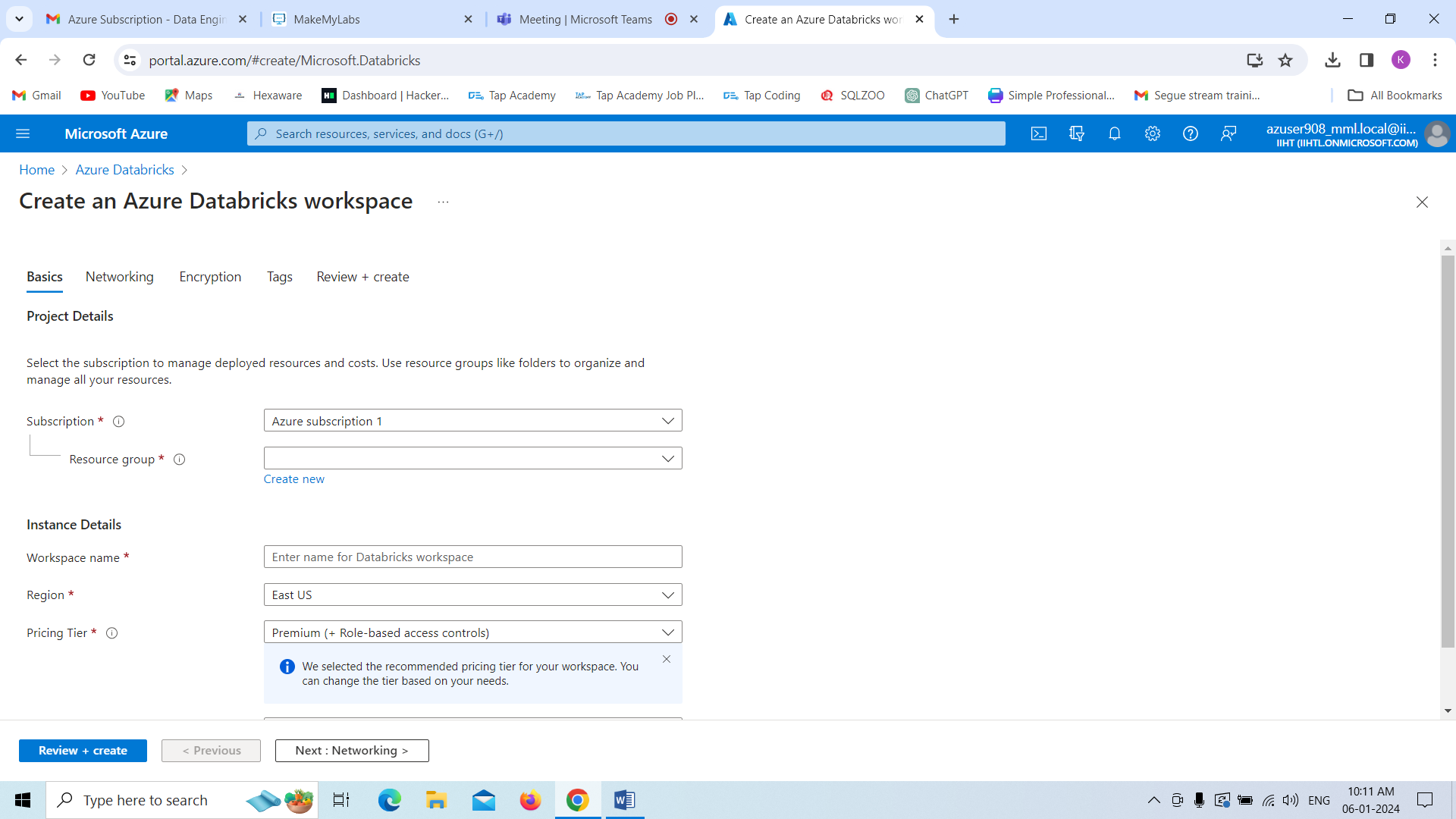
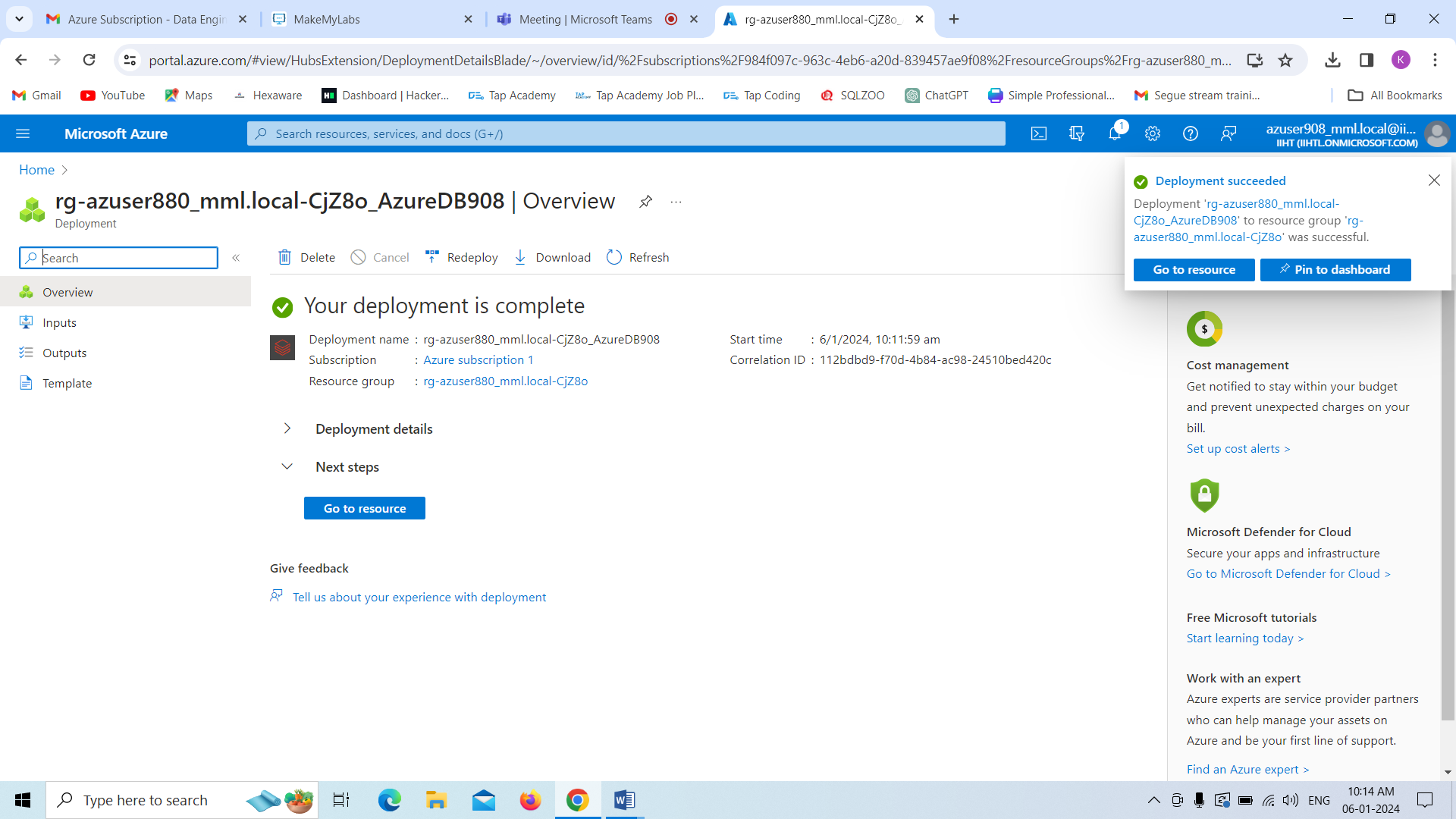
**Azure Databricks Coding Assessment(06-01-2024)**

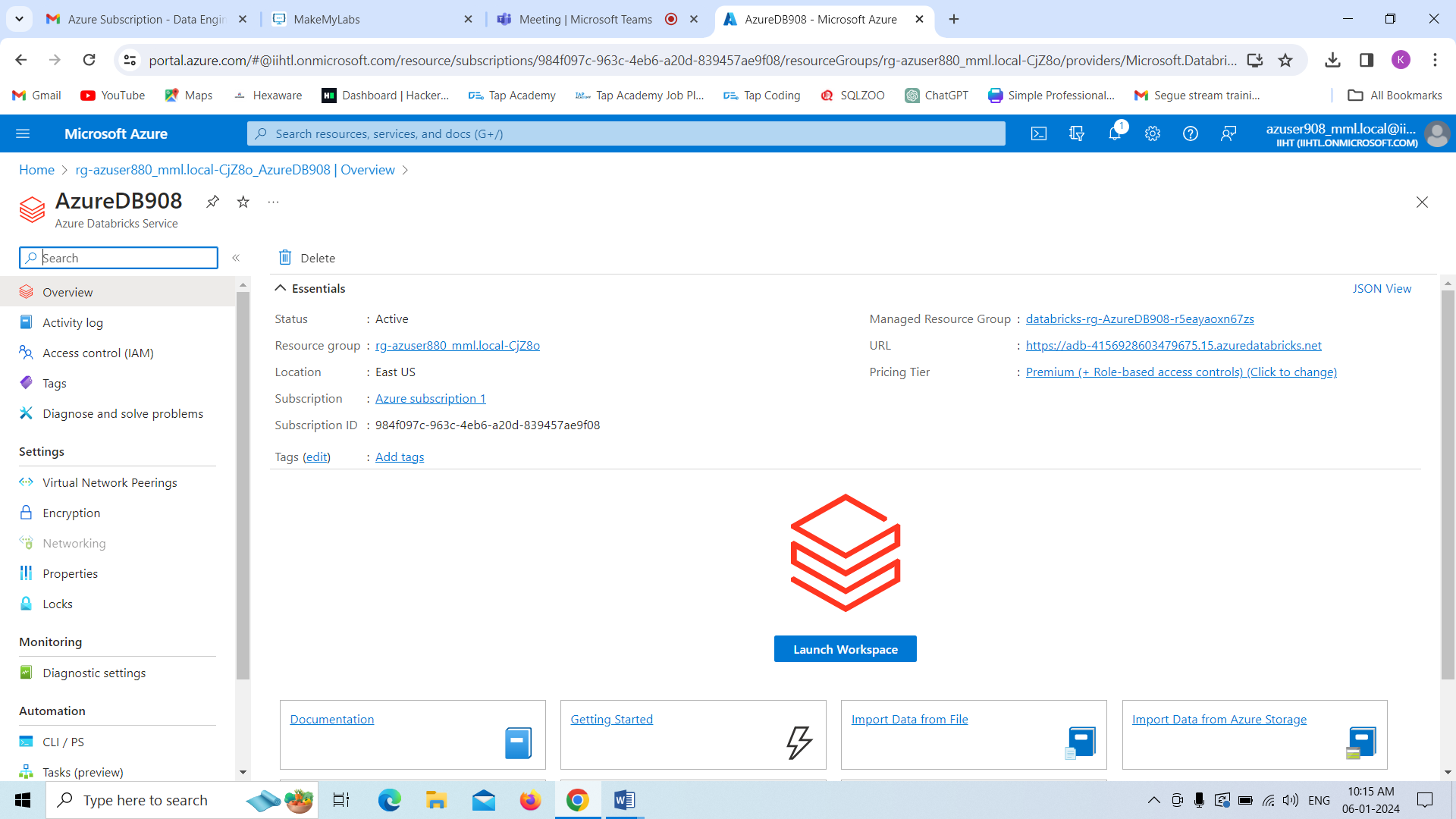
1.Create a cluster &Attach the notebook to the cluster and run all commands in the notebook& creates a DataFrame from a Databricks dataset& Create a Visualizations in Databricks notebooks  
&Rename, duplicate, or remove a visualization or data profile.  
2.Explain the copy activity in Azure data factory.  
First of all,we need Azure Databricks WorkSpace, we can create as followsclick on **review+create.**



