Experiment 4: Map-Reduce Program

Program:-

```java
package supreetha;

import java.io.IOException;
import java.util.*;
import org.apache.hadoop.fs.Path;
import org.apache.hadoop.io.*;
import org.apache.hadoop.mapred.*;
public class TransactionCount {
//MAPPER CODE
public static class Map extends MapReduceBase implements
Mapper<LongWritable, Text, Text, IntWritable> {
private final static IntWritable one = new IntWritable(1);
//private Text word = new Text();
public void map(LongWritable key, Text value, OutputCollector<Text,
IntWritable> output, Reporter reporter) throws IOException {
String myString = value.toString();
String[] userCount = myString.split(",");
output.collect(new Text(userCount[3]), one);
}
}
//REDUCER CODE
public static class Reduce extends MapReduceBase implements
Reducer<Text, IntWritable, Text, IntWritable> {
    public void reduce(Text key, Iterator<IntWritable> values,
            OutputCollector<Text, IntWritable> output, Reporter reporter) throws
            IOException { //{little: {1,1}}
            int finaluserCount = 0 ;
            Text mykey = key ;
            while(values.hasNext()) {
            IntWritable value = values.next();
            finaluserCount += value.get();
            }
            output.collect(mykey, new IntWritable(finaluserCount));
            }
            }
            //DRIVER CODE
            public static void main(String[] args) throws Exception {
            JobConf conf = new JobConf(TransactionCount.class);
            conf.setJobName("transactioncount");
            conf.setOutputKeyClass(Text.class);
            conf.setOutputValueClass(IntWritable.class);
            conf.setMapperClass(Map.class);
            conf.setCombinerClass(Reduce.class);
```
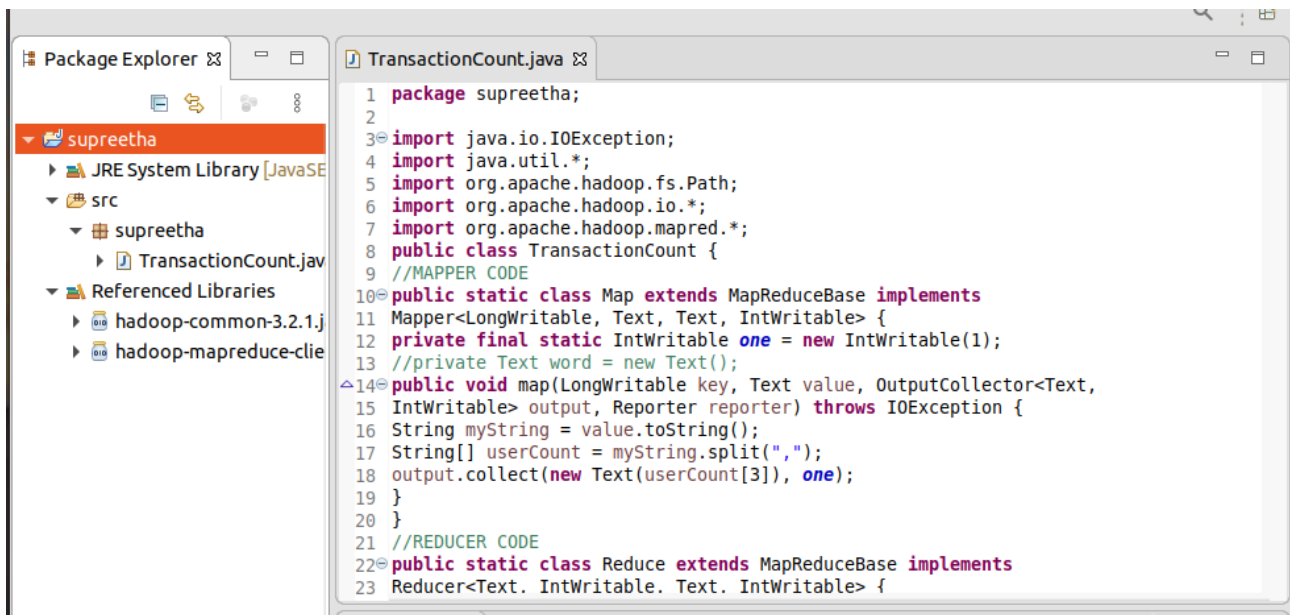
```
        conf.setReducerClass(Reduce.class);
        conf.setInputFormat(TextInputFormat.class);
        conf.setOutputFormat(TextOutputFormat.class);
        FileInputFormat.setInputPaths(conf, new Path(args[0]));
        FileOutputFormat.setOutputPath(conf, new Path(args[1]));
        JobClient.runJob(conf);
        }
    }
```



```
1  package supreetha;
2
3  import java.io.IOException;
4  import java.util.*;
5  import org.apache.hadoop.fs.Path;
6  import org.apache.hadoop.io.*;
7  import org.apache.hadoop.mapred.*;
8  public class TransactionCount {
9  //MAPPER CODE
10 public static class Map extends MapReduceBase implements
11 Mapper<LongWritable, Text, Text, IntWritable> {
12 private final static IntWritable one = new IntWritable(1);
13 //private Text word = new Text();
14 public void map(LongWritable key, Text value, OutputCollector<Text,
15 IntWritable> output, Reporter reporter) throws IOException {
16 String myString = value.toString();
17 String[] userCount = myString.split(",");
18 output.collect(new Text(userCount[3]), one);
19 }
20 }
21 //REDUCER CODE
22 public static class Reduce extends MapReduceBase implements
23 Reducer<Text. IntWritable. Text. IntWritable> {
```

