



Incremental Data



Incremental Loading in Azure Synapse Workspace



Small Project

Incremental Loading in Azure

Incremental loading refers to the process of loading only new or changed data from a data source into a destination storage or processing system.

Why do we need this concept?

It efficiently updates databases with only new or changed data, reducing load times and resource usage, thereby optimizing system performance and minimizing costs.

Want to see small example?

Yes? Let's go!



Incremental Loading in Azure Synapse Workspace

Create 5 tables in SSMS, two tables with int column and three tables with datetime column as delta column.

- Suppliers
- Products
- Employees
- Sales
- Inventory Logs

Run this **Create Table queries**, in SSMS:

```
CREATE TABLE Suppliers (  
    supplier_id INT PRIMARY KEY, -----Delta Column  
    name VARCHAR(255),  
    contact_name VARCHAR(255),  
    phone VARCHAR(50),  
    address Varchar(100)  
);
```

```
CREATE TABLE Products (  
    product_id INT PRIMARY KEY, -----Delta Column  
    name VARCHAR(255),  
    category VARCHAR(100),  
    price DECIMAL(10, 2),  
    stock INT,  
    supplier_id INT,  
    FOREIGN KEY (supplier_id) REFERENCES Suppliers(supplier_id)  
);
```

```
CREATE TABLE Employees (  
    employee_id INT PRIMARY KEY,  
    first_name VARCHAR(255),  
    last_name VARCHAR(255),  
    hire_date DATETIME,  
    last_review_date DATETIME, -----Delta Column  
    role VARCHAR(100)  
);
```

```
CREATE TABLE Sales (  
    sale_id INT PRIMARY KEY,  
    sale_date DATETIME, -----Delta Column  
    product_id INT,  
    quantity INT,  
    total_amount DECIMAL(10, 2),  
    cashier_id INT,  
    FOREIGN KEY (product_id) REFERENCES Products(product_id),  
    FOREIGN KEY (cashier_id) REFERENCES Employees(employee_id)  
);
```

```
CREATE TABLE Inventory_Logs (  
    log_id INT PRIMARY KEY,  
    product_id INT,
```

```
log_date DATETIME, -----Delta Column
change_quantity INT,
remaining_stock INT,
FOREIGN KEY (product_id) REFERENCES Products(product_id)
```

);

```
1 CREATE TABLE Suppliers (
2     supplier_id INT PRIMARY KEY, -----Delta Column
3     name VARCHAR(255),
4     contact_name VARCHAR(255),
5     phone VARCHAR(50),
6     address Varchar(100)
7 );
8
9 CREATE TABLE Products (
10     product_id INT PRIMARY KEY, -----Delta Column
11     name VARCHAR(255),
12     category VARCHAR(100),
13     price DECIMAL(10, 2),
14     stock INT,
15     supplier_id INT,
16     FOREIGN KEY (supplier_id) REFERENCES Suppliers(supplier_id)
17 );
18
19 CREATE TABLE Employees (
20     employee_id INT PRIMARY KEY,
21     first_name VARCHAR(255),
22     last_name VARCHAR(255),
23     hire_date DATETIME,
24     last_review_date DATETIME, -----Delta Column
25     role VARCHAR(100)
26 );
27
28 CREATE TABLE Sales (
29     sale_id INT PRIMARY KEY,
30     sale_date DATETIME, -----Delta Column
31     product_id INT,
32     quantity INT,
33     total_amount DECIMAL(10, 2),
34     cashier_id INT,
35     FOREIGN KEY (product_id) REFERENCES Products(product_id),
36     FOREIGN KEY (cashier_id) REFERENCES Employees(employee_id)
37 );
38
39 CREATE TABLE Inventory_Logs (
40     log_id INT PRIMARY KEY,
41     product_id INT,
42     log_date DATETIME, -----Delta Column
43     change_quantity INT,
44     remaining_stock INT,
45     FOREIGN KEY (product_id) REFERENCES Products(product_id)
```

Insert Data into those 5 tables as below:

----Insert data into the Supplier Table

```
INSERT INTO Suppliers (supplier_id, name, contact_name, phone, address) VALUES
(1, 'Fresh Farms', 'John Doe', '555-3489', '123 Farm Lane'),
(2, 'Healthy Beverages Co.', 'Emily Stone', '555-7623', '47 Beverage Blvd'),
(3, 'Premium Meats', 'Alan Smith', '555-9876', '233 Meat St'),
(4, 'Bakers Delight', 'Nora Bates', '555-4532', '88 Baker Rd');
```

----Insert data into the Products Table

```
INSERT INTO Products (product_id, name, category, price, stock, supplier_id) VALUES
(1, 'Organic Apples', 'Fruits', 2.99, 150, 1),
(2, 'Almond Milk', 'Beverages', 3.49, 85, 2),
(3, 'Chicken Breast', 'Meat', 7.99, 60, 3),
(4, 'Whole Wheat Bread', 'Bakery', 2.50, 120, 4);
```

----Insert data into Employee Table

```
INSERT INTO Employees (employee_id, first_name, last_name, hire_date, last_review_date, role) VALUES
(1, 'Raj', 'Sharma', '2022-01-05 09:00:00', '2023-09-10 00:00:00', 'Cashier'),
(2, 'Harpal', 'Vaghela', '2022-05-15 09:00:00', '2023-09-20 00:00:00', 'Cashier'),
(3, 'Amit', 'Singh', '2023-03-23 09:00:00', '2023-09-30 00:00:00', 'Stock Manager'),
(4, 'Anjali', 'Patel', '2023-04-10 09:00:00', '2023-10-01 00:00:00', 'Sales Manager');
```

----Insert data into the Sales Table

```
INSERT INTO Sales (sale_id, sale_date, product_id, quantity, total_amount, cashier_id) VALUES
(1, '2023-10-01 14:00:00', 1, 10, 29.90, 1),
(2, '2023-10-01 14:15:00', 2, 5, 17.45, 2),
(3, '2023-10-01 15:00:00', 3, 4, 31.96, 1),
(4, '2023-10-01 15:30:00', 4, 6, 15.00, 2);
```

----Insert data into Inventory_Logs Table

```
INSERT INTO Inventory_Logs (log_id, product_id, log_date, change_quantity, remaining_stock) VALUES
(1, 1, '2023-10-01 08:00:00', 20, 170),
(2, 2, '2023-10-01 09:00:00', -10, 75),
(3, 3, '2023-10-01 10:00:00', 30, 90),
(4, 4, '2023-10-01 11:00:00', -5, 115);
```

```
47
48 ----Insert data into Supplier Table
49 INSERT INTO Suppliers (supplier_id, name, contact_name, phone, address) VALUES
50 (1, 'Fresh Farms', 'John Doe', '555-3489', '123 Farm Lane'),
51 (2, 'Healthy Beverages Co.', 'Emily Stone', '555-7623', '47 Beverage Blvd'),
52 (3, 'Premium Meats', 'Alan Smith', '555-9876', '233 Meat St'),
53 (4, 'Bakers Delight', 'Nora Bates', '555-4532', '88 Baker Rd');
54
55 ----Insert data into Products Table
56 INSERT INTO Products (product_id, name, category, price, stock, supplier_id) VALUES
57 (1, 'Organic Apples', 'Fruits', 2.99, 150, 1),
58 (2, 'Almond Milk', 'Beverages', 3.49, 85, 2),
59 (3, 'Chicken Breast', 'Meat', 7.99, 60, 3),
60 (4, 'Whole Wheat Bread', 'Bakery', 2.50, 120, 4);
61
62 ----Insert data into Employee Table
63 INSERT INTO Employees (employee_id, first_name, last_name, hire_date, last_review_date, role) VALUES
64 (1, 'Raj', 'Sharma', '2022-01-05 09:00:00', '2023-09-10 00:00:00', 'Cashier'),
65 (2, 'Harpal', 'Vaghela', '2022-05-15 09:00:00', '2023-09-20 00:00:00', 'Cashier'),
66 (3, 'Amit', 'Singh', '2023-03-23 09:00:00', '2023-09-30 00:00:00', 'Stock Manager'),
67 (4, 'Anjali', 'Patel', '2023-04-10 09:00:00', '2023-10-01 00:00:00', 'Sales Manager');
68
69 ----Insert data into Sales Table
70 INSERT INTO Sales (sale_id, sale_date, product_id, quantity, total_amount, cashier_id) VALUES
71 (1, '2023-10-01 14:00:00', 1, 10, 29.90, 1),
72 (2, '2023-10-01 14:15:00', 2, 5, 17.45, 2),
73 (3, '2023-10-01 15:00:00', 3, 4, 31.96, 1),
74 (4, '2023-10-01 15:30:00', 4, 6, 15.00, 2);
75
76 ----Insert data into Inventory_Logs Table
77 INSERT INTO Inventory_Logs (log_id, product_id, log_date, change_quantity, remaining_stock) VALUES
78 (1, 1, '2023-10-01 08:00:00', 20, 170),
79 (2, 2, '2023-10-01 09:00:00', -10, 75),
80 (3, 3, '2023-10-01 10:00:00', 30, 90),
81 (4, 4, '2023-10-01 11:00:00', -5, 115);
82
```

Check the output of tables:

Select * From Suppliers
Select * From Products

Select * From Employees

Select * From Sales

Select * From Inventory_Logs

Results Messages

	supplier_id	name	contact_name	phone	address
1	1	Fresh Farms	John Doe	555-3489	123 Farm Lane
2	2	Healthy Beverages Co.	Emily Stone	555-7623	47 Beverage Blvd
3	3	Premium Meats	Alan Smith	555-9876	233 Meat St
4	4	Bakers Delight	Nora Bates	555-4532	88 Baker Rd

	product_id	name	category	price	stock	supplier_id
1	1	Organic Apples	Fruits	2.99	150	1
2	2	Almond Milk	Beverages	3.49	85	2
3	3	Chicken Breast	Meat	7.99	60	3
4	4	Whole Wheat Bread	Bakery	2.50	120	4

	employee_id	first_name	last_name	hire_date	last_review_date	role
1	1	Raj	Sharma	2022-01-05 09:00:00.000	2023-09-10 00:00:00.000	Cashier
2	2	Harpal	Vaghela	2022-05-15 09:00:00.000	2023-09-20 00:00:00.000	Cashier
3	3	Amit	Singh	2023-03-23 09:00:00.000	2023-09-30 00:00:00.000	Stock Manager
4	4	Anjali	Patel	2023-04-10 09:00:00.000	2023-10-01 00:00:00.000	Sales Manager

	sale_id	sale_date	product_id	quantity	total_amount	cashier_id
1	1	2023-10-01 14:00:00.000	1	10	29.90	1
2	2	2023-10-01 14:15:00.000	2	5	17.45	2
3	3	2023-10-01 15:00:00.000	3	4	31.96	1
4	4	2023-10-01 15:30:00.000	4	6	15.00	2

	log_id	product_id	log_date	change_quantity	remaining_stock
1	1	1	2023-10-01 08:00:00.000	20	170
2	2	2	2023-10-01 09:00:00.000	-10	75
3	3	3	2023-10-01 10:00:00.000	30	90
4	4	4	2023-10-01 11:00:00.000	-5	115

So we have data in those 5 tables now.

