**Load file**

**Soln:**

**path = "/FileStore/tables/Baby\_Names\_\_Beginning\_2007.csv"**

**pathLoad=spark.read.load(path,format="com.databricks.spark.csv",header="true",**

**inferSchema="true")**

**pathLoad.show()**

**Q-1) How many different types of country**

**Soln:**

**pathLoad.select("County").distinct().count()**

**Q-2) Show distinct name**

**Soln:**

**pathLoad.select("First Name").distinct().orderBy("First Name").show()**

**Q3) Show sex of “DAISY”**

**Soln:**

**pathLoad.select(pathLoad["First Name"],pathLoad["Sex"]).filter(pathLoad["First Name"]=="DAISY").show()**

**Q4) how many children has “LIAM”**

**Soln:**

**display(pathLoad.select(pathLoad["First Name"],pathLoad["Count"]).groupBy(pathLoad["First Name"]).sum().filter(pathLoad["First Name"]=="LIAM"))**

**Q5) show total number of children’s was birth from 2007-2013**

**Q6) show how many children were birth in “NEW YORK”**

**Q7) Show the name whose sex is “M”**

**Q8) Show the name whose sex is “F”**