ADF Transformation Project  
  
Business Case:

* User wants to convert the data based on user id into single row of products purchased and the total sales for the products for the respective ID.
* And store the data in a csv file

Flow Chart:



Steps:

1.Created 2 containers (Input and Output) in storage account.

Input container to land the input data from initial source.

Output container is to land the output data after transformation as per business case.

A screenshot of a computer

AI-generated content may be incorrect.

Data has been uploaded to the input container.

A screenshot of a computer

AI-generated content may be incorrect.

2. Go to Data Factory and create a data flow 🡪 click on ‘+’ button and select Add Source

A screenshot of a computer

AI-generated content may be incorrect.

3. To insert the dataset, Click New.

A screenshot of a computer

AI-generated content may be incorrect.

4. Select the Input folder path to move the data to the source as below:

A screenshot of a computer

AI-generated content may be incorrect.

A screenshot of a computer

AI-generated content may be incorrect.

Click on Schema tab and Import schema from Connection/Store

A screenshot of a computer

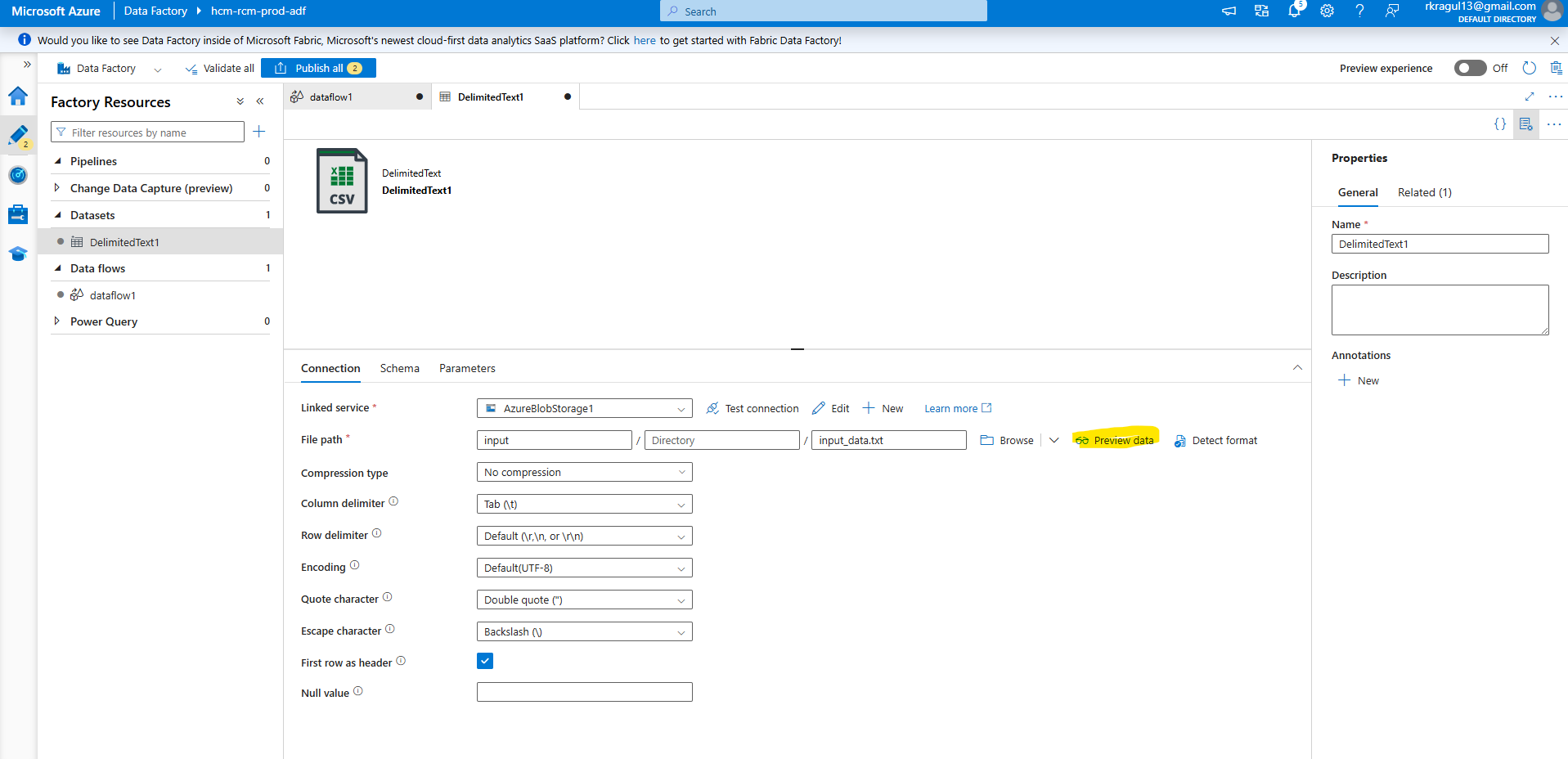
AI-generated content may be incorrect.

5. Click on Open to preview the data that was added to the Source tab.

A screenshot of a computer

AI-generated content may be incorrect.

Click on ‘Preview Data’



A screenshot of a computer

AI-generated content may be incorrect.

Click on ‘+’ icon and select Pivot

A screenshot of a computer

AI-generated content may be incorrect.

6.Select the below column to transform the data as per use case:

I, Group by ‘Id’

A screenshot of a computer

AI-generated content may be incorrect.

A screenshot of a computer

AI-generated content may be incorrect.

II. Add Product name as Pivot key

A screenshot of a computer

AI-generated content may be incorrect.

A screenshot of a computer

AI-generated content may be incorrect.

III, Using the Dataflow Expression Builder, calculate the sum of total sales as below:

A screenshot of a computer

AI-generated content may be incorrect.

A screenshot of a computer

AI-generated content may be incorrect.

7. Create Sink flow to load the transformed data, click on ‘+’ and select 'Sink’

A screenshot of a computer

AI-generated content may be incorrect.

8. Repeat steps for Source to create Sink flow,

A screenshot of a computer

AI-generated content may be incorrect.

Select the output path to extract the data from Sink

A screenshot of a computer

AI-generated content may be incorrect.

Go to Optimize tab Select ‘Single partition’ so the data will be downloaded in a single file.

A screenshot of a computer

AI-generated content may be incorrect.

9. Go to Pipeline and create a new pipeline and add the dataflow into that,

A screenshot of a computer

AI-generated content may be incorrect.

10.Validate all the steps and Publish it before adding trigger to the pipeline.

A screenshot of a computer

AI-generated content may be incorrect.

11. Once the trigger is activated we can use the, you can see the status of the pipeline in Manage 🡪 Pipeline runs.

A screenshot of a computer

AI-generated content may be incorrect.

12. The pipeline has ran successfully.

A screenshot of a computer

AI-generated content may be incorrect.

13. To view the final output in preview mode, select preview data.

A screenshot of a computer

AI-generated content may be incorrect.

A white grid with many squares

AI-generated content may be incorrect.

14. The Output container has stored the final data from Sink dataflow.

A screenshot of a computer

AI-generated content may be incorrect.

Output File:

Null values are placed with the blank space.

A screenshot of a computer

AI-generated content may be incorrect.