databricks

AI/BI for Data Analysts

Databricks Academy

Agenda

Modules in this Course	Time
Dashboards and Visualizations in Databricks	2 hr 15 mins
AI/BI Genie	1 hr 15 mins

Course Learning Objectives

By the end of this course, you'll be able to use Databricks AI/BI tools to:

- Design dashboards for business insights.
- Share business intelligence assets with collaborators and stakeholders.
- Periodically revise data assets in accordance with best practices and new information.
- Create data assets for self-service analytics.
- Manage data assets for business intelligence in Databricks.

Lab Exercise Environment

Technical Details

- Your lab environment is provided by Vocareum.
- It will open in a new tab.
- It has been configured with the permissions and resources required to accomplish the tasks outlined in the lab exercise.
- Third party cookies must be enabled in your browser for Vocareum's user experience to work properly.
- Make sure to enable pop ups!

Before we get started...

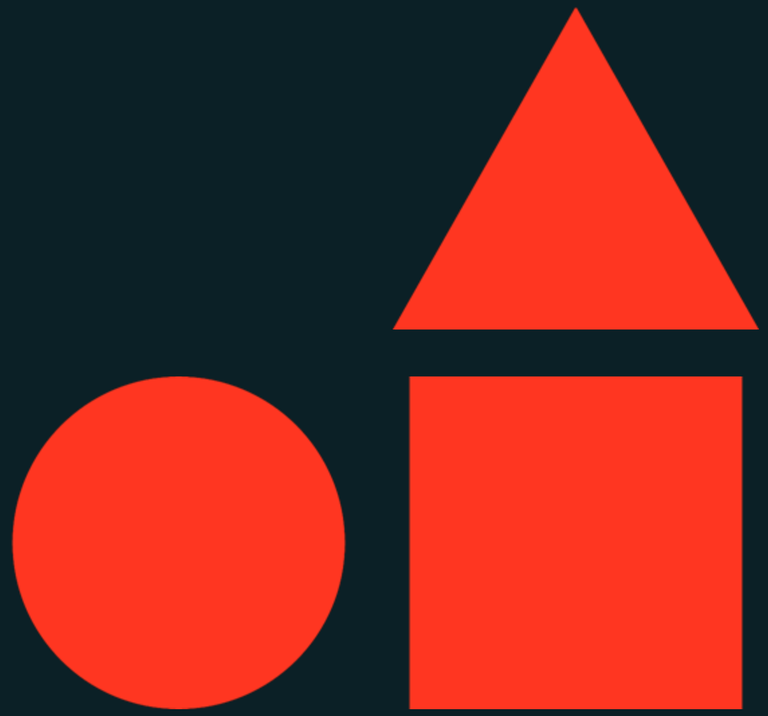
Just a quick note

- Be aware, Databricks is a highly innovative company and through our courses we try to give you the latest and greatest.
- If you encounter inconsistencies in product naming, UI environments, or other areas of this content - let us know!
  Visit help.databricks.com to submit your feedback.

Enjoy the course!

databricks

Dashboards and Visualizations in Databricks



AI/BI for Data Analysts

Agenda

Dashboards and Visualizations in Databricks	Time	Lecture	Demo	Activity
AI/BI Dashboards	15 mins	✓		
Just enough SQL	10 mins	✓		
Designing Datasets for Dashboards	20 mins		✓	
Creating Visualizations and adding Summary Statistics to Dashboards	20 mins		✓	
AI Enhanced Features	8 mins		✓	
Filters	8 mins		✓	
Sharing Dashboards with Stakeholders and Others	8 mins		✓	
Managing Dashboards in Production	8 mins		✓	
Dashboard and Visualization Lab Activity	30 mins			✓

Learning Objectives (Part 1)

Design dashboards for business insights.

- Create datasets specifically for AI/BI Dashboards.
- Create the visualizations and widgets necessary for displaying the results of summary statistics.
- Use AI to create visualizations and widgets for a dashboard.
- Add dynamic features to a dashboard (parameters, filters).

Learning Objectives (Part 2)

Securely share data assets created for business intelligence purposes.

- Share business intelligence assets with collaborators and stakeholders.
- Use embedded credentials with a Dashboard to share with account-level Databricks users.
- Subscribe users to dashboards.

Periodically revise data assets in accordance with best practices and new information.

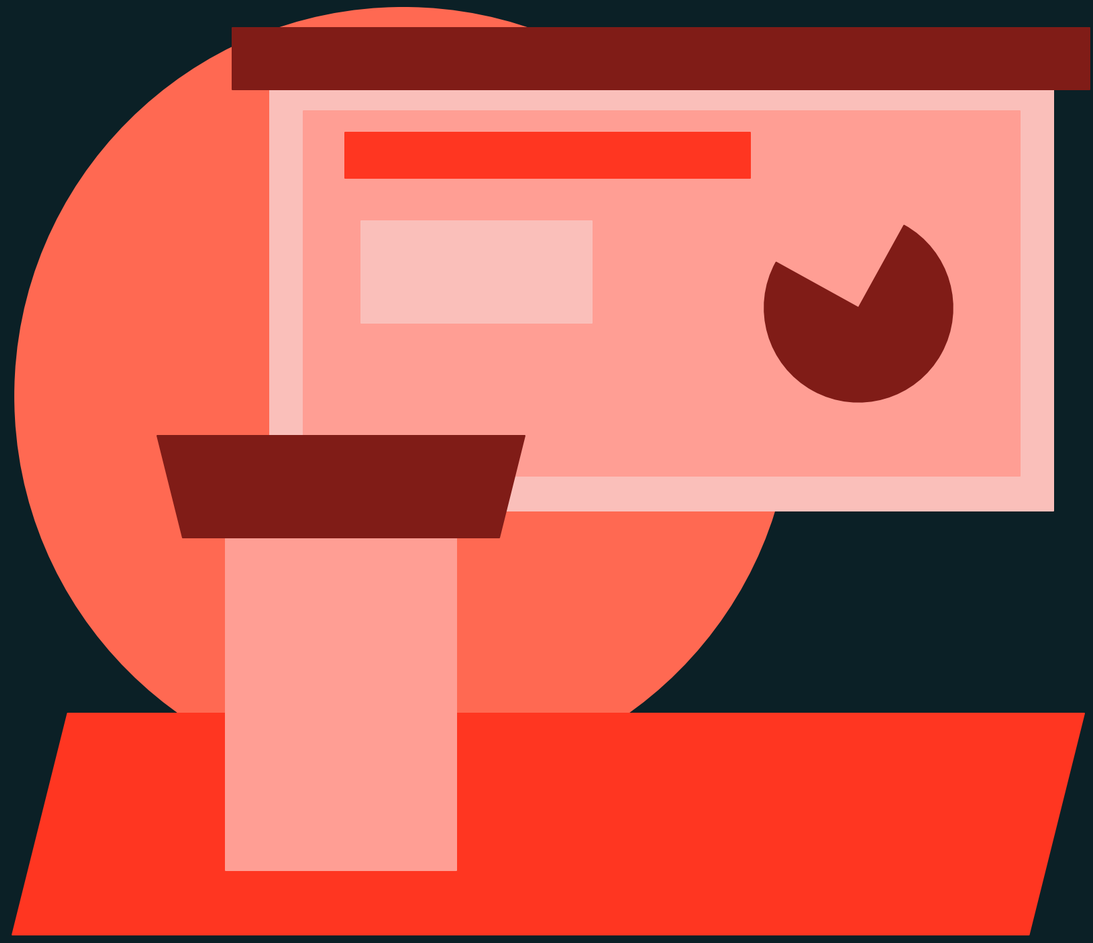
- Set refresh schedules on data assets used for business intelligence.

databricks

Dashboards and Visualizations in Databricks

LECTURE

AI/BI Dashboards



Dashboards bridge the gap between data and information

Data

Dashboards

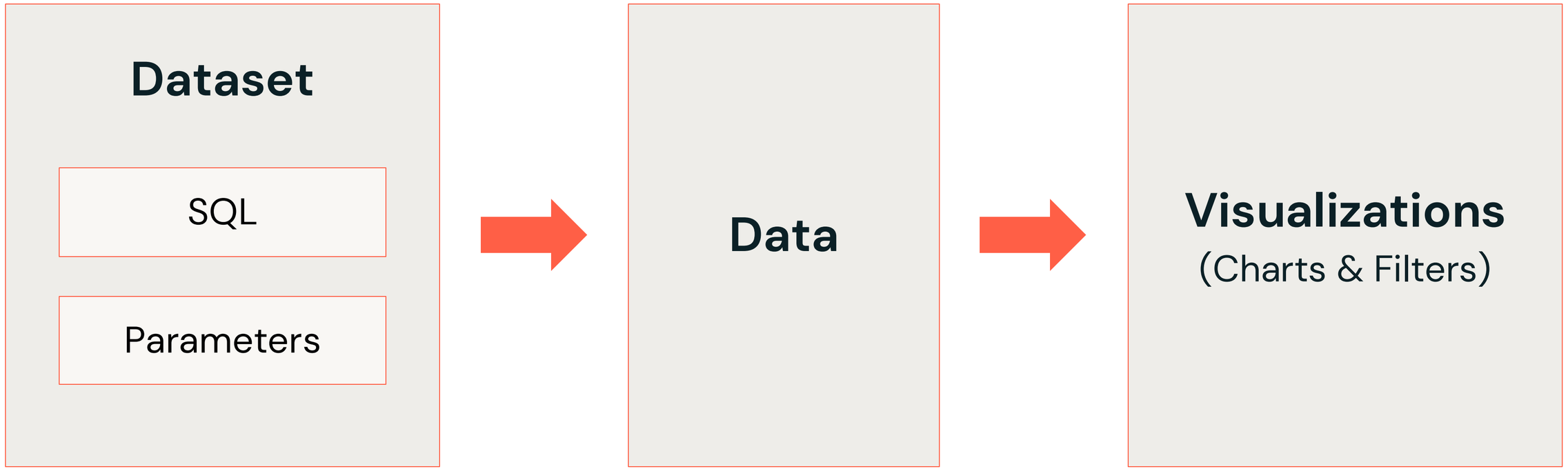
Business insights

What is an AI/BI Dashboard?

It is a modern BI interface with AI-assisted authoring and real-time visualizations

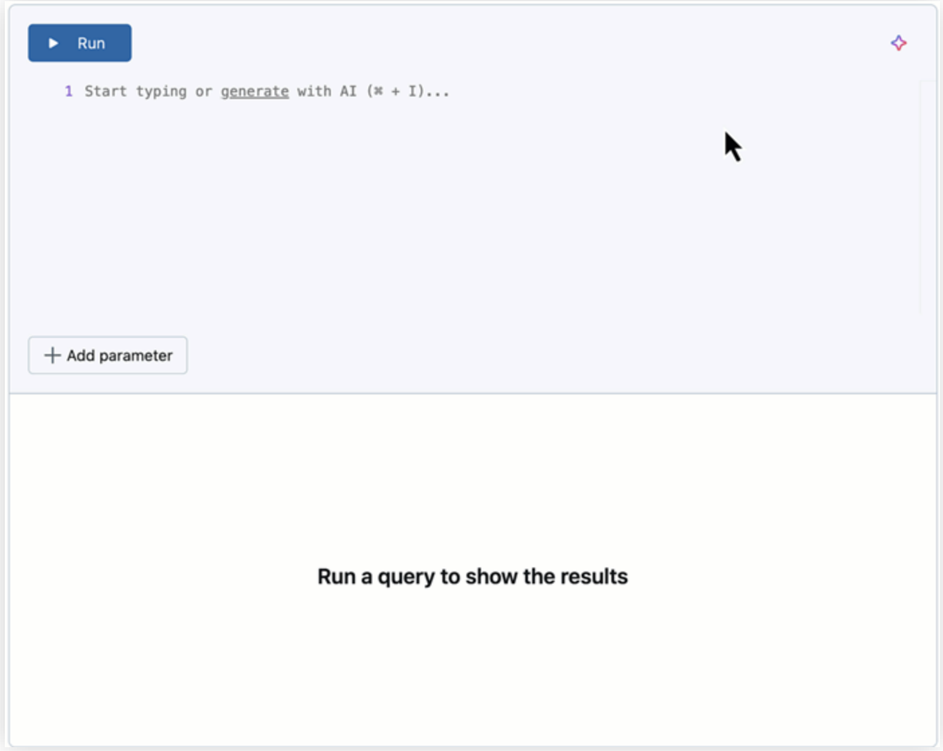
It is designed for collaboration and sharing insights with anyone in your organization

It is integrated with AI/BI Genie to allow users to explore data via conversational analytics

