**Name:Suraj Kumar**

**Usn no:1NT19IS167**

**Experiment 4: Map Reduce Program**

package suraj;

mport java.io.IOException;

import java.util.\*;

import org.apache.hadoop.fs.Path;

import org.apache.hadoop.io.\*;

import org.apache.hadoop.mapred.\*;

public class WordCount {

//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(WordCount.class);

conf.setJobName("wordcount");

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);

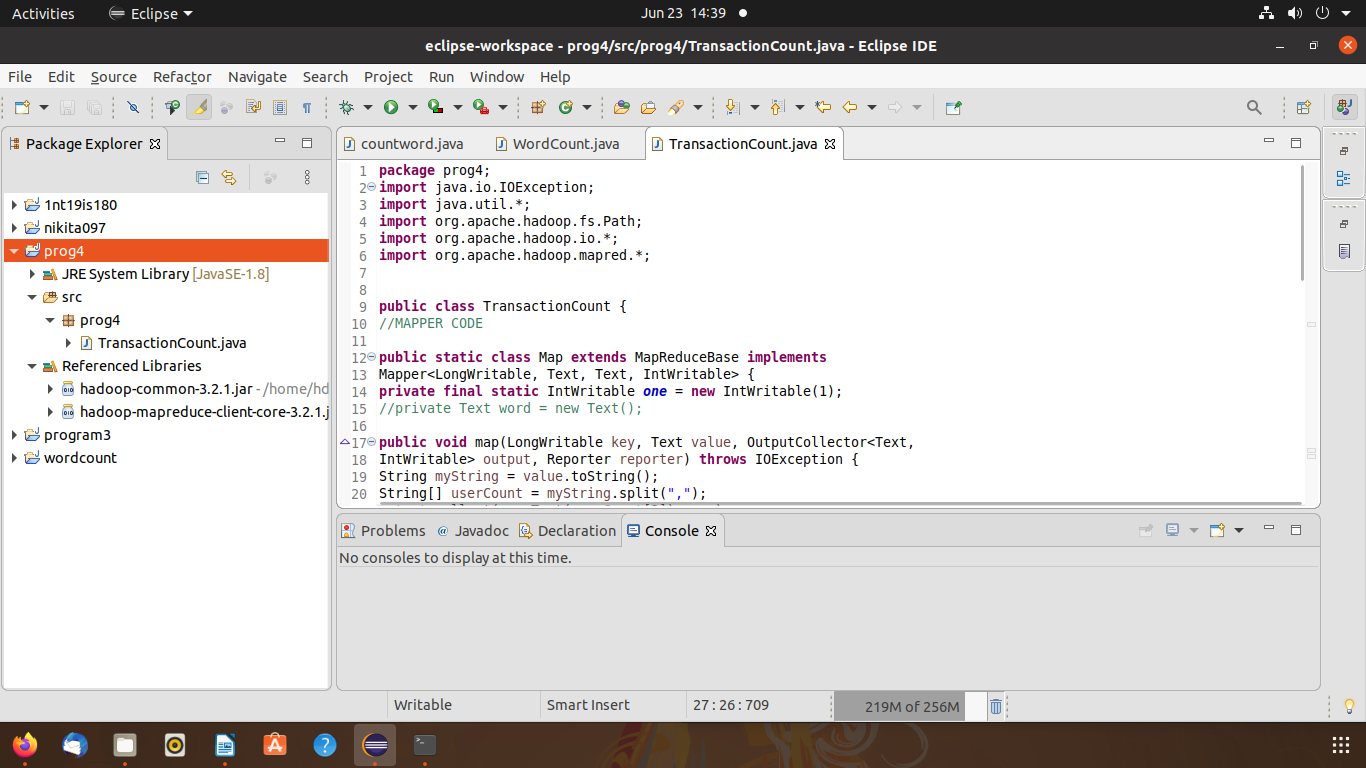
FileInputFormat.setInputPaths(conf, new Path(args[0]));

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

JobClient.runJob(conf);

