SESSION 5 ASSIGNMENT 2

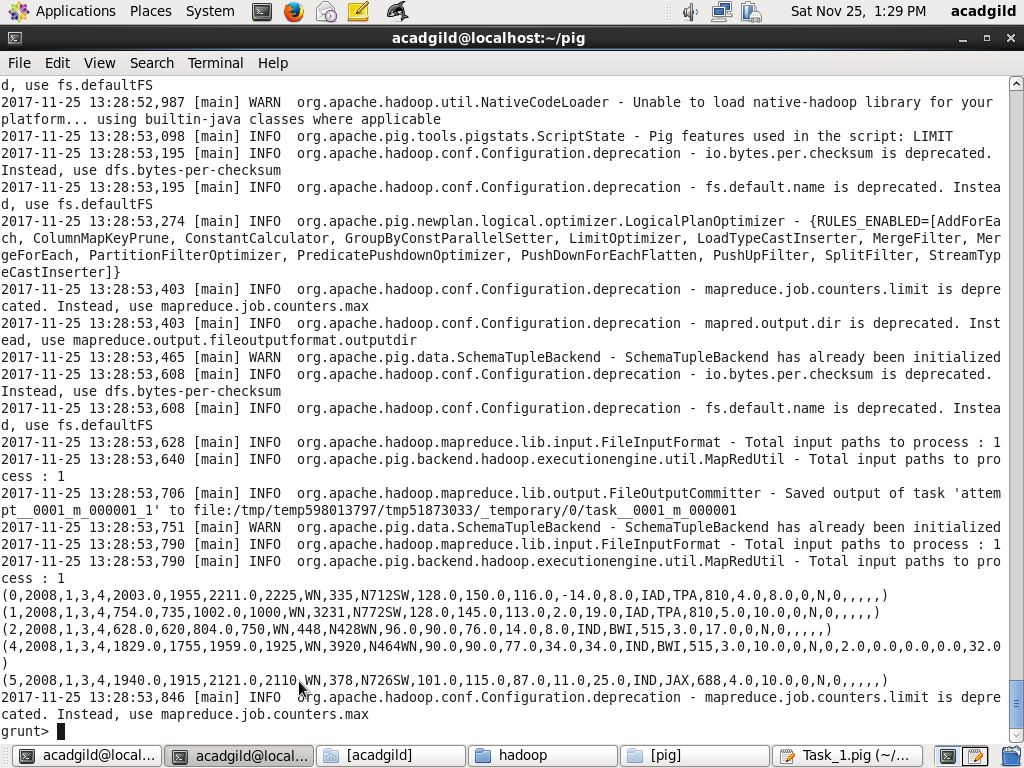
## PROBLEM 1

### Q1.Find out the top 5 most visited destinations.

REGISTER '/usr/local/pig/lib/piggybank.jar';

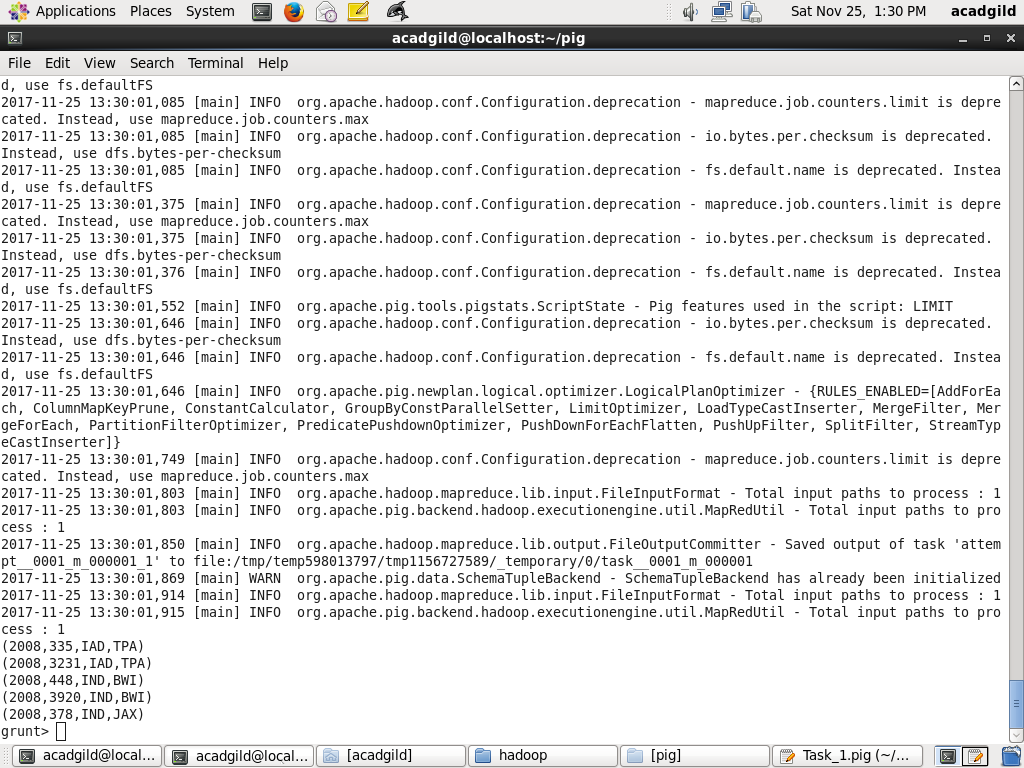
A = load 'DelayedFlights.csv' USING org.apache.pig.piggybank.storage.CSVExcelStorage(',','NO\_MULTILINE','UNIX','SKIP\_INPUT\_HEADER');

