**1. What are the total number of gold medal winners every year**

val textFile = sc.textFile("/home/acadgild/hadoop/Sports\_data.txt")

val schemaStr = "firstname:string,lastname:string,sports:string,medal:string,age:integer,year:integer,country:string"

val schema = StructType(schemaStr.split(",").map(x => StructField(x.split(":")(0), if (x.split(":")(1).equals("string")) StringType else IntegerType,true)))

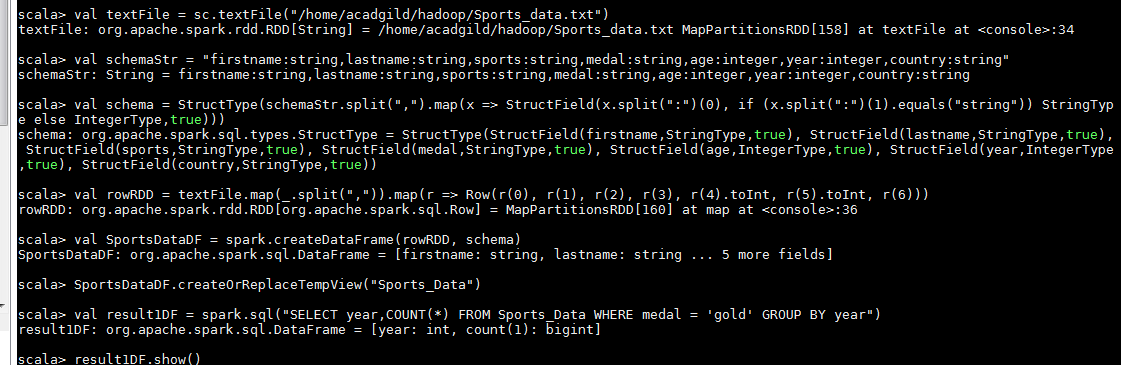
val rowRDD = textFile.map(\_.split(",")).map(r => Row(r(0), r(1), r(2), r(3), r(4).toInt, r(5).toInt, r(6)))

val SportsDataDF = spark.createDataFrame(rowRDD, schema)

SportsDataDF.createOrReplaceTempView("Sports\_Data")

val result1DF = spark.sql("SELECT year,COUNT(\*) FROM Sports\_Data WHERE medal = 'gold' GROUP BY year")

result1DF.show()





**2. How many silver medals have been won by USA in each sport**

val textFile = sc.textFile("/home/acadgild/hadoop/Sports\_data.txt")

val schemaString = "firstname:string,lastname:string,sports:string,medal:string,age:integer,year:integer,country:string"

val schema = StructType(schemaString.split(",").map(x => StructField(x.split(":")(0), if (x.split(":")(1).equals("string")) StringType else IntegerType, true)))

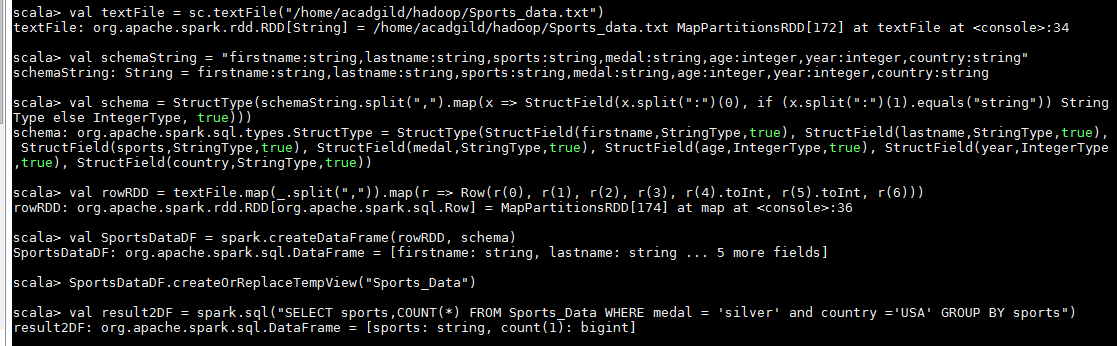
val rowRDD = textFile.map(\_.split(",")).map(r => Row(r(0), r(1), r(2), r(3), r(4).toInt, r(5).toInt, r(6)))

val SportsDataDF = spark.createDataFrame(rowRDD, schema)

SportsDataDF.createOrReplaceTempView("Sports\_Data")

val result2DF = spark.sql("SELECT sports,COUNT(\*) FROM Sports\_Data WHERE medal = 'silver' and country ='USA' GROUP BY sports")

result2DF.show()



****