Problem Statement 4 :

We have a dataset of sales of different TV sets across different locations.
Records look like:
Samsung|Optima|14|Madhya Pradesh|132401|14200
The fields are arranged like:
Company Name|Product Name|Size in inches|State|Pin Code|Price
There are some invalid records which contain 'NA' in either Company Name or Product Name.

Task 1:
Write a Map Reduce program to filter out the invalid records. Map only job will fit for this context.

Solution1 :

Eclipse Code:

```java
package com.acadgild.mapreduce.assignment.task1;

import java.io.IOException;
import org.apache.hadoop.conf.Configuration;
import org.apache.hadoop.fs.Path;
import org.apache.hadoop.io.LongWritable;
import org.apache.hadoop.io.Text;
import org.apache.hadoop.mapreduce.Job;
import org.apache.hadoop.mapreduce.Mapper;
import org.apache.hadoop.mapreduce.lib.input.FileInputFormat;
import org.apache.hadoop.mapreduce.lib.output.FileOutputFormat;

public class FilterTelevision {
        public static class FilterMapper
        extends Mapper<LongWritable, Text, Text, Text>{

    public void map(LongWritable key, Text value, Context context
                    ) throws IOException, InterruptedException {

        String valueString = value.toString();
        String[] SingleTelevisionData = valueString.split("\\|");

        if(!SingleTelevisionData[0].matches("NA") && !
SingleTelevisionData[1].matches("NA")) {
        context.write(value,new Text());
    }
  }
}
```

```java
public static void main(String[] args) throws Exception {
  Configuration conf = new Configuration();
  Job job = Job.getInstance(conf, "filter television");
  job.setJarByClass(FilterTelevision.class);
  job.setMapperClass(FilterMapper.class);
  job.setOutputKeyClass(LongWritable.class);
  job.setOutputValueClass(Text.class);
  job.setNumReduceTasks(0);
  FileInputFormat.addInputPath(job, new Path(args[0]));
  FileOutputFormat.setOutputPath(job, new Path(args[1]));
  System.exit(job.waitForCompletion(true) ? 0 : 1);
}
}
```

Converted to jar : filtertele.jar

Execution:

[acadgild@localhost ~]$ jps
3921 Jps
3196
[acadgild@localhost ~]$ sudo service sshd start
[sudo] password for acadgild:
[acadgild@localhost ~]$ start-all.sh
This script is Deprecated. Instead use start-dfs.sh and start-yarn.sh
18/06/27 03:40:03 WARN util.NativeCodeLoader: Unable to load native-hadoop library for your
platform... using builtin-java classes where applicable
Starting namenodes on [localhost]
localhost: starting namenode, logging to
/home/acadgild/install/hadoop/hadoop-2.6.5/logs/hadoop-acadgild-namenode-localhost.localdomain.ou
t
localhost: starting datanode, logging to
/home/acadgild/install/hadoop/hadoop-2.6.5/logs/hadoop-acadgild-datanode-localhost.localdomain.out
Starting secondary namenodes [0.0.0.0]
0.0.0.0: starting secondarynamenode, logging to
/home/acadgild/install/hadoop/hadoop-2.6.5/logs/hadoop-acadgild-secondarynamenode-localhost.local
domain.out
18/06/27 03:40:37 WARN util.NativeCodeLoader: Unable to load native-hadoop library for your
platform... using builtin-java classes where applicable
starting yarn daemons
starting resourcemanager, logging to
/home/acadgild/install/hadoop/hadoop-2.6.5/logs/yarn-acadgild-resourcemanager-localhost.localdomai
n.out
localhost: starting nodemanager, logging to
/home/acadgild/install/hadoop/hadoop-2.6.5/logs/yarn-acadgild-nodemanager-localhost.localdomain.ou

