**CASE STUDY – DELTA TABLE OPERATIONS**

**Step 1: Create Example Data**

from pyspark.sql.types import StructType, StructField, IntegerType, StringType

# Example data (10 records)

data = [

(1, "Lokanya", "IT"),

(2, "Gowrishankar", "CSE"),

(3, "Arun", "ECE"),

(4, "Priya", "IT"),

(5, "Kumar", "EEE"),

(6, "Meena", "CSE"),

(7, "Ravi", "MECH"),

(8, "Divya", "IT"),

(9, "Suresh", "CIVIL"),

(10, "Aarthi", "CSE")

]

# Define schema

schema = StructType([

StructField("id", IntegerType(), True),

StructField("name", StringType(), True),

StructField("department", StringType(), True)

])

# Create DataFrame

df = spark.createDataFrame(data, schema)

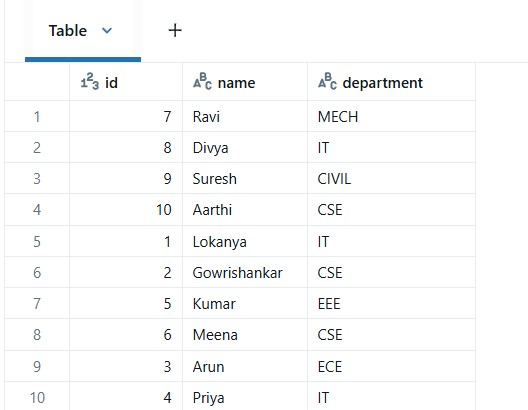
# Write as Delta Table

df.write.format("delta").mode("overwrite").save("/mnt/rawdata/employees")

# Register as SQL table

spark.sql("DROP TABLE IF EXISTS employees")

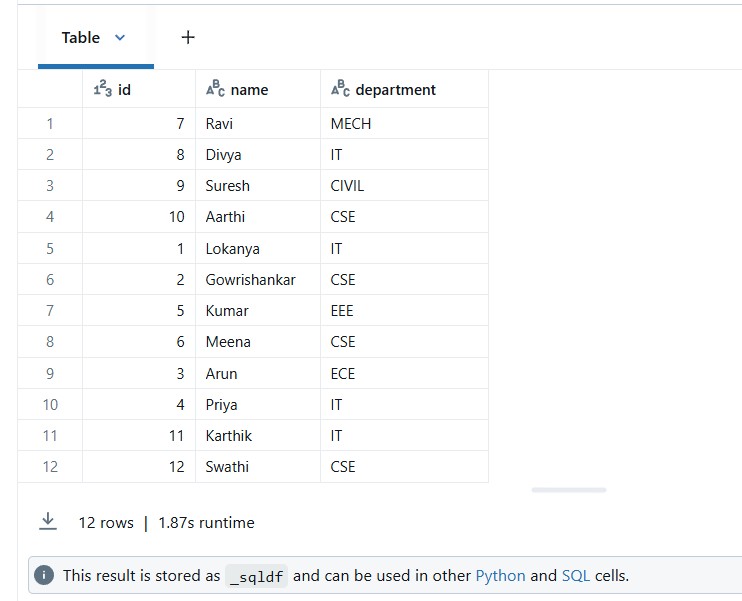
spark.sql("CREATE TABLE employees USING DELTA LOCATION '/mnt/delta/employees'")



**Step 2: Insert New Data**

INSERT INTO employees VALUES (11, "Karthik", "IT");

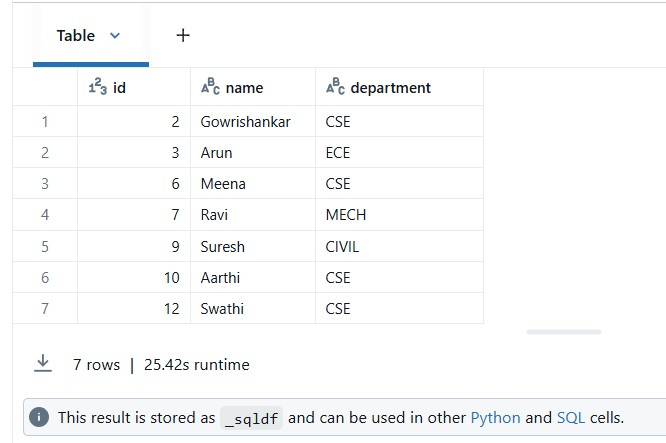
INSERT INTO employees VALUES (12, "Swathi", "CSE");



**Step 3: Delete Operation**

-- Delete employee from IT department

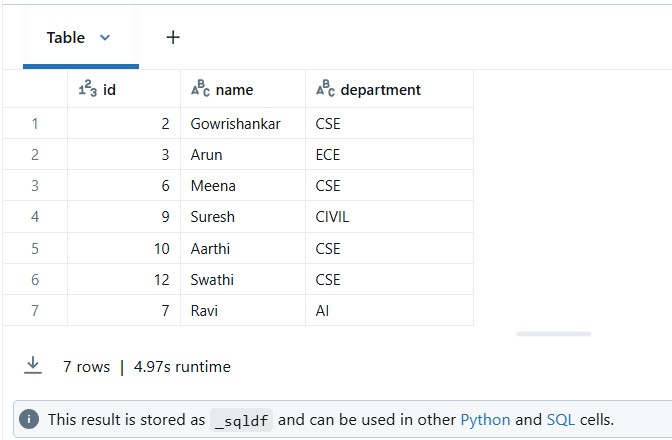
DELETE FROM employees WHERE department = 'IT';



**Step 4: Update Operation**

-- Update department of Ravi from MECH → AI

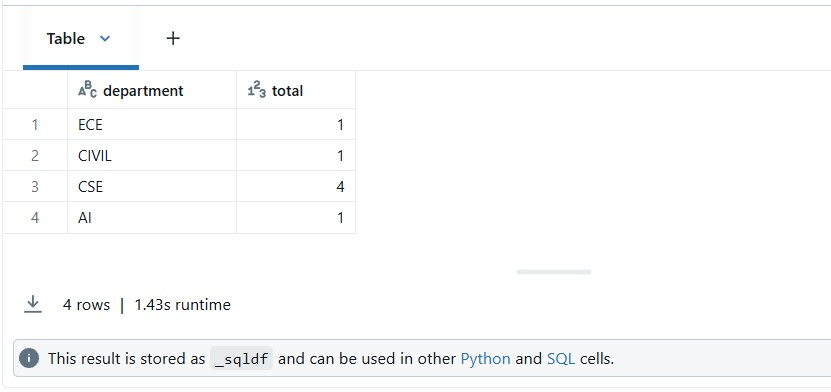
UPDATE employees SET department = 'AI' WHERE name = 'Ravi';



**Step 5: Query Operations**

-- 1. Count by department

SELECT department, COUNT(\*) AS total FROM employees GROUP BY department;



-- 2. Get employees from IT department

SELECT \* FROM employees WHERE department = 'IT';