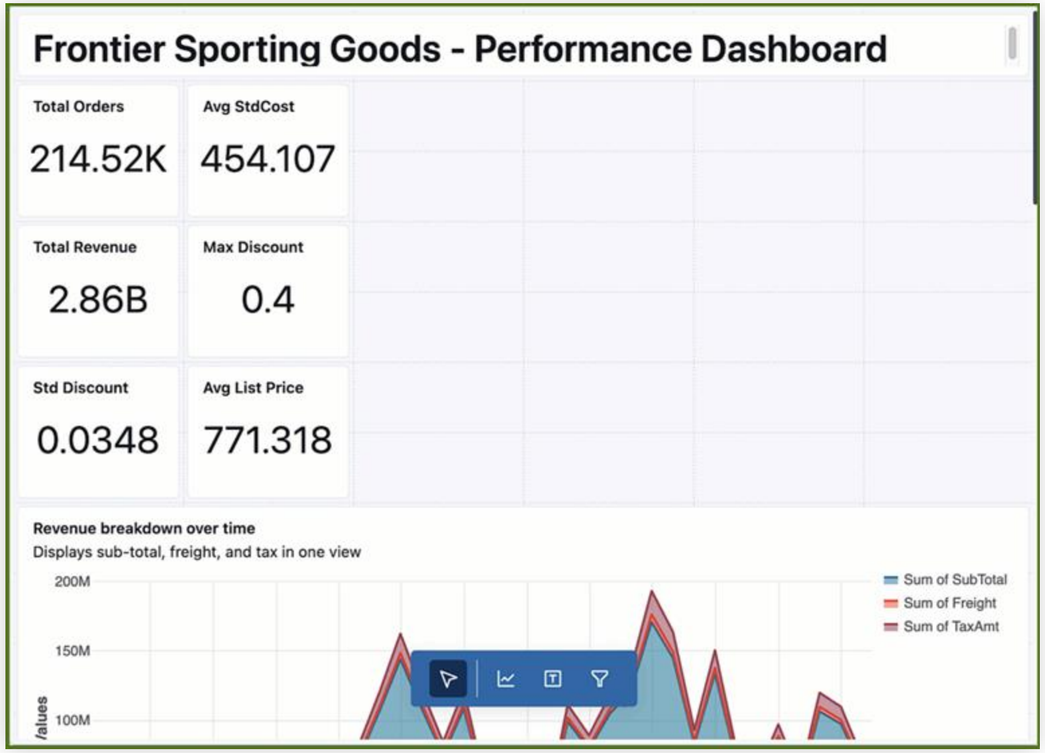


Databricks AI/BI Dashboards offer developers AI assistance

With defining data sets



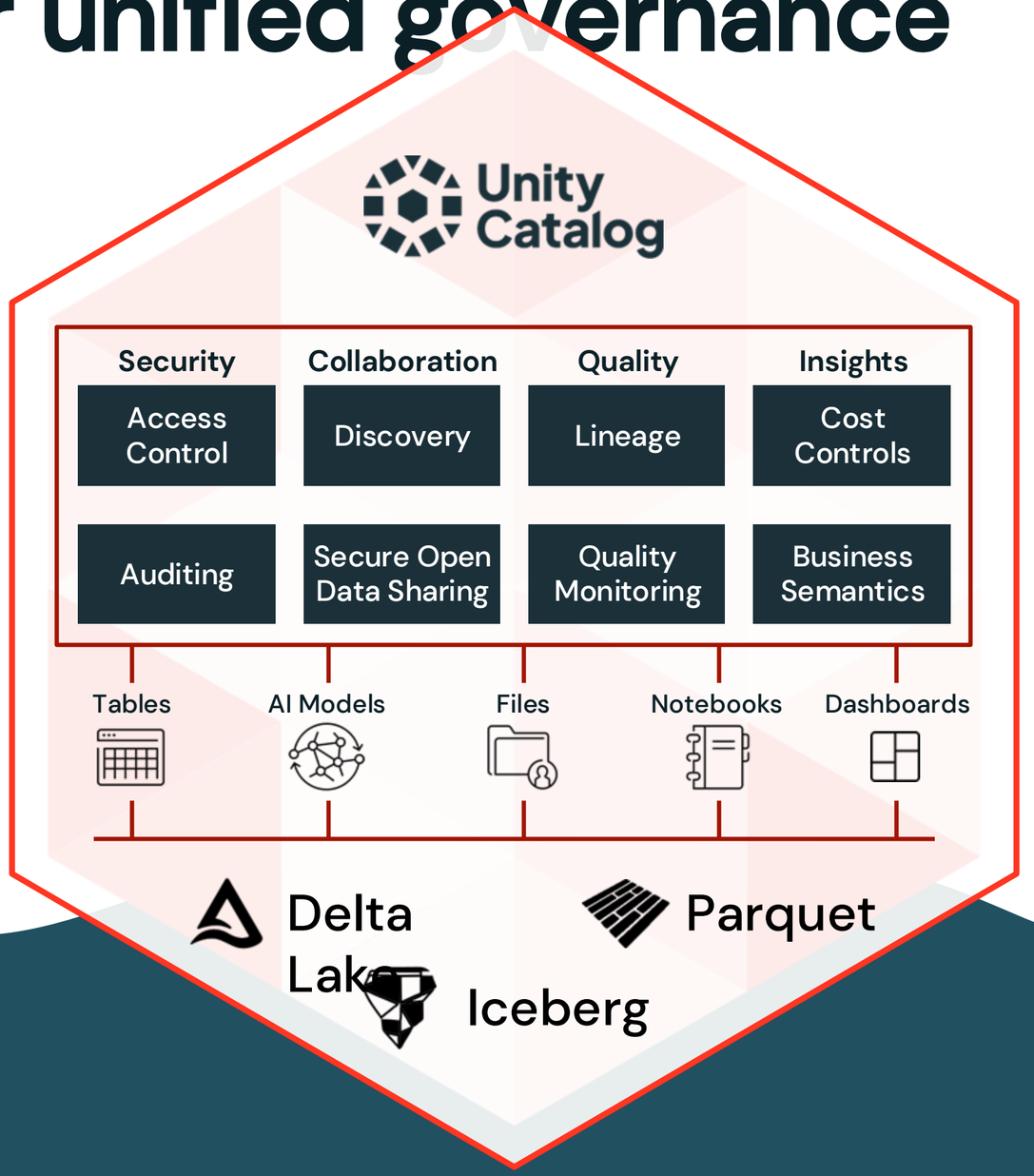
And with making visualizations



Built on Unity Catalog for unified governance

Unified capabilities for **every use case**

Unified governance for **all assets**

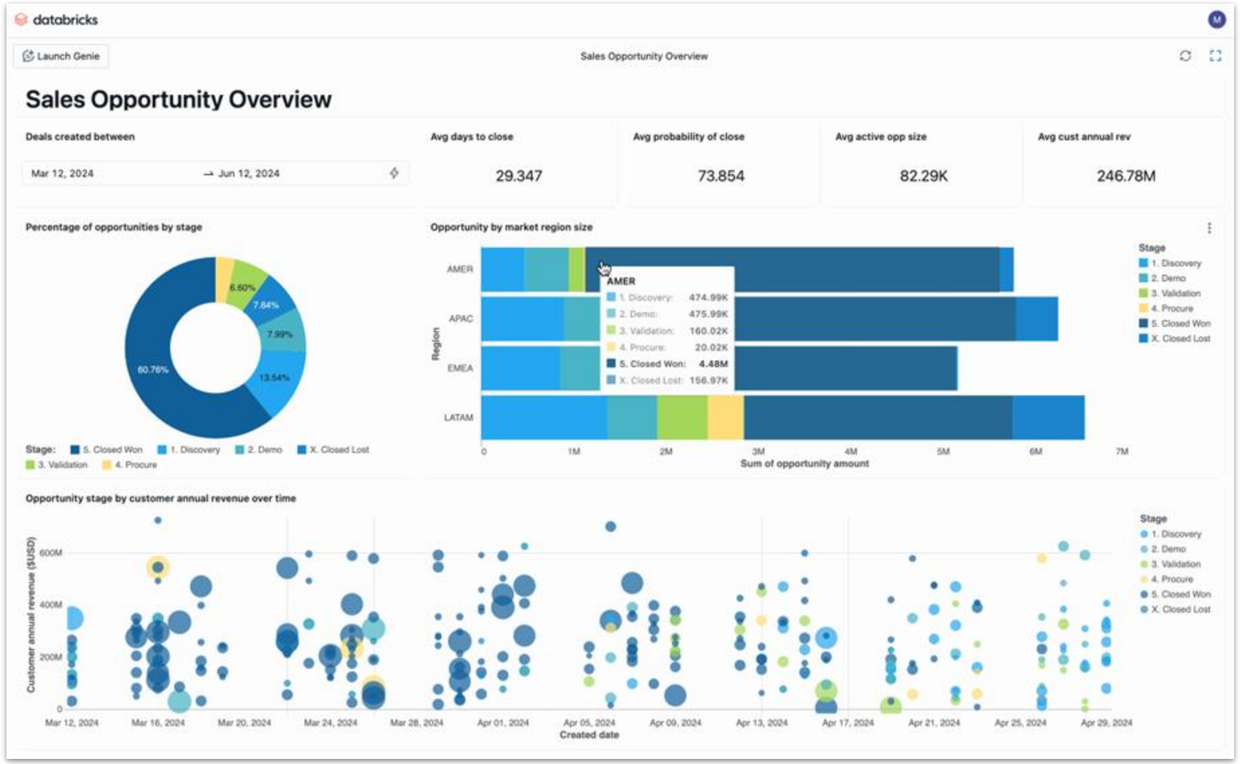


Dashboards offer instant, interactive insights at scale

Data

```
SELECT name,
national_id,
       country, fee_paid
FROM fee_transactions
WHERE country IN
       ('US','CA','MX');
```

Queries



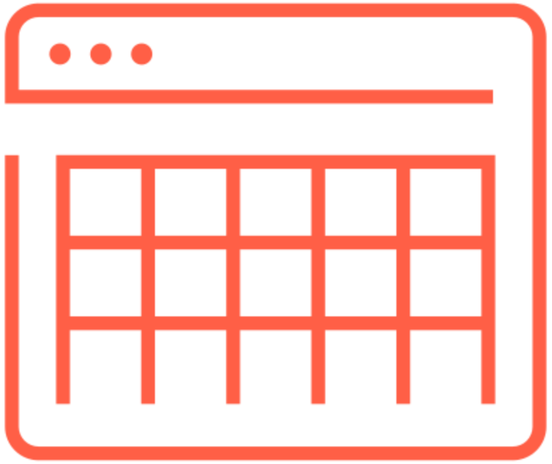
Visualization

AI

What do you need to create a dashboard?

A business purpose

A target audience



Data

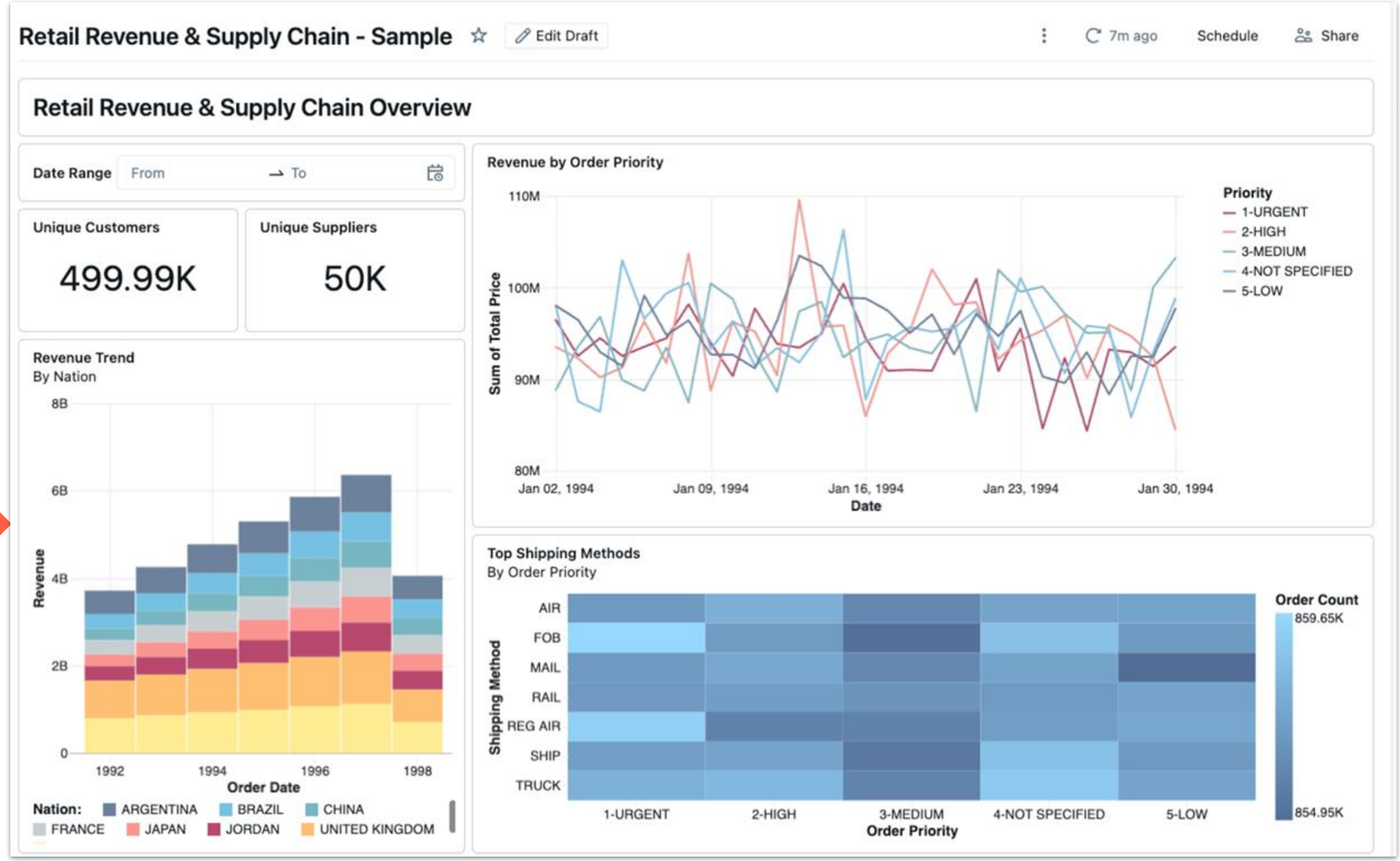
What makes up a dashboard?