t
[acadgild@localhost ~]$ jps
4755 Jps
4504 ResourceManager
4072 NameNode
4346 SecondaryNameNode
4171 DataNode
3196
4605 NodeManager
[acadgild@localhost ~]$ hdfs dfs -ls /user/acadgild
18/06/27 03:42:39 WARN util.NativeCodeLoader: Unable to load native-hadoop library for your platform... using builtin-java classes where applicable
Found 1 items
drwxr-xr-x   - acadgild supergroup   0 2018-06-25 00:55 /user/acadgild/hadoopdata
[acadgild@localhost ~]$ hdfs dfs -ls /user/acadgild/hadoopdata
18/06/27 03:43:04 WARN util.NativeCodeLoader: Unable to load native-hadoop library for your platform... using builtin-java classes where applicable
Found 3 items
drwxr-xr-x   - acadgild supergroup   0 2018-06-25 01:33 /user/acadgild/hadoopdata/assignment3
-rw-r--r--   1 acadgild supergroup   1147 2018-06-24 19:35 /user/acadgild/hadoopdata/file.txt
drwxr-xr-x   - acadgild supergroup   0 2018-06-24 19:39 /user/acadgild/hadoopdata/hdfs
[acadgild@localhost ~]$ hdfs dfs -mkdir /user/acadgild/hadoopdata/assignment/mapreduce/task1
18/06/27 03:44:28 WARN util.NativeCodeLoader: Unable to load native-hadoop library for your platform... using builtin-java classes where applicable
mkdir: `/user/acadgild/hadoopdata/assignment/mapreduce/task1': No such file or directory
[acadgild@localhost ~]$ hdfs dfs -mkdir /user/acadgild/hadoopdata/assignment
18/06/27 03:44:53 WARN util.NativeCodeLoader: Unable to load native-hadoop library for your platform... using builtin-java classes where applicable
[acadgild@localhost ~]$ hdfs dfs -mkdir /user/acadgild/hadoopdata/assignment/mapreduce
18/06/27 03:45:14 WARN util.NativeCodeLoader: Unable to load native-hadoop library for your platform... using builtin-java classes where applicable
[acadgild@localhost ~]$ hdfs dfs -mkdir /user/acadgild/hadoopdata/assignment/mapreduce/task1
18/06/27 03:45:32 WARN util.NativeCodeLoader: Unable to load native-hadoop library for your platform... using builtin-java classes where applicable
[acadgild@localhost ~]$ hdfs dfs -mkdir /user/acadgild/hadoopdata/assignment/mapreduce/task1/input
18/06/27 03:45:49 WARN util.NativeCodeLoader: Unable to load native-hadoop library for your platform... using builtin-java classes where applicable
[acadgild@localhost ~]$ hdfs dfs -mkdir /user/acadgild/hadoopdata/assignment/mapreduce/task1/output
18/06/27 03:45:59 WARN util.NativeCodeLoader: Unable to load native-hadoop library for your platform... using builtin-java classes where applicable
[acadgild@localhost ~]$ hdfs dfs -ls /user/acadgild/hadoopdata/assignment/mapreduce/task1/
18/06/27 03:46:26 WARN util.NativeCodeLoader: Unable to load native-hadoop library for your platform... using builtin-java classes where applicable
Found 2 items
drwxr-xr-x   - acadgild supergroup   0 2018-06-27 03:45 /user/acadgild/hadoopdata/assignment/mapreduce/task1/input
drwxr-xr-x   - acadgild supergroup   0 2018-06-27 03:46 /user/acadgild/hadoopdata/assignment/mapreduce/task1/output

