

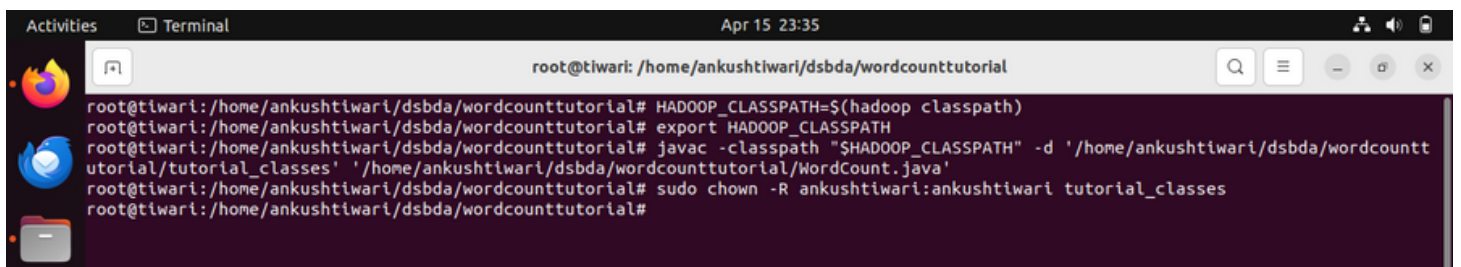
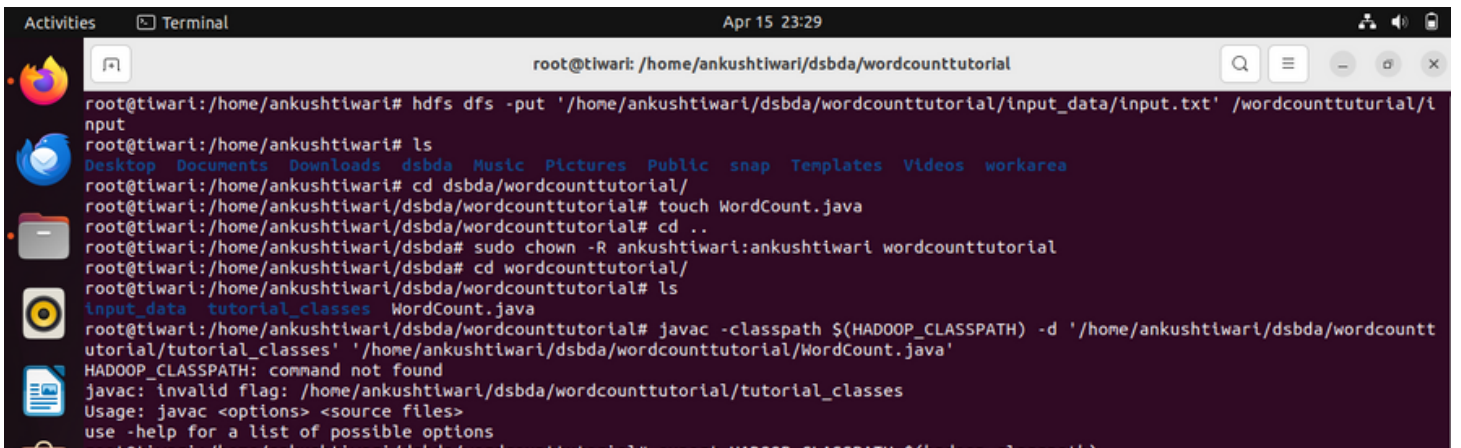
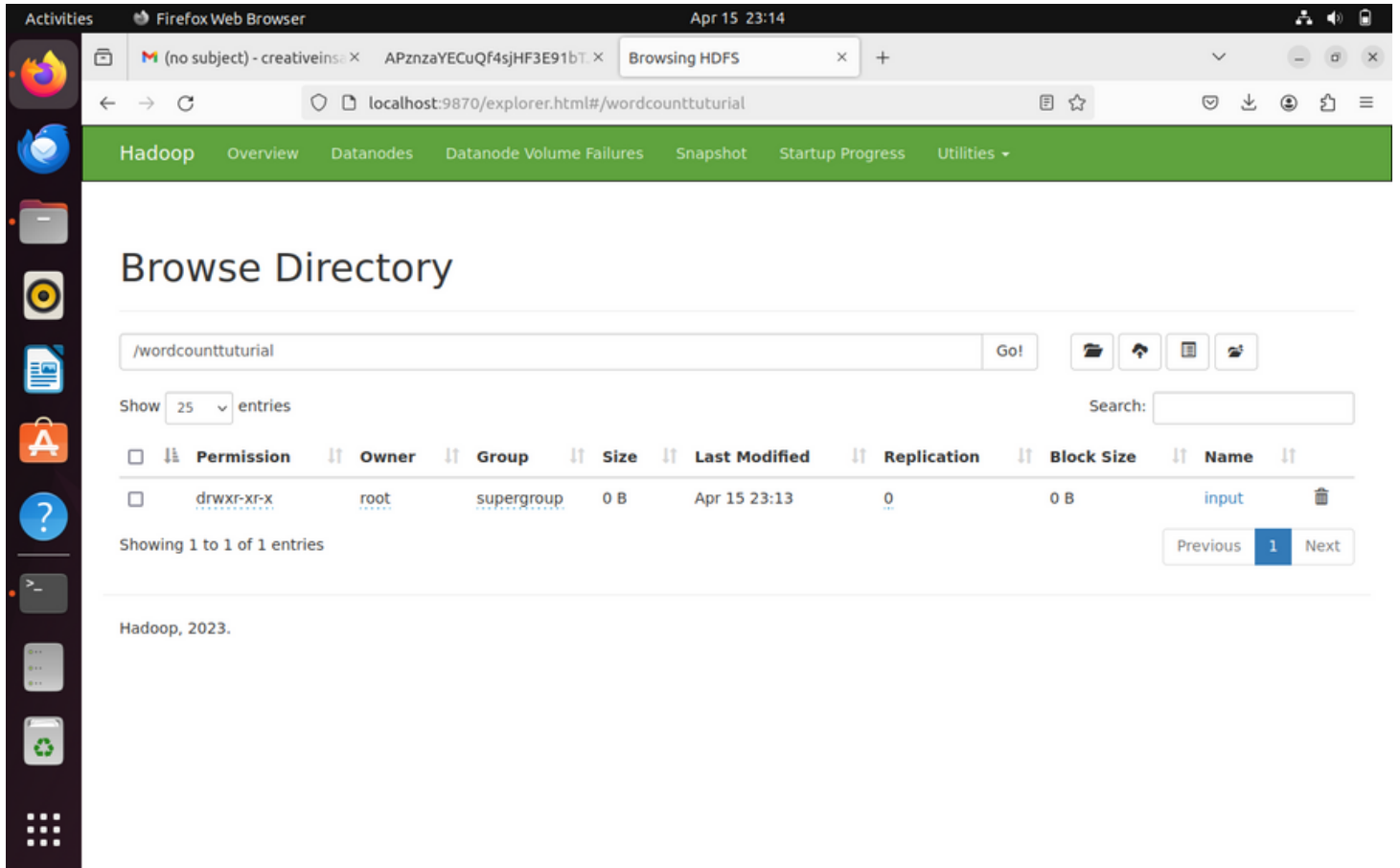
**AIM: Write a code in JAVA for a implement WordCount application that counts the number of occurrence of each word in a given input set using the Hadoop MapReduce framework on local-standalone set-up.**

