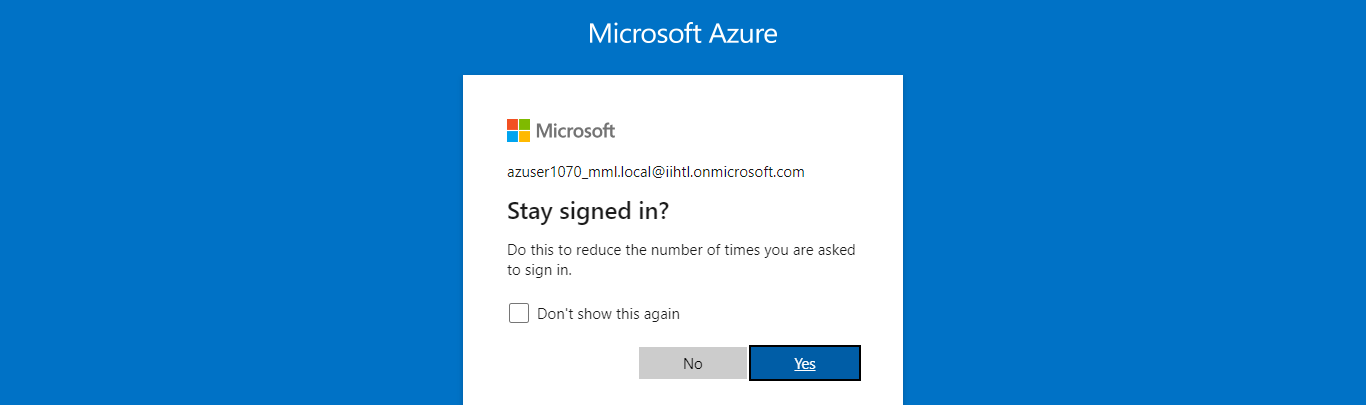
**NAME :- ANIKET SANJAYKUMAR BIYANI**

**DATA ENGINEERING BATCH -1**

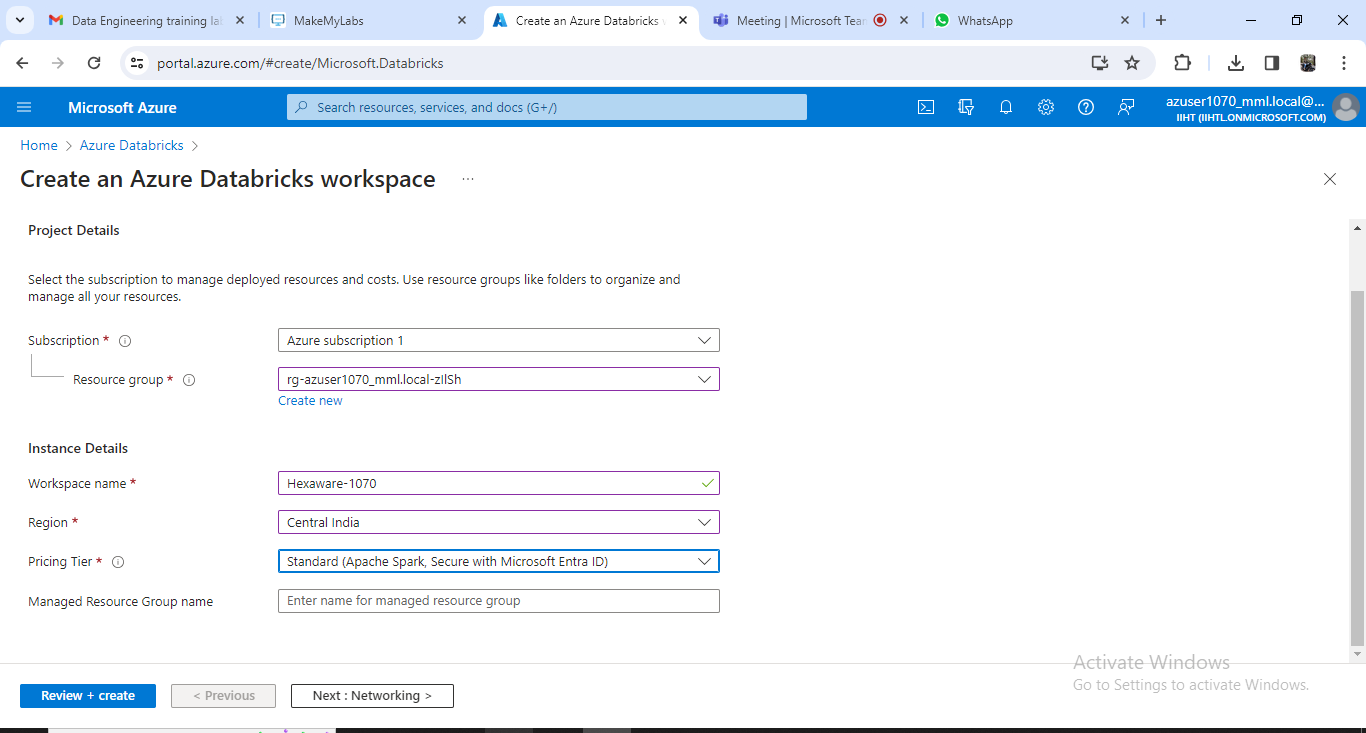