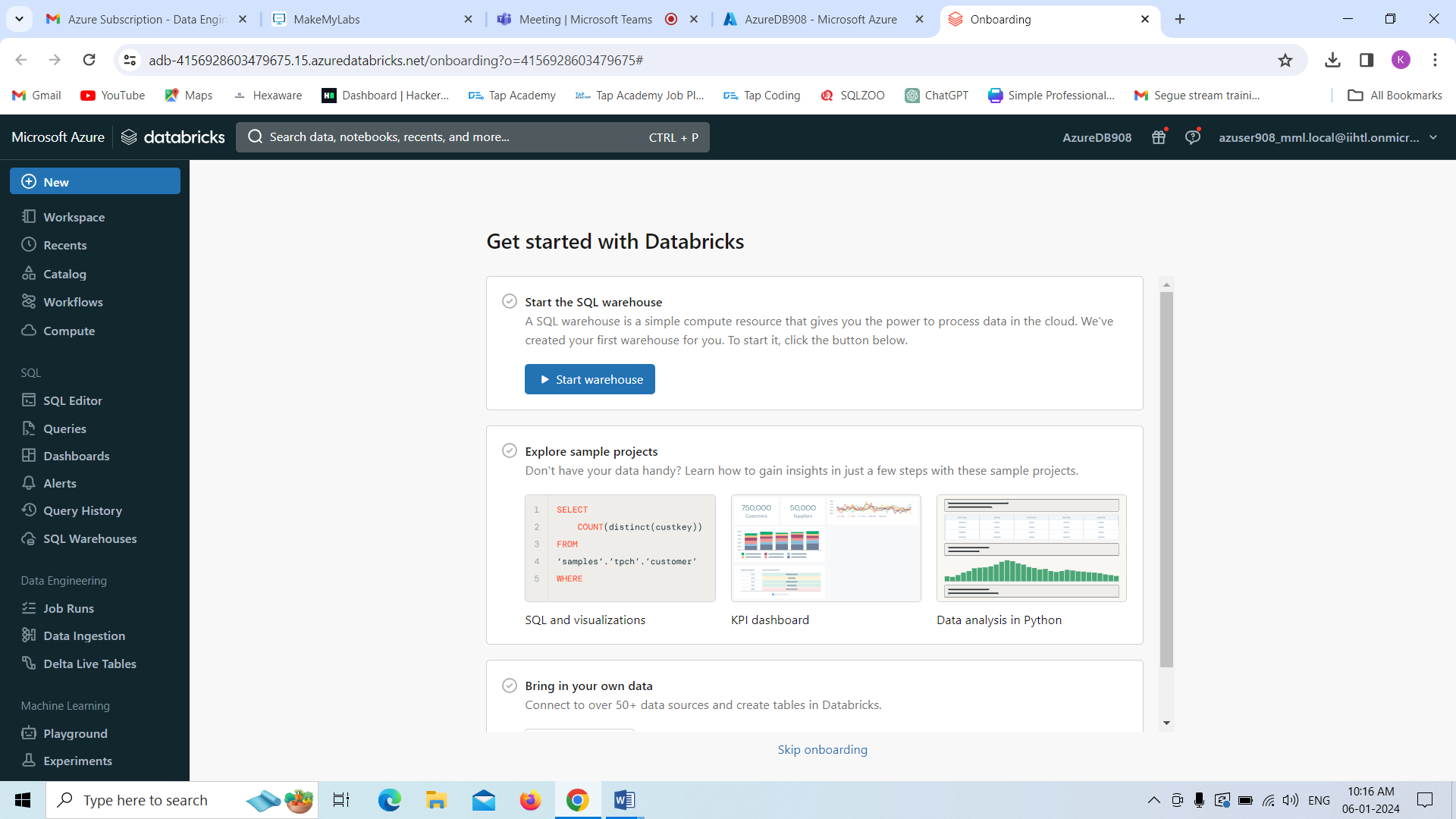
Our Deployment for workspace is completed.click on **Go to Resource.**



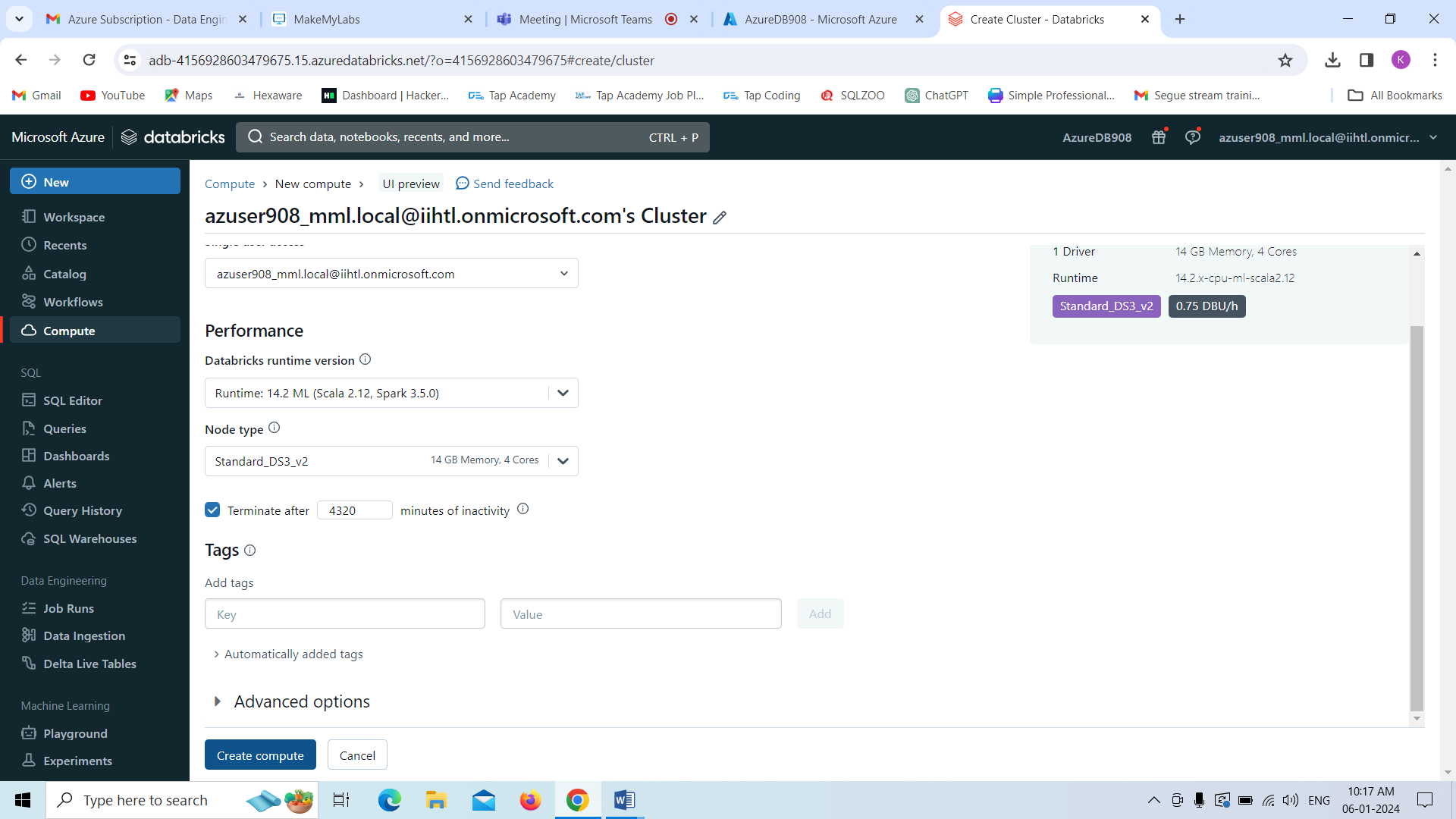
Now you will see the page as below,we have to launch the workspace in order to go forward.click on **Launch Workspace.**



Now our workspace is launched.



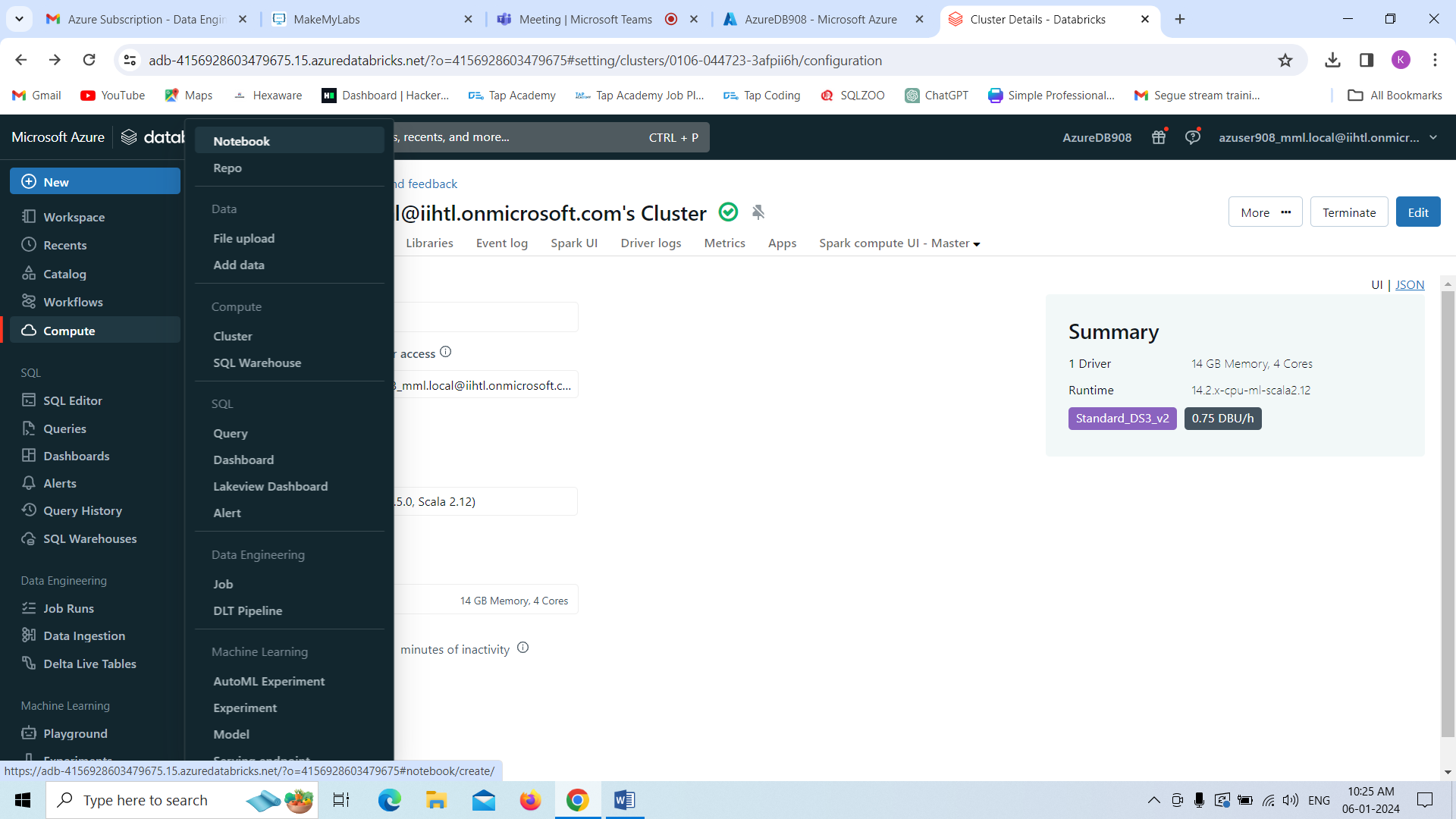
Now we are creating cluster ,for that select **compute** click on **create compute**