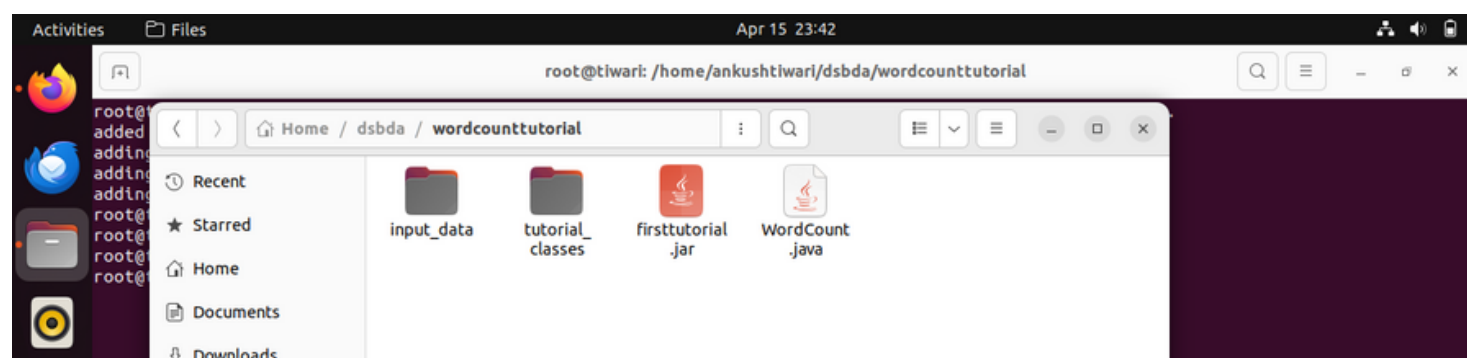
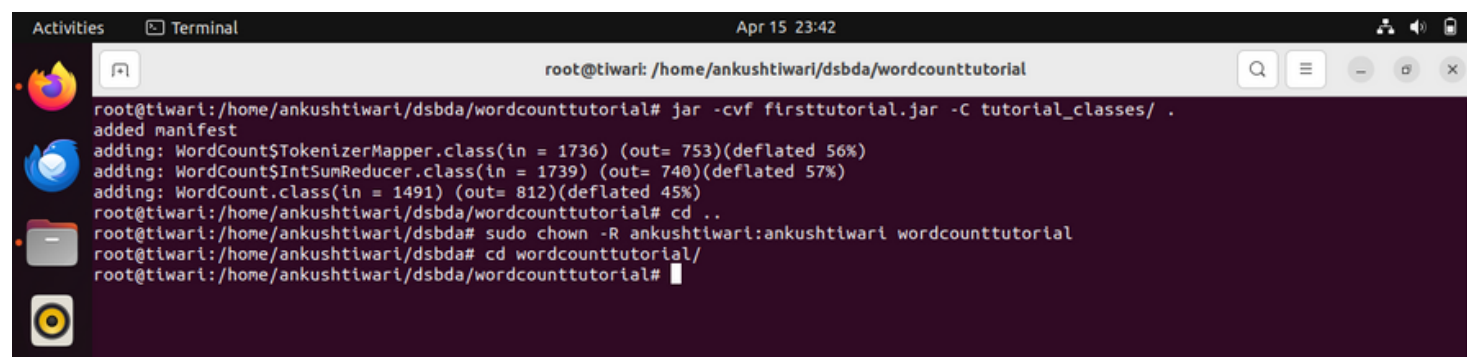
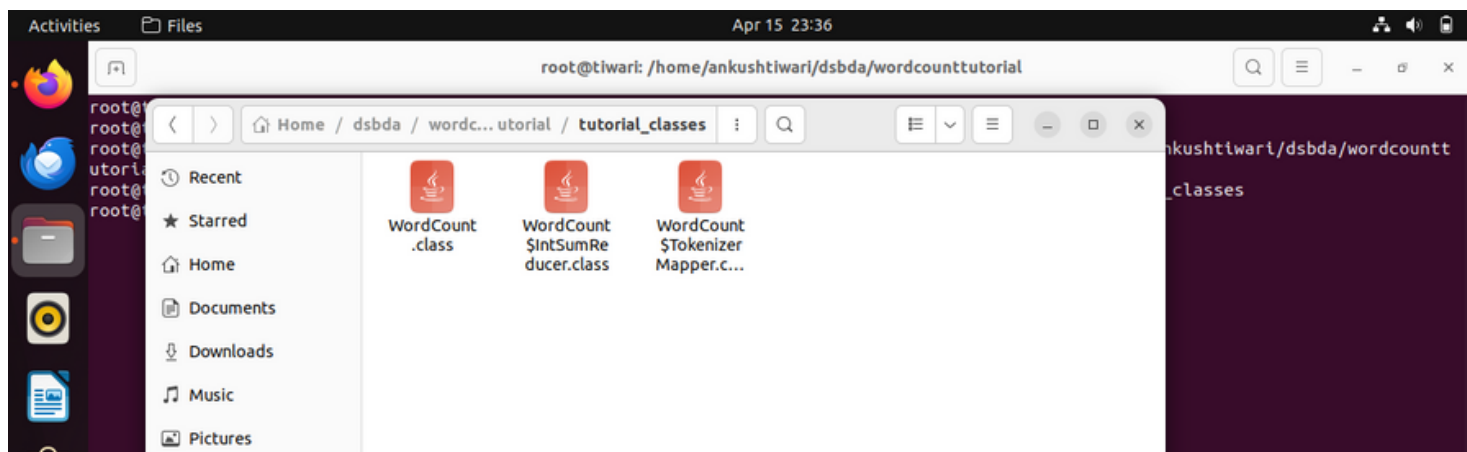
```
Activities Terminal Apr 15 23:05
root@tiwari: /home/ankushtiwari

root@tiwari:/home/ankushtiwari# java -version
openjdk version "1.8.0_402"
OpenJDK Runtime Environment (build 1.8.0_402-8u402-ga-2ubuntu1~22.04-b06)
OpenJDK 64-Bit Server VM (build 25.402-b06, mixed mode)
root@tiwari:/home/ankushtiwari# hadoop version
Hadoop 3.3.6
Source code repository https://github.com/apache/hadoop.git -r 1be78238728da9266a4f88195058f08fd012bf9c
Compiled by ubuntu on 2023-06-18T08:22Z
Compiled on platform linux-x86_64
Compiled with protoc 3.7.1
From source with checksum 5652179ad55f76cb287d9c633bb53bbd
This command was run using /home/ankushtiwari/workarea/software/hadoop-3.3.6/share/hadoop/common/hadoop-common-3.3.6.jar
root@tiwari:/home/ankushtiwari# cd dsbda/
root@tiwari:/home/ankushtiwari/dsbda# mkdir wordcounttutorial
root@tiwari:/home/ankushtiwari/dsbda/wordcounttutorial# touch input.txt
root@tiwari:/home/ankushtiwari/dsbda/wordcounttutorial# cd ..
root@tiwari:/home/ankushtiwari/dsbda# sudo chown -R ankushtiwari:ankushtiwari wordcounttutorial
root@tiwari:/home/ankushtiwari/dsbda# cd wordcounttutorial/
root@tiwari:/home/ankushtiwari/dsbda/wordcounttutorial# mkdir input_data
root@tiwari:/home/ankushtiwari/dsbda/wordcounttutorial# sudo chown -R ankushtiwari:ankushtiwari input_data
root@tiwari:/home/ankushtiwari/dsbda/wordcounttutorial# cd ..
root@tiwari:/home/ankushtiwari/dsbda# cd ..
root@tiwari:/home/ankushtiwari# export HADOOP_CLASSPATH=$(
Display all 2009 possibilities? (y or n)
root@tiwari:/home/ankushtiwari# export HADOOP_CLASSPATH=$(hadoop classpath)
root@tiwari:/home/ankushtiwari# echo $HADOOP_CLASSPATH
/home/ankushtiwari/workarea/software/hadoop-3.3.6/etc/hadoop:/home/ankushtiwari/workarea/software/hadoop-3.3.6/share/hadoop/common/lib/*:/home/ankushtiwari/workarea/software/hadoop-3.3.6/share/hadoop/common/*:/home/ankushtiwari/workarea/software/hadoop-3.3.6/share/hadoop/hdfs:/home/ankushtiwari/workarea/software/hadoop-3.3.6/share/hadoop/hdfs/lib/*:/home/ankushtiwari/workarea/software/hadoop-3.3.6/share/hadoop/hdfs/*:/home/ankushtiwari/workarea/software/hadoop-3.3.6/share/hadoop/mapreduce/*:/home/ankushtiwari/workarea/software/hadoop-3.3.6/share/hadoop/mapreduce/lib/*:/home/ankushtiwari/workarea/software/hadoop-3.3.6/share/hadoop/yarn:/home/ankushtiwari/workarea/software/hadoop-3.3.6/share/hadoop/yarn/lib/*:/home/ankushtiwari/workarea/software/hadoop-3.3.6/share/hadoop/yarn/*