**AZURE DATABRICKS CODING ASSESMENT**

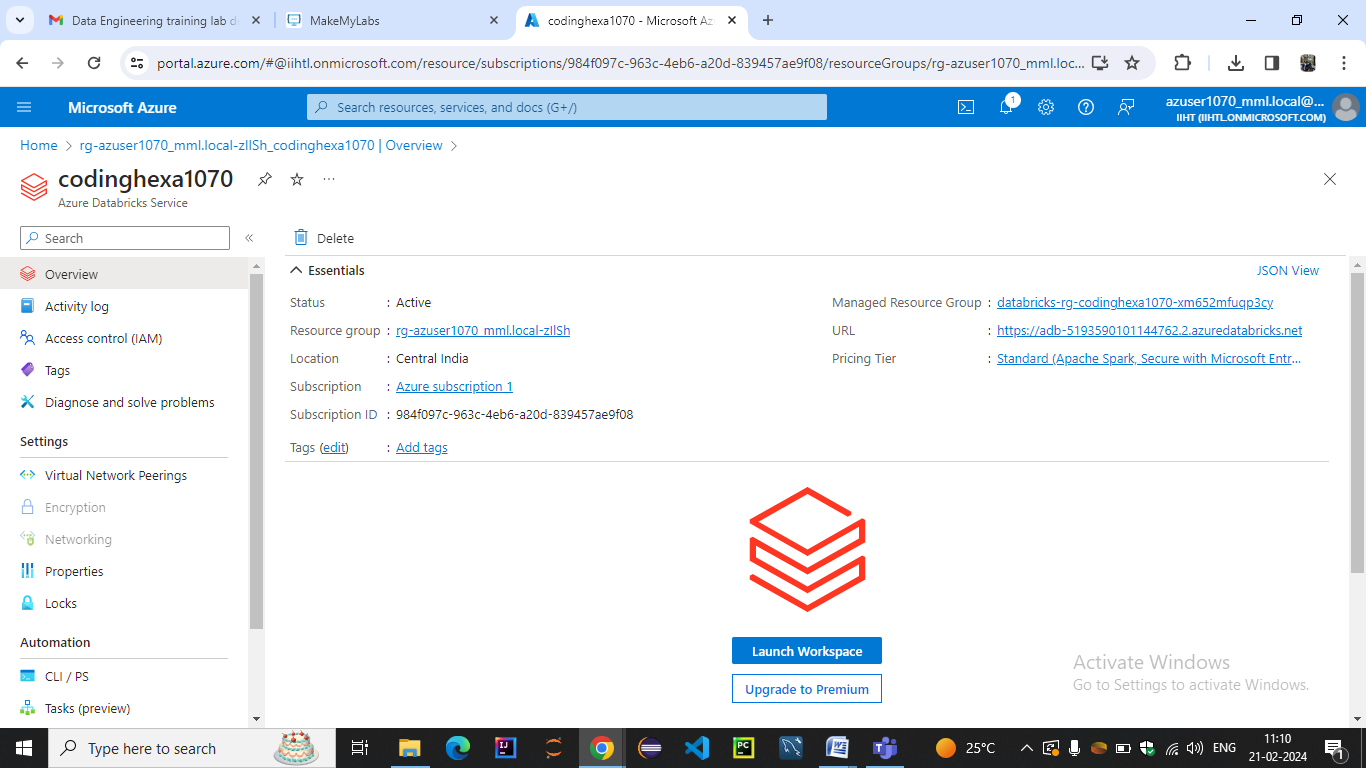
**Exploratory data analysis (EDA) in Databricks &Visualizing data in Databricks**

