

Assignment 4

- Create television.txt in local File system and put the file to Hadoop file system.

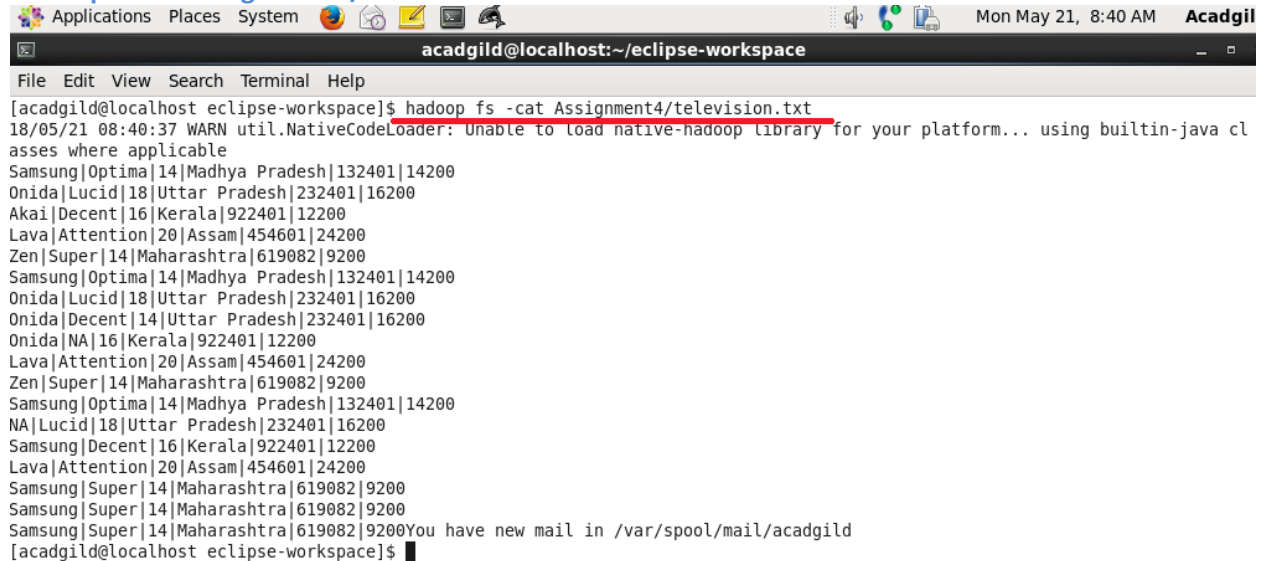
Command:

Hadoop fs -put television.txt Assignment4/television.txt

- Display the contents of television.txt

Command:

Hadoop fs -cat Assignment4/television.txt



```
acadmild@localhost: ~/eclipse-workspace
File Edit View Search Terminal Help
[acadmild@localhost eclipse-workspace]$ hadoop fs -cat Assignment4/television.txt
18/05/21 08:40:37 WARN util.NativeCodeLoader: Unable to load native-hadoop library for your platform... using builtin-java cl
asses 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|9200You have new mail in /var/spool/mail/acadmild
[acadmild@localhost eclipse-workspace]$
```

Task 1:

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

Step 1:

Create jar file "Invalid.jar"

Step 2:

