0.36.0 -q  
dbutils.library.restartPython()
```

Passo 12 – Copiar o ID do "usuário aplicação"

Dentro do App, clique na aba **Authorization** e copie o **Service Principal**:

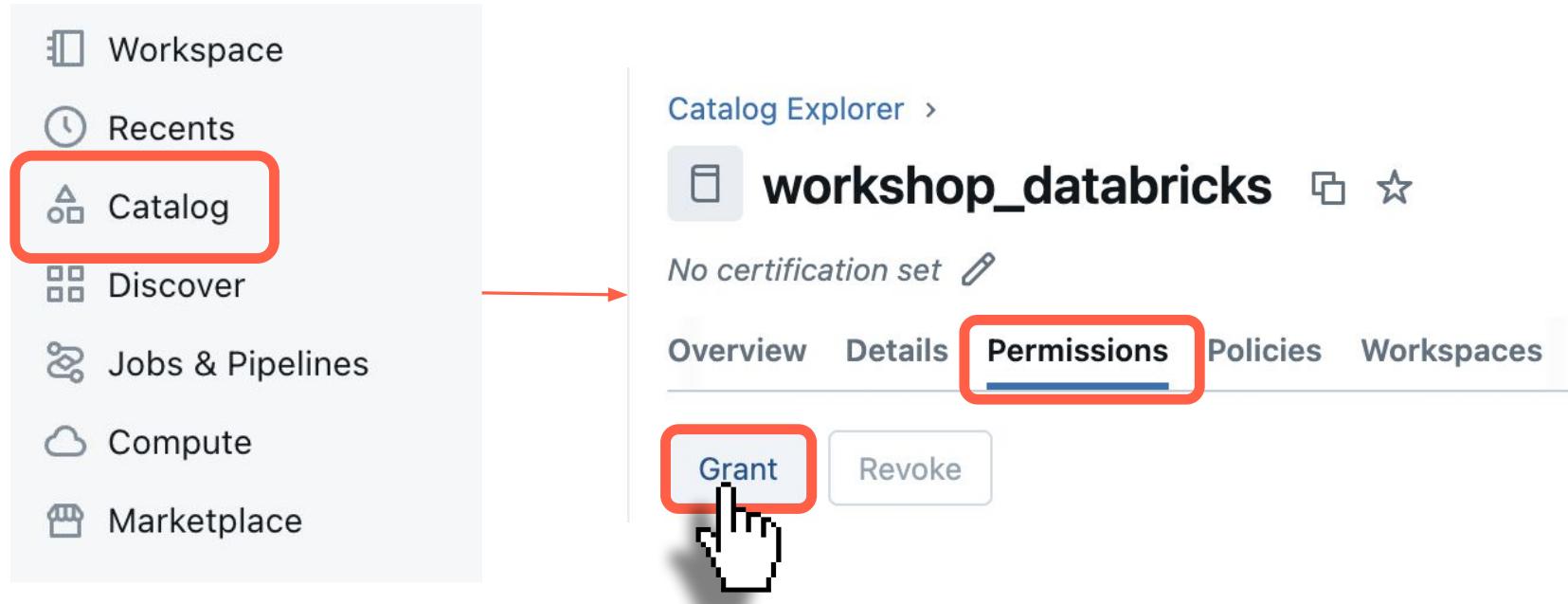
Compute > Apps >

The screenshot shows the Databricks Compute > Apps interface. An application named "geospatial-viewer-demo" is selected. The navigation bar includes tabs for Overview, Authorization, Deployments, Logs, and Environment. The "Authorization" tab is highlighted with a red box and a blue underline. Below the tabs, the "App authorization" section displays the service principal information: "Service principal: app-40zbx9 geospatial-viewer-demo". To the right of this text is a "Copy" button enclosed in a red box, with a hand cursor icon pointing at it. A large red box also surrounds the "Authorization" tab itself.

