

## ASSIGNMENT-7.1

Create a sample dataset and implement the below Pig commands on the same dataset.

### 1) Concat:

```
grunt> con1 = LOAD '/user/acadgild/set' using PigStorage(',') as (i1:chararray, i2:chararray, i3:chararray);
2017-05-15 13:53:39,201 [main] INFO org.apache.hadoop.conf.Configuration.deprecation - mapreduce.job.counters.limit is deprecated. Instead, use mapreduce.job.counters.max
2017-05-15 13:53:39,205 [main] INFO org.apache.hadoop.conf.Configuration.deprecation - io.bytes.per.checksum is deprecated. Instead, use dfs.bytes-per-checksum
2017-05-15 13:53:39,205 [main] INFO org.apache.hadoop.conf.Configuration.deprecation - fs.default.name is deprecated. Instead, use fs.defaultFS
```

```
grunt> dump con1;
2017-05-15 13:53:45,749 [main] INFO org.apache.pig.tools.pigstats.ScriptState - Pig features used in the script: UNKNOWN
2017-05-15 13:53:45,972 [main] INFO org.apache.hadoop.conf.Configuration.deprecation - fs.default.name is deprecated. Instead, use fs.defaultFS
2017-05-15 13:53:45,972 [main] INFO org.apache.hadoop.conf.Configuration.deprecation - mapreduce.job.counters.limit is deprecated. Instead, use mapreduce.job.counters.max
2017-05-15 13:53:45,976 [main] INFO org.apache.hadoop.conf.Configuration.deprecation - io.bytes.per.checksum is deprecated. Instead, use dfs.bytes-per-checksum
```

```
(John,Hannah,Baker)
(Peter,Robert,Clive)
(Scientist,Abdul,Kalam)
```

```
grunt> res = FOREACH con1 GENERATE CONCAT(i2,i3);
```

```
grunt> dump res;
2017-05-13 03:53:05,438 [main] INFO org.apache.pig.tools.pigstats.ScriptState - Pig features used in the script: UNKNOWN
2017-05-13 03:53:05,735 [main] INFO org.apache.hadoop.conf.Configuration.deprecation - fs.default.name is deprecated. Instead, use fs.defaultFS
2017-05-13 03:53:05,735 [main] INFO org.apache.hadoop.conf.Configuration.deprecation - mapreduce.job.counters.limit is deprecated. Instead, use mapreduce.job.counters.max
```

### OUTPUT:

```
(HannahBaker)
(RobertClive)
(AbdulKalam)
```

## 2) Tokenize:

```
grunt> tok = LOAD '/user/acadgild/sample' using PigStorage('\n') AS (j1:chararray);
2017-05-15 14:13:46,106 [main] INFO org.apache.hadoop.conf.Configuration.deprecation - mapreduce.job.counters.limit is deprecated. Instead, use mapreduce.job.counters.max
2017-05-15 14:13:46,110 [main] INFO org.apache.hadoop.conf.Configuration.deprecation - io.bytes.per.checksum is deprecated. Instead, use dfs.bytes-per-checksum
2017-05-15 14:13:46,110 [main] INFO org.apache.hadoop.conf.Configuration.deprecation - fs.default.name is deprecated. Instead, use fs.defaultFS
```

```
grunt> dump tok;
2017-05-15 14:13:58,168 [main] INFO org.apache.pig.tools.pigstats.ScriptState - Pig features used in the script: UNKNOWN
2017-05-15 14:13:58,452 [main] INFO org.apache.hadoop.conf.Configuration.deprecation - fs.default.name is deprecated. Instead, use fs.defaultFS
```

```
(This is a sample document.)
(This is an example document.)
(This is a model document.)
(This document is for reference.)
```

```
grunt> res1 = FOREACH tok GENERATE TOKENIZE(j1);
```

```
grunt> dump res1;
2017-05-15 14:15:58,760 [main] INFO org.apache.pig.tools.pigstats.ScriptState - Pig features used in the script: UNKNOWN
2017-05-15 14:15:59,061 [main] INFO org.apache.hadoop.conf.Configuration.deprecation - fs.default.name is deprecated. Instead, use fs.defaultFS
2017-05-15 14:15:59,061 [main] INFO org.apache.hadoop.conf.Configuration.deprecation - mapreduce.job.counters.limit is deprecated. Instead, use mapreduce.job.counters.max
```

