

# Azure Data Factory

## Lab: Trigger a pipeline using Tumbling window trigger

### Pre-requisites:

- Azure Pass subscription
- Azure Data Lake Storage Gen2 storage account
- Azure SQL Database

### Lab Objective:

#### After completing this lab, you will be able to:

- Copy data from SQL Database to Data Lake Storage
- Create and use Tumbling window trigger in Azure Data Factory.

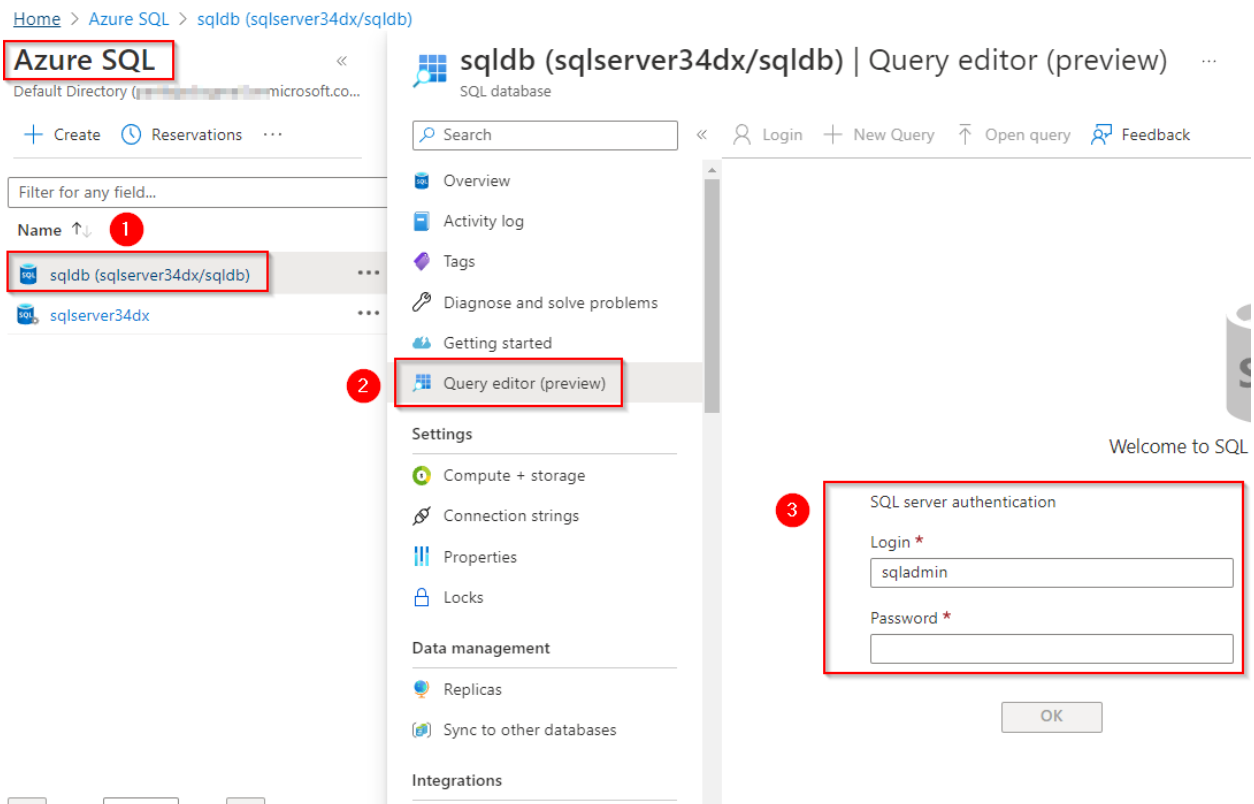
# Exercise 1: Create a Tumbling window trigger to execute the pipeline.

The main task for this exercise are as follows:

- Create a copy activity to copy data from SQL DB to Data Lake.
- Create a Tumbling Window trigger to run the 2 pipelines at a time.

