TASK 2

Configure FTP/SFTP Server and Create an ADF Pipeline for Data Extraction: Set up access to FTP/SFTP servers and build a pipeline in Azure Data Factory to extract files or datasets from these sources. This facilitates integration with external systems or partners.

(A sample SFTP is used using: https://sftpcloud.io/tools/free-sftp-server, it allows, 1GB free storage of cloud based SFTP server for one hour)

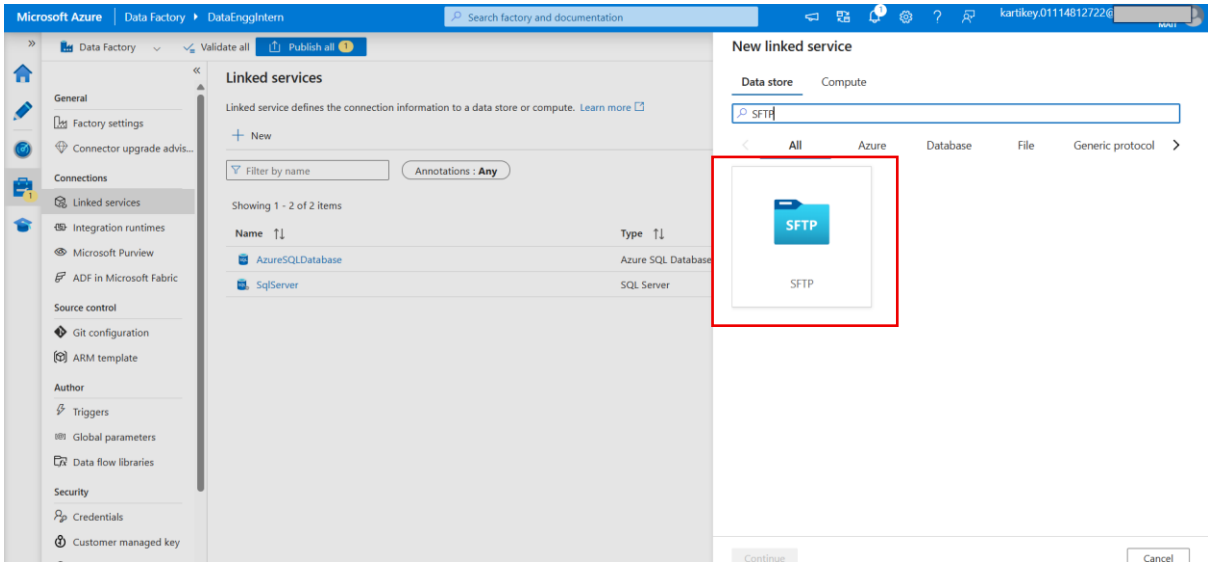
Step 1: Create Linked Services (Connections)

We'll create two linked services:

- One for your **SFTP**
- One for **Azure SQL Database**

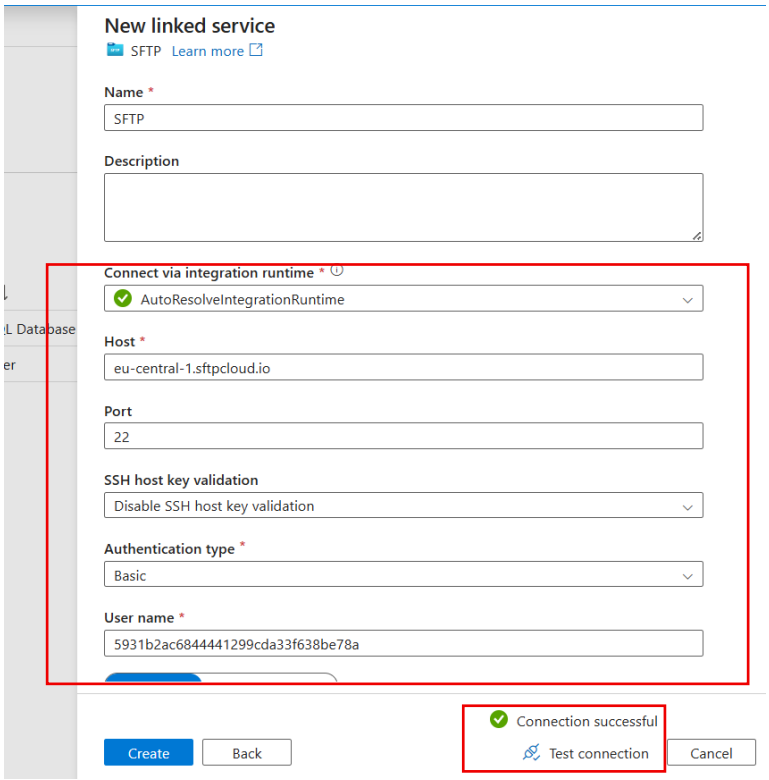
A. Creating Linked Service for SFTP

1. Go to **Manage** > **Linked Services** > Click + **New**.
2. In the search box, type **SFTP**, then select it and click **Continue**.



