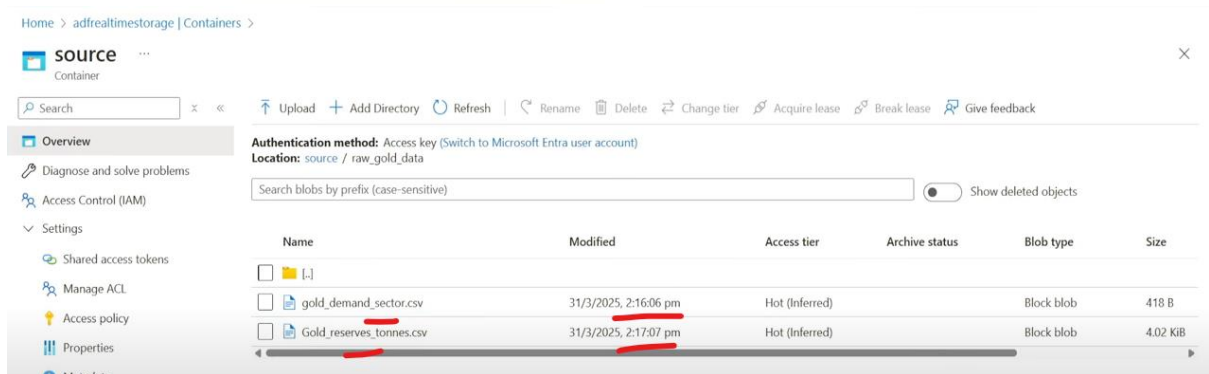
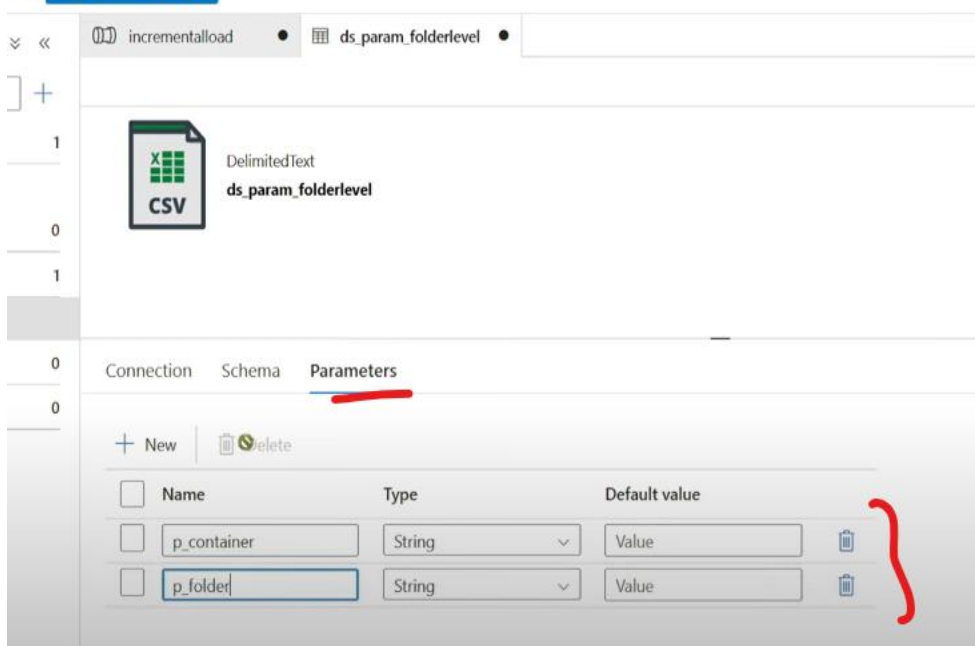


Scenario - 1


Incremental Loading



Create a LS and a DS(make it parameterized so that it can be used for both source and sink)



incrementalload • ds_param_folderlevel •

 DelimitedText
ds_param_folderlevel

Connection Schema Parameters

Linked service * ls_datalake Test connection Edit New Learn more

File path @dataset().p_container / @dataset().p_folder / File name

Compression type No compression

Column delimiter Comma (,)

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

Factory Resources Validate all Publish all Preview experience Off

incrementalload • ds_param_folderlevel •

Activities cop

Move and transform

Copy data

Copy data

Copy data1

General Source Sink Mapping Settings User properties

Dataset properties

Name	Value	Type
p_container	source	string
p_folder	raw_gold_data	string

File path type File path in dataset ☒ Wildcard file path List of files

Wildcard paths {source container} / Wildcard folder path / *.csv Add dynamic content

Start time (UTC) End time (UTC)

