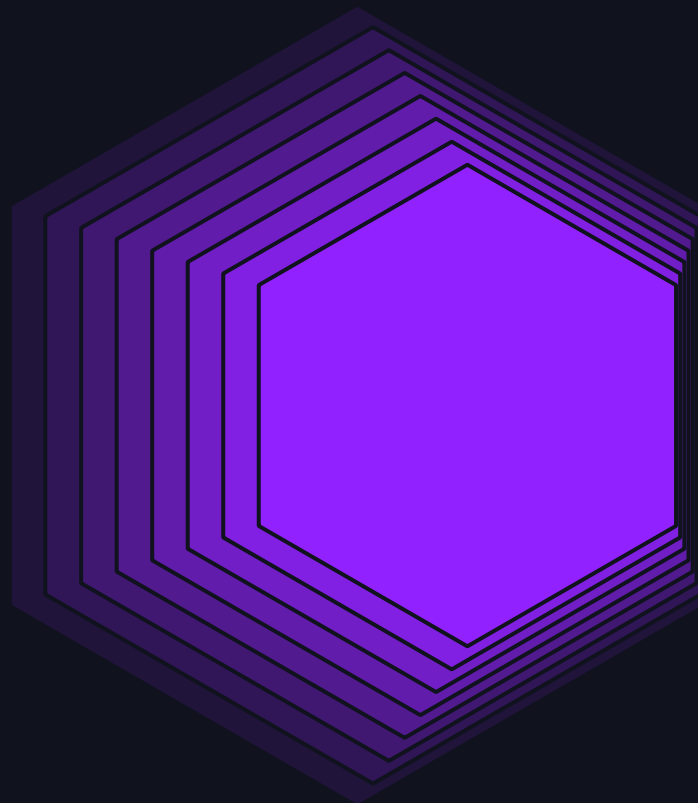


SIX CORE COMPONENTS FOR DATA MODELING SUCCESS

 **SqlIDBM™**


Hazal Sener
Senior Solutions Architect, SqlIDBM



INTRODUCTION

Hazal Sener, Senior Solutions Architect

PREVIOUSLY

Business Intelligence at a major B2B Bedbank, specializing as a Data Warehouse Modeler, building data transformation pipelines, and optimizing SQL queries for efficiency.

TODAY

I leverage this background to ensure SqlDBM users can unlock the platform's full potential and most effectively leverage it for their data modeling needs.



AGENDA

- What is a data model?
- Layers of relational data modeling
- Six components
 - Supporting Technology
 - Model Management
 - Communication & Collaboration
 - Roles & Responsibilities
 - Governance
 - Quality

WHAT IS A DATA MODEL?

WHAT IS A DATA MODEL?

Catalog Explorer `sqldbm_us_east_1_metastore` Send feedback

Catalog

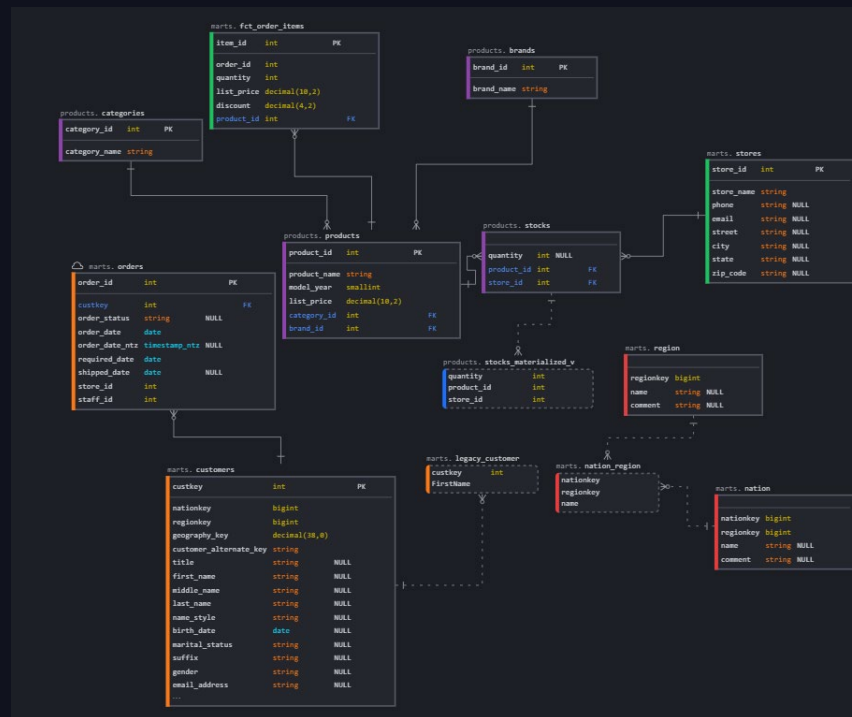
Catalogs > `sqldbm_demo` > `sqldbm_demo.marts`

