

3_Copy_Data_Between_2_storages

Thursday, February 12, 2026 12:22 PM

Steps to Move CSV Files from One ADLS Folder to Another (Using ADF)

1. Create Linked Services

- o One for **source** ADLS
- o One for **destination** ADLS

The screenshot shows the Microsoft Azure Data Factory interface. The left sidebar navigation menu includes 'Data Factory', 'Connections', and 'Linked services'. The 'Linked services' section is currently selected. It displays two entries: 'AzureDataLakeSink' (Type: Azure Data Lake Storage Gen2) and 'AzureDataLakeSource' (Type: Azure Data Lake Storage Gen2). The top right corner shows the user's email 'saibabburi1999@gmail.com' and a 'DEFAULT DIRECTORY' indicator.

2. Create Datasets

- o Source dataset → DelimitedText → points to source folder

The screenshot shows the configuration for the 'adls_source_dataset' dataset. The 'Connection' tab is selected, showing the 'AzureDataLakeSource' linked service. The 'File path' is set to 'empdata / Directory / employees.csv'. Other settings include 'No compression', 'Comma (,) as column delimiter', 'Default (\r,\n, or \r\n)' as row delimiter, 'Default(UTF-8)' as encoding, 'Double quote (")' as quote character, and 'Backslash (\)' as escape character. The 'First row as header' checkbox is checked. The 'Schema' and 'Parameters' tabs are also visible.

- o Sink dataset → DelimitedText → points to destination folder

The screenshot shows the Azure Data Factory pipeline editor. A pipeline named "copy_adls_source2..." is open. In the top navigation bar, three datasets are listed: "adls_source_dataset", "adls_sink_dataset", and "copy_adls_source2...". The main area displays the configuration for the "copy_adls_source2..." activity. On the left, there's a sidebar with tabs for "Connection", "Schema", and "Parameters". Under "Connection", the "Linked service" is set to "AzureDataLakeSink", "File path" is "emptyout / Directory / File name", and "Compression type" is "No compression". Other settings like "Column delimiter", "Row delimiter", "Encoding", "Quote character", "Escape character", and "First row as header" are also visible. To the right is a "Properties" panel with tabs for "General" and "Related (1)". The "General" tab shows the "Name" as "adls_sink_dataset" and an empty "Description" field. Below the properties are sections for "Annotations" and a "Cancel" button.

3. Add a Copy Activity

- Source: select source dataset

The screenshot shows the Azure Data Factory pipeline editor with a pipeline named "copy_adls_source2...". The left sidebar shows activities like "Move and transform", "Synapse", "Azure Data Explorer", etc. Under "a.", there are sections for "General", "HDInsight", "Iteration & conditionals", "Machine Learning", and "Power Query". The main area shows a "Copy data" activity named "Copy data1". The "Source" tab is selected, showing "Source dataset" as "adls_source_dataset", "File path type" as "File path in dataset" (selected), and "Filter by last modified" and "Recursively" options. The "Sink" tab is the active tab, showing "Sink dataset" as "adls_sink_dataset", "Copy behavior" as "Select...", and other sink-related settings like "Max concurrent connections", "Block size (MB)", "Metadata", "Quote all text", and "File extension".

- Sink: select sink dataset

This screenshot is identical to the previous one, showing the configuration for the "copy_adls_source2..." activity. It highlights the "Sink" tab and the configuration for the "adls_sink_dataset" sink, including "Copy behavior", "Max concurrent connections", "Block size (MB)", "Metadata", "Quote all text", and "File extension".

4. Publish and Trigger the Pipeline

- Run the pipeline
- Monitor execution in ADF Monitor tab

The screenshot shows the Azure Data Factory pipeline editor interface. On the left, a sidebar titled "Activities" lists various data integration services: Move and transform, Synapse, Azure Data Explorer, Azure Function, Batch Service, Databricks, Data Lake Analytics, General, HDInsight, Iteration & conditionals, Machine Learning, and Power Query. The main workspace displays a single activity named "Copy data1". Below the activity, the "Output" tab is selected, showing details of the most recent pipeline run. The run information includes:

- Pipeline run ID: 1d7dd27b1-dfa5-4372-8048-dc6f2e
- Pipeline status: Succeeded
- Run start: 2/12/2026, 1:40:56 PM
- Duration: 19s
- Integration runtime: AutoResolveIntegrationRuntime (East US)

At the bottom of the output table, there are navigation arrows for viewing more items.