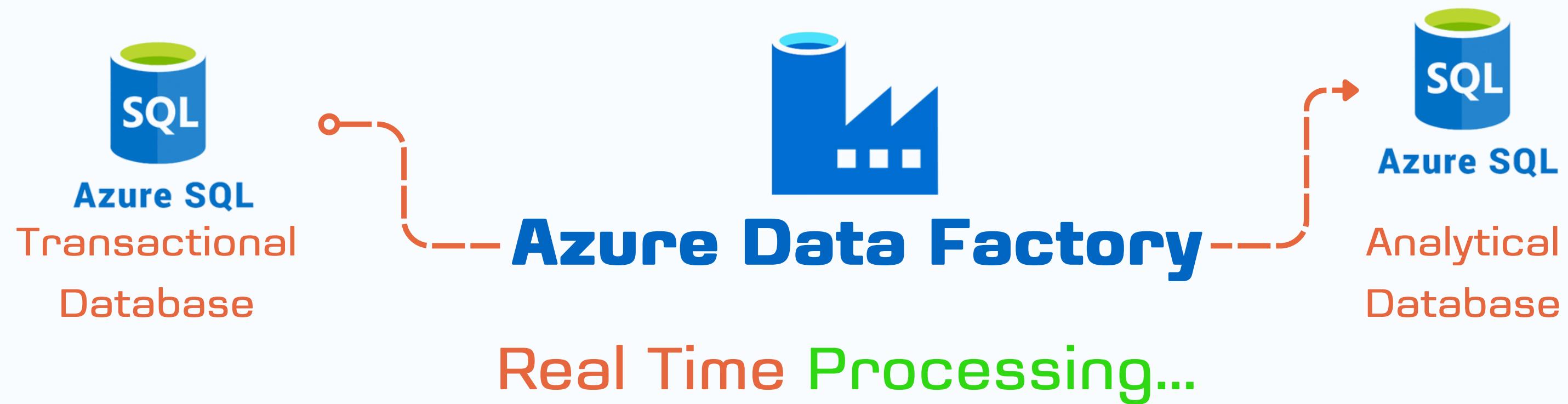


Incremental Load...



[in/prayagv](https://www.linkedin.com/in/prayagv)
[github/prayag-verma](https://github.com/prayag-verma)



Prayag Verma
Data Engineer

Step 1

Login To Azure SQL DB

Go to Azure Portal, then select an existing one or create a new resource group, then next, create a new database or open an existing database, then click on ‘Query editor (preview)’ option from the left side bar.

The screenshot shows the Microsoft Azure portal interface. At the top, there's a blue header bar with the Microsoft Azure logo, an 'Upgrade' button, a search bar ('Search resources, services, and docs (G+/)'), and a Copilot icon. On the far right, it shows the user's email (Prayag.Verma@student...) and a 'STUDENT AMBASSADORS (STDN...)' badge. Below the header, the main content area shows the 'AzurePractice (azure-dbserver-test/AzurePractice)' database details. The left sidebar has a 'Overview' tab selected, followed by 'Activity log', 'Tags', 'Diagnose and solve problems', 'Query editor (preview)' (which is highlighted with a red box), 'Mirror database in Fabric (preview)', 'Settings', 'Data management', 'Integrations', 'Power Platform', 'Security', 'Intelligent performance', 'Monitoring', 'Automation', and 'Help'. The main panel displays the database's resource group (practice_dev_env_rg), status (Online), location (Central India), subscription (Azure subscription 1), and various server settings like server name (azure-dbserver-test.database.windows.net), connection strings, pricing tier (Free - General Purpose - Serverless: Gen5, 2 vCores), overage billing (Disabled), free monthly vCore amount (99,671 vCore seconds remaining), and earliest restore point (2025-02-23 22:09 UTC). It also shows tags (Project: Incremental Load, Environment: Development). Below this, there are sections for 'Getting started', 'Monitoring', 'Properties', 'Features', 'Notifications (0)', 'Integrations', and 'Tutorials'. A 'Start working with your database' section at the bottom provides links to 'Configure access', 'Connect to application', 'Start developing', and 'Mirror database in Fabric'.



/ prayagv



/ prayag-verma

Prayag Verma | Data Engineer

Step 1

Login To Azure SQL DB

Next, Login using your SQL server authentication credentials (username, and Password)

The screenshot shows the Microsoft Azure portal interface. At the top, there's a blue header bar with the Microsoft Azure logo, an 'Upgrade' button, a search bar ('Search resources, services, and docs (G+/)'), a Copilot button, and a user profile for 'Prayag.Verma@student... STUDENT AMBASSADORS (STDN...)'. Below the header, the URL 'Home > Recent > AzurePractice (azure-dbserver-test/AzurePractice)' is visible. The main title is 'AzurePractice (azure-dbserver-test/AzurePractice) | Query editor (preview)'. On the left, a sidebar lists various database management options: Overview, Activity log, Tags, Diagnose and solve problems, Query editor (preview) (which is highlighted with a red box), Mirror database in Fabric (preview), Settings, Data management, Integrations, Power Platform, Security, Intelligent performance, Monitoring, Automation, and Help. The central area is titled 'Welcome to SQL Database Query Editor' and features a large 'SQL' icon. It displays the 'SQL server authentication' section with 'Login *' set to 'sqluser' and 'Password *' masked. A blue 'OK' button is at the bottom right of the login form.

Step 2

Creating Source Tables

Now, create three tables ‘customers, products, and orders’, using the sql queries given below!

The screenshot shows the Microsoft Azure portal interface, specifically the Query editor (preview) for the database 'AzurePractice'.

Left Sidebar: Shows navigation links including Overview, Activity log, Tags, Diagnose and solve problems, Query editor (preview), Mirror database in Fabric (preview), Settings, Data management, Integrations, Power Platform, Security, Intelligent performance, Monitoring, Automation, and Help. A note indicates limited object explorer here; for full capability, click here to open Azure Data Studio.

Top Bar: Microsoft Azure logo, Upgrade button, Search bar ('Search resources, services, and docs (G+)'), Copilot button, and user information ('Prayag.Verma@student... STUDENT AMBASSADORS (STDN...)').

Query Editor: Title bar says 'AzurePractice (azure-dbserver-test/AzurePractice) | Query editor (preview)'. The main area contains two tabs: 'Query 1' (selected) and 'Query 3'. Below the tabs are buttons for Run, Cancel query, Save query, Export data as, Show only Result, and Open Copilot. The code area contains the following SQL script:

```
1  DROP TABLE IF EXISTS customers;
2  CREATE TABLE customers (
3      customer_id INT PRIMARY KEY,
4      first_name NVARCHAR(50),
5      last_name NVARCHAR(50),
6      email NVARCHAR(100),
7      phone NVARCHAR(15),
8      Address NVARCHAR(255)
9  );
10 DROP TABLE IF EXISTS products;
11 CREATE TABLE products (
12     product_id INT PRIMARY KEY,
13     product_name NVARCHAR(100),
14     price DECIMAL(10, 2),
15     stock_quantity INT
16 );
17 DROP TABLE IF EXISTS orders;
18 CREATE TABLE orders (
19     order_id INT PRIMARY KEY,
20     customer_id INT,
21     product_id INT,
22     quantity INT,
23     order_date DATETIME NOT NULL DEFAULT GETDATE(),
24     FOREIGN KEY (customer_id) REFERENCES customers(customer_id),
25     FOREIGN KEY (product_id) REFERENCES products(product_id)
26 );
```

A yellow status bar at the bottom right of the code area says 'Query succeeded | 1s'.



/ prayagv



/ prayag-verma

Prayag Verma | Data Engineer

Step 3

Creating TGT Tables

Also create three dimension tables 'customers_dim, products_dim, and orders_dim', using the same sql queries given below!

The screenshot shows the Microsoft Azure portal with the Query editor (preview) open for the database 'AzurePractice'. The left sidebar shows various database management options like Overview, Activity log, Tags, and Query editor (preview), which is currently selected. The main area displays the following SQL script:

```
1 DROP TABLE IF EXISTS customers_dim;
2 CREATE TABLE customers_dim (
3     customer_id INT PRIMARY KEY,
4     first_name NVARCHAR(50),
5     last_name NVARCHAR(50),
6     email NVARCHAR(100),
7     phone NVARCHAR(15),
8     Address NVARCHAR(255)
9 );
10 DROP TABLE IF EXISTS products_dim;
11 CREATE TABLE products_dim (
12     product_id INT PRIMARY KEY,
13     product_name NVARCHAR(100),
14     price DECIMAL(10, 2),
15     stock_quantity INT
16 );
17 DROP TABLE IF EXISTS orders_dim;
18 CREATE TABLE orders_dim (
19     order_id INT PRIMARY KEY,
20     customer_id INT,
21     product_id INT,
22     quantity INT,
23     order_date DATETIME NOT NULL DEFAULT GETDATE(),
24     FOREIGN KEY (customer_id) REFERENCES customers(customer_id),
25     FOREIGN KEY (product_id) REFERENCES products(product_id)
26 );
```

A message in the center states: "Showing limited object explorer here. For full capability please click here to open Azure Data Studio." At the bottom, a yellow bar indicates "Query succeeded | 0s".



