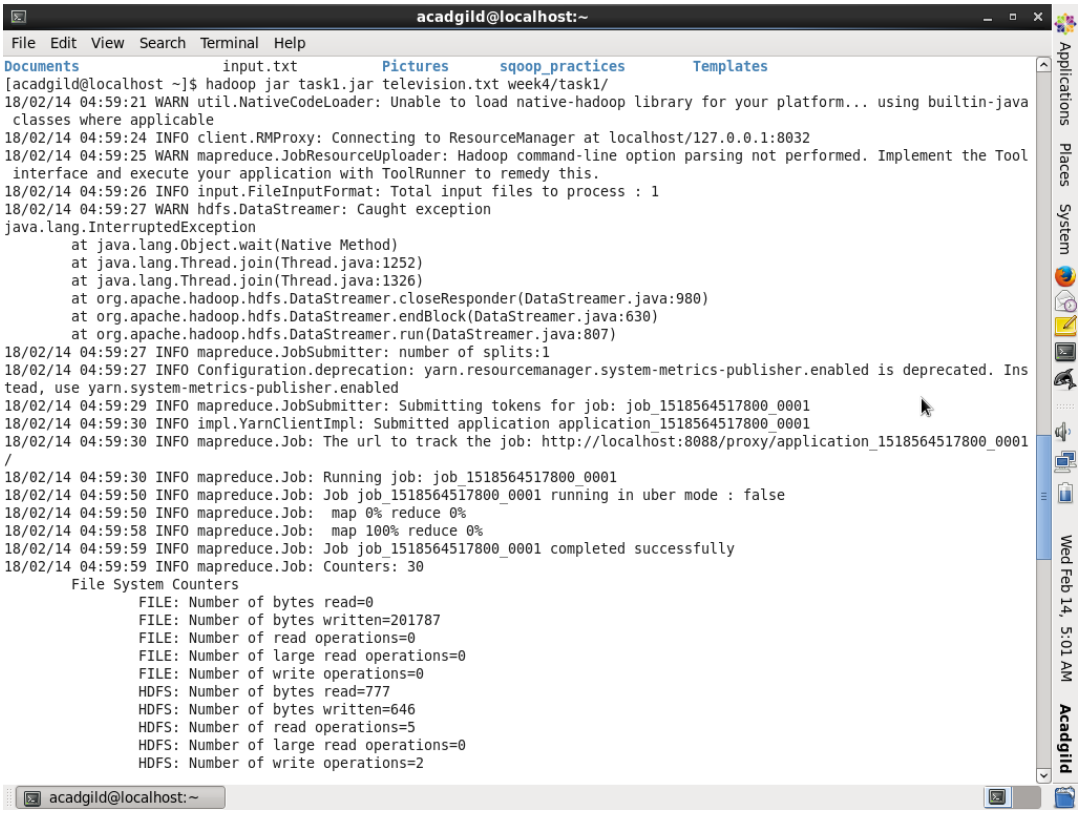
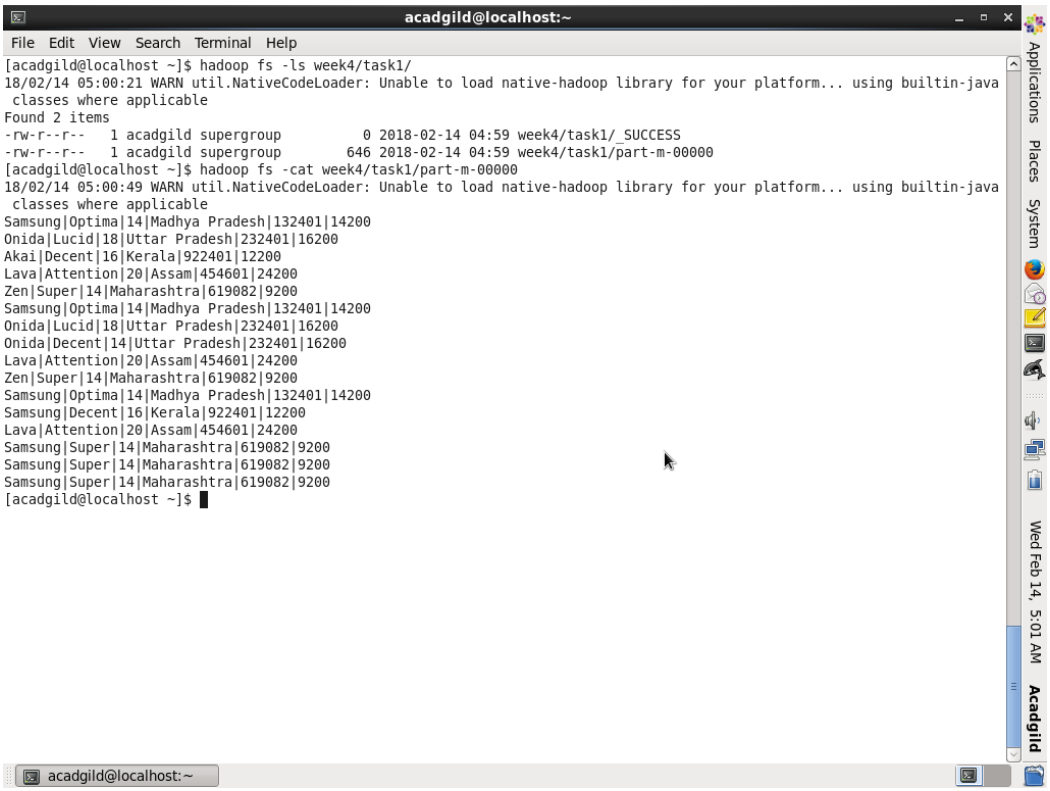
