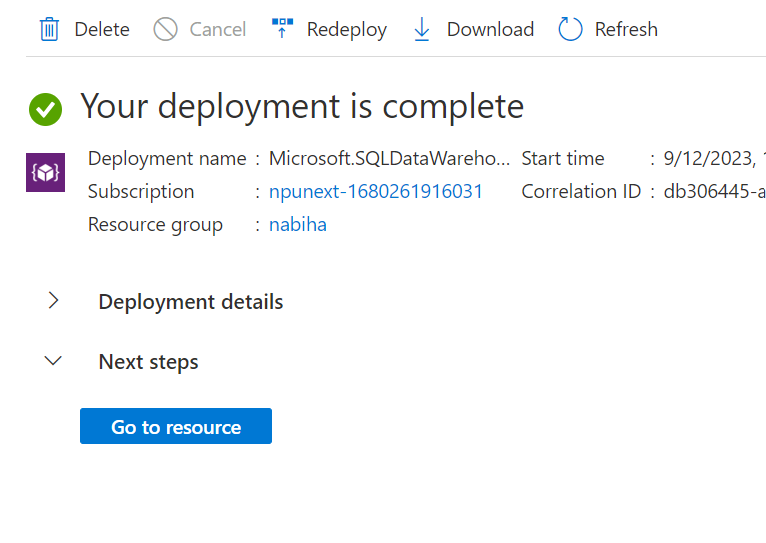
Azure Synapse Analytics

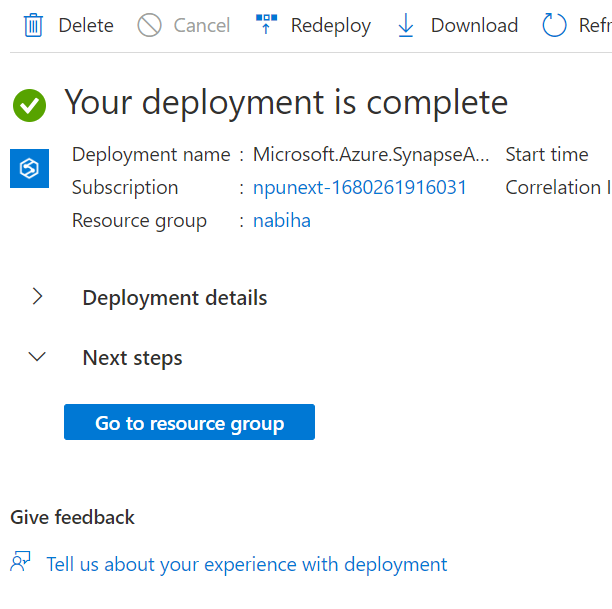
* Dedicated SQL Pool

DW100



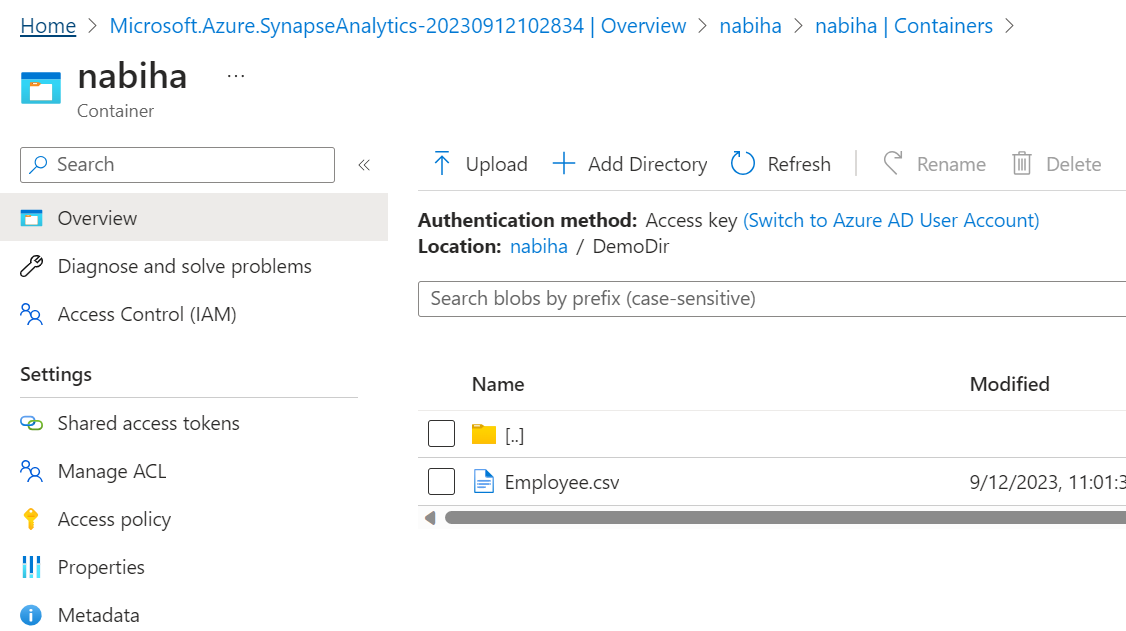
* Azure Synapse Workspace

Create the resource