Let's create a **Watermark2 table** in SSMS. (I took watermark 2 as I already have the watermark table in the database)

```
CREATE TABLE dbo.WATERMARK2
(
    ID INT IDENTITY(1,1),
    TABLE_NAME VARCHAR(100),
    SCHEMA_NAME VARCHAR(100),
    FOLDER_NAME VARCHAR(100),
    LPV VARCHAR(100),
    DELTA_COLUMN VARCHAR(100)
)
```

```

97
98 CREATE TABLE dbo.WATERMARK2
99 (
100 ID INT IDENTITY(1,1),
101 TABLE_NAME VARCHAR(100),
102 SCHEMA_NAME VARCHAR(100),
103 FOLDER_NAME VARCHAR(100),
104 LPV VARCHAR(100),
105 DELTA_COLUMN VARCHAR(100)
106 )
107
108 Select * From WATERMARK2
109

```

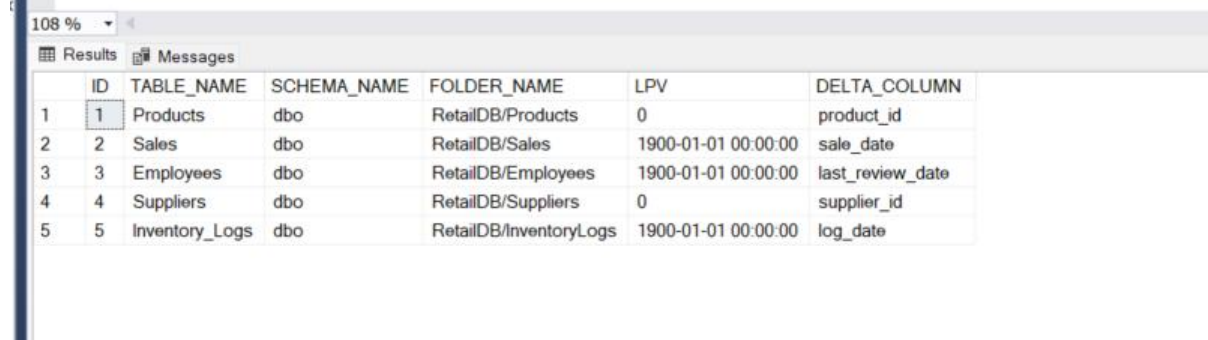
Let's insert data into the watermark2 table:

```

INSERT INTO dbo.WATERMARK2 VALUES ('Products', 'dbo', 'RetailDB/Products', 0, 'product_id');
INSERT INTO dbo.WATERMARK2 VALUES ('Sales', 'dbo', 'RetailDB/Sales', '1900-01-01 00:00:00',
'sale_date');
INSERT INTO dbo.WATERMARK2 VALUES ('Employees', 'dbo', 'RetailDB/Employees', '1900-01-01 00:00:00',
'last_review_date');
INSERT INTO dbo.WATERMARK2 VALUES ('Suppliers', 'dbo', 'RetailDB/Suppliers', 0, 'supplier_id');
INSERT INTO dbo.WATERMARK2 VALUES ('Inventory_Logs', 'dbo', 'RetailDB/InventoryLogs', '1900-01-01
00:00:00', 'log_date');

```

Watermark2 table output:



ID	TABLE_NAME	SCHEMA_NAME	FOLDER_NAME	LPV	DELTA_COLUMN
1	Products	dbo	RetailDB/Products	0	product_id
2	Sales	dbo	RetailDB/Sales	1900-01-01 00:00:00	sale_date
3	Employees	dbo	RetailDB/Employees	1900-01-01 00:00:00	last_review_date
4	Suppliers	dbo	RetailDB/Suppliers	0	supplier_id
5	Inventory_Logs	dbo	RetailDB/InventoryLogs	1900-01-01 00:00:00	log_date

Create a Stored Procedure to update the value of LPV as shown below:

```

18 -----Stored Procedure-----
19
20
21 CREATE PROC USP_WATERMARK2_UPDATE
22 @Table_Name VARCHAR(100),
23 @LPV_Value VARCHAR(50)
24 AS
25 UPDATE WATERMARK2
26 SET LPV = @LPV_Value
27 WHERE TABLE_NAME = @Table_Name
28
29

```

CREATE PROC USP_WATERMARK2_UPDATE

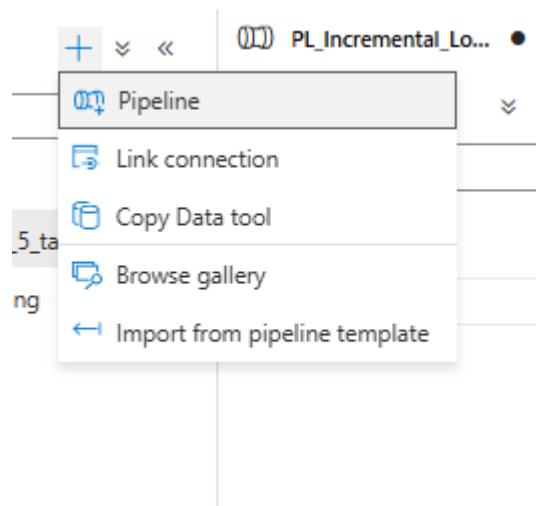
Harpalsinh Vaghela

```
@Table_Name VARCHAR(100),  
@LPV_Value VARCHAR(50)  
AS  
UPDATE WATERMARK2  
SET LPV = @LPV_Value  
WHERE TABLE_NAME = @Table_Name
```

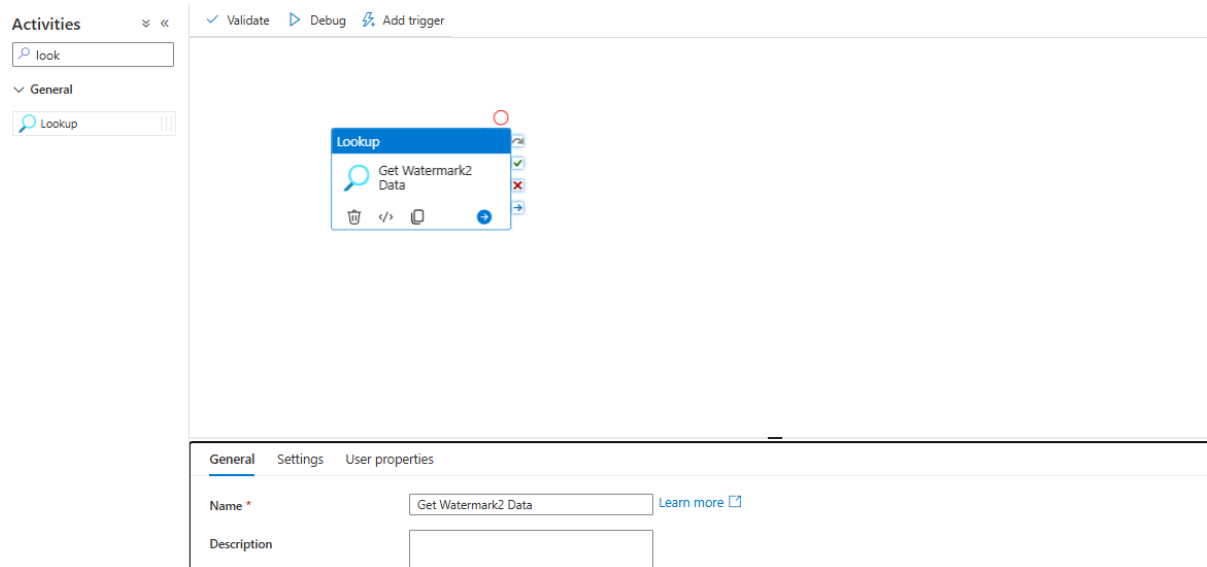
Let's Design the Pipeline

Open Azure Portal and go to **Azure Synapse Workspace**.

Integrate Tab -> Create New Pipeline



Drag and drop **Lookup Activity** and rename it “**Get Watermark2 Data**”



Go to the **Settings** tab in the lookup activity and select **the source dataset** if you already have it or select new.

Select **Azure SQL Database** and Create **Linked Service** as shown below:

Name *

ls_azuresql_metadata

Description

Connect via integration runtime * ⓘ

✓ AutoResolveIntegrationRuntime

Version

☐ Recommended ☒ Legacy

Connection string

Azure Key Vault

Account selection method ⓘ

☐ From Azure subscription ☒ Enter manually

Fully qualified domain name *

sqlserver-harpal.database.windows.net

Database name *

sqldatabase

Authentication type *

SQL authentication

User name *

adminharpal

Password

Azure Key Vault

Password *

.....

Always encrypted ⓘ

☐

Additional connection properties

+ New

✓ Connection successful

Test the Connection.

Go to the **Parameters tab** and write 2 parameters

- **SchemaName**
- **TableName**

Connection Schema **Parameters**

[+ New](#) | [Delete](#)

<input type="checkbox"/>	Name	Type	Default value	
<input type="checkbox"/>	SchemaName	String	Value	Delete
<input type="checkbox"/>	TableName	String	Value	Delete

Go to the **Connection Tab** and Assign those 2 parameters

Click on **Enter Manually** and Select those parameters in Table files as shown below in the connection tab, step by step:

Pipeline expression builder

Add dynamic content below using any combination of [expressions](#), [functions](#) and [system variables](#)

@dataset().SchemaName

[Clear contents](#)

Parameters Functions

[Search](#)

SchemaName

TableName

Pipeline expression builder

Add dynamic content below using any combination of [expressions](#), [functions](#) and [system variables](#)

@dataset().TableName

[Clear contents](#)

Parameters Functions

[Search](#)

SchemaName

TableName

After assigning it will look like this:

Connection Schema Parameters

Linked service * Is_azureql_metadata [Test connection](#) [Edit](#) [+ New](#) [Learn more](#)

Integration runtime * AutoResolveIntegrationRuntime [Edit](#)

Table @dataset().SchemaName . @dataset().TableName [Preview data](#)

☒ Enter manually

Go back to the **Settings tab** of the Lookup activity

Here, **Uncheck the First Row Only option** and Provide the Schema Name and Table name of the Watermark table which is in SSMS.

