Below is the script used to find top5 destinations-

**REGISTER '/home/acadgild/airline\_usecase/piggybank.jar';**

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

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

**C = filter B by dest is not null;**

**D = group C by dest;**

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

**F = order E by $1 DESC;**

**Result = LIMIT F 5;**

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

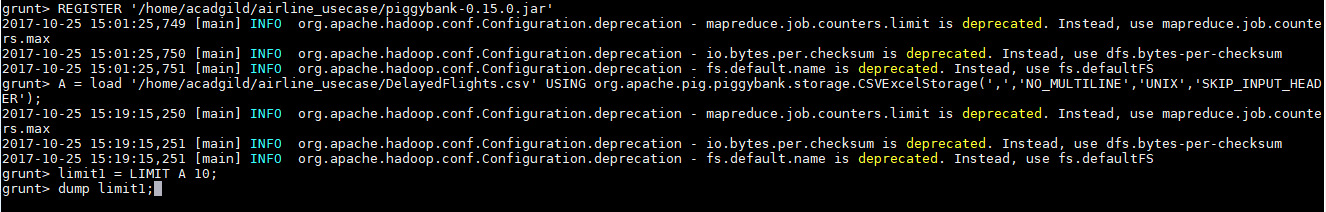
**A2 = foreach A1 generate (chararray)$0 as dest, (chararray)$2 as city, (chararray)$4 as country;**

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

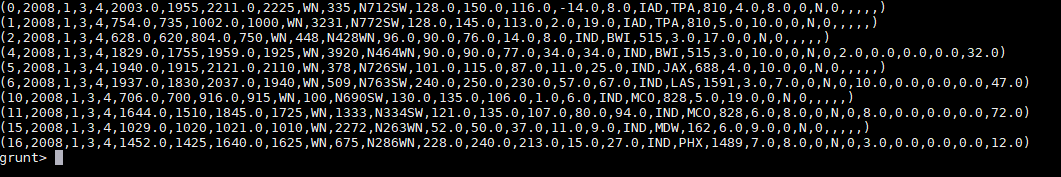
**dump joined\_table;**

Each and every script has been explained below using intermediate outputs-

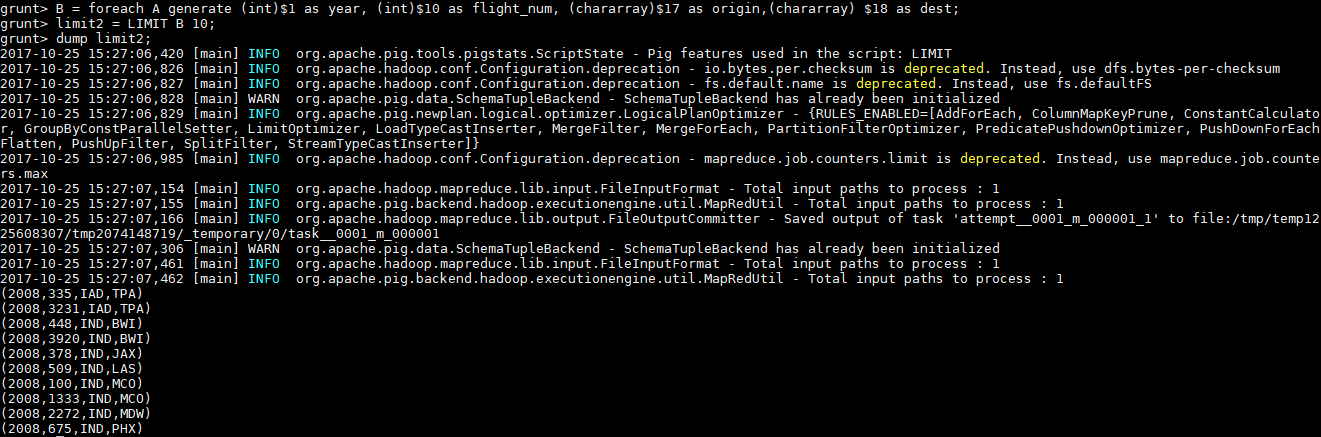
Registering piggybank jar in order to use the CSVExcelStorage class-



