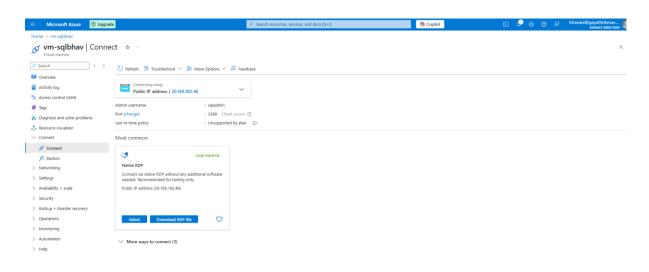
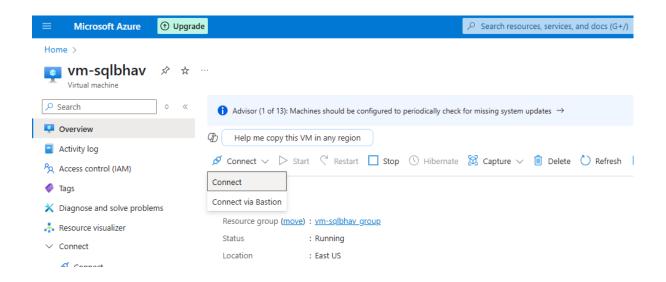
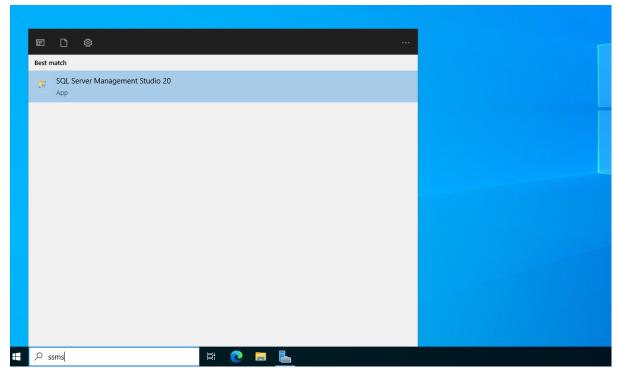
Homework | Cycle 29 | Session 16

 Set up the self hosted IR and copy data from onpremise sql to cloud adlsgen2 and from adls gen2 to c drive of the machine(onpremise)

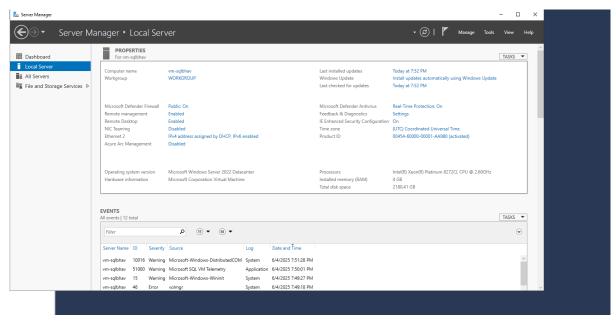
To open VM first we have to start the vm and click on connect and open SSMS inside VM





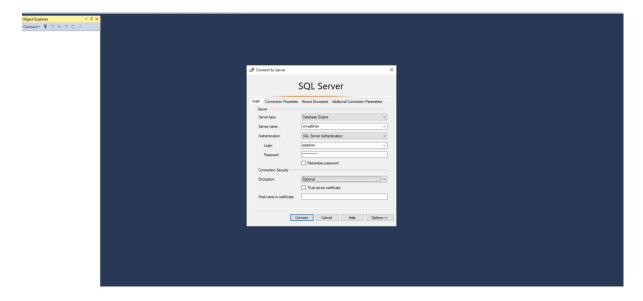


Make turnoff IE Enhancement security turnoff



Select server name as yours and select aunthication as sql server authentication, give username and password that used for for VM creation

Note: Your servername is nothing but your VM name

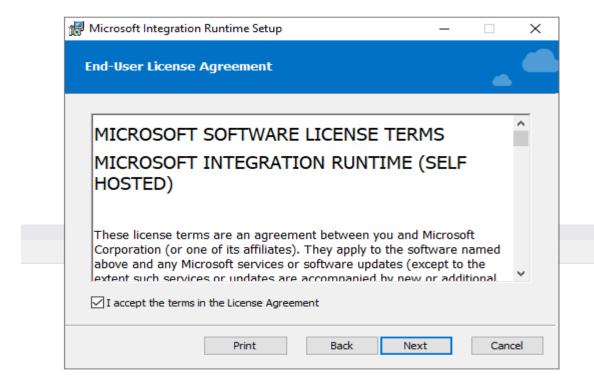


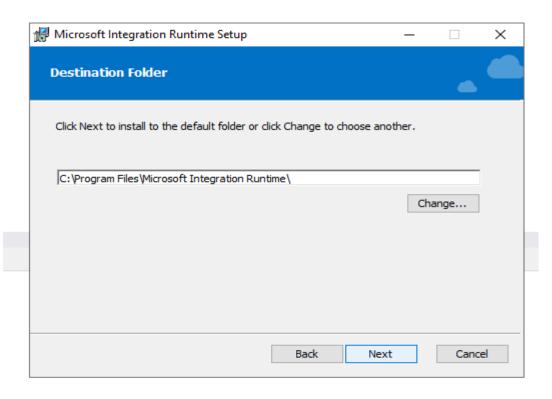
Before we create new database and we have to install Self hosted integration runtime.

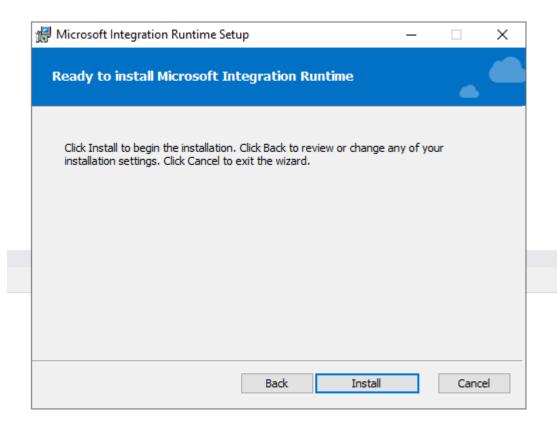
Using below link try to download self hosted IR

<u>Download Microsoft Integration Runtime from Official Microsoft Download Center</u>

After it is downloaded, open the downloaded one and click on next and install to install self hosted ir as below snaps.