/ prayagv



/ prayag-verma

Prayag Verma | Data Engineer

Step 4

Insert Sample Data into source

Next, Insert sample data into those newly created tables ‘customers, products, and orders’.

The screenshot shows the Microsoft Azure Query editor interface. On the left, there's a sidebar with various database management options like Overview, Activity log, Tags, Diagnose and solve problems, and a prominent 'Query editor (preview)' section which is selected. Below this, there's a tree view of tables: 'Tables' (dbo.customers, dbo.customers_dim, dbo.orders, dbo.orders_dim, dbo.products, dbo.products_dim), 'Views', and 'Stored Procedures'. The main area contains two queries. Query 1 is an 'INSERT INTO customers' statement with 7 rows of sample data. Query 2 is an 'INSERT INTO products' statement with 15 rows of sample data. Query 3 is an 'INSERT INTO orders' statement with 3 rows of sample data. A message box in the center says 'Showing limited object explorer here. For full capability please click here to open Azure Data Studio.' At the bottom, a yellow bar indicates 'Query succeeded | 1s'.

```
1 INSERT INTO customers (customer_id, first_name, last_name, email, phone, Address)
2 VALUES
3 (1, 'John', 'Doe', 'john.doe@example.com', '123-456-7890', '123 Main St, CityA, CountryX'),
4 (2, 'Jane', 'Smith', 'jane.smith@example.com', '234-567-8901', '456 Elm St, CityB, CountryY'),
5 (3, 'Alice', 'Johnson', 'alice.johnson@example.com', '345-678-9012', '789 Oak St, CityC, CountryZ'),
6 (4, 'Bob', 'Brown', 'bob.brown@example.com', '456-789-0123', '101 Pine St, CityD, CountryX'),
7 (5, 'Charlie', 'Davis', 'charlie.davis@example.com', '567-890-1234', '202 Maple St, CityE, CountryY');
8
9 INSERT INTO products (product_id, product_name, price, stock_quantity)
10 VALUES
11 (101, 'Laptop', 1200.00, 10),
12 (102, 'Smartphone', 800.00, 20),
13 (103, 'Tablet', 500.00, 15),
14 (104, 'Headphones', 150.00, 30),
15 (105, 'Smartwatch', 300.00, 25);
16
17 INSERT INTO orders (order_id, customer_id, product_id, quantity, order_date)
18 VALUES
19 (1006, 1, 101, 2, '2023-10-10 14:30:00'),
20 (1007, 2, 102, 1, '2023-10-11 09:15:00'),
21 (1008, 3, 103, GETDATE()); -- Current system timestamp (default)
```

Execute simple ‘select’ statement to preview inserted records into ‘orders’ table, you can check for other tables too and make sure data is inserted.

The screenshot shows the Microsoft Azure Query editor interface. The sidebar and tree view are identical to the previous screenshot. The main area contains a single query: 'SELECT * FROM [dbo].[orders];'. Below the query, there are tabs for 'Results' and 'Messages'. The 'Results' tab displays a table with three rows of data from the 'orders' table. The columns are 'order_id', 'customer_id', 'product_id', 'quantity', and 'order_date'. The data is as follows:

order_id	customer_id	product_id	quantity	order_date
1006	1	101	2	2023-10-10T14:30:00.0000000
1007	2	102	1	2023-10-11T09:15:00.0000000
1008	3	103	3	2025-02-24T02:48:48.7770000

Step 5

Insert Sample Data into TGT

Next, Insert sample data into those newly created tables ‘customers_dim, products_dim, and orders_dim’.

AzurePractice (Prayag.Verma@studenta...)

Query 1 × Query 3 × Query 4 ×

Run Cancel query Save query Export data as Show only Result Open Copilot

```
1 INSERT INTO [dbo].[orders_dim] (order_id, customer_id, product_id, quantity, order_date)
2 VALUES
3 (1006, 1, 101, 2, '2023-10-10 14:30:00'), -- Custom timestamp
4 (1007, 2, 102, 1, '2023-10-11 09:15:00'); -- Custom timestamp
5
6 INSERT INTO [dbo].[products_dim] (product_id, product_name, price, stock_quantity)
7 VALUES
8 (101, 'Laptop', 1200.00, 10),
9 (102, 'Smartphone', 800.00, 20),
10 (103, 'Tablet', 500.00, 15),
11 (104, 'Headphones', 150.00, 30),
12 (105, 'Smartwatch', 300.00, 25);
13
14 INSERT INTO [dbo].[customers_dim] (customer_id, first_name, last_name, email, phone, Address)
15 VALUES
16 (1, 'John', 'Doe', 'john.doe@example.com', '123-456-7890', '123 Main St, CityA, CountryX'),
17 (2, 'Jane', 'Smith', 'jane.smith@example.com', '234-567-8901', '456 Elm St, CityB, CountryY'),
18 (3, 'Alice', 'Johnson', 'alice.johnson@example.com', '345-678-9012', '789 Oak St, CityC, CountryZ'),
19 (4, 'Bob', 'Brown', 'bob.brown@example.com', '456-789-0123', '101 Pine St, CityD, CountryX'),
20 (5, 'Charlie', 'Davis', 'charlie.davis@example.com', '567-890-1234', '202 Maple St, CityE, CountryY');
```

Execute simple ‘select’ statement to preview inserted records into ‘orders_dim’ table, you can check for other tables too and make sure data is inserted.

AzurePractice (Prayag.Verma@studenta...)

Query 1 × Query 3 × Query 4 × Query 5 ×

Run Cancel query Save query Export data as Show only Editor Open Copilot

```
1 SELECT * FROM [dbo].[orders_dim];
```

Tables

Results Messages

order_id	customer_id	product_id	quantity	order_date
1006	1	101	2	2023-10-10T14:30:00.0000000
1007	2	102	1	2023-10-11T09:15:00.0000000

Step 6

Azure Data Factory

Next, return to the “Azure Portal” and navigate to “Azure Data Factory”, to set up a new pipeline design to handle our incremental data loading in analytical database.

Create a new ADF env if not already, then click on “Launch studio”

The screenshot shows the Microsoft Azure portal interface for managing Data Factories. On the left, a sidebar lists 'Data factories' under 'Student Ambassadors (stdnpartners.onmicrosoft.com)'. A specific factory named 'ADFincremntalLoad' is selected. The main content area displays the 'Overview' tab for this factory, showing its status as 'Succeeded' and providing details about its resource group ('practice dev env rg'), location ('East US'), and subscription ('Azure subscription 1'). Below the overview, there's a large blue button labeled 'Launch studio'. At the bottom of the page, there are links to 'Quick Starts', 'Tutorials', 'Template Gallery', and 'Training Modules'.