## Task 1: Create a table in SQL DB

### 1. Open query editor in SQL DB



### 2. Add following code to create a table

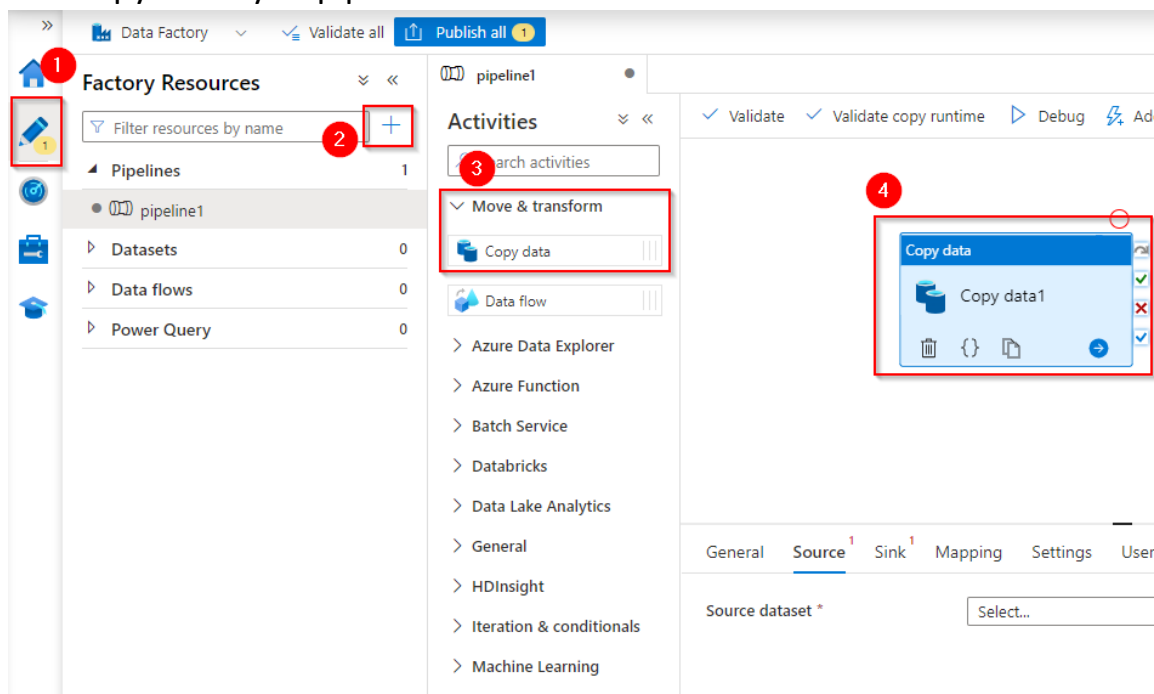
```
CREATE TABLE [dbo].[cars2] (  
    [Make] [nvarchar](4000) NULL,  
    [Model] [nvarchar](4000) NULL,  
    [Type] [nvarchar](4000) NULL,  
    [Origin] [nvarchar](4000) NULL,  
    [DriveTrain] [nvarchar](4000) NULL,  
    [Length] [int] NULL,  
    [ManufactureDate] [datetime2](7) NULL  
)
```

