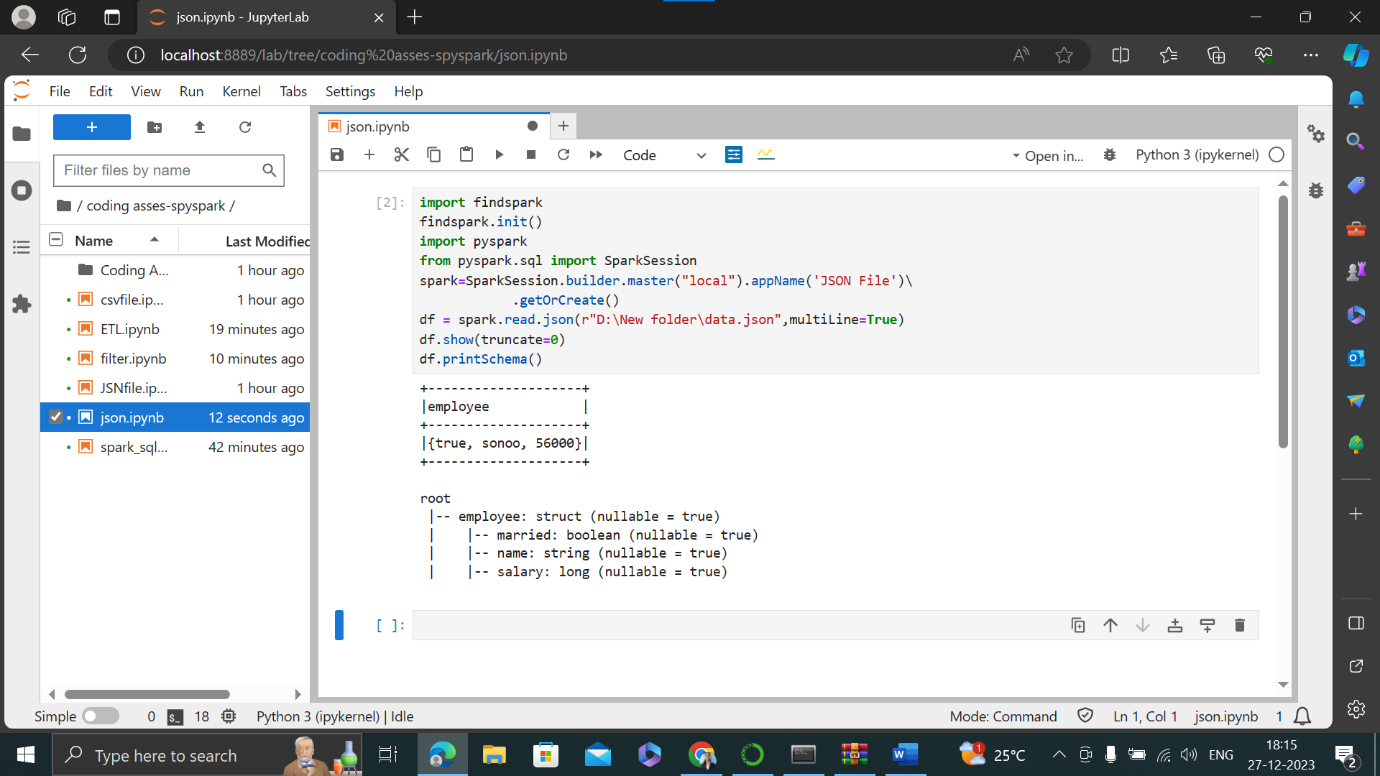
CODING ASSESSMENT SPARK AND PYSPARK Vinutha s