General **Settings** User properties

Source dataset * azuresqllookup [Open](#) [New](#) [Preview data](#) [Learn more](#)

Dataset properties ⓘ

Name	Value
SchemaName	dbo
TableName	Watermark2

First row only ☐

Use query ☒ Table ☐ Query ☐ Stored procedure

Query timeout (minutes) ⓘ

Isolation level ⓘ

Click on Preview Data

Preview data

Linked service: ls_azuresql_metadata

Object: @dataset().SchemaName.@dataset().TableName

ID	TABLE_NAME	SCHEMA_NAME	FOLDER_NAME	LPV	DELTA_COLUMN
1	Products	dbo	RetailDB/Products	0	product_id
2	Sales	dbo	RetailDB/Sales	1900-01-01 00:00:00	sale_date
3	Employees	dbo	RetailDB/Employees	1900-01-01 00:00:00	last_review_date
4	Suppliers	dbo	RetailDB/Suppliers	0	supplier_id
5	Inventory_Logs	dbo	RetailDB/InventoryLogs	1900-01-01 00:00:00	log_date

Here, we are done with this first lookup activity.

Now, Let's **ForEach Activity** in the pipeline design.

Connect Lookup with ForEach as shown below and Go to the **Settings** tab of **ForEach Activity**

General **Settings** Activities (0) User properties

Sequential ☐

Batch count ⓘ

Items *

This property should be parameterized.

[Add dynamic content \[Alt+Shift+D\]](#)

Check on **Items** and then there will be blue text as “**Add dynamic content**” Click on it and **assign the value array** as shown below:

Pipeline expression builder

Add dynamic content below using any combination of [expressions](#), [functions](#) and [system variables](#)

```
@activity('Get Watermark2 Data').output.value
```

[Clear contents](#)


Activity outputs

Parameters

System variables

Functions

Variables

 Search

Get Watermark2 Data

Get Watermark2 Data activity output

Get Watermark2 Data count

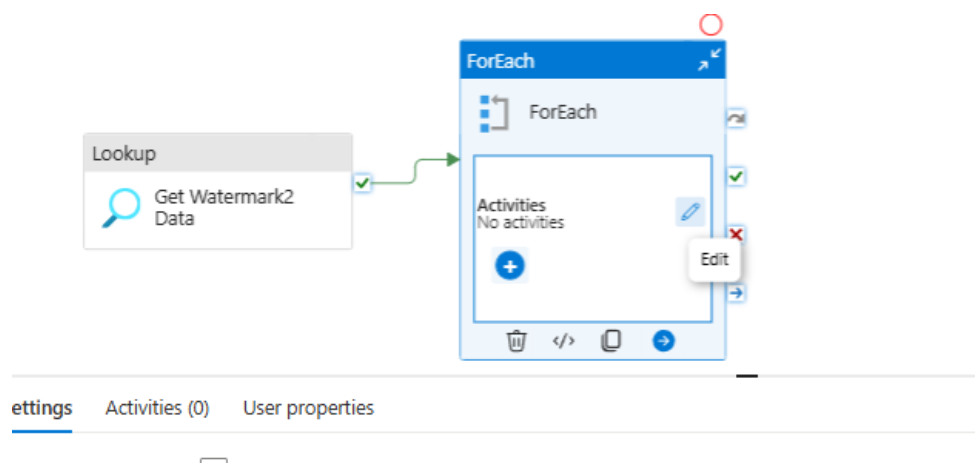
Count of the rows

Get Watermark2 Data value array

Array of row data

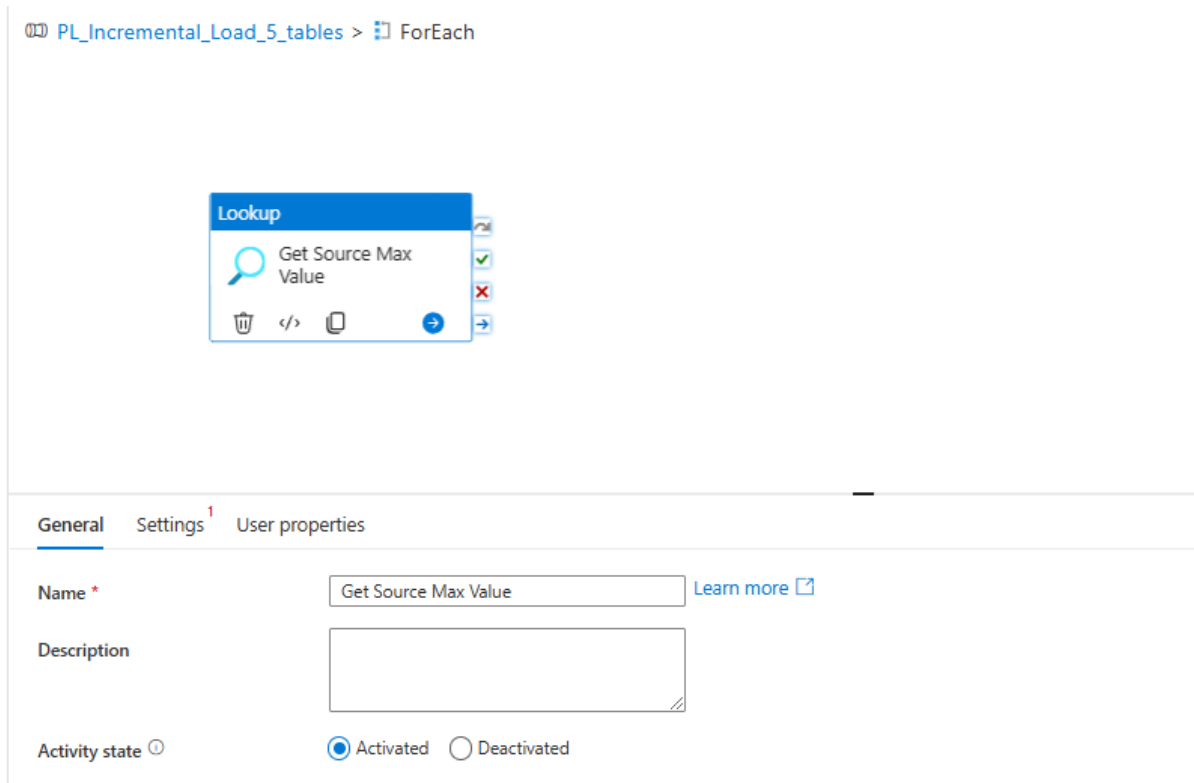
Click OK.

Now Click on the **Edit icon** of **ForEach Activity**



As we can see we are inside the ForEach Activity now.

Drag and drop New **Lookup Activity** and rename it as **Get Source Max Value**.



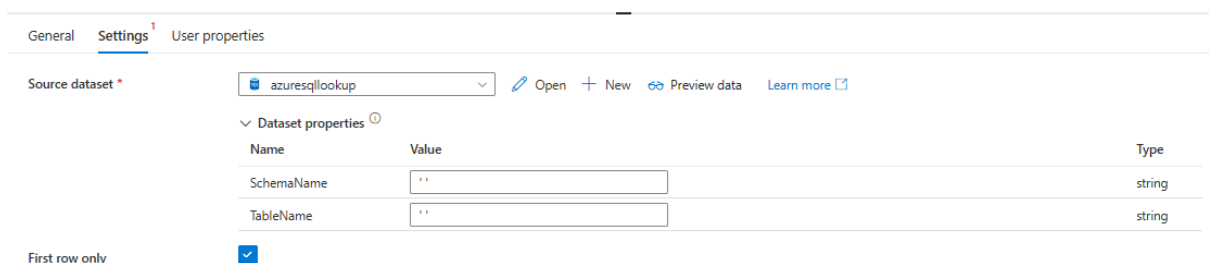
The screenshot shows the 'Lookup' activity configuration pane in Azure Data Studio. The activity is named 'Get Source Max Value' and is currently in the 'General' tab. The 'Name' field is set to 'Get Source Max Value'. The 'Description' field is empty. The 'Activity state' is set to 'Activated'. The 'Settings' tab is also visible, showing the 'Source dataset' as 'azuresqllookup' and the 'First row only' checkbox checked.

Name	Value	Type
SchemaName	' '	string
TableName	' '	string

Go to the **Settings tab** and select the source dataset which we have selected before for the first lookup activity from the dropdown list.

Enter a **single quotation** in the value fields for Schema and Table Name as indicated below. This effectively assigns a space value, signifying that no actual Schema and Table Names are needed. **Our objective is solely to obtain the maximum value for our LPV column, thus these fields can be left unspecified:**

Make sure to tick mark on the **First row only** checkbox



The screenshot shows the 'Settings' tab of the 'Lookup' activity configuration pane. The 'Source dataset' is set to 'azuresqllookup'. The 'Dataset properties' table shows 'SchemaName' and 'TableName' both set to ' ' (single space). The 'First row only' checkbox is checked.

Name	Value	Type
SchemaName	' '	string
TableName	' '	string

Go to Use query options and select the **Query** radio button

Click on **Add Dynamic Content** and write this query in the dialogue opened.

Select MAX(@{item().DELTA_COLUMN}) as MaxValue from
@{item().SCHEMA_NAME}.@{item().TABLE_NAME}

Pipeline expression builder

Add dynamic content below using any combination of [expressions](#), [functions](#) and [system variables](#).

```
Select MAX(@{item().DELTA_COLUMN}) as MaxValue from @{item().SCHEMA_NAME}.@{item().TABLE_NAME}
```

[Clear contents](#)

ForEach iterator Activity outputs Parameters System variables ...

Search

ForEach
Current item

After that, the settings tab of the lookup activity in ForEach Activity will look like this:

General **Settings** User properties

Source dataset * azuresqllookup [Open](#) [New](#) [Preview data](#) [Learn more](#)

Dataset properties

Name	Value
SchemaName	
TableName	

First row only ☒

Use query ☐ Table ☒ Query ☐ Stored procedure

Query Select MAX(@{item().DELTA_COLUMN...