### 3. Insert following rows

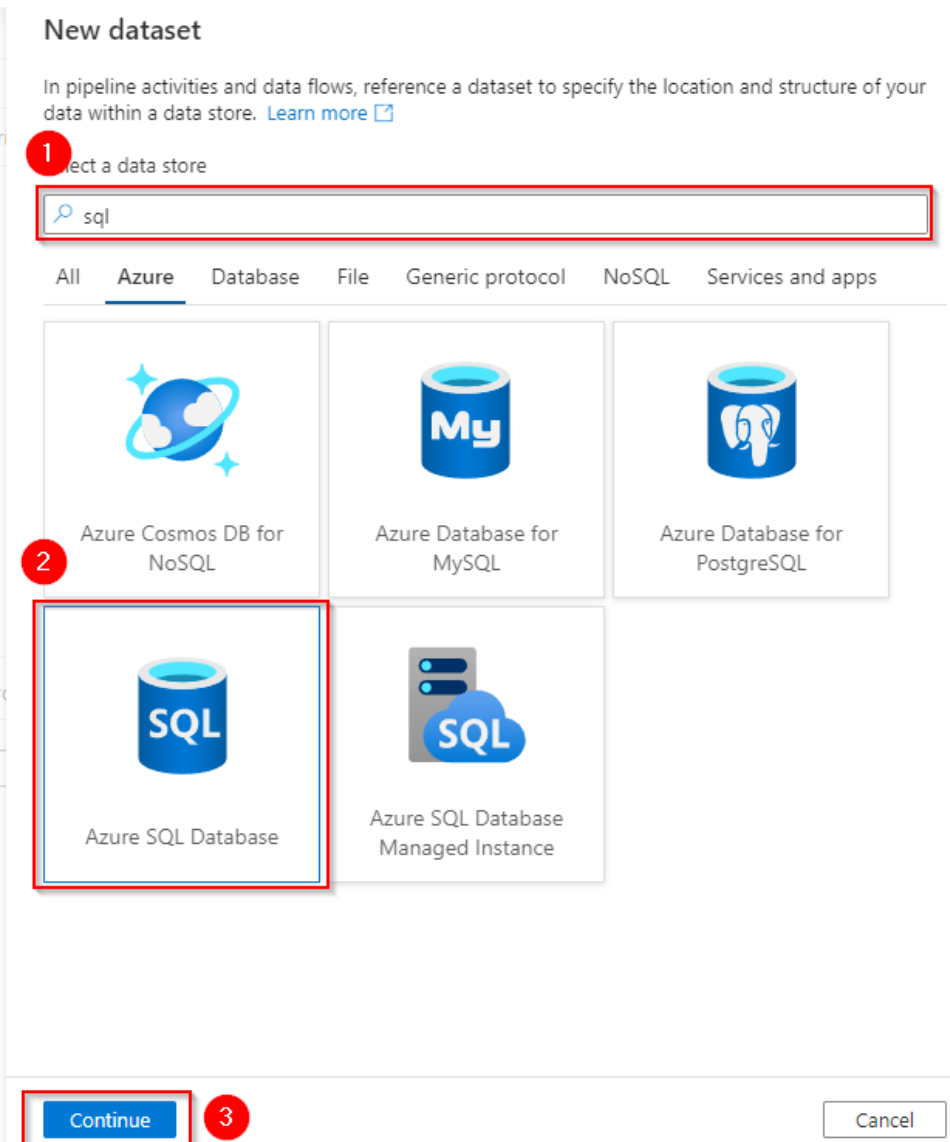
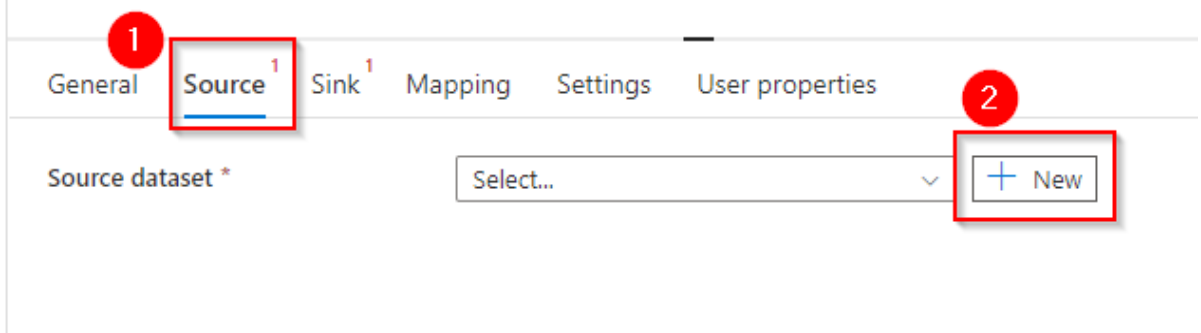
```
INSERT INTO cars2 VALUES('2001','MAX-1','MIN','ASIA','XXX',2889, '2012-01-01T00:02:00')
INSERT INTO cars2 VALUES('2002','MAX-2','MIN','ASIA','XXX',4889, '2012-01-01T00:03:00')
INSERT INTO cars2 VALUES('2003','MAX-3','MIN','ASIA','XXX',5589, '2012-01-01T00:05:00')
INSERT INTO cars2 VALUES('2005','MAX-4','MIN','ASIA','XXX',7879, '2012-01-01T00:07:00')
INSERT INTO cars2 VALUES('2006','MAX-5','MIN','ASIA','XXX',9889, '2012-01-01T00:09:00')
INSERT INTO cars2 VALUES('2004','MAX-1','MIN','ASIA','XXX',5589, '2012-01-01T00:04:00')
INSERT INTO cars2 VALUES('2005','MAX-2','MIN','ASIA','XXX',4889, '2012-01-01T00:06:00')
INSERT INTO cars2 VALUES('2007','MAX-3','MIN','ASIA','XXX',5839, '2012-01-01T00:05:00')
INSERT INTO cars2 VALUES('2008','MAX-4','MIN','ASIA','XXX',7189, '2012-01-01T00:07:00')
