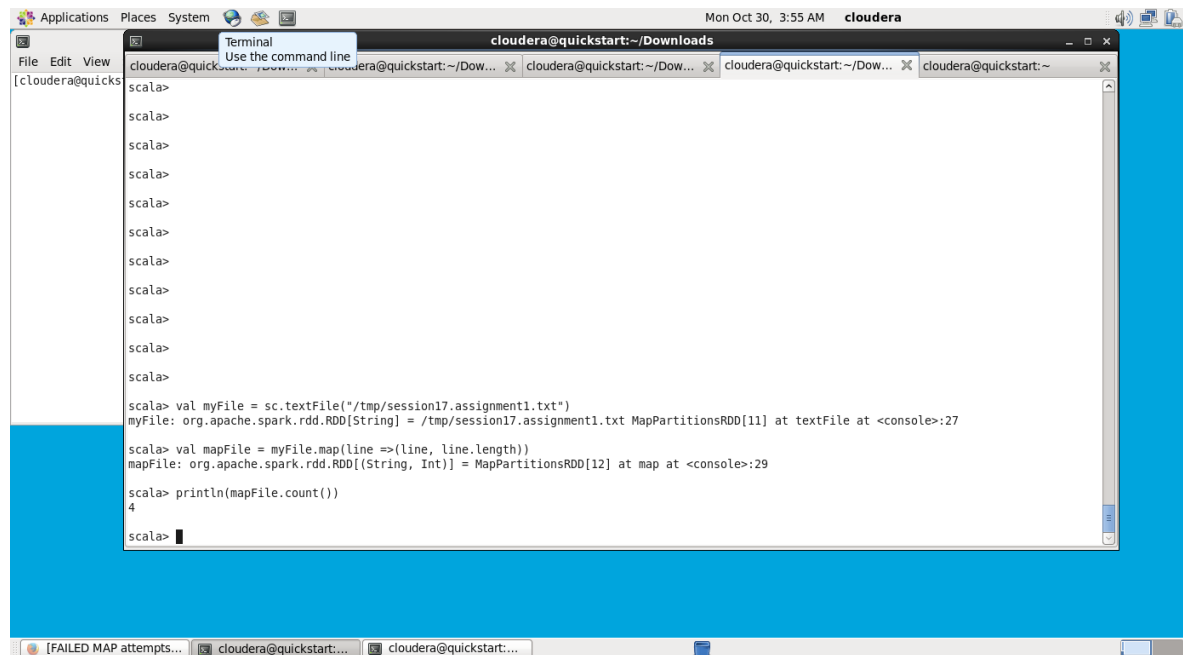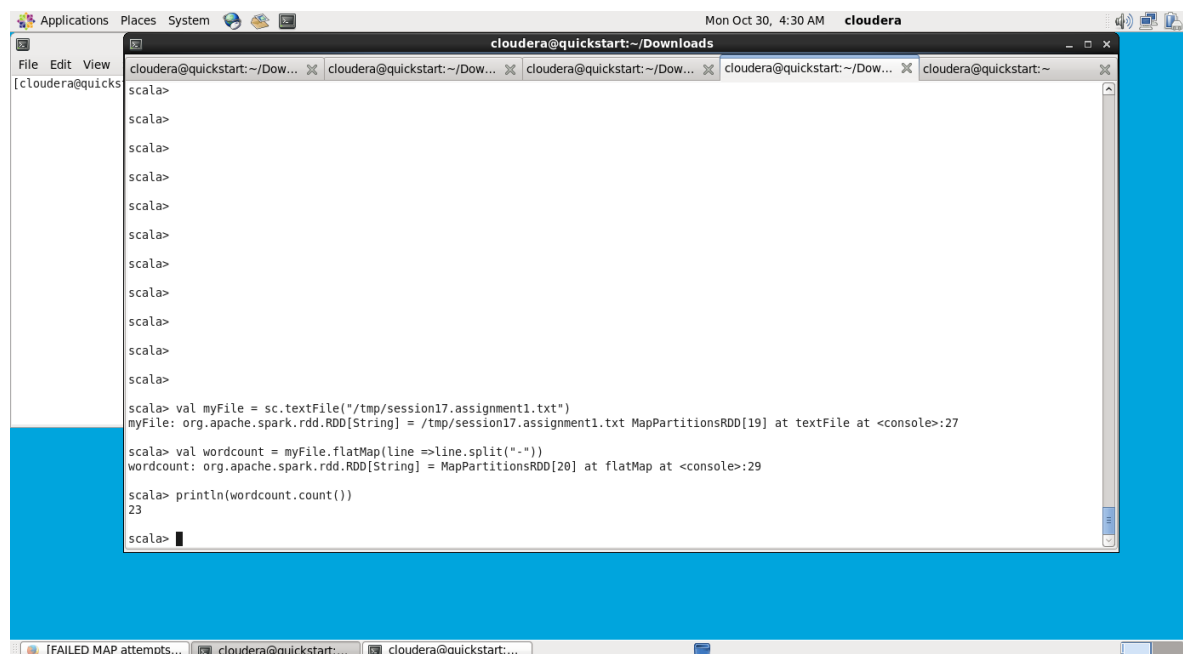1. Write a program to read a text file and print the number of rows of data in the document.


```
scala>
scala>
scala>
scala>
scala>
scala>
scala>
scala>
scala>
scala>
scala>
scala> val myFile = sc.textFile("/tmp/session17.assignment1.txt")
myFile: org.apache.spark.rdd.RDD[String] = /tmp/session17.assignment1.txt MapPartitionsRDD[11] at textFile at <console>:27

scala> val mapFile = myFile.map(line =>(line, line.length))
mapFile: org.apache.spark.rdd.RDD[(String, Int)] = MapPartitionsRDD[12] at map at <console>:29

scala> println(mapFile.count())
4

scala>
```

2. Write a program to read a text file and print the number of words in the document. Or
We have a document where the word separator is '-', instead of space. Write a spark code, to obtain the count of the total number of words present in the document


```
scala>
scala>
scala>
scala>
scala>
scala>
scala>
scala>
scala>
scala>
scala>
scala> val myFile = sc.textFile("/tmp/session17.assignment1.txt")
myFile: org.apache.spark.rdd.RDD[String] = /tmp/session17.assignment1.txt MapPartitionsRDD[19] at textFile at <console>:27

scala> val wordcount = myFile.flatMap(line =>line.split("-"))
wordcount: org.apache.spark.rdd.RDD[String] = MapPartitionsRDD[20] at flatMap at <console>:29

scala> println(wordcount.count())
23

scala>
```