42 • Scenario-1 Incremental Loading in ADF from Azure Data Lake

Validate all Publish all Preview experience Off

incrementalload • ds_param_folderlevel •

Activities cop

Move and transform

Copy data

Copy data

Copy data1

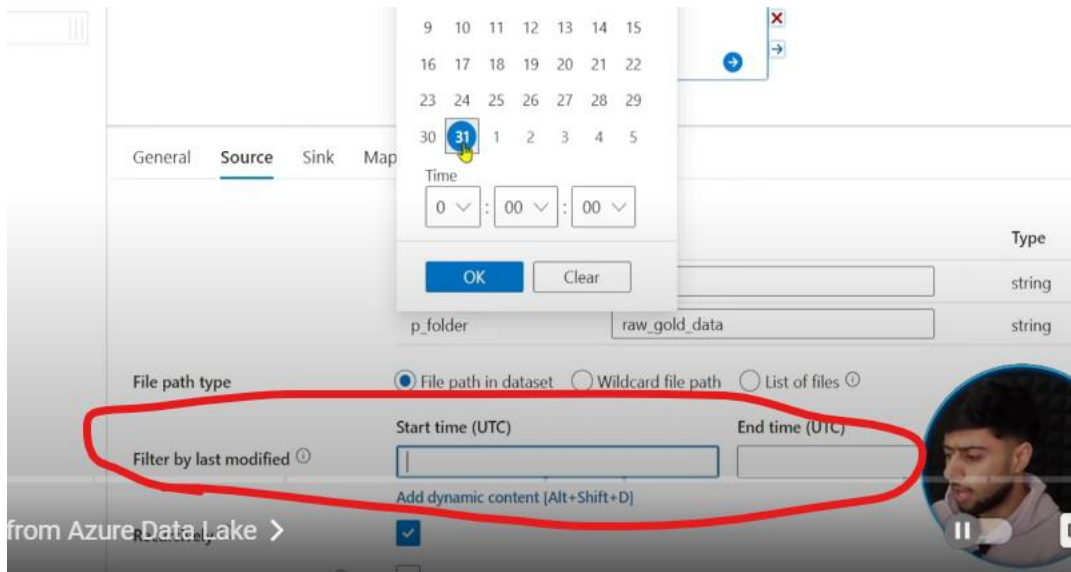
General Source Sink Mapping Settings User properties

Sink dataset * ds_param_folderlevel Open New Learn more

Dataset properties

Name	Value	Type
p_container	destination	string
p_folder	raw_gold_data	string

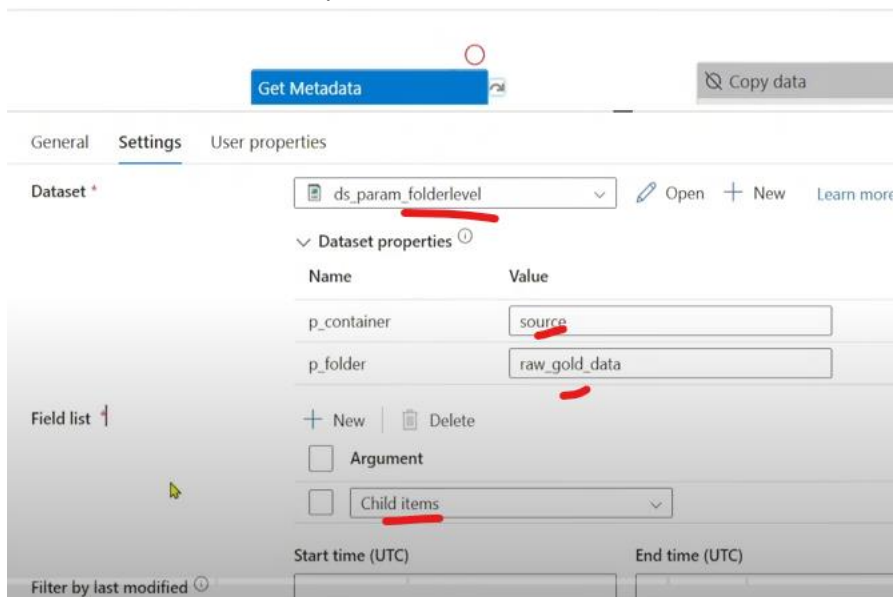
Add dynamic content (Alt+Shift+D)



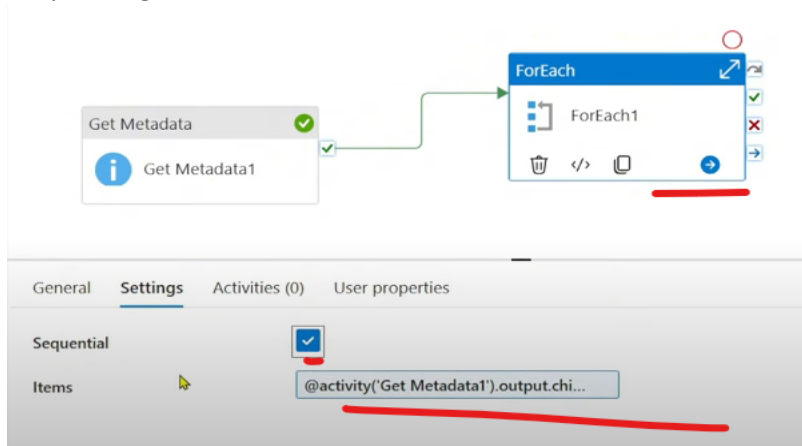
Have to play around with last modified filter

Now lets start designing the pipeline

Use a Get Metadata activity to retrieve all the files



Loop through all the childitems



Use a Get Metadata activity inside For-Each to retrieve the last modified date of each files
Create a new dataset



Connection Schema Parameters

+ New Delete

<input type="checkbox"/>	Name	Type	Default value	
<input type="checkbox"/>	p_container	String	Value	
<input type="checkbox"/>	p_folder	String	Value	
<input type="checkbox"/>	p_file	String	Value	



Connection Schema Parameters

Linked service * ls_datalake Test connection Edit + New Learn more

File path @dataset().p_container / @dataset().p_folder / @dataset().p_file

Compression type No compression

Column delimiter Comma (,)