Overview Details Permissions

Tables 8 Volumes 0 Models 0 Functions 2

Filter tables

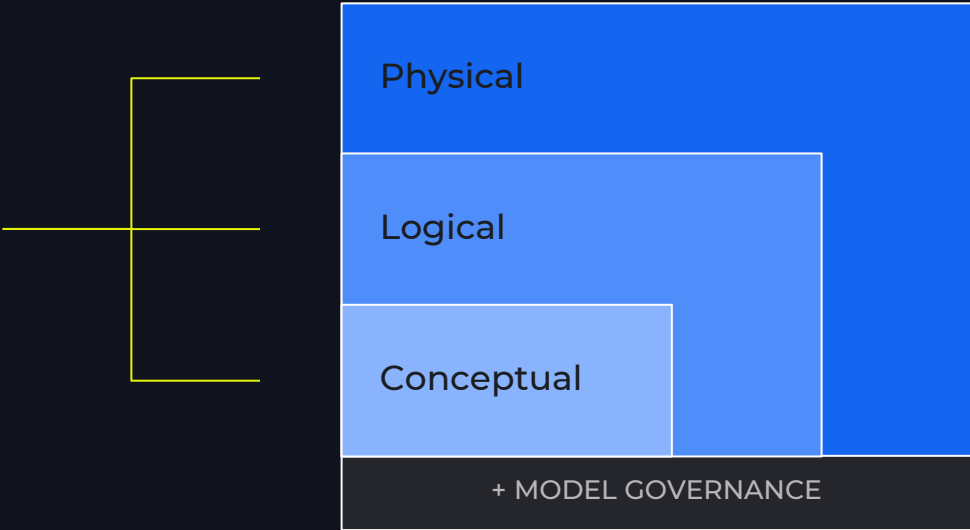
Name	Owner	Created at
customers	dev1@sqldbm.com	2023-04-18 21:47:17
fct_order_items	dev1@sqldbm.com	2023-04-18 21:47:23
legacy_customer	dev1@sqldbm.com	2023-09-07 19:02:41
nation	dev1@sqldbm.com	2023-04-18 21:45:31
nation_region	dev1@sqldbm.com	2023-04-18 21:57:46
orders	dev1@sqldbm.com	2024-03-04 14:30:49
region	dev1@sqldbm.com	2023-04-18 21:45:10
stores	dev1@sqldbm.com	2023-04-20 14:38:47



LAYERS OF RELATIONAL MODELING

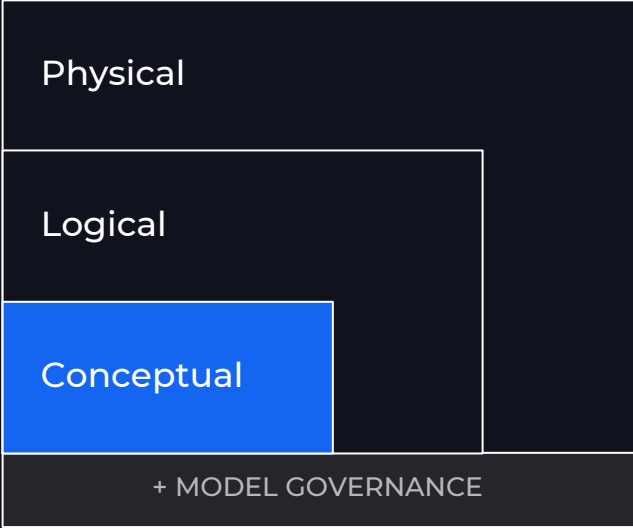
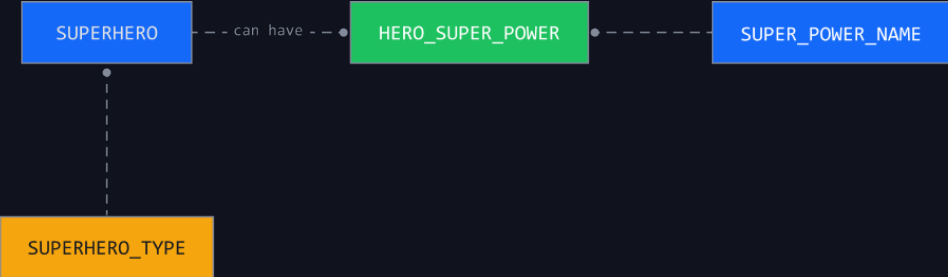
RELATIONAL DATA MODELING LAYERS

