**CASE-STUDY (NUMBER - 9) – SOLUTION SUBMISSION**

BY

NAME – Shaik Mohammed Raihan

ROLL NUMBER – DXC262AB12036

BATCH – DXC-262-ANALYTICS-B12-AZURE

COMPANY – DXC TECHNOLOGY EMPLOYEE

DOMAIN –DATA-ANALYTICS-AZURE

TRAINING UNDER – MANIPAL PRO LEARN

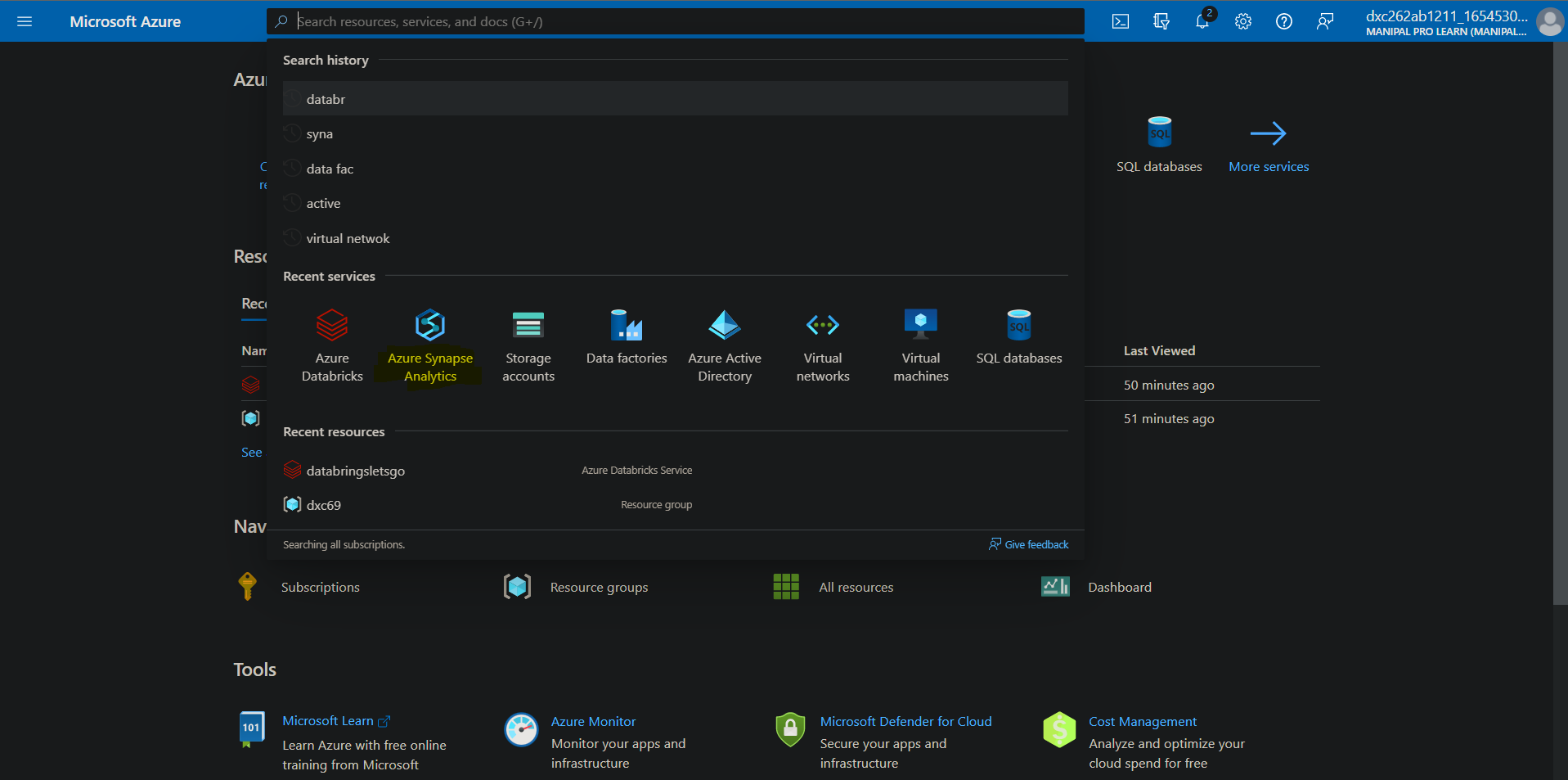
TRAINER NAME – MR. AJAY KUMAR

DATE OF SUBMISSION – 09 JUN 2022

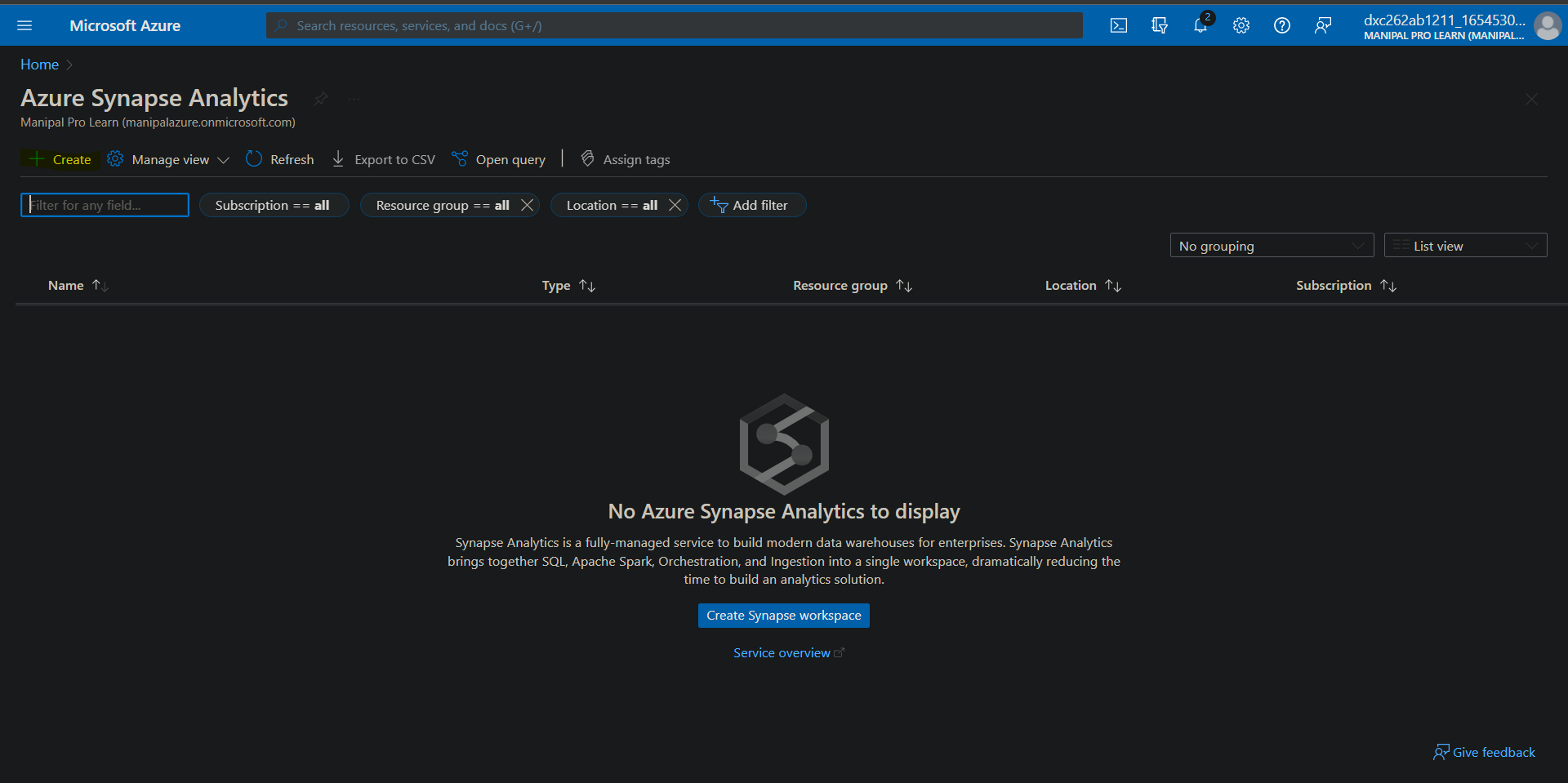
**TOTAL NUMBER OF QUESTION :-11**

1. Explain the steps with screenshots how to AzureSynapse analytics?

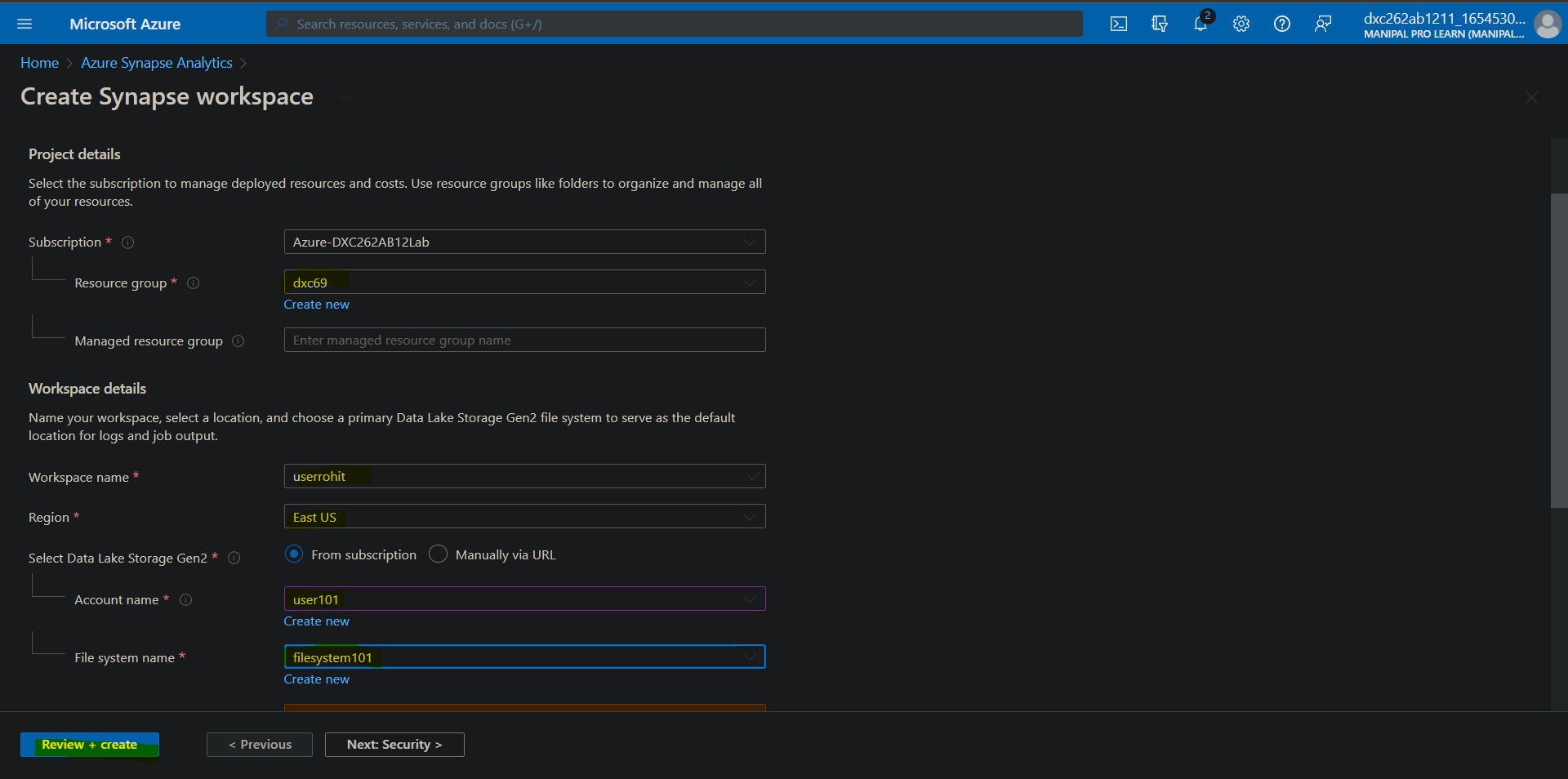
For creating AzureSynapse account on Azure first open the azure portal and search for Azure Synapse in the search bar



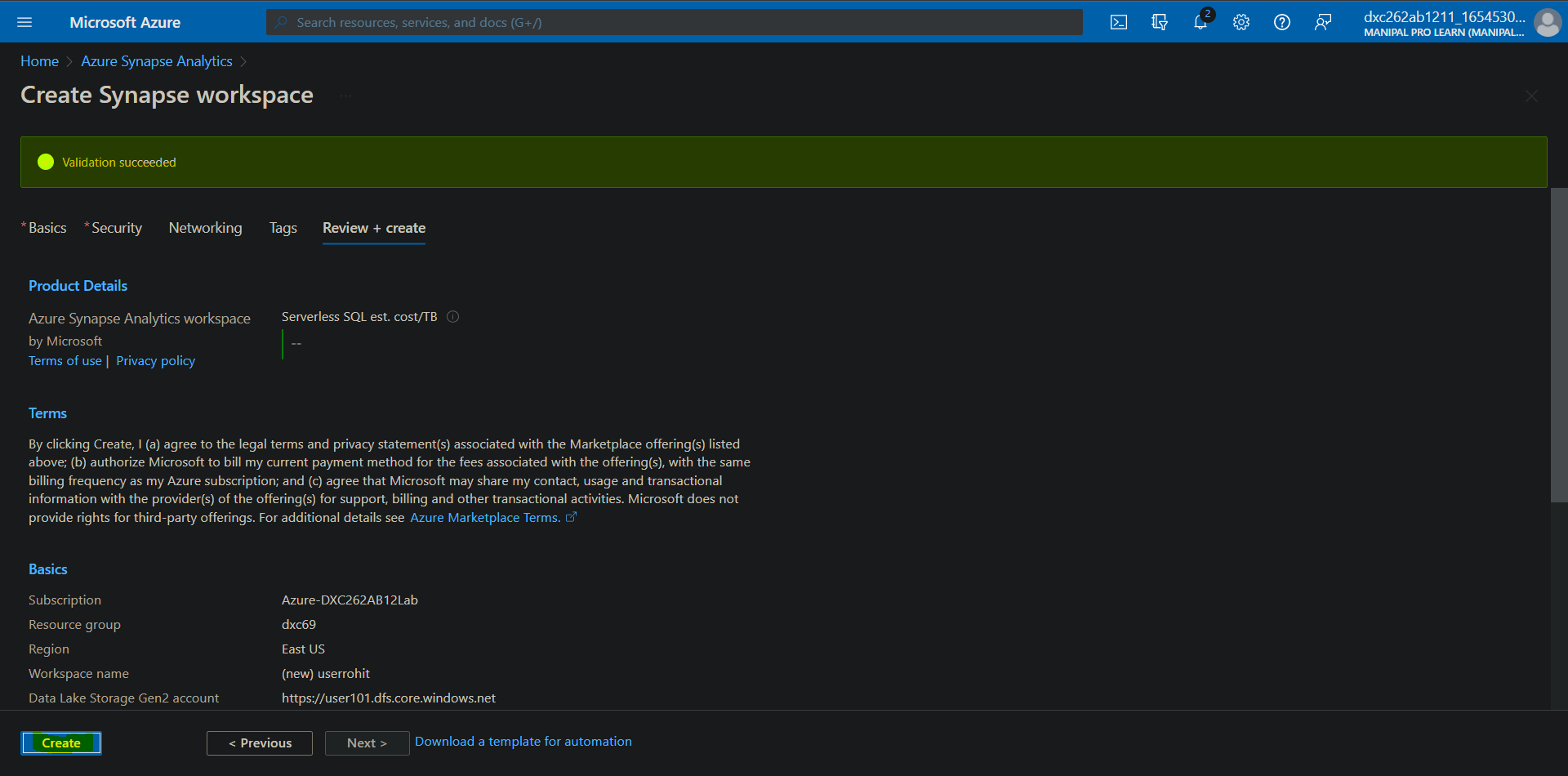
After synapse is opened click on + create



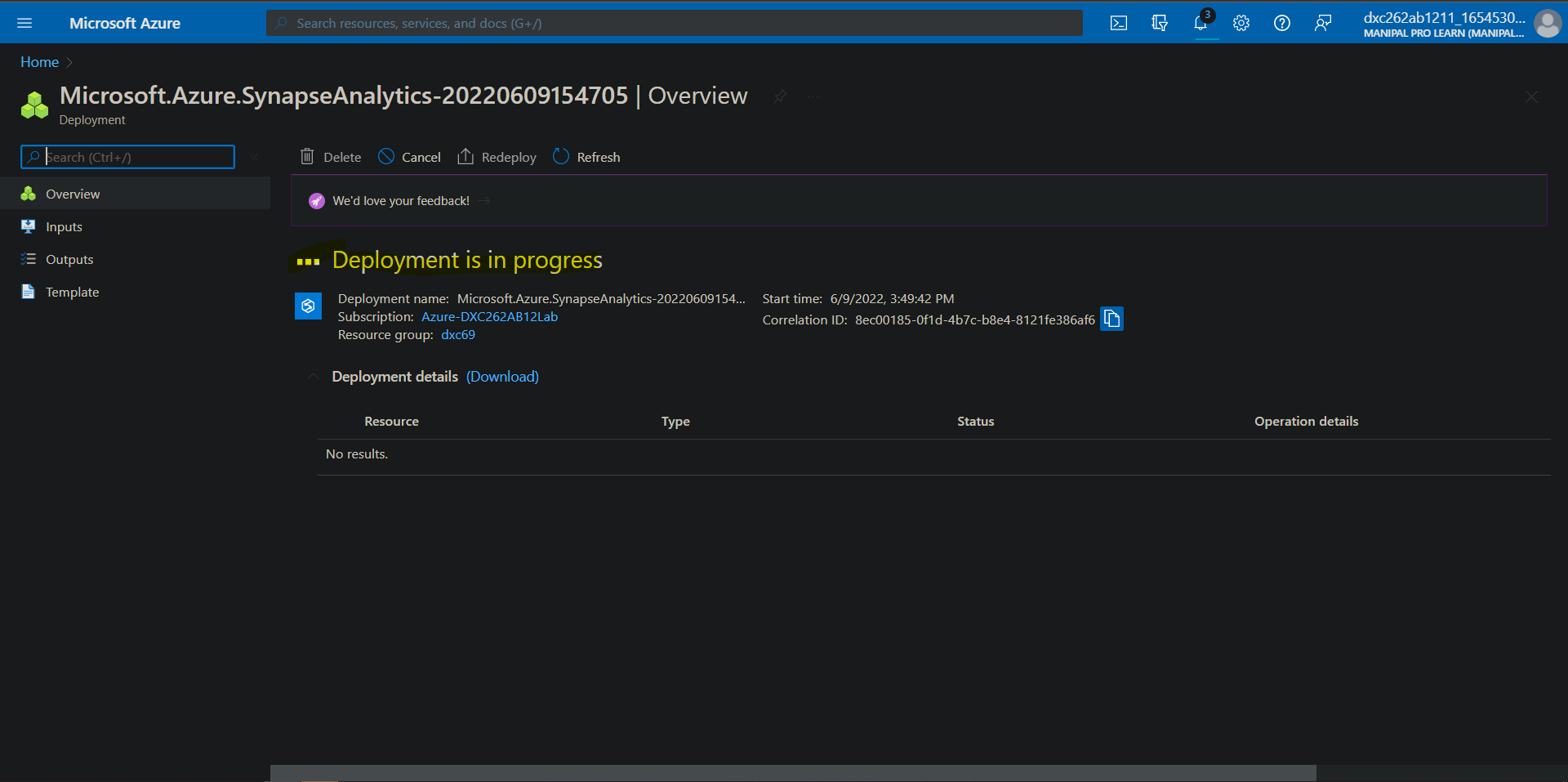
A window will be opened feed in the details of the azzure synapse account



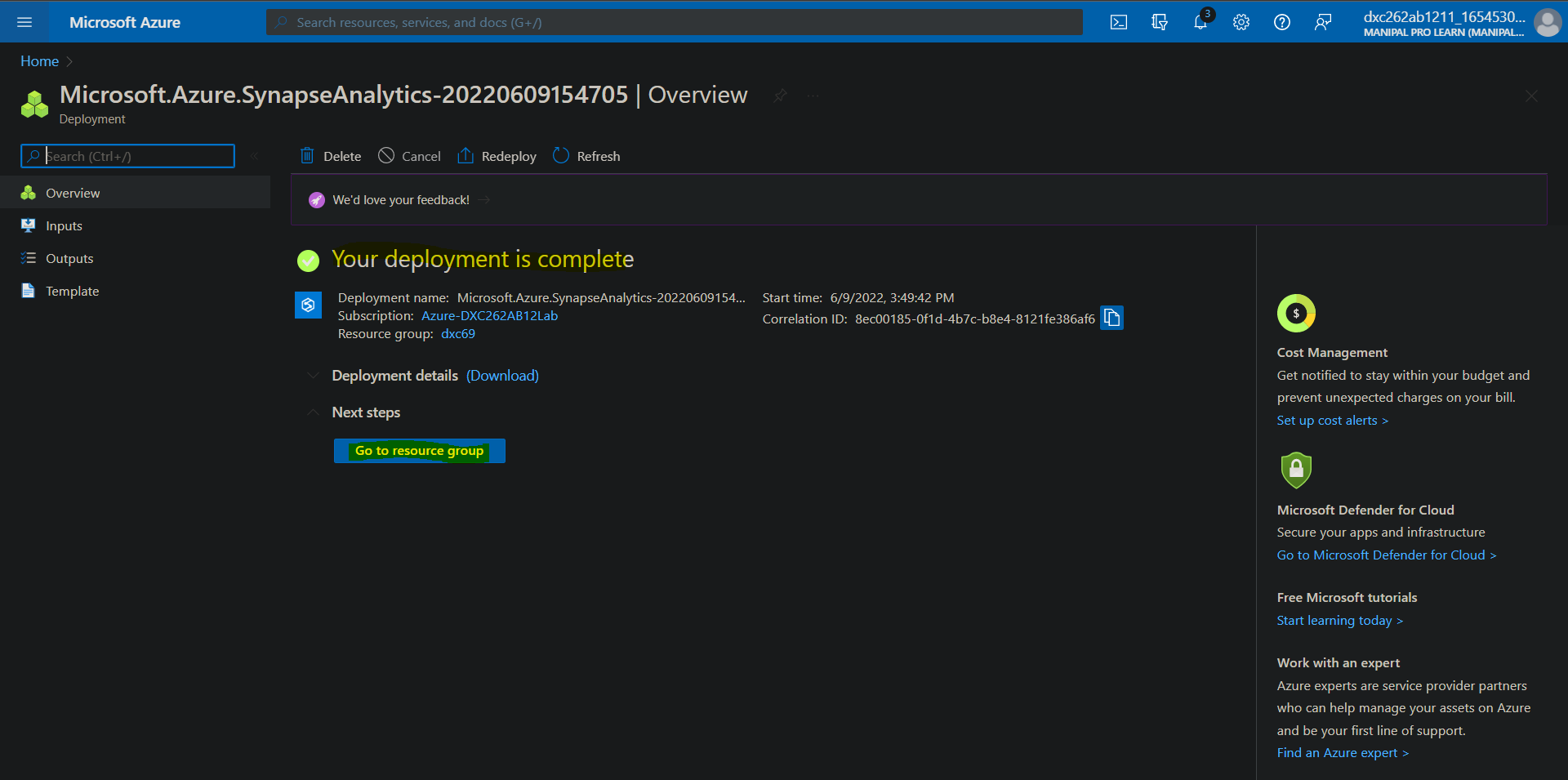
After feeding in the details click on validate to validate all the details you have filled.