Validate Debug Add trigger

incrementalload > ForEach1

Get Metadata2

General Settings User properties

p_container source

p_folder raw_gold_data

p_file @item().name

Field list

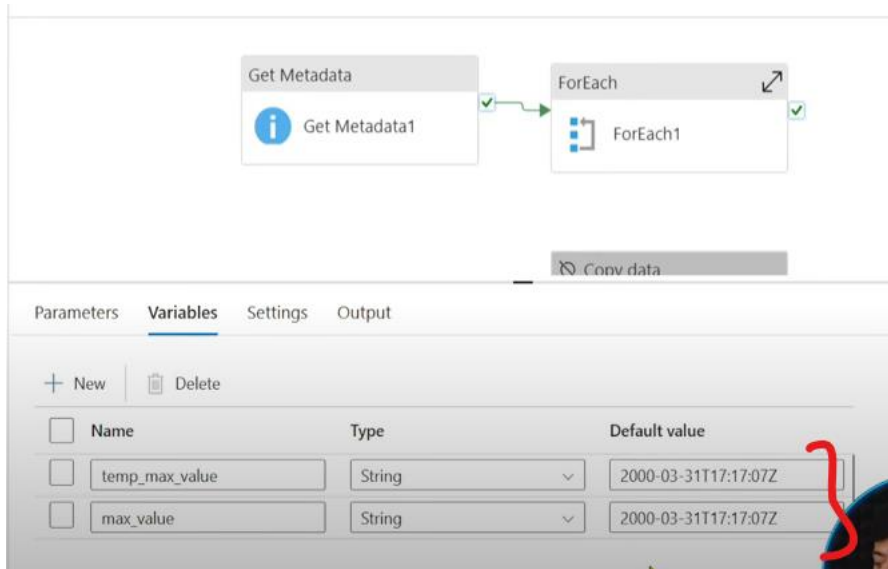
+ New Delete

☐ Argument

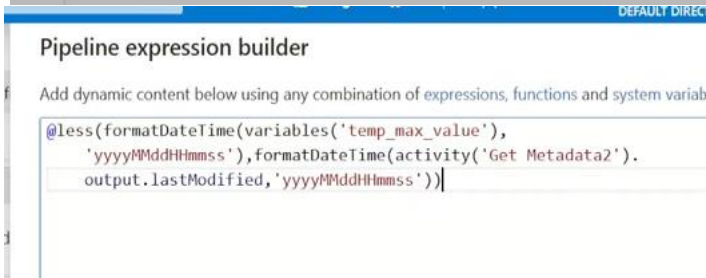
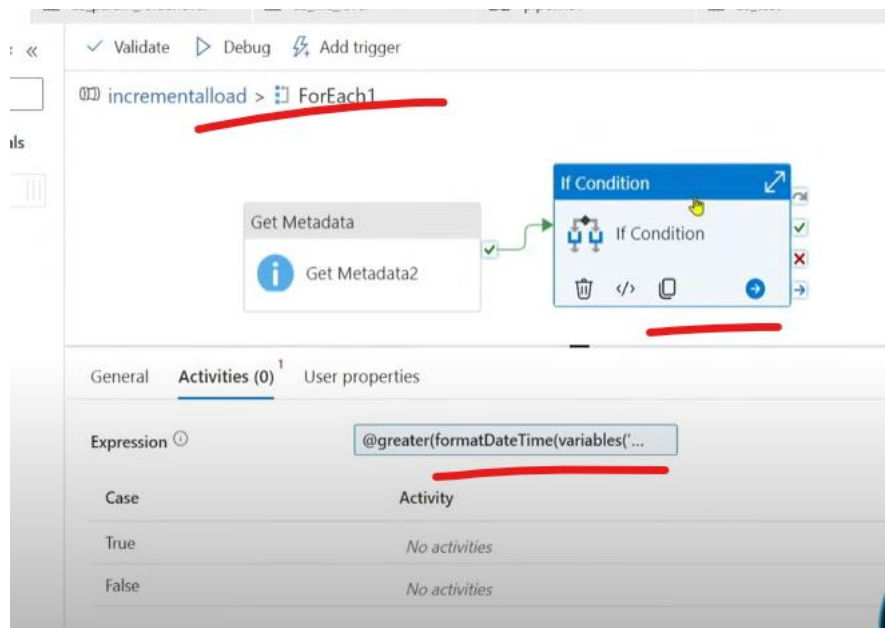
☐ Last modified

Start time (UTC) End time (UTC)