Execute the jar file using the below command

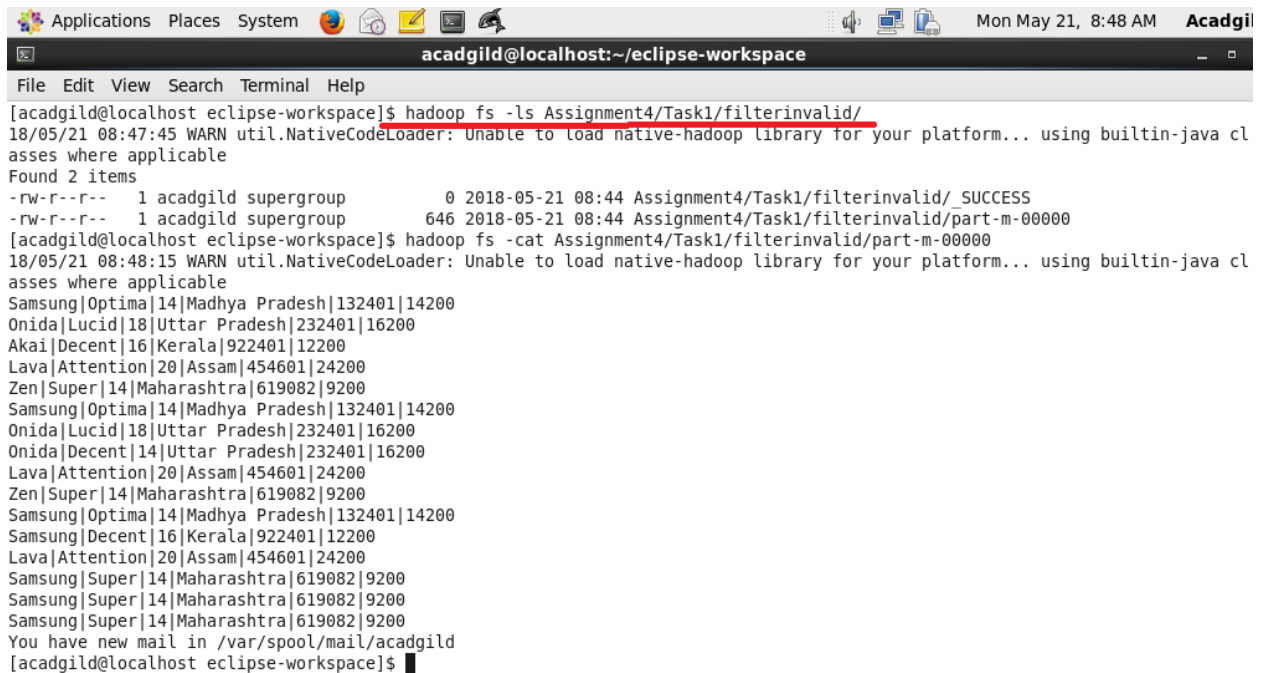
Hadoop jar Invalid.jar com.acadmild.FilterInvalid.Invalid Assignment4/television.txt
Assignmnet4/Task1/filterinvalid

```
Applications Places System Mon May 21, 8:44 AM Acadgild
acadgild@localhost:~/eclipse-workspace
File Edit View Search Terminal Help
[acadgild@localhost eclipse-workspace]$ hadoop jar Invalid.jar com/acadgild/FilterInvalid/Invalid Assignment4/television.txt
Assignment4/Task1/filterinvalid
18/05/21 08:43:15 WARN util.NativeCodeLoader: Unable to load native-hadoop library for your platform... using builtin-java cl
asses where applicable
18/05/21 08:43:17 INFO client.RMPProxy: Connecting to ResourceManager at localhost/127.0.0.1:8032
18/05/21 08:43:21 WARN mapreduce.JobResourceUploader: Hadoop command-line option parsing not performed. Implement the Tool in
terface and execute your application with ToolRunner to remedy this.
18/05/21 08:43:22 INFO input.FileInputFormat: Total input paths to process : 1
18/05/21 08:43:23 INFO mapreduce.JobSubmitter: number of splits:1
18/05/21 08:43:24 INFO mapreduce.JobSubmitter: Submitting tokens for job: job_1526774583418_0043
18/05/21 08:43:26 INFO impl.YarnClientImpl: Submitted application application_1526774583418_0043
18/05/21 08:43:26 INFO mapreduce.Job: The url to track the job: http://localhost:8088/proxy/application_1526774583418_0043/
18/05/21 08:43:26 INFO mapreduce.Job: Running job: job_1526774583418_0043
18/05/21 08:43:56 INFO mapreduce.Job: Job job_1526774583418_0043 running in uber mode : false
18/05/21 08:43:56 INFO mapreduce.Job: map 0% reduce 0%
18/05/21 08:44:19 INFO mapreduce.Job: map 100% reduce 0%
18/05/21 08:44:21 INFO mapreduce.Job: Job job_1526774583418_0043 completed successfully
18/05/21 08:44:22 INFO mapreduce.Job: Counters: 30
File System Counters
  FILE: Number of bytes read=0
  FILE: Number of bytes written=107043
  FILE: Number of read operations=0
  FILE: Number of large read operations=0
  FILE: Number of write operations=0
  HDFS: Number of bytes read=860
  HDFS: Number of bytes written=646
  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)=20343
  Total time spent by all reduces in occupied slots (ms)=0
  Total time spent by all map tasks (ms)=20343
  Total vcore-milliseconds taken by all map tasks=20343
  Total megabyte-milliseconds taken by all map tasks=20831232
Map-Reduce Framework
  Map input records=18
```

```
Applications Places System Mon May 21, 8:44 AM Acadgild
acadgild@localhost:~/eclipse-workspace
File Edit View Search Terminal Help
18/05/21 08:44:21 INFO mapreduce.Job: Job job_1526774583418_0043 completed successfully
18/05/21 08:44:22 INFO mapreduce.Job: Counters: 30
File System Counters
  FILE: Number of bytes read=0
  FILE: Number of bytes written=107043
  FILE: Number of read operations=0
  FILE: Number of large read operations=0
  FILE: Number of write operations=0
  HDFS: Number of bytes read=860
  HDFS: Number of bytes written=646
  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)=20343
  Total time spent by all reduces in occupied slots (ms)=0
  Total time spent by all map tasks (ms)=20343
  Total vcore-milliseconds taken by all map tasks=20343
  Total megabyte-milliseconds taken by all map tasks=20831232
Map-Reduce Framework
  Map input records=18
  Map output records=16
  Input split bytes=127
  Spilled Records=0
  Failed Shuffles=0
  Merged Map outputs=0
  GC time elapsed (ms)=99
  CPU time spent (ms)=1020
  Physical memory (bytes) snapshot=92962816
  Virtual memory (bytes) snapshot=2056757248
  Total committed heap usage (bytes)=32571392
File Input Format Counters
  Bytes Read=733
File Output Format Counters
  Bytes Written=646
You have new mail in /var/spool/mail/acadgild
[acadgild@localhost eclipse-workspace]$
```