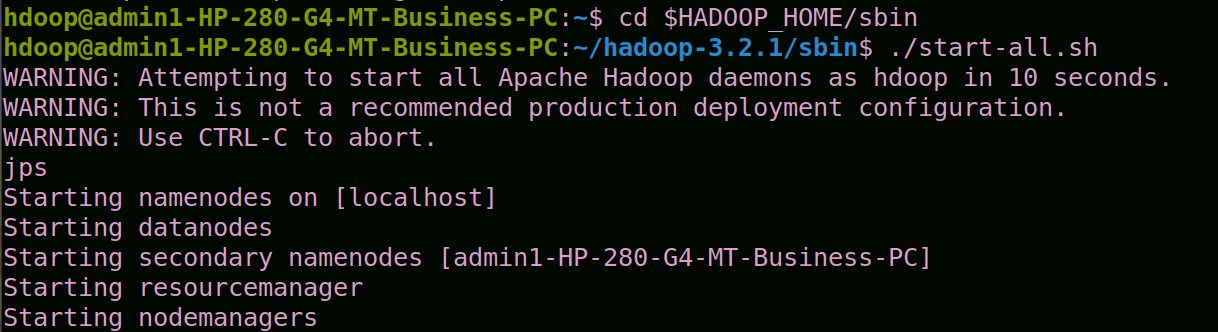
}

}



$ cd $HADOOP\_HOME/sbin

./start-all.sh

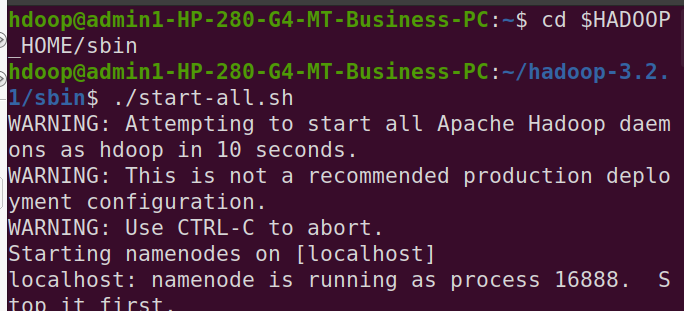


hdfs dfs -mkdir -p /suraj

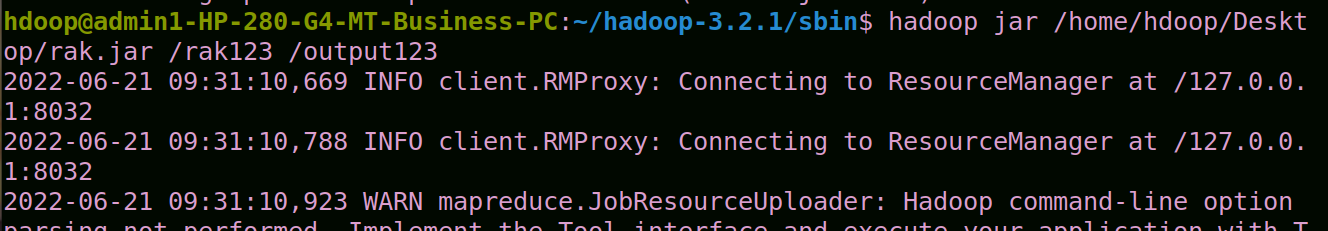
hdfs dfs -copyFromLocal /home/hdoop/Documents/suraj.csv /suraj

cd $HADOOP\_HOME/sbin

./start-all.sh



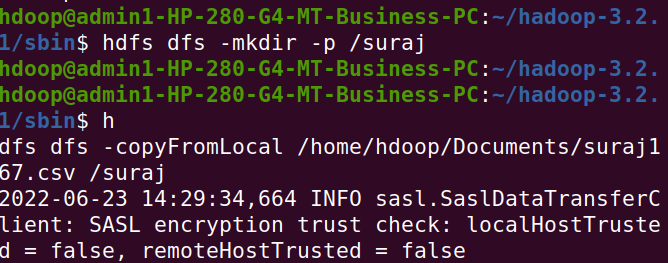
hadoop jar /home/hdoop/Desktop/rak.jar /rak123 /output



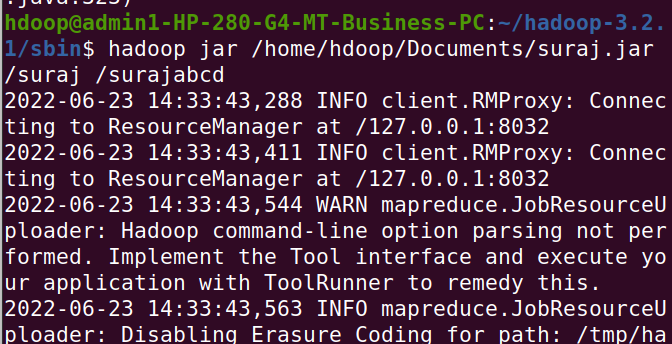
hdfs dfs -cat /output123/part\*

hdfs dfs -mkdir -p /suraj

hdfs dfs -copyFromLocal /home/hdoop/Documents/suraj167.csv /suraj



hadoop jar /home/hdoop/Documents/suraj.jar /suraj /surajabcd



hdfs dfs -cat /surajabcd/part\*