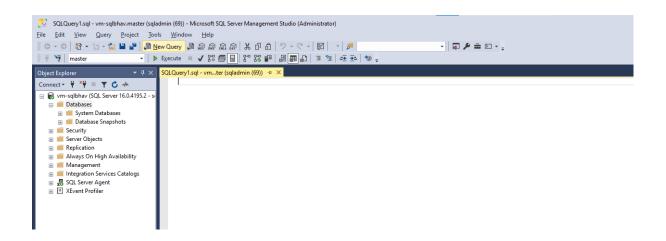


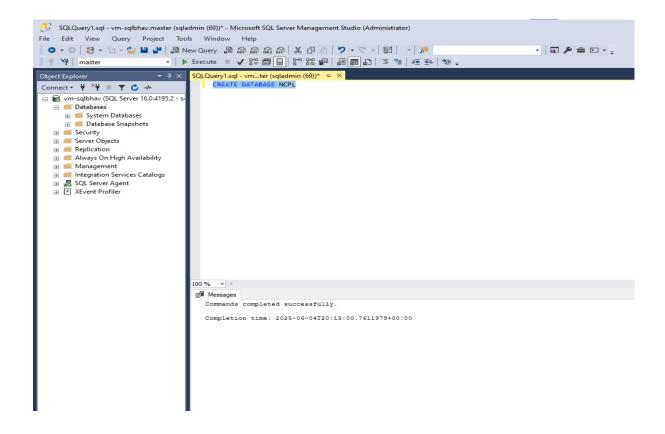


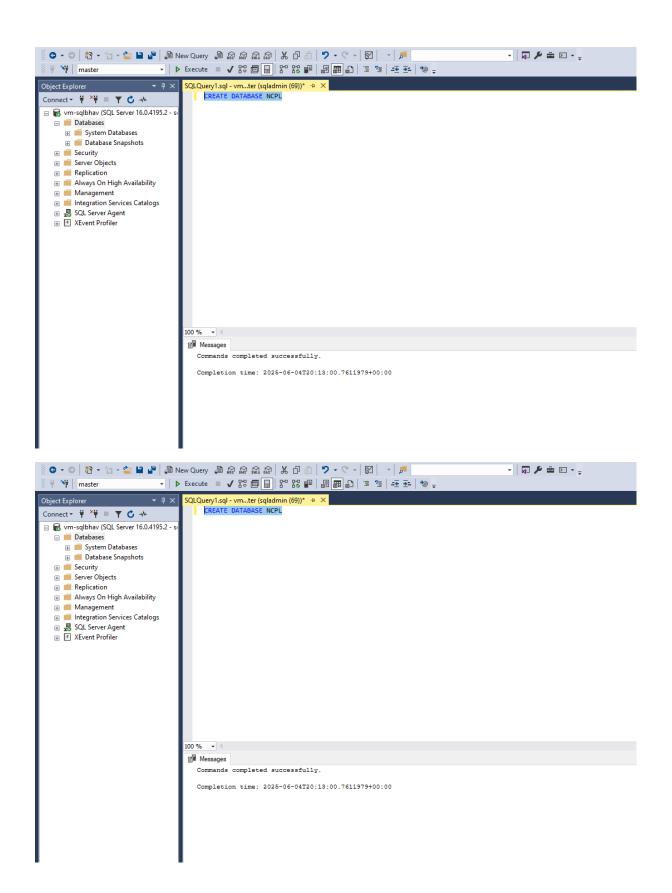
Click on Nwe query to create new database as we have all system created databases in ssms, we are creating a database using below query

CREATE DATABASE NCPL and to use the created database use command as

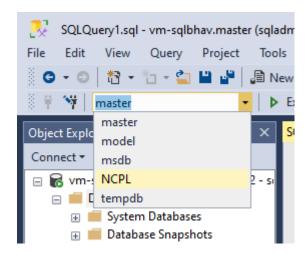
'use database name'







OR else you can select from change connection option to use your database and select NCPL from list as shown below



Craete a sample table using the below query and insert values to it.

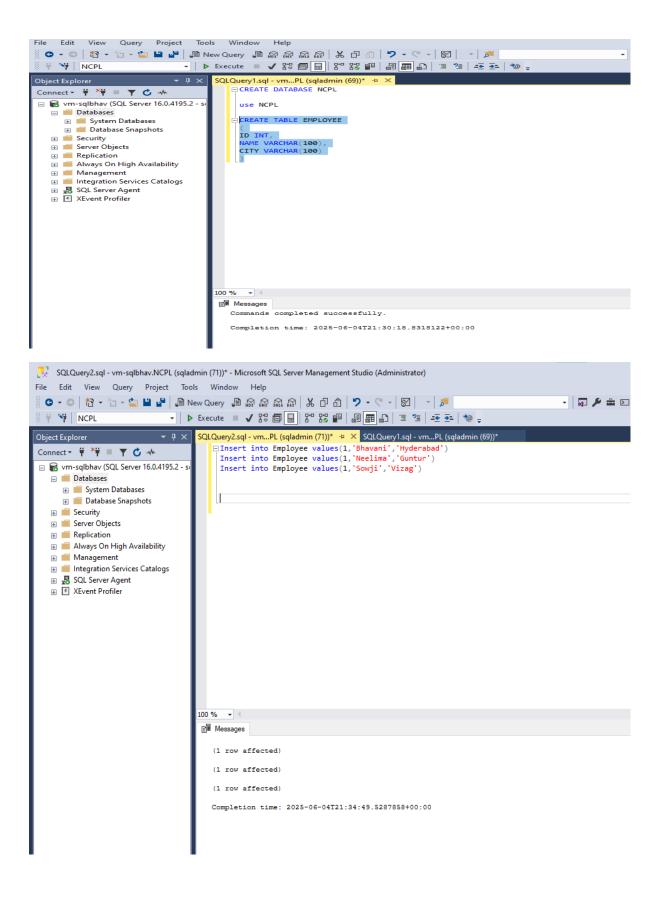
```
CREATE TABLE EMPLOYEE
(
ID INT,
NAME VARCHAR(100),
CITY VARCHAR(100)
)
```

For Insert query as below

Insert into Employee values(1, 'Bhavani', 'Hyderabad')

Insert into Employee values(1,'Neelima','Guntur')

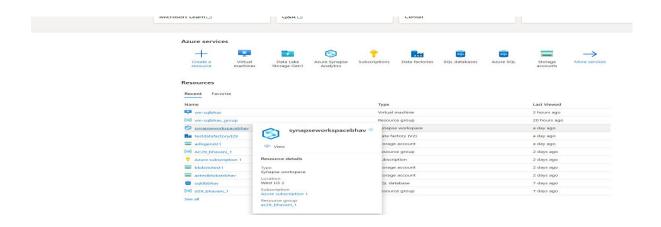
Insert into Employee values(1, 'Sowji', 'Vizag')

