

QUERY FILES USING A SERVERLESS SQL POOL

BY NITIN

OBJECTIVES

- Learned how to query data files in a serverless SQL pool.
 - Learned how to use SQL to query CSV files.
 - Learned how to use SQL to query parquet file
 - Learned how to use SQL to query JSON file
-

labs.xtremelabs.io/LabViewerConnection/DetachLabMa... Microsoft Azure Portal | Microsoft Home - Microsoft Azure Instance ID:6731016 labs.xtremelabs.io/LabViewerConnection/DetachLabManual?labl... portal.azure.com/#home Microsoft Azure Search resources, services, and docs (G+) Copilot ...

Instance ID:6731016
XtremeLabs: DP-203T00-A-CEP [DP-203T00-A-M02-CEP] Module 02: Query files using a serverless SQL pool

PowerShell Requesting a Cloud Shell. Succeeded. Connecting terminal... Welcome to Azure Cloud Shell Type "az" to use Azure CLI Type "help" to learn about Cloud Shell MOTD: To Connect and Manage Exchange Online; Connect-EXOPSSession VERBOSE: Authenticating to Azure ... VERBOSE: Building your Azure drive ... PS /home/xlab-rpm-910> []

3. Note that you can resize the cloud shell by dragging the separator bar at the top of the pane, or by using the —, □, and X icons at the top right of the pane to minimize, maximize, and close the pane. For more information about using the Azure Cloud Shell, see the [Azure Cloud Shell documentation](#).

4. In the PowerShell pane, manually enter the following commands to clone this repo:

Paste Content Paste Content

```
rm -r dp-203 -f
git clone https://github.com/MicrosoftLearning/dp-203-azure-data-engineer dp-203
```

5. After the repo has been cloned, enter the following commands to change to the folder for this lab and run the **setup.ps1** script it contains:

Paste Content Paste Content

Page: 3/11

MOTD: SqlServer has been updated to Version 22!

VERBOSE: Authenticating to Azure ...
VERBOSE: Building your Azure drive ...

```
PS /home/xlab-rpm-910> rm -r dp-203 -f
PS /home/xlab-rpm-910> git clone https://github.com/MicrosoftLearning/dp-203-azure-data-engineer dp-203
Cloning into 'dp-203'...
remote: Enumerating objects: 1506, done.
remote: Counting objects: 100% (179/179), done.
remote: Compressing objects: 100% (88/88), done.
remote: Total 1506 (delta 88), reused 120 (delta 57), pack-reused 1327 (from 1)
Receiving objects: 100% (1506/1506), 28.73 MiB | 20.68 MiB/s, done.
Resolving deltas: 100% (784/784), done.
PS /home/xlab-rpm-910> [ ]
```

Microsoft Azure Search resources, services, and docs (G+) Copilot ...

Azure services

Create a resource Quickstart Center Azure AI services Kubernetes services Virtual machines App Services Storage accounts SQL databases

Azure Cosmos DB More services

Resources

Switch to Bash Restart Manage files New session Editor Web preview Settings Help

ENG IN 2:12 PM 10/22/2024

labs.xtremelabs.io/LabViewerConnection/DetachLabMa... Microsoft Azure Portal | Microsoft Home - Microsoft Azure portal.azure.com/#home Microsoft Azure Search resources, services, and docs (G+) Copilot ...

Instance ID:6731016
XtremeLabs: DP-203T00-A-CEP [DP-203T00-A-M02-CEP] Module 02: Query files using a serverless SQL pool

following commands to change to the folder for this lab and run the **setup.ps1** script it contains:

Paste Content Paste Content

```
cd dp-203/Allfiles/labs/02  
../setup.ps1
```

6. If prompted, choose which subscription you want to use (this will only happen if you have access to multiple Azure subscriptions).

7. When prompted, enter a suitable password to be set for your Azure Synapse SQL pool.

Note: Be sure to remember this password!

8. Wait for the script to complete - this typically takes around 10 minutes, but in some cases may take longer. While you are waiting, review the [Serverless SQL pool in Azure Synapse Analytics](#) article in the Azure Synapse Analytics documentation.

