**Azure Databricks Case Study**

**Submitted By-**

Subrat Shukla, DE-1

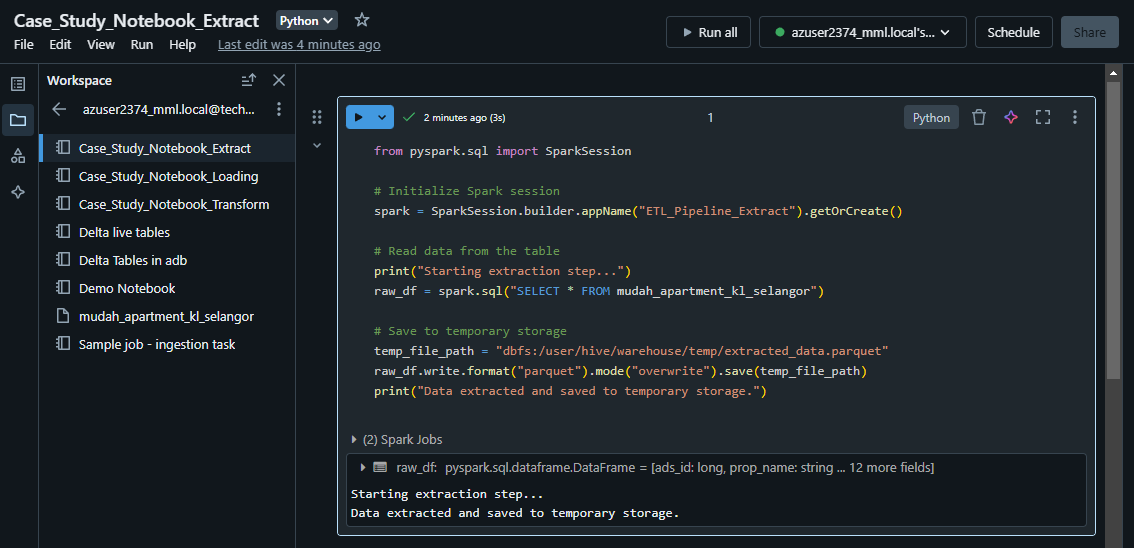
**Task given: Create an ETL pipeline of ingestion & transform and load queries on any data set and initiate the pipeline from workflow using the notebook.**

**Steps:**

* Create a notebook with ETL queries.
* Run the notebook from workflow pipeline in azure databricks workspace.

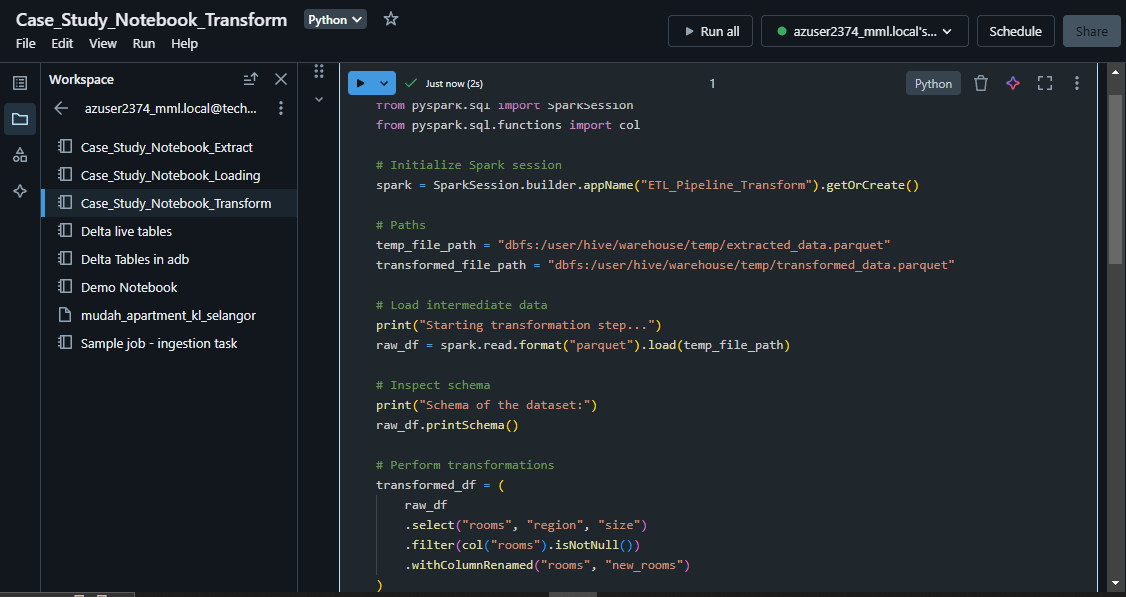
1. **Notebook for Ingestion (Extract):**