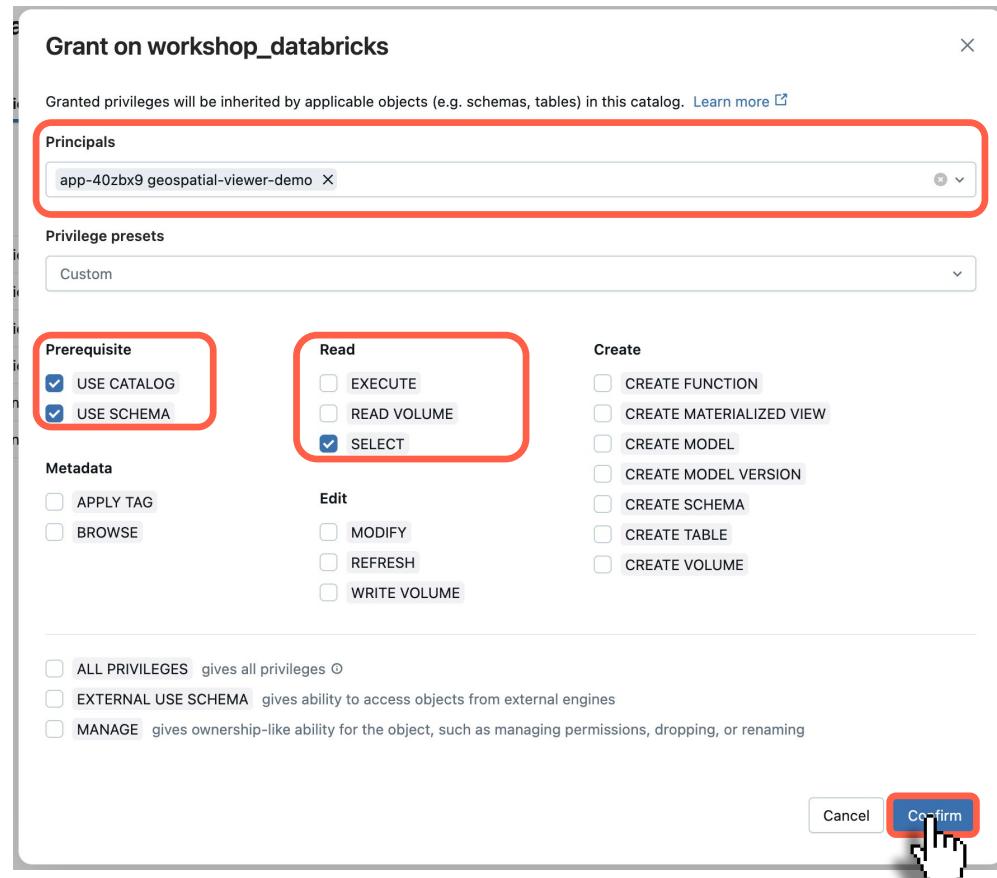


Passo 13 – Liberar o acesso ao Catálogo (Dados)

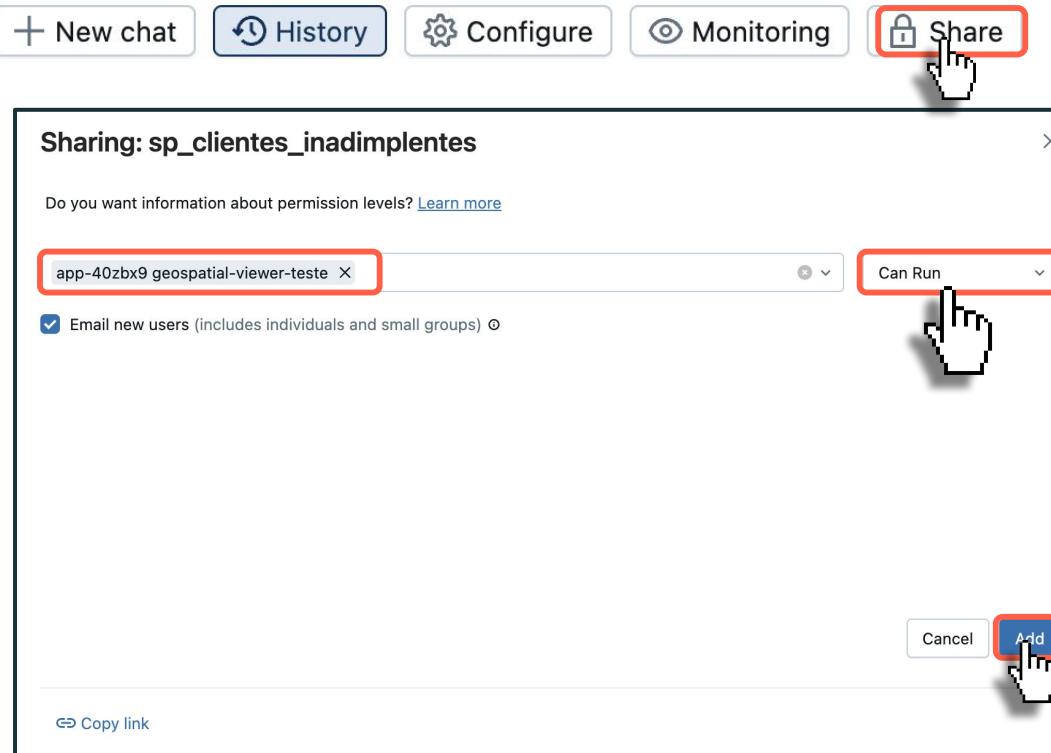
Agora, precisamos **liberar os acessos** para o Catálogo (**USE CATALOG**), Schema (**USE SCHEMA**) e Tabela (**SELECT**).



Passo 13 – Liberar o acesso ao Catálogo (Dados)



Passo 14 – Liberar o acesso ao GENIE



Passo 15 – Executar o APP

Dentro do App, na aba **Overview**, clique no **link do lado de Running**.

Compute > Apps >

geospatial-viewer-demo

Send feedback

Overview Authorization Deployments Logs Environment

● Running <https://geospatial-viewer-demo-1444828305810485.aws.databricksapps.com> 

Deployment

Source: /Workspace/Users/vinicius.fialho@databricks.com/Hunter/2025_Evento_Apps/apps/lakehouse_app

✓ Last deploy completed 16 minutes ago Inspect deployed code

Aug 19, 2025, 04:02:41 PM GMT-3 ✓ Stopped app successfully

Aug 19, 2025, 04:02:41 PM GMT-3 ✓ Source code downloaded successfully

Aug 19, 2025, 04:02:41 PM GMT-3 ✓ App spec loaded successfully

Aug 19, 2025, 04:03:24 PM GMT-3 ✓ Packages installed successfully

