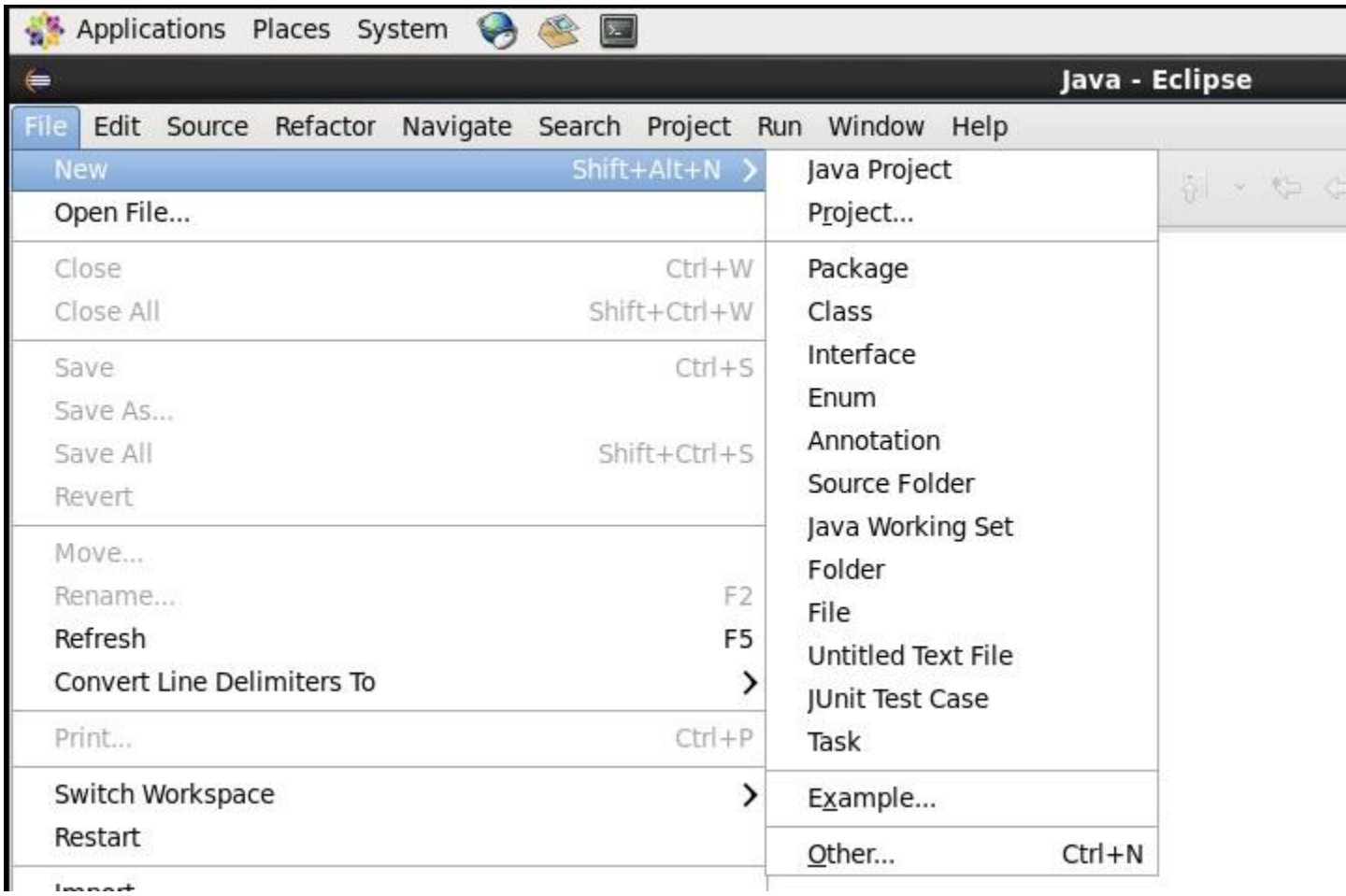