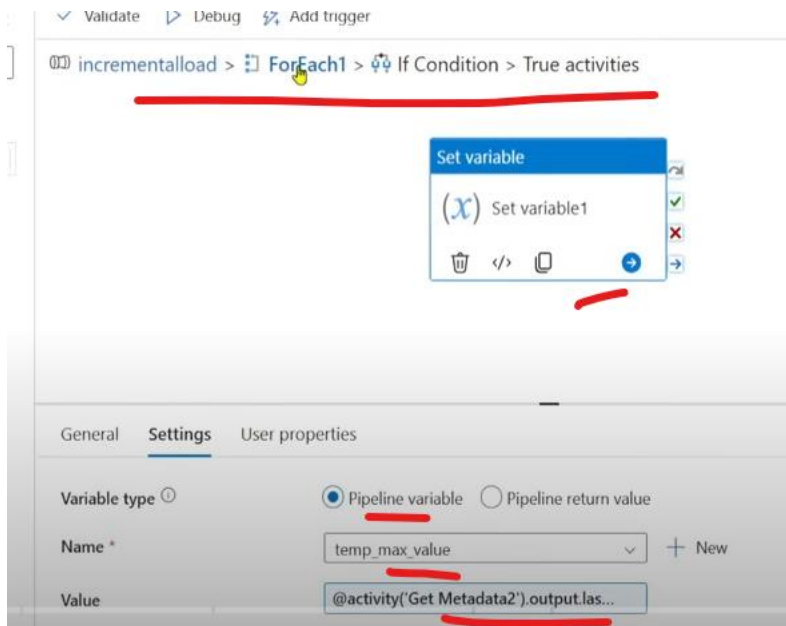
We need to save the last modified date in a variable so that we can compare the dates.



If output of last modified is greater than temp_max_value then replace value of max_value with last modified date, if not keep it as is.



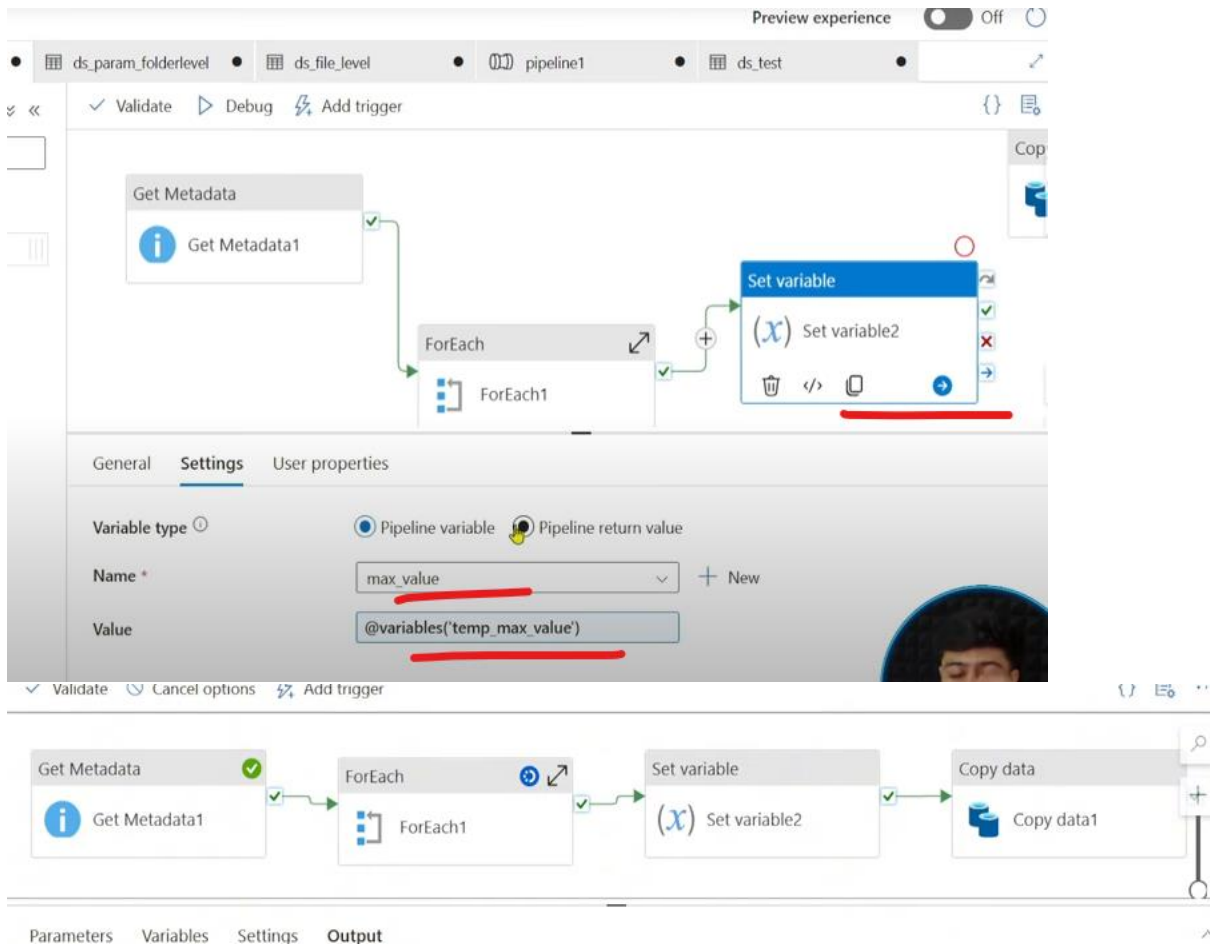
You can use greater also, just swap the order



Pipeline expression builder

Add dynamic content below using any combination of expressions, functions, and operators.