3. Fill in the details:
    - **Name**: <name>
    - **Host**: sftp.xxxxx.com
    - **Database name**: <db_name>
    - **Authentication**: **SSH or Basic authentication**(enter username/password)

4. Click **Test Connection** (make sure it's green).
5. Click **Create**.



B. Creating Linked Service for Azure SQL Database

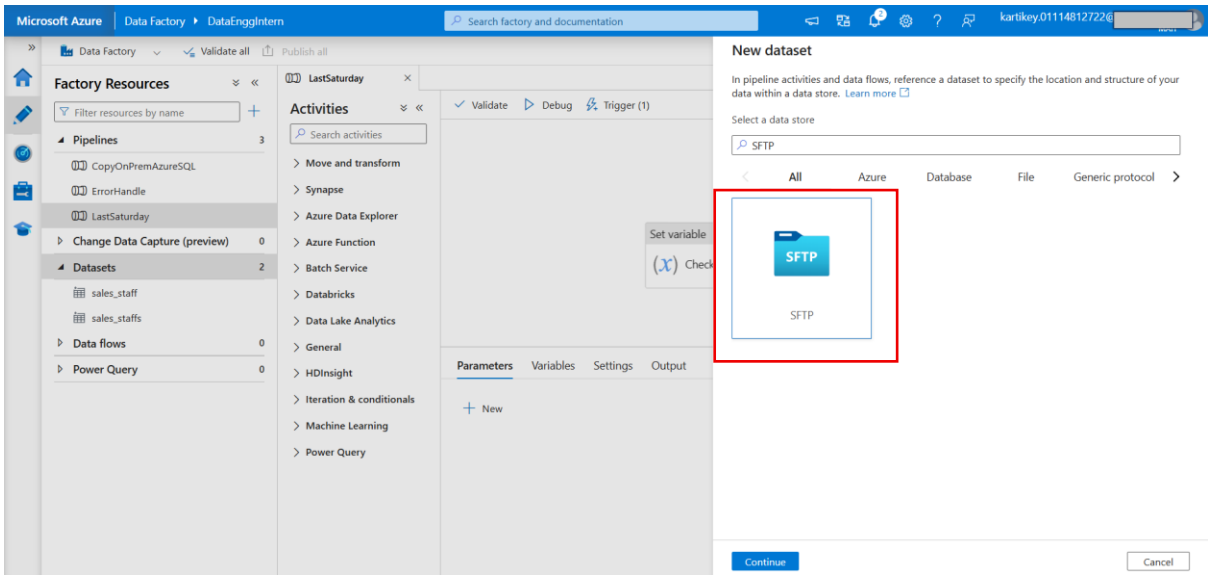
1. We created Azure SQL Database in Task 1

Step 2: Create Datasets

These datasets define the structure and source/target of the data being moved.

A. Create SFTP Dataset

1. Go to **Author** > Datasets > + **New dataset**.



2. Choose:
    - **SFTP**
3. Choose file format
4. Choose:
    - **File name**
    - Linked Service
    - File Path
5. Click **OK**.