* Exploratory Data Analysis (EDA) in Databricks is a crucial step in understanding the structure, relationships, and patterns within your dataset.
* It is used to discover trends, patterns, or to check assumptions with the help of statistical summary and graphical representations.
* Exploratory Data Analysis is a fundamental and indispensable stage in the data analysis process, serving as a bridge between raw data and actionable insights
* EDA plays a pivotal role in guiding subsequent modeling, hypothesis testing, and decision-making processes.
* Azure Databricks provides a powerful platform for performing EDA and visualizing data efficiently.
* It provides a collaborative and interactive workspace for data transformation, exploration, and innovation, and supports data science, data engineering, and business use cases
* **Step 1: Setting up Azure Databricks**
* ***Sign in to the Azure portal (***[***https://portal.azure.com/***](https://portal.azure.com/)***).***



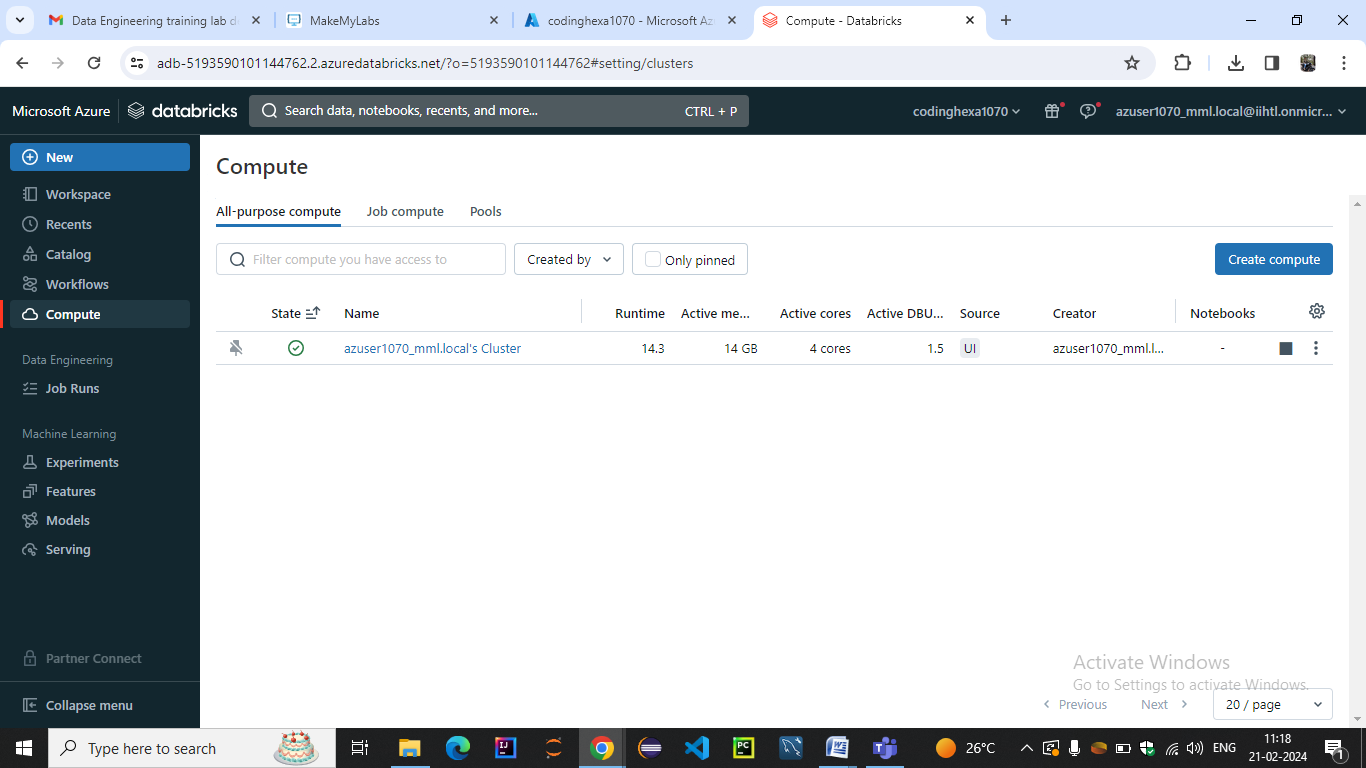
* ***Create an Azure Databricks Workspace***: Set up an Azure Databricks workspace in your Azure portal



