Name: Krushnakumar Patle

Email: krishnapatle128@gmail.com

Batch: Data Engineering Batch-1

Azure Databricks Coding Challenge

Q2. Explain Overview of 3 level namespace and creating Unity Catalog objects.

Unity Catalog provides centralized access control, auditing, lineage, and data discovery capabilities across Azure Databricks workspaces.

- A three-level namespace in Unity Catalogue consists of three levels of hierarchy — **catalog, schema, and objects.**
- Catalogs are like databases, schemas are like folders or directories, and objects can be tables, views, functions, or other entities.
- For instance, consider a data warehouse containing sales data. You may create a catalog called “sales_data”, and within this catalog, you may create schemas for each sales region, such as “north”, “south”, “east”, and “west”. Within each schema, you may create tables for each quarter, such as “Q1_2020”, “Q2_2020”, “Q3_2020”, and “Q4_2020”. Each table can contain data for a specific region and quarter. This three-level namespace structure makes it easier to organize and access data based on its purpose and location.

The Unity Catalog object model

In Unity Catalog, the hierarchy of primary data objects flows from metastore to table or volume:

• Metastore: The top-level container for metadata. Each metastore exposes a three-level namespace (catalog.schema.table) that organizes your data.

• Catalog: The first layer of the object hierarchy, used to organize your data assets.

• Schema: Also known as databases, schemas are the second layer of the object hierarchy and contain tables and views.

• Tables, views, and volumes: At the lowest level in the data object hierarchy are tables, views, and volumes. Volumes provide governance for non-tabular data.

• Models: Although they are not, strictly speaking, data assets, registered models can also be managed in Unity Catalog and reside at the lowest level in the object hierarchy.