One cloud platform for all data modeling layers



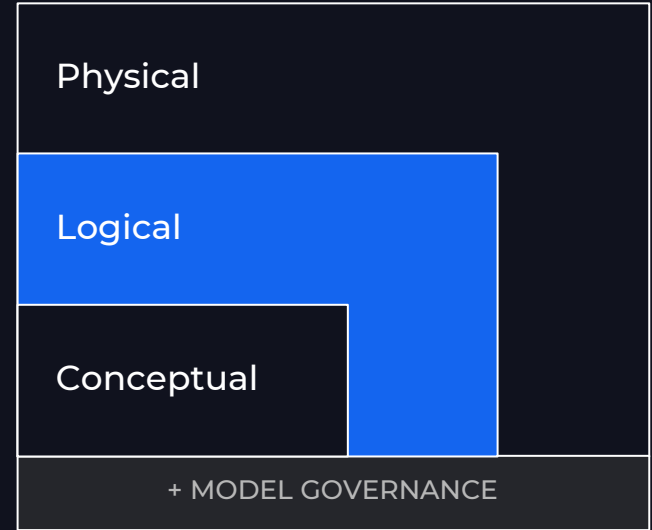
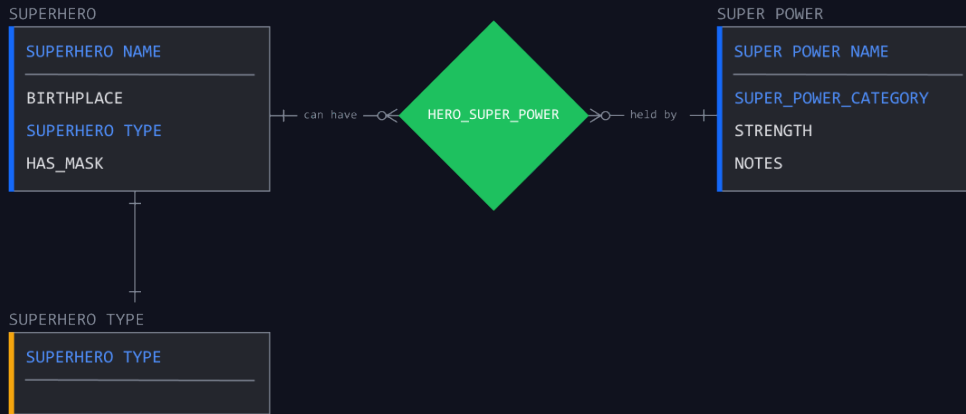
CONCEPTUAL DATA MODELING

A high-level snapshot to support the conversation between business and IT stakeholders



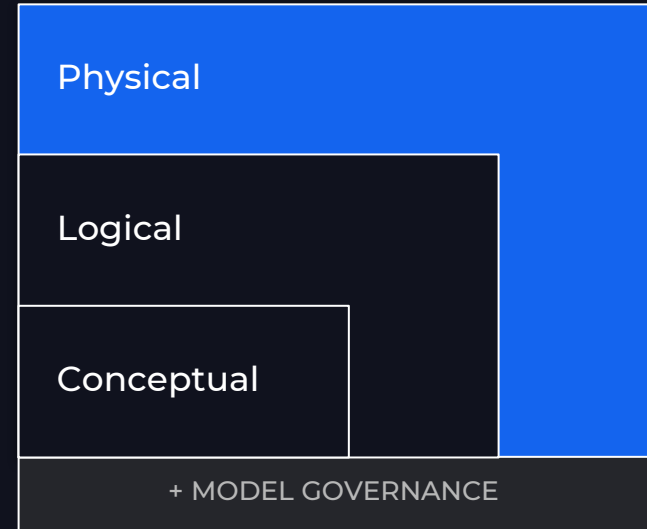
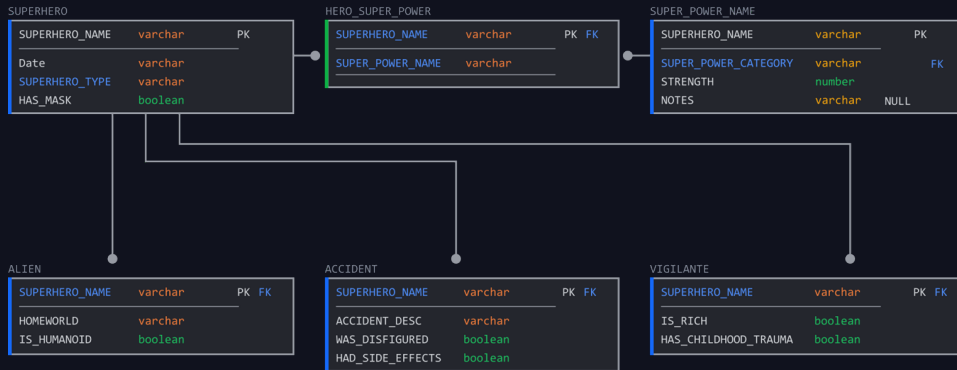
LOGICAL DATA MODELING