## OUTPUT:

```
((This),(is),(a),(sample),(document.)))
((This),(is),(an),(example),(document.)))
((This),(is),(a),(model),(document.)))
((This),(document),(is),(for),(reference.)))
```

### 3) Sum:

```
grunt> ssum = LOAD '/user/acadgild/datas' using PigStorage(',') AS (name:chararray, num:int);
2017-05-15 14:36:52,450 [main] INFO org.apache.hadoop.conf.Configuration.deprecation - mapreduce.job.counters.limit is deprecated. Instead, use mapreduce.job.counters.max
2017-05-15 14:36:52,451 [main] INFO org.apache.hadoop.conf.Configuration.deprecation - io.bytes.per.checksum is deprecated. Instead, use dfs.bytes-per-checksum
2017-05-15 14:36:52,451 [main] INFO org.apache.hadoop.conf.Configuration.deprecation - fs.default.name is deprecated. Instead, use fs.defaultFS
```

```
grunt> dump ssum;
2017-05-15 14:37:25,379 [main] INFO org.apache.pig.tools.pigstats.ScriptState - Pig features used in the script: UNKNOWN
2017-05-15 14:37:25,531 [main] INFO org.apache.hadoop.conf.Configuration.deprecation - fs.default.name is deprecated. Instead, use fs.defaultFS
2017-05-15 14:37:25,531 [main] INFO org.apache.hadoop.conf.Configuration.deprecation - mapreduce.job.counters.limit is deprecated. Instead, use mapreduce.job.counters.max
2017-05-15 14:37:25,532 [main] INFO org.apache.hadoop.conf.Configuration.deprecation - io.bytes.per.checksum is deprecated. Instead, use dfs.bytes-per-checksum
2017-05-15 14:37:25,533 [main] INFO org.apache.pig.data.SchemaTupleBackend - Key [pig.schematuple] was not set... will not generate code.
```

```
(aaa,4)
(bbb,5)
(ccc,3)
(aaa,2)
(aaa,1)
(ccc,3)
(bbb,1)
```

```
grunt> grp = GROUP ssum BY name;
```

```
grunt> dump grp;
2017-05-15 14:40:39,102 [main] INFO org.apache.pig.tools.pigstats.ScriptState - Pig features used in the script: GROUP BY
2017-05-15 14:40:39,318 [main] INFO org.apache.hadoop.conf.Configuration.deprecation - fs.default.name is deprecated. Instead, use fs.defaultFS
2017-05-15 14:40:39,318 [main] INFO org.apache.hadoop.conf.Configuration.deprecation - mapreduce.job.counters.limit is deprecated. Instead, use mapreduce.job.counters.max
2017-05-15 14:40:39,324 [main] INFO org.apache.hadoop.conf.Configuration.deprecation - io.bytes.per.checksum is deprecated. Instead, use dfs.bytes-per-checksum
```

```
(aaa,{(aaa,1),(aaa,2),(aaa,4)})
(bbb,{(bbb,1),(bbb,5)})
(ccc,{(ccc,3),(ccc,3)})
```

```
grunt> res2 = FOREACH grp GENERATE group, SUM(ssum.num);
```

```
grunt> dump res2;
2017-05-15 14:45:16,446 [main] INFO org.apache.pig.tools.pigstats.ScriptState - Pig features used in the script: GROUP BY
2017-05-15 14:45:16,688 [main] INFO org.apache.hadoop.conf.Configuration.deprecation - fs.default.name is deprecated. Instead, use fs.defaultFS
2017-05-15 14:45:16,688 [main] INFO org.apache.hadoop.conf.Configuration.deprecation - mapreduce.job.counters.limit is deprecated. Instead, use mapreduce.job.counters.max
2017-05-15 14:45:16,692 [main] INFO org.apache.hadoop.conf.Configuration.deprecation - io.bytes.per.checksum is deprecated. Instead, use dfs.bytes-per-checksum
2017-05-15 14:45:16,698 [main] INFO org.apache.pig.data.SchemaTupleBackend - Key [pig.schematuple] was not set... will not generate code.
```

