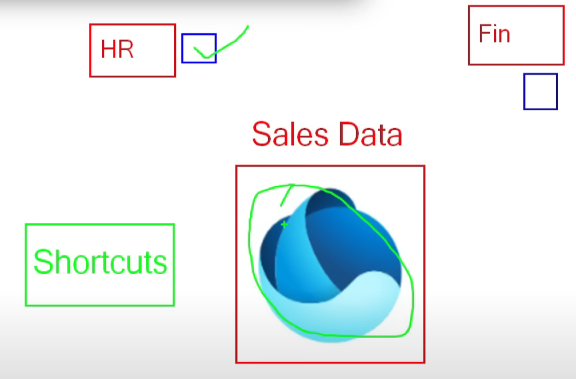
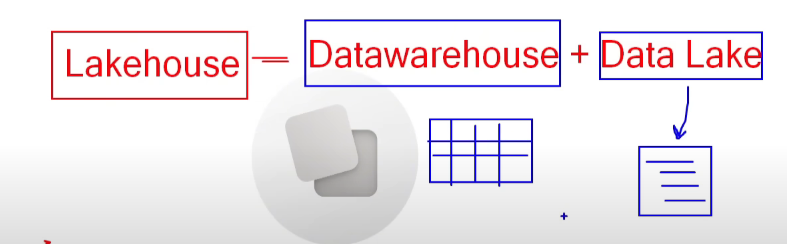
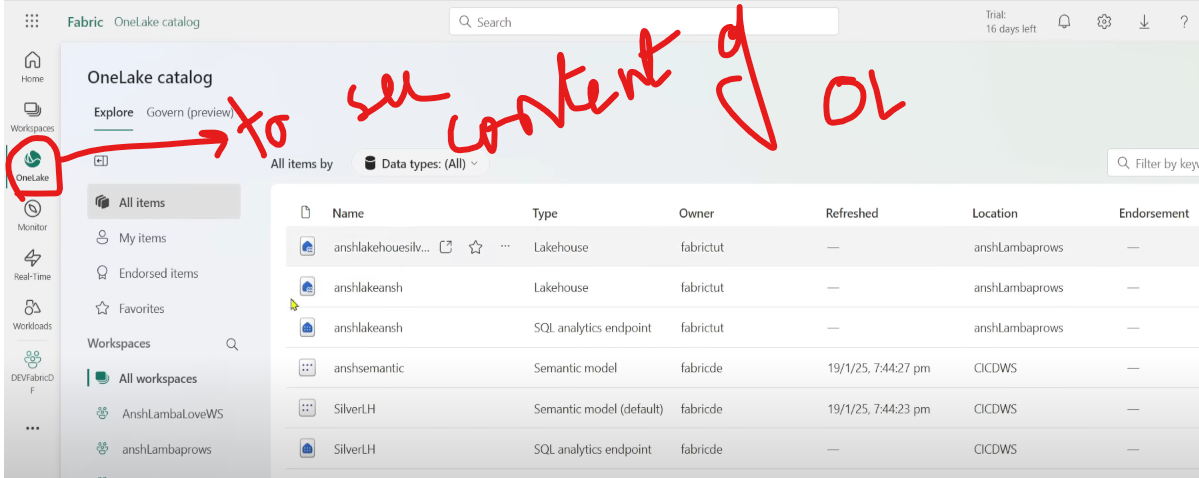
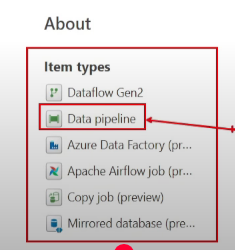
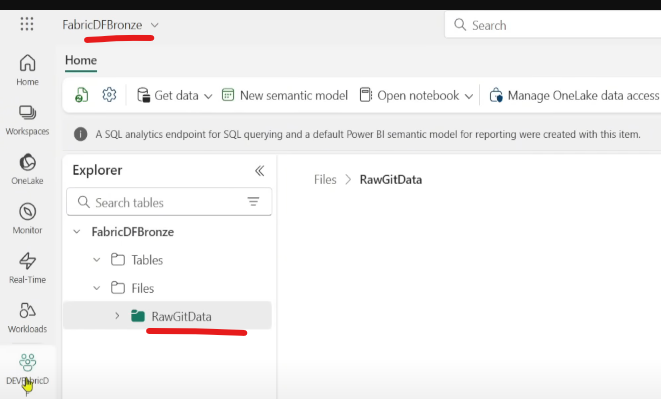
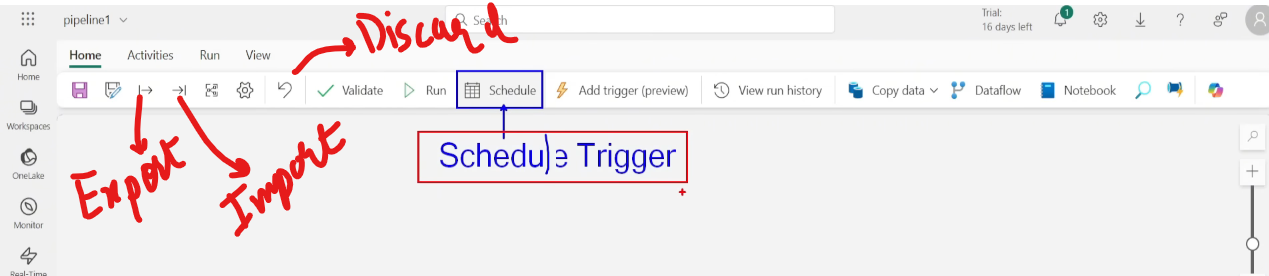
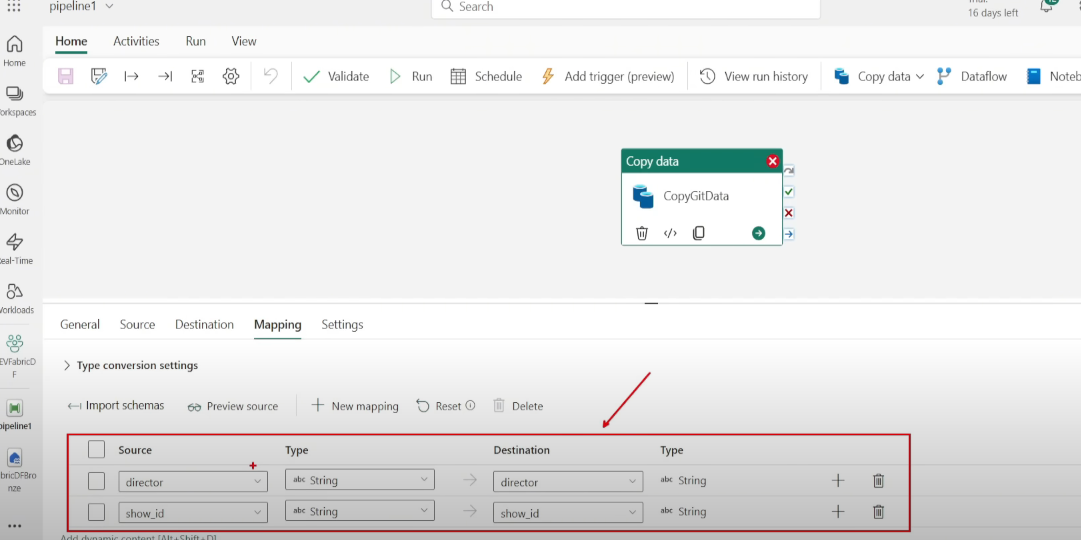
