Azure Data Engineer: Batch 6

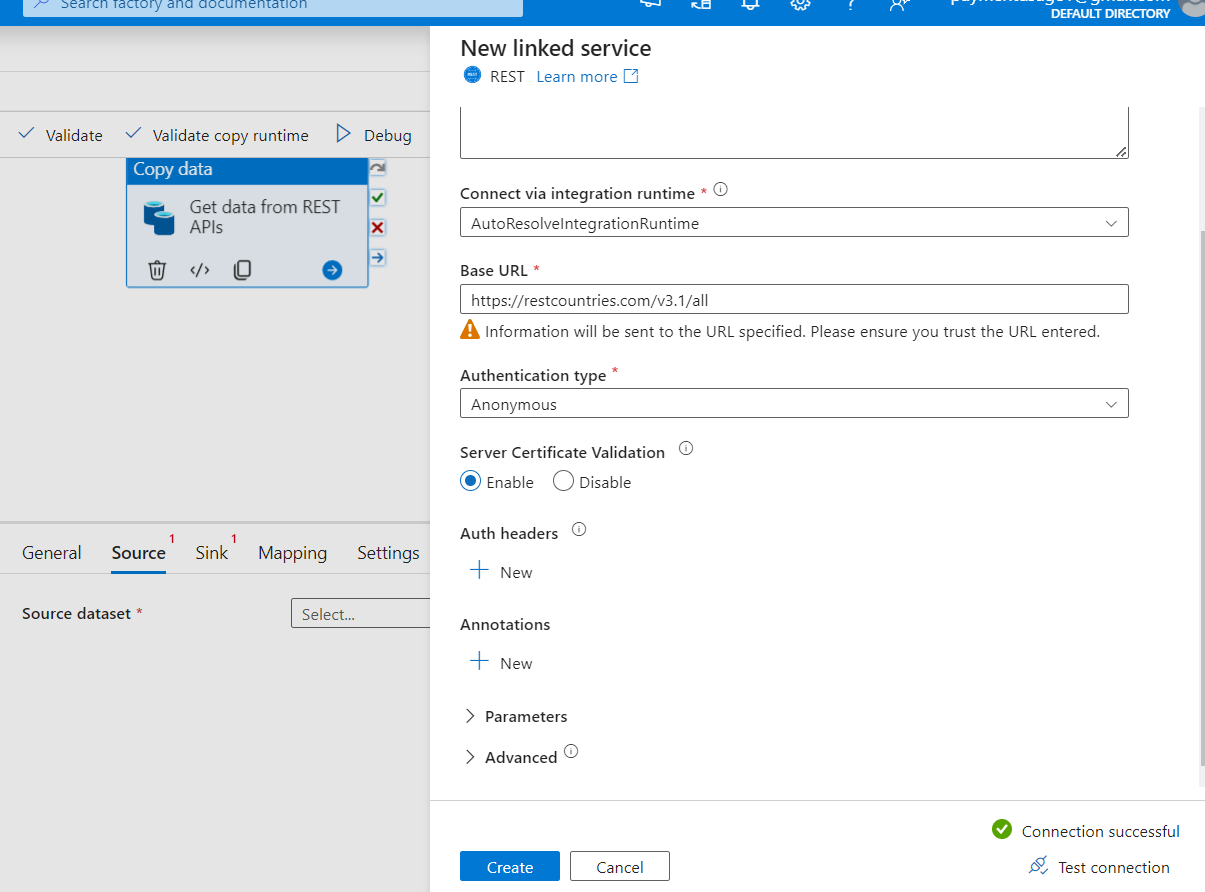
(October 2023)

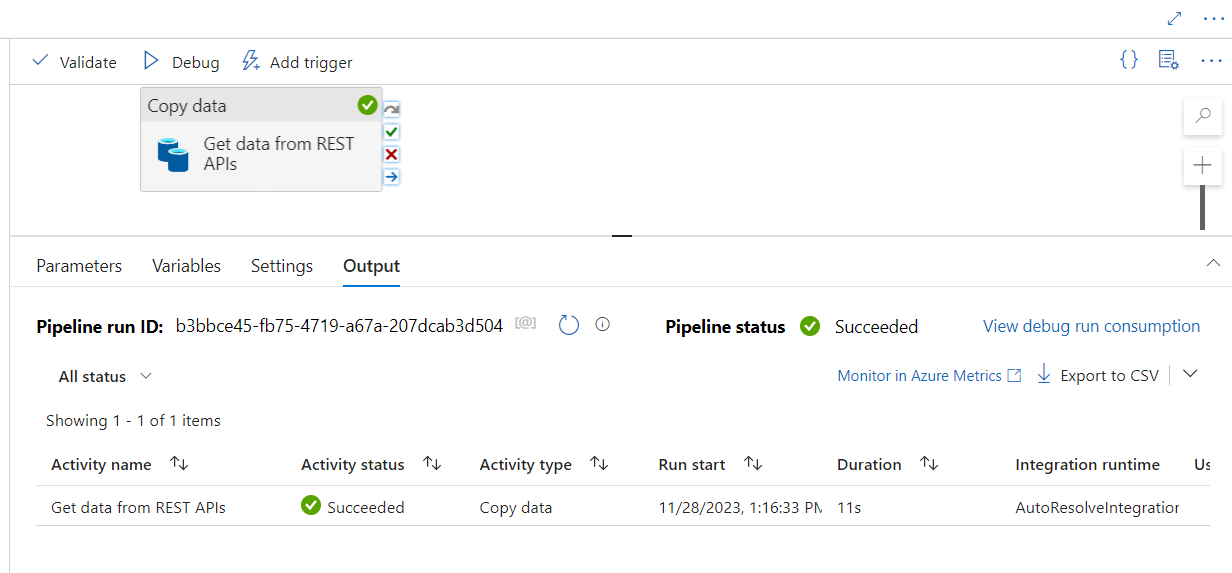
**Submitted by: Roopmathi Gunna**

# Assignment 10: ADF Data Flow questions

## Create a pipeline to fetch the All countries data from Rest API (<https://restcountries.com/v3.1/all)> and save it as parquet file in ADLS.