Title

Filters

Counters

Stacked bar chart

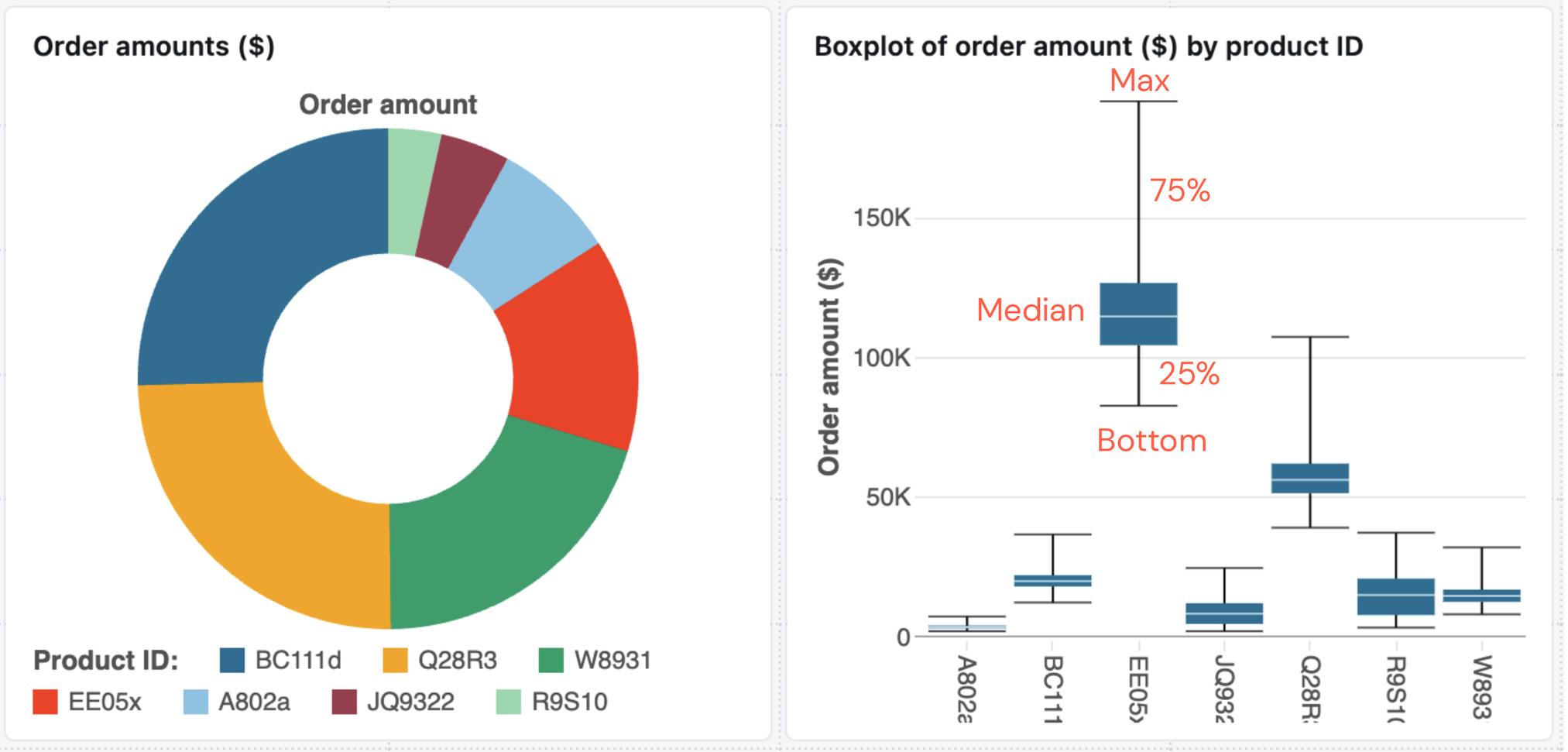


Text Box

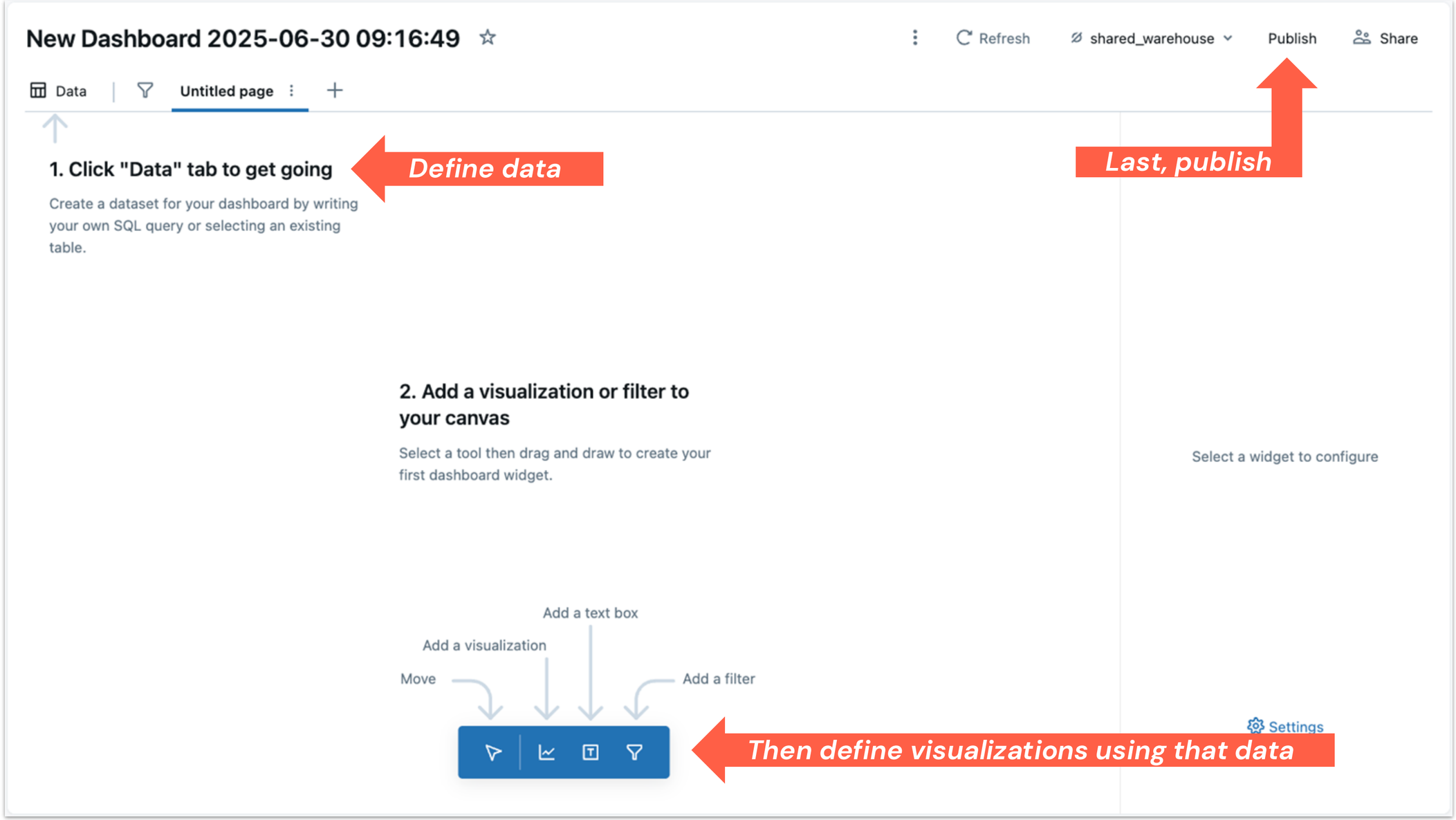
Line chart

Heat map

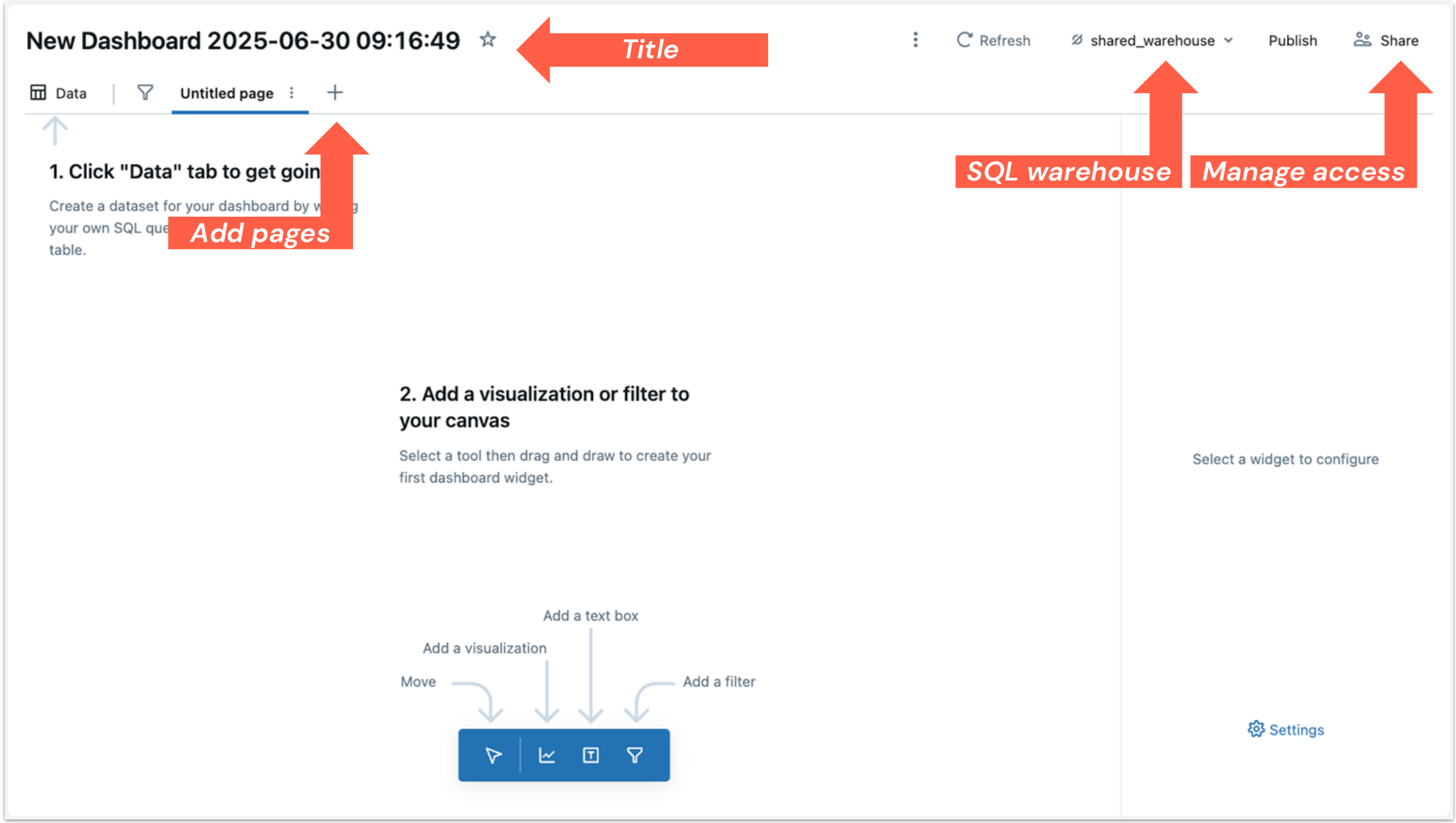
Dashboards can show summary statistics



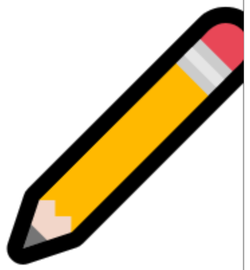
It all starts with Create Dashboard



What options are available in the UI?



Permissions: what can you do with a dashboard?

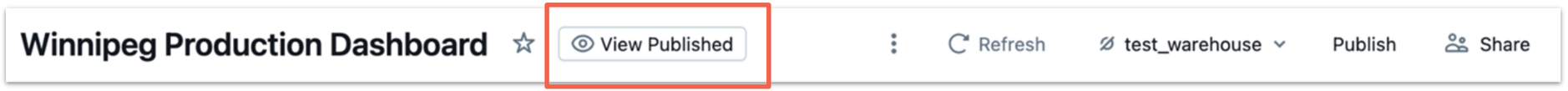
CAN VIEW or **CAN RUN**		✓ Can look at and refresh the dashboard ✓ Can click on widgets in the dashboard and interact with them ✓ Can clone the dashboard
CAN EDIT		All the above, plus ✓ Can edit the dashboard ✓ Can publish the dashboard
CAN MANAGE		All the above, plus ✓ Can change users' permissions on the dashboard ✓ Can delete the dashboard

Your dashboard's publication cycle

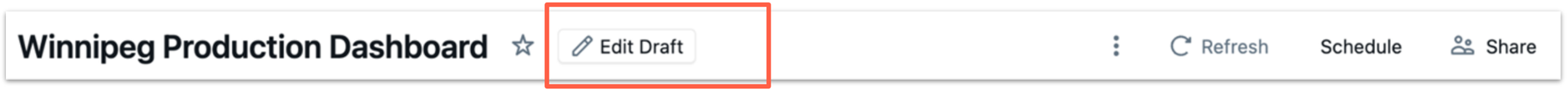
First creation

After naming

Published (Edit View)

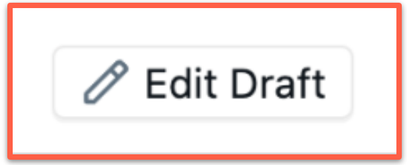


Published

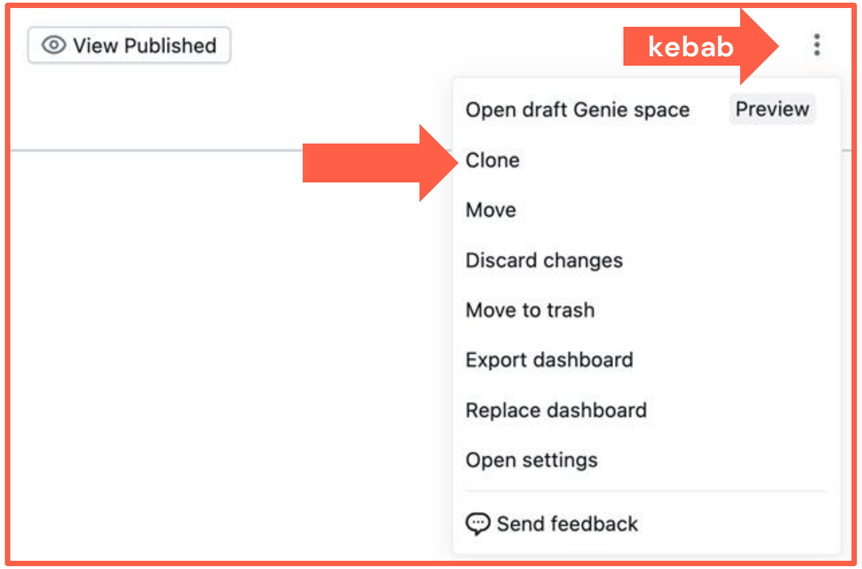


How can you make a copy of an existing dashboard?

