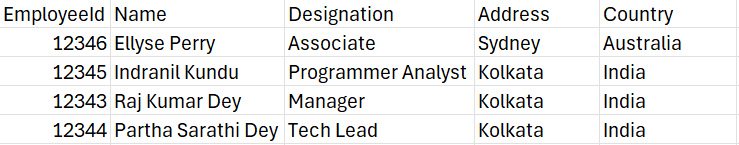
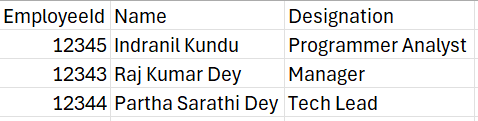
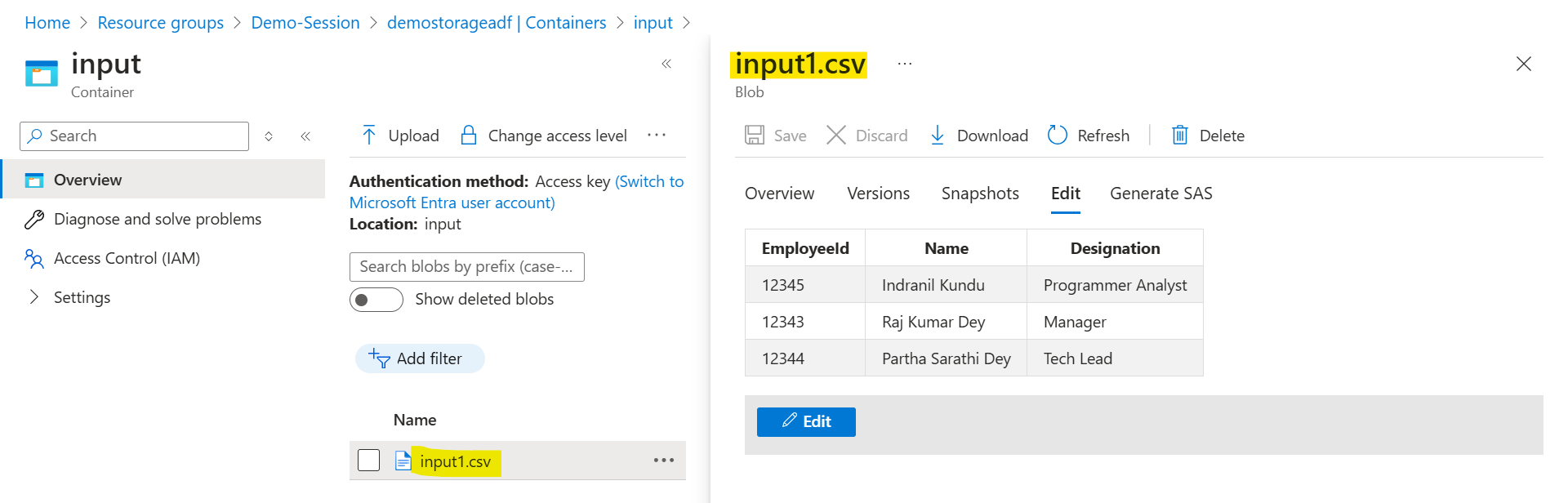
**Azure POC 4 (Azure Data Factory)**