Aug 19, 2025, 04:03:24 PM GMT-3 ✓ App built successfully

Aug 19, 2025, 04:03:26 PM GMT-3 ✓ App started successfully

App pronto para uso



Monitoramento das áreas de inadimplência

Total de clientes inadimplentes: 1,050

Mapa de Inadimplência

Escolha uma página:

Mapa Interativo

Links Úteis

Repositório Git

Referências

Contato



© OpenStreetMap

© 2025 Databricks Brasil - Mapa de Inadimplência | Powered by Streamlit, Kepler.gl, PyDeck & Databricks

PERGUNTAS ?



DATA+AI
WORLD TOUR

3 de setembro de 2025

Data Intelligence for All

O Data + AI World Tour leva a Databricks Data Intelligence Platform para a comunidade global de dados. As principais empresas estão assumindo o controle de seus dados e aplicando inteligência artificial para transformar seus negócios. Venha descobrir como você também pode fazer isso.

[Inscreva-se](#)

São Paulo, Brasil

PRESENÇAS CONFIRMADAS:

- Santander
- iFood
- Nubank
- Prodesp
- Natura
- Bradesco
- Smiles
- Webmotors
- Bemol
- Serasa
- Amil



SCAN ME



Obrigado!



databricks_br



Luis Assunção
DATABRICKS





databricks