#### OUTPUT:

```
(aaa, 7)
(bbb, 6)
(ccc, 6)
```

#### 4) Min:

```
grunt> dump grp;
2017-05-15 14:40:39,102 [main] INFO org.apache.pig.tools.pigstats.ScriptState - Pig features used in the script: GROUP BY
2017-05-15 14:40:39,318 [main] INFO org.apache.hadoop.conf.Configuration.deprecation - fs.default.name is deprecated. Instead, use fs.defaultFS
2017-05-15 14:40:39,318 [main] INFO org.apache.hadoop.conf.Configuration.deprecation - mapreduce.job.counters.limit is deprecated. Instead, use mapreduce.job.counters.max
2017-05-15 14:40:39,324 [main] INFO org.apache.hadoop.conf.Configuration.deprecation - io.bytes.per.checksum is deprecated. Instead, use dfs.bytes-per-checksum
```

```
(aaa, {(aaa, 1), (aaa, 2), (aaa, 4)})
(bbb, {(bbb, 1), (bbb, 5)})
(ccc, {(ccc, 3), (ccc, 3)})
```

\\n\\n\\n

```
grunt> res3= FOREACH grp GENERATE group, MIN(ssum.num);
```

```
grunt> dump res3;
2017-05-15 14:50:18,766 [main] INFO org.apache.pig.tools.pigstats.ScriptState - Pig features used in the script: GROUP BY
2017-05-15 14:50:18,944 [main] INFO org.apache.hadoop.conf.Configuration.deprecation - fs.default.name is deprecated. Instead, use fs.defaultFS
2017-05-15 14:50:18,948 [main] INFO org.apache.hadoop.conf.Configuration.deprecation - mapreduce.job.counters.limit is deprecated. Instead, use mapreduce.job.counters.max
```

#### OUTPUT:

```
(aaa, 1)
(bbb, 1)
(ccc, 3)
```

#### 5) Max

```
grunt> dump grp;
2017-05-15 14:40:39,102 [main] INFO org.apache.pig.tools.pigstats.ScriptState - Pig features used in the script: GROUP_BY
2017-05-15 14:40:39,318 [main] INFO org.apache.hadoop.conf.Configuration.deprecation - fs.default.name is deprecated. Instead, use fs.defaultFS
2017-05-15 14:40:39,318 [main] INFO org.apache.hadoop.conf.Configuration.deprecation - mapreduce.job.counters.limit is deprecated. Instead, use mapreduce.job.counters.max
2017-05-15 14:40:39,324 [main] INFO org.apache.hadoop.conf.Configuration.deprecation - io.bytes.per.checksum is deprecated. Instead, use dfs.bytes-per-checksum
```

```
(aaa,{(aaa,1),(aaa,2),(aaa,4)})
(bbb,{(bbb,1),(bbb,5)})
(ccc,{(ccc,3),(ccc,3)})
```

```
grunt> res4= FOREACH grp GENERATE group, MAX(ssum.num);
```

```
grunt> dump res4;
2017-05-15 14:53:44,421 [main] INFO org.apache.pig.tools.pigstats.ScriptState - Pig features used in the script: GROUP_BY
2017-05-15 14:53:44,571 [main] INFO org.apache.hadoop.conf.Configuration.deprecation - fs.default.name is deprecated. Instead, use fs.defaultFS
2017-05-15 14:53:44,571 [main] INFO org.apache.hadoop.conf.Configuration.deprecation - mapreduce.job.counters.limit is deprecated. Instead, use mapreduce.job.counters.max
2017-05-15 14:53:44,571 [main] INFO org.apache.hadoop.conf.Configuration.deprecation - io.bytes.per.checksum is deprecated. Instead, use dfs.bytes-per-checksum
```

## OUTPUT:

```
(aaa,4)
(bbb,5)
(ccc,3)
```

## 6) Limit:

```

grunt> dump ssum;
2017-05-15 14:37:25,379 [main] INFO org.apache.pig.tools.pigstats.ScriptState - Pig features used in the script: UNKNOWN
2017-05-15 14:37:25,531 [main] INFO org.apache.hadoop.conf.Configuration.deprecation - fs.default.name is deprecated. Instead, use fs.defaultFS
2017-05-15 14:37:25,531 [main] INFO org.apache.hadoop.conf.Configuration.deprecation - mapreduce.job.counters.limit is deprecated. Instead, use mapreduce.job.counters.max
2017-05-15 14:37:25,532 [main] INFO org.apache.hadoop.conf.Configuration.deprecation - io.bytes.per.checksum is deprecated. Instead, use dfs.bytes-per-checksum
2017-05-15 14:37:25,533 [main] INFO org.apache.pig.data.SchemaTupleBackend - Key [pig.schematuple] was not set... will not generate code.

