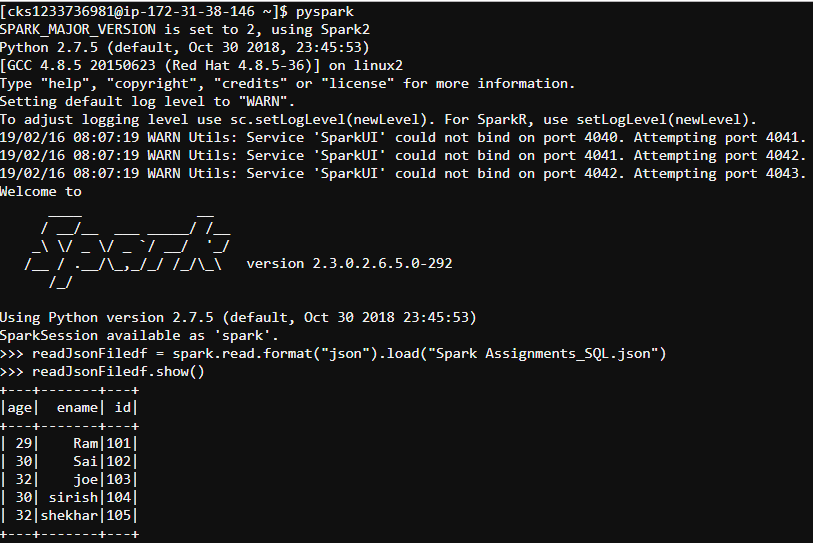
1. From the above json file answer the following.

a) Display file data using data frame.

readJsonFiledf = spark.read.format("json").load("Spark Assignments\_SQL.json")

readJsonFiledf.show()



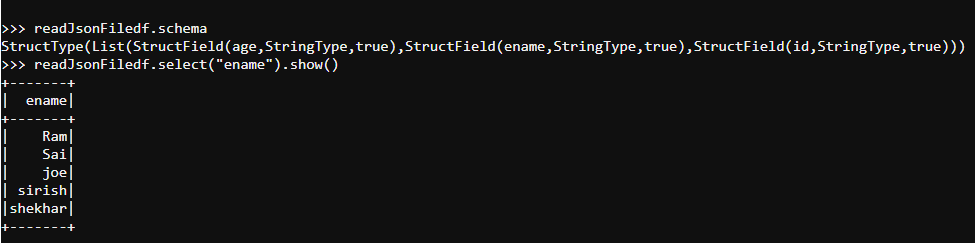
b) Display the structure of data frame .

readJsonFiledf.schema

StructType(List(StructField(age,StringType,true),StructField(ename,StringType,true),StructField(id,StringType,true)))

c) Display only enames from the given file using select.

readJsonFiledf.select("ename").show()



2. From the json file given, answer the following.

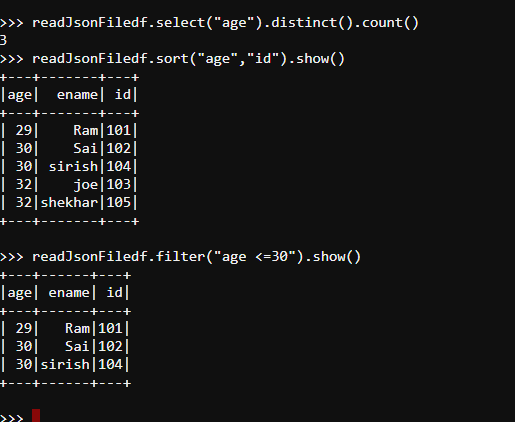
a) Provide the count of each age.

readJsonFiledf.select("age").distinct().count()

b) Provide all records in sorted order.

readJsonFiledf.sort("id","age").show()

readJsonFiledf.sort("age","id").show()



c) Display records whose age <=30.

readJsonFiledf.filter("age <=30").show()

readJsonFiledf.where("col(age) <=30").show()

from pyspark.sql.functions import col

readJsonFiledf.where(col("age") <=30).show()