-- Lab - Using External tables

-- First we need to create a database in the serverless pool

CREATE DATABASE [appdb]

-- Here we are creating a database master key. This key will be used to protect the Shared Access Signature which is specified in the next step

-- Ensure to switch the context to the new database first

CREATE MASTER KEY ENCRYPTION BY PASSWORD = 'P@ssw0rd@123';

-- Here we are using the Shared Access Signature to authorize the use of the Azure Data Lake Storage account

CREATE DATABASE SCOPED CREDENTIAL SasToken

WITH IDENTITY='SHARED ACCESS SIGNATURE'

, SECRET = 'sv=2021-06-08&ss=b&srt=sco&sp=rlx&se=2022-08-10T23:39:19Z&st=2022-08-10T15:39:19Z&spr=https&sig=gecsUxC6Uo3UQ6r%2FKogsKsfT1PeuEO%2F0%2FdMPOBhEBNQ%3D';

-- This defines the source of the data.

CREATE EXTERNAL DATA SOURCE log\_data

WITH (    LOCATION   = 'https://datalakerojalin.dfs.core.windows.net/data',

          CREDENTIAL = SasToken

)

/\* This creates an External File Format object that defines the external data that can be

present in Hadoop, Azure Blob storage or Azure Data Lake Store

Here with FIRST\_ROW, we are saying please skip the first row because this contains header information

\*/

CREATE EXTERNAL FILE FORMAT TextFileFormat WITH (

      FORMAT\_TYPE = DELIMITEDTEXT,

    FORMAT\_OPTIONS (

        FIELD\_TERMINATOR = ',',

        FIRST\_ROW = 2))

-- Here we define the external table

CREATE EXTERNAL TABLE [logdata]

(

    [Id] [int] NULL,

    [Correlationid] [varchar](200) NULL,

    [Operationname] [varchar](200) NULL,

    [Status] [varchar](100) NULL,

    [Eventcategory] [varchar](100) NULL,

    [Level] [varchar](100) NULL,

    [Time] [datetime] NULL,

    [Subscription] [varchar](200) NULL,

    [Eventinitiatedby] [varchar](1000) NULL,

    [Resourcetype] [varchar](1000) NULL,

    [Resourcegroup] [varchar](1000) NULL)

WITH (

 LOCATION = '/Log.csv',

    DATA\_SOURCE = log\_data,

    FILE\_FORMAT = TextFileFormat

)

-- If you made a mistake with the table, you can drop the table and recreate it again

DROP EXTERNAL TABLE [logdata]

/\*

Common errors

1. External table 'logdata' is not accessible because location does not exist or it is used by another process.

Here your Shared Access Siganture is an issue. Ensure to create the right Shared Access Siganture

2. Msg 16544, Level 16, State 3, Line 34

The maximum reject threshold is reached.

This happens when you try to select the rows of data from the table. This can happen if the rows are not matching the schema defined for the table

\*/

SELECT \* FROM [logdata]

SELECT [Operationname] , COUNT([Operationname]) as [Operation Count]

FROM [logdata]

GROUP BY [Operationname]

ORDER BY [Operation Count]