27/12/2023

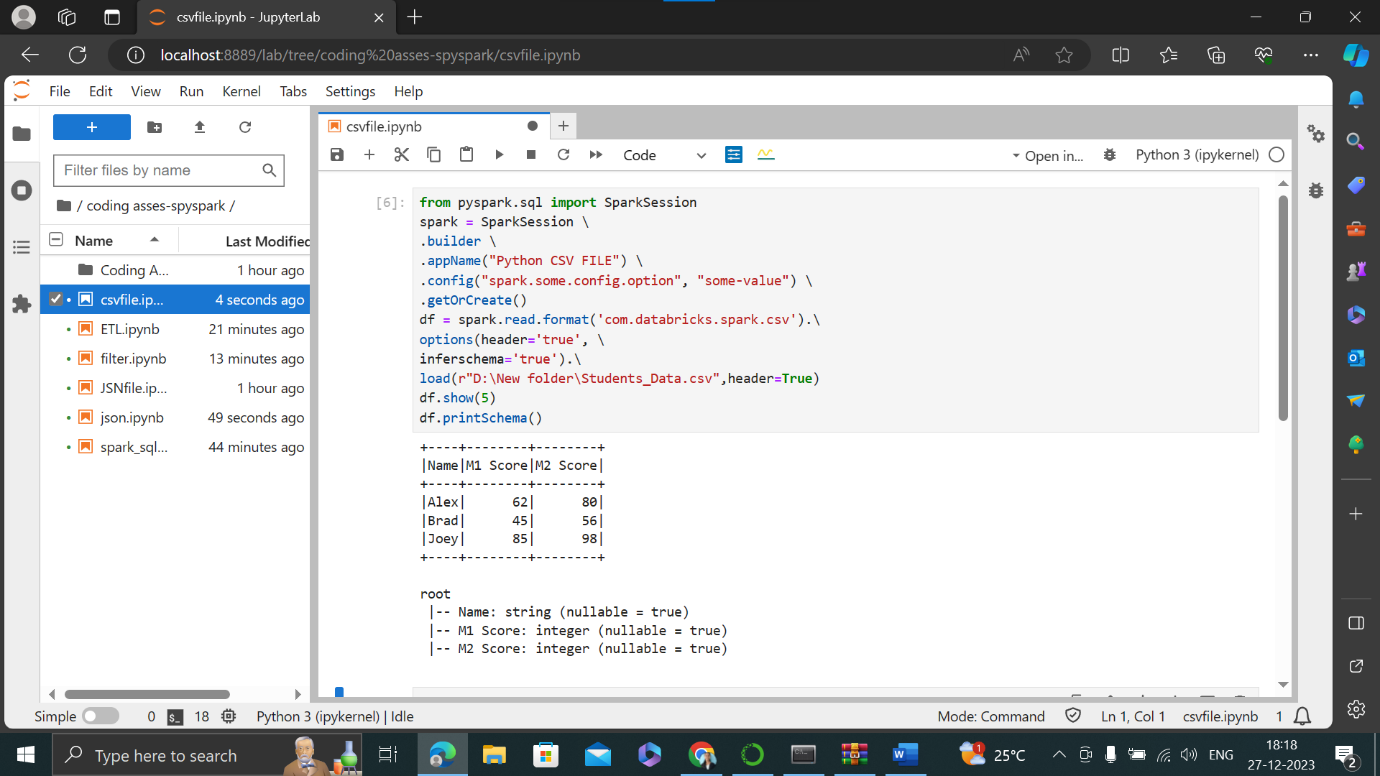
**Coding Assesment**

1. **Implement Processing JSON and CSV data with PySpark**

**JSON FILE IS EXTRACTED:**



**CSV FILE IS EXTRACTED:**

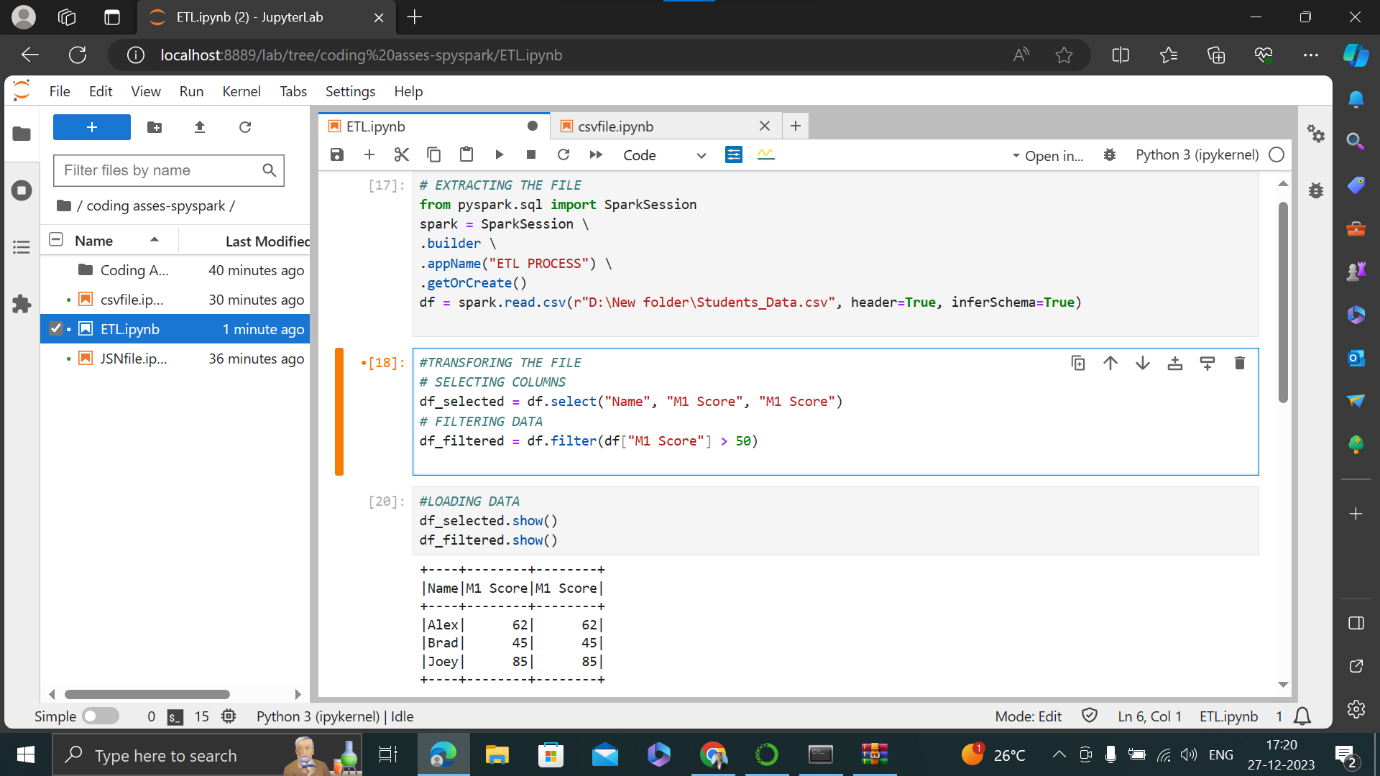


1. **Explain ETL (Extract, Transform, Load) with PySpark**