< >

Page: 3/11

Support

Microsoft Azure Services

Create a resource Quickstart Center Azure AI services Kubernetes services Virtual machines App Services Storage accounts SQL databases

Switch to Bash Restart Manage files New session Editor Web preview Settings Help

S043705.json

```
ICloudBlob : Microsoft.Azure.Storage.Blob.CloudBlockBlob
BlobType : BlockBlob
Length : 284
IsDeleted : False
BlobClient : Azure.Storage.Blobs.BlobClient
BlobBaseClient : Azure.Storage.Blobs.Specialized.BlockBlobClient
BlobProperties : Azure.Storage.Blobs.Models.BlobProperties
RemainingDaysBeforePermanentDelete :
ContentType : application/octet-stream
LastModified : 10/22/2024 8:18:45 PM +00:00
SnapshotTime :
ContinuationToken :
VersionId :
IsLatestVersion :
AccessTier : Hot
TagCount : 0
Tags :
ListBlobProperties :
Context : Microsoft.WindowsAzure.Commands.Storage.Common.AzureStorageContext
Name : sales/json/S043705.json
```

Script completed at 10/22/2024 20:18:46
PS /home/xlab-rpm-910/dp-203/Allfiles/labs/02>

2:21 PM 10/22/2024 ENG IN

Instance ID:6731016
XtremeLabs: DP-203T00-A-CEP [DP-203T00-A-M02-CEP] Module 02: Query files using a serverless SQL pool

1. After the script has completed, in the Azure portal, go to the **dp203-xxxxxx** resource group that it created, and select your Synapse workspace.

2. In the **Overview** page for your Synapse workspace, in the **Open Synapse Studio** card, select **Open** to open Synapse Studio in a new browser tab; signing in if prompted.

3. On the left side of Synapse Studio, use the **>>** icon to expand the menu - this reveals the different pages within Synapse Studio that you'll use to manage resources and perform data analytics tasks.

4. On the **Data** page, view the **Linked** tab and verify that your workspace includes a link to your Azure Data Lake Storage Gen2 storage account, which should have a name similar to **synapsexxxxxxxx (Primary) - datalakexxxxxxxx**.

5. Expand your storage account and verify that it contains a file system container named **files**.

6. Select the **files** container, and note that it contains a folder named **sales**. This folder contains the data files you are going to query.

Page: 4/11

Support

Microsoft Azure | Synapse Analytics | web.azure-synapse.net/en/authoring/explore/linked?workspace=%2Fsubscriptions%2F4d335bb4-c496-4c... | Microsoft Azure Portal | Resource groups - Microsoft | synapse2wifjt - Microsoft | synapse2wifjt - Azure Search | + | Microsoft Edge | Search | Home | Back | Forward | Home | Refresh | Help | Settings | ? | More options | XLab-rPM-910@xtemelabs.us | LAB DIRECTORY

Microsoft Azure | Synapse Analytics | web.azure-synapse.net/en/authoring/explore/linked?workspace=%2Fsubscriptions%2F4d335bb4-c496-4c... | Microsoft Azure Portal | Resource groups - Microsoft | synapse2wifjt - Microsoft | synapse2wifjt - Azure Search | + | Microsoft Edge | Search | Home | Back | Forward | Home | Refresh | Help | Settings | ? | More options | XLab-rPM-910@xtemelabs.us | LAB DIRECTORY

We use optional cookies to provide a better experience. [Learn more](#)

Synapse live | Validate all | Publish all

Home | Data | Develop | Integrate | Monitor | Manage

Data | Workspace | Linked | Filter resources by name | Azure Data Lake Storage Gen2 | 2

Select an item

Use the resource explorer to select or create a new item

2:24 PM 10/22/2024

Instance ID:6731016
XtremeLabs: DP-203T00-A-CEP [DP-203T00-A-M02-CEP] Module 02: Query files using a serverless SQL pool

7. Open the **sales** folder and the **csv** folder it contains, and observe that this folder contains .csv files for three years of sales data.

8. Right-click any of the files and select **Preview** to see the data it contains. Note that the files do not contain a header row, so you can unselect the option to display column headers.

9. Close the preview, and then use the ↑ button to navigate back to the **sales** folder.

10. In the **sales** folder, open the **json** folder and observe that it contains some sample sales orders in .json files. Preview any of these files to see the JSON format used for a sales order.

11. Close the preview, and then use the ↑ button to navigate back to the **sales** folder.

12. In the sales folder, open the **parquet** folder and observe that it contains a subfolder for each year (2019-2021), in each of which a file named **orders.snappy.parquet** contains the order data for that year.