The next step. Transitioning from business concepts to data entities for more definitive relationships between them



PHYSICAL DATA MODELING

Introduces database specific context to data structures to ensure accurate translation of business logic into database schemas, tables, & other objects



PHYSICAL DATA MODELING

Conceptual and Logical Modeling determine the “what” while Physical Modeling determines the “how”. Having all three synchronized in one platform allows a seamless translation of business concepts to data structures.

```
3
4 ALTER TABLE marts.orders
5 DROP CONSTRAINT fk_3_2;
6 ALTER TABLE marts.customers
7   SET TBLPROPERTIES (delta.minWriterVersion = 5);
8 ALTER TABLE marts.customers
9   RENAME COLUMN GeographyKey TO geography_key;
10 ALTER TABLE marts.customers
11   RENAME COLUMN CustomerAlternateKey TO customer_id;
12 ALTER TABLE marts.customers
13   RENAME COLUMN FirstName TO first_name;
14 ALTER TABLE marts.customers
15   RENAME COLUMN MiddleName TO middle_name;
16 ALTER TABLE marts.customers
17   RENAME COLUMN LastName TO last_name;
18 ALTER TABLE marts.customers
19   RENAME COLUMN NameStyle TO name_style;
20 ALTER TABLE marts.customers
21   RENAME COLUMN BirthDate TO birth_date;
22 ALTER TABLE marts.customers
23   RENAME COLUMN MaritalStatus TO marital_status;
24 ALTER TABLE marts.customers
25   RENAME COLUMN EmailAddress TO email_address;
26 ALTER TABLE marts.customers
27   RENAME COLUMN YearlyIncome TO yearly_income;
```

```
104 CREATE TABLE products.brands
105 (
106   brand_id int NOT NULL CONSTRAINT pk_3 PRIMARY KEY,
107   brand_name string NOT NULL
108 )
109 USING DELTA
110 COMMENT 'table stores the brand's information of bikes, for example, Electra, Haro, a
111 TBLPROPERTIES (delta.minWriterVersion = 5, delta.minReaderVersion = 2, delta.checkpoi
112
113 -- ***** products.products
114
115 CREATE TABLE products.products
116 (
117   product_id int NOT NULL CONSTRAINT pk_4 PRIMARY KEY,
118   product_name string NOT NULL,
119   model_year smallint NOT NULL,
120   list_price decimal(10,2) NOT NULL,
121   category_id int NOT NULL CONSTRAINT fk_4 REFERENCES products.categories,
122   brand_id int NOT NULL CONSTRAINT fk_3 REFERENCES products.brands
123 )
124 USING DELTA
125 COMMENT 'table stores the product's information such as name, brand, category, model
126 TBLPROPERTIES (delta.minWriterVersion = 5, delta.minReaderVersion = 2, delta.checkpoi
127
128 -- ***** products.stocks
129
130 CREATE TABLE products.stocks
131 (
132   quantity int,
133   product_id int NOT NULL CONSTRAINT fk_10 REFERENCES products.products,
134   store_id int NOT NULL CONSTRAINT fk_11 REFERENCES marts.stores
135 )
136 USING DELTA
137 COMMENT 'table stores the inventory information i.e. the quantity of a particular pro
138
```

