**Assignment 12**

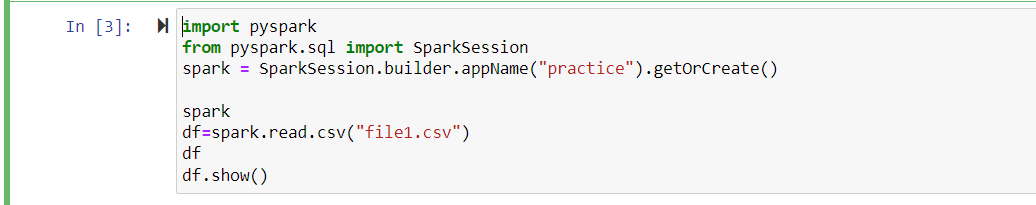
**Name**: Jeevan Sai Badana

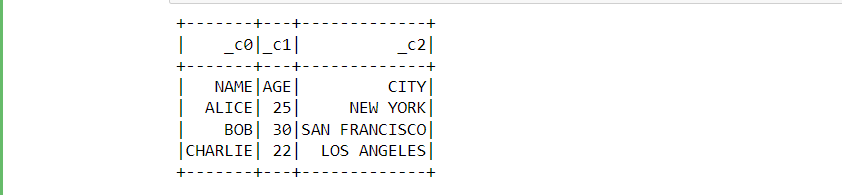
**Mail**: [jeevansai100@gmail.com](mailto:jeevansai100@gmail.com)

**Date**: 05-02-2024

Here we read csv file using Spark for this first we import the PySpark module

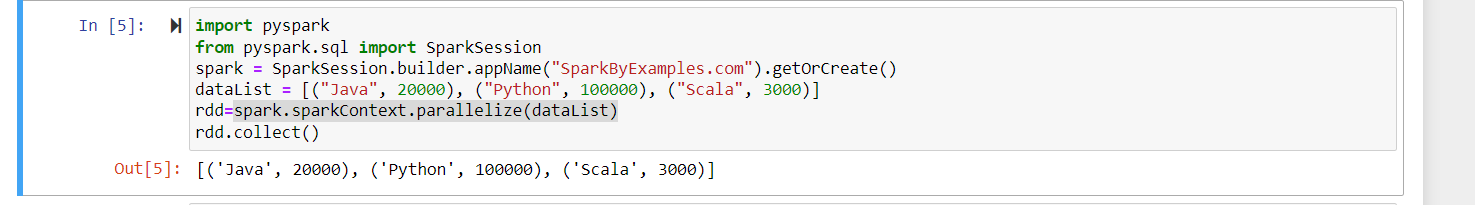
* Later we import SparkSession and with read.csv we read and we print the data using df.show()





Now we can create RDD and print data in it

* First we import PySpark and then import SparkSession
* Now we give spark variable and then we write our datalist
* Now we give rdd with spark.sparkContext.parallelize(dataList)
* Now we print data using rdd.collect()



**Running notes :-**

