**Task 1:**

Execute **WordMedian** , **WordMean** , **WordStandardDeviation** programs using

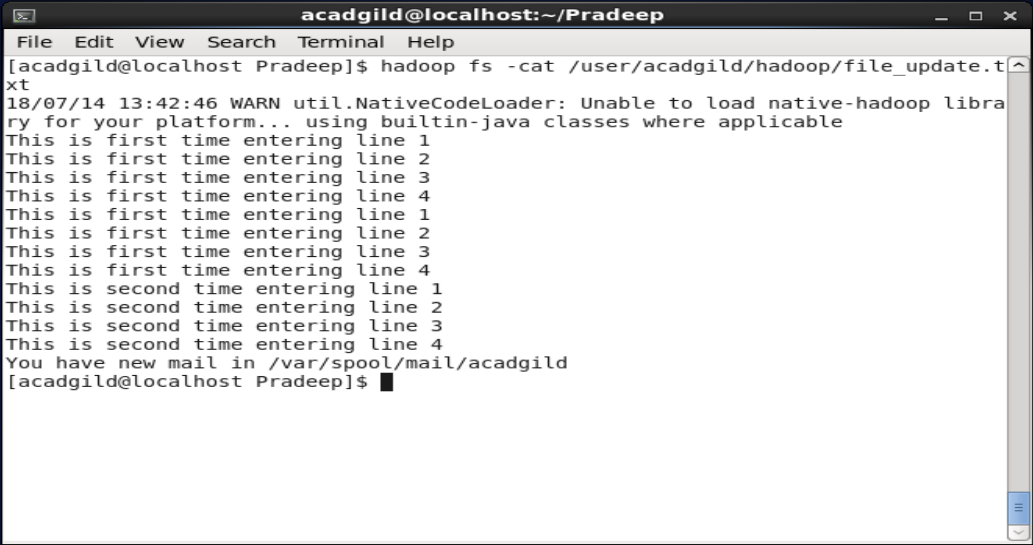
hadoop-mapreduce-examples-2.9.0.jar file present in your AcadGild VM.

Refer path below.

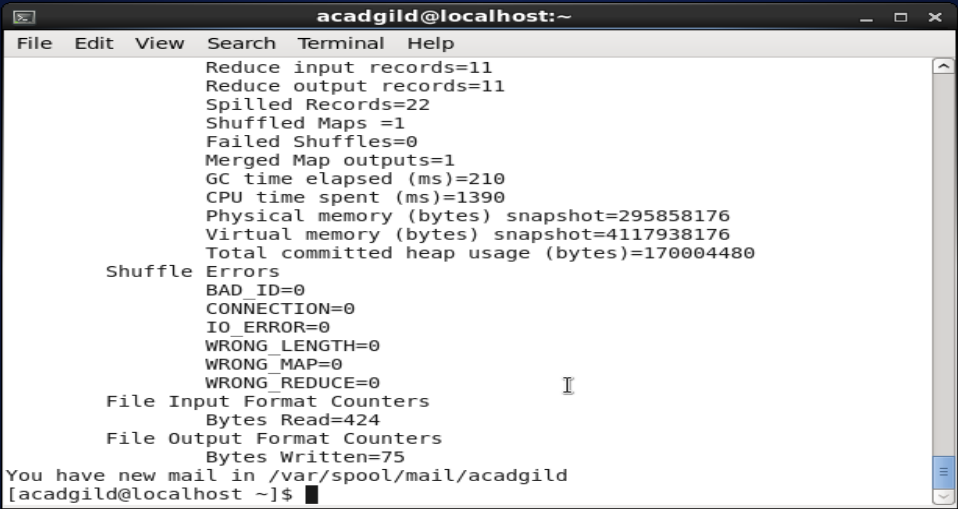
/home/acadgild/install/hadoop/hadoop-**2.9.0**/share/hadoop/mapreduce

**OUTPUT:**

1. I will use the following hadoop file system file file\_update.txt for execute the jar file:



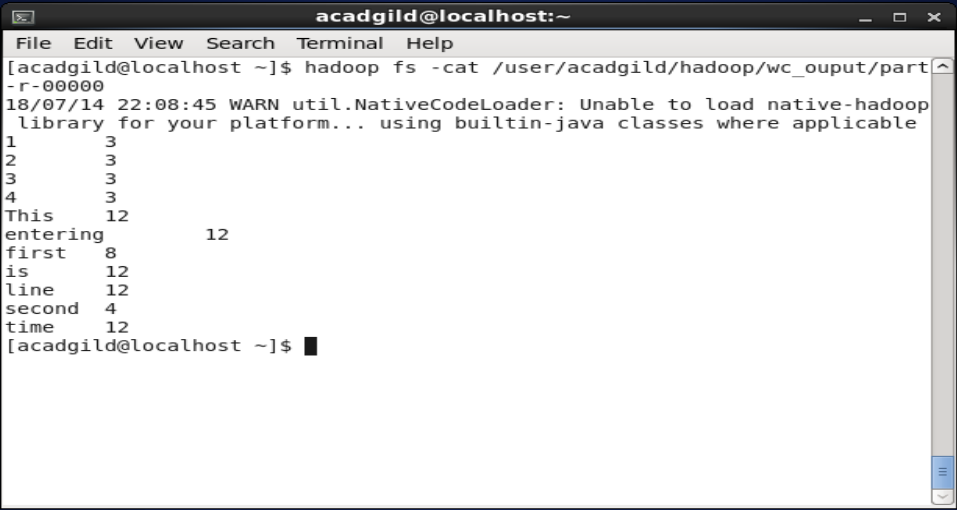
1. Then I have executed the following command:
   1. Hadoop jar /home/acadgild/install/hadoop/hadoop-2.6.5/share/hadoop/mapreduce/ hadoop-mapreduce-examples-2.6.5.jar wordcount /user/acadgild/hadoop/file\_update.txt /user/acadgild/hadoop/wc\_output



