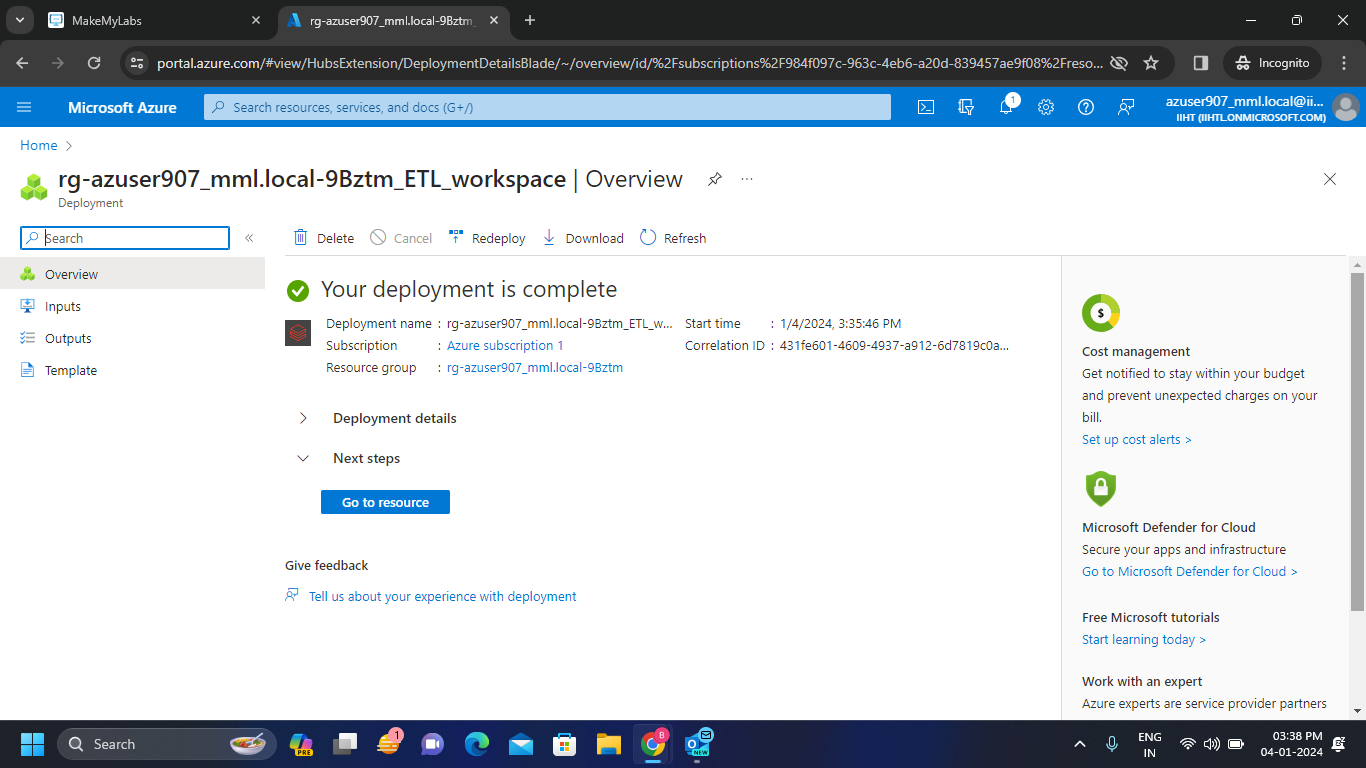
**Alekhya Krishna Balivada Azure Databricks Assignment -6 (04-01-2024)**