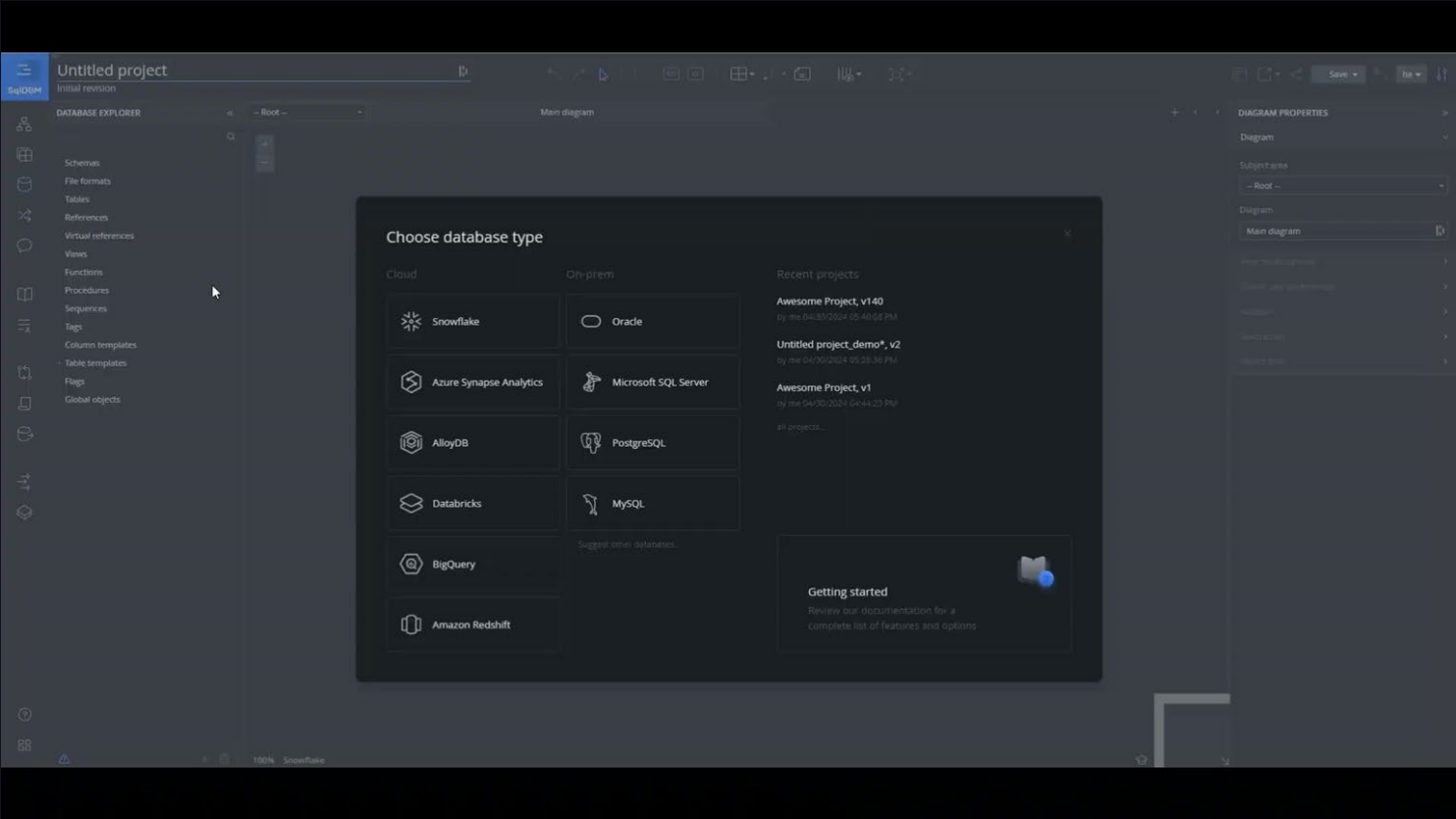
SIX CORE COMPONENTS OF RELATIONAL DATA MODELING