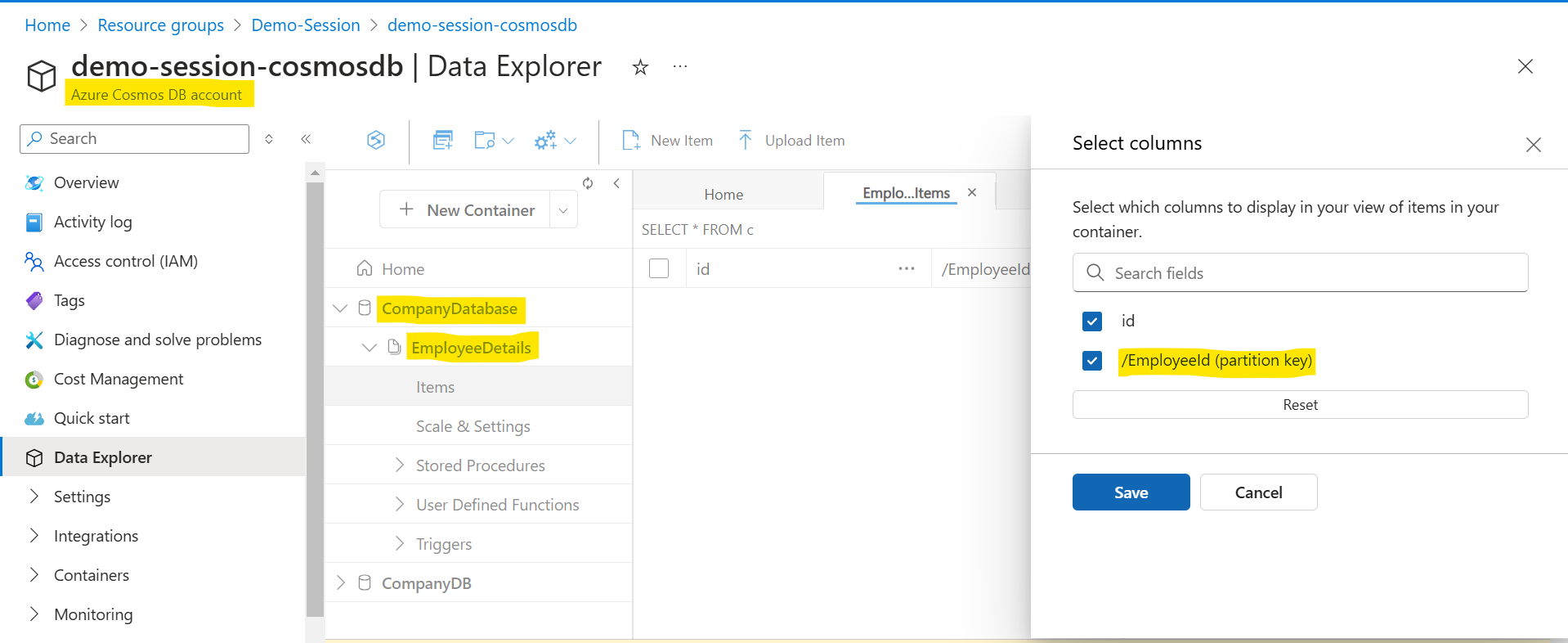
Scenario – We are getting some information about employee details with 3 columns now. In future we may get more additional column of information. We need to store all the information in the database.

Input 1: Input2: (more columns can be added in future)

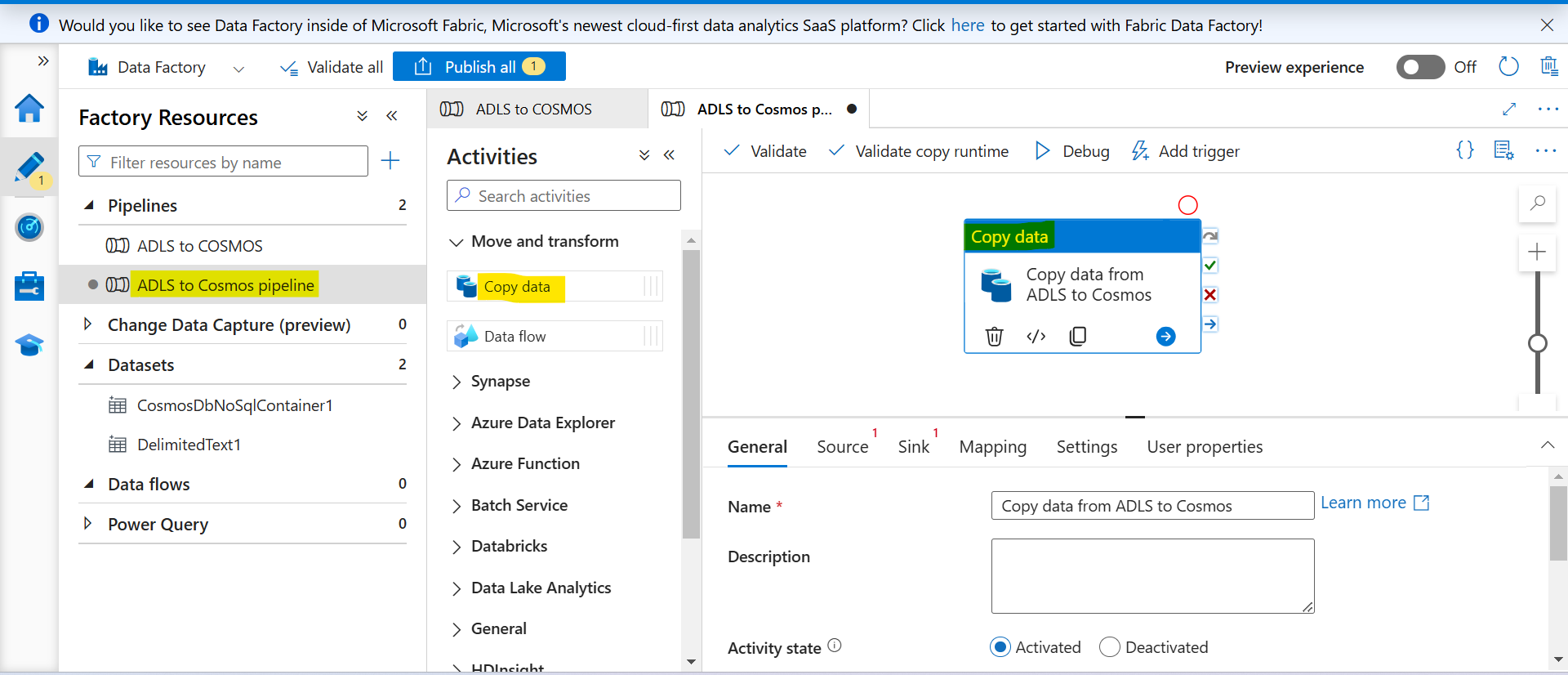
1. Create a blob storage. (input1 as first source file)