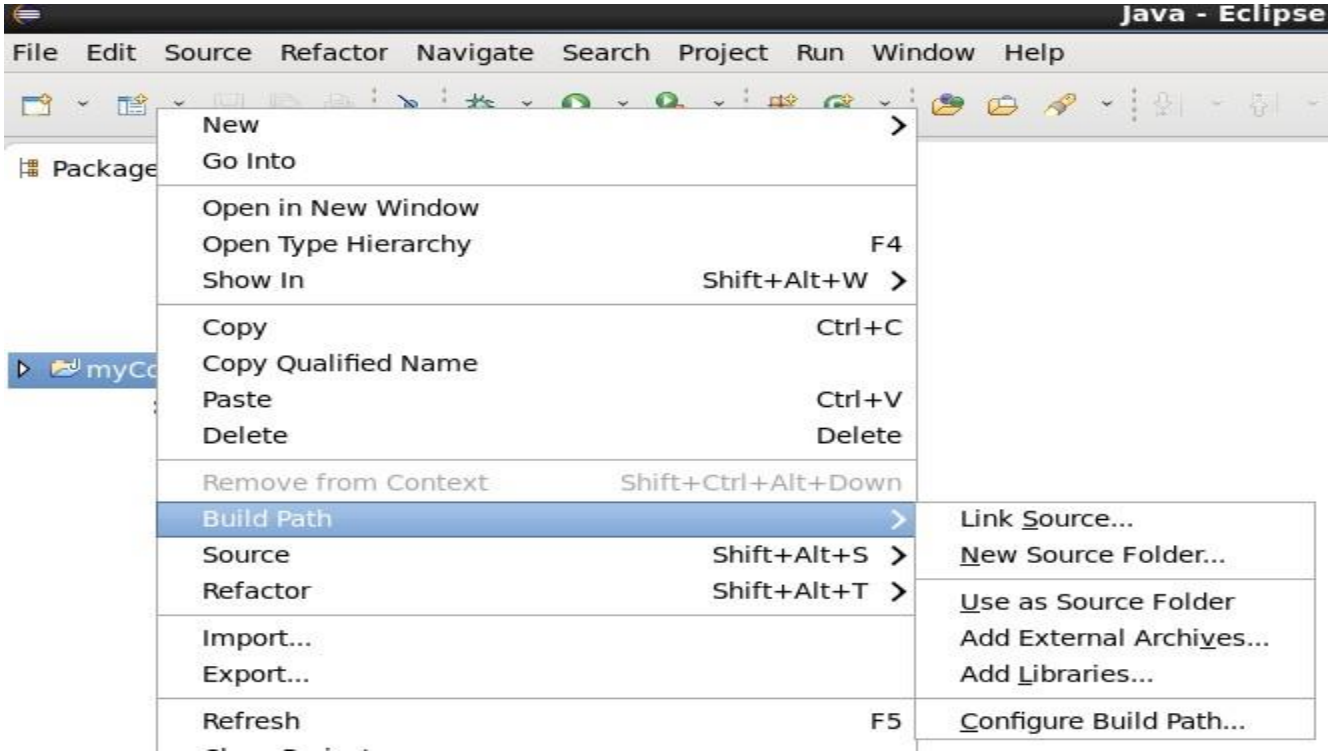
ExperimentNo:11