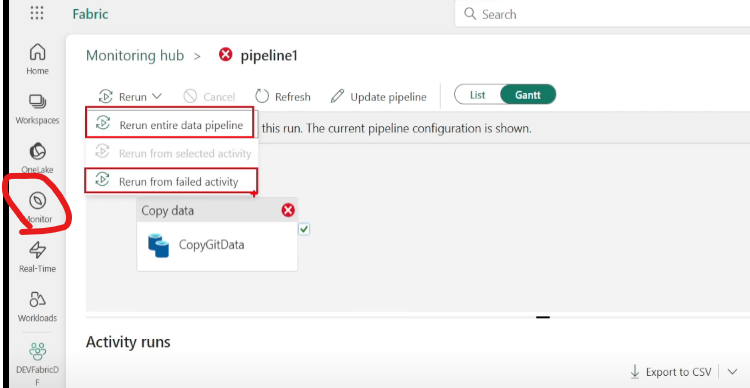
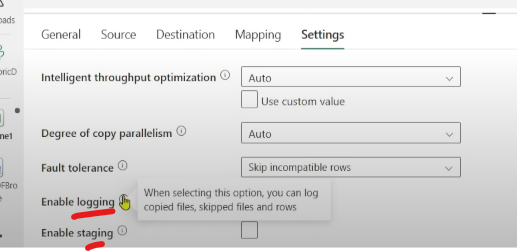
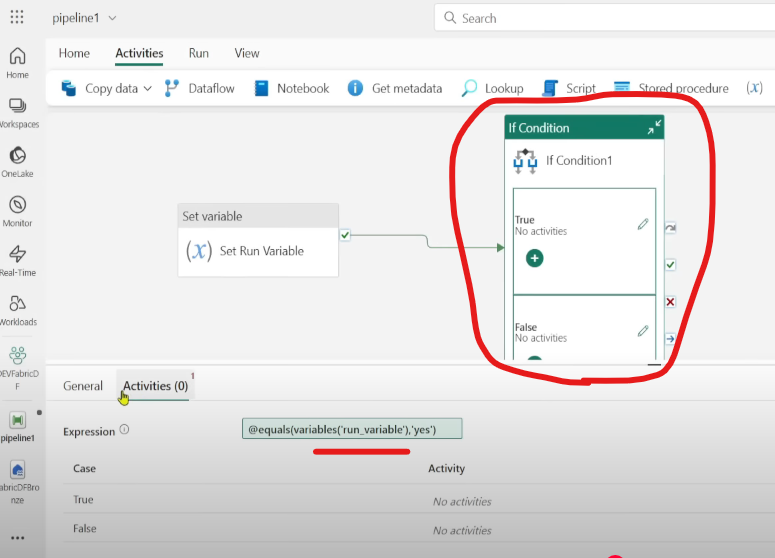
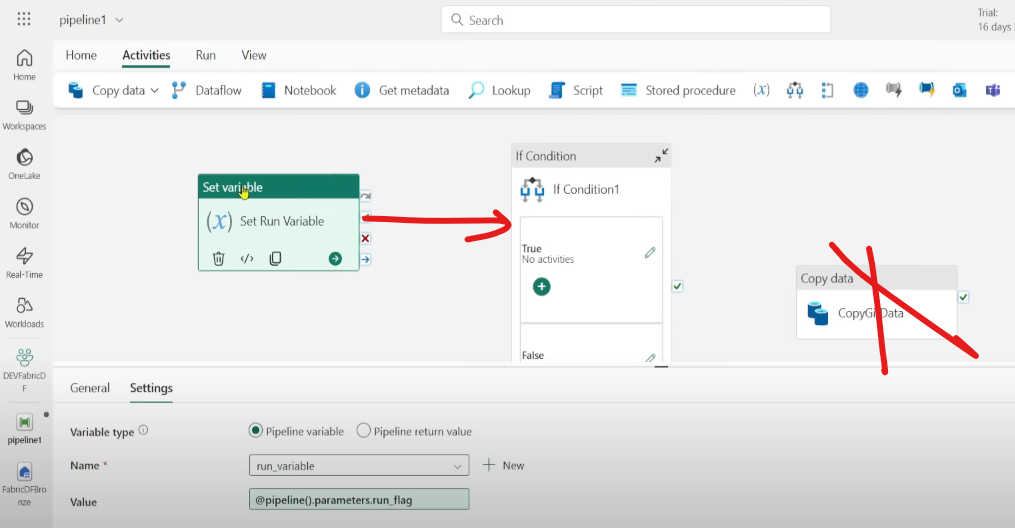
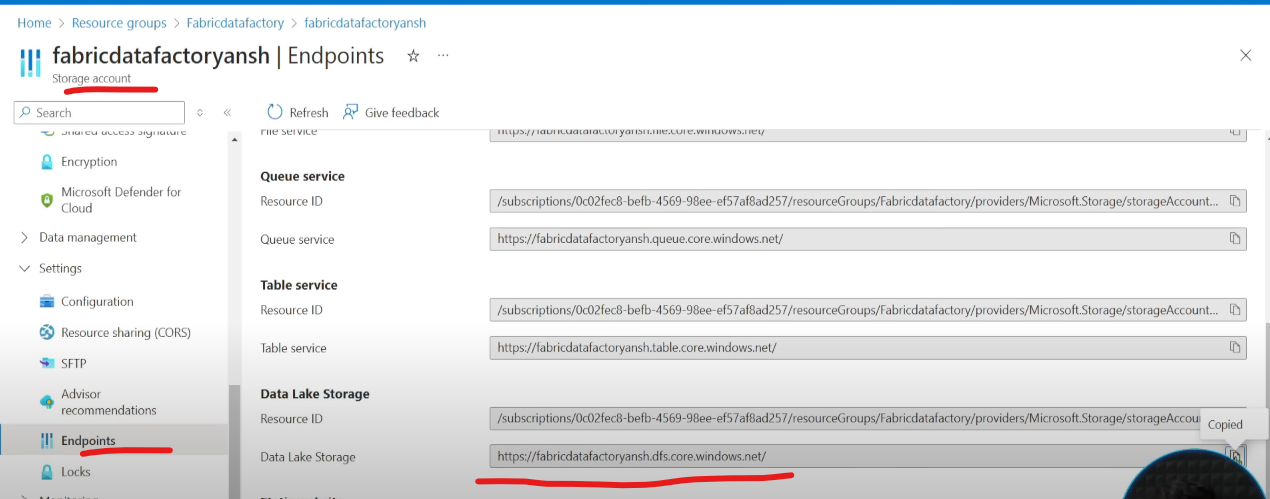
**FABRIC DATA FACTORY**  
MF is End to End Unified Data Solution (from data ingestion to deployment)  
Best part of MS is that all the services are connected/integrated(we don’t have to worry about connections)  
Fabric Data Factory 🡪 ETL/ELT solution of Fabric, helps in Data Movement, Transformation, Orchestration and Integration  
  