@activity('Get Metadata2').output.lastModified



Validate Validate copy runtime Debug Add trigger

General Source Sink Mapping Settings User properties

Source dataset * ds_param_folderlevel Open + New Preview data Learn more

Dataset properties

Name	Value
p_container	source
p_folder	raw_gold_data

File path type ☐ File path in dataset ☒ Wildcard file path ☐ List of files

Wildcard paths {source container} / Wildcard folder path / *.csv

incrementalload ds_param_folderlevel ds_file_level pipeline1 ds_test

Validate Validate copy runtime Debug Add trigger

General Source Sink Mapping Settings User properties

File path type ☐ File path in dataset ☒ Wildcard file path ☐ List of files

Wildcard paths {source container} / Wildcard folder path / *.csv

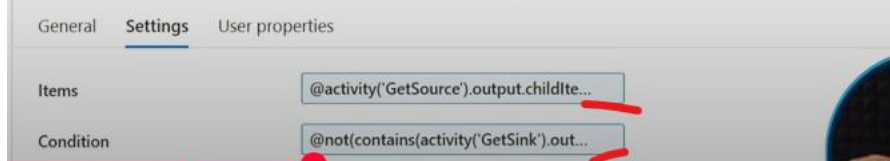
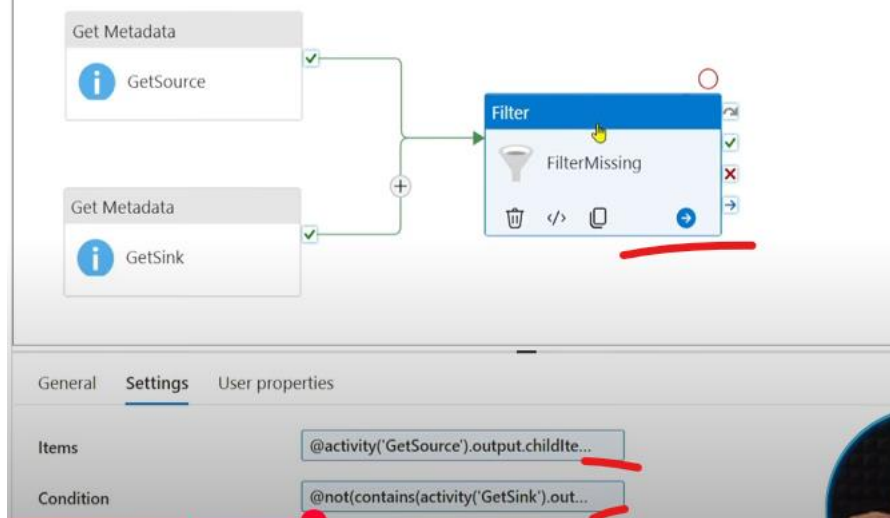
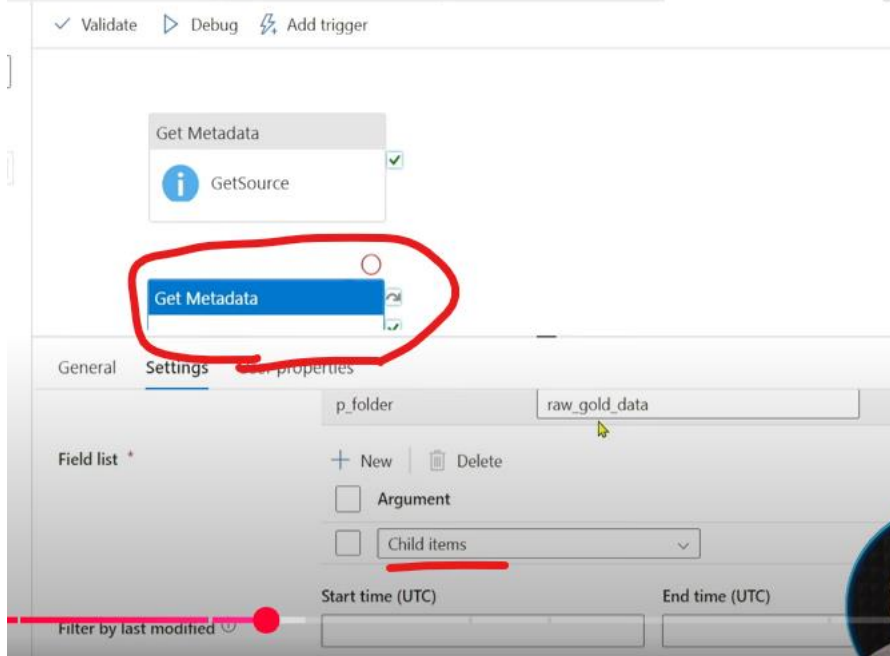
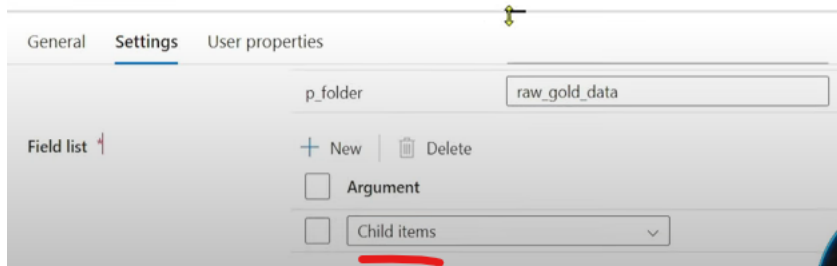
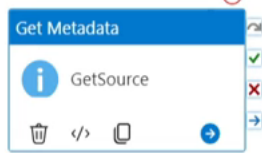
Filter by last modified Start time (UTC) End time (UTC)

Recursively ☒

Enable partitions discovery ☐

Scenario - 2

Copy only those files which are not available in the destination (previously done)



Pipeline expression builder

Add dynamic content below using any combination of [express](#)

```
@activity('GetSource').output.childItems
```

Pipeline expression builder

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

```
@not(contains(activity('GetSink').output.childItems,item()))
```

The screenshot shows a pipeline editor with a visual flow. Two source activities feed into a 'Filter' activity. The 'Filter' activity has a 'FilterMissing' icon. Its output is connected to a 'ForEach' activity. The 'ForEach' activity has a 'ForEach1' icon. Below the visual flow, the 'Settings' tab is active, showing the 'Items' property set to '@activity('FilterMissing').output'. A red circle highlights this expression.