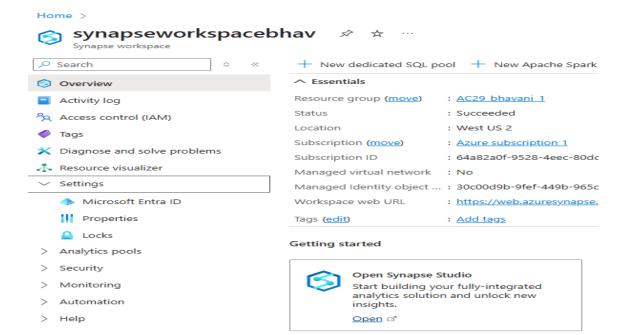


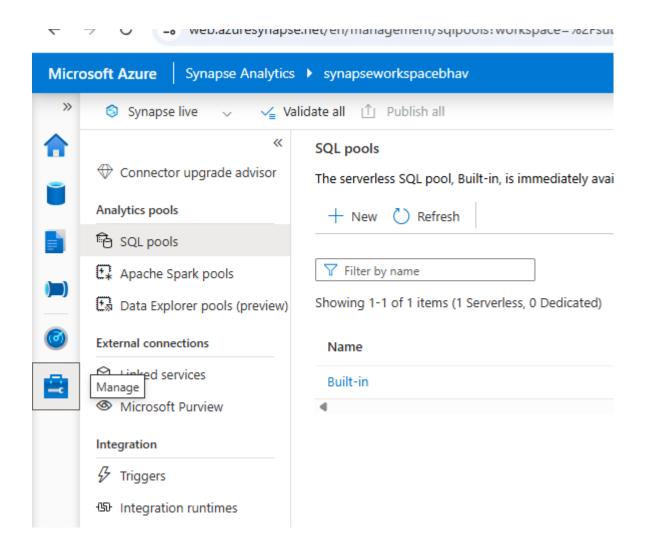


Create self hosted runtime from synapse workspace for connecting onpremises to cloud

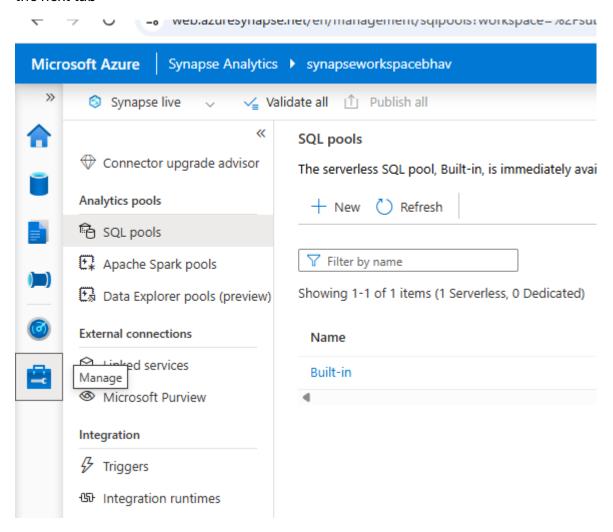
- Go to your synapse workspace, and then go to managetab and click on open synapse studio.
- Go to managetab and Integration runtime

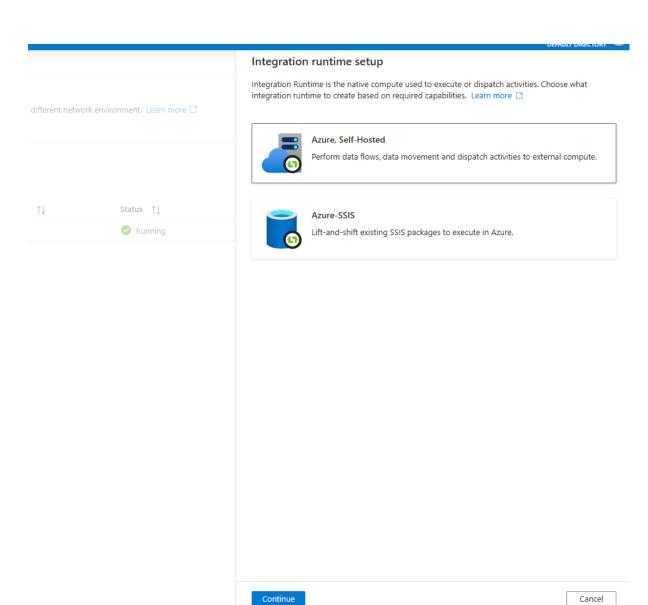






To create IR create 'NEW' button and select Azure self hosted and select self hosted in the next tab





Integration runtime setup

Network environment:

Choose the network environment of the data source / destination or external compute to which the integration runtime will connect to for data flows, data movement or dispatch activities:



Azure

Use this for running data flows, data movement, external and pipeline activities in a fully managed, serverless compute in Azure.



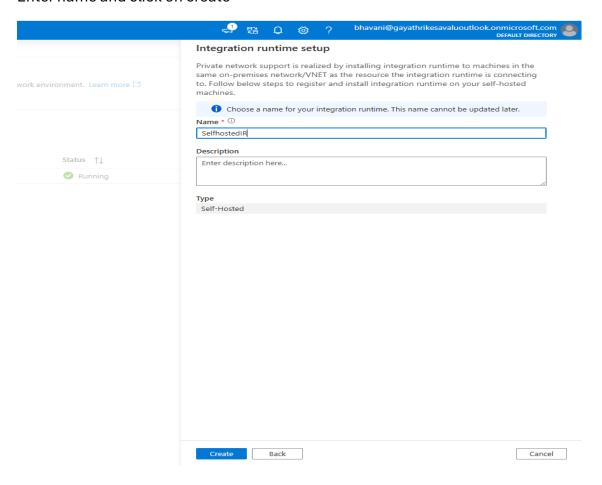
Self-Hosted

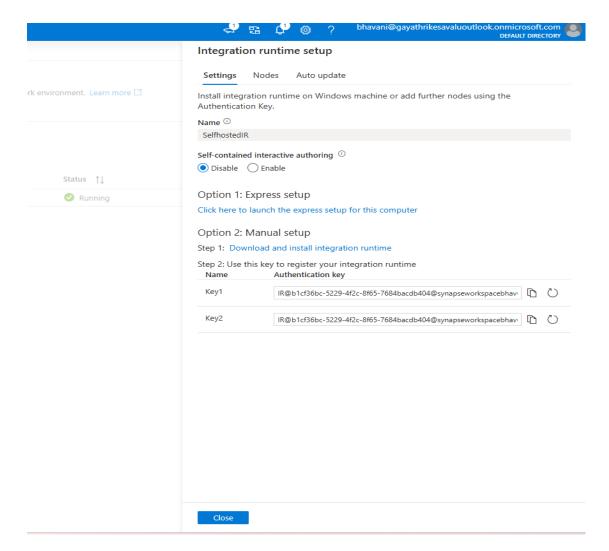
Use this for running data movement, external and pipeline activities in an onpremises / private network by installing the integration runtime.

Note: Data flows are only supported on Azure integration runtime. You can use self-hosted integration runtime to stage the data on cloud storage and then use data flows to transform it.

