

ASSIGNMENT – 5 (Azure Databricks)

TRAINEE NAME: Swathi Baskaran

Creating a container in Azure Storage Account

Home > storageaccountsynapsep2

storageaccountsynapsep2 | Containers

Search containers by prefix

Only show active containers

Name	Last modified	Anonymous access level	Lease state
slogs	27/08/2025, 15:28:37	Private	Available
datalake	27/08/2025, 15:29:27	Private	Available
employeeedata	29/08/2025, 12:18:32	Private	Available

Uploading the csv file in the container

Home > storageaccountsynapsep2 | Containers > employeeedata

Overview

Authentication method: Access key (Switch to Microsoft Entra user account)

Search blobs by prefix (case-sensitive)

Only show active objects

Name	Last modified	Access tier	Blob type	Size	Lease state
employees.csv	29/08/2025, 12:21:26	Hot (Inferred)	Block blob	3.69 KiB	Available

Creating a Data Factory in Azure Portal

The screenshot shows the Azure Portal interface for a Data Factory (V2) named 'EmployeesData'. The top navigation bar includes the Microsoft Azure logo, a search bar, and user information. The left sidebar contains a navigation menu with options like Overview, Activity log, Access control (IAM), Tags, Diagnose and solve problems, Resource visualizer, Settings, Getting started, Monitoring, Automation, and Help. The main content area displays the 'Overview' tab for the Data Factory, showing essential details such as Resource group, Status, Location, Subscription, and Subscription ID. Below this, there is a large 'Azure Data Factory Studio' logo and a 'Launch studio' button. At the bottom, there are four tiles for 'Quick Starts', 'Tutorials', 'Template Gallery', and 'Training Modules'. A small note at the bottom left indicates 'Add or remove favorites by pressing Ctrl+Shift+F'.

Click on Ingest in the Home Page

The screenshot shows the 'EmployeesData' Data Factory home page in the Azure Portal. The top navigation bar includes the Microsoft Azure logo, the breadcrumb 'Data Factory > EmployeesData', a search bar, and user information. The left sidebar contains a navigation menu with options like Home, Activity log, Access control (IAM), Tags, Diagnose and solve problems, Resource visualizer, Settings, Getting started, Monitoring, Automation, and Help. The main content area displays the 'Overview' tab for the Data Factory, showing essential details such as Resource group, Status, Location, Subscription, and Subscription ID. Below this, there is a large 'Azure Data Factory Studio' logo and a 'Launch studio' button. At the bottom, there are four tiles for 'Quick Starts', 'Tutorials', 'Template Gallery', and 'Training Modules'. A small note at the bottom left indicates 'Add or remove favorites by pressing Ctrl+Shift+F'.

Click on Built-in copy task

Microsoft Azure | Data Factory | EmployeesData | Search

Copy Data tool

1 Properties
2 Source
3 Destination
4 Settings
5 Review and finish

Use Copy Data Tool to perform a one-time or scheduled data load from over 100 data sources. Follow the wizard experience to specify your data loading settings, and let the Copy Data Tool generate the artifacts for you, including pipelines, datasets, and linked services. [Learn more](#)

Properties

Select copy data task type and configure task schedule

Task type

Built-in copy task

You will get a single pipeline which is capable of smoothly copying data from over 100 different data sources.

Metadata-driven copy task

You will get parameterized pipelines which can read metadata from an external store to load data at a large scale.

You will get single pipeline to quickly copy objects from data source store to destination in a very intuitive manner.

Task cadence or task schedule *

☒ Run once now ☐ Schedule ☐ Tumbling window

< Previous Next > Cancel

Moving to the source data source page

Microsoft Azure | Data Factory | EmployeesData | Search

Copy Data tool

1 Properties
2 Source
3 Dataset
4 Configuration
5 Destination
6 Settings
7 Review and finish

Source data store

Specify the source data store for the copy task. You can use an existing data store connection or specify a new data store.


