**Sunkari Manohar**

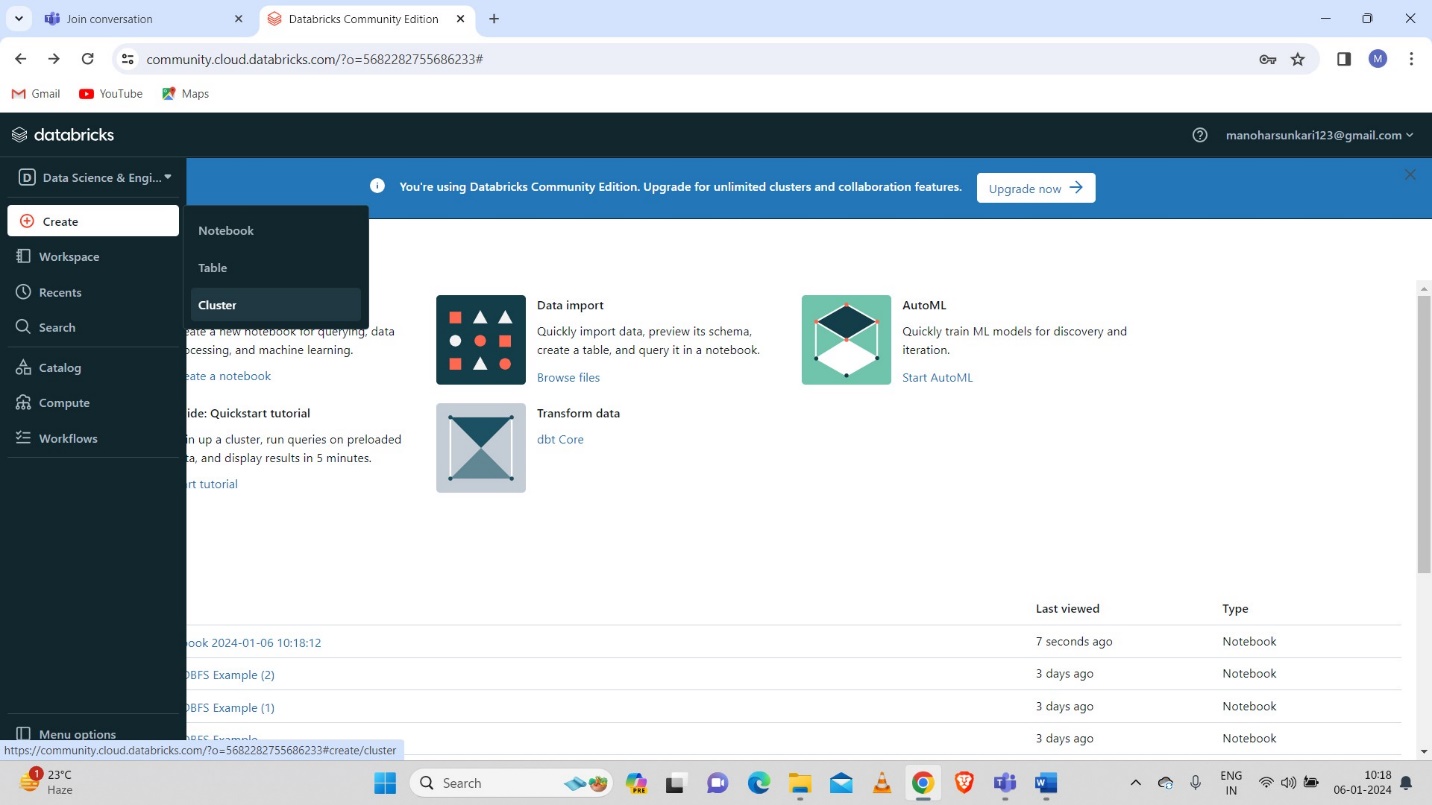
**Weekly Coding Assessment - 4**

**Azure Databricks**

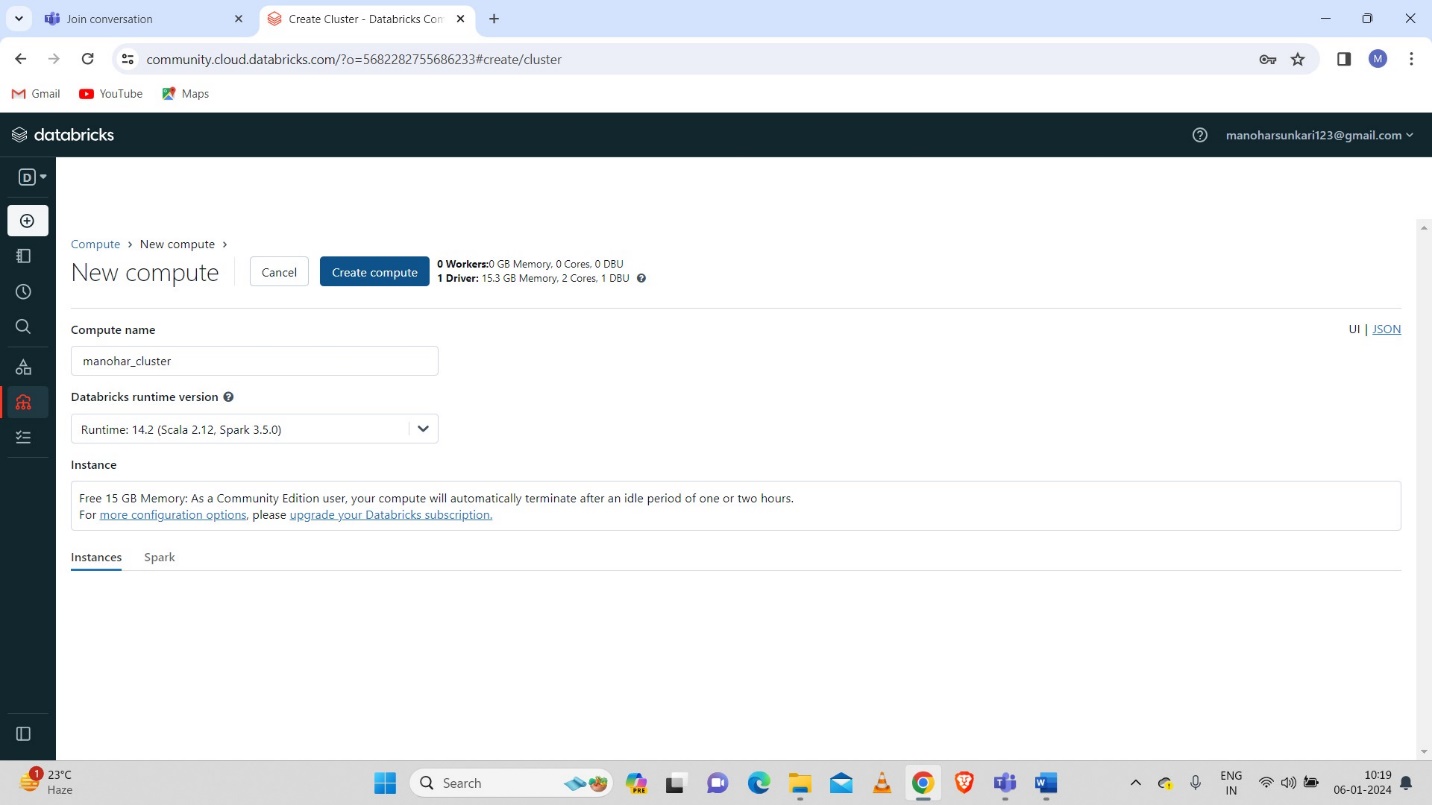
1.Create a cluster & Attach the notebook to the cluster and run all commands in the notebook & creates a DataFrame from a Databricks dataset & Create a Visualizations in Databricks notebooks  
&Rename, duplicate, or remove a visualization or data profile.