INSERT INTO cars2 VALUES('2009','MAX-5','MIN','ASIA','XXX',9789, '2012-01-01T00:01:00')
INSERT INTO cars2 VALUES('2011','MAX-1','MIN','ASIA','XXX',2889, '2012-01-01T00:08:00')
INSERT INTO cars2 VALUES('2012','MAX-2','MIN','ASIA','XXX',4879, '2012-01-01T00:08:00')
INSERT INTO cars2 VALUES('2013','MAX-3','MIN','ASIA','XXX',5399, '2012-01-01T00:02:00')
INSERT INTO cars2 VALUES('2025','MAX-4','MIN','ASIA','XXX',7489, '2012-01-01T00:11:00')
INSERT INTO cars2 VALUES('2036','MAX-5','MIN','ASIA','XXX',9869, '2012-01-01T00:12:00')
INSERT INTO cars2 VALUES('2021','MAX-1','MIN','ASIA','XXX',2449, '2012-01-01T00:15:00')
INSERT INTO cars2 VALUES('2042','MAX-2','MIN','ASIA','XXX',4969, '2012-01-01T00:12:00')
INSERT INTO cars2 VALUES('2023','MAX-3','MIN','ASIA','XXX',5259, '2012-01-01T00:13:00')
INSERT INTO cars2 VALUES('2065','MAX-4','MIN','ASIA','XXX',7819, '2012-01-01T00:14:00')
INSERT INTO cars2 VALUES('2026','MAX-5','MIN','ASIA','XXX',9579, '2012-01-01T00:14:00')
```