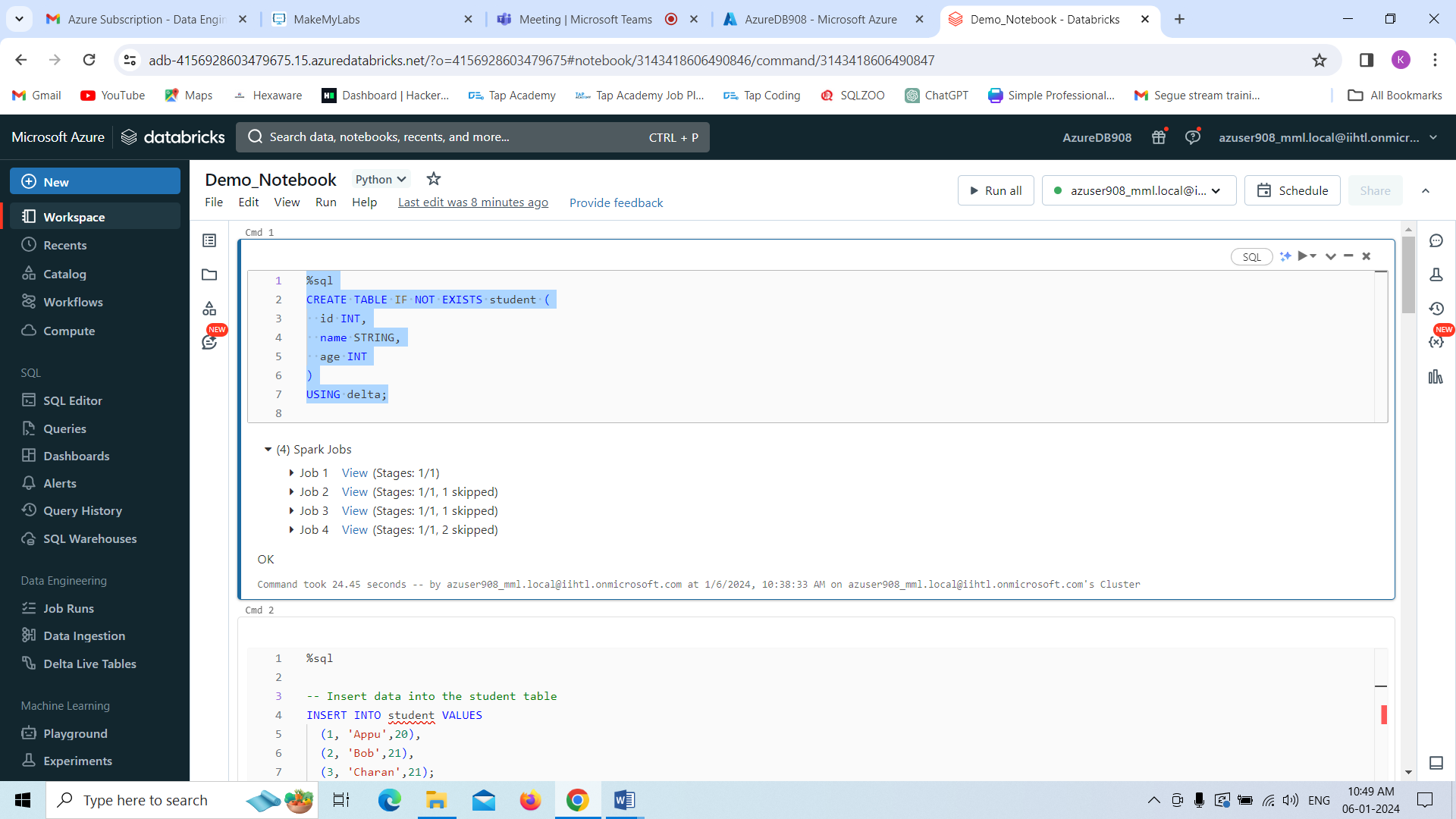
Cluster is successfully created



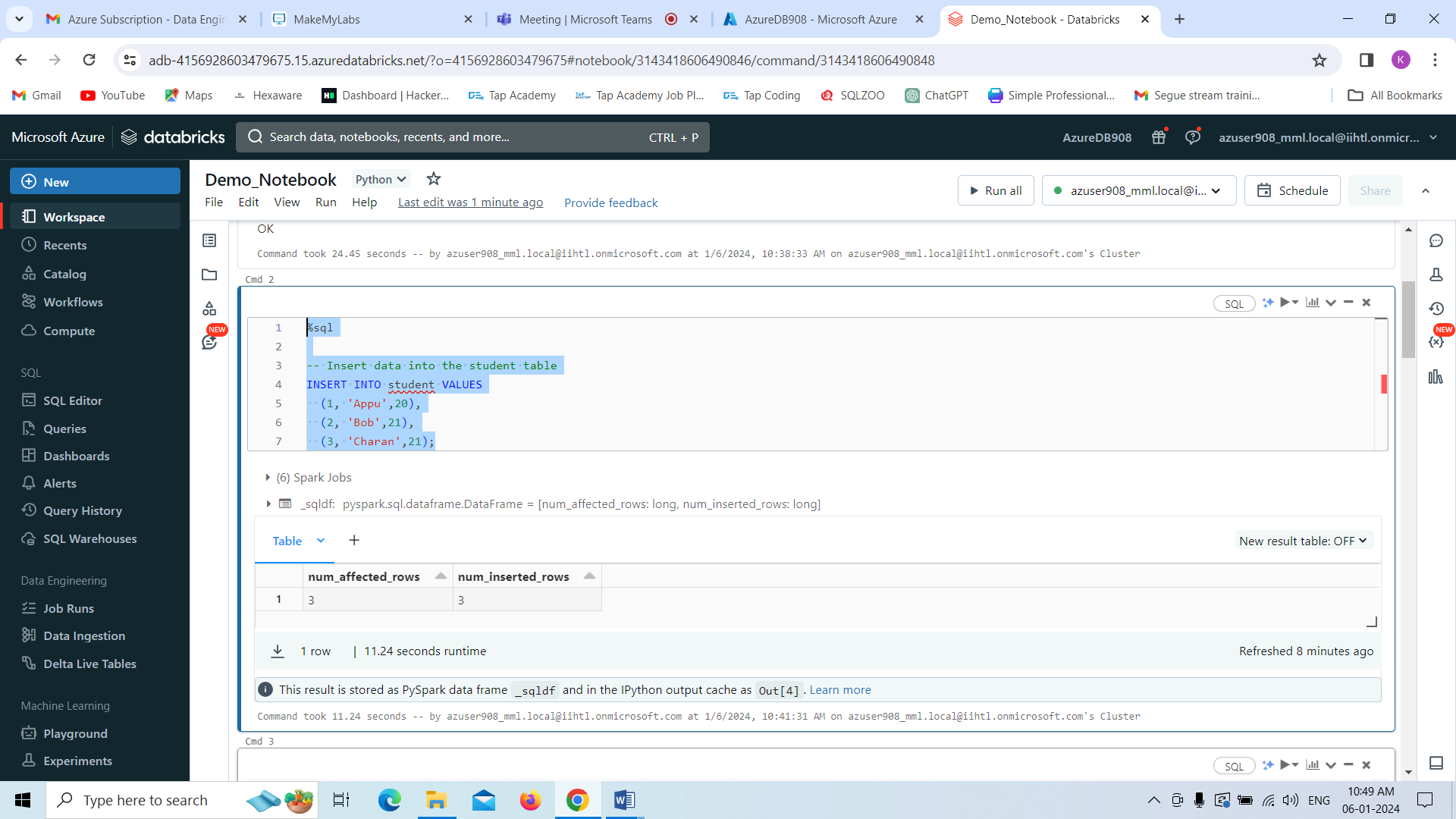
Now we are creating the notebook



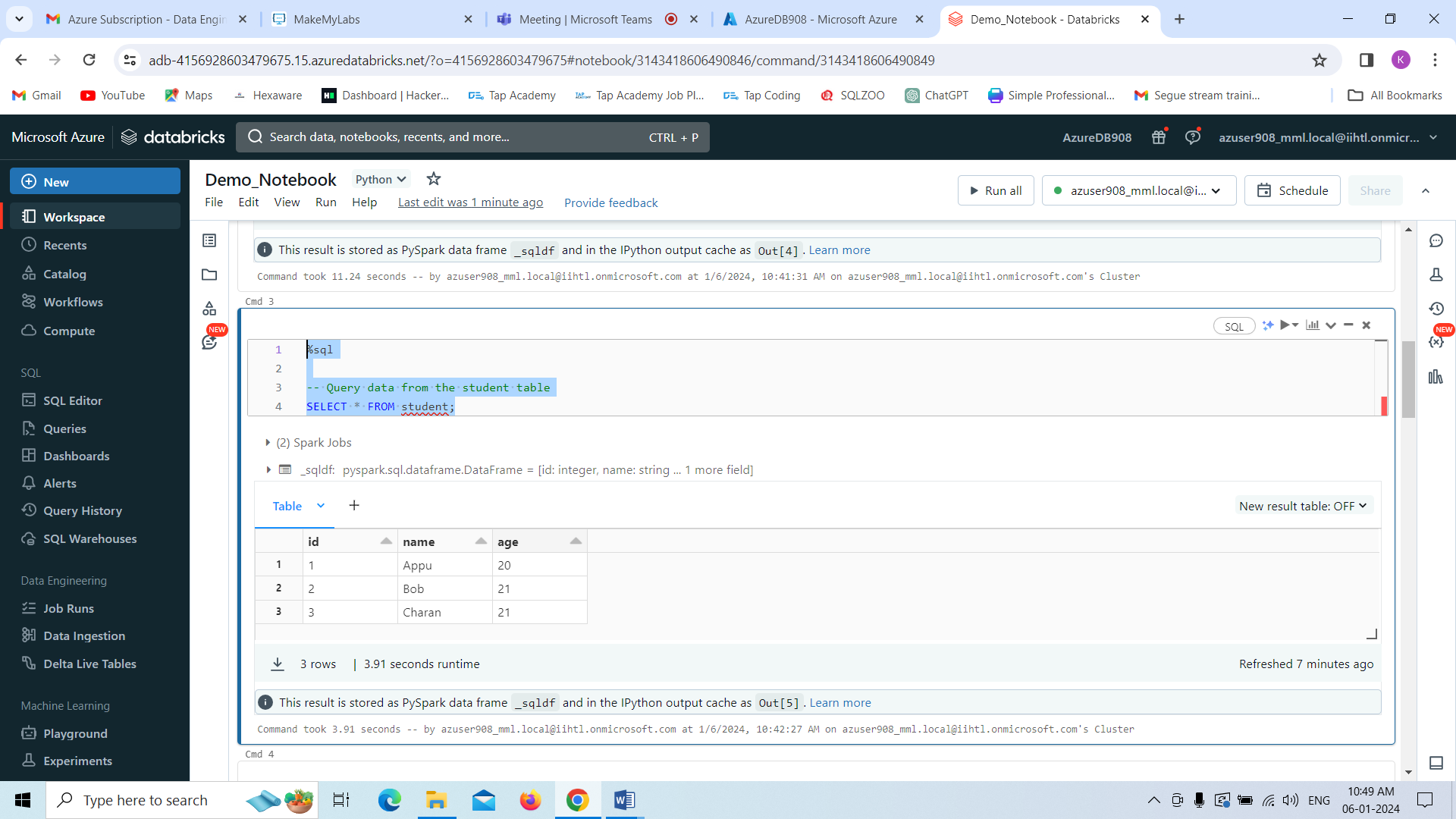
Now I am running all the commands in the notebook by **using SQL** as base. Here I am creating the table as shown in below.



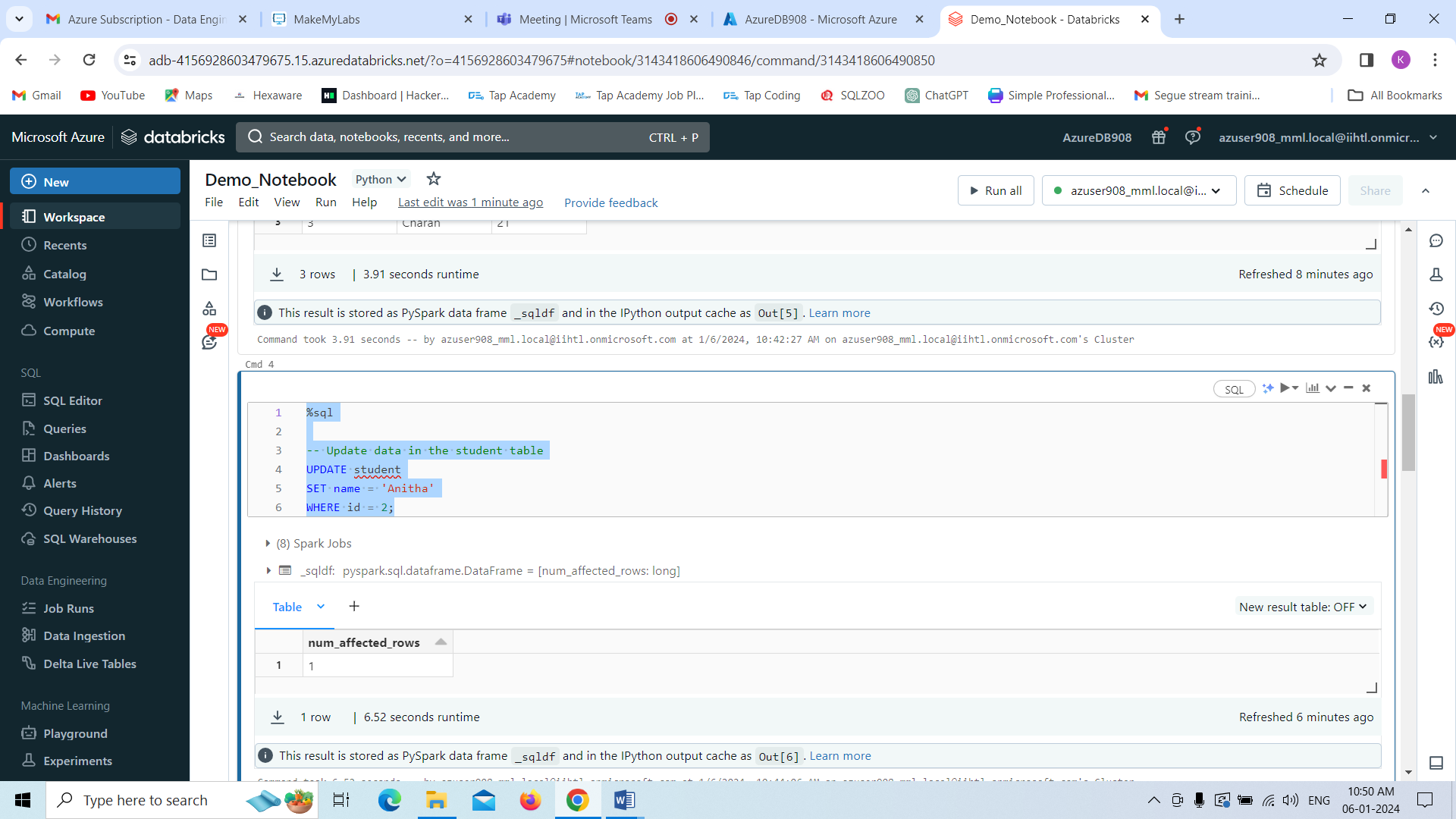
After that I am just inserting the values into it by using **insert** command.