Step 3:

Display the contents of the output file. The records containing “NA” are not present.



The screenshot shows a terminal window titled 'Acadgild@localhost: ~/eclipse-workspace'. The terminal displays the following commands and output:

```
[acadgild@localhost eclipse-workspace]$ hadoop fs -ls Assignment4/Task1/filterinvalid/
18/05/21 08:47:45 WARN util.NativeCodeLoader: Unable to load native-hadoop library for your platform... using builtin-java cl
asses where applicable
Found 2 items
-rw-r--r-- 1 acadgild supergroup 0 2018-05-21 08:44 Assignment4/Task1/filterinvalid/_SUCCESS
-rw-r--r-- 1 acadgild supergroup 646 2018-05-21 08:44 Assignment4/Task1/filterinvalid/part-m-00000
[acadgild@localhost eclipse-workspace]$ hadoop fs -cat Assignment4/Task1/filterinvalid/part-m-00000
18/05/21 08:48:15 WARN util.NativeCodeLoader: Unable to load native-hadoop library for your platform... using builtin-java cl
asses 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
You have new mail in /var/spool/mail/acadgild
[acadgild@localhost eclipse-workspace]$
```

Code:

```
package com.acadgild.FilterInvalid;
import java.io.IOException;
import org.apache.hadoop.conf.Configuration;
import org.apache.hadoop.fs.Path;
import org.apache.hadoop.io.*;
import org.apache.hadoop.mapreduce.lib.input.FileInputFormat;
import org.apache.hadoop.mapreduce.Job;
import org.apache.hadoop.mapreduce.Mapper;
import org.apache.hadoop.mapreduce.lib.output.FileOutputFormat;

public class Invalid {
    //Mapper Class
    public static class filtermap extends Mapper<LongWritable,Text,Text,Text>
    {
        //override map function
        public void map(LongWritable key, Text value,
            Context context)
```

```

throws IOException,InterruptedException
{
    //Convert Value to string
    String valueString = value.toString();
    //split the data on the basis of the separator "|"
    String[] finstr= valueString.split("\\|");
    //Assign Company Name and Product Name to the variables cname and pname
    String cname = finstr[0].trim();
    String pname = finstr[1].trim();
    /*if cname or pname does not contain "NA" then write the data present in value parameter.
    context.Write is used to write the data in the output file*/
    if(!cname.equals("NA") && !pname.equals("NA"))
        context.write(value, null);
}
}

```

//Driver code is present in main()

```

public static void main(String[] args) throws IllegalArgumentException, IOException,
ClassNotFoundException, InterruptedException
{
    Configuration conf = new Configuration();
    //Creates a new job with name "MapOnlyJob"
    Job job = new Job(conf, "MapOnlyJob");
    //SetJarname to Invalid
    job.setJarByClass(Invalid.class);
    //Set Mapper Class to filtermap
    job.setMapperClass(filtermap.class);
    //sets the type in which Output Key and Value will be written
    job.setOutputKeyClass(Text.class);
    job.setOutputValueClass(Text.class);
    // Sets reducer tasks to 0 as reducer is not used in the program
    job.setNumReduceTasks(0);
    //Sets Input and output paths from the arguments provided while executing jar file
    FileInputFormat.addInputPath(job, new Path(args[0]));
    FileOutputFormat.setOutputPath(job, new Path(args[1]));
    boolean result = job.waitForCompletion(true);
    System.exit(result ? 0 : 1);
}
}

```

Task 2:

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

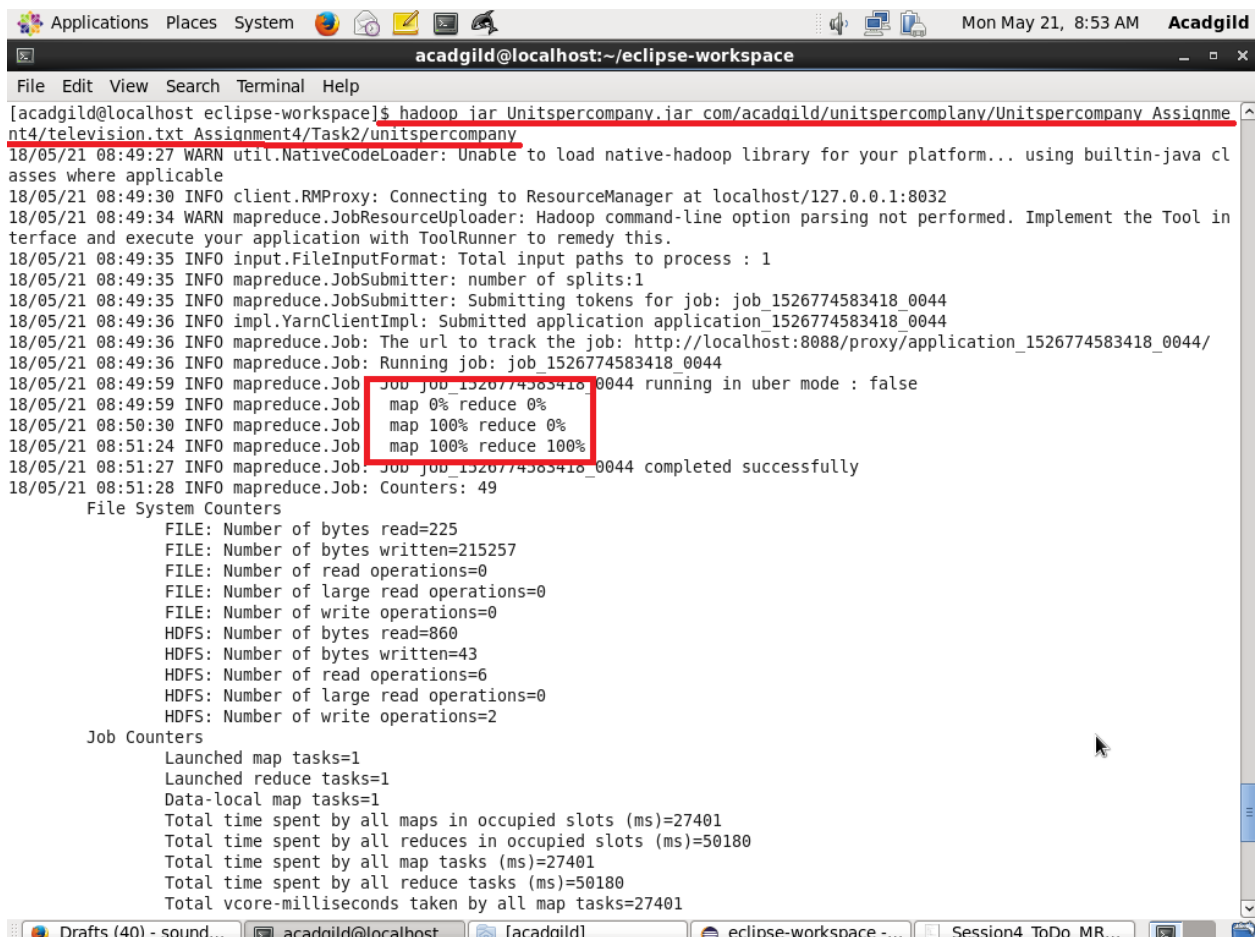
Step 1:

Create jar file "Unitspercompany.jar"

Step 2:

Execute the jar file using the below command

Hadoop jar Invalid.jar com.acadgild. unitspercompany. Unitspercompany
Assignment4/television.txt Assignmnet4/Task2/unitspercompany



```
[acadgild@localhost eclipse-workspace]$ hadoop jar Unitspercompany.jar com.acadgild.unitspercompany.Unitspercompany Assignment4/television.txt Assignment4/Task2/unitspercompany
18/05/21 08:49:27 WARN util.NativeCodeLoader: Unable to load native-hadoop library for your platform... using builtin-java classes where applicable
18/05/21 08:49:30 INFO client.RMProxy: Connecting to ResourceManager at localhost/127.0.0.1:8032
18/05/21 08:49:34 WARN mapreduce.JobResourceUploader: Hadoop command-line option parsing not performed. Implement the Tool interface and execute your application with ToolRunner to remedy this.
18/05/21 08:49:35 INFO input.FileInputFormat: Total input paths to process : 1
18/05/21 08:49:35 INFO mapreduce.JobSubmitter: number of splits:1
18/05/21 08:49:35 INFO mapreduce.JobSubmitter: Submitting tokens for job: job_1526774583418_0044
18/05/21 08:49:36 INFO impl.YarnClientImpl: Submitted application application_1526774583418_0044
18/05/21 08:49:36 INFO mapreduce.Job: The url to track the job: http://localhost:8088/proxy/application_1526774583418_0044/
18/05/21 08:49:36 INFO mapreduce.Job: Running job: job_1526774583418_0044
18/05/21 08:49:59 INFO mapreduce.Job: Job job_1526774583418_0044 running in uber mode : false
18/05/21 08:49:59 INFO mapreduce.Job: map 0% reduce 0%
18/05/21 08:50:30 INFO mapreduce.Job: map 100% reduce 0%
18/05/21 08:51:24 INFO mapreduce.Job: map 100% reduce 100%
18/05/21 08:51:27 INFO mapreduce.Job: Job job_1526774583418_0044 completed successfully
18/05/21 08:51:28 INFO mapreduce.Job: Counters: 49
File System Counters
  FILE: Number of bytes read=225
  FILE: Number of bytes written=215257
  FILE: Number of read operations=0
  FILE: Number of large read operations=0
  FILE: Number of write operations=0
  HDFS: Number of bytes read=860
  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)=27401
  Total time spent by all reduces in occupied slots (ms)=50180
  Total time spent by all map tasks (ms)=27401
  Total time spent by all reduce tasks (ms)=50180
  Total vcore-milliseconds taken by all map tasks=27401
```

```
Applications Places System Mon May 21, 8:53 AM Acadgild
acadgild@localhost:~/eclipse-workspace
File Edit View Search Terminal Help
Total time spent by all reduce tasks (ms)=50180
Total vcore-milliseconds taken by all map tasks=27401
Total vcore-milliseconds taken by all reduce tasks=50180
Total megabyte-milliseconds taken by all map tasks=28058624
Total megabyte-milliseconds taken by all reduce tasks=51384320
Map-Reduce Framework
  Map input records=18
  Map output records=18
  Map output bytes=183
  Map output materialized bytes=225
  Input split bytes=127
  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)=671
  CPU time spent (ms)=5210
  Physical memory (bytes) snapshot=287592448
  Virtual memory (bytes) snapshot=4117913600
  Total committed heap usage (bytes)=170004480
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
You have new mail in /var/spool/mail/acadgild
[acadgild@localhost eclipse-workspace]$
```

Step 3:

Display the contents of the output file.

```
Applications Places System Mon May 21, 8:55 AM Acadgild
acadgild@localhost:~/eclipse-workspace
File Edit View Search Terminal Help
You have new mail in /var/spool/mail/acadgild
[acadgild@localhost eclipse-workspace]$ hadoop fs -ls Assignment4/Task2/unitspercompany
18/05/21 08:54:23 WARN util.NativeCodeLoader: Unable to load native-hadoop library for your platform... using builtin-java cl
asses where applicable
Found 2 items
-rw-r--r-- 1 acadgild supergroup          0 2018-05-21 08:51 Assignment4/Task2/unitspercompany/_SUCCESS
-rw-r--r-- 1 acadgild supergroup        43 2018-05-21 08:51 Assignment4/Task2/unitspercompany/part-r-000000
[acadgild@localhost eclipse-workspace]$ hadoop fs -cat Assignment4/Task2/unitspercompany/part-r-000000
18/05/21 08:55:04 WARN util.NativeCodeLoader: Unable to load native-hadoop library for your platform... using builtin-java cl
asses where applicable
Akai      1
Lava      3
NA        1
Onida     4
Samsung   7
Zen       2
You have new mail in /var/spool/mail/acadgild
[acadgild@localhost eclipse-workspace]$
```

