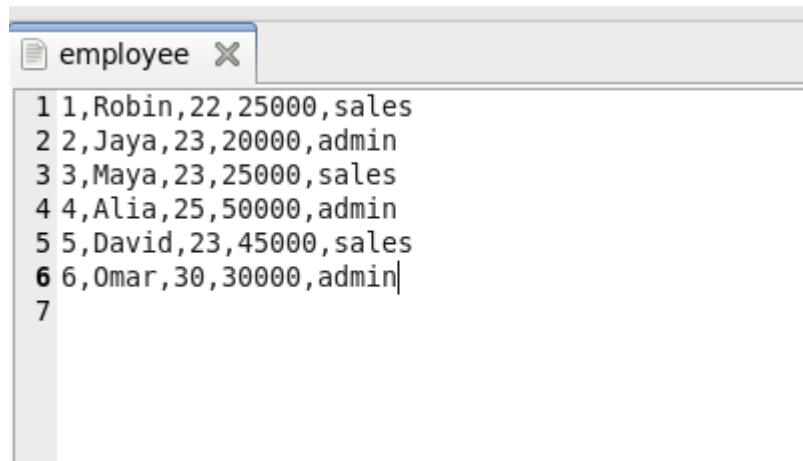# Assignment13.2

load an input dataset from Pig to Hbase table

input dataset:

```
employee  ✕

1 1,Robin,22,25000,sales
2 2,Jaya,23,20000,admin
3 3,Maya,23,25000,sales
4 4,Alia,25,50000,admin
5 5,David,23,45000,sales
6 6,Omar,30,30000,admin
7
```

Creating a table employee in Hbase

```
hbase(main):001:0> create 'employee', 'emp_info'
SLF4J: Class path contains multiple SLF4J bindings.
SLF4J: Found binding in [jar:file:/usr/local/hbase/lib/slf4j-log4j12-1.6.4.jar!/
org/slf4j/impl/StaticLoggerBinder.class]
SLF4J: Found binding in [jar:file:/usr/local/hadoop-2.6.0/share/hadoop/common/li
b/slf4j-log4j12-1.7.5.jar!/org/slf4j/impl/StaticLoggerBinder.class]
SLF4J: See http://www.slf4j.org/codes.html#multiple_bindings for an explanation.
2017-05-28 14:05:10,691 WARN  [main] util.NativeCodeLoader: Unable to load nativ
e-hadoop library for your platform... using builtin-java classes where applicabl
e
0 row(s) in 10.7180 seconds

=> Hbase::Table - employee
```

listing the tables in hbase

```
hbase(main):002:0> list
TABLE
Acadgild
Details
PersonDetails
customer
employee
htest
song-artist-map
station-geo-map
subscribed-users
9 row(s) in 0.1340 seconds
```

exporting the required jar files

```
[acadgild@localhost ~]$ export PIG_CLASSPATH=/usr/local/hbase/hbase-client-0.98.
14-hadoop2.jar:/usr/local/hbase/hbase-server-0.98.14-hadoop2.jar:/usr/local/hbas
e/zookeeper-3.4.6-.jar;
[acadgild@localhost ~]$
```

loading the data using pig

```
grunt> emp = LOAD '/user/acadgild/employee' USING PigStorage(',') as (sno:int, name:chararray, age:int, salary:int, dept:char
array);
2017-05-28 14:30:53,885 [main] INFO  org.apache.hadoop.conf.Configuration.deprecation - mapred.job.tracker.persist.jobstatus.
hours is deprecated. Instead, use mapreduce.jobtracker.persist.jobstatus.hours
```

storing the data into hbase table

```
grunt> store emp INTO 'hbase://employee' USING org.apache.pig.backend.hadoop.hbase.HBaseStorage('emp_info:sno, emp_info:name,
emp_info:age, emp_info:salary, emp_info:dept');
```

output:

```
hbase(main):065:0> scan 'employee'
ROW                      COLUMN+CELL
 1                       column=emp_info:age, timestamp=1495971028986, value=22
 1                       column=emp_info:dept, timestamp=1495971031243, value=sales
 1                       column=emp_info:name, timestamp=1495971028558, value=Robin
 1                       column=emp_info:salary, timestamp=1495971029265, value=25000
 1                       column=emp_info:sno, timestamp=1495971028136, value=1
 2                       column=emp_info:age, timestamp=1495971063704, value=23
 2                       column=emp_info:dept, timestamp=1495971064239, value=admin
 2                       column=emp_info:name, timestamp=1495971063511, value=Jaya
 2                       column=emp_info:salary, timestamp=1495971063889, value=20000
 2                       column=emp_info:sno, timestamp=1495971063238, value=2
 3                       column=emp_info:age, timestamp=1495971065054, value=23
 3                       column=emp_info:dept, timestamp=1495971065568, value=sales
 3                       column=emp_info:name, timestamp=1495971064798, value=Maya
 3                       column=emp_info:salary, timestamp=1495971065310, value=25000
 3                       column=emp_info:sno, timestamp=1495971064453, value=3
 4                       column=emp_info:age, timestamp=1495971066146, value=25
 4                       column=emp_info:dept, timestamp=1495971066596, value=admin
 4                       column=emp_info:name, timestamp=1495971065988, value=Alia
 4                       column=emp_info:salary, timestamp=1495971066369, value=50000
 4                       column=emp_info:sno, timestamp=1495971065805, value=4
 5                       column=emp_info:age, timestamp=1495971067249, value=23
 5                       column=emp_info:dept, timestamp=1495971067701, value=sales
 5                       column=emp_info:name, timestamp=1495971067055, value=David
 5                       column=emp_info:salary, timestamp=1495971067485, value=45000
 5                       column=emp_info:sno, timestamp=1495971066798, value=5
 6                       column=emp_info:age, timestamp=1495971068417, value=30
 6                       column=emp_info:dept, timestamp=1495971068776, value=admin
 6                       column=emp_info:name, timestamp=1495971068253, value=Omar
 6                       column=emp_info:salary, timestamp=1495971068614, value=30000
 6                       column=emp_info:sno, timestamp=1495971067996, value=6
6 row(s) in 0.5000 seconds
```