

Assignment No 12

SalesCountryDriver.java

```
package SalesCountry;
```

```
import org.apache.hadoop.fs.Path;
import org.apache.hadoop.io.*;
import org.apache.hadoop.mapred.*;
```

```
public class SalesCountryDriver {
    public static void main(String[] args) {
        JobClient my_client = new JobClient();
        // Create a configuration object for the job
        JobConf job_conf = new JobConf(SalesCountryDriver.class);

        // Set a name of the Job
        job_conf.setJobName("SalePerCountry");

        // Specify data type of output key and value
        job_conf.setOutputKeyClass(Text.class);
        job_conf.setOutputValueClass(IntWritable.class);

        // Specify names of Mapper and Reducer Class
        job_conf.setMapperClass(SalesCountry.SalesMapper.class);
        job_conf.setReducerClass(SalesCountry.SalesCountryReducer.class);

        // Specify formats of the data type of Input and output
        job_conf.setInputFormat(TextInputFormat.class);
        job_conf.setOutputFormat(TextOutputFormat.class);

        // Set input and output directories using command line arguments,
        //arg[0] = name of input directory on HDFS, and arg[1] = name of output
        //directory to be created to store the output file.

        FileInputFormat.setInputPaths(job_conf, new Path(args[0]));
        FileOutputFormat.setOutputPath(job_conf, new Path(args[1]));

        my_client.setConf(job_conf);
        try {
            // Run the job
            JobClient.runJob(job_conf);
        }
    }
}
```

```
        } catch (Exception e) {
            e.printStackTrace();
        }
    }}
}
```

SalesCountryReducer.java

```
package SalesCountry;
```

```
import java.io.IOException;
import java.util.*;
```

```
import org.apache.hadoop.io.IntWritable;
import org.apache.hadoop.io.Text;
import org.apache.hadoop.mapred.*;
```

```
public class SalesCountryReducer extends MapReduceBase implements Reducer<Text,
IntWritable, Text, IntWritable> {
```

```
    public void reduce(Text t_key, Iterator<IntWritable> values,
OutputCollector<Text,IntWritable> output, Reporter reporter) throws IOException {
        Text key = t_key;
        int frequencyForCountry = 0;
        while (values.hasNext()) {
            // replace type of value with the actual type of our value
            IntWritable value = (IntWritable) values.next();
            frequencyForCountry += value.get();
        }
        output.collect(key, new IntWritable(frequencyForCountry));
    }
}
```

SalesMapper.java

```
package SalesCountry;
```

```
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.*;
```

```

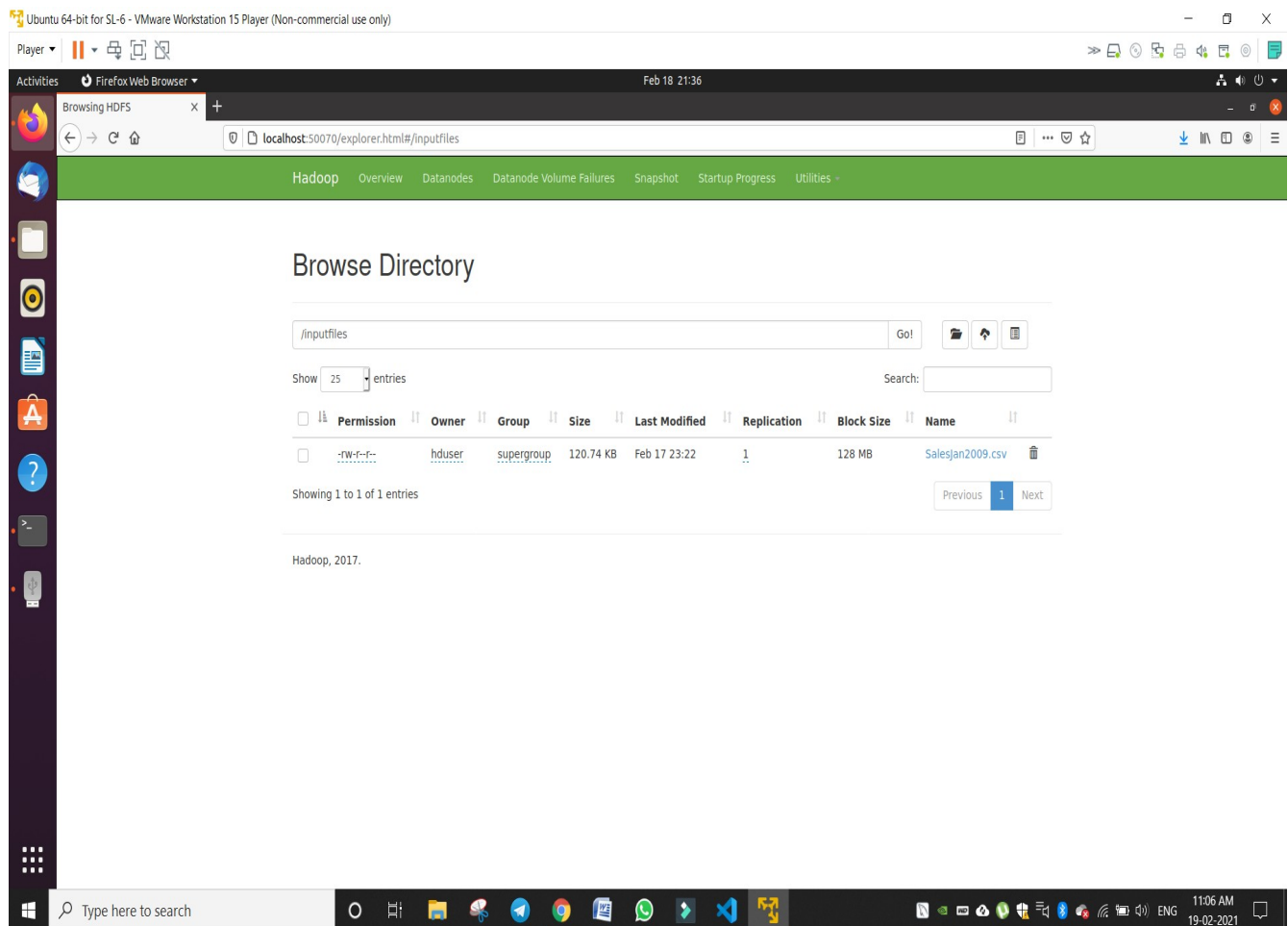
public class SalesMapper extends MapReduceBase implements Mapper<LongWritable, Text,
Text, IntWritable> {
    private final static IntWritable one = new IntWritable(1);

    public void map(LongWritable key, Text value, OutputCollector<Text, IntWritable>
output, Reporter reporter) throws IOException {

        String valueString = value.toString();
        String[] SingleCountryData = valueString.split(",");
        output.collect(new Text(SingleCountryData[7]), one);
    }
}