```

```

(aaa,4)
(bbb,5)
(ccc,3)
(aaa,2)
(aaa,1)
(ccc,3)
(bbb,1)

```

```

grunt> res5= LIMIT ssum 3;

```

```

grunt> dump res5;
2017-05-15 14:58:50,628 [main] INFO org.apache.pig.tools.pigstats.ScriptState - Pig features used in the script: LIMIT
2017-05-15 14:58:50,868 [main] INFO org.apache.hadoop.conf.Configuration.deprecation - fs.default.name is deprecated. Instead, use fs.defaultFS
2017-05-15 14:58:50,868 [main] INFO org.apache.hadoop.conf.Configuration.deprecation - mapreduce.job.counters.limit is deprecated. Instead, use mapreduce.job.counters.max
2017-05-15 14:58:50,872 [main] INFO org.apache.hadoop.conf.Configuration.deprecation - io.bytes.per.checksum is deprecated. Instead, use dfs.bytes-per-checksum
2017-05-15 14:58:50,876 [main] INFO org.apache.pig.data.SchemaTupleBackend - Key [pig.schematuple] was not set... will not generate code.
2017-05-15 14:58:50,887 [main] INFO org.apache.pig.newplan.logical.optimizer.LogicalPlanOptimizer - {RULES_ENABLED=[AddForEach, ColumnMapKeyPrune, ConstantCalculator, GroupByConstParallelSetter, LimitOptimizer, LoadTypeCastInserter, MergeFilter, MergeForEach, PartitionFilterOptimizer, PredicatePushdownOptimizer, PushDownForEachFlatten, PushUpFilter, SplitFilter, StreamTypeCastInserter]}

```

## OUTPUT:

```

(aaa,4)
(bbb,5)
(ccc,3)

```

## 7) Store:

```
grunt> st = LOAD '/user/acadgild/edata' using PigStorage(',') AS (a1:int,a2:int,a3:int);
2017-05-15 15:05:45,269 [main] INFO org.apache.hadoop.conf.Configuration.deprecation - mapreduce.job.counters.limit is deprecated. Instead, use mapreduce.job.counters.max
2017-05-15 15:05:45,269 [main] INFO org.apache.hadoop.conf.Configuration.deprecation - io.bytes.per.checksum is deprecated. Instead, use dfs.bytes-per-checksum
2017-05-15 15:05:45,270 [main] INFO org.apache.hadoop.conf.Configuration.deprecation - fs.default.name is deprecated. Instead, use fs.defaultFS
```

```
grunt> dump st;
2017-05-15 15:01:12,501 [main] INFO org.apache.pig.tools.pigstats.ScriptState - Pig features used in the script: UNKNOWN
2017-05-15 15:01:12,910 [main] INFO org.apache.hadoop.conf.Configuration.deprecation - fs.default.name is deprecated. Instead, use fs.defaultFS
2017-05-15 15:01:12,910 [main] INFO org.apache.hadoop.conf.Configuration.deprecation - mapreduce.job.counters.limit is deprecated. Instead, use mapreduce.job.counters.max
2017-05-15 15:01:12,918 [main] INFO org.apache.hadoop.conf.Configuration.deprecation - io.bytes.per.checksum is deprecated. Instead, use dfs.bytes-per-checksum
2017-05-15 15:01:12,957 [main] INFO org.apache.pig.data.SchemaTupleBackend - Key [pig.schematuple] was not set... will not generate code.
```

```
(8,3,4)
(1,2,3)
(4,3,3)
(4,3,3)
(1,2,3)
```

```
grunt> STORE st INTO 'output' USING PigStorage('*');
2017-05-15 15:13:29,005 [main] INFO org.apache.hadoop.conf.Configuration.deprecation - mapreduce.job.counters.limit is deprecated. Instead, use mapreduce.job.counters.max
2017-05-15 15:13:29,008 [main] INFO org.apache.hadoop.conf.Configuration.deprecation - io.bytes.per.checksum is deprecated. Instead, use dfs.bytes-per-checksum
2017-05-15 15:13:29,008 [main] INFO org.apache.hadoop.conf.Configuration.deprecation - fs.default.name is deprecated. Instead, use fs.defaultFS
2017-05-15 15:13:29,337 [main] INFO org.apache.hadoop.conf.Configuration.deprecation - mapred.textoutputformat.separator is deprecated. Instead, use mapreduce.output.textoutputformat.separator
```

