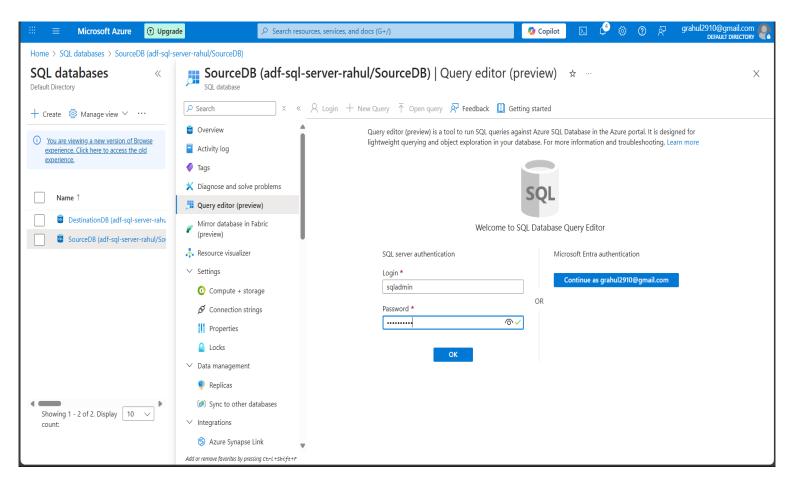
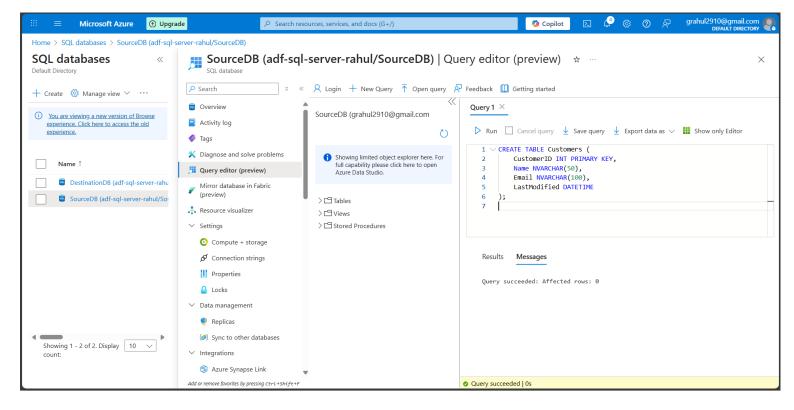
# **Incremental Load Pipeline in Azure Data Factory**

1. Created SQL Tables with Watermark Column
In both SourceDB and DestinationDB:
sql
CopyEdit
CREATE TABLE Customers (
 CustomerID INT PRIMARY KEY,
 Name NVARCHAR(50),
 Email NVARCHAR(100),
 LastModified DATETIME

);





2. Inserted Initial Sample Data into SourceDB

sql

CopyEdit

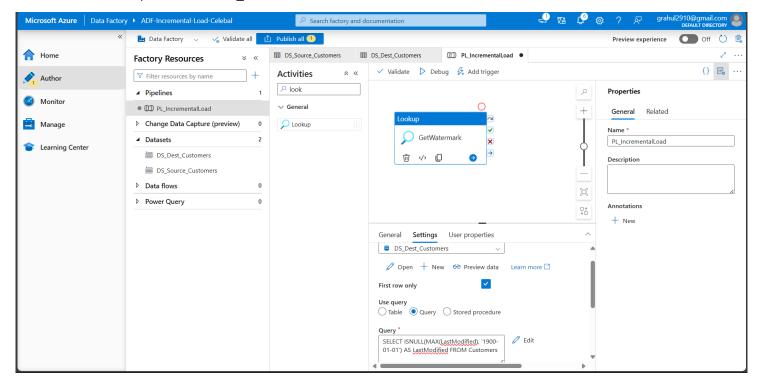
INSERT INTO Customers (CustomerID, Name, Email, LastModified)

