Azure Data Pipeline Lab

Outline

- Lab1: Setup & query data from Azure SQL Database
- Lab2: ETL pipeline in Azure Data Factory (ADF) with Azure Data Lake (ADL)
  - Lab2-1: Create Dataflow “IMDBDataflow”
  - Lab2-2: Create Pipeline “IMDBPipeline”
    - Add “Web node” to send email
    - Need to setup LogicApps
  - Lab2-3: Set triggers (scheduler)

Lab1: Setup & query data from Azure SQL Database

Steps

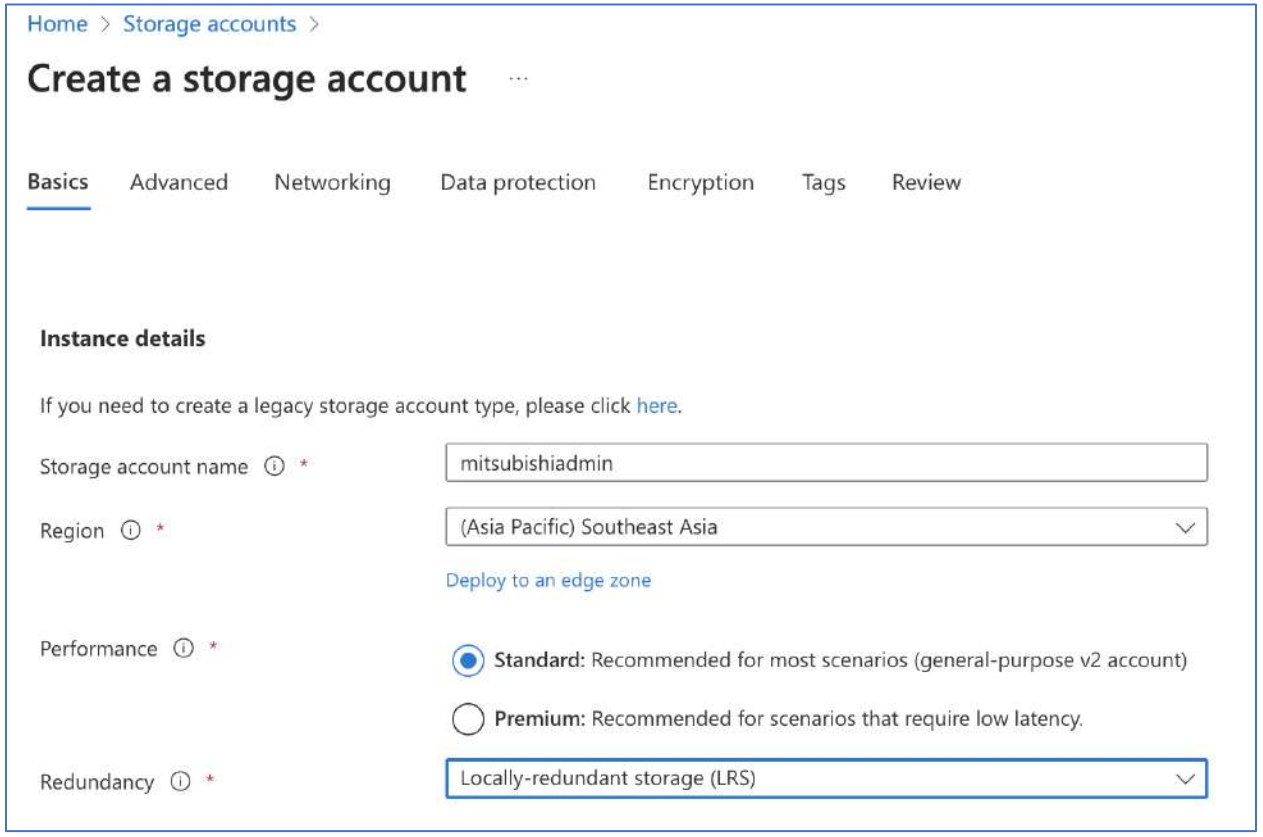
- Create “Storage account” (lake)
- Upload data source “AllElectronics.bacpac” to Container “data”