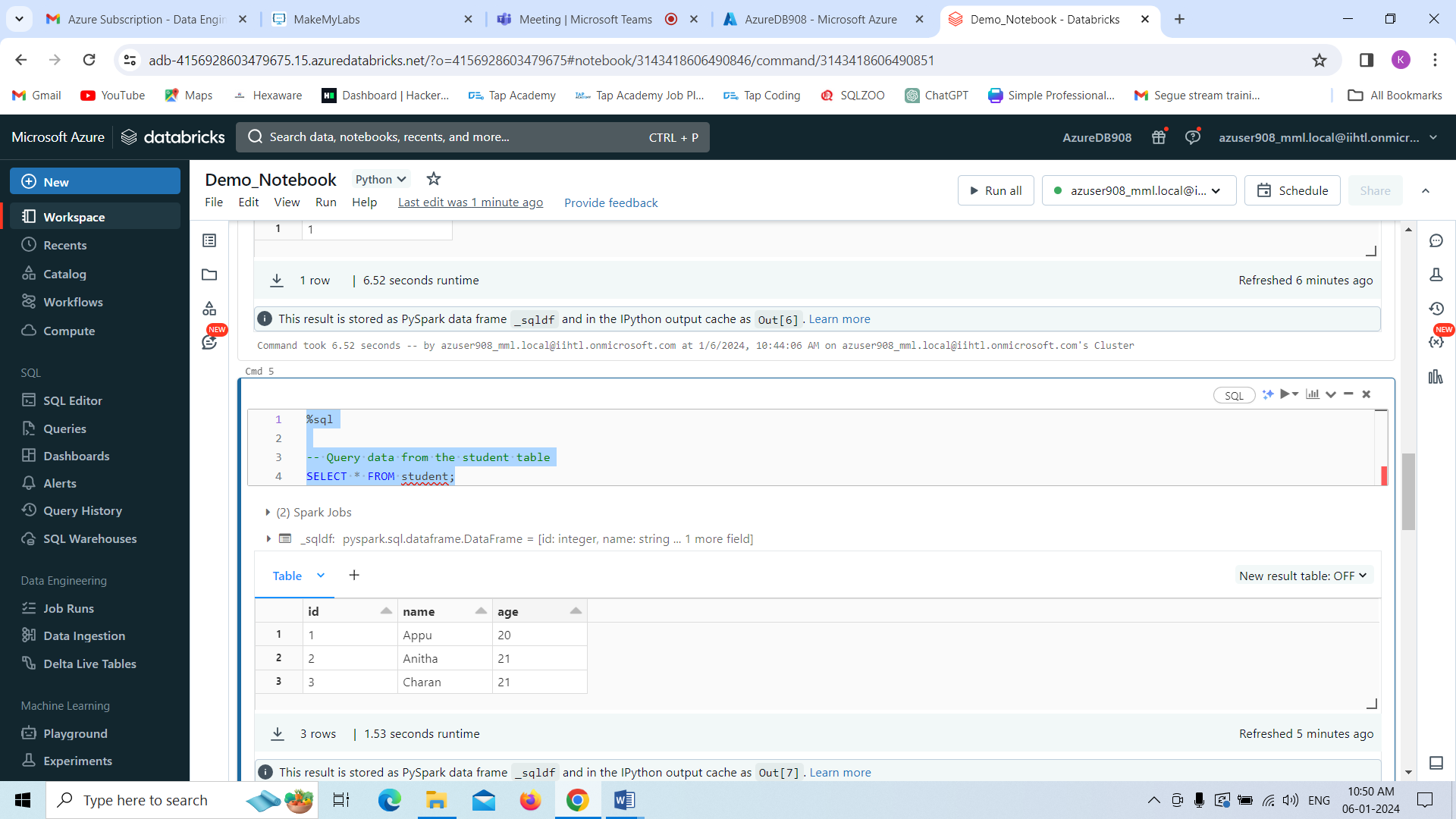
Now we are displaying the table content by using **select** command.



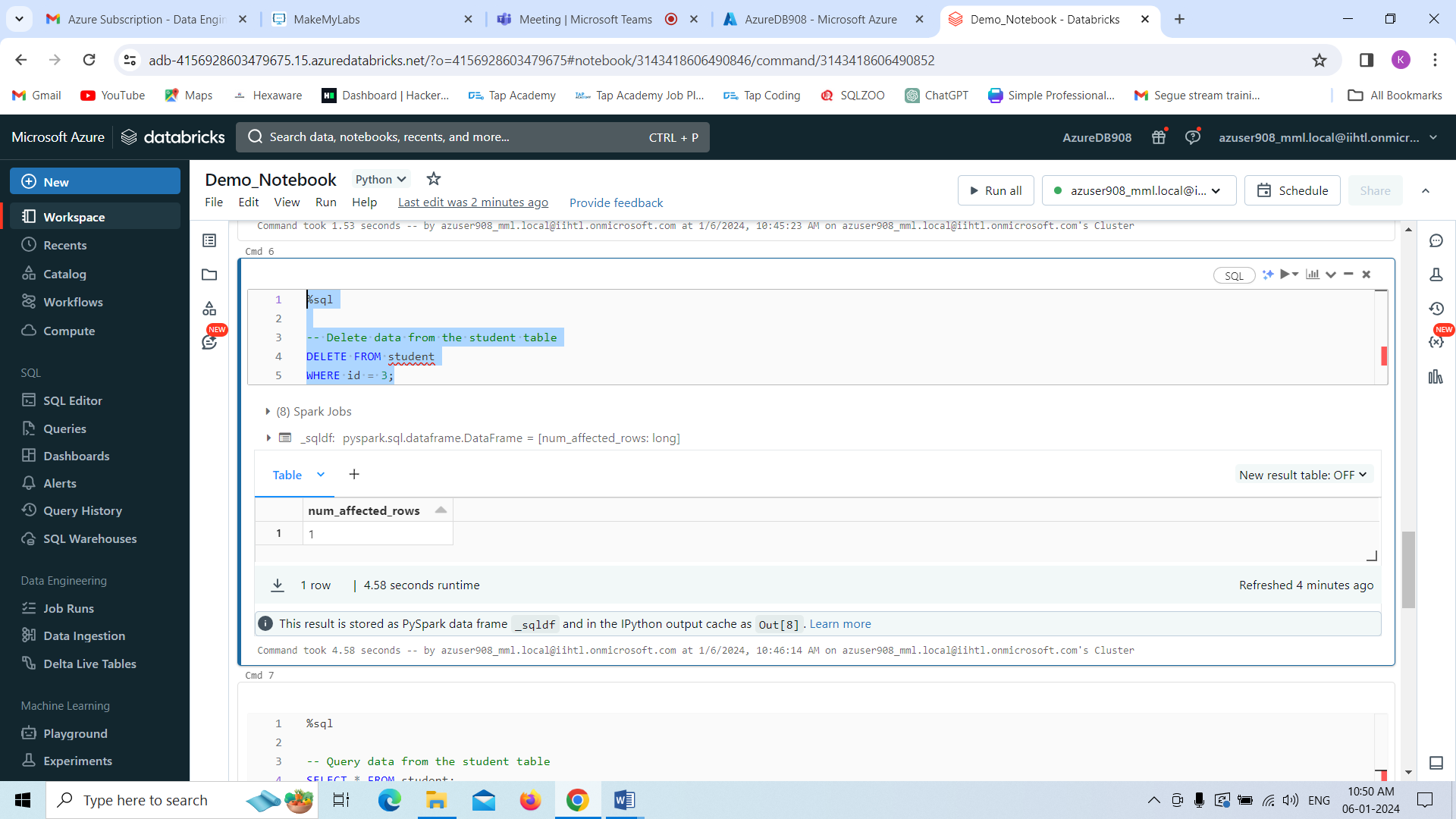
Now I am updating the records by using **update** command.



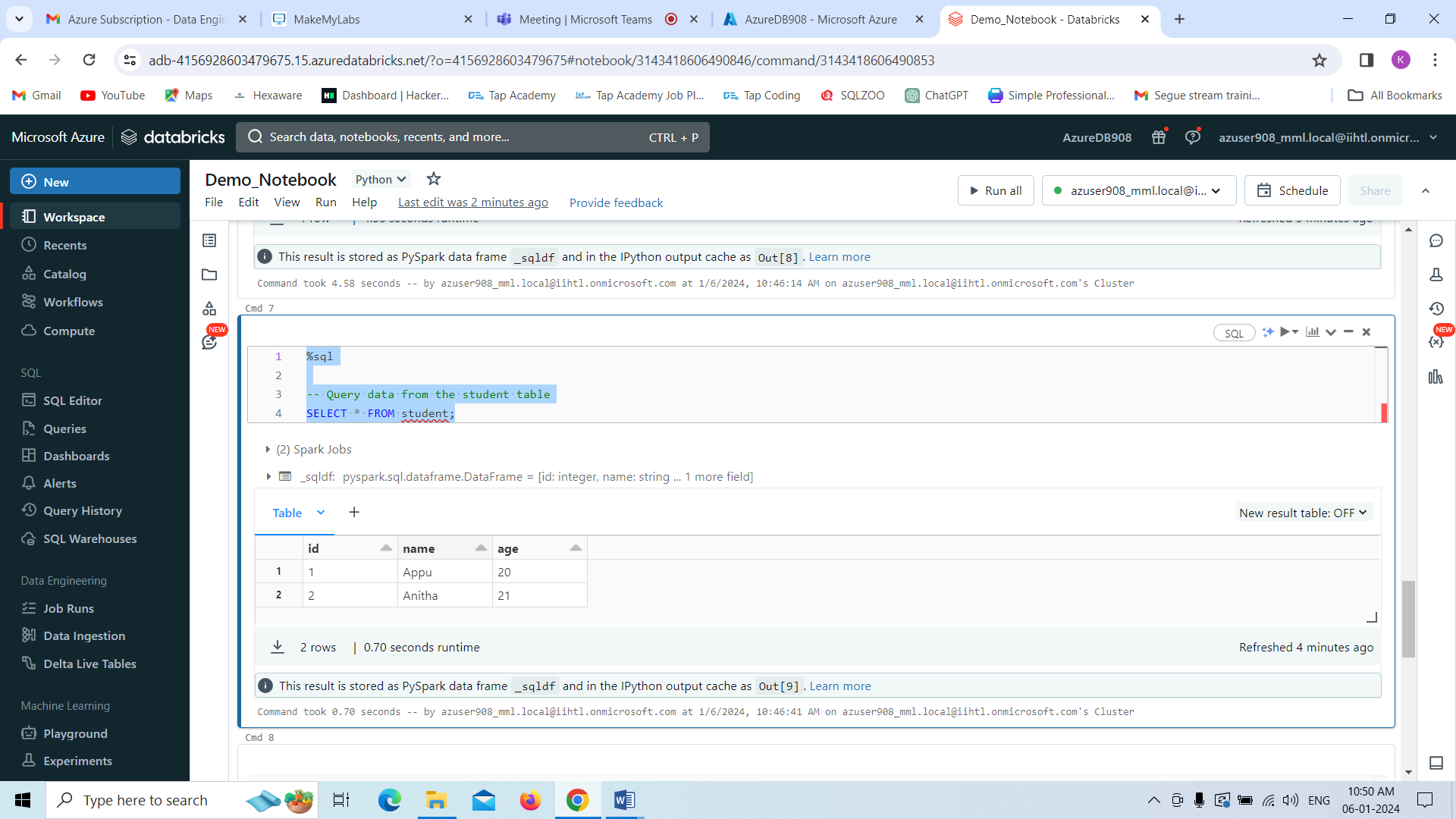
Below we can see the table ,it has been updated.