13. Return to the **sales** folder so you can see the **csv**, **json**, and **parquet** folders.

Page: 4/11

Support

Microsoft Azure | Synapse Analytics | Search | Accept | Reject | More options

We use optional cookies to provide a better experience. Learn more

Synapse live | Validate all | Publish all

Data | Workspace | Linked | Filter resources by name

Azure Data Lake Storage Gen2 | synapsep2wifjt (Primary - datalake...) | files (Primary) | (Attached Containers)

files

New SQL script | New data flow | New integration dataset | More

files > sales

Name	Last Modified	Content Type	Size
csv	10/22/2024, 2:18:39 PM	Folder	
json	10/22/2024, 2:18:43 PM	Folder	
parquet	10/22/2024, 2:18:42 PM	Folder	

Showing 1 to 3 of 3 cached items

2:33 PM 10/22/2024

Instance ID:6731016
XtremeLabs: DP-203T00-A-CEP [DP-203T00-A-M02-CEP] Module 02: Query files using a serverless SQL pool

```
TOP 100 *
FROM
OPENROWSET(
    BULK 'https://datalakexxxxxxx.dfs.core.windows.net/files/sales/csv'
    FORMAT = 'CSV',
    PARSE_VERSION='2.0'
) AS [result]
```

This code uses the OPENROWSET to read data from the CSV files in the sales folder and retrieves the first 100 rows.

- 5. In the **Connect to** list, ensure **Built-in** is selected - this represents the built-in connection to the workspace.
- 6. On the toolbar, use the **Run** button to run the SQL code, and review the results.

C1	C2	C3	C4	C5	C6
SO49171	1	2021-01-01	Mariah Foster	mariah21@adventure-works.com	Road-250 Black, 48
...

7. Note the results consist of columns named C1, C2, and so on. In this example it's possible to work with the data using the generic column names that have been generated. To understand the data if you define a tabular schema. To accomplish this, add the schema here (replacing datalakexxxxxxx with the name of your data lake storage account).

Paste Content

Paste Content

```
SELECT
TOP 100 *
FROM
OPENROWSET(
    BULK 'https://datalakexxxxxxx.dfs.core.windows.net/files/sales/csv'
    FORMAT = 'CSV'.
```

Page: 5/11

Support

Microsoft Azure | Resource groups - | synapsep2wifjt - M | synapsep2wifjt - A | +

web.azure-synapse.net/en/authoring/explore/linked/sqlscripts/SQL%20script%201... | XLab-rPM-910@xtemelabs.us LAB DIRECTORY

synapsep2wifjt Search

We use optional cookies to provide a better experience. Learn more

Synapse live Validate all Publish all 1

files Sales CSV query

Other users in your workspace may have access to modify this item. Do not use this item unless you trust all users who may have access to the workspace.

Run Undo Publish Query plan Connect to Built-in ...

```
1 -- This is auto-generated code
2 SELECT
3 TOP 100 *
4 FROM
5 OPENROWSET(
6     BULK 'https://datalakep2wifjt.dfs.core.windows.net/files/sales/csv'
7     FORMAT = 'CSV',
8     PARSE_VERSION = '2.0'
9 ) AS [result]
10
```

Properties

General Related (0)

Name * Sales CSV query

Description

Type .sql script

Size 228 bytes

Results settings per query ⓘ

First 5000 rows (default)

All rows

Results Messages

View Table Chart Export results

Search

C1	C2	C3	C4
SO49171	1	2021-01-01	Mariah Foster
SO49172	1	2021-01-01	Brian Howard

00:00:09 Query executed successfully.

2:47 PM 10/22/2024 ENG IN

labs.xtremelabs.io/LabViewerConnection/DetachLabManual?labInstanceGuid=56b90206-af08-493c-b... Instance ID:673106 XtremeLabs: DP-203T00-A-CEP [DP-203T00-A-M02-CEP] Module 02: Query files using a serverless SQL pool

```
SELECT
    TOP 100 *
FROM
    OPENROWSET(
        BULK 'https://datalakexxxxxxx.dfs.core.windows.net/files/sales/csv/**',
        FORMAT = 'CSV',
        PARSER_VERSION='2.0'
    )
    WITH (
        SalesOrderNumber VARCHAR(10) COLLATE Latin1_General_100_BIN2_UTF8,
        SalesOrderLineNumber INT,
        OrderDate DATE,
        CustomerName VARCHAR(25) COLLATE Latin1_General_100_BIN2_UTF8,
        EmailAddress VARCHAR(50) COLLATE Latin1_General_100_BIN2_UTF8,
        Item VARCHAR(30) COLLATE Latin1_General_100_BIN2_UTF8,
        Quantity INT,
        UnitPrice DECIMAL(18,2),
        TaxAmount DECIMAL (18,2)
    ) AS [result]
```

Now the results look like this:

SalesOrderNumber	SalesOrderLineNumber	OrderDate	CustomerName	EmailAddress	Item	Quantity	Un
SO49171	1	2021-01-01	Mariah Foster	mariah21@adventure-works.com	Road-Black,	250 48	1 1
...

