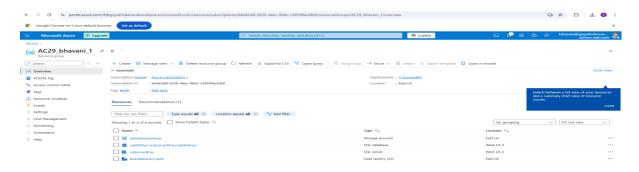
## Homework | Cycle 29 | Session 13

- Extracts data dynamically from a SQL database.
- Loads the extracted data into Azure Data Lake Storage Gen2 (ADLS Gen2) in Parquet format.
- Converts the Parquet files into CSV format within ADLS Gen2.
- Login into portal.azure.com and go to your resource group



#### Go to your Data factory and open



### Steps:

- In this activity we are bring data for multiple tables using single pipeline
- For this create 1<sup>st</sup> METADATA table name as TestTable in Azure Sqldb and insert data into created meta data table with detaqils of ID, Tablename, foldername, filename.
- Create a lookup activity in ADF which brings metadata table information details and displays into Json format file as output.
- In the settings tab uncheck the First row only check box and provide the source dataset details by clickig on New button.
- Select the azure sql db from dataset and click on continue in the **Set Properties** window provide the dataset name and linked service details.

- Select linked service and don't select the table name from the list as we are passing dynamic values and click on ok
- To create parameters i should go to parameters level so click on open button from dataset, it will open a tab of connection, schema and Parameters tabs.
- Go to Parameters session 1. Table Name, Schema Name, then go to connection tab as we have to pass those parameters whatever we created in Parameters tab.
- Click enter manually, click on Table name box, you will be able to see dynamic option and select schema and table name parameters as shown in snaps.
- Then go to your pipeline and provide the table name and schema name and click on debug.
- You will be able to see the input and output files after the pipeline runs completed successfully
- So, lookup will bring the metadata whatever we created azure sql database in the Json format.
- The Json output we will be passing in the similar way to next activity (for each)
- Step2: Take foreach activity and pass look up activity output to foreach input
