ASSIGNMENT 3

YARN

PROBLEM STATEMENT –

- Execute WordMedian , WordMean , WordStandardDeviation programs usinghadoop-mapreduce-examples-2.9.0.jar file present in your AcadGild VM.

Input file – test.txt

```
[acadgild@localhost ~]$ hadoop fs -cat /test.txt
18/09/22 11:58:56 WARN util.NativeCodeLoader: Unable to load native-hadoop
welcome to acadgild
```

1. **Wordmean**

Command to run wordmean program-

hadoop jar /home/acadgild/install/hadoop/hadoop-2.6.5/share/hadoop/mapreduce/hadoop-mapreduce-examples-2.6.5.jar wordmean /test.txt /meanoutput

Input file – test.txt

Output directory – meanoutput

Output –

```
The mean is: 5.666666666666667
You have new mail in /var/spool/mail/acadgild
[acadgild@localhost ~]$ hadoop fs -cat /meanoutput/part-r-00000
18/09/22 12:19:10 WARN util.NativeCodeLoader: Unable to load native-hadoop
count   3
length  17
```

Explanation-

Input is "welcome to acadgild", number of words is 3 and number of letters is 17.
Mean = Number of letters/Number of words, i.e, 17/3 = 5.6666

2. WordMedian

Command to run wordmean program-

hadoop jar /home/acadgild/install/hadoop/hadoop-
2.6.5/share/hadoop/mapreduce/hadoop-mapreduce-examples-2.6.5.jar wordmedian
/test.txt /medianoutput

Input file – test.txt

Output directory – medianoutput

Output –

```
The median is: 7
You have new mail in /var/spool/mail/acadgild
[acadgild@localhost ~]$ hadoop fs -cat /medianoutput/part-r-00000
18/09/22 12:12:16 WARN util.NativeCodeLoader: Unable to load native-hadoop
2       1
7       1
8       1
```

Explanation-

Input is "welcome to acadgild", after counting number of letters and sorting,

Word	Number of letters	Count
to	2	1
welcome	7	1
acadgild	8	1

Median = The midpoint or middle value of a set of numbers, i.e, 7
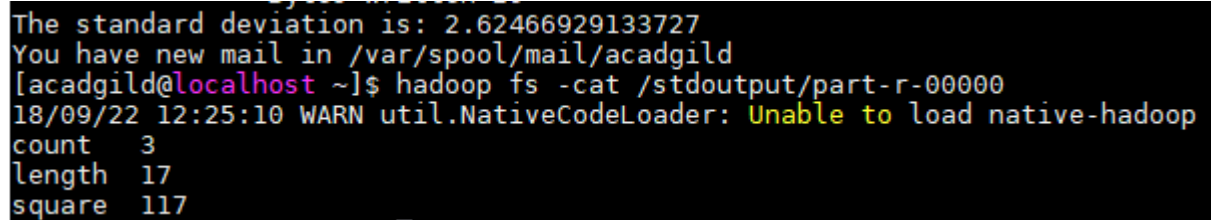
3. WordStandardDeviation

Command to run wordmean program-

hadoop jar /home/acadgild/install/hadoop/hadoop-
2.6.5/share/hadoop/mapreduce/hadoop-mapreduce-examples-2.6.5.jar
wordstandarddeviation /test.txt /stdoutput

Input file - test.txt

Output directory - stdoutput

Output -

```
The standard deviation is: 2.62466929133727
You have new mail in /var/spool/mail/acadgild
[acadgild@localhost ~]$ hadoop fs -cat /stdoutput/part-r-00000
18/09/22 12:25:10 WARN util.NativeCodeLoader: Unable to load native-hadoop
count   3
length  17
square  117
```

Explanation-

Standard deviation = 2.6246