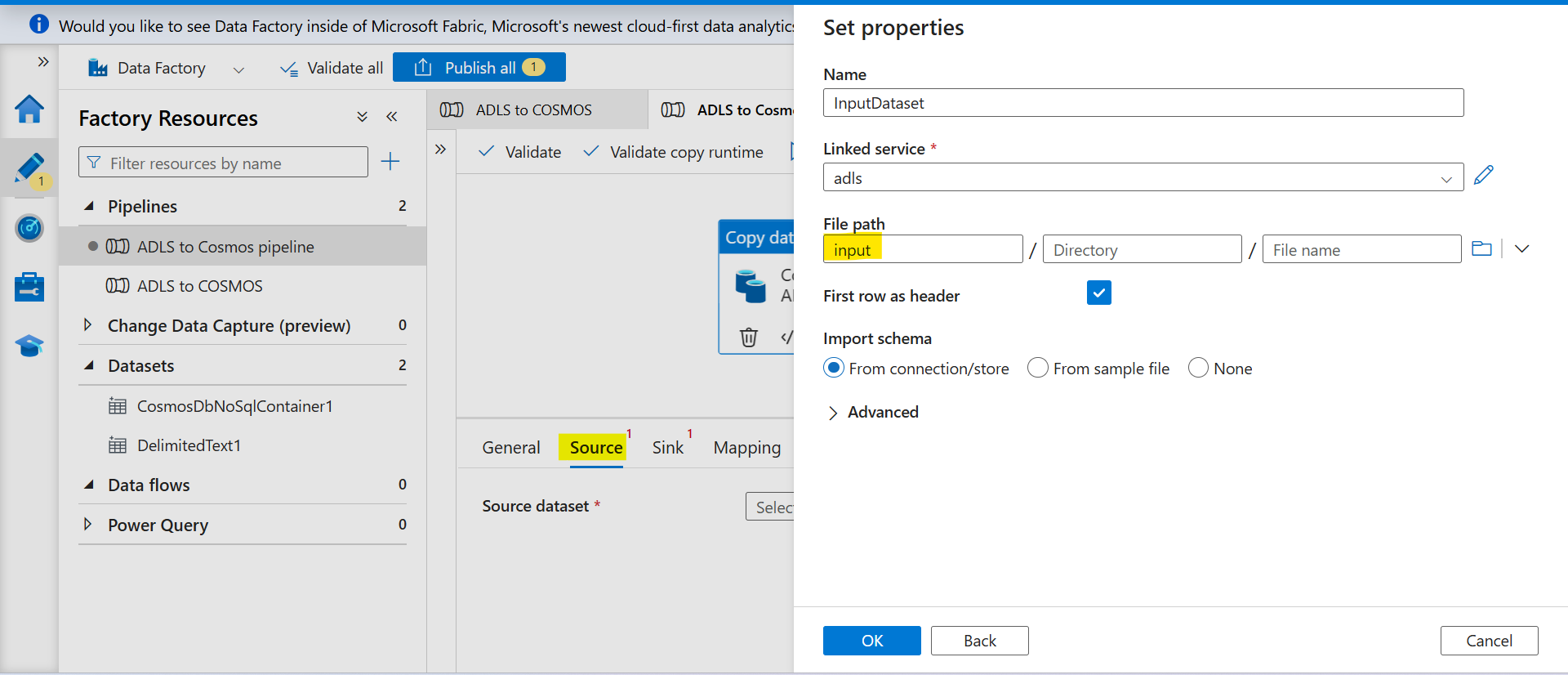
1. Create an Azure Cosmos DB account for scalable database.