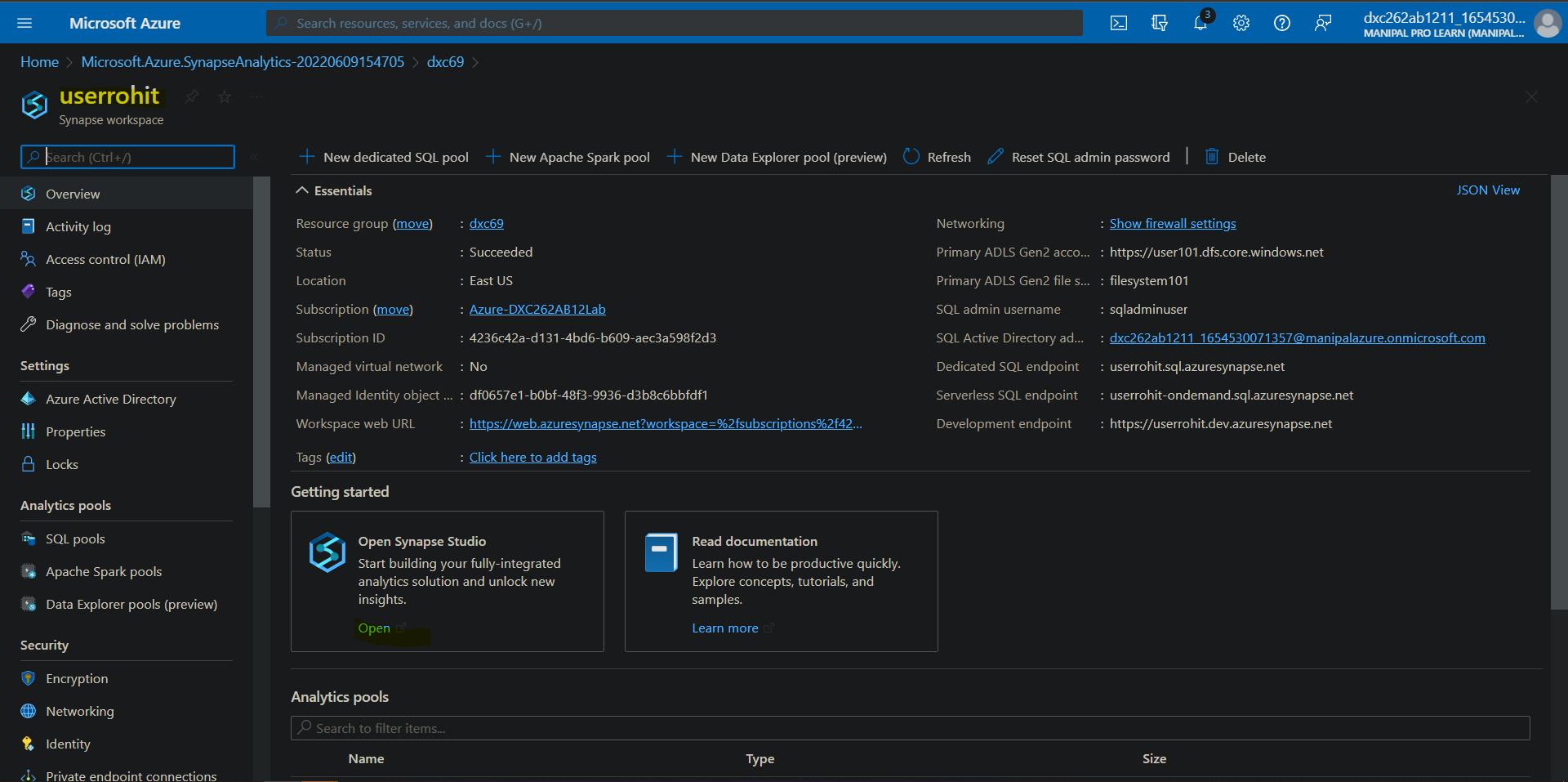
After validation is done successfully click on create



Wait for the deployment process to complete

2. Explain the steps with

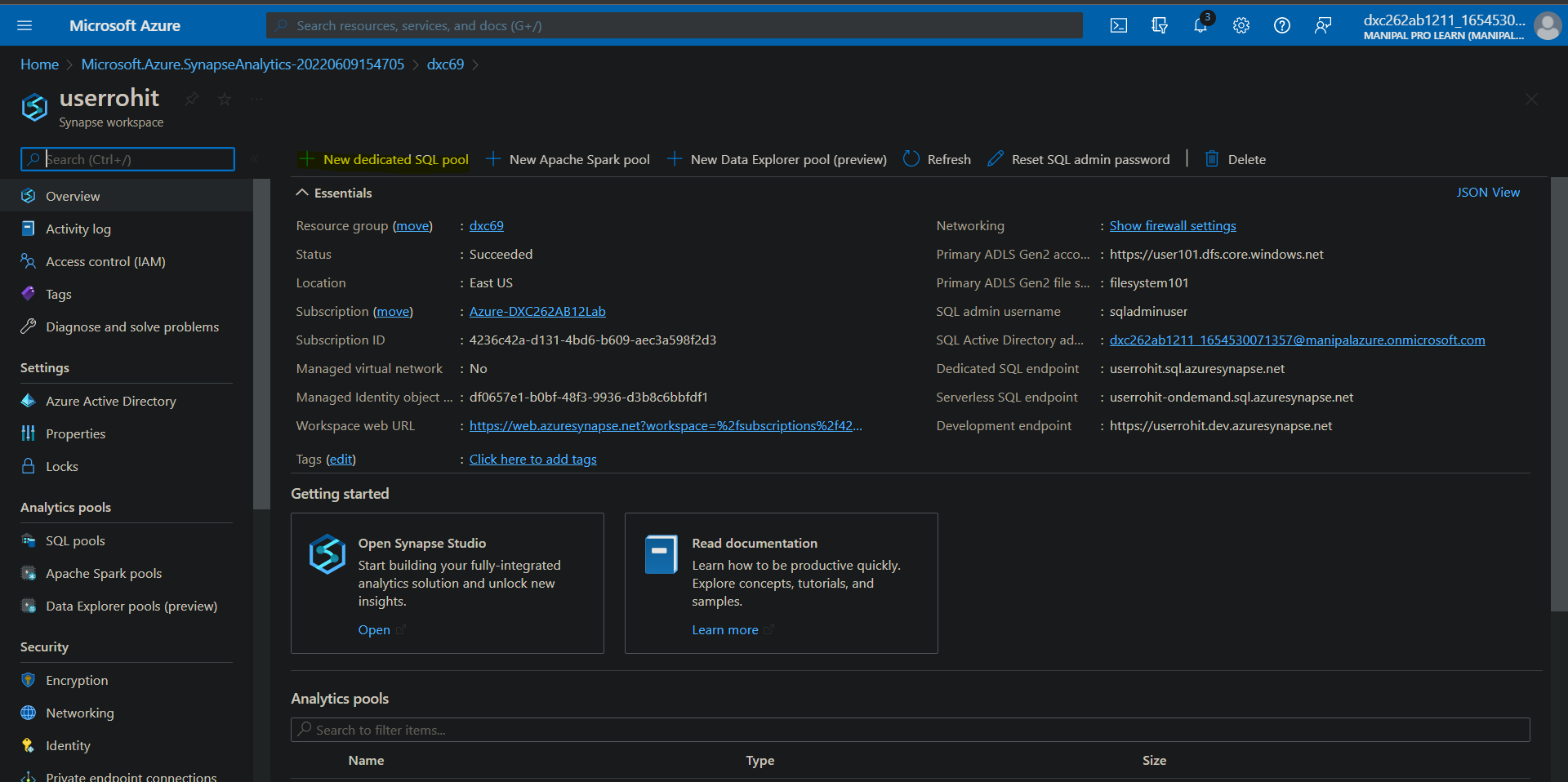
Click on Go to Resource group to see the deployment for the azure synapse



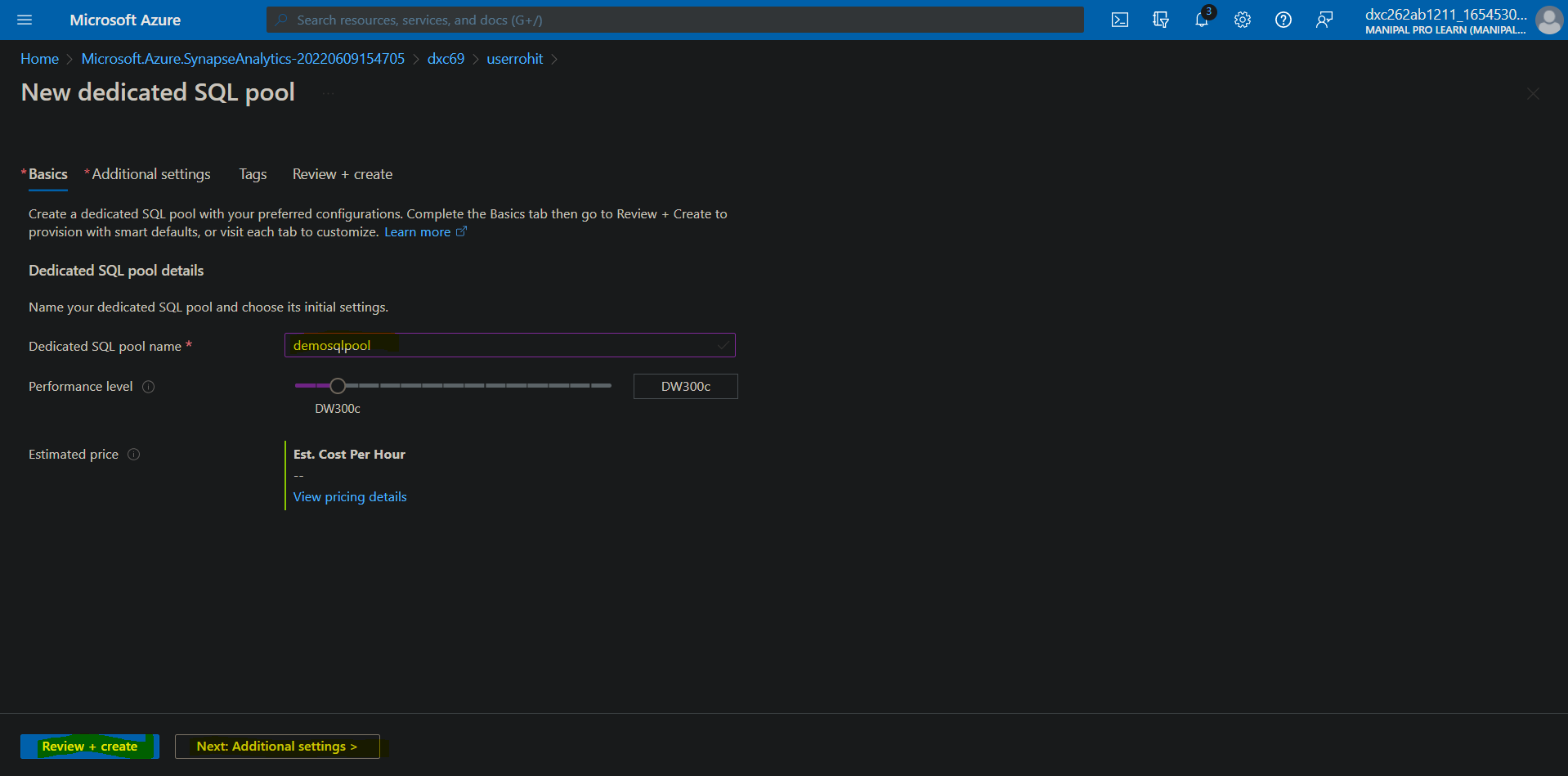
There you have created a new Azure Synapse account and you can navigate ‘Open’ to start working on it.

2. Explain the steps with screenshots how to SQL Pool in AzureSynapse analytics?

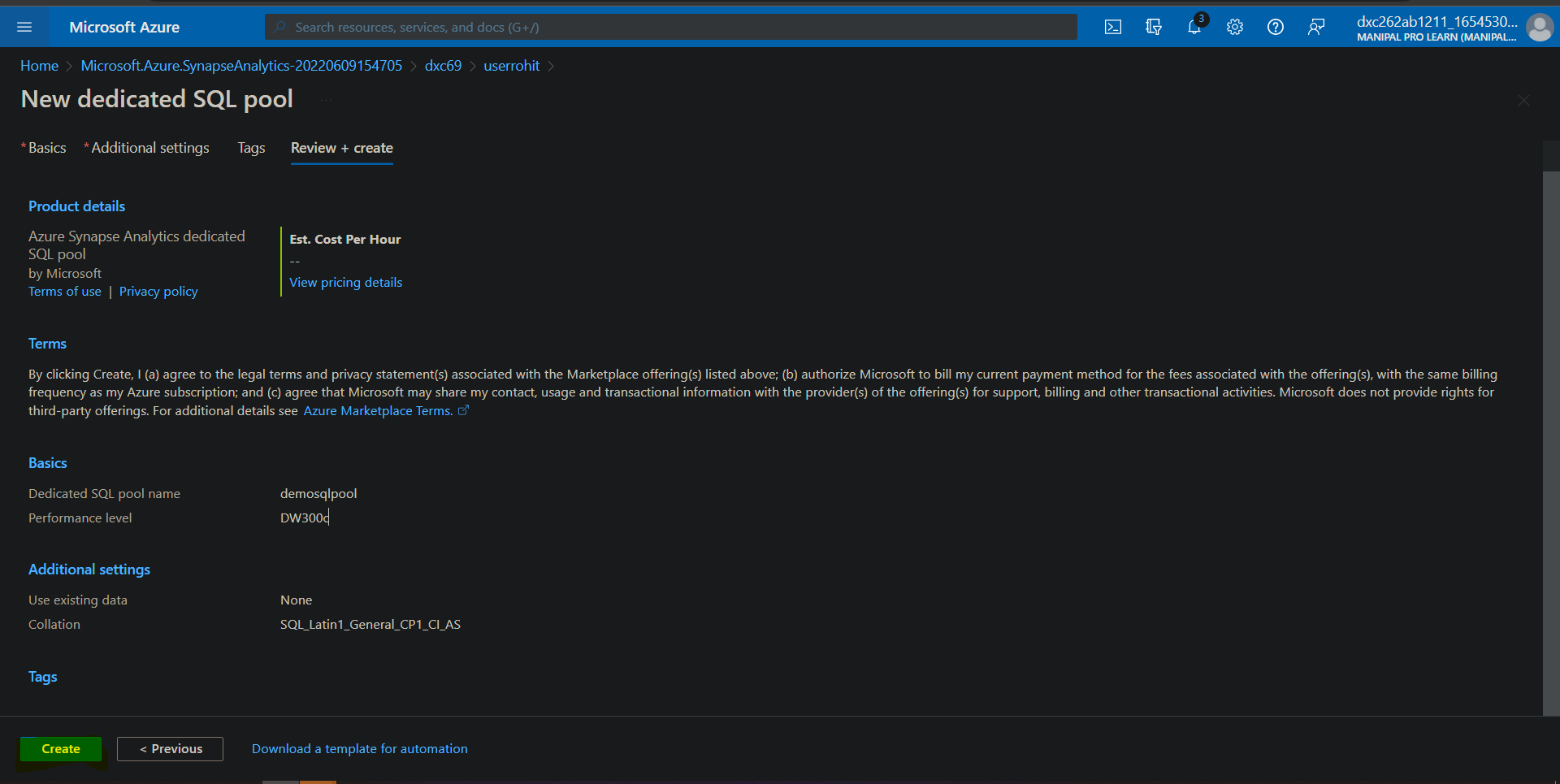
To make a SQL pool in synapse go to the resource group of azuresynapse and choose the account as we did in previous answer , Click on new dedicated SQL pool



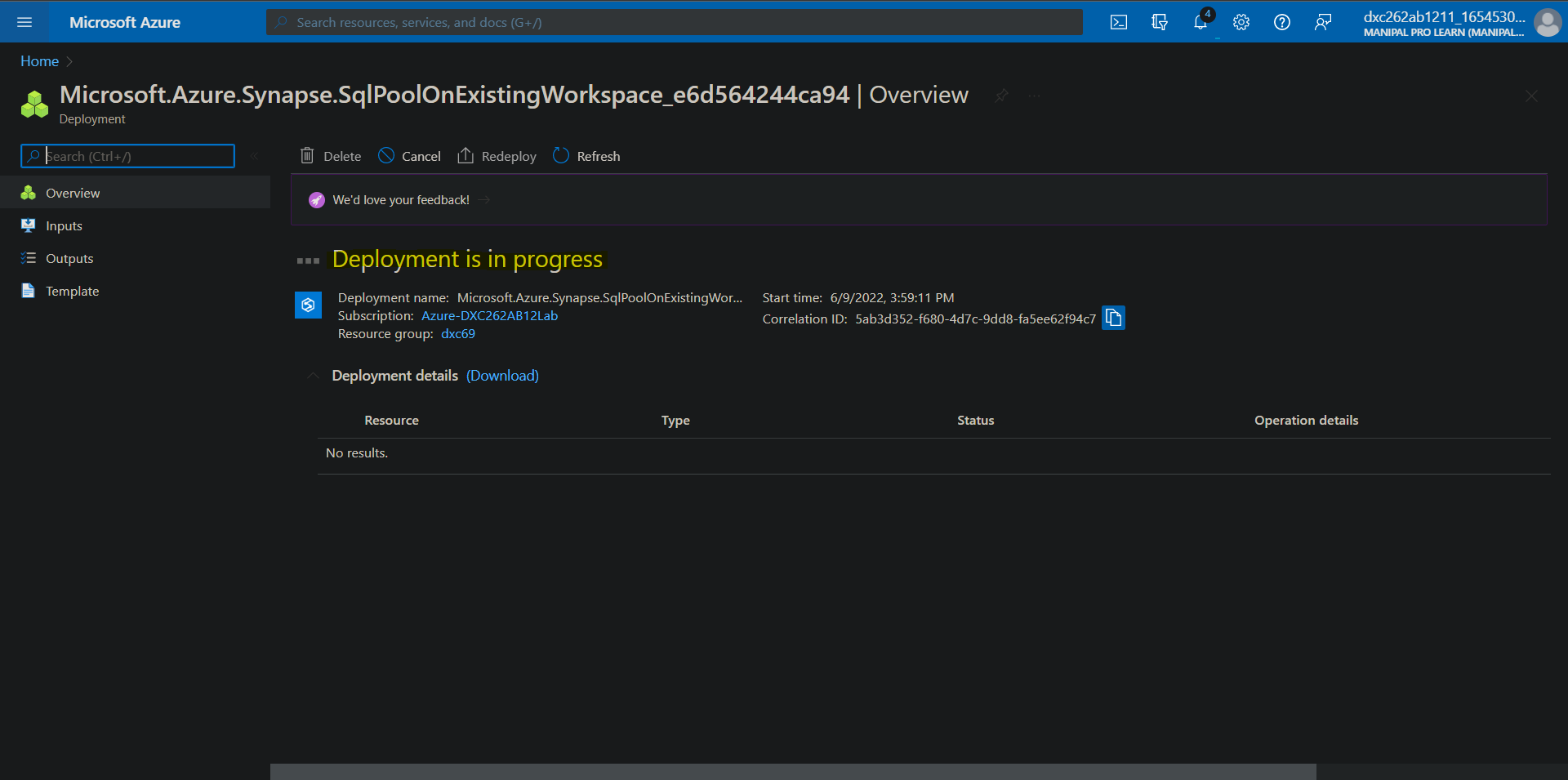
Feed in the details about the SQL pool and click on review + create option