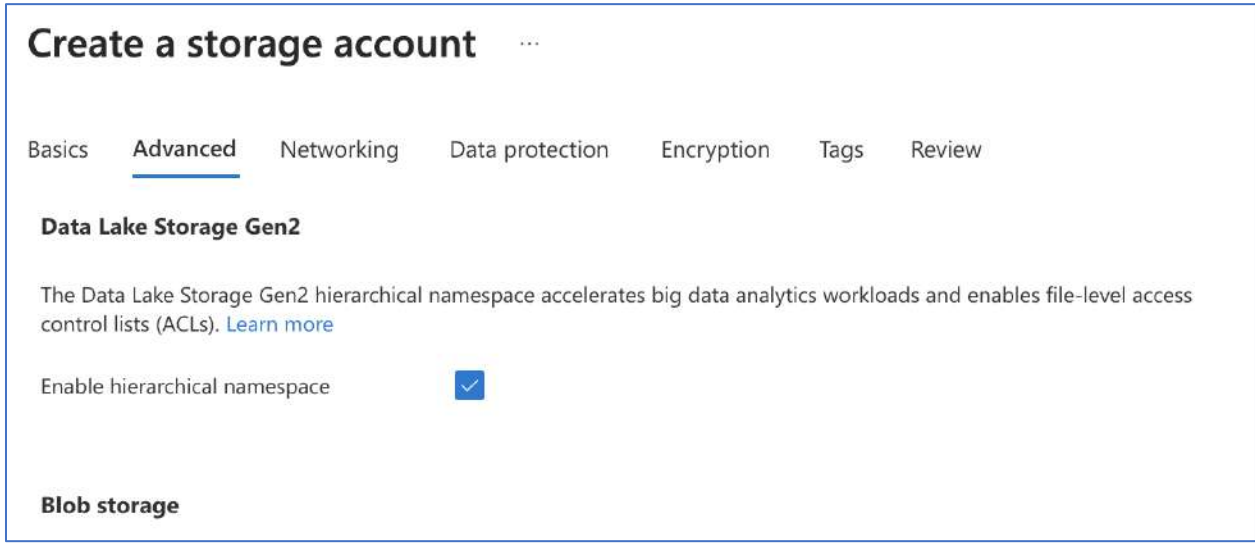
- Create “Azure SQL server” (serverless)
- Create “Azure SQL database”
- Import database (file.bacpac) from the Storage account (lake)

- Now ready to query data using “Query Editor”

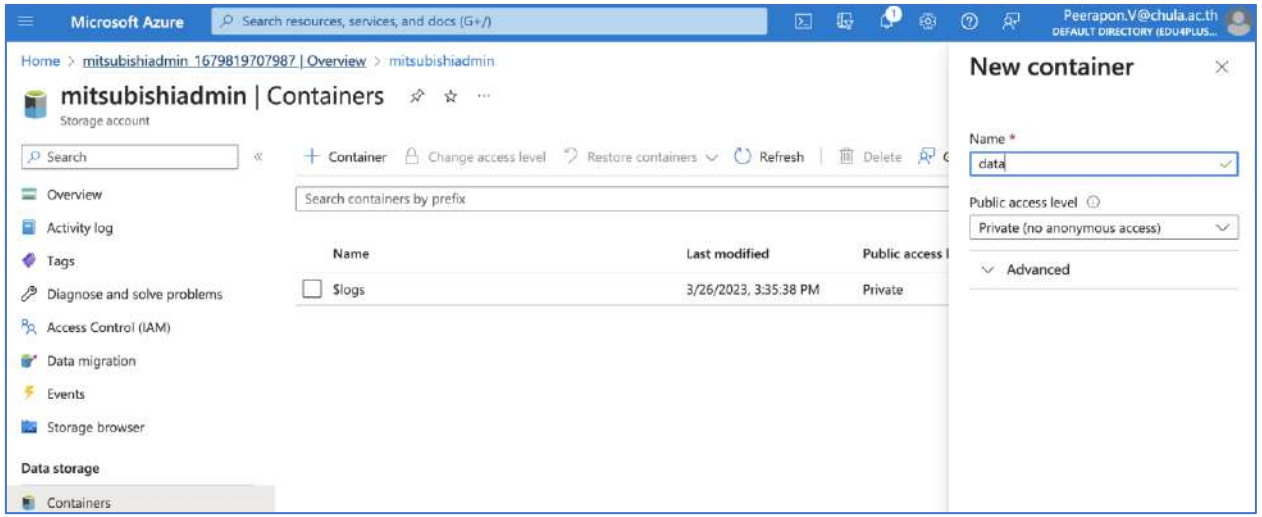
Create “Storage account”