OneLake is cheaper than ADLS Gen2 since we have only one copy of data, no need to manage networking between storage accounts since its only one lake.  
  
Let’s say we have Sales data, now we have 2 workspaces HR and Finance, how the Sales data made available to both the workspaces? OL bts creates Shortcuts  
Simply put files in Data Lake, DWH is a logical DWH over here, it will create a metadata layer on top of files and it will behave like a table  
  
  
Go to app.fabric.microsoft.com  
Create a Workspace  
Create a Lakehouse  
Create a Data Pipeline  
A screenshot of a computer

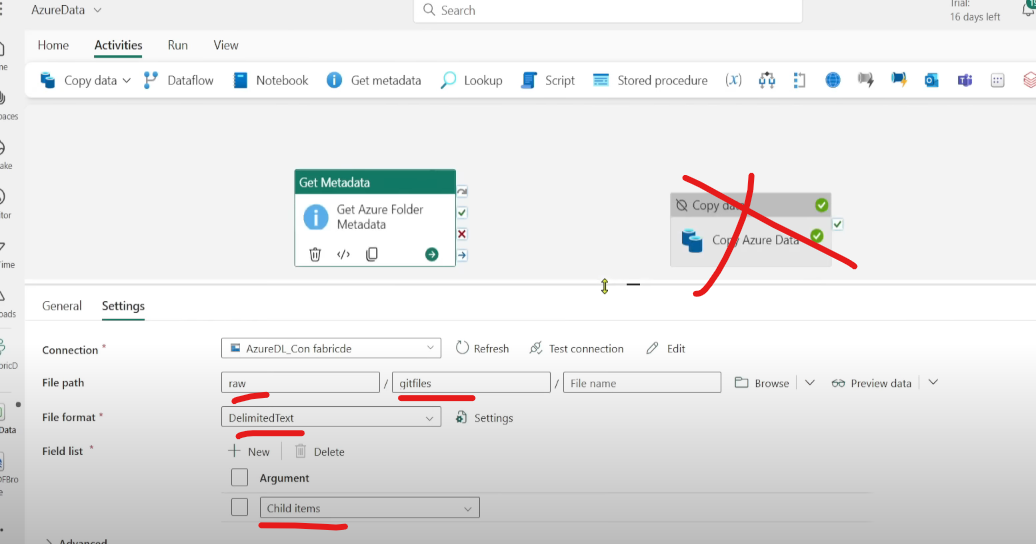
AI-generated content may be incorrect.  
Get data from GitHub and store in LHMapping is required for Schema EnforcementA screenshot of a computer

AI-generated content may be incorrect.  
  
Problem statement1: Copy data from GitHub only if the pipeline parameter is yes  
Create a pipeline parameter & a Variable  
A screenshot of a computer program

AI-generated content may be incorrect.  
A screenshot of a computer

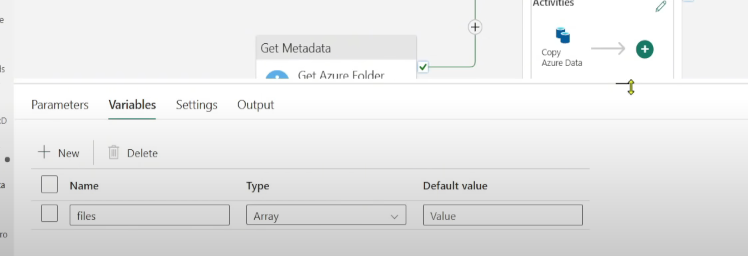
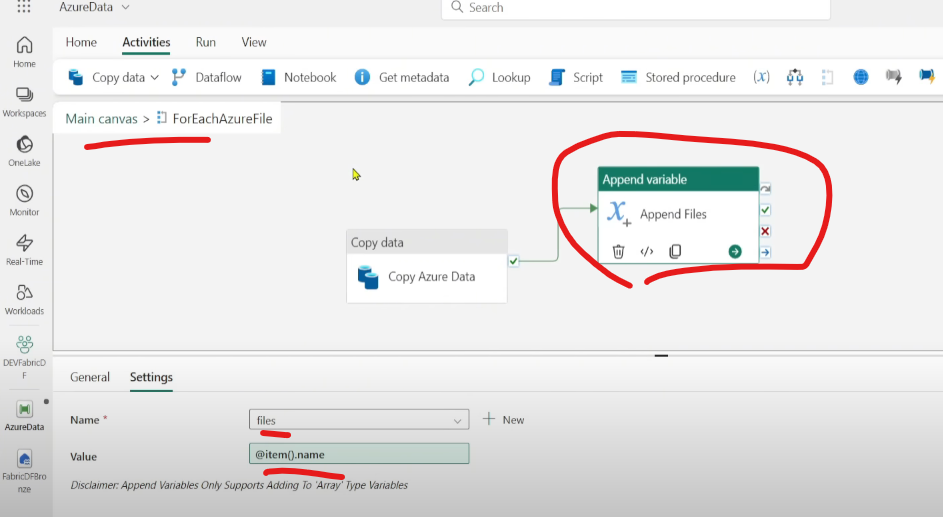
AI-generated content may be incorrect.A screenshot of a computer

AI-generated content may be incorrect.  
  
Note : We cant use For-Each inside If condition in Fabric (in ADF its supported)  
  
Problem statement2: Get a list of files from ADLS and keep it in Fabric LH  
Give Storage Blob Data Contributor access of ADLS to Fabric  


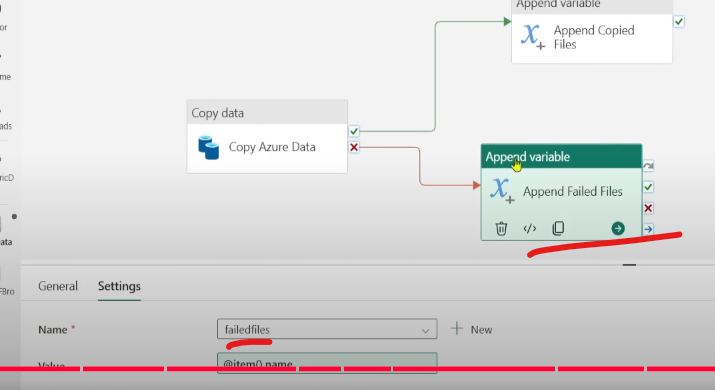
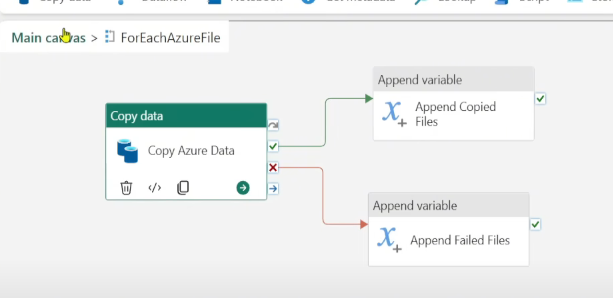
Use a Get Metadata activity to get all the File names   
 A screenshot of a computer