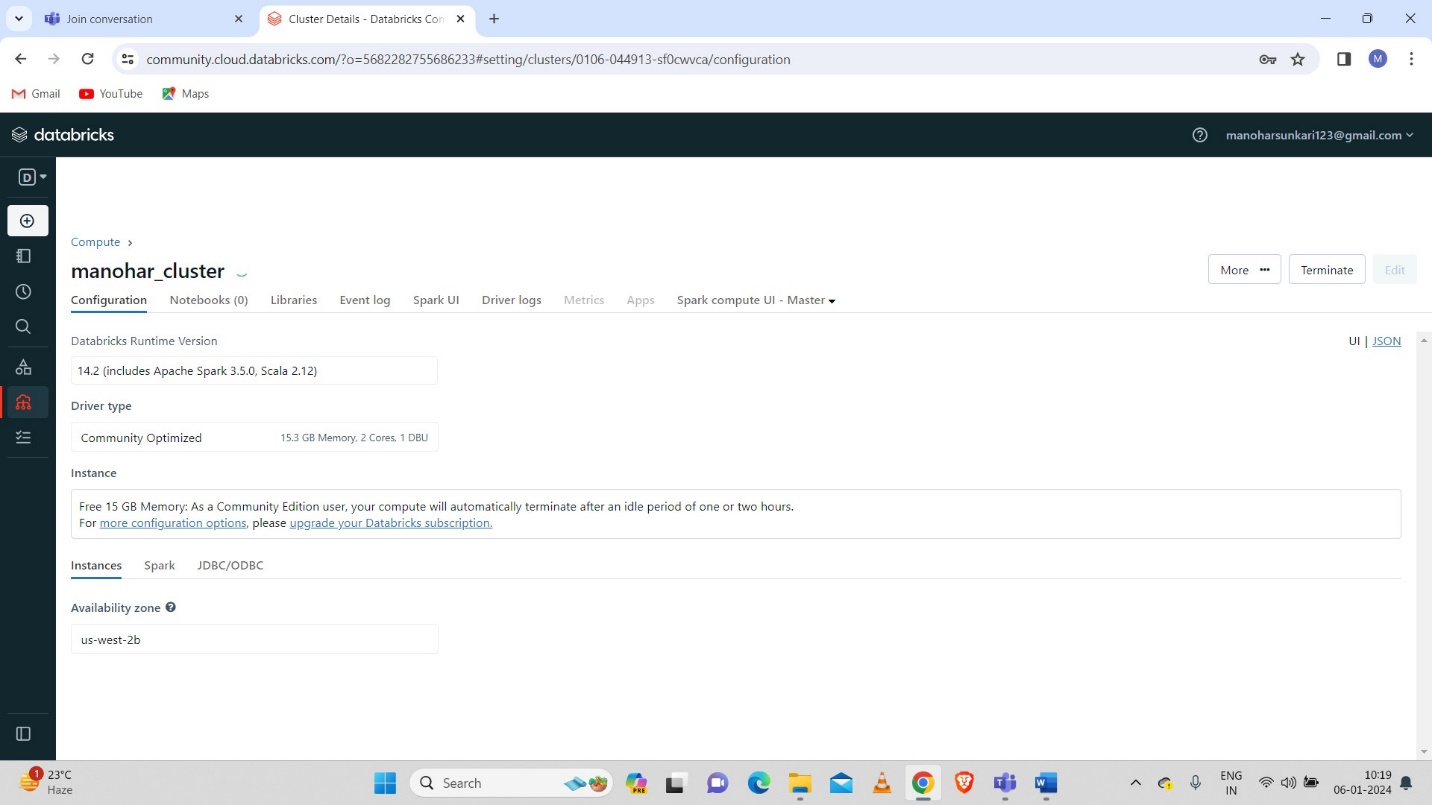
* **Creation of Cluster**

To create a new cluster we need to click on “+new” tab and select cluster.

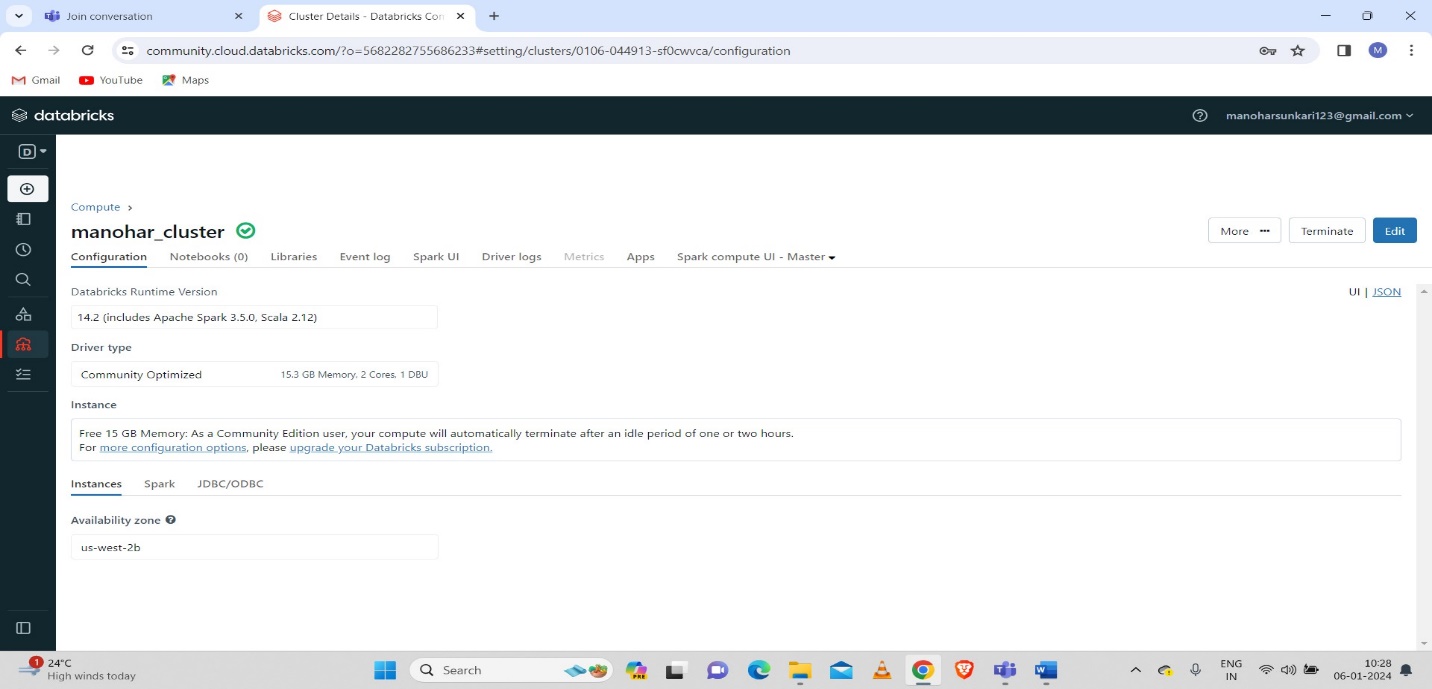


Now we have to give the cluster name and click on create.