Step 1 Create the Linked Service to connect to the countries REST API

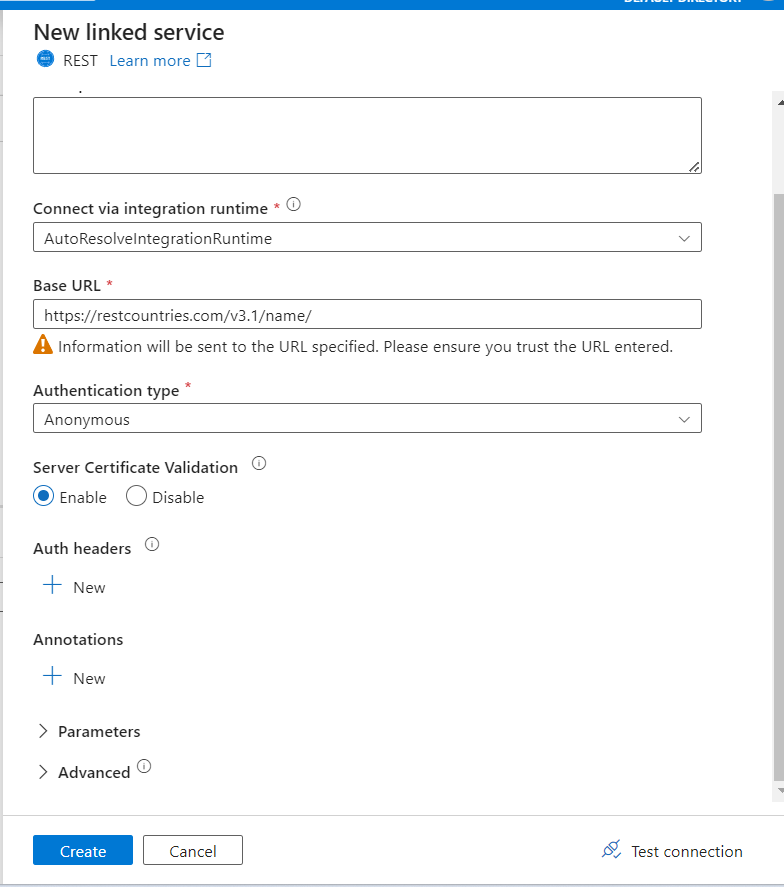




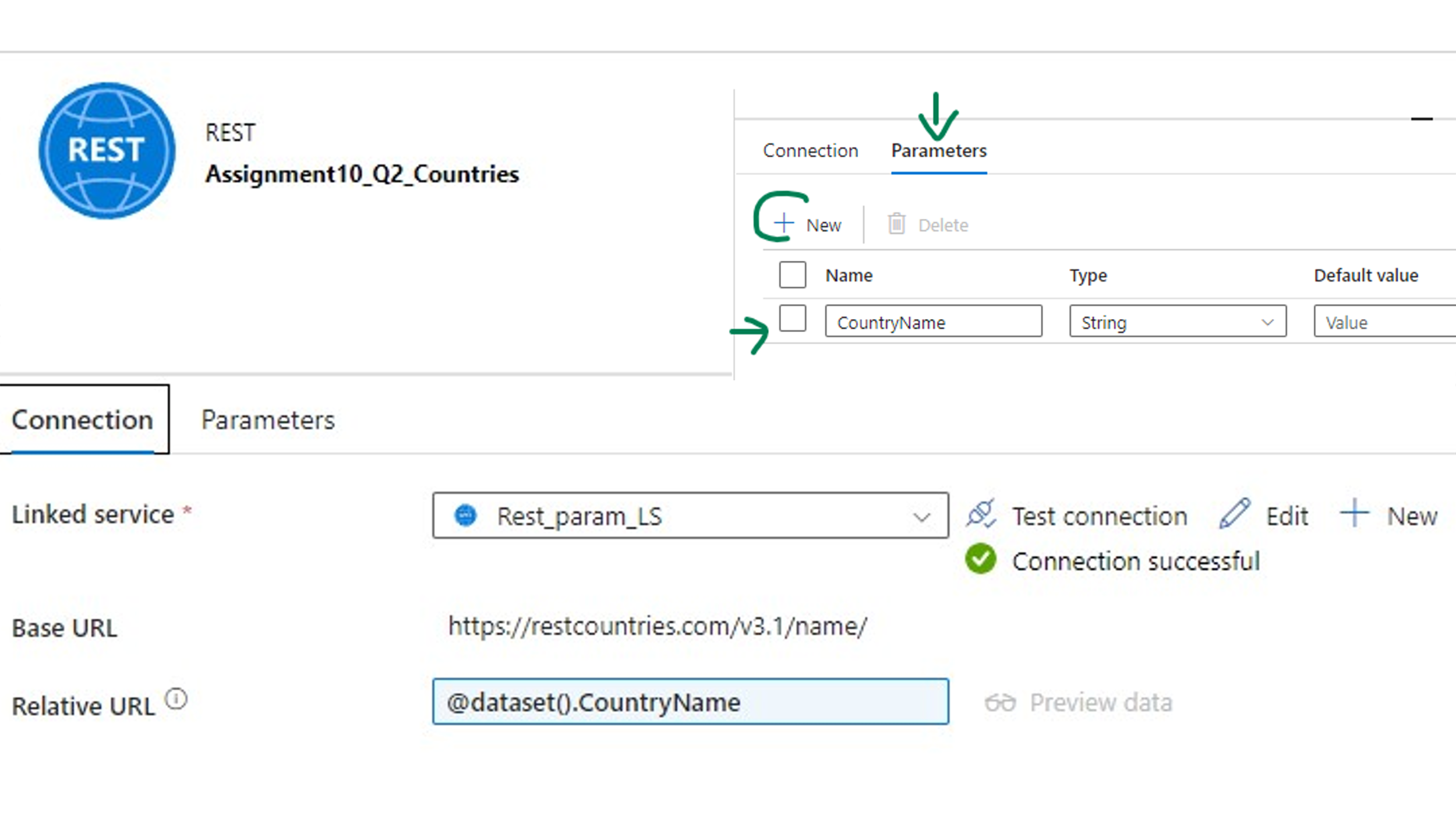
## **Tricky**:

## Create a pipeline to fetch the 5 countries (India,us,uk,china,russia) data from Rest API ([https://restcountries.com/v3.1/name/{name}](https://restcountries.com/v3.1/name/%7Bname%7D)  here replace the {name} with a Country name like <https://restcountries.com/v3.1/name/us)> and save it in a separate file as JSON with File name equal to Country name.

Step 1 Create a new LinkedIn service.



1. **Base URL and Relative URL Structure:**
   * We have a REST API with a base URL that remains constant (**https://restcountries.com/v3.1/name**), and for each country, I need to append a dynamic part, known as the relative URL.



1. **Dynamic Country Names:**
   * To dynamically insert country names into the relative URL, I have used a Lookup Activity.
     + The Lookup Activity is responsible for retrieving the list of country names. This is a CSV dataset from my ADLS location that contains the names of the countries I want to fetch data for.

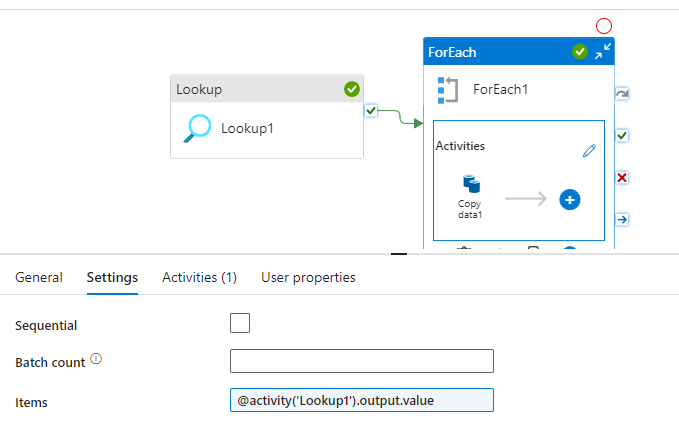
A screenshot of a computer

Description automatically generated

1. **For Each Activity:**

Following the Lookup Activity, I have e implemented a For Each Activity.

* + The For Each Activity iterates through the list of country names obtained from the Lookup Activity. For each iteration, it passes the current country name as a parameter to the next step.



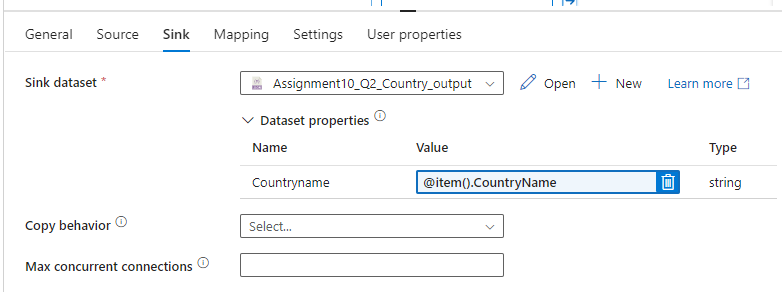
1. **Copy Activity with Dynamic Relative URL:**

Inside the For Each loop, I have a Copy Activity.

