SCALA :

// Importing SparkSession

**import** org.apache.spark.sql.SparkSession

// Creating SparkSession object

**val** sparkSession **=** SparkSession.builder()

                   .appName("My First Spark Application")

                   .master("local").getOrCreate()

// Loading sparkContext

**val** sparkContext **=** sparkSession.sparkContext

// Creating an RDD

**val** intArray **=** Array(1, 2, 3, 4, 5, 6, 7, 8, 9, 10)

// parallelize method creates partitions, which additionally

// takes integer argument to specifies the number of partitions.

// Here we are using 3 partitions.

**val** intRDD **=** sparkContext.parallelize(intArray, 3)

// Printing number of partitions

println(s"Number of partitions in intRDD : ${intRDD.partitions.size}")

// Printing first element of RDD

println(s"First element in intRDD : ${intRDD.first}")

// Creating string from RDD

// take(n) function is used to fetch n elements from

// RDD and returns an Array.

// Then we will convert the Array to string using

// mkString function in scala.

**val** strFromRDD **=** intRDD.take(intRDD.count.toInt).mkString(", ")

println(s"String from intRDD : ${strFromRDD}")

// Printing contents of RDD

// collect function is used to retrieve all the data in an RDD.

println("Printing intRDD: ")

intRDD.collect().foreach(println)

Output:

Number of partitions in intRDD : 3

First element in intRDD : 1

String from intRDD : 1, 2, 3, 4, 5, 6, 7, 8, 9, 10

Printing intRDD:

1

2

3

4

5

6

7

8

9

10