Now I am deleting a record by using delete command.



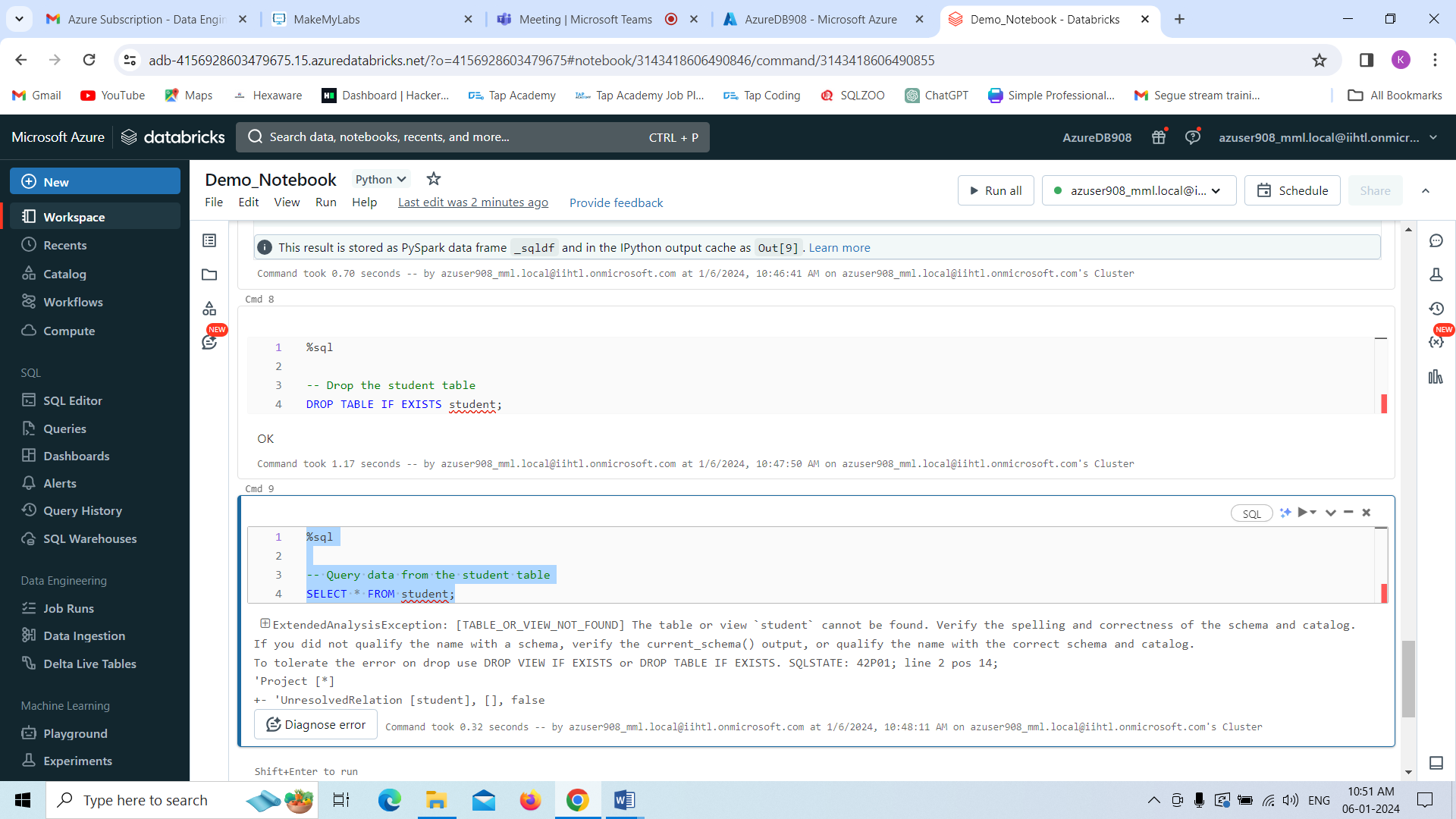
Here you can see ,only two records are showing in the table.



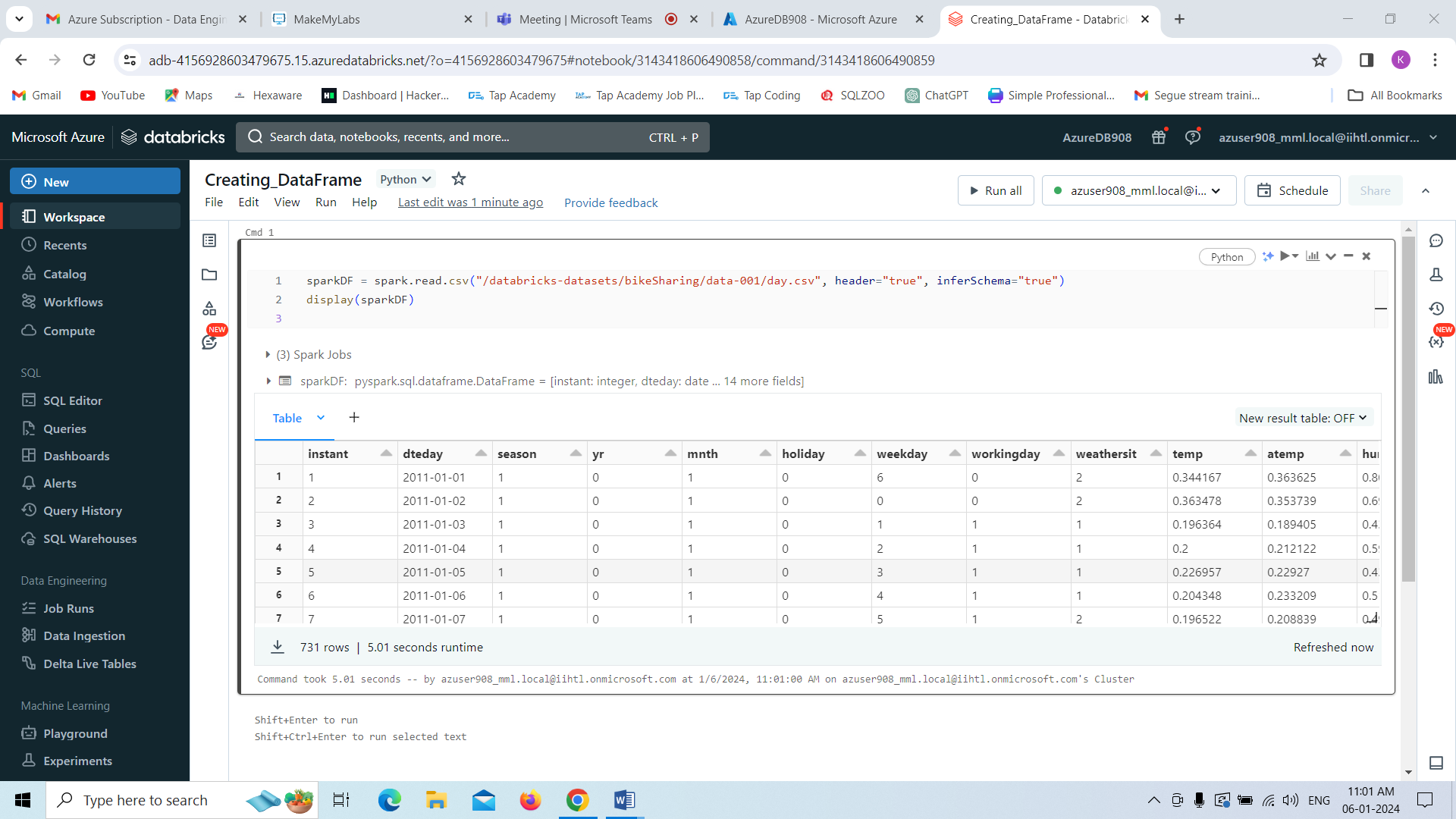
Now I am dropping the table by using dropping command that means entire table got created.



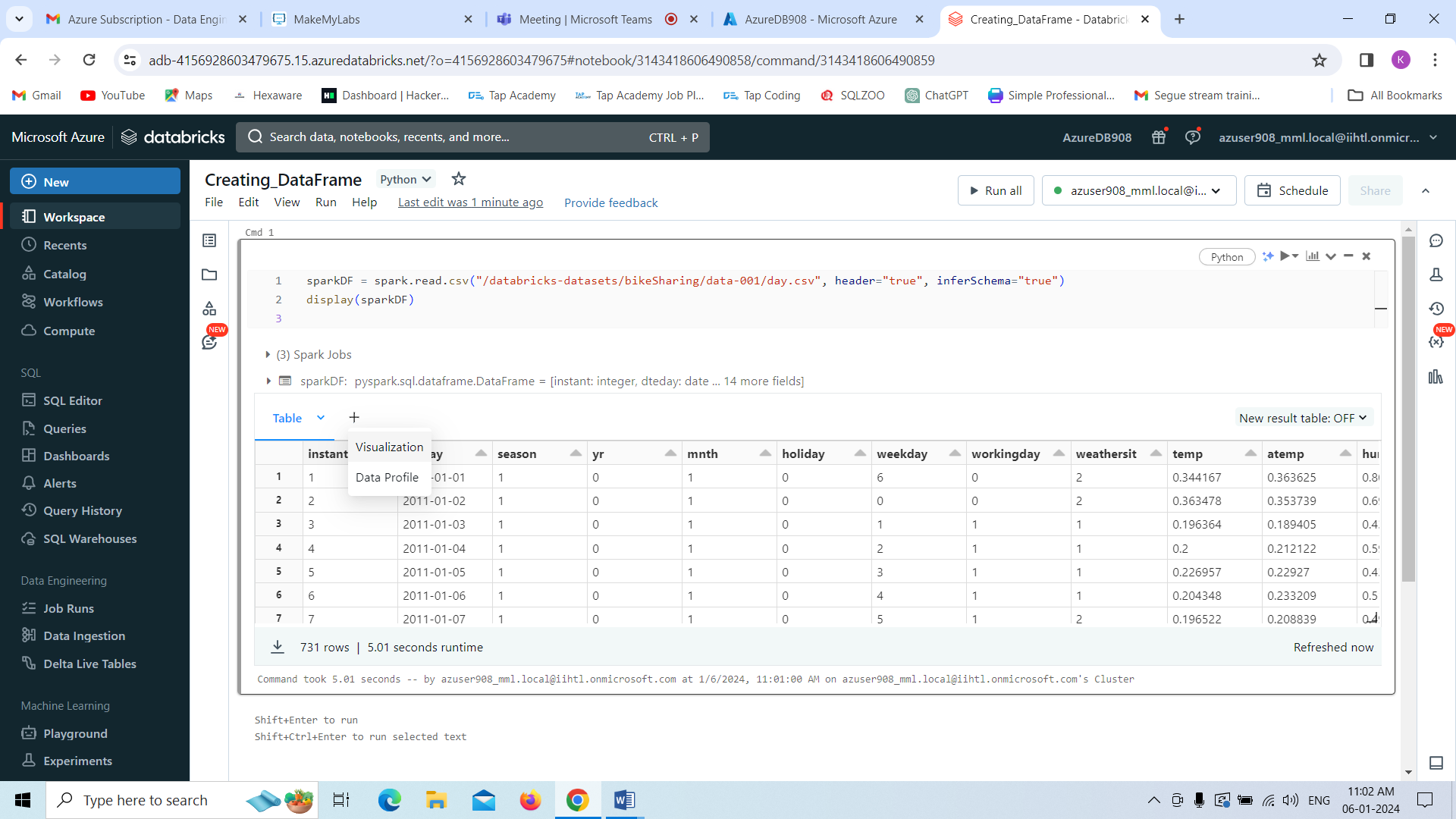
Here you can see the result as follows.