[acadgild@localhost ~]$ hdfs dfs -put /home/acadgild/Desktop/television.txt /user/acadgild/hadoopdata/assignment/mapreduce/task1/input
18/06/27 03:47:34 WARN util.NativeCodeLoader: Unable to load native-hadoop library for your platform... using builtin-java classes where applicable
[acadgild@localhost ~]$ hdfs dfs -ls /user/acadgild/hadoopdata/assignment/mapreduce/task1/
18/06/27 03:48:00 WARN util.NativeCodeLoader: Unable to load native-hadoop library for your platform... using builtin-java classes where applicable
Found 2 items
drwxr-xr-x   - acadgild supergroup   0 2018-06-27 03:47 /user/acadgild/hadoopdata/assignment/mapreduce/task1/input
drwxr-xr-x   - acadgild supergroup   0 2018-06-27 03:46 /user/acadgild/hadoopdata/assignment/mapreduce/task1/output
[acadgild@localhost ~]$ hdfs dfs -ls /user/acadgild/hadoopdata/assignment/mapreduce/task1/input
18/06/27 03:48:19 WARN util.NativeCodeLoader: Unable to load native-hadoop library for your platform... using builtin-java classes where applicable
Found 1 items
-rw-r--r--   1 acadgild supergroup   733 2018-06-27 03:47 /user/acadgild/hadoopdata/assignment/mapreduce/task1/input/television.txt
[acadgild@localhost ~]$ hdfs dfs -ls /user/acadgild/hadoopdata/assignment/mapreduce/task1/input/^C
[acadgild@localhost ~]$ hdfs dfs -cat /user/acadgild/hadoopdata/assignment/mapreduce/task1/input/television.txt
18/06/27 03:49:11 WARN util.NativeCodeLoader: Unable to load native-hadoop library for your platform... using builtin-java classes where applicable
Samsung|Optima|14|Madhya Pradesh|132401|14200
Onida|Lucid|18|Uttar Pradesh|232401|16200
Akai|Decent|16|Kerala|922401|12200
Lava|Attention|20|Assam|454601|24200
Zen|Super|14|Maharashtra|619082|9200
Samsung|Optima|14|Madhya Pradesh|132401|14200
Onida|Lucid|18|Uttar Pradesh|232401|16200
Onida|Decent|14|Uttar Pradesh|232401|16200
Onida|NA|16|Kerala|922401|12200
Lava|Attention|20|Assam|454601|24200
Zen|Super|14|Maharashtra|619082|9200
Samsung|Optima|14|Madhya Pradesh|132401|14200
NA|Lucid|18|Uttar Pradesh|232401|16200
Samsung|Decent|16|Kerala|922401|12200
Lava|Attention|20|Assam|454601|24200
Samsung|Super|14|Maharashtra|619082|9200
Samsung|Super|14|Maharashtra|619082|9200
Samsung|Super|14|Maharashtra|619082|9200[acadgild@localhost ~]$

[acadgild@localhost ~]$ hadoop jar /home/acadgild/Desktop/filtertele.jar com.acadgild.mapreduce.assignment.task1.FilterTelevision /user/acadgild/hadoopdata/assignment/mapreduce/task1/input/television.txt /user/acadgild/hadoopdata/assignment/mapreduce/task1/outputfilter
18/06/27 05:29:22 WARN util.NativeCodeLoader: Unable to load native-hadoop library for your platform... using builtin-java classes where applicable
18/06/27 05:29:23 INFO client.RMProxy: Connecting to ResourceManager at localhost/127.0.0.1:8032

18/06/27 05:29:25 WARN mapreduce.JobResourceUploader: Hadoop command-line option parsing not performed. Implement the Tool interface and execute your application with ToolRunner to remedy this.
18/06/27 05:29:25 INFO input.FileInputFormat: Total input paths to process : 1
18/06/27 05:29:26 INFO mapreduce.JobSubmitter: number of splits:1
18/06/27 05:29:26 INFO mapreduce.JobSubmitter: Submitting tokens for job: job_1530051047296_0009
18/06/27 05:29:26 INFO impl.YarnClientImpl: Submitted application application_1530051047296_0009
18/06/27 05:29:27 INFO mapreduce.Job: The url to track the job: http://localhost:8088/proxy/application_1530051047296_0009/
18/06/27 05:29:27 INFO mapreduce.Job: Running job: job_1530051047296_0009
18/06/27 05:29:44 INFO mapreduce.Job: Job job_1530051047296_0009 running in uber mode : false
18/06/27 05:29:44 INFO mapreduce.Job: map 0% reduce 0%
18/06/27 05:29:57 INFO mapreduce.Job: map 100% reduce 0%
18/06/27 05:29:58 INFO mapreduce.Job: Job job_1530051047296_0009 completed successfully
18/06/27 05:29:58 INFO mapreduce.Job: Counters: 30
        File System Counters
                FILE: Number of bytes read=0
                FILE: Number of bytes written=107137
                FILE: Number of read operations=0
                FILE: Number of large read operations=0
                FILE: Number of write operations=0
                HDFS: Number of bytes read=892
                HDFS: Number of bytes written=662
                HDFS: Number of read operations=5
                HDFS: Number of large read operations=0
                HDFS: Number of write operations=2
        Job Counters
                Launched map tasks=1
                Data-local map tasks=1
                Total time spent by all maps in occupied slots (ms)=9384
                Total time spent by all reduces in occupied slots (ms)=0
                Total time spent by all map tasks (ms)=9384
                Total vcore-milliseconds taken by all map tasks=9384
                Total megabyte-milliseconds taken by all map tasks=9609216
        Map-Reduce Framework
                Map input records=18
                Map output records=16
                Input split bytes=159
                Spilled Records=0
                Failed Shuffles=0
                Merged Map outputs=0
                GC time elapsed (ms)=97
                CPU time spent (ms)=710
                Physical memory (bytes) snapshot=65511424
                Virtual memory (bytes) snapshot=339922944
                Total committed heap usage (bytes)=16252928
        File Input Format Counters
                Bytes Read=733