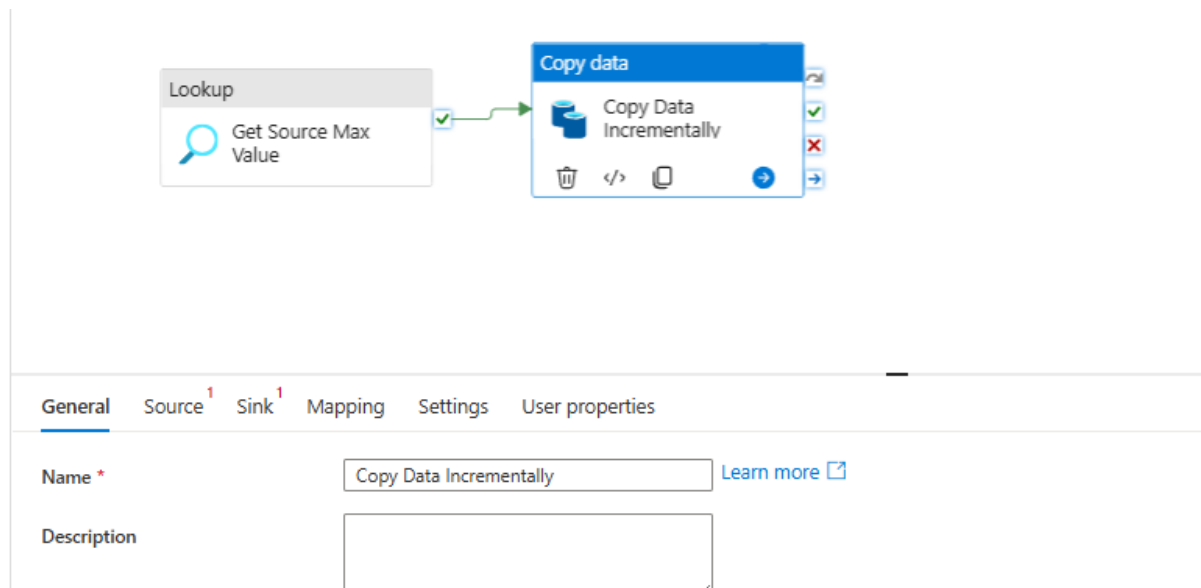
Query timeout (minutes) 120

Isolation level Select...

Partition option ☒ None ☐ Physical partitions of table ☐ Dynamic range

Please preview data to validate the partition settings.

Now, drag and drop the **Copy Data Activity** in the design area and **connect it from the Lookup Activity**

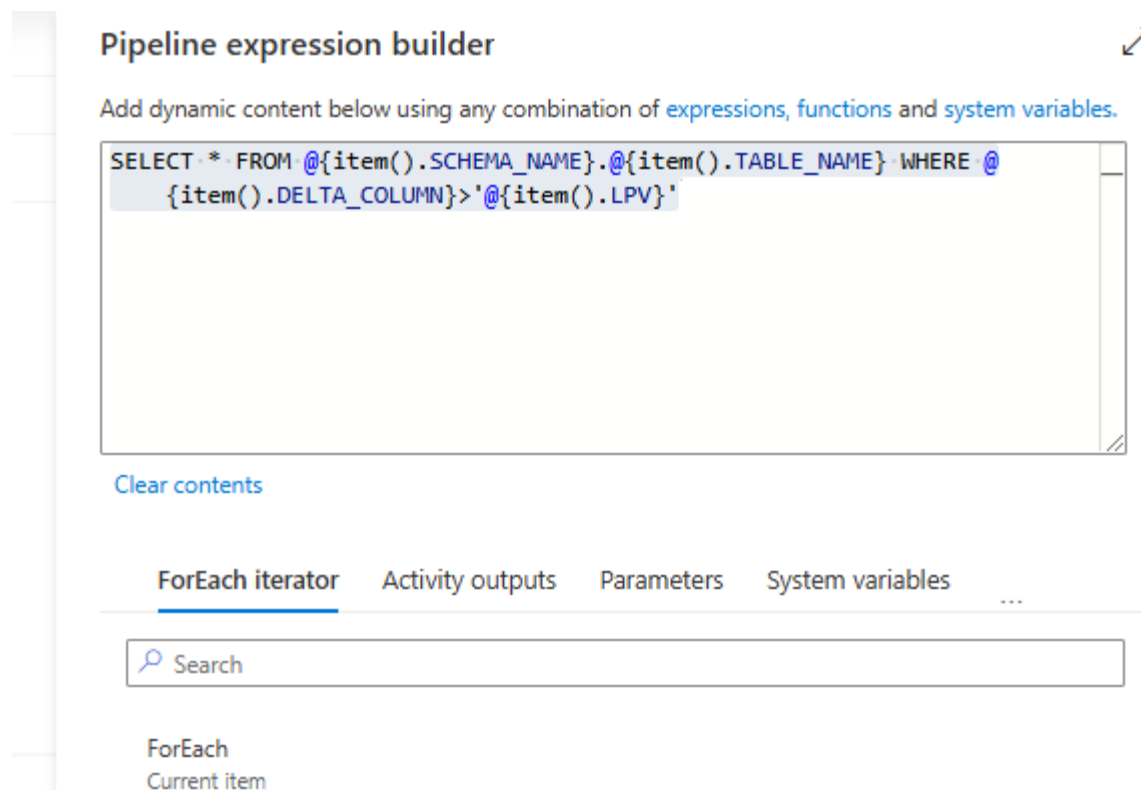


Go to the **Source Tab** and **select the source dataset** which we selected before in the lookup activity (Azure SQL Database)

- Provide a single quotation in schema name and table name (')
- Go to Use query options and select the **Query** radio button
- Click on Add Dynamic Content and the dialog will open

Write this below query in that dialog:

**SELECT * FROM @item().SCHEMA_NAME}@item().TABLE_NAME} WHERE @
@item().DELTA_COLUMN}>'@item().LPV}'**



Click OK.

After that, the Source tab will look like this:

General **Source** Sink¹ Mapping Settings User properties

Source dataset * azuresqllookup [Open](#) [+ New](#) [Preview data](#) [Learn more](#)

Dataset properties [ⓘ]

Name	Value
SchemaName	''
TableName	''

Use query ☐ Table ☒ Query ☐ Stored procedure

Query SELECT * FROM @({item()}.SCHEMA_N...

Query timeout (minutes) [ⓘ] 120

Isolation level [ⓘ] Select...

Partition option [ⓘ] ☒ None ☐ Physical partitions of table [ⓘ] ☐ Dynamic range [ⓘ]

i Please preview data to validate the partition settings.

Additional columns [ⓘ] [+ New](#)

Go to the **Sink Tab** and Select **ADLS Gen 2** as the **sink dataset** and **DelimitedFile** for the destination file format

Go to Parameters Tab

Create two parameters

- **FolderName**
- **FileName**

Connection Schema **Parameters**

[+ New](#) | [Delete](#)

<input type="checkbox"/> Name	Type	Default value	
<input type="checkbox"/> FolderName	String	Value	Delete
<input type="checkbox"/> FileName	String	Value	Delete

Go to the Connection Tab

Assign those parameters by clicking on **Add Dynamic Content**

Connection tab will look like this.

Connection Schema Parameters

Linked service * AzureDataLakeStorage1 [Test connection](#) [Edit](#) [+ New](#) [Learn more](#)

Integration runtime * AutoResolveIntegrationRuntime [Edit](#)

File path csvfiles / @dataset().FolderName / @dataset().FileName [Browse](#) [Preview data](#) [Detect format](#)

Compression type No compression

Column delimiter Comma (,)

Row delimiter Default (\r\n or \n)

Encoding Default(UTF-8)

Quote character Double quote (")

Escape character Backslash (\)

First row as header ☒

Null value

Go back to the Sink Tab and provide Values for FolderName and FileName

So, for FolderName we will write as shown below by clicking on **Add Dynamic Content**

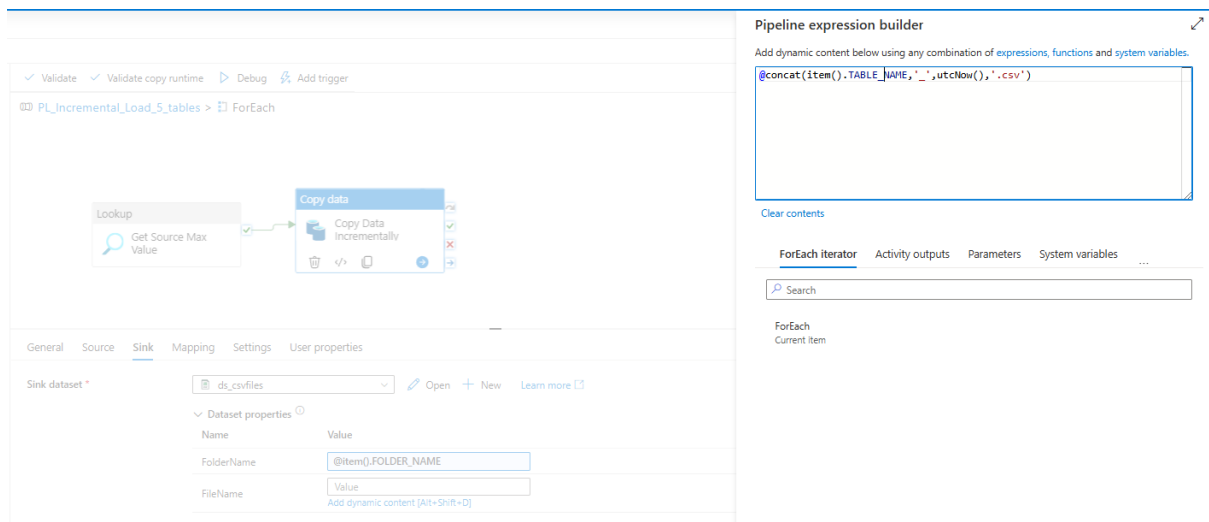
The screenshot shows the 'Sink' tab configuration for a 'Copy data' activity. The 'Sink dataset' is 'ds_csvfiles'. Under 'Dataset properties', 'FolderName' is set to '@item().FOLDER_NAME' and 'FileName' is set to 'Value'. The 'Copy behavior' is 'Select...'. The 'Pipeline expression builder' on the right shows the expression '@item().FOLDER_NAME'.

