**Mock Test 1 - Big Data Question no. 6**

**6.** Retrieve all the customers from the "Customers" table whose age is greater than 25 and have made at least one purchase.

**Sol.**

from pyspark.sql import SparkSession

spark = SparkSession.builder \

.appName("CustomerQuery") \

.getOrCreate()

customers.createOrReplaceTempView("Customers")

# Execute the SQL query

query = """

SELECT \*

FROM Customers

WHERE age > 25 AND customer\_id IN (

SELECT DISTINCT customer\_id

FROM Purchases

)

"""

result = spark.sql(query)

# Show the result

result.show()

# Stop the SparkSession

spark.stop()