AI-generated content may be incorrect.  A screenshot of a computer

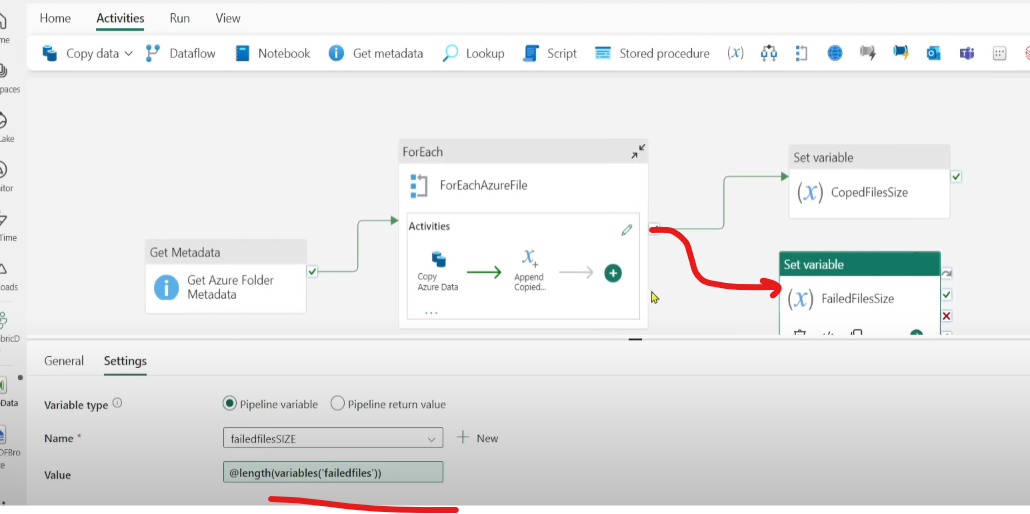
AI-generated content may be incorrect.  
Copy activity 🡪 Source 🡪 connection for ADLS Gen2  
 A screenshot of a computer

AI-generated content may be incorrect.  
Lets enhance the pipeline by adding logging mechanism  
Create a variable, everytime the files are getting copied, the information needs to be stored in this variable  
  
Add a Append variable activity after Copy to hold all the names   
  
Lets enhance it more, need to track both successful and failed files  
create 2 variables A screenshot of a computer

AI-generated content may be incorrect.  
A screenshot of a computer

AI-generated content may be incorrect.    
the data of 2 arrays inside the loop also needs to be stored, create 2 variables for that A screenshot of a computer

AI-generated content may be incorrect. A screenshot of a computer

AI-generated content may be incorrect. Now add Notifications for these activities A screenshot of a computer

AI-generated content may be incorrect. A screenshot of a computer

AI-generated content may be incorrect.  
Just a change, link both the Set Variable activities with On completion with For\_each (not with success)  
Delete the last activity and run the pipeline  
 A screenshot of a computer

AI-generated content may be incorrect. Because of bad data  
A screenshot of a computer

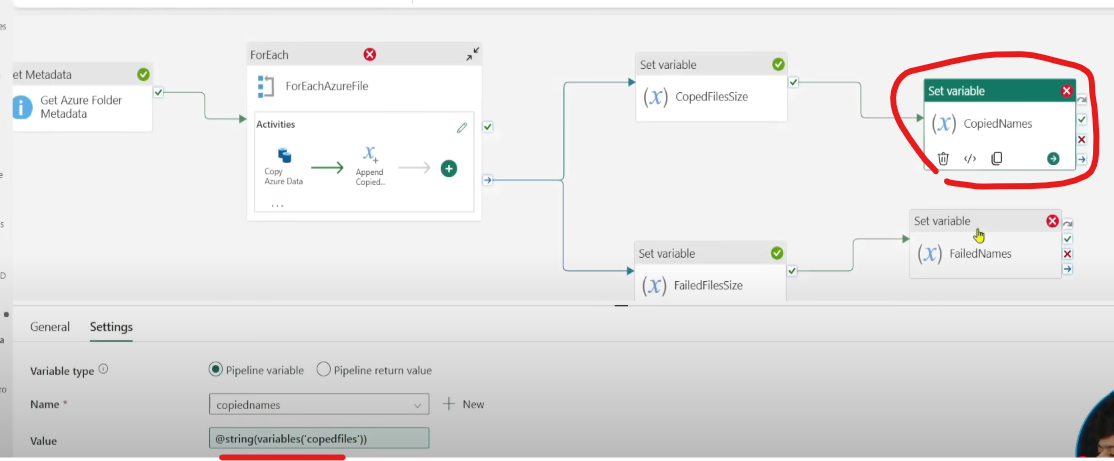
AI-generated content may be incorrect. A screenshot of a computer

AI-generated content may be incorrect.  
I also want to store the details of failed names, create 2 more variablesA screenshot of a computer