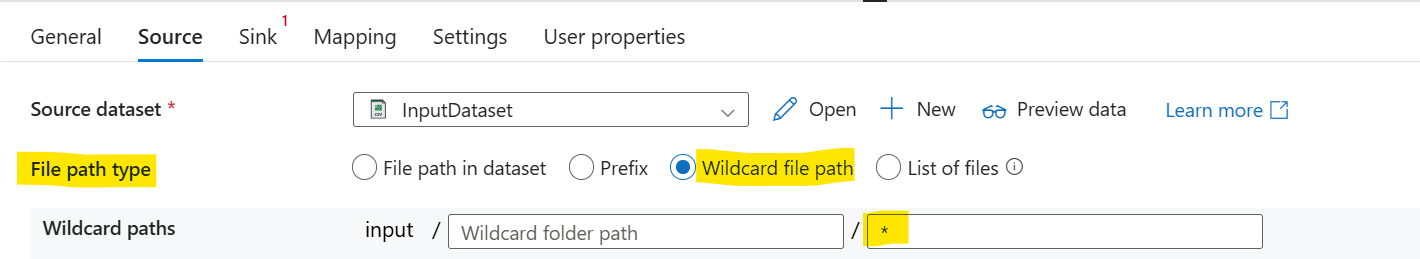


1. Create an Azure Data Factory pipeline and take Copy data activities.



Set Source where input data will come -> choose File path type as Wildcard file path (\* for all files) -> set id as EmployeeId for unique identifier.