8. Publish the changes to your script, and then close the script pane.

Previous Exercise < Next Exercise >

Page: 5/11

Microsoft Az... | Resource gro... | synapsep2wif... | synapsep2wif... +

web.azuresynapse.net/en/authoring/explore/linked/sqlscripts/SQ... ☆

synapsep2wif... Search 1 1 ?

We use optional cookies to provide a better experience. Learn more

Accept Reject More options

Synapse live Validate all Publishing 1

files Sales CSV query

Run Undo Publish... Query plan ...

10 SalesOrderNumber VARCHAR(10) COLLATE Latin1_General_100_BIN2_UTF8
11 SalesOrderLineNumber INT,
12 OrderDate DATE,
13 CustomerName VARCHAR(25) COLLATE Latin1_General_100_BIN2_UTF8
14 EmailAddress VARCHAR(50) COLLATE Latin1_General_100_BIN2_UTF8
15 Item VARCHAR(30) COLLATE Latin1_General_100_BIN2_UTF8
16 Quantity INT,
17 UnitPrice DECIMAL(18,2),
18 TaxAmount DECIMAL (18,2)
19 AS [result]
20

Properties

General Related (0)

Name * Sales CSV query

Description

Type .sql script

Size 228 bytes

Results settings per query ⓘ

First 5000 rows (default)

All rows

Results Messages

View Table Chart Export results

Search

SalesOrderNu...	SalesOrderLine...	OrderDate
SO49171	1	2021-01-01

00:00:02 Query executed successfully.

labs.xtremelabs.io/LabViewerConnection/DetachLabManual?la... Instance ID:6731016

labs.xtremelabs.io/LabViewerConnection/DetachLabManual?labIns... Search

XtremeLabs: DP-203T00-A-CEP [DP-203T00-A-M02-CEP] Module 02: Query files using a serverless SQL pool

it.

Paste Content

Paste Content

```
SELECT YEAR(OrderDate) AS OrderYear,  
       COUNT(*) AS OrdredItems  
  FROM  
    OPENROWSET(  
      BULK 'https://datalakexxxxxxx.dfs.core.windows.net/files/sales/parquet/year=**/*',  
      FORMAT = 'PARQUET'  
    ) AS [result]  
 WHERE [result].filepath(1) IN ('2019', '2020')  
 GROUP BY YEAR(OrderDate)  
 ORDER BY OrderYear
```

8. Review the results and note that they include only the sales counts for 2019 and 2020. This filtering is achieved by including a wildcard for the partition folder value in the BULK path (year=*) and a WHERE clause based on the *filepath* property of the results returned by OPENROWSET (which in this case has the alias *[result]*).

9. Name your script **Sales Parquet query**, and publish it. Then close the script pane.

< Previous Exercise Next Exercise >

⊕ Page: 6/11 ⊕

Support

Microsoft Azure Portal | Resource groups - Micro | synapsep2wifjt - Micro | synapsep2wifjt - Azure +

web.azuresynapse.net/en/authoring/explore/linked/sqlscripts/SQL%20script%201?workspace=%2...

Synapse Analytics > synapsep2wifjt Search 1 2 ? ? Accept Reject More options

We use optional cookies to provide a better experience. Learn more Accept Reject More options

Synapse live Validate all Publish all

files Sales Parquet query ... Connect to Built-in

Other users in your workspace may have access to modify this item. Do not use this item unless you trust all users who may have access to the workspace.

```
1 SELECT YEAR(OrderDate) AS OrderYear,  
2       COUNT(*) AS OrdredItems  
3  FROM  
4    OPENROWSET(  
5      BULK 'https://datalakep2wifjt.dfs.core.windows.net/files/sales/parquet/year=**/*',  
6      FORMAT = 'PARQUET'  
7    ) AS [result]  
8 WHERE [result].filepath(1) IN ('2019', '2020')  
9 GROUP BY YEAR(OrderDate)  
10 ORDER BY OrderYear  
11
```

Properties

General Related (0)

Name *

Description

Type .sql script

Size 203 bytes

Results settings per query First 5000 rows (default) All rows

Results Messages

View Table Chart Export results Export results

Search

OrderYear	OrdredItems
2019	1201
2020	2733

00:00:01 Query executed successfully.

labs.xtremelabs.io/LabViewerConnection/DetachLabManual?labIns... Microsoft Azure Portal | Resource groups - Micro | synapsep2wifjt - Micro | synapsep2wifjt - Azure + Instance ID:6731016 labs.xtremelabs.io/LabViewerConnection/DetachLabManual?labIns... Search

Search

Page: 7/11

Support

Microsoft Edge

3:27 PM 10/22/2024 ENG US

a JSON document for each order.