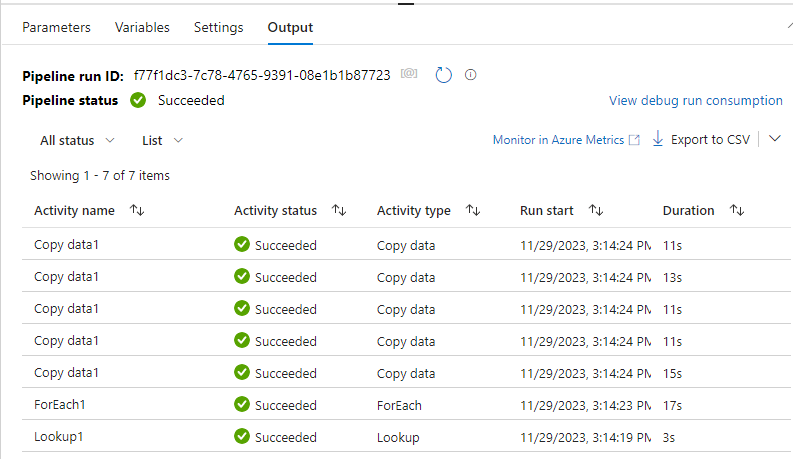
* + The Copy Activity is configured to call the REST API using the base URL and a relative URL that includes the dynamic country name parameter.
  + This ensures that the REST API is called for each country with a unique URL tailored to that specific country.

A screenshot of a computer

Description automatically generated



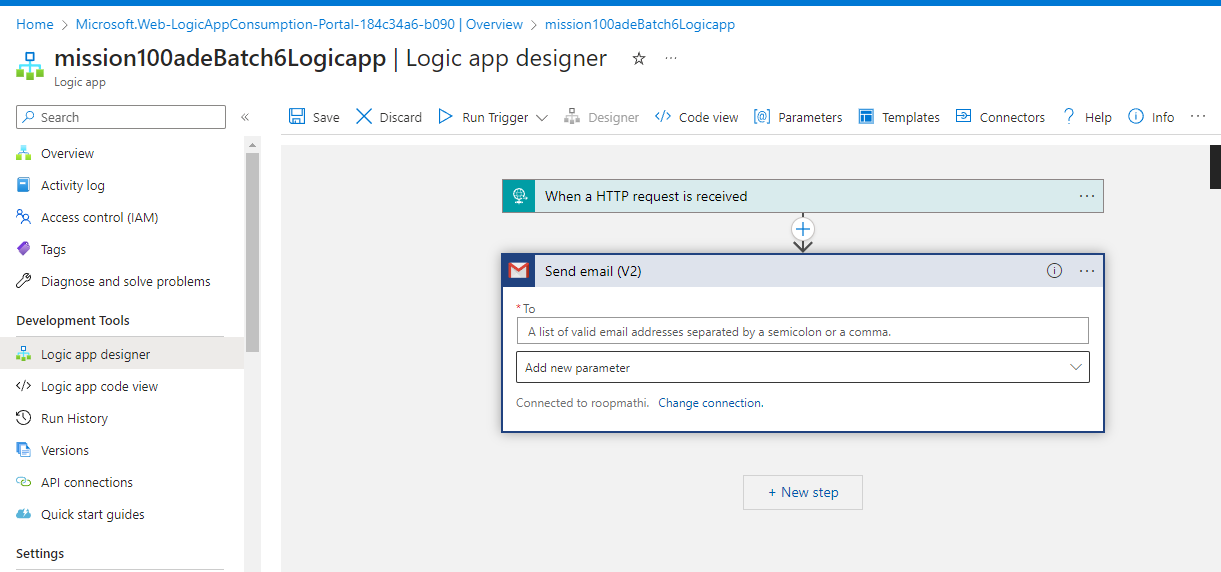
My pipeline executed successfully as seen below:



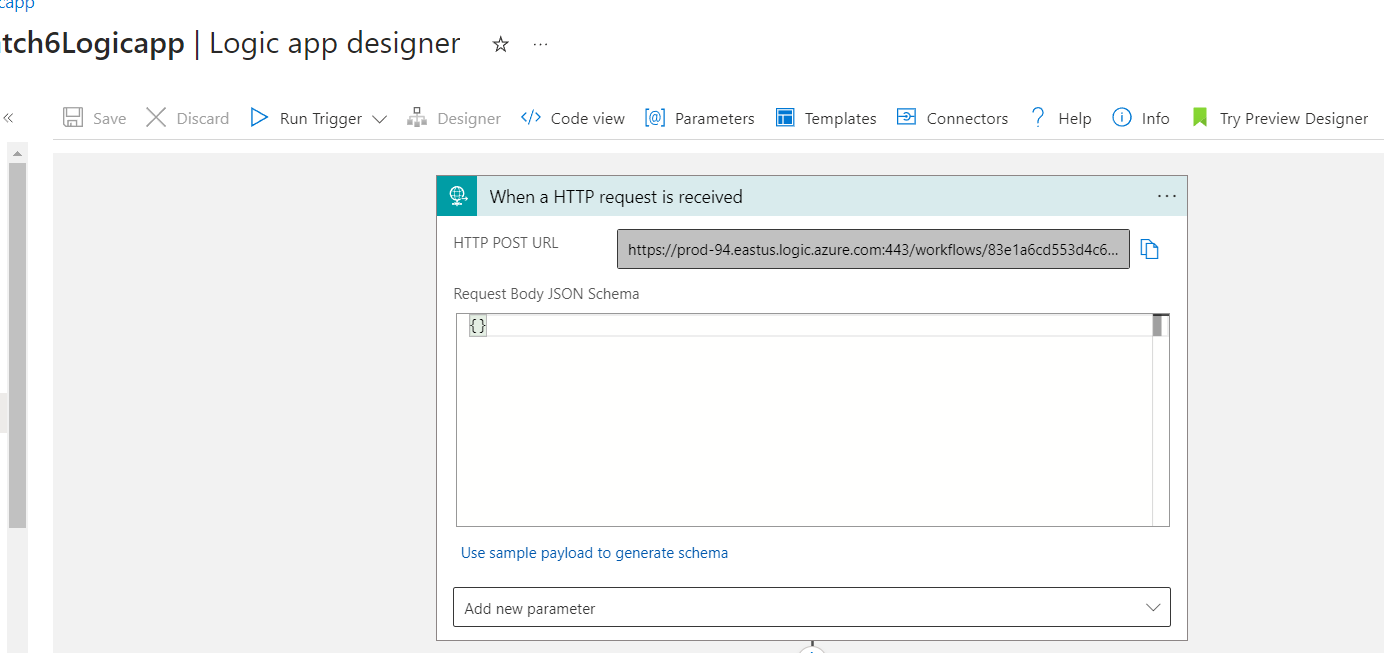
A screenshot of a computer

Description automatically generated

## Let’s first create Logic App example taught by Deepak in class before we dive into Trigger.



[URL to call that Logic App:](https://prod-94.eastus.logic.azure.com:443/workflows/83e1a6cd553d4c64bd7d7c88349baaa9/triggers/manual/paths/invoke?api-version=2016-10-01&sp=%2Ftriggers%2Fmanual%2Frun&sv=1.0&sig=PXNuOr0k4lfk99vEzNSy2Qip6AW7BIxMNA3iKClqw5w)



A screenshot of a computer

Description automatically generated

## This is my basic Logic App Data Pipeline (no-dynamic parameters):

A screenshot of a computer

Description automatically generated

**Copy activity** settings:

A screenshot of a computer

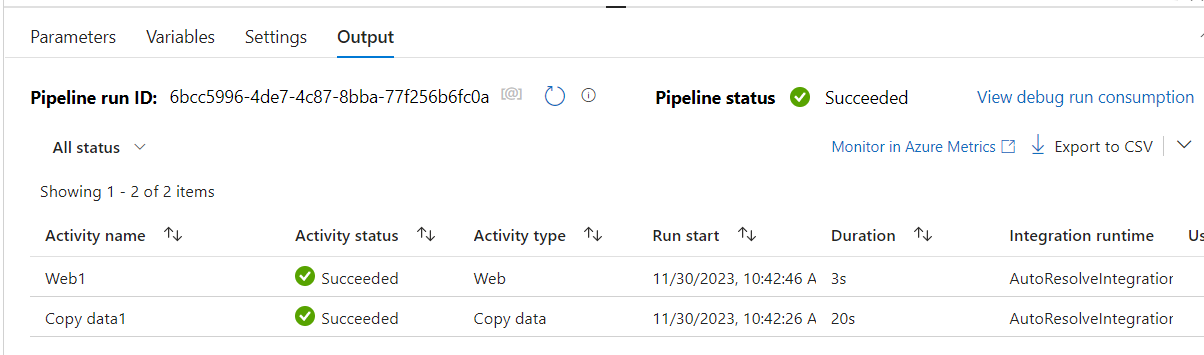
Description automatically generated

**Web activity** Settings:

Use the URL generated in Logic App here in the URL settings.

A screenshot of a computer screen

Description automatically generated

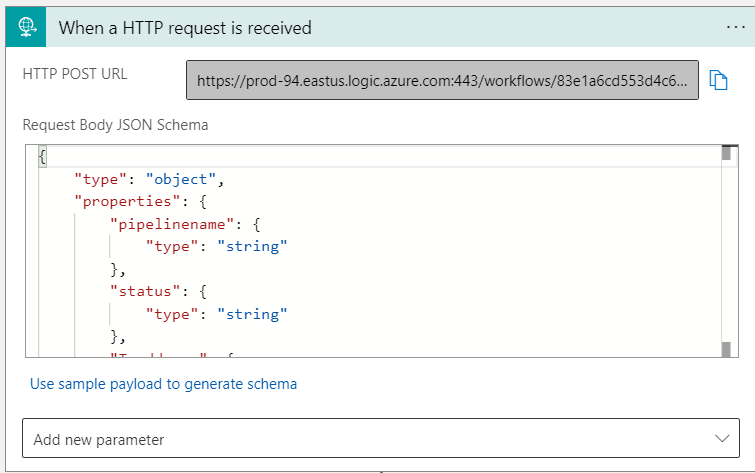


## Got the email! I chose Gmail service in the Logic app

A white background with black and white clouds

Description automatically generated

Now repeating this example using Dynamic Email, Subject and Body of the email in Logic App



