

Simplify ingestion with the Copy Activity

Note:

You are not required to complete the processes, tasks, activities, or steps presented in this example. The various samples provided are for illustrative purposes only and it's likely that if you try this out you will encounter issues in your system.

The broad capabilities of the Copy Activity allow you to quickly and easily move data into SQL Pools from a variety of sources.

The Copy Activity supports a large range of data sources and sinks on-premises and in the cloud. It facilitates the efficient, yet flexible parsing and transfer of data or files between systems in an optimized fashion as well as giving you capability of easily converting datasets into other formats.

In the following example, you can load data from a public storage account. Here the COPY statement's defaults match the format of the line item csv file.

```
COPY INTO dbo.[lineitem] FROM 'https://unsecureaccount.blob.core.windows.net/customerdatasets/folder1/lineitem.csv'
```

The default values for csv files of the COPY command are:

- DATEFORMAT = Session DATEFORMAT
- MAXERRORS = 0
- COMPRESSION default is uncompressed
- FIELDQUOTE = “”
- FIELDTERMINATOR = “,”
- ROWTERMINATOR = ‘\n’
- FIRSTROW = 1
- ENCODING = ‘UTF8’
- FILE_TYPE = ‘CSV’
- IDENTITY_INSERT = ‘OFF’

This example loads files specifying a column list with default values.

1

1
2
3
4
5
6
7
8

9
10
11
12
13
14

```
--Note when specifying the column list, input field numbers start from 1
COPY INTO test_1 (Col_one default 'myStringDefault' 1, Col_two default 1 3)
FROM 'https://myaccount.blob.core.windows.net/myblobcontainer/folder1/'
WITH (
    FILE_TYPE = 'CSV',
    CREDENTIAL=(IDENTITY= 'Storage Account Key', SECRET='<Your_Account_Key>'),
    --CREDENTIAL should look something like this:
    --CREDENTIAL=(IDENTITY= 'Storage Account Key', SECRET='x6RWv4It5F2msnjelv3H4
DA80n0PQW0daPdw43jM0nyetx4c6CpDkdj3986DX5AHFMIf/YN4y6kkCnU8lb+Wx0Pj+6MDw=='),
    FIELDQUOTE = '',
    FIELDTERMINATOR=',',
    ROWTERMINATOR='0x0A',
    ENCODING = 'UTF8',
    FIRSTROW = 2
)
```

The following example loads files that use the line feed as a row terminator such as a UNIX output. This example also uses a SAS key to authenticate to Azure blob storage.

1
2
3
4
5
6
7
8
9
10
11
12
13
14
15
16

```
COPY INTO test_1
FROM 'https://myaccount.blob.core.windows.net/myblobcontainer/folder1/'
WITH (
    FILE_TYPE = 'CSV',
```

```
CREDENTIAL=(IDENTITY= 'Shared Access Signature', SECRET='<Your_SAS-Token>'),
--CREDENTIAL should look something like this:
--CREDENTIAL=(IDENTITY= 'Shared Access Signature', SECRET='?sv=2018-03-
28&ss=bfqt&srt=sco&sp=rl&st=2016-10-17T20%3A14%3A55Z&se=2021-10-
18T20%3A19%3A00Z&sig=IEoOdmeYnE9%2FKiJDSHFSYsz4AkNa%2F%2BTx61FuQ%2FfKHefqoBE
%3D'),
FIELDQUOTE = '',
FIELDTERMINATOR=';',
ROWTERMINATOR='0X0A',
ENCODING = 'UTF8',
DATEFORMAT = 'ymd',
MAXERRORS = 10,
ERRORFILE = '/errorsfolder',--path starting from the storage container
IDENTITY_INSERT = 'ON'
)
```