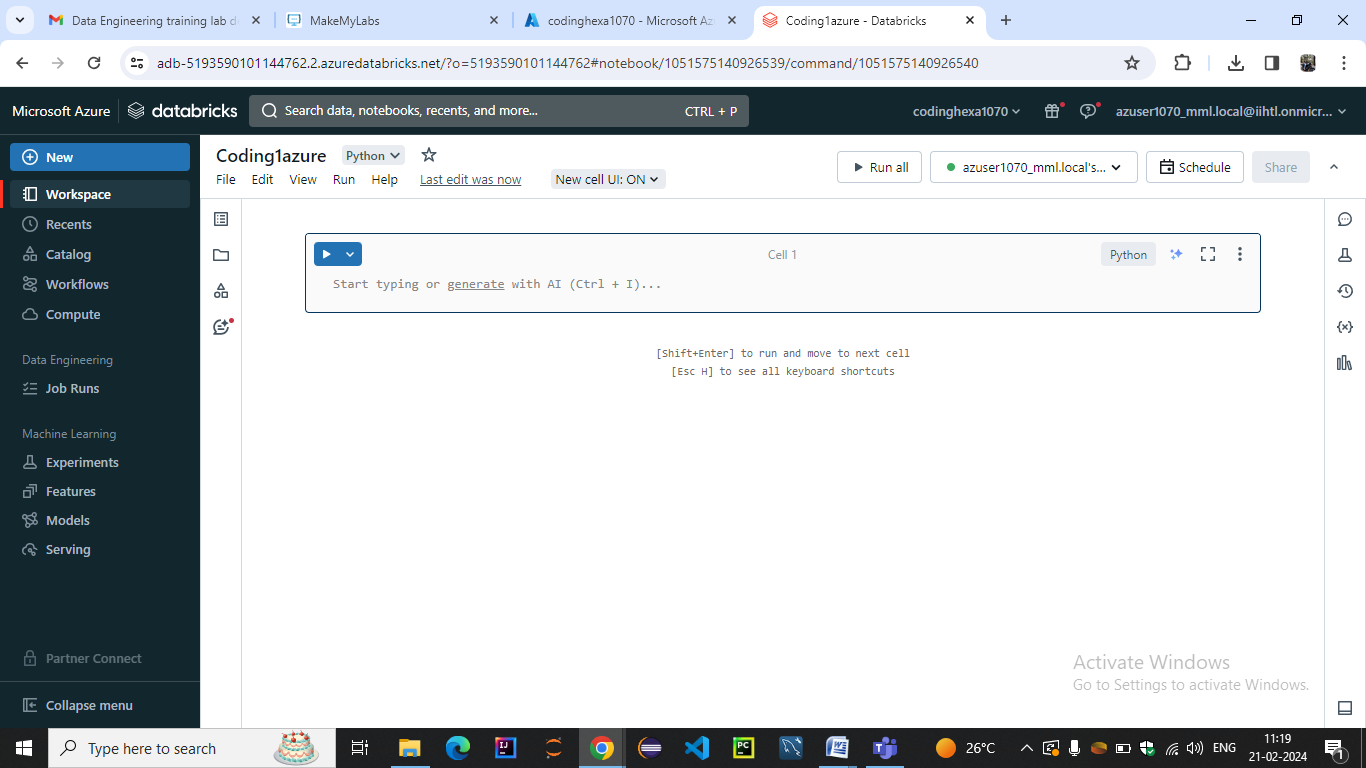


***LAUNCH WORKSPACE AND CREATE CLUSTER***

* ***Create a Cluster***: Create a Databricks cluster from your Azure Databricks workspace.