/ prayagv



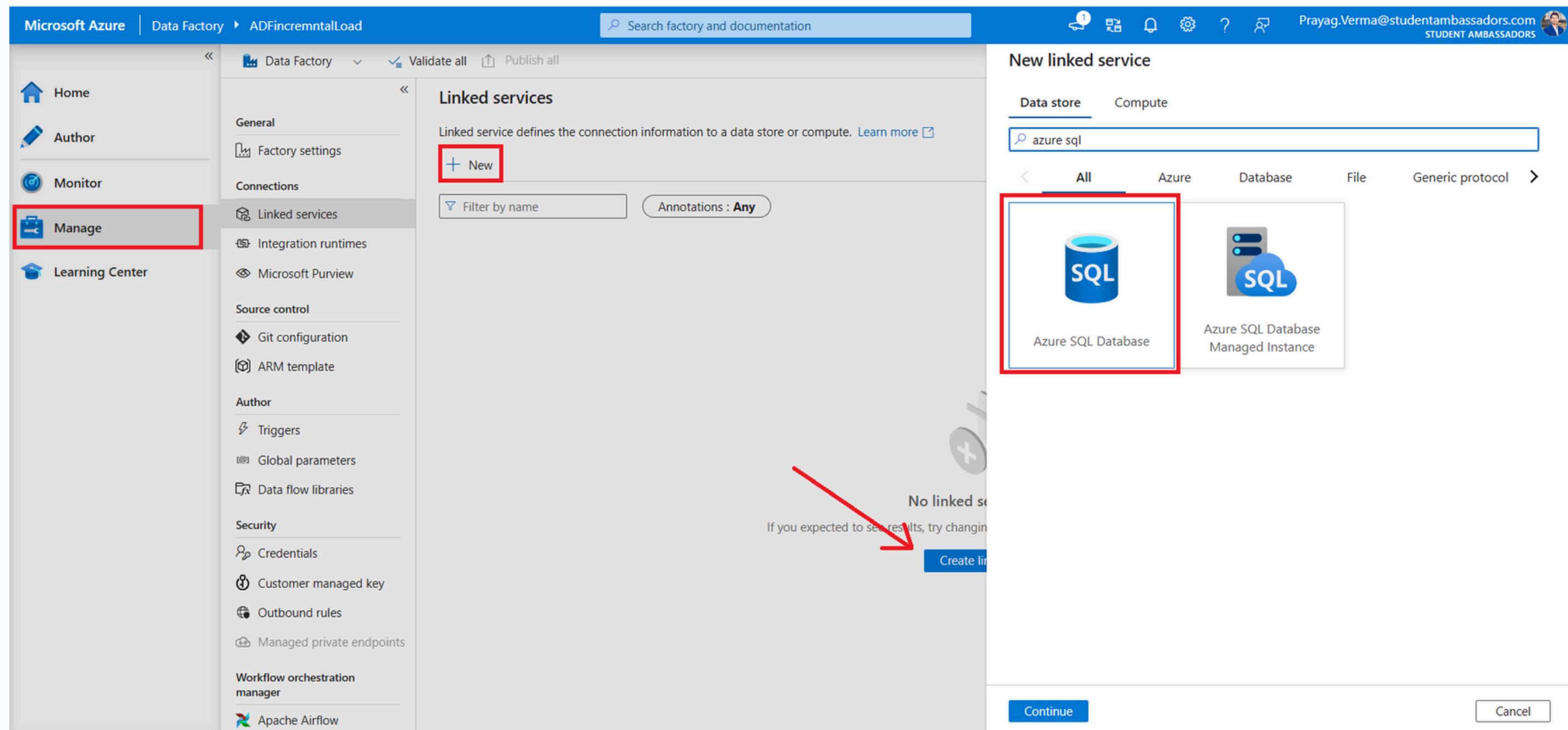
/ prayag-verma

Prayag Verma | Data Engineer

Step 7

Create Linked services

"Once the Azure Portal is loaded, click on 'Manage', then 'Linked Services'. From there, either select '+ New' or 'Create Linked Service', which will appear in the center of the dashboard. then, In the search bar, type 'Azure SQL' and choose the 'Azure SQL Database' option. Then, click 'Continue' and follow the remaining steps to create a new Linked Service.