Source type: All

Connection *: Select... + New connection

< Previous Next > Cancel

Creating a new connection

New connection

 Azure Blob Storage [Learn more](#)

Name *

Description

Connect via integration runtime * ⓘ

☒ AutoResolveIntegrationRuntime

Authentication type

Account key

Connection string **Azure Key Vault**

Account selection method ⓘ

☒ From Azure subscription ☐ Enter manually

Azure subscription ⓘ

Select all


Storage account name *

Additional connection properties

+ New

Create


Back

 Test connection

Cancel

Testing the connection after filling all the details

New connection

 Azure Blob Storage [Learn more](#)

Name *

Description

Connect via integration runtime * ⓘ

☒ AutoResolveIntegrationRuntime

Authentication type

Account key

Connection string **Azure Key Vault**

Account selection method ⓘ

☒ From Azure subscription ☐ Enter manually

Azure subscription ⓘ

MML Learners (2a3c6418-97b9-4d96-a24b-2c2d7633d375)

Storage account name *


hexadatastoragegen3

Additional connection properties

Create

Back

☒ Connection successful

 Test connection

Cancel

Filling the other details on Source Data Source Page

Microsoft Azure | Data Factory | EmployeeData

Copy Data tool

Source data store

Specify the source data store for the copy task. You can use an existing data store connection or specify a new data store.

Source type: All

Connection: EmployeeBS

File or folder: employeeData/input/employees.csv

Options:

- ☒ Binary copy
- Compression type: None
- ☒ Recursively
- ☐ Delete files after completion

Max concurrent connections: 1

Filter by last modified

Start time (UTC): End time (UTC)

< Previous Next > Cancel

Filling the details on the Destination Data Source Page

Microsoft Azure | Data Factory | EmployeeData

Copy Data tool

Destination data store

Specify the destination data store for the copy task. You can use an existing data store connection or specify a new data store.

Destination type: All

Connection: EmployeeBS

Folder path: employeeData/output

File name:

Compression type: None

Copy behavior: Select...

Max concurrent connections: 1

Block size (MB): 1

< Previous Next > Cancel

Validating the Process

Microsoft Azure | Data Factory | EmployeeData

Copy Data tool

Validating

Deployment step: Validating copy runtime environment

Status: In progress

Skip validation

Completion of Deployment

Microsoft Azure | Data Factory | EmployeesData

Copy Data tool

Deployment complete

Deployment step	Status
Validating copy runtime environment	Succeeded
> Creating datasets	Succeeded
> Creating pipelines	Succeeded
> Running pipelines	Succeeded

Datasets and pipelines have been created. You can now monitor and edit the copy pipelines or click finish to close Copy Data Tool.

[Finish](#) [Edit pipeline](#) [Monitor](#)

Monitoring the execution of Pipeline

Microsoft Azure | Data Factory | EmployeesData

Pipeline runs

Triggered Debug Rerun Cancel options Refresh Edit columns List Gantt

Filter by run ID or name Chennai, Kolkata, Ma... Last 24 hours Pipeline name: PL_EmployeesCopy Status: All

Runs: Latest runs Triggered by: All Add filter

Showing 1 - 1 items

Pipeline name	Run start	Run end	Duration	Triggered by	Status	Run	Parameters
PL_EmployeesCopy	8/29/2025, 12:34:10 PM	8/29/2025, 12:34:35 PM	26s	Manual trigger	Succeeded	Original	

Last refreshed 0 minutes ago

Viewing the pipeline run

Microsoft Azure | Data Factory | EmployeesData

All pipeline runs > PL_EmployeesCopy - Activity runs

Rerun Cancel Refresh Update pipeline List Gantt

Copy data

Copy_to1

Activity runs

Pipeline run ID: 5e5520ba-8f30-42ce-a363-22b78770dcaa

All status Monitor in Azure Metrics Export to CSV

Showing 1 - 1 items

Activity name	Activity st...	Activit...	Run start	Duration	Integration runtime	User prop...	Destination
Copy_to1	Succeeded	Copy data	8/29/2025, 12:34:19 PM	16s	AutoResolveIntegrationRuntime (Central India)		employeeData/oupu

Viewing the output generated in the container

Microsoft Azure

Search resources, services, and docs (G+V)

Copilot

azuser4030.mml.local@TECHADEMY LEARNING SOLUTIONS

Home > storageaccountsynapse2 | Containers >

employeeedata

Container

Search

+ Add Directory

Upload

Refresh

Delete

Copy

Paste

Rename

Acquire lease

Break lease

Edit columns

Overview

Diagnose and solve problems

Access Control (IAM)

Settings

Shared access tokens

Manage ACL

Access policy

Properties

Metadata

employeeedata > output

Authentication method: Access key (Switch to Microsoft Entra user account)

Search blobs by prefix (case-sensitive)

Only show active objects

Showing all 1 items

<input type="checkbox"/>	Name	Last modified	Access tier	Blob type	Size	Lease state
<input type="checkbox"/>	[-]					...
<input type="checkbox"/>	employees.csv	29/08/2025, 12:34:31	Hot (Inferred)	Block blob	3.69 KiB	Available

Add or remove favorites by pressing Ctrl+Shift+F