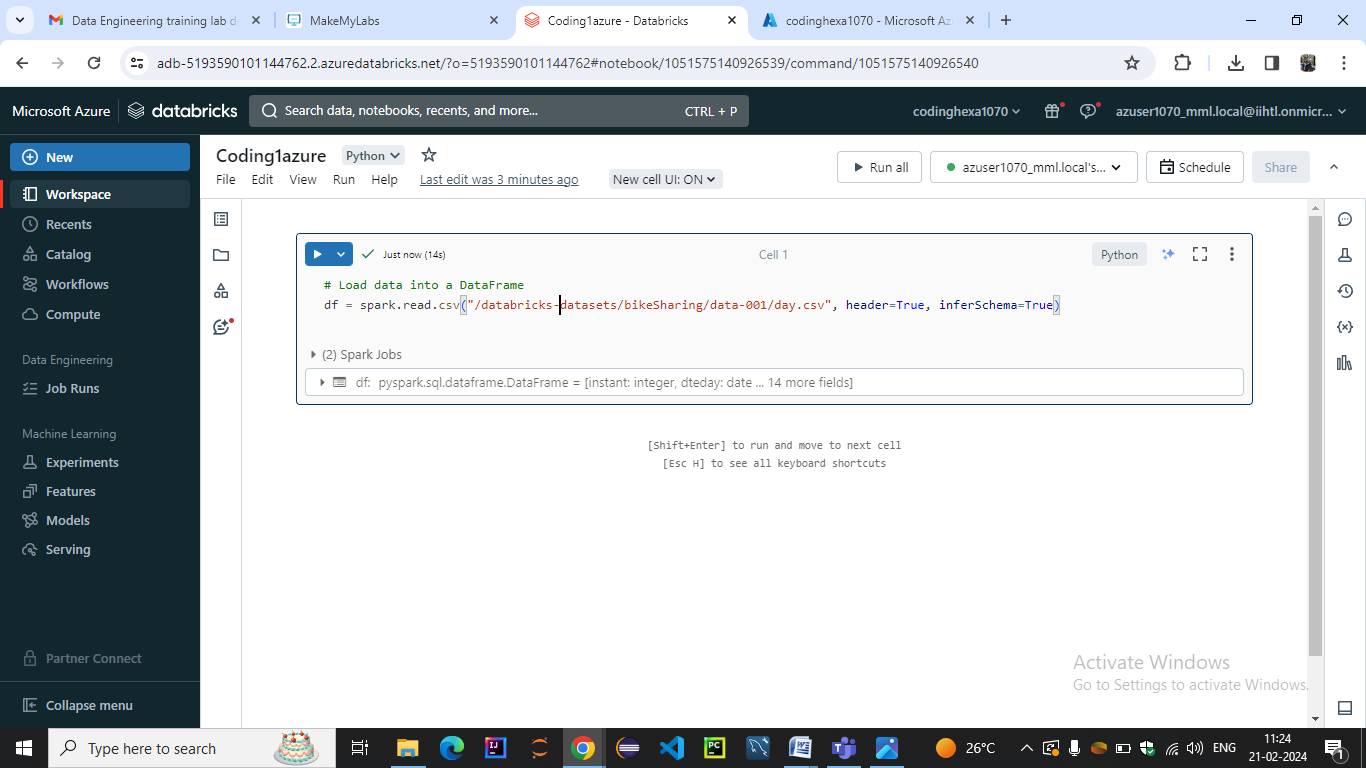


* ***Create a Notebook***: In the Databricks workspace, create a new notebook where you will perform your EDA and visualization tasks.

