

```
Oct 6 12:10
hadoop@ab4cseibm133: ~
hadoop@ab4cseibm133: ~
(file:/home/hadoop/Downloads/books.txt,&#8220;treasures&#8221;,1.7120864696215674E-5)
(file:/home/hadoop/Downloads/books.txt,978-1-4555-5433-1still,1.7120864696215674E-5)
(file:/home/hadoop/Downloads/books.txt,Grant....well...what.,1.7120864696215674E-5)
(file:/home/hadoop/Downloads/books.txt,acvermen.blogspot.com,1.7120864696215674E-5)
(file:/home/hadoop/Downloads/books.txt,behind-the-scenes-look,1.7120864696215674E-5)
(file:/home/hadoop/Downloads/books.txt,loud.Characters:Ellie,1.7120864696215674E-5)
(file:/home/hadoop/Downloads/books.txt,subconscious&#8221;and,1.7120864696215674E-5)
(file:/home/hadoop/Downloads/books.txt,universe...&#34;Ashley,1.7120864696215674E-5)
(file:/home/hadoop/Downloads/books.txt,&#34;Dyscolution&#34;,1.7120864696215674E-5)
(file:/home/hadoop/Downloads/books.txt,&#34;Haddenarch.&#34;,1.7120864696215674E-5)
(file:/home/hadoop/Downloads/books.txt,&#8220;messiness&#8221;,1.7120864696215674E-5)
(file:/home/hadoop/Downloads/books.txt,SallyReadsTooMuchSoWhat,1.7120864696215674E-5)
(file:/home/hadoop/Downloads/books.txt,foresight/foreknowledge,1.7120864696215674E-5)
(file:/home/hadoop/Downloads/books.txt,know.&#8221;&#8220;They,1.7120864696215674E-5)
(file:/home/hadoop/Downloads/books.txt,teenager.Heaven&#8217;s,1.7120864696215674E-5)
(file:/home/hadoop/Downloads/books.txt,&#8220;unromantic&#8221;,1.7120864696215674E-5)
(file:/home/hadoop/Downloads/books.txt,Sherwin&#34;dragon&#34;,1.7120864696215674E-5)
(file:/home/hadoop/Downloads/books.txt,mInutes.&#34;&#34;That's,1.7120864696215674E-5)
(file:/home/hadoop/Downloads/books.txt,ms-t-plan-for-everything,1.7120864696215674E-5)
(file:/home/hadoop/Downloads/books.txt,relationship.Recommended,1.7120864696215674E-5)
(file:/home/hadoop/Downloads/books.txt,resistance.&#8221;Seeling,1.7120864696215674E-5)
(file:/home/hadoop/Downloads/books.txt,worse.&#8220;Let&#8217;s,1.7120864696215674E-5)
(file:/home/hadoop/Downloads/books.txt,&#34;scientifically.&#34;,1.7120864696215674E-5)
(file:/home/hadoop/Downloads/books.txt,characters.Hunter&#8217;s,1.7120864696215674E-5)
(file:/home/hadoop/Downloads/books.txt,characters.&#34;Onton&#34;,1.7120864696215674E-5)
(file:/home/hadoop/Downloads/books.txt,Author.recommended!!Iplease,1.7120864696215674E-5)
(file:/home/hadoop/Downloads/books.txt,I..Loved..This..Book...This,1.7120864696215674E-5)
(file:/home/hadoop/Downloads/books.txt,978-1-6888-835-6&#8220;This,1.7120864696215674E-5)
(file:/home/hadoop/Downloads/books.txt,jacquelinesreads.blogspot.com,1.7120864696215674E-5)
(file:/home/hadoop/Downloads/books.txt,Author/illustrator/editor/rev...,1.7120864696215674E-5)
(file:/home/hadoop/Downloads/books.txt,Supercalafragilisticpealadocuous,1.7120864696215674E-5)
(file:/home/hadoop/Downloads/books.txt,dBrownstein[dot]directfrompublisher[dot]iconalso,1.7120864696215674E-5)
(file:/home/hadoop/Downloads/books.txt,she's over it and they're getting married.THE.END,1.7120864696215674E-5)
(file:/home/hadoop/Downloads/books.txt,1.7120864696215674E-5)
grunt>
```

```
Oct 6 12:10
hadoop@ab4cseibm133: ~
hadoop@ab4cseibm133: ~
(file:/home/hadoop/Downloads/books.txt,top-performing,1.7120864696215674E-5)