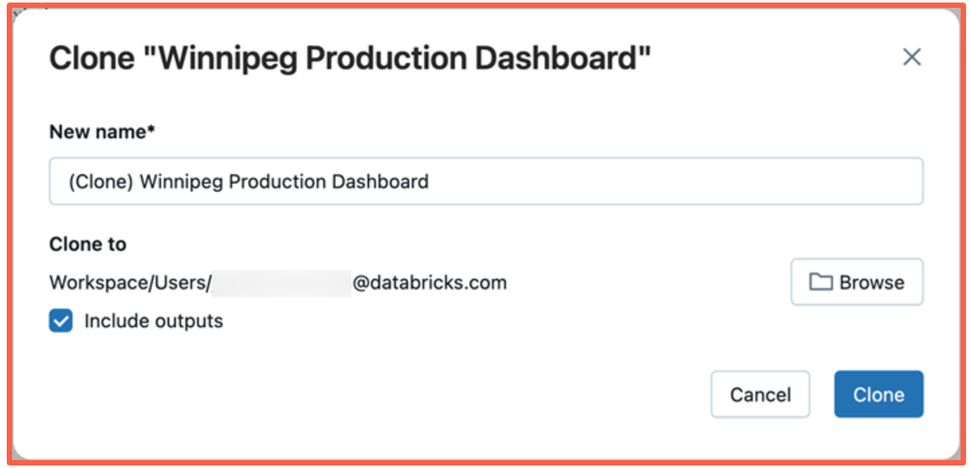
Move to draft mode



Choose Clone from the kebab menu



Specify new name and location



Caution: **Dashboards are subject to the inherited permissions of the locations they are stored**

Publishing and Sharing an AI/BI Dashboard

Managing the dashboard using Git/Version control

Initial commit

- Move the dashboard into a Git folder in the workspace
- Commit and push to the remote Git repository

Development

- Developers create Git folders on separate dev branches, typically in their home folders
- Commit changes to the development branch
- Merge development branches to main using pull requests

Deployment

- Create a Git folder on the deployment branch in a non-user top-level folder. See CI/CD techniques with Git and Databricks Git folders (Repos)
- Pull changes into the deployment folder
- Publish dashboards from this folder
- Remove edit access and restrict updates to Git
- Share dashboards with consumers

databricks

Dashboards and Visualizations in Databricks

LECTURE

Just enough SQL

Databricks SQL lets you define custom datasets

- Can filter data values to focus on subsets of the data
- Can bridge together related tables
- SQL queries are stored as part of the dashboard's infrastructure
- Are refreshed on demand and on defined schedules

SQL queries are especially useful for lookups

OrderID	Product	Qty	CustID
B2104	Avocado	41	8ad2750b
B2105	Grapefruit	84	451e98f5
B2106	Cabbage	8	8ad2750b
B2107	Avocado	55	bb4a23a2
B2108	Pear	50	9302b293
more...			

Transactions table

CustID	Name	Province	*more...*
451e98f5	Ted	ON	
bb4a23a2	Emma	BC	
8ad2750b	Maya	AB	
9302b293	Jo	ON	
3aa9402e	Yumiko	QC	
more...			

Customer table

The simplest kind of query

```
SELECT *
    FROM fruteria.sales.transactions ;
```

Add a WHERE clause to focus on relevant rows

```
SELECT *
       FROM fruteria.sales.transactions
       WHERE franchiseID = "3000047" ;
```

Add a WHERE clause to focus on relevant rows

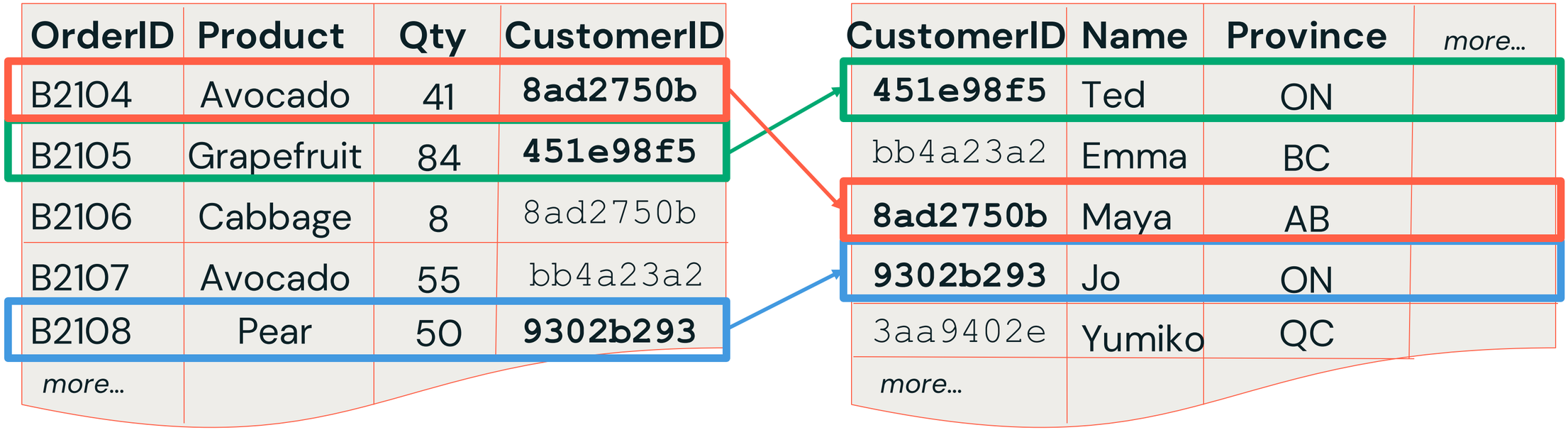
```
SELECT *
       FROM fruteria.sales.transactions
       WHERE franchiseID IN
             ("3000047","3000022","3000093") ;
```

Name only relevant fields

```
SELECT product, quantity
       FROM fruteria.sales.transactions
       WHERE franchiseID = "3000047" ;
```

But what if we want to bring together fields from more than one table? (JOINS)

- ❏ Must specify the table each desired field comes from
- ❏ Must tell the query how to tie the tables together



OrderID	Product	Qty	CustomerID
B2104	Avocado	41	**8ad2750b**
B2105	Grapefruit	84	**451e98f5**
B2106	Cabbage	8	8ad2750b
B2107	Avocado	55	bb4a23a2
B2108	Pear	50	**9302b293**
more...			

Transactions table

CustomerID	Name	Province	*more...*
451e98f5	Ted	ON	
bb4a23a2	Emma	BC	
8ad2750b	Maya	AB	
9302b293	Jo	ON	
3aa9402e	Yumiko	QC	
more...			

Customer table

First, let's name the two tables…

```
SELECT product, quantity
       FROM fruteria.sales.transactions
            AS t
       JOIN fruteria.sales.customers
                  AS c
     [...] ;
```

Next let's specify the linking condition…

```
SELECT product, quantity
       FROM fruteria.sales.transactions
            AS t
       JOIN fruteria.sales.customers
                  AS c
       ON t.customerID = c.customerID ;
```

Now we can use fields from both tables

```
SELECT t.product, t.quantity, c.province
       FROM fruteria.sales.transactions
            AS t
       JOIN fruteria.sales.customers
                  AS c
       ON t.customerID = c.customerID ;
```

And we can further filter

```
SELECT t.product, t.quantity, c.province
       FROM fruteria.sales.transactions
            AS t
       JOIN fruteria.sales.customers
                  AS c
       ON t.customerID = c.customerID
       WHERE c.province IN ("ON", "BC") ;
```

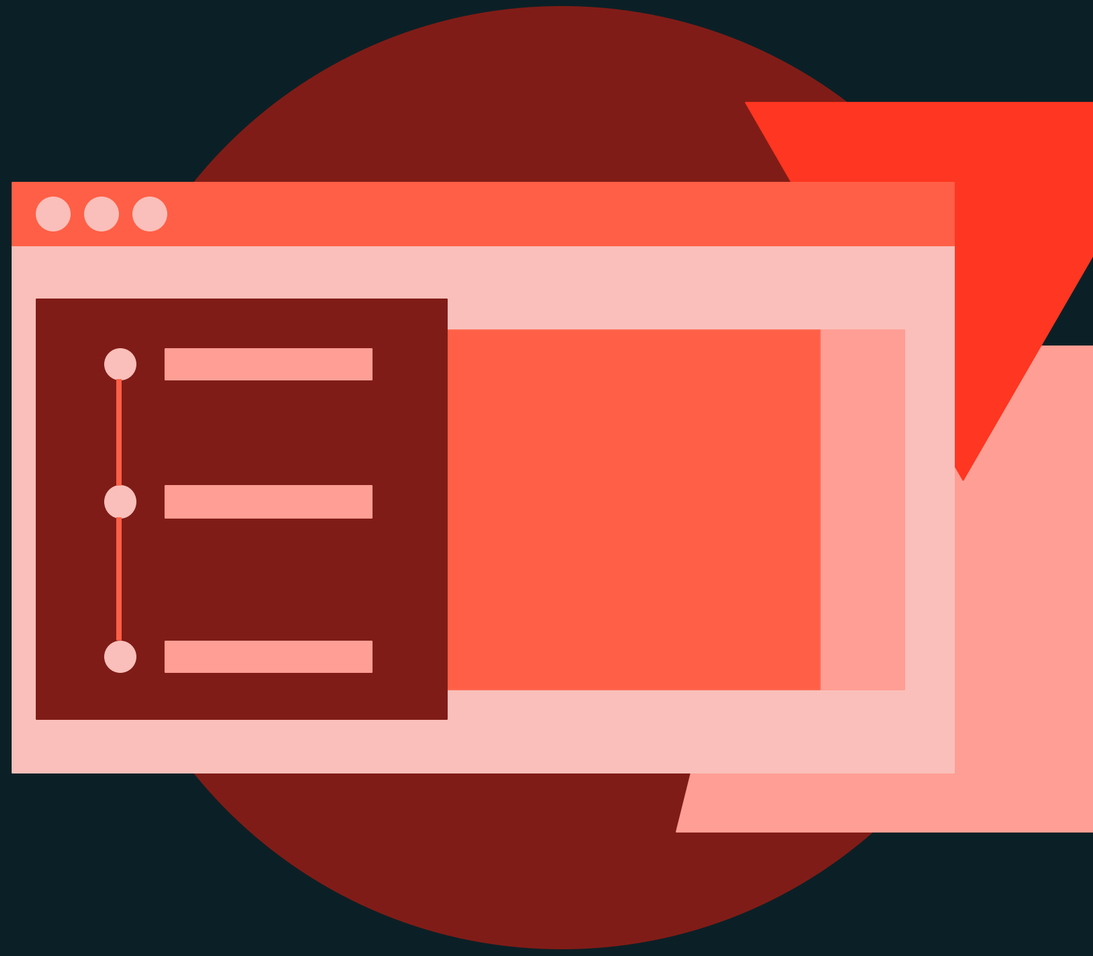
And we can further filter

```
SELECT t.product, t.quantity, c.province
       FROM fruteria.sales.transactions
            AS t
       JOIN fruteria.sales.customers
                  AS c
       ON t.customerID = c.customerID
       WHERE c.province IN ("ON", "BC")
                  AND
             t.product = "Grapefruit" ;
```

