**Analyze data with a serverless SQL pool**

**The Built-in serverless SQL pool**

Serverless SQL pools let you use SQL without having to reserve capacity. Billing for a serverless SQL pool is based on the amount of data processed to run the query and not the number of nodes used to run the query.

Every workspace comes with a pre-configured serverless SQL pool called **Built-in**.

**Analyze NYC Taxi data with a serverless SQL pool**

**Note**

Make sure you have **placed the sample data into the primary storage account**

1. In Synapse Studio, go to the **Develop** hub
2. Create a new SQL script.
3. Paste the following code into the script.

SQLCopy

SELECT

TOP 100 \*

FROM

OPENROWSET(

BULK 'https://ckctestdbmount1.dfs.core.windows.net/demo/NYCTripSmall.parquet',

FORMAT='PARQUET'

) AS [result]

1. Click **Run**.

Data exploration is just a simplified scenario where you can understand the basic characteristics of your data.

**Create data exploration database**

You can browse the content of the files directly via master database. For some simple data exploration scenarios, you don't need to create a separate database. However, as you continue data exploration, you might want to create some utility objects, such as:

* External data sources that represent the named references for storage accounts.
* Database scoped credentials that enable you to specify how to authenticate to external data source.
* Database users with the permissions to access some data sources or database objects.
* Utility views, procedures, and functions that you can use in the queries.

1. Use the master database to create a separate database for custom database objects. Custom database objects, cannot be created in the master database.

SQLCopy

CREATE DATABASE DataExplorationDB

COLLATE Latin1\_General\_100\_BIN2\_UTF8

**Important**

Use a collation with \_UTF8 suffix to ensure that UTF-8 text is properly converted to VARCHAR columns. Latin1\_General\_100\_BIN2\_UTF8 provides the best performance in the queries that read data from Parquet files and cosmos Db containers.

1. Switch from master to DataExplorationDB using the following command. You can also use the UI control **use database** to switch your current database:

SQLCopy

USE DataExplorationDB

1. From the 'DataExplorationDB', create utility objects such as credentials and data sources.

SQLCopy

CREATE EXTERNAL DATA SOURCE ContosoLake

WITH ( LOCATION = ‘https://ckctestdbmount1.dfs.core.windows.net/’)

**Note**

An external data source can be created without a credential. If a credential does not exist, the caller's identity will be used to access the external data source.

1. Optionally, use the newly created 'DataExplorationDB' database to create a login for a user in DataExplorationDB that will access external data:

SQLCopy

CREATE LOGIN data\_explorer WITH PASSWORD = 'My Very Strong Password 1234!';

Next create a database user in 'DataExplorationDB' for the above login and grant the ADMINISTER DATABASE BULK OPERATIONS permission.

SQLCopy

CREATE USER data\_explorer FOR LOGIN data\_explorer;

GO

GRANT ADMINISTER DATABASE BULK OPERATIONS TO data\_explorer;

GO

1. Explore the content of the file using the relative path and the data source:

SQLCopy

SELECT

TOP 100 \*

FROM

OPENROWSET(

BULK '/users/NYCTripSmall.parquet',

DATA\_SOURCE = 'ContosoLake',

FORMAT='PARQUET'

) AS [result]

Data exploration database is just a simple placeholder where you can store your utility objects. Synapse SQL pool enables you to do much more and create a Logical Data Warehouse - a relational layer built on top of Azure data sources