6. Modify the query as follows (replacing `datalakexxxxxx` with the name of your data lake storage account) so that it uses the `JSON_VALUE` function to extract individual field values from the JSON data.

Paste Content

Paste Content

```
SELECT JSON_VALUE(Doc, '$.SalesOrderNumber') AS OrderNumber,
       JSON_VALUE(Doc, '$.CustomerName') AS Customer,
       Doc
  FROM OPENROWSET(
      BULK 'https://datalakexxxxxx.dfs.core.windows.net/files/sales/json/',
      FORMAT = 'CSV',
      FIELDTERMINATOR = '0x0b',
      FIELDQUOTE = '0x0b',
      ROWTERMINATOR = '0x0b'
    ) WITH (Doc NVARCHAR(MAX)) as rows
```

7. Name your script **Sales JSON query**, and publish it. Then close the script pane.

Previous Exercise

Next Exercise >

Synapse Analytics > synapsep2wifjt

We use optional cookies to provide a better experience. Learn more

Accept Reject More options

Synapse live Validate all Publish all

files Sales Parquet query Sales JSON query

Run Undo Publish Query plan Connect to Built-in ...

1 SELECT JSON_VALUE(Doc, '\$.SalesOrderNumber') AS OrderNumber,
2 JSON_VALUE(Doc, '\$.CustomerName') AS Customer,
3 Doc
4 FROM OPENROWSET(
5 BULK 'https://datalakep2wifjt.dfs.core.windows.net/files/sales/json',
6 FORMAT = 'CSV',
7 FIELDTERMINATOR = '0x0b',
8 FIELDQUOTE = '0x0b',
9 ROWTERMINATOR = '0x0b'
10) WITH (Doc NVARCHAR(MAX)) as rows

Properties

General Related (0)

Name * Sales JSON query

Description

Type .sql script

Size 546 bytes

Results settings per query ⓘ

First 5000 rows (default)

All rows

Results Messages

View Table Chart Export results

Search

OrderNumber	Customer	Doc
SO43700	Ruben Prasad	{ "SalesOrderNum...
SO43701	Christy Zhu	{ "SalesOrderNum...
SO43705	Curtis Lu	{ "SalesOrderNum...

00:00:02 Query executed successfully.

labs.xtremelabs.io/LabViewerConnection/DetachLabManual?labInsta... Microsoft Azure Porta Resource groups - Mi synapsep2wifjt - Mic synapsep2wifjt - Azur + labs.xtremelabs.io/LabViewerConnection/DetachLabManual?labInsta... web.azure-synapse.net/en/authoring/explore/workspace/sqlscripts/SQL%20script%201?works... XLab-rPM-910@xtremelabs.us LAB DIRECTORY

Instance ID:6731016
XtremeLabs: DP-203T00-A-CEP [DP-203T00-A-M02-CEP] Module 02: Query files using a serverless SQL pool

following query:

Paste Content Paste Content

```
SELECT *  
FROM  
OPENROWSET(  
    BULK 'csv/*.csv',  
    DATA_SOURCE = 'sales_data',  
    FORMAT = 'CSV',  
    PARSER_VERSION = '2.0'  
) AS orders
```

The query uses the external data source to connect to the data lake, and the OPENROWSET function now only need to reference the relative path to the .csv files.

8. Modify the code as follows to query the parquet files using the data source.

Paste Content Paste Content

```
SELECT *  
FROM  
OPENROWSET(  
    BULK 'parquet/year=/*/*.snappy.parquet',  
    DATA_SOURCE = 'sales_data',  
    FORMAT='PARQUET'  
) AS orders  
WHERE orders.filepath(1) = '2019'
```

Previous Exercise Next Exercise

Page: 8/11

Support

Synapse Analytics > synapsep2wifjt Search

We use optional cookies to provide a better experience. [Learn more](#)

Synapse live Publish all

Create Sales DB files SQL script 1

Other users in your workspace may have access to modify this item. Do not use this item unless you trust all users who may have access to the workspace.