A\_LIMIT = LIMIT A 5;



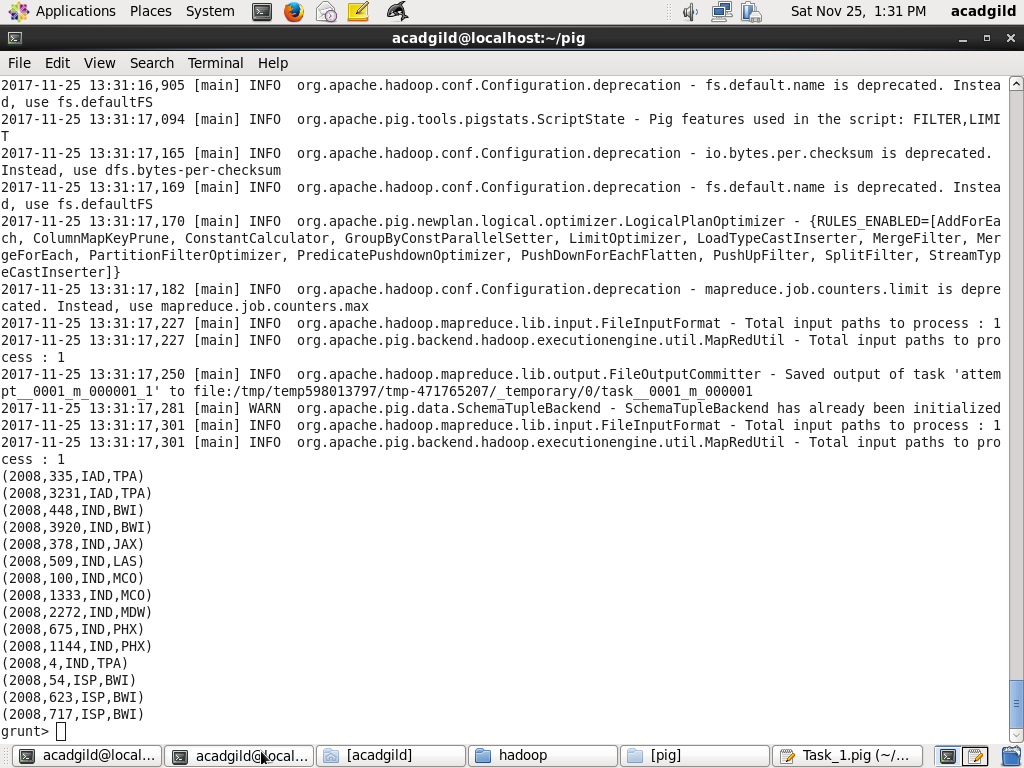
B = foreach A generate (int)$1 as year, (int)$10 as flight\_num, (chararray)$17 as origin,(chararray) $18 as dest;

B\_LIMIT = LIMIT B 5;