Dashboards and Visualizations in Databricks

DEMONSTRATION

Designing Datasets for Dashboards



Follow-along Instructions

Estimated Time: 20 minutes

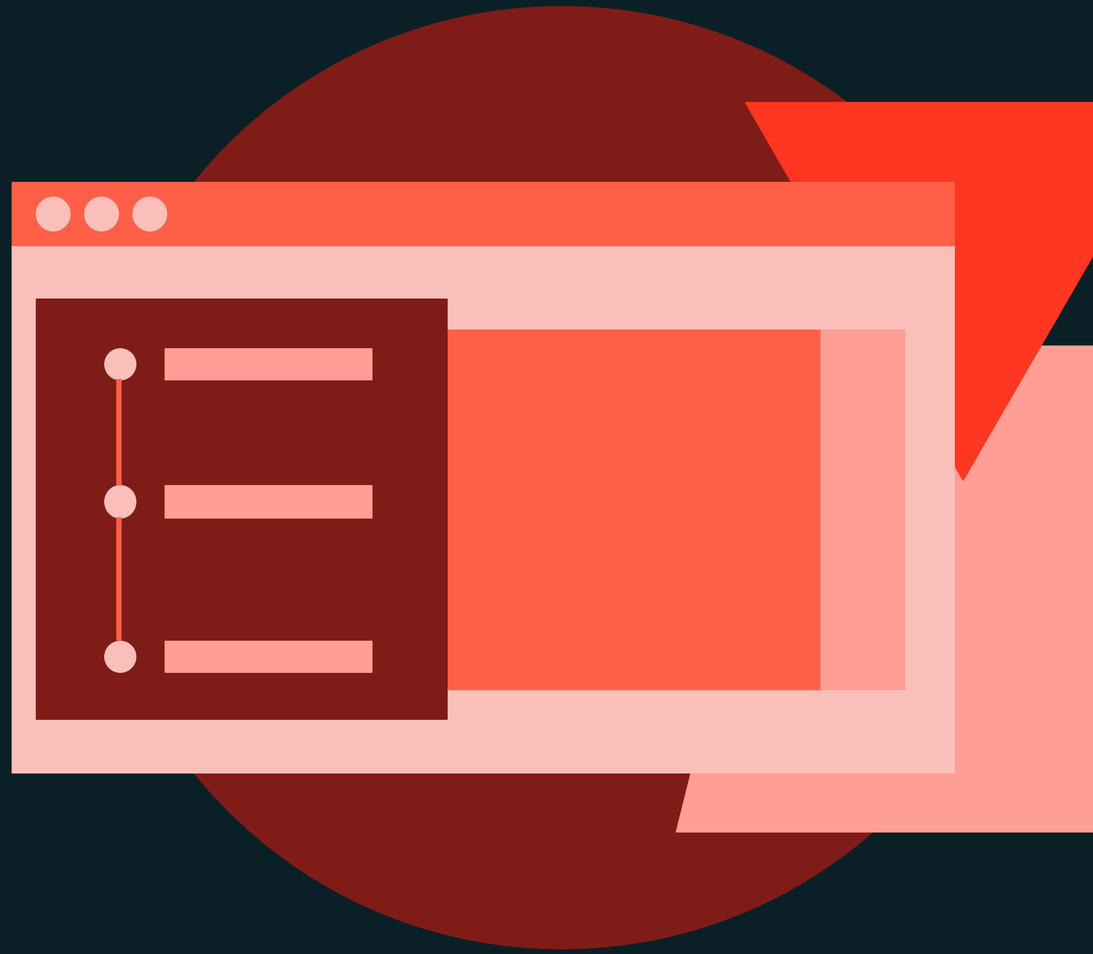
For this demonstration, the instructor will walk you through the process of creating or selecting data assets for AI/BI Dashboards. You will also learn the process for creating a AI/BI Dashboard within the platform.

If you have access to the Vocareum lab environment, feel free to follow-along.

Dashboards and Visualizations in Databricks

DEMONSTRATION

Creating Visualizations and adding Summary Statistics to Dashboards



Follow-along Instructions

Estimated Time: 20 minutes

For this demonstration, the instructor will show you how to add widgets to your dashboard for various purposes including for the display of text and graphics, numerical values, and illustrative charts.

If you have access to the Vocareum lab environment, feel free to follow-along.

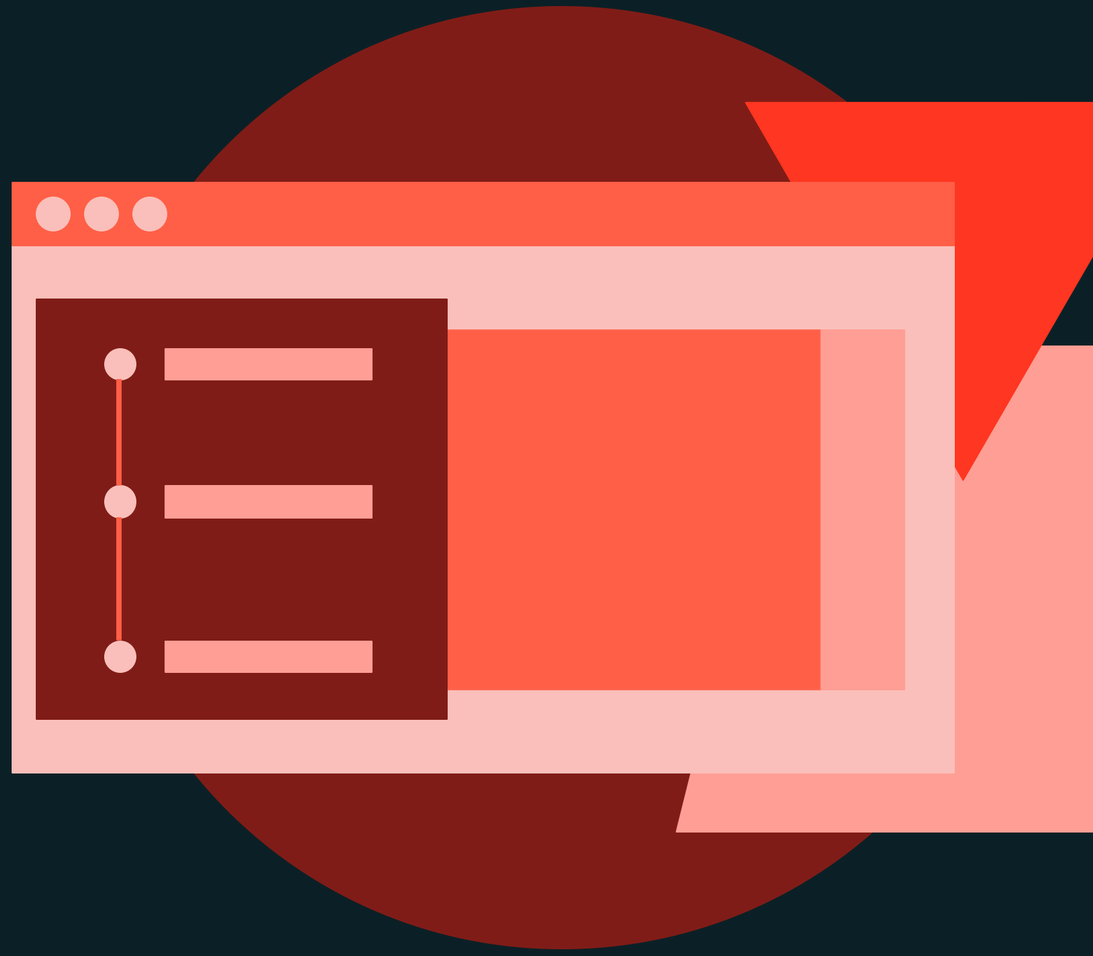
Note: The steps in this demonstration require you to have completed the steps in the *Designing Datasets for Dashboards* demonstration.

databricks

Dashboards and Visualizations in Databricks

DEMONSTRATION

AI Enhanced Features



Follow-along Instructions

Estimated Time: 8 minutes

For this demonstration, the instructor will walk you through how to use Databricks' AI features to create visualizations and support your work through the use of natural language inputs.

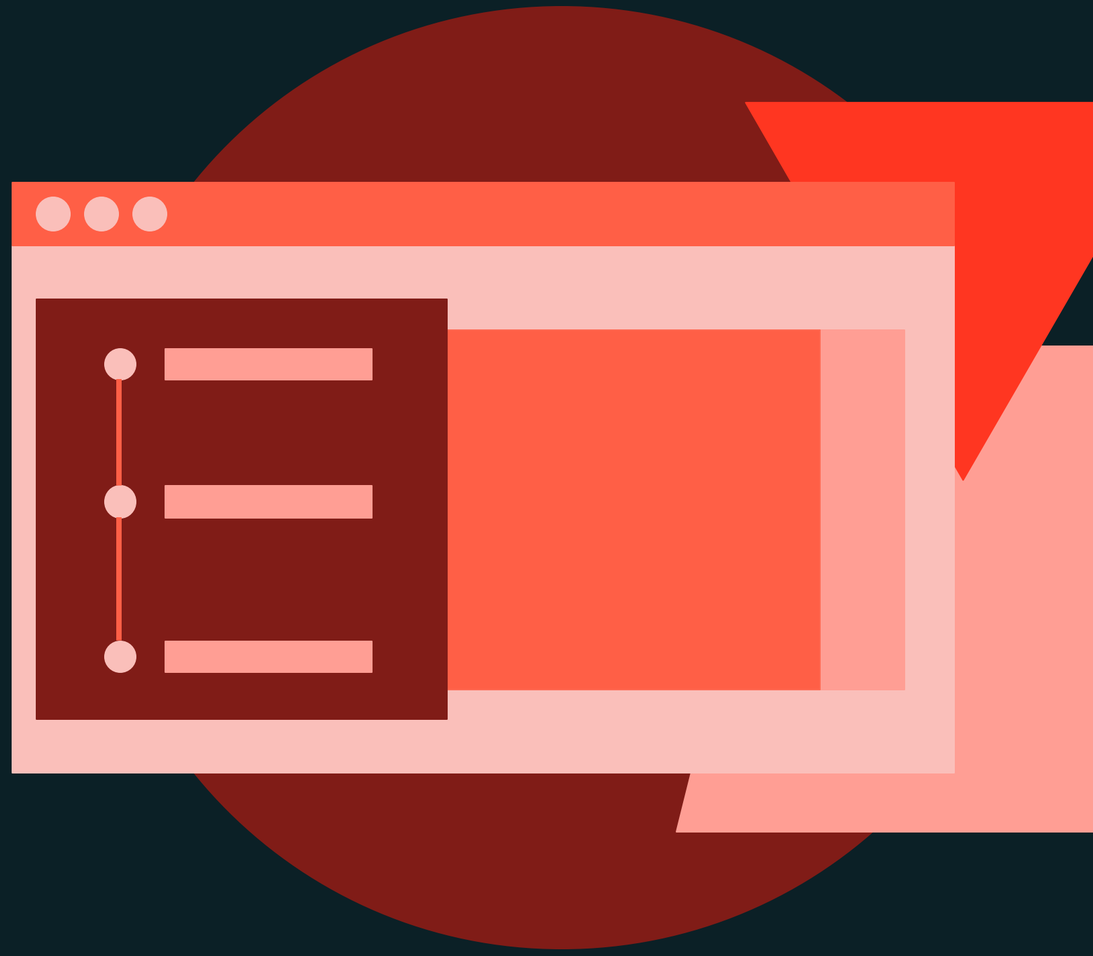
If you have access to the Vocareum lab environment, feel free to follow-along.

databricks

Dashboards and Visualizations in Databricks

DEMONSTRATION

Filters



Follow-along Instructions

Estimated Time: 8 minutes

For this demonstration, the instructor will demonstrate how to add dynamic features to dashboards, such as filters. You'll also learn how to use cross filtering to drill down on multiple visualizations simultaneously.

If you have access to the Vocareum lab environment, feel free to follow-along.

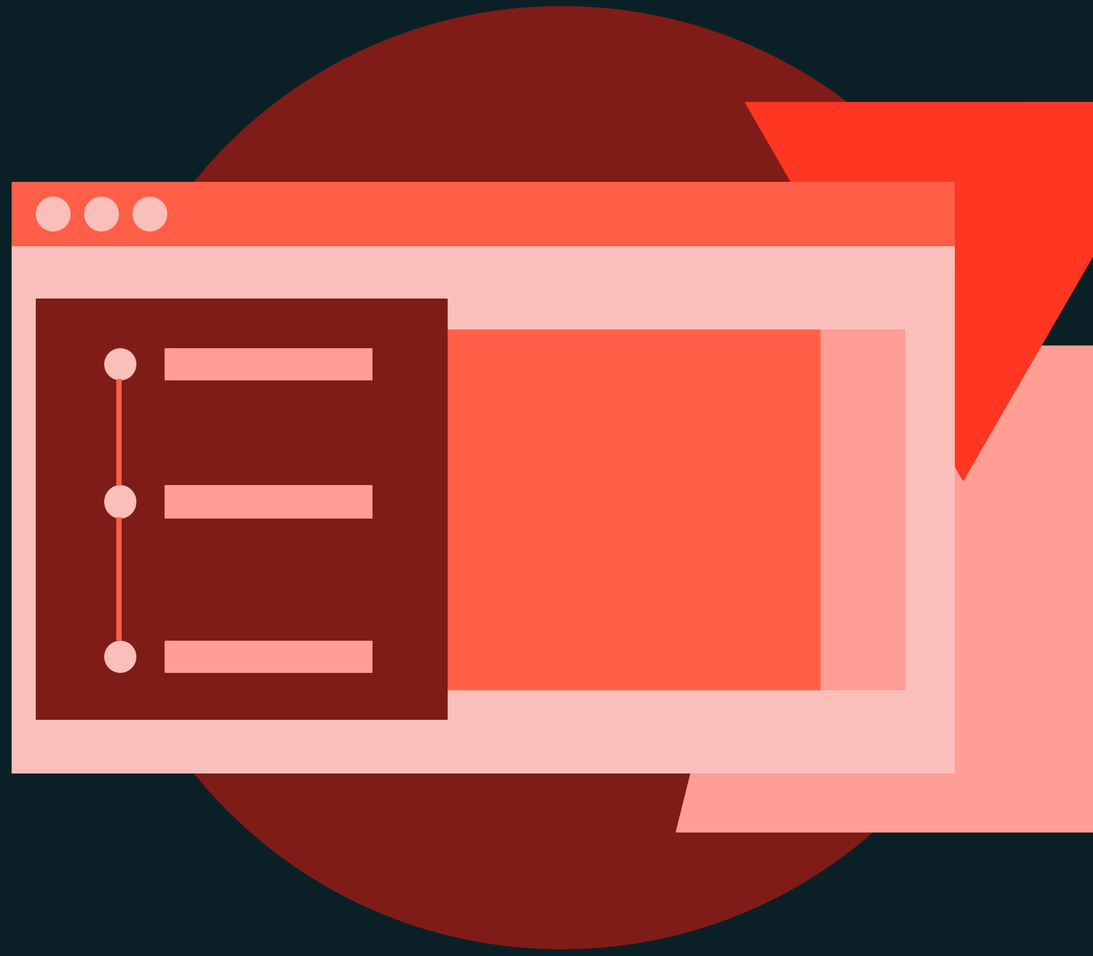
Note: The steps in this demonstration require you to have completed the steps in the *Creating Visualizations and adding Summary Statistics to Dashboards* demonstration.

databricks

Dashboards and Visualizations in Databricks

DEMONSTRATION

Sharing Dashboards with Stakeholders and Others



Follow-along Instructions

Estimated Time: 8 minutes

For this demonstration, the instructor will walk you through how to securely share and publish dashboards for others to view and use. You'll see how to adjust the permissions settings for your dashboards for appropriate access.

