Spark SQL: Inferring Schemas using case classes

1. Create new Scala notebook and import the following packages

Chart

Description automatically generated

1. Create Customer case class

A picture containing chart

Description automatically generated

1. Use the file which we already loaded into hive(activity – 1)

A picture containing graphical user interface

Description automatically generated

1. Split each line into an array of string, based on comma as a delimiter



1. Convert the RDD of Array [String] into the RDD of Customer case objects



1. Convert the customerRDD into customerDF data frame

Shape

Description automatically generated

1. Register the customerDF as a view

A picture containing logo

Description automatically generated

1. Run a SQL query against it

A picture containing graphical user interface

Description automatically generated

1. Get the output values from persons

A picture containing graphical user interface

Description automatically generated