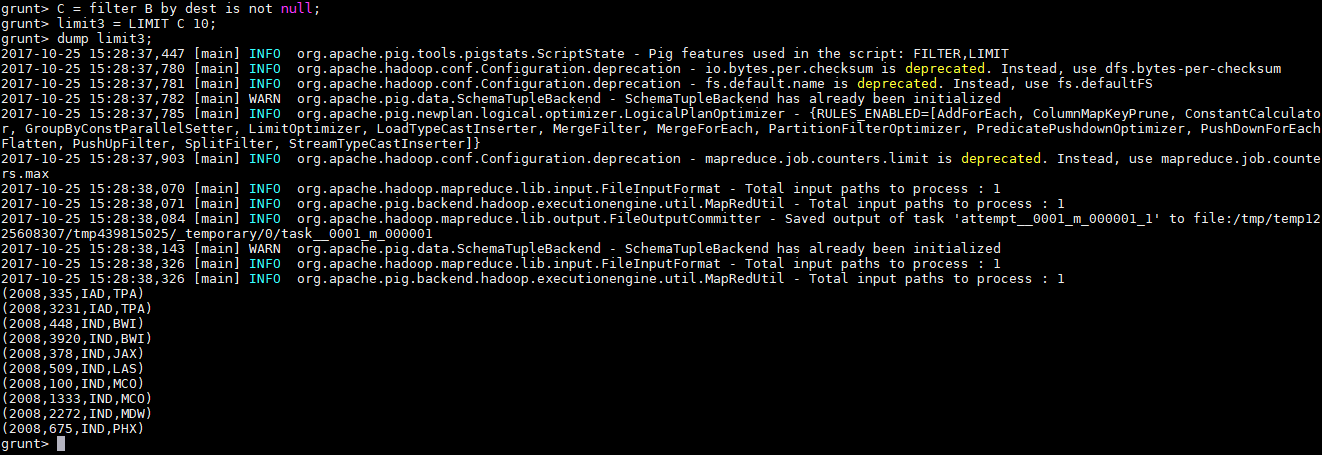
Contents of file-



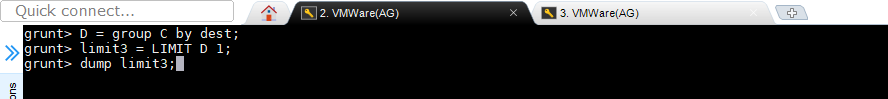
Extracting columns for year, flight number, origin and destination-

