|  |  |  |
| --- | --- | --- |
|  |  | Azure Data Ecosystem  Automating ELT with Data Bricks, Data Lake & Data Factory to SQL DB/DW. |

# Description

Most of the data movement and data migration requires a lot of manual intervention and it requires a lot of engineering effort in developing and designing the solutions which can handle data from multiple sources and the data can be of any type . This tutorial considers automating 90% of the requirement of ELT / ETL from any sources and it can be implemented for any requirement.



### Part -2:

### Copying Data from Landing Data Lake to Output Data Lake Folder using Data Bricks.

After completing the copy activity from Blob to Landing Zone in Data Lake. The next most important thing is to utilize the power of data bricks to apply transformation and compression techniques to make it the best way to process and store the data into Output Zone in Data Lake. Data bricks is not part of this use case scope but this is the place where we utilize data bricks or Data Lake Analytics for data processing

1. Next step is to open the Data factory pipeline and drag and drop the Copy Activity as below and give a meaningful name:

A screenshot of a cell phone

Description automatically generated

1. If observed we have 2 important and mandatory tabs to be updated . Lets update those fields which require Source & Sink Paths.

Source WildCard FileName : @{concat(item().FileName,'\_',item().EmpId,'.csv')}

A screenshot of a cell phone

Description automatically generated

Create a new dataset and add Sink FilePath : @concat(item().OUTPUTPATH,'/',item().EmpId,item().DatePath),

@concat(item().FileName,'\_',item().EmpId,'.csv')

A screenshot of a social media post

Description automatically generated

Make Header checkbox enabled :

A screenshot of a cell phone

Description automatically generated

1. After successfully executing all the steps you should be able to see the file in the Output Zone in data lake.
2. Now we can define the Copy Activity to copy the data from Data Lake Output Zone to SQL DB using copy activity.

A screenshot of a cell phone

Description automatically generated

1. Source Connection in CopyActivity :

A screenshot of a cell phone

Description automatically generated

1. Create Sink Connection Dataset for Copy Activity :

TableName: @concat(item().OUTPUTPATH,'/',item().EmpId,item().DatePath)

A screenshot of a cell phone

Description automatically generated

A screenshot of a cell phone

Description automatically generated

1. Publish after creating the copy activity .
2. Trigger the pipeline to execute the whole pipeline .

A screenshot of a cell phone

Description automatically generated

1. Go to SQL DB and see if the data is present in all the tables.

Awesome !!! .

You have automated the complete ETL/ELT Processing. The same approach can be implemented for any customer or any requirement for ETL Processing and can also include Data Bricks and Data Lake Analytics activities which we can see in next Hands on Labs.