1)

REGISTER /home/acadgild/pig/piggybank-0.15.0.jar;

log\_data = LOAD 'sample\_log' USING org.apache.pig.piggybank.storage.apachelog.CombinedLogLoader

AS (ipaddr:chararray,logname:chararray,temp0:chararray,time:chararray,temp1:chararray,temp2:chararray,temp3:chararray,temp4:int,temp5:int,page\_link:chararray,temp6:chararray);

grp\_log\_data = GROUP log\_data by page\_link;

page\_hits = FOREACH grp\_log\_data GENERATE flatten($0), COUNT($1) as page\_count;

page\_hits\_ordered = ORDER page\_hits by page\_count DESC;

most\_viewed\_page = LIMIT page\_hits\_ordered 1;

DUMP most\_viewed\_page;

2)

DEFINE DATE\_EXTRACT org.apache.pig.piggybank.evaluation.util.apachelogparser.DateExtractor('dd/MMM/yyyy:HH:mm:ss Z','dd');

grp\_log\_time = GROUP log\_data BY DATE\_EXTRACT(time);

hits\_per\_day = FOREACH grp\_log\_time GENERATE flatten($0), COUNT($1);

DUMP hits\_per\_day;