Assignment 1 Report

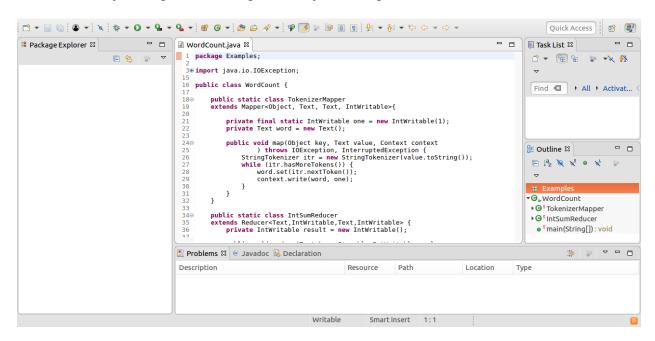
1. Screenshot below of Hadoop command running from Hadoop/bin/

```
jimmy@jimmy-VirtualBox:-$ ls
data Documents examples.desktop Music Public
Desktop Downloads Hadoop Pictures Templates
jimmy@jimmy-VirtualBox:-$ Hadoop/bin/hadoop
Usage: hadoop [OPTIONS] SUBCOMMAND [SUBCOMMAND OPTIONS]
or hadoop [OPTIONS] CLASSNAME [CLASSNAME OPTIONS]
where CLASSNAME is a user-provided Java class
     OPTIONS is none or any of:
 buildpaths
                                                                                          attempt to add class files from build tree
    -config dir
-debug
-help
                                                                                      Hadoop config directory
turn on shell script debug mode
usage information
hosts to use in slave mode
list of hosts to use in slave mode
set the log4] level for this command
turn on worker mode
 - netp
hostnames list[,of,host,names]
hosts filename
loglevel level
workers
     SUBCOMMAND is one of:
          Admin Commands:
 daemonlog
                                     get/set the log level for each daemon
          Client Commands:
                                   create a Hadoop archive check native Hadoop and compression libraries availability prints the class path needed to get the Hadoop jar and the required libraries validate configuration XML files interact with credential providers distributed metadata changer conviding or distributed security is provided to the convictories recursively.
 archive
checknative
  classpath
  conftest
 credential
distch
 distcp
dtutil
                                      copy file or directories recursively operations related to delegation tokens
```

2) File peterpan.txt and test.txt uploaded to HDFS successfully

```
🔊 🖃 📵 🛾 jimmy@jimmy-VirtualBox: ~/hadoop
/*********************
SHUTDOWN_MSG: Shutting down NameNode at jimmy-VirtualBox/127.0.1.1
******************
jimmy@jimmy-VirtualBox:~$ sbin/start-dfs.sh
-bash: sbin/start-dfs.sh: No such file or directory
jimmy@jimmy-VirtualBox:~$ cd hadoop
jimmy@jimmy-VirtualBox:~/hadoop$ sbin/start-dfs.sh
Starting namenodes on [localhost]
Starting datanodes
Starting secondary namenodes [jimmy-VirtualBox]
jimmy@jimmy-VirtualBox:~/hadoop$ hdfs dfs -mkdir /user
jimmy@jimmy-VirtualBox:~/hadoop$ hdfs dfs -mkdir /user/jimmy
jimmy@jimmy-VirtualBox:~/hadoop$ hdfs dfs -mkdir /user/jimmy/data
jimmy@jimmy-VirtualBox:~/hadoop$ hdfs dfs -put /home/jimmy/data/peterpan.txt /us
er/jimmy/data
jimmy@jimmy-VirtualBox:~/hadoop$ hdfs dfs -put /home/jimmy/data/test.txt /user/j
immy/data
jimmy@jimmy-VirtualBox:~/hadoop$ hdfs dfs -ls /user/jimmy/data
Found 2 items
                                  291927 2018-01-29 16:16 /user/jimmy/data/pet
- FW- F-- F--
            1 jimmy supergroup
erpan.txt
- - W - C - - C - -
                                      66 2018-01-29 16:16 /user/jimmy/data/tes
            1 jimmy supergroup
t.txt
jimmy@jimmy-VirtualBox:~/hadoop$
```

3) WordCount.java imported and exported to .jar in Eclipse.



4) WordCount.jar run on peterpan.txt and test.txt

```
jimmy@jimmy-VirtualBox: ~/hadoop
                Merged Map outputs=1
                GC time elapsed (ms)=200
                Total committed heap usage (bytes)=266608640
        Shuffle Errors
                BAD_ID=0
                CONNECTION=0
                IO ERROR=0
                WRONG LENGTH=0
                WRONG MAP=0
                WRONG REDUCE=0
        File Input Format Counters
                Bytes Read=291927
        File Output Format Counters
                Bytes Written=95576
jimmy@jimmy-VirtualBox:~/data$ hadoop fs -get /user/jimmy/peterpan_output/* ~/da
ta/
get: `/home/jimmy/data/_SUCCESS': File exists
jimmy@jimmy-VirtualBox:~/data$ cd ...
jimmy@jimmy-VirtualBox:~$ cd hadoop
jimmy@jimmy-VirtualBox:~/hadoop$ sbin/stop-dfs.sh
Stopping namenodes on [localhost]
Stopping datanodes
Stopping secondary namenodes [jimmy-VirtualBox]
jimmy@jimmy-VirtualBox:~/hadoop$
```