Pipeline expression builder

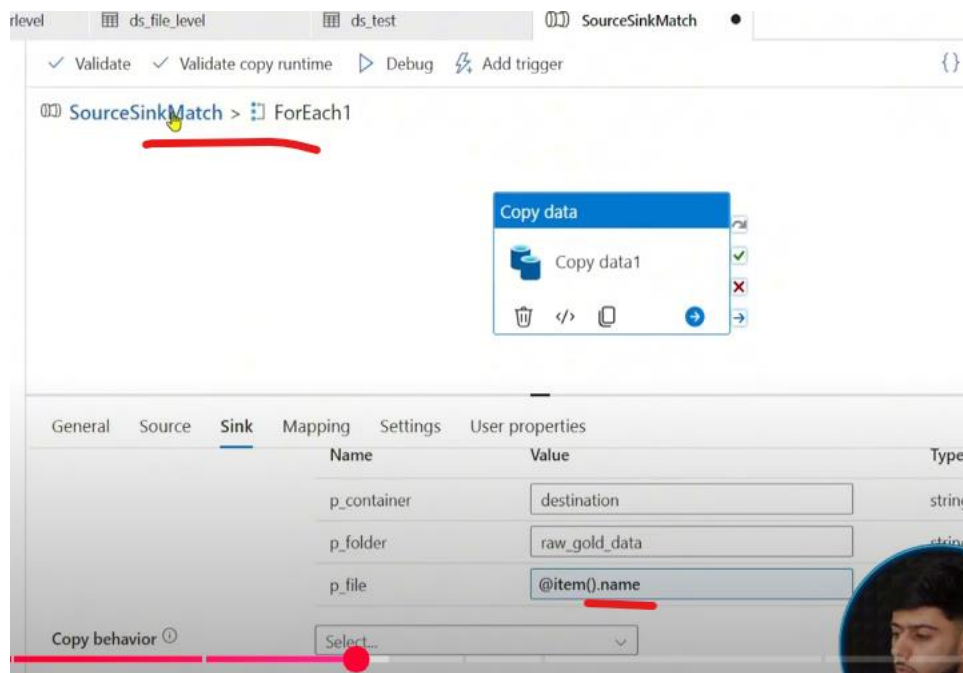
Add dynamic content below using any combination of [expressions](#), [fun](#)

```
@activity('FilterMissing').output.Value
```

SourceSinkMatch > ForEach1

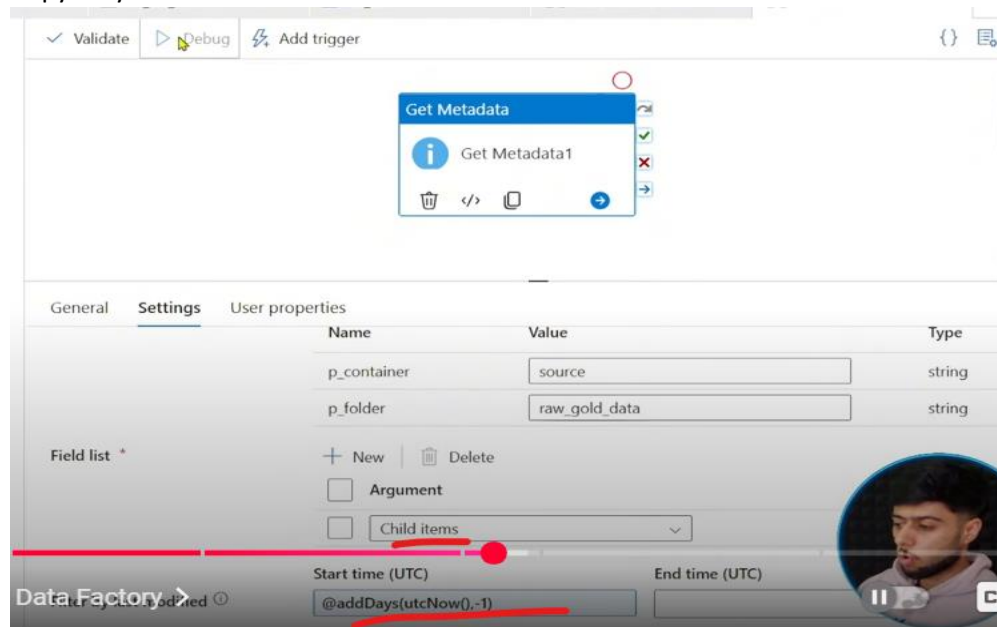
The 'Copy data' activity settings are shown. The 'Dataset properties' section is expanded, showing a table with columns 'Name', 'Value', and 'Type'. The 'p_file' property is highlighted with a red circle and has its value '@item().name' underlined in red.