Set properties

Name

DelimitedText2

Linked service *

SFTP

File path

eu-central-1.sftpcloud.io / sales / customers

First row as header

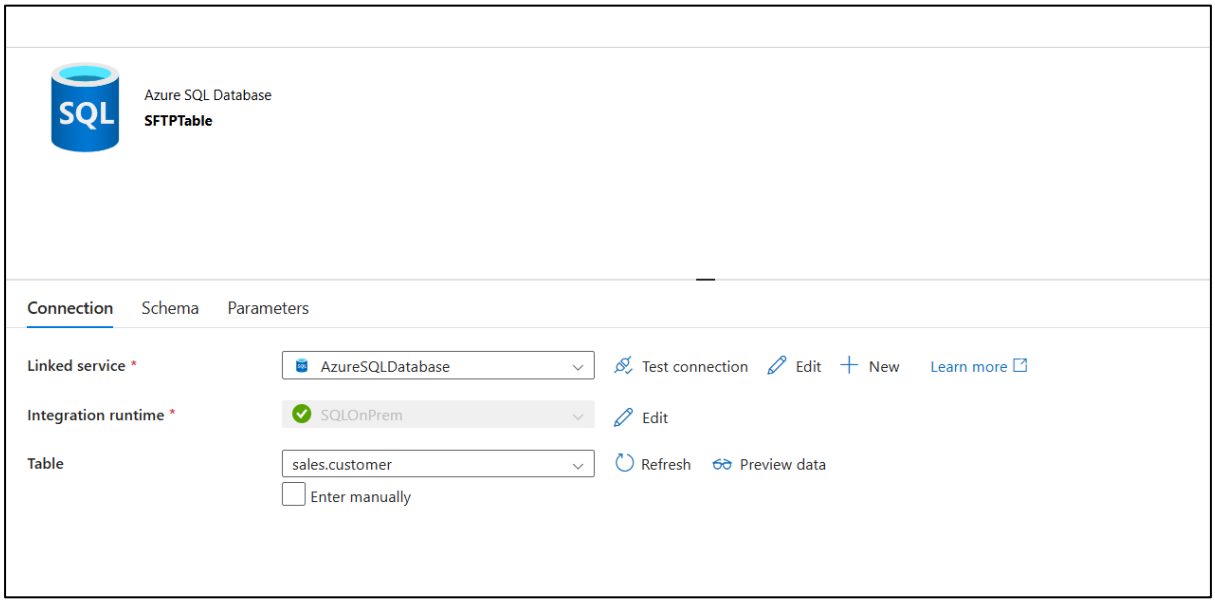
Import schema

From connection/store | From sample file | None

OK | Back | Cancel

B. Create Sink Dataset (Azure SQL Database)

1. Click + **New dataset**.
2. Choose:
    - **Azure SQL Database**
    - Linked Service: <service_name>
3. Name it: <table_name>
4. Select the target table name (can be pre-created or we can let ADF create it).
5. Click **OK**.



