We have Sample\_log file. Use MR mode while running Pig and write Pig Latin script to get below results:

1)Find out the most viewed page

2)Find total hits per unique day

Ans 1 :

Register /usr/local/pig/lib/piggybank.jar  
  
A = Load 'sample\_log' Using org.apache.pig.piggybank.storage.apachelog.CombinedLogLoader As (IP:chararray, logname:chararray, UserID:chararray, time:chararray, request1:chararray, request2:chararray, request3:chararray, temp1:int, temp2:int, response:chararray, browser:chararray);  
  
B = Group A BY response;  
  
C = Foreach B Generate flatten($0),COUNT($1) AS page\_count;  
  
D = Order C BY page\_count DESC;  
  
E = Limit D 2;  
  
Dump E;

  
  
  
  
Ans 2 :

DEFINE Date\_Extract org.apache.pig.piggybank.evaluation.util.apachelogparser.DateExtractor('dd/MMM/yyyy:HH:mm:ss Z','dd');  
  
A = Load 'sample\_log' Using org.apache.pig.piggybank.storage.apachelog.CombinedLogLoader As (IP:chararray, logname:chararray, UserID:chararray, time:chararray, request1:chararray, request2:chararray, request3:chararray, temp1:int, temp2:int, response:chararray, browser:chararray);  
  
findday = Group A By Date\_Extract(time);  
  
hitsperday = Foreach findday Generate flatten($0), COUNT($1);  
  
Dump hitsperday;