1. SUPPORTING TECHNOLOGY



+ Hive Metastore
and Unity Catalog

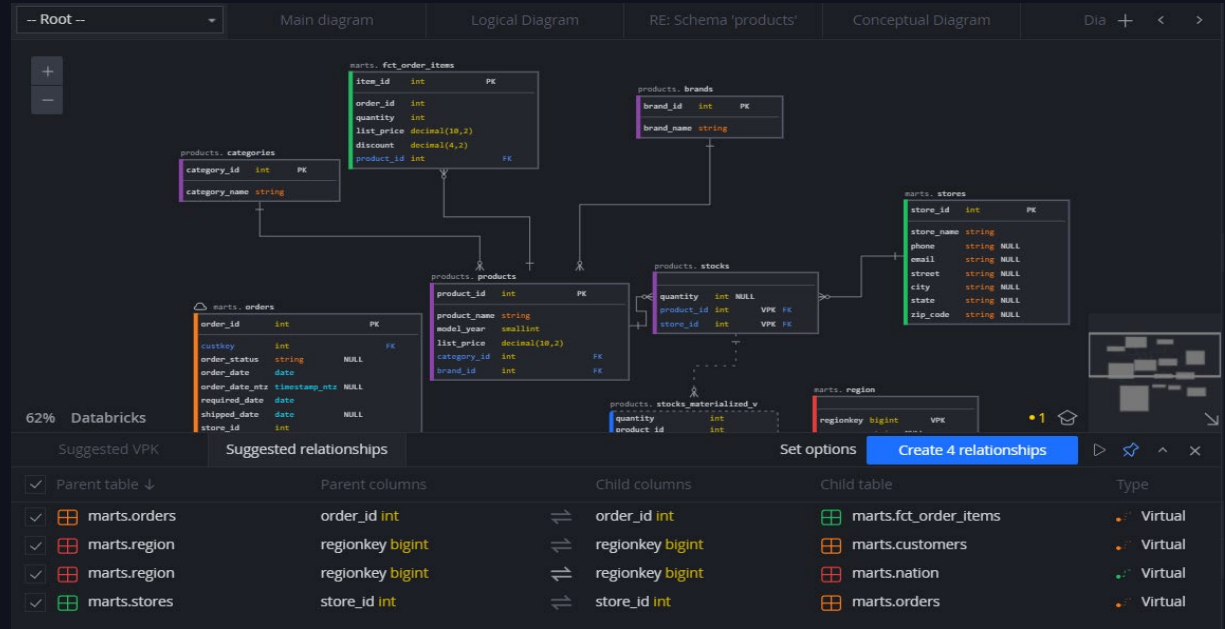
CONNECT DIRECTLY TO THE UNITY CATALOG



2. MODEL MANAGEMENT

Suggested Relationships

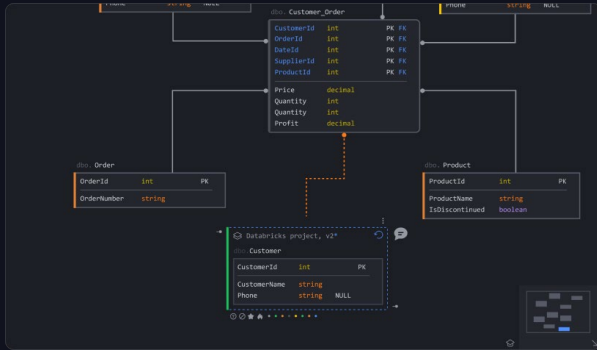
- Physical constraints for Unity Catalog
- Virtual constraints for Hive Metastore



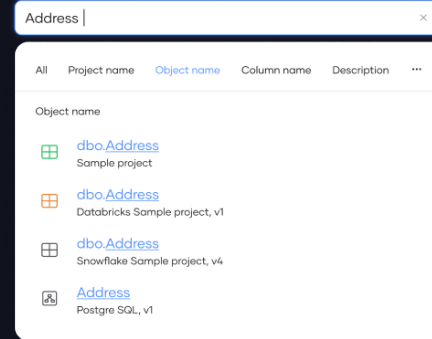
2. MODEL MANAGEMENT

Global Modeling

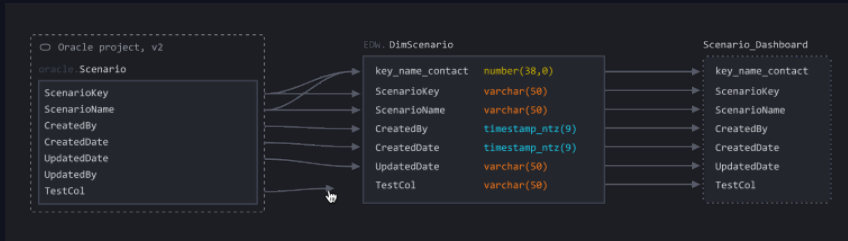
Reference



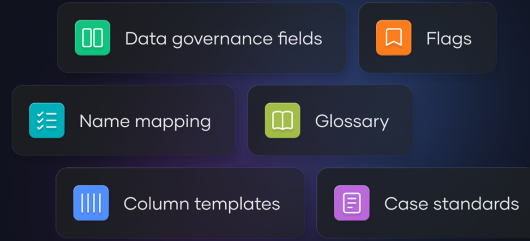
Search (coming soon)



Lineage (coming soon)