Steps :

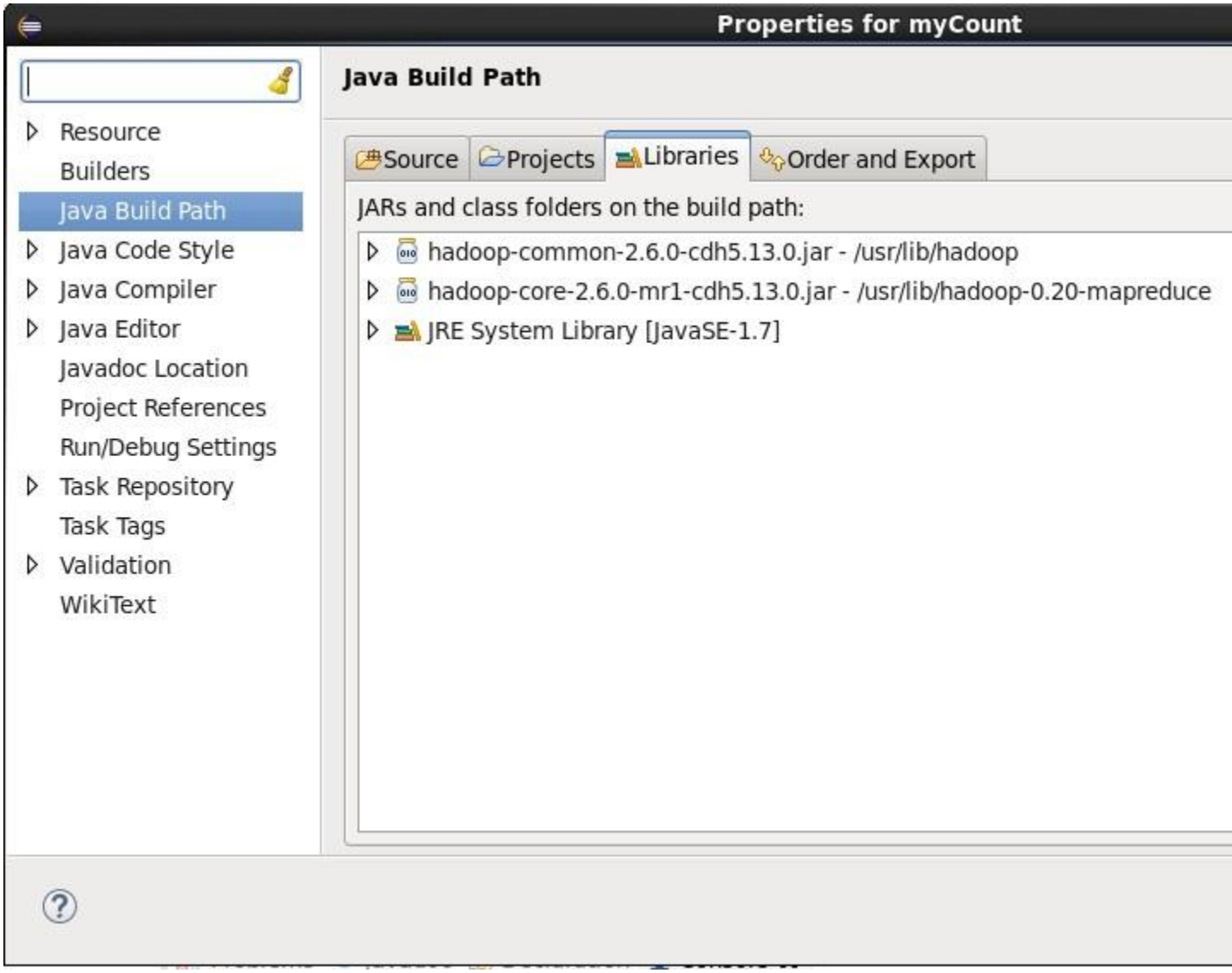
- FirstOpen**Eclipse**->thenselect**File**->**New**->**JavaProject**->Nameit**WordCount**->then**Finish**.



- CreateThreeJavaClassesintotheproject.Namethem**WCDriver**(havingthemainfunction),**WCMapper**,**WCReducer**.
- YouhavetoincludetwoReferenceLibrariesforthat: RightClickon**Project**->thenselect**BuildPath**->Clickon**ConfigureBuildPath**



- IntheabovefigureyoucanseetheAddExternalJARsoptionontheRightHand Side. Click on it and add the below mention files. You can find thesefilesin*/usr/lib/*
  1. /usr/lib/hadoop-0.20-mapreduce/hadoop-core-2.6.0-mr1-cdh5.13.0.jar
  2. /usr/lib/hadoop/hadoop-common-2.6.0-cdh5.13.0.jar



- **MapperCode:**YouhavetocopypastethisprogramintotheWCMapperJavaClassfile.