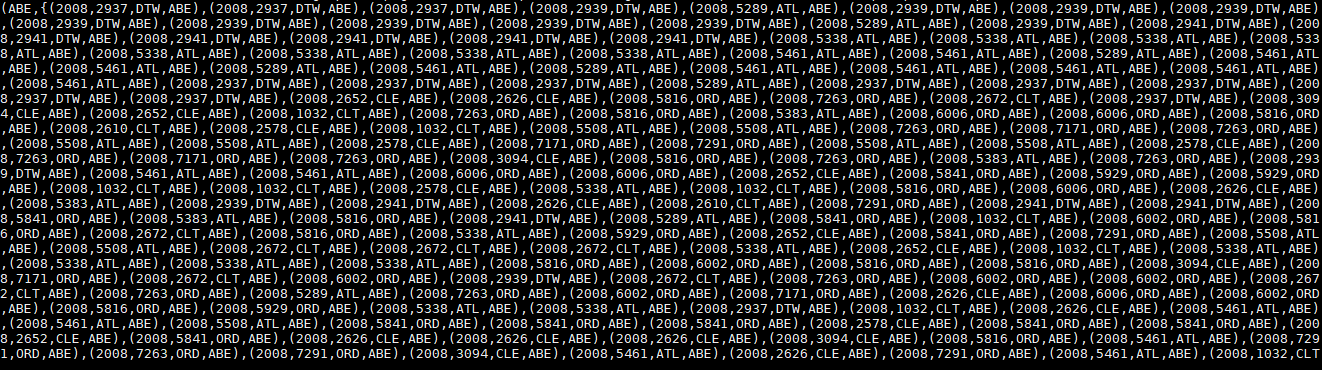


Filter records which have destination as NULL-

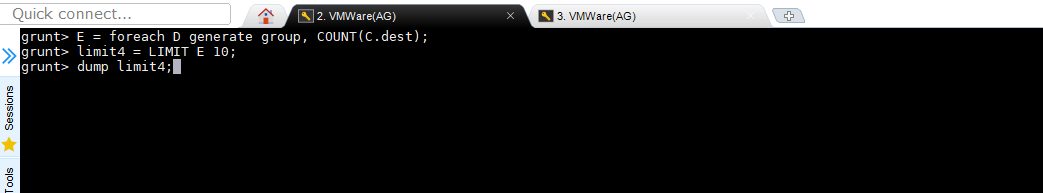


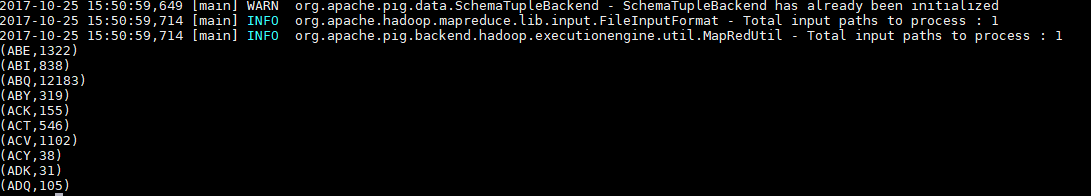
Group records by destination-



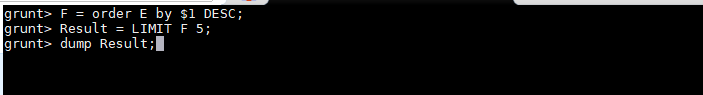


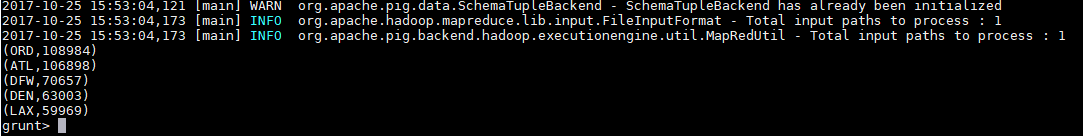
Taking number of count for each destination to find the frequency of destination visited-



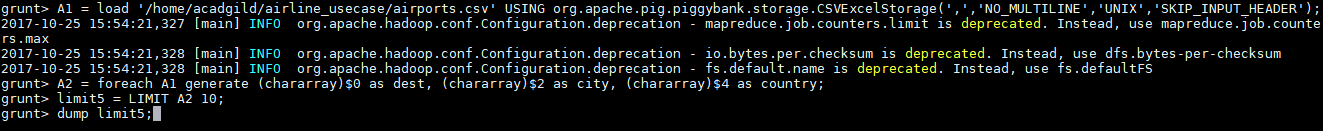


Sorting the count for destination

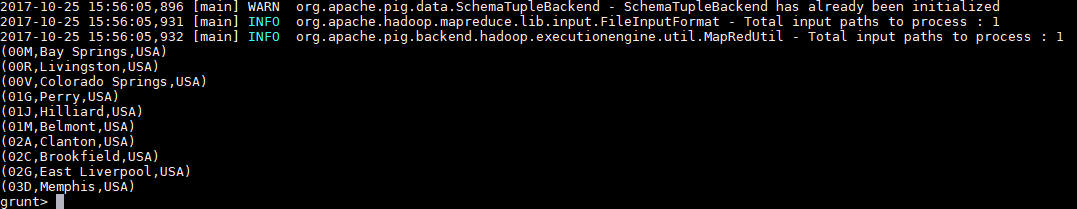




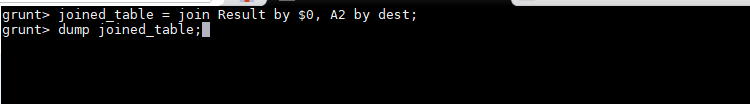
Loading the airports.csv file to find other fields-



Contents of file-



Joining result from 1st file and 2nd file with destination id-



**Final Output-**

