

WCMapper.java

```
import java.io.IOException;
import org.apache.hadoop.io.IntWritable;
import org.apache.hadoop.io.LongWritable;
import org.apache.hadoop.io.Text;
import org.apache.hadoop.mapred.MapReduceBase;
import org.apache.hadoop.mapred.Mapper;
import org.apache.hadoop.mapred.OutputCollector;
import org.apache.hadoop.mapred.Reporter;
public class WCMapper extends MapReduceBase implements
Mapper<LongWritable,Text, Text, IntWritable> {
    // Map function
    public void map(LongWritable key, Text value, OutputCollector<Text,
IntWritable> output, Reporter rep) throws IOException

{
    String line = value.toString();
    // Splitting the line on spaces
    for (String word : line.split(" "))
    {
        if (word.length() > 0)
        {
            output.collect(new Text(word), new IntWritable(1));
        }
    }
}
}
```

WCReducer.java

```
import java.io.IOException;
import java.util.Iterator;
import org.apache.hadoop.io.IntWritable;
import org.apache.hadoop.io.Text;
import org.apache.hadoop.mapred.MapReduceBase;
import org.apache.hadoop.mapred.OutputCollector;
import org.apache.hadoop.mapred.Reducer;
import org.apache.hadoop.mapred.Reporter;

public class WCReducer extends MapReduceBase implements
Reducer<Text,IntWritable, Text, IntWritable> {
    // Reduce function
    public void reduce(Text key, Iterator<IntWritable> value,
OutputCollector<Text, IntWritable> output,

Reporter rep) throws IOException

{
    int count = 0;
    // Counting the frequency of each words
    while (value.hasNext())
    {
        IntWritable i = value.next();
        count += i.get();
    }
    output.collect(key, new IntWritable(count));
}
```

```
}
```

WCDriver.java

```
import java.io.IOException;
import org.apache.hadoop.conf.Configured;
import org.apache.hadoop.fs.Path;
import org.apache.hadoop.io.IntWritable;
import org.apache.hadoop.io.Text;
import org.apache.hadoop.mapred.FileInputFormat;
import org.apache.hadoop.mapred.FileOutputFormat;

import org.apache.hadoop.mapred.JobClient;
import org.apache.hadoop.mapred.JobConf;
import org.apache.hadoop.util.Tool;
import org.apache.hadoop.util.ToolRunner;
public class WCDriver extends Configured implements Tool {
public int run(String args[]) throws IOException
{
if (args.length < 2)
{
System.out.println("Please give valid inputs");
return -1;
}
JobConf conf = new JobConf(WCDriver.class);
FileInputFormat.setInputPaths(conf, new Path(args[0]));
FileOutputFormat.setOutputPath(conf, new Path(args[1]));
conf.setMapperClass(WCMapper.class);
conf.setReducerClass(WCReducer.class);
conf.setMapOutputKeyClass(Text.class);
conf.setMapOutputValueClass(IntWritable.class);
conf.setOutputKeyClass(Text.class);
conf.setOutputValueClass(IntWritable.class);
JobClient.runJob(conf);
return 0;
}
// Main Method
public static void main(String args[]) throws Exception
{
int exitCode = ToolRunner.run(new WCDriver(), args);
System.out.println(exitCode);
}
}
```

```
hduser@bmsce-Precision-T1700:~$ jps
8625 NodeManager
7924 DataNode
8292 ResourceManager
8133 SecondaryNameNode
10280 Jps
7753 NameNode
hduser@bmsce-Precision-T1700:~$ sample.txt
sample.txt: command not found
hduser@bmsce-Precision-T1700:~$ hadoop fs -ls /
Found 2 items
drwxrwxr-x - hduser supergroup 0 2019-08-01 16:19 /tmp
drwxr-xr-x - hduser supergroup 0 2019-08-01 16:03 /user
hduser@bmsce-Precision-T1700:~$ hadoop fs -mkdir /rgs
hduser@bmsce-Precision-T1700:~$ hadoop fs -copyFromLocal D:/sample.txt /rgs/test.txt
copyFromLocal: '/sample.txt': No such file or directory
hduser@bmsce-Precision-T1700:~$ hadoop fs -copyFromLocal /home/bmsce/Desktop/sample.txt /rgs/test.txt
hduser@bmsce-Precision-T1700:~$ hdfs dfs -ls /rgs
Found 1 items
-rw-r--r-- 1 hduser supergroup 96 2022-06-14 10:08 /rgs/test.txt
copyFromLocal: '/sample.txt': No such file or directory
hduser@bmsce-Precision-T1700:~$ hadoop fs -copyFromLocal /home/bmsce/Desktop/sample.txt /rgs/test.txt
hduser@bmsce-Precision-T1700:~$ hdfs dfs -ls /rgs
Found 1 items
-rw-r--r-- 1 hduser supergroup 96 2022-06-14 10:08 /rgs/test.txt
```

```

hduser@bmsce-Precision-T1700:~$ hadoop jar /home/bmsce/eclipse-workspace/Wordcount134.jar/ WCDriver /rgs/test.txt /output/
22/06/14 10:14:47 INFO Configuration.deprecation: session.id is deprecated. Instead, use dfs.metrics.session-id
22/06/14 10:14:47 INFO jvm.JvmMetrics: Initializing JVM Metrics with processName=JobTracker, sessionId=
22/06/14 10:14:47 INFO jvm.JvmMetrics: Cannot initialize JVM Metrics with processName=JobTracker, sessionId= - already initialized
22/06/14 10:14:47 WARN mapreduce.JobSubmitter: Hadoop command-line option parsing not performed. Implement the Tool interface and execute your application with ToolRunner to remedy this.
22/06/14 10:14:47 INFO mapred.FileInputFormat: Total input paths to process : 1
22/06/14 10:14:47 INFO mapreduce.JobSubmitter: number of splits:1
22/06/14 10:14:47 INFO mapreduce.JobSubmitter: Submitting tokens for job: job_local1145327495_0001
22/06/14 10:14:47 INFO mapreduce.Job: The url to track the job: http://localhost:8080/
22/06/14 10:14:47 INFO mapred.LocalJobRunner: OutputCommitter set in config null
22/06/14 10:14:47 INFO mapreduce.Job: Running job: job_local1145327495_0001
22/06/14 10:14:47 INFO mapred.LocalJobRunner: OutputCommitter is org.apache.hadoop.mapred.FileOutputCommitter
22/06/14 10:14:47 INFO mapred.LocalJobRunner: Waiting for map tasks
22/06/14 10:14:47 INFO mapred.LocalJobRunner: Starting task: attempt_local1145327495_0001_m_000000_0
22/06/14 10:14:47 INFO mapred.Task: Using ResourceCalculatorProcessTree : [ ]
22/06/14 10:14:47 INFO mapred.MapTask: Processing split: hdfs://localhost:54310/rgs/test.txt:0+96
22/06/14 10:14:47 INFO mapred.MapTask: numReduceTasks: 1
22/06/14 10:14:47 INFO mapred.MapTask: (EQUATOR) 0 kvi 26214396(104857584)
22/06/14 10:14:47 INFO mapred.MapTask: mapreduce.task.io.sort.mb: 100
22/06/14 10:14:47 INFO mapred.MapTask: soft limit at 83886080
22/06/14 10:14:47 INFO mapred.MapTask: bufstart = 0; bufvoid = 104857600

```

```

0
hduser@bmsce-Precision-T1700:~$ hdfs dfs -cat /output/*
Balaji 1
are 1
brother 1
family 1
hi 1
how 5
is 4
job 1
sister 1
you 1
your 4
hduser@bmsce-Precision-T1700:~$

```