- Check to enable “Data Lake Storage Gen2” in Advanced tab



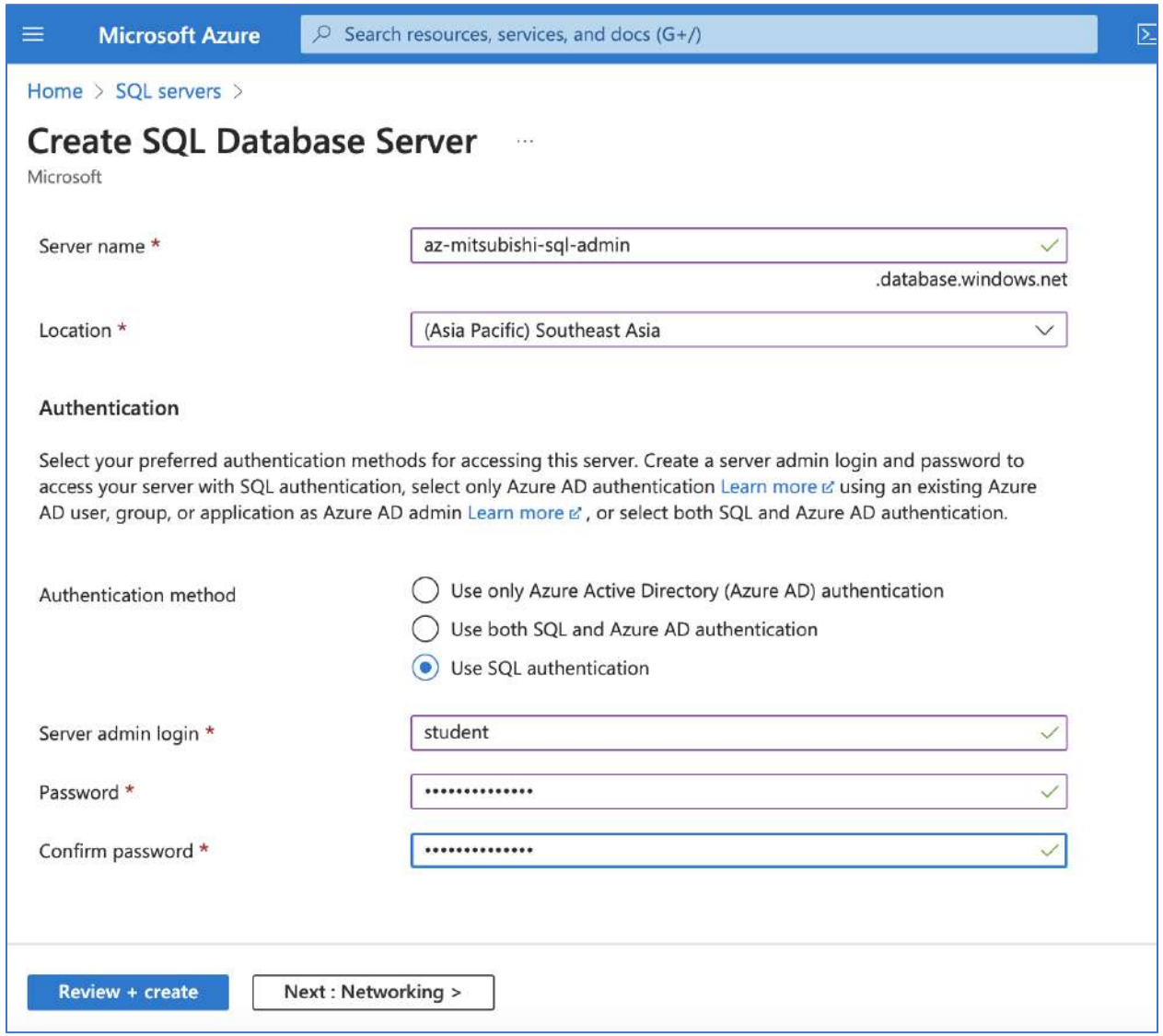


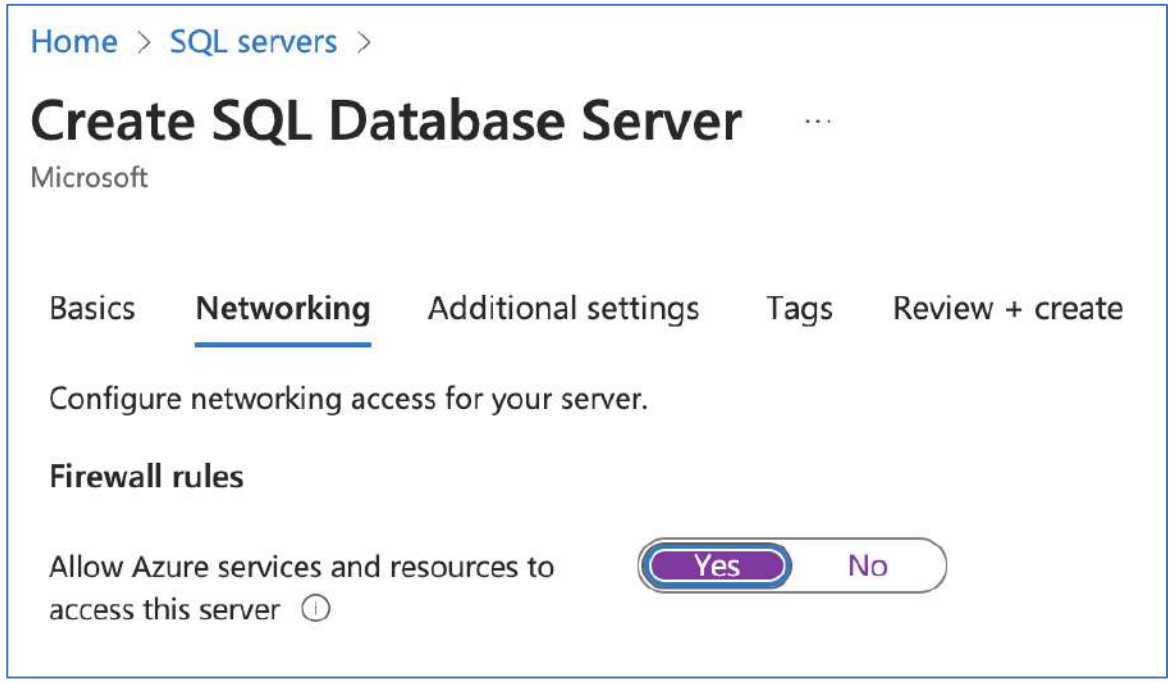
Upload data source “AllElectronics.bacpac” to Container “data”