/ prayagv



/ prayag-verma

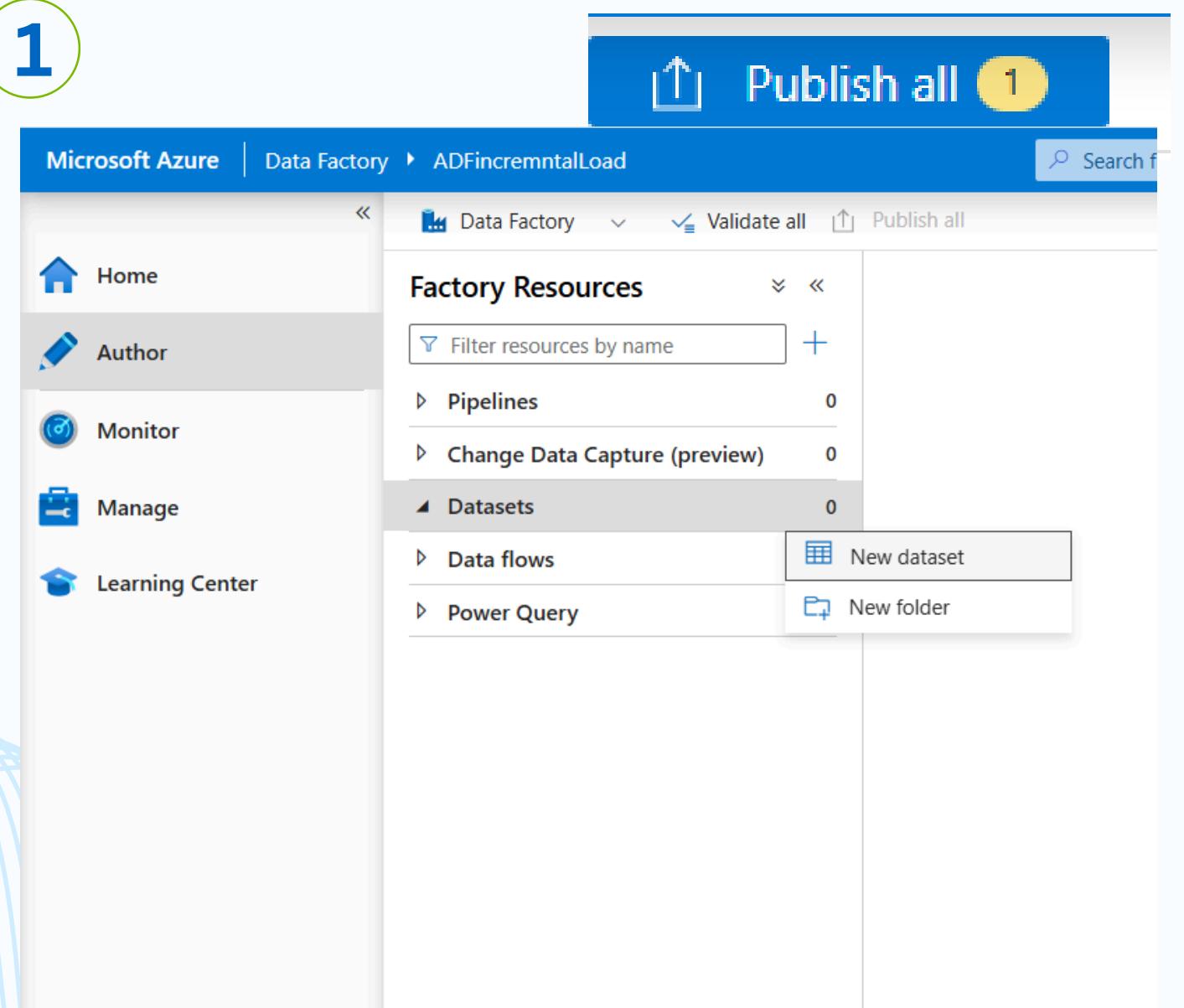
Prayag Verma | Data Engineer

Step 8

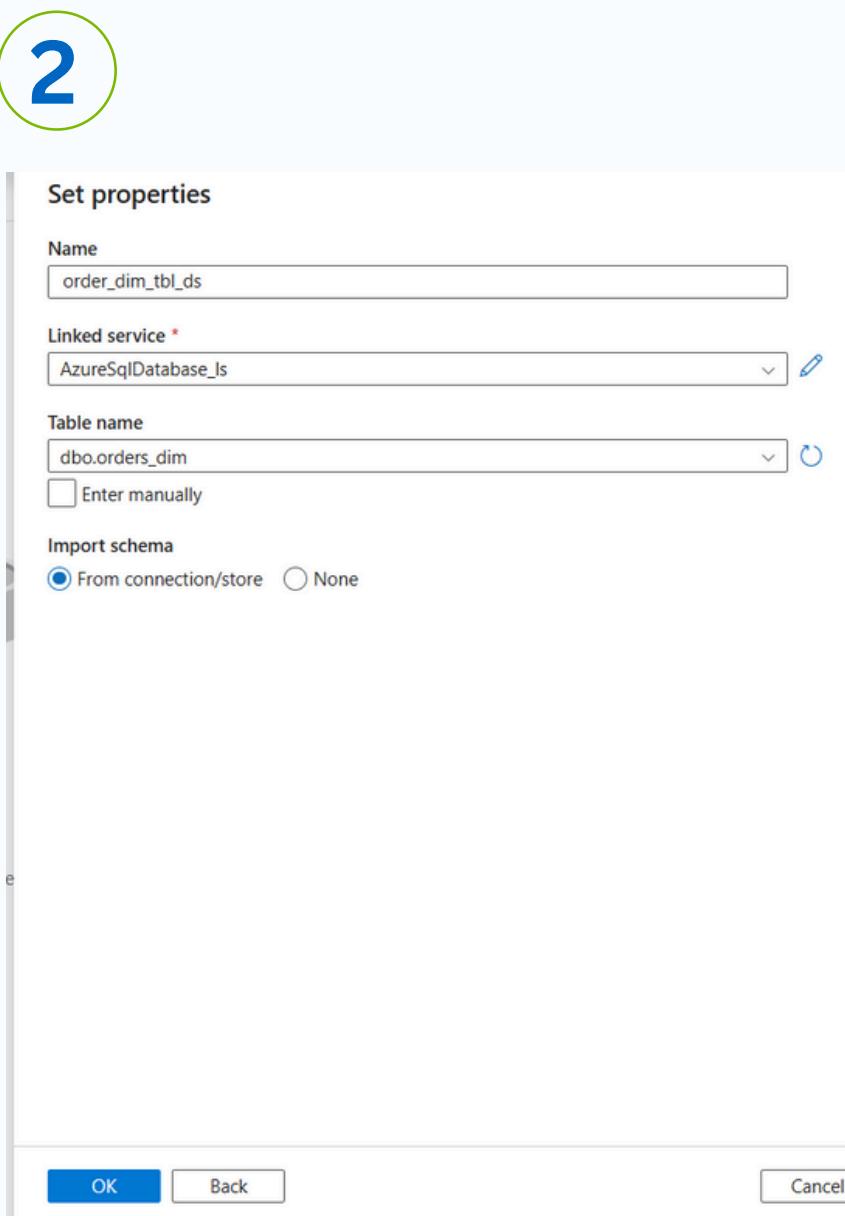
Create Datasets

Once the Linked Service is set up, go to the ‘Author’ section and click on ‘Datasets’. In the search bar, type ‘Azure SQL’ and choose ‘Azure SQL Database’. Next, name the dataset as ‘order_dim_tbl_ds’, select the Linked Service you created earlier, and pick the ‘dbo.order_dim’ table from the dropdown. Then select the ‘Import schema’ option to ‘From connection/store’, then click ‘OK’ to confirm. Finally, save your changes by clicking ‘Publish all’ option visible at the top!

1

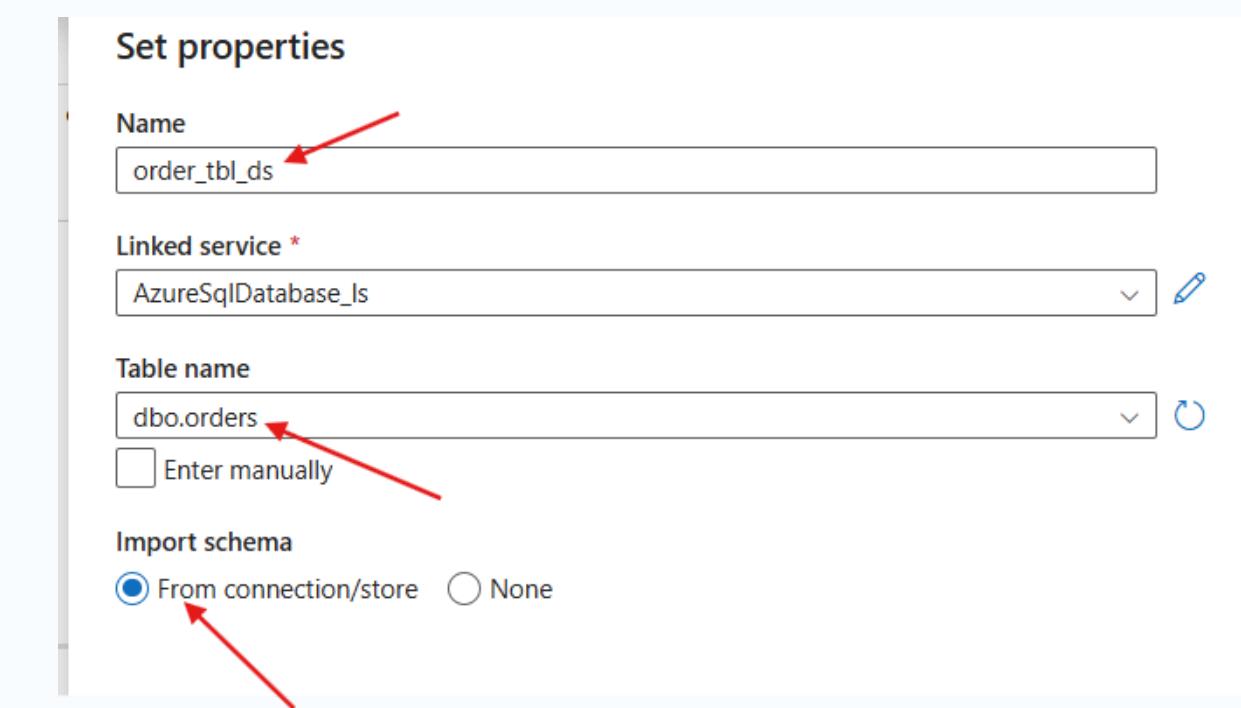


2



3

Follow the same steps and create another datasets name it as ‘order_tbl_ds’



/ prayagv



/ prayag-verma

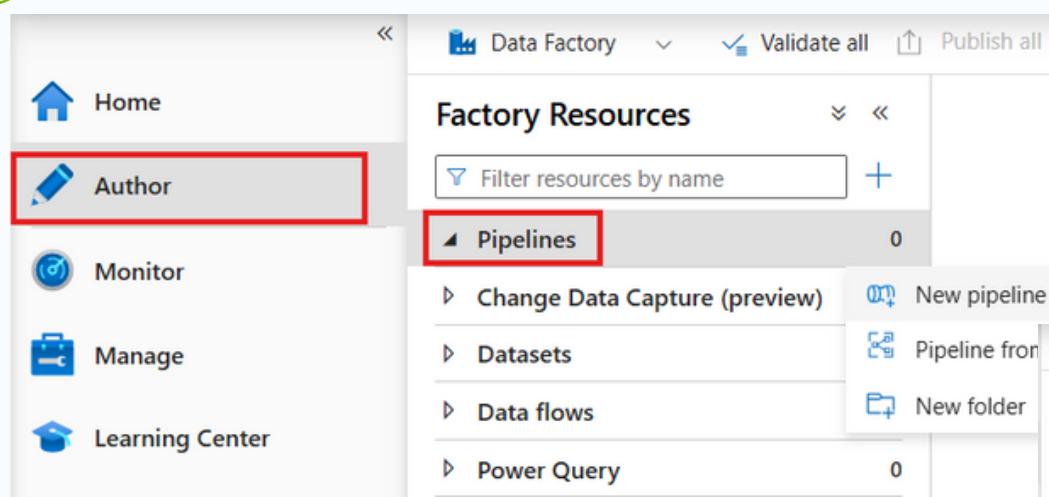
Prayag Verma | Data Engineer

Step 9

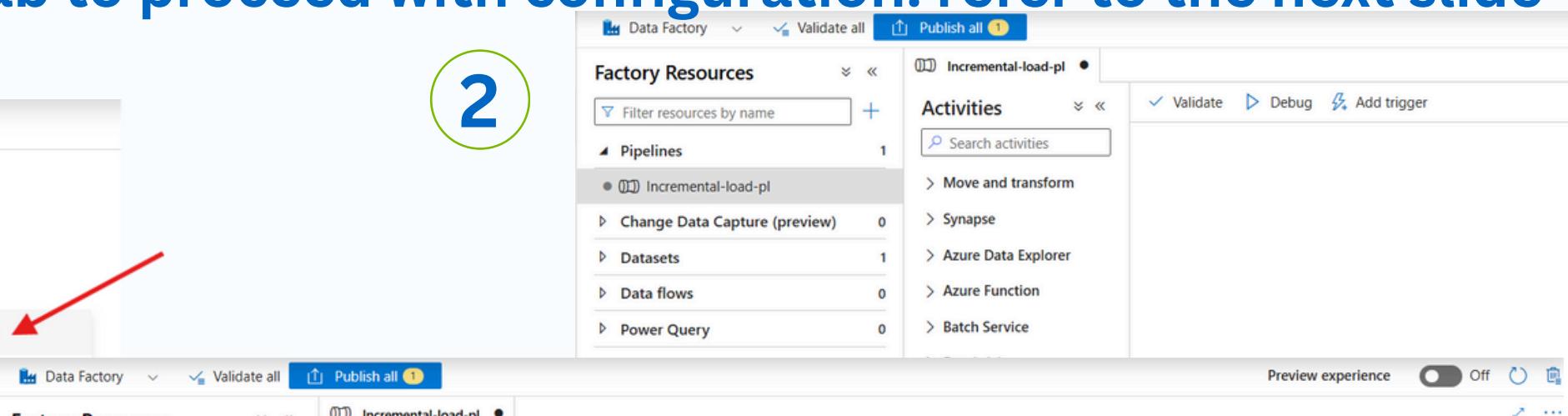
Create Pipeline & Lookup Activities

Next, hover your mouse over ‘Pipelines’ section. Three dots (...) will appear—click on it and choose the ‘New pipeline’ option. Name the pipeline ‘Incremental-load-pl’. After that, search for ‘Lookup’ activity in the Activities panel, drag it to the design area on the right, and rename this activity to ‘lookup-order_dim-tbl’. Finally, click on the ‘Settings’ tab to proceed with configuration. refer to the next slide-->

