

```
root@student: /home/student/DSBDAL

24/04/17 15:28:00 WARN Utils: Set SPARK_LOCAL_IP if you need to bind to another address
Setting default log level to "WARN".
To adjust logging level use sc.setLogLevel(newLevel). For SparkR, use setLogLevel(newLevel).
24/04/17 15:28:07 WARN NativeCodeLoader: Unable to load native-hadoop library for your platform... using builtin-java classes where applicable
Spark context Web UI available at http://10.11.5.92:4040
Spark context available as 'sc' (master = local[*], app id = local-1713347889284).
Spark session available as 'spark'.
Welcome to

  ____      __
 /    |     / 
|    |    /  |
|    |   /   |
|    |  /    |
|    | /     |
|    |/      |
|____/_____| version 3.3.1

Using Scala version 2.12.15 (OpenJDK 64-Bit Server VM, Java 11.0.22)
Type in expressions to have them evaluated.
Type :help for more information.

scala> val inputFile = sc.textFile("wordcount.txt")
inputFile: org.apache.spark.rdd.RDD[String] = wordcount.txt MapPartitionsRDD[1] at textFile at <console>:23

scala> val counts = inputFile.flatMap(line => line.split(" ")).map(word => (word, 1)).reduceByKey(_+_);
counts: org.apache.spark.rdd.RDD[(String, Int)] = ShuffledRDD[4] at reduceByKey at <console>:23

scala> counts.toDebugString
res0: String =
(2) ShuffledRDD[4] at reduceByKey at <console>:23 []
+- (2) MapPartitionsRDD[3] at map at <console>:23 []
    | MapPartitionsRDD[2] at flatMap at <console>:23 []
    | wordcount.txt MapPartitionsRDD[1] at textFile at <console>:23 []
    | wordcount.txt HadoopRDD[0] at textFile at <console>:23 []

scala> counts.cache()
res1: counts.type = ShuffledRDD[4] at reduceByKey at <console>:23

scala> counts.saveAsTextFile("wordcount_output")

scala>
```

```
root@student: /home/student/DSBDAL/wordcount_output

root@student:/home/student/DSBDAL# cd wordcount_output
root@student:/home/student/DSBDAL/wordcount_output# cat part-00000
(capabilities,1)
(provides,1)
(is,3)
(API,1)
(its,1)
(can,1)
(Unofficial,1)
(Java,,1)
(Clojure,,1)
(as,1)
(This,1)
(multiple,1)
(cluster,1)
(form,1)
(any,1)
(Scala,,2)
(computing,1)
(Python,,1)
(application,3)
(R,,1)
root@student:/home/student/DSBDAL/wordcount_output# cat part-00001
(Spark,3)
(interface,1)
(languages,,1)
(it,1)
(languages,,1)
(a,2)
(You,1)
(such,1)
(library,1)
(available,,1)
(to,1)
(available,2)
(written,1)
(in,5)
(develop,1)
(a,2)
```

```
root@student: /home/student/DSBDAL/wordcount_output
(any,1)
(Scala,,2)
(computing,1)
(Python,,1)
(application,3)
(R,,1)
root@student:/home/student/DSBDAL/wordcount_output# cat part-00001
(Spark,3)
(interface,1)
(languages.,1)
(it,1)
(languages,,1)
(a,2)
(You,1)
(such,1)
(library,1)
(available.,1)
(to,1)
(available,2)
(written,1)
(in,5)
(develop,1)
(of,2)
(also,1)
(for,1)
(makes,1)
(library.,1)
(an,2)
((API),1)
(programming,1)
(but,1)
(and,1)
(support,1)
(languages.,,1)
(these,1)
(additional,1)
(the,2)
root@student: /home/student/DSBDAL/wordcount_output#
```