View less ^

Enter name and click on create





Copy the key and paste it in the Register integration runtime from VM and paste the copied key and click on Register

Option 2. Manual setup

Step 1: Download and install integration runtime

Step 2: Use this key to register your integration runtime

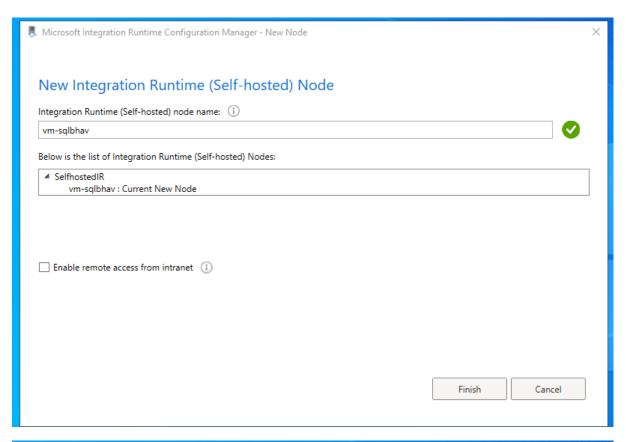
Name Authentication key

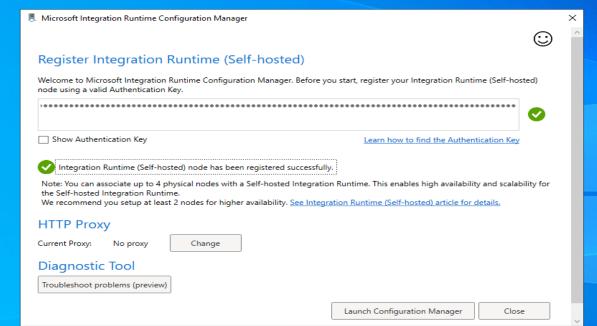
Key1 IR@b1cf36bc-5229-4f2c-8f65-7684bacdb404@synapseworkspacebhavi



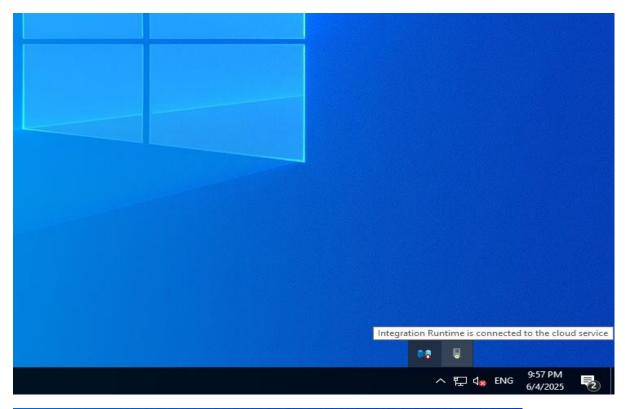


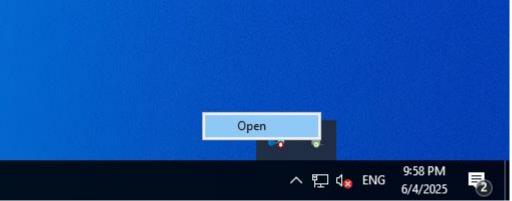
Click on finish and close



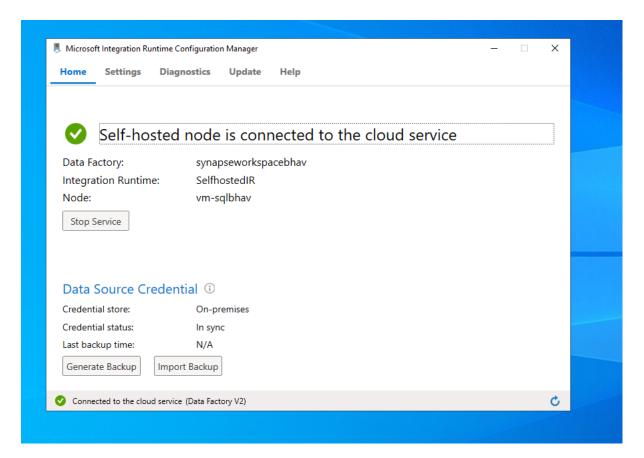


Click on Integration runtime from taskbar icon and right click then open as below showed





It shows your synapse workspace and your vm is connected, self hosted node is connected to cloud service

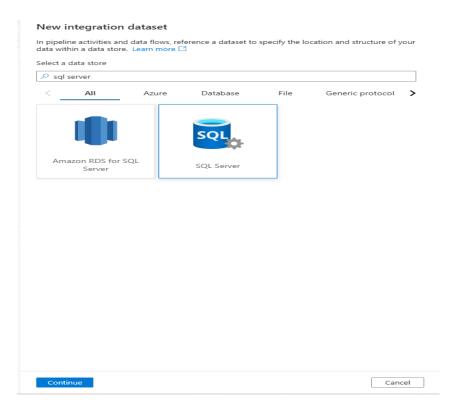


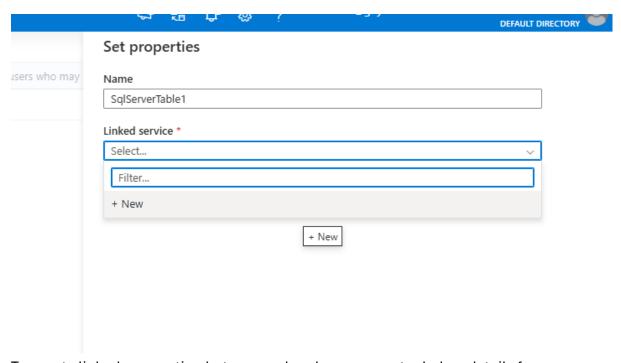
Go to synapse workspace and it clearly shows IR status is showing Running, which means it perfectly integrated with my onpremise vm and synapse.



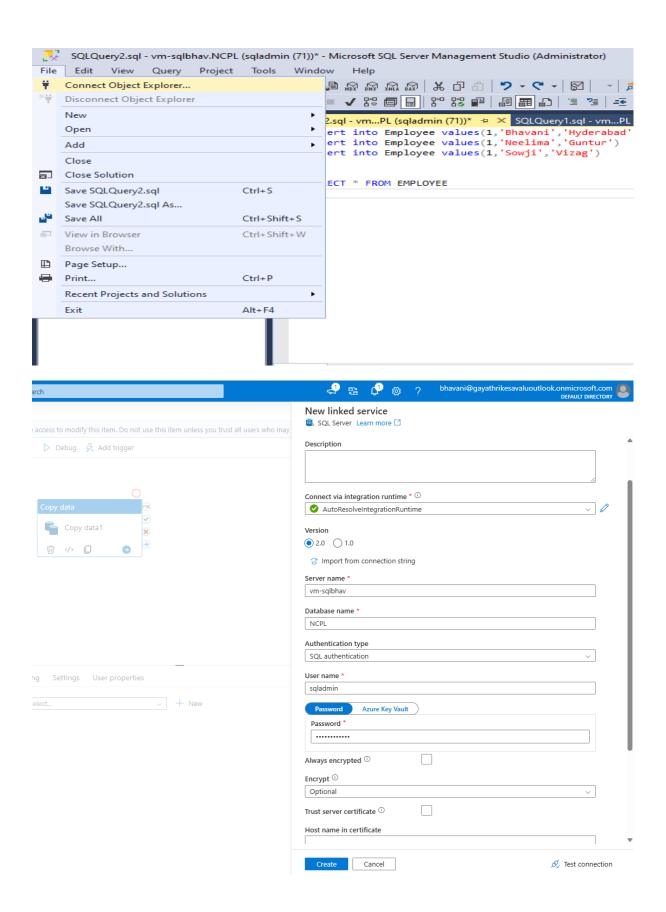
Now create a pipeline to bring data from Onpremises to cloud

