WEATHER ANALYSIS :

import java.io.IOException;

import java.util.ArrayList;

import java.util.Iterator;

import java.util.List;

import java.util.StringTokenizer;

import org.apache.hadoop.conf.Configuration;

import org.apache.hadoop.conf.Configured;

import org.apache.hadoop.fs.Path;

import org.apache.hadoop.io.LongWritable;

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.mapred.KeyValueTextInputFormat;

import org.apache.hadoop.mapred.MapReduceBase;

import org.apache.hadoop.mapred.Mapper;

import org.apache.hadoop.mapred.OutputCollector;

import org.apache.hadoop.mapred.Reducer;

import org.apache.hadoop.mapred.Reporter;

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

public class WeatherAnalysis extends Configured implements Tool {

final long DEFAULT\_SPLIT\_SIZE = 128 \* 1024 \* 1024;

public static class MapClass extends MapReduceBase

implements Mapper<LongWritable, Text, Text, Text> {

private Text word = new Text();

private Text values = new Text();

public void map(LongWritable key, Text value,

OutputCollector<Text, Text> output,

Reporter reporter) throws IOException {

String line = value.toString();

StringTokenizer itr = new StringTokenizer(line);

int counter = 0;

String key\_out = null;

String value\_str = null;

boolean skip = false;

loop:while (itr.hasMoreTokens() && counter<13) {

String str = itr.nextToken();

switch (counter) {

case 0:

key\_out = str;

if(str.contains("STN")){

skip = true;

break loop;

}else{

break;

}

case 2:

int hour = Integer.valueOf(str.substring(str.lastIndexOf("\_")+1, str.length()));

str = str.substring(4,str.lastIndexOf("\_")-2);

if(hour>4 && hour<=10){

str = str.concat("\_section1");

}else if(hour>10 && hour<=16){

str = str.concat("\_section2");

}else if(hour>16 && hour<=22){

str = str.concat("\_section3");

} else{ str = str.concat("\_section4");

}

key\_out = key\_out.concat("\_").concat(str);

break;

case 3:

if(str.equals("9999.9")){

skip = true;

break loop;

}else{

value\_str = str.concat(" ");

break;

}

case 4:

if(str.equals("9999.9")){

skip = true;

break loop;

}else{

value\_str = value\_str.concat(str).concat(" ");

break;

}

case 12:

if(str.equals("999.9")){

skip = true;

break loop;

}else{ value\_str = value\_str.concat(str).concat(" ");

break;

}

default: break;

}

counter++;

}

if(!skip){

word.set(key\_out);

values.set(value\_str);

output.collect(word, values);

}

}

}

public static class Reduce extends MapReduceBase

implements Reducer<Text, Text, Text, Text> {

private Text value\_out\_text = new Text();

public void reduce(Text key, Iterator<Text> values,

OutputCollector<Text, Text> output, Reporter reporter) throws IOException {

double sum\_temp = 0;

double sum\_dew = 0;

double sum\_wind = 0;

int count = 0;

while (values.hasNext()) {

String str = values.next().toString();

StringTokenizer itr = new StringTokenizer(str);

int count\_vector = 0;

while (itr.hasMoreTokens()) {

String nextToken = itr.nextToken(" ");

if(count\_vector==0){

sum\_temp += Double.valueOf(nextToken);

}

if(count\_vector==1){

sum\_dew += Double.valueOf(nextToken);

}

if(count\_vector==2){

sum\_wind += Double.valueOf(nextToken);

}

count\_vector++;

}

count++;

}

double avg\_tmp = sum\_temp / count;

double avg\_dew = sum\_dew / count;

double avg\_wind = sum\_wind / count;

System.out.println(key.toString()+" count is "+count+" sum of temp is "+sum\_temp+" sum of dew is "+sum\_dew+" sum of wind is "+sum\_wind+"\n");

String value\_out = String.valueOf(avg\_tmp).concat(" ").concat(String.valueOf(avg\_dew)).concat(" ").concat(String.valueOf(avg\_wind));

value\_out\_text.set(value\_out);

output.collect(key, value\_out\_text);

}

}

static int printUsage() {

System.out.println("weather [-m <maps>] [-r <reduces>] <job\_1 input> <job\_1 output> <job\_2 output>");

ToolRunner.printGenericCommandUsage(System.out);

return -1;

}

public int run(String[] args) throws Exception {

Configuration config = getConf();

