

```

import java.io.*;
import org.apache.hadoop.conf.Configuration;
import org.apache.hadoop.fs.Path;
import org.apache.hadoop.io.Text;
import org.apache.hadoop.fs.FileSystem;
import org.apache.hadoop.io.LongWritable;
import org.apache.hadoop.mapreduce.Job;
import org.apache.hadoop.mapreduce.Mapper;
import org.apache.hadoop.mapreduce.Reducer;
import org.apache.hadoop.mapreduce.lib.input.MultipleInputs;
import org.apache.hadoop.mapreduce.lib.input.FileInputFormat;
import org.apache.hadoop.mapreduce.lib.input.TextInputFormat;
import org.apache.hadoop.mapreduce.lib.output.FileOutputFormat;

public class StateRegionWiseTotalSalesPrice {
    public static class StateRegionWiseTotalSalesPriceMapper extends Mapper<LongWritable,
Text, Text,Text> {
        public void map(LongWritable key ,Text value, Context con)throws
IOException,InterruptedException{

            String line = value.toString();
            String[] record = line.split(",");
            con.write(new Text(record[9]),new Text("sum\t" + record[15]));
        }
    }

    public static class StateRegionWiseTotalSalesPriceReducer extends Reducer <Text,Text
,Text,Text> {
        public void reduce(Text key,Iterable<Text> values, Context con)throws IOException,
InterruptedException {

            double Totalprice=0.0;
            for(Text t:values) {
                String[] parts =t.toString().split("\t");
                Totalprice +=Float.parseFloat(parts[1]);
            }

            String str =String.format("%.f",Totalprice);
            con.write(key, new Text(str));
        }
    }

    public static void main(String args[]) throws Exception {

        boolean recursive=true;
        Configuration conf=new Configuration();
        FileSystem fs=FileSystem.get(conf);
    }
}