Code:

```
package com.acadgild.unitspercomplanly;

import java.io.IOException;
import java.util.Iterator;
import org.apache.hadoop.conf.Configuration;
import org.apache.hadoop.fs.Path;
import org.apache.hadoop.io.*;
import org.apache.hadoop.mapreduce.lib.input.FileInputFormat;
import org.apache.hadoop.mapreduce.lib.input.TextInputFormat;
import org.apache.hadoop.mapreduce.Job;
import org.apache.hadoop.mapreduce.Mapper;
import org.apache.hadoop.mapreduce.Reducer;
import org.apache.hadoop.mapreduce.lib.output.FileOutputFormat;
import org.apache.hadoop.mapreduce.lib.output.TextOutputFormat;

public class Unitspercompany {
    //Mapper Class
    public static class Unitsmapper extends Mapper<LongWritable, Text, Text, IntWritable>
    {
        //Assign One = 1
        private final static IntWritable one = new IntWritable(1);
        //override map function
        public void map(LongWritable key, Text value, Context context) throws
        IOException, InterruptedException
        {
            //Convert Value to string
            String finstr = value.toString();
            //split the data on the basis of the separator "|"
            String[] Tdata = finstr.split("\\|");
            //Write the value of Tdata[0] which contains company name.
            context.write(new Text(Tdata[0]), one);
        }
    }

    //Reducer Class
    public static class Unitsreducer extends Reducer<Text, IntWritable, Text, IntWritable>
    {
        //Override reduce function
        public void reduce(Text key, Iterable<IntWritable> values, Context context)
            throws IOException, InterruptedException
    }
}
```

```

    {
        int cname = 0;
        //Iterate through the values and generate the count of each company
        for (IntWritable value : values)
        {
            cname = cname + value.get();
        }
        /*write key and value pair into the output file : Key contains the company name, cname
contains the count*/
        context.write(key, new IntWritable(cname));
    }
}

```

//Driver Code

```

public static void main(String[] args) throws IllegalArgumentException, IOException,
ClassNotFoundException, InterruptedException
{
    Configuration conf = new Configuration();
    //Creates a new job with name "Unitspercompany"
    Job job = new Job(conf, "Unitspercompany");
    //SetJarname to Unitspercompany
    job.setJarByClass(Unitspercompany.class);
    //Set Mapper Class to Unitsmapper
    job.setMapperClass(Unitsmapper.class);

    //set Reducer class to Unitsreducer
    job.setReducerClass(Unitsreducer.class);
    //sets the type in which Output Key and Value will be written
    job.setOutputKeyClass(Text.class);
    job.setOutputValueClass(IntWritable.class);
    //Sets Input and output paths from the arguments provided while executing jar file
    FileInputFormat.addInputPath(job, new Path(args[0]));
    FileOutputFormat.setOutputPath(job, new Path(args[1]));
    boolean result = job.waitForCompletion(true);
    System.exit(result ? 0 : 1);
}
}

```


Task 3:

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

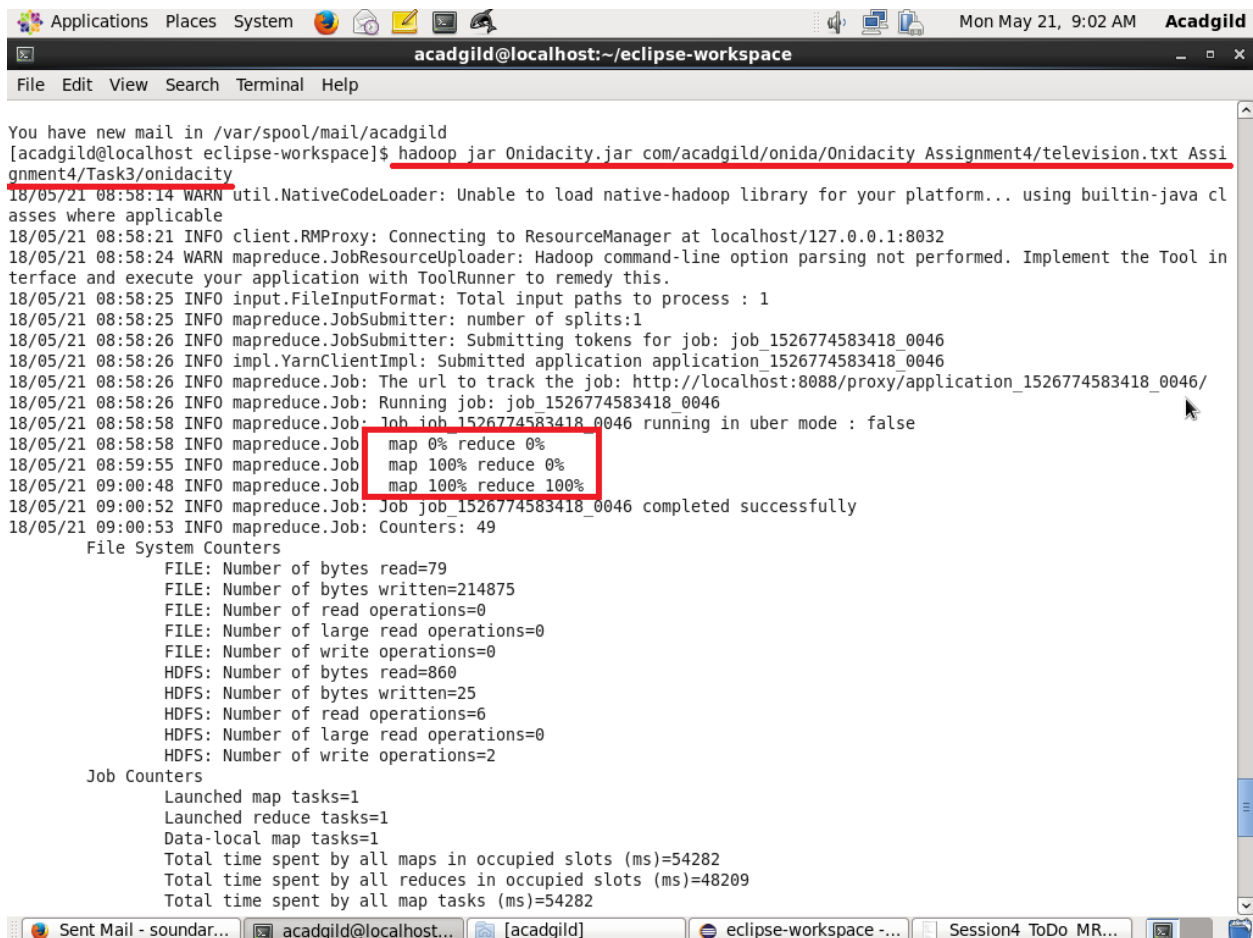
Step 1:

Create jar file "Invalid.jar"

Step 2:

Execute the jar file using the below command

Hadoop jar Invalid.jar com.acadgild. onida.Onida Assignment4/television.txt
Assignmnet4/Task3/onida



The screenshot shows a terminal window titled "acadgild@localhost: ~/eclipse-workspace". The command executed is `hadoop jar Onidacity.jar com/acadgild/onida/Onidacity Assignment4/television.txt Assignment4/Task3/onidacity`. The output shows the job running successfully with various status messages and counters. A red box highlights the progress of the map and reduce tasks.

```
You have new mail in /var/spool/mail/acadgild
[acadgild@localhost eclipse-workspace]$ hadoop jar Onidacity.jar com/acadgild/onida/Onidacity Assignment4/television.txt Assignment4/Task3/onidacity
18/05/21 08:58:14 WARN util.NativeCodeLoader: Unable to load native-hadoop library for your platform... using builtin-java classes where applicable
18/05/21 08:58:21 INFO client.RMProxy: Connecting to ResourceManager at localhost/127.0.0.1:8032
18/05/21 08:58:24 WARN mapreduce.JobResourceUploader: Hadoop command-line option parsing not performed. Implement the Tool in interface and execute your application with ToolRunner to remedy this.
18/05/21 08:58:25 INFO input.FileInputFormat: Total input paths to process : 1
18/05/21 08:58:25 INFO mapreduce.JobSubmitter: number of splits:1
18/05/21 08:58:26 INFO mapreduce.JobSubmitter: Submitting tokens for job: job_1526774583418_0046
18/05/21 08:58:26 INFO impl.YarnClientImpl: Submitted application application_1526774583418_0046
18/05/21 08:58:26 INFO mapreduce.Job: The url to track the job: http://localhost:8088/proxy/application_1526774583418_0046/
18/05/21 08:58:26 INFO mapreduce.Job: Running job: job_1526774583418_0046
18/05/21 08:58:58 INFO mapreduce.Job: Job job_1526774583418_0046 running in uber mode : false
18/05/21 08:58:58 INFO mapreduce.Job: map 0% reduce 0%
18/05/21 08:59:55 INFO mapreduce.Job: map 100% reduce 0%
18/05/21 09:00:48 INFO mapreduce.Job: map 100% reduce 100%
18/05/21 09:00:52 INFO mapreduce.Job: Job job_1526774583418_0046 completed successfully
18/05/21 09:00:53 INFO mapreduce.Job: Counters: 49
  File System Counters
    FILE: Number of bytes read=79
    FILE: Number of bytes written=214875
    FILE: Number of read operations=0
    FILE: Number of large read operations=0
    FILE: Number of write operations=0
    HDFS: Number of bytes read=860
    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)=54282
    Total time spent by all reduces in occupied slots (ms)=48209
    Total time spent by all map tasks (ms)=54282
```

```
Applications Places System Mon May 21, 9:02 AM Acadgild
acadgild@localhost:~/eclipse-workspace
File Edit View Search Terminal Help
Total time spent by all reduce tasks (ms)=48209
Total vcore-milliseconds taken by all map tasks=54282
Total vcore-milliseconds taken by all reduce tasks=48209
Total megabyte-milliseconds taken by all map tasks=55584768
Total megabyte-milliseconds taken by all reduce tasks=49366016
Map-Reduce Framework
  Map input records=18
  Map output records=4
  Map output bytes=65
  Map output materialized bytes=79
  Input split bytes=127
  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)=998
  CPU time spent (ms)=6460
  Physical memory (bytes) snapshot=283975680
  Virtual memory (bytes) snapshot=4117913600
  Total committed heap usage (bytes)=170004480
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
You have new mail in /var/spool/mail/acadgild
[acadgild@localhost eclipse-workspace]$
```