JobConf conf = new JobConf(config, WeatherAnalysis.class);

conf.setJobName("Weather Job1");

conf.setOutputKeyClass(Text.class);

conf.setOutputValueClass(Text.class);

conf.setMapOutputKeyClass(Text.class);

conf.setMapOutputValueClass(Text.class);

conf.setMapperClass(MapClass.class);

conf.setReducerClass(Reduce.class);

List<String> other\_args = new ArrayList<String>();

for(int i=0; i < args.length; ++i) {

try {

if ("-m".equals(args[i])) {

conf.setNumMapTasks(Integer.parseInt(args[++i]));

} else if ("-r".equals(args[i])) {

conf.setNumReduceTasks(Integer.parseInt(args[++i]));

} else {

other\_args.add(args[i]);

}

} catch (NumberFormatException except) {

System.out.println("ERROR: Integer expected instead of " + args[i]);

return printUsage();

} catch (ArrayIndexOutOfBoundsException except) {

System.out.println("ERROR: Required parameter missing from " +

args[i-1]);

return printUsage();

}

}

FileInputFormat.setInputPaths(conf, other\_args.get(0));

FileOutputFormat.setOutputPath(conf, new Path(other\_args.get(1)));

JobClient.runJob(conf);

return 0;

}

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

int res = ToolRunner.run(new Configuration(), new WeatherAnalysis(), args);

System.exit(res);