For the FileName, we will use this:

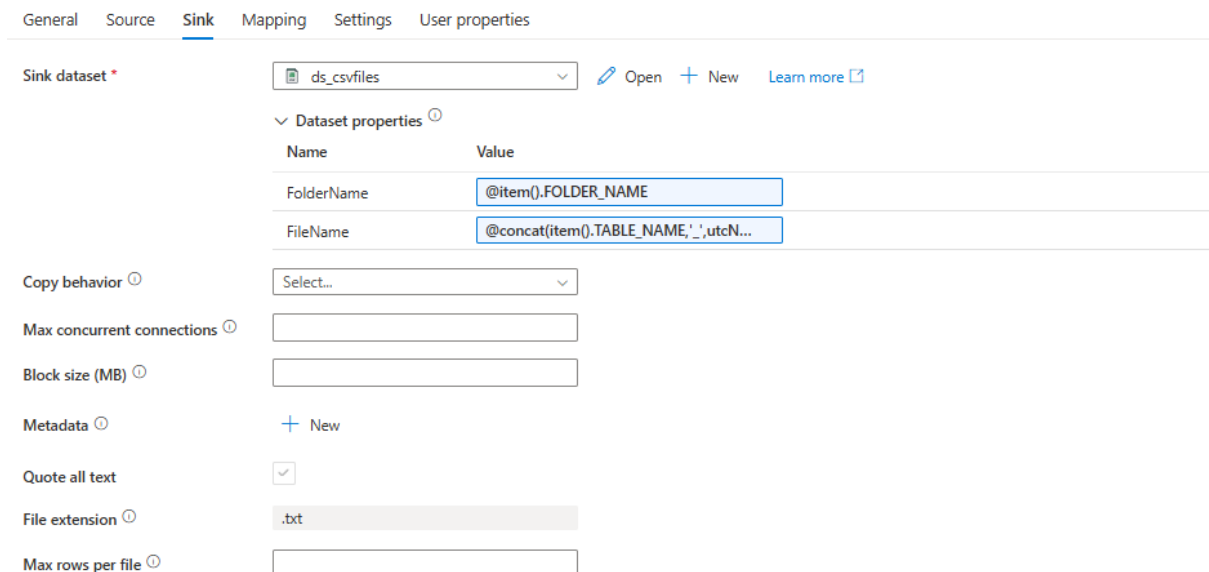
Concat function:

@concat(item().TABLE_NAME,'_',utcNow(),'.csv')

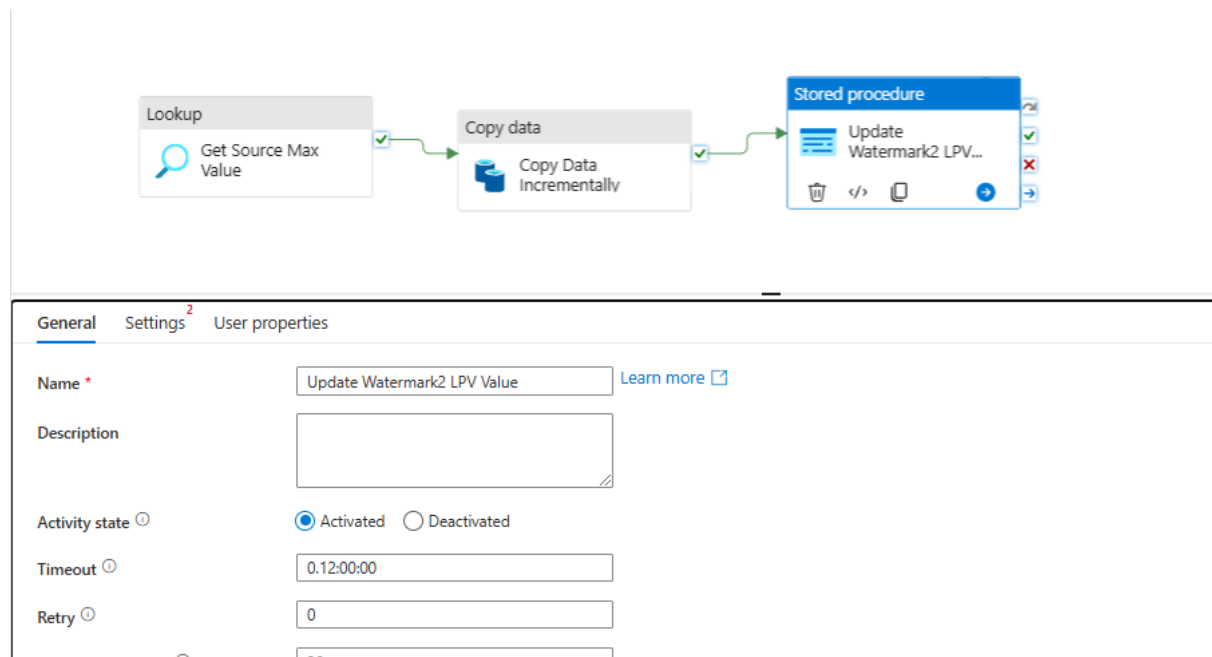
Why this: The expression in Azure Synapse dynamically constructs a CSV filename combining the table name with the current UTC timestamp, ensuring unique and time-stamped file names for data exports or logs.



After that, it will look like this in the Sink tab:

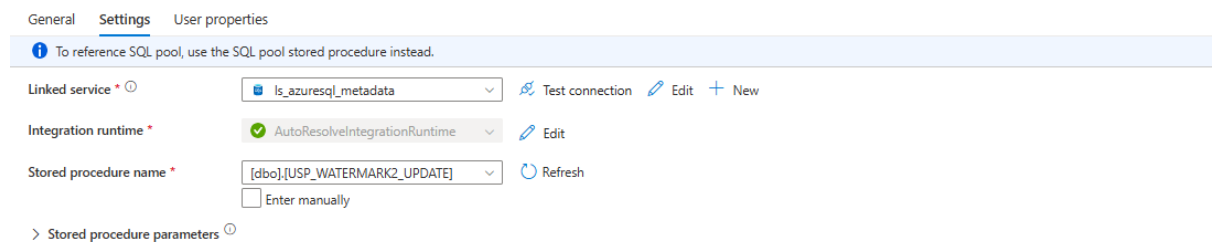


Now, drag and drop the **Stored Procedure** activity in the Design area and connect it from the Copy Data Activity as shown below:



Go to the **Settings** tab and select the **linked service** as selected before in the source/lookup activity (ADLS Gen 2)

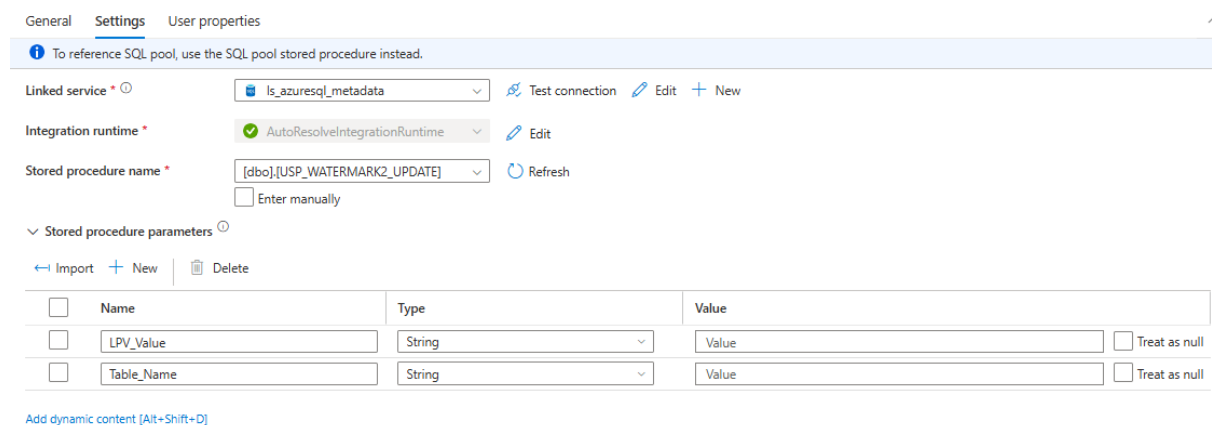
Select the **Stored Procedure** name from the dropdown



Click on **Stored Procedure Parameters**

And Click on **Import**

It will automatically take parameters from our stored procedure:



Now, Click on the value field for LPV_Value

Go to **Activity Output** -> Select **Get Source Max Value first row**

And it will automatically fill the above textbox with a query and we need to mention **MaxValue** (we did it in the second lookup activity inside ForEach) at the end of that query as shown below.

@activity('Get Source Max Value').output.firstRow.MaxValue

For the **Table name parameter**, select the ForEach current item from the dropdown, and match the SSMS Column name (TABLE_NAME in this case) of the Watermark2 table here, as shown below:

The screenshot displays the 'Stored procedure' activity settings in the 'Settings' tab of the 'Update Watermark2 LPV...' activity. The 'Linked service' is 'ls_azuresql_metadata', and the 'Integration runtime' is 'AutoResolveIntegrationRuntime'. The 'Stored procedure name' is '[dbo].[USP_WATERMARK2_UPDATE]'. Under 'Stored procedure parameters', two parameters are defined: 'LPV_Value' (String) and 'Table_Name' (String). The 'Value' for 'Table_Name' is '@activity('Get Source Max Value').ou...'. The 'Pipeline expression builder' on the right shows the expression '@item().TABLE_NAME'.

Stored procedure activity's settings tab It will look like this:

PL_Incremental_Load_5_tables > ForEach

General Settings User properties

To reference SQL pool, use the SQL pool stored procedure instead.

Linked service * Test connection Edit + New

Integration runtime * Edit

Stored procedure name * Refresh

☐ Enter manually

Stored procedure parameters

Import + New Delete

Name	Type	Value
LPV_Value	String	@activity('Get Source Max Value').ou...
Table_Name	String	@item().TABLE_NAME

Debug the pipeline.

PL_Incremental_Load_5_tables

Activities Validate Debug Add trigger

Synapse SQL pool stored proc... General Stored procedure

Parameters Variables Settings Output

Pipeline run ID: 295d3020-0cf6-46d6-b8f2-de0002b0cc3a Pipeline status: Succeeded View debug run consumption