A screenshot of a computer

Description automatically generated

A screenshot of a computer

Description automatically generated

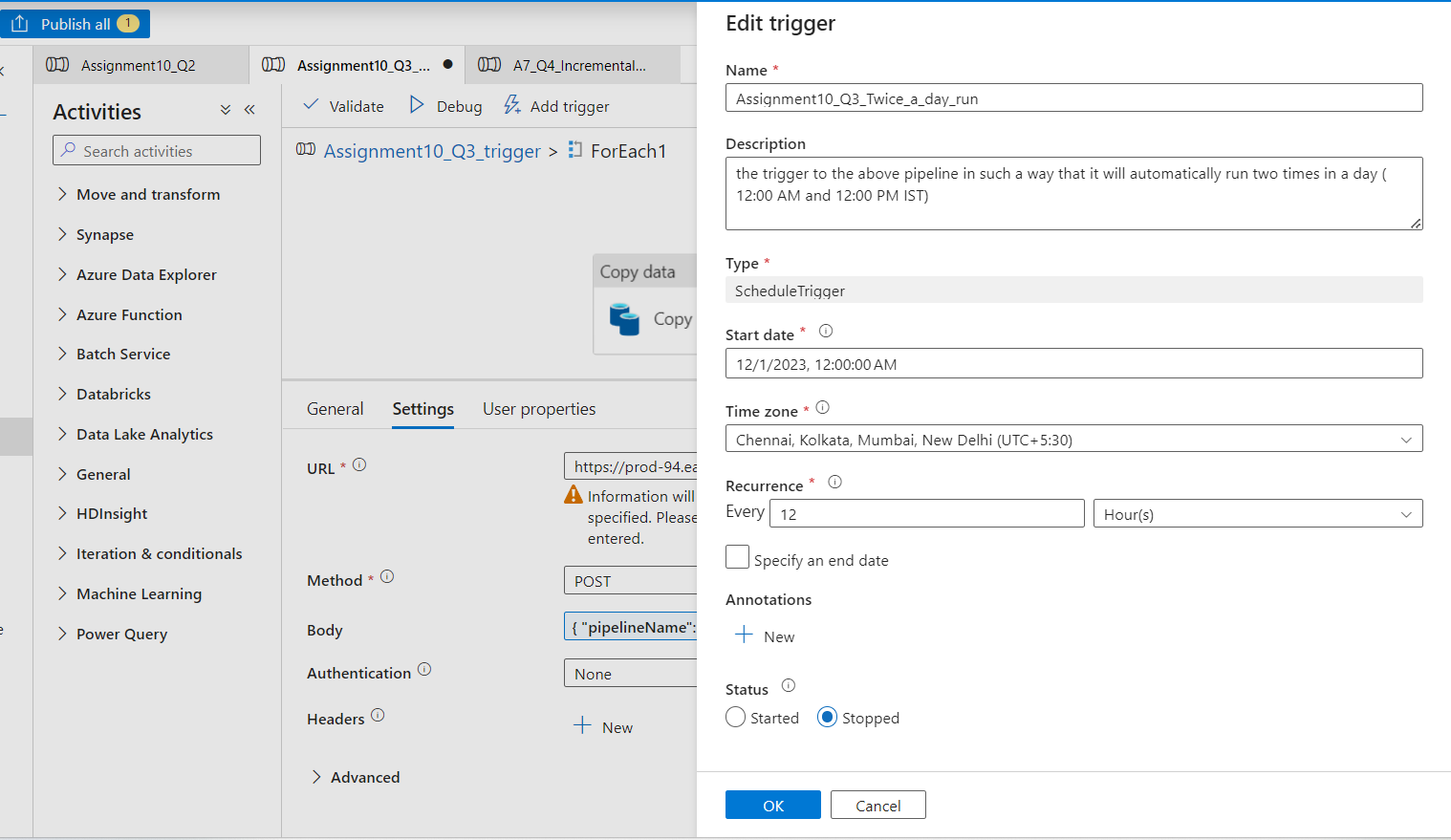
A screenshot of a computer

Description automatically generated

## A screenshot of a computer Description automatically generated

## 3. Add the trigger to the above pipeline in such a way that it will automatically run two times in a day ( 12:00 AM and 12:00 PM IST)

## Tip: Incase you can't able to do Q2, do this on Q1



A screenshot of a computer

Description automatically generated

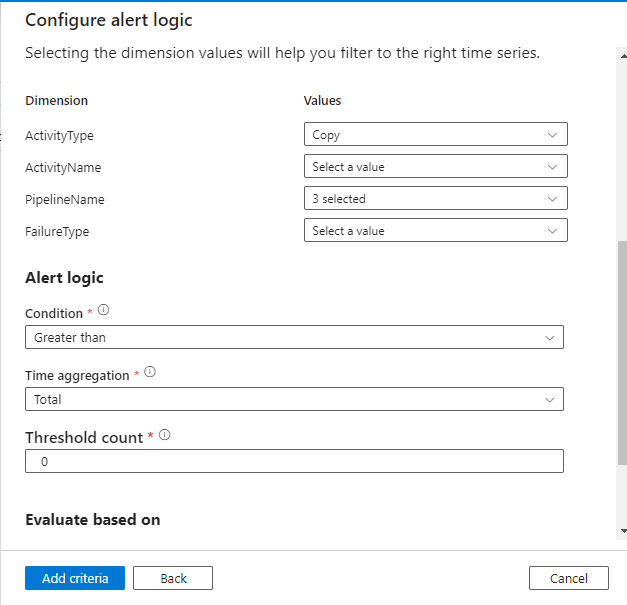
A screenshot of a computer

Description automatically generated

4. Now to make it more realistic, as soon as any copy activity fails in the above pipeline send an automated email to notify user.