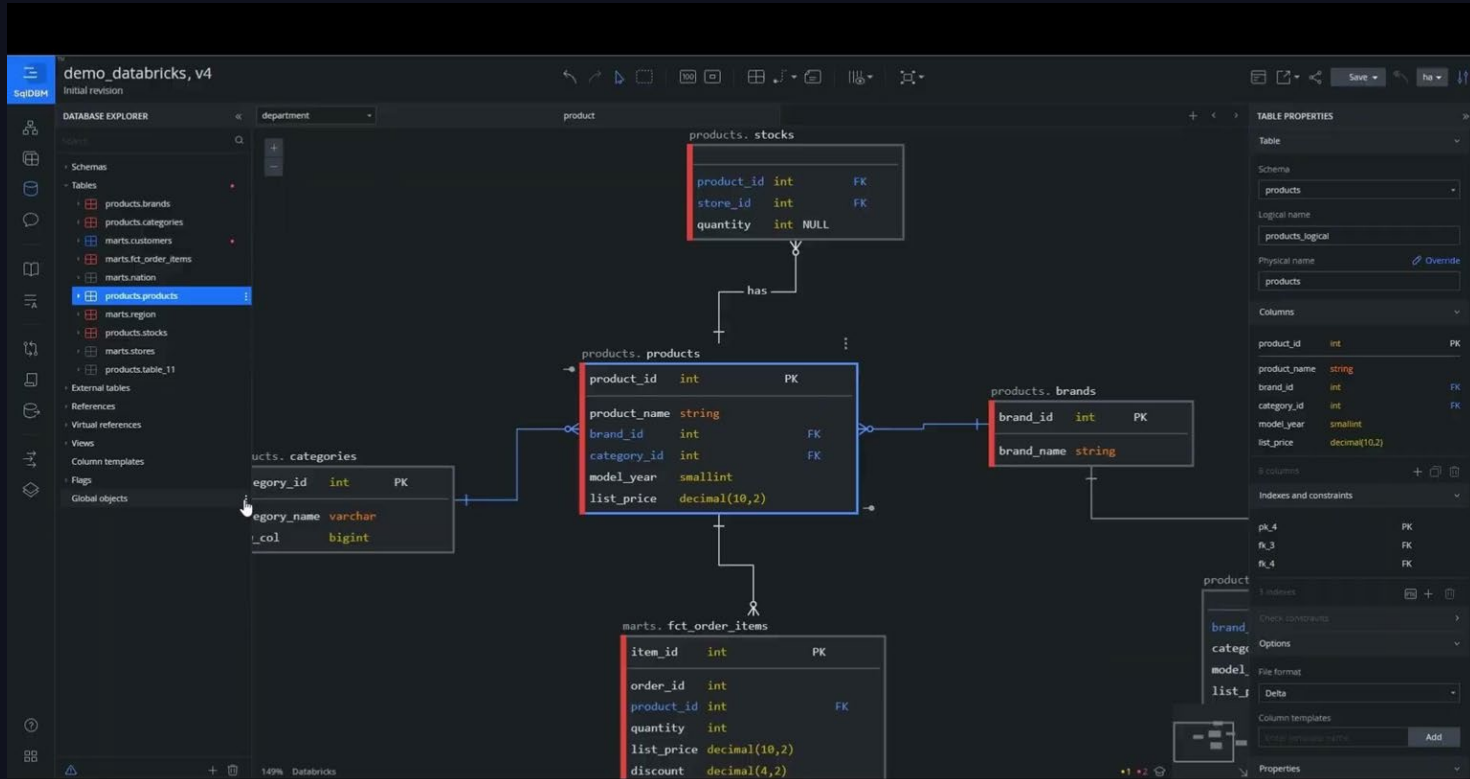


Standards (coming soon)



2. MODEL MANAGEMENT

Global Reference



3. COMMUNICATION & COLLABORATION

The screenshot displays the Databricks SQL interface for a 'Databricks Sample Project, v17'. The main area shows a 'Level_of_Detail' diagram with various tables like 'part', 'partsupp', 'supplier', 'nation', and 'region' connected by relationship lines. A comment thread is visible on the left side of the diagram.

Project Collaborators

Project Team

Email Send Email

Allow to Edit

Em Emily.Thomas@sqldb.com Consumer

Mi Michael Page 26 minutes ago, v1 ✓

In the Find and Replace dialogue box of text documents, you can select to include the comments texts in your searches.

16 minutes ago, v2

Comments on the text are given in italic text within brackets.

Mi Cameron Williamson 2 minutes ago, v1

- The draft text was submitted to the Venice Commission last year and comments on the text are expected soon.