If you have access to the Vocareum lab environment, feel free to follow-along.

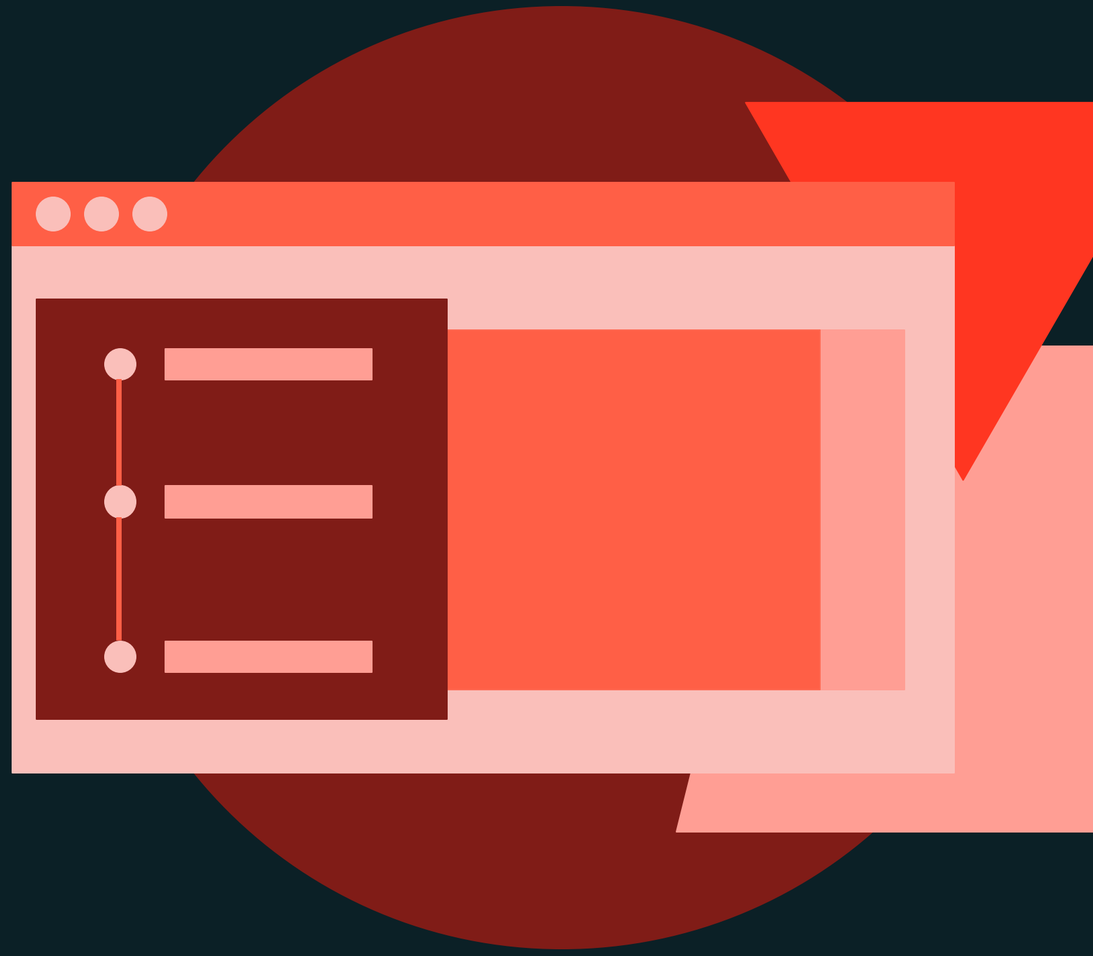
Note: The steps in this demonstration require you to have completed the steps in the *Filters* demonstration.

databricks

Dashboards and Visualizations in Databricks

DEMONSTRATION

Managing Dashboards in Production



Follow-along Instructions

Estimated Time: 8 minutes

For this demonstration, the instructor will demonstrate how to update dashboards once they've been published.

If you have access to the Vocareum lab environment, feel free to follow-along.

Note: The steps in this demonstration require you to have completed the steps in the *Sharing Dashboards with Stakeholders and Others* demonstration.

databricks

Dashboards and Visualizations in Databricks

LAB EXERCISE

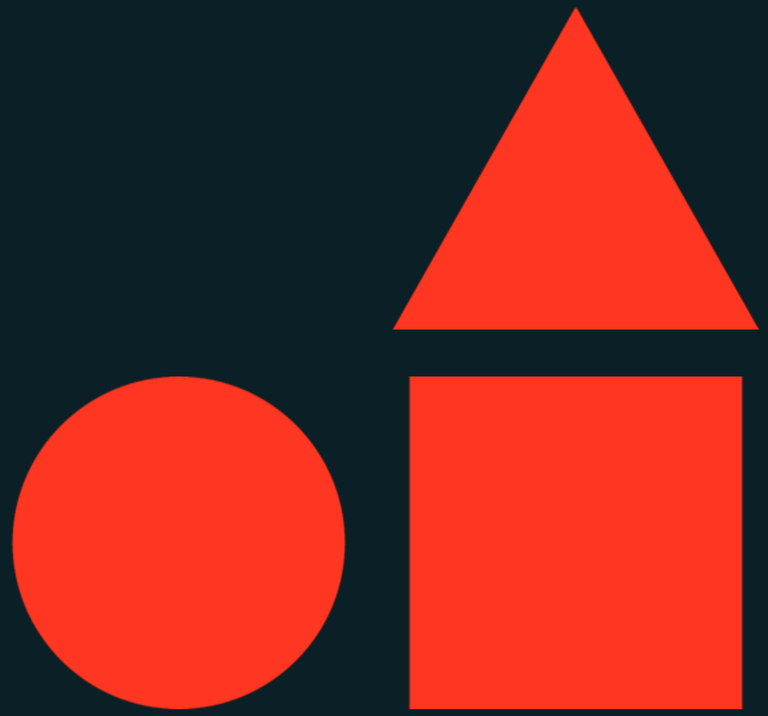
Dashboard and Visualization Lab Activity

What’s in this lab activity?

- Add Datasets to a Dashboard
- Add several visualizations to a dashboard.
- Add multiple tabs to a dashboard
- Add additional (non-visualization) widgets to a dashboard.
- Add filters to a dashboard.
- Establish Refresh schedules.
- Share Dashboards securely.

databricks

AI/BI Genie



AI/BI for Data Analysts

Agenda

AI/BI Genie	Time	Lecture	Demo	Activity
AI/BI Genie	10 mins	✓		
Developing Genie Spaces	15 mins		✓	
Sharing Genie Spaces	5 mins		✓	
Maintaining Genie Spaces	10 mins		✓	
AI/BI Genie Space Development Activity Lab	30 mins			✓

Learning Objectives

Create data assets for self-service analytics.

- Establish AI/BI Genie spaces associated with a dashboard.
- Establish stand-alone AI/BI Genie spaces.
- Share AI/BI Genie spaces with stakeholders.
- Solicit feedback from stakeholders and non-technical users.

Manage the data assets for business intelligence in Databricks.

- Make edits to the settings of an AI/BI Genie Space.
- Review and respond to feedback on AI/BI Genie Spaces from stakeholders and business users.

databricks

AI/BI Genie

LECTURE

AI/BI Genie

Genie lets users converse with data in natural language

How does Genie use AI?

HUMAN NATURAL LANGUAGE ENTRY

Chat responses

```
SELECT name, national_id,
       country, fee_paid
FROM fee_transactions
WHERE country IN
       ('US','CA','MX');
```