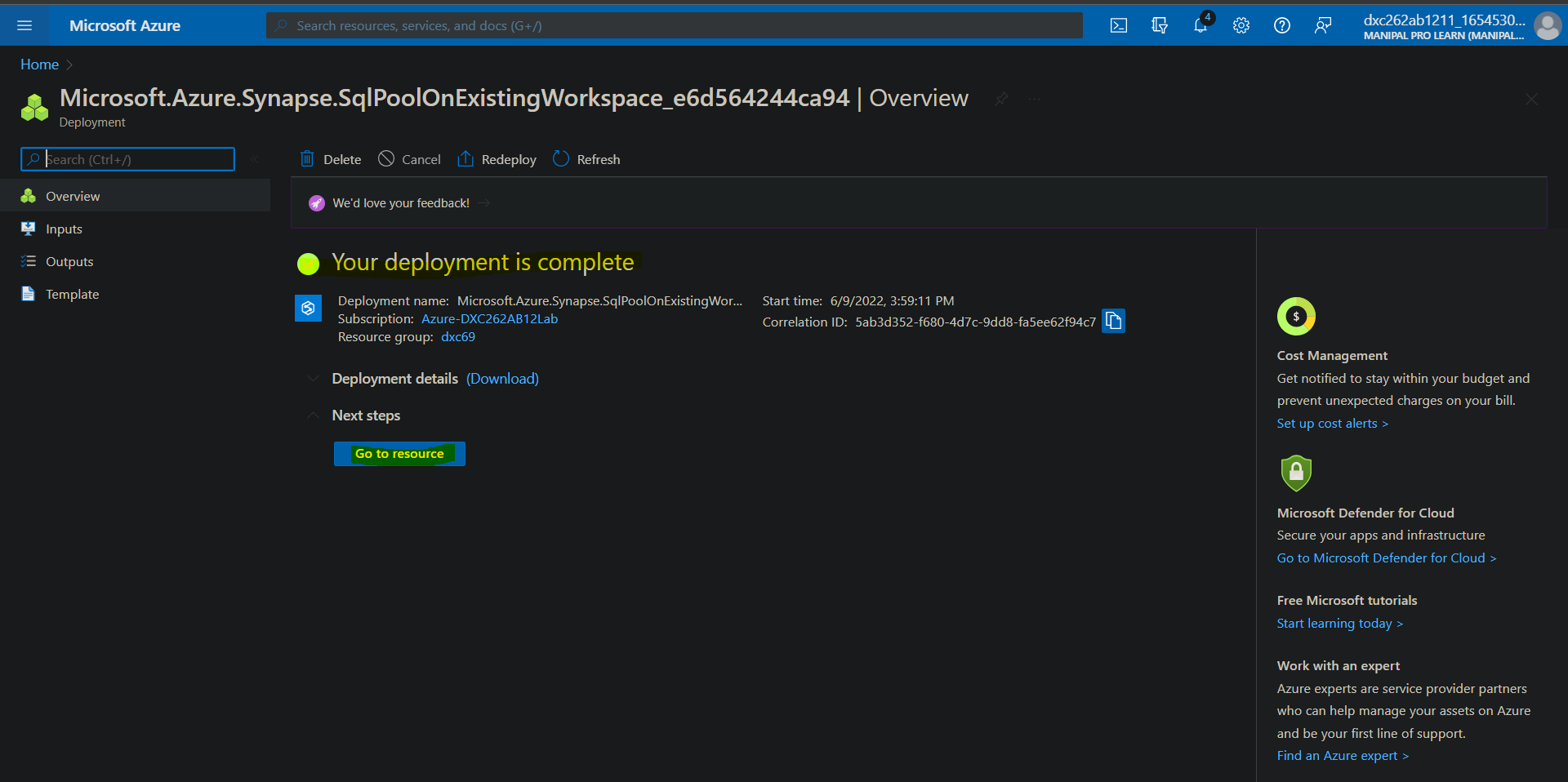
After the validation process is successfully complete , click on create



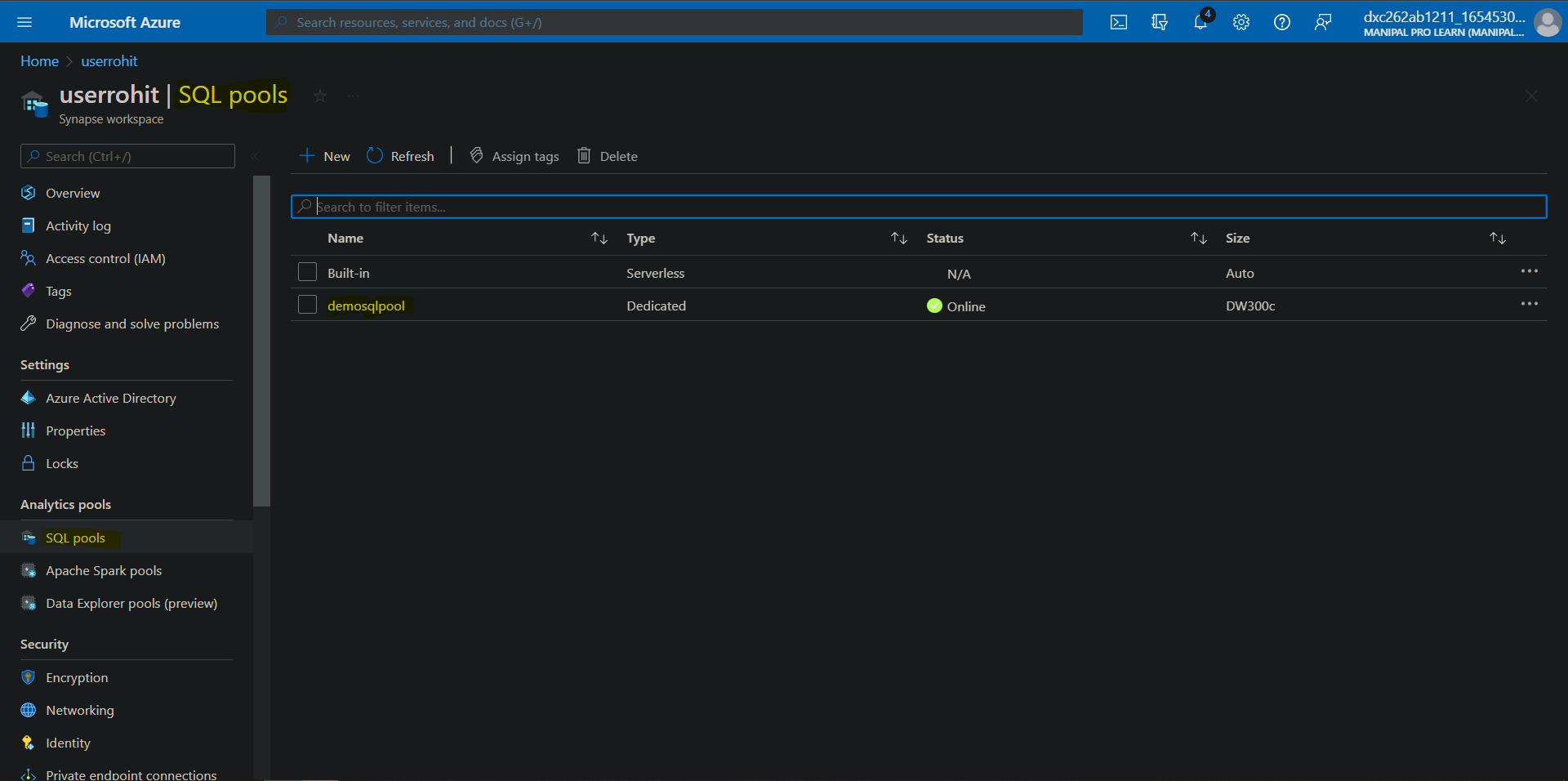
Let’s the deployment process finish



After the deployment process is finished click on go to resource

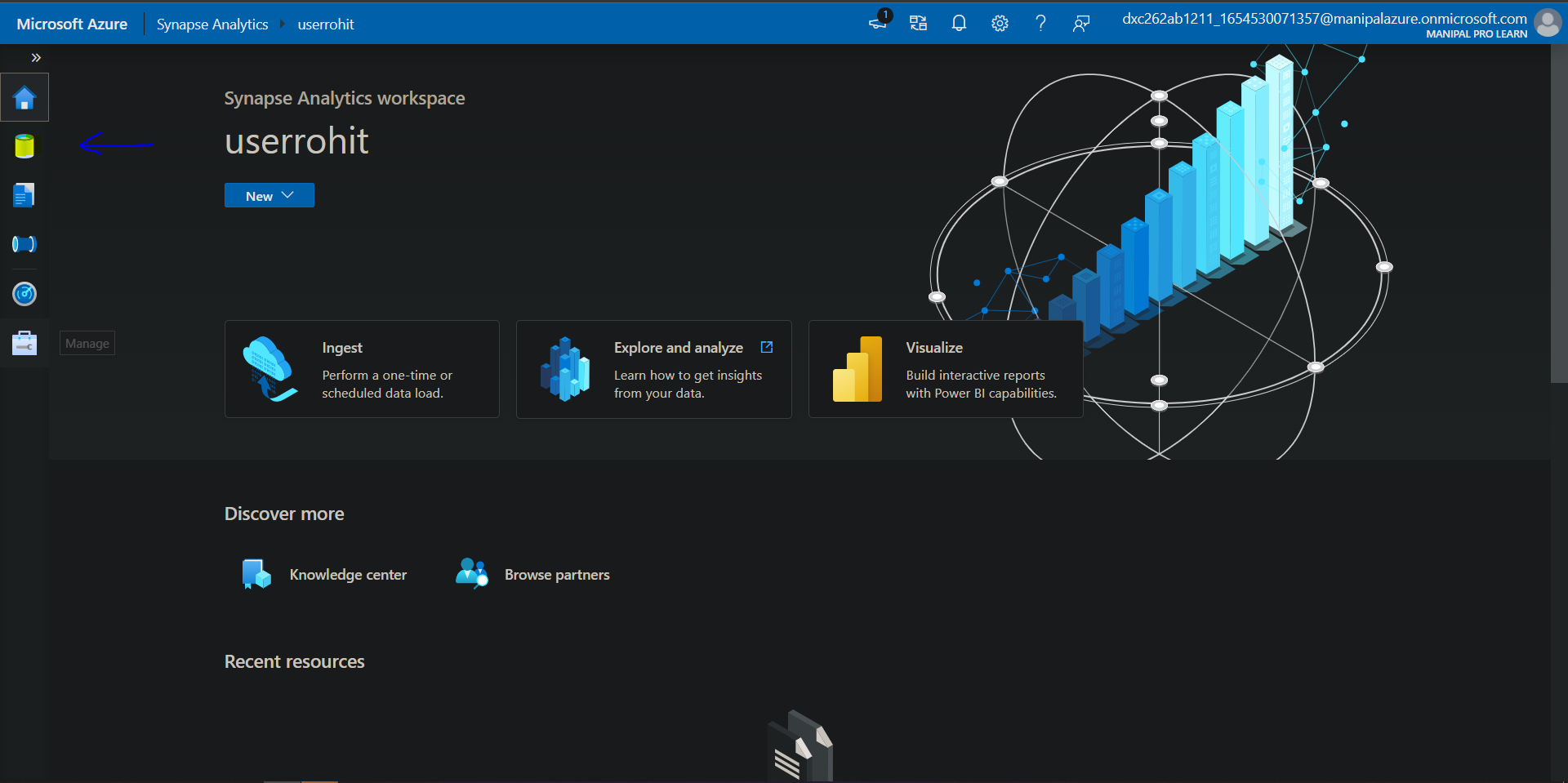


And here we have successfully created a SQL pool

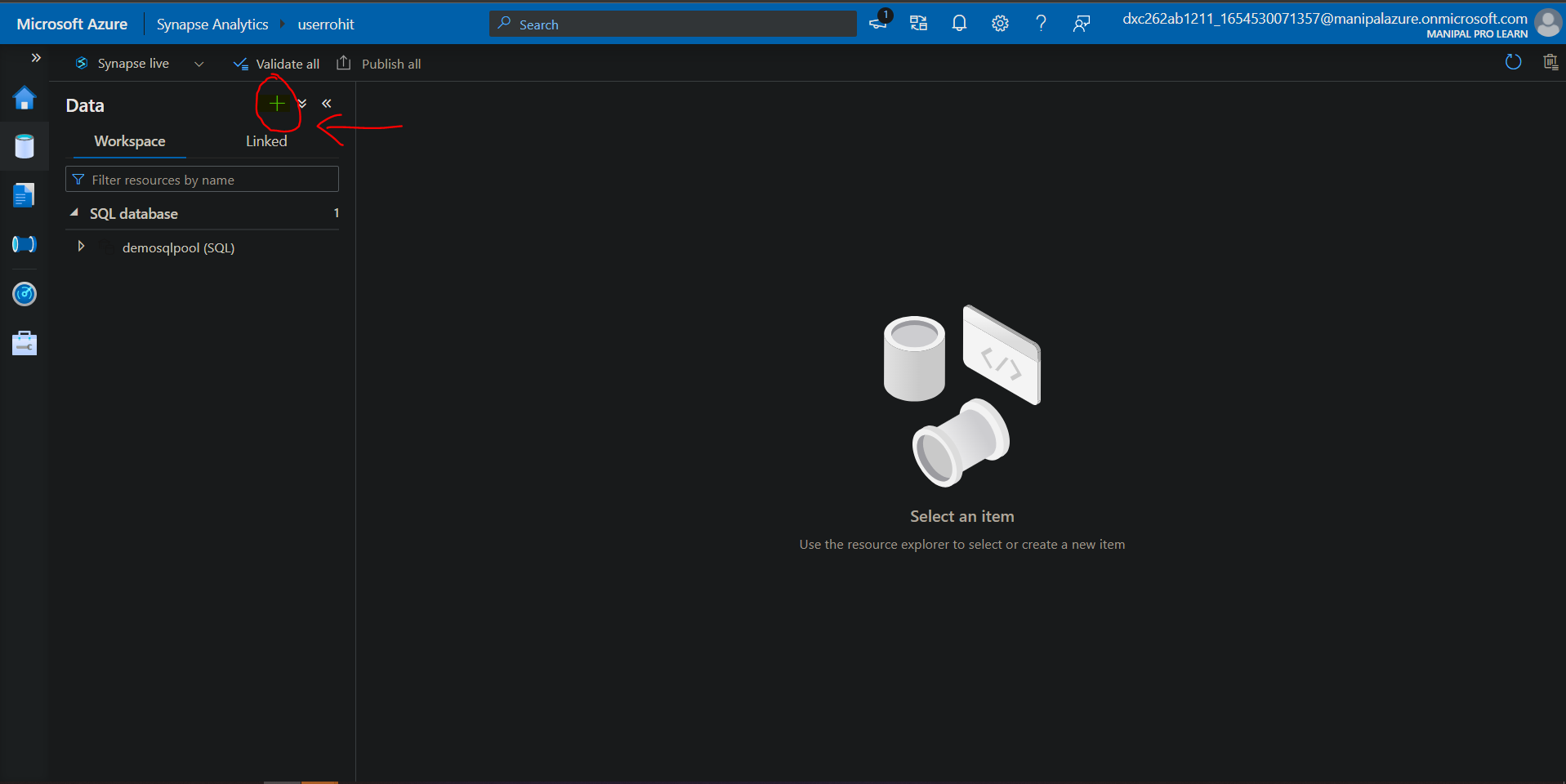


3. Explain the steps with screenshots how to import COVID19 dataset in AzureSynapse analytics and run sample 500 rows & display the output?

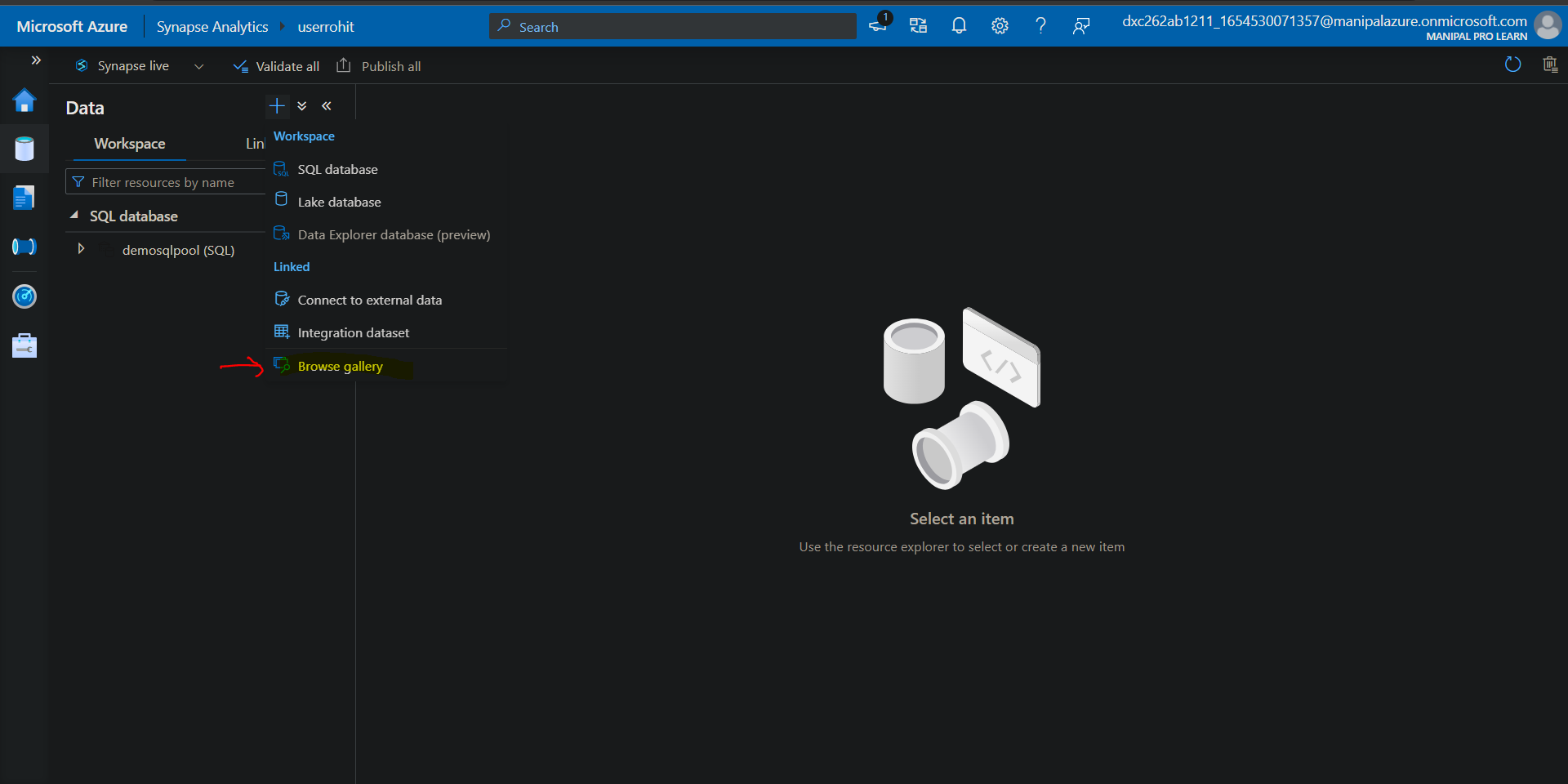
Firstly open Azure Synapse studio and navigate to the data section