Name	Value	Type
p_container	source	
p_folder	raw_gold_data	
p_file	@item().name	



Scenario - 3

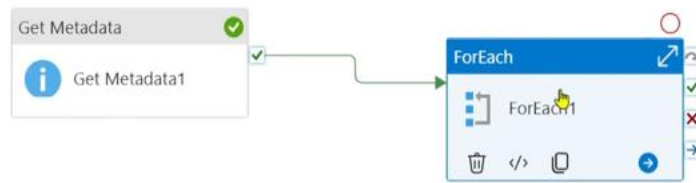
Copy only those files which are submitted within 24 hours



Pipeline expression builder

Add dynamic content below using any combin

```
@addDays(utcNow(),-1)
```



Complete the flow

Now lets say we have to copy the files which are older than 7 days, same method, instead of Start time provide End time → @addDays(utcNow(),-7)

Scenario - 4

Delete all the files that are not older than 7 days

M1: same as last approach, provide end time in Get Metadata and inside For-Each provide Delete activity

M2:

Create a pipeline parameter to pass number of days dynamically

The screenshot shows the 'Parameters' tab in a pipeline editor. It features a table with columns for 'Name', 'Type', and 'Default value'. A single parameter is defined: 'p_last7days' of type 'Int' with a default value of 'Value'. There are '+ New' and 'Delete' buttons at the top left of the table.

Name	Type	Default value
p_last7days	Int	Value

The screenshot shows the 'Settings' tab in a pipeline editor. It displays fields for 'p_container' (set to 'source') and 'p_folder' (set to 'raw_gold_data'). Below these fields is a 'Field list' section with '+ New' and 'Delete' buttons, and a table with columns for 'Argument' and 'Child items'.

Argument	Child items
----------	-------------

The image shows a sequence of steps to configure a 'deleteFiles' pipeline activity:

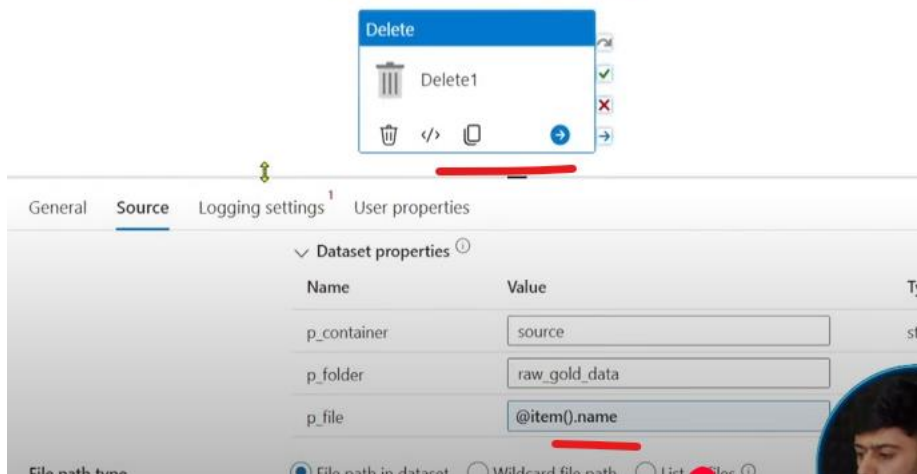
- Step 1:** A 'Get Metadata' activity (Get Metadata1) is connected to a 'ForEach' loop (ForEach1).
- Step 2:** The 'Settings' tab for 'ForEach1' is shown. 'Sequential' is checked. The 'Items' field is set to `@activity('Get Metadata1').output.chi...`.
- Step 3:** The 'Field list' for the 'deleteFiles' activity is shown. The 'p_file' field is set to `@item().name`. The 'Last modified' checkbox is checked. A filter is applied: 'Filter by last modified'.
- Step 4:** The 'Activities (0)' tab is shown. The 'Expression' field contains the code: `@greater(activity('Get Metadata2').output.lastModified, addDays(utcNow(), pipeline().parameters.p_last7days))`.

Pipeline expression builder

Add dynamic content below using any combination of expressions, functions and system va

```
@greater(activity('Get Metadata2').output.lastModified, addDays(
  utcNow(), pipeline().parameters.p_last7days))
```