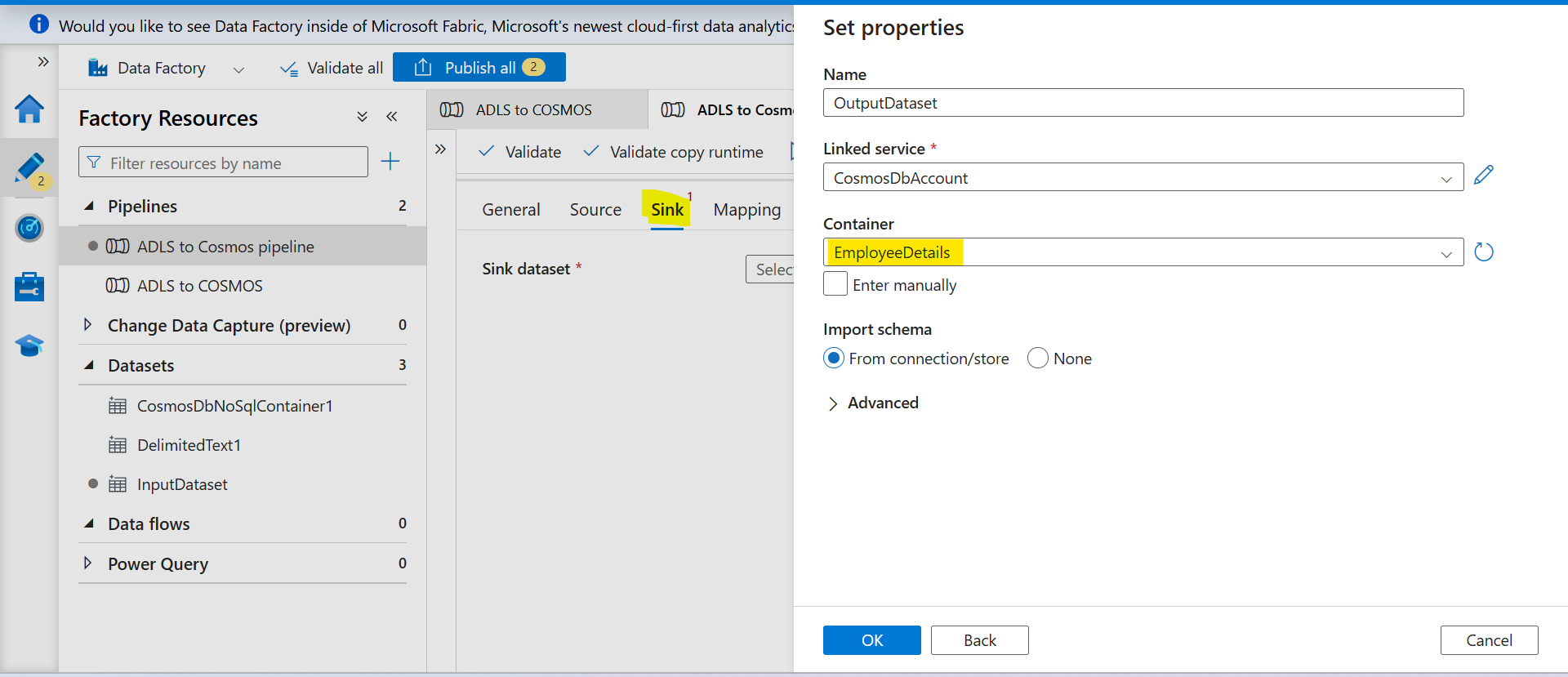


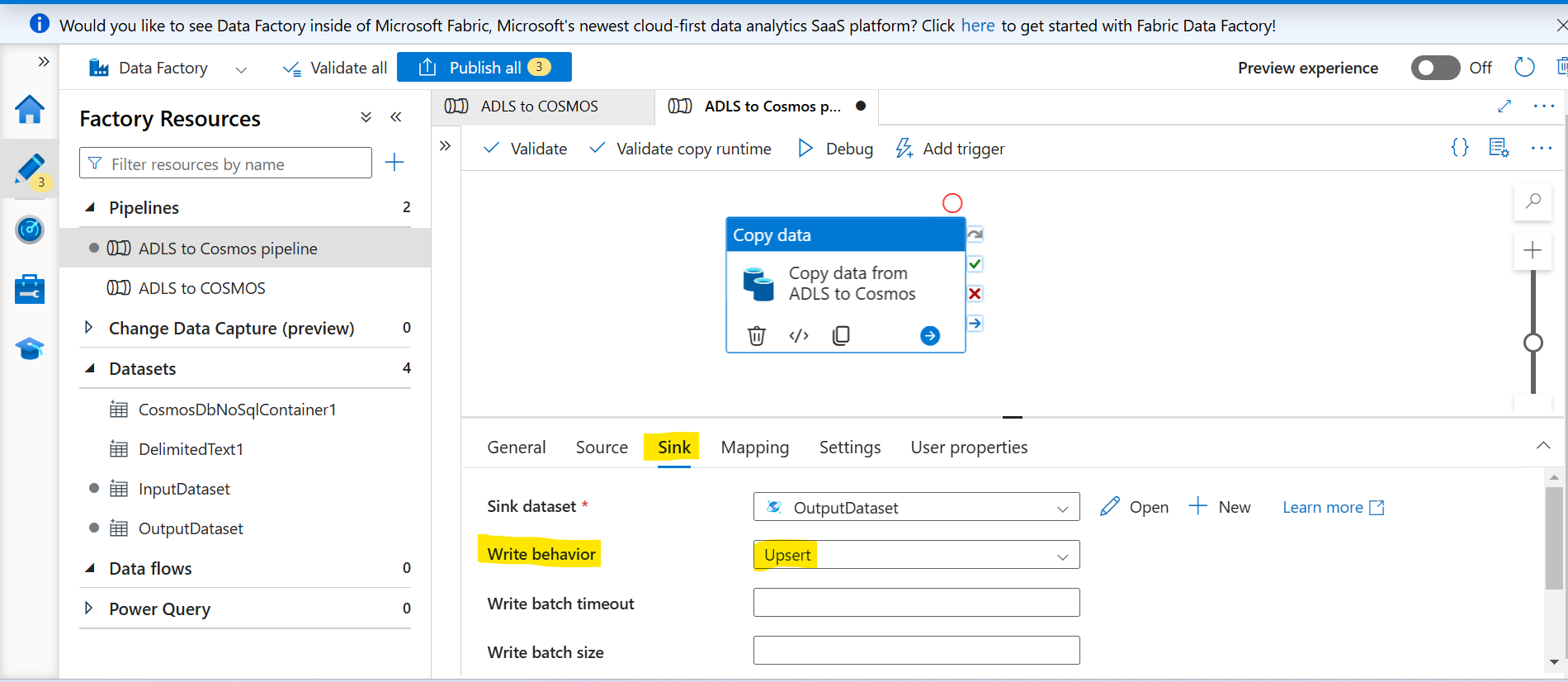


A screenshot of a computer

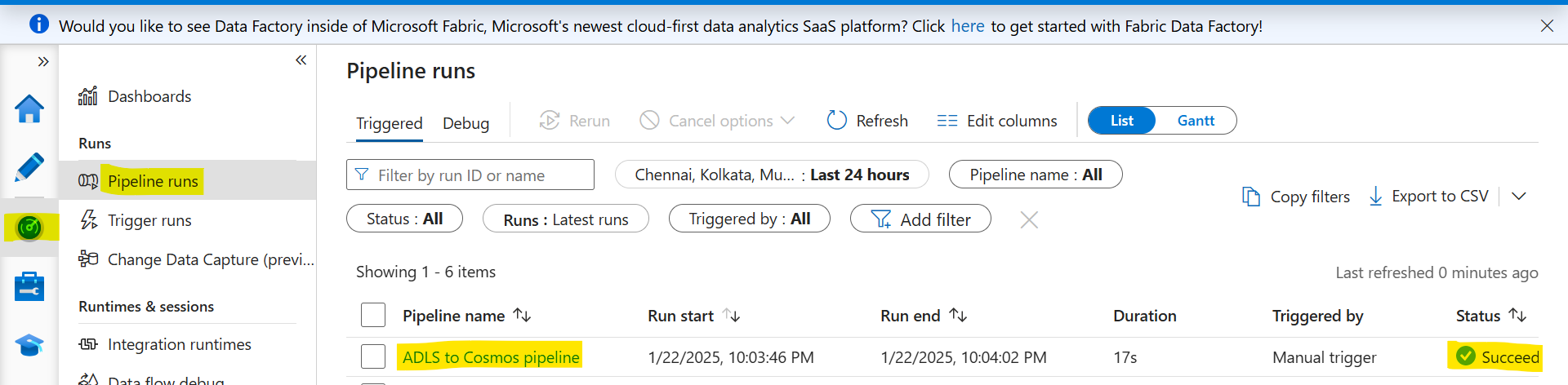
Description automatically generated

Set Sink to CosmosDB as destination -> set Write behavior to Upsert(insert/update).

