Hints for Installing and Running Hadoop in Windows

Installation:

- When setting the environment variables JAVA_HOME, PATH, and HADOOP_HOME use Progra~1 to include path names with spaces (legacy names)
- You will probably also need the files hadoop.dll and winutils.exe in the bin\ directory (can be downloaded from http://www.srccodes.com/p/article/39/error-util-shell-failed-locate-winutils-binary-hadoop-binary-path)
- You need to set the variable in file etc\hadoop\mapred-site.xml called mapreduce.framework.name to local
- Visualizing results on localhost:50070 (make sure that you check the almost transparent little icon at the bottom of the Overview screen called LegacyUI)

Running WordCount:

- >cd %HADOOP_HOME%
- >sbin\start-all
- >hadoop fs -mkdir /demo
- >hadoop fs –put D:\Hadoop\Homework\demoinput.txt /demo/demoinput.txt
- >hadoop fs –cat /demo/demoinput.txt
- >cd share\hadoop\mapreduce
- >hadoop jar hadoop-mapreduce-examples-2.3.0.jar wordcount /demo/demoinput.txt /demo/demooutput
- >hadoop fs –get \demo\demoouput D:\Hadoop\Homework\
- The sources are available from \share\hadoop\mapreduce\sources\hadoop-mapreduce-examples-2.3.0-sources.jar, which you need to rename as zip in order to view the sources
- >hadoop fs –rm /demo/demoinput.txt
- >hadoop fs –rm /demo/demooutput/*
- >hadoop fs –rmdir /demo/demoouput