```

```
if(fs.exists(new Path(args[1])))  
fs.delete(new Path(args[1]),recursive);  
  
Job job=Job.getInstance(conf, " State Region wise Total Sales price ");  
  
job.setJarByClass(StateRegionWiseTotalSalesPrice.class);  
  
job.setMapperClass(StateRegionWiseTotalSalesPriceMapper.class);  
job.setReducerClass(StateRegionWiseTotalSalesPriceReducer.class);  
  
job.setNumReduceTasks(1);  
  
job.setOutputKeyClass(Text.class);  
job.setOutputValueClass(Text.class);  
  
FileInputFormat.addInputPath(job,new Path(args[0]));  
FileOutputFormat.setOutputPath(job,new Path(args[1]));  
  
job.waitForCompletion(true);  
  
}  
  
}
```

```
Activities Terminal Sep 1 11:43
hadoop@abi1-cse136: ~/Documents/purchase2

File Edit View Search Terminal Help

.2.3/share/hadoop/yarn/hadoop-yarn-services-core-3.2.3.jar:/home/hadoop/hadoop-3.2.3/share/hadoop/yarn/hadoop-yarn-submarine-3.2.3.jar:/home/h
adoop/hadoop-3.2.3/share/hadoop/yarn/lib:/home/hadoop/hadoop-3.2.3/share/hadoop/yarn/sources:/home/hadoop/hadoop-3.2.3/share/hadoop/yarn/test:
/home/hadoop/hadoop-3.2.3/share/hadoop/yarn/timeline-service:/home/hadoop/hadoop-3.2.3/share/hadoop/yarn/webapps:/home/hadoop/hadoop-3.2.3/shar
e/hadoop/yarn/yarn-service-examples
hadoop@abi1-cse136:~$ hadoop fs -mkdir hdfs://localhost:9000/purchase2
hadoop@abi1-cse136:~$ hadoop fs -mkdir hdfs://localhost:9000/purchase2/input
hadoop@abi1-cse136:~$ hadoop fs -put '/home/hadoop/Documents/purchase2/input_data/superstore.csv' /purchase2/input
hadoop@abi1-cse136:~$ ls
bddata Downloads javax.activation-api-1.2.0.jar marvel Public Templates weather
code ex8.java javax.activation-api-1.2.0.jar.1 mepco purchase test WordCountTutorial
desktop g3 javax.activation-api-1.2.0.jar.2 Mustic purchase1 TestDFSIO_results.log
dfsdata hadoop-3.2.3 kcet2 Pictures snap tnpdata
Documents hadoop-3.2.3.tar.gz kcet2 prem stopword Videos
hadoop@abi1-cse136:~$ javac -classpath $(HADOOP_CLASSPATH) -d '/home/hadoop/Documents/purchase2/data_classes' '/home/hadoop/Documents/purchase2
/StateRegionWiseTotalSalesPrice.java'
hadoop@abi1-cse136:~$ cd '/home/hadoop/Documents/purchase2'
hadoop@abi1-cse136:~/Documents/purchase2$ jar -cvf firstTutorial.jar -C data_classes/ .
added manifest
adding: StateRegionWiseTotalSalesPrice.class(in = 1904) (out= 913)(deflated 52%)
adding: StateRegionWiseTotalSalesPrice$StateRegionWiseTotalSalesPriceReducer.class(in = 2092) (out= 868)(deflated 58%)
adding: StateRegionWiseTotalSalesPrice$StateRegionWiseTotalSalesPriceMapper.class(in = 1765) (out= 675)(deflated 61%)
hadoop@abi1-cse136:~/Documents/purchase2$ hadoop jar '/home/hadoop/Documents/purchase2/firstTutorial.jar' StateRegionWiseTotalSalesPrice /purch
ase2/input /purchase2/output
2022-09-01 11:27:27,903 INFO client.RMProxy: Connecting to ResourceManager at /127.0.0.1:8032
2022-09-01 11:27:28,093 WARN mapreduce.JobResourceUploader: Hadoop command-line option parsing not performed. Implement the Tool interface and
execute your application with ToolRunner to remedy this.
2022-09-01 11:27:28,138 INFO mapreduce.JobResourceUploader: Disabling Erasure Coding for path: /tnp/hadoop-yarn/staging/hadoop/.staging/job_16
62004055166_0004
2022-09-01 11:27:28,381 INFO input.FileInputFormat: Total input files to process : 1
2022-09-01 11:27:28,526 INFO mapreduce.JobSubmitter: number of splits:1
2022-09-01 11:27:28,705 INFO mapreduce.JobSubmitter: Submitting tokens for job: job_1662004055166_0004
2022-09-01 11:27:28,707 INFO mapreduce.JobSubmitter: Executing with tokens: []
2022-09-01 11:27:28,831 INFO conf.Configuration: resource-types.xml not found
2022-09-01 11:27:28,832 INFO resource.ResourceUtils: Unable to find 'resource-types.xml'.
2022-09-01 11:27:28,886 INFO impl.YarnClientImpl: Submitted application 1662004055166_0004
2022-09-01 11:27:28,915 INFO mapreduce.Job: The url to track the job: http://abi1-cse136:8088/proxy/application_1662004055166_0004/
2022-09-01 11:27:28,916 INFO mapreduce.Job: Running job: job_1662004055166_0004
```

```
Activities Terminal Sep 1 11:43
hadoop@abi1-cse136: ~/Documents/purchase2

File Edit View Search Terminal Help

CPU time spent (ms)=1860
Physical memory (bytes) snapshot=512290816
Virtual memory (bytes) snapshot=5460537344
Total committed heap usage (bytes)=369698752
Peak Map Physical memory (bytes)=299315200
Peak Map Virtual memory (bytes)=2726006784
Peak Reduce Physical memory (bytes)=212975616
Peak Reduce Virtual memory (bytes)=2734530560

Shuffle Errors
BAD_ID=0
CONNECTION=0
IO_ERROR=0
WRONG_LENGTH=0
WRONG_MAP=0
WRONG_REDUCE=0

File Input Format Counters
Bytes Read=1843488
File Output Format Counters
Bytes Written=1104

hadoop@abi1-cse136:~/Documents/purchase2$ hadoop dfs -cat /purchase2/output/*
WARNING: Use of this script to execute dfs is deprecated.
WARNING: Attempting to execute replacement "hdfs dfs" instead.

Alabama 19510.639920
Arizona 35282.000965
Arkansas 11678.129952
California 457687.630219
Colorado 32108.118134
Connecticut 13384.357016
Delaware 27451.068752
District of Columbia 2865.020073
Florida 89473.708429
Georgia 40095.640078
Idaho 4302.405970
Illinois 80166.100892
Indiana 53555.358967
Iowa 4579.759923
```

Activities Firefox Web Browser Sep 1 11:43

localhost:9870/explorer.html#/purchase2/output

### Browse Directory

/purchase2/output

Show 25 entries

Permission Owner

Showing 1 to 2 of 2 entries

Hadoop, 2022.

File information - part-r-00000

Download Head the file (first 32K) Tail the file (last 32K)

Block information - Block 0

Block ID: 1073742006  
Block Pool ID: BP-689235202-172.16.7.136-1659521372740  
Generation Stamp: 1182  
Size: 1104  
Availability:  
• ab1-cse136

File contents

Alabama 19510.639920  
Arizona 35282.000965  
Arkansas 11678.129952  
California 457687.630219  
Colorado 32108.118134  
Connecticut 13384.357016  
Delaware 27451.068752  
District of Columbia 2865.020073

Activities Firefox Web Browser Sep 1 11:44

localhost:9870/explorer.html#/purchase2

### Browse Directory

/purchase2

Show 25 entries

Search:

Permission	Owner	Group	Size	Last Modified	Replication	Block Size	Name
drwxr-xr-x	hadoop	supergroup	0 B	Sep 01 11:24	0	0 B	input
drwxr-xr-x	hadoop	supergroup	0 B	Sep 01 11:27	0	0 B	output

Showing 1 to 2 of 2 entries

Previous 1 Next

Hadoop, 2022.