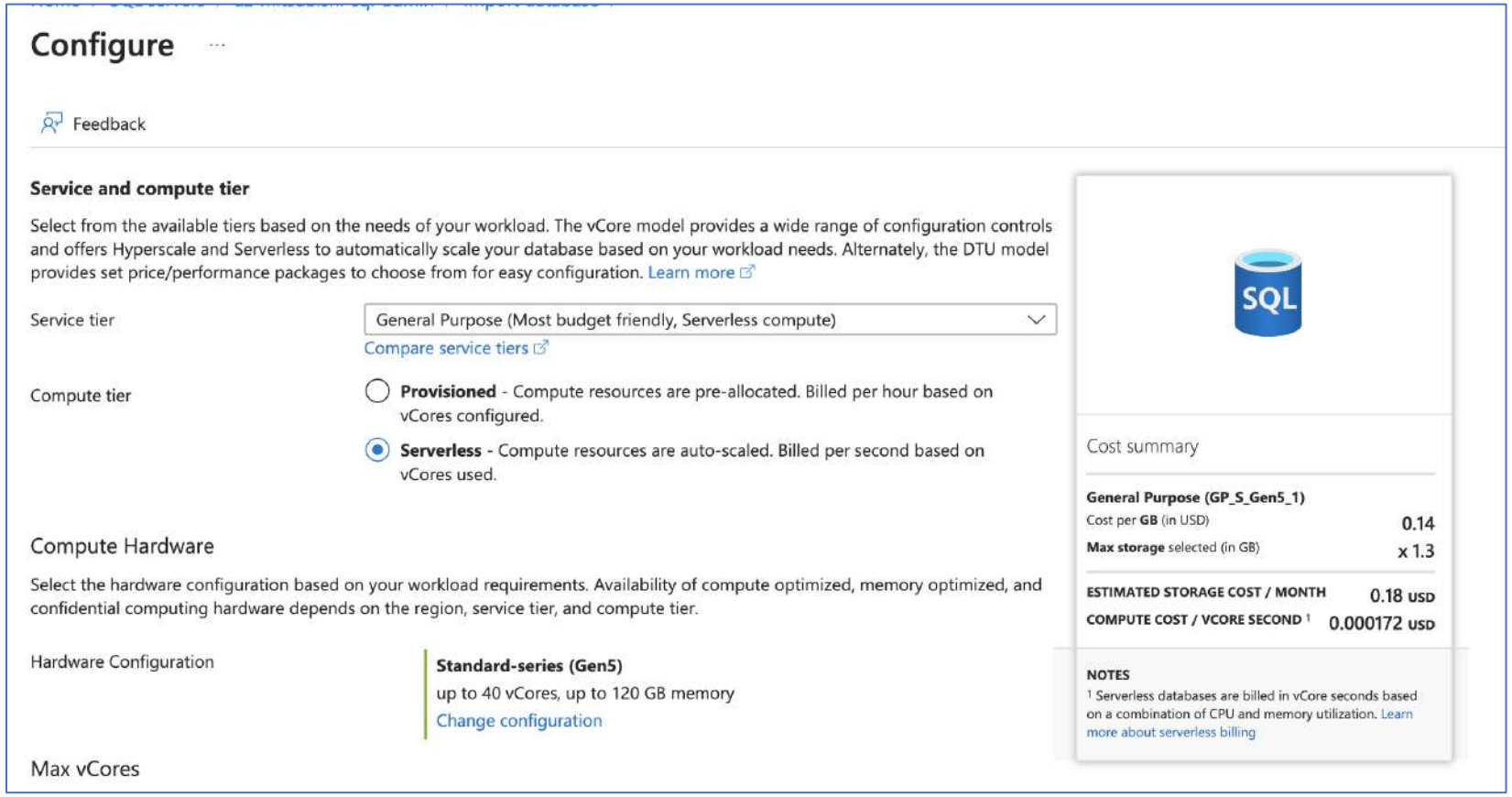
Create Azure SQL Server

User: student
Password: Mitsubishi2023

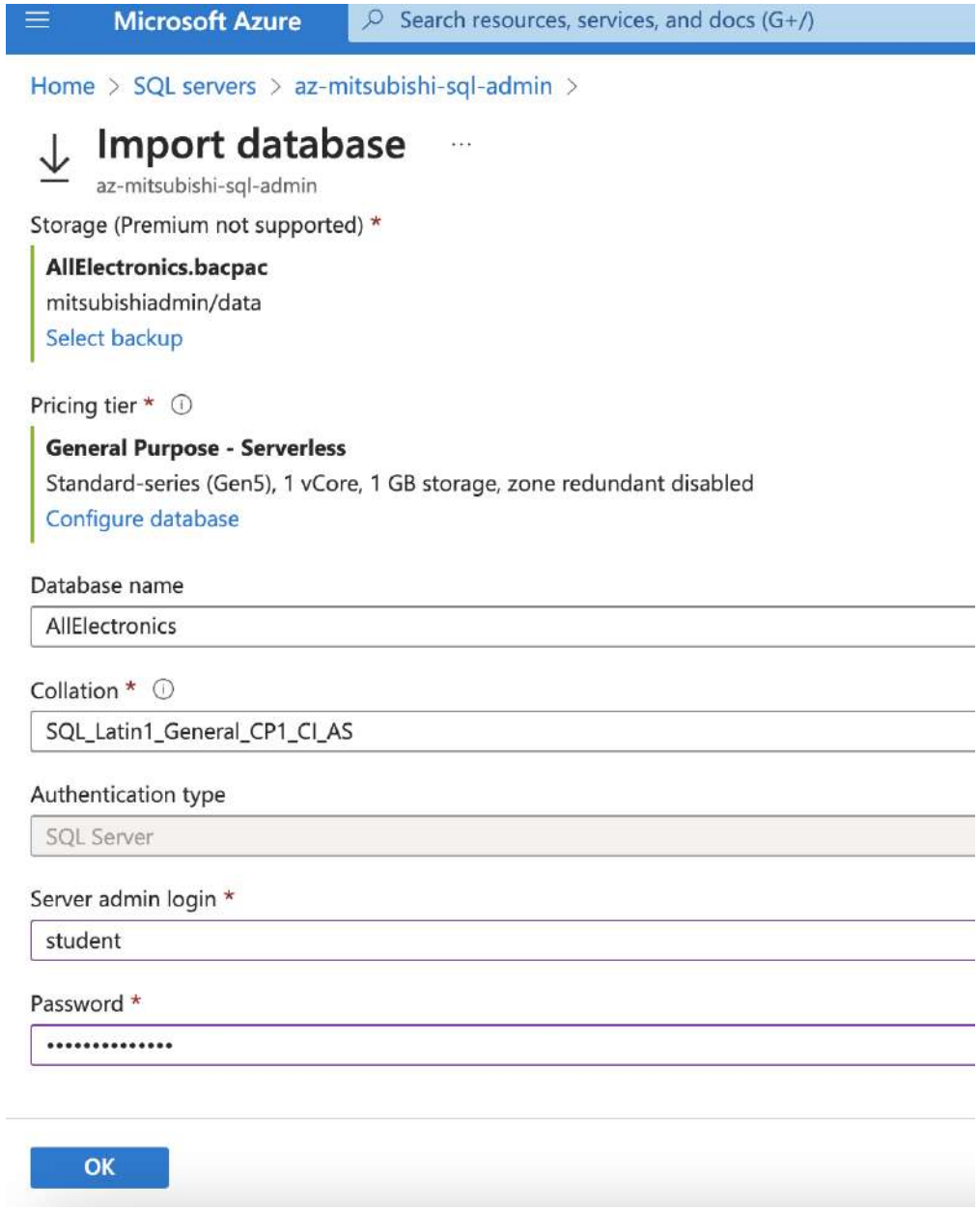




Set SQL Server to be “serverless”



Import database



Lab2: ETL pipeline in Azure Data Factory (ADF) with Azure Data Lake (ADL)

