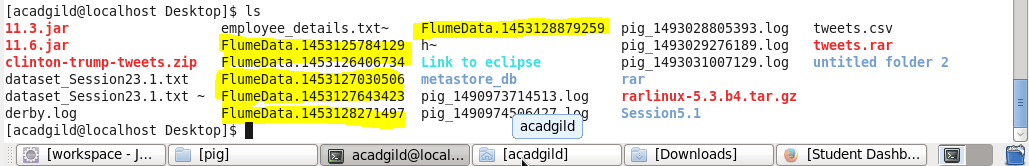
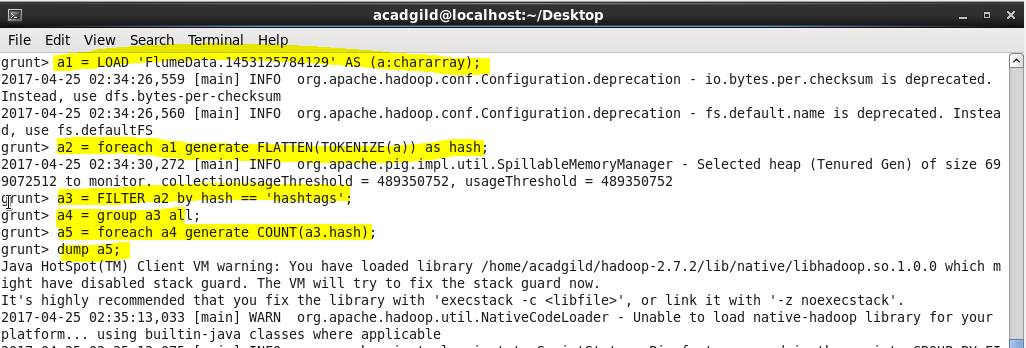


After doing unrar of our tweet.rar dataset we got following files :

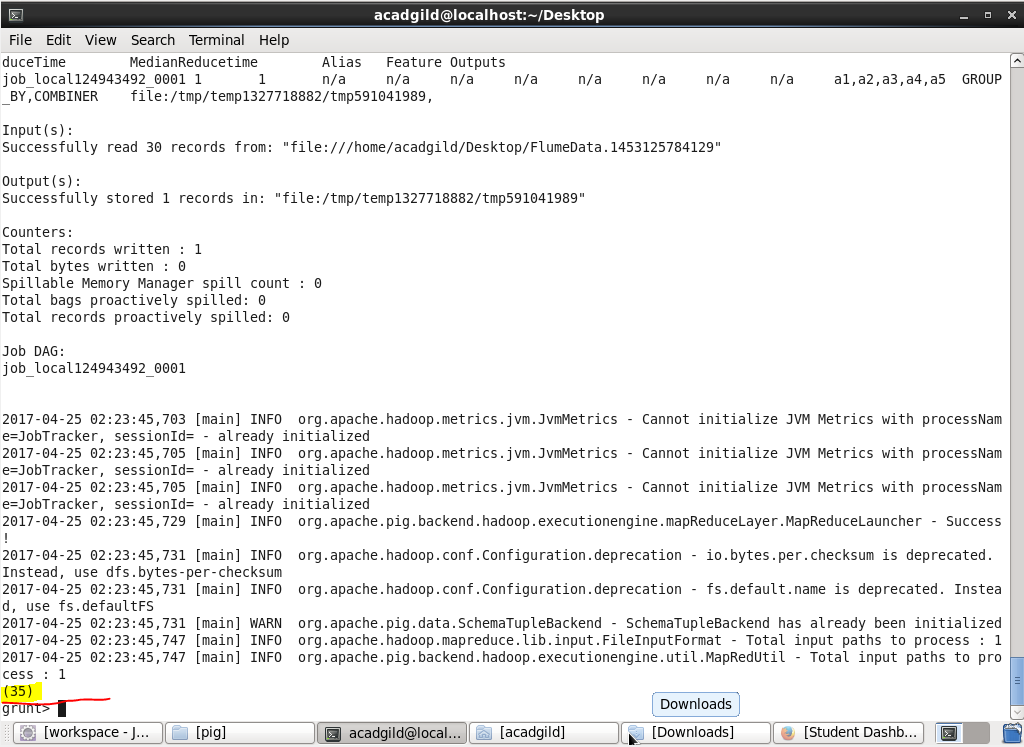


pig script

1. For file FlumeData.1453125784129

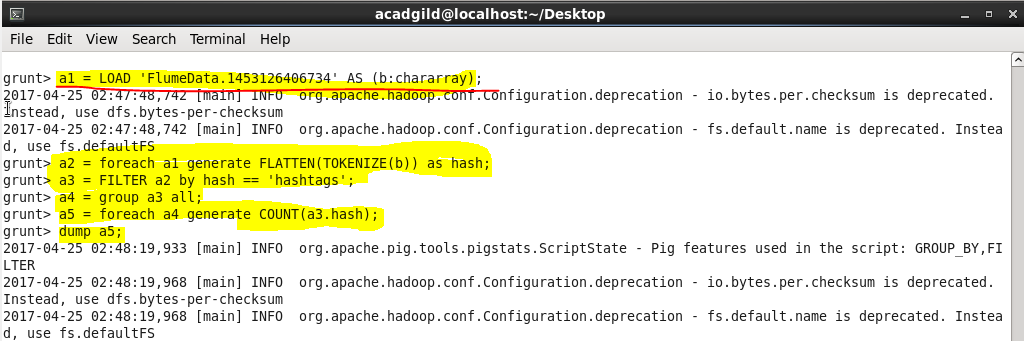


Output :