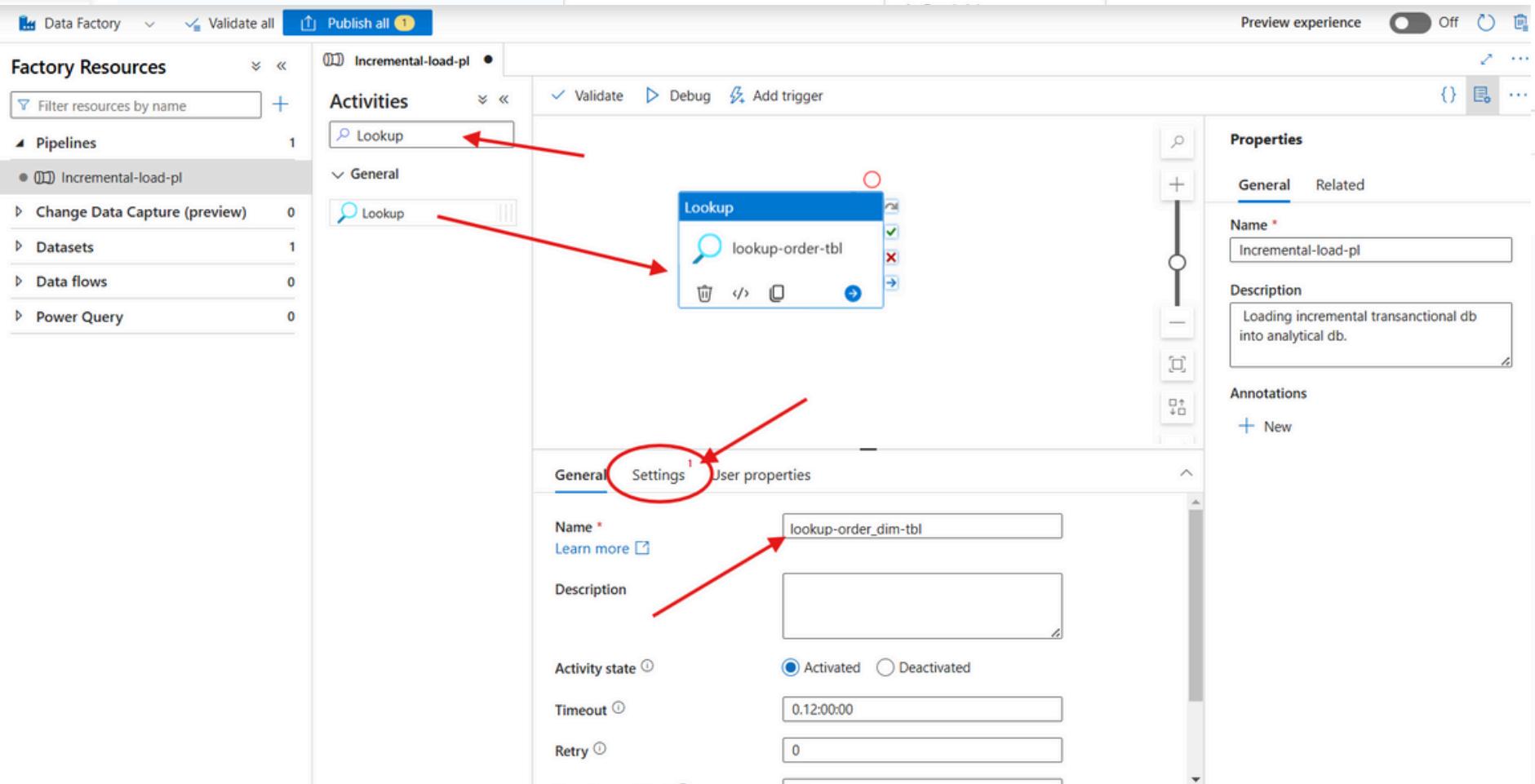
1



2



3



/ prayagv



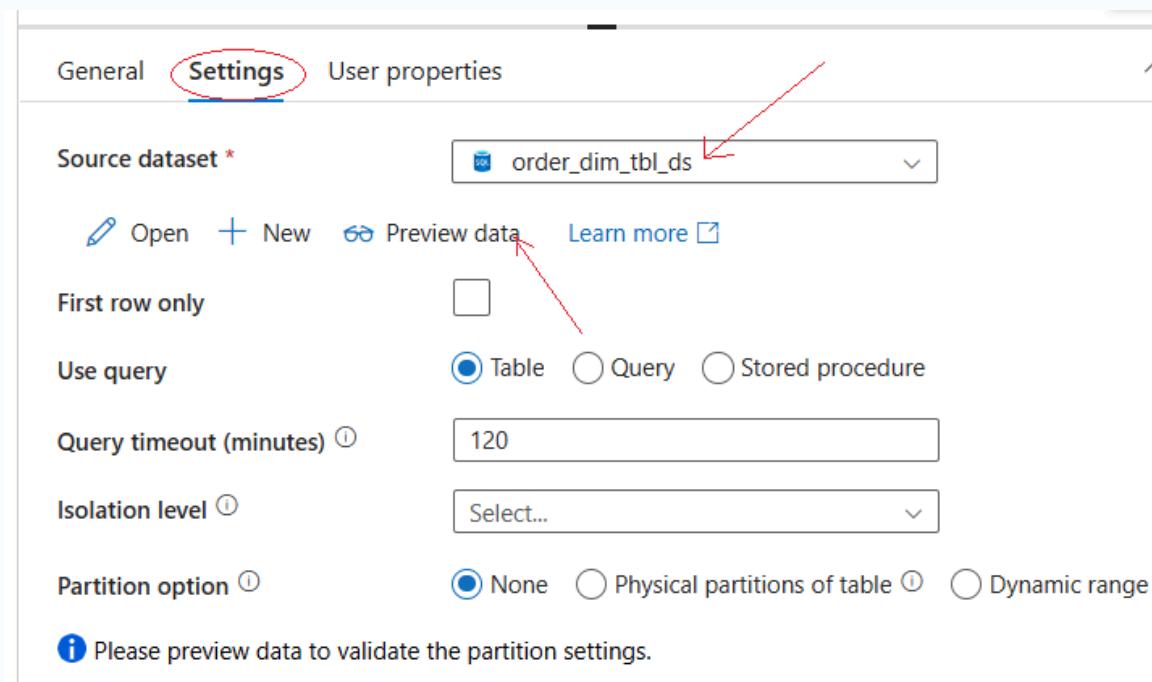
/ prayag-verma

Prayag Verma | Data Engineer

Step 9

Create Pipeline & Lookup Activities

Once you click on ‘Settings,’ choose the ‘order_dim_tbl_ds’ dataset that we created earlier. Then, click on ‘Preview data’ to display two records from the ‘order_dim’ table.



The 'Preview data' modal is open, showing the following table:

	order_id	customer_id	product_id	quantity	order_date
1	1006	1	101	2	2023-10-10T14:30:00
2	1007	2	102	1	2023-10-11T09:15:00

Since we need the most recent order_date (i.e., the MAX(order_date)), select the ‘Query’ option and enter the given query in the text field, as shown.

The screenshot shows the 'Settings' tab selected. Under 'Use query', the 'Query' radio button is selected and highlighted with a red circle and a green circle labeled '3'. The 'Query' text input field contains the following SQL code: 'SELECT MAX(order_date) as most_recent_ord_date FROM [dbo].[orders_dim];'. Other settings remain the same as in step 1.

Step 10

Create Copy Activity

Click 'Preview data' again and this time, latest or max(order_date) value will be displayed from 'order_dim' table.

The screenshot shows the 'Settings' tab of the dataset configuration. The 'Source dataset' dropdown is set to 'order_dim_tbl_ds'. A red arrow points to the 'Preview data' button. Other settings include 'First row only' (unchecked), 'Use query' (radio button selected), 'Query timeout (minutes)' (120), 'Isolation level' (Select...), and 'Partition option' (None). A note at the bottom says 'Please preview data to validate the partition settings.'

