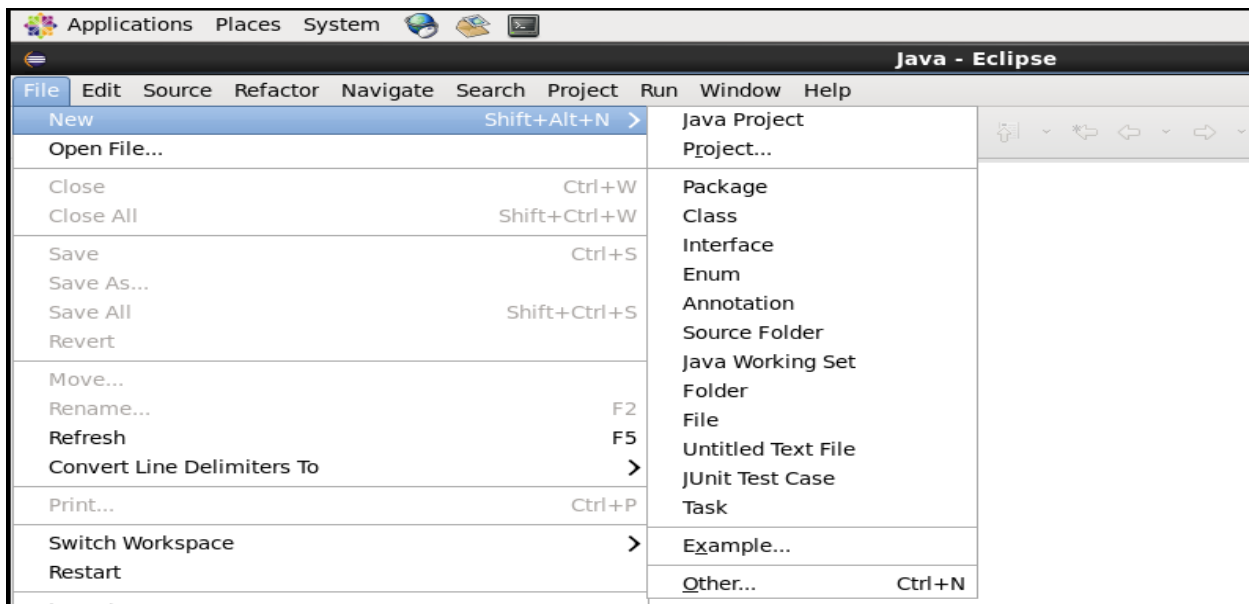


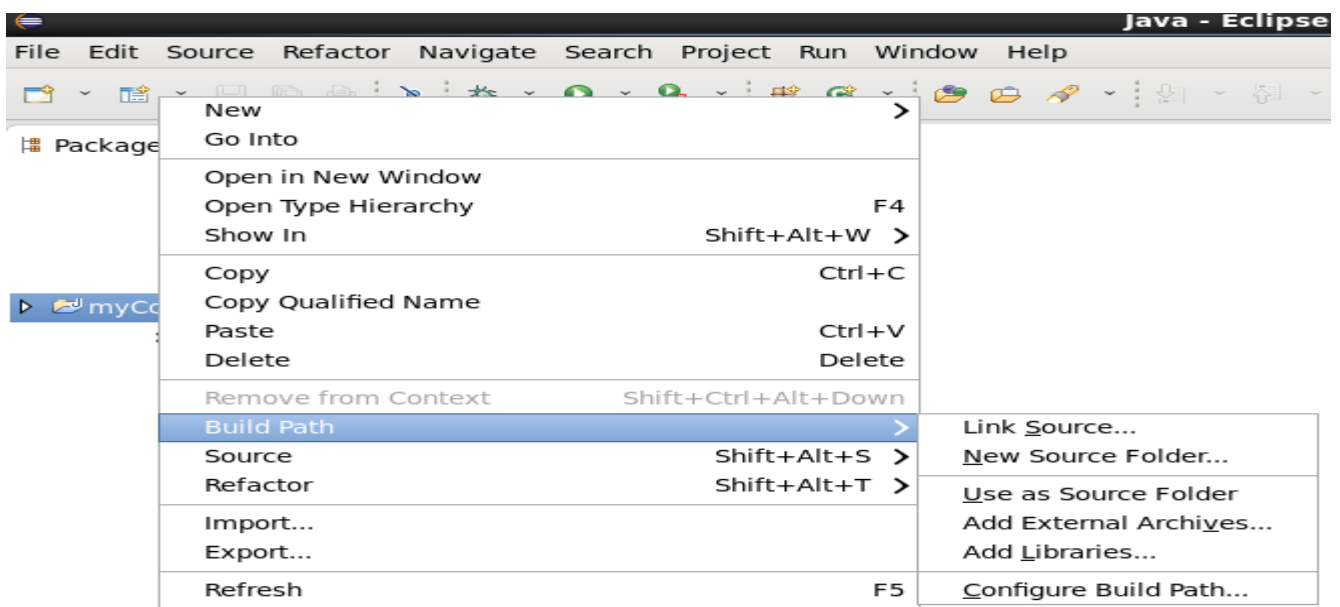
Jar File Creation Steps:

1. First Open Eclipse -> then select File->New->Java Project->Name it WordCount-> then Finish.



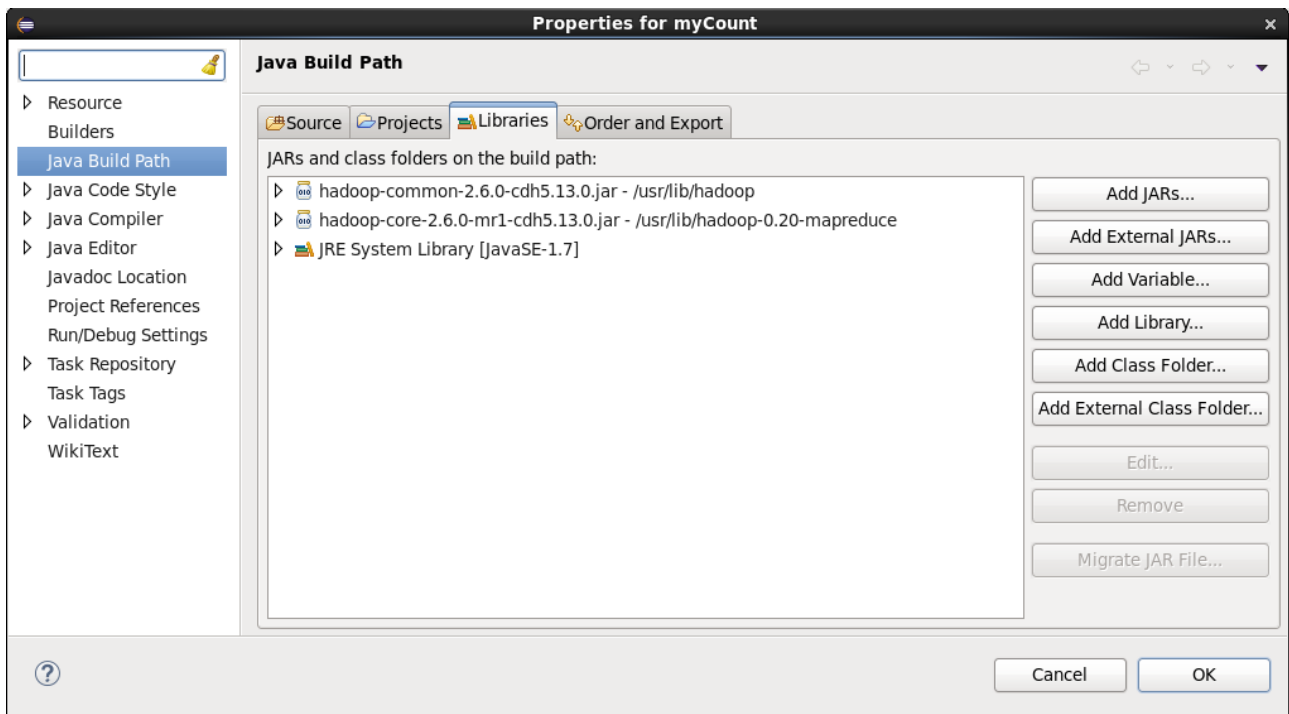
2. You have to include Reference Libraries for that:

Right Click on Project-> then select Build Path-> Click on Configure Build Path

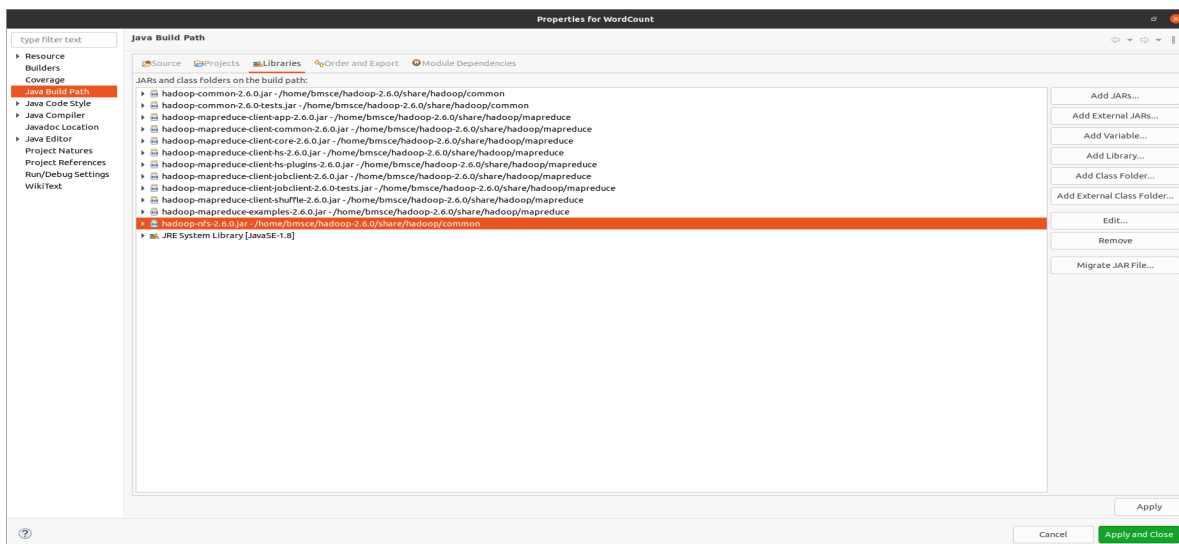


In the above figure, you can see the Add External JARs option on the Right Hand Side. Click on it and add the below mention files

1. /home/bmsce/hadoop-2.6.0/share/hadoop/mapreduce/
2. /home/bmsce/hadoop-2.6.0/share/common/



3. Select JRE System Library- goto edit- select JavaSE-1.8 and click apply



4. Create Three Java Classes into the project. Name them WCDriver(having the main function), WCMapper, WCReducer.

```
// Importing libraries
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));
            }
        }
    }
}

```

Reducer Code: You have to copy paste this program into the WCReducer Java Class file.

```

// Importing libraries
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));
    }
}