## Task 2: Create a copy activity to copy data from SQL Table

### 1. Add copy activity to pipeline canvas



2. Source dataset: Add new SQL DB data source. (Create linked service if not exists for SQL DB)



3. Select the linked Service and select the dbo.cars2 table,

**Set properties**

1 **Name**  
AzureSqlTable1

**Linked service \***  
SQLDB 2

**Table name**  
dbo.cars2 3

☐ Edit

4 **Import schema**  
☒ From connection/store ☐ None

> Advanced

**Note: Create a linked service if not exists otherwise skip this step.**

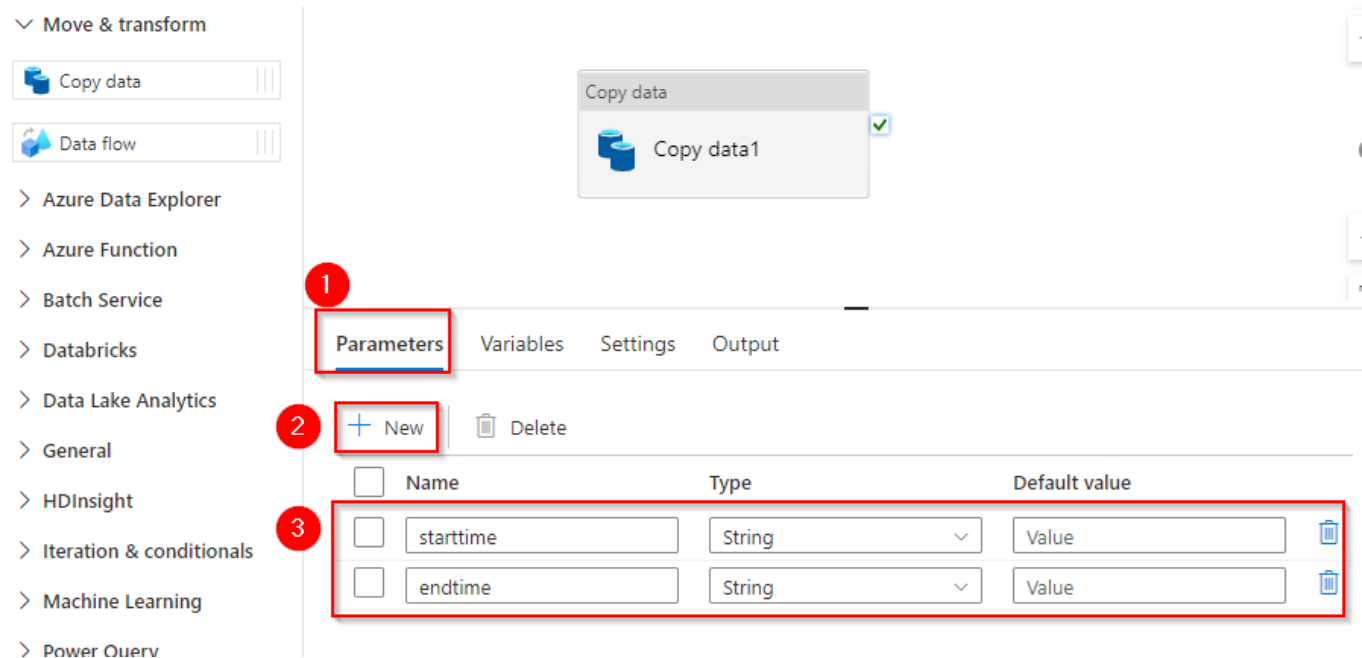
**Set properties**

1 **Name**  
AzureSqlTable1

**Linked service \***  
Select... 2  
Filter...  
Select...  
+ New 3

+ New

4. Create two pipeline parameter as follows:
- starttime
  - endtime

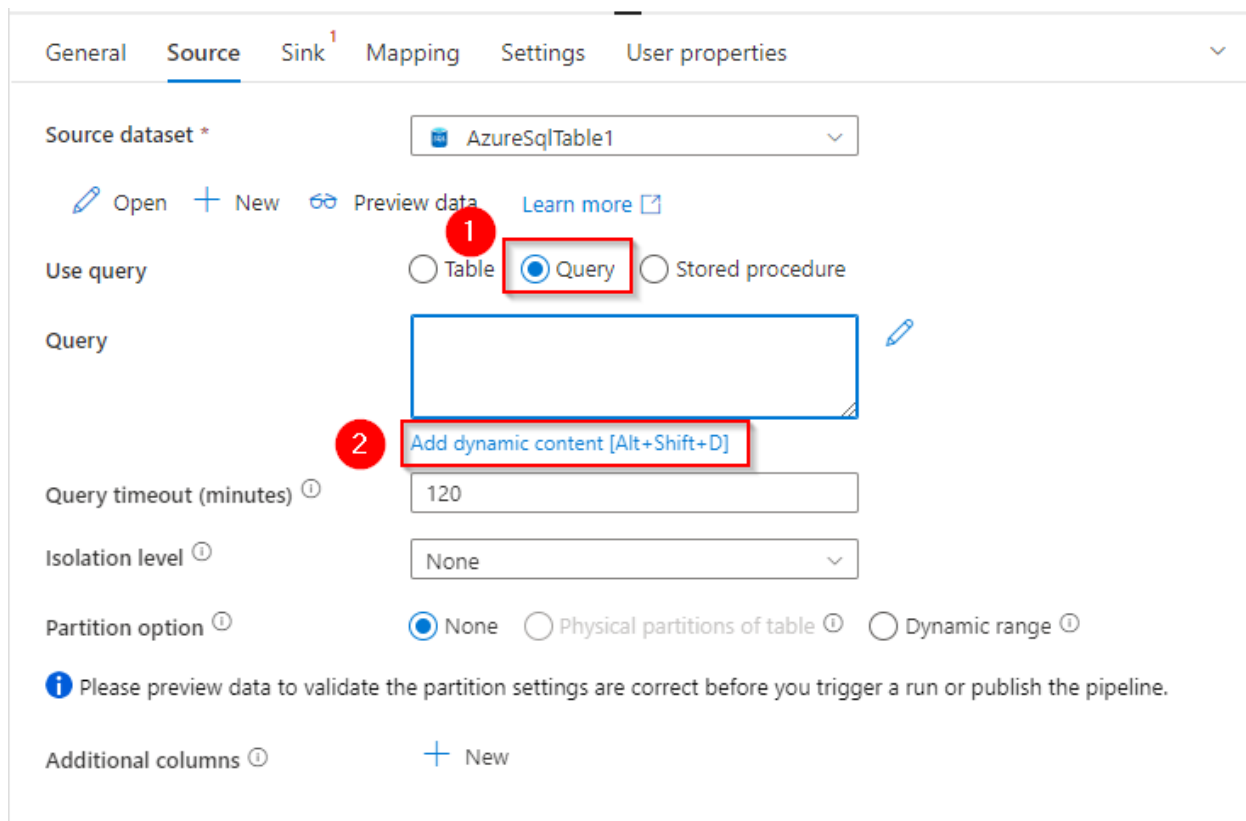


5. Configure the source of copy activity as show in figure:

Use query: Query

Query: Click Add dynamic content and add the following query

select \* from cars2 where manufacturedate > '{@pipeline().parameters.starttime}'  
and manufacturedate <= '{@pipeline().parameters.endtime}'



6. Configure the sink of copy activity, select new dataset.

