

## Web Log Analysis - Partitioner Assignment

Date : 01-09-22

For Odd Roll No:

**Find Out the hit count per month**

### CODE:

```
import java.io.*;
import org.apache.hadoop.io.*;
import org.apache.hadoop.mapreduce.*;
import org.apache.hadoop.conf.*;
import org.apache.hadoop.fs.*;
import org.apache.hadoop.mapreduce.lib.input.*;
import org.apache.hadoop.mapreduce.lib.output.*;
import org.apache.hadoop.util.*;

public class multiFilesPart extends Configured implements Tool{
//map class
public static class Map extends Mapper<LongWritable, Text, Text, IntWritable> {
@Override public void map (LongWritable key, Text value, Context context)
throws IOException, InterruptedException {
String line = value.toString();
String[] word =line.split("/") ;
String code=word[1];
context.write(new Text (code), new IntWritable (1));}}
//Reducer class
public static class SumReducer extends Reducer<Text, IntWritable, Text, IntWritable> {
@Override public void reduce (Text key, Iterable<IntWritable> values, Context context)
throws IOException, InterruptedException {
int wordCount = 0;
for (IntWritable value : values) {
wordCount += value.get();
}
context.write(key, new IntWritable (wordCount));}}
//Partitioner class
public static class CaderPartitioner extends Partitioner<Text,IntWritable>{
public int getPartition ( Text key , IntWritable value , int numReduceTasks )
{
if ( numReduceTasks == 0 )
{
return 0 ;
}
if ( key.toString().equals("Jan"))
{
return 1 % numReduceTasks ;
}
else if (key.toString().equals("Feb"))
```

```
{
    return 2% numReduceTasks;
}
else if ( key.toString().equals("Mar") )
{
    return 3% numReduceTasks;
}
else if ( key.toString().equals("Apr") )
{
    return 4% numReduceTasks;
}
else if ( key.toString().equals("May") )
{
    return 5% numReduceTasks;
}
else if ( key.toString().equals("Jun") )
{
    return 6% numReduceTasks;
}
else if ( key.toString().equals("Jul") )
{
    return 7% numReduceTasks;
}
else if ( key.toString().equals("Aug") )
{
    return 8% numReduceTasks;
}
else if ( key.toString().equals("Sep") )
{
    return 9% numReduceTasks;
}
else if ( key.toString().equals("Oct") )
{
    return 10% numReduceTasks;
}
else if ( key.toString().equals("Nov") )
{
    return 11% numReduceTasks;
}
else if ( key.toString().equals("Dec") )
{
    return 12% numReduceTasks;
}
else{
```

```

    return 0;}}}
public static void main(String[] args) throws Exception{
int res=ToolRunner.run(new Configuration(),new multiFilesPart(), args);
System.exit(0);}@Override
public int run(String[] arg) throws Exception {
Configuration conf = getConf() ;
    Job job = new Job (conf ,"Partitioner2" ) ;
    job.setJarByClass(multiFilesPart.class ) ;
    FileInputFormat.setInputPaths (job ,new Path ( arg [ 0 ]));
    FileOutputFormat.setOutputPath (job ,new Path ( arg [ 1 ]));
    job.setMapperClass(Map.class ) ;
    job.setMapOutputKeyClass ( Text.class ) ;
    job.setMapOutputValueClass (IntWritable.class ) ;
    // set partitioner statement
    job.setPartitionerClass (CaderPartitioner.class ) ;
    job.setReducerClass(SumReducer.class);
    job.setNumReduceTasks(13) ;
    job.setInputFormatClass (TextInputFormat.class ) ;
    job.setOutputFormatClass(TextOutputFormat.class ) ;
    job.setOutputKeyClass ( Text.class ) ;
    job.setOutputValueClass ( Text.class ) ;
    System.exit(job.waitForCompletion ( true) ? 0 : 1);
    return 0 ;
}

```

File: [/user/training/vaibhav/access\\_log](#)

Goto :

[Go back to dir listing](#)

[Advanced view/download options](#)

