

```

package it.polito.bigdata.spark.exercise35;

import org.apache.spark.api.java.*;
import java.util.List;
import org.apache.spark.SparkConf;

public class SparkDriver {

    public static void main(String[] args) {

        String inputPath;
        String outputPath;

        inputPath=args[0];
        outputPath=args[1];

        // Create a configuration object and set the name of the application
        SparkConf conf=new SparkConf().setAppName("Spark Exercise #34");

        // Create a Spark Context object
        JavaSparkContext sc = new JavaSparkContext(conf);

        // Read the content of the input file
        JavaRDD<String> readingsRDD = sc.textFile(inputPath);

        // Extract the PM10 values
        // It can be implemented by using the map transformation
        JavaRDD<Double> pm10ValuesRDD = readingsRDD.map(new
ExtractPM10Value());

        // Select the top-1 value
        List<Double> topPM10Values= pm10ValuesRDD.top(1);
        Double topValue=topPM10Values.get(0);

        // Filter the content of readingsRDD
        // Select only the line(s) associated with the topValue
        JavaRDD<String> selectedRecordsRDD=readingsRDD.filter(new
LinesTopPM10Value(topValue));

        // Extract the dates from the selected records
        JavaRDD<String> datesRDD=selectedRecordsRDD.map(new ExtractDate());

        // Remove duplicates, if any
        JavaRDD<String> distinctDatesRDD=datesRDD.distinct();

        // Store the result in the output folder
        distinctDatesRDD.saveAsTextFile(outputPath);

        // Close the Spark context
        sc.close();
    }
}

```