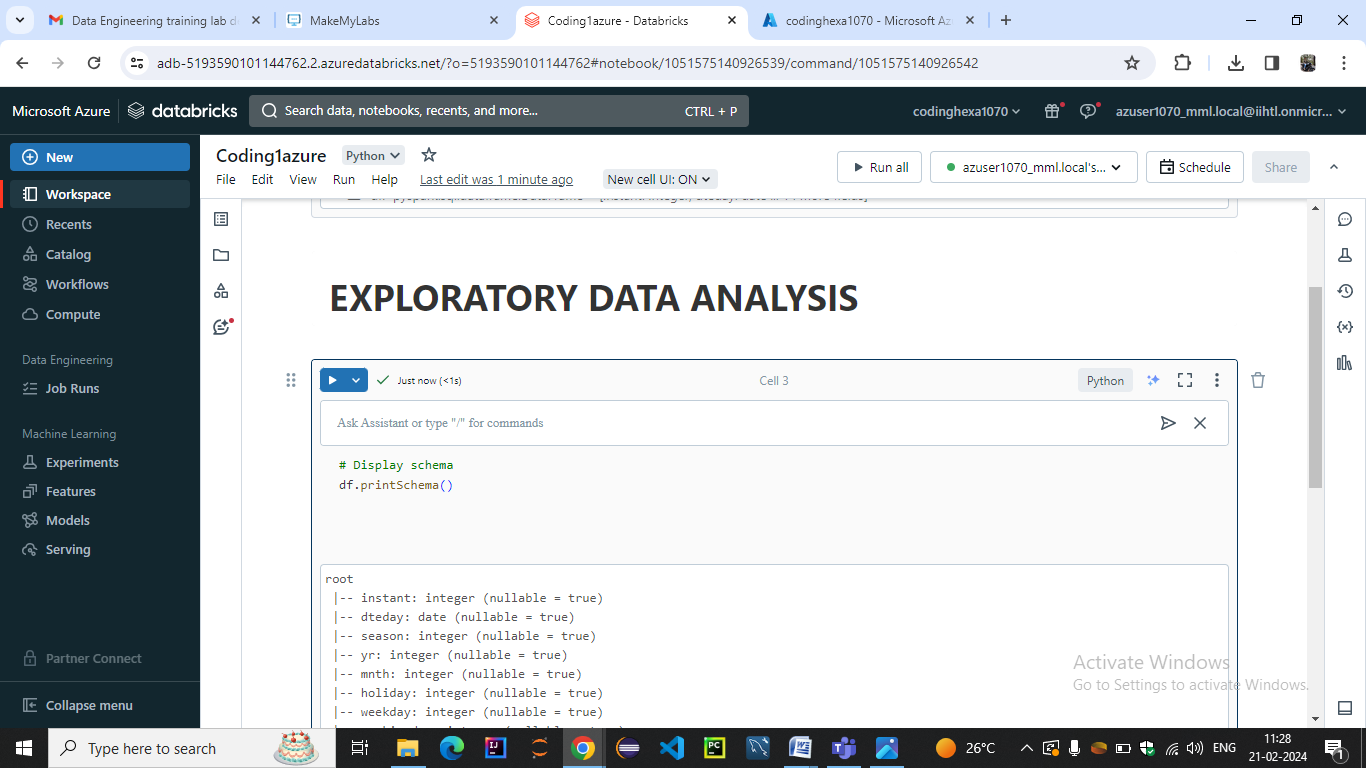


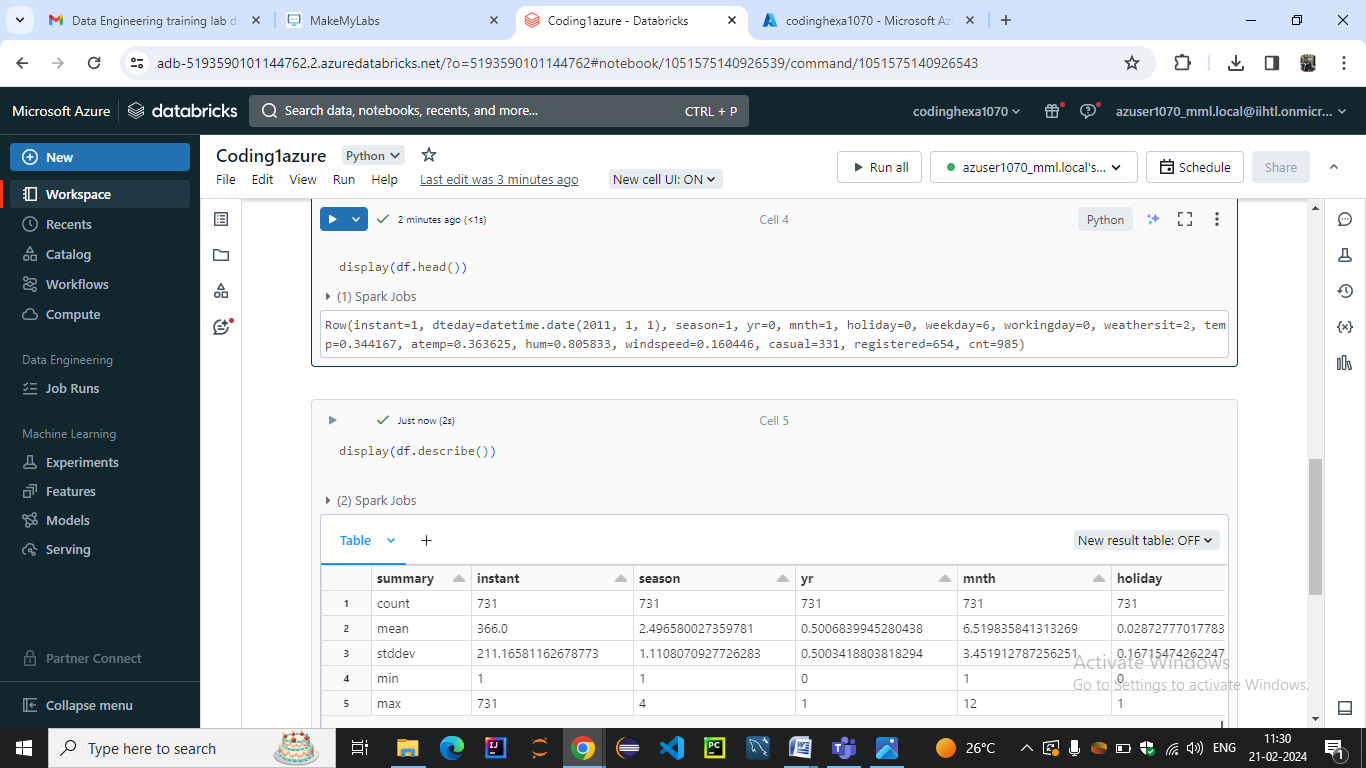
**Step 2: Importing Data**

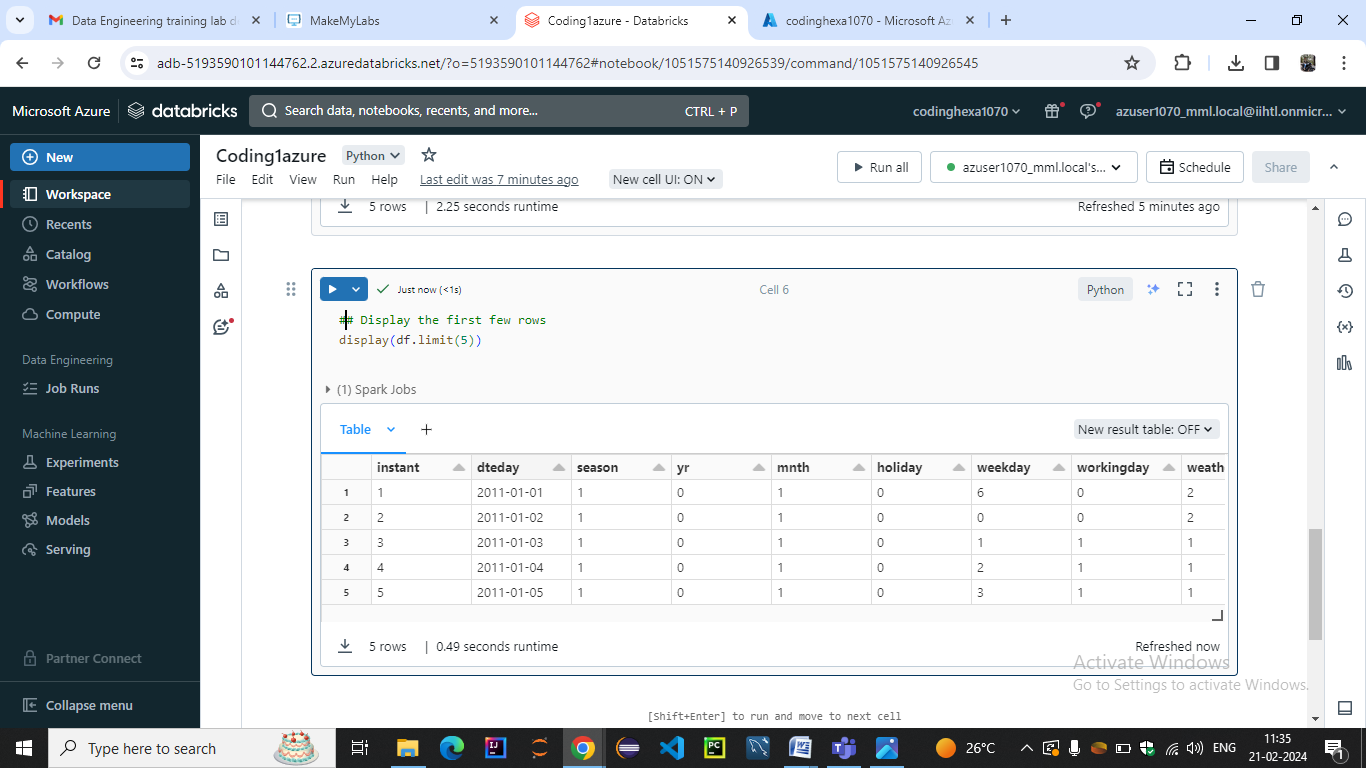
You can use various methods such as uploading files, connecting to external data sources, or using Azure Data Lake Storage.