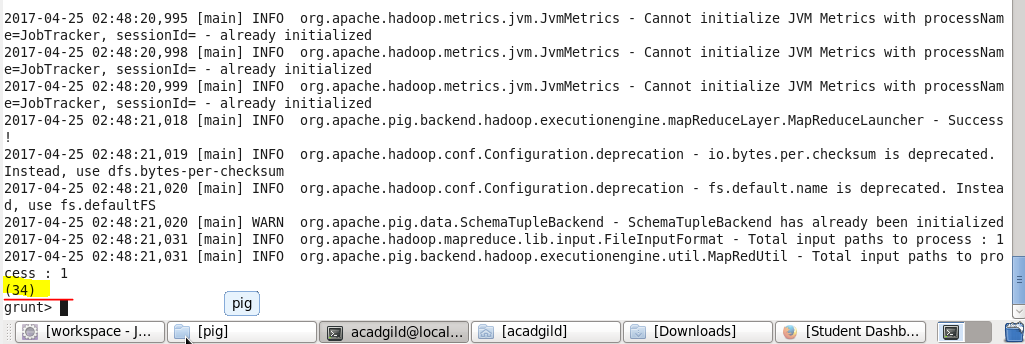


So in total 35 hashtags in file 1

1. For file FlumeData.1453126406734

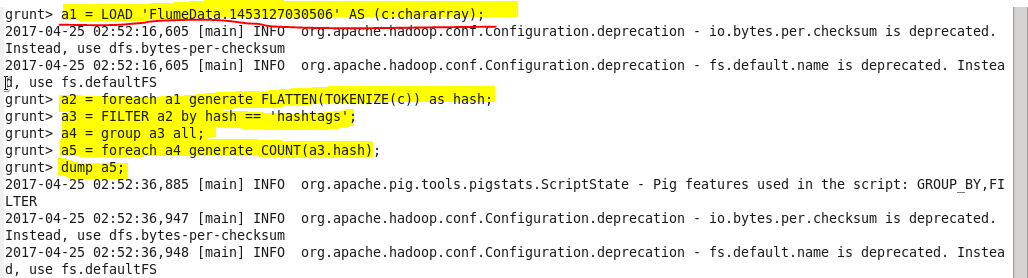


Output :

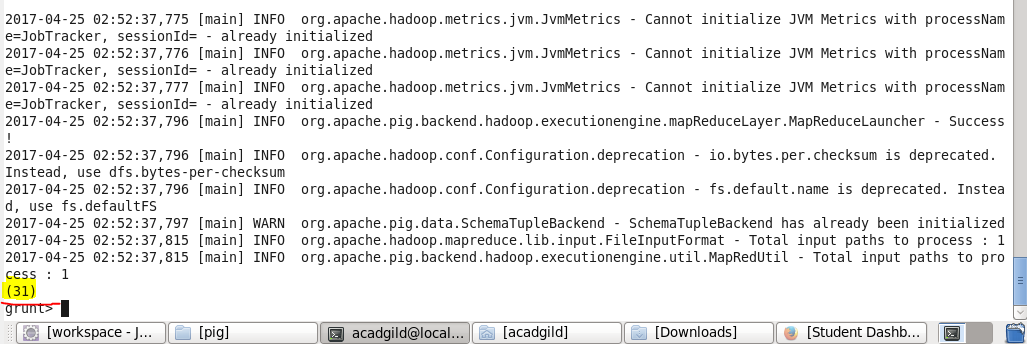


So in total 34 hashtags in file 2

1. For file FlumeData.1453127030506

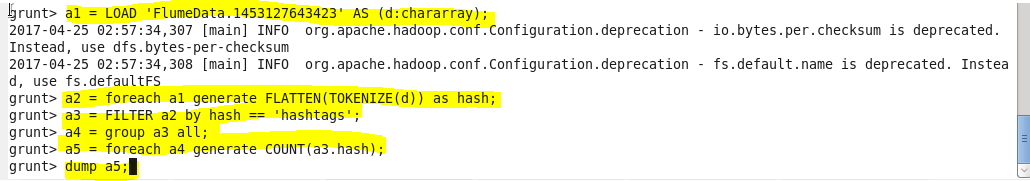


Output :