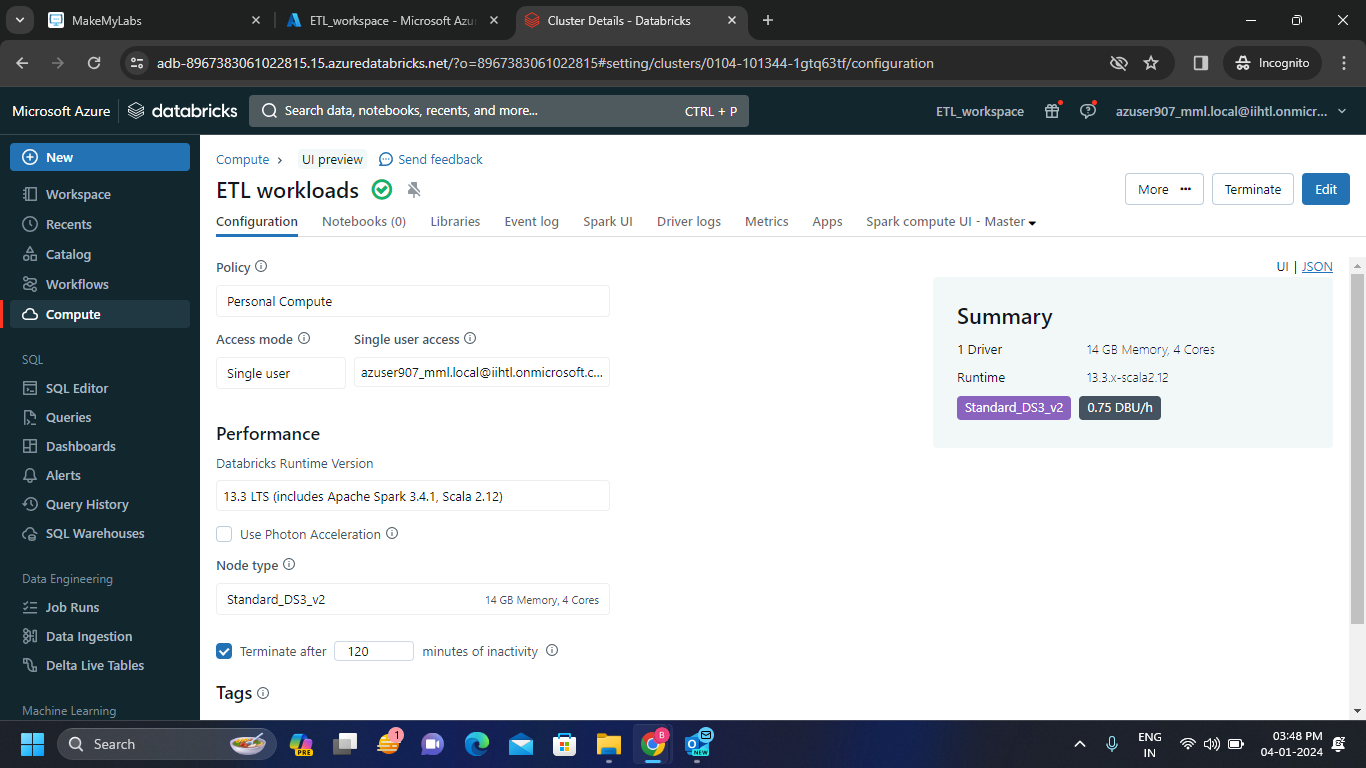
|  |  |
| --- | --- |
| Name | Alekhya Krishna Balivada |
| Date | 04-01-2024 |
| Type | Daily Assignment |
| Topic | Running First ETL workload in Azure Data bricks. |

