**Data Factory Coding Challenge**

**Name: Vikas Reddy Gorantla**

**Date: 10-12-2024**

**Steps to Copy Data Between Azure Data Lake Storages Using Azure Data Factory**

**1. Create a Resource Group**

* Navigate to the Azure portal.
* Create a new resource group for the Azure Data Factory.

A computer screen shot of a computer

Description automatically generated

**2. Open Azure Data Factory and Select "Ingest"**

* Inside the created Data Factory, click on the **Ingest** option.

A screenshot of a computer

Description automatically generated

**3. Choose Task Type**

* Select **Built-in copy task** as the task type.
* Choose **Run Once** for task cadence (since this is a tutorial).

A screenshot of a computer

Description automatically generated

**4. Configure Source Data Store**

* Select **Azure Data Lake Storage Gen 2** as the source type.
* Click **New Connection** if no connection exists:
  1. Provide a connection name.
  2. Select **Integration Runtime** and **Authentication Type** (auto-selected).
  3. Use **Azure Subscription** as the account selection method.
  4. Choose your **Azure Subscription** and specify the **Storage Account Name**.
  5. Test the connection and click **Create**.
* Select the source file or folder from the container.
* Enable **Binary Copy** and **Recursive** options.

A screenshot of a computer

Description automatically generated

**A screenshot of a computer

Description automatically generated**

**5. Configure Destination Data Store**

* Choose **Azure Data Lake Storage Gen 2** again as the destination type.
* Repeat the **New Connection** process:
  1. Provide a name, runtime, authentication, and account details.
  2. Test and create the connection.
* Specify the destination folder path in the new Data Lake container.

**A screenshot of a computer

Description automatically generated**

**6. Set Pipeline Settings**

* Provide a **Task Name** (e.g., CopyPipeline2).

**A screenshot of a computer

Description automatically generated**

**7. Review and Deploy**

* Review the summary in the **Review and Finish** step.
* Proceed to the deployment dashboard, where the following will occur:
  + **Create Datasets**.
  + **Create Pipelines**.
  + **Validate and Run the Pipeline**.
* Once all steps succeed, click **Finish**.

**A screenshot of a computer

Description automatically generated**

**A screenshot of a computer

Description automatically generated**

**8. Debug the Pipeline**

* Go to the **Author** option on the Azure Data Factory homepage.
* Locate your pipeline by the name you provided (e.g., CopyPipeline2).
* Click **Debug** in the pipeline dashboard.
* Wait for the pipeline execution to complete.
* Confirm the activity status is marked **Succeeded**.

**A computer screen shot of a computer

Description automatically generated**

**9. Verify Data Transfer**

* Confirm that the data has been successfully copied from the source Data Lake to the destination Data Lake.

**A computer screen shot of a computer

Description automatically generated**