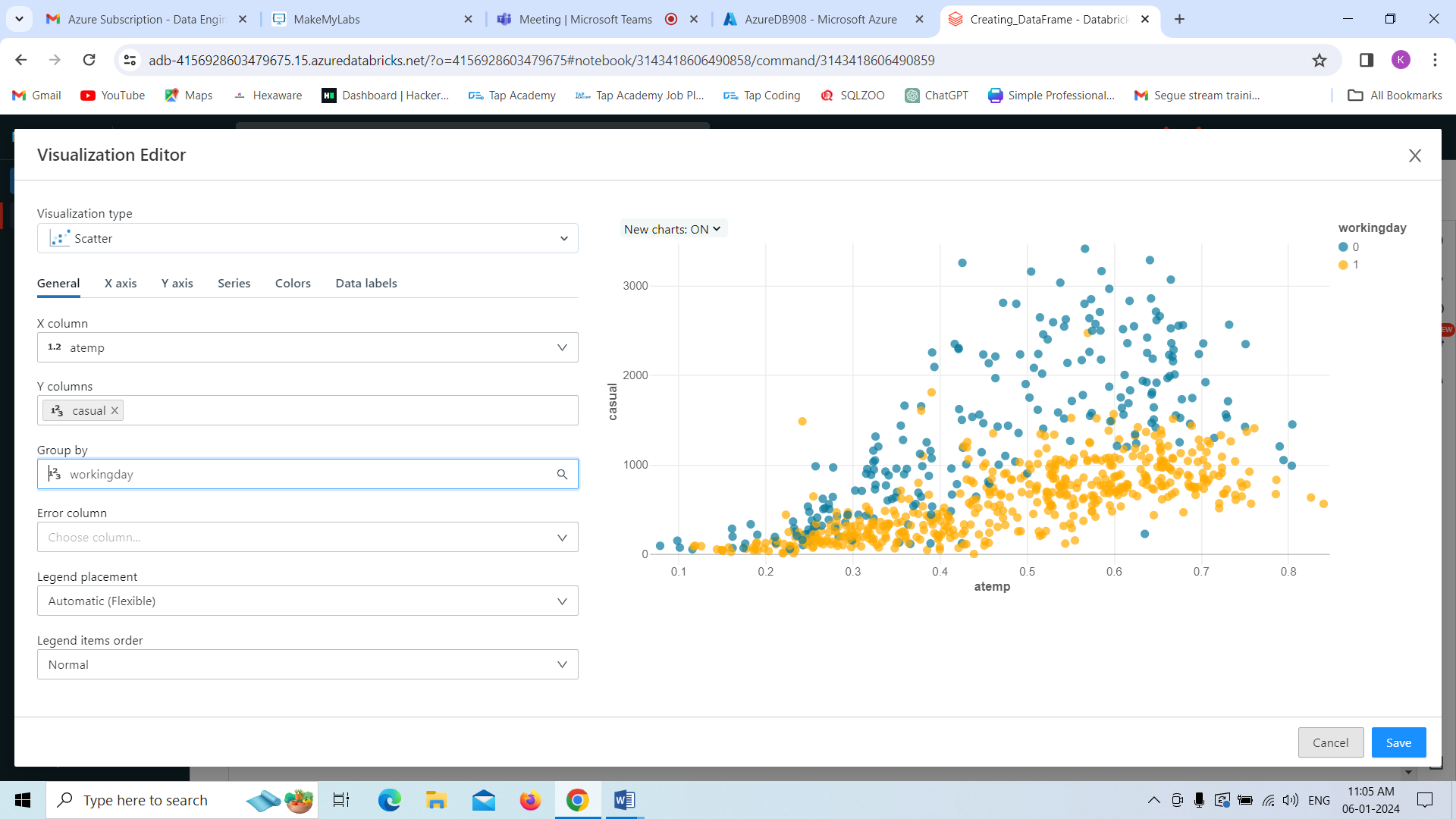
creates a Data Frame from a Databricks dataset& Create a Visualizations in Databricks notebooks  
&Rename, duplicate, or remove a visualization or data profile.

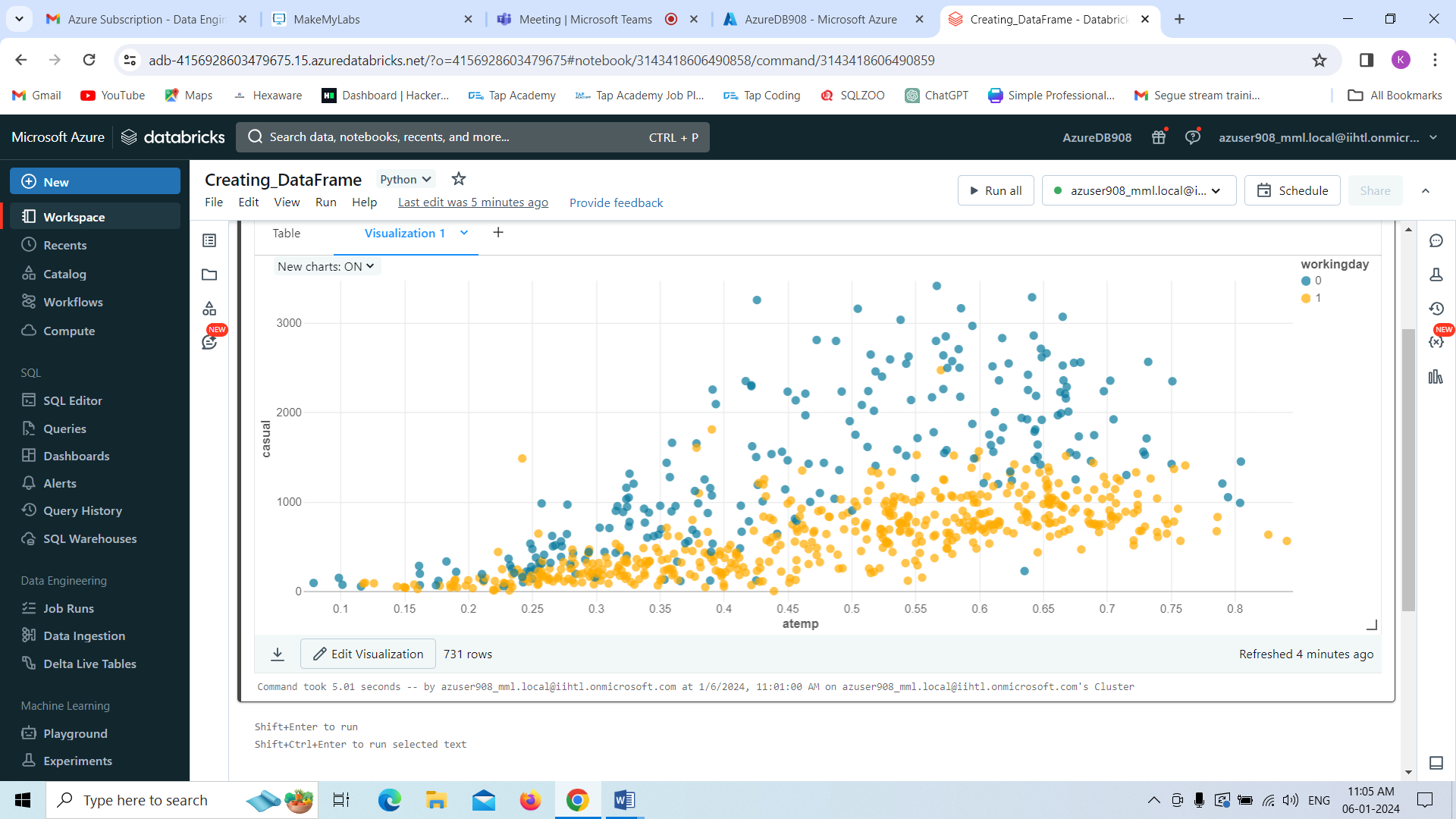
Below,I just created a dataframe from the existing Dataset only. That is showing as output.  


Now I am creating the visualisation by clicking Visualisation.

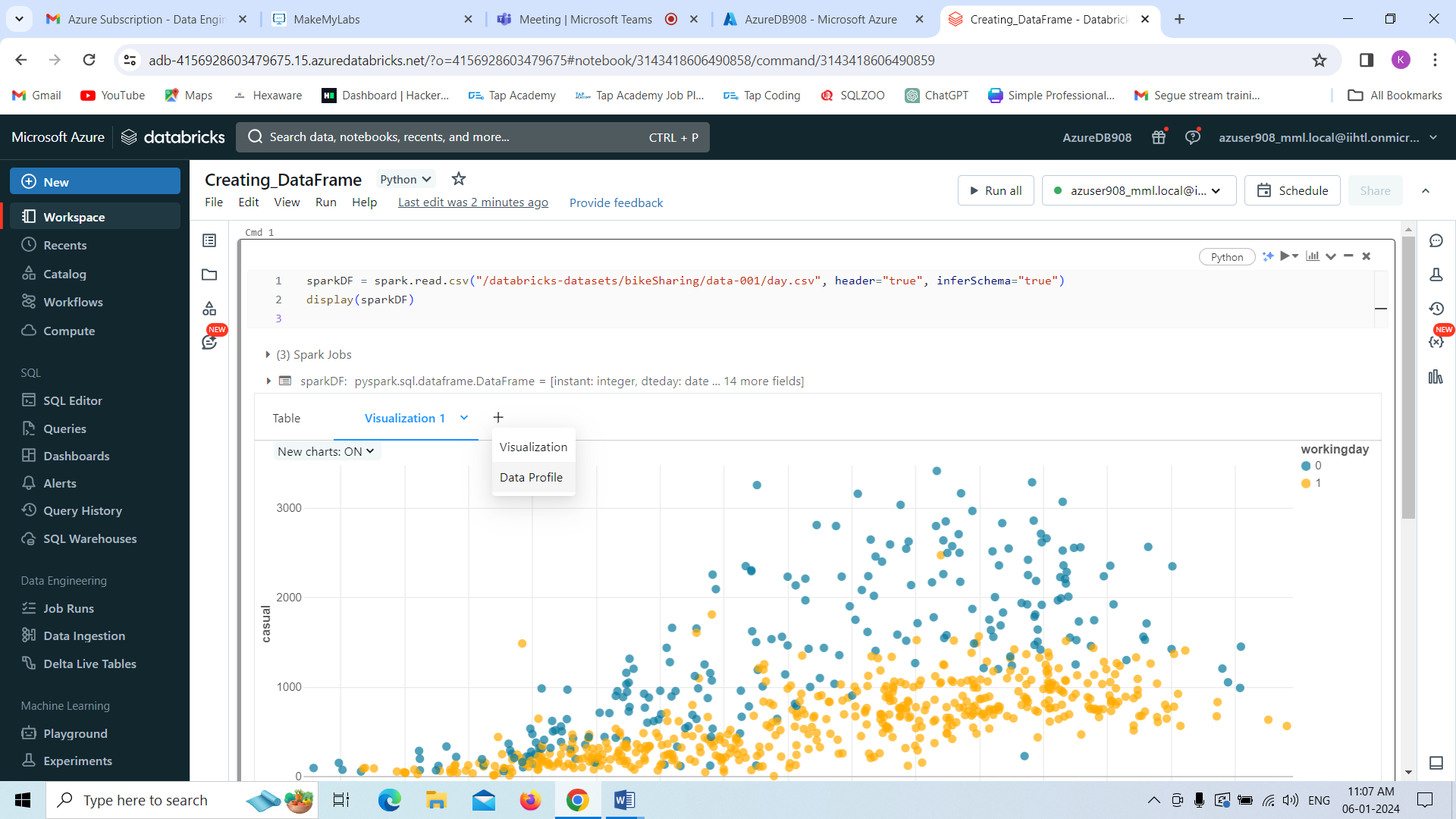
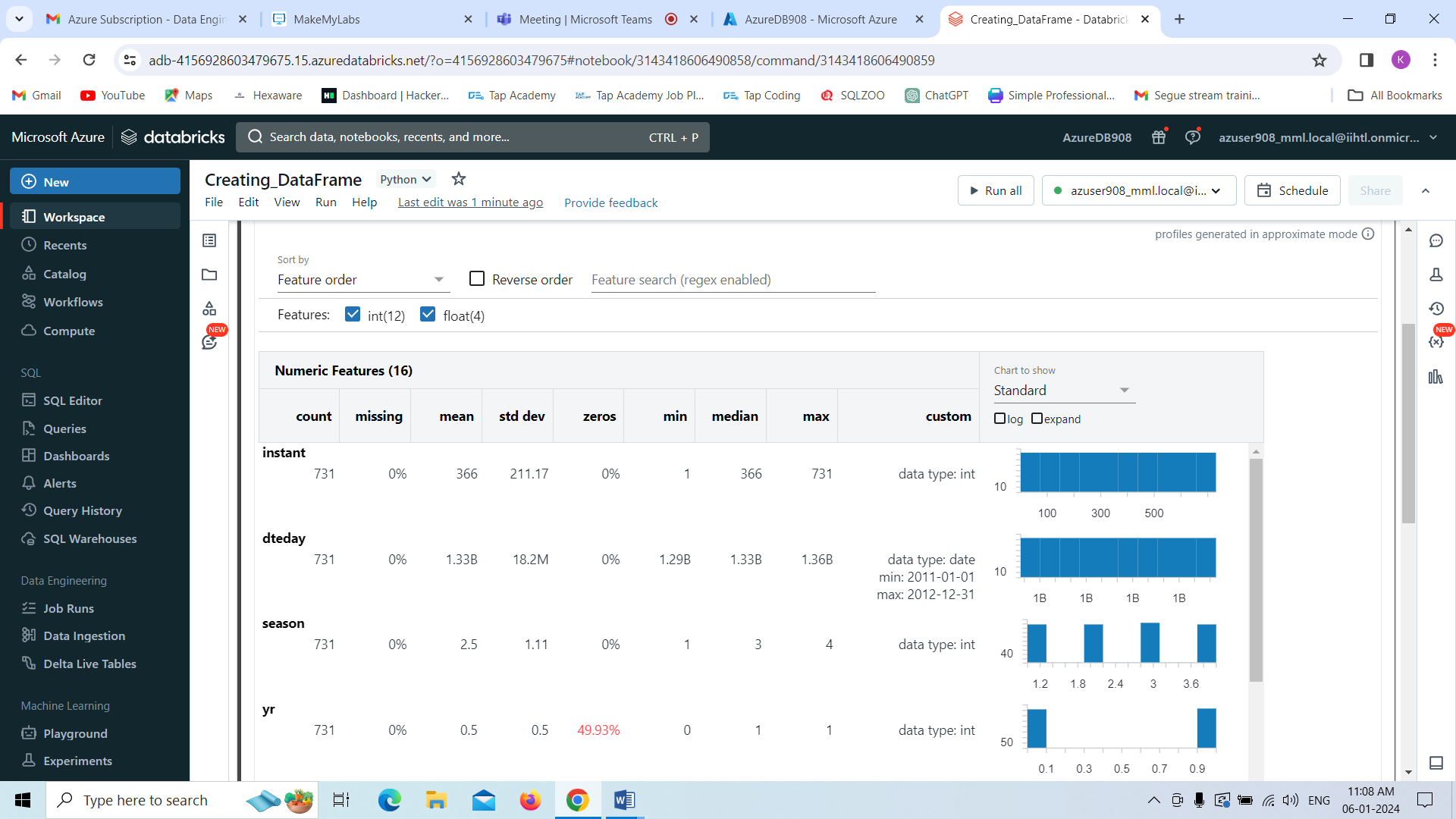