Go to Integrate tab and create a pipeline and from source tab select sql server and continue then click on new for linked service



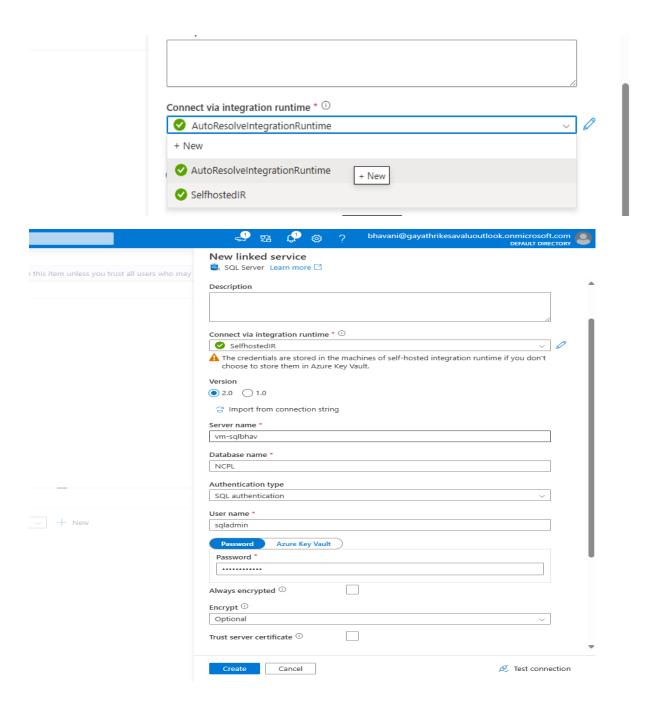


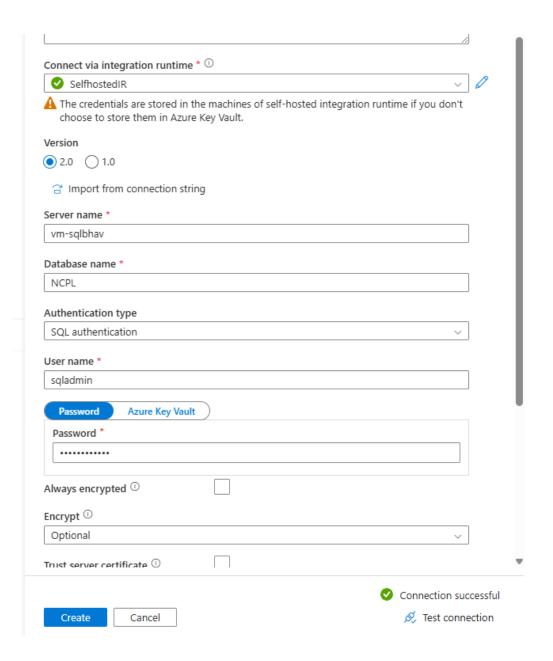
To create linked connection between sql and synapse enter below details from ssms server details like copuy server name as vm-sqlbhav and database name as NCPL and username and password as below



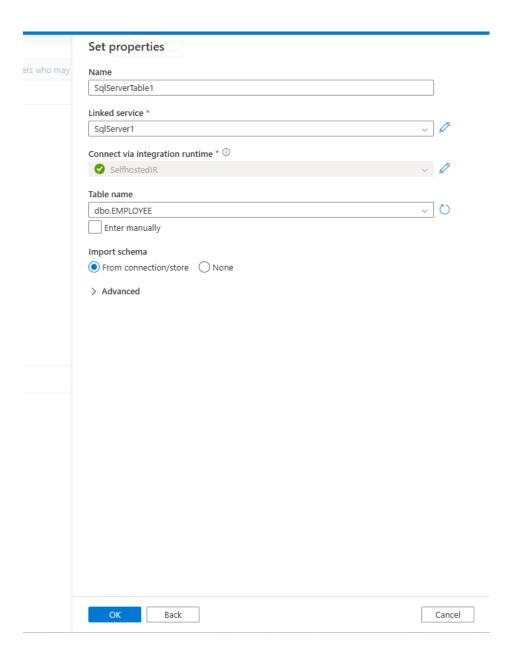
And finally select the integartion Runtime as self hosted on because we are bringing data from onpremis to cloud and also if you select auto resolvce it will fail because the

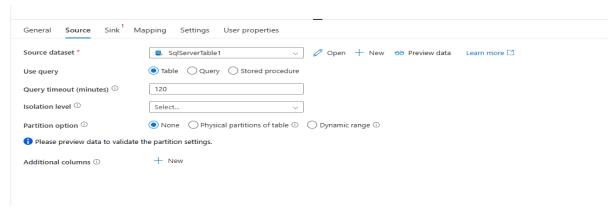
vm is azure but inside that we are connecting to the onpremises sql server, so we have to select self hosted IR. Then Test Connection