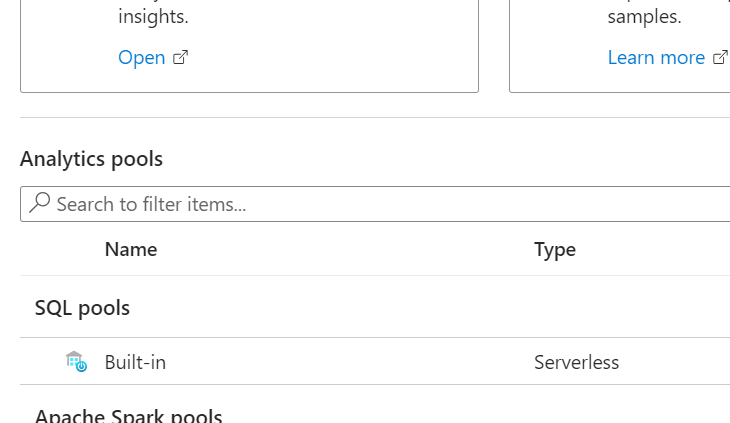


SERVERLESS SQL APPROACH

Go to storage > DLS> container> new dir> upload csv

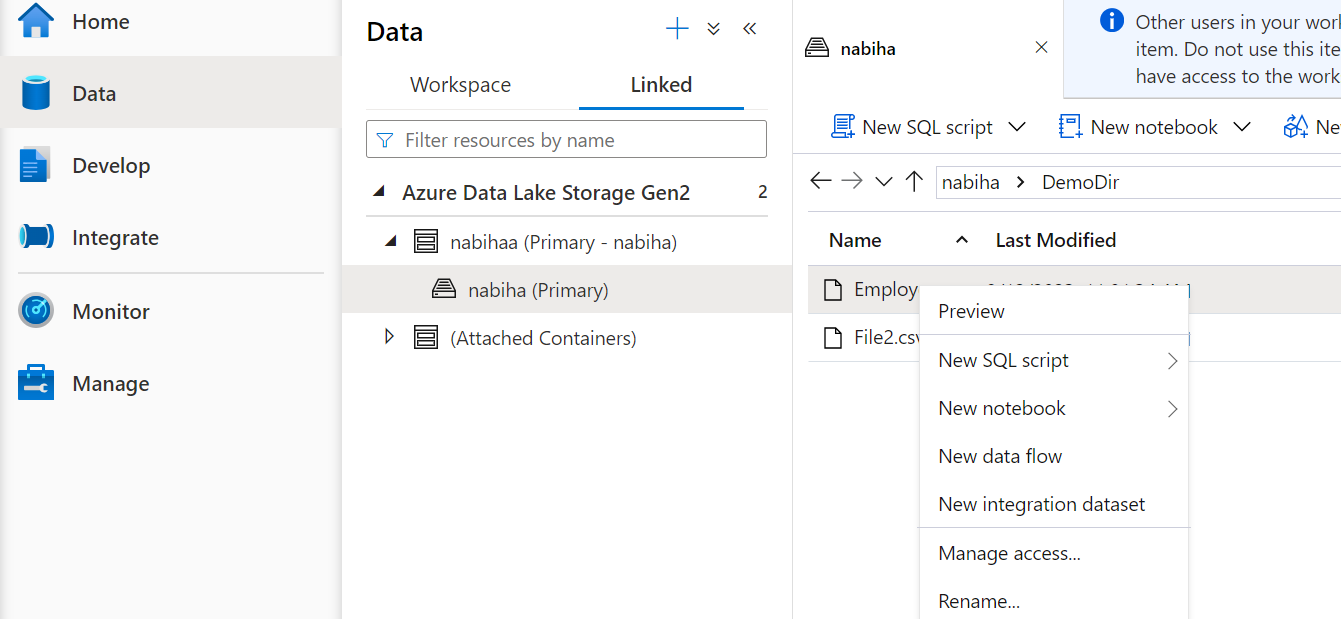


Open synapse studio from all resources