Run Undo Publish Query plan Connect to Built-in ...

```
1 SELECT *  
2 FROM  
3 OPENROWSET(  
4     BULK 'csv/*.csv',  
5     DATA_SOURCE = 'sales_data',  
6     FORMAT = 'CSV',  
7     PARSER_VERSION = '2.0'  
8 ) AS orders
```

Properties

General Related (0)

Name * SQL script 1

Description

Type .sql script

Size 0 bytes

Results settings per query
 First 5000 rows (default)
 All rows

Results Messages

View Table Chart Export results

Search

C1	C2	C3	C4	C5
SO49171	1	2021-01-01	Mariah Foster	mariyah210

00:00:08 Query executed successfully.

3:36 PM 10/22/2024

labs.xtremelabs.io/LabViewerConnection/DetachLabManual?labInsta... Microsoft Azure Porta Resource groups - Mi synapsep2wifjt - Mic synapsep2wifjt - Azur + labs.xtremelabs.io/LabViewerConnection/DetachLabManual?labInstance... web.azure.synapse.net/en/authoring/explore/workspace/sqlscripts/SQL%20script%201?works... XLab-rPM-910@xtremelabs.us LAB DIRECTORY

Instance ID:6731016
XtremeLabs: DP-203T00-A-CEP [DP-203T00-A-M02-CEP] Module 02: Query files using a serverless SQL pool

79% Completed
Lab Time Left: 00:27:13

Lab Actions

9. Query files using a serverless SQL pool:
Create an external table:

The external data source makes it easier to access the files in the data lake, but most data analysts using SQL are used to working with tables in a database. Fortunately, you can also define external file formats and external tables that encapsulate rowsets from files in database tables.

1. Replace the SQL code with the following statement to define an external data format for CSV files, and an external table that references the CSV files, and run it:

Paste Content
Paste Content

```
CREATE EXTERNAL FILE FORMAT CsvFormat
    WITH (
        FORMAT_TYPE = DELIMITEDTEXT,
        FORMAT_OPTIONS(
            FIELD_TERMINATOR = ',',
            STRING_DELIMITER = ''
        )
    );
GO;

CREATE EXTERNAL TABLE dbo.orders
(
    SalesOrderNumber VARCHAR(10),
    SalesOrderLineNumber INT,
    OrderDate DATE,
    CustomerName VARCHAR(25),
    EmailAddress NVARCHAR(40)
);
```

Page: 9/11

Synapse Analytics > synapsep2wifjt Search 1 5 ? Publish all Accept Reject More options

We use optional cookies to provide a better experience. Learn more

Synapse live Validate all Publish all 1

Create Sales DB files SQL script 1

Other users in your workspace may have access to modify this item. Do not use this item unless you trust all users who may have access to the workspace.

Run Undo Publish Query plan Connect to Built-in ...

```
1 SELECT *
2 FROM
3 OPENROWSET(
4     BULK 'parquet/year=/*/*.snappy.parquet',
5     DATA_SOURCE = 'sales_data',
6     FORMAT='PARQUET'
7 ) AS orders
8 WHERE orders.filepath(1) = '2019'
```

Properties

General Related (0)

Name * SQL script 1

Description

Type .sql script

Size 0 bytes

Results settings per query ⓘ

First 5000 rows (default)

All rows

Results Messages

View Table Chart Export results

Search

SalesOrderNu...	SalesOrderLine...	OrderDate	CustomerName	EmailAddres...
SO43701	1	2019-07-01	Christy Zhu	christy12@ad...
SO43704	1	2019-07-01	Julio Ruiz	julio1@ad...

00:00:03 Query executed successfully.

Support

Search

3:37 PM 10/22/2024

Instance ID:6731016
XtremeLabs: DP-203T00-A-CEP [DP-203T00-A-M02-CEP] Module 02: Query files using a serverless SQL pool

```

SalesOrderNumber VARCHAR(10),
SalesOrderLineNumber INT,
OrderDate DATE,
CustomerName VARCHAR(25),
EmailAddress VARCHAR(50),
Item VARCHAR(30),
Quantity INT,
UnitPrice DECIMAL(18,2),
TaxAmount DECIMAL (18,2)
)
WITH
(
    DATA_SOURCE =sales_data,
    LOCATION = 'csv/*.csv',
    FILE_FORMAT = CsvFormat
);
GO

```

2. Refresh and expand the **External tables** folder in the **Data** pane and confirm that a table named **dbo.orders** has been created in the **Sales** database.

3. In the ... menu for the **dbo.orders** table, select **New SQL script > Select TOP 100 rows**.

4. Run the SELECT script that has been generated, and verify that it retrieves the first 100 rows of data from the table, which in turn references the files in the data lake.