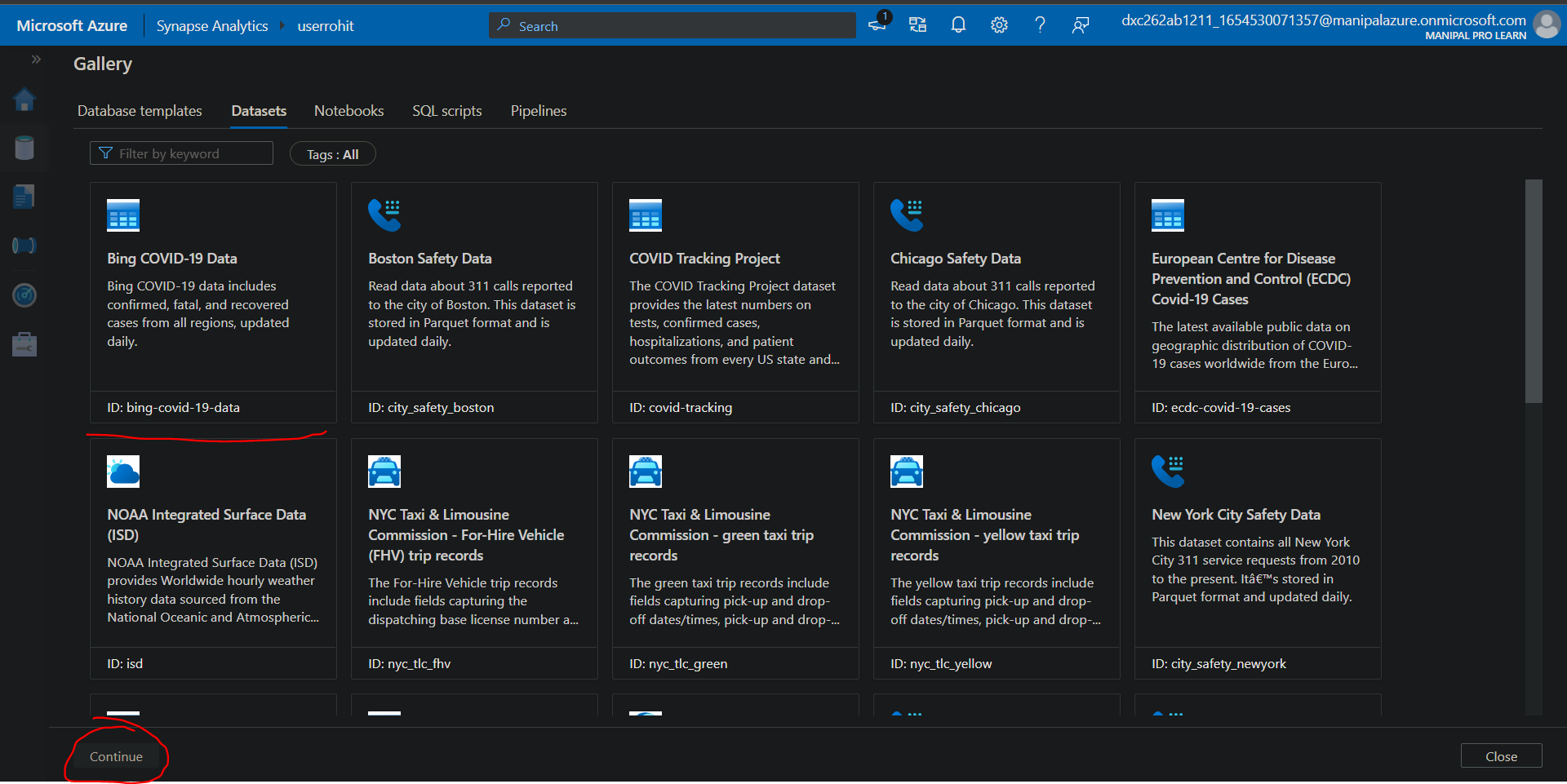
Click on + sign to add the sample data



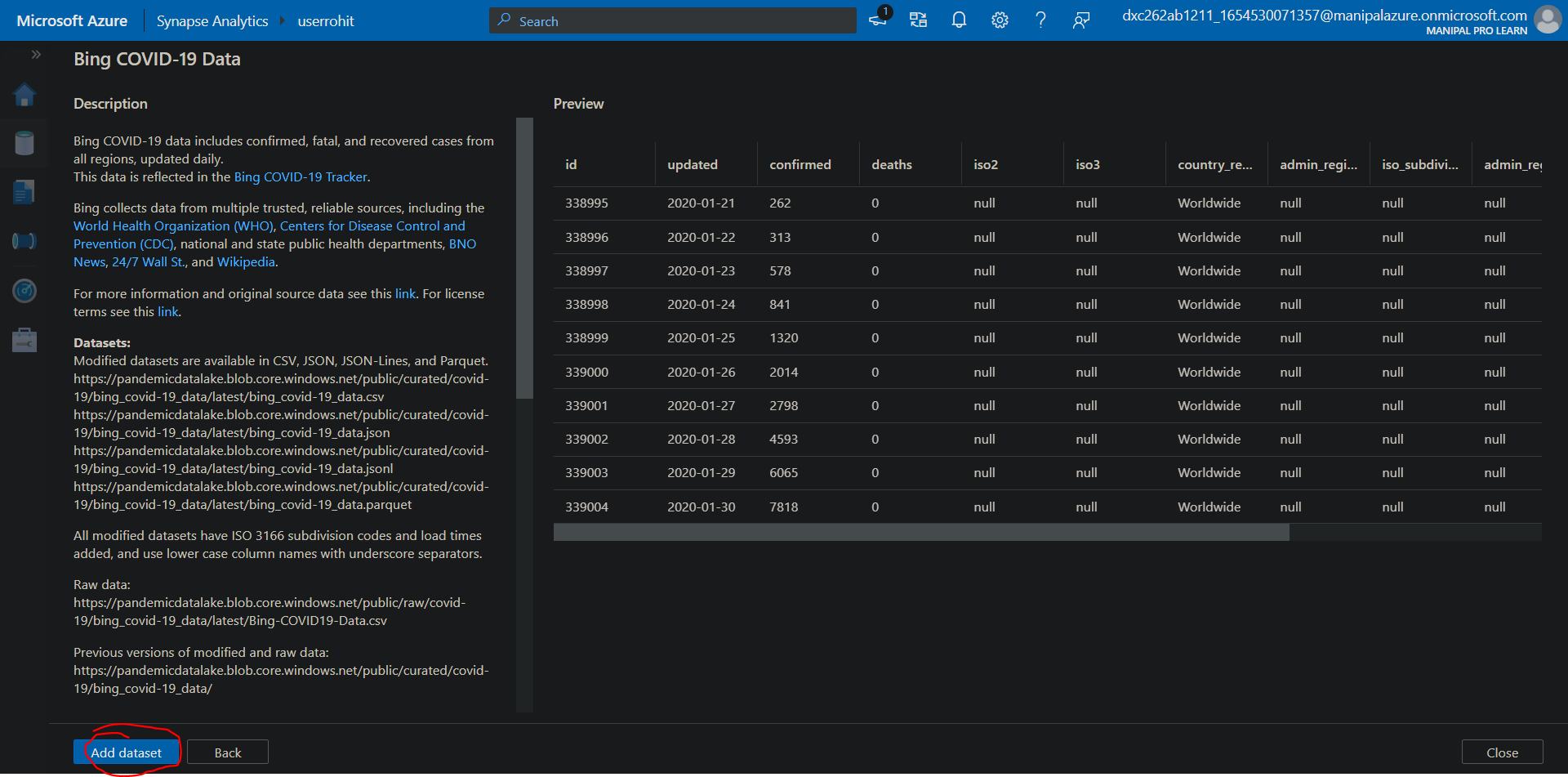
Go to the browse gallery



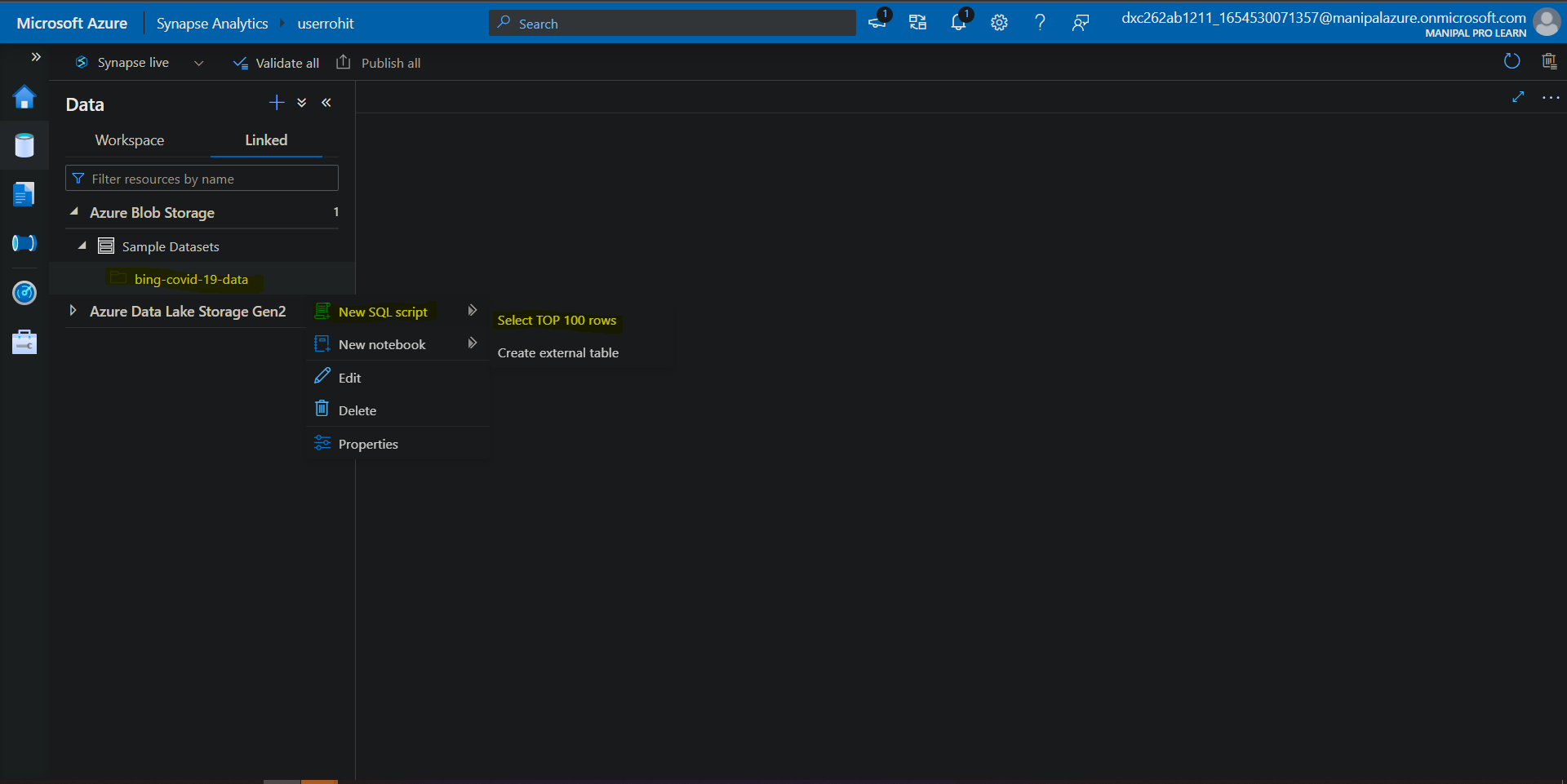
Choose the dataset you want to work on here we will choose covid 19 data



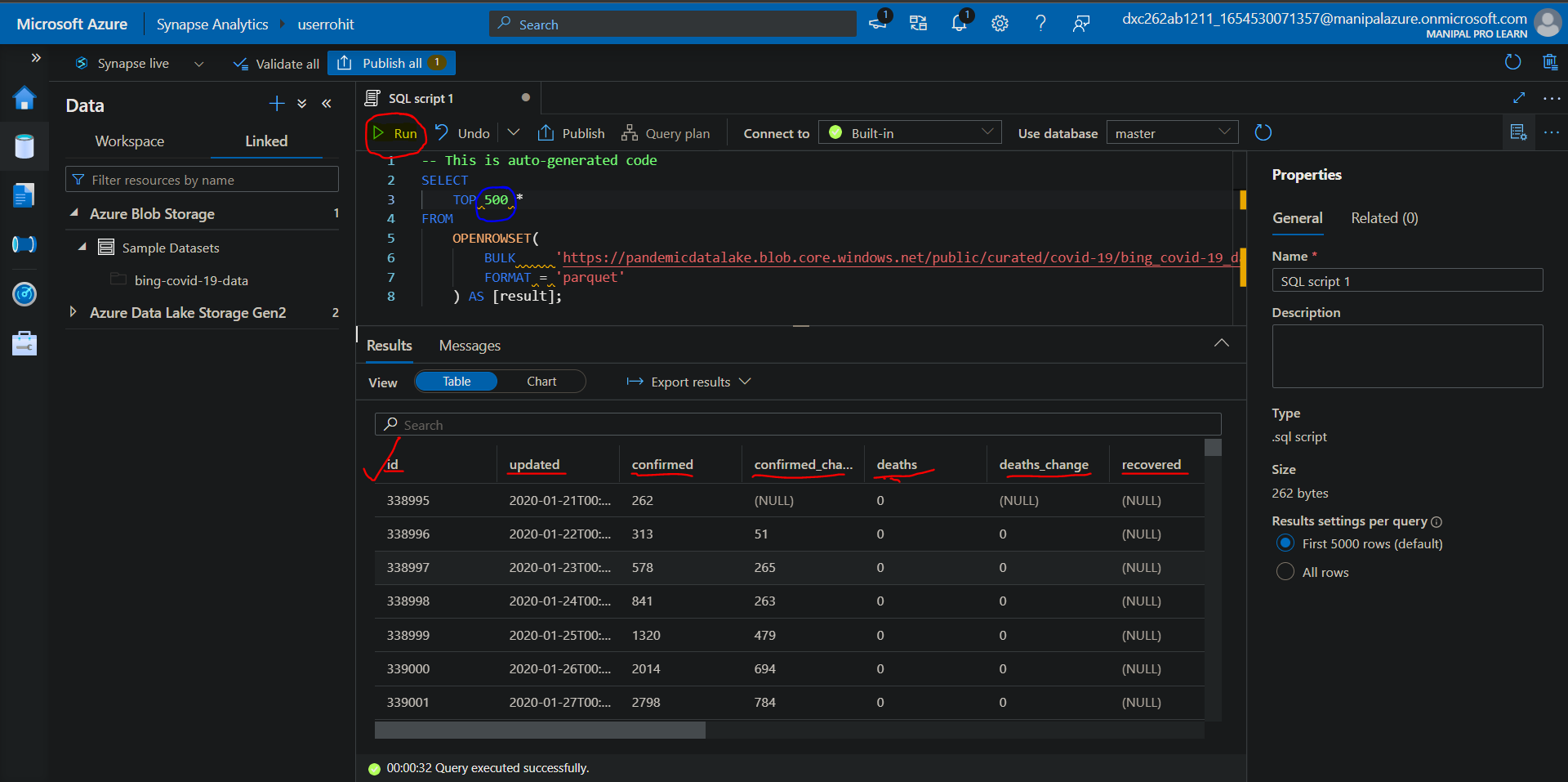
Preview the data and click on Add dataaset



After the dataset is added navigate over creating a new SQL script , dataset > new SQL script , Select top 100 rows

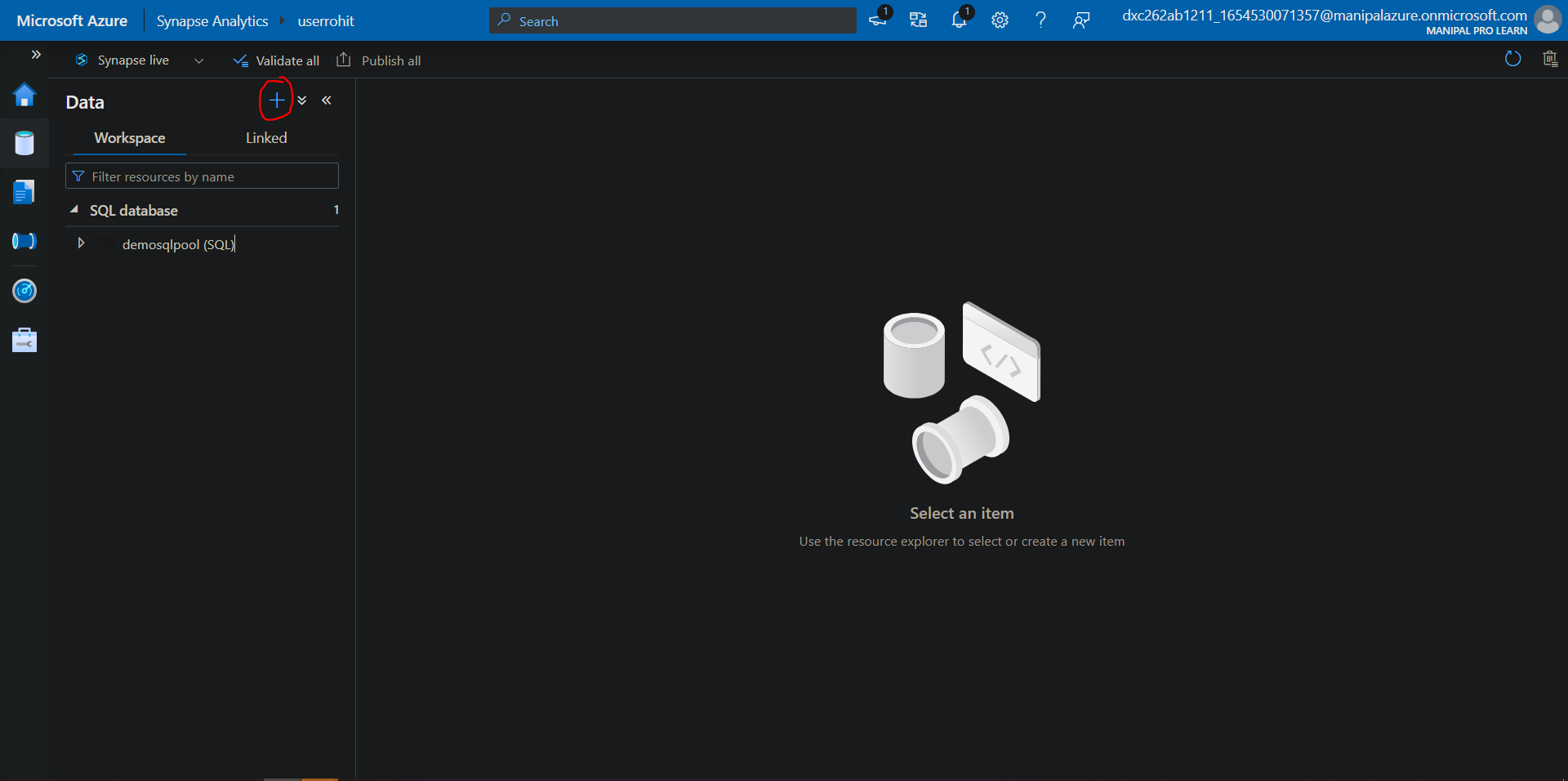


This will create a sample query to print top 100 rows of the data and as we create a SQL pool in the last question we can use that pool to provide session in order to run the SQL script, and moreover we need to print 500 therefore manipulate the number to 500 and here we have our results.