File Output Format Counters
		Bytes Written=662
[acadgild@localhost ~]$ hdfs dfs -ls
/user/acadgild/hadoopdata/assignment/mapreduce/task1/outputfilter
18/06/27 05:30:27 WARN util.NativeCodeLoader: Unable to load native-hadoop library for your
platform... using builtin-java classes where applicable
Found 2 items
-rw-r--r--   1 acadgild supergroup   0 2018-06-27 05:29
/user/acadgild/hadoopdata/assignment/mapreduce/task1/outputfilter/_SUCCESS
-rw-r--r--   1 acadgild supergroup   662 2018-06-27 05:29
/user/acadgild/hadoopdata/assignment/mapreduce/task1/outputfilter/part-m-00000
[acadgild@localhost ~]$ hdfs dfs -cat
/user/acadgild/hadoopdata/assignment/mapreduce/task1/outputfilter/part-m-0000018/06/27 05:30:49
WARN util.NativeCodeLoader: Unable to load native-hadoop library for your platform... using
builtin-java classes where applicable
Samsung|Optima|14|Madhya Pradesh|132401|14200
Onida|Lucid|18|Uttar Pradesh|232401|16200
Akai|Decent|16|Kerala|922401|12200
Lava|Attention|20|Assam|454601|24200
Zen|Super|14|Maharashtra|619082|9200
Samsung|Optima|14|Madhya Pradesh|132401|14200
Onida|Lucid|18|Uttar Pradesh|232401|16200
Onida|Decent|14|Uttar Pradesh|232401|16200
Lava|Attention|20|Assam|454601|24200
Zen|Super|14|Maharashtra|619082|9200
Samsung|Optima|14|Madhya Pradesh|132401|14200
Samsung|Decent|16|Kerala|922401|12200
Lava|Attention|20|Assam|454601|24200
Samsung|Super|14|Maharashtra|619082|9200
Samsung|Super|14|Maharashtra|619082|9200
Samsung|Super|14|Maharashtra|619082|9200

Task 2:
Write a Map Reduce program to calculate the total units sold for each Company.

Solution2 :

Eclipse Code:

```java
package com.acadgild.mapreduce.assignment.task2;

import java.io.IOException;
import org.apache.hadoop.conf.Configuration;
import org.apache.hadoop.fs.Path;
import org.apache.hadoop.io.IntWritable;
import org.apache.hadoop.io.LongWritable;
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 CompanyUnit {
      public static class UnitMapper
    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[] SingleTelevisionData = valueString.split("\\|");

         context.write(new Text(SingleTelevisionData[0]),one);

 }
}

      public static class UnitSumReducer
    extends Reducer<Text,IntWritable,Text,IntWritable> {
 private IntWritable result = new IntWritable();

 public void reduce(Text key, Iterable<IntWritable> values,
                    Context context
                    ) throws IOException, InterruptedException {
   int sum = 0;
   for (IntWritable val : values) {
     sum += val.get();
   }
   result.set(sum);
   context.write(key, result);
 }
}
```

```java
        public static void main(String[] args) throws Exception {
                Configuration conf = new Configuration();
                Job job = Job.getInstance(conf, "filter television");
                job.setJarByClass(CompanyUnit.class);
                job.setMapperClass(UnitMapper.class);
                job.setReducerClass(UnitSumReducer.class);
                job.setOutputKeyClass(Text.class);
                job.setOutputValueClass(IntWritable.class);
                FileInputFormat.addInputPath(job, new Path(args[0]));
                FileOutputFormat.setOutputPath(job, new Path(args[1]));
                System.exit(job.waitForCompletion(true) ? 0 : 1);
        }
}
```