< Previous Exercise Next Exercise >

⊕ Page: 9/11 ⊕

Support

Synapse Analytics > synapse < Search

Synapse live Validate all Publish all 2

Data Workspace Linked

Filter resources by name

SQL database 1

Sales (SQL)

External tables

dbo.orders

Columns

- SalesOrderNumber (va...)
- SalesOrderLineNumbe...
- OrderDate (date, null)
- CustomerName (varchar...)
- EmailAddress (varchar(...))
- Item (varchar(30), null)
- Quantity (int, null)
- UnitPrice (decimal(18,...))
- TaxAmount (decimal(1...

Run Undo Publish Query plan ...

Properties

General Related (0)

Name * SQL script 2

Description

Type .sql script

Results Messages

View Table Chart ...

Search

00:00:02 Query executed successfully.

Microsoft Azure Portal | Microsoft Azure Resource groups - Microsoft Azure synapse2wifjt - Microsoft Azure synapse2wifjt - Azure Synapse

web.azure-synapse.net/en/authoring/analyze/sqlscripts/SQL%20script%203?workspace=%2Fsubscriptions%2F4d335bb4-c496-4cbf-adb0-6ef1a022b73d%2FresourceGroups%2Fdp2...

Microsoft Azure | Synapse Analytics > synapse2wifjt

Search

Synapse live Validate all Publish all 3

Create Sales DB files SQL script 1 SQL script 2 SQL script 3

Run Undo Publish Query plan Connect to Built-in Use database Sales

1 SELECT YEAR(OrderDate) AS OrderYear,
2 SUM((UnitPrice * Quantity) + TaxAmount) AS GrossRevenue

Results Messages

View Table Chart Save as image

12.5M
10M
7.5M
5M
2.5M
0

OrderYear GrossRevenue

0 1 2

Chart type: Line
Category column: (none)
Legend (series) columns: OrderYear, GrossRevenue
Legend position: bottom - center
Legend (series) label:
Category label:

Properties

General Related (0)

Name * SQL script 3
Description
Type .sql script
Size 0 bytes
Results settings per query
 First 5000 rows (default)
 All rows

00:00:02 Query executed successfully.

Search

3:47 PM 10/22/2024

Microsoft Azure Portal | Microsoft Azure Resource groups - Microsoft Azure synapse2wifjt - Microsoft Azure synapse2wifjt - Azure Synapse

web.azure-synapse.net/en/authoring/analyze/sqlscripts/SQL%20script%203?workspace=%2Fsubscriptions%2F4d335bb4-c496-4cbf-adb0-6ef1a022b73d%2FresourceGroups%2Fdp2...

Microsoft Azure | Synapse Analytics > synapse2wifjt

Search

Synapse live Validate all Publish all 3

Create Sales DB files SQL script 1 SQL script 2 SQL script 3

Run Undo Publish Query plan Connect to Built-in Use database Sales

```
1 SELECT YEAR(OrderDate) AS OrderYear,
2      SUM((UnitPrice * Quantity) + TaxAmount) AS GrossRevenue
3 FROM dbo.orders
4 GROUP BY YEAR(OrderDate)
5 ORDER BY OrderYear;
```

Results Messages

View Table Chart Save as image

15M
10M
5M
0

2019 2020 2021

GrossRevenue

Legend (series) columns: GrossRevenue
Legend position: bottom - center
Legend (series) label:
Legend (series) minimum value:
Legend (series) maximum value:
Category label:

Properties

General Related (0)

Name * SQL script 3
Description

Type .sql script
Size 0 bytes
Results settings per query ⓘ
 First 5000 rows (default)
 All rows

00:00:02 Query executed successfully.

Search

3:49 PM 10/22/2024 ENG IN

labs.xtremelabs.io/LabViewerConnection/DetachLab... Microsoft Azure Portal | A Resource groups - Micros A synapse2wifjt - Microso A synapse2wifjt - Azure Sy + labs.xtremelabs.io/LabViewerConnection/DetachLabMan... web.azure-synapse.net/en/authoring/analyze?workspace=%2Fsubscriptions%2F4d335bb4-c496-4cbf-adb0-... XLab-rPM-910@xtremelabs.us LAB DIRECTORY

Instance ID:6731016
XtremeLabs: DP-203T00-A-CEP [DP-203T00-A-M02-CEP] Module 02: Query files using a serverless SQL pool

95% Completed
Lab Time Left: 00:13:36

Lab Actions ▾

11. Query files using a serverless SQL pool:

Delete Azure resources:

If you've finished exploring Azure Synapse Analytics, you should delete the resources you've created to avoid unnecessary Azure costs.

