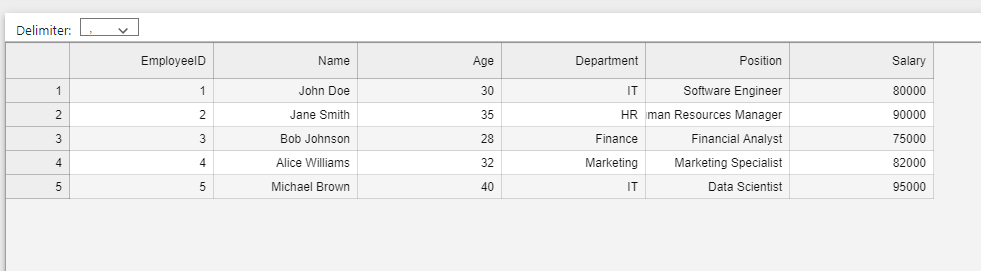
**Pyspark Assignment**

CSV file



Create Spark session

SparkSession is a unified entry point for Spark applications; it was introduced in Spark 2.0. It acts as a connector to all Spark’s underlying functionalities, including RDDs, DataFrames, and Datasets, providing a unified interface to work with structured data processing. It is one of the very first objects you create while developing a Spark SQL application. you create a SparkSession using the SparkSession.builder() method

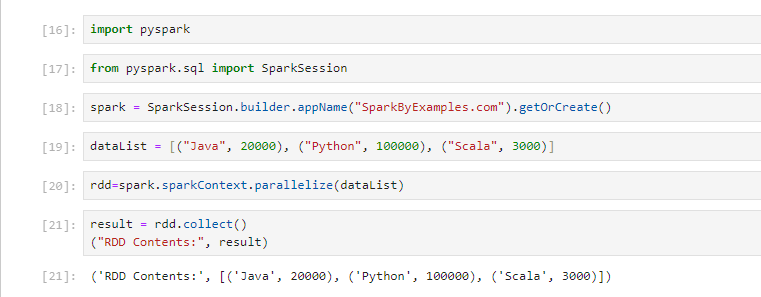
SparkSession consolidates several previously separate contexts, such as SQLContext, HiveContext, and StreamingContext, into one entry point, simplifying the interaction with Spark and its different APIs. It enables users to perform various operations like reading data from various sources, executing SQL queries, creating DataFrames and Datasets, and performing actions on distributed datasets efficiently.



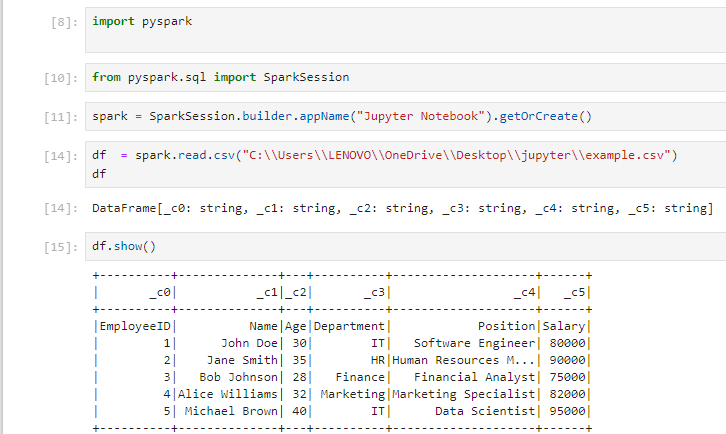
Read Text file



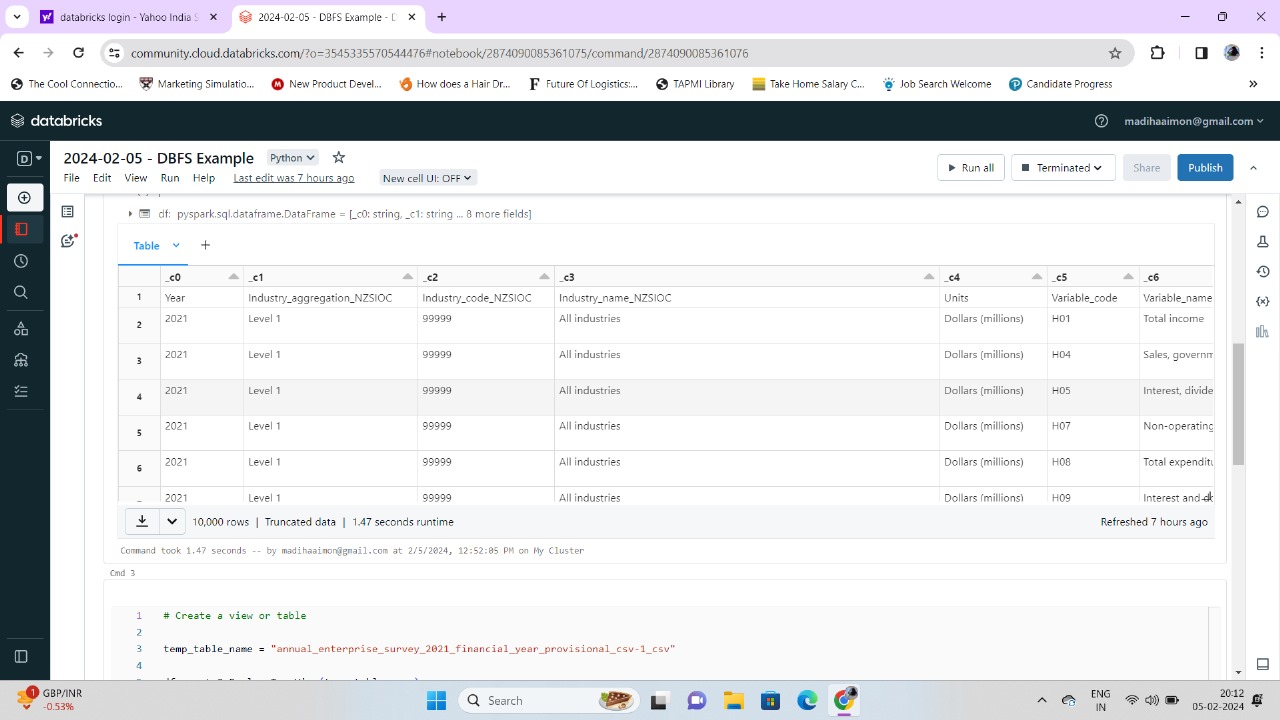
Read List



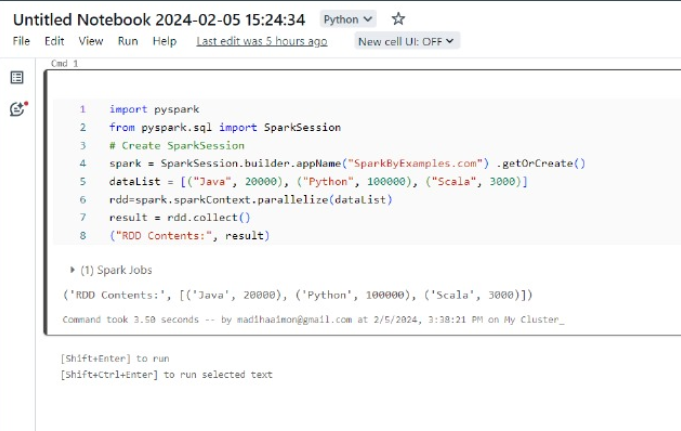
Read CSV data



Insert Table into DataBricks



Read List



Notes