By selecting visualization types and below options we can create the visualization as follows.

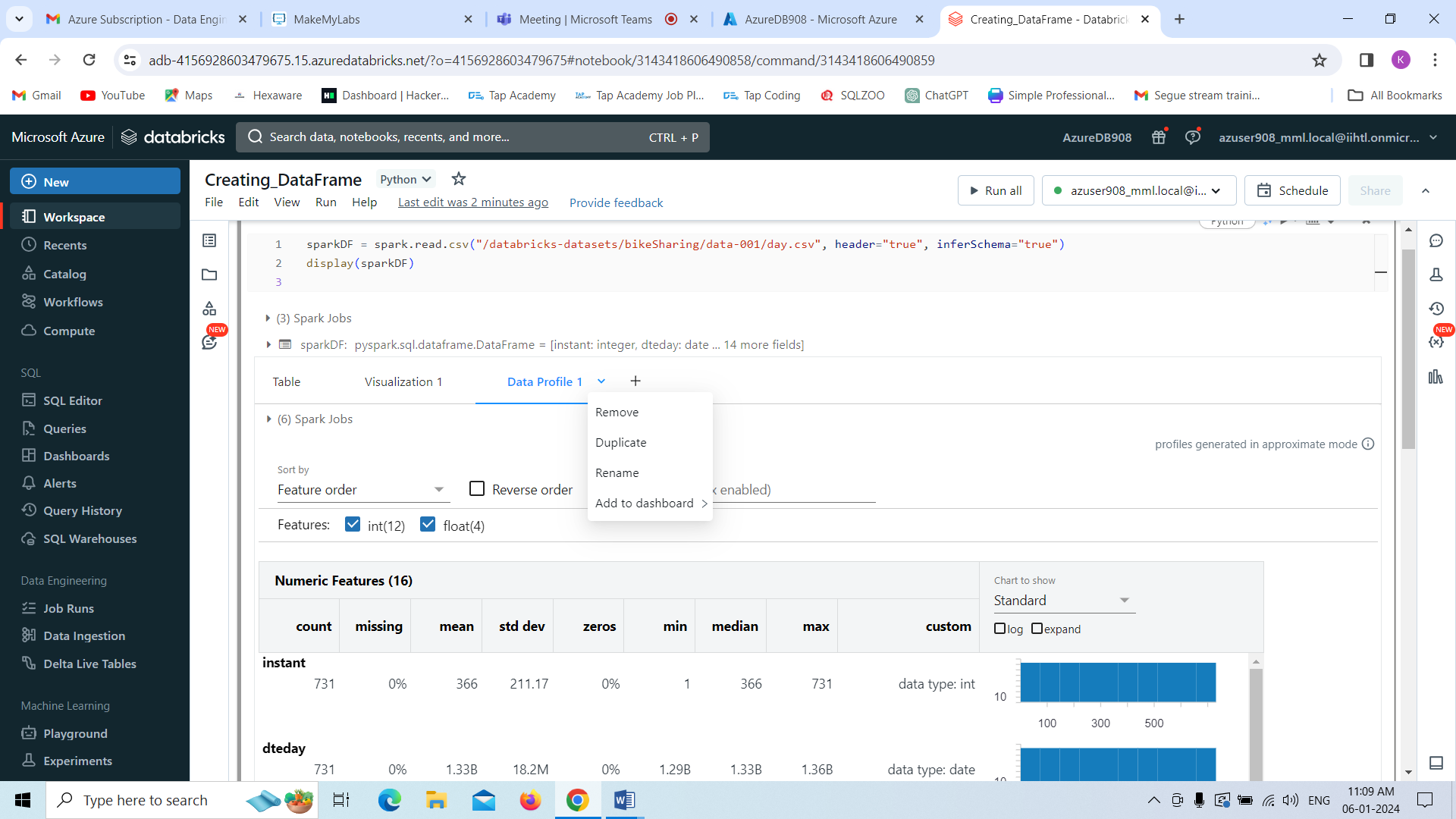




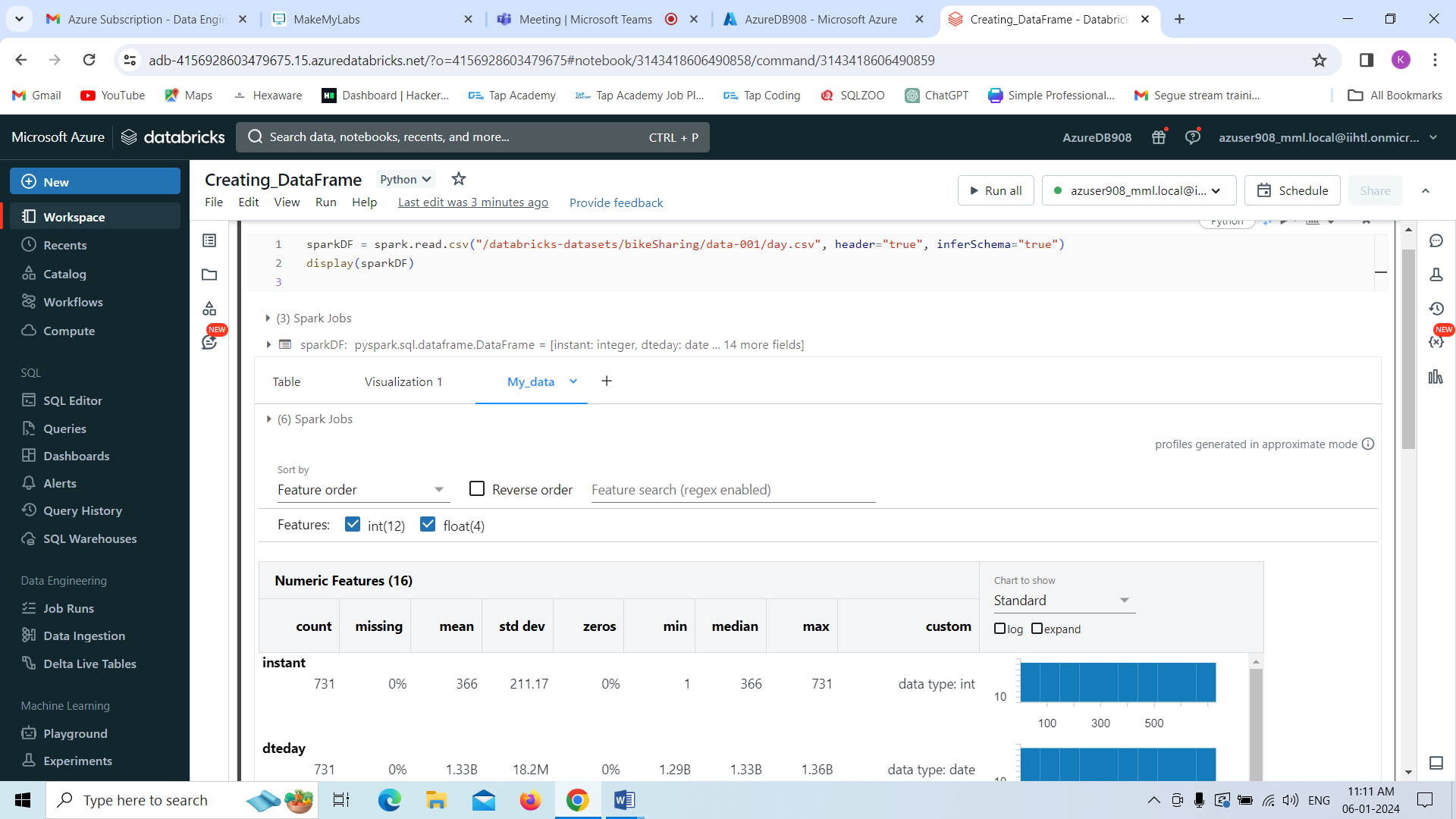
Now we are creating the Data Profile by clicking the **Data Profile Option**.