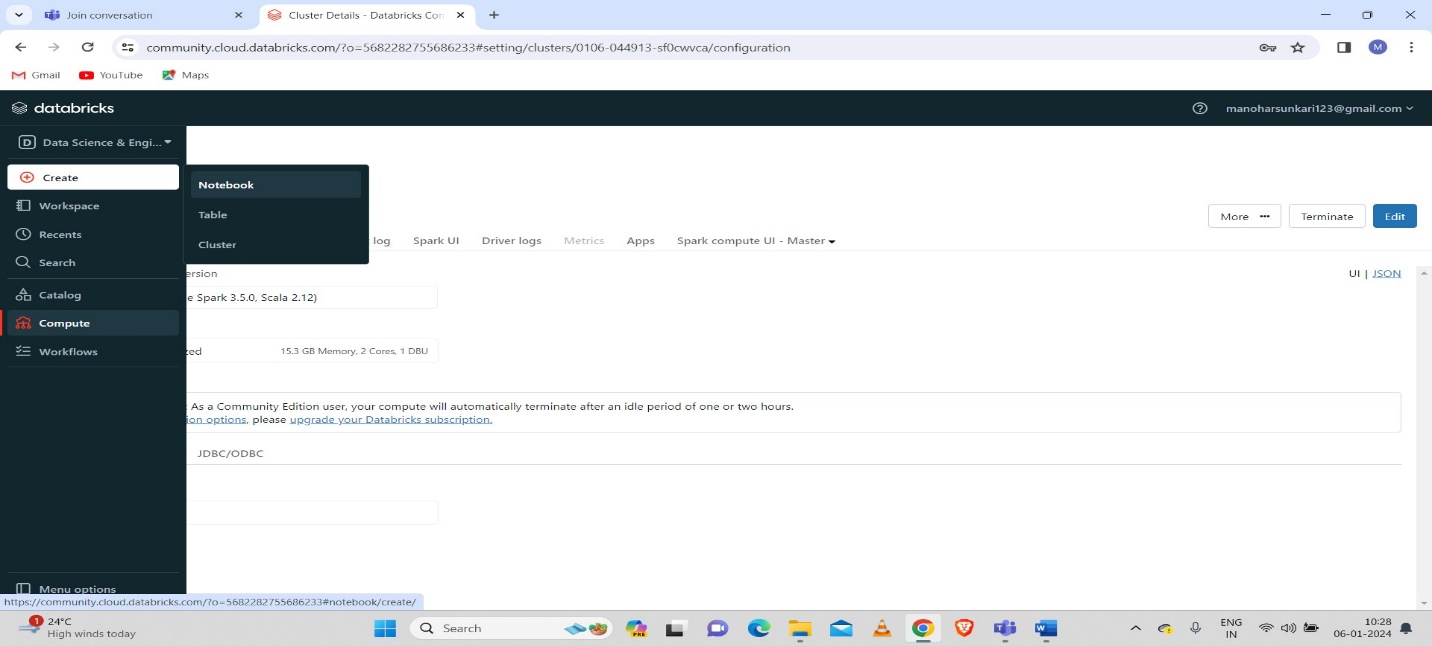
* **Created Cluster**

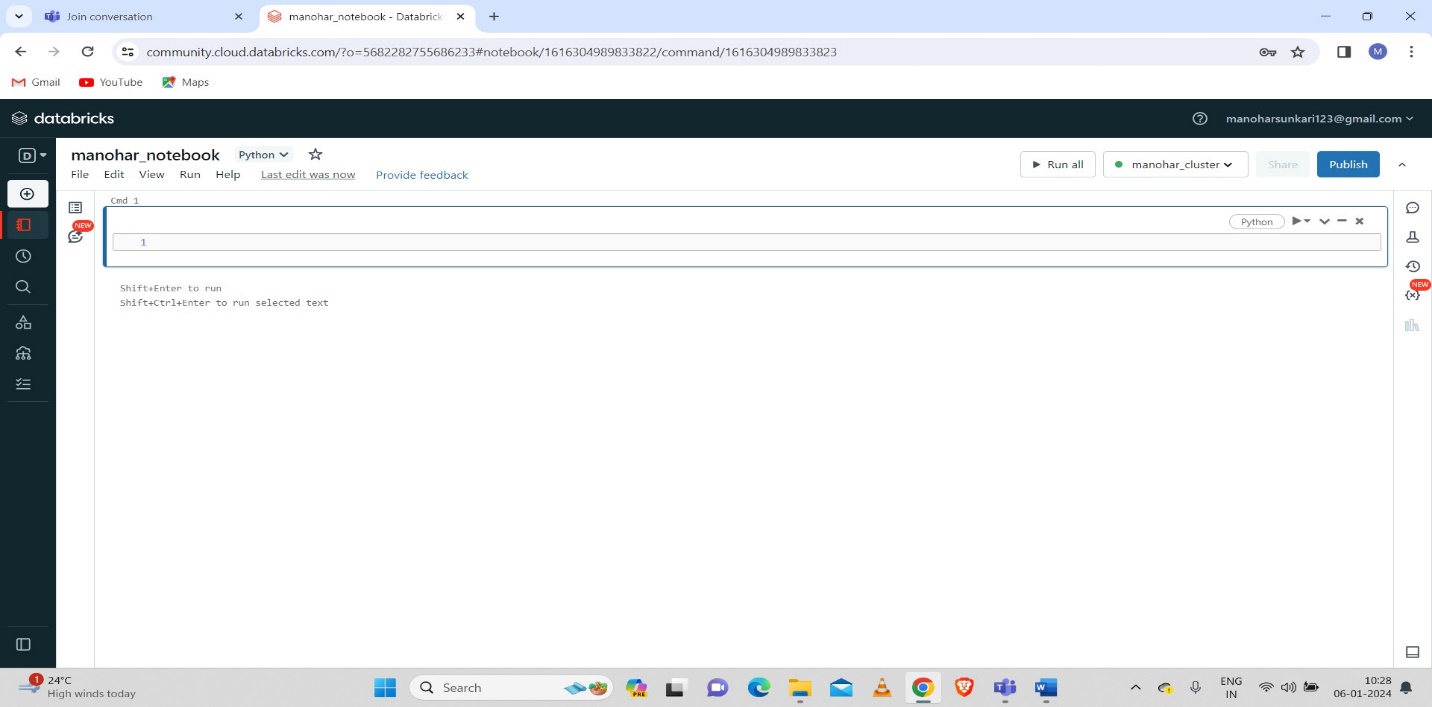


* **Creation of Notebook**

After Creation of cluster we need to create a Notebook to execute the commands.

To create a cluster we need to click on “+new” tab and select the notebook.

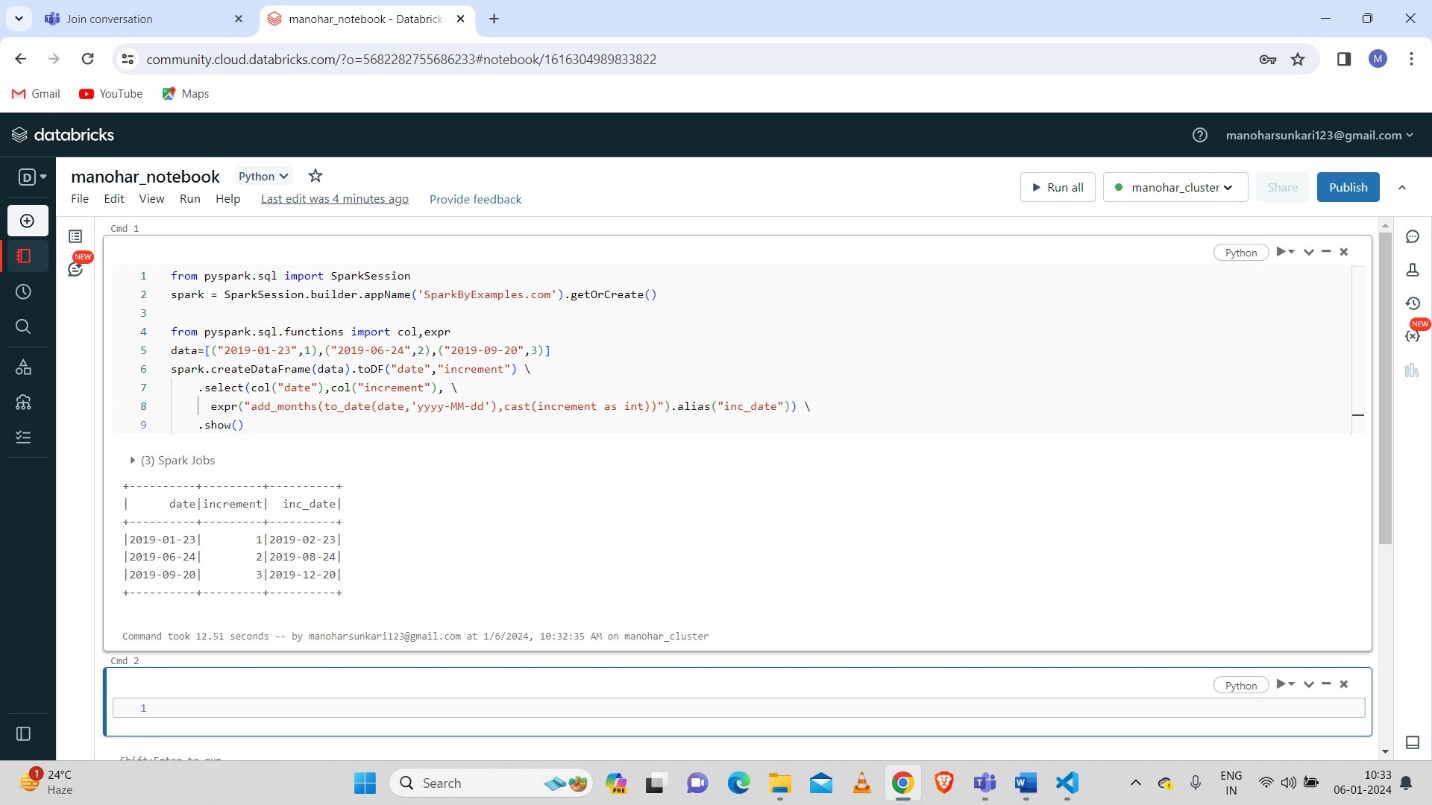




* Executed some commands in databricks notebook:

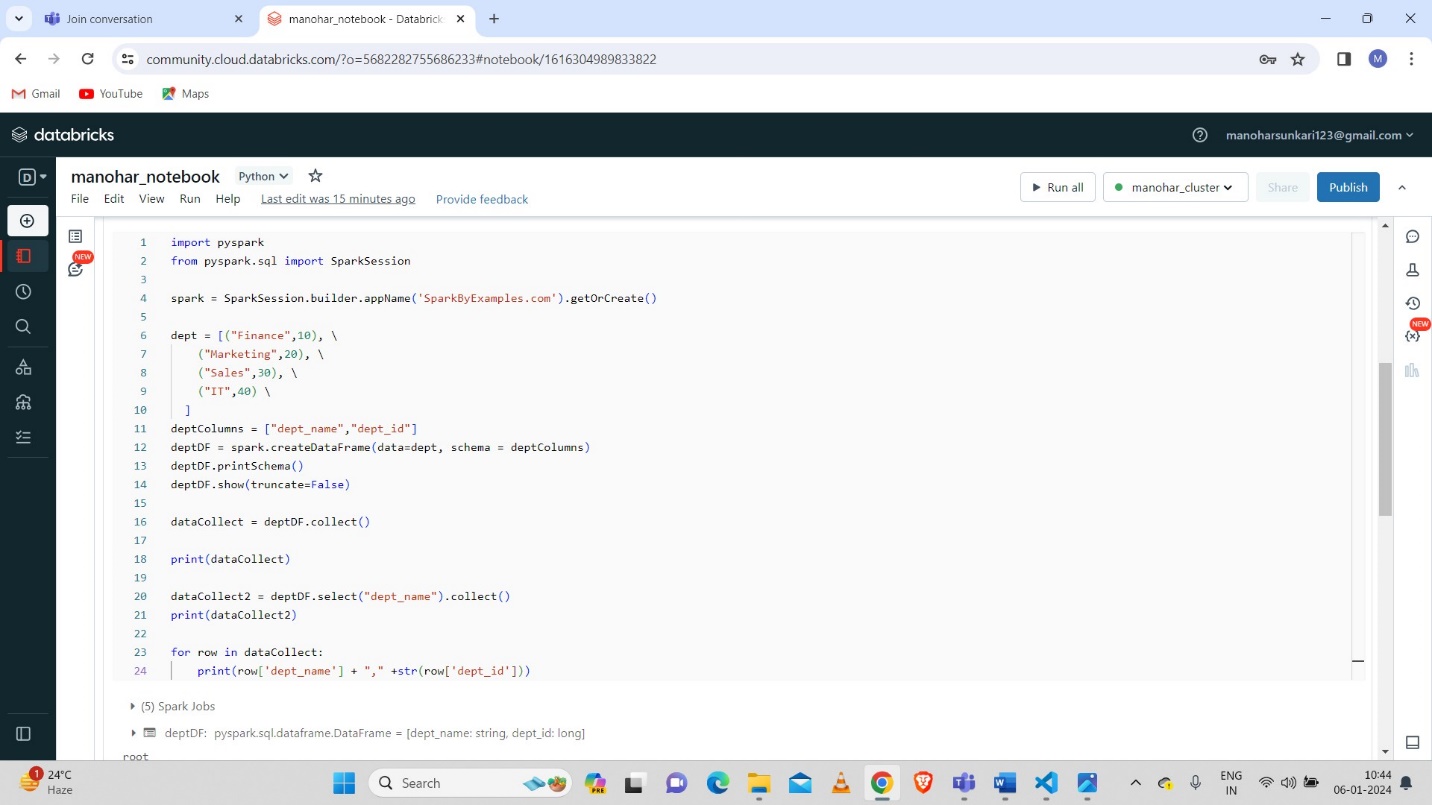
**Example 1:**

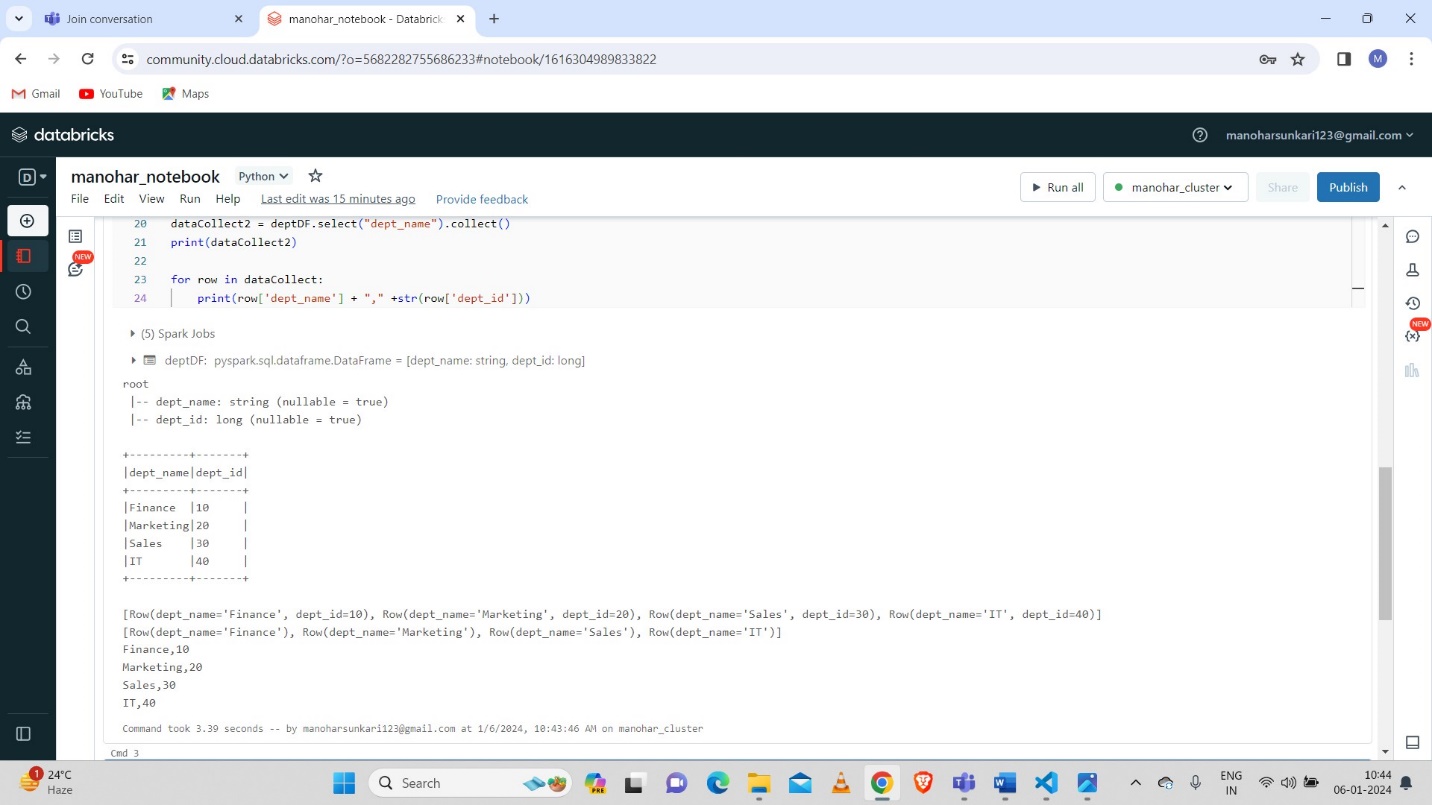
* **Adding month**



**Example 2:**

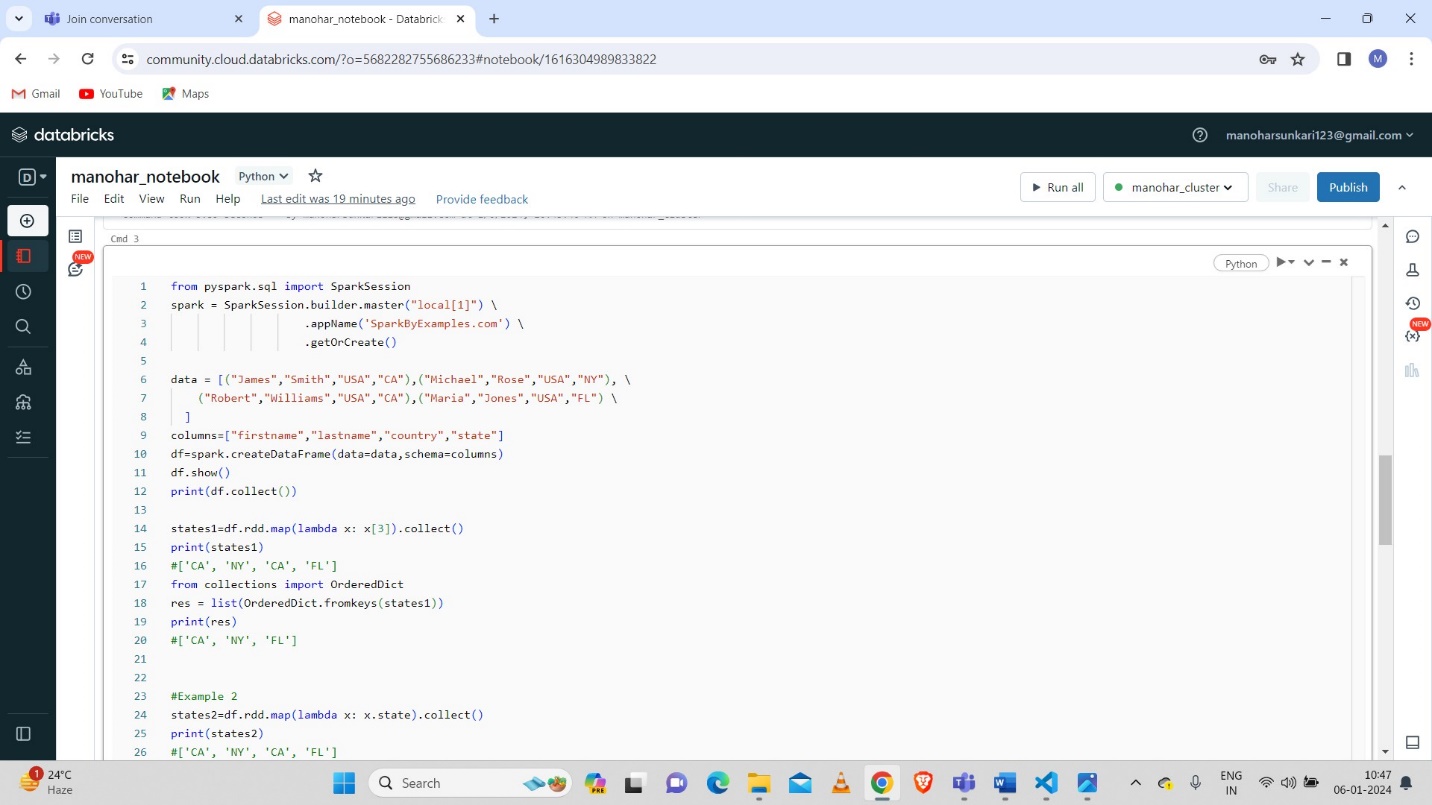
* **Data Collect**

