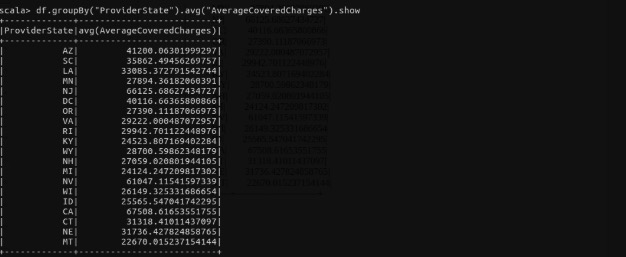
1. Loading CSV file into spark

val session = org.apache.spark.sql.SparkSession.builder.master("local").appName("Spark CSV Reader").getOrCreate; val df = session.read.format("com.databricks.spark.csv").option("header", "true").option("inferSchema", "true").load(<file:///home/kiran/Documents/datasets/inpatientCharges.csv>);

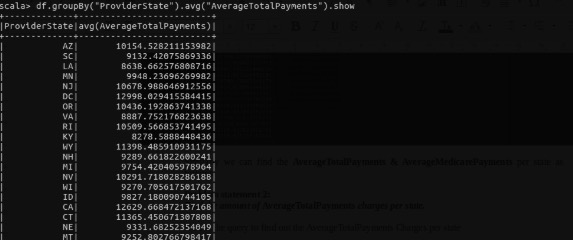
1. Finding AverageCoveredCharges

Scala>df.groupBy(“ProviderState”).avg(“AverageCoveredCharges”).show



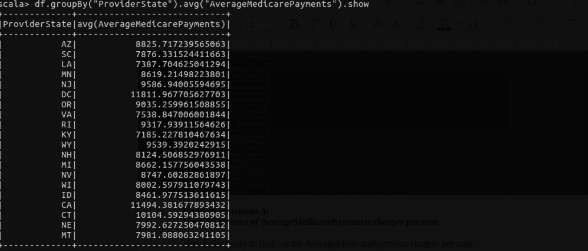
3) Find the amount of**Average Total Payments**charges per state**.**

Scala>df.groupBy(“ProviderState”).avg(“AverageTotalPayments”).show



4) Find the amount of **Average Medicare Payments**charges per state**.**

Scala>df.groupBy(“ProviderState”).avg(“AverageMedicarePayments”).show



5) Find out the total number of **Discharges** per state and for each disease.

df.groupBy((“ProviderState”),(“DRGefinition”)).sum(“TotalDischarges”).sort(desc(sum(“TotalDischarges”).toString)).show