- Click on lookup activity on **success button** as shown in picture and drag it to foreach activity input.
- Go to foreach settings tab and click on Items box and click on Add dynamic content there you will be able to see lookup activity outputs and select last option Look up value array
- Click on Edit button from inside foreach activity (pencil icon) as shown below
- And select your Copy activity inside foreach as below
- Go to source tab and select your dataset from source dataset dropdown
- Now select parameters from foreach (we will get parameters from foreach) and click on table name and click dynamic content and select foreach option
- As my main parameters from json file is TABLENAME and SCHEMANAME, i will add beside to item as .TABLENAME for TableName and .SCHEMANAME for Schema name and click on OK.
- Go to Sink tab, as we are loading data to Adls gen2, so will select the connector and select CSV format
- and click on edit to provide the parameters as foldername and filename then go to connection tab
- Provide @concat(dataset().FileName,'.parquet') as format
- Open copydata pipeline and give the values as parameters of foldername and filename as shown below
- Save the pipeline and publish it and Click on debug
- Files are placed to GEN2
- Next Convert the .parquet files into .csv files

```
CREATE TABLE TestTable

(ID INT,
TABLENAME VARCHAR(50),
SCHEMANAME VARCHAR(50),
FOLDERNAME VARCHAR(50),
FOLDERNAME VARCHAR(50)

FILENAME VARCHAR(50)

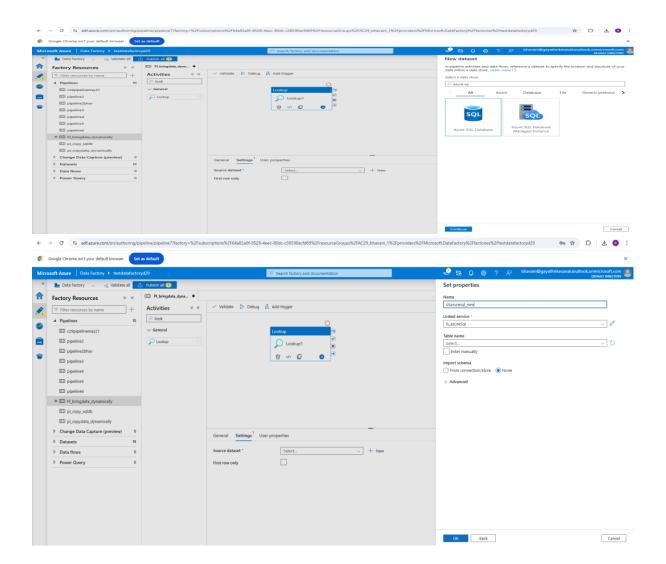
TINSERT INTO TestTable VALUES(1, 'ProductModel','saleslt','d29/saleslt/ProductModel','ProductModel')

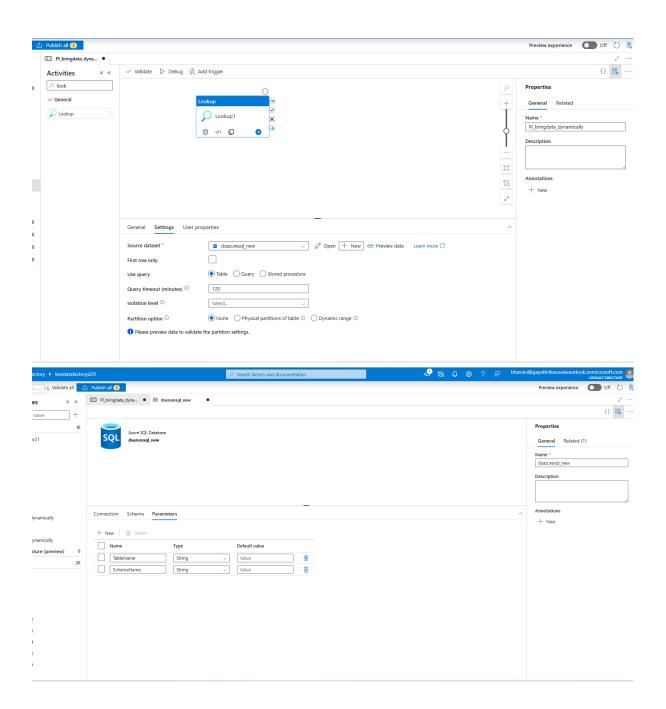
INSERT INTO TestTable VALUES(2, 'Customer','saleslt','d29/saleslt/Customer','Customer')

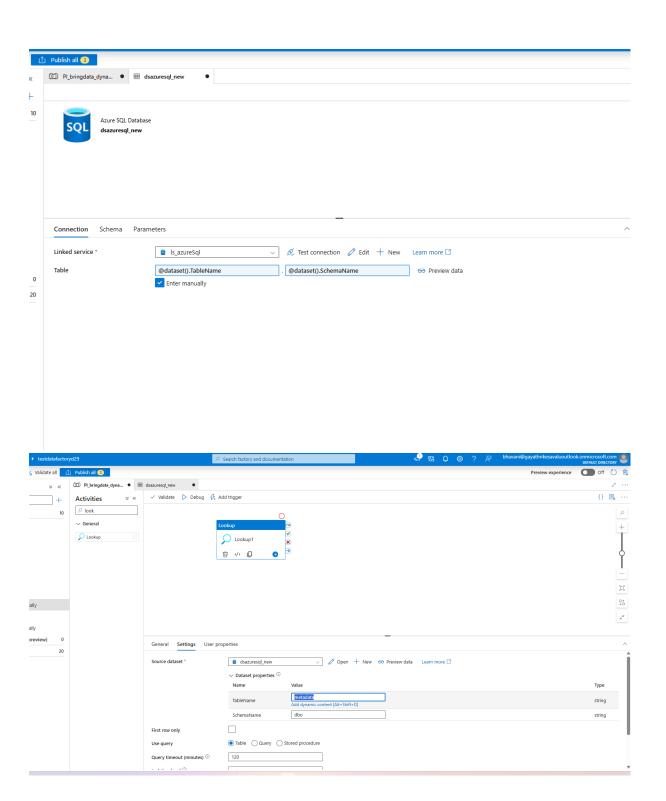
INSERT INTO TestTable VALUES(3, 'CustomerAddress','saleslt','d29/saleslt/CustomerAddress','CustomerAddress')

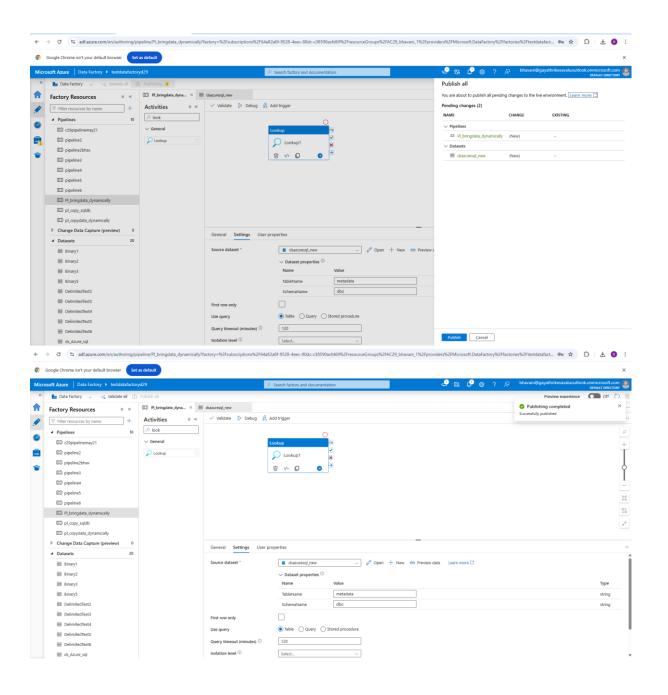
SELECT * FROM TestTable

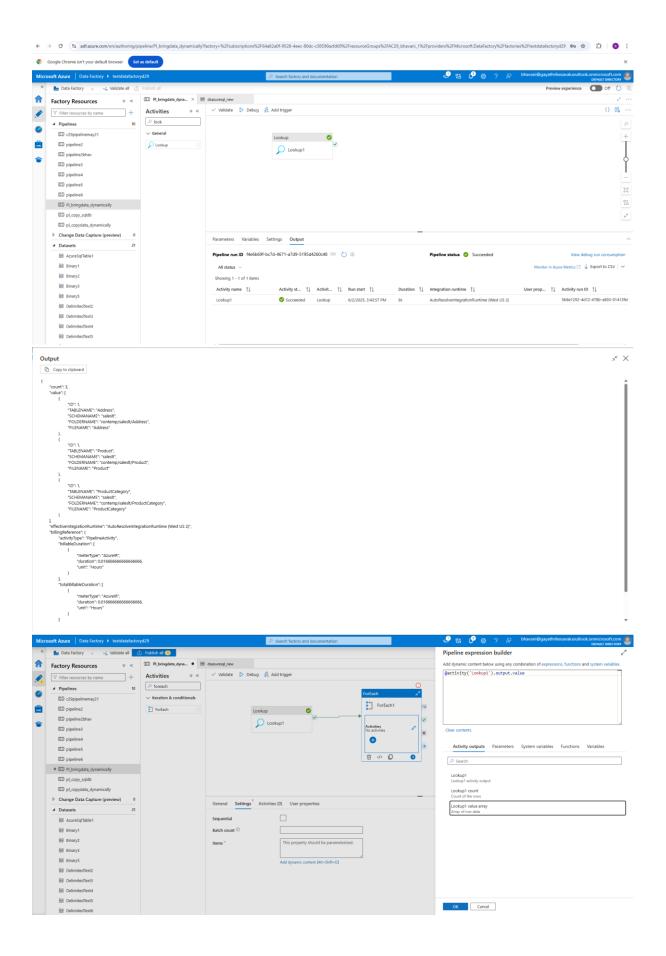
SELECT * FROM TestTable
```

