

# Unable to Load CSV with duplicate columns

Last updated by | Veena Pachauri | Mar 8, 2023 at 11:22 PM PST

Contents

- [Issue](#)
- [Symptoms](#)
- [Cause](#)
- [Resolution](#)

## Issue

Customer getting below exception:

```
{
  "errorCode": "2200",
  "message": "ErrorCode=UserErrorInvalidColumnMappingColumnNotFound,'Type=Microsoft.DataTransfer.Common.Shared.HybridDeliveryException,Message=Column 'EmpID0',",
  "failureType": "UserError",
  "target": "Copy_data1",
  "details": []
}
```

## Symptoms

When user trying to upload a CSV file to the database using copy-activity, The CSV data has some columns duplicated and used don't have control over the column names.

In copy activity, user set the source dataset as the CSV file (delimited text dataset) and sink dataset as Azure SQLDB table with all columns NVARCHAR(MAX) type.

When user import schema in source dataset, it sources them with the column number prefix (for the duplicated ones). For eg, if Column EmpID features twice, on position 0 and 6, then it shows as **'EmpID0'** and **'EmpID6'**.

Example: Below Screenshot explains

Preview data

Linked service: AzureBlobStorage3  
Object: Sankalp sample.csv

EmpID0	Name1	Email	Number	Name4	Dept	EmpID6
1	Mark	m@bing.com	123	Smith	Finance	Smith
2	Chris	c@bing.com	123	Smith	Finance	Smith

When using this schema and user manually mapping the columns to the schema in the Azure Table, it seems fine. For eg, **'EmpID0'** map column to column **'EmpID'** in the Azure Table.

Import schemas Preview source + New mapping Clear Delete

Source	Type	Destination
EmpID0	abc String	EmpID (Int32)
Name1	abc String	Name (String)
Email	abc String	Email (String)
Number	abc String	Number (Int32)
Dept	abc String	Dept (String)

But, when user run it, ADF throws a below error

```
{
  "errorCode": "2200",
  "message": "ErrorCode=UserErrorInvalidColumnMappingColumnNotFound,'Type=Microsoft.DataTransfer.Common.Shared.HybridDeliveryException,Message=Column 'EmpID0',",
  "failureType": "UserError",
  "target": "Copy_data1",
  "details": []
}
```

## Cause

Source dataset CSV have duplicate column names

## Resolution

CSS tried again using sample of data with the help of below ICM, CSS able to copy data successfully even if data have duplicate columns names

If data have duplicate column names, user can follow the below steps to copy.

1. User has to uncheck "First row as header" in source dataset. and set "Skip line count" as 1 in copy source

The screenshot shows the 'Connection' tab of the Azure Data Factory interface. The 'First row as header' checkbox is checked and circled in red. The 'Skip line count' field is set to 1 and also circled in red. Other settings include 'Column delimiter' as Comma (,), 'Row delimiter' as Auto detect (\r,\n, or \r\n), 'Encoding' as Default(UTF-8), 'Escape character' as Backslash (\), 'Quote character' as Double quote ("), and 'Null value' as an empty field.

2. In mapping page, do column mapping using column number.

The screenshot shows the 'Mapping' tab of the Azure Data Factory interface. It displays a table with 5 source columns (all 'abc' type) mapped to 5 destination columns (EmpID, Name, Email, Number, Dept). The mapping is as follows:

Source (Column number)	Type	Destination
1	abc: String	EmpID (int32)
2	abc: String	Name (String)
3	abc: String	Email (String)
4	abc: String	Number (int32)
5	abc: String	Dept (String)

## Additional Information:

- Icm References: [Icm Link](#)

How good have you found this content?

