**Pair RDDs**

Comandos:

**ReduceByKey()** Combina valores con la misma clave

rdd.reduceByKey((x, y) => x + y)

**GroupByKey()** Agrupa valores con la misma clave

Rdd.groupByKey()

**CombineByKey()** combina valores con la misma clave utilizando un tipo de resultado diferente

Rdd.groupByKey()

**MapValues(func)** aplica una función a cada valor de un pair rdd sin cambiar la clave

Rdd.mapValues(x => x+1)

**flatMapValues(func)** aplica una función que devuelve un iterador a cada valor de un pair rdd, y para cada elemento devuelto, produce una entrada clave/valor con esa clave antigua.

Rdd.flatMapValue(x => (x to 5))

**Keys()** devuelve una rdd de sólo las claves

Rdd.keys()

**SubstractByKey()** elimina el elemento con clave presente en la otra rdd

Rdd.substractByKey(othe)

**Join** Realiza un inner join entre dos rdds

Rdd.join(other)

**RightOuterJoin** Realiza un inner join entre dos rdds donde la clave debe estar presente en la primera rdd

Rdd.rightOuterJoin(other)

**LeftOuterJoin** Realiza un inner join entre dos rdds donde la clave debe estar presente en la otrardd

Rdd.leftOuterJoin(other)

**Cogroup** Agrupa los datos de ambas rdds que comparten la misma clave

Rdd.group(other)

**Values()** devuelve una rdd de sólo los valores

Rdd.values()

**SortByKey()** devuelve una rdd ordenada por la clave

Rdd.sortByKey()

**CountByKey()** cuenta el número de elementos de cada clave

Rdd.countByKey()

**CollectAsMap()** recoge los resultados en forma de mapa para facilitar la búsqueda

Rdd.collectAsMap()

**Lookup(key)** devuelve todos los valores asociados a la clave proporcionada

Rdd.lookup(4)