All status List Showing 1 - 17 of 17 items

Activity name	Activity status	Activity type	Run start	Duration	Integration runtime	User properties	Activity run ID
Update Watermark2 LPV Value	Succeeded	Stored procedure	2/25/2025, 10:54:00 PM	7s	AutoResolveIntegrationRuntime (Australia East)		a809afcb-e4b9-488
Update Watermark2 LPV Value	Succeeded	Stored procedure	2/25/2025, 10:52:28 PM	9s	AutoResolveIntegrationRuntime (Australia East)		091b38cd-3cd9-44
Update Watermark2 LPV Value	Succeeded	Stored procedure	2/25/2025, 10:52:27 PM	16s	AutoResolveIntegrationRuntime (Australia East)		17e18ff3-9ed1-4df
Copy Data Incrementally	Succeeded	Copy data	2/25/2025, 10:52:21 PM	17s	AutoResolveIntegrationRuntime (Australia East)		d753c409-8529-4fe
Update Watermark2 LPV Value	Succeeded	Stored procedure	2/25/2025, 10:52:20 PM	11s	AutoResolveIntegrationRuntime (Australia East)		2396b220-751e-4ff
Update Watermark2 LPV Value	Succeeded	Stored procedure	2/25/2025, 10:51:17 PM	17s	AutoResolveIntegrationRuntime (Australia East)		d586bab6-6751-4a
Copy Data Incrementally	Succeeded	Copy data	2/25/2025, 10:50:08 PM	19s	AutoResolveIntegrationRuntime (Australia East)		6326b1ab-edb4-49
Copy Data Incrementally	Succeeded	Copy data	2/25/2025, 10:50:03 PM	22s	AutoResolveIntegrationRuntime (Australia East)		5c6a69ea-443a-40
Copy Data Incrementally	Succeeded	Copy data	2/25/2025, 10:49:59 PM	20s	AutoResolveIntegrationRuntime (Australia East)		095a922e-9edf-4c4
Copy Data Incrementally	Succeeded	Copy data	2/25/2025, 10:48:58 PM	18s	AutoResolveIntegrationRuntime (Australia East)		cc81df0d-15a1-4ac
Get Source Max Value	Succeeded	Lookup	2/25/2025, 10:48:48 PM	13s	AutoResolveIntegrationRuntime (Australia East)		5119042a-a71b-447
Get Source Max Value	Succeeded	Lookup	2/25/2025, 10:48:48 PM	9s	AutoResolveIntegrationRuntime (Australia East)		1188983d-9950-4c
Get Source Max Value	Succeeded	Lookup	2/25/2025, 10:48:48 PM	31s	AutoResolveIntegrationRuntime (Australia East)		d1dc44a8-2710-4a
Get Source Max Value	Succeeded	Lookup	2/25/2025, 10:48:48 PM	10s	AutoResolveIntegrationRuntime (Australia East)		d92ff455-3638-4c1
Get Source Max Value	Succeeded	Lookup	2/25/2025, 10:48:48 PM	19s	AutoResolveIntegrationRuntime (Australia East)		79529d39-675b-4a
ForEach	Succeeded	ForEach	2/25/2025, 10:48:48 PM	1m 3s	AutoResolveIntegrationRuntime (Australia East)		069a170a-7eb1-43
Get Watermark2 Data	Succeeded	Lookup	2/25/2025, 10:47:37 PM	10s	AutoResolveIntegrationRuntime (Australia East)		a676bc0d-ee4f-4f3

It was executed successfully.

Publish the Pipeline.

Now, let's check the Watermark2 table data

Results		Messages				
	ID	TABLE_NAME	SCHEMA_NAME	FOLDER_NAME	LPV	DELTA_COLUMN
1	1	Products	dbo	RetailDB/Products	4	product_id
2	2	Sales	dbo	RetailDB/Sales	2023-10-01T15:30:00	sale_date
3	3	Employees	dbo	RetailDB/Employees	2023-10-01T00:00:00	last_review_date
4	4	Suppliers	dbo	RetailDB/Suppliers	4	supplier_id
5	5	Inventory_Logs	dbo	RetailDB/InventoryLogs	2023-10-01T11:00:00	log_date

The **Watermark2** table has been updated correctly.

Now, let's check the data in the ADLS Gen2 storage

Synapse live Validate all Publish all

Data + <<

Workspace Linked

Filter resources by name

Azure Data Lake Storage Gen2 2

wsp-synapse-harpal (Primary - adls...

csvfiles (Primary)

samplecsvfiles

(Attached Containers)

Integration datasets 2

PL_Incremental_Loa... ds_csvfiles csvfiles x

New SQL script New data flow New integration dataset Upload Download New folder Select all Copy link

← → ↕ ↗ ↘ ↙ ↚ ↛ ↜ ↝ ↞ ↠ ↡ ↢ ↣ ↤ ↥ ↦ ↧ ↨ ↩ ↪ ↫ ↬ ↭ ↮ ↯ ↰ ↱ ↲ ↳ ↴ ↵ ↶ ↷ ↸ ↹ ↺ ↻ ↼ ↽ ↾ ↿ ⇀ ⇁ ⇂ ⇃ ⇄ ⇅ ⇆ ⇇ ⇈ ⇉ ⇊ ⇋ ⇌ ⇍ ⇎ ⇏ ⇐ ⇑ ⇒ ⇓ ⇔ ⇕ ⇖ ⇗ ⇘ ⇙ ⇚ ⇛ ⇜ ⇝ ⇞ ⇟ ⇠ ⇡ ⇢ ⇣ ⇤ ⇥ ⇦ ⇧ ⇨ ⇩ ⇪ ⇫ ⇬ ⇭ ⇮ ⇯ ⇰ ⇱ ⇲ ⇳ ⇴ ⇵ ⇶ ⇷ ⇸ ⇹ ⇺ ⇻ ⇼ ⇽ ⇾ ⇿ ⇿ ⇿ ⇿ ⇿ ⇿ ⇿ ⇿ ⇿ ⇿

Name	Last Modified
DBO	2/25/2025, 10:04:23 AM
RetailDB	2/25/2025, 1:05:13 PM
cities.csv	2/21/2025, 8:18:06 AM
day.csv	2/21/2025, 8:18:06 AM
Employee_DB.csv	2/21/2025, 8:18:06 AM
Employee_Details.csv	2/21/2025, 8:18:06 AM
industry.csv	2/21/2025, 8:18:06 AM
month.csv	2/21/2025, 8:18:06 AM
time.csv	2/21/2025, 8:18:06 AM

Go to ADLS Gen 2-> csvfiles -> RetailDB

Name	Last Modified	Content Type
Employees	2/25/2025, 1:05:14 PM	Folder
InventoryLogs	2/25/2025, 1:05:20 PM	Folder
Products	2/25/2025, 1:05:35 PM	Folder
Sales	2/25/2025, 1:05:13 PM	Folder
Suppliers	2/25/2025, 1:05:23 PM	Folder

Employee Data:

New SQL script New notebook New data flow New integration dataset Upload Download New folder Select all Rename Manage access Properties

← → ↕ ↗ ↘ ↙ ↚ ↛ ↜ ↝ ↞ ↠ ↡ ↢ ↣ ↤ ↥ ↦ ↧ ↨ ↩ ↪ ↫ ↬ ↭ ↮ ↯ ↰ ↱ ↲ ↳ ↴ ↵ ↶ ↷ ↸ ↹ ↺ ↻ ↼ ↽ ↾ ↿ ⇀ ⇁ ⇂ ⇃ ⇄ ⇅ ⇆ ⇇ ⇈ ⇉ ⇊ ⇋ ⇌ ⇍ ⇎ ⇏ ⇐ ⇑ ⇒ ⇓ ⇔ ⇕ ⇖ ⇗ ⇘ ⇙ ⇚ ⇛ ⇜ ⇝ ⇞ ⇟ ⇠ ⇡ ⇢ ⇣ ⇤ ⇥ ⇦ ⇧ ⇨ ⇩ ⇪ ⇫ ⇬ ⇭ ⇮ ⇯ ⇰ ⇱ ⇲ ⇳ ⇴ ⇵ ⇶ ⇷ ⇸ ⇹ ⇺ ⇻ ⇼ ⇽ ⇾ ⇿ ⇿ ⇿ ⇿ ⇿ ⇿ ⇿ ⇿ ⇿ ⇿

← → ↕ ↗ ↘ ↙ ↚ ↛ ↜ ↝ ↞ ↠ ↡ ↢ ↣ ↤ ↥ ↦ ↧ ↨ ↩ ↪ ↫ ↬ ↭ ↮ ↯ ↰ ↱ ↲ ↳ ↴ ↵ ↶ ↷ ↸ ↹ ↺ ↻ ↼ ↽ ↾ ↿ ⇀ ⇁ ⇂ ⇃ ⇄ ⇅ ⇆ ⇇ ⇈ ⇉ ⇊ ⇋ ⇌ ⇍ ⇎ ⇏ ⇐ ⇑ ⇒ ⇓ ⇔ ⇕ ⇖ ⇗ ⇘ ⇙ ⇚ ⇛ ⇜ ⇝ ⇞ ⇟ ⇠ ⇡ ⇢ ⇣ ⇤ ⇥ ⇦ ⇧ ⇨ ⇩ ⇪ ⇫ ⇬ ⇭ ⇮ ⇯ ⇰ ⇱ ⇲ ⇳ ⇴ ⇵ ⇶ ⇷ ⇸ ⇹ ⇺ ⇻ ⇼ ⇽ ⇾ ⇿ ⇿ ⇿ ⇿ ⇿ ⇿ ⇿ ⇿ ⇿ ⇿

csvfiles RetailDB Employees

Name

Employees_2025-02-25T18:04:59.8444368Z.csv

Path https://adlsgen2stgharpal.dfs.core.windows.net/csvfiles/RetailDB/Emplo

Modified 2/25/2025, 1:05:15 PM

With column header On

EMPLOYEE_ID	FIRST_NAME	LAST_NAME	HIRE_DATE
1	Raj	Sharma	2022-01-05 09:
2	Harpal	Vaghela	2022-05-15 09:
3	Amit	Singh	2023-03-23 09:
4	Anjali	Patel	2023-04-10 09:
NULL	NULL	NULL	NULL

Inventory Logs Data:

SQL scriptNew notebookNew data flowNew integration datasetUploadDownloadNew folderSelect allRenameManage access

↑ csvfiles > RetailDB > InventoryLogs

InventoryLogsInventory_Logs_2025-02-25T18:05:03.7724618Z.csv

Inventory_Logs_2025-02-25T18:05:03.7724618Z.csv

Pathhttps://adlsgen2stgharpal.dfs.core.windows.net/csvfiles/RetailDB/Inventory_Logs_2025-02-25T18:05:03.7724618Z.csv

Modified2/25/2025, 1:05:21 PM

With column headerOn

LOG_ID	PRODUCT_ID	LOG_DATE	CHANGE_QL
1	1	2023-10-01 08:00:00	20
2	2	2023-10-01 09:00:00	-10
3	3	2023-10-01 10:00:00	30
4	4	2023-10-01 11:00:00	-5
NULL	NULL	NULL	NULL

Content Type

Products Data:

PL_Incremental_Load.csvds_csvfilescsvfiles

New SQL scriptNew notebookNew data flowNew integration datasetUploadDownloadNew folderSelect allRenameManage accessPropertiesMore

← → ↑ csvfiles > RetailDB > Products

NameProducts_2025-02-25T18:05:21.8345890Z.csv

Products_2025-02-25T18:05:21.8345890Z.csv

Pathhttps://adlsgen2stgharpal.dfs.core.windows.net/csvfiles/RetailDB/Products_2025-02-25T18:05:21.8345890Z.csv

Modified2/25/2025, 1:05:36 PM

With column headerOn

PRODUCT_ID	NAME	CATEGORY	PRICE
1	Organic Apples	Fruits	2.99
2	Almond Milk	Beverages	3.49
3	Chicken Breast	Meat	7.99
4	Whole Wheat Br...	Bakery	2.50
NULL	NULL	NULL	NULL

Content Type

Sales Data:

New SQL scriptNew notebookNew data flowNew integration datasetUploadDownloadNew folderSelect allRenameManage accessPropertiesMore