Converted to jar : companyunit.jar

Execution:

[acadgild@localhost ~]$ hadoop jar /home/acadgild/Desktop/companyunit.jar
com.acadgild.mapreduce.assignment.task2.CompanyUnit
/user/acadgild/hadoopdata/assignment/mapreduce/task1/input/television.txt
/user/acadgild/hadoopdata/assignment/mapreduce/task1/outputcompanyunit
18/06/27 06:05:01 WARN util.NativeCodeLoader: Unable to load native-hadoop library for your
platform... using builtin-java classes where applicable
18/06/27 06:05:03 INFO client.RMProxy: Connecting to ResourceManager at localhost/127.0.0.1:8032
18/06/27 06:05:04 WARN mapreduce.JobResourceUploader: Hadoop command-line option parsing not
performed. Implement the Tool interface and execute your application with ToolRunner to remedy this.
18/06/27 06:05:05 INFO input.FileInputFormat: Total input paths to process : 1
18/06/27 06:05:05 INFO mapreduce.JobSubmitter: number of splits:1
18/06/27 06:05:05 INFO mapreduce.JobSubmitter: Submitting tokens for job:
job_1530051047296_0011
18/06/27 06:05:06 INFO impl.YarnClientImpl: Submitted application
application_1530051047296_0011
18/06/27 06:05:06 INFO mapreduce.Job: The url to track the job:
http://localhost:8088/proxy/application_1530051047296_0011/
18/06/27 06:05:06 INFO mapreduce.Job: Running job: job_1530051047296_0011
18/06/27 06:05:17 INFO mapreduce.Job: Job job_1530051047296_0011 running in uber mode : false
18/06/27 06:05:17 INFO mapreduce.Job: map 0% reduce 0%
18/06/27 06:05:25 INFO mapreduce.Job: map 100% reduce 0%
18/06/27 06:05:36 INFO mapreduce.Job: map 100% reduce 100%
18/06/27 06:05:36 INFO mapreduce.Job: Job job_1530051047296_0011 completed successfully
18/06/27 06:05:36 INFO mapreduce.Job: Counters: 49
        File System Counters
                FILE: Number of bytes read=225
                FILE: Number of bytes written=215401
                FILE: Number of read operations=0

```
            FILE: Number of large read operations=0
            FILE: Number of write operations=0
            HDFS: Number of bytes read=892
            HDFS: Number of bytes written=43
            HDFS: Number of read operations=6
            HDFS: Number of large read operations=0
            HDFS: Number of write operations=2
    Job Counters
            Launched map tasks=1
            Launched reduce tasks=1
            Data-local map tasks=1
            Total time spent by all maps in occupied slots (ms)=6066
            Total time spent by all reduces in occupied slots (ms)=6955
            Total time spent by all map tasks (ms)=6066
            Total time spent by all reduce tasks (ms)=6955
            Total vcore-milliseconds taken by all map tasks=6066
            Total vcore-milliseconds taken by all reduce tasks=6955
            Total megabyte-milliseconds taken by all map tasks=6211584
            Total megabyte-milliseconds taken by all reduce tasks=7121920
    Map-Reduce Framework
            Map input records=18
            Map output records=18
            Map output bytes=183
            Map output materialized bytes=225
            Input split bytes=159
            Combine input records=0
            Combine output records=0
            Reduce input groups=6
            Reduce shuffle bytes=225
            Reduce input records=18
            Reduce output records=6
            Spilled Records=36
            Shuffled Maps =1
            Failed Shuffles=0
            Merged Map outputs=1
            GC time elapsed (ms)=279
            CPU time spent (ms)=1810
            Physical memory (bytes) snapshot=236269568
            Virtual memory (bytes) snapshot=681746432
            Total committed heap usage (bytes)=139972608
    Shuffle Errors
            BAD_ID=0
            CONNECTION=0
            IO_ERROR=0
            WRONG_LENGTH=0
            WRONG_MAP=0
            WRONG_REDUCE=0
    File Input Format Counters
            Bytes Read=733
```