1

2

3

Preview data

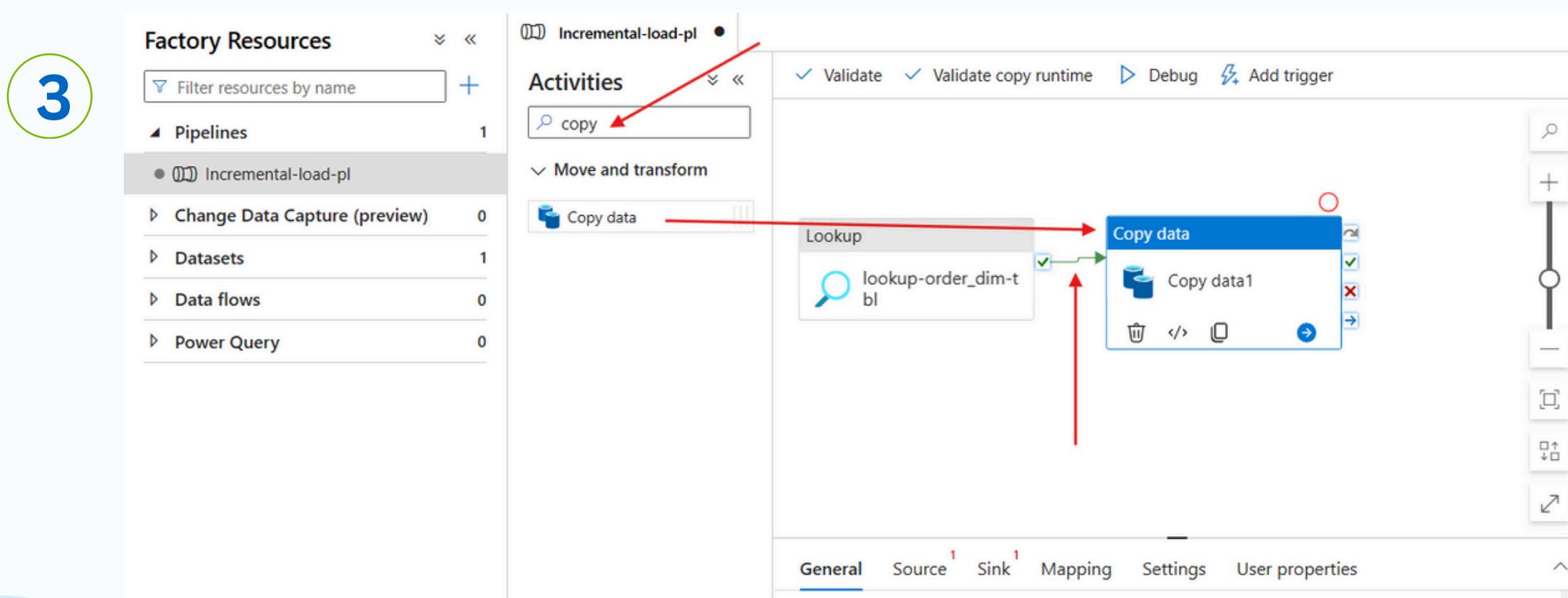
Linked service: AzureSqlDatabase_ls

Object:

most_recent_ord_date

1 2023-10-11T09:15:00

Go to the activities section, find 'Copy Data', drag it next to the Lookup activity, and connect the two as shown.