```

Output



Ubuntu 64-bit for SL-6 - VMware Workstation 15 Player (Non-commercial use only)

Player ▾ | Feb 18 21:35

Activities | Firefox Web Browser | Browsing HDFS | localhost:50070/explorer.html#/

Hadoop Overview Datanodes Datanode Volume Failures Snapshot Startup Progress Utilities ▾

Browse Directory

/

Show 25 entries Search:

<input type="checkbox"/>	Permission	Owner	Group	Size	Last Modified	Replication	Block Size	Name	<input type="button" value="🗑"/>
<input type="checkbox"/>	drwxr-xr-x	hduser	supergroup	0 B	Feb 18 06:03	0	0 B	input3000	<input type="button" value="🗑"/>
<input type="checkbox"/>	drwxr-xr-x	hduser	supergroup	0 B	Feb 17 23:22	0	0 B	inputfiles	<input type="button" value="🗑"/>
<input type="checkbox"/>	drwxr-xr-x	hduser	supergroup	0 B	Feb 18 06:05	0	0 B	output3000	<input type="button" value="🗑"/>
<input type="checkbox"/>	drwxr-xr-x	hduser	supergroup	0 B	Feb 17 23:34	0	0 B	outputfiles	<input type="button" value="🗑"/>
<input type="checkbox"/>	drwx-----	hduser	supergroup	0 B	Feb 17 23:20	0	0 B	tmp	<input type="button" value="🗑"/>

Showing 1 to 5 of 5 entries **1**

Hadoop, 2017.

Type here to search | 11:05 AM 19-02-2021

Ubuntu 64-bit for SL-6 - VMware Workstation 15 Player (Non-commercial use only)

Player ▾ | Feb 18 21:37

Activities | Firefox Web Browser | Browsing HDFS | localhost:50070/explorer.html#/outputfiles

Hadoop Overview Datanodes Datanode Volume Failures Snapshot Startup Progress Utilities ▾

Browse Directory

/outputfiles

Show 25 entries Search:

<input type="checkbox"/>	Permission	Owner	Group	Size	Last Modified	Replication	Block Size	Name	<input type="button" value="🗑"/>
<input type="checkbox"/>	-rw-r--r--	hduser	supergroup	0 B	Feb 17 23:34	1	128 MB	_SUCCESS	<input type="button" value="🗑"/>
<input type="checkbox"/>	-rw-r--r--	hduser	supergroup	661 B	Feb 17 23:34	1	128 MB	part-00000	<input type="button" value="🗑"/>

Showing 1 to 2 of 2 entries **1**

Hadoop, 2017.

Type here to search | 11:07 AM 19-02-2021