       File Output Format Counters
           Bytes Written=43
[acadgild@localhost ~]$ hdfs dfs -ls /user/acadgild/hadoopdata/assignment/mapreduce/task1/outputcompanyunit
18/06/27 06:06:06 WARN util.NativeCodeLoader: Unable to load native-hadoop library for your platform... using builtin-java classes where applicable
Found 2 items
-rw-r--r--   1 acadgild supergroup   0 2018-06-27 06:05 /user/acadgild/hadoopdata/assignment/mapreduce/task1/outputcompanyunit/_SUCCESS
-rw-r--r--   1 acadgild supergroup   43 2018-06-27 06:05 /user/acadgild/hadoopdata/assignment/mapreduce/task1/outputcompanyunit/part-r-00000
[acadgild@localhost ~]$ hdfs dfs -cat /user/acadgild/hadoopdata/assignment/mapreduce/task1/outputcompanyunit/part-r-0000018/06/27 06:06:57 WARN util.NativeCodeLoader: Unable to load native-hadoop library for your platform... using builtin-java classes where applicable
Akai   1
Lava   3
NA   1
Onida 4
Samsung   7
Zen   2

Task 3:
Write a Map Reduce program to calculate the total units sold in each state for Onida company.

Solution3 :

Eclipse Code:

```java
package com.acadgild.mapreduce.assignment.task3;

import java.io.IOException;

import org.apache.hadoop.conf.Configuration;
import org.apache.hadoop.fs.Path;
import org.apache.hadoop.io.IntWritable;
import org.apache.hadoop.io.LongWritable;
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 OnidaStateUnit {
    public static class OnidaUnitMapper
```

```java
        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[] SingleTelevisionData = valueString.split("\\|");

        if(SingleTelevisionData[0].matches("Onida"))
         context.write(new Text(SingleTelevisionData[3]),one);

 }
}

      public static class OnidaUnitSumReducer
    extends Reducer<Text,IntWritable,Text,IntWritable> {
 private IntWritable result = new IntWritable();

 public void reduce(Text key, Iterable<IntWritable> values,
                    Context context
                    ) throws IOException, InterruptedException {
   int sum = 0;
   for (IntWritable val : values) {
     sum += val.get();
   }
   result.set(sum);
   context.write(key, result);
 }
}


      public static void main(String[] args) throws Exception {
            Configuration conf = new Configuration();
            Job job = Job.getInstance(conf, "filter television");
            job.setJarByClass(OnidaStateUnit.class);
            job.setMapperClass(OnidaUnitMapper.class);
            job.setReducerClass(OnidaUnitSumReducer.class);
            job.setOutputKeyClass(Text.class);
            job.setOutputValueClass(IntWritable.class);
            FileInputFormat.addInputPath(job, new Path(args[0]));
            FileOutputFormat.setOutputPath(job, new Path(args[1]));
            System.exit(job.waitForCompletion(true) ? 0 : 1);
          }
}
```

Converted to jar : onidaunit.jar

Execution:

[acadgild@localhost ~]$ hadoop jar /home/acadgild/Desktop/onidaunit.jar
com.acadgild.mapreduce.assignment.task3.OnidaStateUnit