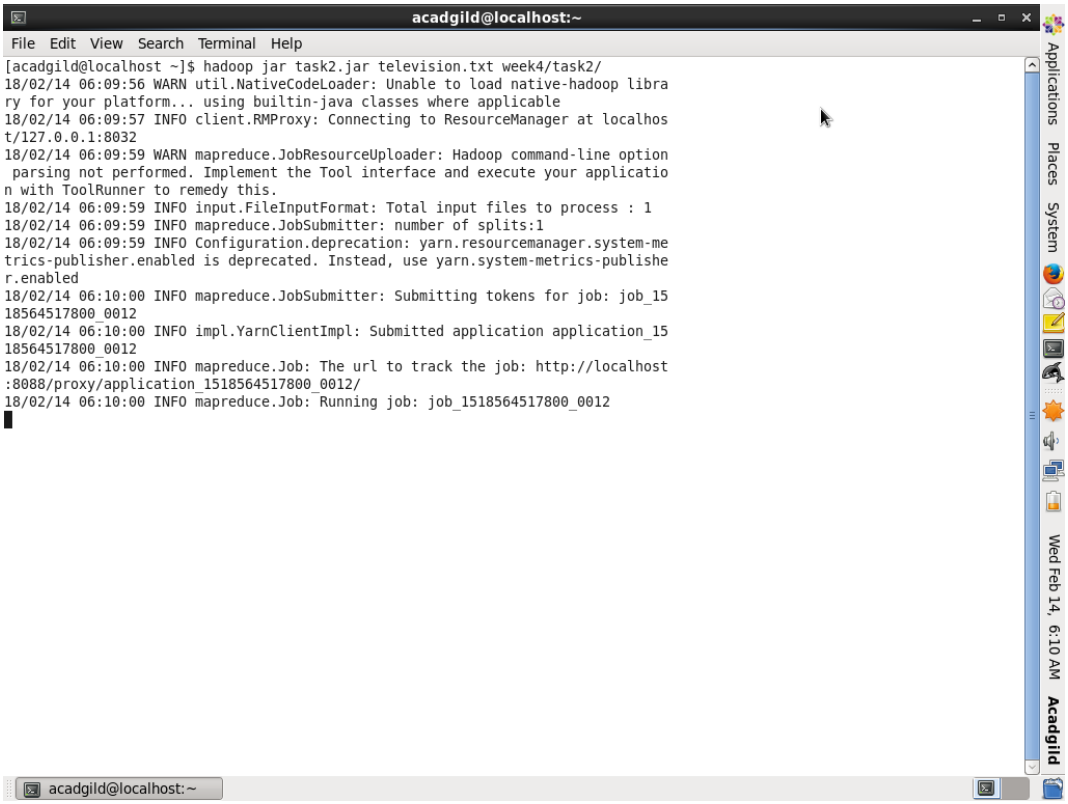
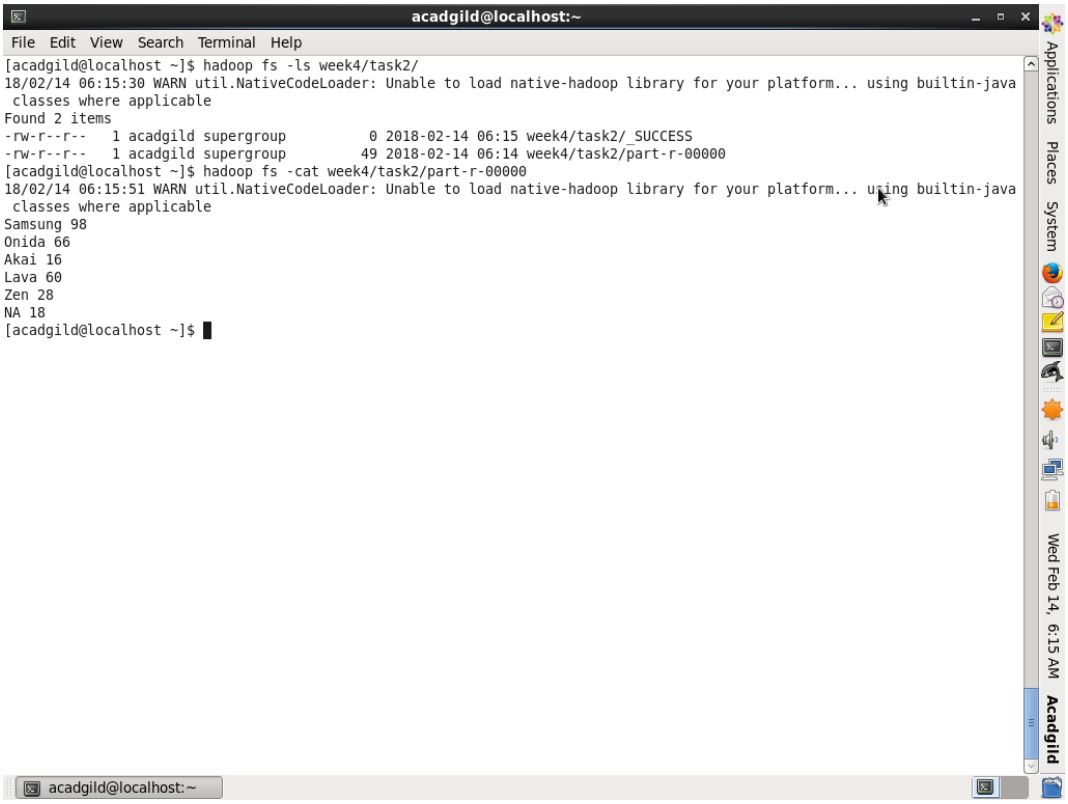