[View Next chunk](#)

```

10.223.157.186 - - [15/Jul/2009:14:58:59 -0700] "GET / HTTP/1.1" 403 202
10.223.157.186 - - [15/Jul/2009:14:58:59 -0700] "GET /favicon.ico HTTP/1.1" 404 209
10.223.157.186 - - [15/Jul/2009:15:50:35 -0700] "GET / HTTP/1.1" 200 9157
10.223.157.186 - - [15/Jul/2009:15:50:35 -0700] "GET /assets/js/lowpro.js HTTP/1.1" 200 10469
10.223.157.186 - - [15/Jul/2009:15:50:35 -0700] "GET /assets/css/reset.css HTTP/1.1" 200 1014
10.223.157.186 - - [15/Jul/2009:15:50:35 -0700] "GET /assets/css/960.css HTTP/1.1" 200 6206
10.223.157.186 - - [15/Jul/2009:15:50:35 -0700] "GET /assets/css/the-associates.css HTTP/1.1" 200 15779
10.223.157.186 - - [15/Jul/2009:15:50:35 -0700] "GET /assets/js/the-associates.js HTTP/1.1" 200 4492
10.223.157.186 - - [15/Jul/2009:15:50:35 -0700] "GET /assets/js/lightbox.js HTTP/1.1" 200 25960
10.223.157.186 - - [15/Jul/2009:15:50:36 -0700] "GET /assets/img/search-button.gif HTTP/1.1" 200 168
10.223.157.186 - - [15/Jul/2009:15:50:36 -0700] "GET /assets/img/dummy/secondary-news-3.jpg HTTP/1.1" 200 5604
10.223.157.186 - - [15/Jul/2009:15:50:36 -0700] "GET /assets/img/dummy/primary-news-1.jpg HTTP/1.1" 200 10556
10.223.157.186 - - [15/Jul/2009:15:50:36 -0700] "GET /assets/img/dummy/primary-news-2.jpg HTTP/1.1" 200 9925
10.223.157.186 - - [15/Jul/2009:15:50:36 -0700] "GET /assets/img/closetlabel.gif HTTP/1.1" 200 979
10.223.157.186 - - [15/Jul/2009:15:50:36 -0700] "GET /assets/img/home-logo.png HTTP/1.1" 200 3892
10.223.157.186 - - [15/Jul/2009:15:50:36 -0700] "GET /assets/img/dummy/secondary-news-2.jpg HTTP/1.1" 200 5397
10.223.157.186 - - [15/Jul/2009:15:50:36 -0700] "GET /assets/img/loading.gif HTTP/1.1" 200 2767
10.223.157.186 - - [15/Jul/2009:15:50:36 -0700] "GET /assets/img/dummy/secondary-news-4.jpg HTTP/1.1" 200 5766
10.223.157.186 - - [15/Jul/2009:15:50:36 -0700] "GET /assets/img/home-media-block-placeholder.jpg HTTP/1.1" 200 68831
10.223.157.186 - - [15/Jul/2009:15:50:37 -0700] "GET /assets/img/dummy/secondary-news-1.jpg HTTP/1.1" 200 5766
10.223.157.186 - - [15/Jul/2009:15:50:37 -0700] "GET /assets/swf/home-media-block.swf HTTP/1.1" 200 123884
10.223.157.186 - - [15/Jul/2009:15:50:51 -0700] "GET / HTTP/1.1" 200 9157
10.223.157.186 - - [15/Jul/2009:15:50:51 -0700] "GET /assets/css/960.css HTTP/1.1" 304 -
10.223.157.186 - - [15/Jul/2009:15:50:51 -0700] "GET /assets/css/the-associates.css HTTP/1.1" 304 -
10.223.157.186 - - [15/Jul/2009:15:50:51 -0700] "GET /assets/js/lowpro.js HTTP/1.1" 304 -
10.223.157.186 - - [15/Jul/2009:15:50:51 -0700] "GET /assets/js/lightbox.js HTTP/1.1" 304 -

```

```
[training@localhost ~]$ hadoop jar /home/training/part.jar /user/training/vaibha
v/access_log /user/training/partoutput
22/09/01 11:20:56 WARN mapred.JobClient: Use GenericOptionsParser for parsing th
e arguments. Applications should implement Tool for the same.
22/09/01 11:20:56 INFO input.FileInputFormat: Total input paths to process : 1
22/09/01 11:20:56 WARN snappy.LoadSnappy: Snappy native library is available
22/09/01 11:20:56 INFO snappy.LoadSnappy: Snappy native library loaded
22/09/01 11:20:56 INFO mapred.JobClient: Running job: job_202208300919_0001
22/09/01 11:20:57 INFO mapred.JobClient: map 0% reduce 0%
22/09/01 11:21:07 INFO mapred.JobClient: map 18% reduce 0%
22/09/01 11:21:10 INFO mapred.JobClient: map 25% reduce 0%
22/09/01 11:21:19 INFO mapred.JobClient: map 33% reduce 1%
22/09/01 11:21:20 INFO mapred.JobClient: map 46% reduce 1%
22/09/01 11:21:22 INFO mapred.JobClient: map 50% reduce 1%
22/09/01 11:21:25 INFO mapred.JobClient: map 50% reduce 2%
22/09/01 11:21:29 INFO mapred.JobClient: map 62% reduce 2%
22/09/01 11:21:30 INFO mapred.JobClient: map 73% reduce 2%
22/09/01 11:21:31 INFO mapred.JobClient: map 75% reduce 2%
22/09/01 11:21:32 INFO mapred.JobClient: map 75% reduce 3%
22/09/01 11:21:37 INFO mapred.JobClient: map 87% reduce 3%
22/09/01 11:21:38 INFO mapred.JobClient: map 100% reduce 3%
22/09/01 11:21:40 INFO mapred.JobClient: map 100% reduce 15%
22/09/01 11:21:44 INFO mapred.JobClient: map 100% reduce 30%
22/09/01 11:21:47 INFO mapred.JobClient: map 100% reduce 38%
22/09/01 11:21:48 INFO mapred.JobClient: map 100% reduce 46%
22/09/01 11:21:51 INFO mapred.JobClient: map 100% reduce 61%
22/09/01 11:21:55 INFO mapred.JobClient: map 100% reduce 69%
22/09/01 11:21:56 INFO mapred.JobClient: map 100% reduce 76%
22/09/01 11:21:59 INFO mapred.JobClient: map 100% reduce 84%
22/09/01 11:22:00 INFO mapred.JobClient: map 100% reduce 92%
22/09/01 11:22:02 INFO mapred.JobClient: map 100% reduce 100%
22/09/01 11:22:02 INFO mapred.JobClient: Job complete: job_202208300919_0001
22/09/01 11:22:02 INFO mapred.JobClient: Counters: 32
22/09/01 11:22:02 INFO mapred.JobClient: File System Counters
22/09/01 11:22:02 INFO mapred.JobClient: FILE: Number of bytes read=90232994
22/09/01 11:22:02 INFO mapred.JobClient: FILE: Number of bytes written=138205541
22/09/01 11:22:02 INFO mapred.JobClient: FILE: Number of read operations=0
22/09/01 11:22:02 INFO mapred.JobClient: FILE: Number of large read operations=0
22/09/01 11:22:02 INFO mapred.JobClient: FILE: Number of write operations=0
}
```

