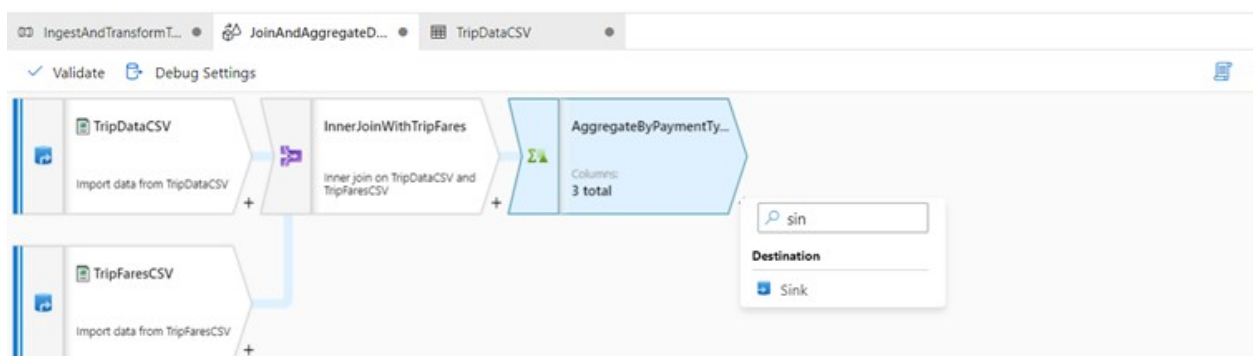


# Exercise: Sink dataset into Azure Synapse Analytics with Azure Data Factory

## Configure Azure Synapse Analytics sink

We have finished our transformation logic in the previous exercise, so now we are ready to sink our data in an Azure Synapse Analytics table.

Add a sink transformation under the Destination section.



Add Sink Transformation

Name your sink 'SynapseSink'.

Click New next to the sink dataset field to create a new Azure Synapse Analytics dataset.

Sink
Settings
Mapping
Optimize
Inspect
Data preview

Output stream name \*
[Learn more](#)

Incoming stream \*

Sink type \*

Dataset \*
+ New

Options
☒ Allow schema drift ⓘ
☐ Validate schema ⓘ

Add dataset to Sink Transformation

**Select the Azure Synapse Analytics (formerly SQL DW) tile and click continue.**

Validate all
Refresh
Discard all
Data flow debug
ARM template

IngestAndTransformT...
JoinAndAggregateD...
TripDataCSV

Validate
Debug Settings

TripDataCSV
Import data from TripDataCSV

InnerJoinWithTripFares
Inner join on TripDataCSV and TripFaresCSV

Aggregatel
Aggregating i payment typ columns 'avel total\_trip\_dist

TripFaresCSV
Import data from TripFaresCSV

Sink
Settings
Mapping
Optimize
Inspect
Data preview

Output stream name \*
[Learn more](#)

Incoming stream \*

Sink type \*

Dataset \*
+ New


Options
☒ Allow schema drift ⓘ
☐ Validate schema ⓘ

New dataset

In pipeline activities and data flows, reference a dataset to specify the location and structure of your data within a data store. [Learn more](#)

Select a data store

All
Azure
Database
File
Generic protocol
NoSQL
Services and apps



Azure Synapse Analytics (formerly SQL DW)

Continue
Cancel

Select Azure Synapse Analytics for Sink Transformations

**Call your dataset 'AggregatedTaxiData'.**

**Select 'Synapse' as your linked service.**

Select Create new table and name the new table dbo.AggregateTaxiData.

Click OK when finished

### Set properties

Name

AggregatedTaxiData

Linked service \*

Synapse



☐ Select from existing table ☒ Create new table

Schema and table name

dbo

. AggregatedTaxiData

▸ Advanced

OK

Back

Cancel

Set properties for Azure Synapse Analytics dataset

**After creating your dataset, there are no further changes you need to make your sink!**

**In the Settings tab, you will notice that by default the only update method selected is Allow insert, which means all rows will be inserted to the table.**

**In conjunction with an alter row transformation, you can tag certain rows as delete, upsert, or update as well.**

**Enable staging is also enabled by default, which allows for the data flow to utilize Synapse's PolyBase technology for performant loading.**

The screenshot displays the Synapse Data Flow Designer interface. At the top, a breadcrumb trail shows the path: IngestAndTransformT... > JoinAndAggregateD... > TripDataCSV. Below this, a toolbar includes 'Validate' and 'Debug Settings' buttons. The main canvas shows a data flow diagram with four components: a 'Reference' component (1 row, 14 columns), an 'InnerJoinWithTripFares' component (inner join on TripDataCSV and TripFaresCSV), an 'AggregateByPaymentTy...' component (aggregating data by 'payment\_type' producing columns: 'average\_fare', 'total\_trip\_distance'), and a 'SynapseSink' component (3 columns total). A 'TripFaresCSV' component (import data from TripFaresCSV) is also shown. The bottom panel is the 'Settings' tab for the 'SynapseSink' component. It includes fields for 'Output stream name' (SynapseSink), 'Incoming stream' (AggregateByPaymentType), 'Sink type' (Dataset), and 'Dataset' (AggregatedTaxiData). There are also buttons for 'Test connection', 'Open', and 'New'. Under 'Options', the 'Allow schema drift' checkbox is checked, and 'Validate schema' is unchecked.

Overview Synapse Sink

**You have successfully created your data flow.**