Assignment No 11

```java
import org.apache.hadoop.fs.Path;

import org.apache.hadoop.io.IntWritable;

import org.apache.hadoop.io.Text;

import org.apache.hadoop.mapreduce.lib.input.FileInputFormat;

import org.apache.hadoop.mapreduce.lib.output.FileOutputFormat;

import org.apache.hadoop.mapreduce.Job;

import org.apache.hadoop.conf.Configured;

import org.apache.hadoop.util.Tool;

import org.apache.hadoop.util.ToolRunner;

public class WordCount extends Configured implements Tool {

 public int run(String[] args) throws Exception {

  if (args.length != 2) {

   System.out.printf(

      "Usage: %s [generic options] <input dir> <output dir>\n", getClass()

         .getSimpleName());

   ToolRunner.printGenericCommandUsage(System.out);

   return -1;

  }

  Job job = new Job(getConf());

  job.setJarByClass(WordCount.class);

  job.setJobName(this.getClass().getName());

  FileInputFormat.setInputPaths(job, new Path(args[0]));

  FileOutputFormat.setOutputPath(job, new Path(args[1]));

  job.setMapperClass(WordMapper.class);
```

```java
      job.setReducerClass(SumReducer.class);

      job.setMapOutputKeyClass(Text.class);

      job.setMapOutputValueClass(IntWritable.class);

      job.setOutputKeyClass(Text.class);

      job.setOutputValueClass(IntWritable.class);


      if (job.waitForCompletion(true)) {

        return 0;

      }

      return 1;

    }

    public static void main(String[] args) throws Exception {

      int exitCode = ToolRunner.run(new WordCount(), args);

      System.exit(exitCode);

    }

}

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.mapreduce.Mapper;

public class WordMapper extends Mapper<LongWritable, Text, Text, IntWritable> {

  public void map(LongWritable key, Text value, Context context)

      throws IOException, InterruptedException {

    String s = value.toString();

    for (String word : s.split("\\W+")) {
```

```java
    if (word.length() > 0) {

      context.write(new Text(word), new IntWritable(1));

    }

   }

 }

}

import java.io.IOException;

import org.apache.hadoop.io.IntWritable;

import org.apache.hadoop.io.Text;

import org.apache.hadoop.mapreduce.Reducer;

public class SumReducer extends Reducer<Text, IntWritable, Text, IntWritable> {

 public void reduce(Text key, Iterable<IntWritable> values, Context context)

    throws IOException, InterruptedException {

  int wordCount = 0;

  for (IntWritable value : values) {

   wordCount += value.get();

  }

  context.write(key, new IntWritable(wordCount));

 }

}
```