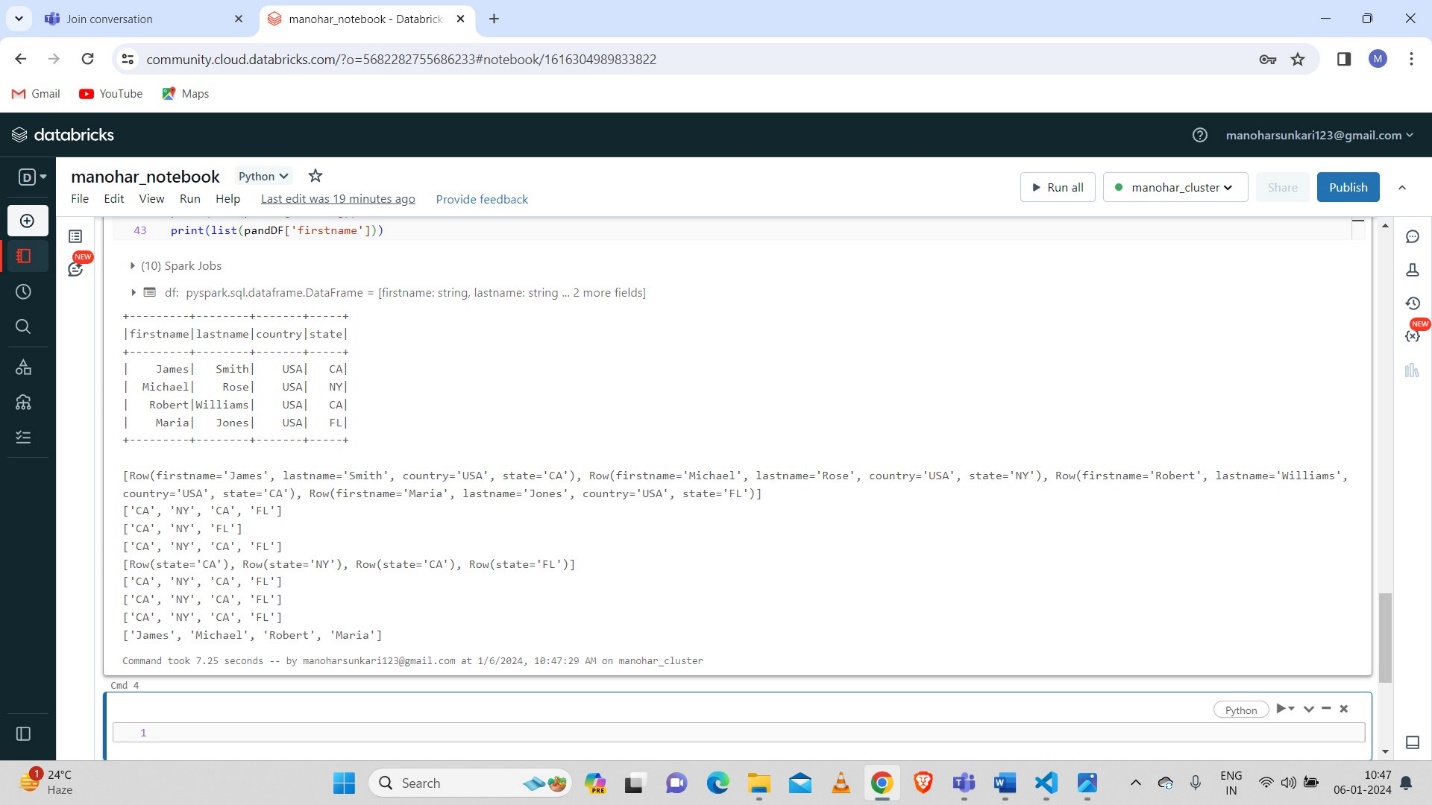




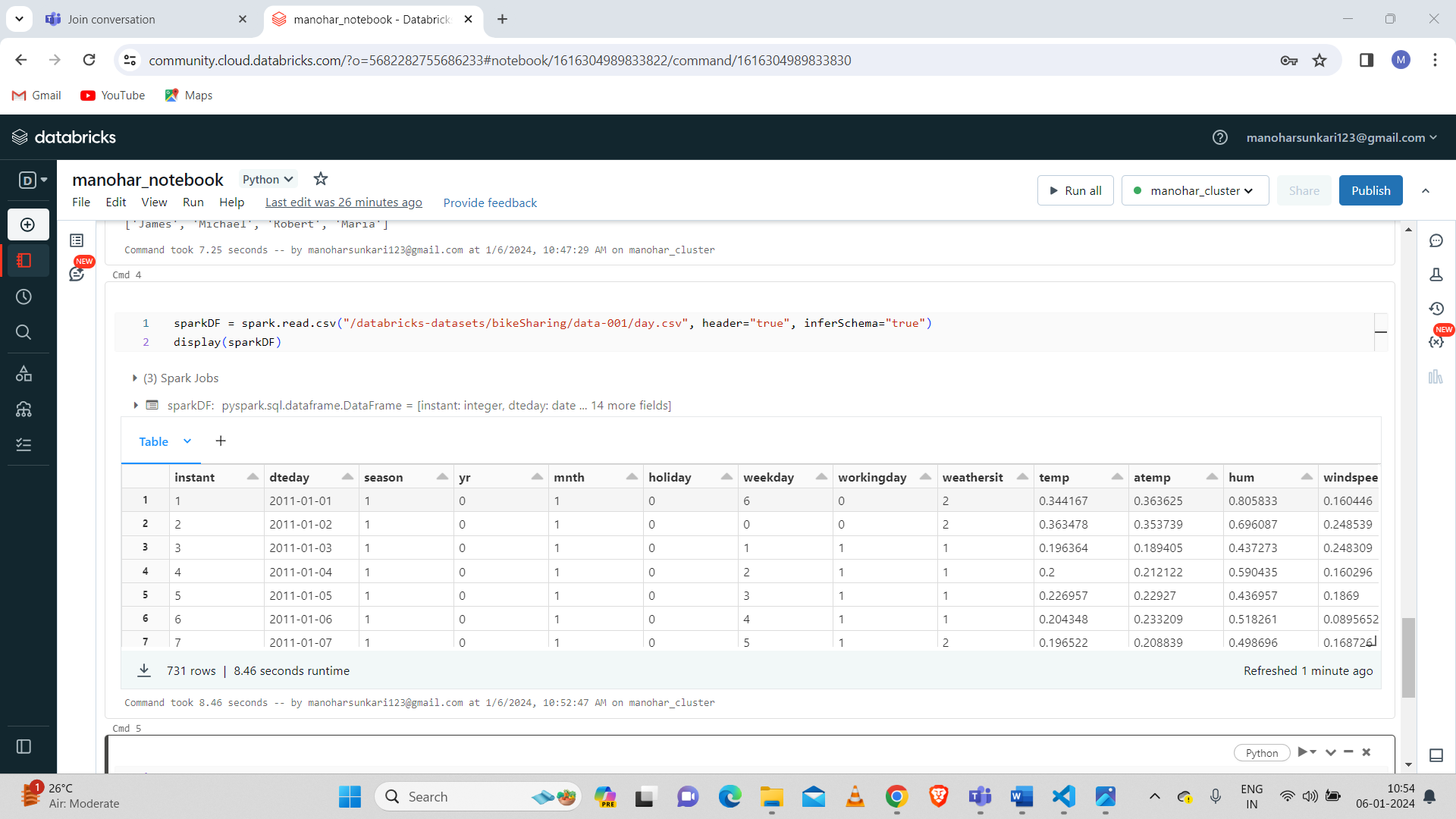
**Example 3:**

* **Converting column into python list:**



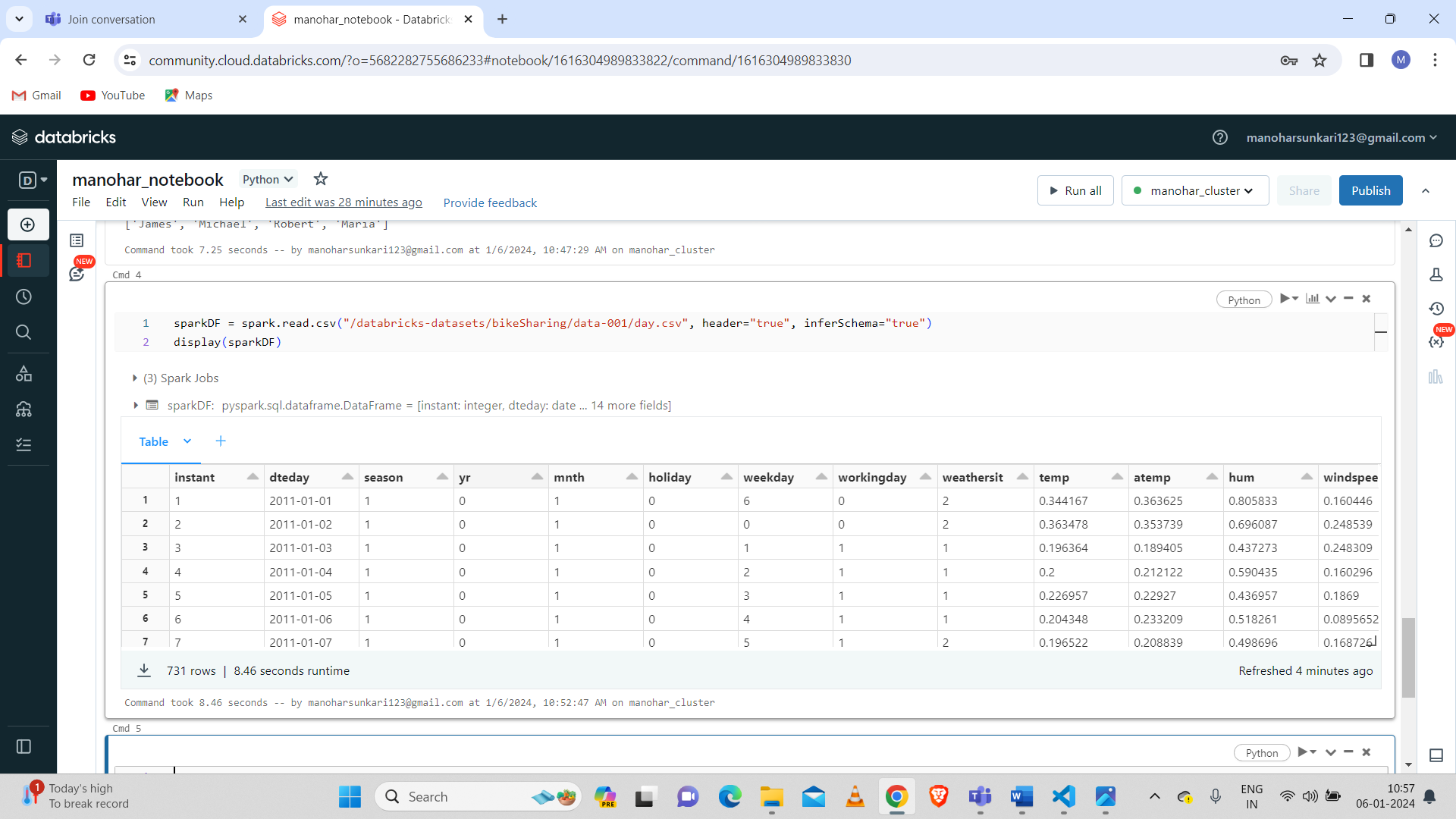


* **Creation of Dataframe:**