(file:/home/hadoop/Downloads/books.txt,transportation,1.7120864696215674E-5)
(file:/home/hadoop/Downloads/books.txt,unclisptlers,1.7120864696215674E-5)
(file:/home/hadoop/Downloads/books.txt,unconventional,1.7120864696215674E-5)
(file:/home/hadoop/Downloads/books.txt,understandable,3.424172939243135E-5)
(file:/home/hadoop/Downloads/books.txt,understanding.,1.7120864696215674E-5)
(file:/home/hadoop/Downloads/books.txt,uninteresting.,1.7120864696215674E-5)
(file:/home/hadoop/Downloads/books.txt,unlrvetwime,0.81712086469621567)
(file:/home/hadoop/Downloads/books.txt,useful.Chapter,1.7120864696215674E-5)
(file:/home/hadoop/Downloads/books.txt,vulnerability.,1.7120864696215674E-5)
(file:/home/hadoop/Downloads/books.txt,well-developed,1.7120864696215674E-5)
(file:/home/hadoop/Downloads/books.txt,wholeheartedly,1.7120864696215674E-5)
(file:/home/hadoop/Downloads/books.txt,withhistorical,1.7120864696215674E-5)
(file:/home/hadoop/Downloads/books.txt,wouldn&#8217;t,3.424172939243135E-5)
(file:/home/hadoop/Downloads/books.txt,&#34;Oh.&#34;In,1.7120864696215674E-5)
(file:/home/hadoop/Downloads/books.txt,&#34;click&#34;,1.7120864696215674E-5)
(file:/home/hadoop/Downloads/books.txt,&#34;green&#34;,1.7120864696215674E-5)
(file:/home/hadoop/Downloads/books.txt,&#34;known&#34;,1.7120864696215674E-5)
(file:/home/hadoop/Downloads/books.txt,&#34;magic&#34;,1.7120864696215674E-5)
(file:/home/hadoop/Downloads/books.txt,&#34;paranormal,1.7120864696215674E-5)
(file:/home/hadoop/Downloads/books.txt,&#34;potty&#34;,1.7120864696215674E-5)
(file:/home/hadoop/Downloads/books.txt,&#34;rides&#34;,1.7120864696215674E-5)
(file:/home/hadoop/Downloads/books.txt,&#34;story&#34;,1.7120864696215674E-5)
(file:/home/hadoop/Downloads/books.txt,&#34;thing&#34;,1.7120864696215674E-5)
(file:/home/hadoop/Downloads/books.txt,&#34;vell.&#34;,1.7120864696215674E-5)
(file:/home/hadoop/Downloads/books.txt,&#8220;barbatic,1.7120864696215674E-5)
(file:/home/hadoop/Downloads/books.txt,&#8220;material,1.7120864696215674E-5)
(file:/home/hadoop/Downloads/books.txt,&#8220;pictures,1.7120864696215674E-5)
(file:/home/hadoop/Downloads/books.txt,AH-MAZING!!It's,1.7120864696215674E-5)
(file:/home/hadoop/Downloads/books.txt,Accountability,1.7120864696215674E-5)
(file:/home/hadoop/Downloads/books.txt,BookAddictMumma,1.7120864696215674E-5)
(file:/home/hadoop/Downloads/books.txt,BookReviewsus,1.7120864696215674E-5)
(file:/home/hadoop/Downloads/books.txt,Braydon&#8217;s,1.7120864696215674E-5)
(file:/home/hadoop/Downloads/books.txt,cardsByCarolAnn,1.7120864696215674E-5)
(file:/home/hadoop/Downloads/books.txt,Christian-based,1.7120864696215674E-5)
(file:/home/hadoop/Downloads/books.txt,Cordella.Portia,1.7120864696215674E-5)
```

```
A = load '/tmp/books_small' using PigStorage('\n','-tagPath');
B = foreach A generate $0 as file, flatten(TOKENIZE((chararray)$1)) as word;
-- Job 1
C = group B by (word, file);
D = foreach C generate COUNT(B) as n, group.word, group.file;
-- Job 2
E = group D by file;
F = foreach E generate SUM(D.n) as N, flatten(D);
-- Job 3
G = group F by word;
H = foreach G generate COUNT(F.file) as m, flatten(F);
-- Job 4
R = foreach H generate file, word, ((1.0*n)/N)*LOG(10.0/m) as tfidf;
```