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)$2 as month,(int)$10 as flight\_num,(int)$22 as cancelled,(chararray)$23 as cancel\_code;**

**C = filter B by cancelled == 1 AND cancel\_code =='B';**

**D = group C by month;**

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

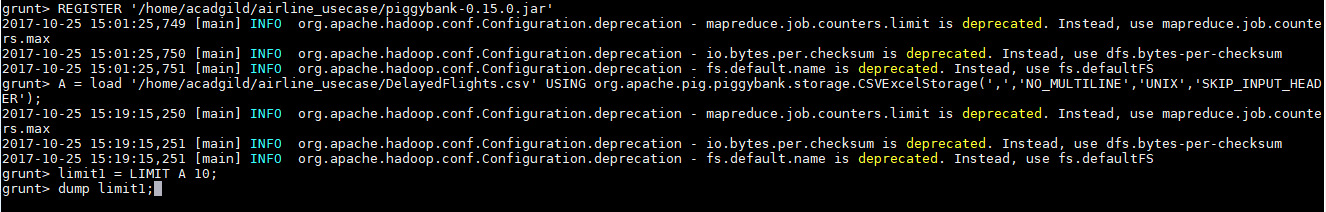
**F= order E by $1 DESC;**

**Result = limit F 1;**

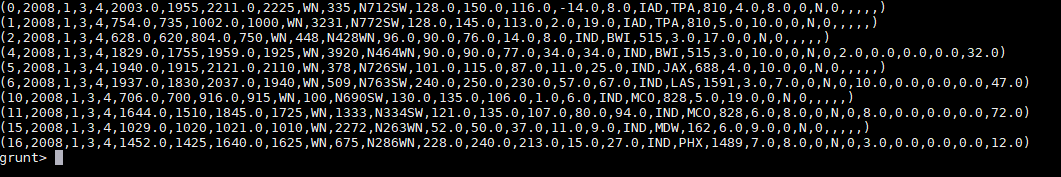
**dump Result;**

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

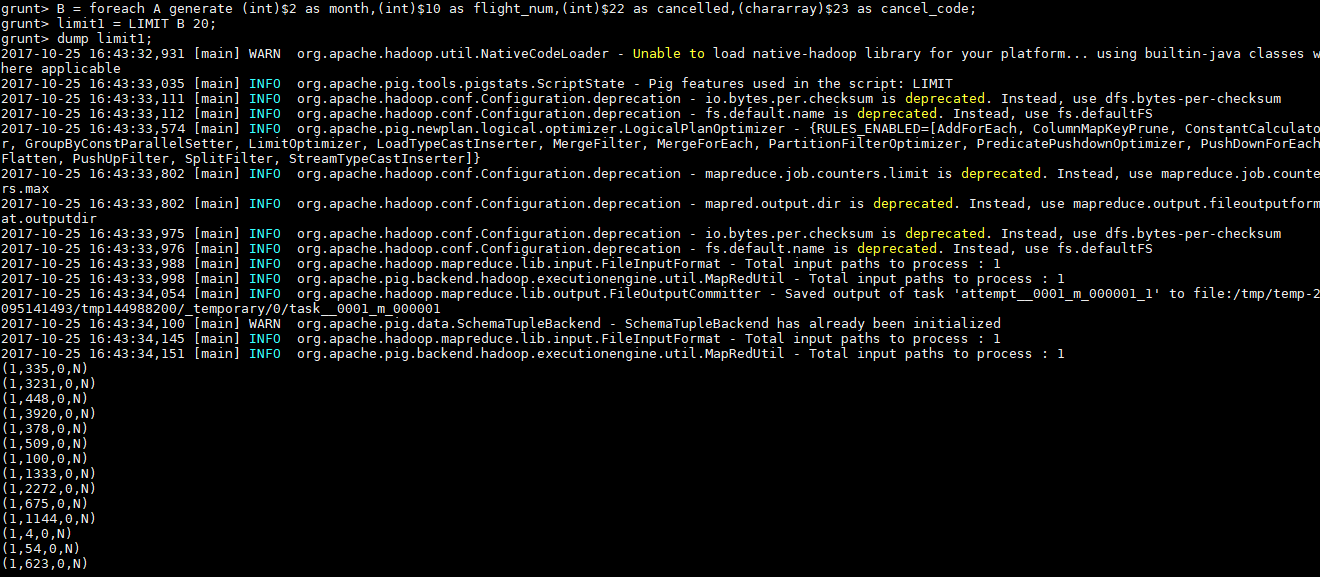
Registering piggybank jar in order to use the CSVExcelStorage class-



