Hands-On Lab 1: Join Transformation in Azure Data Factory

Lab Created by : Dr. Sandeep Kumar Sharma

What This Lab Will Do

This very simple lab will show you **only one concept**: how to perform a **Join transformation** in Azure Data Factory Mapping Data Flows.
You will join two small CSV files — **Orders** and **Customers** — based on `customer_id` and produce a combined enriched output.

Step 1 — Prepare Storage and Upload Files

Upload these two files into a folder named `lab1/` in your Storage Account.

1. orders.csv

```
order_id,customer_id,product,quantity
1,101,Keyboard,2
2,102,Mouse,1
3,101,Monitor,1
```

2. customers.csv

```
customer_id,customer_name,customer_tier
101,Sandeep,Gold
102,varun,Silver
```

Step 2 — Create Linked Service in ADF

1. Open **Azure Data Factory → Author & Monitor**.
2. Go to **Manage → Linked Services → New**.
3. Choose **Azure Blob Storage / ADLS Gen2**.
4. Authenticate and test the connection.

Step 3 — Create Datasets

Create two datasets:

- `ds_orders_lab1` → points to `lab1/orders.csv`
- `ds_customers_lab1` → points to `lab1/customers.csv`

Use **Delimited Text**, First Row Header = True.

Step 4 — Build the Data Flow

1. Go to **Data Flows → New Mapping Data Flow**.
2. Name it: **df_join_lab1**.

Add Sources

- Source 1: `src_orders` → dataset: `ds_orders_lab1`
- Source 2: `src_customers` → dataset: `ds_customers_lab1`

Add Join Transformation

1. From `src_orders`, add **Join**.
2. Join settings:
3. Left stream: `src_orders`
4. Right stream: `src_customers`
5. Join type: **Inner Join**
6. Condition: `src_orders.customer_id == src_customers.customer_id`

You will now get combined data: orders + customer_name + customer_tier.

Step 5 — Add Sink

1. Add **Sink**.
2. Create a new dataset: `ds_output_lab1` → folder `lab1/output/join/`.
3. Map all fields.

Step 6 — Debug & Run

1. Enable **Data Flow Debug**.
2. Preview output in the Join transformation.
3. Publish and **Trigger Now**.

Output should look like:

```
order_id,customer_id,product,quantity,customer_name,customer_tier
1,101,Keyboard,2,Sandeep,Gold
2,102,Mouse,1,varun,Silver
3,101,Monitor,1,Sandeep,Gold
```

Lab Completed

This lab covered ONLY the **Join transformation**, nothing else.