**package** com.boffins.hashtag;

**import** java.io.IOException;

**import** java.util.StringTokenizer;

**import** org.apache.hadoop.conf.Configuration;

**import** org.apache.hadoop.fs.Path;

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

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

**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.FileInputFormat;

**import** org.apache.hadoop.mapreduce.lib.output.FileOutputFormat;

**public** **class** Mapper\_HashTag **extends** Mapper<Object, Text, Text, IntWritable>

{

**private** **final** **static** IntWritable ***one*** = **new** IntWritable(1);

**private** Text word = **new** Text();

**private** String hashRegExp = "(?:\\s|\\A|^)[##]+([A-Za-z0-9-\_]+)";

**public** **void** map(Object key, Text value, Context context)

**throws** IOException, InterruptedException {

String[] words = value.toString().split(" ") ;

**for** (String str: words)

{

**if** (str.matches(hashRegExp)) {

word.set(str);

context.write(word, ***one***);

}

}

}

}