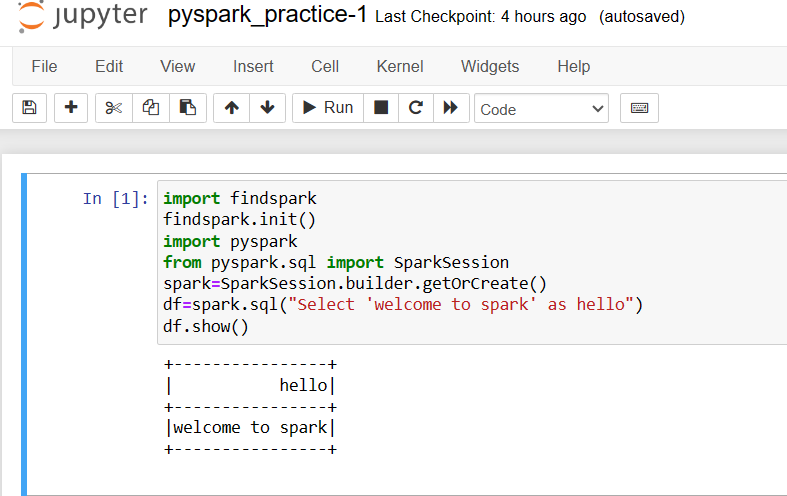
**Name : Hemalatha Karumuri**

**Date : 26-12-23**

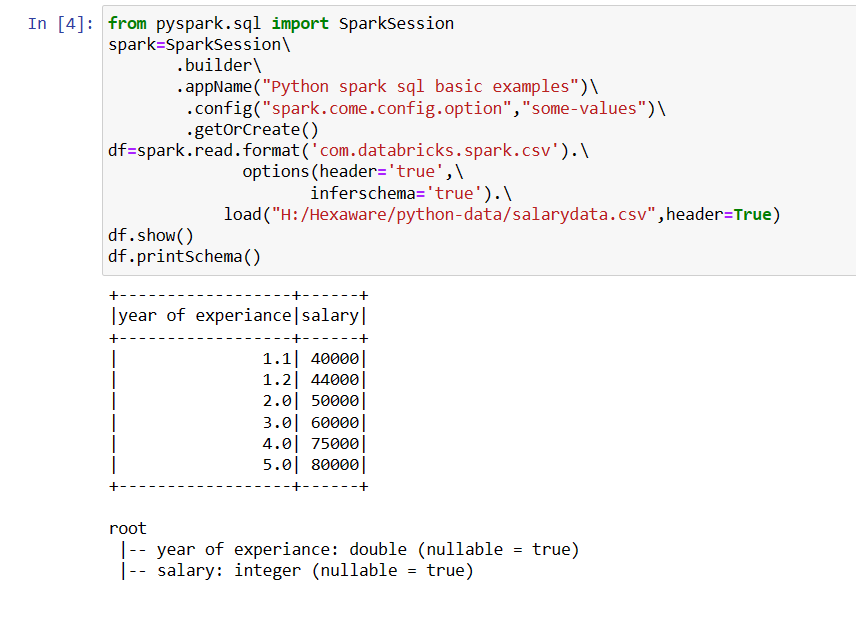
1)Initializing the spark and displaying the text



2)Taking the random number and performing operations



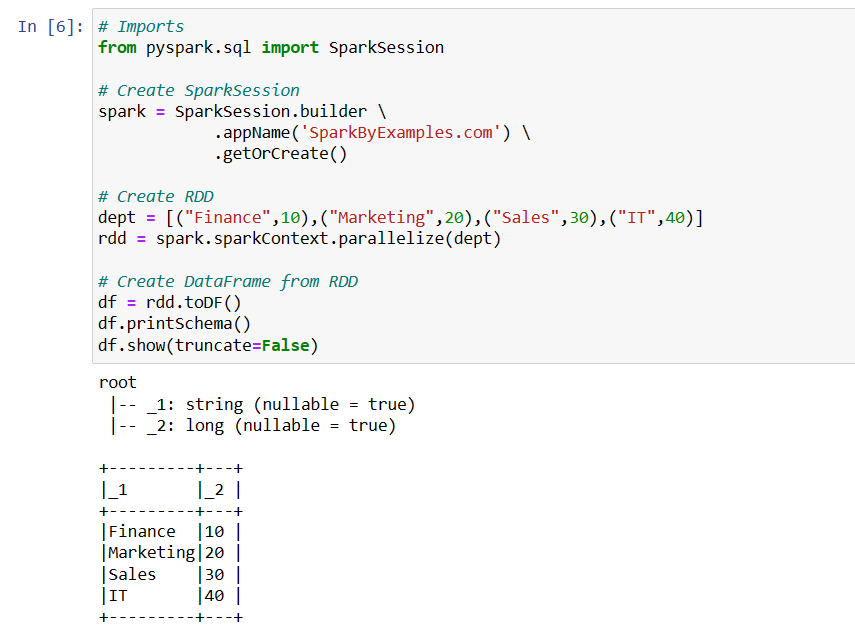
3)Loading the csv file and displaying the content of the file



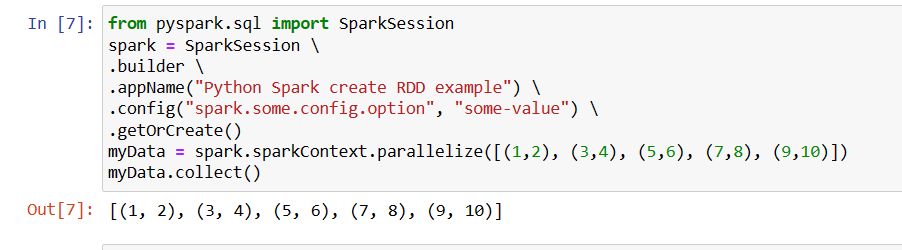
4)Creating the RDD using parallelize() method



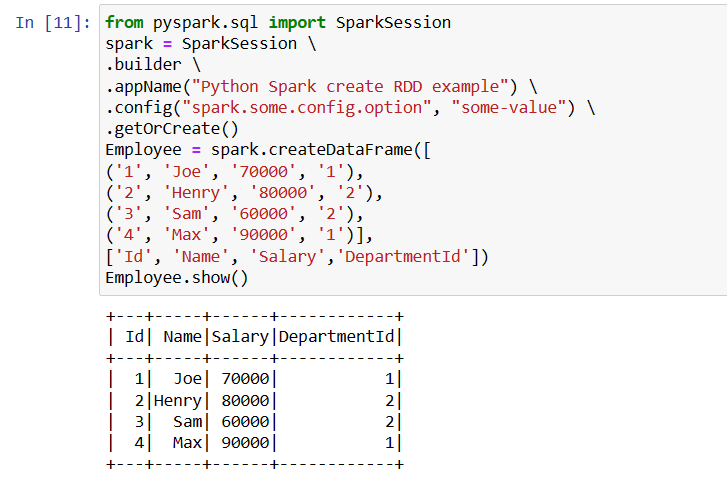
5)creating a dataframe and printing the schema



6)Taking the list of data and displaying the data using collect() method



7) create dataframe By using createDataFrame( ) function



8)Using ETL to retrieve the data from source path and applying some filter and adding a column to the file and reading random values to the added column,perform order by clause,concatenation of two columns etc.