← → ↑ csvfiles > RetailDB > Sales

NameSales_2025-02-25T18:04:58.7998580Z.csv

Sales_2025-02-25T18:04:58.7998580Z.csv

Pathhttps://adlsgen2stgharpal.dfs.core.windows.net/csvfiles/RetailDB/Sales/Sales_2025-02-25T18:04:58.7998580Z.csv

Modified2/25/2025, 1:05:14 PM

With column headerOn

SALE_ID	SALE_DATE	PRODUCT_ID	QUANTITY
1	2023-10-01 14:00:00	1	10
2	2023-10-01 14:10:00	2	5
3	2023-10-01 15:00:00	3	4
4	2023-10-01 15:30:00	4	6
NULL	NULL	NULL	NULL

Content Type

Suppliers Data:

New SQL script

New notebook

New data flow

New integration dataset

Upload

Download

New folder

Select all

Rename

Manage access

Properties

More

← → ↶ ↷

csvfiles

RetailDB

Suppliers

Name

Suppliers_2025-02-25T18:05:08.8652181Z.csv

Suppliers_2025-02-25T18:05:08.8652181Z.csv

Path

https://adlsgen2stgharpal.dfs.core.windows.net/csvfiles/RetailDB/Suppl02-25T18:05:08.8652181Z.csv

Modified

2/25/2025, 1:05:24 PM

With column header

On

SUPPLIER_ID	NAME	CONTACT_NA...	PHONE
1	Fresh Farms	John Doe	555-3489
2	Healthy Beverag...	Emily Stone	555-7623
3	Premium Meats	Alan Smith	555-9876
4	Bakers Delight	Nora Bates	555-4532
NULL	NULL	NULL	NULL

Content Type

Now, let's insert some new records into tables:

----Insert new data into product table

INSERT INTO Products (product_id, name, category, price, stock, supplier_id) VALUES
(5, 'Vegetable Oil', 'Groceries', 4.50, 100, 2);

----Insert new data into Sales table

INSERT INTO Sales (sale_id, sale_date, product_id, quantity, total_amount, cashier_id) VALUES
(5, '2025-02-25 10:30:00', 5, 20, 90.00, 1);

----Insert new data into Employee table

INSERT INTO Employees (employee_id, first_name, last_name, hire_date, last_review_date, role) VALUES
(5, 'Nayan', 'Vaghela', '2025-01-01 09:00:00', '2025-02-20 00:00:00', 'Inventory Specialist');

132	INSERT INTO Products (product_id, name, category, price, stock, supplier_id) VALUES
133	(5, 'Vegetable Oil', 'Groceries', 4.50, 100, 2);
134	
135	----Insert new data into Sales table
136	
137	INSERT INTO Sales (sale_id, sale_date, product_id, quantity, total_amount, cashier_id) VALUES
138	(5, '2025-02-25 10:30:00', 5, 20, 90.00, 1);
139	
140	----Insert new data into Employee table
141	
142	INSERT INTO Employees (employee_id, first_name, last_name, hire_date, last_review_date, role) VALUES
143	(5, 'Nayan', 'Vaghela', '2025-01-01 09:00:00', '2025-02-20 00:00:00', 'Inventory Specialist');
144	
145	
146	Select * From Products
147	Select * From Employees

product_id	name	category	price	stock	supplier_id
1	Organic Apples	Fruits	2.99	150	1
2	Almond Milk	Beverages	3.49	85	2
3	Chicken Breast	Meat	7.99	60	3
4	Whole Wheat Bread	Bakery	2.50	120	4
5	Vegetable Oil	Groceries	4.50	100	2

employee_id	first_name	last_name	hire_date	last_review_date	role
1	Raj	Sharma	2022-01-05 09:00:00.000	2023-09-10 00:00:00.000	Cashier
2	Harpal	Vaghela	2022-05-15 09:00:00.000	2023-09-20 00:00:00.000	Cashier
3	Amit	Singh	2023-03-23 09:00:00.000	2023-09-30 00:00:00.000	Stock Manager
4	Anjali	Patel	2023-04-10 09:00:00.000	2023-10-01 00:00:00.000	Sales Manager
5	Nayan	Vaghela	2025-01-01 09:00:00.000	2025-02-20 00:00:00.000	Inventory Specialist

sale_id	sale_date	product_id	quantity	total_amount	cashier_id
1	2023-10-01 14:00:00.000	1	10	29.90	1
2	2023-10-01 14:15:00.000	2	5	17.45	2
3	2023-10-01 15:00:00.000	3	4	31.96	1
4	2023-10-01 15:30:00.000	4	6	15.00	2

Now, let's run the pipeline again.

✓ Validate ▶ Debug ⚙ Add trigger

Parameters Variables Settings Output

Pipeline run ID: 6745fc96-4fff-4205-aad3-d0909500dcb2 Pipeline status: Succeeded View debug run consumption

All status List

Showing 1 - 17 of 17 items

Activity name	Activity status	Activity...	Run start	Duration	Integration runtime	User prope...	Activity run ID	Log
Update Watermark2 LPV Value	Succeeded	Stored procedur	2/25/2025, 1:17:39 PM	6s	AutoResolveIntegrationRuntime (Australia East)		4737ee5d-878b-41c3-b5db-cfec9a38d5a7	
Update Watermark2 LPV Value	Succeeded	Stored procedur	2/25/2025, 1:17:29 PM	4s	AutoResolveIntegrationRuntime (Australia East)		3f8bb0ef-0fdf-4a0c-8b6b-f14af91f15fd	
Update Watermark2 LPV Value	Succeeded	Stored procedur	2/25/2025, 1:17:21 PM	12s	AutoResolveIntegrationRuntime (Australia East)		64b587bf-c9d-42d3-aec5-43d3e824bc1	
Update Watermark2 LPV Value	Succeeded	Stored procedur	2/25/2025, 1:17:16 PM	5s	AutoResolveIntegrationRuntime (Australia East)		642408de-1c87-41ee-9985-2bd0e73719d2	
Update Watermark2 LPV Value	Succeeded	Stored procedur	2/25/2025, 1:17:14 PM	7s	AutoResolveIntegrationRuntime (Australia East)		00e04b0e-9cab-42a1-947f-3712d5209ab4	
Copy Data Incrementally	Succeeded	Copy data	2/25/2025, 1:17:03 PM	25s	AutoResolveIntegrationRuntime (Australia East)		93b11d24-ca0c-44b1-b43f-a5c83df24bad	
Copy Data Incrementally	Succeeded	Copy data	2/25/2025, 1:17:03 PM	36s	AutoResolveIntegrationRuntime (Australia East)		9dfb917f-a1f7-4477-95a1-938340b98a51	
Copy Data Incrementally	Succeeded	Copy data	2/25/2025, 1:17:00 PM	20s	AutoResolveIntegrationRuntime (Australia East)		e041f0e-b21d-4ac1-97db-287dced78c71	
Copy Data Incrementally	Succeeded	Copy data	2/25/2025, 1:16:56 PM	19s	AutoResolveIntegrationRuntime (Australia East)		ed1592c1-b784-4a94-901a-ca144cb34187	
Copy Data Incrementally	Succeeded	Copy data	2/25/2025, 1:16:52 PM	20s	AutoResolveIntegrationRuntime (Australia East)		c7919ec4-63e2-4e68-b288-4440f015446	
Get Source Max Value	Succeeded	Lookup	2/25/2025, 1:16:43 PM	11s	AutoResolveIntegrationRuntime (Australia East)		ea8f0099-be59-40c7-b3ee-bf9e68c9fe3c	
Get Source Max Value	Succeeded	Lookup	2/25/2025, 1:16:42 PM	19s	AutoResolveIntegrationRuntime (Australia East)		2b322790-f085-4b06-9e0c-acde9c10fef0	
Get Source Max Value	Succeeded	Lookup	2/25/2025, 1:16:42 PM	9s	AutoResolveIntegrationRuntime (Australia East)		5b2e72bd-7b55-4170-a980-9d1fe6d50386	
Get Source Max Value	Succeeded	Lookup	2/25/2025, 1:16:42 PM	20s	AutoResolveIntegrationRuntime (Australia East)		49fdcd1f-5925-4e83-9336-3365fc1b2bb7	
Get Source Max Value	Succeeded	Lookup	2/25/2025, 1:16:42 PM	15s	AutoResolveIntegrationRuntime (Australia East)		e81b4444-4373-480f-a773-da8be3e2c775	
ForEach	Succeeded	ForEach	2/25/2025, 1:16:42 PM	1m 7s	AutoResolveIntegrationRuntime (Australia East)		af214660-15a5-472a-b11d-81d0ce25d216	
Get Watermark2 Data	Succeeded	Lookup	2/25/2025, 1:16:21 PM	20s	AutoResolveIntegrationRuntime (Australia East)		f7bb1176-a80d-44ce-bf4a-0027f688e897	

Let's check the data and watermark2 table:

119 %

Results Messages

	ID	TABLE_NAME	SCHEMA_NAME	FOLDER_NAME	LPV	DELTA_COLUMN
1	1	Products	dbo	RetailDB/Products	5	product_id
2	2	Sales	dbo	RetailDB/Sales	2023-10-01T15:30:00	sale_date
3	3	Employees	dbo	RetailDB/Employees	2025-02-20T00:00:00	last_review_date
4	4	Suppliers	dbo	RetailDB/Suppliers	4	supplier_id
5	5	Inventory_Logs	dbo	RetailDB/InventoryLogs	2023-10-01T11:00:00	log_date

The **Watermark2** table has been updated correctly.

Now, let's check the data in the ADLS Gen2 storage

Go to ADLS Gen 2-> csvfiles -> RetailDB

Data

Workspace Linked

Filter resources by name

Azure Data Lake Storage Gen2 2

vsp-synapse-hargal (Primary - adl...

csvfiles (Primary)

samplecsvfiles

Integration datasets 2

Name	Last Modified	Content Type	Size
Employees	2/25/2025, 1:05:14 PM	Folder	
InventoryLogs	2/25/2025, 1:05:20 PM	Folder	
Products	2/25/2025, 1:05:35 PM	Folder	
Sales	2/25/2025, 1:05:13 PM	Folder	
Suppliers	2/25/2025, 1:05:23 PM	Folder	

Employee Data:

← → ↕

csvfiles > RetailDB > Employees

Name	La
Employees_2025-02-25T18:04:59.8444368Z.csv	2/25
Employees_2025-02-25T18:16:56.4928692Z.csv	2/25

Employees_2025-02-25T18:16:56.4928692Z.csv

Path

https://adlsgen2stgharpal.dfs.core.windows.net/csvfiles/RetailDB/Emplc02-25T18:16:56.4928692Z.csv

Modified

2/25/2025, 1:17:13 PM

With column header

☒ On

EMPLOYEE_ID	FIRST_NAME	LAST_NAME	HIRE_DATE
5	Nayan	Vaghela	2025-01-01 09:
NULL	NULL	NULL	NULL

Content Type

The Employee record is updated successfully in destination storage.