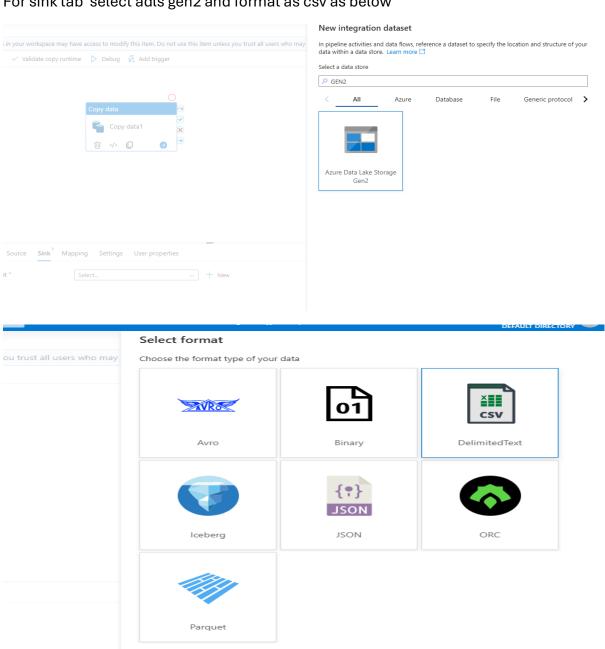


Next is selecttable and clcik OK

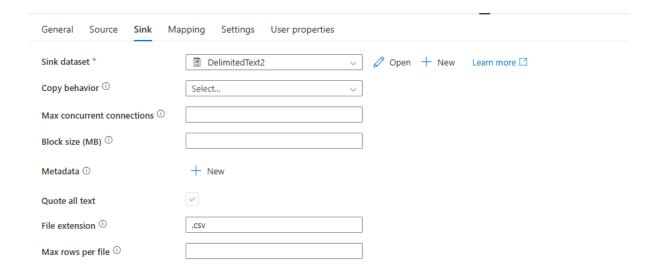




For sink tab select adls gen2 and format as csv as below

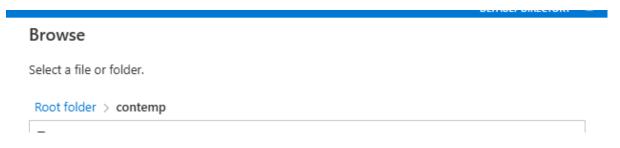


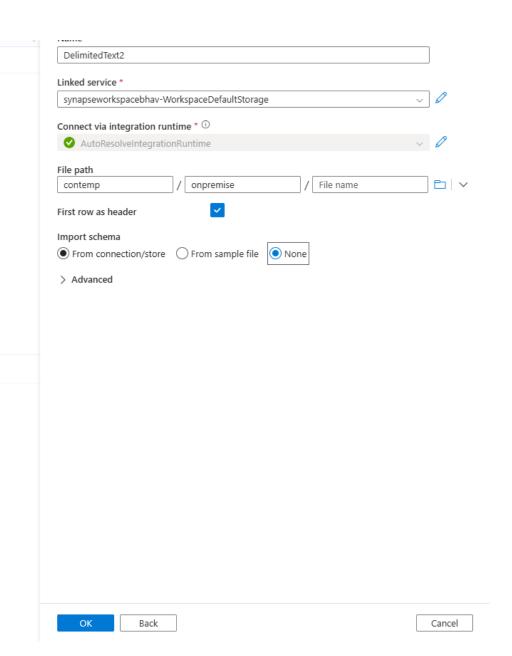
Select .csv instead of .txt from file extension feild as below



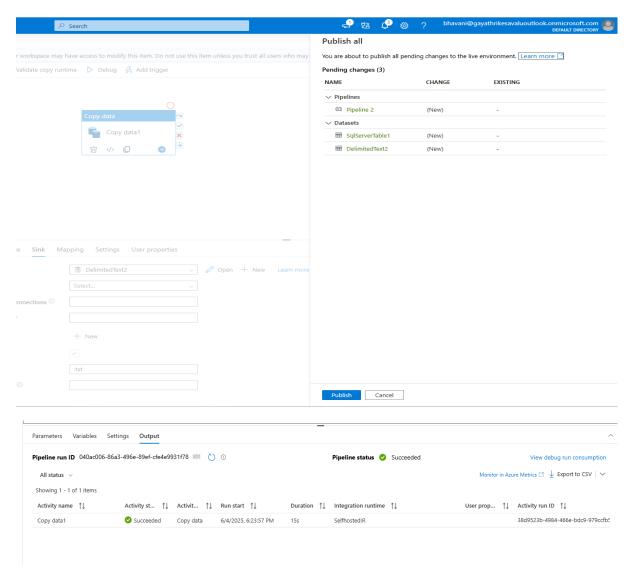
I will select my container and click ok. Then give onpremise and click none as it has to craete new folder in fen2 for copy data.

NBote: Here in sink tab we are selecting auto resolve IR because it is a cloud

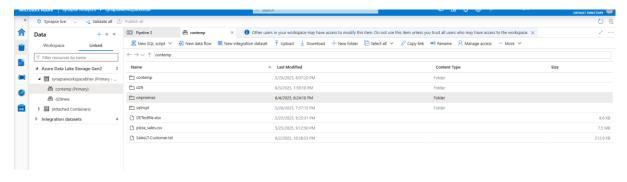


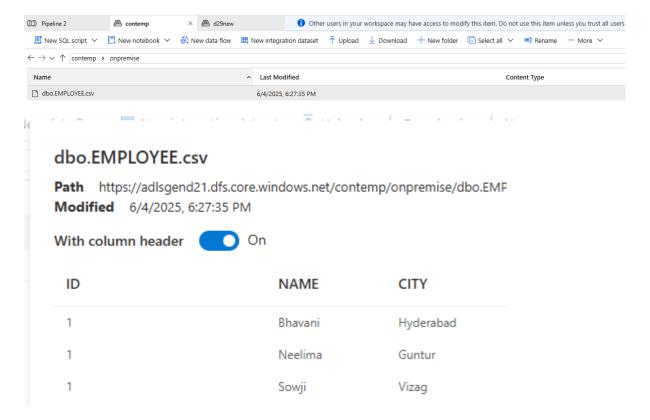


Finally do Publish and debug



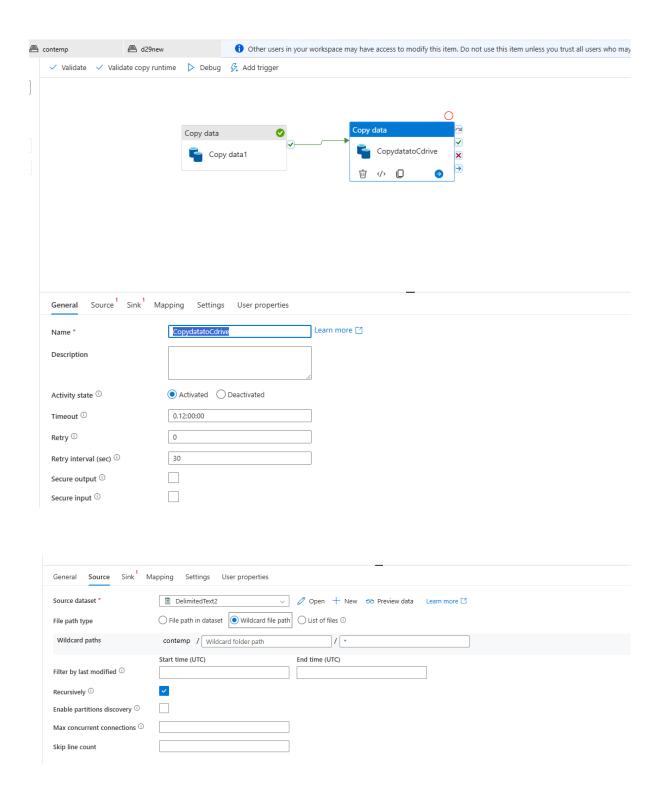
To verify whether 'onpremise' folder is craeted and data has copied, go to data tab then select your synapse workspace then go to your container