1. Write a Map Reduce program to filter out the invalid records. Map only job will fit for this context.



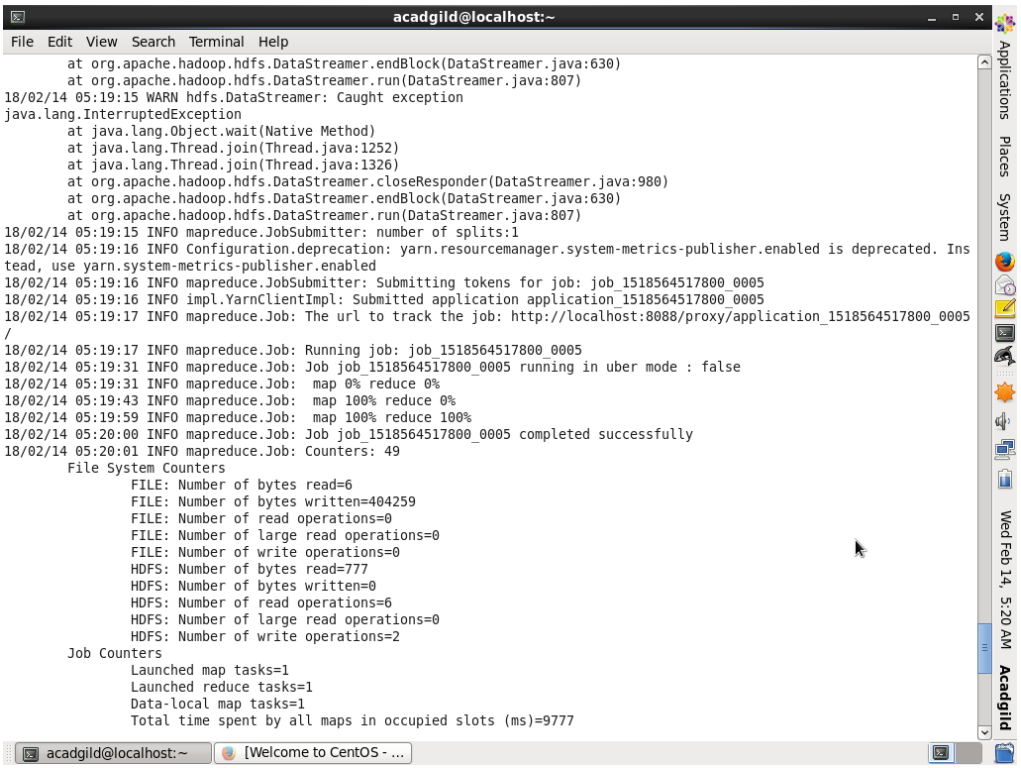