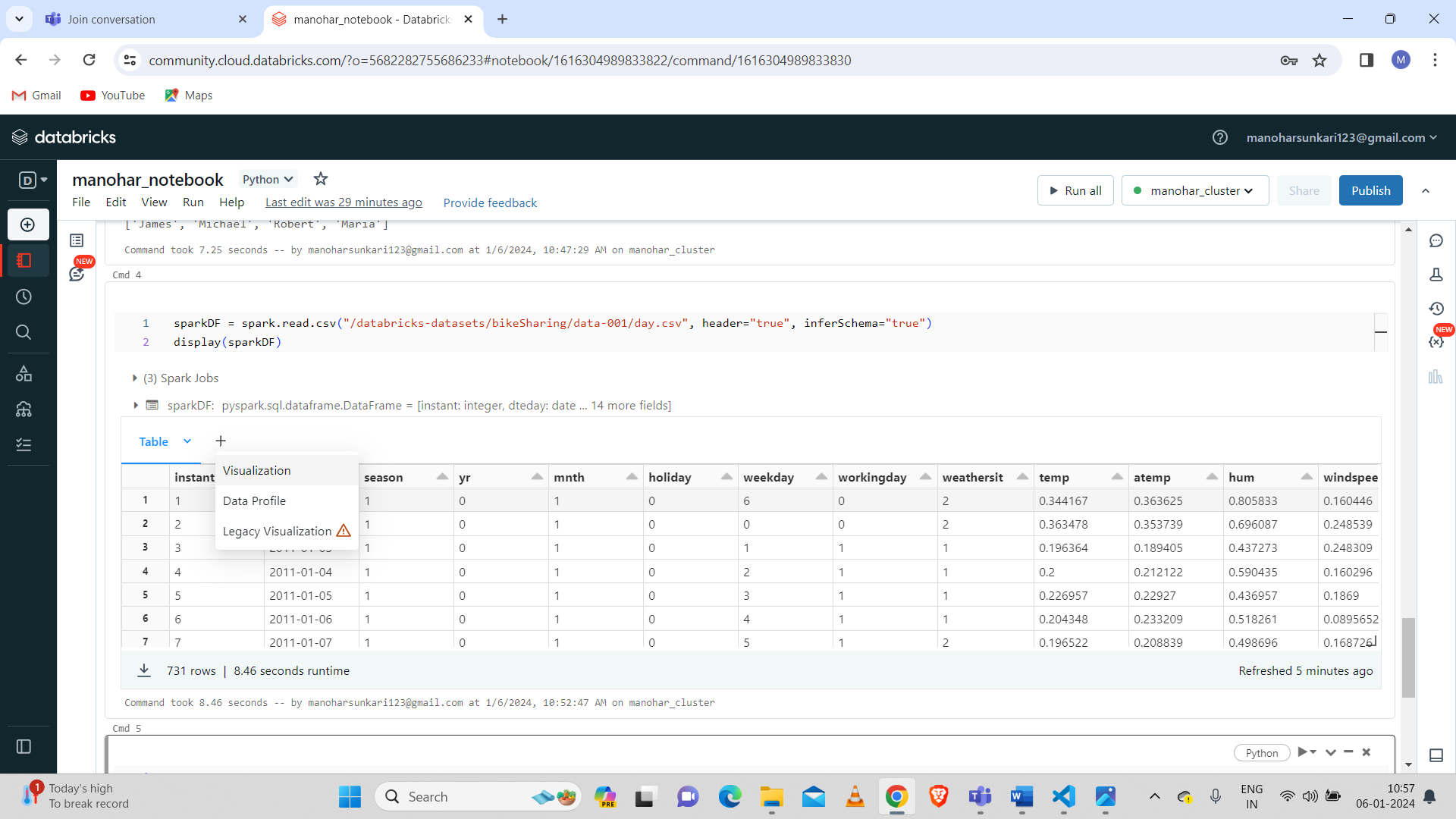


* **Visualization:**

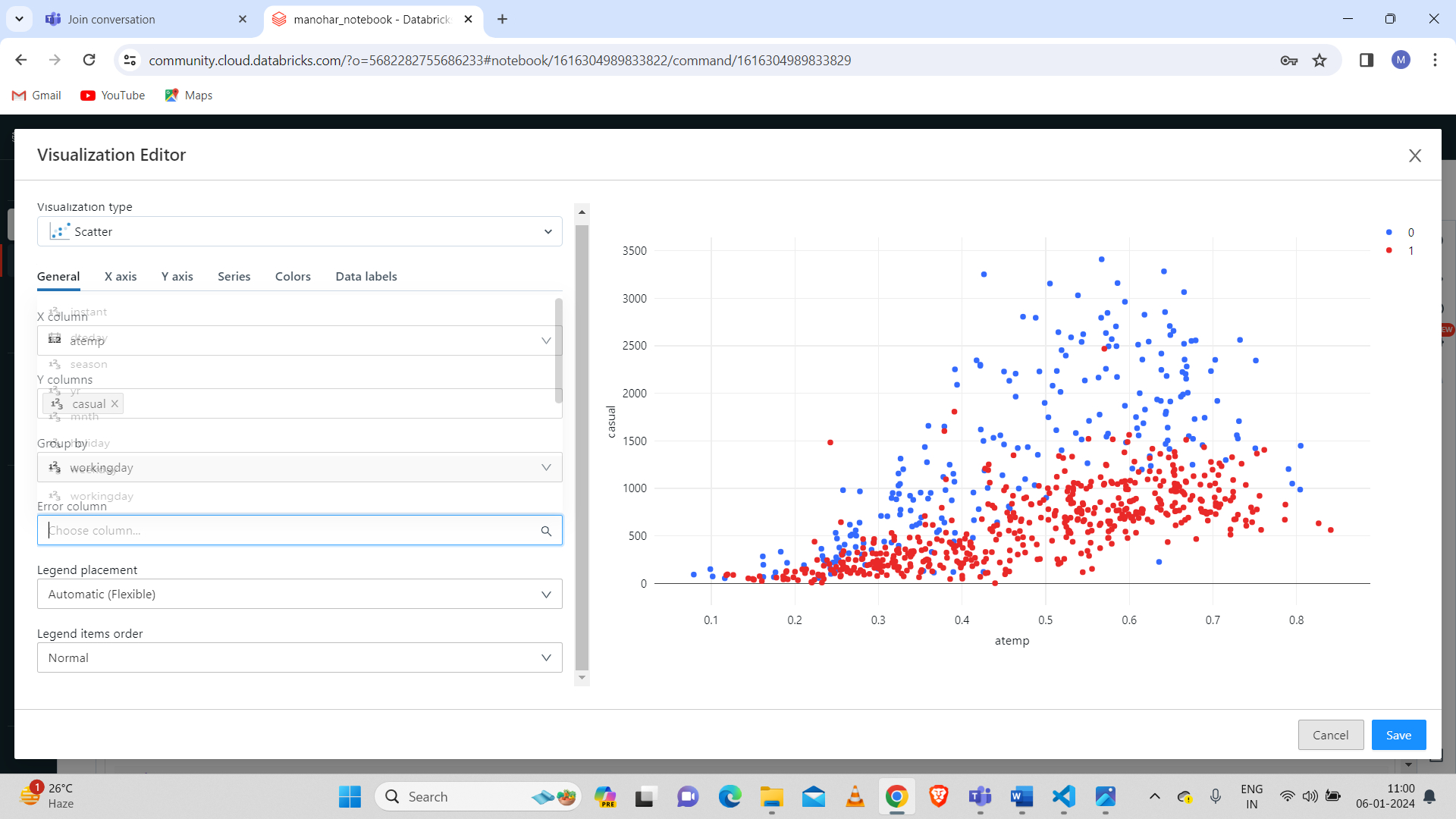
To visualize the dataframe we need to click on “+” icon which is above the table output.



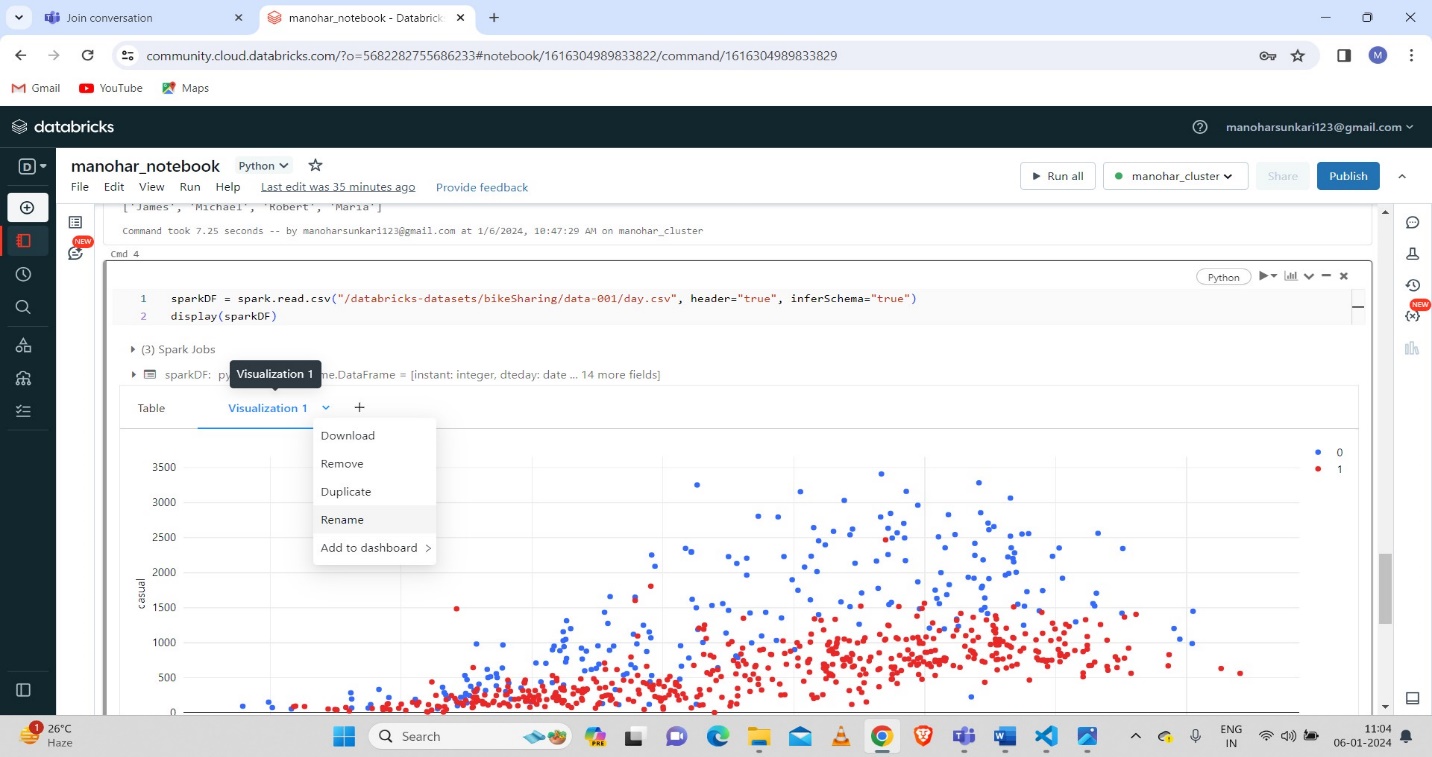
And click on Visualization to visualize the dataframe.



* **Visualization of Dataframe**:



To edit the visualization or data profile we need to click the down arrow beside the visualization.

