Azure Coding Challenge

-Sarthak Niranjan Kulkarni (Maverick)

- [sarthakkul2311@gmail.com](mailto:sarthakkul2311@gmail.com) - (+91) 93256 02791

**10/12/2024 (Tuesday)**

**Copy activity from datalake to datalake in azure datafactory: -**

**1. Opening Data Factory**

* Create Data Factory

**A screenshot of a computer

Description automatically generated**

* Then launch Azure Data Factory Studio

A screenshot of a computer

Description automatically generated

**2. Ingesting the data**

* Click on Ingest

A screenshot of a computer

Description automatically generated

**3. Setting up Copy data tool**

* Properties

Select your property and proceed

A screenshot of a computer

Description automatically generated

* Source 🡪 Dataset

Select your source file and check binary copy checkbox

A screenshot of a computer

Description automatically generated

* Source 🡪 Dataset 🡪 Configuration

Select your desired configurations and click on next

A screenshot of a computer

Description automatically generated

* Destination 🡪 Dataset

Create a new connection of Data Lake Storage

A screenshot of a computer

Description automatically generated

Select destination path and give a name to the file

A screenshot of a computer

Description automatically generated

* Destination 🡪 Dataset 🡪 Configuration

Select your desired configurations and click on next

A screenshot of a computer

Description automatically generated

* Settings

Give some name to the task and proceed

A screenshot of a computer

Description automatically generated

**4. Deploying the Copy Data Pipeline**

* Review all the configuration and click on next

A screenshot of a computer

Description automatically generated

* This window would ensure the deployment of Copy data Pipeline

A screenshot of a computer

Description automatically generated

**5. Validating the Pipeline**

**A screenshot of a computer

Description automatically generated**

**6. Debugging the pipeline**

* Debugging process would start wait for some time to complete

**A screenshot of a computer

Description automatically generated**

**A screenshot of a computer

Description automatically generated**

**7. Verify the whole process by visiting Data Lake storage account**

**A screenshot of a computer

Description automatically generated**

**Data Factory:** Azure Data Factory is a cloud-based data integration service that enables you to create, schedule, and orchestrate data workflows. It supports data movement and transformation across a wide range of sources, making it ideal for ETL processes and data automation.

**Copy Data Pipeline:** The Copy Data Pipeline in Azure Data Factory is used to move data from one source to a destination efficiently. It supports diverse data formats and can handle complex transformations during the data transfer process.