For this we will use the Alerts and Metrics feature in ADF > Monitor

Lets first create an Alert titled “ FailAlert\_on\_Copy\_Activity”



A screenshot of a computer program

Description automatically generated

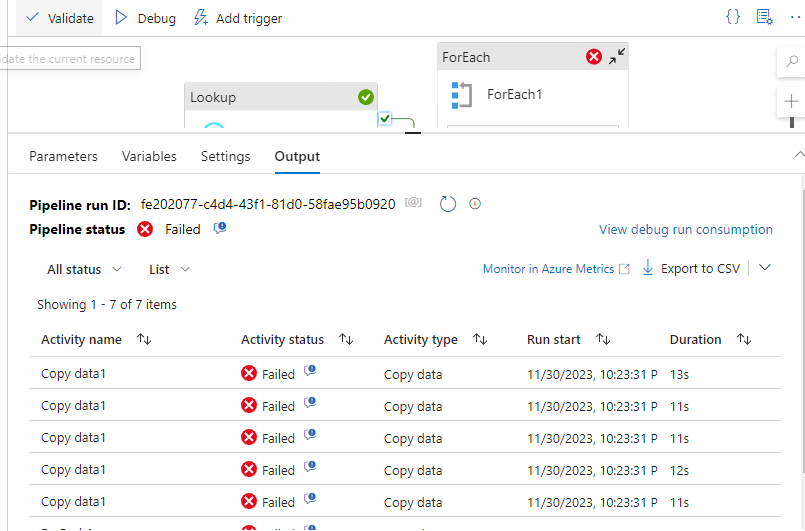
A screenshot of a computer

Description automatically generated

A screenshot of a computer

Description automatically generated

Ok so now lets tests this, intentionally made an error in copy activity on Source such that my Copy activity will fail and we see here the pipeline has failed:



At 12:00 AM IST, the scheduled pipeline failed, triggering an email alert for Copy Activity failure. With a failure alert evaluation window set to the last 30 minutes, a second alert was received at 12:30 AM .

A screenshot of a computer error

Description automatically generated

A screenshot of a computer error

Description automatically generated

1. Make the mail more generic by sending the Pipeline name, Status, SenderMailiD, and Name within the mail. Ensure mail will come in case of both success and Fail.

In this task, I customized email notifications within a Logic App triggered by an Azure Data Factory (ADF) pipeline. The goal was to send a more generic email that includes crucial details such as the pipeline name, status, sender email ID, recipient name, and specific country name as Iam using a For Each activity in my Pipeline to fetch corresponding REST APIs for countries India,us,uk,china,russia.

**Step 1: HTTP Request Trigger Configuration**

1. **HTTP Request Trigger**: In the Logic App, I configured an HTTP Request trigger to expect parameters such as Name, CountryName, PipelineName, and PipelineStatus. The schema was set up to include these properties.



**Step 2: Fetching Data in ADF Pipeline**

1. **ADF Pipeline Setup**: In the parent ADF pipeline, I added a Web Activity following a Copy Activity running in a For Each loop. The Web Activity calls a Logic App using a dynamic Body, passing parameters like PipelineName, status, ToAddress, countryName, and Name.

A screenshot of a computer

Description automatically generated A screenshot of a computer code

Description automatically generated



**Step 3: Send Email Action Configuration- Dynamic Content and Expressions**

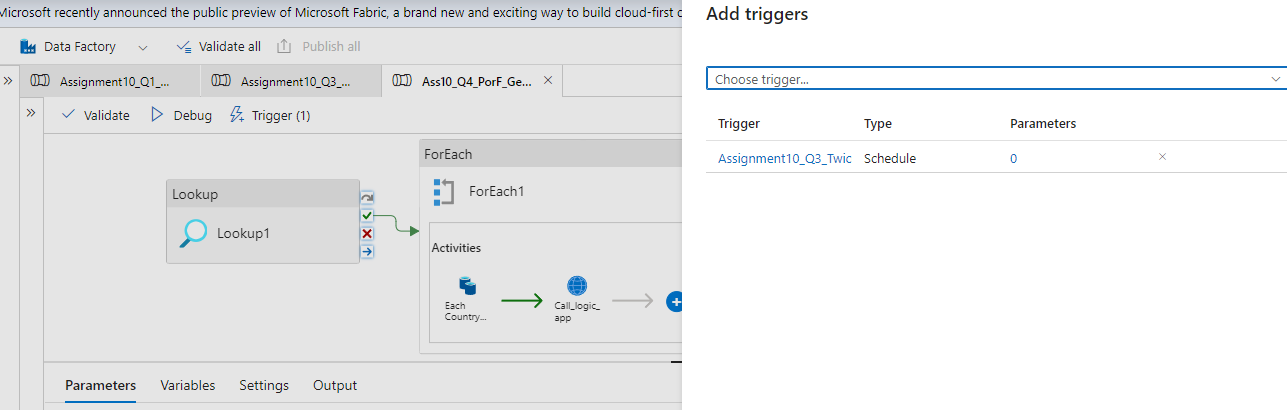
While configuring the Send Email action, I utilized dynamic content from the trigger body, ensuring the correct extraction of values like recipient name, country name, pipeline name, and pipeline status. Correct usage of dynamic content prevented any evaluation errors.