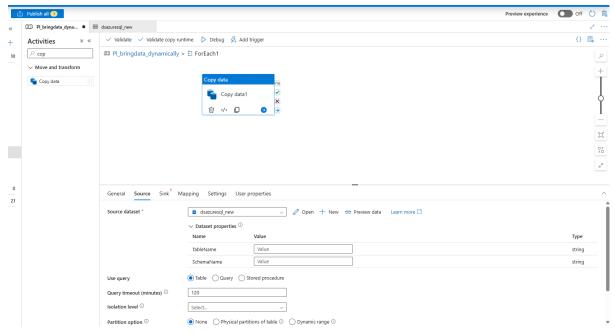




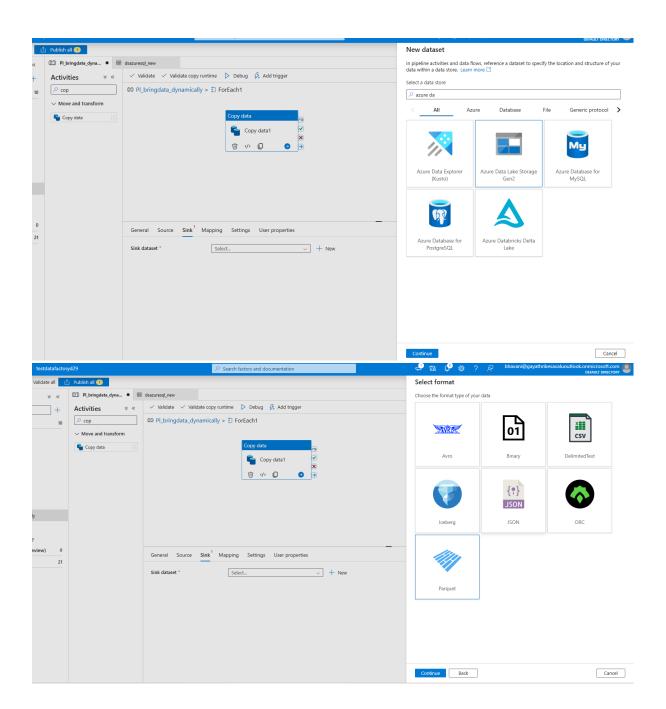




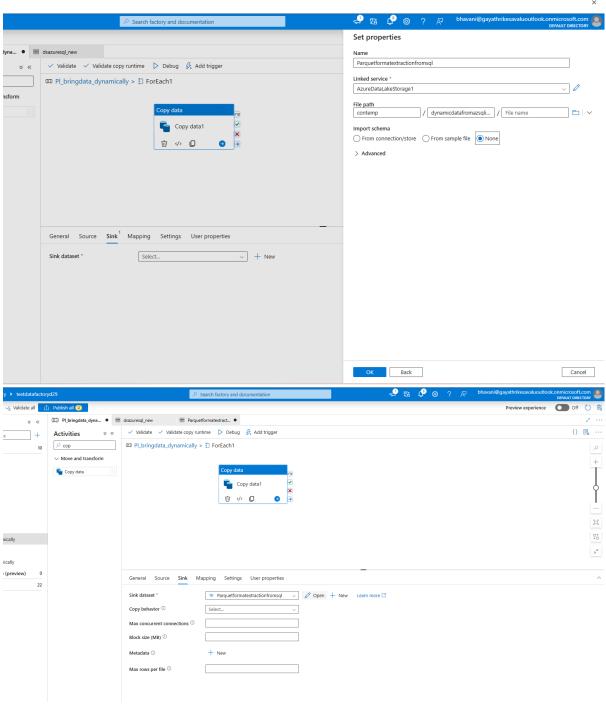


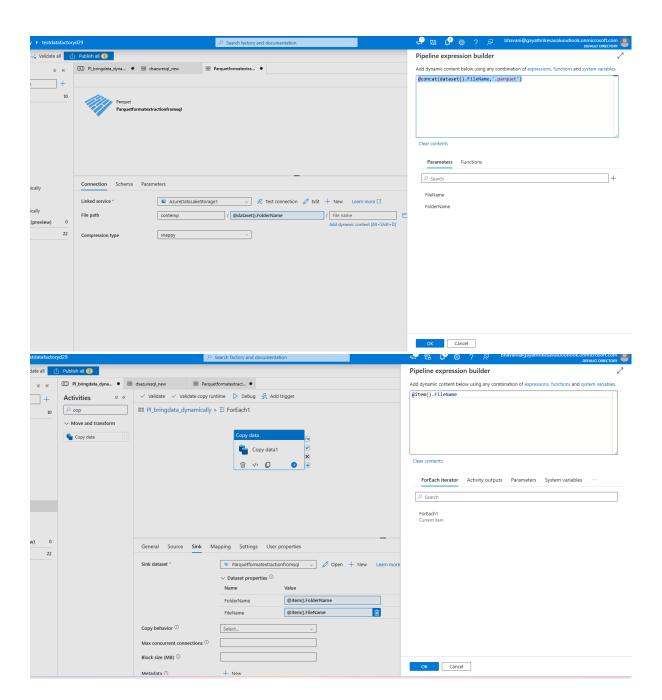


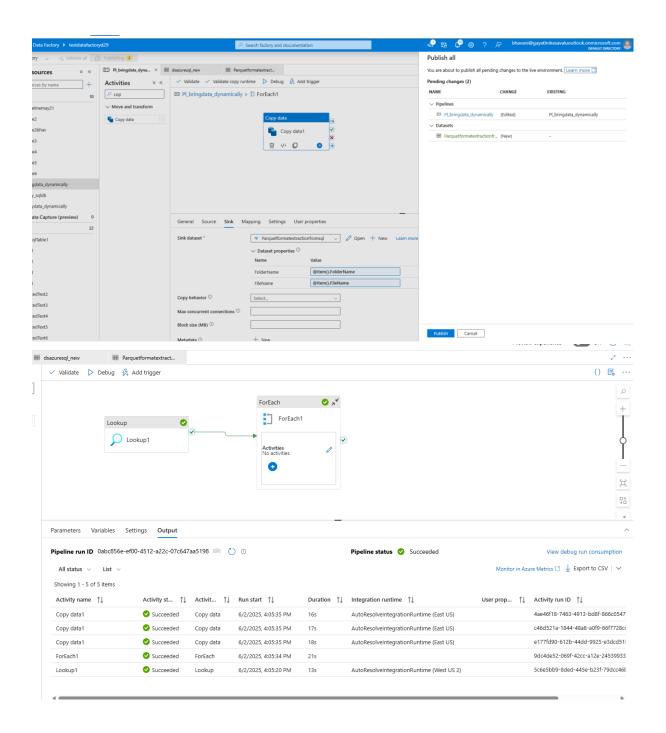
Storeadl2gen2 into parquet format







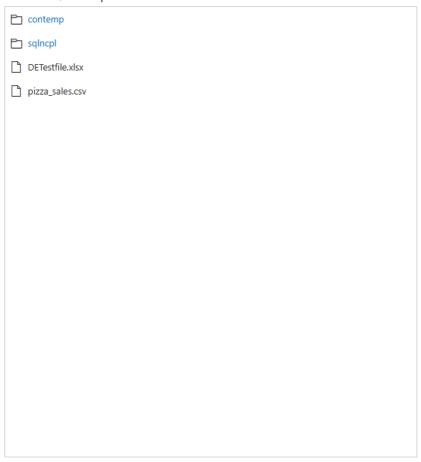




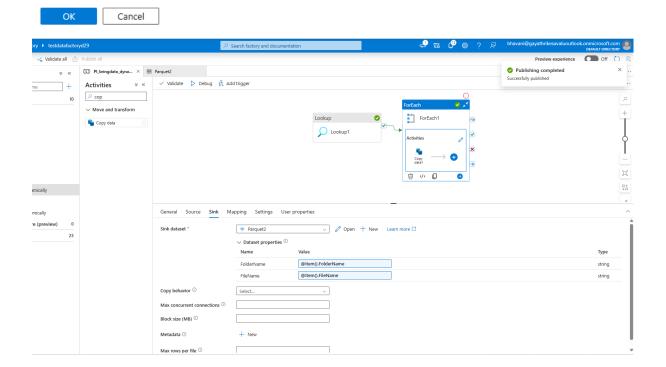
#### **Browse**

Select a file or folder.

## Root folder > contemp



Showing 1 - 4 of 4 items



# Input

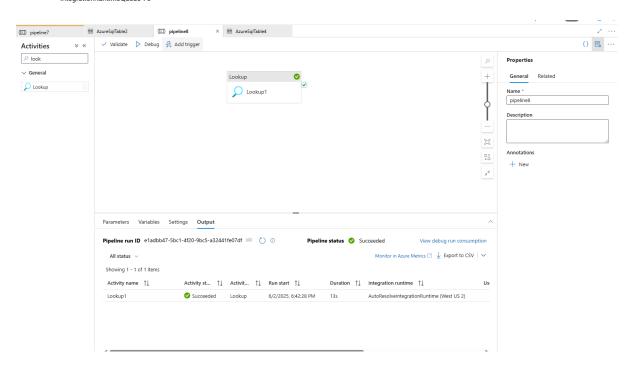
}

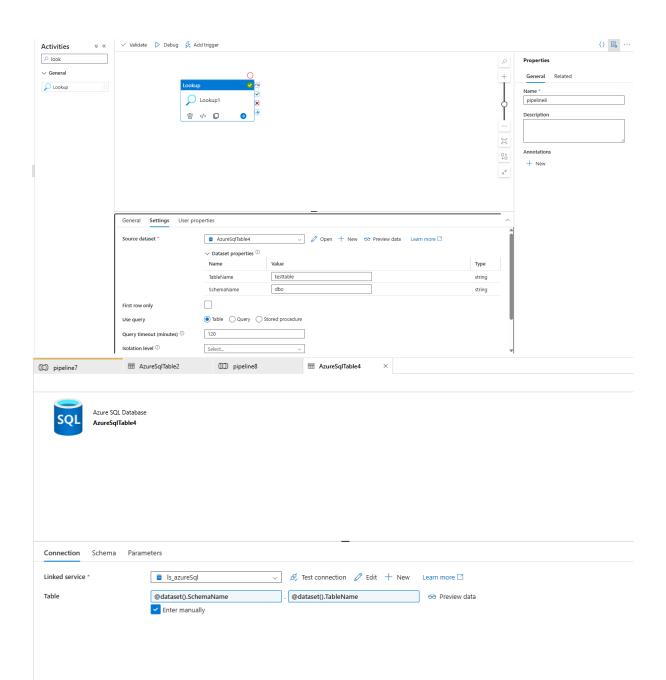
```
Copy to clipboard

{
    "source": {
        "type": "AzureSqlSource",
        "queryTimeout": "02:00:00",
        "partitionOption": "None"
},
    "dataset": {
        "referenceName": "AzureSqlTable4",
        "type": "DatasetReference",
        "parameters": {
            "TableName": "testtable",
            "SchemaName": "dbo"
        }
},
"firstRowOnly": false
```

#### Output

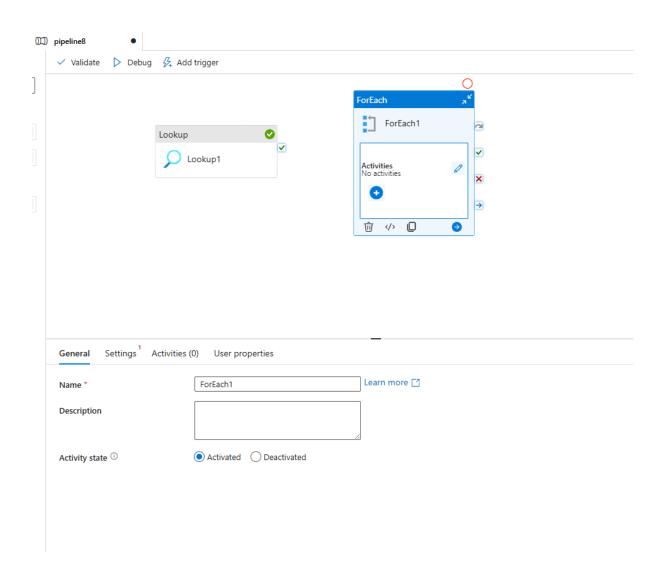
```
Copy to clipboard
   "count": 3,
   "value": [
             "ID": 1.
                "TABLENAME": "ProductModel",
               "SCHEMANAME": "salesIt",
"FOLDERNAME": "d29/salesIt/ProductModel",
                "FILENAME": "ProductModel"
                "ID": 2.
                "TABLENAME": "Customer",
               "SCHEMANAME": "salesit",
"FOLDERNAME": "d29/salesit/Customer",
"FILENAME": "Customer"
               "ID": 3,
"TABLENAME": "CustomerAddress",
               "SCHEMANIAME": "salesit",
"FOLDERNAME": "d29/salesit/CustomerAddress",
"FILENAME": "CustomerAddress"
   "effectiveIntegrationRuntime": "AutoResolveIntegrationRuntime (West US 2)",
   "billingReference": {
    "activityType": "PipelineActivity",
         "billableDuration": [
              {
                     "meterType": "AzureIR",
"duration": 0.0166666666666666666,
                      "unit": "Hours"
              }
         "totalBillableDuration": [
              {
                     "meterType": "AzureIR",
"duration": 0.016666666666666666,
"unit": "Hours"
        ]
   "durationInOueue": {
         "integrationRuntimeQueue": 0
```