.csv File:

	A	B	C	D	E
1	sl no	card name	user name	amount	
2	1	abc	mno	1000	
3	2	def	pqr	1000	
4	3	ghi	stu	2000	
5	4	jkl	vwx	2000	
6					
7					
8					
9					
10					

Commands:

$ cd $HADOOP_HOME

$ cd sbin

./start-all.sh

```
hdoop@admin1-HP-280-G4-MT-Business-PC:~$ cd $HADOOP_HOME
hdoop@admin1-HP-280-G4-MT-Business-PC:~/hadoop-3.2.1$ cd sbin
hdoop@admin1-HP-280-G4-MT-Business-PC:~/hadoop-3.2.1/sbin$ start-all.sh
WARNING: Attempting to start all Apache Hadoop daemons as hdoop in 10 seconds.
WARNING: This is not a recommended production deployment configuration.
WARNING: Use CTRL-C to abort.
Starting namenodes on [localhost]
```

hdfs dfs -mkdir -p /supreetha

hdfs dfs -copyFromLocal /home/hdoop/Desktop/1nt19is166_supreetha/exp_4.csv /supreetha

```
hdoop@admin1-HP-280-G4-MT-Business-PC:~/hadoop-3.2.1/sbin$ hdfs dfs -mkdir -p /supreetha
hdoop@admin1-HP-280-G4-MT-Business-PC:~/hadoop-3.2.1/sbin$ hdfs dfs -copyFromLocal /home/hdoop/Desktop/1nt19is166_supreetha/exp_4.csv /supreetha
2022-06-23 14:09:46,882 INFO sasl.SaslDataTransferClient: SASL encryption trust check: localHostTrusted = false, remoteHostTrusted = false
```

hadoop jar /home/hdoop/Desktop/supreetha1.jar /supreetha /supreethaop

```
hdoop@admin1-HP-280-G4-MT-Business-PC:~/hadoop-3.2.1/sbin$ hadoop jar /home/hdoop/Desktop/supreetha1.jar /supreetha /supreethaop
2022-06-23 14:11:05,288 INFO client.RMProxy: Connecting to ResourceManager at /127.0.0.1:8032
2022-06-23 14:11:05,415 INFO client.RMProxy: Connecting to ResourceManager at /127.0.0.1:8032
2022-06-23 14:11:05,549 WARN mapreduce.JobResourceUploader: Hadoop command-line option parsing not performed. Implement the Tool interface and execute your applicat
ion with ToolRunner to remedy this.
2022-06-23 14:11:05,573 INFO mapreduce.JobResourceUploader: Disabling Erasure Coding for path: /tmp/hadoop-yarn/staging/hdoop/.staging/job_1655973226068_0001
2022-06-23 14:11:05,680 INFO sasl.SaslDataTransferClient: SASL encryption trust check: localHostTrusted = false, remoteHostTrusted = false
2022-06-23 14:11:05,780 INFO mapred.FileInputFormat: Total input files to process : 1
2022-06-23 14:11:05,830 INFO sasl.SaslDataTransferClient: SASL encryption trust check: localHostTrusted = false, remoteHostTrusted = false
2022-06-23 14:11:05,873 INFO sasl.SaslDataTransferClient: SASL encryption trust check: localHostTrusted = false, remoteHostTrusted = false
2022-06-23 14:11:05,897 INFO mapreduce.JobSubmitter: number of splits:2
2022-06-23 14:11:06,039 INFO sasl.SaslDataTransferClient: SASL encryption trust check: localHostTrusted = false, remoteHostTrusted = false
2022-06-23 14:11:06,055 INFO mapreduce.JobSubmitter: Submitting tokens for job: job_1655973226068_0001
2022-06-23 14:11:06,055 INFO mapreduce.JobSubmitter: Executing with tokens: []
```

```
2022-06-21 09:31:18,662 INFO mapreduce.Job: Job job_1655783592787_0001 running in uber
mode : false
2022-06-21 09:31:18,664 INFO mapreduce.Job:  map 0% reduce 0%
2022-06-21 09:31:22,742 INFO mapreduce.Job:  map 100% reduce 0%
2022-06-21 09:31:25,763 INFO mapreduce.Job:  map 100% reduce 100%
2022-06-21 09:31:26,782 INFO mapreduce.Job: Job job_1655783592787_0001 completed succes
sfully
2022-06-21 09:31:26,838 INFO mapreduce.Job: Counters: 54
```

hdfs dfs -cat /supreethaop/part*

```
hdoop@admin1-HP-280-G4-MT-Business-PC:~/hadoop-3.2.1/sbin$ hdfs dfs -cat /supreethaop/part*
2022-06-23 14:12:21,902 INFO sasl.SaslDataTransferClient: SASL encryption trust check: localHostTrusted = false, remoteHostTrusted = false
1000    2
2000    2
amount  1
hdoop@admin1-HP-280-G4-MT-Business-PC:~/hadoop-3.2.1/sbin$
```