Step 3: Create a Copy Pipeline

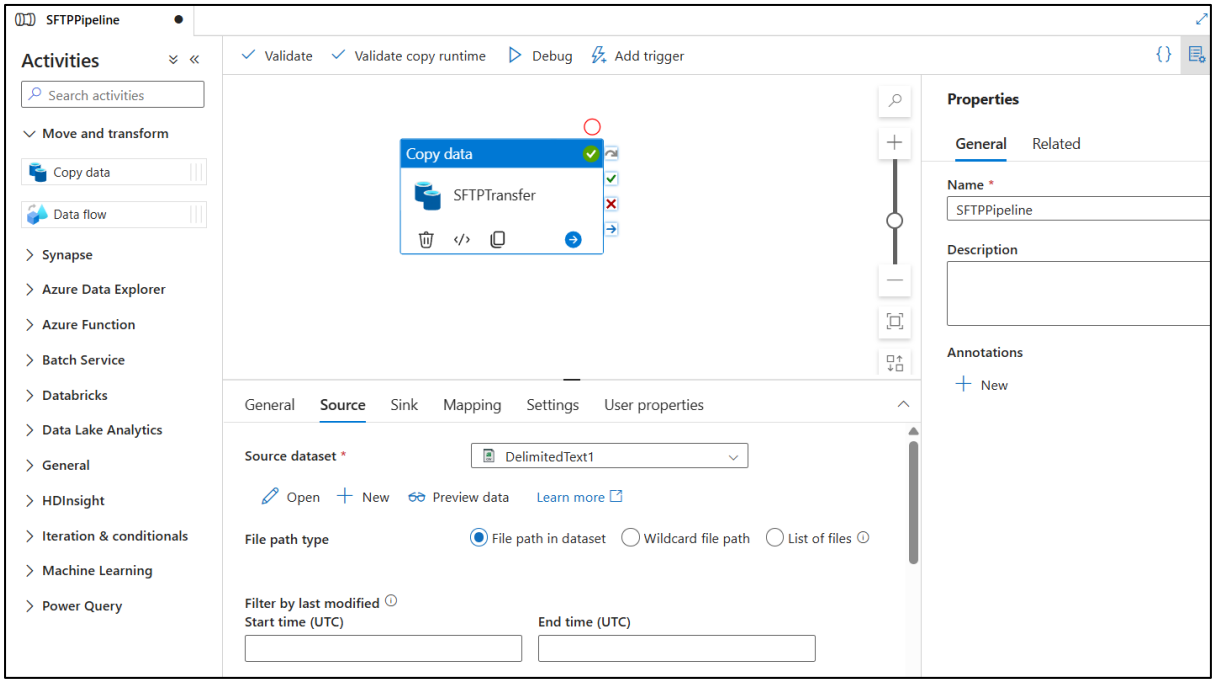
A. Creating the Pipeline

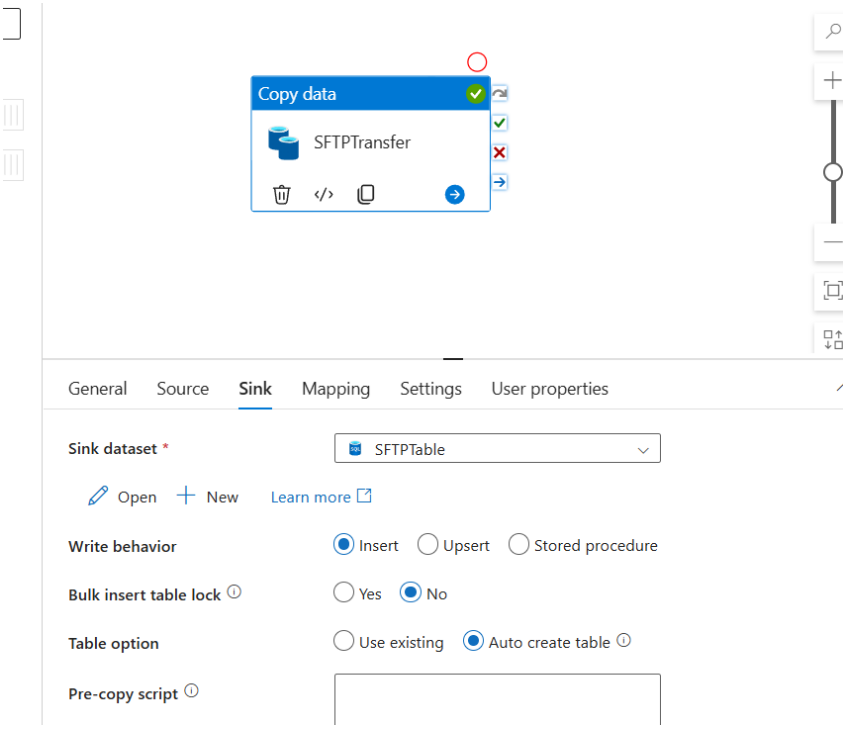
1. In **Author**, click **Pipeline**.
2. Name the pipeline: <pipeline_name>.

B. Add a Copy Data Activity

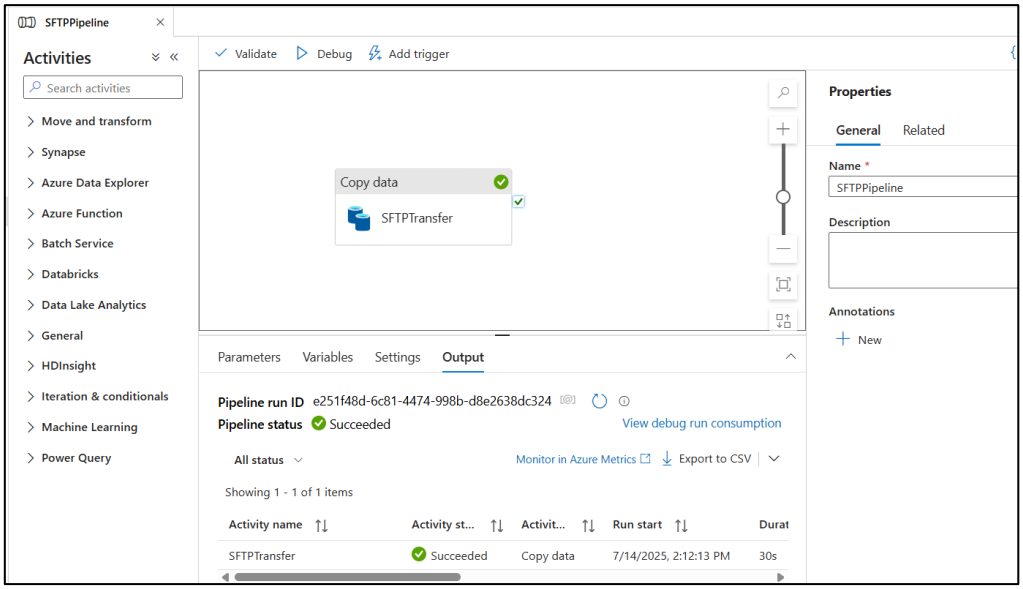
1. Drag **Copy Data** from the Activities pane to the pipeline canvas.
2. Select it and configure:
    - **Source tab**:
        - Dataset: <table_name>
        - Query/table: double-check settings
    - **Sink tab**:

- Dataset: <table_name>
- We can auto-create the table or map columns manually in the **Mapping tab**





TESTING THE PIPELINE



DATA COPIED SUCCESSFULLY IN AZURE SQL