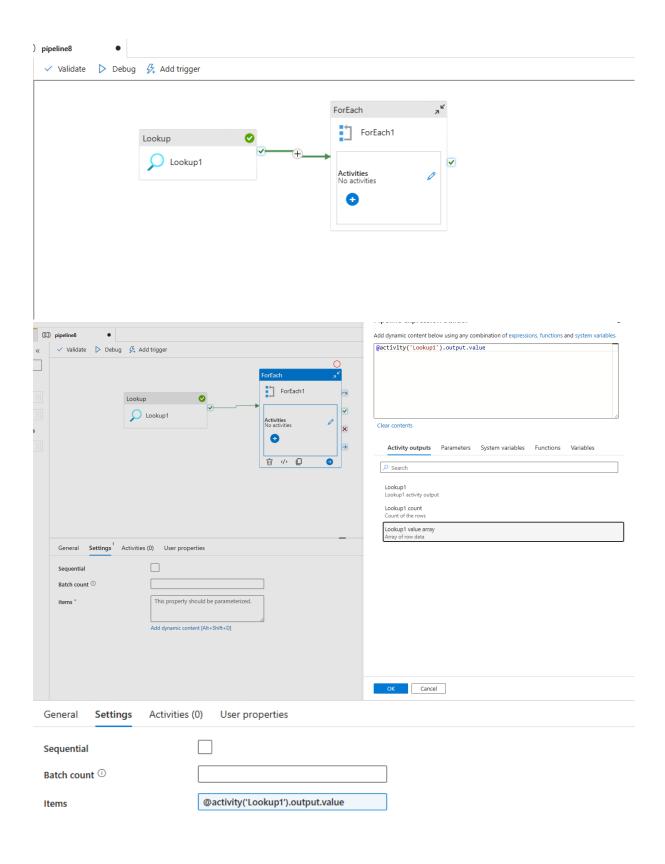


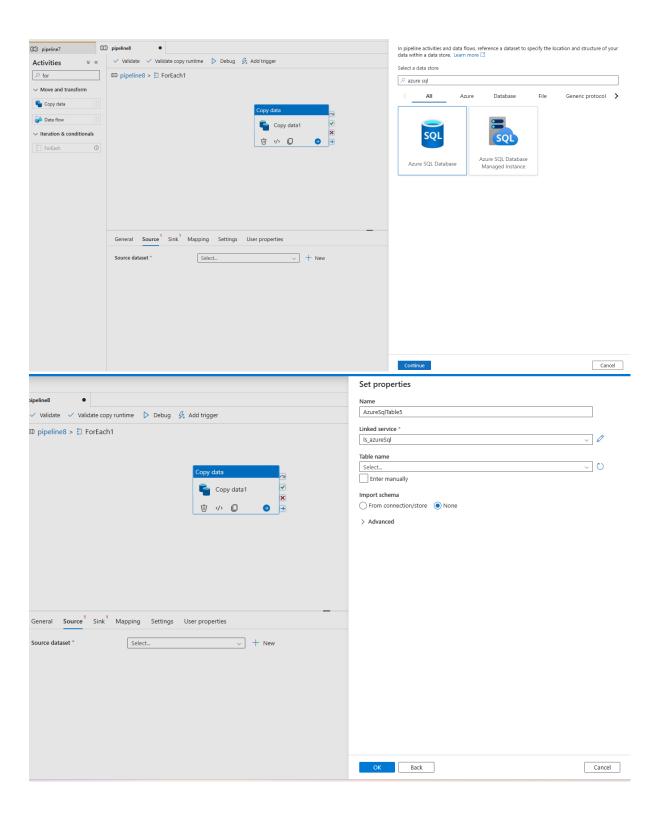


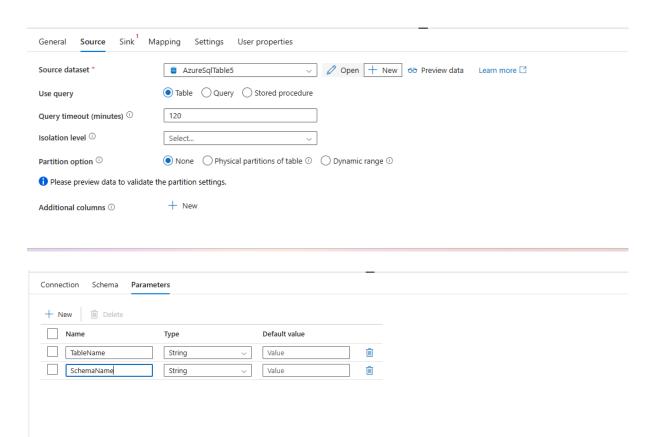


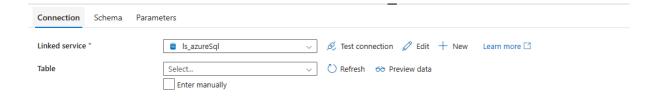


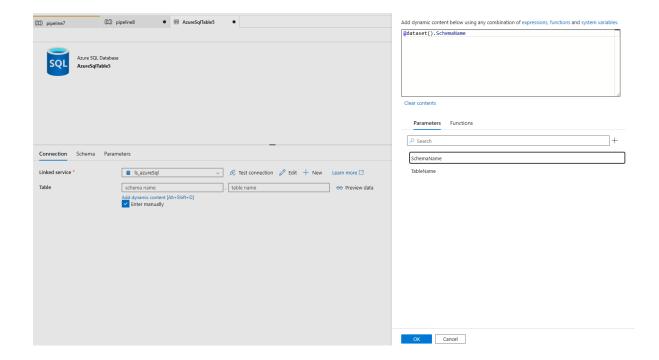


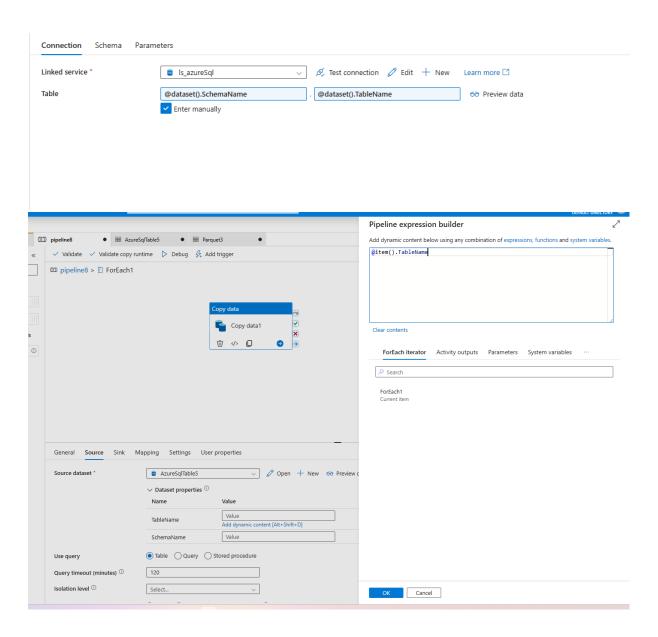