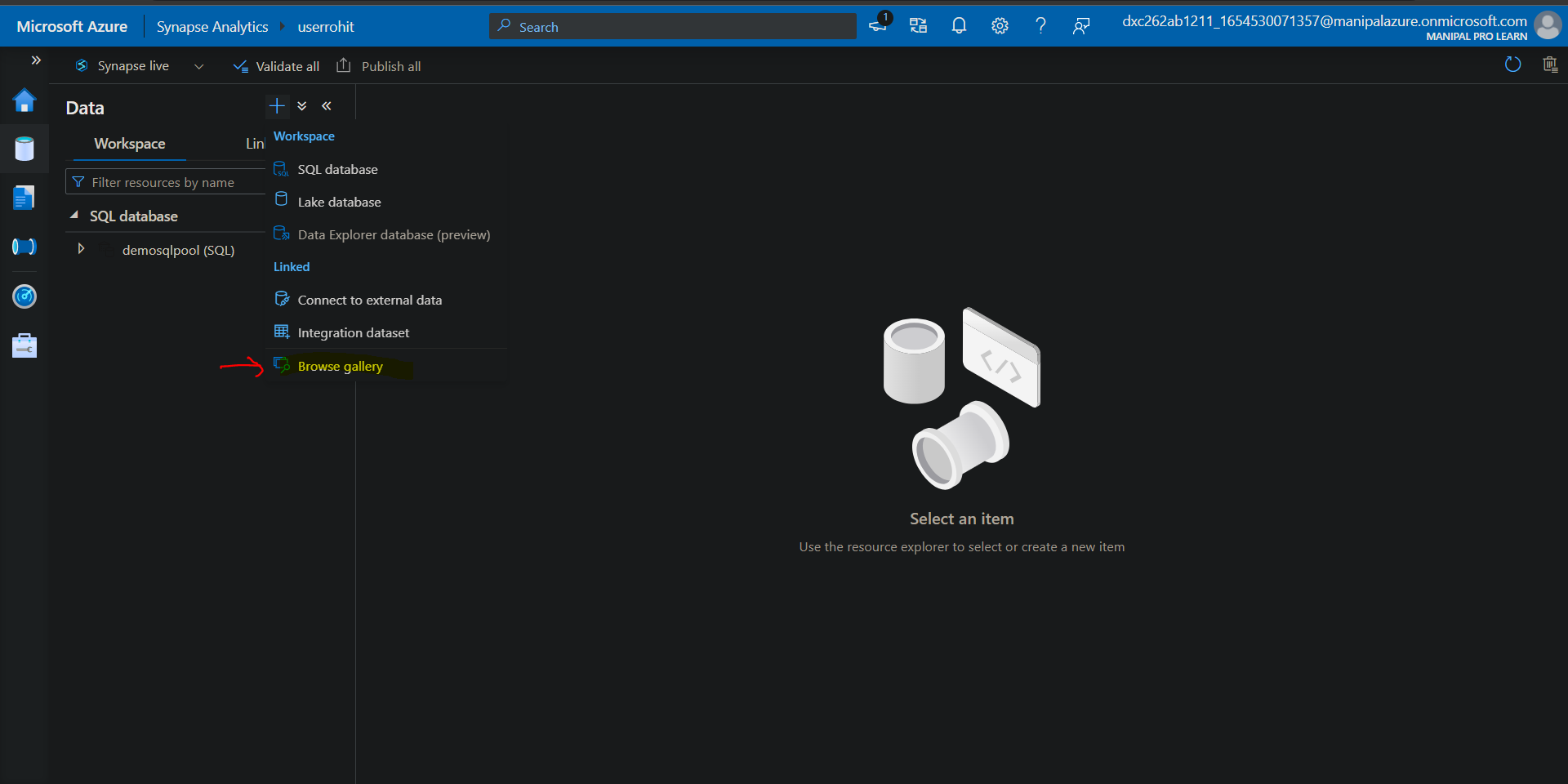


4. Explain the steps with screenshots how to input Boston Safety datasets into AzureSynapse analytics using Notebooks ?

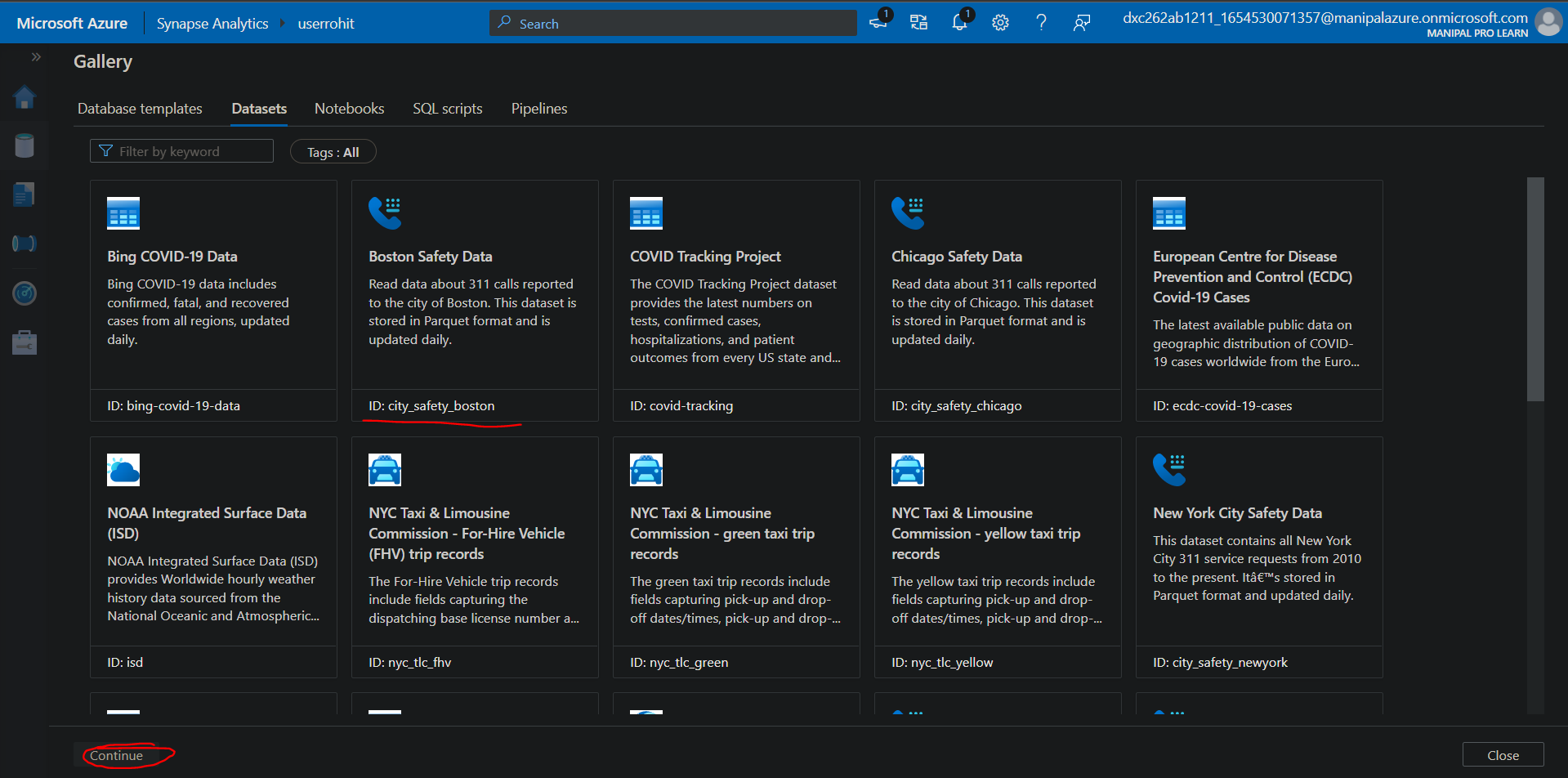
In order to use notebook for analyzing the data we need to open the Synapse sudio > data and after data window is opened click on +



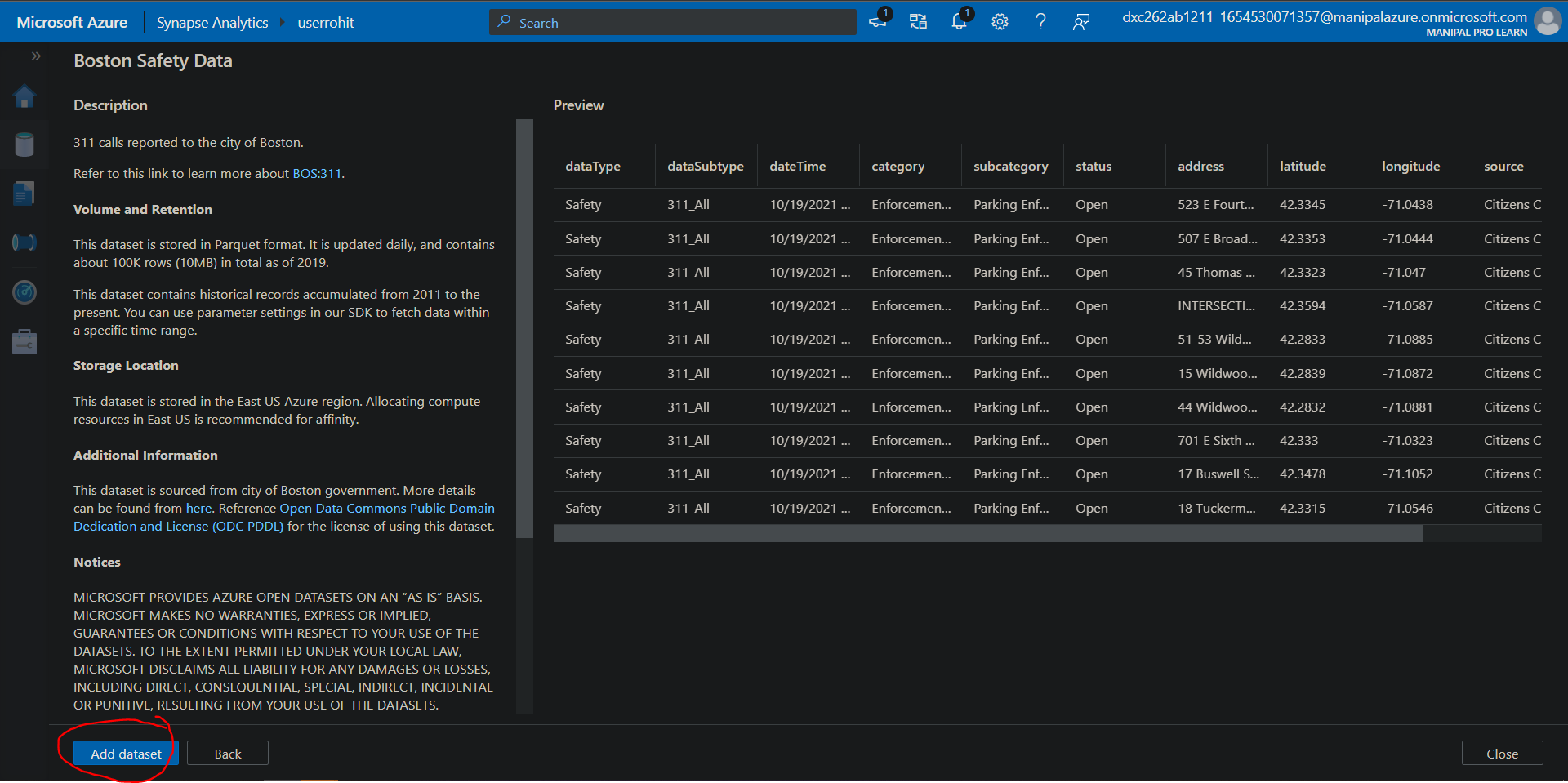
Go to Browse gallery in order to load the boston data



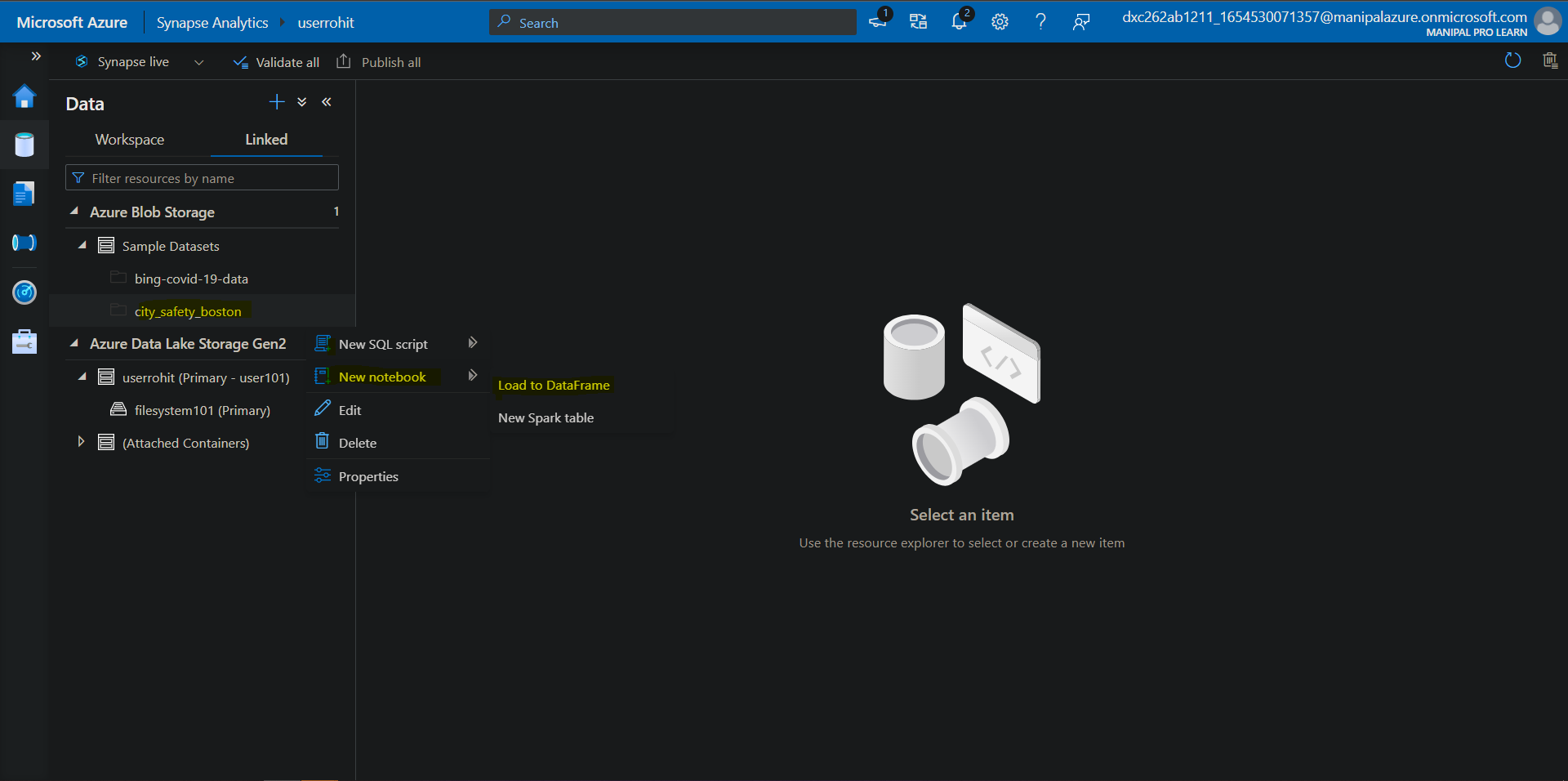
Choose the data to load and click on continue



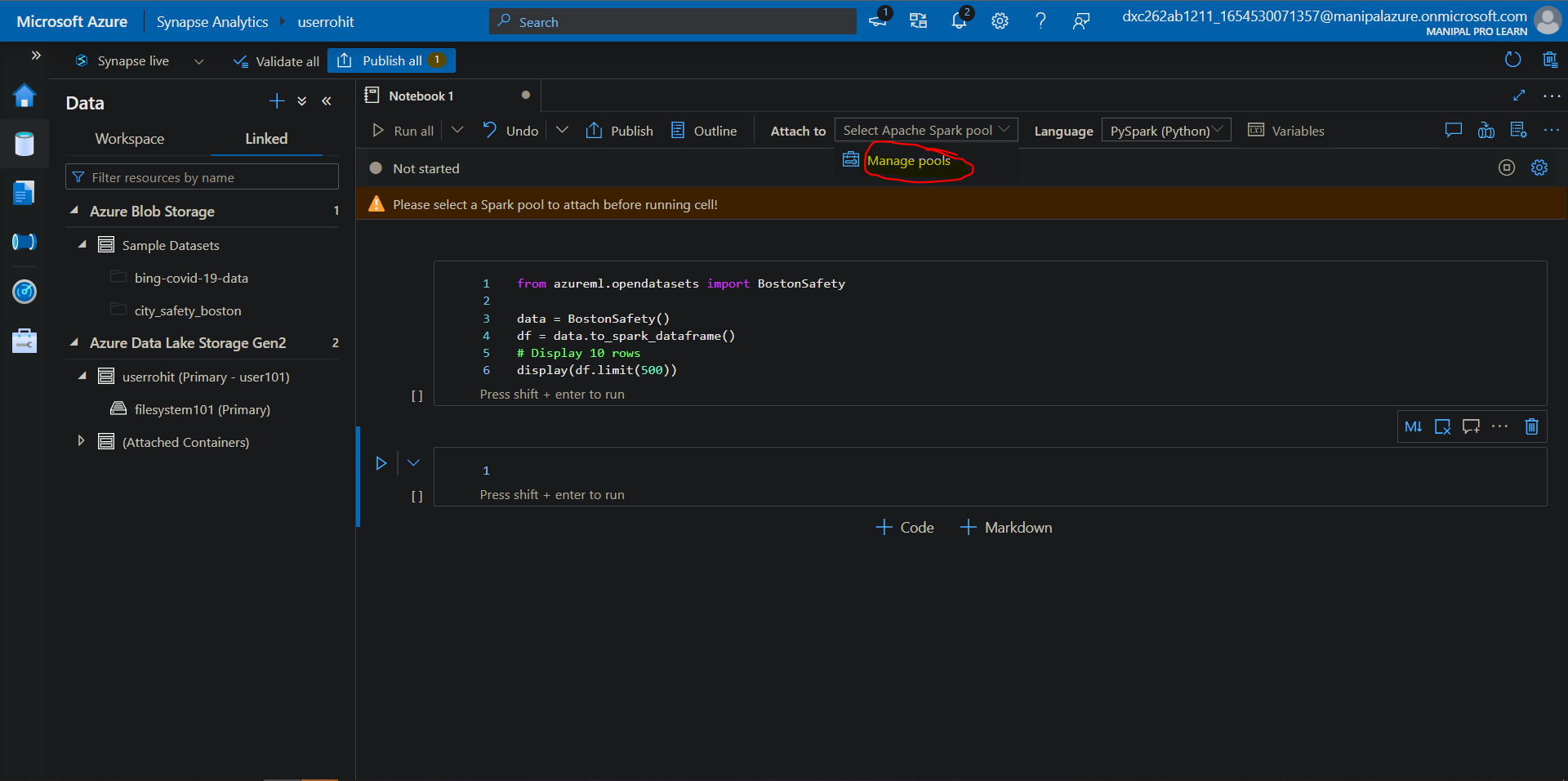
After getting the preview of data click on add dataset



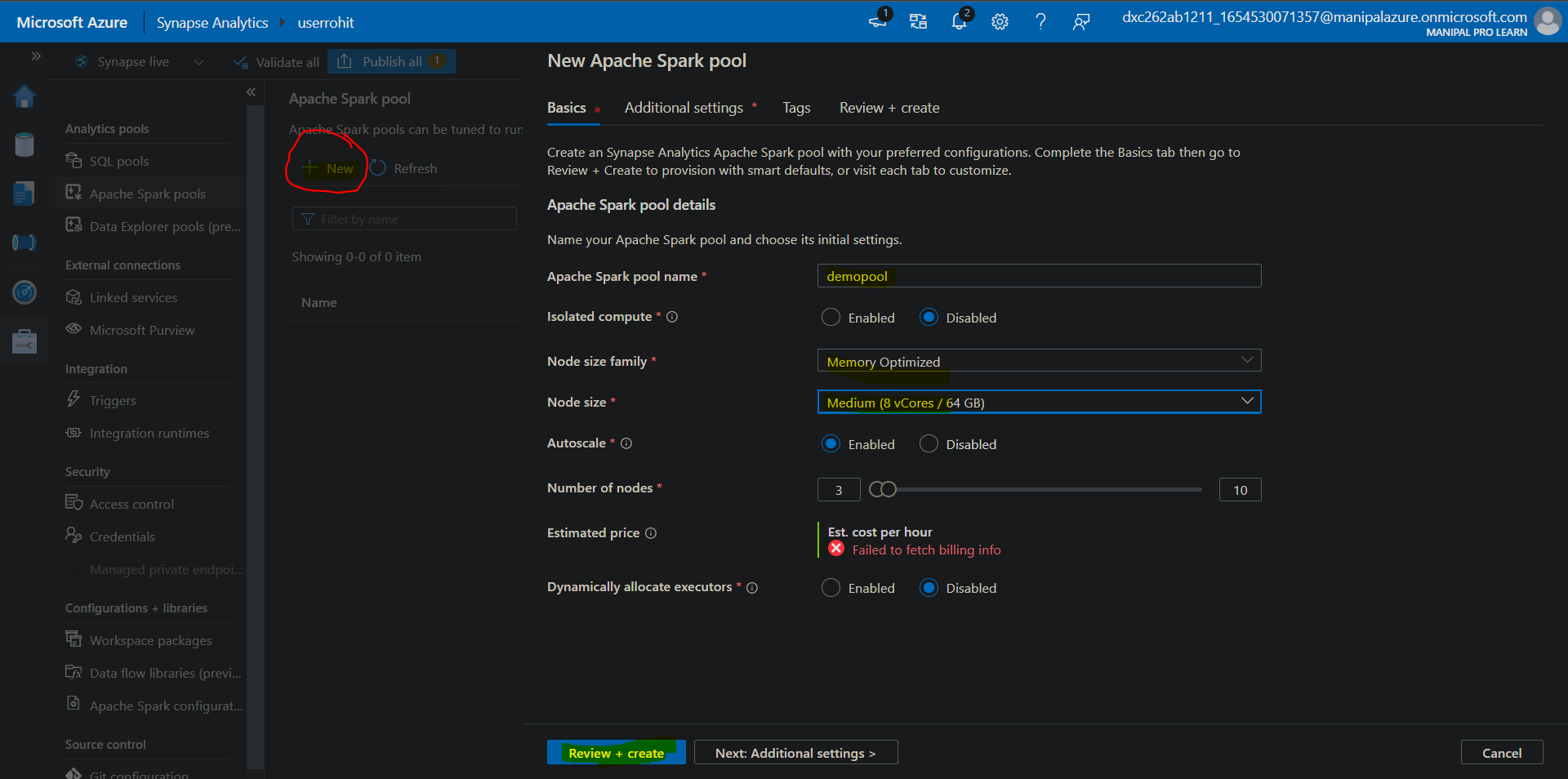
After the dataset is added navigate towards dataset>new Notebook >Load to DataFrame