**Type:** Azure Data Lake Gen2 Storage Account

**File Format:** Delimited Text

**Linked Service:** Select from the drop down if exists other wise create a new one (2).

**File Path:** Select data container (3) and enter directory name output (4). Leave file name empty. Click Ok.

**Set properties**

1 Name  
DelimitedText1

Linked service \*  
AzureDataLakeStorage 2

File path  
data / output 3 / File name 4

First row as header ☐

Import schema  
☒ From connection/store ☐ From sample file ☐ None

> Advanced

7. Open sink dataset (created in previous step)

General Sink 1 Mapping Settings User properties

Sink dataset \*  
DelimitedText1

Open 2 New Learn more

Copy behavior ⓘ  
None

8. Add dataset parameter, **FileName**.

The screenshot shows the 'Parameters' tab for a dataset named 'DelimitedText1'. The interface includes a 'Connection' tab, a 'Schema' tab, and a 'Parameters' tab. The 'Parameters' tab is active, showing a table with columns 'Name', 'Type', and 'Default value'. A new parameter 'FileName' of type 'String' is being added, with a default value of 'Value'. The 'New' button is highlighted with a red box and a red circle with the number 3. The 'FileName' parameter row is highlighted with a red box and a red circle with the number 4. The 'DelimitedText1' dataset is highlighted with a red box and a red circle with the number 1. The 'Parameters' tab is highlighted with a red box and a red circle with the number 2.

