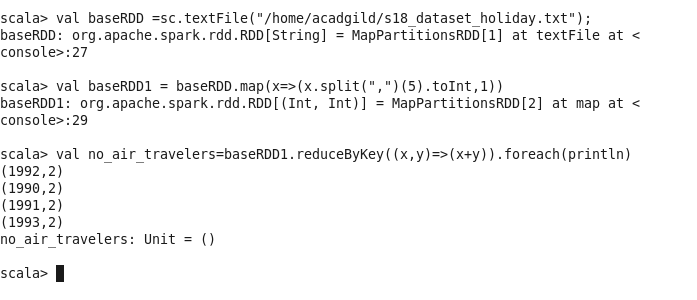
1. What is the distribution of the total number of air-travelers per year

Val baseRDD =sc.textFile(“/home/acadgild/s18\_dataset\_holiday.txt”)

val baseRDD1 =baseRDD.map(x=>(x.split(“,”)(5).toInt,1))

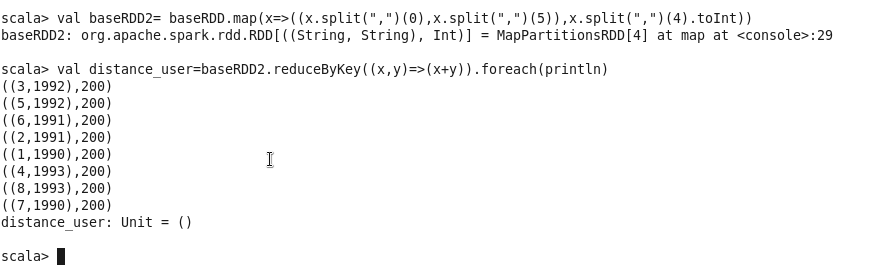
val no\_air\_travelers=baseRDD1.reduceByKey((x,y)=>(x+y)).foreach(println)



1. What is the total air distance covered by each user per year

Val baseRDD2=baseRDD.map(x=>((x.split(“,”)(0),x.split(“,”)(5)),x.split(“,”)(4).toInt))

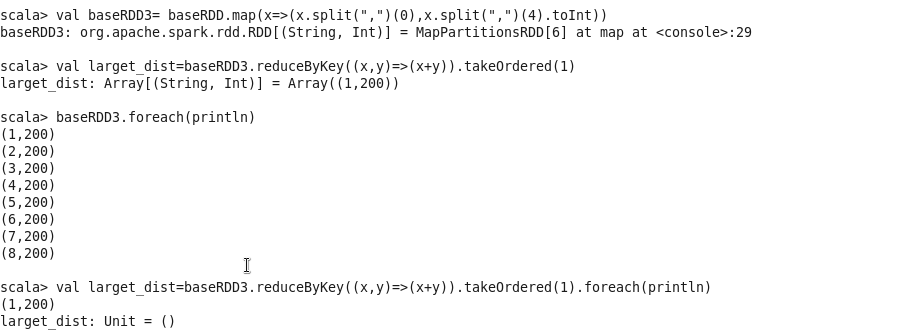
Val distanct\_user= baseRDD2.reduceByKey((x,y)=>(x+y)).foreach(println)



1. Which user has travelled the largest distance till date

Val baseRDD3=baseRDD.map(x=>(x.split(“,”)(0),x.split(“,”)(4). toInt))

Val larget\_dist=baseRDD3.reduceByKey((x,y=>(x+y)).takeOrdered(1).foreach(println)



1. What is the most preferred destination for all users.