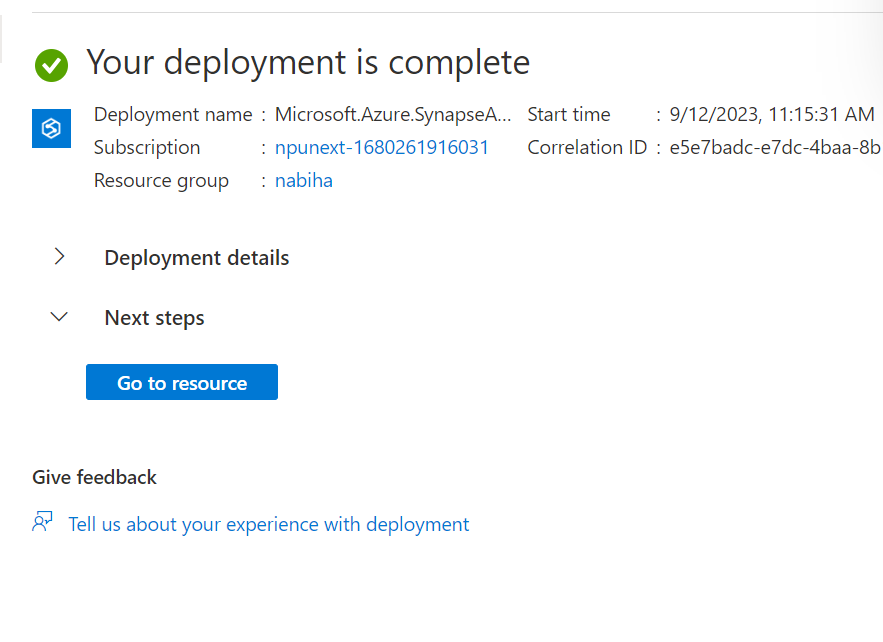
Expand>data> workspace, linked

Linked > expand adls> primary> dwtemp primary> demo dir> csv> right click> New Sql script -> serverless sql pool, new notebook -> apache spark synapse session