}

}

SAMPLE WEATHER

690190 13910 20060201\_0 51.75 33.0 24 1006.3 24 943.9 24 15.0 24 10.7 24 22.0 28.9 0.00I 999.9 000000

690190 13910 20060201\_1 54.74 33.0 24 1006.3 24 943.9 24 15.0 24 10.7 24 22.0 28.9 0.00I 999.9 000000

690190 13910 20060201\_2 50.59 33.0 24 1006.3 24 943.9 24 15.0 24 10.7 24 22.0 28.9 0.00I 999.9 000000

690190 13910 20060201\_3 51.67 33.0 24 1006.3 24 943.9 24 15.0 24 10.7 24 22.0 28.9 0.00I 999.9 000000

690190 13910 20060201\_4 65.67 33.0 24 1006.3 24 943.9 24 15.0 24 10.7 24 22.0 28.9 0.00I 999.9 000000

690190 13910 20060201\_5 55.37 33.0 24 1006.3 24 943.9 24 15.0 24 10.7 24 22.0 28.9 0.00I 999.9 000000

690190 13910 20060201\_6 49.26 33.0 24 1006.3 24 943.9 24 15.0 24 10.7 24 22.0 28.9 0.00I 999.9 000000

690190 13910 20060201\_7 55.44 33.0 24 1006.3 24 943.9 24 15.0 24 10.7 24 22.0 28.9 0.00I 999.9 000000

690190 13910 20060201\_8 64.05 33.0 24 1006.3 24 943.9 24 15.0 24 10.7 24 22.0 28.9 0.00I 999.9 000000

690190 13910 20060201\_9 68.77 33.0 24 1006.3 24 943.9 24 15.0 24 10.7 24 22.0 28.9 0.00I 999.9 000000

690190 13910 20060201\_10 48.93 33.0 24 1006.3 24 943.9 24 15.0 24 10.7 24 22.0 28.9 0.00I 999.9 000000

690190 13910 20060201\_11 65.37 33.0 24 1006.3 24 943.9 24 15.0 24 10.7 24 22.0 28.9 0.00I 999.9 000000

690190 13910 20060201\_12 69.45 33.0 24 1006.3 24 943.9 24 15.0 24 10.7 24 22.0 28.9 0.00I 999.9 000000

690190 13910 20060201\_13 52.91 33.0 24 1006.3 24 943.9 24 15.0 24 10.7 24 22.0 28.9 0.00I 999.9 000000

690190 13910 20060201\_14 53.69 33.0 24 1006.3 24 943.9 24 15.0 24 10.7 24 22.0 28.9 0.00I 999.9 000000

690190 13910 20060201\_15 53.30 33.0 24 1006.3 24 943.9 24 15.0 24 10.7 24 22.0 28.9 0.00I 999.9 000000

690190 13910 20060201\_16 66.17 33.0 24 1006.3 24 943.9 24 15.0 24 10.7 24 22.0 28.9 0.00I 999.9 000000

690190 13910 20060201\_17 53.83 33.0 24 1006.3 24 943.9 24 15.0 24 10.7 24 22.0 28.9 0.00I 999.9 000000

690190 13910 20060201\_18 50.54 33.0 24 1006.3 24 943.9 24 15.0 24 10.7 24 22.0 28.9 0.00I 999.9 000000

690190 13910 20060201\_19 50.27 33.0 24 1006.3 24 943.9 24 15.0 24 10.7 24 22.0 28.9 0.00I 999.9 000000

690190 13910 20060201\_20 59.08 33.0 24 1006.3 24 943.9 24 15.0 24 10.7 24 22.0 28.9 0.00I 999.9 000000

690190 13910 20060201\_21 53.05 33.0 24 1006.3 24 943.9 24 15.0 24 10.7 24 22.0 28.9 0.00I 999.9 000000

690190 13910 20060201\_22 57.97 33.0 24 1006.3 24 943.9 24 15.0 24 10.7 24 22.0 28.9 0.00I 999.9 000000

690190 13910 20060201\_23 48.23 33.0 24 1006.3 24 943.9 24 15.0 24 10.7 24 22.0 28.9 0.00I 999.9 000000

690190 13910 20060202\_0 47.16 28.5 24 1003.2 24 940.6 24 15.0 24 5.7 24 12.0 999.9 0.00I 999.9 000000

690190 13910 20060202\_1 69.72 28.5 24 1003.2 24 940.6 24 15.0 24 5.7 24 12.0 999.9 0.00I 999.9 000000

690190 13910 20060202\_2 62.71 28.5 24 1003.2 24 940.6 24 15.0 24 5.7 24 12.0 999.9 0.00I 999.9 000000

690190 13910 20060202\_3 46.34 28.5 24 1003.2 24 940.6 24 15.0 24 5.7 24 12.0 999.9 0.00I 999.9 000000

690190 13910 20060202\_4 53.15 28.5 24 1003.2 24 940.6 24 15.0 24 5.7 24 12.0 999.9 0.00I 999.9 000000

690190 13910 20060202\_5 64.59 28.5 24 1003.2 24 940.6 24 15.0 24 5.7 24 12.0 999.9 0.00I 999.9 000000

690190 13910 20060202\_6 58.26 28.5 24 1003.2 24 940.6 24 15.0 24 5.7 24 12.0 999.9 0.00I 999.9 000000

690190 13910 20060202\_7 53.27 28.5 24 1003.2 24 940.6 24 15.0 24 5.7 24 12.0 999.9 0.00I 999.9 000000