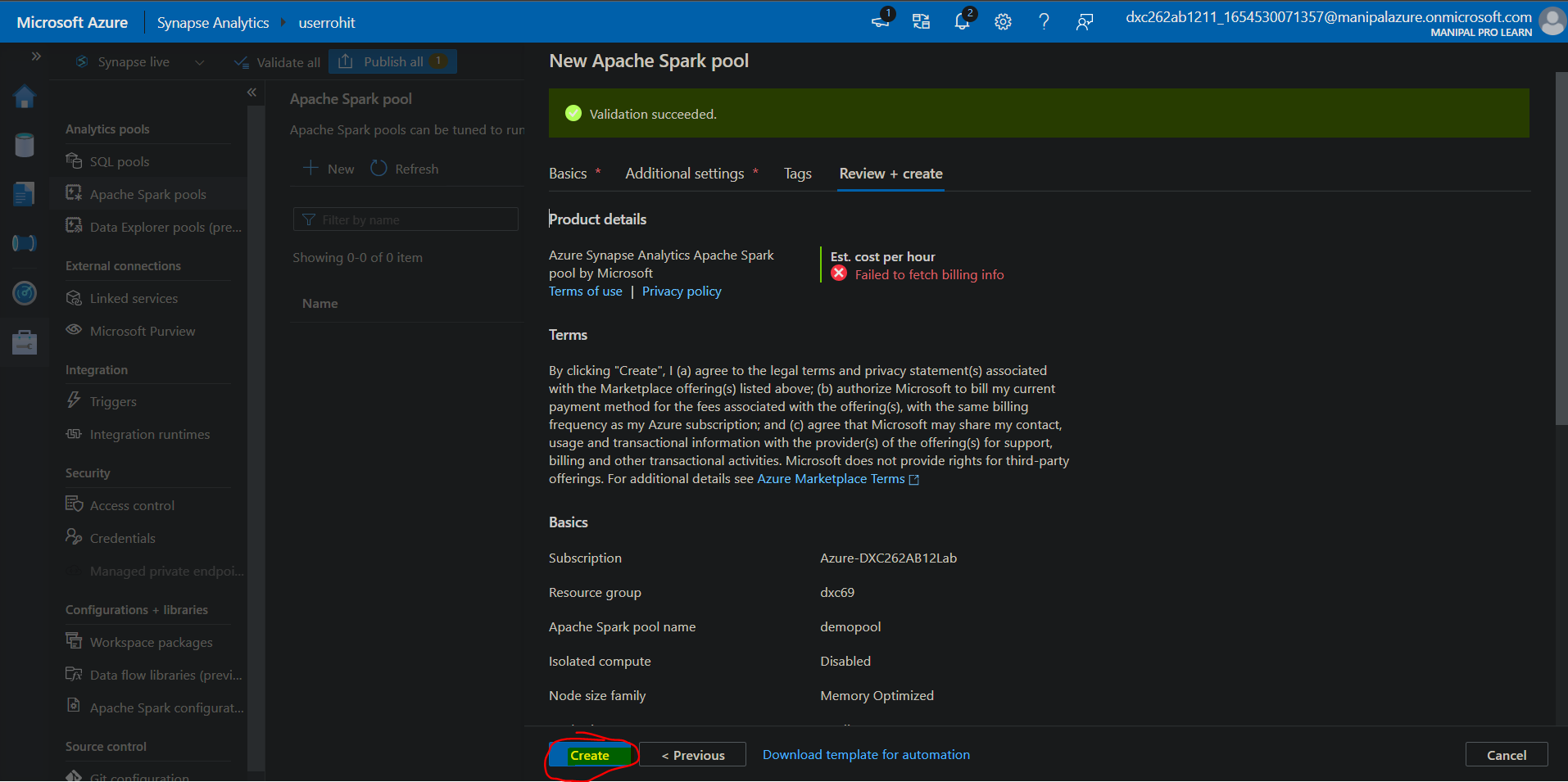
After clicking that a sample notebook will be opened where we can now load the dataset using Spark pool , for creating a spark pool click on manage pool



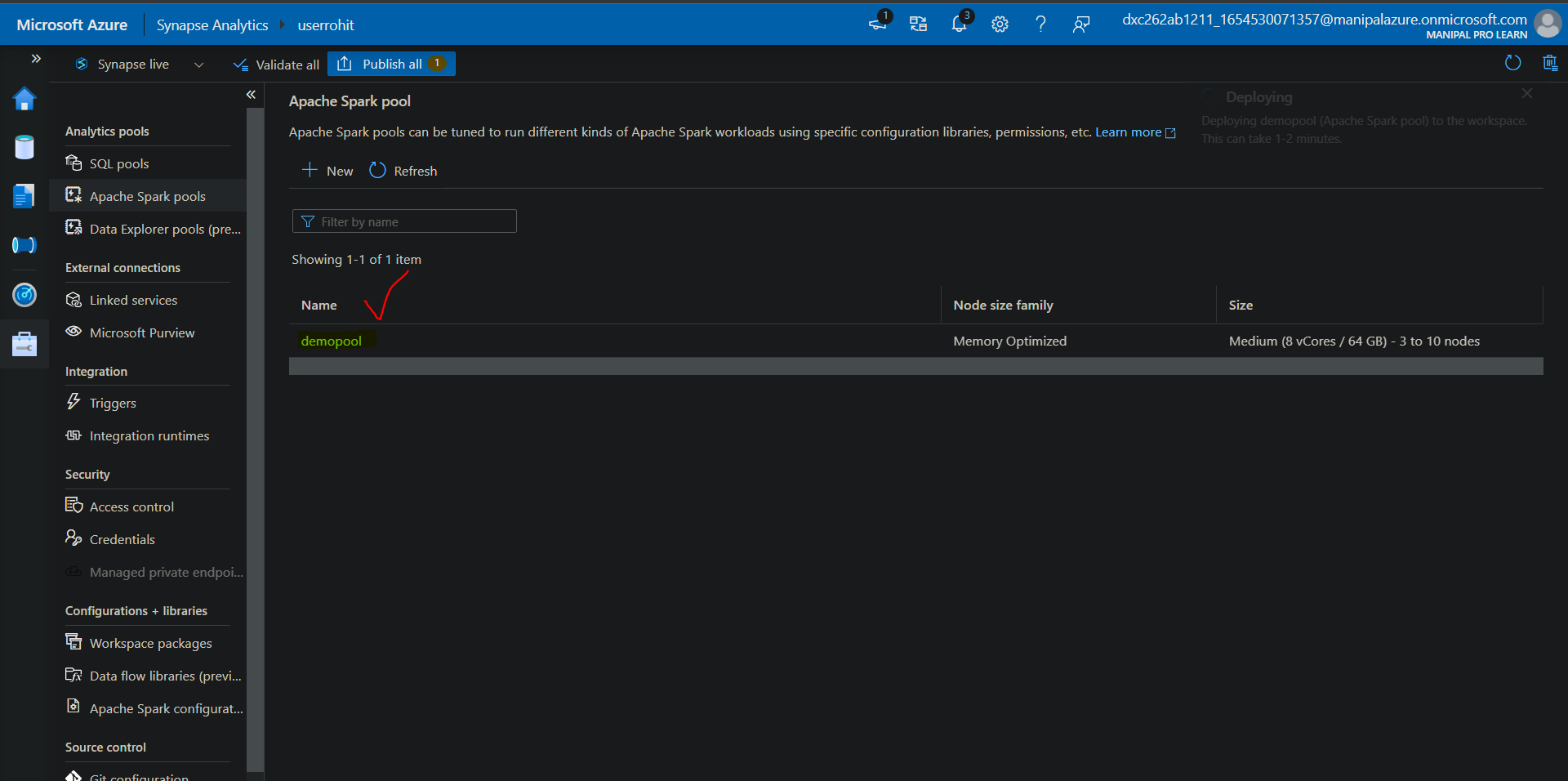
Give the details regarding the pool and click on review and create



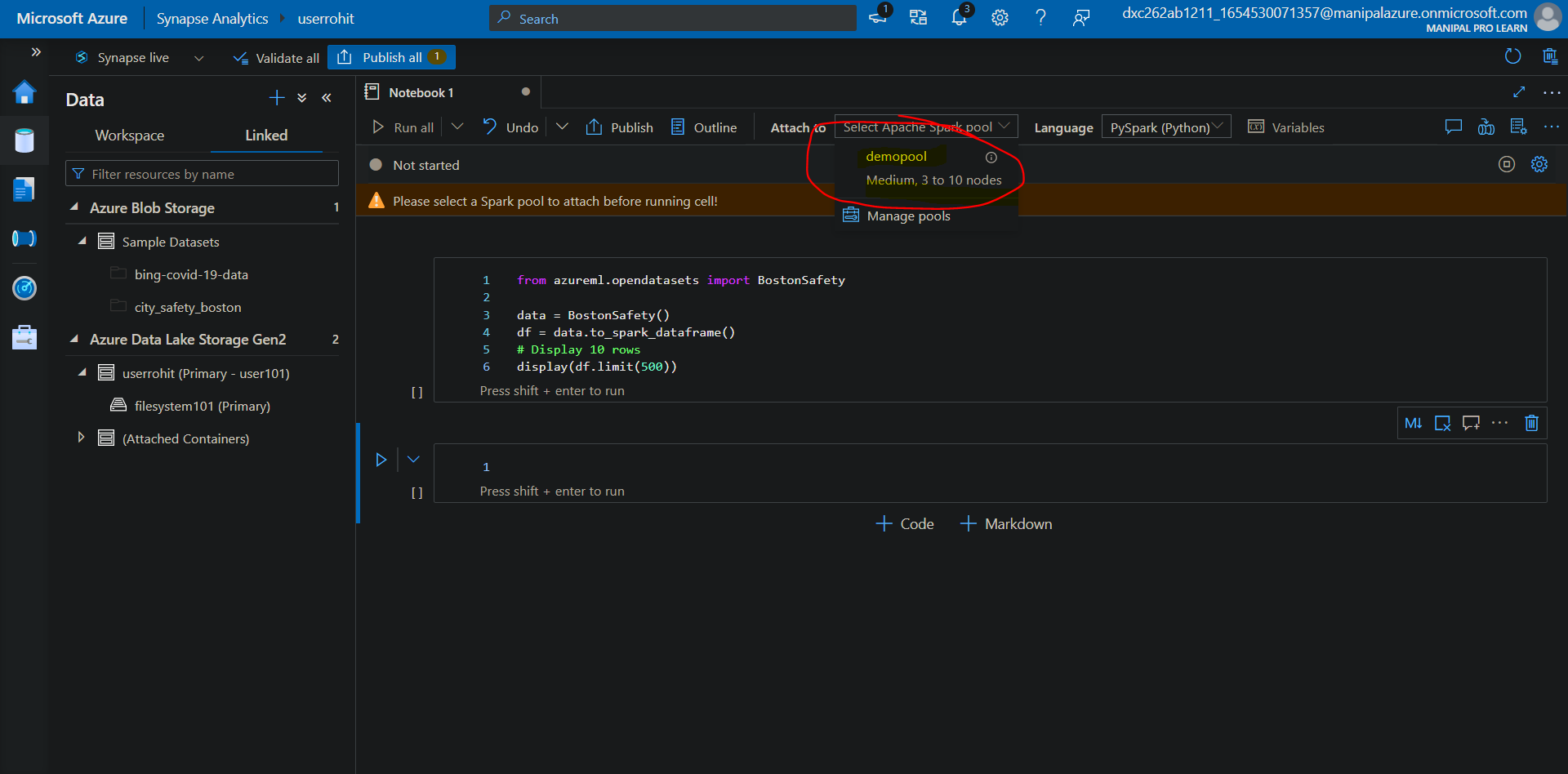
After the validation process is done click on create



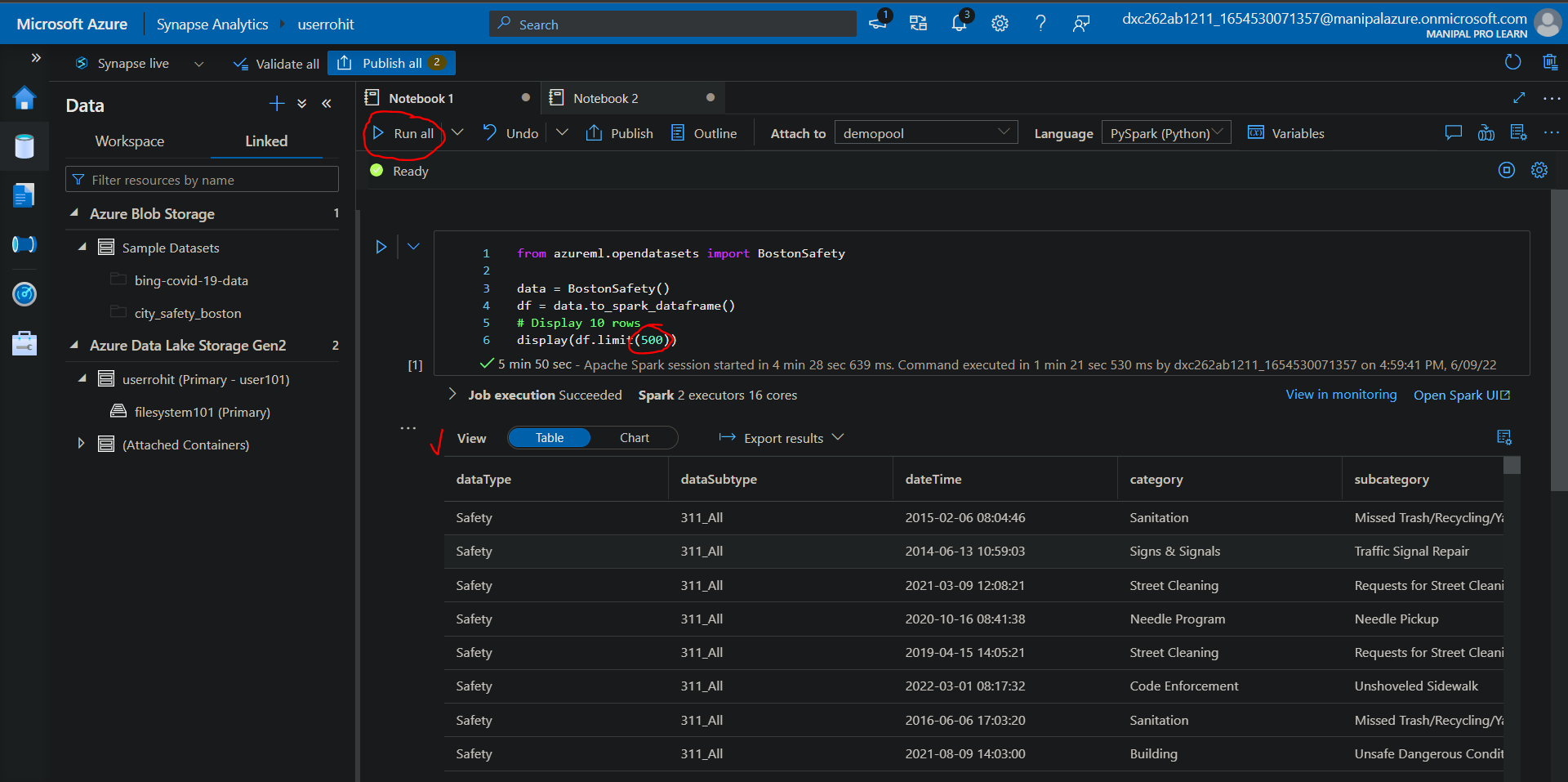
Now you can see the Spark pool is created and we can use it provide the resources and session to our dataloading notebook



Jump back to the notebook and choose the pool we just created



Clicking that now press shift + enter on the cell we want to run by manipulating the number and here we have our results.