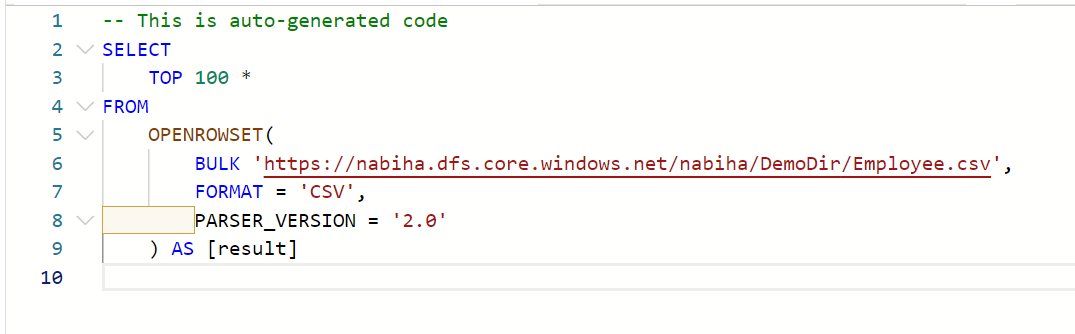
New Apache Spark tool from synapse workspace

All default



ASA

* **New Sql Script> top 100 rows**

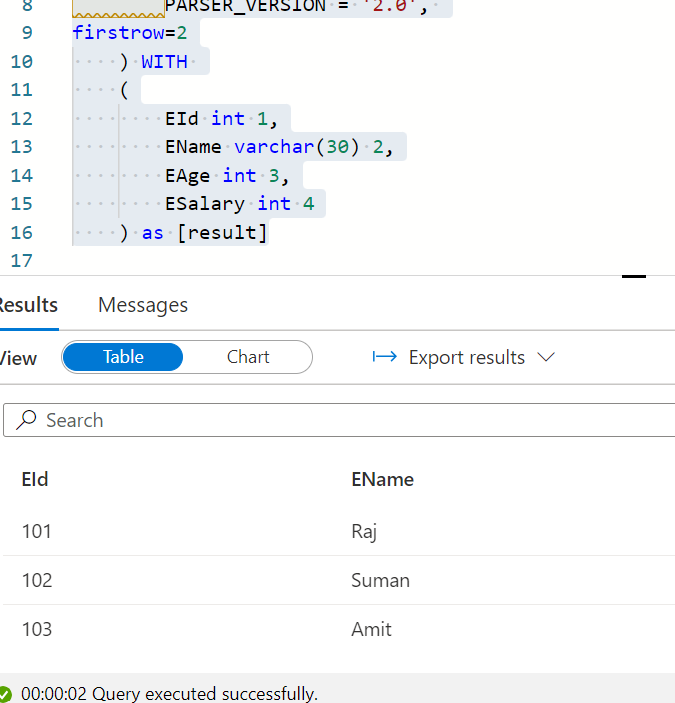


Openrowset -> reading data from external storage, ie blob storage

BULK -> container, sub directory, file name

FORMAT -> specified

PARSER VERSION -> parse from storage



SELECT

    TOP 5 \*

FROM

    OPENROWSET(

        BULK 'https://nabiha.dfs.core.windows.net/nabiha/DemoDir/Employee.csv',

        FORMAT = 'CSV',

        PARSER\_VERSION = '2.0',

firstrow=2

) WITH

    (

        EId int 1,

        EName varchar(30) 2,

        EAge int 3,

        ESalary int 4

    ) as [result]

create database demodbsqlpool

create external data source covid

with (

    location ='https://pandemicdatalake.blob.core.windows.net/public/curated/covid-19/ecdc\_cases'

);

PARQUET

select top 10 \*

from openrowset(

bulk 'https://pandemicdatalake.blob.core.windows.net/public/curated/covid-19/ecdc\_cases/latest/ecdc\_cases.parquet',

format = 'parquet') as rows

JSON FORMAT

select top 10 \*

from

openrowset(

bulk 'https://pandemicdatalake.blob.core.windows.net/public/curated/

covid-19/ecdc\_cases/latest/ecdc\_cases.jsonl',

format = 'csv',

fieldterminator ='0x0b',

fieldquote = '0x0b'

) with (doc nvarchar(max)) as rows

* Create external database objects
* Develop> sql script

Master key -> customized credentials

Shared access signature -> storage acc> shared access sig> allow access> blob service sas rule>