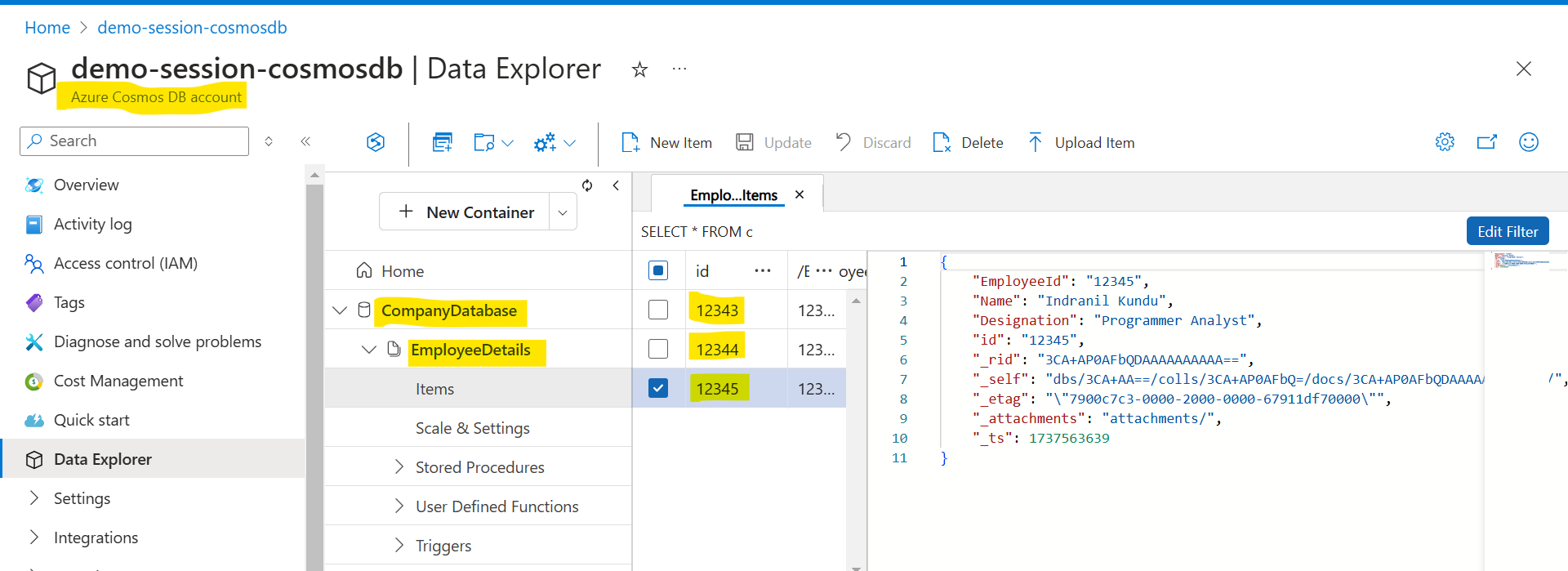




Click on Validate all -> Publish all -> trigger the pipeline and Monitor it.



Check the database that all the data updated correctly or not.

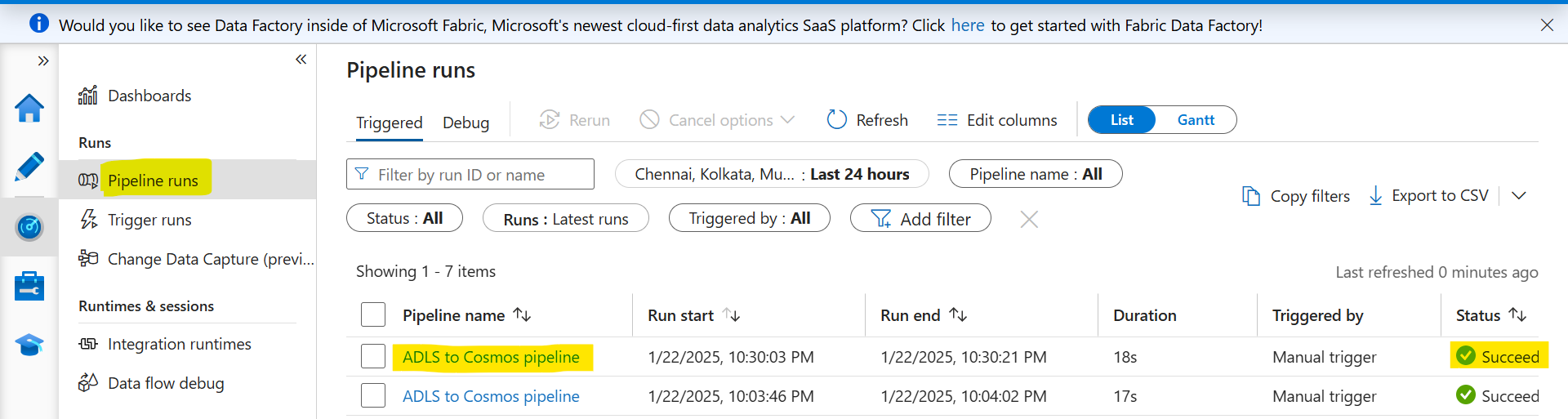


Now upload input2 file which has more additional information.

A screenshot of a computer

Description automatically generated

Trigger the pipeline.



Again, check the database that all the additional information updated correctly or not.

