



Loading data into a Table Using COPY Command

1. In this lab, we are going to create a persisted table for our CSV and Parquet files where the data will be persisted inside our tables. The data will not be fetched from the data lake it will be stored inside the table.
2. In the last lab, we created a table by copying everything from an external table using the Polybase function but in this lab first we will create a table and then we will copy the data from data lake into the table using the COPY command.
3. First, we will create a new SQL Script and then we will use the commands to create a table. In the 1st step, we will drop the table because we will use the same name for the table. Also change the database to Dedicated SQL Pool.

Pool-CSV Table

Run Undo Publish Query plan Connect to DataPool

```
1 DROP TABLE PoolActivityLog;
2
3 CREATE TABLE PoolActivityLog
4 (
5     [Correlationid] varchar(200),
6     [Operationname] varchar(300),
7     [Status] varchar(100),
8     [Eventcategory] varchar(100),
9     [Level] varchar(100),
10    [Time] varchar(100),
11    [Subscription] varchar(200),
12    [Eventinitiatedby] varchar(1000),
13    [Resourcetype] varchar(300),
14    [Resourcegroup] varchar(1000),
15    [Resource] varchar(2000)
16 )
17 WITH(
18     DISTRIBUTION=ROUND_ROBIN
19 )
```

4. Then we will create our table whose name is Pool Activity Log. Also, if you pay attention, you will see that we are creating a normal table with a round-robin distribution.

5. Now we will use the COPY command to copy the data from our CSV file to the table. Here we are using the Link of our CSV File and the SAS Token.

```
21 COPY INTO PoolActivityLog
22 FROM 'https://thestorageaccount1201.blob.core.windows.net/data/ActivityLog01.csv'
23 WITH(
24     FILE_TYPE='CSV',
25     FIRSTROW=2,
26     CREDENTIAL=
27     (
28         IDENTITY='Shared Access Signature',
29         SECRET='sv=2022-11-02&ss=b&srt=sco&sp=r1&se=2024-12-24T14:53:02Z&st=2024-12-24T06:53:02Z'
30     )
31 );
```

6. We will see the data in our persisted table if we run the Select query.

```
33 SELECT * FROM PoolActivityLog
34
```

Results Messages

View Table Chart Export results

Search

Correlationid	Operationname	Status	Eventcategory	Level	Time
c40ce711-5d0f-...	Sync Web Apps...	Started	Administrative	Informational	2024-05-15T12:...
50591832-2222...	Update website	Started	Administrative	Informational	2024-05-07T10:...
225486ae-cbd9...	Delete Public Ip...	Succeeded	Administrative	Informational	2024-04-30T13:...
42717fb3-4419...	'audit' Policy ac...	Succeeded	Policy	Warning	2024-04-23T04:...
c5cf5ce4-d26f-	Delete Storane	Started	Administrative	Informational	2024-05-30T14...

00:00:05 Query executed successfully.

😄 For the Parquet file

7. The process is the same as before first, we will drop our table and then recreate it because we will use the same name for the table.

Pool-Parquet Table

Run Undo Publish Query plan Connect to DataPool

```
1 DROP TABLE PoolActivityLog;
2
```

8. Then create your table and if we see the data inside the table using the Select query you will see that our table is empty.

```

3  CREATE TABLE PoolActivityLog
4  (
5      [Correlationid] varchar(200),
6      [Operationname] varchar(300),
7      [Status] varchar(100),
8      [Eventcategory] varchar(100),
9      [Level] varchar(100),
10     [Time] varchar(100),
11     [Subscription] varchar(200),
12     [Eventinitiatedby] varchar(1000),
13     [Resourcetype] varchar(300),
14     [Resourcegroup] varchar(1000),
15     [Resource] varchar(2000)
16 )
17 WITH(
18     DISTRIBUTION=ROUND_ROBIN
19 )

```

```

21 SELECT * FROM PoolActivityLog;
22

```

Results Messages

View Table Chart Export results

Search

Correlationid	Operationname	Status	Eventcategory	Level	Time
---------------	---------------	--------	---------------	-------	------

9. Now we will run the COPY command to copy the data from the parquet file to the table.

```

23 COPY INTO PoolActivityLog
24 FROM 'https://thestorageaccount1201.blob.core.windows.net/data/ActivityLog01.parquet'
25 WITH(
26     FILE_TYPE='PARQUET',
27     CREDENTIAL=
28     (
29         IDENTITY='Shared Access Signature',
30         SECRET='sv=2022-11-02&ss=b&srt=sco&sp=r1&se=2024-12-24T14:53:02Z&st=2024-12-24T06:53:02Z'
31     )
32 )

```

10. Now we can see that our table has the data.

```
21 SELECT * FROM PoolActivityLog;
22
```

Results Messages

View

Table

Chart

Export results

Search					
Correlationid	Operationname	Status	Eventcategory	Level	Time
bedc87a0-bce9...	CheckNameAva...	Started	Administrative	Informational	2024-05-10T09:...
0e67cb2c-9ae4...	Delete resource...	Started	Administrative	Informational	2024-05-10T07:...
8502f407-181c...	Validate Deploy...	Started	Administrative	Informational	2024-04-26T06:...
1da5af26-4c2a...	Update website	Started	Administrative	Informational	2024-05-07T11:...
604c28d9-2d95...	List Web Apps ...	Succeeded	Administrative	Informational	2024-05-07T10:...
a0d378b0-4527...	Get Connection...	Started	Administrative	Informational	2024-05-07T09:...

00:00:03 Query executed successfully.

11. Once you are done then you can publish the changes to save everything. Also, if you want to pause the dedicated SQL pool you can also do that.
12. Just go to the Manage tab and click on Pause button to pause your SQL Pool.

The screenshot shows the Synapse live interface. On the left, the 'SQL pools' tab is selected. The main area displays a table of SQL pools. The 'DataPool' is highlighted, and the 'Pause' button is visible next to it.

Name	Type	Status	Size
Built-in	Serverless	Online	Auto
DataPool	Dedicated	Online	DW100c