Name	Type	Default value
FileName	String	Value

9. In dataset, go to connection tab and give the file name as show:

The screenshot shows the 'Connection' tab for a dataset named 'DelimitedText1'. The interface includes a 'Connection' tab, a 'Schema' tab, and a 'Parameters' tab. The 'Connection' tab is active, showing the 'Linked service' as 'AzureDataLakeStorage', the 'File path' as 'data / output', and the 'Column delimiter' as 'Comma (,)'. The 'File path' field is highlighted with a red box and a red circle with the number 3. The '@dataset().FileName' expression is highlighted with a red box and a red circle with the number 3. The 'DelimitedText1' dataset is highlighted with a red box and a red circle with the number 1. The 'Connection' tab is highlighted with a red box and a red circle with the number 2.

Linked service \* AzureDataLakeStorage Test connection Edit + New Learn mc

File path \* data / output @dataset().FileName

Compression type None

Column delimiter ① Comma (,)

Edit



10. Go back the sink configuration in pipeline and enter the following using add dynamic content for **FileName** properties of Dataset.

```
@concat('cars2_', pipeline().parameters.endtime, '.csv')
```

General Source **Sink** Mapping Settings User properties

Sink dataset \* DelimitedText1 Open + New Learn

Dataset properties ⓘ

Name	Value
FileName	@concat('cars2_', pipeline().parameters.endtime, '.csv')

Copy behavior ⓘ None

Max concurrent connections ⓘ

11. Add a trigger to pipeline

pipeline1 DelimitedText1

Validate Validate copy runtime Debug **Add trigger**

Trigger now

**New/Edit**

Copy data

Copy data1

### Add triggers

Choose trigger... ▼

Search

**+ New**

12. Create new Tumbling window trigger as follows:

Trigger Name: trigger1

Type: Tumbling Window

Start Date: 1/1/12 00:00:00

End Date: 1/1/12 00:15:00

New trigger

Name \*

trigger1

Description

Type \*

Tumbling window

Start Date (UTC) \*

1/1/12 00:00:00

Recurrence \*

Every 5

Minute(s)

☒ Specify an end date

End On (UTC) \*

1/1/12 00:15:00

In Advanced: Max Concurrency: 2

Advanced

Add dependencies



New

Delete



Trigger

Offset

Window ...

No records found

Delay

00:00:00

Max concurrency \*

2

13. Click ok, on the next page enter Trigger Run Parameter  
starttime: @trigger().outputs.windowStartTime  
endtime: @trigger().outputs.windowEndTime

## New trigger

### Trigger Run Parameters


 Parameters that are not provided a value will not be included in the trigger.

Name	Type	Value	
starttime	string	@trigger().outputs.windowSt...	1
endtime	string	@trigger().outputs.windowEn...	2

Then click Ok.

14. Click on Publish all to save the pipeline, dataset, linked service etc.

## Publish all

You are about to publish all pending changes to the live environment. [Learn more](#) 

### Pending changes (4)

NAME	CHANGE	EXISTING
▼ Pipelines		
 pipeline1	(New)	-
▼ Datasets		
 AzureSqlTable1	(New)	-
 DelimitedText1	(New)	-
▼ Triggers		
 trigger1	(New)	-

Click on Publish.

15. When success fully published, go to Monitor Tab and see the pipeline run.

» <<

Home Dashboards

Runs

**Pipeline runs**

Trigger runs

Runtimes & sessions

Integration runtimes

Data flow debug

Notifications

Alerts & metrics

### Pipeline runs

Triggered Debug Rerun Cancel options Refresh Edit columns List Gantt

Filter by run ID or name Chennai, Kolkata, Mu... : Last 24 hours Pipeline name : All Status : All Runs : La

Triggered by : All Add filter X

Showing 1 - 3 items

<input type="checkbox"/> Pipeline name ↑↓	Run start ↑↓	Run end ↑↓	Duration	Triggered by
<input type="checkbox"/> pipeline1	12/13/2022, 7:07:53 PM	12/13/2022, 7:08:05 PM	00:00:12	trigger1
<input type="checkbox"/> pipeline1	12/13/2022, 7:07:27 PM	12/13/2022, 7:07:40 PM	00:00:13	trigger1
<input type="checkbox"/> pipeline1	12/13/2022, 7:07:27 PM	12/13/2022, 7:07:41 PM	00:00:14	trigger1

16. Verify the files in the storage account.