690190 13910 20060202\_8 43.68 28.5 24 1003.2 24 940.6 24 15.0 24 5.7 24 12.0 999.9 0.00I 999.9 000000

690190 13910 20060202\_9 65.70 28.5 24 1003.2 24 940.6 24 15.0 24 5.7 24 12.0 999.9 0.00I 999.9 000000

690190 13910 20060202\_10 66.27 28.5 24 1003.2 24 940.6 24 15.0 24 5.7 24 12.0 999.9 0.00I 999.9 000000

690190 13910 20060202\_11 53.05 28.5 24 1003.2 24 940.6 24 15.0 24 5.7 24 12.0 999.9 0.00I 999.9 000000

690190 13910 20060202\_12 68.45 28.5 24 1003.2 24 940.6 24 15.0 24 5.7 24 12.0 999.9 0.00I 999.9 000000

690190 13910 20060202\_13 49.03 28.5 24 1003.2 24 940.6 24 15.0 24 5.7 24 12.0 999.9 0.00I 999.9 000000

690190 13910 20060202\_14 66.59 28.5 24 1003.2 24 940.6 24 15.0 24 5.7 24 12.0 999.9 0.00I 999.9 000000

690190 13910 20060202\_15 63.12 28.5 24 1003.2 24 940.6 24 15.0 24 5.7 24 12.0 999.9 0.00I 999.9 000000

690190 13910 20060202\_16 49.13 28.5 24 1003.2 24 940.6 24 15.0 24 5.7 24 12.0 999.9 0.00I 999.9 000000

690190 13910 20060202\_17 62.85 28.5 24 1003.2 24 940.6 24 15.0 24 5.7 24 12.0 999.9 0.00I 999.9 000000

690190 13910 20060202\_18 64.67 28.5 24 1003.2 24 940.6 24 15.0 24 5.7 24 12.0 999.9 0.00I 999.9 000000

690190 13910 20060202\_19 55.73 28.5 24 1003.2 24 940.6 24 15.0 24 5.7 24 12.0 999.9 0.00I 999.9 000000

690190 13910 20060202\_20 56.42 28.5 24 1003.2 24 940.6 24 15.0 24 5.7 24 12.0 999.9 0.00I 999.9 000000

690190 13910 20060202\_21 53.83 28.5 24 1003.2 24 940.6 24 15.0 24 5.7 24 12.0 999.9 0.00I 999.9 000000

690190 13910 20060202\_22 45.14 28.5 24 1003.2 24 940.6 24 15.0 24 5.7 24 12.0 999.9 0.00I 999.9 000000

690190 13910 20060202\_23 68.18 28.5 24 1003.2 24 940.6 24 15.0 24 5.7 24 12.0 999.9 0.00I 999.9 000000

690190 13910 20060203\_0 48.41 26.5 24 1008.7 24 946.0 24 15.0 24 9.6 24 17.1 22.0 0.00I 999.9 000000

690190 13910 20060203\_1 55.12 26.5 24 1008.7 24 946.0 24 15.0 24 9.6 24 17.1 22.0 0.00I 999.9 000000

690190 13910 20060203\_2 46.48 26.5 24 1008.7 24 946.0 24 15.0 24 9.6 24 17.1 22.0 0.00I 999.9 000000

690190 13910 20060203\_3 54.99 26.5 24 1008.7 24 946.0 24 15.0 24 9.6 24 17.1 22.0 0.00I 999.9 000000

690190 13910 20060203\_4 50.62 26.5 24 1008.7 24 946.0 24 15.0 24 9.6 24 17.1 22.0 0.00I 999.9 000000

690190 13910 20060203\_5 55.81 26.5 24 1008.7 24 946.0 24 15.0 24 9.6 24 17.1 22.0 0.00I 999.9 000000

690190 13910 20060203\_6 59.28 26.5 24 1008.7 24 946.0 24 15.0 24 9.6 24 17.1 22.0 0.00I 999.9 000000

690190 13910 20060203\_7 60.55 26.5 24 1008.7 24 946.0 24 15.0 24 9.6 24 17.1 22.0 0.00I 999.9 000000

690190 13910 20060203\_8 52.62 26.5 24 1008.7 24 946.0 24 15.0 24 9.6 24 17.1 22.0 0.00I 999.9 000000

690190 13910 20060203\_9 62.27 26.5 24 1008.7 24 946.0 24 15.0 24 9.6 24 17.1 22.0 0.00I 999.9 000000

690190 13910 20060203\_10 49.90 26.5 24 1008.7 24 946.0 24 15.0 24 9.6 24 17.1 22.0 0.00I 999.9 000000

690190 13910 20060203\_11 45.12 26.5 24 1008.7 24 946.0 24 15.0 24 9.6 24 17.1 22.0 0.00I 999.9 000000

690190 13910 20060203\_12 62.85 26.5 24 1008.7 24 946.0 24 15.0 24 9.6 24 17.1 22.0 0.00I 999.9 000000

690190 13910 20060203\_13 50.90 26.5 24 1008.7 24 946.0 24 15.0 24 9.6 24 17.1 22.0 0.00I 999.9 000000

690190 13910 20060203\_14 49.73 26.5 24 1008.7 24 946.0 24 15.0 24 9.6 24 17.1 22.0 0.00I 999.9 000000

690190 13910 20060203\_15 47.28 26.5 24 1008.7 24 946.0 24 15.0 24 9.6 24 17.1 22.0 0.00I 999.9 000000