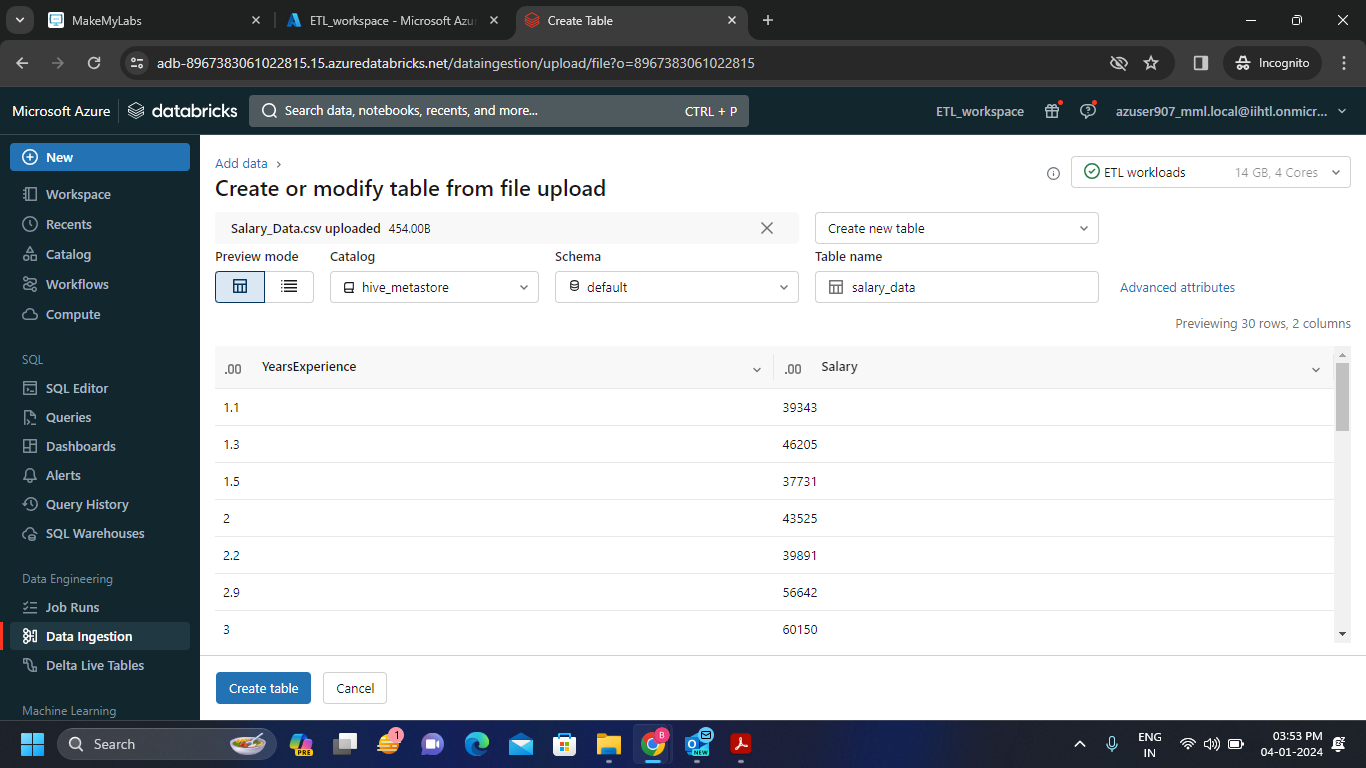
**ETL:**

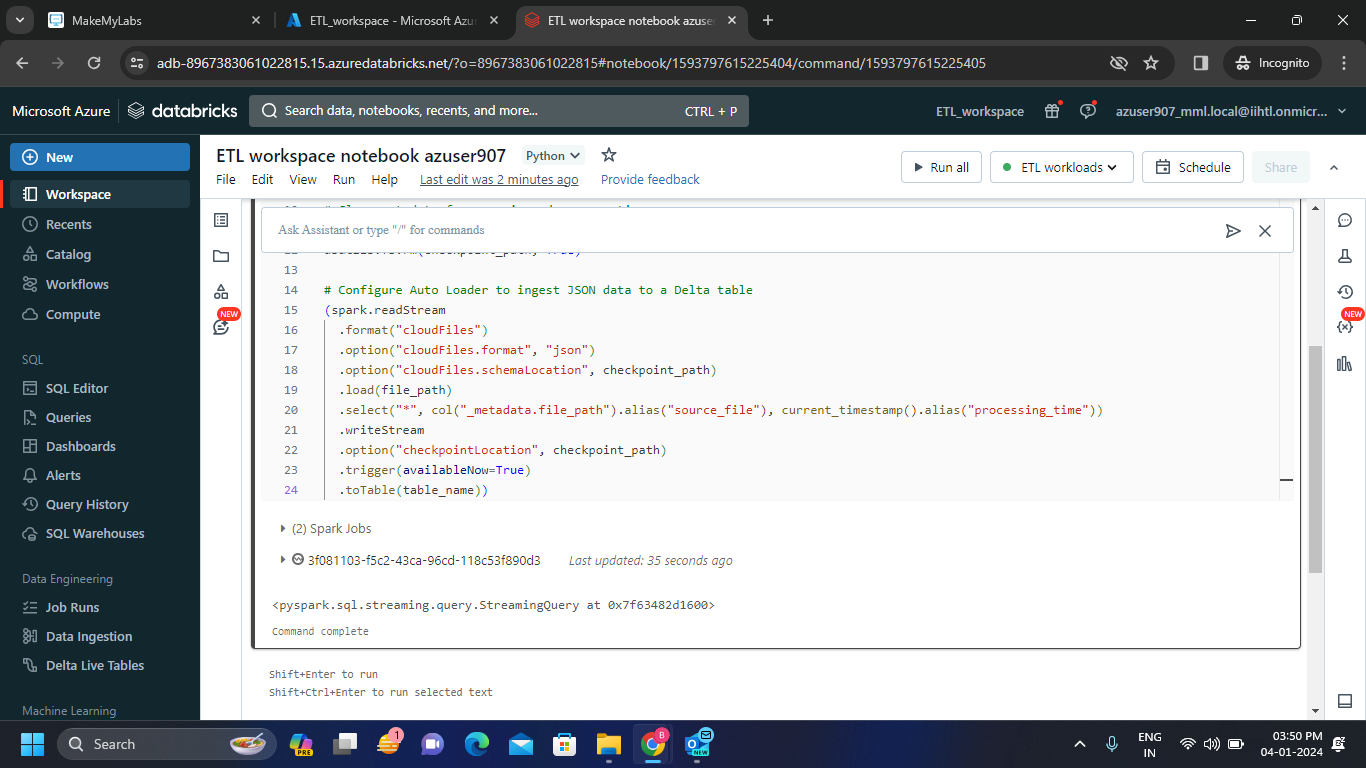
ETL stands for Extract, Transform and Load. It’s a process used in data integration and data warehousing to collect data from various sources, transform it into a format that’s suitable for analysis, and load it into a target database or data warehouse for storage and retrieval.