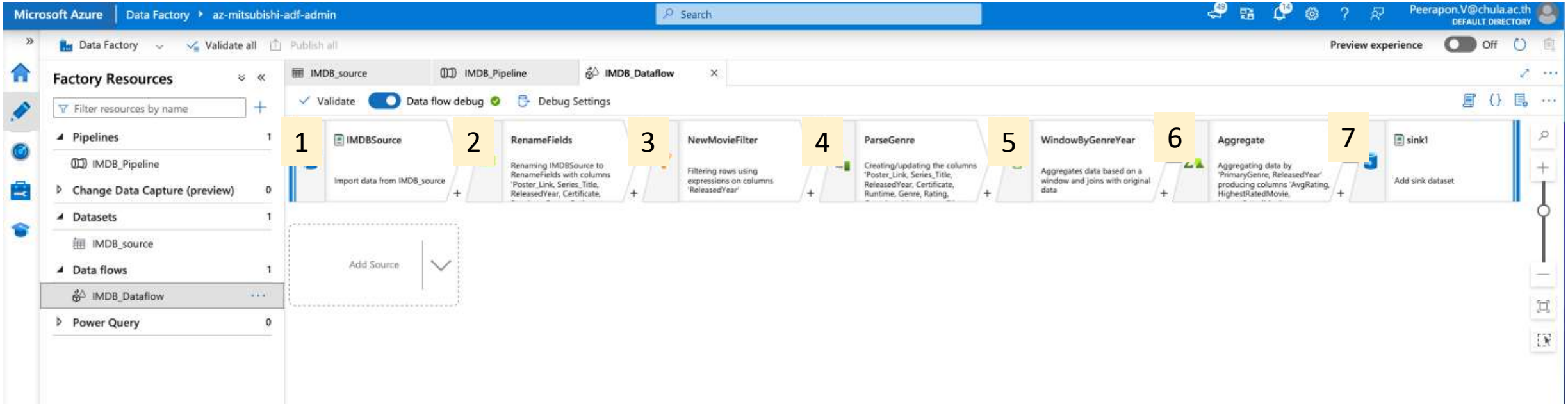
Steps

- Lab2-0: Upload “imdb-top-1000.csv” into ADL (Storage account)
- Lab2-1: Create Dataflow “IMDBDataflow”
- Lab2-2: Create Pipeline “IMDBPipeline”
  - Add “Web node” to send email
  - Need to setup LogicApps
- Lab2-3: Set triggers (scheduler)