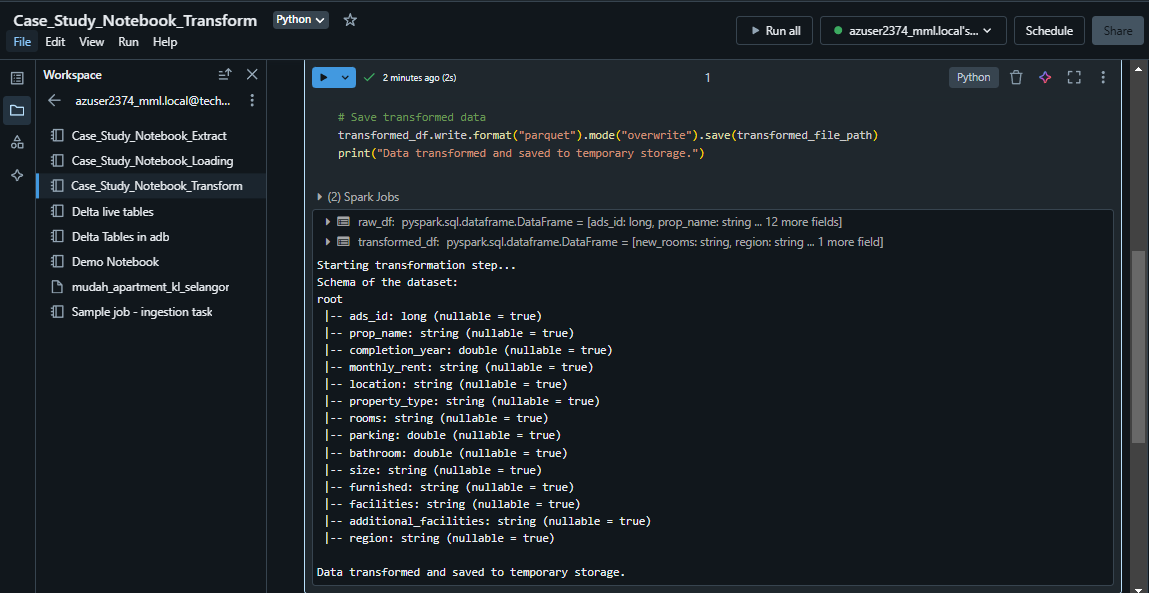
This notebook will load the raw data into the Databricks environment, either from a file or a table.



1. **Notebook for Transformation**

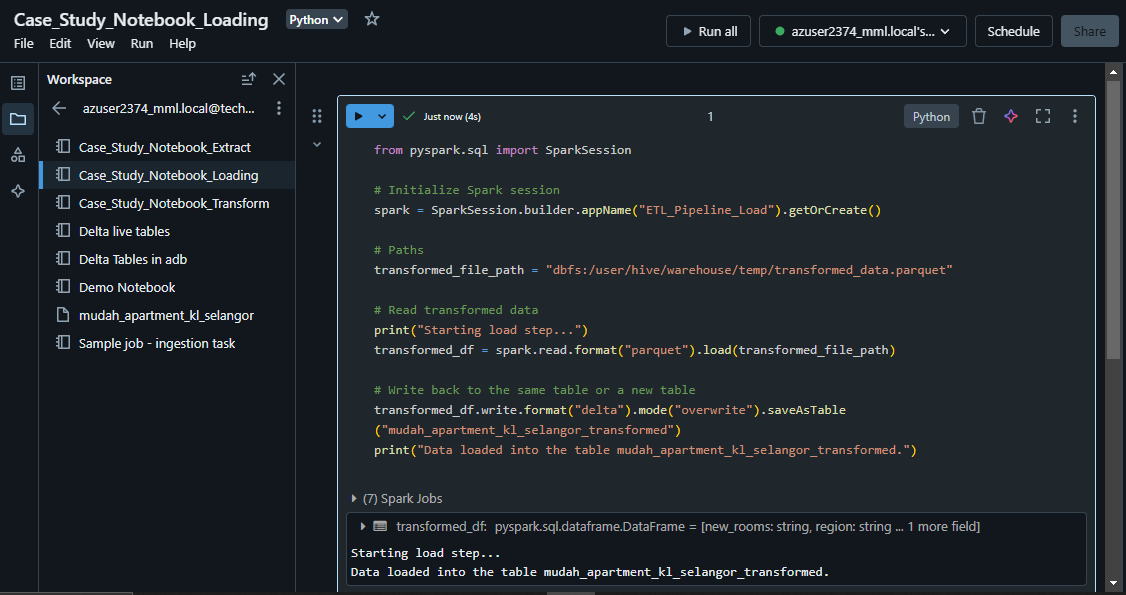
This notebook will read the extracted data, apply transformations, and prepare it for loading.