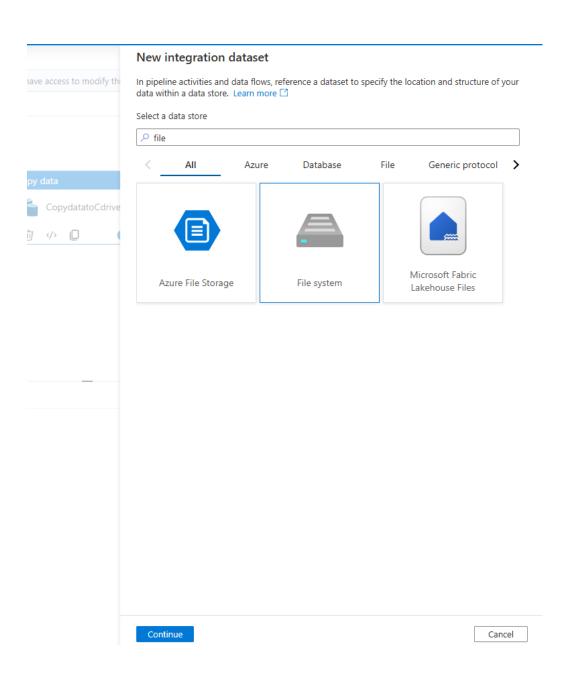


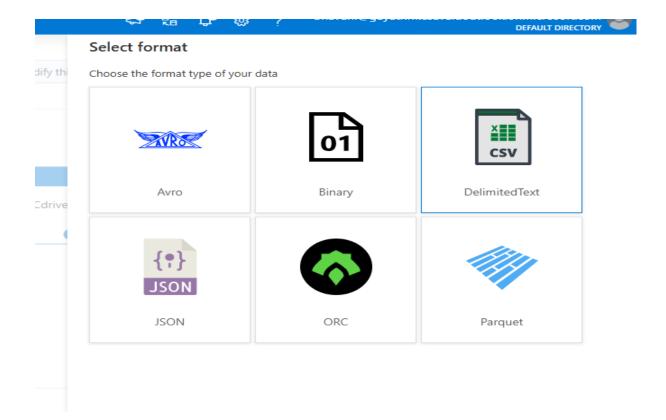


Now copy data from Cloud to Onpremise local folder c drive folder

- Take another pipeline and do onsuccess connection of 1st pipeline output to new pipeline and enter the name as copydatatoCdrive from General tab of pipeline2
- And also select same source dataset delimitedText2 which we used in pipeline1 sink tab as i want to copy same loaded onpremise data to VM C drive folder
- Note: in the sink tab for loading into another machine c drive that is called as File system connector we have to use

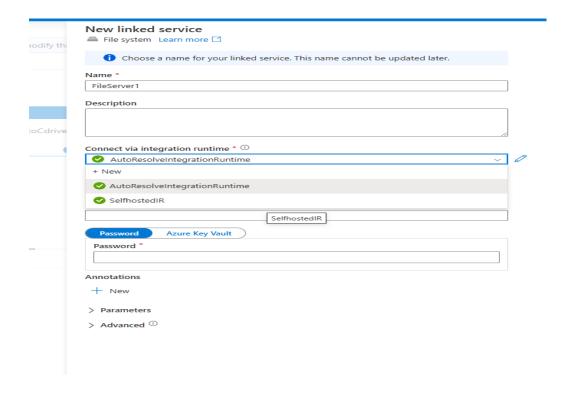


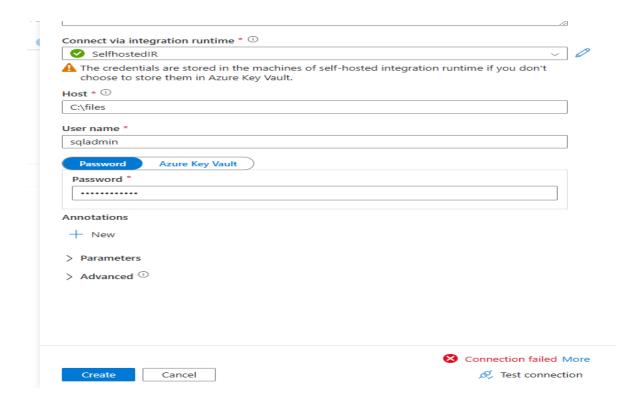




Select IR as Self Hosted IR and Test connection.

It has to fail as there is a bug which is not fixed by microsoft





Go to below link and copy the commands and execute them in your VM power shell as below to resolve this issue

Command 1 and enter

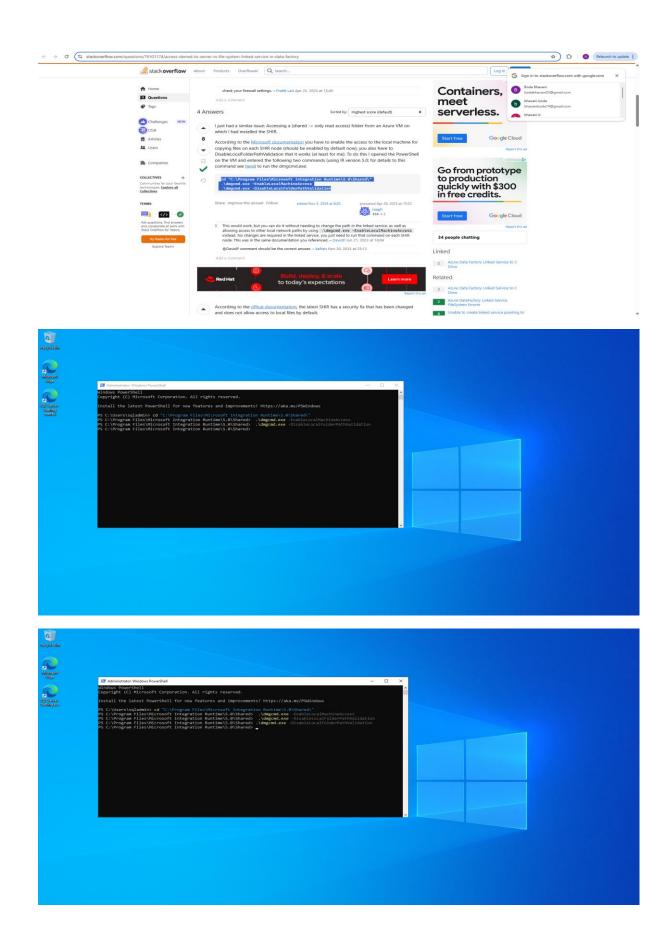
cd "C:\Program Files\Microsoft Integration Runtime\5.0\Shared\" .\dmgcmd.exe -EnableLocalMachineAccess .\dmgcmd.exe -DisableLocalFolderPathValidation