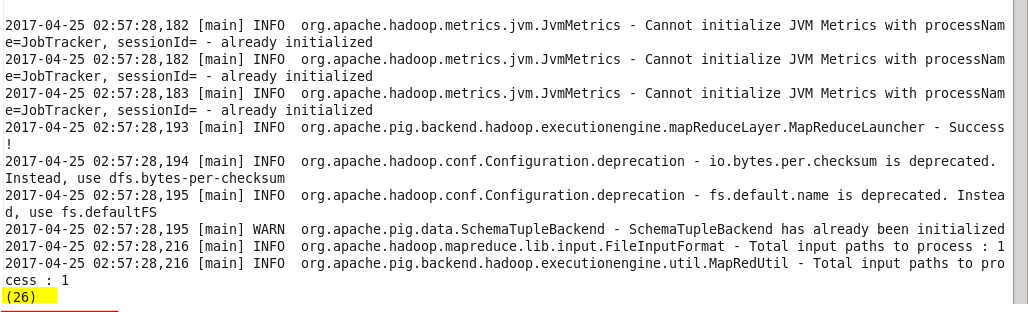


So in total 31 hashtags in file 3

1. For file FlumeData.1453127643423

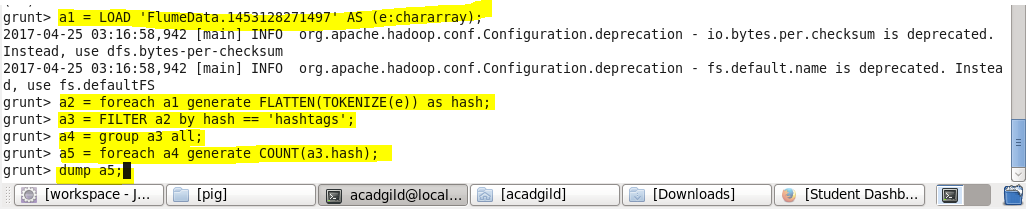


Output :

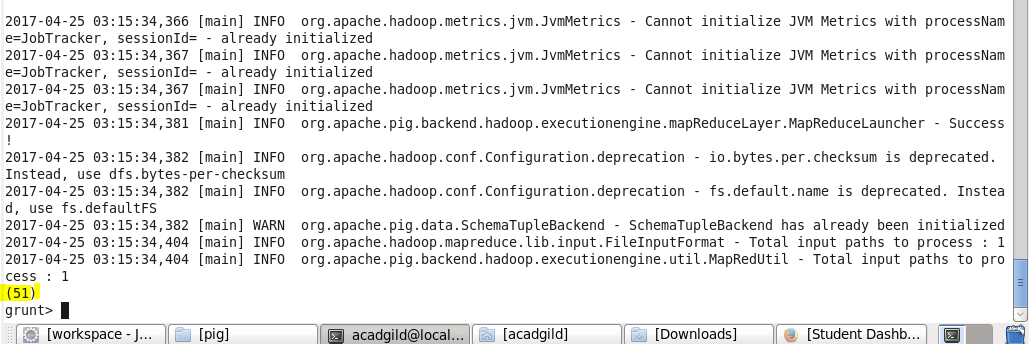


So in total 26 hashtags in file 4

1. For file FlumeData.1453128271497

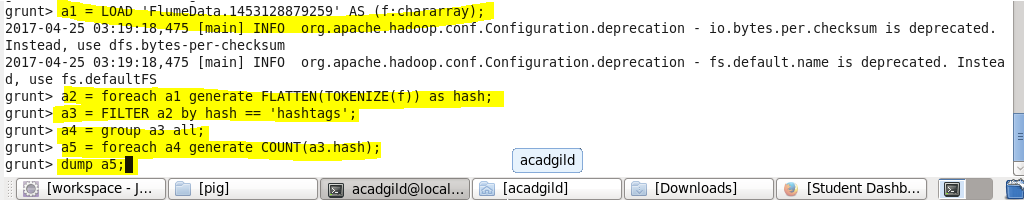


Output :

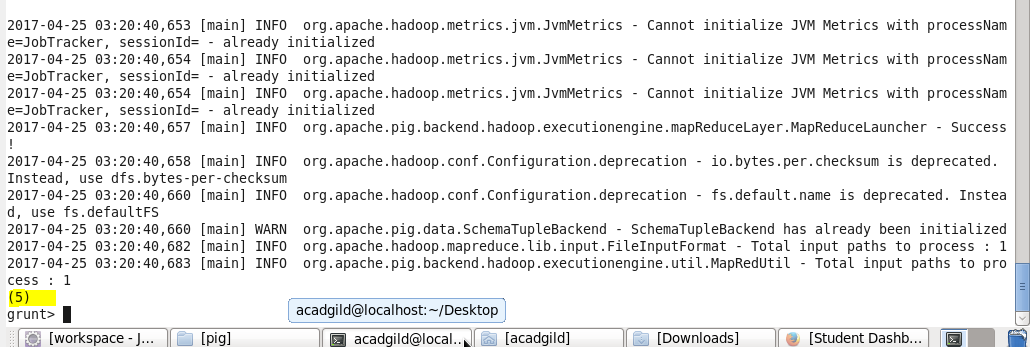


So in total 51 hashtags in file 5

1. For file FlumeData.1453128879259



Output :



So in total 5 hashtags in file 6