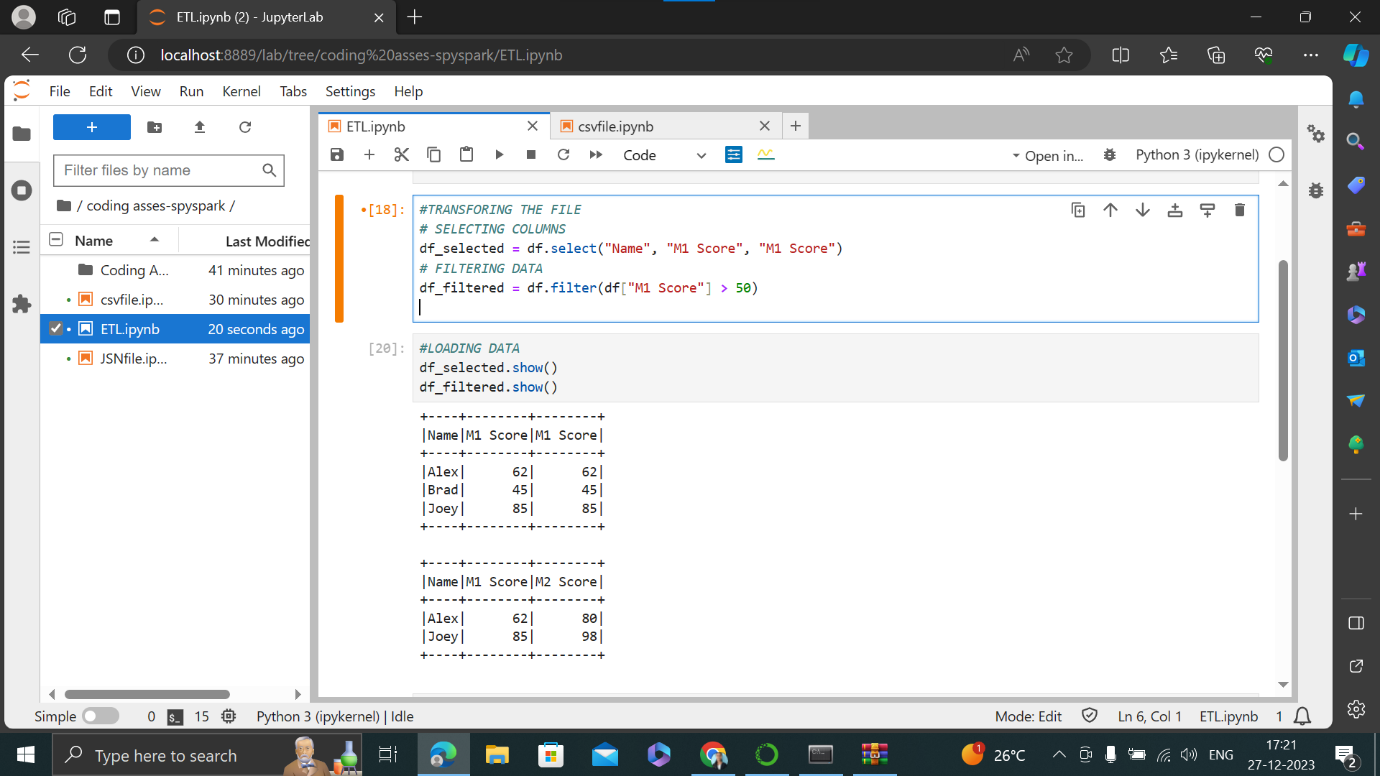
* Extracting data from various sources (databases, files, APIs, etc.).
* Transforming the data to clean, standardize, and prepare it for analysis.
* Loading the processed data into a target destination (data warehouse, database, data lake, etc.)