- Java

```
// Importing

librariesimportjava.io.IOEx

ception;
```

```
import
org.apache.hadoop.io.IntWritable;importorg.apache.h
adoop.io.LongWritable;importorg.apache.hadoop.io.Te
xt;
importorg.apache.hadoop.mapred.MapReduceBase;import
org.apache.hadoop.mapred.Mapper;
importorg.apache.hadoop.mapred.OutputCollector;importorg.ap
ache.hadoop.mapred.Reporter;


publicclassWCMapperextendsMapReduceBaseimplementsMapper<LongWritable,
                                                    Text,Text,IntWritable>{

    //Mapfunction
    publicvoidmap(LongWritablekey,Textvalue,OutputCollector<Text,IntWritable>output,Report
                  errep)throwsIOException
    {

        Stringline=value.toString();

        //Splittingthelineonspaces
```

```
        for(Stringword:line.split(""))
        {
            if(word.length() >0)
            {
                output.collect(newText(word),newIntWritable(1));
            }
        }
    }
}
```

ReducerCode: YouhavetocopypastethisprogramintotheWCReducerJavaClassfile.

- Java

```
// Importing
librariesimportjava.io.IOEx
ception;importjava.util.Ite
rator;
importorg.apache.hadoop.io.IntWritable;im
portorg.apache.hadoop.io.Text;
importorg.apache.hadoop.mapred.MapReduceBase;
```

```
importorg.apache.hadoop.mapred.OutputCollector;import
org.apache.hadoop.mapred.Reducer;
importorg.apache.hadoop.mapred.Reporter;

publicclassWCReducerextendsMapReduceBaseimplementsReducer<Text,
                                    IntWritable,Text,IntWritable>{

    //Reducefunction
    publicvoidreduce(Textkey,
                    Iterator<IntWritable>value,OutputCollecto
                    r<Text,IntWritable>output,
                              Reporterrep)throwsIOException
    {

        intcount=0;

        //Countingthefrequencyofeachwordswhile
        (value.hasNext())
        {
            IntWritablei=value.next();
```

```
            count+=i.get();

        }



        output.collect(key,newIntWritable(count));

    }

}
```

DriverCode: YouhavetocopypastethisprogramintotheWCDriverJavaClassfile.

- Java

```
// Importing

librariesimportjava.io.IOEx

ception;

importorg.apache.hadoop.conf.Configured;import

org.apache.hadoop.fs.Path;

importorg.apache.hadoop.io.IntWritable;im

portorg.apache.hadoop.io.Text;

import

org.apache.hadoop.mapred.FileInputFormat;importorg.ap

ache.hadoop.mapred.FileOutputFormat;importorg.apache.

hadoop.mapred.JobClient;
```