AI-generated content may be incorrect. A screenshot of a computer

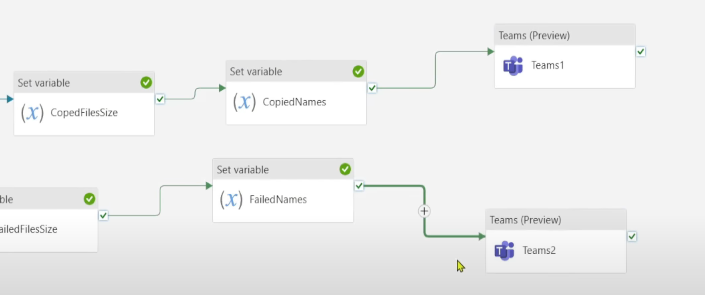
AI-generated content may be incorrect. A screenshot of a computer

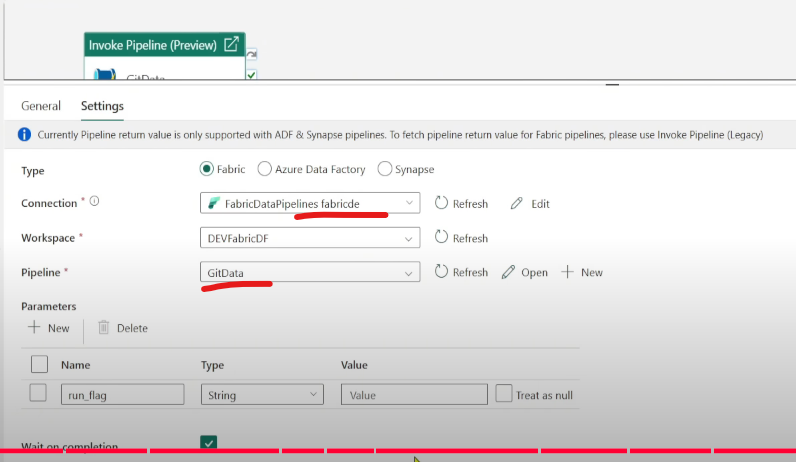
AI-generated content may be incorrect. A screenshot of a computer

AI-generated content may be incorrect.   
add string to hold array variable  
A screenshot of a computer

AI-generated content may be incorrect. A screenshot of a computer

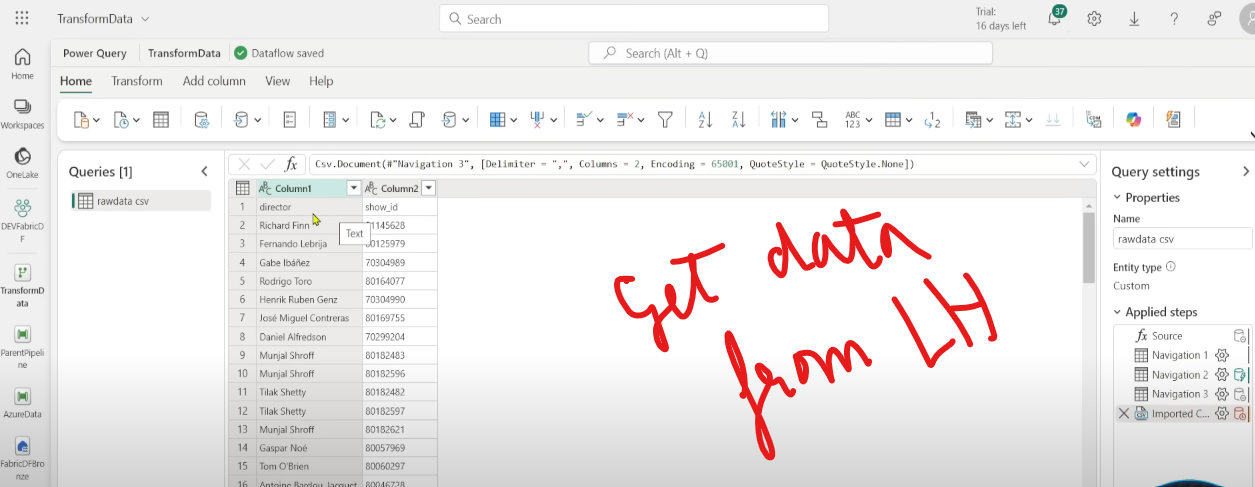
AI-generated content may be incorrect. A screenshot of a computer

AI-generated content may be incorrect.  
Lets add Teams notification activity at the end  


Create a parent pipeline 🡪 use Invoke pipeline activity   
create a pipeline parameter and feed it to child pipeline A screenshot of a computer

