

-- copy this to a file (or 2 files 1 line each) , if you want to match the input files in the wordcount script

-- you might have these as testfile1 and testfile2 from the previous course

-- BUT TAKE OUT these comment lines. In cloudera vm you can use gedit

A long time ago in a galaxy far far away

Another episode of Star Wars

There is a cheat sheet for PIG at this site that seems concise and useful:

<https://www.qubole.com/resources/cheatsheet/pig-function-cheat-sheet/>

Apache org has detailed documentation

[https://\*\*pig.apache.org/docs\*\*/r0.12.0/](https://pig.apache.org/docs/r0.12.0/)

The O'Reilly Hadoop Definitive Guide has a chapter on PIG, and a book 'Programming Pig'.