**ETL workload in Azure Databricks:**

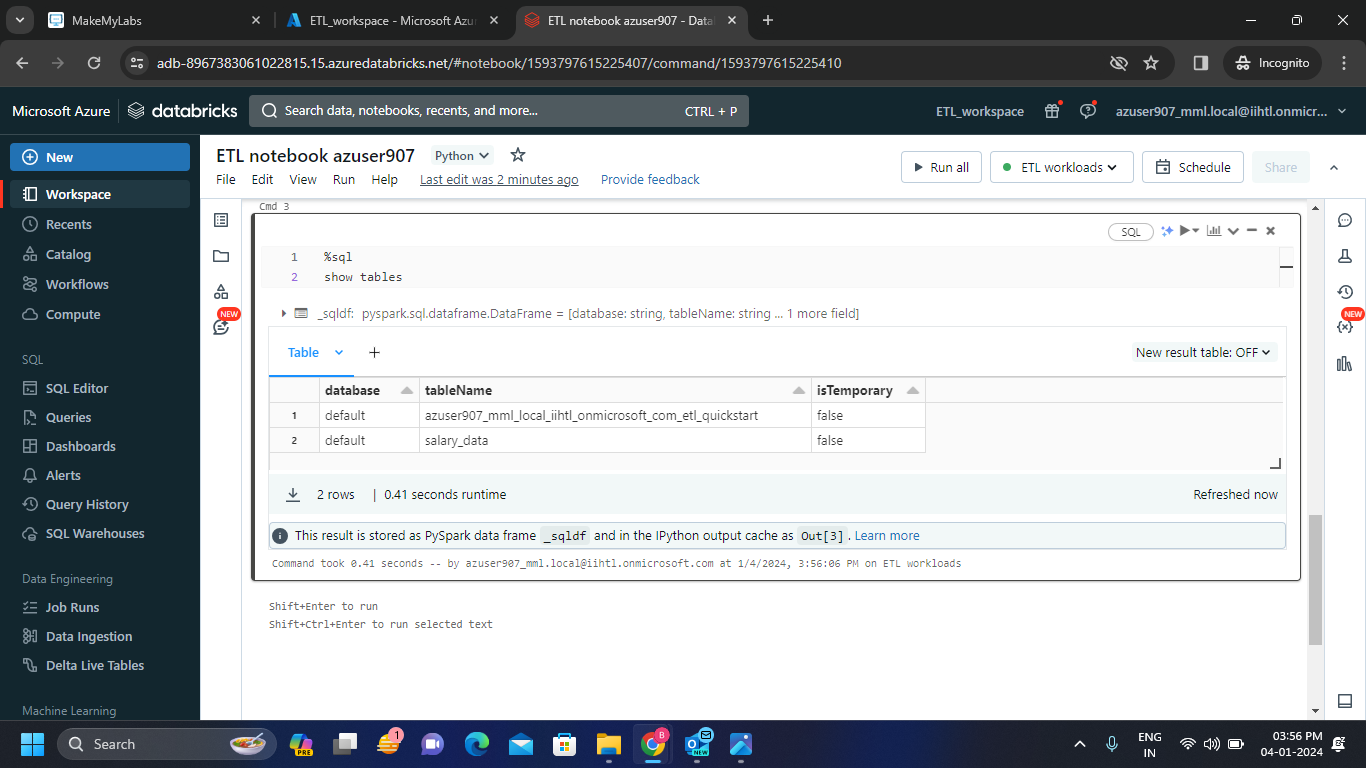
**Step-1:** Create a workspace and create a cluster in the workspace.