--Creating master key

create master key

-- Create db scoped credentials

CREATE DATABASE SCOPED CREDENTIAL[sqlcommand]

WITH IDENTITY='SHARED ACCESS SIGNATURE',

SECRET='?sv=2022-11-02&ss=bfqt&srt=sco&sp=rwdlacupyx&se=2023-09-12T16:11:45Z&st=2023-09-12T08:11:45Z&spr=https&sig=u1fWnwD5NkE95OviH51DmY1HXXRtVYuno4xaJ%2Btywvo%3D'

--drop DATABASE SCOPED CREDENTIAL sqlcommand

--drop EXTERNAL DATA SOURCE sqlondemand

--External data source creation

CREATE EXTERNAL DATA SOURCE sqlondemand WITH(

    LOCATION='https://nabiha.blob.core.windows.net',

    CREDENTIAL=[sqlcommand]

)

--DROP EXTERNAL FILE FORMAT sqlfilen

--create extrenal file format

CREATE EXTERNAL FILE FORMAT sqlfilen

WITH

(

    FORMAT\_TYPE=DELIMITEDTEXT,

    FORMAT\_OPTIONS(FIELD\_TERMINATOR=',',STRING\_DELIMITER='"',FIRST\_ROW=2)

)

--drop external TABLE emptablen

--creating external table

CREATE EXTERNAL TABLE EmpTablen

(

    EId int,

    EName varchar(30),

    EAge int,

    ESalary int

)

WITH

(

    LOCATION='nabiha/DemoDir/Employee.csv',

    DATA\_SOURCE=sqlondemand,

    FILE\_FORMAT=sqlfilen

)