## OUTPUT

```
grunt> CAT output;
8*3*4
1*2*3
4*3*3
4*3*3
1*2*3
```

## 8) Distinct:

```
(8,3,4)
(1,2,3)
(4,3,3)
(4,3,3)
(1,2,3)
```

```
grunt> res6 = DISTINCT st;
2017-05-15 15:16:30,744 [main] INFO org.apache.hadoop.conf.Configuration.deprecation - fs.default.name is deprecated. Instead, use fs.defaultFS
2017-05-15 15:16:30,744 [main] INFO org.apache.hadoop.conf.Configuration.deprecation - mapreduce.job.counters.limit is deprecated. Instead, use mapreduce.job.counters.max
2017-05-15 15:16:30,745 [main] INFO org.apache.hadoop.conf.Configuration.deprecation - io.bytes.per.checksum is deprecated. Instead, use dfs.bytes-per-checksum
```

```
grunt> dump res6;
2017-05-15 15:16:41,451 [main] INFO org.apache.pig.tools.pigstats.ScriptState - Pig features used in the script: DISTINCT
2017-05-15 15:16:41,629 [main] INFO org.apache.hadoop.conf.Configuration.deprecation - fs.default.name is deprecated. Instead, use fs.defaultFS
```

## OUTPUT

```
Process : 1
(1,2,3)
(4,3,3)
(8,3,4)
```

### 9) Flatten:

```
grunt> a1 = load '/user/acadgild/test' using PigStorage(',') as (a:int, b:int);
2017-05-15 15:25:29,749 [main] INFO org.apache.hadoop.conf.Configuration.deprecation - mapreduce.job.counters.limit is deprecated. Instead, use mapreduce.job.counters.max
2017-05-15 15:25:29,749 [main] INFO org.apache.hadoop.conf.Configuration.deprecation - io.bytes.per.checksum is deprecated. Instead, use dfs.bytes-per-checksum
2017-05-15 15:25:29,750 [main] INFO org.apache.hadoop.conf.Configuration.deprecation - fs.default.name is deprecated. Instead, use fs.defaultFS
```



```
grunt> dump a1;
2017-05-15 15:25:36,855 [main] INFO org.apache.pig.tools.pigstats.ScriptState - Pig features used in the script: UNKNOWN
2017-05-15 15:25:37,515 [main] INFO org.apache.hadoop.conf.Configuration.deprecation - fs.default.name is deprecated. Instead, use fs.defaultFS
2017-05-15 15:25:37,515 [main] INFO org.apache.hadoop.conf.Configuration.deprecation - mapreduce.job.counters.limit is deprecated. Instead, use mapreduce.job.counters.max
2017-05-15 15:25:37,516 [main] INFO org.apache.hadoop.conf.Configuration.deprecation - io.bytes.per.checksum is deprecated. Instead, use dfs.bytes-per-checksum
2017-05-15 15:25:37,518 [main] INFO org.apache.pig.data.SchemaTupleBackend - Key [pig.schematuple] was not set... will not generate code.
```

```
(1,2)
(1,3)
(1,4)
(2,5)
(2,6)
(2,7)
```

```
grunt> a2 = group a1 by a;
grunt> a3 = foreach a2 generate group as a, a1 as b;
grunt> a4 = foreach a3 generate a, FLATTEN(BagToTuple(b.b));
```

```
grunt> dump a4;
2017-05-15 15:29:00,554 [main] INFO org.apache.pig.tools.pigstats.ScriptState - Pig features used in the script: GROUP_BY
2017-05-15 15:29:00,679 [main] INFO org.apache.hadoop.conf.Configuration.deprecation - fs.default.name is deprecated. Instead, use fs.defaultFS
2017-05-15 15:29:00,680 [main] INFO org.apache.hadoop.conf.Configuration.deprecation - mapreduce.job.counters.limit is deprecated. Instead, use mapreduce.job.counters.max
2017-05-15 15:29:00,681 [main] INFO org.apache.hadoop.conf.Configuration.deprecation - io.bytes.per.checksum is deprecated. Instead, use dfs.bytes-per-checksum
```

