**READING NESTED JSON DATA USING SPARK**

val tagsDF2=spark.read.option("multiLine", true).option("delimiter",",").option("inferSchema", true).json("/user/edureka\_918210/JsonFile/employees.json")

scala> tagsDF2.printSchema

root

|-- addresses: array (nullable = true)

| |-- element: struct (containsNull = true)

| | |-- city: string (nullable = true)

| | |-- postal\_code: string (nullable = true)

| | |-- state: string (nullable = true)

| | |-- street\_name: string (nullable = true)

| | |-- street\_number: string (nullable = true)

|-- email: string (nullable = true)

|-- first\_name: string (nullable = true)

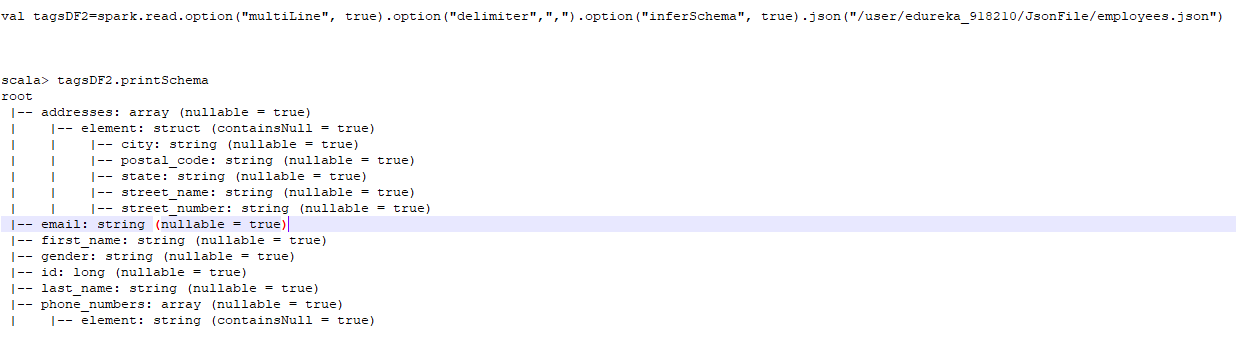
|-- gender: string (nullable = true)

|-- id: long (nullable = true)

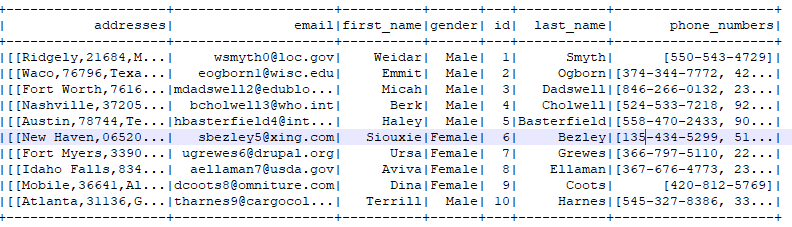
|-- last\_name: string (nullable = true)

|-- phone\_numbers: array (nullable = true)

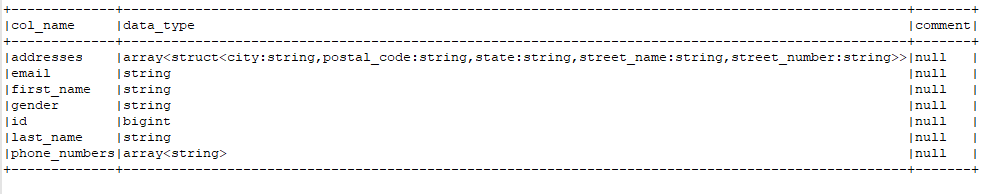
| |-- element: string (containsNull = true)



scala> tagsDF2.show

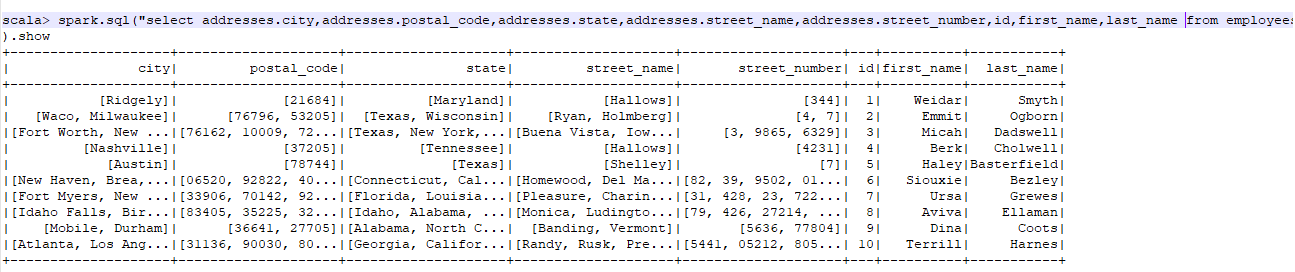


scala> spark.sql("Describe employees").show(false)



scala> spark.sql("select addresses.city,addresses.postal\_code,addresses.state,addresses.street\_name,addresses.street\_number,id,first\_name,last\_name from employees"

).show



df3.repartition(5).write.option("header","true").json("/user/edureka\_918210/JsonToCSVFile")

hadoop fs -cat /user/edureka\_918210/JsonToCSVFile/\*