**The database is created and table is created in the below picture.**

**And also loaded the table for select command.**

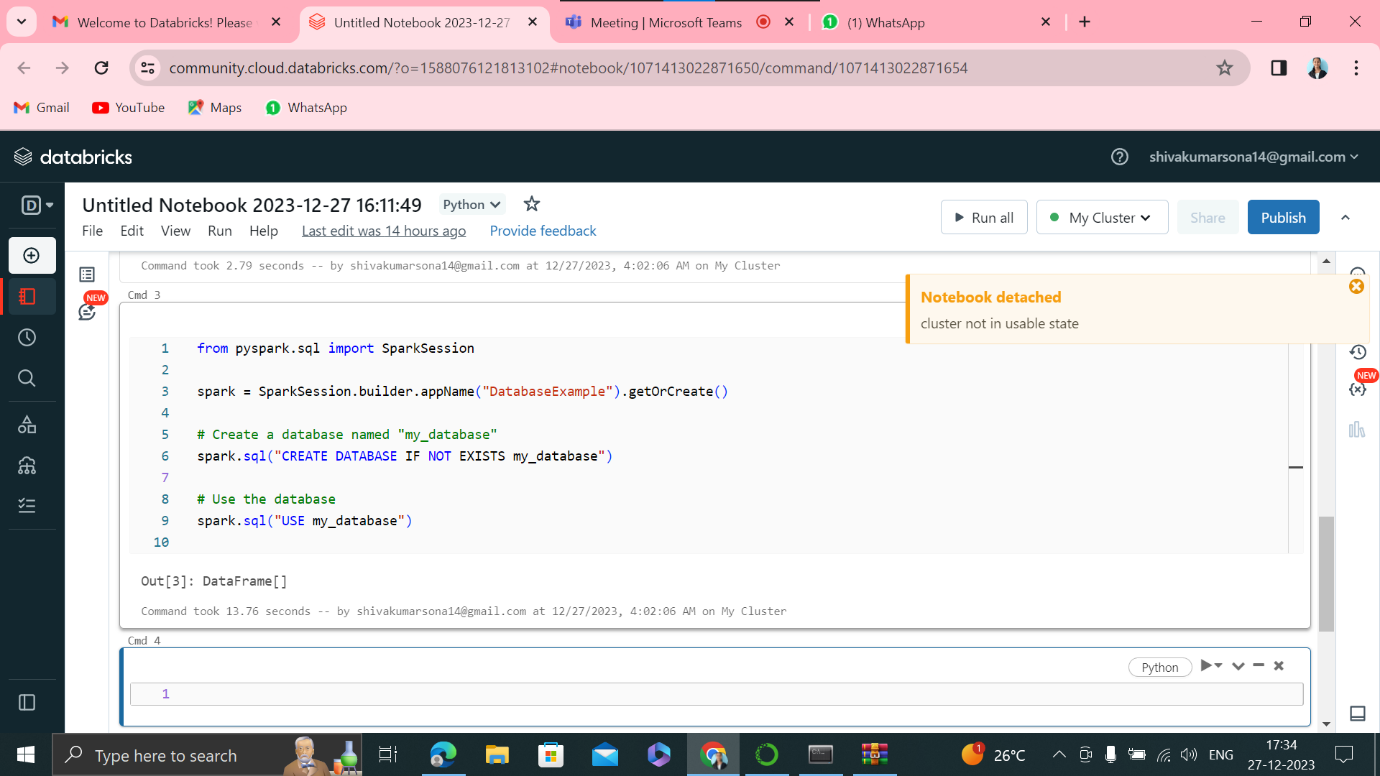


**For loading the data filter command is used in the below picture.**



1. **Using Spark SQL - Creating databases, tables**

**Creating database: A Database is created in the below picture.**



**Created the table and also displayed the output of the table in the below picture.**