```

Driver Code: You have to copy paste this program into the WCDriver Java Class file.

```

// Importing libraries
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 {

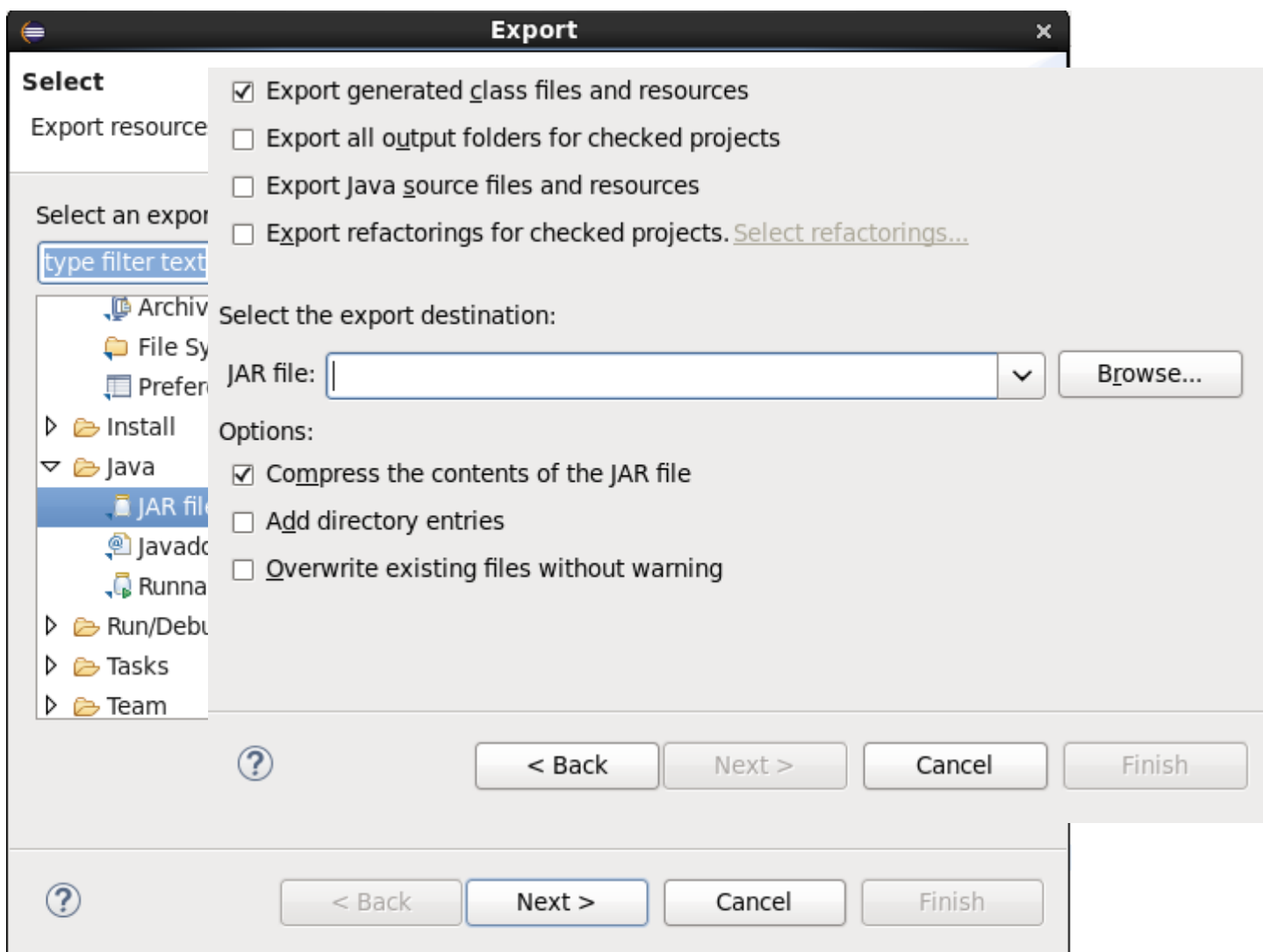
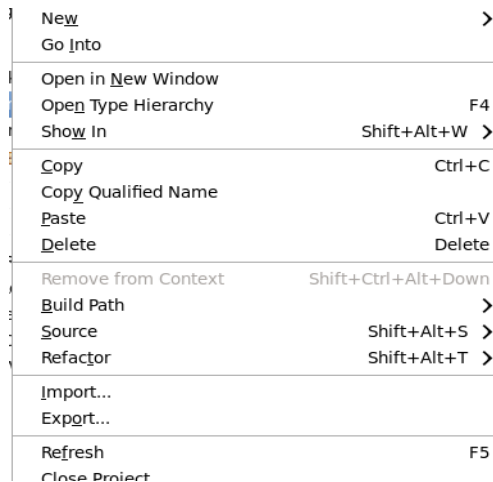
    public int run(String args[]) throws IOException
    {
        if (args.length < 2)
        {
            System.out.println("Please give valid inputs");
            return -1;
        }

        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);
    }
}

```

Now you have to make a jar file. Right Click on Project-> Click on Export-> Select export destination as Jar File-> Name the jar File(WordCount.jar) -> Click on next -> at last Click on Finish.



To run Mapreduce Program

1. `hduser@bmsce-Precision-T1700:/$ su hduser`

Password:

`hduser@bmsce-Precision-T1700:/$ cd /`

`hduser@bmsce-Precision-T1700:/$ cd /usr/local/hadoop/sbin`

2. `hduser@bmsce-Precision-T1700:/usr/local/hadoop/sbin$ start-all.sh`

3. `hduser@bmsce-Precision-T1700:/usr/local/hadoop/sbin$ jps`

4. create a file on Desktop(sample.txt) and type the below lines:

```
hi how are you
how is your job
how is your family
how is your brother
how is your sister
```

save your file

5. View the directory content

```
hadoop fs -ls /
```

6. Create a directory using the following command. If any directory existing, use the same directory for the command

```
hadoop fs -mkdir /rgs
```

7. Copy the file into HDFS

```
hadoop fs -copyFromLocal D:/sample.txt /rgs/test.txt
```

8. Run the Map Reduce Program

```
hadoop jar /home/hduser/Desktop/Jwordcount.jar WCDriver input output
```

9. `hadoop fs -ls /output/`

10. `hadoop fs -cat /output/part-00000`