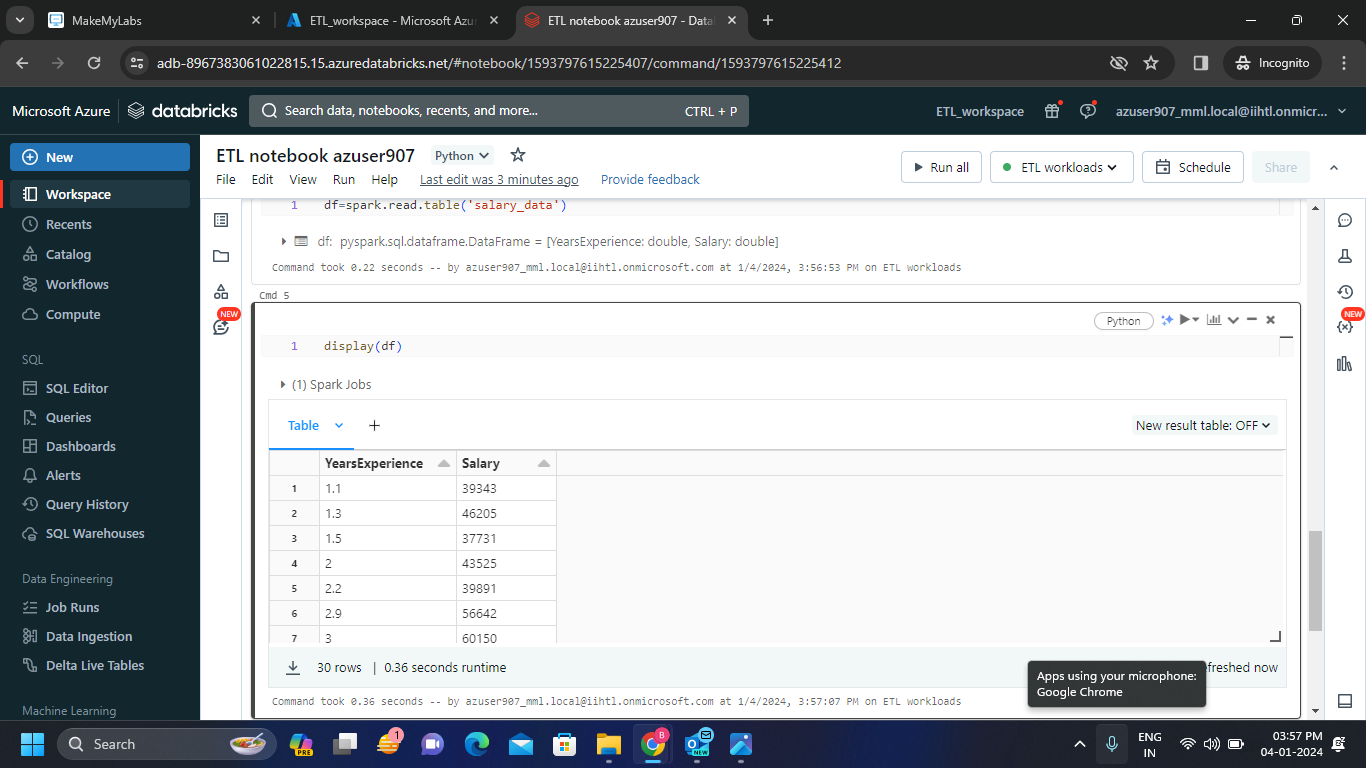


**Step-2:** Create a table by uploading a csv file.

**Step-3:** Create a notebook and connect the notebook with the cluster we created before.

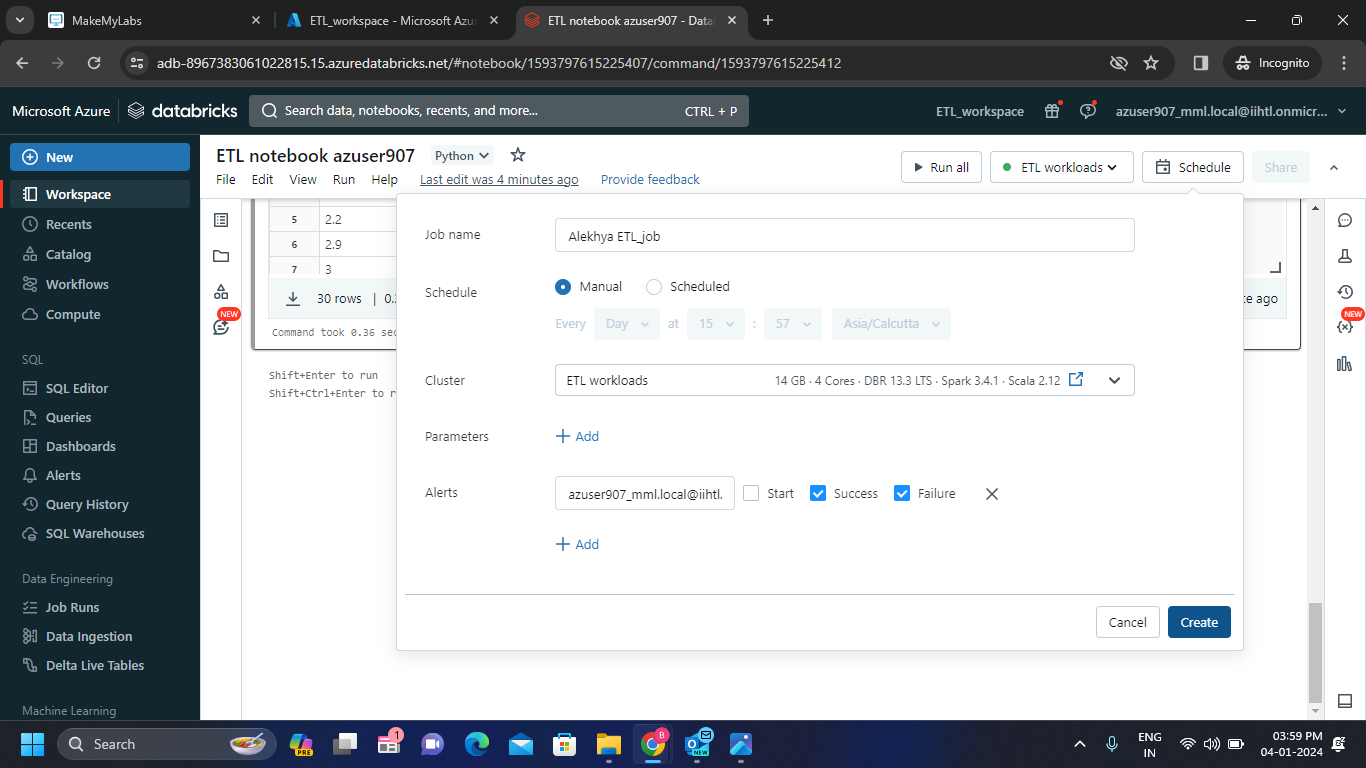
* Run the code as above.