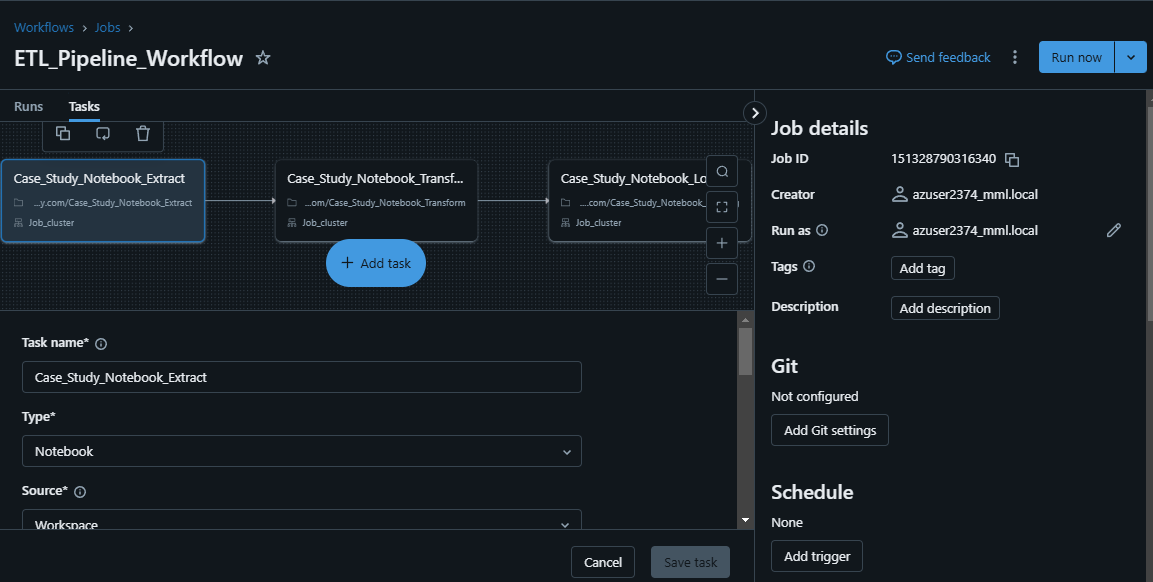
1. **Notebook for Loading**

This notebook will load the transformed data into the final destination, such as a Delta table or another storage format.