/ prayagv



/ prayag-verma

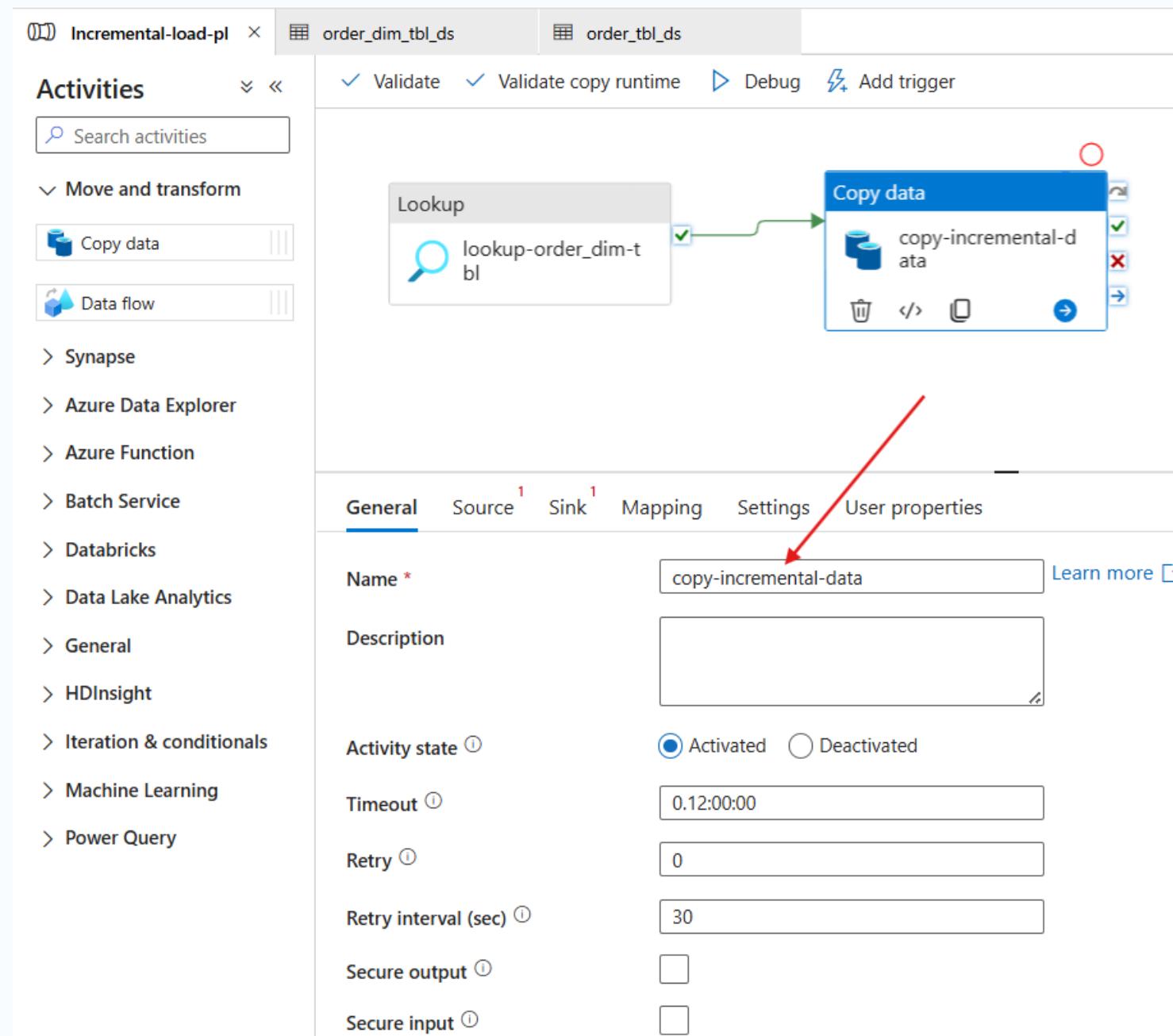
Prayag Verma | Data Engineer

Step 11

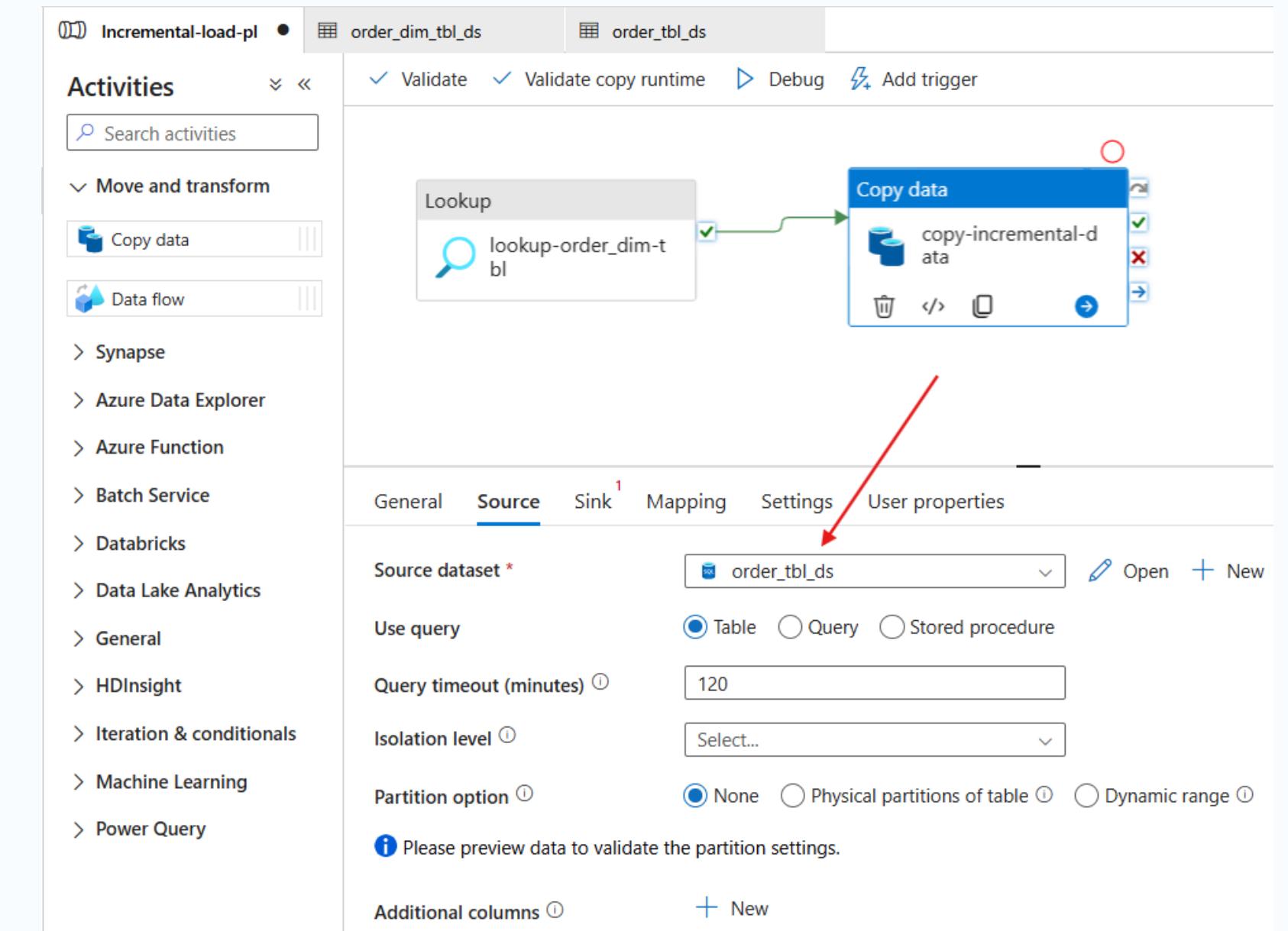
Configure Copy Activity

Rename the copy activity to ‘copy-incremental-data’ under ‘General’ tab, then go to ‘Source’ tab, and select ‘order_tbl_ds’ because, we’ll copy new records from the ‘order’ table (transactional) into the ‘order_dim’ table on a daily basis.

1



2



/ prayagv



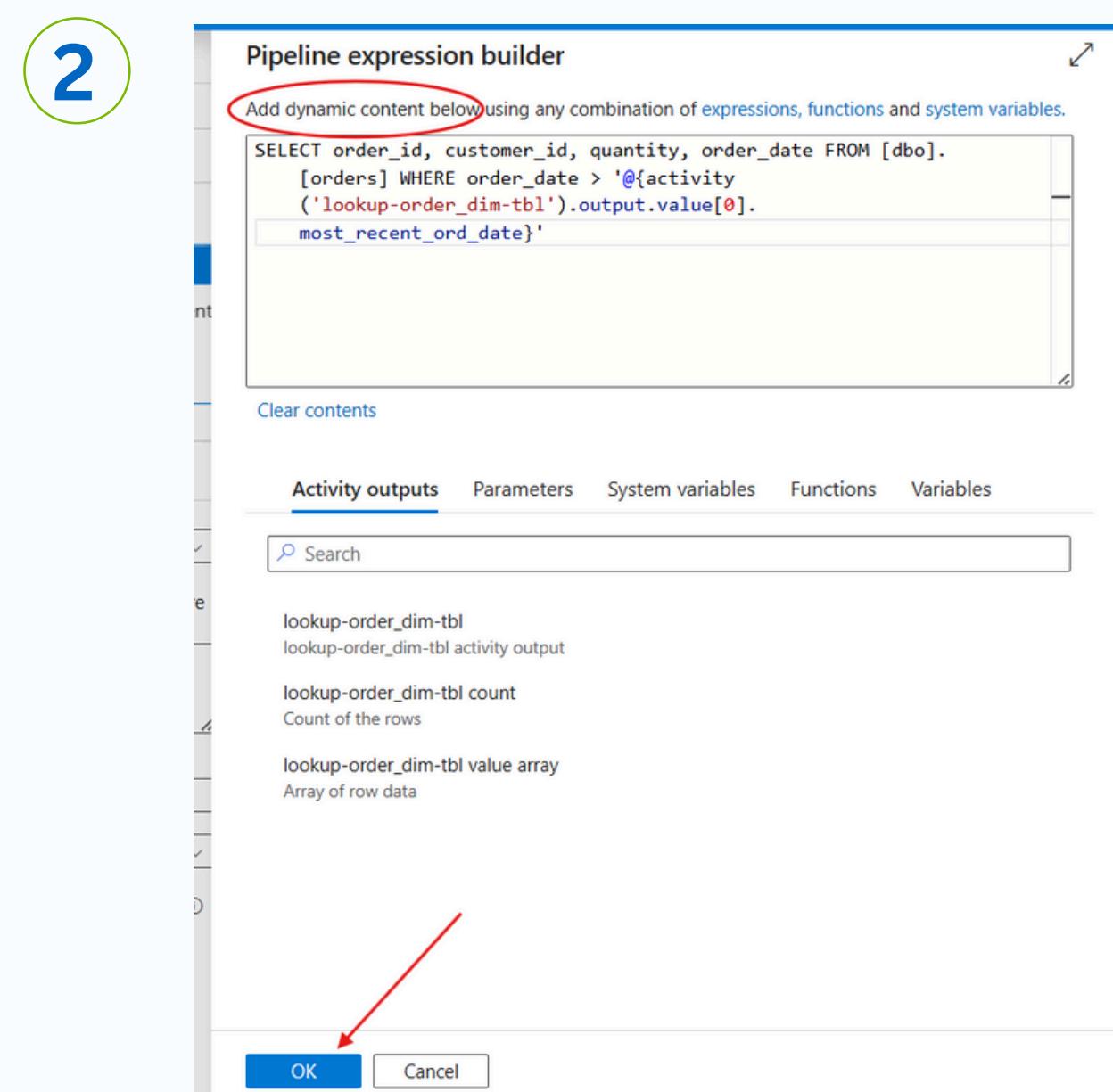
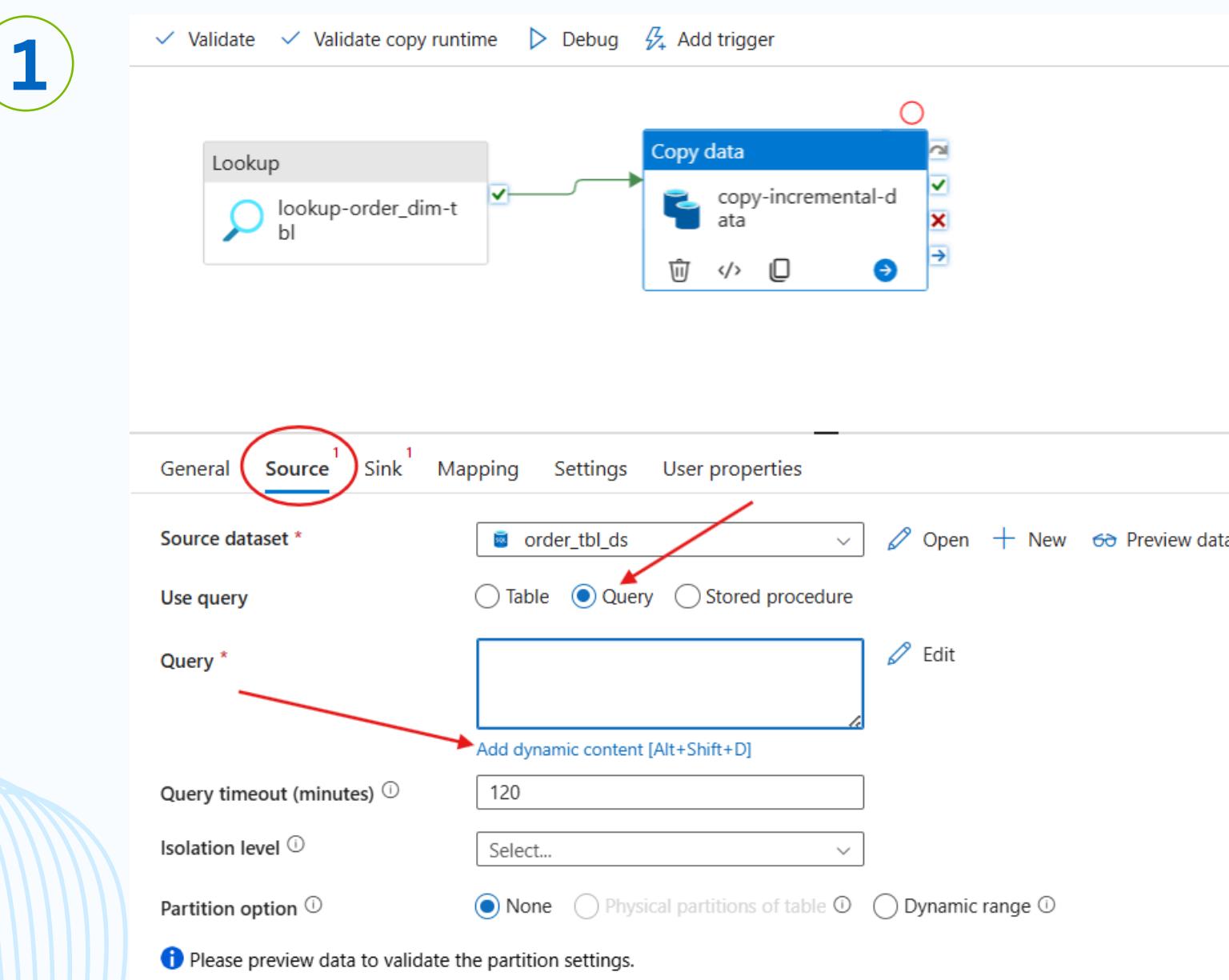
/ prayag-verma