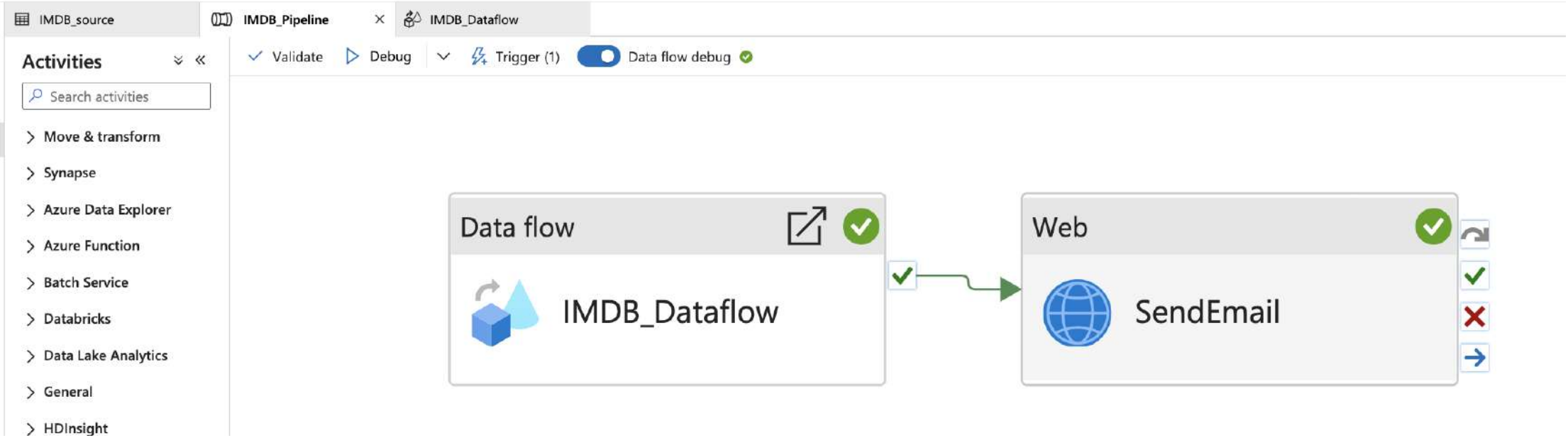
Lab2-1: Create Dataflow “IMDBDataflow”

- Create Link
  - AzureDataLakeStorage
- Create Data Set
  - IMDBsource “imdb-top-1000.csv” into ADL

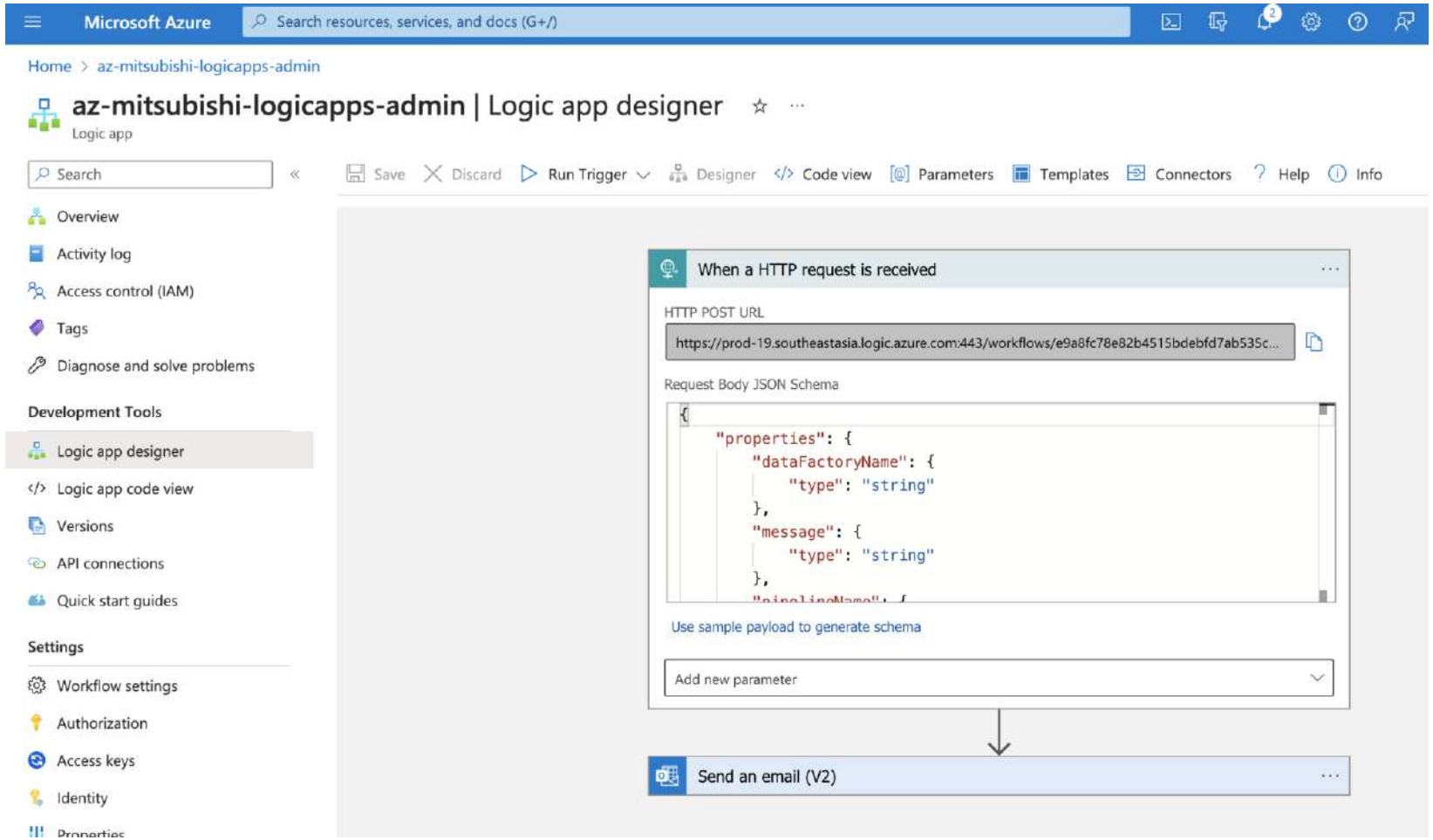
Lab2-1: Create Dataflow “IMDBDataflow” (cont.)