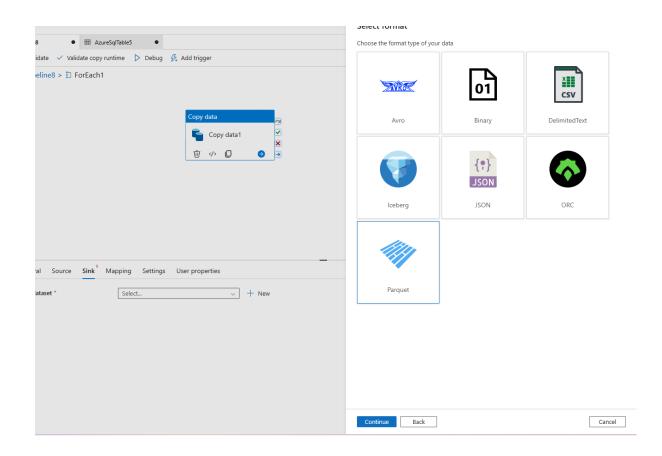


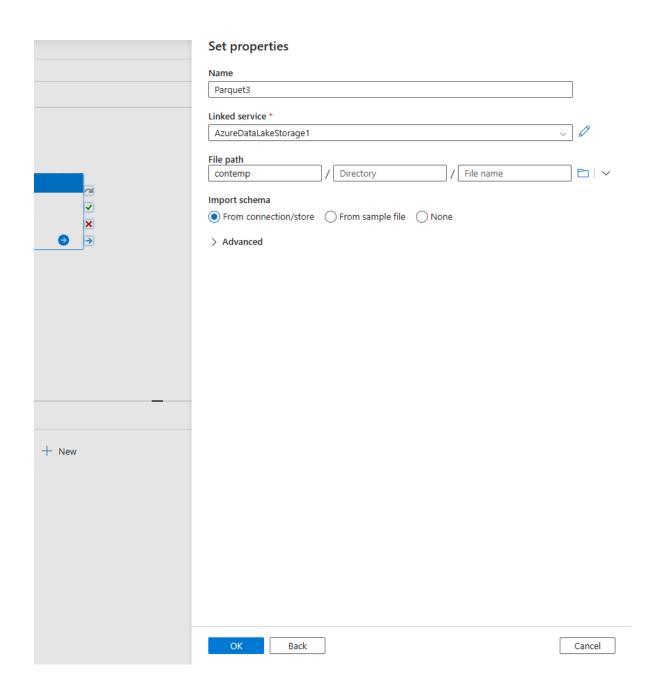


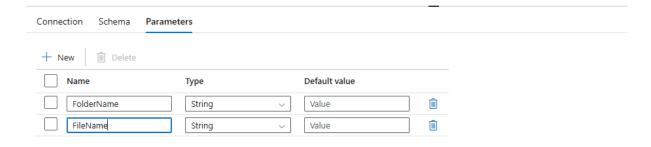


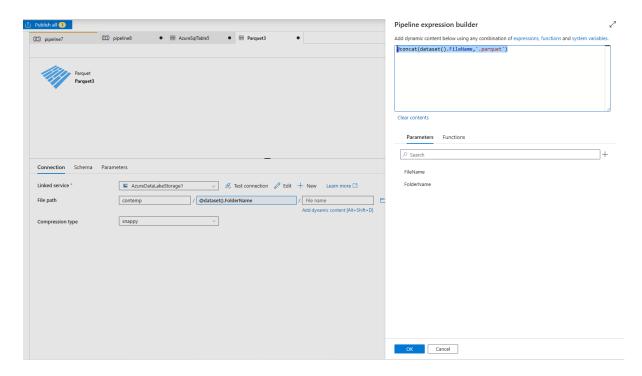


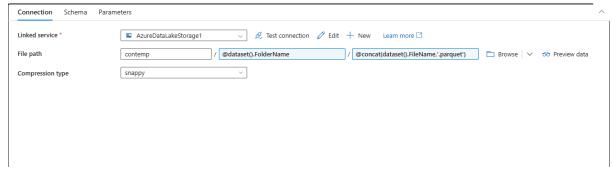


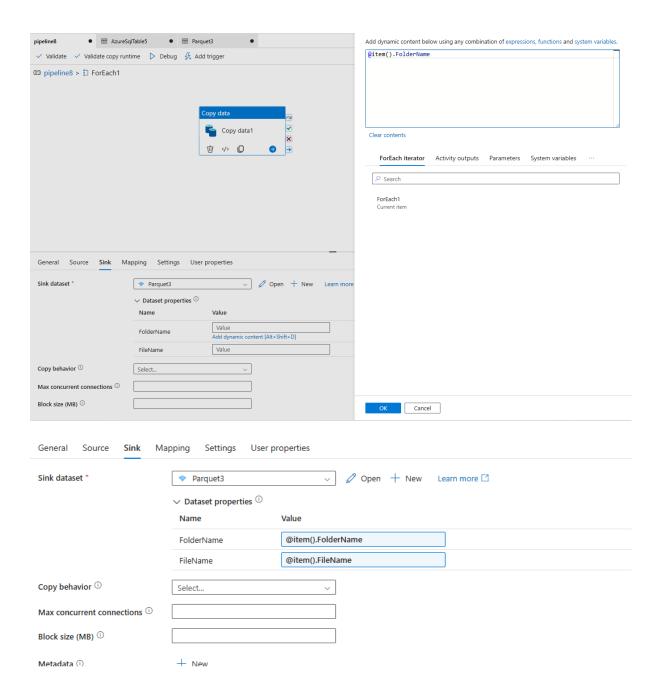


