Lab 22 – Databricks ADF

1. Provision Azure Resources

A screenshot of a computer

AI-generated content may be incorrect.

1. Import a notebook

A screenshot of a computer

AI-generated content may be incorrect.

1. Generate an access token

A screenshot of a computer

AI-generated content may be incorrect.

1. Create a linked service in Azure Data Factory

A screenshot of a computer

AI-generated content may be incorrect.

1. Create a pipeline

A screenshot of a computer

AI-generated content may be incorrect.

1. Run the pipeline

A screenshot of a computer

AI-generated content may be incorrect.

1. Delete Azure Resources

A screenshot of a computer

AI-generated content may be incorrect.

Summary:

To orchestrate tasks as part of a data engineering pipeline , Azure Data Factory can be used along with Databricks. In this lab, we will be focusing on automating an Azure Databricks Notebook with Azure Data Factory. We begin our lab by provisioning Azure resources. Once the resources are allocated, we open the databricks workspace and import a notebook from Git. The notebook uses a parameter folder which stores the location of the data. To integrate Azure databricks with Azure Databricks we create an access token and linked service. Once the linked service is created, we can set up a pipeline and run the pipeline automatically through ADF. The lab ends with deleting Azure resources.