deleteFiles > ForEach1 > If Condition1 - True activities



Pass -7 in parameter as we havent considered minus in the code

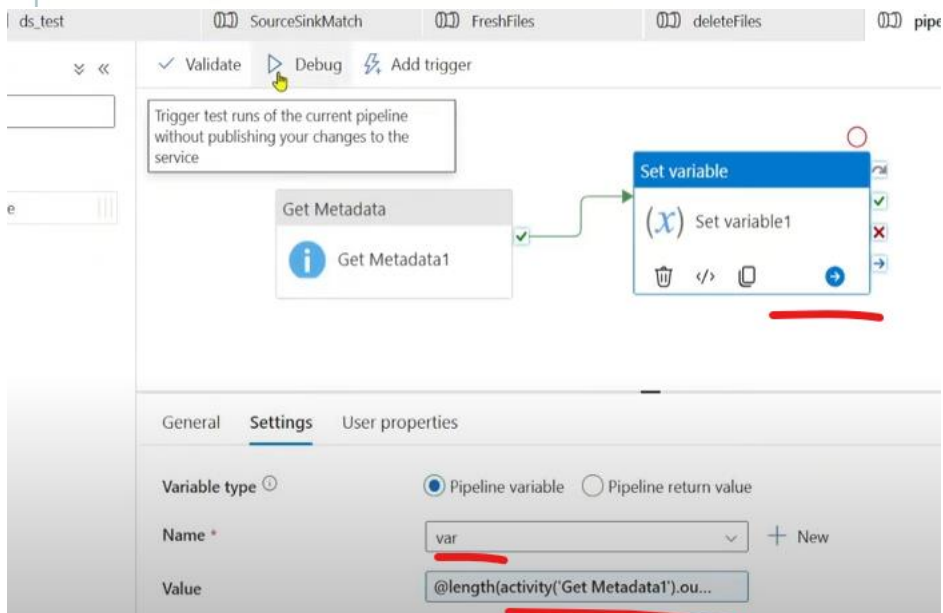
Scenario - 5

Count the number of files in a container

Pipeline expression builder

Add dynamic content below using any combination of expressions, functions and s

@length(activity('Get Metadata1').output.childItems)|

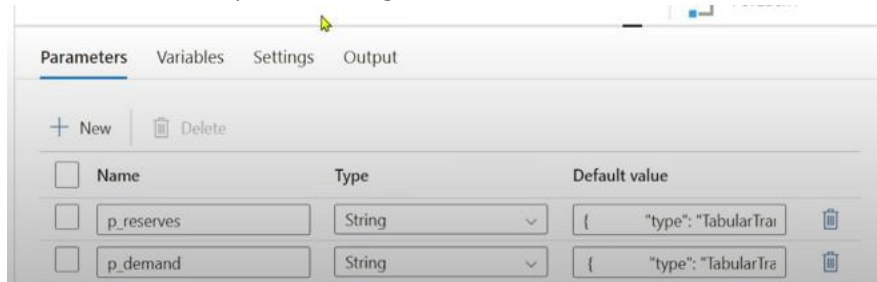


Scenario - 6

Dynamic Schema Mapping

Refer previous documentation : you can see we have passed the parameters as Object

Extra: You can also pass as String

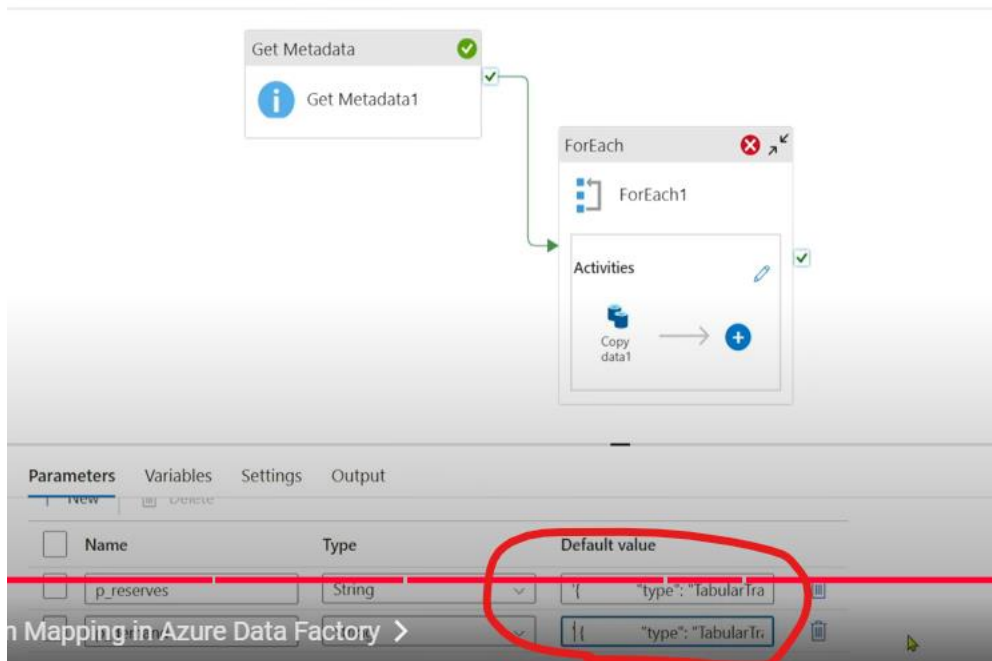


Pipeline expression builder

Add dynamic content below using any combination of expressions, functions and system variables

```
@if(contains(item().name,'demand'),json(pipeline().parameters.p_demand),json(pipeline().parameters.p_reserves))
```

M1:



Pass it like a string by encapsulating within a string

M2: Remove the encapsulated string

ParametersVariablesSettingsOutput

+ New

Delete

<input type="checkbox"/>	Name	Type	Default value	
<input type="checkbox"/>	p_reserves	String	{ "type": "TabularTrai	
<input type="checkbox"/>	p_demand	String	{ "type": "TabularTra	

time-ws

Search factory and documentation

anishamdaaz@gmail.com

DEFAULT DIRECTORY

Publish all

SourceSinkMatchFreshFiles

ValidateValidate copy runtimeDebugAd

6DynamicMap > ForEach1

0

GeneralSourceSinkMappingSettingsUser

Pipeline expression builder

Add dynamic content below using any combination of expressions, functions and system variables.

@if(contains(item().name,'demand'),json(string(pipeline().parameters.p_demand)),json(string(pipeline().parameters.p_reserves)))

Position 4 Syntax error: Missing enclosing parenthesis

Clear contents

ForEach iteratorActivity outputsParametersSystem variables