AI-generated content may be incorrect. A screenshot of a computer

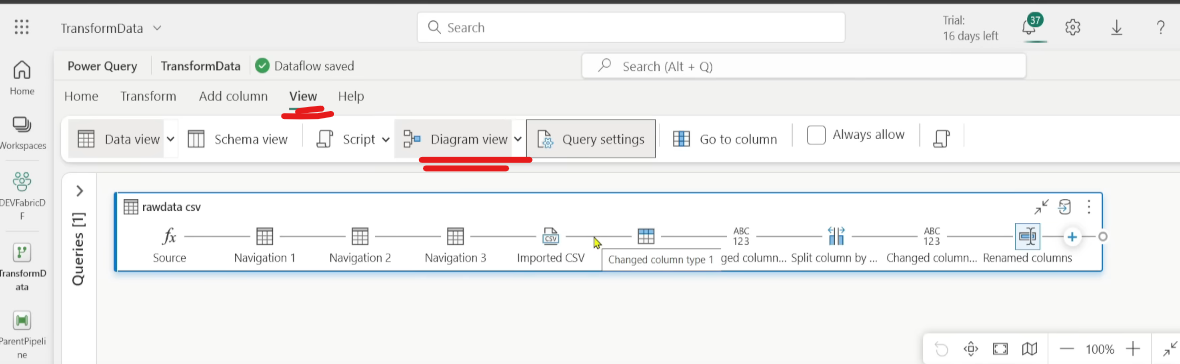
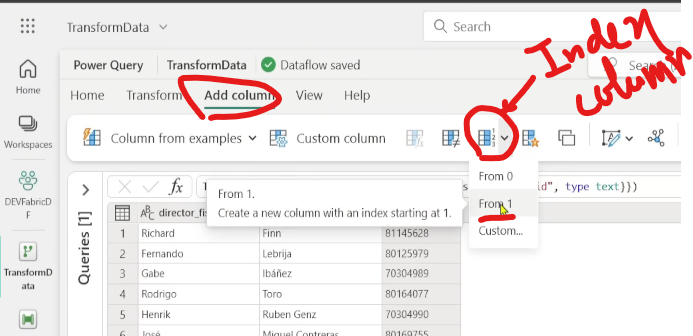
AI-generated content may be incorrect.

**Transformation**  
Data Flow Gen1 (backdated, use Gen2)  
  
Create Data Flow Gen2 (Power Query online integrated with Fabric)🡪 Low code/no code manner  
You can perform ETL in DFG2 but not ELT, but in Data Pipelines you can do ELT.  
  
 A screenshot of a computer

AI-generated content may be incorrect. Wanna split the directors name  
A screenshot of a computer

AI-generated content may be incorrect.  
A screenshot of a computer

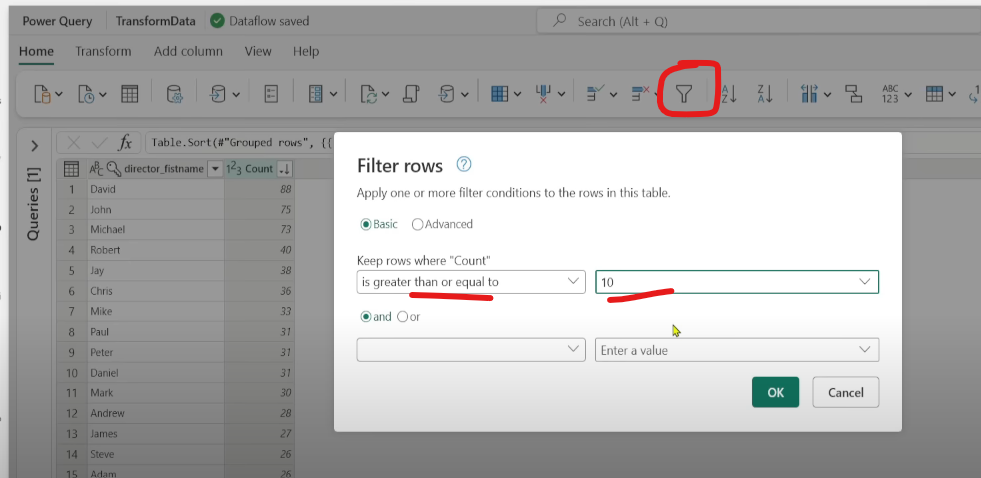
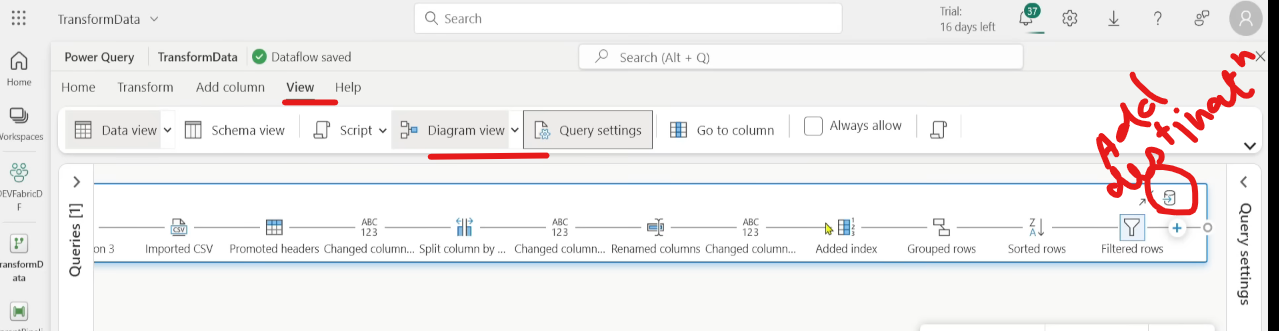
AI-generated content may be incorrect. A screenshot of a computer

AI-generated content may be incorrect.  
Rename the headers   
Add a new column containing row number   
A screenshot of a computer