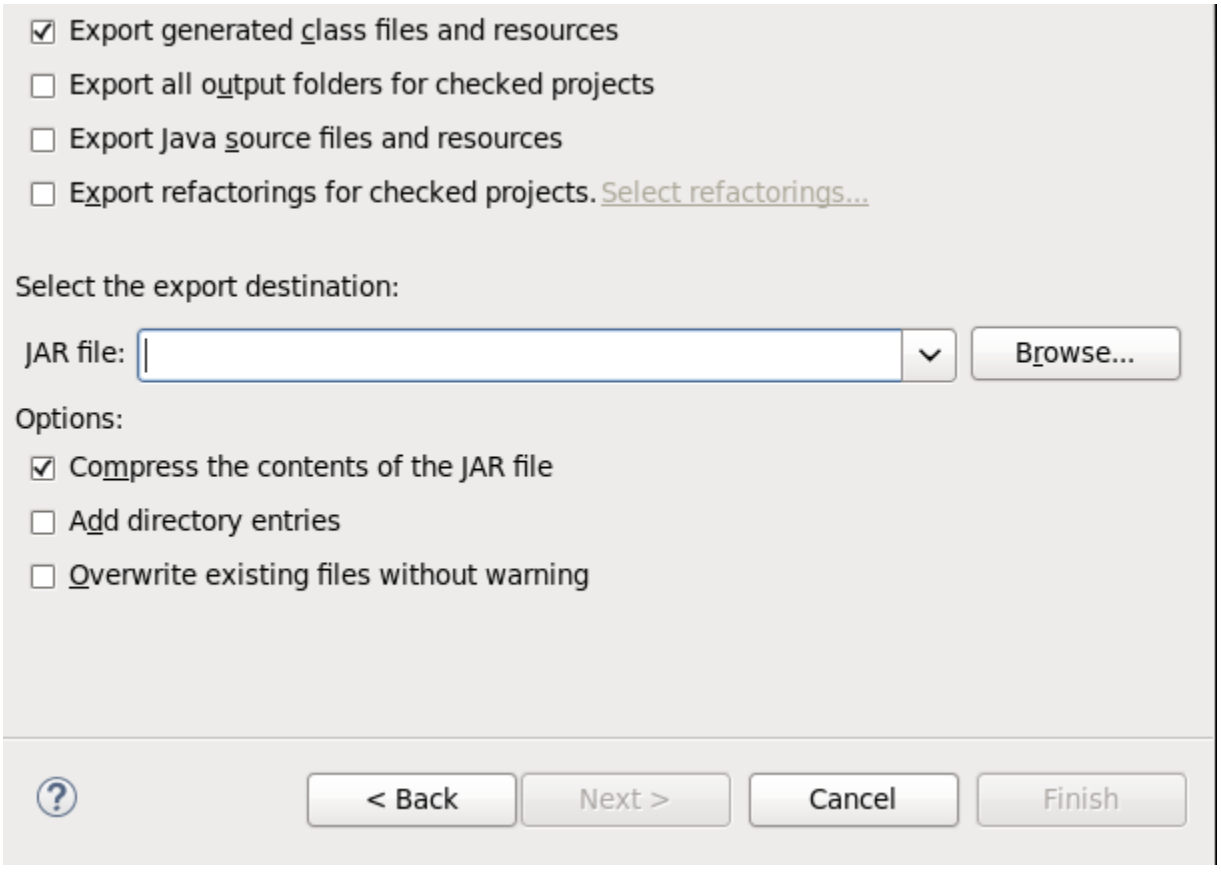
```
import
org.apache.hadoop.mapred.JobConf;import
org.apache.hadoop.util.Tool;importorg.ap
ache.hadoop.util.ToolRunner;

publicclassWCDriverextendsConfiguredimplementsTool{

    publicintrun(Stringargs[])throwsIOException
    {
        if(args.length <2)
        {
            System.out.println("Pleasegivevalidinputs");re
            turn-1;
        }

        JobConf conf = new
        JobConf(WCDriver.class);FileInputFormat.setInputPaths(conf,
        newPath(args[0]));
        FileOutputFormat.setOutputPath(conf,newPath(args[1]));conf.setMapperCla
        ss(WCMapper.class);conf.setReducerClass(WCReducer.class);
conf.setMapOutputKeyClass(Text.class);conf.setMapOutputValu
```

```
eClass(IntWritable.class);conf.setOutputKeyClass(Text.class
);conf.setOutputValueClass(IntWritable.class);JobClient.run
Job(conf);
            return0;
        }
        //MainMethod
        publicstaticvoidmain(Stringargs[])throwsException
        {
            intexitCode=ToolRunner.run(newWCDriver(),args);System.out.println(exitC
            ode);
        }
    }
```



```
cloudera@quickstart:~/workspace
File  Edit  View  Search  Terminal  Help
[cloudera@quickstart workspace]$ hadoop fs -put WCFile.txt WCFile.txt
[cloudera@quickstart workspace]$
```

- Nowtorunthejarfilebywritingthecodeasshowninthescreenshot.

```
cloudera@quickstart:~/workspace
File  Edit  View  Search  Terminal  Help
[cloudera@quickstart workspace]$ hadoop jar wordCount.jar WCDriver WCFile.txt WCOutput
19/05/06 22:43:22 INFO client.RMProxy: Connecting to ResourceManager at /0.0.0.0:8032
19/05/06 22:43:22 INFO client.RMProxy: Connecting to ResourceManager at /0.0.0.0:8032
```

- AfterExecutingthecode,youcanseetheresultin*WCOutput*fileorbywritingfollowing commandonterminal.

```
hadoop fs -cat WCOutput/part-00000
```

-

```
cloudera@quickstart:~/workspace
File  Edit  View  Search  Terminal  Help
[cloudera@quickstart workspace]$ hadoop fs -cat WCOutput/part-00000
GeeksforGeeks   1
Hello   2
I       2
Intern  1
am      2
an      1
```