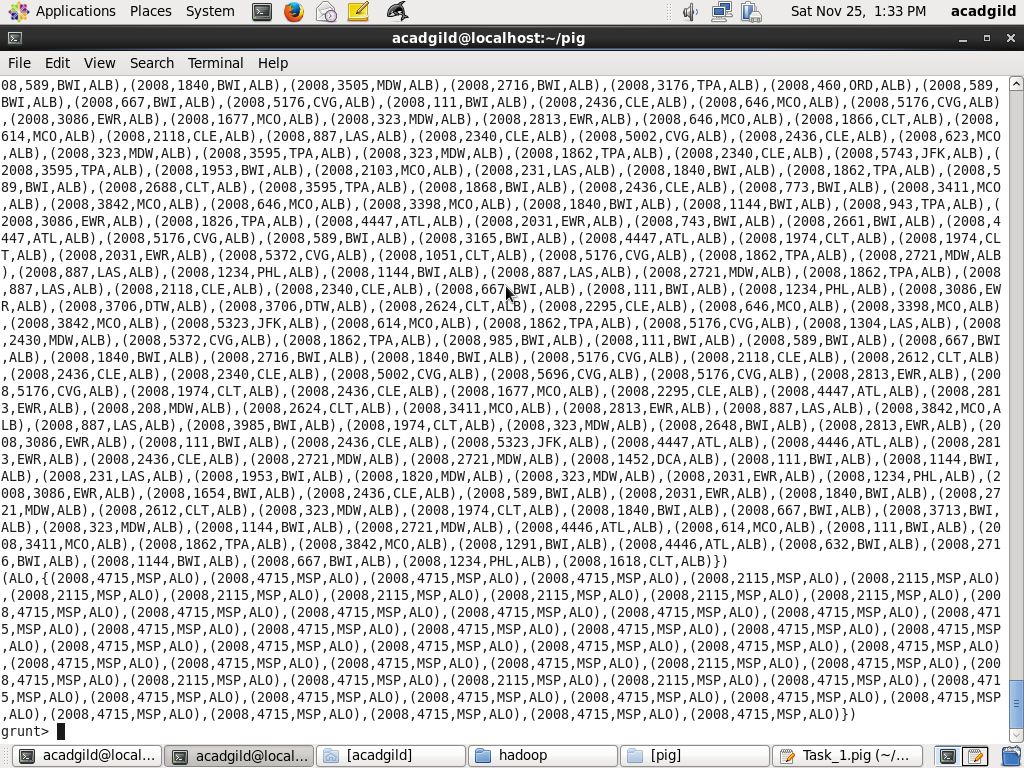
C = filter B by dest is not null;

C\_LIMIT = LIMIT C 15;



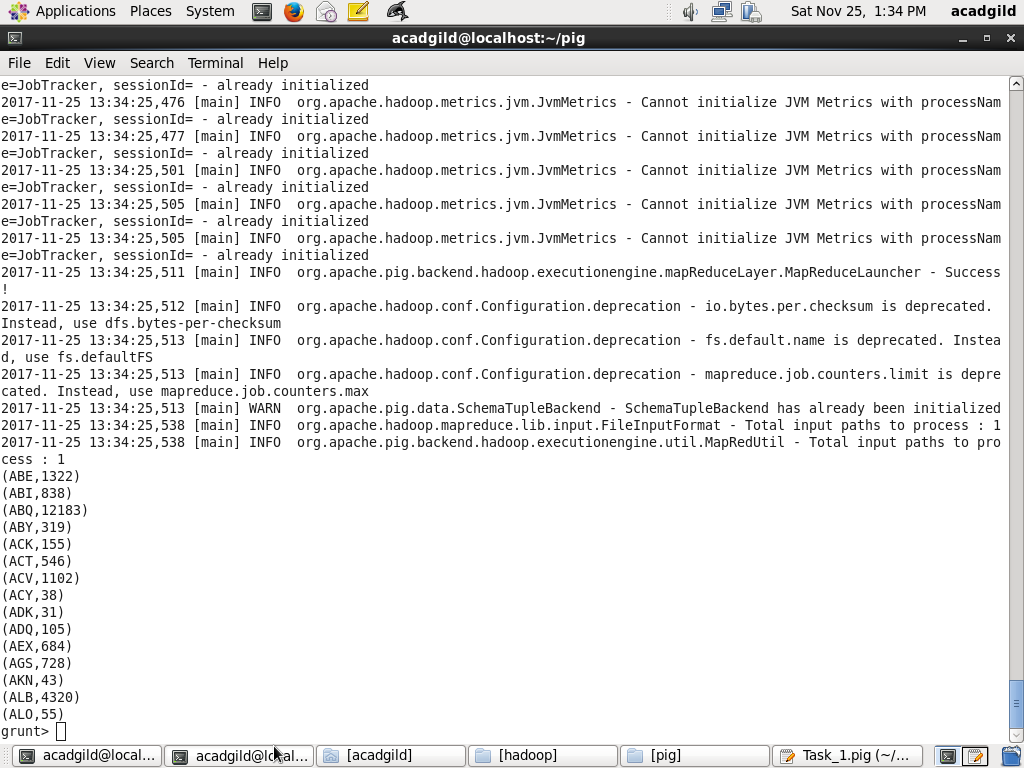
D = group C by dest;