Step 3:

Display the contents of the output file.

```
Applications Places System Mon May 21, 9:04 AM Acadgild
acadgild@localhost:~/eclipse-workspace
File Edit View Search Terminal Help
You have new mail in /var/spool/mail/acadgild
[acadgild@localhost eclipse-workspace]$ hadoop fs -ls Assignment4/Task3/onidacity
18/05/21 09:03:21 WARN util.NativeCodeLoader: Unable to load native-hadoop library for your platform... using builtin-java cl
asses where applicable
Found 2 items
-rw-r--r-- 1 acadgild supergroup          0 2018-05-21 09:00 Assignment4/Task3/onidacity/_SUCCESS
-rw-r--r-- 1 acadgild supergroup        25 2018-05-21 09:00 Assignment4/Task3/onidacity/part-r-000000
[acadgild@localhost eclipse-workspace]$ hadoop fs -cat Assignment4/Task3/onidacity/part-r-000000
18/05/21 09:04:23 WARN util.NativeCodeLoader: Unable to load native-hadoop library for your platform... using builtin-java cl
asses where applicable
Kerala 1
Uttar Pradesh 3
You have new mail in /var/spool/mail/acadgild
[acadgild@localhost eclipse-workspace]$
```

Code:

```
package com.acadgild.onida;
import java.io.IOException;
import java.util.Iterator;
import org.apache.hadoop.conf.Configuration;
import org.apache.hadoop.fs.Path;
import org.apache.hadoop.io.*;
import org.apache.hadoop.mapreduce.lib.input.FileInputFormat;
import org.apache.hadoop.mapreduce.lib.input.TextInputFormat;
import org.apache.hadoop.mapreduce.Job;
import org.apache.hadoop.mapreduce.Mapper;
import org.apache.hadoop.mapreduce.Reducer;
import org.apache.hadoop.mapreduce.lib.output.FileOutputFormat;
import org.apache.hadoop.mapreduce.lib.output.TextOutputFormat;

public class Onidacity {

    //Mapper Class
    public static class Onidamapper extends Mapper<LongWritable, Text, Text, IntWritable>
    {
        private final static IntWritable one = new IntWritable(1);
        //Override map function
        public void map(LongWritable key, Text value,
            Context context) throws IOException, InterruptedException
        {
            //Convert Value to string
            String finstr = value.toString();
            //split the data on the basis of the separator "|"
            String[] Tdata = finstr.split("\\|");
            //assign company name to the variable company. Tdata[0] contains company name
            String company = Tdata[0];
            //write the state name which is present in Tdata[3] only if company name is "Onida"
            if(company.equals("Onida"))
                context.write(new Text(Tdata[3]), one);
        }
    }

    //Reducer Class
    public static class Onidareducer extends Reducer<Text, IntWritable, Text, IntWritable>
    {
        public void reduce(Text key, Iterable<IntWritable> values, Context context)
            throws
            IOException, InterruptedException
        {
            int cname = 0;
            //Iterate through the values and generate the count of each company
            for (IntWritable value : values)
```

```

    {
        cname = cname + value.get();
    }
    /*write key and value pair into the output file : Key contains the company name, cname
contains the count*/
    context.write(key, new IntWritable(cname));
}
}

//Driver Code
public static void main(String[] args) throws IllegalArgumentException, IOException,
ClassNotFoundException, InterruptedException
{
    Configuration conf = new Configuration();

    //Creates a new job with name "OnidaCompany"
    Job job = new Job(conf, "OnidaCompany");

    //Set Jarname to Onidacity
    job.setJarByClass(Onidacity.class);

    //Set Mapper Class to Onidamapper
    job.setMapperClass(Onidamapper.class);

    //sets the type in which Output Key and Value will be written
    job.setOutputKeyClass(Text.class);
    job.setOutputValueClass(IntWritable.class);

    //set Reducer class to Onidareducer
    job.setReducerClass(Onidareducer.class);

    //Sets Input and output paths from the arguments provided while executing jar file
    FileInputFormat.addInputPath(job, new Path(args[0]));
    FileOutputFormat.setOutputPath(job, new Path(args[1]));
    boolean result = job.waitForCompletion(true);
    System.exit(result ? 0 : 1);
}
}

```