

PairRDDfunctions_2arg

```
// @Author : Maniram Yadav
```

```
print(sc)
```

Took: 1 second 82 milliseconds, at 2018-1-16 15:22

```
val rdd1 = sc.parallelize(Seq(
    ("key1",2),
    ("key2",1),
    ("key3",4),
    ("key4",7),
    ("key5",2)
))
```

```
val rdd2 = sc.parallelize(Seq(
    ("key1",3),
    ("key2",9),
    ("key3",10)
))
```

Took: 1 second 70 milliseconds, at 2018-1-16 15:22

```
// RDD Join Function
```

```
println("Join of RDD1 to RDD2 ")
rdd1.join(rdd2) foreach println
println("Join of RDD2 to RDD1 ")
rdd2.join(rdd1) foreach println
println("Left Outer Join of RDD1 to RDD2 ")
rdd1.leftOuterJoin(rdd2) foreach println
println("Right Outer Join of RDD1 to RDD2 ")
rdd1.rightOuterJoin(rdd1) foreach println
```

Took: 1 second 840 milliseconds, at 2018-1-16 15:22

```
// RDD cogroup method
```

```
rdd1.cogroup(rdd2) foreach println
```

Took: 1 second 282 milliseconds, at 2018-1-16 15:25

```
// Adding Cogrouped Key values
val rdd3 = rdd1.cogroup(rdd2)

val rdd4 = rdd3.map{ x =>
  {
    var key = x._1
    var values = x._2

    var sum1,sum2=0
    //values.foreach(v=> sum += v(0))
    //(key,sum)
    values._1.map(x=>sum1=x)
    values._2.map(x=>sum2=x)
    (key,sum1+sum2)
  }
}
rdd4 foreach println
```

Took: 1 second 239 milliseconds, at 2018-1-16 15:41

Build: | **buildTime**-Thu Dec 21 13:04:00 UTC 2017 | **formattedShaVersion**-0.7.0-743db69157361e7f476c80b9507060482add9383 | **sbtVersion**-0.13.8 | **scalaVersion**-2.11.7 | **sparkNotebookVersion**-0.7.0 | **hadoopVersion**-2.7.2 | **jets3tVersion**-0.7.1 | **jlineDef**-(jline,2.12) | **sparkVersion**-1.6.3 | **withHive**-false | **withParquet**-false |.