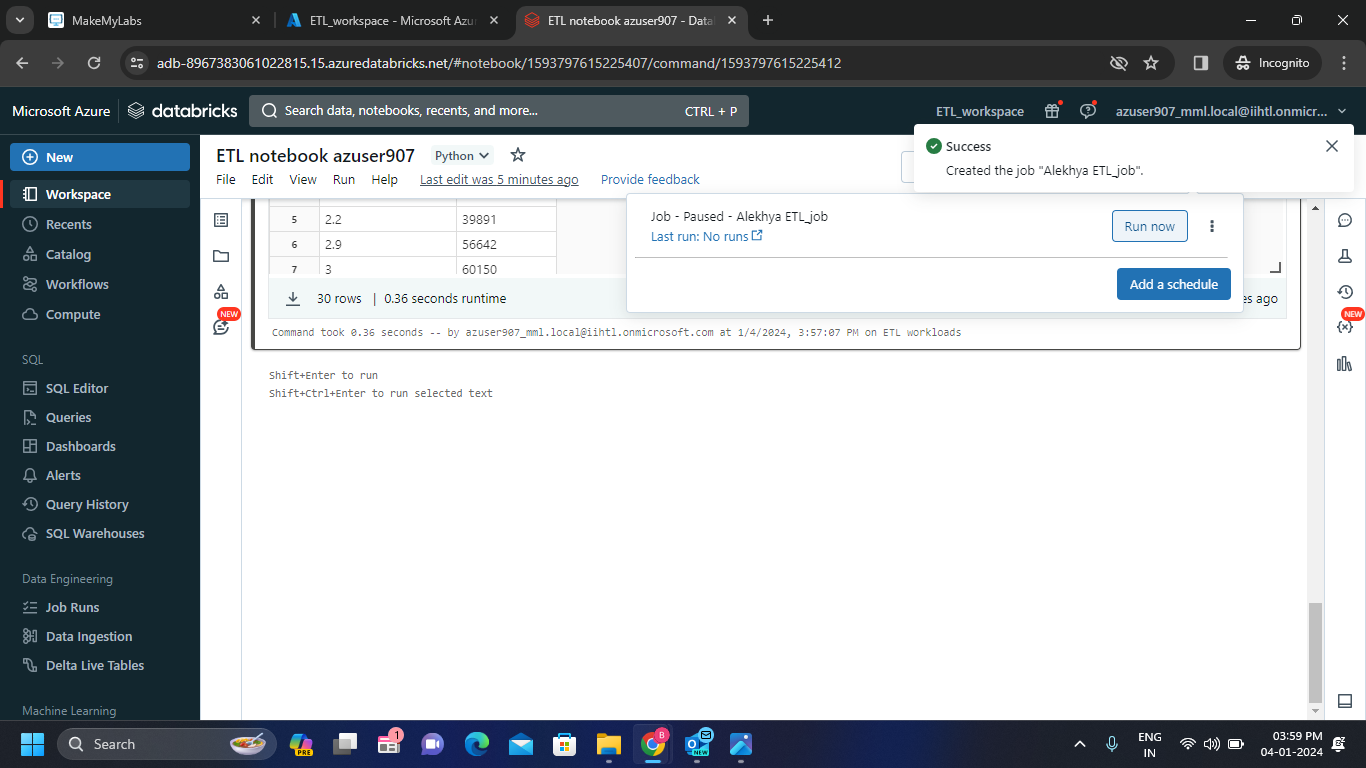


****

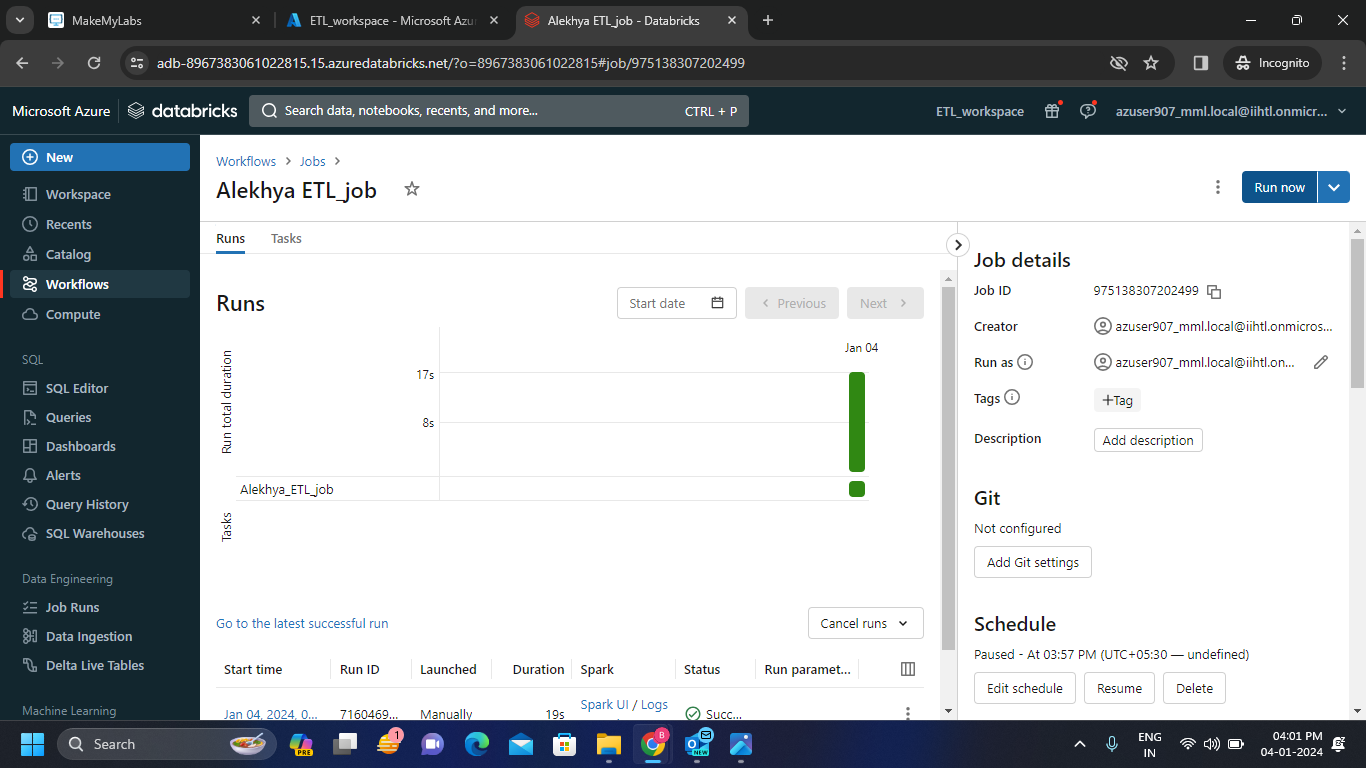
**Scheduling a job:**

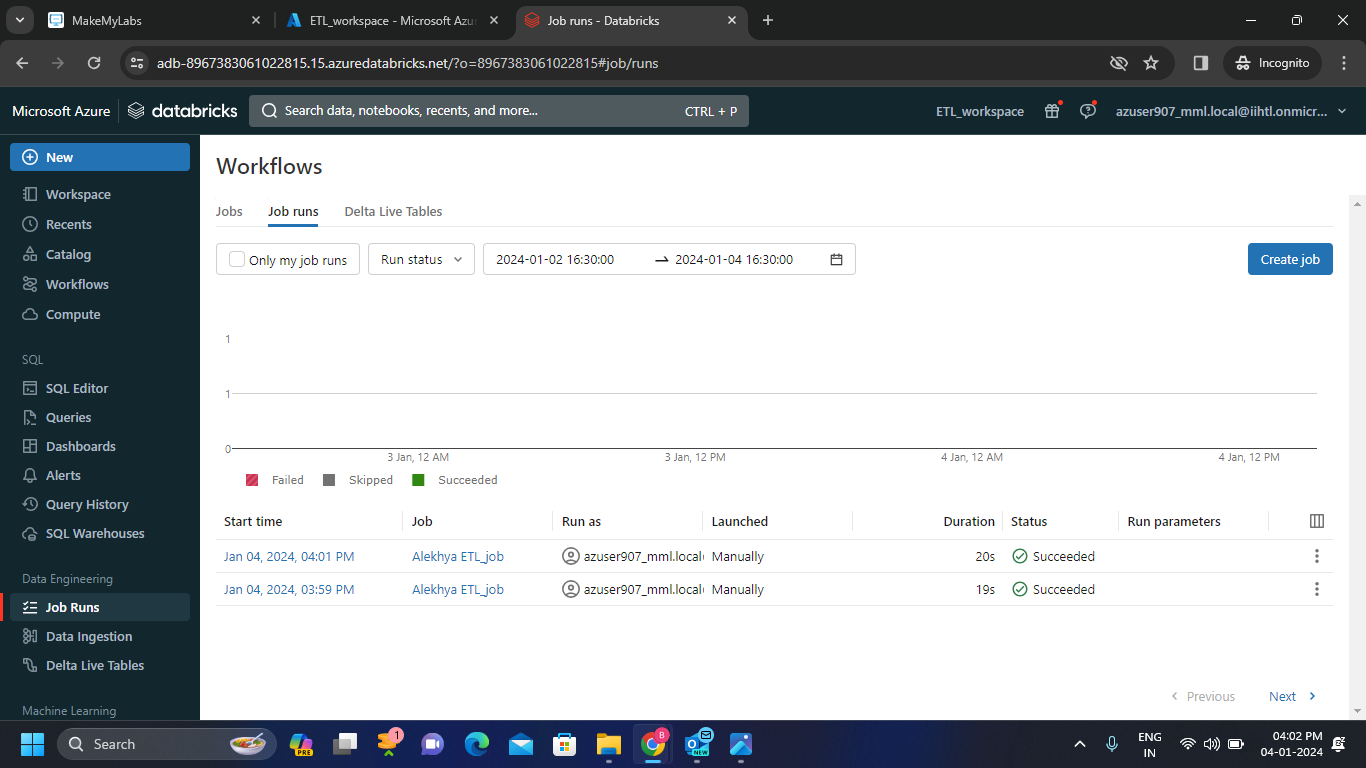
**Job scheduling:** Job scheduling involves the organization and execution of individual tasks or jobs at specific times or under defined conditions.

**Step-1:** Schedule the job for the created table as follows and create the job.

**Step-2:** After scheduling a job is successful we get the alert that the job is created.

Now click on Run now to run the job.

****

**Step-3: **We get the job run is successful in the status column.