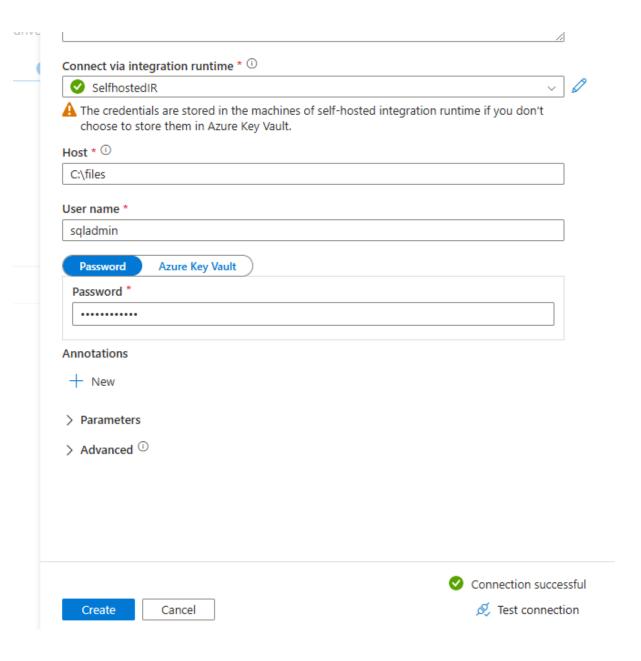
Command 2 enter

.\dmgcmd.exe -DisableLocalFolderPathValidation

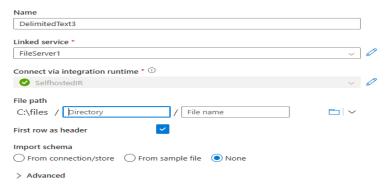
Then now go to your synapse worspace and Test connection now it will success and click on create to craete LS

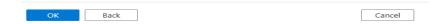
Change .csv from .txt in the sink tab file extension field and clcik on publish then debug

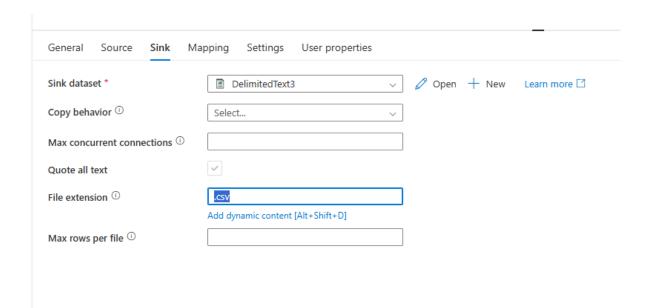




Set properties





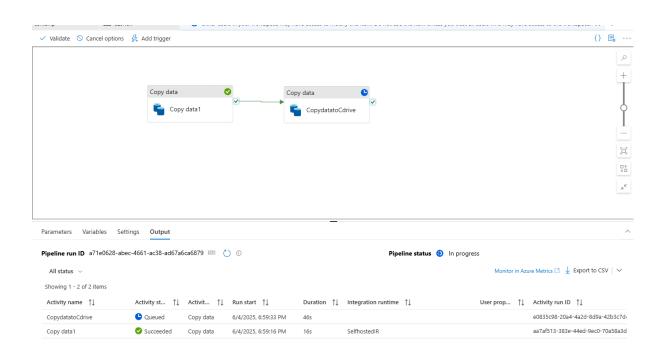


Publish all

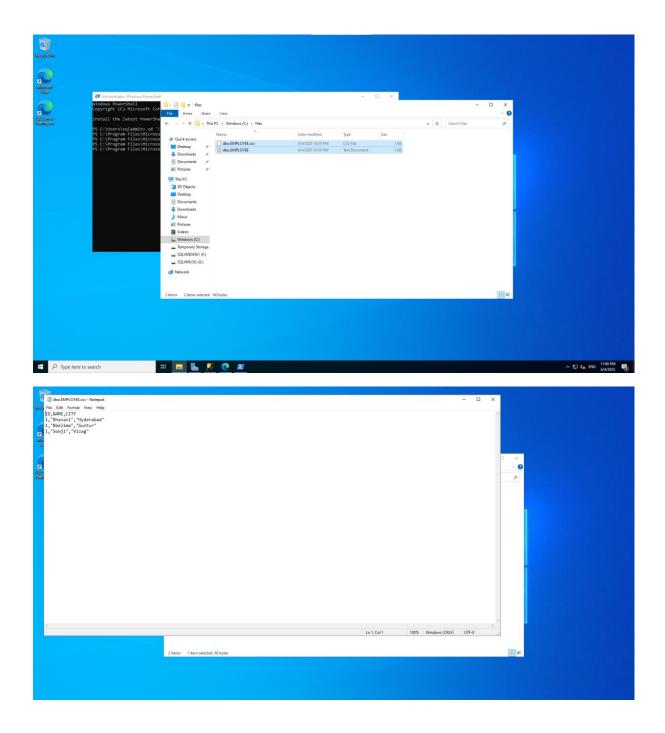
You are about to publish all pending changes to the live environment. Learn more

Pending changes (2)

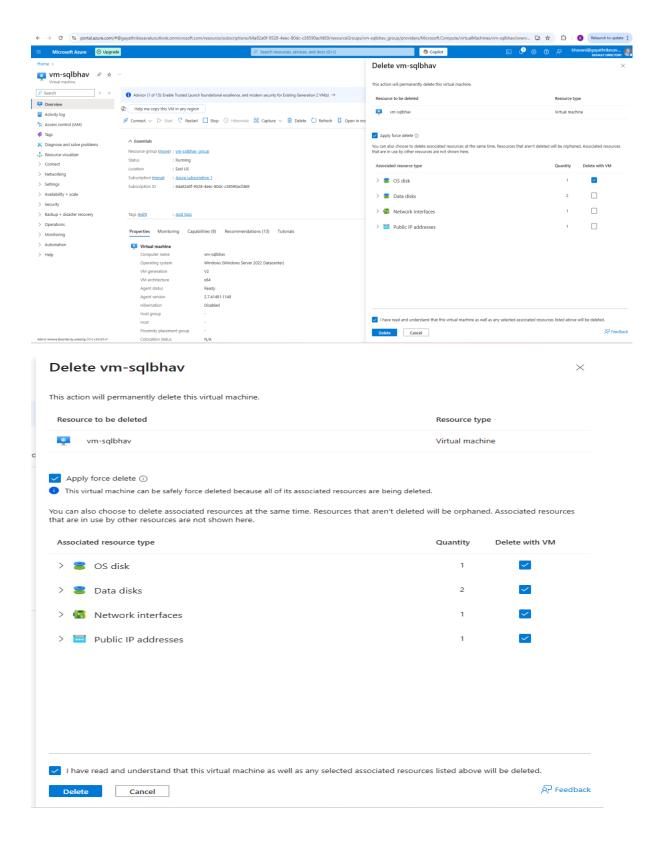
NAME	CHANGE	EXISTING	
∨ Pipelines			
OD Pipeline 2	(Edited)	Pipeline 2	
∨ Datasets			
■ DelimitedText3	(New)	-	



Files are loaded to VM C drive files folder successfully



Now for cost savings delete your VM, go to your VM from Azure and click on delete and do apply force delete, select all options and delete



Successfully deleted VM and to verify go and check on your resource group as below