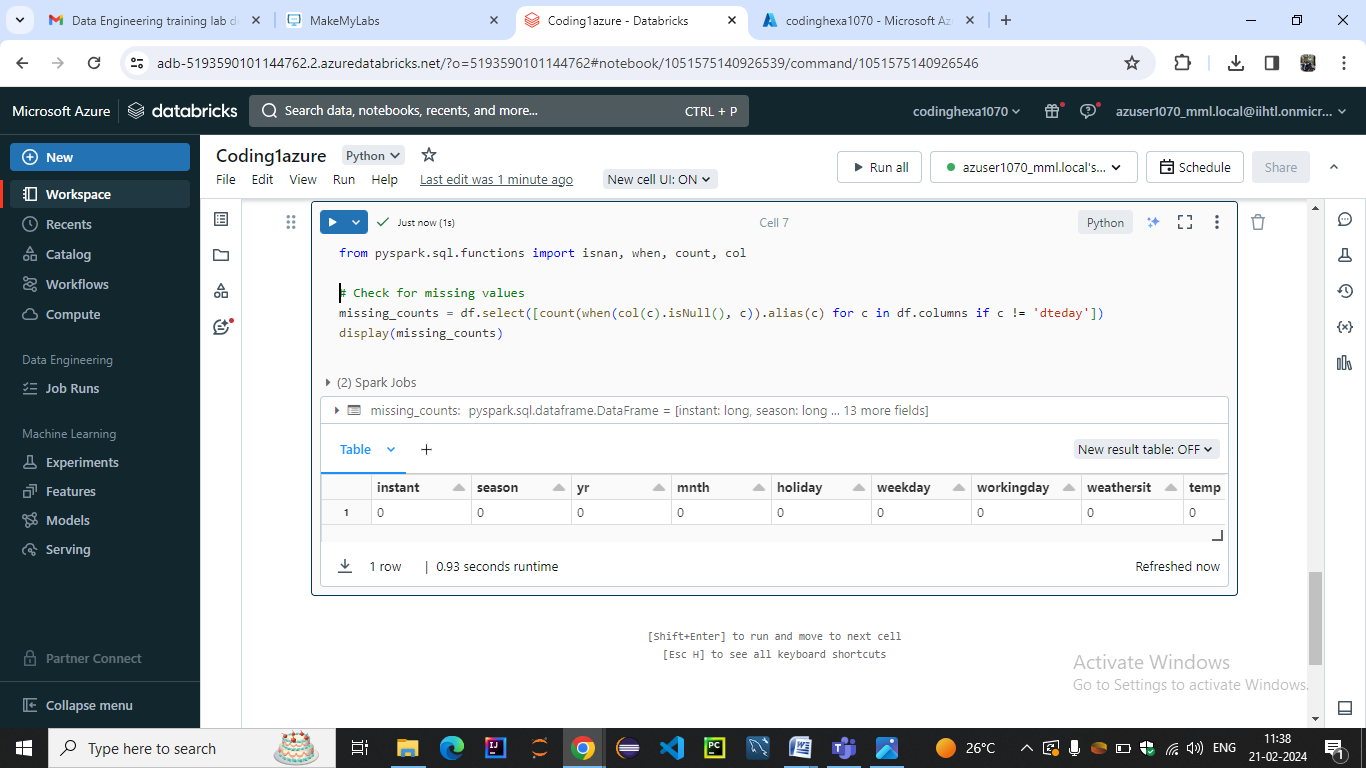


**Step 3: Exploratory Data Analysis**

****







DATA VISUALISATION:-