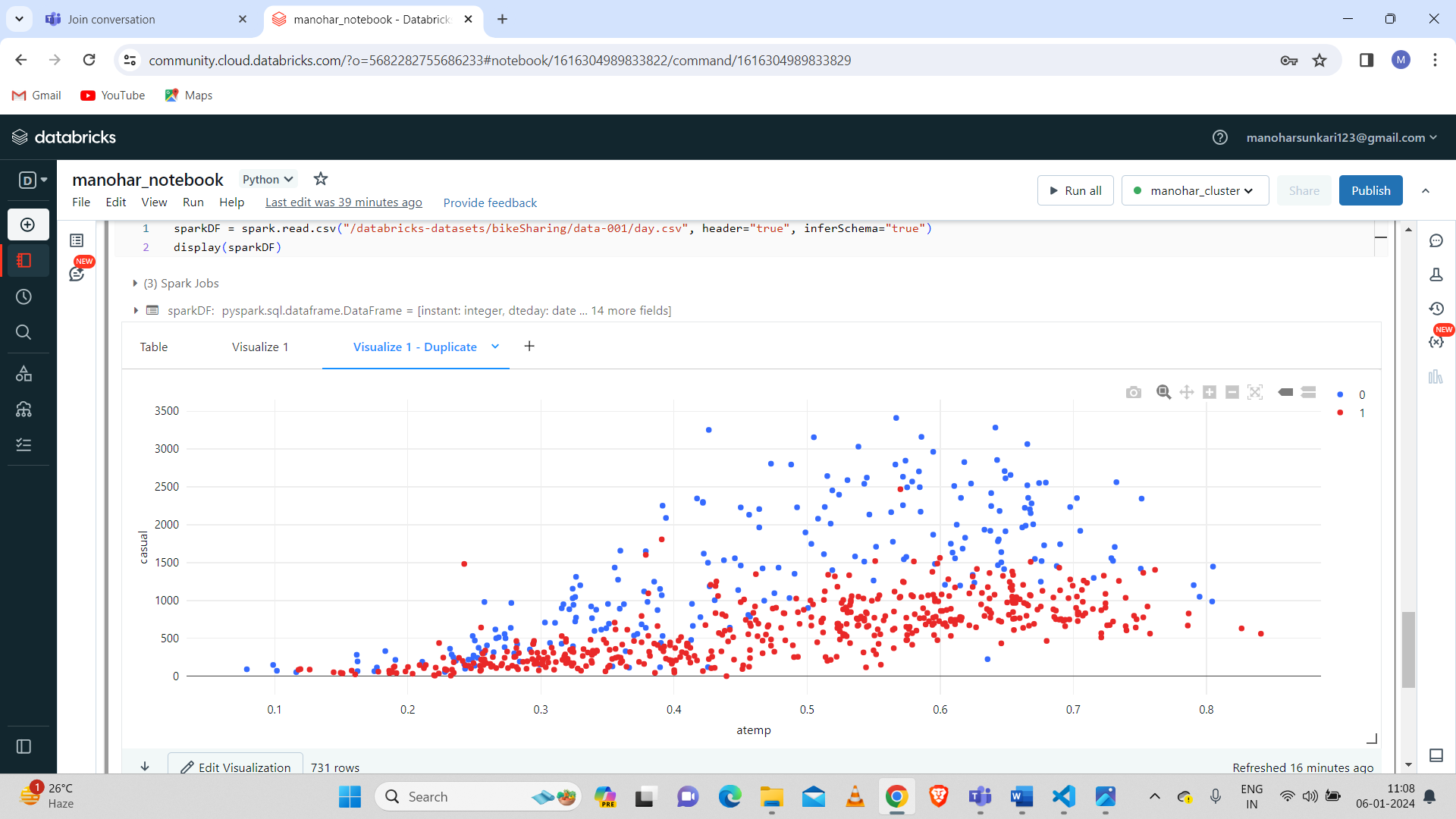


* **Renaming a Visualization:**

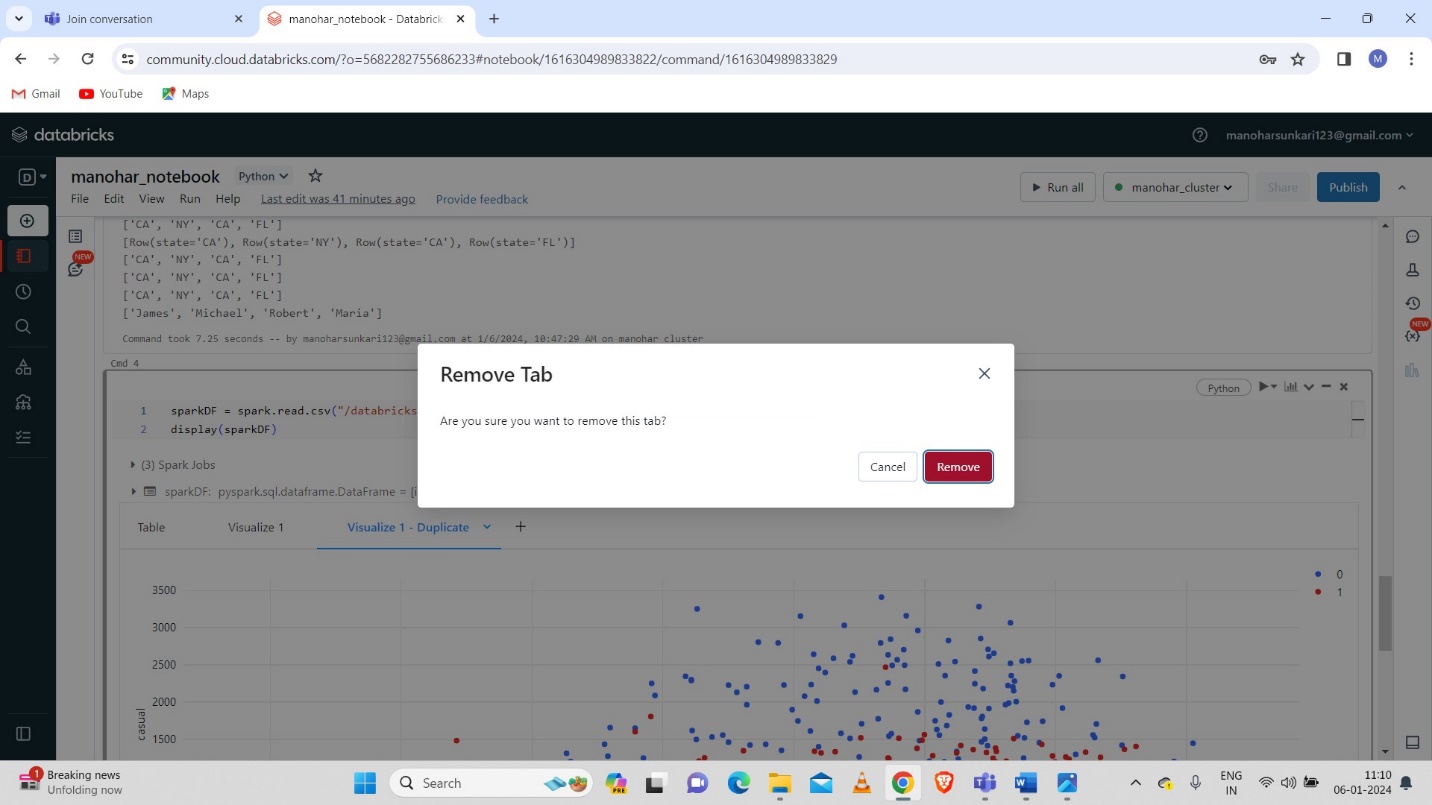
Here renamed the “visualization 1” to “visualize 1”.



* **Duplication of Visualization:**



* **Removal of Visualization Duplicate:**





2.Explain the copy activity in Azure data factory.

* **Copy Activity of Azure Data factory:**

**Step-1: Create a Data factory.**

* We can use our existing data factory if already created otherwise create a new data factory.
* Create a Data factory by using Azure Portal.

**Step-2: Use the copy data tool to copy the data.**

* By clicking on copy data tool to proceed for copying the data.

**Step-3: Start the Copy data tool.**

* On Home page of Azure data factory we have to select the ingestion to start the copy data tool.
* On the properties page of copy data tool we need to choose the built-in copy task which is under task type and click on “next”.

**Step-4: Complete source configuration.**

* Click + Create new connection to establish a connection.
* Select the linked service type that you want to create for the source connection. In this tutorial, we use Azure Blob Storage. Select it from the gallery, and then select Continue.
* On the New connection (Azure Blob Storage) page, specify a name for your connection. Select your Azure subscription from the Azure subscription list and your storage account from the Storage account name list, test connection, and then select Create.
* Select the newly created connection in the Connection block.
* In the File or folder section, select Browse to navigate to the adftutorial/input folder, select the emp.txt file, and then click OK.
* Click on the Binary copy checkbox to copy file as-is, and then select Next.

**Step-5: Complete destination configuration.**

* Click on the Azure Blob Storage connection that you created in the Connection block.
* In Folder path section, enter adftutorial/output for the folder path.
* And click on next.

**Step-6: Review all settings and deployment.**

* On the Settings page, specify a name for the pipeline and its description, then select Next to use other default configurations.
* On the Summary page, review all settings, and select Next.
* On the Deployment complete page, select Monitor to monitor the pipeline that you created.

**Step-7: Monitor the running results.**