A screenshot of a email

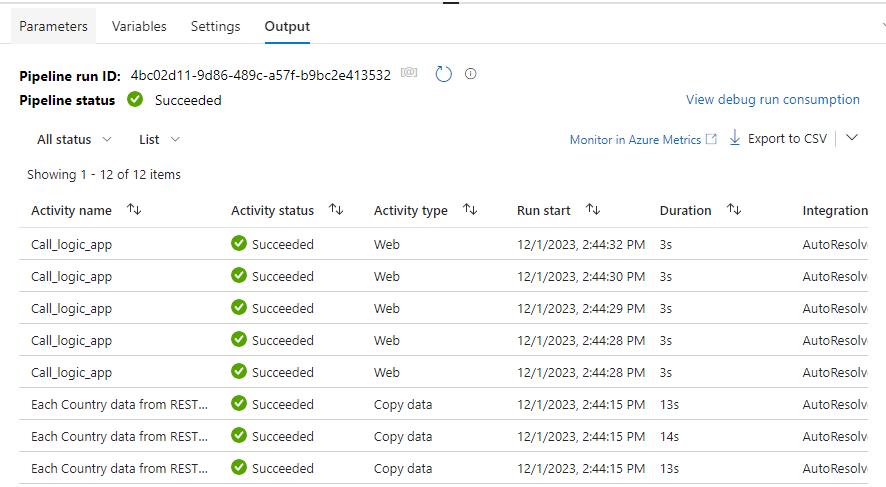
Description automatically generated

**Step 4: ADF Web Activity Body Configuration**

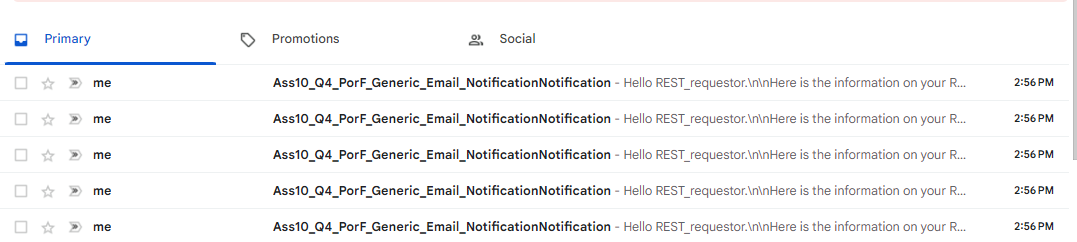
1. **ADF Web Activity Configuration**: In the ADF pipeline, the body field of the calling Web Activity was updated. The dynamic content in the body now includes the pipeline name, status, to address, country name, and a static value for the name field. Also, Add the run twice a day Trigger to this pipeline



**Step 7: Successful Execution**

1. **Successful Execution**: Upon successful execution, the Logic App sent a well-crafted email to the specified recipient, incorporating personalized details for a more informative notification experience. Here is a screenshot that shows the pipeline was executed successfully:

Here is a screenshot of the email notifications I received for each country data is copied



And here is the body of the email:

A screenshot of a computer

Description automatically generated

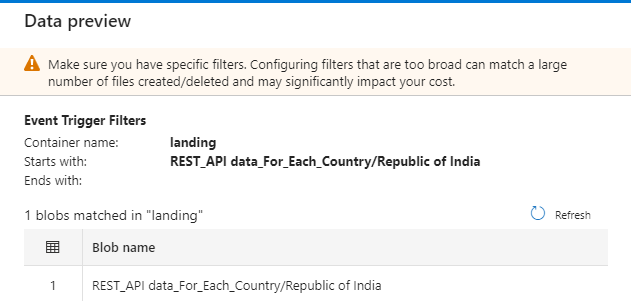
Question #6 Once a file for India gets generated trigger a pipeline that will copy the India.json file as CSV into the ADLS account.

Create Storage Event Trigger:

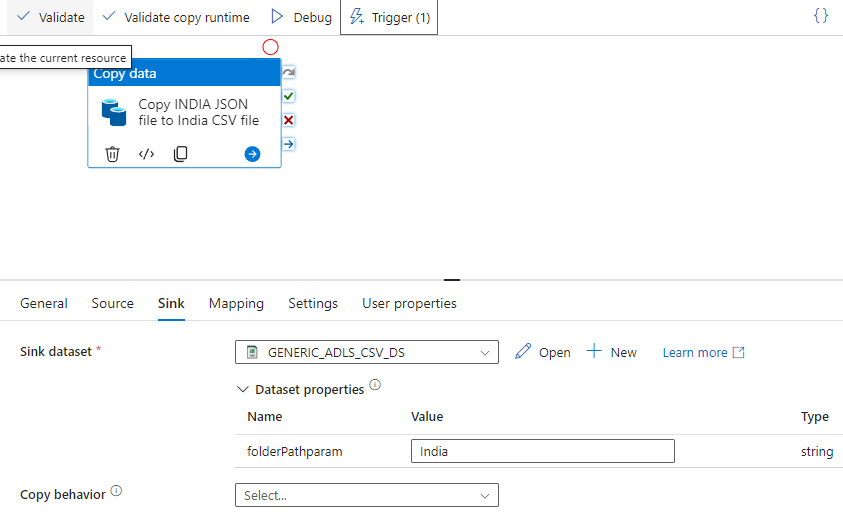
A screenshot of a computer

Description automatically generatedA screenshot of a computer

Description automatically generated



Step 2: Create a new pipeline that copies India. JSON files that get created in the main file into a CSV file in another ADLS location. Add the Storage Event Trigger to this pipeline so that when India.JSON file is created it will be automatically copied into CSV file



Step 3: Execute our Main RESTAPI by country pipeline , we can see the storage Event trigger is fired

A screenshot of a computer

Description automatically generated

And as a result, Assignment10\_Q6\_ Pipeline gets executed :

A screenshot of a computer

Description automatically generated

And India File is copies in our different ADLS location:A screenshot of a computer

Description automatically generated

A screenshot of a computer

Description automatically generated