690190 13910 20060203\_16 46.67 26.5 24 1008.7 24 946.0 24 15.0 24 9.6 24 17.1 22.0 0.00I 999.9 000000

690190 13910 20060203\_17 49.48 26.5 24 1008.7 24 946.0 24 15.0 24 9.6 24 17.1 22.0 0.00I 999.9 000000

690190 13910 20060203\_18 59.53 26.5 24 1008.7 24 946.0 24 15.0 24 9.6 24 17.1 22.0 0.00I 999.9 000000

690190 13910 20060203\_19 59.49 26.5 24 1008.7 24 946.0 24 15.0 24 9.6 24 17.1 22.0 0.00I 999.9 000000

690190 13910 20060203\_20 52.25 26.5 24 1008.7 24 946.0 24 15.0 24 9.6 24 17.1 22.0 0.00I 999.9 000000

690190 13910 20060203\_21 57.67 26.5 24 1008.7 24 946.0 24 15.0 24 9.6 24 17.1 22.0 0.00I 999.9 000000

690190 13910 20060203\_22 54.04 26.5 24 1008.7 24 946.0 24 15.0 24 9.6 24 17.1 22.0 0.00I 999.9 000000

690190 13910 20060203\_23 58.76 26.5 24 1008.7 24 946.0 24 15.0 24 9.6 24 17.1 22.0 0.00I 999.9 000000

690190 13910 20060204\_0 48.60 15.6 24 1020.0 24 955.9 24 15.0 24 5.1 24 11.1 15.9 0.00I 999.9 000000

690190 13910 20060204\_1 50.54 15.6 24 1020.0 24 955.9 24 15.0 24 5.1 24 11.1 15.9 0.00I 999.9 000000

690190 13910 20060204\_2 38.55 15.6 24 1020.0 24 955.9 24 15.0 24 5.1 24 11.1 15.9 0.00I 999.9 000000

690190 13910 20060204\_3 50.14 15.6 24 1020.0 24 955.9 24 15.0 24 5.1 24 11.1 15.9 0.00I 999.9 000000

690190 13910 20060204\_4 34.86 15.6 24 1020.0 24 955.9 24 15.0 24 5.1 24 11.1 15.9 0.00I 999.9 000000

690190 13910 20060204\_5 35.38 15.6 24 1020.0 24 955.9 24 15.0 24 5.1 24 11.1 15.9 0.00I 999.9 000000

690190 13910 20060204\_6 41.52 15.6 24 1020.0 24 955.9 24 15.0 24 5.1 24 11.1 15.9 0.00I 999.9 000000

690190 13910 20060204\_7 42.28 15.6 24 1020.0 24 955.9 24 15.0 24 5.1 24 11.1 15.9 0.00I 999.9 000000

690190 13910 20060204\_8 45.11 15.6 24 1020.0 24 955.9 24 15.0 24 5.1 24 11.1 15.9 0.00I 999.9 000000

690190 13910 20060204\_9 51.01 15.6 24 1020.0 24 955.9 24 15.0 24 5.1 24 11.1 15.9 0.00I 999.9 000000

690190 13910 20060204\_10 43.60 15.6 24 1020.0 24 955.9 24 15.0 24 5.1 24 11.1 15.9 0.00I 999.9 000000

690190 13910 20060204\_11 50.12 15.6 24 1020.0 24 955.9 24 15.0 24 5.1 24 11.1 15.9 0.00I 999.9 000000

690190 13910 20060204\_12 55.07 15.6 24 1020.0 24 955.9 24 15.0 24 5.1 24 11.1 15.9 0.00I 999.9 000000

690190 13910 20060204\_13 43.24 15.6 24 1020.0 24 955.9 24 15.0 24 5.1 24 11.1 15.9 0.00I 999.9 000000

690190 13910 20060204\_14 49.84 15.6 24 1020.0 24 955.9 24 15.0 24 5.1 24 11.1 15.9 0.00I 999.9 000000

690190 13910 20060204\_15 44.42 15.6 24 1020.0 24 955.9 24 15.0 24 5.1 24 11.1 15.9 0.00I 999.9 000000

690190 13910 20060204\_16 58.77 15.6 24 1020.0 24 955.9 24 15.0 24 5.1 24 11.1 15.9 0.00I 999.9 000000

690190 13910 20060204\_17 35.32 15.6 24 1020.0 24 955.9 24 15.0 24 5.1 24 11.1 15.9 0.00I 999.9 000000

690190 13910 20060204\_18 34.83 15.6 24 1020.0 24 955.9 24 15.0 24 5.1 24 11.1 15.9 0.00I 999.9 000000

690190 13910 20060204\_19 51.72 15.6 24 1020.0 24 955.9 24 15.0 24 5.1 24 11.1 15.9 0.00I 999.9 000000

690190 13910 20060204\_20 52.39 15.6 24 1020.0 24 955.9 24 15.0 24 5.1 24 11.1 15.9 0.00I 999.9 000000

690190 13910 20060204\_21 57.54 15.6 24 1020.0 24 955.9 24 15.0 24 5.1 24 11.1 15.9 0.00I 999.9 000000

690190 13910 20060204\_22 50.37 15.6 24 1020.0 24 955.9 24 15.0 24 5.1 24 11.1 15.9 0.00I 999.9 000000

690190 13910 20060204\_23 51.75 15.6 24 1020.0 24 955.9 24 15.0 24 5.1 24 11.1 15.9 0.00I 999.9 000000