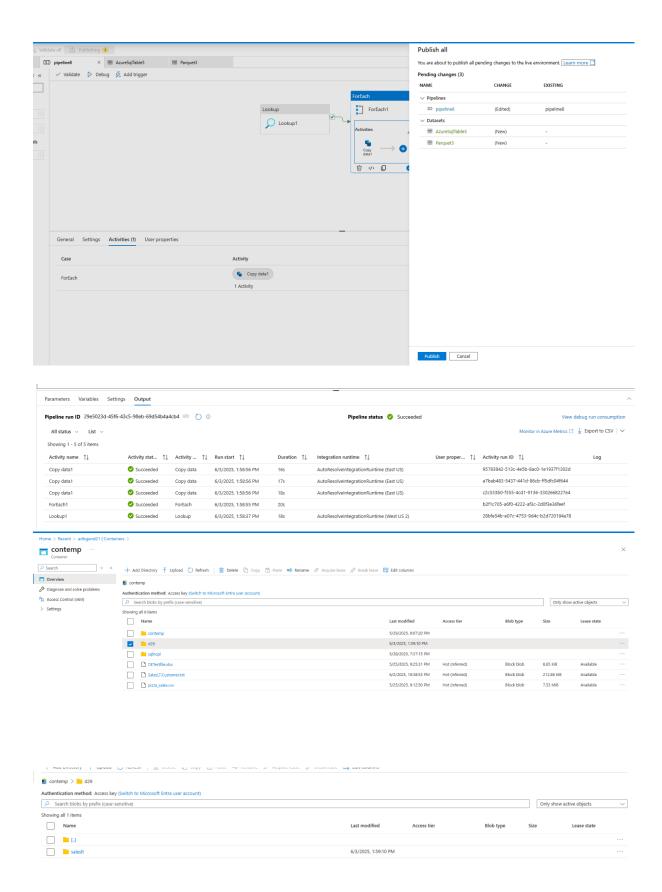


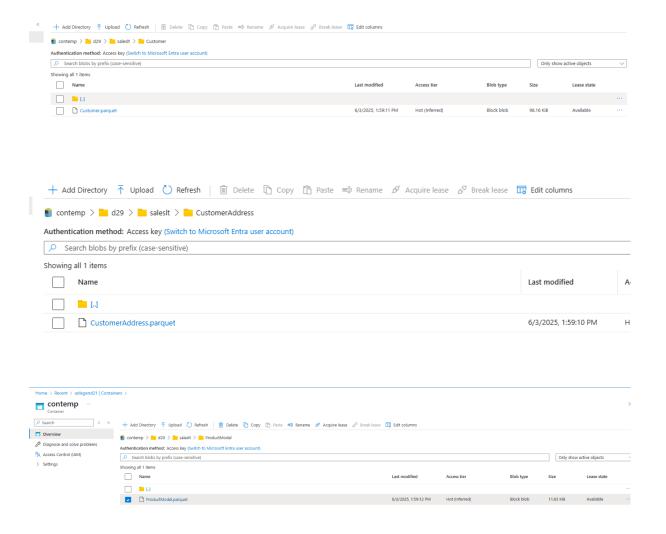




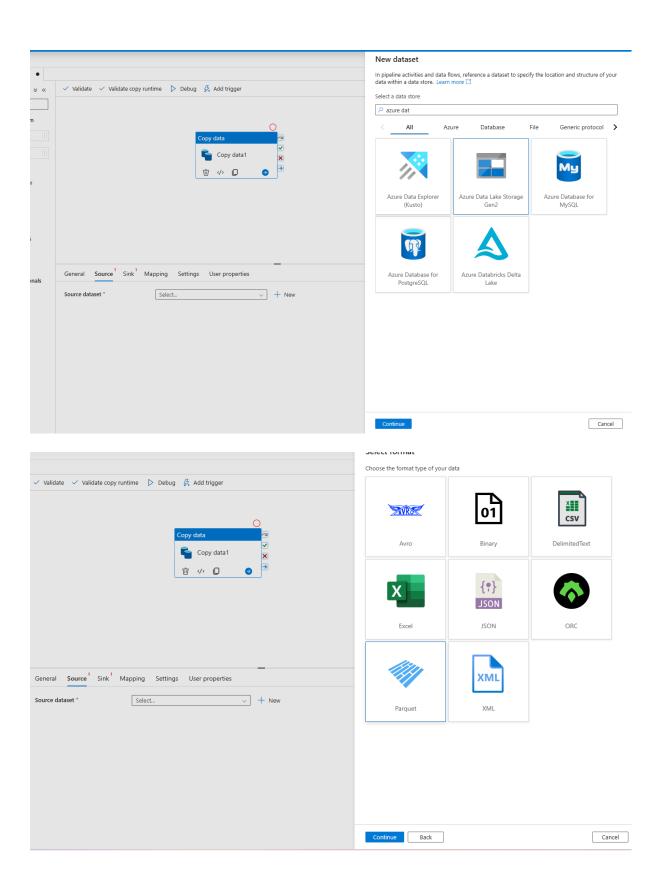


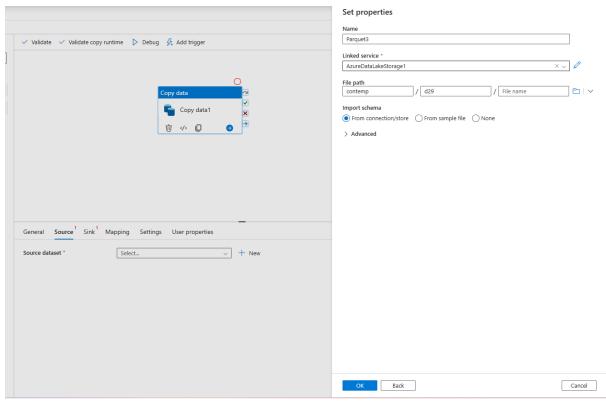


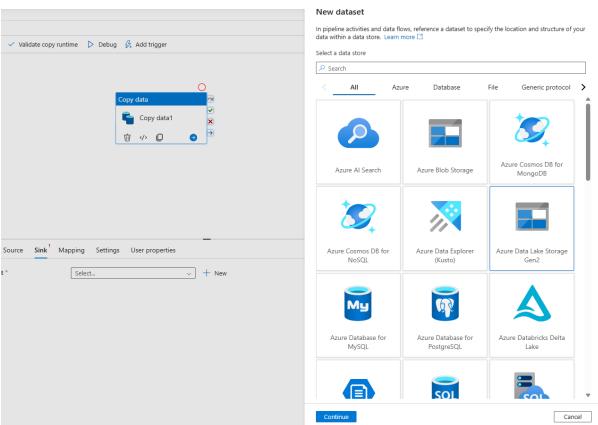


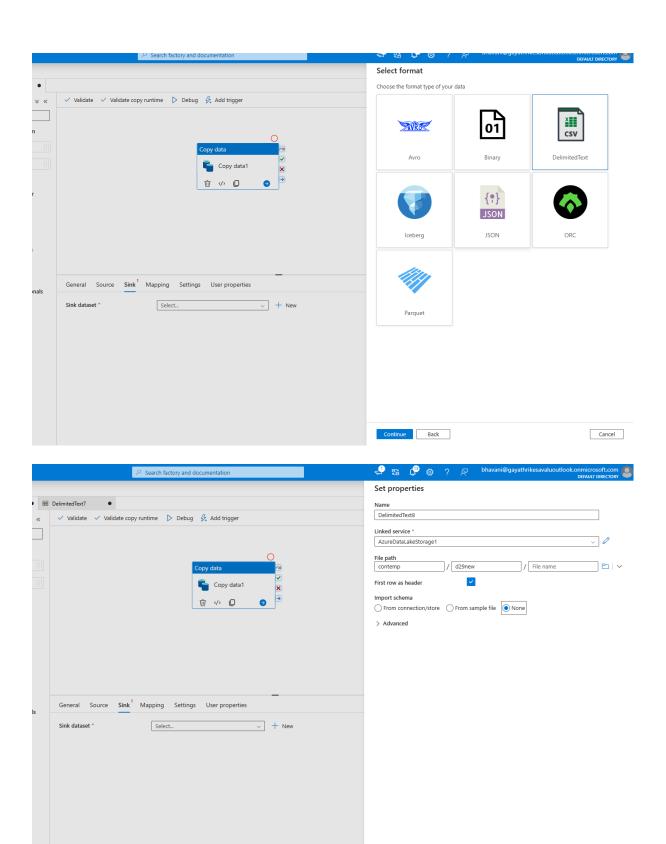


