

```
[cloudera@quickstart ~]$ hadoop fs -mkdir raaghu
[cloudera@quickstart ~]$ hadoop fs -ls
[cloudera@quickstart ~]$ hadoop fs -copyFromLocal
/home/cloudera/Desktop raaghu/dept.csv
[cloudera@quickstart ~]$ hadoop fs -copyFromLocal
/home/cloudera/Desktop raaghu/emp.csv
[cloudera@quickstart ~]$ hadoop fs -copyFromLocal
/home/cloudera/Desktop raaghu/wordcount.txt
[cloudera@quickstart ~]$ hadoop fs -ls raaghu
[cloudera@quickstart ~]$ pig
```

PIG

```
grunt> cd raaghu
grunt> pwd
grunt> A = load '/user/cloudera/raaghu/emp.csv' using
PigStorage(',') as (eid:int, ename:chararray, esal:int, ecom:int,
edpno:int);
grunt> dump A;
grunt> B = filter A by ecom==30;
grunt> dump B;
grunt> C = filter A by esal>=1500;
grunt> dump C;
grunt> D = filter A by esal==30 and ename=="SMITH";
grunt> dump D;
grunt> store A into '/user/cloudera/raaghu/Pigout1' using
PigStorage(',');
grunt> cd Pigout1
grunt> ls
grunt> cat part-m-00000
grunt> E = foreach A generate esal;
grunt> dump E;
grunt> F = foreach A generate *, ecome*2 as bonus, esal*5 as
incentive;
grunt> dump F;
grunt> G = foreach A generate SUBSTRING(ename,0,4);
grunt> dump G;
grunt> H=foreach A generate $0,$1,$2;
grunt> dump H;
grunt> I = group A by edpno;
grunt> dump I;
WORD COUNT
grunt> lines = load '/user/cloudera/wordcount.txt' as
(line:chararray);
grunt> dump lines;
grunt> Token = foreach lines generate TOKENIZE(line);
grunt> dump Token;
grunt> flats = foreach lines generate FLATTEN(line);
grunt> dump flats;
grunt> group_words = group flats by $0;
grunt> dump group_words;
grunt> count_words = foreach group_words generate grouo as word,
COUNT($1) as word_occurance;
grunt> dump count_words;
```

Hive

```
[cloudera@quickstart ~]$ hive
hive> show databases;
hive> create database raghu;
hive> create table emp(eid int, ename string, epos string, esal
int, ecom int, edpno int) row format delimited fields terminated
by ',';
hive> load data local inpath '/home/cloudera/Desktop/emp.csv' into
table emp;
hive> select * from emp;
hive> create table dept (edpno int, epos string, ecity string) row
format delimited fields terminated by ',';
OK
Time taken: 0.706 seconds
hive> load data local inpath '/home/cloudera/Desktop/dept.csv'
into table dept;
select * from dept;
```