1. Which route is generating the most revenue per year

**case class Holidays(userid:String,source:String,destination:String,transportationmode:String,amount:Int,year:Int)**

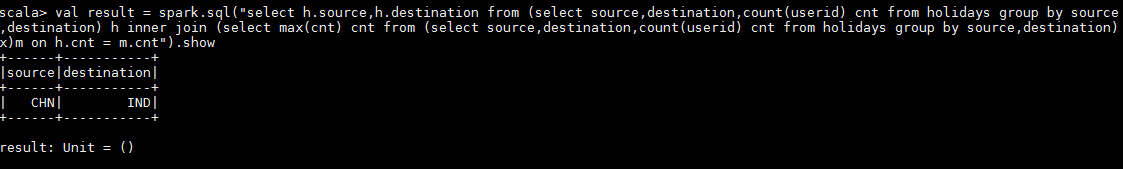
**val holidaydsRDD = sc.textFile("/home/acadgild/inputdir/S18\_Dataset\_Holidays.txt").map(line => line.split(","))**

**val holidaysmapRDD = holidaydsRDD.map(h=>Holidays(h(0),h(1),h(2),h(3),h(4).toInt,h(5).toInt))**

**val holidaysDF = holidaysmapRDD.toDF**

**holidaysDF.registerTempTable("holidays")**

**val result = spark.sql("select h.source,h.destination from (select source,destination,count(userid) cnt from holidays group by source,destination) h inner join (select max(cnt) cnt from (select source,destination,count(userid) cnt from holidays group by source,destination)x)m on h.cnt = m.cnt").show**



1. What is the total amount spent by every user on air-travel per year

**case class Holidays(userid:String,source:String,destination:String,transportationmode:String,amount:Int,year:Int)**

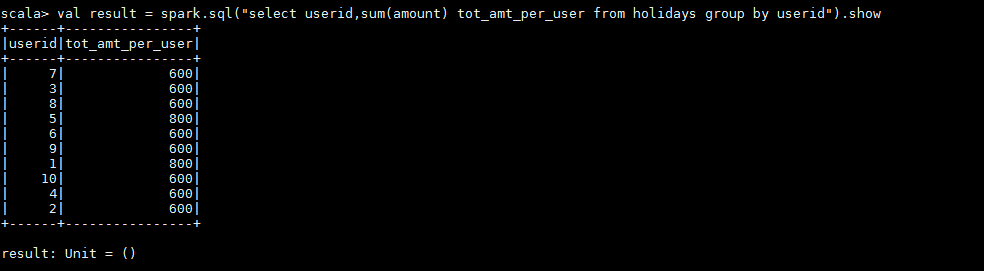
**val holidaydsRDD = sc.textFile("/home/acadgild/inputdir/S18\_Dataset\_Holidays.txt").map(line => line.split(","))**

**val holidaysmapRDD = holidaydsRDD.map(h=>Holidays(h(0),h(1),h(2),h(3),h(4).toInt,h(5).toInt))**

**val holidaysDF = holidaysmapRDD.toDF**

**holidaysDF.registerTempTable("holidays")**

**val result = spark.sql("select userid,sum(amount) tot\_amt\_per\_user from holidays group by userid").show**



3) Considering age groups of < 20 , 20-35, 35 > ,Which age group is travelling the most

every year.

**case class Holidays(userid:String,source:String,destination:String,transportationmode:String,amount:Int,year:Int)**

**val holidaydsRDD = sc.textFile("/home/acadgild/inputdir/S18\_Dataset\_Holidays.txt").map(line => line.split(","))**

**val holidaysmapRDD = holidaydsRDD.map(h=>Holidays(h(0),h(1),h(2),h(3),h(4).toInt,h(5).toInt))**

**val holidaysDF = holidaysmapRDD.toDF**

**holidaysDF.registerTempTable("holidays")**

**case class UserDetails(userid:String,name:String,age:Int)**

**val userdetailsRDD = sc.textFile("/home/acadgild/inputdir/S18\_Dataset\_User\_details.txt").map(line => line.split(","))**

**val userdetailssmapRDD = userdetailsRDD.map(u=>UserDetails(u(0),u(1),u(2).toInt))**

**val userdetailsDF = userdetailssmapRDD.toDF**

**userdetailsDF.registerTempTable("userdetails")**

**val result = spark.sql("select a.age\_group,count(a.age\_group) cnt from (select h.userid,u.name,case when u.age between 1 and 20 then '20' when u.age between 21 and 35 then '21-35' else '35' end age\_group from holidays h inner join userdetails u on h.userid=u.userid)a group by a.age\_group order by cnt desc limit 1").show**