Converts the Parquet files into CSV format within ADLS Gen2.



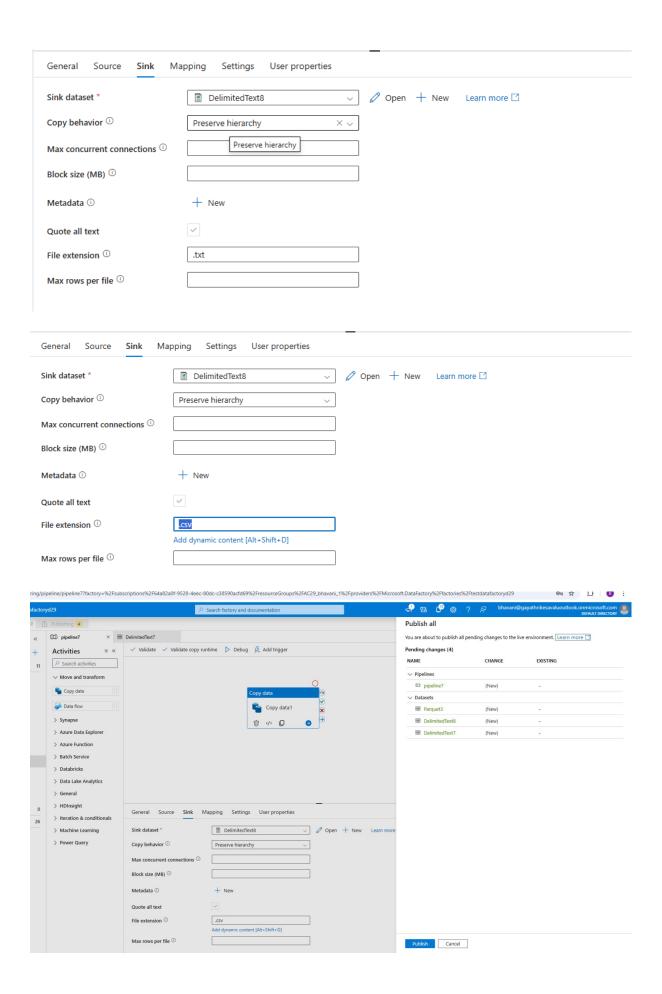


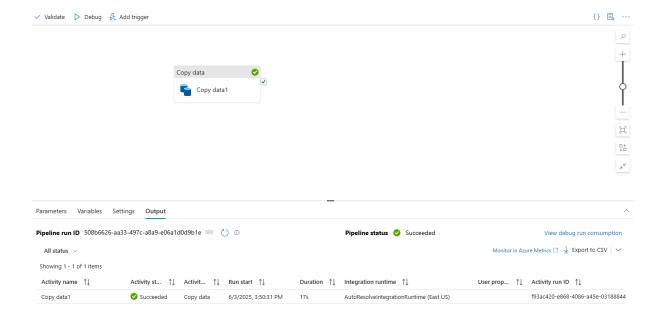




OK Back

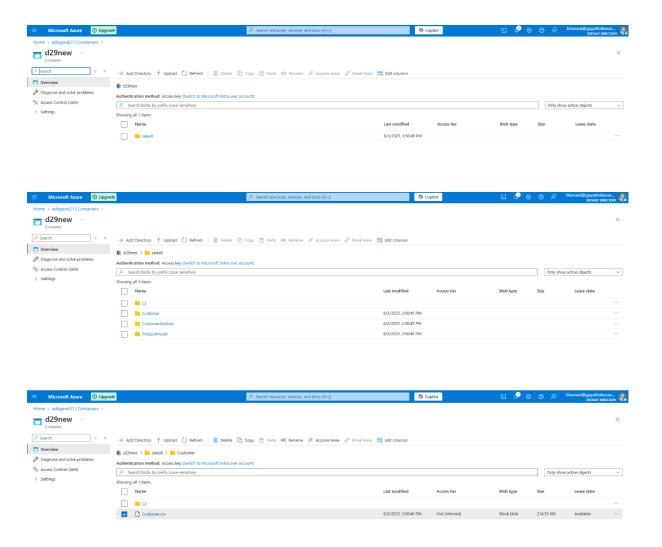
Cancel





## With Preserve Hierarchy





#### With Flatten Hierarchy

