

Azure Databricks Coding Challenge

Pratik Wani

Question 2

❖ 3 level namespace and creating Unity Catalog objects.:

- A three-level namespace in Unity Catalogue consists of three levels of hierarchy
 - Catalog (Databases)
 - Schema (Folders or Directories)
 - Objects (Views)
- It is a powerful feature for organizing and managing data in Unity Catalogue
- First created a workspace, using Access Connector I attached a Unity Catalog metastore (created in the same region) to a workspace.
- Then Enabled the admin role for Databricks account
- After successful connection, the metastore name is visible at the top in the catalog

❖ Setup:

○ Created Metastore

Microsoft AzuredatabricksAccount

pratikwani116@outlook.com

Data

Metastores

A metastore is the top-level container for data in Unity Catalog. Within a metastore, Unity Catalog provides a 3-level namespace for organizing data: catalogs, schemas (also called databases), and tables / views. [Learn More](#)

Create metastore

Name	Region	Path	Created at	Updated at
metastore_azure_no...	northeurope		today at 10:23 AM	today at 10:23 AM
hexa-deb-1056-met...	centralindia	abfss://hexadeb1056container@hexadeb1056storageacc.dfs.core.windows.net/c683c395-47d7-4261-8892-d5ec12677566	last Monday at 5:17 ...	last Monday at 5:17 ...

○ Enabled Admin role

Microsoft AzuredatabricksAccount

pratikwani116@outlook.com

pratik Wani >

Wani

InformationRoles

Role name	Description
Account admin	Can manage workspaces, users & groups, cloud resources and settings.
Marketplace admin	Can manage exchanges and listings on the Marketplace.

- Unity Catalog Badging can be seen here

The screenshot displays the Databricks interface for configuring a cluster named "Pratik Wani's Cluster". The left sidebar contains navigation options like Workspace, Recents, Catalog, Workflows, Compute, SQL, SQL Editor, Queries, Dashboards, Alerts, Query History, SQL Warehouses, Data Engineering, Job Runs, Data Ingestion, Delta Live Tables, Machine Learning, Experiments, Features, Models, and Serving. The main content area is divided into sections: Policy (Personal Compute), Access mode (Single user access, Single user, Pratik Wani), Performance (Databricks Runtime Version 14.3 LTS ML, Use Photon Acceleration), Node type (Standard_DS3_v2, 14 GB Memory, 4 Cores), Tags (No custom tags), and Advanced options. A summary box on the right provides a quick overview of the cluster configuration.

Pratik Wani's Cluster Send feedback

Configuration | Notebooks (0) | Libraries | Event log | Spark UI | Driver logs | Metrics | Apps | Spark compute UI - Master

Policy

Personal Compute

Access mode

Single user access

Single user | Pratik Wani

Performance

Databricks Runtime Version

14.3 LTS ML (includes Apache Spark 3.5.0, Scala 2.12)

☐ Use Photon Acceleration

Node type

Standard_DS3_v2 | 14 GB Memory, 4 Cores

☒ Terminate after 120 minutes of inactivity

Tags

No custom tags

> Automatically added tags

► Advanced options

Summary

1 Driver | 14 GB Memory, 4 Cores

Runtime | 14.3.x-cpu-mi-scala2.12

Unity Catalog | Standard_DS3_v2 | 0.75 DBU/h

❖ Created Unity Catalog Object

Unity Catalog 17:20:17 Python ☆
File Edit View Run Help Last edit was 1 minute ago New cell UI: ON ▾

▶ Run all ● Pratik Wani's Cluster ▾ 📅 Schedule Share ^

Cell 1
05:21 PM (<1s)
from pyspark.sql import SparkSession

Cell 2
05:21 PM (<1s)
spark=SparkSession.Builder().appName('Unity').getOrCreate()

Cell 3
05:21 PM (5s)
spark.sql("CREATE CATALOG IF NOT EXISTS quickstart_catalog")
DataFrame[]

Cell 4
05:25 PM (1s)
spark.sql("USE CATALOG quickstart_catalog")
DataFrame[]

Unity Catalog 17:20:17 Python ☆
File Edit View Run Help Last edit was 2 minutes ago New cell UI: ON ▾

▶ Run all ● Pratik Wani's Cluster ▾ 📅 Schedule

Cell 5
05:25 PM (<1s)
display(spark.sql("SHOW CATALOGS"))

Table ▾ + New result table: OFF ▾

	catalog
1	hive_metastore
2	main
3	quickstart_catalog
4	samples
5	system

5 rows | 0.45 seconds runtime Refreshed 12 minutes ago

Cell 6
05:30 PM (1s)
spark.sql("""
GRANT CREATE, USE CATALOG
ON CATALOG quickstart_catalog
TO `account users`""")
DataFrame[]

Cell 7

05:31 PM (1s)

display(spark.sql("SHOW GRANT ON CATALOG quickstart_catalog"))

Table + New result table: OFF

	Principal	ActionType	ObjectType	ObjectKey
1	account users	CREATE SCHEMA	CATALOG	quickstart_catalog
2	account users	USE CATALOG	CATALOG	quickstart_catalog

2 rows | 0.61 seconds runtime Refreshed 7 minutes ago

Cell 8

05:31 PM (1s)

spark.sql("""
CREATE SCHEMA IF NOT EXISTS quickstart_schema
COMMENT 'A new Unity Catalog schema called quickstart_schema'""")

DataFrame[]

Cell 9

05:31 PM (1s)

display(spark.sql("SHOW SCHEMAS"))

Table + New result table: OFF

	dbName
1	default
2	information_schema
3	quickstart_schema

3 rows | 0.59 seconds runtime Refreshed 7 minutes ago

Cell 10

05:31 PM (<1s)

display(spark.sql("DESCRIBE SCHEMA EXTENDED quickstart_schema"))

Table + New result table: OFF

	database_description_item	database_description_value
1	Catalog Name	quickstart_catalog
2	Namespace Name	quickstart_schema
3	Comment	A new Unity Catalog schema called quickstart_schema
4	Location	
5	Owner	pratikwani116@outlook.com
6	Properties	