/user/acadgild/hadoopdata/assignment/mapreduce/task1/input/television.txt
/user/acadgild/hadoopdata/assignment/mapreduce/task1/outputonidaunit
18/06/27 06:26:39 WARN util.NativeCodeLoader: Unable to load native-hadoop library for your platform... using builtin-java classes where applicable
18/06/27 06:26:40 INFO client.RMProxy: Connecting to ResourceManager at localhost/127.0.0.1:8032
18/06/27 06:26:42 WARN mapreduce.JobResourceUploader: Hadoop command-line option parsing not performed. Implement the Tool interface and execute your application with ToolRunner to remedy this.
18/06/27 06:26:42 INFO input.FileInputFormat: Total input paths to process : 1
18/06/27 06:26:42 INFO mapreduce.JobSubmitter: number of splits:1
18/06/27 06:26:43 INFO mapreduce.JobSubmitter: Submitting tokens for job: job_1530051047296_0012
18/06/27 06:26:43 INFO impl.YarnClientImpl: Submitted application application_1530051047296_0012
18/06/27 06:26:43 INFO mapreduce.Job: The url to track the job: http://localhost:8088/proxy/application_1530051047296_0012/
18/06/27 06:26:43 INFO mapreduce.Job: Running job: job_1530051047296_0012
18/06/27 06:26:56 INFO mapreduce.Job: Job job_1530051047296_0012 running in uber mode : false
18/06/27 06:26:56 INFO mapreduce.Job: map 0% reduce 0%
18/06/27 06:27:06 INFO mapreduce.Job: map 100% reduce 0%
18/06/27 06:27:16 INFO mapreduce.Job: map 100% reduce 100%
18/06/27 06:27:17 INFO mapreduce.Job: Job job_1530051047296_0012 completed successfully
18/06/27 06:27:18 INFO mapreduce.Job: Counters: 49
        File System Counters
                FILE: Number of bytes read=79
                FILE: Number of bytes written=215137
                FILE: Number of read operations=0
                FILE: Number of large read operations=0
                FILE: Number of write operations=0
                HDFS: Number of bytes read=892
                HDFS: Number of bytes written=25
                HDFS: Number of read operations=6
                HDFS: Number of large read operations=0
                HDFS: Number of write operations=2
        Job Counters
                Launched map tasks=1
                Launched reduce tasks=1
                Data-local map tasks=1
                Total time spent by all maps in occupied slots (ms)=8044
                Total time spent by all reduces in occupied slots (ms)=8229
                Total time spent by all map tasks (ms)=8044
                Total time spent by all reduce tasks (ms)=8229
                Total vcore-milliseconds taken by all map tasks=8044
                Total vcore-milliseconds taken by all reduce tasks=8229
                Total megabyte-milliseconds taken by all map tasks=8237056
                Total megabyte-milliseconds taken by all reduce tasks=8426496
        Map-Reduce Framework
                Map input records=18
                Map output records=4
                Map output bytes=65

Map output materialized bytes=79
Input split bytes=159
Combine input records=0
Combine output records=0
Reduce input groups=2
Reduce shuffle bytes=79
Reduce input records=4
Reduce output records=2
Spilled Records=8
Shuffled Maps =1
Failed Shuffles=0
Merged Map outputs=1
GC time elapsed (ms)=369
CPU time spent (ms)=1990
Physical memory (bytes) snapshot=236556288
Virtual memory (bytes) snapshot=680697856
Total committed heap usage (bytes)=139993088
Shuffle Errors
BAD_ID=0
CONNECTION=0
IO_ERROR=0
WRONG_LENGTH=0
WRONG_MAP=0
WRONG_REDUCE=0
File Input Format Counters
Bytes Read=733
File Output Format Counters
Bytes Written=25
[acadgild@localhost ~]$ hdfs dfs -ls
/user/acadgild/hadoopdata/assignment/mapreduce/task1/outputonidaunit
18/06/27 06:29:08 WARN util.NativeCodeLoader: Unable to load native-hadoop library for your
platform... using builtin-java classes where applicable
Found 2 items
-rw-r--r--   1 acadgild supergroup   0 2018-06-27 06:27
/user/acadgild/hadoopdata/assignment/mapreduce/task1/outputonidaunit/_SUCCESS
-rw-r--r--   1 acadgild supergroup   25 2018-06-27 06:27
/user/acadgild/hadoopdata/assignment/mapreduce/task1/outputonidaunit/part-r-00000
[acadgild@localhost ~]$ hdfs dfs -cat
/user/acadgild/hadoopdata/assignment/mapreduce/task1/outputonidaunit/part-r-00000
18/06/27 06:30:13 WARN util.NativeCodeLoader: Unable to load native-hadoop library for your
platform... using builtin-java classes where applicable
Kerala 1
Uttar Pradesh 3