--Fetch

select \* from EmpTablen

--create view using sql query

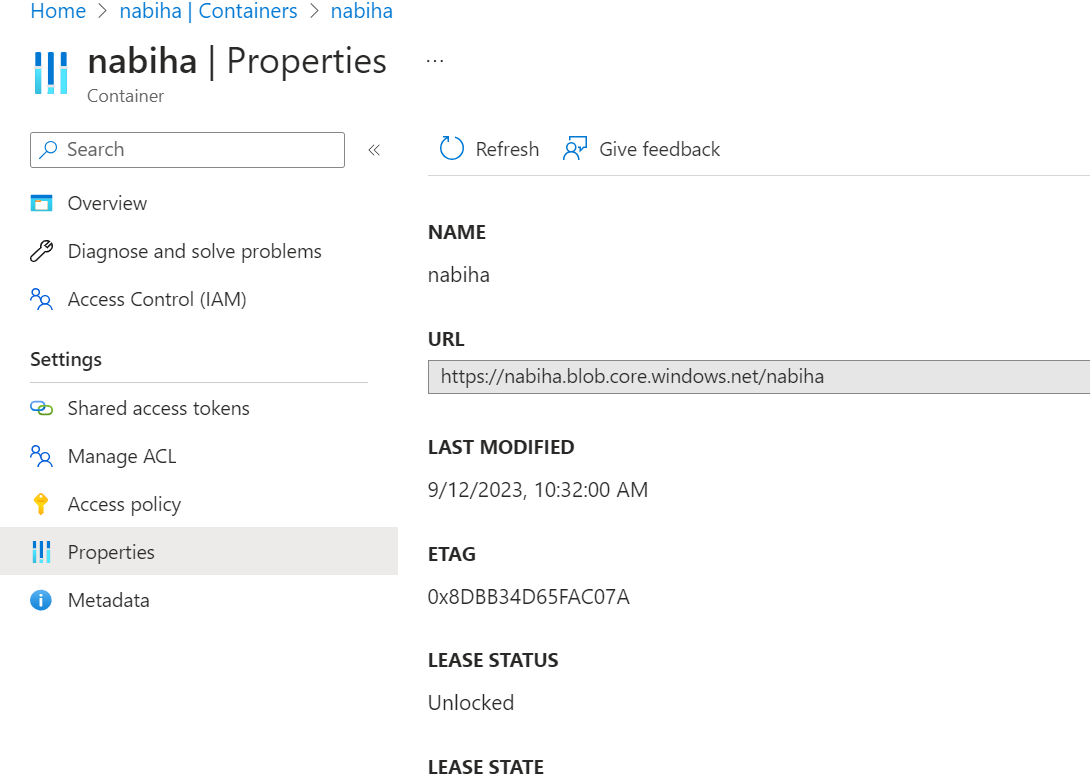
CREATE VIEW viewemployee

AS

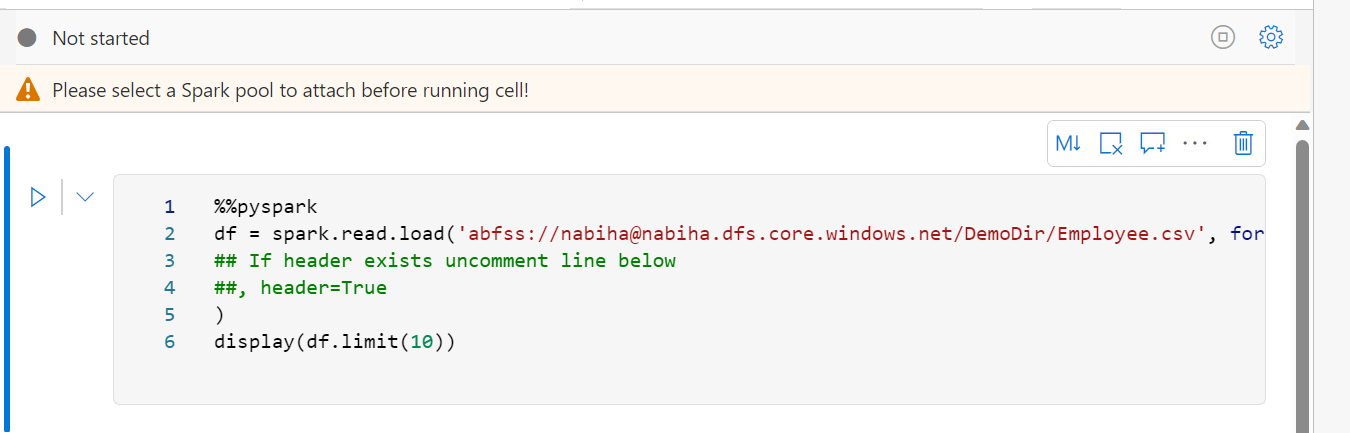
select \* from EmpTablen

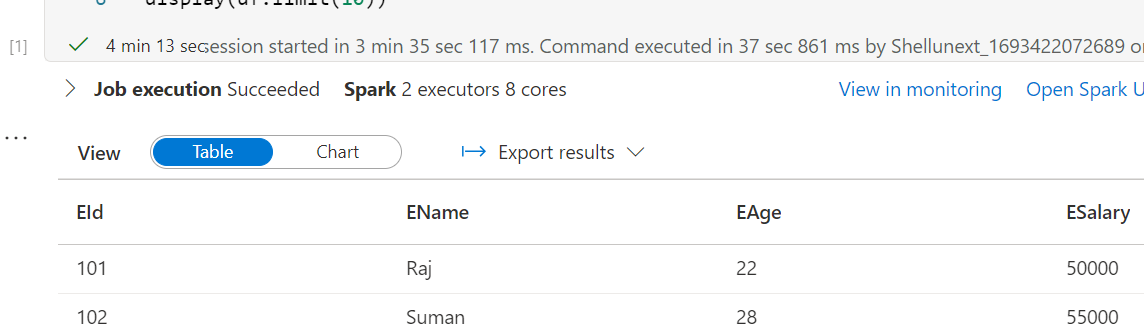
--reading from view

select \* from viewemployee



* **Load to data frame**





* DEDICATED SQL POOL