root@tiwari:/home/ankushtiwari#
```

```
Activities Terminal Apr 15 23:14
root@tiwari: /home/ankushtiwari

root@tiwari:/home/ankushtiwari# start-all.sh
Starting namenodes on [localhost]
Starting datanodes
Starting secondary namenodes [tiwari]
Starting resourcemanager
Starting nodemanagers
root@tiwari:/home/ankushtiwari# hadoop fs -mkdir /wordcounttutorial
root@tiwari:/home/ankushtiwari# hadoop fs -mkdir /wordcounttutorial/input
root@tiwari:/home/ankushtiwari#
```







```
Activities Terminal Apr 16 00:09
root@tiwari: /home/ankushtiwari/dsbda/wordcounttutorial

root@tiwari: /home/ankushtiwari/dsbda/wordcounttutorial
root@tiwari: /home/ankushtiwari/dsbda/wordcounttutorial# start-all.sh
Starting namenodes on [localhost]
Starting datanodes
Starting secondary namenodes [tiwari]
Starting resourcemanager
Starting nodemanagers
root@tiwari: /home/ankushtiwari/dsbda/wordcounttutorial# hadoop jar '/home/ankushtiwari/dsbda/wordcounttutorial/firsttutorial.jar' WordCount /wordcounttutorial/input /wordcounttutorial/output
2024-04-16 00:06:58,682 INFO client.DefaultNoHARMAFailoverProxyProvider: Connecting to ResourceManager at /0.0.0.0:8032
2024-04-16 00:07:05,302 WARN mapreduce.JobResourceUploader: Hadoop command-line option parsing not performed. Implement the Tool interface and execute your application with ToolRunner to remedy this.
2024-04-16 00:07:05,704 INFO mapreduce.JobResourceUploader: Disabling Erasure Coding for path: /tmp/hadoop-yarn/staging/root/.staging/job_1713206100354_0001
2024-04-16 00:07:18,781 INFO input.FileInputFormat: Total input files to process : 1
2024-04-16 00:07:20,101 INFO mapreduce.JobSubmitter: number of splits:1
2024-04-16 00:07:20,943 INFO mapreduce.JobSubmitter: Submitting tokens for job: job_1713206100354_0001
2024-04-16 00:07:20,944 INFO mapreduce.JobSubmitter: Executing with tokens: []
2024-04-16 00:07:21,928 INFO conf.Configuration: resource-types.xml not found
2024-04-16 00:07:21,931 INFO resource.ResourceUtils: Unable to find 'resource-types.xml'.