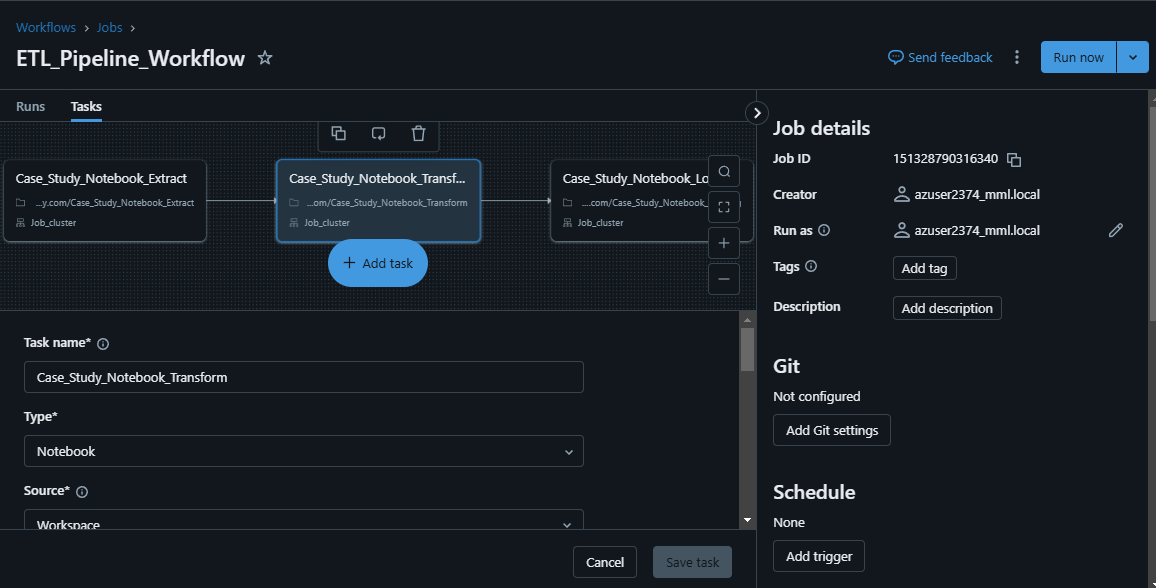


1. **Creating Databricks Workflow:**

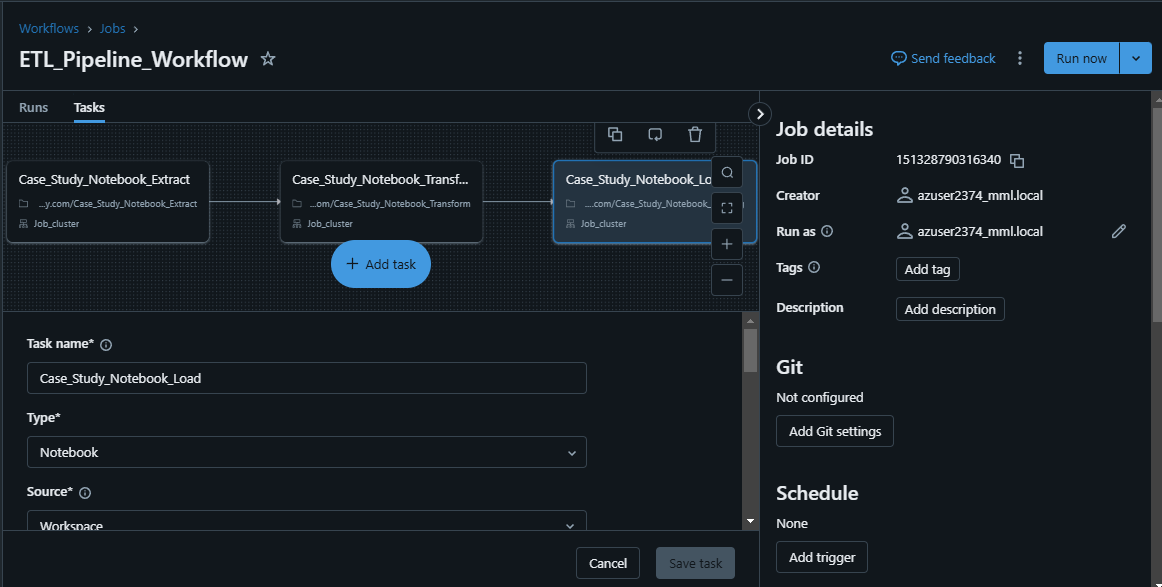
* For Extract process-



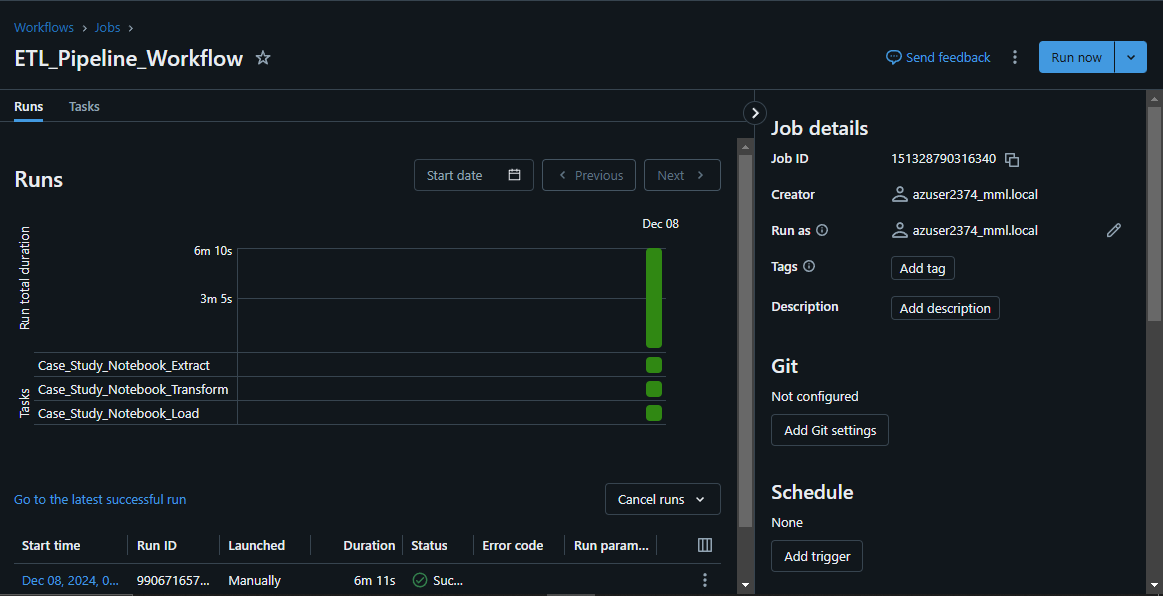
* For Transformation process-



* For Loading process-



1. **Running the Workflows:**



* Hence, The ETL Pipeline consisting processes of Extraction, Transformation, Loading has successfully created and running.

**--Thank You!**