D\_LIMIT = LIMIT D 15;



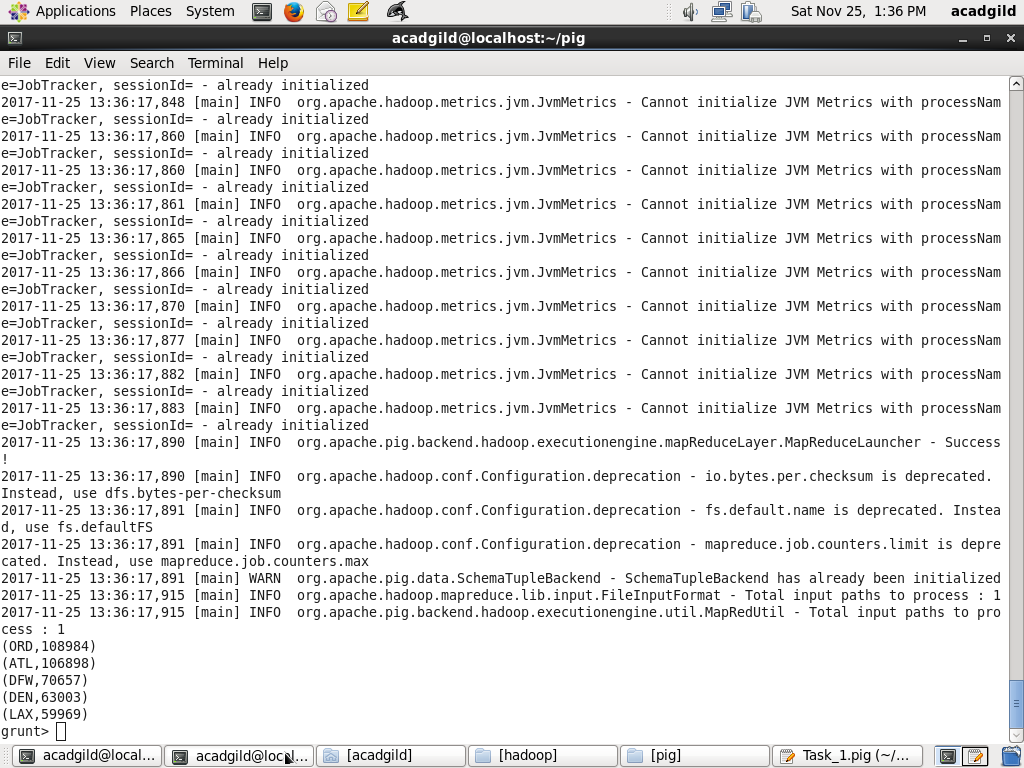
E = foreach D generate group, COUNT(C.dest);

E\_LIMIT = LIMIT E 15;



F = order E by $1 DESC;

Result = LIMIT F 5;



A1 = load 'airports.csv' USING org.apache.pig.piggybank.storage.CSVExcelStorage(',','NO\_MULTILINE','UNIX','SKIP\_INPUT\_HEADER');

A2 = foreach A1 generate (chararray)$0 as dest, (chararray)$2 as city, (chararray)$4 as country;

joined\_table = join Result by $0, A2 by dest;

dump joined\_table;