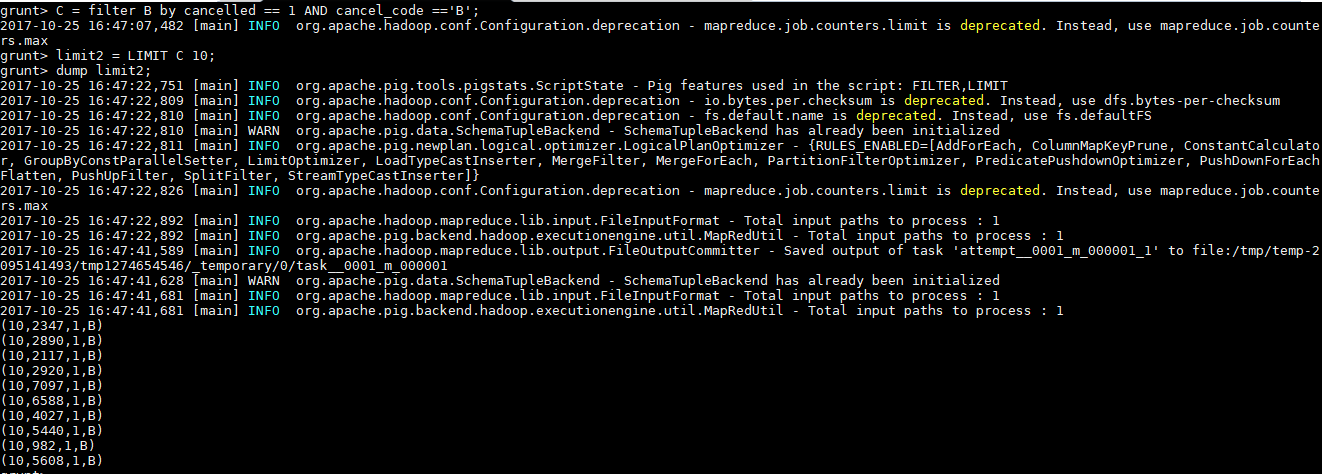
Contents of file-



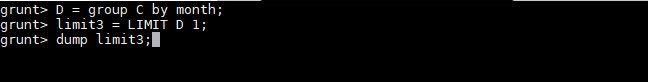
Extract month, flight number, cancelled indicator and cancelled code from th above file-

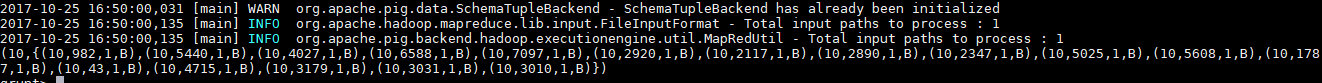


Filter records with cancelled indicator as 1 and cancel code as B-

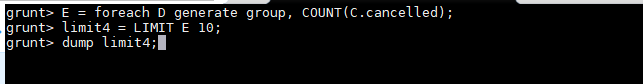


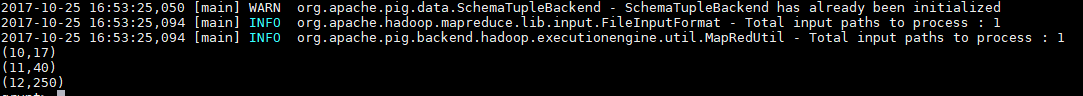
Group each record with month-





Find the number of months in each group-





Arrange the above output in descending order and take the topmost to find the maximum count of cancellations-



FINAL O/P-

