

## Scala Program to find word count

```
scala> var linesRDD = sc.textFile("input1.txt")
linesRDD: org.apache.spark.rdd.RDD[String] = input1.txt MapPartitionsRDD[6] at textFile at
<console>:23
```

```
scala> var wordsRDD = linesRDD.flatMap(_.split(" "))
wordsRDD: org.apache.spark.rdd.RDD[String] = MapPartitionsRDD[7] at flatMap at
<console>:23
```

```
scala> var wordsKvRdd = wordsRDD.map((_, 1))
wordsKvRdd: org.apache.spark.rdd.RDD[(String, Int)] = MapPartitionsRDD[8] at map at
<console>:23
```

```
scala> var wordCounts = wordsKvRdd.reduceByKey(_ + _)
wordCounts: org.apache.spark.rdd.RDD[(String, Int)] = ShuffledRDD[9] at reduceByKey at
<console>:23
```

```
scala> wordCounts.saveAsTextFile("myopt")
```

```
scala> wordCounts.foreach(println)
(hemlo,17)
>Hello,7)
(Hi,16)
>Hello,17)
(,1)
(hola,17)
(helo,17)
```