

Assignment 2

UserLogDriver.java:

```
package mrLogFile_demo;

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.lib.input.FileInputFormat;
import org.apache.hadoop.mapreduce.lib.output.FileOutputFormat;

public class UserLogDriver {

    public static void main(String[] args) throws Exception {
        if (args.length != 2) {
            System.err.println("Usage: UserLogDriver <input path> <output path>");
            System.exit(-1);
        }

        // Create a configuration object for the job
        Configuration conf = new Configuration();

        // Create the Job instance
        Job job = Job.getInstance(conf, "MaxLoggedUsers");

        // Set the jar by class to identify the location of the jar
        job.setJarByClass(UserLogDriver.class);

        // Set Mapper and Reducer classes
        job.setMapperClass(UserLogMapper.class);
        job.setReducerClass(UserLogReducer.class);

        // Set output key and value types
        job.setOutputKeyClass(Text.class);
        job.setOutputValueClass(IntWritable.class);

        // Set input and output formats
        FileInputFormat.addInputPath(job, new Path(args[0]));
        FileOutputFormat.setOutputPath(job, new Path(args[1]));

        // Wait for the job to complete
        System.exit(job.waitForCompletion(true) ? 0 : 1);
    }
}
```

UserLogMapper.java:

```
package mrLogFile_demo;

import java.io.IOException;
import org.apache.hadoop.io.IntWritable;
import org.apache.hadoop.io.LongWritable;
import org.apache.hadoop.io.Text;
import org.apache.hadoop.mapreduce.Mapper;

public class UserLogMapper extends Mapper<LongWritable, Text, Text,
IntWritable> {
    private final static IntWritable one = new IntWritable(1);

    public void map(LongWritable key, Text value, Context context) throws
IOException, InterruptedException {
        String valueString = value.toString();
        String[] SingleUserData = valueString.split("-");
        context.write(new Text(SingleUserData[0]), one);
    }
}
```

UserLogReducer.java

```
package mrLogFile_demo;

import org.apache.hadoop.io.IntWritable;
import org.apache.hadoop.io.Text;
import org.apache.hadoop.mapreduce.Reducer;

import java.io.IOException;

public class UserLogReducer extends Reducer<Text, IntWritable, Text,
IntWritable> {

    private IntWritable result = new IntWritable();

    @Override
    protected void reduce(Text key, Iterable<IntWritable> values, Context
context) throws IOException, InterruptedException {
        int sum = 0;

        // Iterate through the list of values (log counts for the user)
        for (IntWritable val : values) {
            sum += val.get(); // Add the count
        }

        // Set the result to the sum and write it to the context
        result.set(sum);
        context.write(key, result); // Emit (user, total count)
    }
}
```

```
}  
}
```

Output:

```
cat: '/user/prathamesh/output/part-00000': No such file or directory  
prathamesh@prathamesh-HP-Laptop-14s-fq1xxx:~/inputdata$ hdfs dfs -ls /user/prathamesh/output/  
Found 2 items  
-rw-r--r-- 1 prathamesh supergroup          0 2025-04-20 14:37 /user/prathamesh/output/_SUCCESS  
-rw-r--r-- 1 prathamesh supergroup    3838 2025-04-20 14:37 /user/prathamesh/output/part-r-00000  
prathamesh@prathamesh-HP-Laptop-14s-fq1xxx:~/inputdata$ hdfs dfs -cat /user/prathamesh/output/part-r-00000  
10.1.1.236 7  
10.1.181.142 14  
10.1.232.31 5  
10.10.55.142 14  
10.102.101.66 1  
10.103.184.104 1  
10.103.190.81 53  
10.103.63.29 1  
10.104.73.51 1  
10.105.160.183 1  
10.108.91.151 1  
10.109.21.76 1  
10.11.131.40 1  
10.111.71.20 8  
10.112.227.184 6  
10.114.74.30 1  
10.115.118.78 1  
10.117.224.230 1  
10.117.76.22 12  
10.118.19.97 1  
10.118.250.30 7  
10.119.117.132 23  
10.119.33.245 1  
10.119.74.120 1  
10.12.113.198 2  
10.12.219.30 1  
10.120.165.113 1  
10.120.207.127 4  
10.122.124.47 1
```

access log short.csv

```
1 10.223.157.186 - - [15/Jul/2009:20:50:33 -"0700] ""GET /assets/img/home"-logo.png HTTP/1.1"" 304 "-----  
2 10.223.157.186 - - [15/Jul/2009:20:50:33 -"0700] ""GET /assets/img/dummy/primary"-news-2.jpg HTTP/1.1"" 304 "-----  
3 10.223.157.186 - - [15/Jul/2009:20:50:33 -"0700] ""GET /assets/img/dummy/primary"-news-1.jpg HTTP/1.1"" 304 "-----  
4 10.223.157.186 - - [15/Jul/2009:20:50:33 -"0700] ""GET /assets/img/home"-media-block-"placeholder.jpg HTTP/1.1"" 304 "-----  
5 10.223.157.186 - - [15/Jul/2009:20:50:33 -"0700] ""GET /assets/img/dummy/secondary"-news-4.jpg HTTP/1.1"" 304 "-----  
6 10.223.157.186 - - [15/Jul/2009:20:50:33 -"0700] ""GET /assets/img/loading.gif HTTP/1.1"" 304 "-----  
7 10.223.157.186 - - [15/Jul/2009:20:50:33 -"0700] ""GET /assets/img/search"-button.gif HTTP/1.1"" 304 "-----  
8 10.223.157.186 - - [15/Jul/2009:20:50:33 -"0700] ""GET /assets/img/dummy/secondary"-news-3.jpg HTTP/1.1"" 304 "-----  
9 10.223.157.186 - - [15/Jul/2009:20:50:33 -"0700] ""GET /assets/img/dummy/secondary"-news-2.jpg HTTP/1.1"" 304 "-----  
10 10.223.157.186 - - [15/Jul/2009:20:50:33 -"0700] ""GET /assets/img/dummy/secondary"-news-1.jpg HTTP/1.1"" 304 "-----  
11 10.216.113.172 - - [16/Jul/2009:02:51:31 -"0700] ""GET /assets/img/closetlabel.gif HTTP/1.1"" 200 979"-----  
12 10.216.113.172 - - [16/Jul/2009:02:51:31 -"0700] ""GET /favicon.ico HTTP/1.1"" 404 209"-----  
13 10.216.113.172 - - [16/Jul/2009:02:51:31 -"0700] ""GET /assets/swf/home"-media-"block.swf HTTP/1.1"" 200 123884"-----  
14 10.216.113.172 - - [16/Jul/2009:02:51:41 -"0700] ""GET /films/district"-13 HTTP/1.1"" 301 268"-----  
15 10.216.113.172 - - [16/Jul/2009:02:51:41 -"0700] ""GET /films/district"-13/ HTTP/1.1"" 200 12772"-----  
16 10.216.113.172 - - [16/Jul/2009:02:51:41 -"0700] ""GET /assets/img/pdf"-icon.gif HTTP/1.1"" 200 228"-----  
17 10.216.113.172 - - [16/Jul/2009:02:51:41 -"0700] ""GET /assets/img/dummy/momentum"-pictures-"logo.gif HTTP/1.1"" 200 3579"-----  
18 10.216.113.172 - - [16/Jul/2009:02:51:41 -"0700] ""GET /assets/img/dummy/films/district"-13/image-"thumbs/1.jpg HTTP/1.1"" 200 9276"-----
```