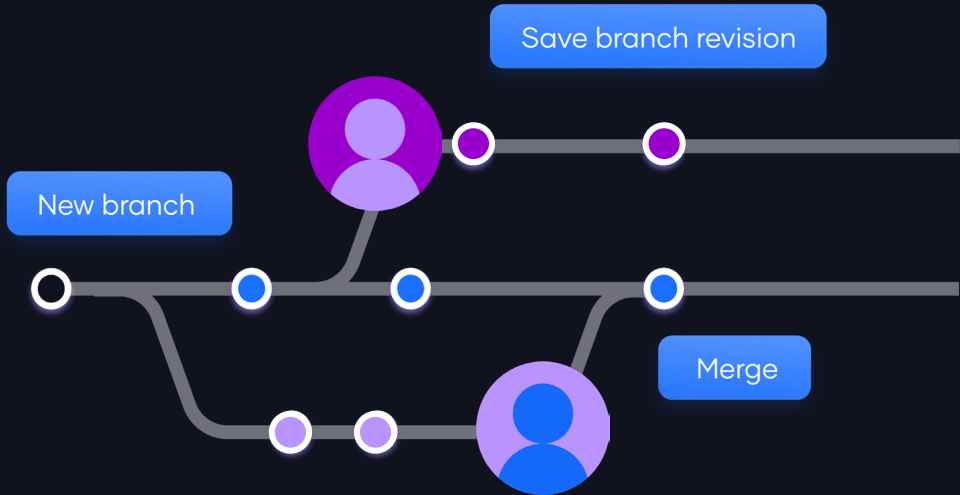
Reply



3. COMMUNICATION & COLLABORATION

Concurrent modeling

The entire project team can work collaboratively and in parallel without having to lock a project. Changes made in any branch can be compared before merging to the main branch and synchronized with any new updates.

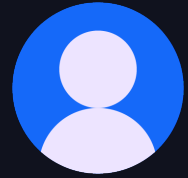


4. ROLES & RESPONSIBILITIES

Define the key roles involved in data modeling within the organization, such as Data Consumers, Data Modelers and Data Stewards and outline their specific responsibilities and contributions to the data modeling process, ensuring clear accountability and effective collaboration.



Marta
Data Consumer



Dave
Data Modeler



Trevor
Governance

5. MODEL GOVERNANCE

Pages

Fields

Databricks Sample Project, v25
My organization has 1 columnar source

Database documentation for mySchema

Column name	Order	Owner	Mediation	Source table	Location	File schema	Dist	Critical
lineItemNumber	1	Danny	S					
orderKey								
partKey								
supply								
customerKey								
quantity								
costPerUnit								
discount								
tax								
returnFlag								
lineStatus								
shipDate								
commitDate								

DB DOCUMENTATION PROPERTIES

Content options

Data governance fields

- Select all
- Owner
- Mediation
- Source table
- Location
- File schema
- Dist
- Critical

3 fields

dbt properties

Select all

- access
- docs
- external
- freshness
- identifier
- latest_version
- loaded_at_field
- model_config

Final DB
Initial Revision

DATA GOVERNANCE PAGES

Search

Pages

- Business terms
- Project goals
- Business goals
- Project team
- Acceptance criteria
- Business rules

BUSINESS TERMS

Header 1

Header 2

Header 3

Click Data Classification button from the ribbon and choose one of the built-in classification functions. Create custom fields

To add custom fields to your repository choose Custom Fields option from the ribbon and add new field. For new field provide: Title - make it something meaningful, for instance a name of regulation or a meaningful keyword, like Sensitive data, Classification Type - choose Drop-down list, open or closed, depending if you want users to be able to add new values as they browse the documentation (closed is advised) Definition - list all classification labels separated with comma @Michail@John

Visibility - choose Column at minimum, you may also add Table/View, Parameter and Term

Business Customer, Business Order, Business Order Item, Business Product, Business Discount, Business Supplier

Reports

AdWorks, v6
Snowflake

DB documentation

- Schemas
 - EDW
 - Tables
 - ChildDemo
 - ChildDemo_Test
 - DimAccount
 - DimCurrency
 - DimCustomer
 - DimDate
 - DimDepartmentGroup
 - DimEmployee
 - DimGeography
 - DimOrganization

EDW.DimAccount

Details Columns Tags & Flows

Name

AccountCodeAlternateKey

AccountDescription

AccountKey

AccountType

ParentAccountCodeAlternateKey

CustomMemberOptions

CustomMembers

Operator

ParentAccountCodeAlternateKey

All	Object name	Column name	Description	Field	Flags	Tags
	AccountKey					EDW.DimAccount
	AccountKey					EDW.DimAccount
	ParentAccountKey					EDW.DimAccount
	ParentAccountKey					EDW.DimAccount
	AccountCodeAlternateKey					EDW.DimAccount
	AccountCodeAlternateKey					EDW.DimAccount
	ParentAccountCodeAlternateKey					EDW.DimAccount
	ParentAccountCodeAlternateKey					EDW.DimAccount
	CustomMembers	CustomMembers				varchar(300)
	Operator	Operator				varchar(50)
	ParentAccountCodeAlternateKey	ParentAccountCodeAlternateKey				number(38,0)



THANK YOU

 **SqlDBM™**