Products Data:

← → ↕

csvfiles > RetailDB > Products

Name	La
Products_2025-02-25T18:05:21.8345890Z.csv	2/25
Products_2025-02-25T18:17:03.6759168Z.csv	2/25

Products_2025-02-25T18:17:03.6759168Z.csv

Path

https://adlsgen2stgharpal.dfs.core.windows.net/csvfiles/RetailDB/Produ02-25T18:17:03.6759168Z.csv

Modified

2/25/2025, 1:17:36 PM

With column header

☒ On

PRODUCT_ID	NAME	CATEGORY	PRICE
5	Vegetable Oil	Groceries	4.50
NULL	NULL	NULL	NULL

Content Type

The Products table record is updated successfully in destination storage.

Sales Data:

Sales_2025-02-25T18:22:04.3564746Z.csv

Path

https://adlsgen2stgharpal.dfs.core.windows.net/csvfiles/RetailDB/Sales/02-25T18:22:04.3564746Z.csv

Modified

2/25/2025, 1:22:17 PM

With column header

☒ On

SALE_ID	SALE_DATE	PRODUCT_ID	QUANTITY
5	2025-02-25 10:3...	5	20
NULL	NULL	NULL	NULL

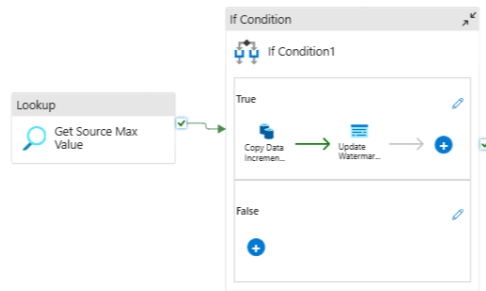
Content Type

The Sales table record is updated successfully in destination storage.

Put If Condition, so no file with empty records will be created.

Let's work on this now,

PL_Incremental_Load_5_tables_If_Condition > ForEach



Drag and drop **if condition** from activities tab.

Activities

if

Iteration & conditionals

If Condition

Validate Debug Trigger (1)

PL_Incremental_Load_5_tables_If_Condition > ForEach

If Condition

If Condition2

True
No activities

False
No activities

+

+

+

Go to Activities in If condition and in Expression field write this expression:

The screenshot shows the Azure Data Factory pipeline editor. On the left, the pipeline structure is visible: 'PL_Incremental_Load_5_tables_If_Condition' > 'ForEach'. The 'If Condition' activity is selected, and its configuration is shown in the main pane. The 'Expression' field contains the following code: `@less(string(item().LPV), string(activity('Get Source Max Value').output.firstRow.MaxValue))`. The 'Case' field is set to 'True', and the 'Activity' field is set to 'Copy Data Incrementally'. The 'True' case has 2 activities, and the 'False' case has no activities. On the right, the 'Pipeline expression builder' pane is open, showing the same expression code. Below the code, there are tabs for 'ForEach iterator', 'Activity outputs', 'Parameters', and 'System variables'. The 'ForEach iterator' tab is selected, and it shows 'ForEach' and 'Current item'.

`@less(string(item().LPV), string(activity('Get Source Max Value').output.firstRow.MaxValue))`

Now, click on **Edit icon of “if condition” activity**,

Here, move previously created copy data and stored procedure activities.

It will look like this:

You can see we're in forEach-> If Condition -> True Activities

The screenshot shows the Azure Data Factory pipeline editor. The pipeline structure is 'PL_Incremental_Load_5_tables_If_Condition' > 'ForEach' > 'If Condition1 - True activities'. The 'True activities' are visible, showing a sequence of two activities: 'Copy Data Incrementally' and 'Update Watermark2 LPV...'. Both activities have a green checkmark next to them, indicating they are successful.

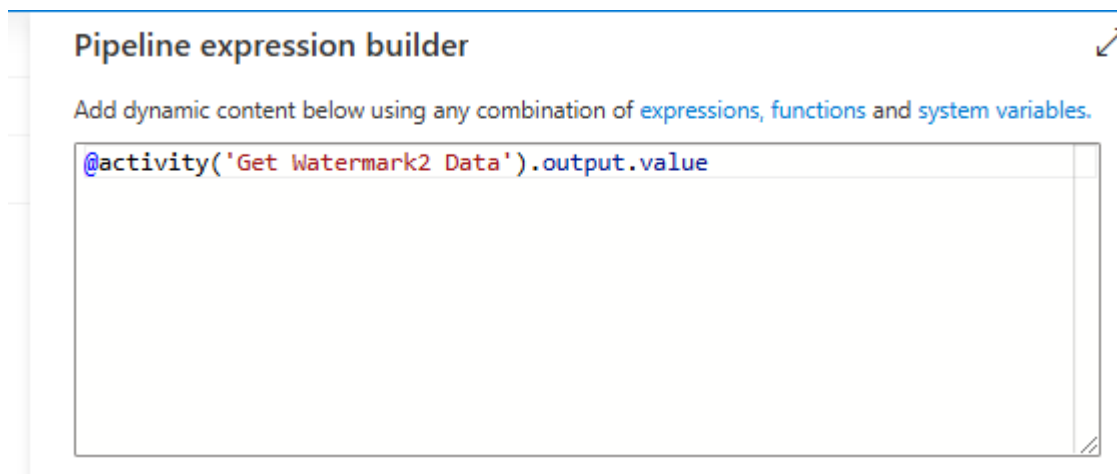
Publish the Pipeline again.

And Check the Data.

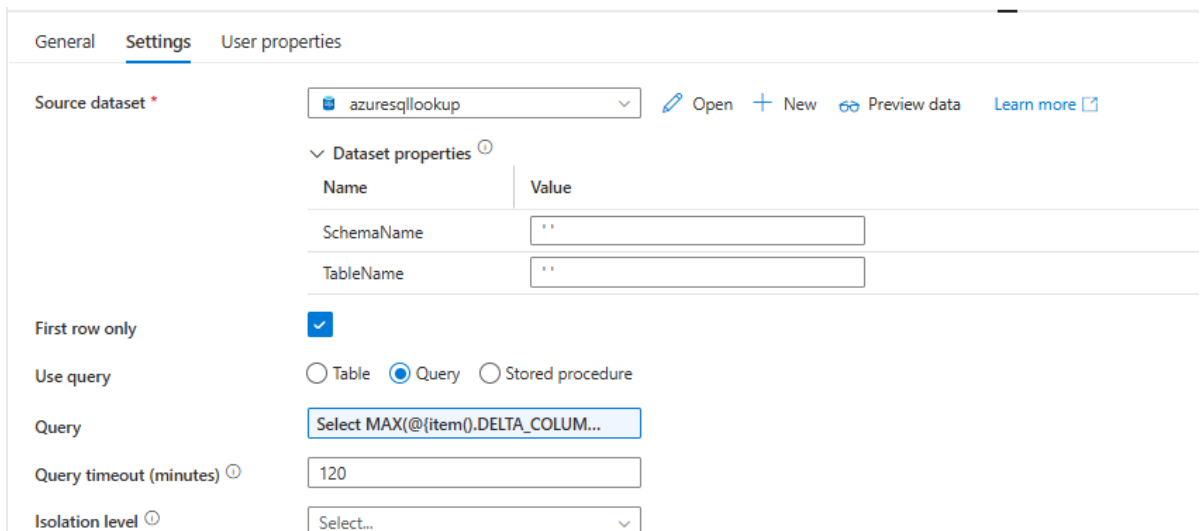
Now, it will not create any csv file with empty data.

Points to Remember:

- Expression in **ForEach** Activity should be a value array like this:



In the ForEach Activity, we utilize a Lookup table to retrieve the MaxValue from a source dataset previously employed, hence the system prompts for schema and table name inputs, which are not necessary in this context. The MaxValue is obtained through a custom query. When employing a new dataset, schema name and table name fields are absent. Conversely, when reusing an older dataset as indicated, it is essential to enclose value fields in single quotes to bypass them and execute the custom query.



The screenshot shows the 'Settings' tab of an activity configuration window. It has three tabs: 'General', 'Settings', and 'User properties'. The 'Source dataset' is set to 'azuresqllookup'. Below it, 'Dataset properties' are shown with 'SchemaName' and 'TableName' both set to empty string values. The 'First row only' checkbox is checked. The 'Use query' radio button is selected, and the 'Query' text box contains 'Select MAX(@{item().DELTA_COLUM...'. The 'Query timeout (minutes)' is set to 120, and the 'Isolation level' is set to 'Select...'. There are links for 'Open', 'New', 'Preview data', and 'Learn more'.

Custom Query:

- Remember this alias name as MaxValue as we will need this variable in the stored procedure to update the LPV value.

DEFAULT DIRECTORY

Pipeline expression builder

Add dynamic content below using any combination of [expressions](#), [functions](#) and [system variables](#).

```
Select MAX(@{item().DELTA_COLUMN}) as MaxValue from @{item().SCHEMA_NAME}.@{item().TABLE_NAME}
```

[Clear contents](#)

- Copy Data Activity inside If Condition
 - Make sure to write table, schema name, LPV and delta column name as same as SSMS watermark column names.

Pipeline expression builder

Add dynamic content below using any combination of [expressions](#), [functions](#) and [system variables](#).

```
SELECT * FROM @{item().SCHEMA_NAME}.@{item().TABLE_NAME} WHERE @
{item().DELTA_COLUMN}>'@{item().LPV}'
```

- In the Source configuration of the Copy Data Activity, we are using an older dataset, which prompts the request for parameter values. We are addressing this by enclosing these parameters in single quotation marks as shown below, effectively bypassing the request:

General **Source** Sink Mapping Settings User properties

Source dataset * azuresqllookup [Open](#) [New](#) [Preview data](#) [Learn more](#)

Dataset properties ⓘ

Name	Value
SchemaName	''
TableName	''

Use query ☐ Table ☒ Query ☐ Stored procedure

Query SELECT * FROM @(item().SCHEMA_N...

- In the Stored Procedure configuration, when parameters are imported, it prompts for the Table name and LPV value entries. For the LPV value, an expression should be used. Ensure that the variable name from the Lookup activity within the ForEach activity (MaxValue) is correctly specified

Pipeline expression builder



Add dynamic content below using any combination of [expressions](#), [functions](#) and [system variables](#).

```
@activity('Get Source Max Value').output.firstRow.MaxValue
```

[Clear contents](#)

Thank you for your time!

You can connect with me on:

LinkedIn: <https://www.linkedin.com/in/harpalvaghela/>