How AI/BI Genie understands the business

Dashboards and data

Curated instructions

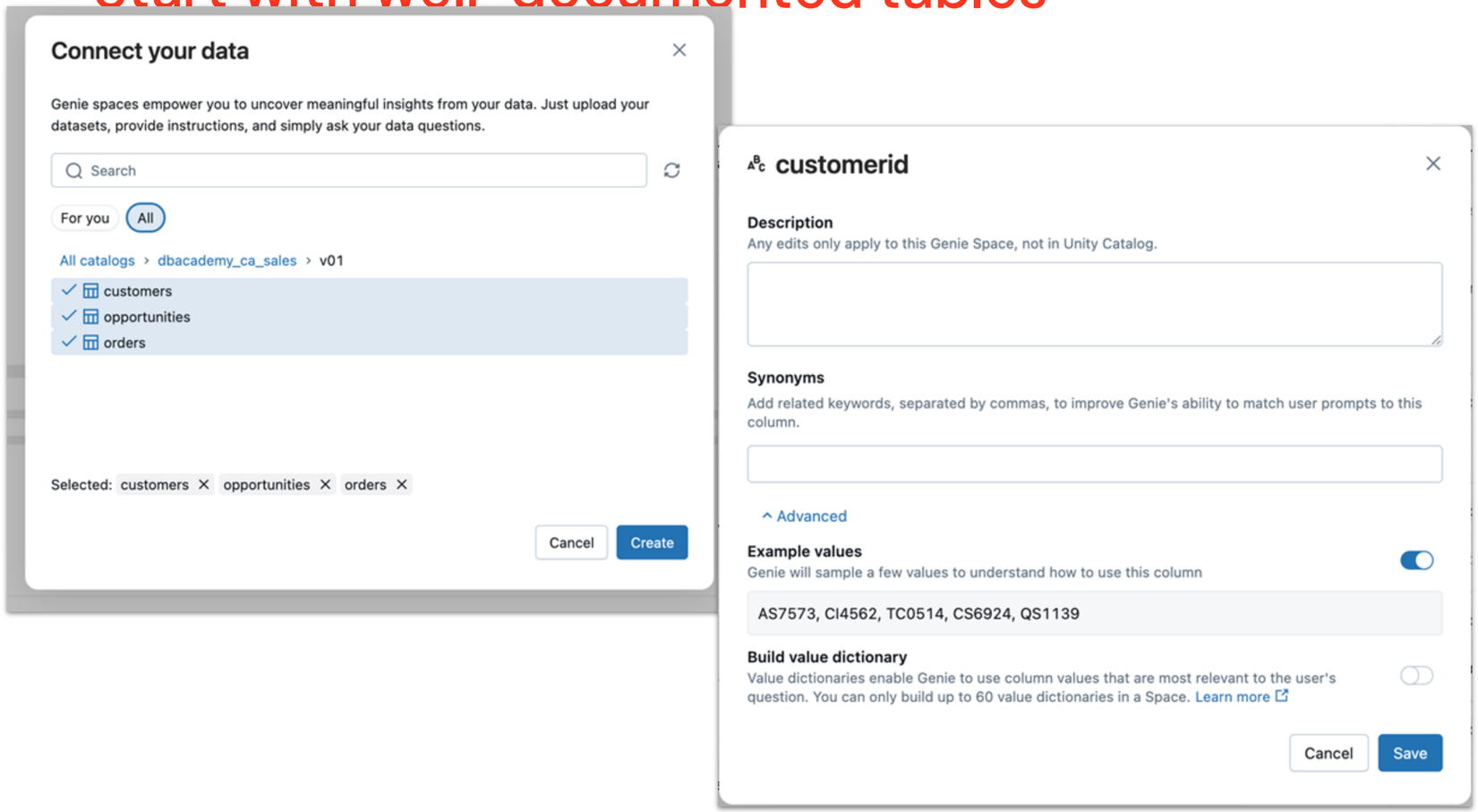
Genie space

Unity Catalog

User feedback

Building your Genie Space

Start with well-documented tables

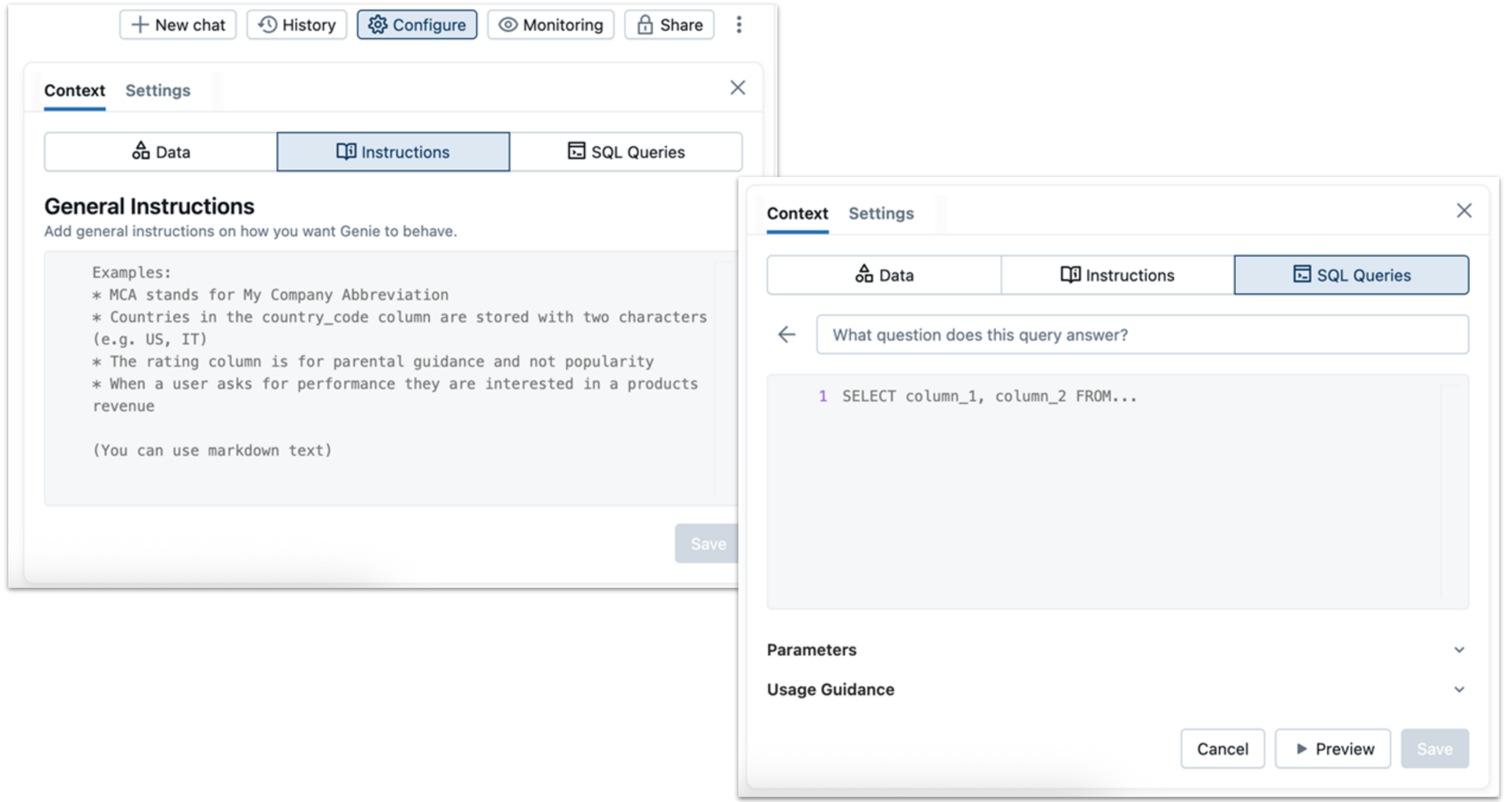


- Begin with UC tables/views and add meaningful comments to columns and table descriptions for clarity
- Clearly define primary keys and foreign keys for better context understanding

Building your Genie Space

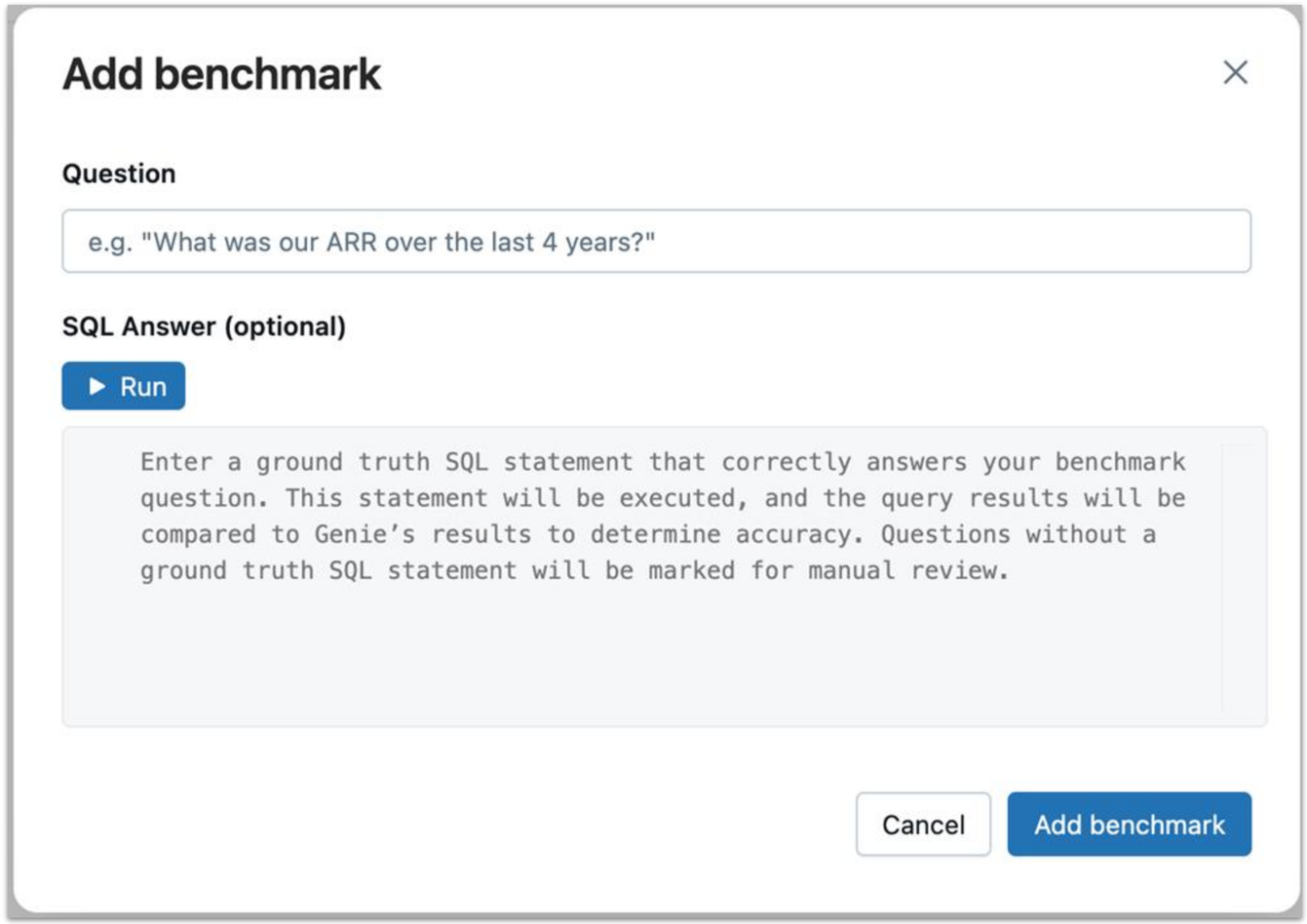
Genie learns from you

- Include relevant instructions that help Genie understand unique jargon, logic or business concepts
- Save as instruction → verified SQL queries can be saved, teaching Genie to provide more accurate answers in the future

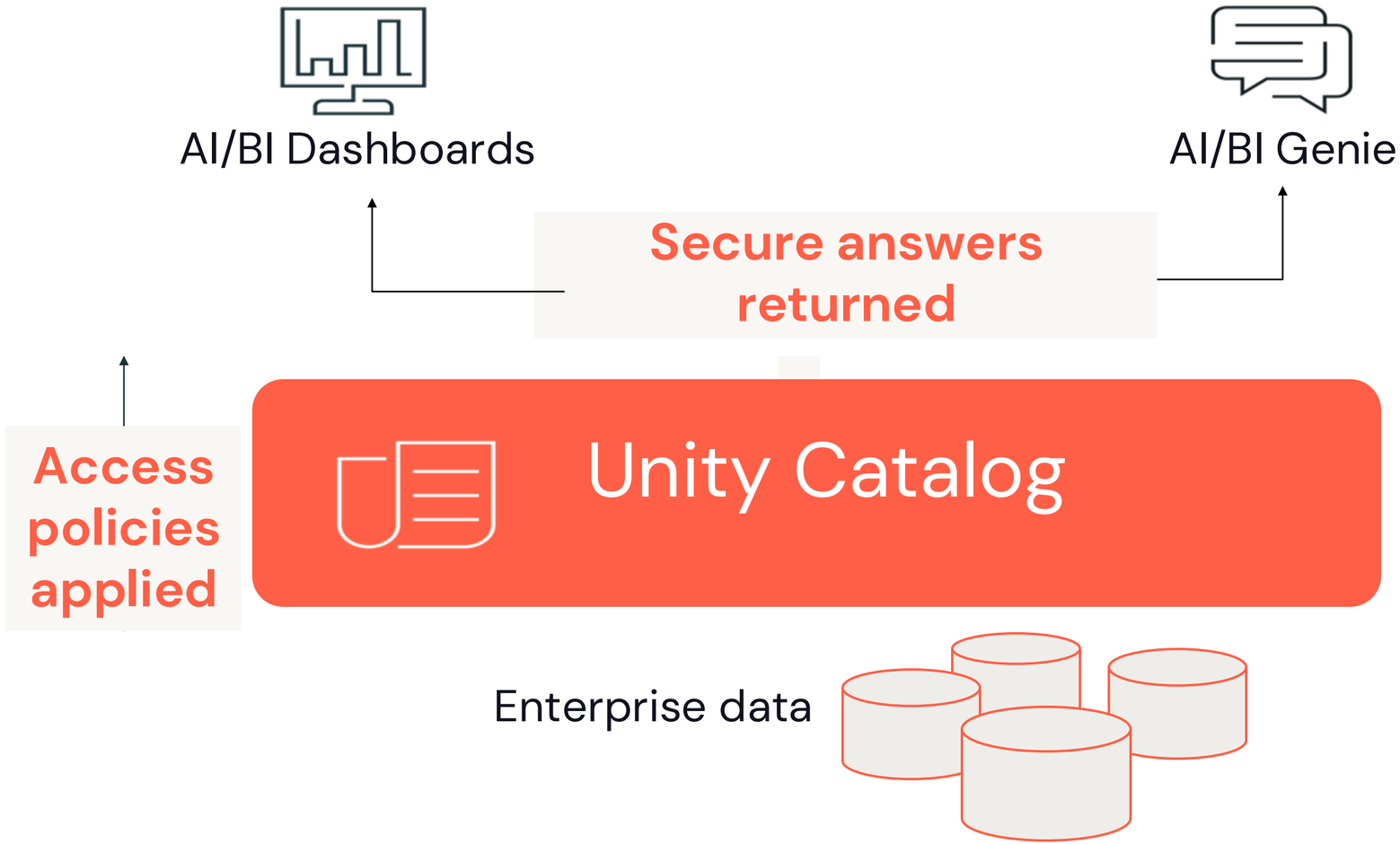


Using Genie Benchmarks

- Benchmarks are used for evaluation, not training
- Set of standardized questions with supplied SQL to answer the question.
- Running a benchmark provides three responses:
  - Good
  - Bad
  - Requires Manual Review



Genie offers governed and secure insights



What are your responsibilities for using AI/BI Genie with stakeholders and business users?

❑ Carefully curate the base datasets

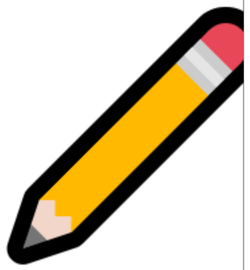
❑ Provide reasonable sample queries

❑ Mark vetted queries as Trusted Assets

❑ Respond to user feedback and add instructions

❑ Ensure that the data is fresh

Permissions: what can you do with a Genie space?

CAN VIEW or **CAN RUN**		✓ Can ask Genie questions and get answers ✓ Can give feedback on answers 👍 👎
CAN EDIT		All the above, plus ✓ Can add or edit saved instructions ✓ Can add or edit sample questions ✓ Can bring in or throw out included data tables
CAN MANAGE		All the above, plus ✓ Can see all the questions users ask in the space, and the answers ✓ Can see all the feedback ✓ Can change the space's permissions ✓ Can delete the space

Governing your Genie Space

Governed Genie space interactions

Unified security and governance

All interactions governed by UC's security policies and data access controls

RLS and CLS ensure data privacy

Trusted interactions

All Genie interactions inherit access policies

Users only see data they are authorized to access

Prepping for governance

Users must have:

SELECT permissions on UC tables, views, functions, etc.