2024-04-16 00:07:27,139 INFO impl.YarnClientImpl: Submitted application application_1713206100354_0001
2024-04-16 00:07:27,330 INFO mapreduce.Job: The url to track the job: http://tiwari.myguest.virtualbox.org:8088/proxy/application_1713206100354_0001/
2024-04-16 00:07:27,339 INFO mapreduce.Job: Running job: job_1713206100354_0001
2024-04-16 00:08:36,109 INFO mapreduce.Job: Job job_1713206100354_0001 running in uber mode : false
2024-04-16 00:08:36,170 INFO mapreduce.Job: map 0% reduce 0%
2024-04-16 00:08:48,736 INFO mapreduce.Job: map 100% reduce 0%
2024-04-16 00:08:59,066 INFO mapreduce.Job: map 100% reduce 100%
2024-04-16 00:09:01,150 INFO mapreduce.Job: Job job_1713206100354_0001 completed successfully
2024-04-16 00:09:02,587 INFO mapreduce.Job: Counters: 54
File System Counters
  FILE: Number of bytes read=1265
  FILE: Number of bytes written=555483
  FILE: Number of read operations=0
  FILE: Number of large read operations=0
  FILE: Number of write operations=0
  HDFS: Number of bytes read=919
  HDFS: Number of bytes written=851
  HDFS: Number of read operations=8
```

```
Activities Terminal Apr 16 00:10
root@tiwari: /home/ankushtiwari/dsbda/wordcounttutorial

root@tiwari: /home/ankushtiwari/dsbda/wordcounttutorial
root@tiwari: /home/ankushtiwari/dsbda/wordcounttutorial#
Map-Reduce Framework
  Map input records=8
  Map output records=139
  Map output bytes=1355
  Map output materialized bytes=1265
  Input split bytes=120
  Combine input records=139
  Combine output records=102
  Reduce input groups=102
  Reduce shuffle bytes=1265
  Reduce input records=102
  Reduce output records=102
  Spilled Records=204
  Shuffled Maps =1
  Failed Shuffles=0
  Merged Map outputs=1
  GC time elapsed (ms)=244
  CPU time spent (ms)=1950
  Physical memory (bytes) snapshot=536207360
  Virtual memory (bytes) snapshot=5133156352
  Total committed heap usage (bytes)=401080320
  Peak Map Physical memory (bytes)=320638976
  Peak Map Virtual memory (bytes)=2560835584
  Peak Reduce Physical memory (bytes)=215568384
  Peak Reduce Virtual memory (bytes)=2572320768
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=799
File Output Format Counters
  Bytes Written=851
root@tiwari: /home/ankushtiwari/dsbda/wordcounttutorial#
```

