**Serverless Querying of CSV Data in S3 Using AWS Athena**

**Objective**

To query structured data stored in an Amazon S3 bucket using SQL — without deploying or managing a traditional database system — by leveraging AWS Athena, a serverless interactive query service.

**Steps Performed**

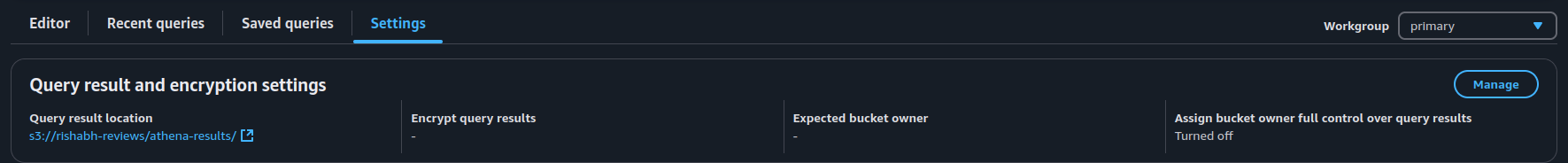
**Upload CSV File to S3**

File uploaded to: s3://rishabh-imdb-reviews/IMDB Dataset.csv

**Step 2: Open AWS Athena Console**

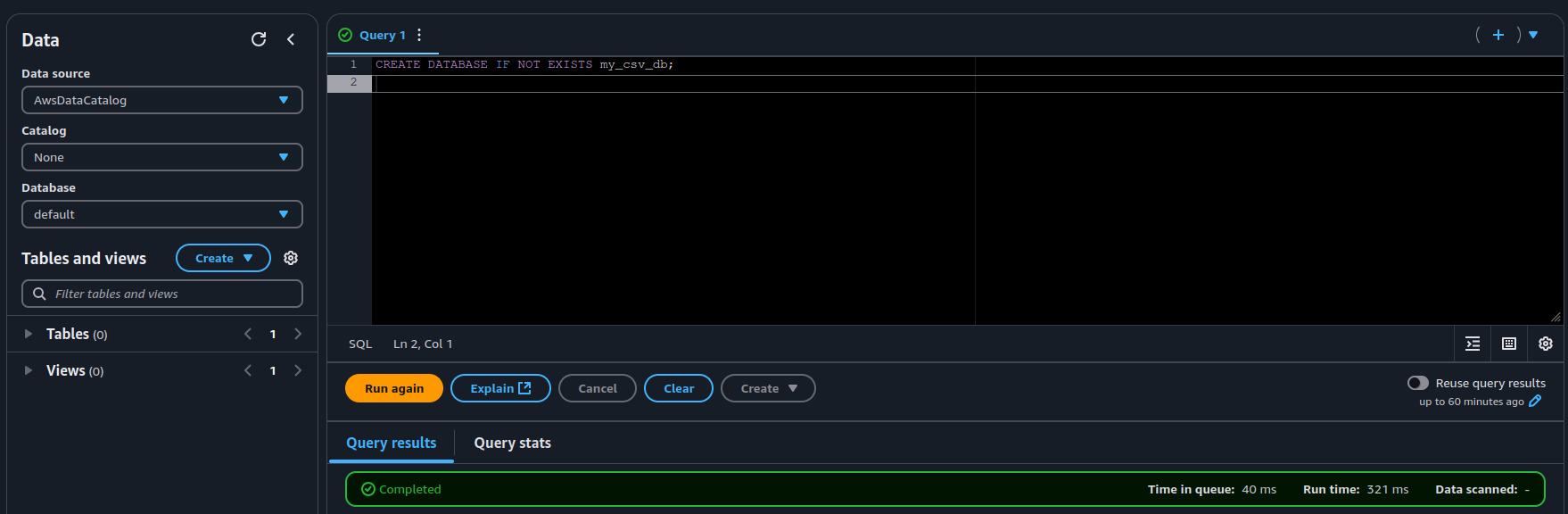
**Step 3: Set Query Result Location**

Set result output to: s3://rishabh-imdb-reviews/athena-results/ under Settings.

**Step 4: Create Database**

Run the following query in Athena:

CREATE DATABASE IF NOT EXISTS my\_csv\_db;



Step 5: Create External Table

Run the following query to create the external table:

CREATE EXTERNAL TABLE IF NOT EXISTS imdb\_reviews (

review STRING,

sentiment STRING

)

ROW FORMAT SERDE 'org.apache.hadoop.hive.serde2.lazy.LazySimpleSerDe'

WITH SERDEPROPERTIES (

'serialization.format' = ',',

'field.delim' = ','

)

STORED AS TEXTFILE

LOCATION 's3://rishabh-imdb-reviews/'

TBLPROPERTIES (

'skip.header.line.count' = '1'

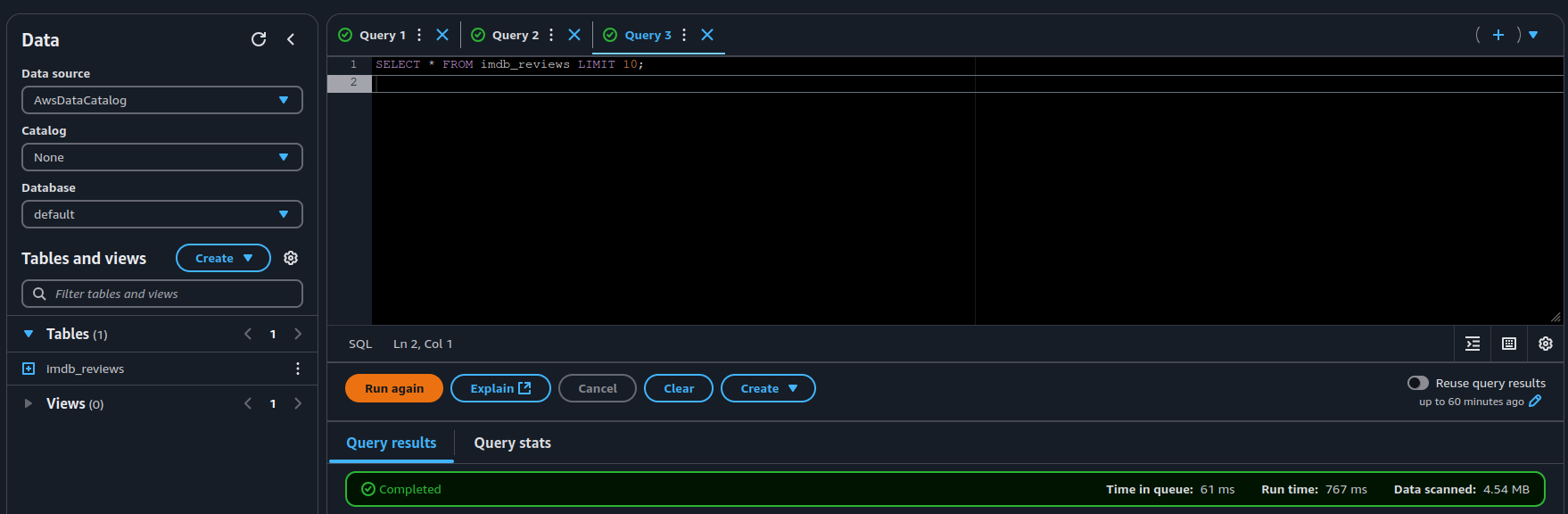
);



Step 6: Query the Table

Run the following query to verify the data:

SELECT \* FROM imdb\_reviews LIMIT 10;

Outcome

Successfully queried structured data from S3 using Athena without provisioning any database engine. The entire process was serverless, simple, and efficient.