## OUTPUT

```
(1,4,3,2)
(2,7,6,5)
```

## 10) isEmpty:

```
grunt> emp_sales1 = LOAD '/user/acadgild/emp_sales' USING PigStorage(',') as (sno:int, name:chararray, age:int, salary:int, dept:chararray);
```

```
2017-05-15 15:32:23,386 [main] INFO org.apache.hadoop.conf.Configuration.deprecation - mapreduce.job.counters.limit is deprecated. Instead, use mapreduce.job.counters.max
```

```
2017-05-15 15:32:23,387 [main] INFO org.apache.hadoop.conf.Configuration.deprecation - io.bytes.per.checksum is deprecated. Instead, use dfs.bytes-per-checksum
```

```
2017-05-15 15:32:23,387 [main] INFO org.apache.hadoop.conf.Configuration.deprecation - fs.default.name is deprecated. Instead, use fs.defaultFS
```

```
grunt> emp_bonus1 = LOAD '/user/acadgild/emp_bonus' USING PigStorage(',') as (sno:int, name:chararray, age:int, salary:int, dept:chararray);
```

```
2017-05-15 15:32:43,653 [main] INFO org.apache.hadoop.conf.Configuration.deprecation - mapreduce.job.counters.limit is deprecated. Instead, use mapreduce.job.counters.max
```

```
2017-05-15 15:32:43,730 [main] INFO org.apache.hadoop.conf.Configuration.deprecation - io.bytes.per.checksum is deprecated. Instead, use dfs.bytes-per-checksum
```

```
2017-05-15 15:32:43,730 [main] INFO org.apache.hadoop.conf.Configuration.deprecation - fs.default.name is deprecated. Instead, use fs.defaultFS
```

```
grunt> dump emp_sales1;
```

```
2017-05-15 15:33:00,836 [main] INFO org.apache.pig.tools.pigstats.ScriptState - Pig features used in the script: UNKNOWN
```

```
2017-05-15 15:33:00,950 [main] INFO org.apache.hadoop.conf.Configuration.deprecation - fs.default.name is deprecated. Instead, use fs.defaultFS
```

```
2017-05-15 15:33:00,950 [main] INFO org.apache.hadoop.conf.Configuration.deprecation - mapreduce.job.counters.limit is deprecated. Instead, use mapreduce.job.counters.max
```

```
2017-05-15 15:33:00,952 [main] INFO org.apache.hadoop.conf.Configuration.deprecation - io.bytes.per.checksum is deprecated. Instead, use dfs.bytes-per-checksum
```

```
2017-05-15 15:33:00,954 [main] INFO org.apache.pig.data.SchemaTupleBackend - Key [pig.schematuple] was not set... will not generate code.
```

```
(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)
```

```
grunt> dump emp_bonus1;
```

```
2017-05-15 15:35:33,469 [main] INFO org.apache.pig.tools.pigstats.ScriptState - Pig features used in the script: UNKNOWN
```

```
2017-05-15 15:35:33,590 [main] INFO org.apache.hadoop.conf.Configuration.deprecation - fs.default.name is deprecated. Instead, use fs.defaultFS
```

```
2017-05-15 15:35:33,590 [main] INFO org.apache.hadoop.conf.Configuration.deprecation - mapreduce.job.counters.limit is deprecated. Instead, use mapreduce.job.counters.max
```

```
2017-05-15 15:35:33,591 [main] INFO org.apache.hadoop.conf.Configuration.deprecation - io.bytes.per.checksum is deprecated. Instead, use dfs.bytes-per-checksum
```

```
2017-05-15 15:35:33,593 [main] INFO org.apache.pig.data.SchemaTupleBackend - Key [pig.schematuple] was not set... will not generate code.
```

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

```
grunt> group_data = COGROUP emp_sales1 by age, emp_bonus1 by age;
grunt> res7= filter group_data by IsEmpty(emp_sales1);
grunt> dump res7;
```

```
2017-05-15 15:39:45,982 [main] INFO org.apache.pig.tools.pigstats.ScriptState - Pig features used in the script: COGROUP,FILTER
```

```
2017-05-15 15:39:46,201 [main] INFO org.apache.hadoop.conf.Configuration.deprecation - fs.default.name is deprecated. Instead, use fs.defaultFS
```

```
2017-05-15 15:39:46,202 [main] INFO org.apache.hadoop.conf.Configuration.deprecation - mapreduce.job.counters.limit is deprecated. Instead, use mapreduce.job.counters.max
```

```
2017-05-15 15:39:46,223 [main] INFO org.apache.hadoop.conf.Configuration.deprecation - io.bytes.per.checksum is deprecated. Instead, use dfs.bytes-per-checksum
```

## OUTPUT

```
((30,{}),{(6,Omar,30,30000,admin)}))
```