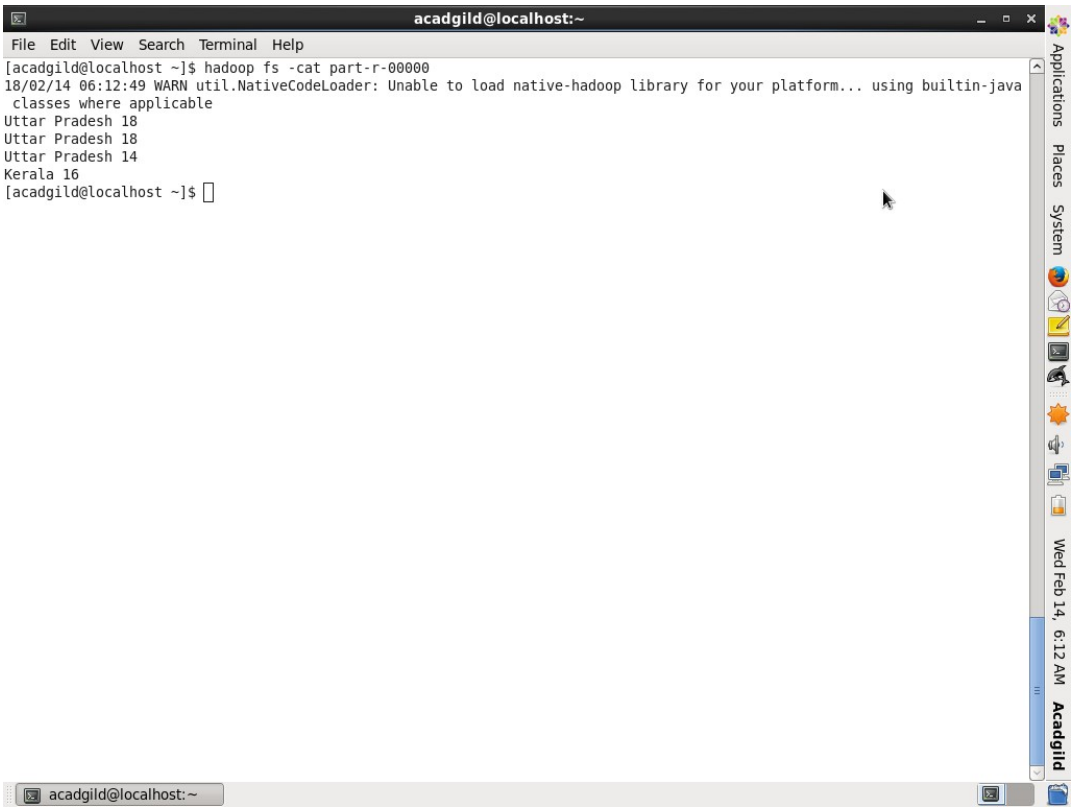
2. Write a Map Reduce program to calculate the total units sold for each Company.





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


acadgild@localhost:~
File Edit View Search Terminal Help
at org.apache.hadoop.hdfs.DataStreamer.endBlock(DataStreamer.java:630)
at org.apache.hadoop.hdfs.DataStreamer.run(DataStreamer.java:807)
18/02/14 05:19:15 WARN hdfs.DataStreamer: Caught exception
java.lang.InterruptedException
at java.lang.Object.wait(Native Method)
at java.lang.Thread.join(Thread.java:1252)
at java.lang.Thread.join(Thread.java:1326)
at org.apache.hadoop.hdfs.DataStreamer.closeResponder(DataStreamer.java:980)
at org.apache.hadoop.hdfs.DataStreamer.endBlock(DataStreamer.java:630)
at org.apache.hadoop.hdfs.DataStreamer.run(DataStreamer.java:807)
18/02/14 05:19:15 INFO mapreduce.JobSubmitter: number of splits:1
18/02/14 05:19:16 INFO Configuration.deprecation: yarn.resourcemanager.system-metrics-publisher.enabled is deprecated. Ins
tead, use yarn.system-metrics-publisher.enabled
18/02/14 05:19:16 INFO mapreduce.JobSubmitter: Submitting tokens for job: job_1518564517800_0005
18/02/14 05:19:16 INFO impl.YarnClientImpl: Submitted application application_1518564517800_0005
18/02/14 05:19:17 INFO mapreduce.Job: The url to track the job: http://localhost:8088/proxy/application_1518564517800_0005
/
18/02/14 05:19:17 INFO mapreduce.Job: Running job: job_1518564517800_0005
18/02/14 05:19:31 INFO mapreduce.Job: Job job_1518564517800_0005 running in uber mode : false
18/02/14 05:19:31 INFO mapreduce.Job: map 0% reduce 0%
18/02/14 05:19:43 INFO mapreduce.Job: map 100% reduce 0%
18/02/14 05:19:59 INFO mapreduce.Job: map 100% reduce 100%
18/02/14 05:20:00 INFO mapreduce.Job: Job job_1518564517800_0005 completed successfully
18/02/14 05:20:01 INFO mapreduce.Job: Counters: 49
File System Counters
FILE: Number of bytes read=6
FILE: Number of bytes written=404259
FILE: Number of read operations=0
FILE: Number of large read operations=0
FILE: Number of write operations=0
HDFS: Number of bytes read=777
HDFS: Number of bytes written=0
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)=9777
Applications Places System
Wed Feb 14, 5:20 AM Acadgild
acadgild@localhost:~
[Welcome to CentOS - ...


acadgild@localhost:~
File Edit View Search Terminal Help
[acadgild@localhost ~]$ hadoop fs -cat part-r-00000
18/02/14 06:12:49 WARN util.NativeCodeLoader: Unable to load native-hadoop library for your platform... using builtin-java
classes where applicable
Uttar Pradesh 18
Uttar Pradesh 18
Uttar Pradesh 14
Kerala 16
[acadgild@localhost ~]$
Applications Places System
Wed Feb 14, 6:12 AM Acadgild
acadgild@localhost:~