CAN USE permissions on the SQL warehouse

Consider using a code-styled callout box for these permissions to make them pop

A few best practices when creating a Genie space

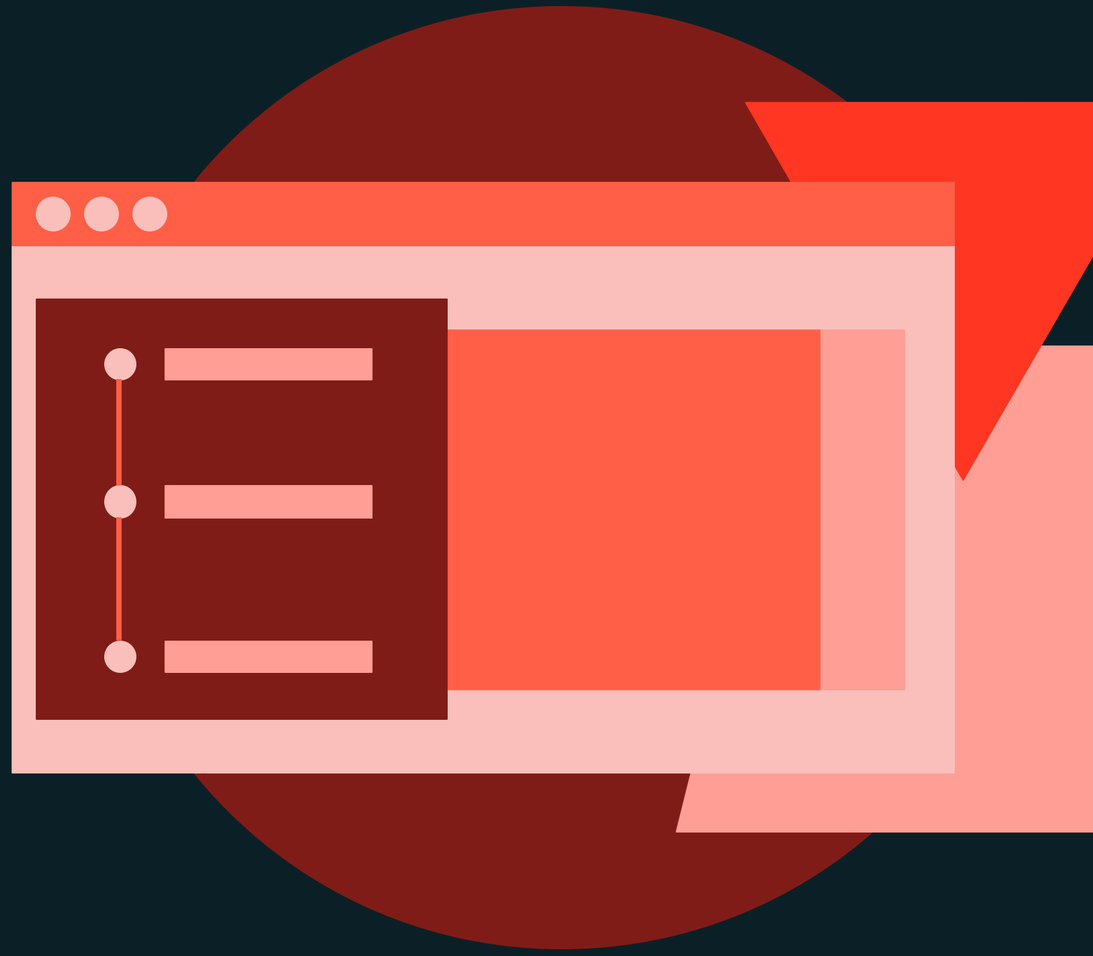
- What should users be able to answer?
- Encourage alignment with business needs
- Test the response quality
- Rephrase the questions; does the model still understand?
- Ask new, related questions; does Genie generalize well?
- Refine your instructions
- Iterate on prompts until answers match expectations
- Encourage fine-tuning based on user testing

databricks

AI/BI Genie

DEMONSTRATION

Developing Genie Spaces



Follow-along Instructions

Estimated Time: 15 minutes

For this demonstration, the instructor will show you how to create a Genie space, including how to configure some of the basic settings.

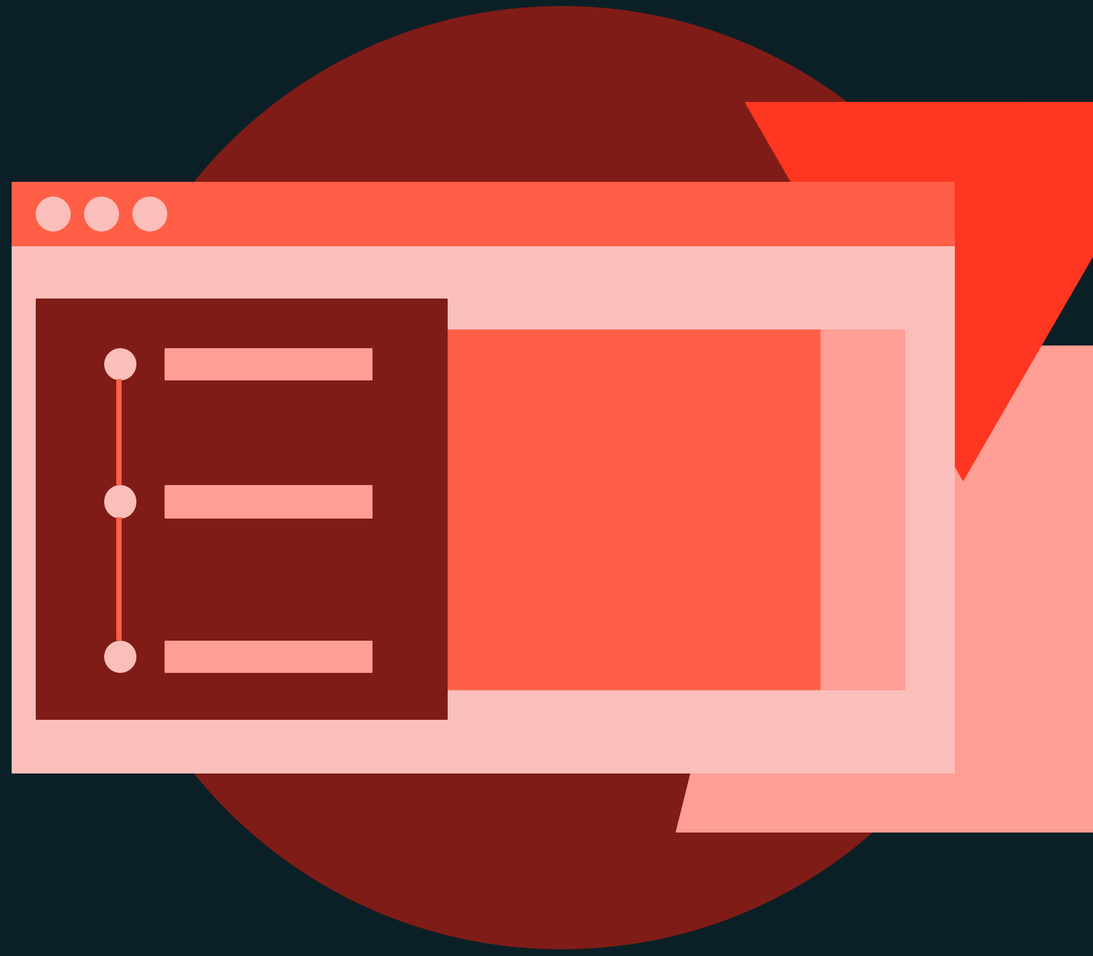
If you have access to the Vocareum lab environment, feel free to follow-along.

databricks

AI/BI Genie

DEMONSTRATION

Sharing Genie Spaces



Follow-along Instructions

Estimated Time: 5 minutes

For this demonstration, the instructor will demonstrate the process for sharing Genie spaces with others within your Databricks account.

If you have access to the Vocareum lab environment, feel free to follow-along.

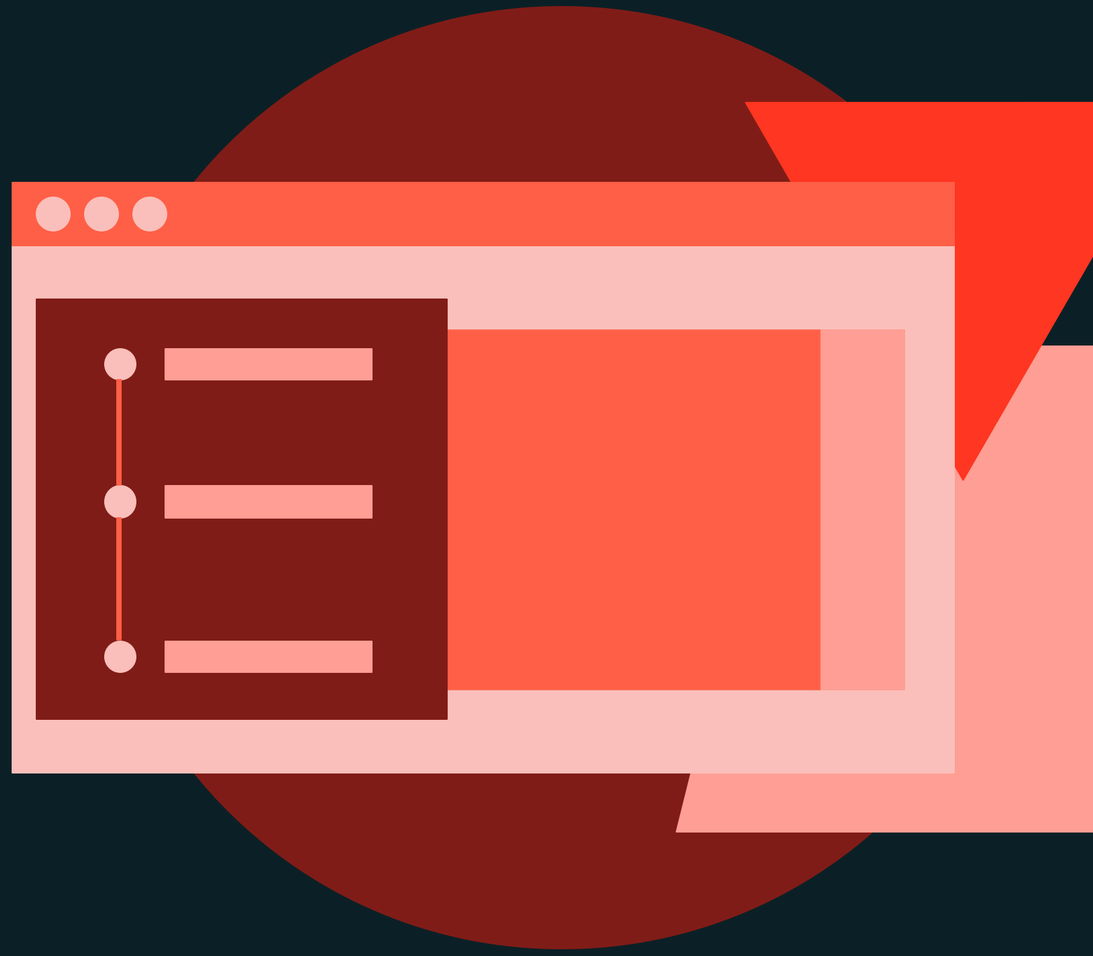
Note: The steps in this demonstration require you to have completed the steps in the *Developing Genie Spaces* demonstration.

databricks

AI/BI Genie

DEMONSTRATION

Maintaining Genie Spaces



Follow-along Instructions

Estimated Time: 10 minutes

For this demonstration, the instructor will demonstrate how to edit and maintain a Genie space based on feedback from end-users, including setting custom instructions, adding benchmark questions, and ensuring datasets are properly notated and described.

If you have access to the Vocareum lab environment, feel free to follow-along.

Note: The steps in this demonstration require you to have completed the steps in the *Sharing Genie Spaces* demonstration.

databricks

AI/BI Genie

LAB EXERCISE

AI/BI Genie Space Development Activity Lab

What’s in this lab activity?

Develop a Genie Space

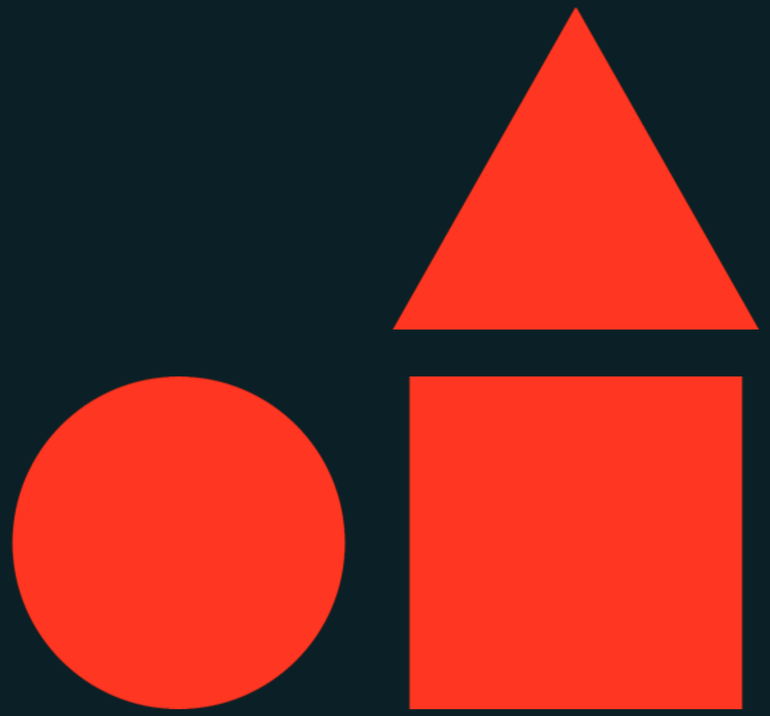
Manage settings and custom instructions.

Practice using a Genie Space as if you were an end user.

Review the feedback provided through the UI as a Genie Space Developer

databricks

Summary and Next Steps



AI/BI for Data Analysts

Course Learning Objectives Recap

In this course, you learned how to use Databricks AI/BI tools to:

- Design dashboards for business insights.
- Share business intelligence assets with collaborators and stakeholders.
- Periodically revise data assets in accordance with best practices and new information.
- Create data assets for self-service analytics.
- Manage the data assets for business intelligence in Databricks.

Earn a Databricks certification!

Certification helps you gain industry recognition, competitive differentiation, greater productivity, and results.

- This course helps you prepare for the Databricks **Certified Data Analyst Associate exam.**
- Recommended Self-Paced Courses
  - Get Started with SQL Analytics and BI on Databricks
  - SQL Analytics on Databricks
- Please see the Databricks Academy for additional prep materials.

For more information visit:
databricks.com/learn/certification