

```
(1,Robin,22,25000,sales )
(2,BOB,23,30000,sales)
(3,Maya,23,25000,sales )
(4,Sara,25,40000,sales)
(5,David,23,45000,sales )
(6,Maggy,22,35000,sales)
```

```
(1,Robin,22,25000,sales )
(2,BOB,23,30000,sales)
(3,Maya,23,25000,sales )
(4,Sara,25,40000,sales)
(5,David,23,45000,sales )
(6,Maggy,22,35000,sales)
```

## Pig Operators: FLATTEN

ACADGILD

- Sometimes we have data in a bag or a tuple and you want to remove that level of nesting.
- Eg: baseball data.
- Because a player can play more than one position and position is stored in a bag.
- This allows us to still have one entry per player in the baseball file.
- But when you want to switch around your data on the fly and group by a particular position, you need a way to pull those entries out of the bag.
- To do this, Pig provides the flatten modifier in foreach:

In this example of an outer join, if the join key is missing from a table it is replaced by null.

```
grunt> cogroup_data = COGROUP emp_sales by age, emp_bonus by age;
grunt> flatten_data = foreach cogroup_data FLATTEN((IsEmpty(emp_sales) ? null : emp_sales)),FLATTEN((IsEmpty(emp_bonus) ? null : emp_bonus));
```

```
(6,Maggy,22,35000,sales,1,Robin,22,25000,sales )
(1,Robin,22,25000,sales ,1,Robin,22,25000,sales )
(5,David,23,45000,sales ,5,David,23,45000,sales )
(5,David,23,45000,sales ,3,Maya,23,25000,sales )
(5,David,23,45000,sales ,2,Jaya,23,20000,admin )
(3,Maya,23,25000,sales ,5,David,23,45000,sales )
(3,Maya,23,25000,sales ,3,Maya,23,25000,sales )
(3,Maya,23,25000,sales ,2,Jaya,23,20000,admin )
(2,BOB,23,30000,sales,5,David,23,45000,sales )
(2,BOB,23,30000,sales,3,Maya,23,25000,sales )
(2,BOB,23,30000,sales,2,Jaya,23,20000,admin )
(4,Sara,25,40000,sales,4,Alia,25,50000,admin )
(6,Omar,30,30000,admin)
```