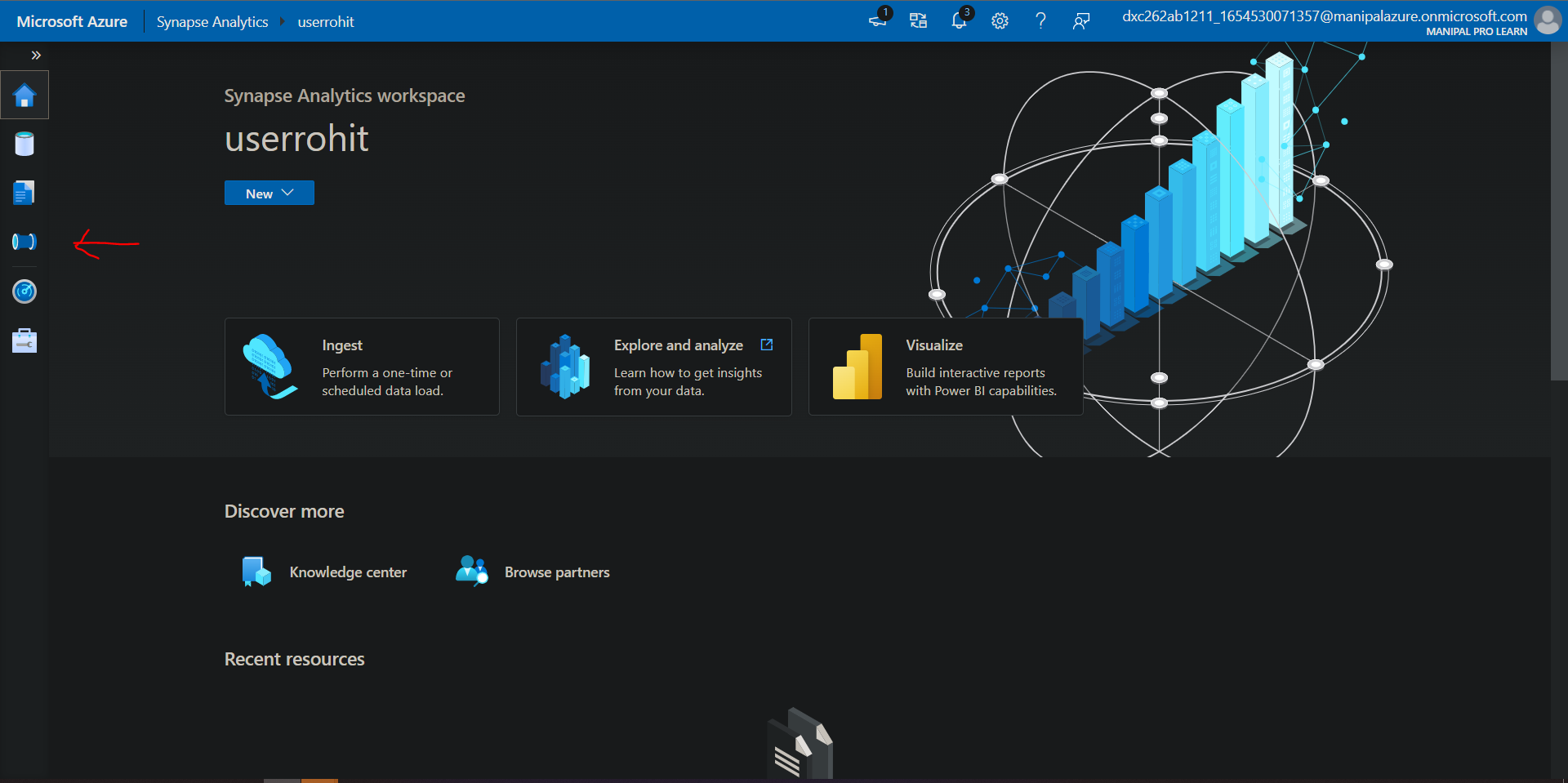


5. Explain the steps with screenshots how to create Spark pool in AzureSynapse analytics?

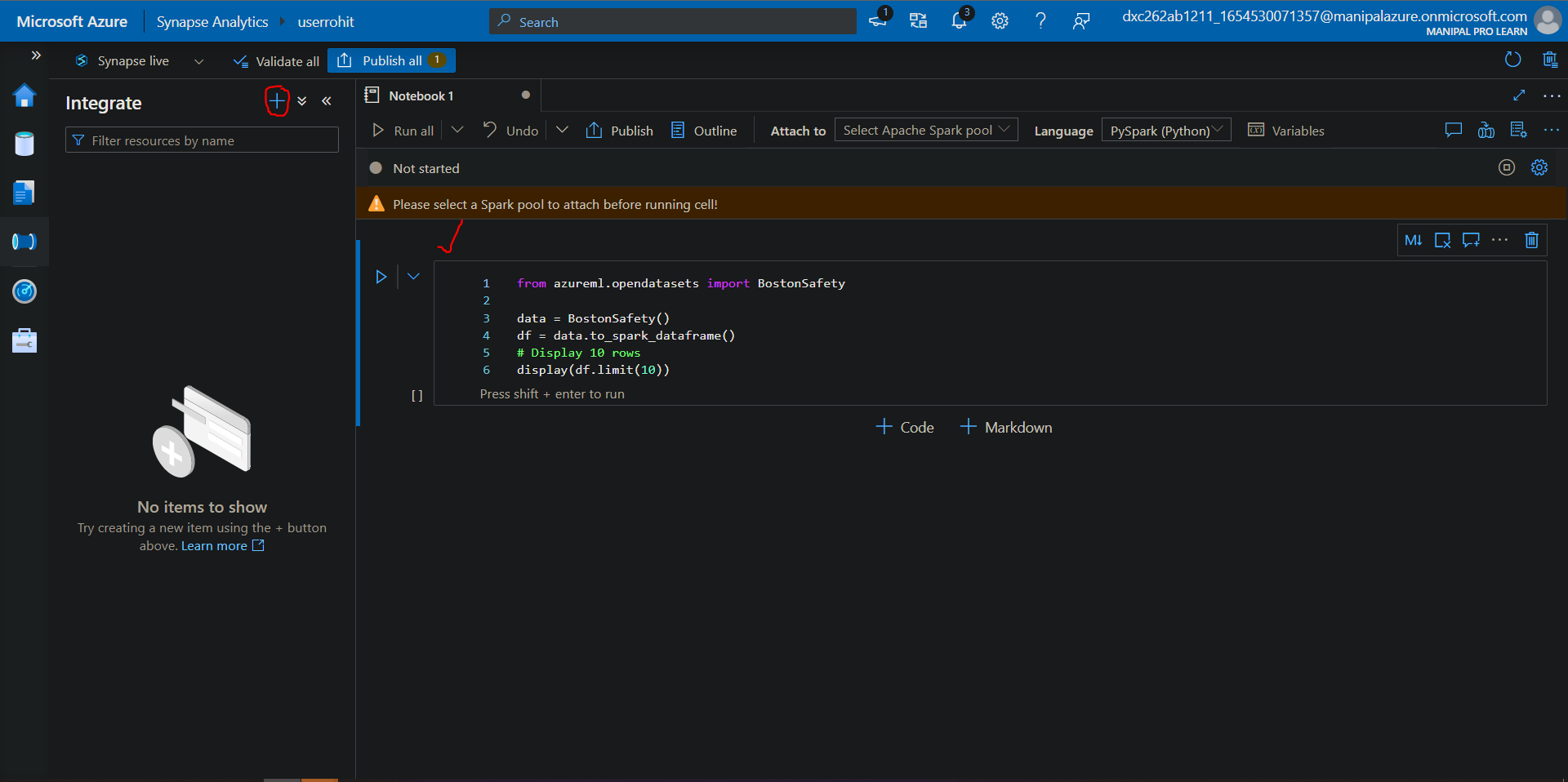
Created in the previous question

6. Explain the steps with screenshots how to create pipeline in AzureSynapse analytics? ?

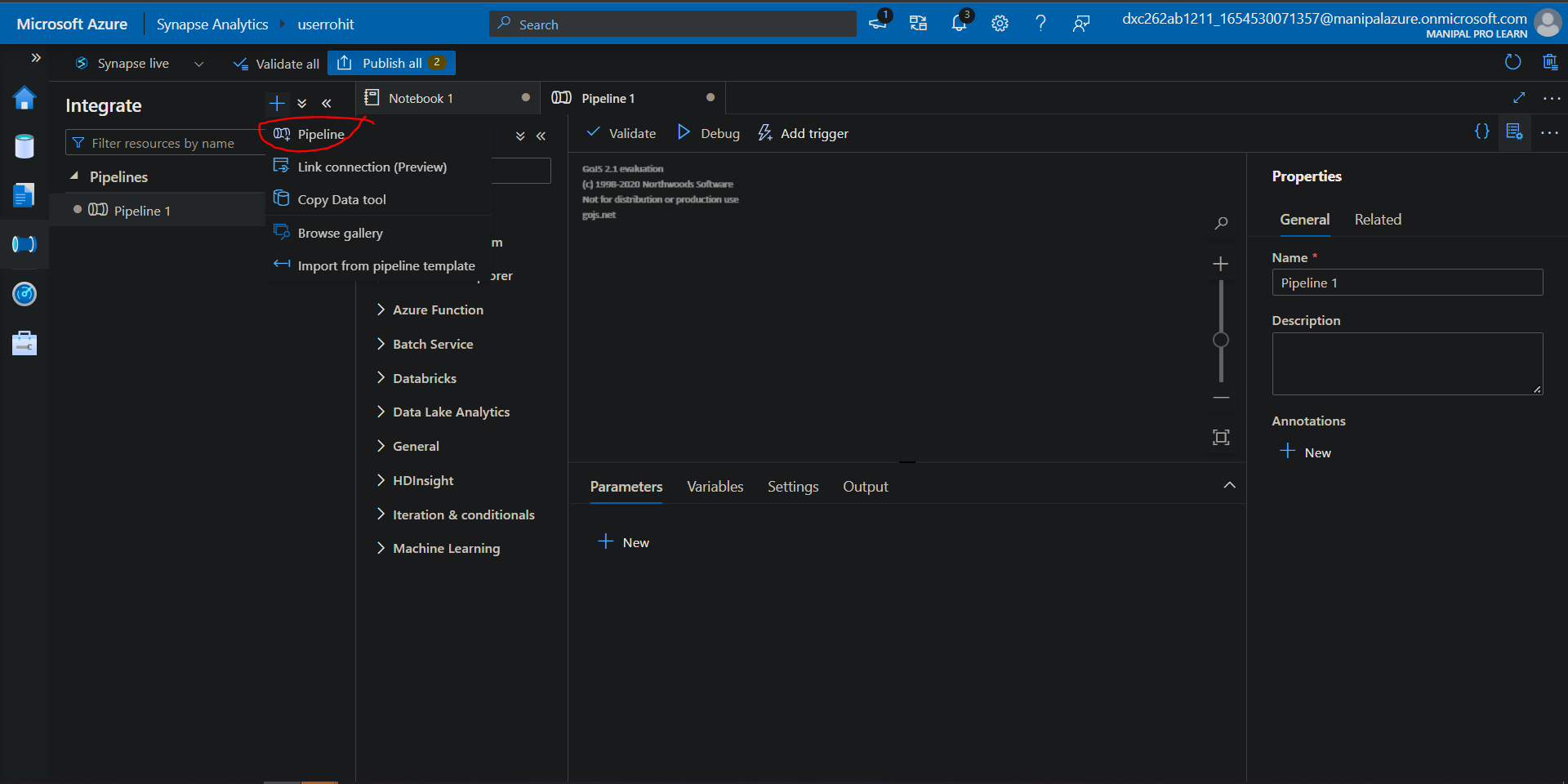
Open Synapse studio first and navigate onto the integrate bar



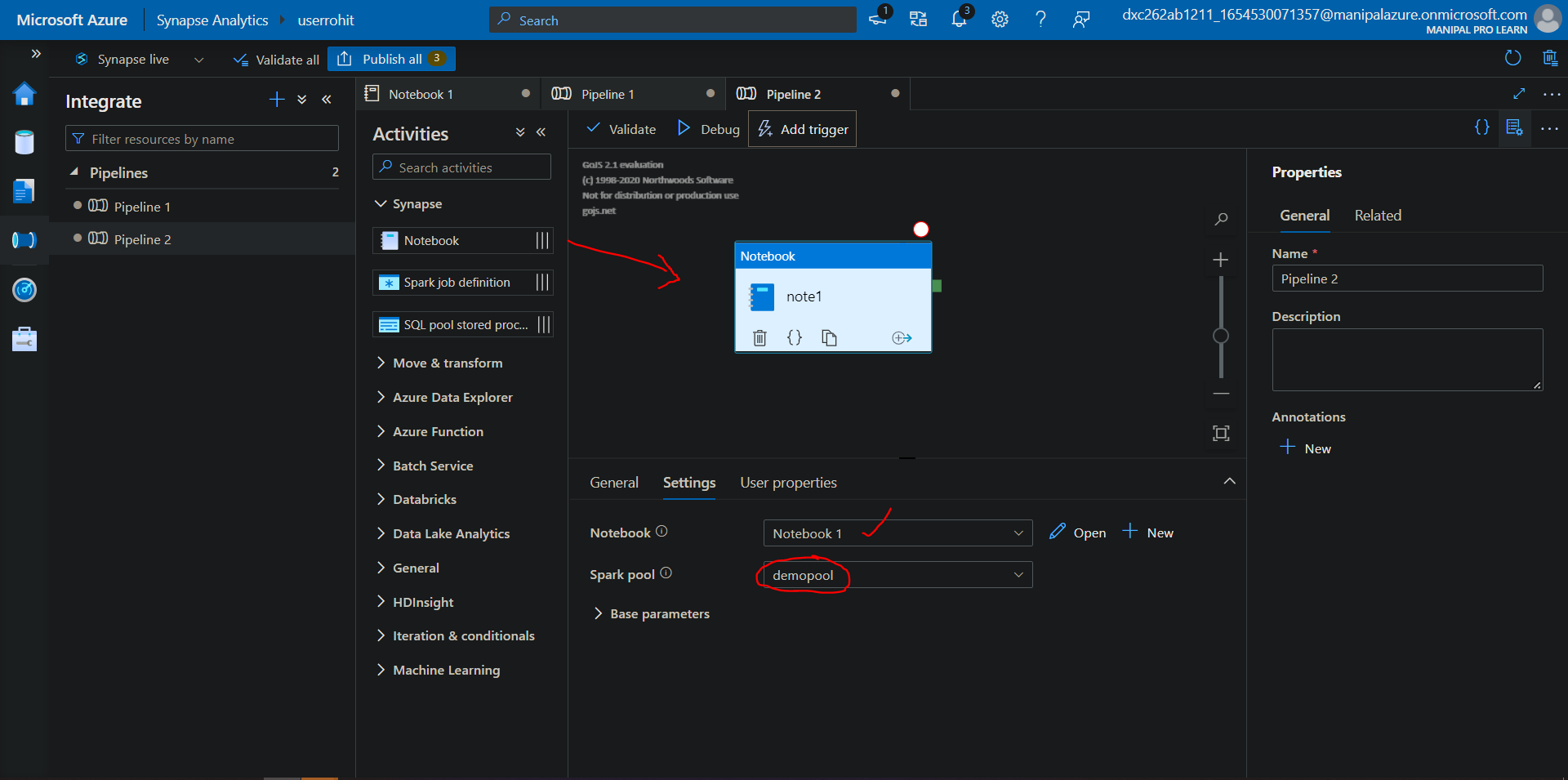
Click on + to create a new pipeline



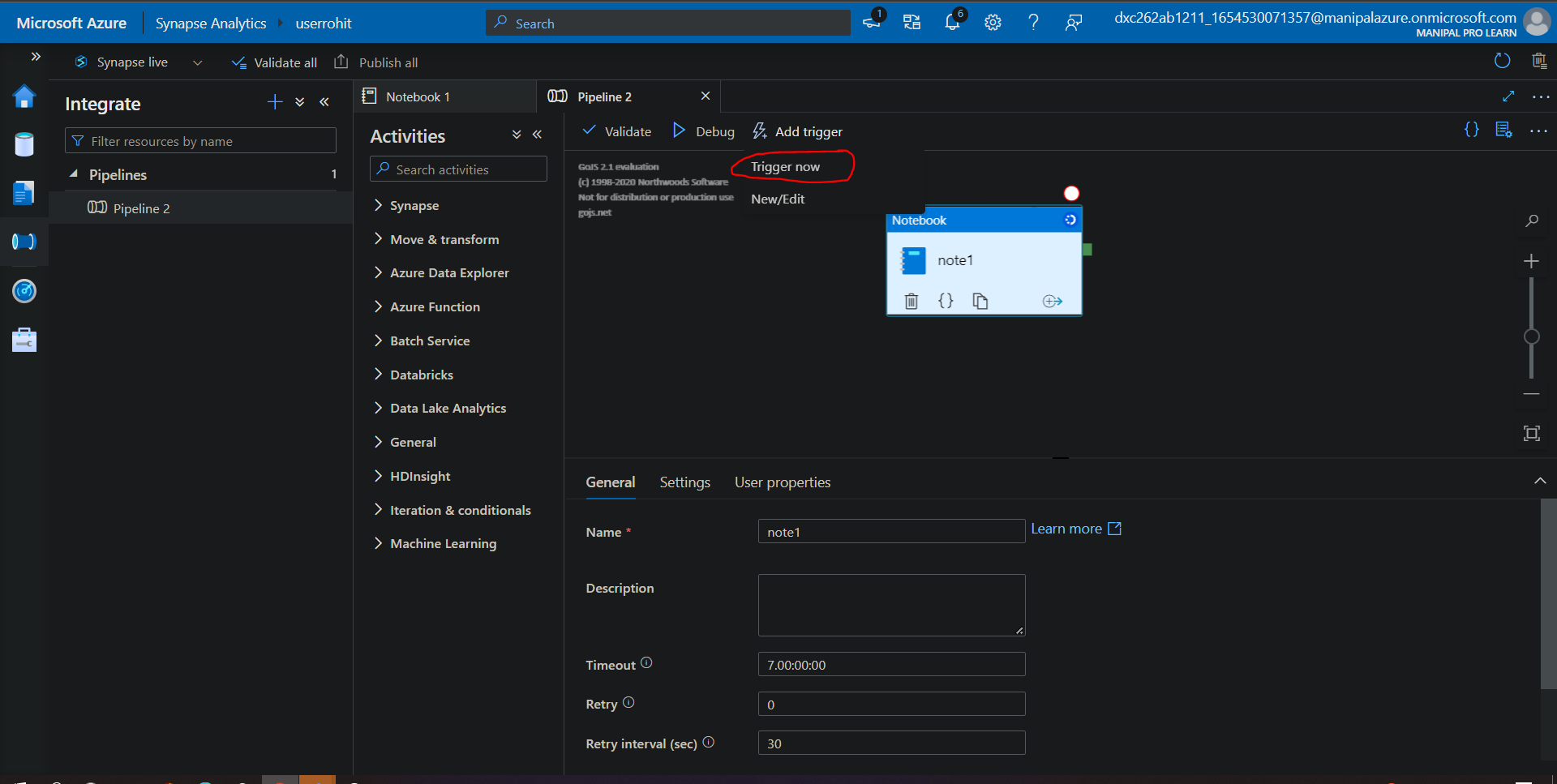
Click on pipeline to add a new pipeline



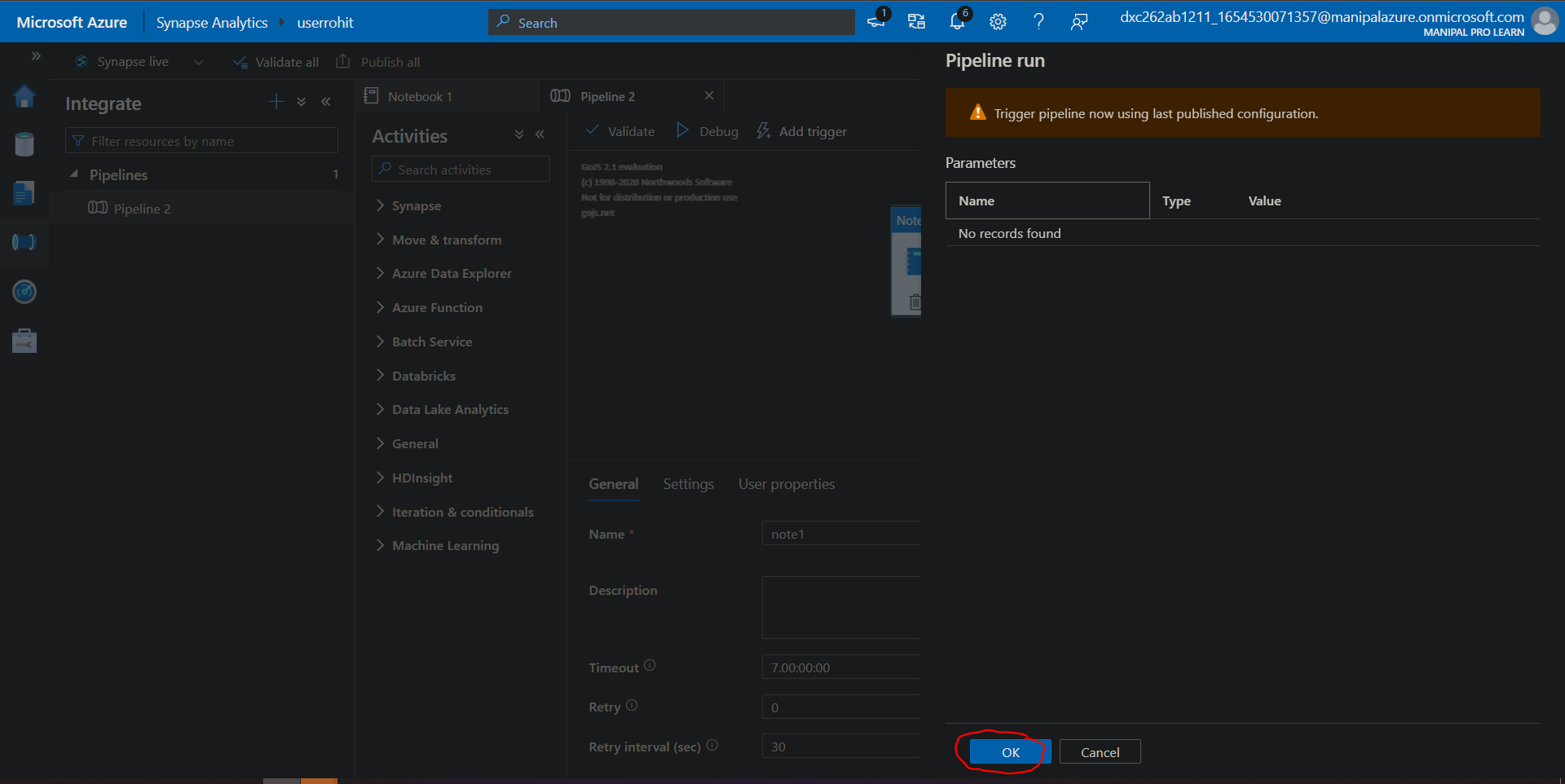
Add the job in our case we will run the notebook by triggering it from pipeline



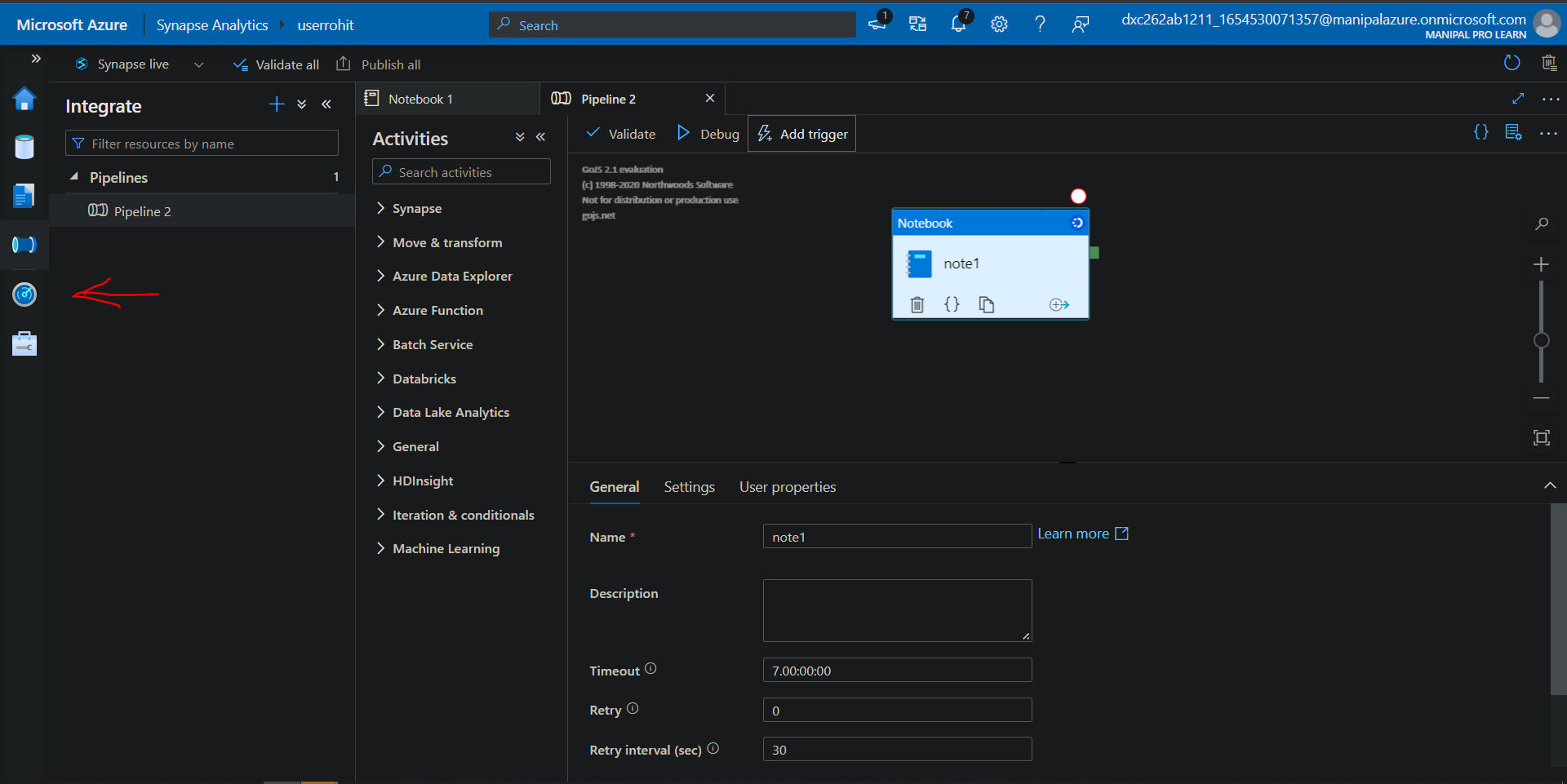
Choose the pool and name and click on trigger now to trigger the pipeline



