

LAB 5

APACHE PIG LAB – PART - 1

Homework Questions

5 points each



1) Finding Word Count using Pig

Create a “.pig” script file and write all the commands needed to run word count example in Pig. Your script must correctly generate word counts for the given input file *InputForWC.txt*

Submit the pig script file along with the output file.

2) Joins in Pig – Top 5 most visited sites

Create some sample data for “users.csv” and “pages.csv” files as discussed during lecture.

Find the top 5 most visited sites by users aged between 18 - 25.

**Submit the pig script file along with your input & output files.
Paste screenshots wherever applicable.**