1. Then I have executed the following command to check whether output directory created or not:
   1. Hadoop fs –ls /user/acadgild/hadoop/wc\_output/



1. Then I have verified the output by executing the following command and by manually counting the file\_update.txt word count which matched same.
   1. Hadoop fs –cat /user/acadgild/hadoop/wc\_output/part-r-00000



1. To execute the remaining classes I **don’t have** **hadoop-mapreduce-examples-2.9.0.jar** file in **AcadGild VM**. But I understood the problem statement and commands to achieve the given classes like WordMedian, WordMean, WordStandardDeviation and store the output in respective output folders like WordMedianOutput, WordMeanOutput, WordStandardDeviationOutput

**Task2:**

Associated Data Files

<https://drive.google.com/file/d/0Bxr27gVaXO5sVjQ5QW0wQ3RCTUU/view?usp=sharing>

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.

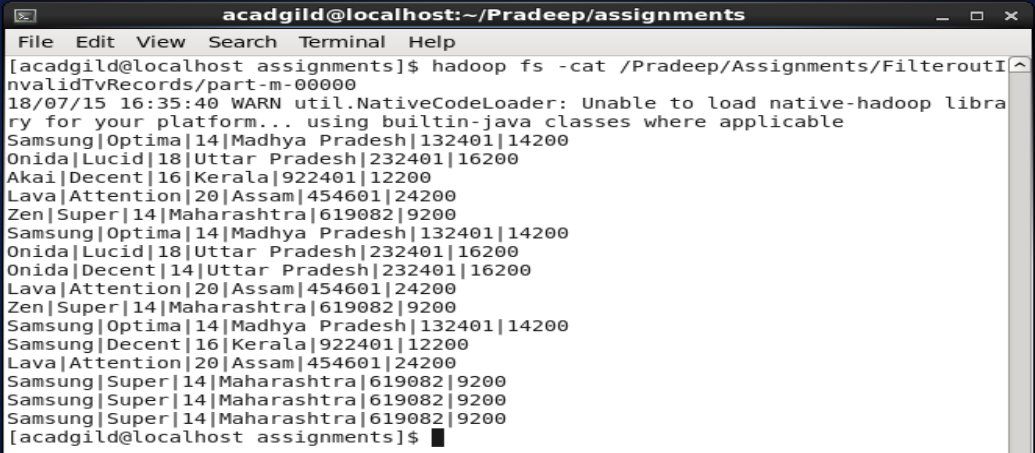
**Task 1.1**

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



Hadoop jar FilterOutNATvRecords.jar /Pradeep/Assignments/tv\_sales\_details.txt /Pradeep/Assignments/FilteroutInvalidTvRecords/

**Output:**



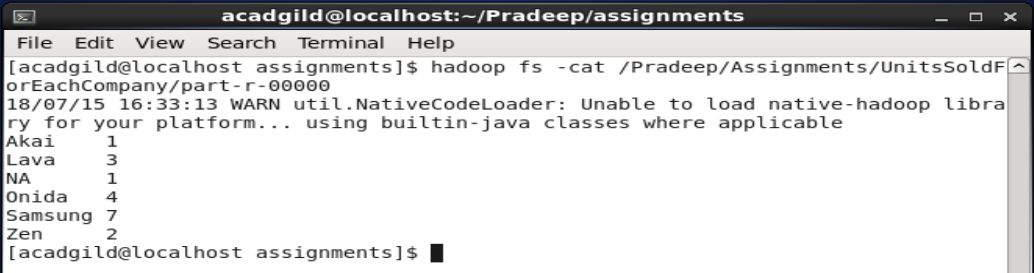
**Task 1.2**

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



Hadoop jar UnitsSoldForEachCompany.jar /Pradeep/Assignments/tv\_sales\_details.txt /Pradeep/Assignments/ UnitsSoldForEachCompany/

**Output:**

****

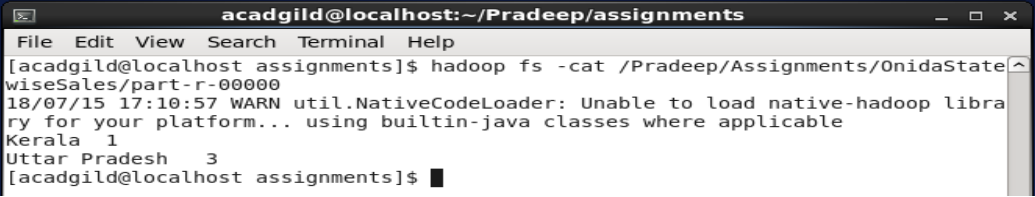
**Task 1.3**

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



Hadoop jar OnidaStatewiseSales.jar /Pradeep/Assignments/tv\_sales\_details.txt /Pradeep/Assignments/ OnidaStatewiseSales/

**Output:**

****