**VALUES** 

(1, 'Rahul Sharma', 'rahul@example.com', GETDATE()),

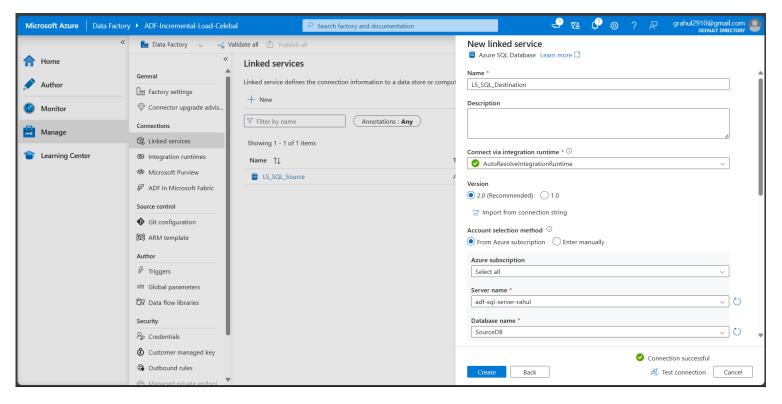
(101, 'New User', 'newuser@example.com', GETDATE());

- 3. Launched ADF Studio and Created the Pipeline
  - Pipeline Name: PL\_IncrementalLoad



#### 4. Created Linked Services

- LS\_SQL\_Source → Connected to SourceDB
- LS\_SQL\_Destination → Connected to DestinationDB



#### 5. Created Datasets

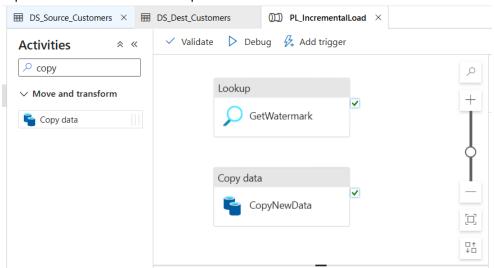
- DS\_Source\_Customers → Table: Customers, from LS\_SQL\_Source
- DS\_Dest\_Customers → Table: Customers, from LS\_SQL\_Destination
- 6. Added Lookup Activity: GetWatermark
  - SQL Query:

sql

# CopyEdit

SELECT ISNULL(MAX(LastModified), '1900-01-01') AS LastModified FROM Customers

• Purpose: Get the latest timestamp from destination table to use as a watermark



- 7. Added Copy Data Activity: CopyNewData
  - Source:
    - o Use query:

sql

CopyEdit

**SELECT \* FROM Customers** 

WHERE LastModified > '@{activity('GetWatermark').output.firstRow.LastModified}'

• Sink:

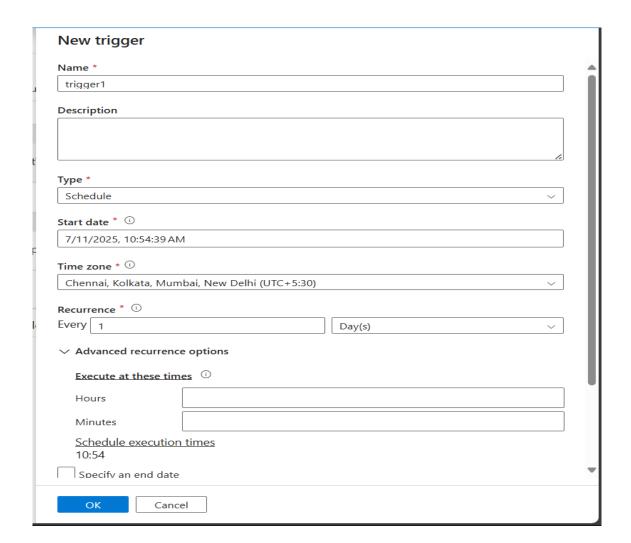
Dataset: DS\_Dest\_Customers

o Write Behavior: Upsert

o Key Column: CustomerID

## 8. Created a Scheduled Trigger

- Type: Recurring (every 1 day)
- Attached to pipeline PL\_IncrementalLoad



### 9. Verified the Pipeline