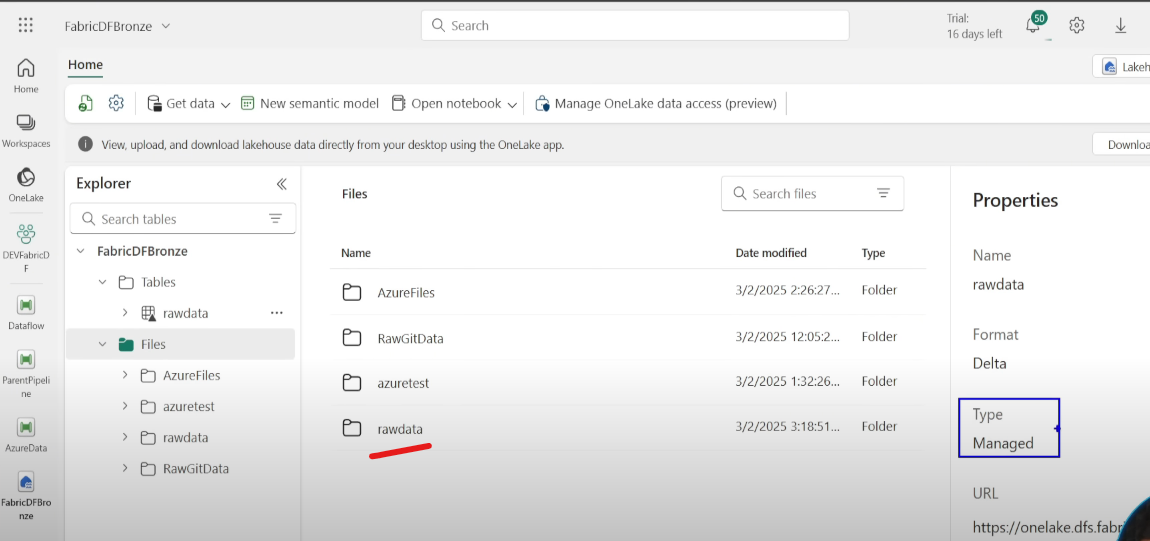
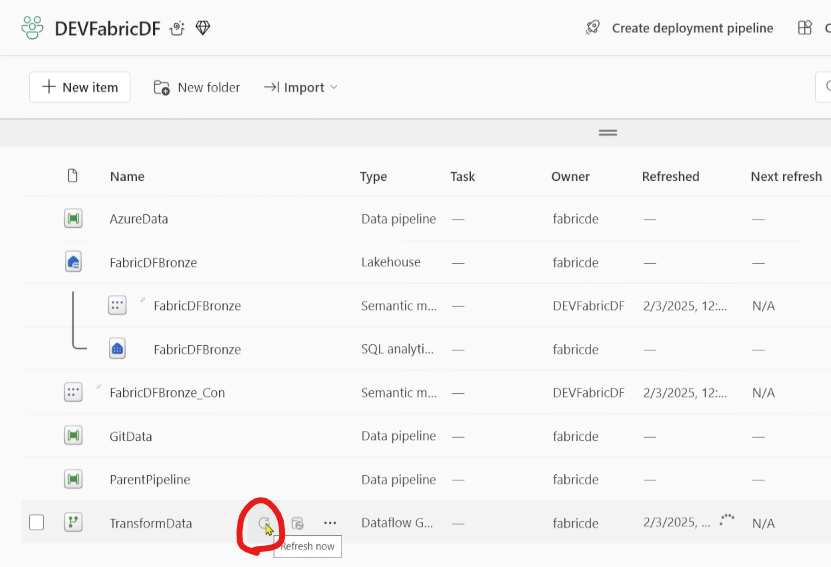
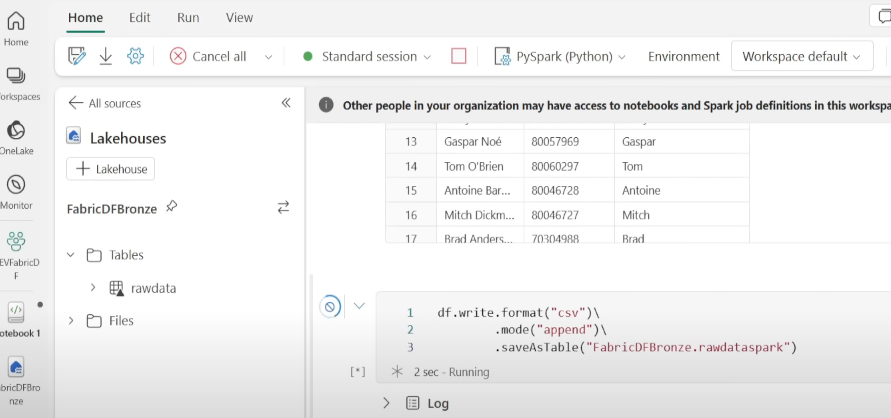
AI-generated content may be incorrect.  
I want to see how many movies are done by the same director, Group by on directors  
 A screenshot of a computer

AI-generated content may be incorrect.  
A screenshot of a computer

AI-generated content may be incorrect. A screenshot of a computer

AI-generated content may be incorrect.  
Filter only those if count more than 10  
  
 A screenshot of a computer

AI-generated content may be incorrect. A screenshot of a computer

AI-generated content may be incorrect.  
Publish the Data Flow, run it   
We can run Data Flows in 2 ways: 1)manual 2)orchestrate with Data Pipeline   
  
  
  
**Data Transformation using Notebook**  
No triangle because it’s a csv table managed  
Use Notebook activity in Data Pipeline to orchestrate the flow  
A screenshot of a computer

AI-generated content may be incorrect.  
Add Schedule Trigger to automate the flow  
A screenshot of a computer

AI-generated content may be incorrect. A screenshot of a computer

AI-generated content may be incorrect.