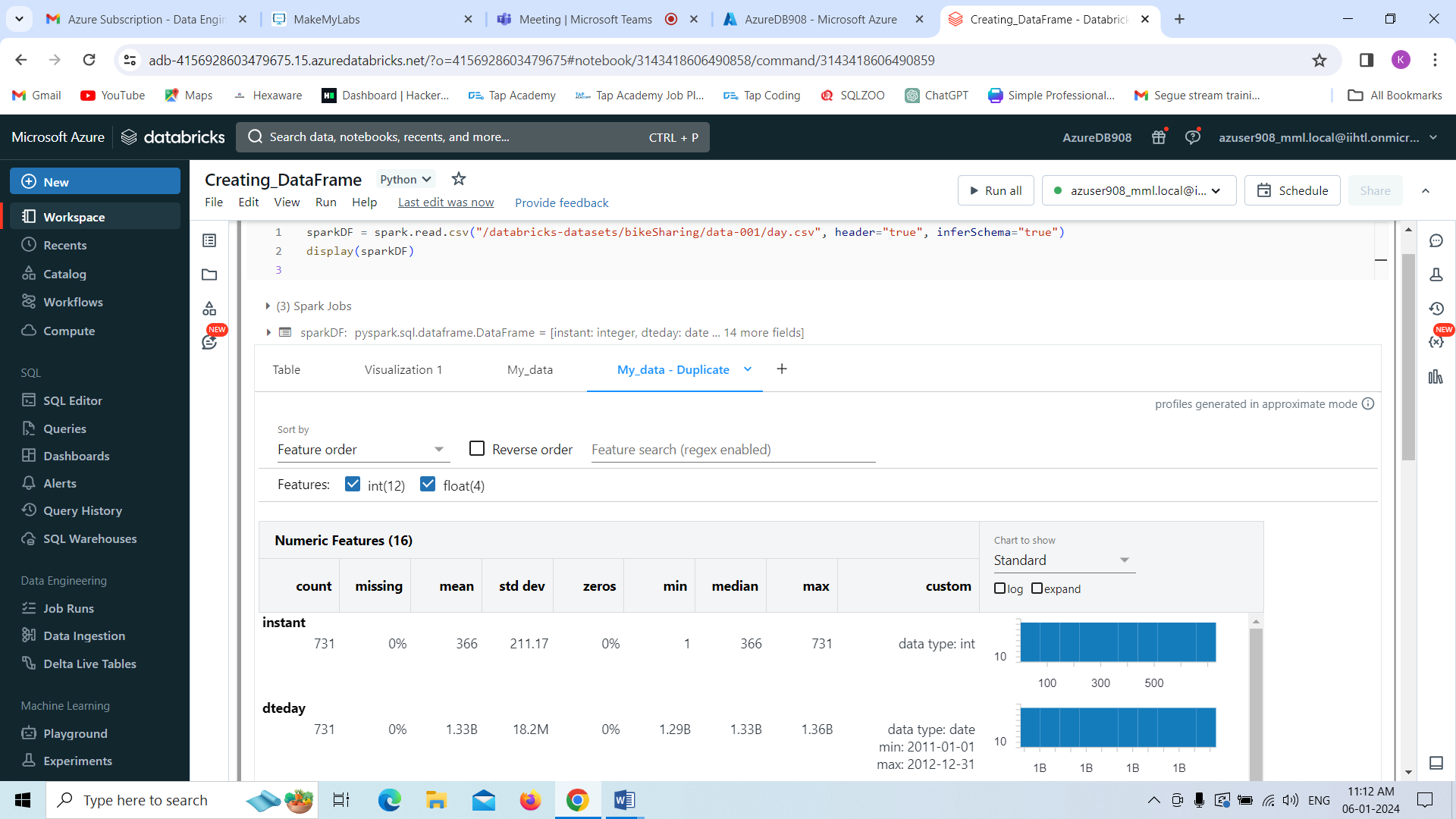
Now we are perfoming the manipulations on the data profile.



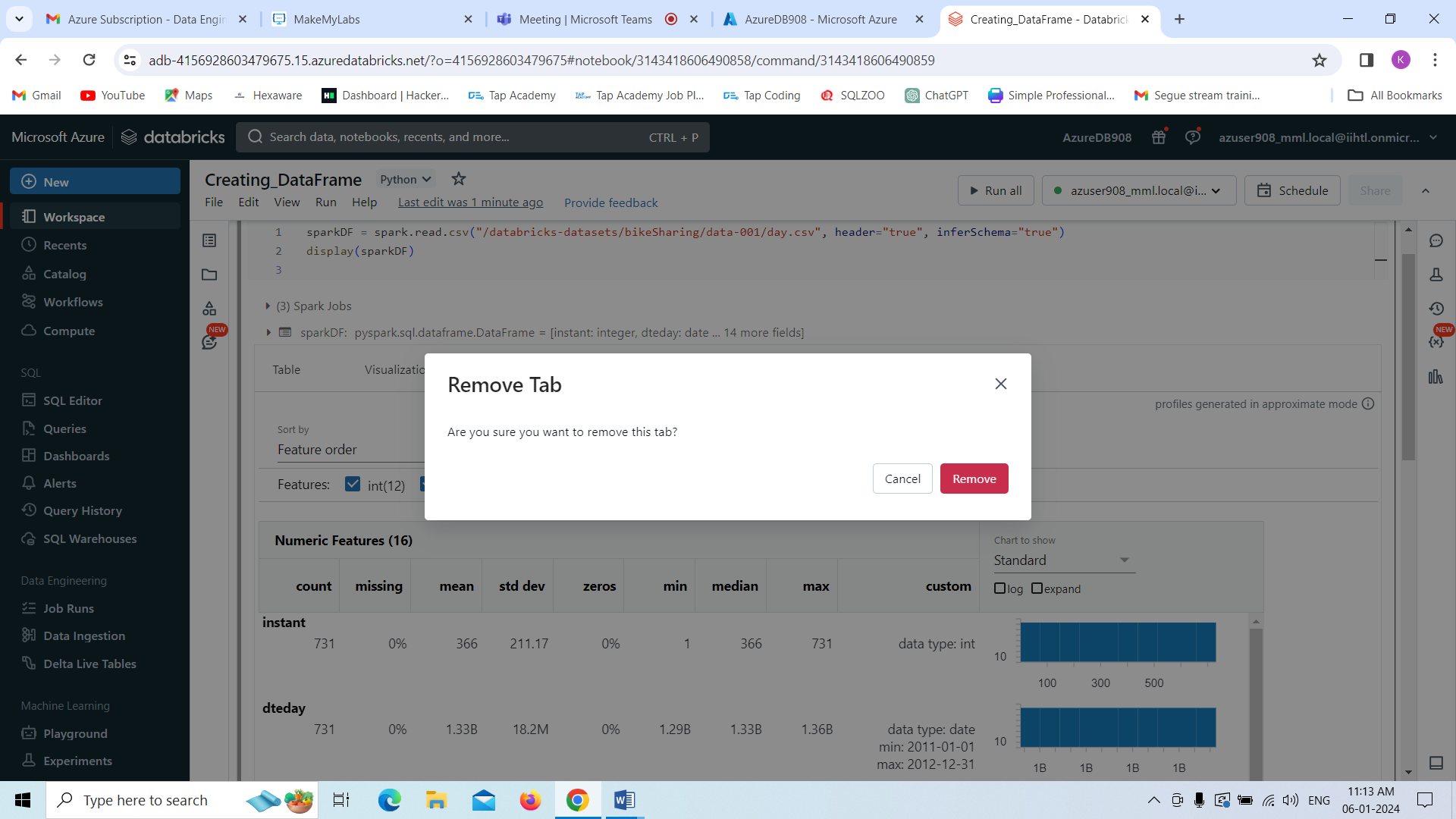
By clicking rename we can rename our data profile as **My\_Data**



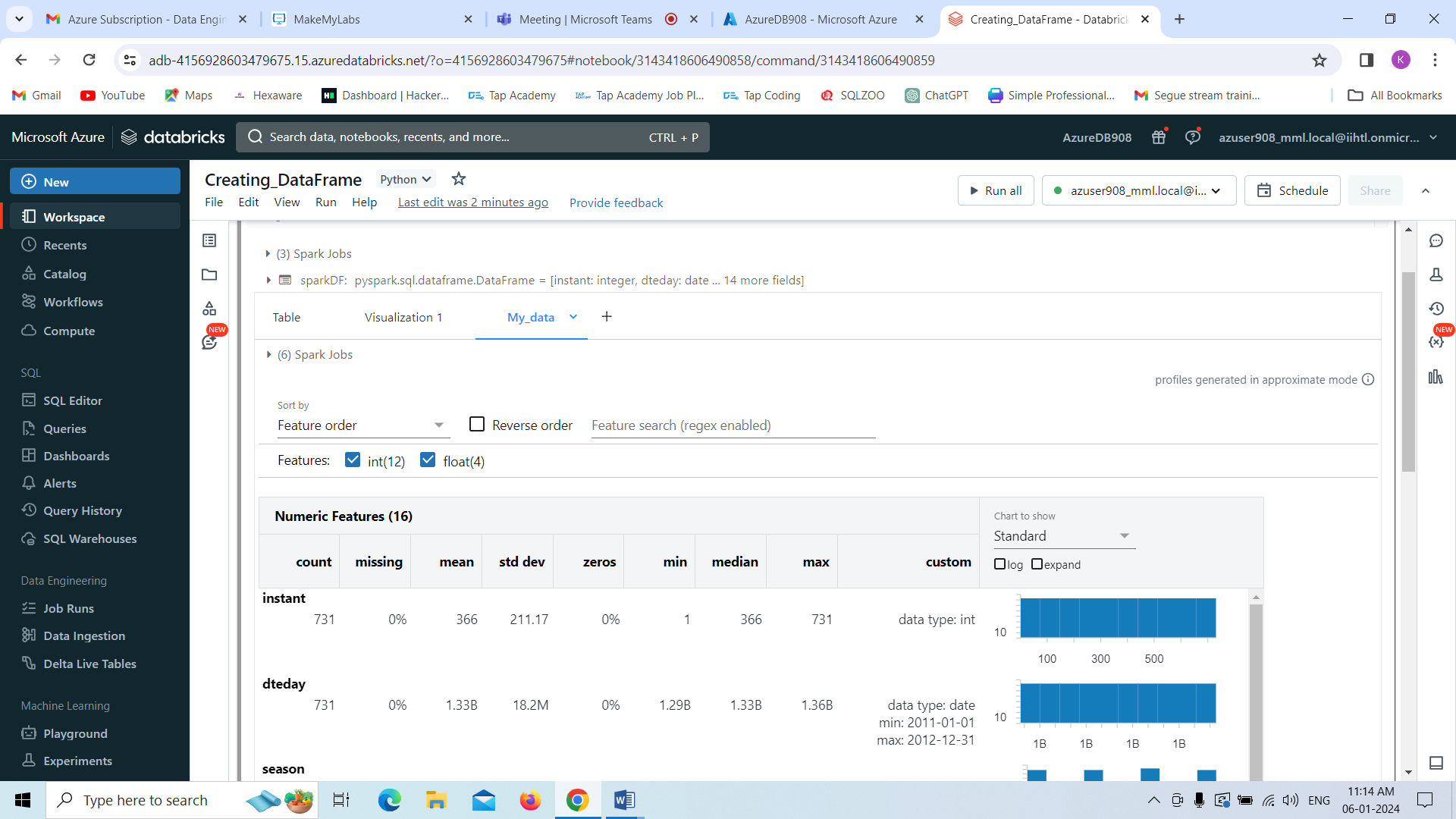
Duplicating the data profile by using **duplicate** option.



And Finally I am removing the Duplicated data profile by using **remove** option.



Now we can see as follows.



2.Explain the copy activity in Azure data factory.

1. If we want to perform copy Activity we need storage account in that we need to have a folder as your wish I choose as **input** for my understanding in the blob Storage.
2. Now we need to create Data Factory in the Azure Portal.
3. We have to use the copy data tool for coping data.
4. In the home page of Azure Data Factory, select the **Ingest** tile to start the Copy Data tool.
5. Now in the Properties page of the Copy Data tool, choose **Built-in copy task** under **Task type**, then select Next.
6. Click **Create new connection** to add a connection in the portal.
7. We need to select the linked service type that you want to create for the source connection.
8. Here we use **Azure Blob Storage**. Select it from the gallery, and then select **Continue**.
9. Now in the New connection (Azure Blob Storage) page, we need to specify a name for your connection.
10. We need to select our Azure subscription from the **Azure subscription** list and our storage account from the **Storage account name** list, test connection, and then select **Create**.
11. Select the newly created connection in the **Connection** block in the workspace environment.
12. In the **File** section, select **Browse** to navigate to the **input** folder, select the **sales.txt** file, and then click **OK**.
13. Select the **Binary copy** checkbox to copy file as-is, and then select **Next**.
14. Select the **Azure Blob Storage** connection that we created in the **Connection** block.
15. In the **Folder path** section, enter **output** for the folder path.
16. We need to leave all other settings as default and then select **Next**.
17. On the **Settings** page, we need to specify a name for the pipeline and its description, then select **Next** to use other default configurations.
18. On the **Summary** page, we need to review all settings, and select **Next**.
19. On the **Deployment complete** page, select **Monitor** to monitor the pipeline that you created.
20. The application switches to the **Monitor** tab. You see the status of the pipeline on this tab. Select **Refresh** to refresh the list. Click the link under **Pipeline name** to view activity run details or rerun the pipeline.
21. On the Activity runs page, select the **Details** link under the **Activity name** column for more details about copy operation.
22. Now we can see the results as green check box,that shows that out files are copied.
23. Now we have to come **to storage Account in Azure** and check that in that output folder **File is created or not or data is copied** or not in the file.