Cloudera-Training-VM-4.1.1.c - VMware Workstation 14 Player (Non-commercial use only)

Player ▾ | Applications | Places | System | |

HDFS:/user/training/partou

File Edit View History Bookmarks Tools Help

HDFS:/user/training/partoutput

localhost.localdomain:50075/browseDirectory.jsp?dir=/user/training/partoutput&namenodeInfoPort=50070&nna

Most Visited ▾ | Hadoop NameNode | Hadoop JobTracker | Hadoop YARN | Hue | HBase Master

### Contents of directory /user/training/partoutput

Goto :

[Go to parent directory](#)

Name	Type	Size	Replication	Block Size	Modification Time	Permission	Owner	Group
<a href="#">_SUCCESS</a>	file	0 KB	1	64 MB	2022-09-01 11:22	rw-r--r--	training	supergroup
<a href="#">_logs</a>	dir				2022-09-01 11:20	rw-r--r--	training	supergroup
<a href="#">part-r-00000</a>	file	0.05 KB	1	64 MB	2022-09-01 11:21	rw-r--r--	training	supergroup
<a href="#">part-r-00001</a>	file	0.01 KB	1	64 MB	2022-09-01 11:21	rw-r--r--	training	supergroup
<a href="#">part-r-00002</a>	file	0.01 KB	1	64 MB	2022-09-01 11:21	rw-r--r--	training	supergroup
<a href="#">part-r-00003</a>	file	0.01 KB	1	64 MB	2022-09-01 11:21	rw-r--r--	training	supergroup
<a href="#">part-r-00004</a>	file	0.01 KB	1	64 MB	2022-09-01 11:21	rw-r--r--	training	supergroup
<a href="#">part-r-00005</a>	file	0.01 KB	1	64 MB	2022-09-01 11:21	rw-r--r--	training	supergroup
<a href="#">part-r-00006</a>	file	0.01 KB	1	64 MB	2022-09-01 11:21	rw-r--r--	training	supergroup
<a href="#">part-r-00007</a>	file	0.01 KB	1	64 MB	2022-09-01 11:21	rw-r--r--	training	supergroup
<a href="#">part-r-00008</a>	file	0.01 KB	1	64 MB	2022-09-01 11:21	rw-r--r--	training	supergroup
<a href="#">part-r-00009</a>	file	0.01 KB	1	64 MB	2022-09-01 11:21	rw-r--r--	training	supergroup
<a href="#">part-r-00010</a>	file	0.01 KB	1	64 MB	2022-09-01 11:21	rw-r--r--	training	supergroup
<a href="#">part-r-00011</a>	file	0.01 KB	1	64 MB	2022-09-01 11:21	rw-r--r--	training	supergroup
<a href="#">part-r-00012</a>	file	0.01 KB	1	64 MB	2022-09-01 11:21	rw-r--r--	training	supergroup

File: [/user/training/partoutput/part-r-00001](#)

Goto :

[Go back to dir listing](#)

[Advanced view/download options](#)

Jan 273096

File: [/user/training/partoutput/part-r-00011](#)

Goto :

[Go back to dir listing](#)

[Advanced view/download options](#)

Nov 421626

File: [/user/training/partoutput/part-r-00009](#)

Goto :

[Go back to dir listing](#)

[Advanced view/download options](#)

Sep 430529

File: [/user/training/partoutput/part-r-00005](#)

Goto :

[Go back to dir listing](#)

[Advanced view/download options](#)

May 339551