WordCount Output

and   1

hadoop 1

is   1

mapreduce   1

powerful   1

simple 1

```
(base) student@composnlab-114:~/Desktop/23232$ su - hadoopuser
Password:
hadoopuser@composnlab-114:~$ cd ~/wordcount
hadoopuser@composnlab-114:~/wordcount$ hadoop fs -rm -r /user/hadoopuser/output
2025-04-17 13:14:40,675 WARN util.NativeCodeLoader: Unable to load native-hadoop library for your platform... using builtin-java classes where applicable
rm: Call From composnlab-114/172.16.33.114 to localhost:9000 failed on connection exception: java.net.ConnectException: Connection refused; For more details see:  http://wiki.apache.org/hadoop/ConnectionRefused
hadoopuser@composnlab-114:~/wordcount$ jps
9505 Jps
hadoopuser@composnlab-114:~/wordcount$ start-dfs.sh
start-yarn.sh
Starting namenodes on [localhost]
Starting datanodes
Starting secondary namenodes [composnlab-114]
2025-04-17 13:15:21,069 WARN util.NativeCodeLoader: Unable to load native-hadoop library for your platform... using builtin-java classes where applicable
Starting resourcemanager
Starting nodemanagers
hadoopuser@composnlab-114:~/wordcount$ hadoop fs -rm -r /user/hadoopuser/output
2025-04-17 13:15:37,422 WARN util.NativeCodeLoader: Unable to load native-hadoop library for your platform... using builtin-java classes where applicable
rm: Cannot delete /user/hadoopuser/output. Name node is in safe mode.
hadoopuser@composnlab-114:~/wordcount$ hdfs dfsadmin -safemode leave
2025-04-17 13:16:01,361 WARN util.NativeCodeLoader: Unable to load native-hadoop library for your platform... using builtin-java classes where applicable
Safe mode is OFF
hadoopuser@composnlab-114:~/wordcount$ hadoop fs -ls /user/hadoopuser/input
2025-04-17 13:16:13,690 WARN util.NativeCodeLoader: Unable to load native-hadoop library for your platform... using builtin-java classes where applicable
Found 1 items
-rw-r--r--   1 hadoopuser supergroup         40 2025-04-09 13:44 /user/hadoopuser/input/sample.txt
hadoopuser@composnlab-114:~/wordcount$ hadoop jar wordcount.jar WordCount /user/hadoopuser/input /user/hadoopuser/output
2025-04-17 13:16:22,076 WARN util.NativeCodeLoader: Unable to load native-hadoop library for your platform... using builtin-java classes where applicable
2025-04-17 13:16:22,438 INFO client.DefaultNoHARMFailoverProxyProvider: Connecting to ResourceManager at localhost/127.0.0.1:8032
Exception in thread "main" org.apache.hadoop.mapred.FileAlreadyExistsException: Output directory hdfs://localhost:9000/user/hadoopuser/output already exists
        at org.apache.hadoop.mapreduce.lib.output.FileOutputFormat.checkOutputSpecs(FileOutputFormat.java:164)
        at org.apache.hadoop.mapreduce.JobSubmitter.checkSpecs(JobSubmitter.java:277)
        at org.apache.hadoop.mapreduce.JobSubmitter.submitJobInternal(JobSubmitter.java:143)
        at org.apache.hadoop.mapreduce.Job$11.run(Job.java:1678)
        at org.apache.hadoop.mapreduce.Job$11.run(Job.java:1675)
        at java.base/java.security.AccessController.doPrivileged(Native Method)
        at java.base/javax.security.auth.Subject.doAs(Subject.java:423)
        at org.apache.hadoop.security.UserGroupInformation.doAs(UserGroupInformation.java:1899)
        at org.apache.hadoop.mapreduce.Job.submit(Job.java:1675)
        at org.apache.hadoop.mapreduce.Job.waitForCompletion(Job.java:1696)
        at WordCount.run(WordCount.java:45)
        at org.apache.hadoop.util.ToolRunner.run(ToolRunner.java:82)
        at org.apache.hadoop.util.ToolRunner.run(ToolRunner.java:97)
        at WordCount.main(WordCount.java:49)
        at java.base/jdk.internal.reflect.NativeMethodAccessorImpl.invoke0(Native Method)
        at java.base/jdk.internal.reflect.NativeMethodAccessorImpl.invoke(NativeMethodAccessorImpl.java:62)
        at java.base/jdk.internal.reflect.DelegatingMethodAccessorImpl.invoke(DelegatingMethodAccessorImpl.java:43)
        at java.base/java.lang.reflect.Method.invoke(Method.java:566)
        at org.apache.hadoop.util.RunJar.run(RunJar.java:328)
        at org.apache.hadoop.util.RunJar.main(RunJar.java:241)
hadoopuser@composnlab-114:~/wordcount$ hadoop fs -cat /user/hadoopuser/output/part-r-00000
2025-04-17 13:16:36,345 WARN util.NativeCodeLoader: Unable to load native-hadoop library for your platform... using builtin-java classes where applicable
and     1
hadoop  1
is      1
mapreduce       1
powerful        1
simple  1
hadoopuser@composnlab-114:~/wordcount$
```