1. **Program to run wordcount on scala shell**
   Note- Create a textfile sparkdata.txt locally and give appropriate path while loading the data using sc.textFile

```
val data=sc.textFile("sparkdata.txt")

data.collect;

val splitdata = data.flatMap(line => line.split(" "));

splitdata.collect;

val mapdata = splitdata.map(word => (word,1));

mapdata.collect;

val reducedata = mapdata.reduceByKey(_+_);

reducedata.collect;
```

2. **Using RDD and FlaMap count how many times each word appears in a file and write out a list of words whose count is strictly greater than 4 using Spark.**

```
  val textFile = sc.textFile("/home/bhoom/Desktop/wc.txt")
val counts = textFile.flatMap(line => line.split(" ")).map(word => (word, 1)).reduceByKey(_ + _)
import scala.collection.immutable.ListMap
val sorted=ListMap(counts.collect.sortWith(_._2 > _._2):_*)// sort in descending order based on values
println(sorted)
for((k,v)<-sorted)
{
  if(v>4)
    {
     print(k+",")
       print(v)
       println()
     }
}
```

3. Execute Hello World Program in SCALA IDE. Follow the steps given in https://www.dataneb.com/post/hello-world-with-scala-ide

4. **Program to run wordcount on scala IDE**

   Note: Follow the Helloworld program steps to create scala object and type in the following program and execute

```
package wordcount


import org.apache.spark.SparkConf
import org.apache.spark.SparkContext


import org.apache.spark.rdd.RDD.rddToPairRDDFunctions
```

```
object WordCount {

def  main(args: Array[String]) = {


//Start the Spark context

   val conf = new SparkConf().setAppName("WordCount")
.setMaster("local")



val sc = new SparkContext(conf)


//Read some example file to a test RDD

val test = sc.textFile("input.txt")

test.flatMap { line => //for each line
line.split(" ") //split
the line in word by word.
}.map {
word => //for each word (word, 1)

//Return
a key/value tuple, with the word as key and 1 as value

}



.reduceByKey(_ + _) //Sum all of the value with same key



.saveAsTextFile("output.txt") //Save to a text file


//Stop the Spark context

sc.stop
}


}
```