Lab2-2: Create Pipeline “IMDBPipeline”

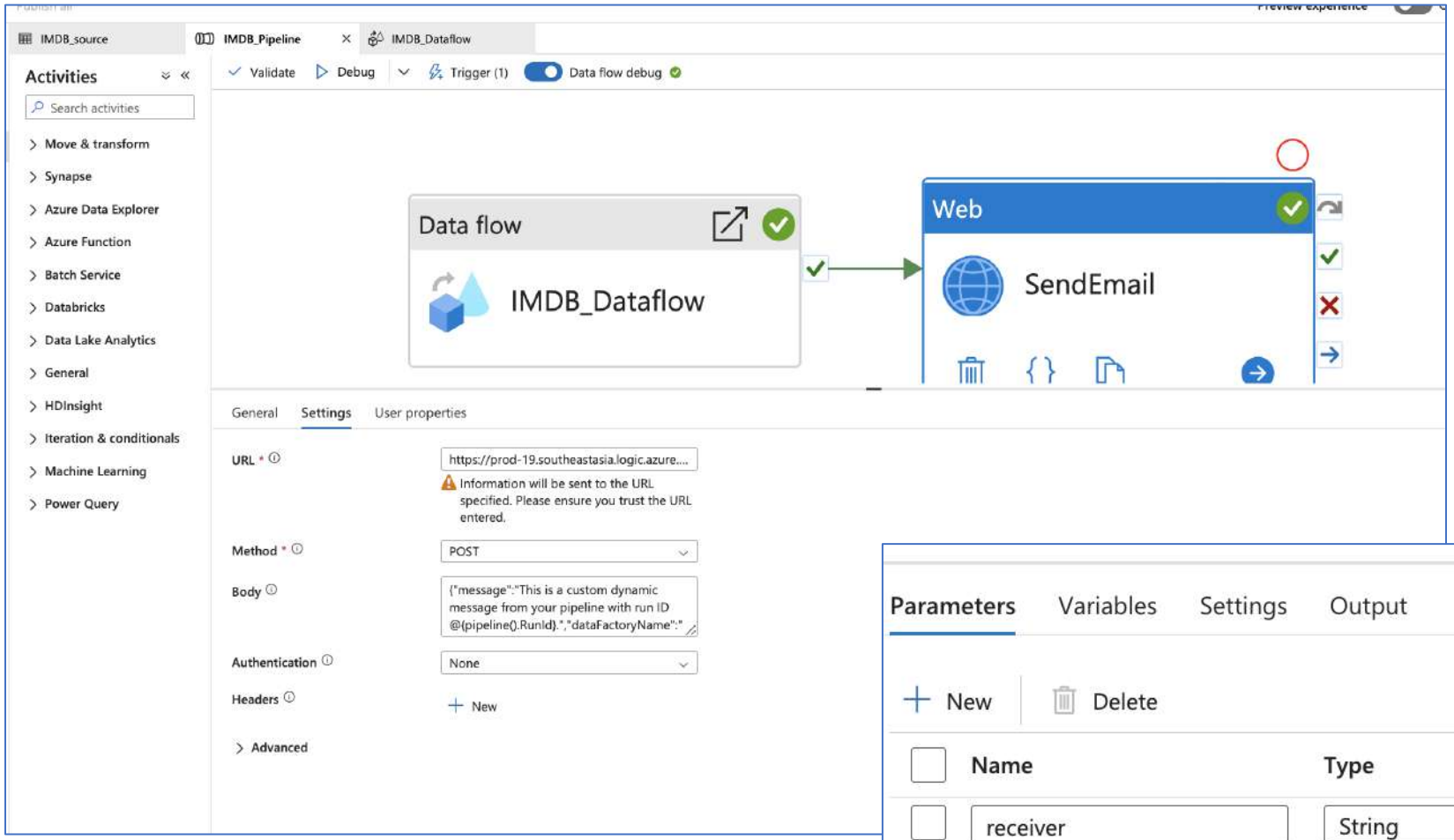


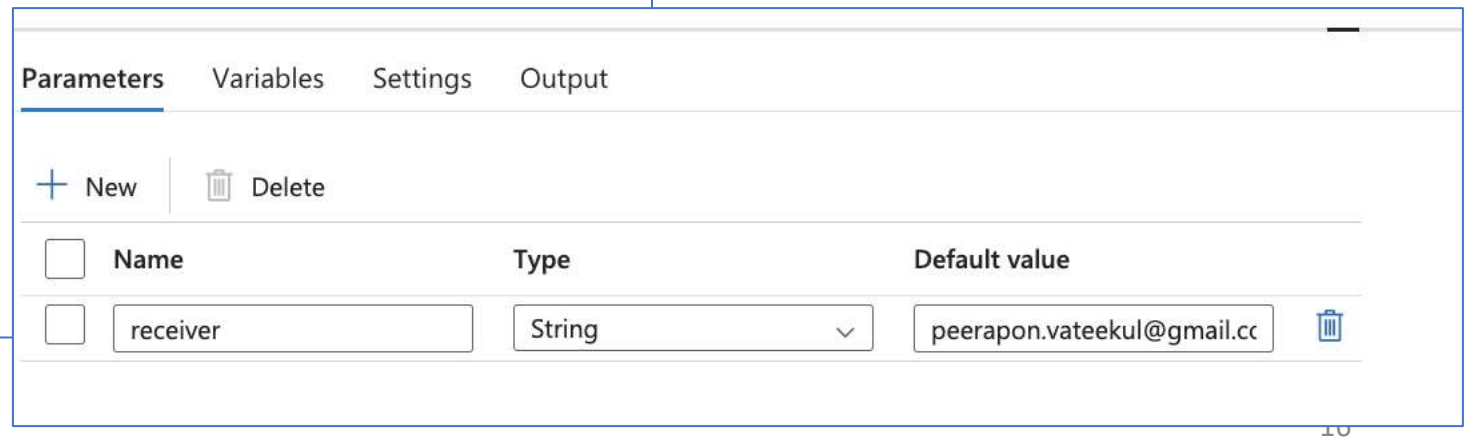
Create LogicApps



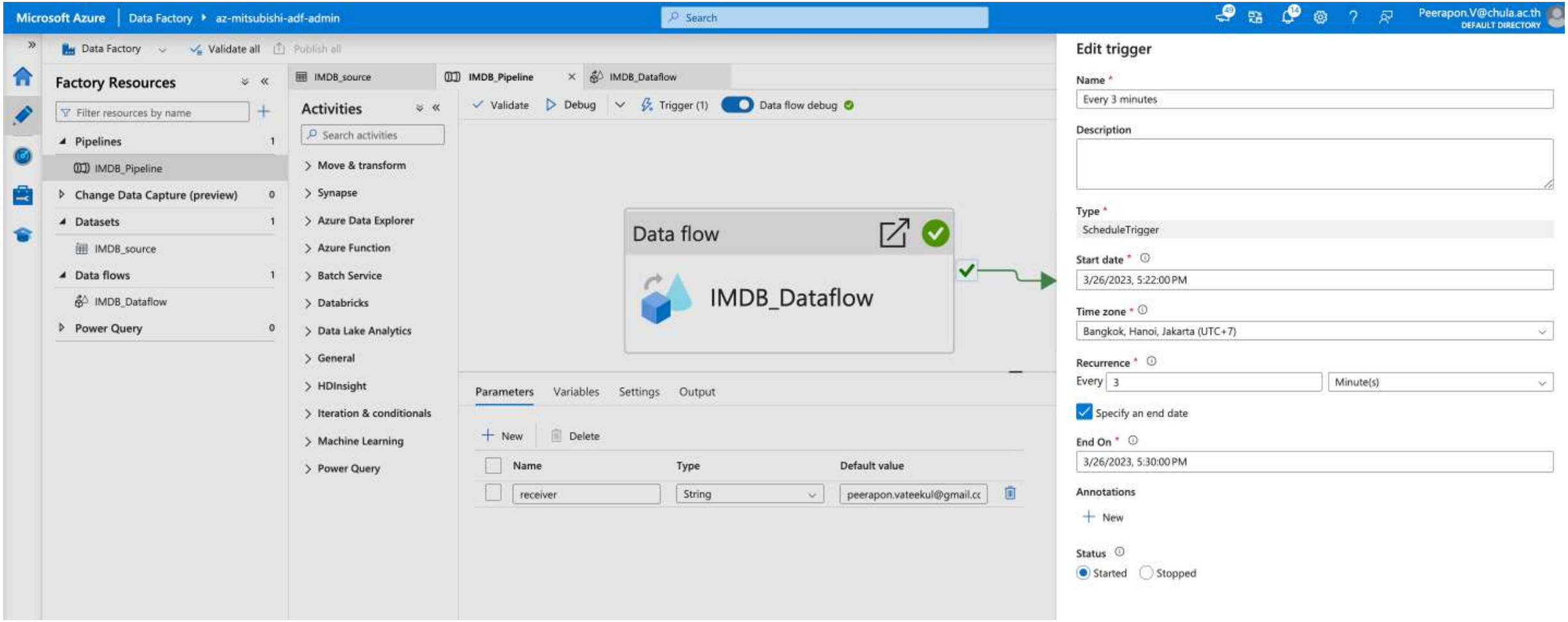
https://learn.microsoft.com/en-us/azure/data-factory/how-to-send-email

Create Web node in Pipeline





Lab2-3: Set triggers (scheduler)



Finally, you can check your trigger here ☺

- Monitor → Trigger runs