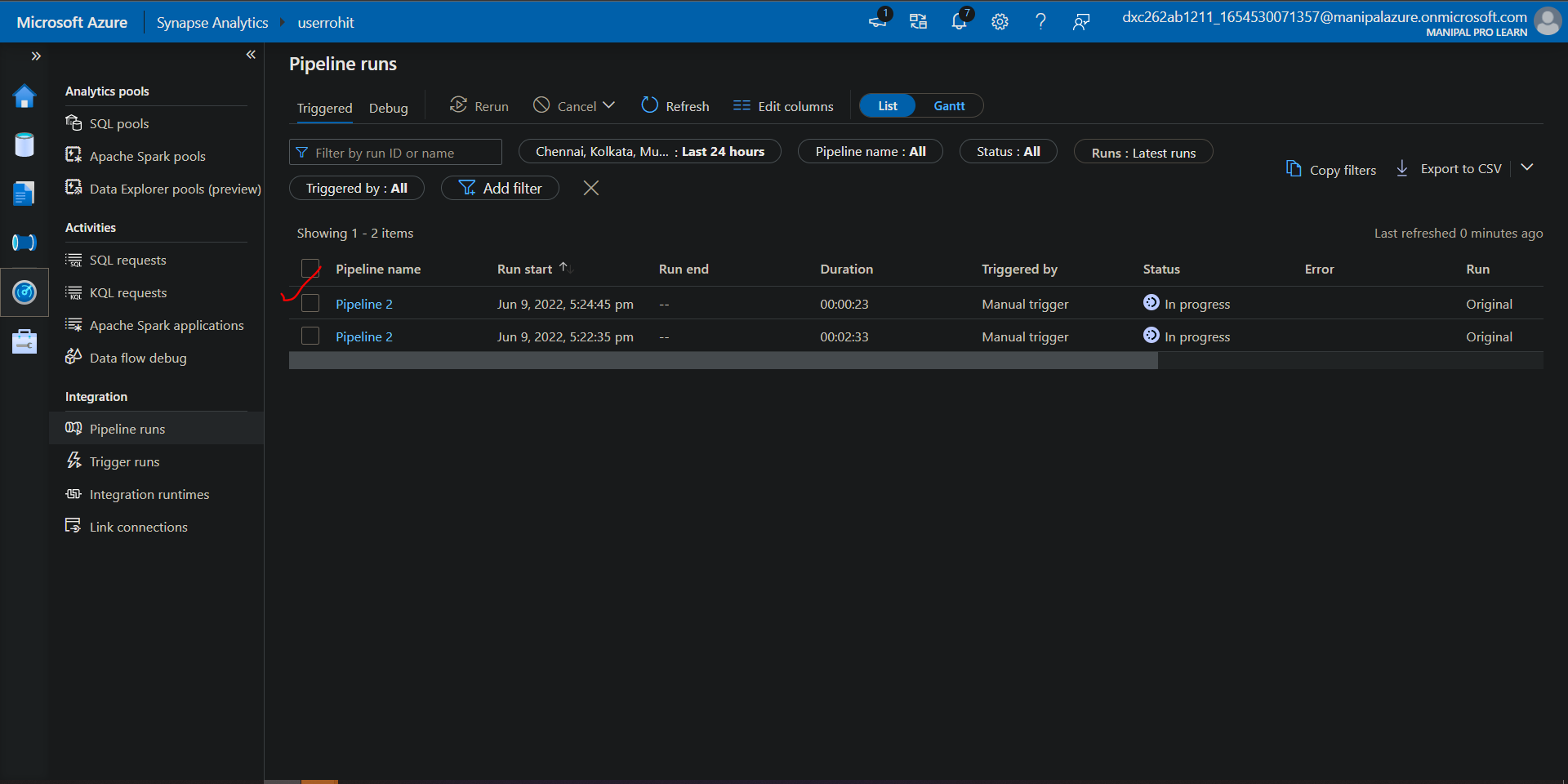
make sure to publish the pipeline before triggering it. Otherwise it won’t be able to find the pipeline . Click on ok and it will take the most recent publication done for the pipeline



To monitor the status of the trigger go to monitor bar

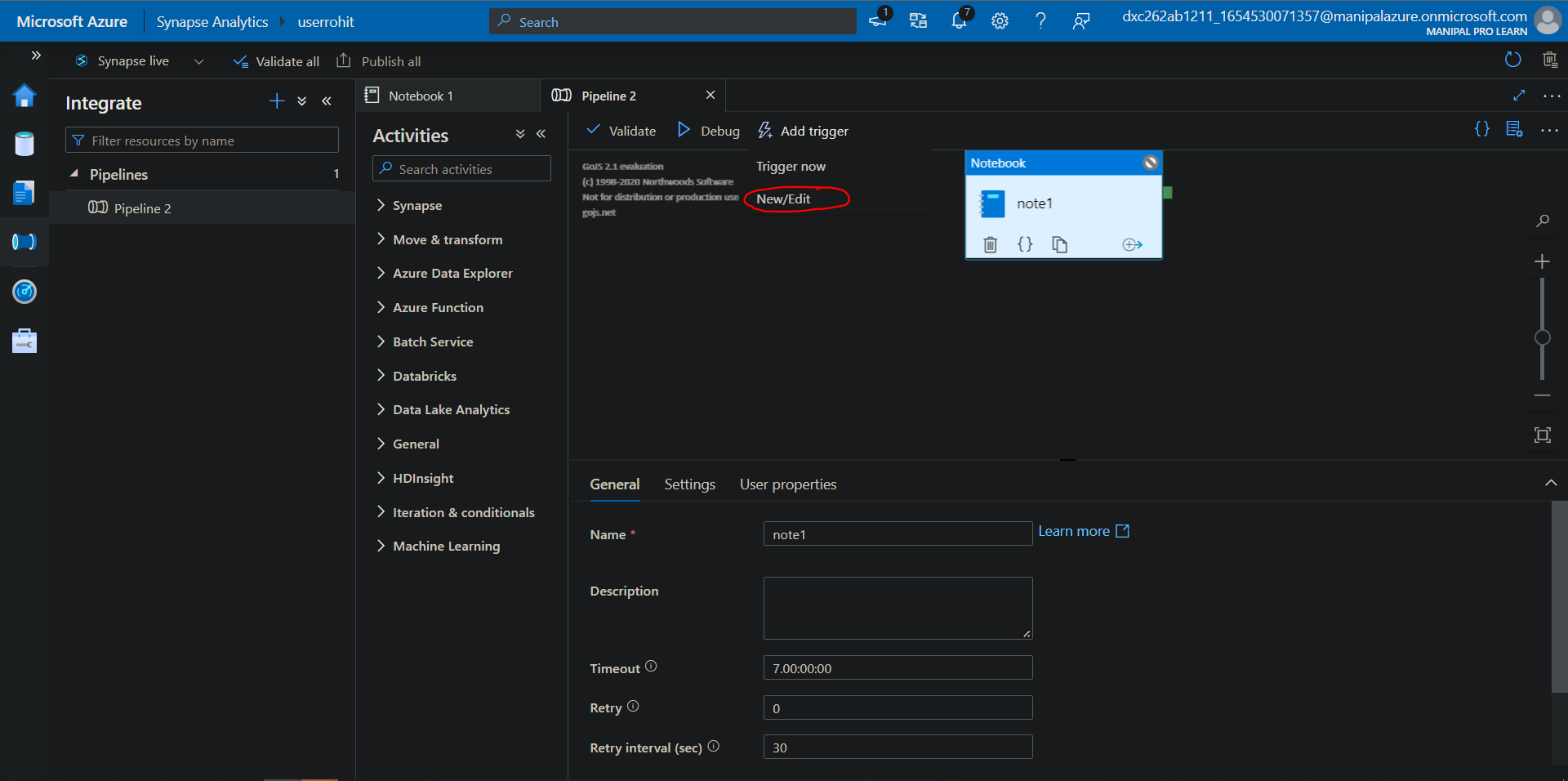


And here we have successfully created a pipeline and triggered it manually

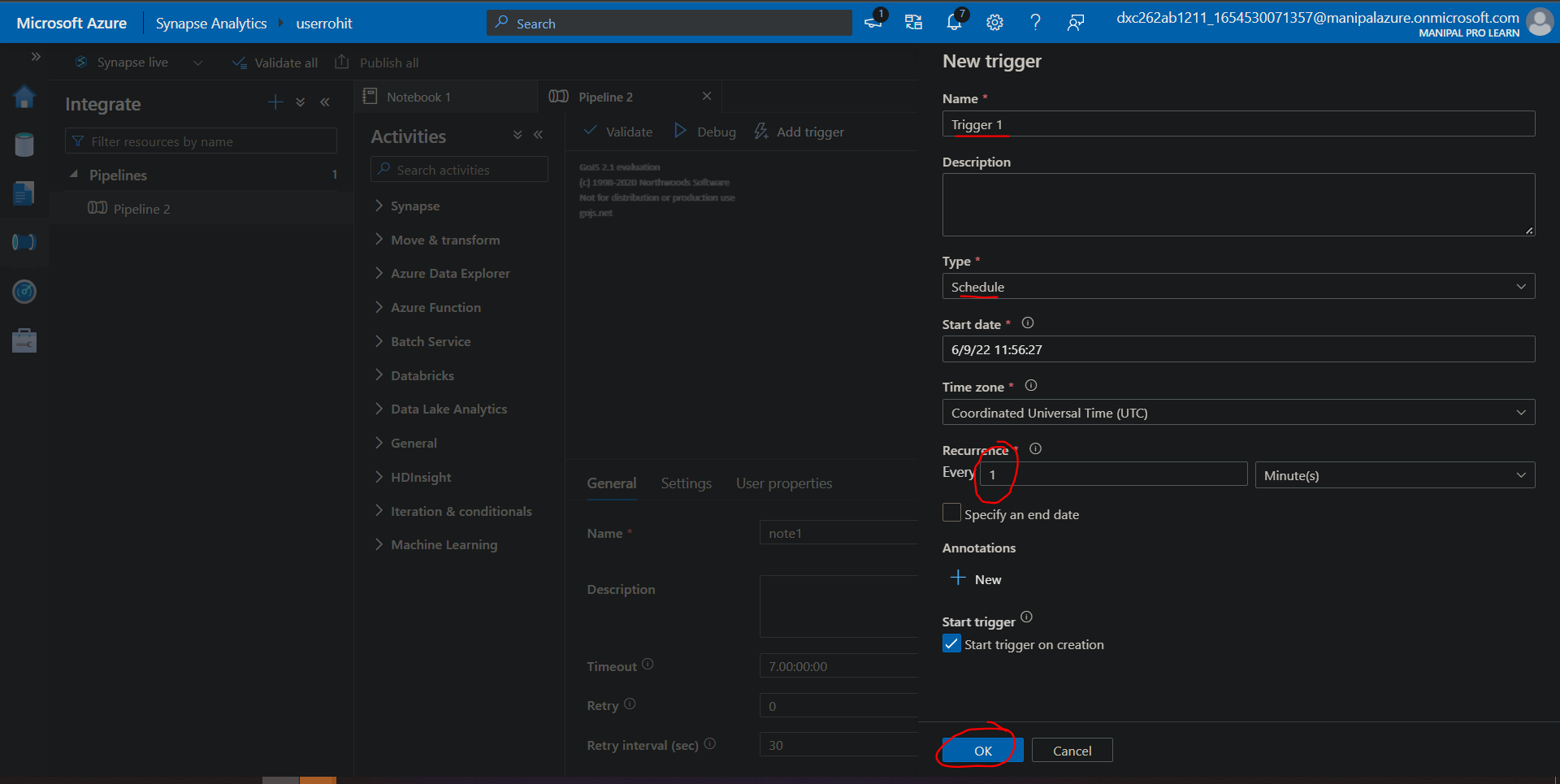


7. Explain the steps with screenshots how to automate the pipelines in AzureSynapse analytics??

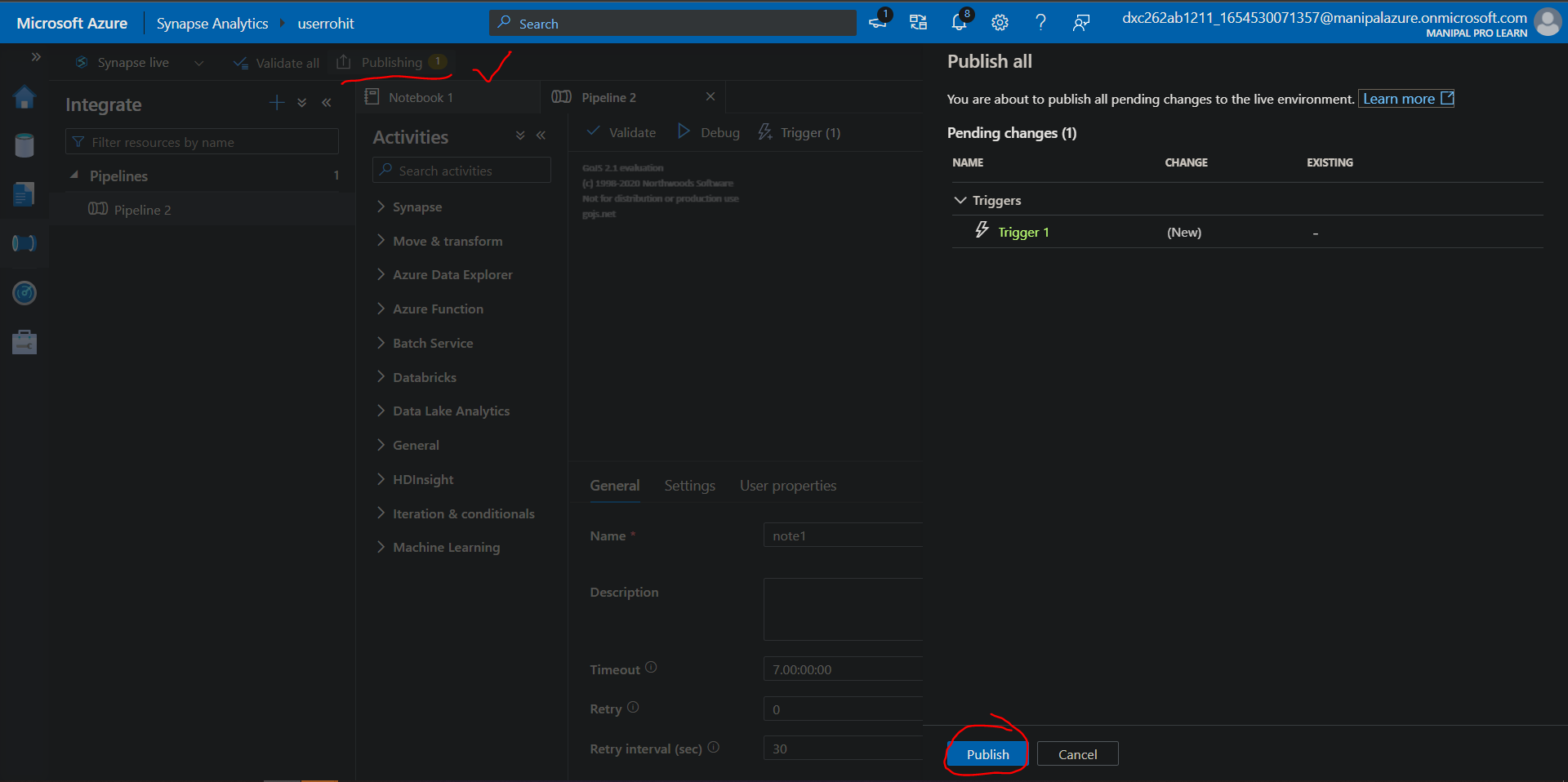
In order to automate a pipeline we will pick the pipeline we created in the previous question and will try to automate it. Open the same pipeline setting and instead of trigger now click on New/edit



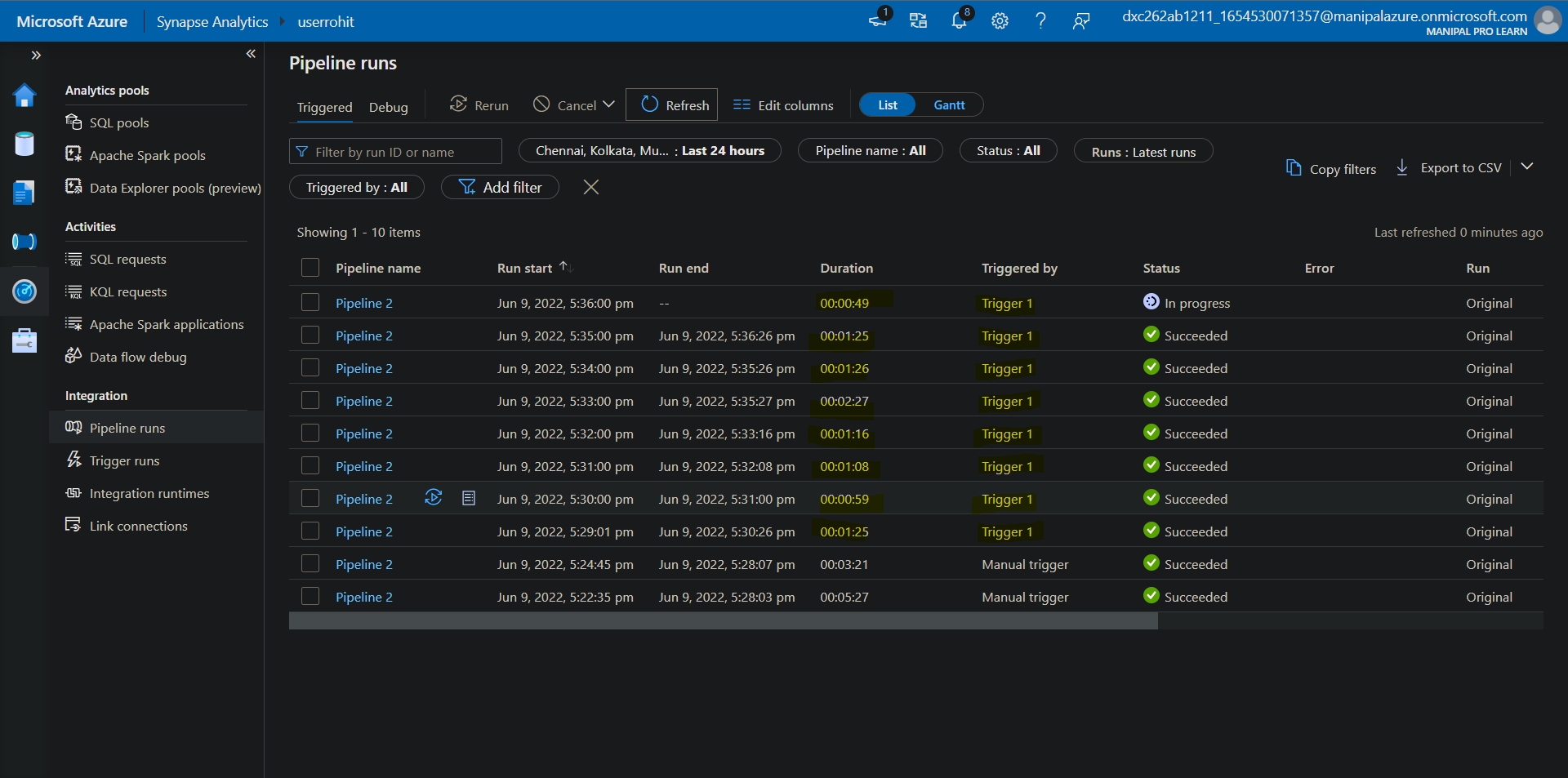
Give the details for the new trigger and we’ll try to schedule it for every 1 minute. After filling details click on ok



Publish the trigger to make it work



Here we have automated a pipeline to run every minute .Navigate to the Monitor to view the results.



8. Explain the steps with screenshots how to Databricks ?

9. Explain the steps with screenshots how to create notebooks in Databricks ?

10. Explain the steps with screenshots how to insert data into databricks notebook & display the result?

11. Explain the steps with screenshots how to create cluster in databricks ?