Prayag Verma | Data Engineer

Step 11

Configure Copy Activity

Under the 'Source' tab in the Copy activity, select 'Query' and click on 'Add dynamic content'. We'll set a condition to load new records from the 'order' table into 'order_dim' only if the 'order_date' in 'order' is greater than the maximum 'order_date' in 'order_dim' (fetched via the Lookup activity using the 'max()' function). This condition is crucial before loading any new records into 'order_dim' to ensure incremental loading only.



Note:

- 1 If the 'First row only' option in the Lookup activity is unchecked, the value of MAX(order_date) as most_recent_ord_date is returned as an array. In this case, you must use '.output.value[0].<column_name>' to access the value.

```
SELECT order_id, customer_id, quantity, order_date FROM [dbo].[orders] WHERE order_date > '@{activity('lookup-order_dim-tbl').output.value[0].most_recent_ord_date}'
```

- 2 If the 'First row only' option in the Lookup activity is checked, the value of MAX(order_date) as most_recent_ord_date is returned as a single object. In this case, you must use '.output.firstRow.<column_name>' to access the value.

```
SELECT order_id, customer_id, quantity, order_date FROM [dbo].[orders] WHERE order_date > '@{activity('lookup-order_dim-tbl').output.firstRow.most_recent_ord_date}'
```

Step 12

Run ADF Pipeline

Now, let's run the pipeline...

The screenshot shows the Microsoft Azure Data Factory pipeline editor interface. On the left, the 'Factory Resources' sidebar lists 'Pipelines' (Incremental-load-pl), 'Datasets' (order_dim_tbl_ds, order_tbl_ds), and other resources like 'Change Data Capture (preview)', 'Data flows', and 'Power Query'. The main workspace displays the 'Incremental-load-pl' pipeline. It consists of two activities: 'Lookup' (lookup-order_dim-tbl) and 'Copy data' (copy-incremental-data). The 'Copy data' activity has a green checkmark indicating success. Below the activities, the 'Output' tab shows the pipeline run details. The 'Pipeline run ID' is 207d4e19-1482-4e22-b5e0-0d6b881f92b5, and the 'Pipeline status' is 'Succeeded'. The table below lists the activities and their statuses:

Activity name	Activity st...	Activit...	Run start	Duration	Integration runtime	User prop...
copy-incremental-data	Succeeded	Copy data	2/24/2025, 1:30:42 AM	20s	AutoResolveIntegrationRuntime (Central India)	
lookup-order_dim-tbl	Succeeded	Lookup	2/24/2025, 1:30:21 AM	20s	AutoResolveIntegrationRuntime (Central India)	



/ prayagv



/ prayag-verma

Prayag Verma | Data Engineer

Step 13

Validate Data Loading...

The pipeline ran successfully, and the third record was loaded into ‘order_dim’, which previously had only two records (refer to slide 8).

The screenshot shows the Microsoft Azure Query editor interface. The left sidebar displays the database structure for 'AzurePractice' with the 'Query editor (preview)' tab selected. The main area shows a query window with 'Query 5' selected, containing the following SQL code:

```
1  SELECT order_id, customer_id, quantity, order_date FROM [dbo].[orders_dim];
```

The results pane displays the following data:

order_id	customer_id	quantity	order_date
1006	1	2	2023-10-10T14:30:00.0000000
1007	2	1	2023-10-11T09:15:00.0000000
1008	3	3	2025-02-24T02:48:48.7770000

A red arrow points from the text 'Try inserting new records into the “orders” table and rerun the pipeline to see the newly loaded records in the “orders_dim” table.' to the '1008' row in the results table.

Try inserting new records into the ‘orders’ table and rerun the pipeline to see the newly loaded records in the ‘orders_dim’ table.

Step 14

Creating Trigger

Go to ‘Manage’, select ‘Triggers’, and click on ‘Create trigger’ or ‘+ New’. Fill in the details as per when you want the trigger to run.

1

The screenshot shows the Microsoft Azure Data Factory interface. In the left sidebar, under the 'Manage' section, the 'Triggers' option is selected. A red circle highlights the '+ New' button in the center of the page. A red arrow points from the 'Triggers' section in the sidebar to the '+ New' button. The main area displays the message 'No triggers to show' and a 'Create trigger' button.

2

The screenshot shows the 'New trigger' configuration dialog. The 'Name' field is filled with 'DailyDelta_Load_order_table'. The 'Description' field contains the text: 'This trigger is meant to load incremental data from 'order' table to 'order_dim' table!'. The 'Type' is set to 'Schedule'. The 'Start date' is '2/23/2025, 2:42:00 AM'. The 'Time zone' is 'Central Time (US & Canada) (UTC-6)'. A note below states: 'This time zone observes daylight savings. Trigger will auto-adjust for one hour difference.' Under 'Recurrence', it is set to 'Every 15 Day(s)'. There is also an 'Advanced recurrence options' section and a 'Execute at these times' section with fields for 'Hours' and 'Minutes'.

Step 15

Trigger Run & Monitor

Go to 'Monitor', select 'Pipeline runs', you can see the pipeline status getting changed from 'In progress' to 'Succeeded'

1

The screenshot shows the Microsoft Azure Data Factory interface. On the left, there's a navigation sidebar with 'Home', 'Author', 'Monitor' (which is highlighted), 'Manage', and 'Learning Center'. The main area is titled 'Pipeline runs' and shows a single run for 'Incremental-load-pl'. The run started at 2/24/2025, 2:56:54 AM and ended at 2/24/2025, 2:57:43 AM, with a duration of 49s. It was triggered by a manual trigger and is marked as 'Succeeded'. A red arrow points from the 'Monitor' button in the sidebar to the 'Pipeline runs' section.

2

The screenshot shows the Microsoft Azure Data Factory interface. On the left, there's a navigation sidebar with 'Home', 'Author', 'Monitor' (which is highlighted), 'Manage', and 'Learning Center'. The main area is titled 'All pipeline runs > Incremental-load-pl - Activity runs'. It displays two activity runs: 'copy-incremental-data' (Copy data activity) and 'lookup-order_dim-tbl' (Lookup activity). Both activities are marked as 'Succeeded'. Below this, there's a table for 'Activity runs' showing the same two items with their details. A red arrow points from the 'Monitor' button in the sidebar to the 'Activity runs' section.



/ prayagv



/ prayag-verma

Prayag Verma | Data Engineer



- THANK YOU -

Prayag Verma | Data Engineer



[/prayagv](https://www.linkedin.com/in/prayagv)



[/prayag-verma](https://github.com/prayag-verma)

Follow Me To Learn More and Grow Together!