Activities Firefox Web Browser Apr 16 00:11

(no subject) - creativeins... APznzaYECuQf4sjHF3E91bT... Browsing HDFS Hadoop Permisi3n de Us... +

localhost:9870/explorer.html#/wordcounttutorial

Hadoop Overview Datanodes Datanode Volume Failures Snapshot Startup Progress Utilities

## Browse Directory

/wordcounttutorial Go!

Show 25 entries Search:

<input type="checkbox"/>	Permission	Owner	Group	Size	Last Modified	Replication	Block Size	Name	
<input type="checkbox"/>	drwxr-xr-x	root	supergroup	0 B	Apr 15 23:19	0	0 B	input	
<input type="checkbox"/>	drwxr-xr-x	root	supergroup	0 B	Apr 16 00:08	0	0 B	output	

Showing 1 to 2 of 2 entries Previous 1 Next

Hadoop, 2023.

Activities Text Editor Apr 16 00:14

(no subject) - creativeins... APznzaYECuQf4sjHF3E91bT... Browsing HDFS Hadoop Permisi3n de Us... +

part-r-00000 ~/Downloads

Save

```
1 "Who 1
2 & 2
3 1931 1
4 1931, 1
5 1949. 1
6 American 1
7 April 1
8 Detective 1
9 Each 1
10 For 1
11 Gibson 1
12 Magazine. 1
13 Map 1
14 Project 1
15 Reduce 1
16 Shadow 2
17 Shadow, 1
18 Smith 1
19 Smith's 1
20 Story 1
21 Street 2
22 The 4
23 Walter 1
24 WordCount 1
25 a 4
26 about 1
27 an 1
28 and 4
29 be 1
30 broadcasts 1
31 by 2
32 can 1
33 contained 1
34 contains 1
35 count 1
36 counts 1
37 crime-fighting 1
38 data 1
```