1. Close the Synapse Studio browser tab and return to the Azure portal.

2. On the Azure portal, on the **Home** page, select **Resource groups**.

3. Select the **dp203-xxxxxx** resource group for your Synapse Analytics workspace (not the managed resource group), and verify that it contains the Synapse workspace and storage account for your workspace.

4. At the top of the **Overview** page for your resource group, select **Delete resource group**.

5. Enter the **dp203-xxxxxx** resource group name to confirm you want to delete it, and select **Delete**.

After a few minutes, your Azure Synapse workspace resource group and the managed workspace resource group associated with it will be deleted.

Page: 11/11

Support

Synapse Analytics > synapse2wifjt

Search

1 5 ? 🔍

Synapse live Validate all Publish all

»

»

Home Storage Synapse Analytics Data Factory Data Lake Analytics Machine Learning

Select an item

Use the resource explorer to select or create a new item

3:51 PM 10/22/2024 ENG IN

labs.xtremelabs.io/LabViewerConnection/DetachLabManu... Microsoft Azure Portal | Microsoft Azure dp203-p2wifjt - Microsoft Azure

Instance ID:6731016
XtremeLabs: DP-203T00-A-CEP [DP-203T00-A-M02-CEP] Module 02: Query files using a serverless SQL pool

97% Completed
Lab Time Left: 00:09:57

Lab Actions

11. Query files using a serverless SQL pool:

Delete Azure resources:

- If you've finished exploring Azure Synapse Analytics, you should delete the resources you've created to avoid unnecessary Azure costs.
- 1. Close the Synapse Studio browser tab and return to the Azure portal.
- 2. On the Azure portal, on the **Home** page, select **Resource groups**.
- 3. Select the **dp203-xxxxxxxx** resource group for your Synapse Analytics workspace (not the managed resource group), and verify that it contains the Synapse workspace and storage account for your workspace.
- 4. At the top of the **Overview** page for your resource group, select **Delete resource group**.
- 5. Enter the **dp203-xxxxxxxx** resource group name to confirm you want to delete it, and select **Delete**.

After a few minutes, your Azure Synapse workspace resource group and the managed workspace resource group associated with it will be deleted.

Page: 11/11

Support

portal.azure.com/#@xtremelabs.onmicrosoft.com/resource/subscriptions/4d335bb4-c496-4cbf-a... Copilot

Microsoft Azure

Home > Resource groups > dp203-p2wifjt

Resource group

Search

+ Create Manage view Delete resource group Refresh Export to CSV Open query

Overview

Activity log Access control (IAM) Tags Resource visualizer Events Settings Cost Management Monitoring Automation Help

Essentials

Resources Recommendations

Filter for any field... Type equals all Location equals all Add filter

Showing 0 to 0 of 0 records. Show hidden types

List view

Name ↑ Type ↑ Location ↑

No resources match your filters

Try changing or clearing your filters.

Create resources Clear filters

Learn more Give feedback

3:54 PM 10/22/2024



This training session brought to you by Southern Alberta Institute of Technology (XTP) Powered by



Welcome to XtremeLabs, Nitin Nitin



Search for courses

DP-203T00-A-CEP

DP-900T00-A-CEP



Course DP-203T00-A-CEP: Data Engineering on Microsoft Azure

[Go To Lesson](#)

Lab Title: [DP-203T00-A-M01-CEP] Module 01: Explore Azure Synapse Analytics
Duration: 240 minutes
Status: **Completed**

[Take Lab](#)

Lab Title: [DP-203T00-A-M02-CEP] Module 02: Query files using a serverless SQL pool
Duration: 120 minutes
Status: **Completed**

[Take Lab](#)

Lab Title: [DP-203T00-A-M03-CEP] Module 03: Transform data using a serverless SQL pool
Duration: 120 minutes
Status: Not Initiated

[Take Lab](#)

Lab Title: [DP-203T00-A-M04-CEP] Module 04: Analyze data in a lake database
Duration: 120 minutes
Status: Not Initiated

[Take Lab](#)

Lab Title: [DP-203T00-A-M05-CEP] Module 05: Analyze data in a data lake with Spark
Duration: 120 minutes
Status: Not Initiated

Lab Title: [DP-203T00-A-M06-CEP] Module 06: Transform data using Spark in Synapse Analytics
Duration: 120 minutes
Status: Not Initiated


