

**BIG DATA DEVELOPMENT**



**Batch - June 16th 2018**

**Session One**

**Assignment One - Solution**

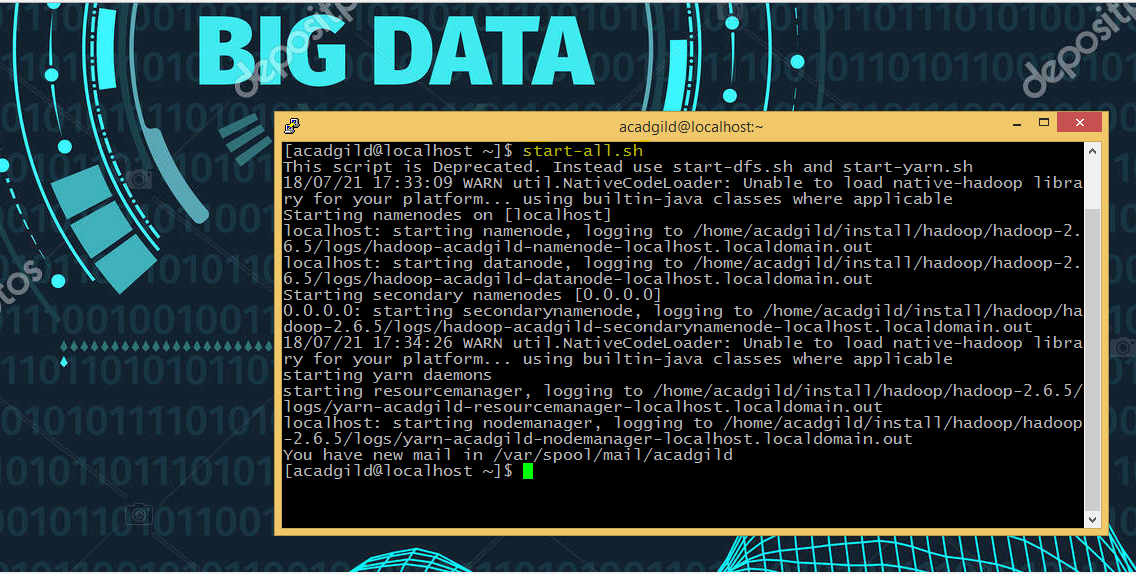
**Name: Venky**

**Task 1**

**1. Start Hadoop single node on AcadGild VM (The command is start-all.sh)**

command: **start-all.sh**

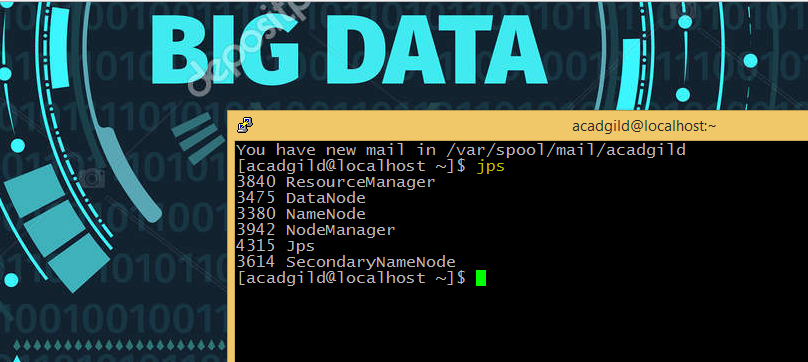
This command starts all the hadoop services like **DataNode**, **NameNode**, **NodeManager** etc

****

**2. Run a JPS command to see if all the Hadoop daemons are running**

command: **jps**

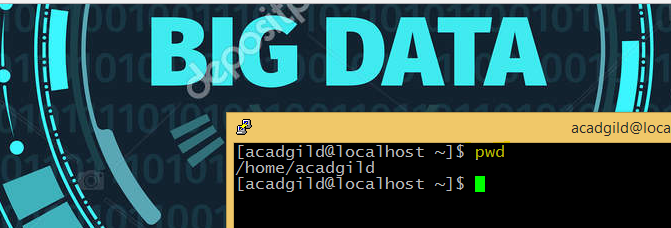
This command lists out all the active services

****

**3. Run a couple of Unix commands like pwd, ls -ls etc.**

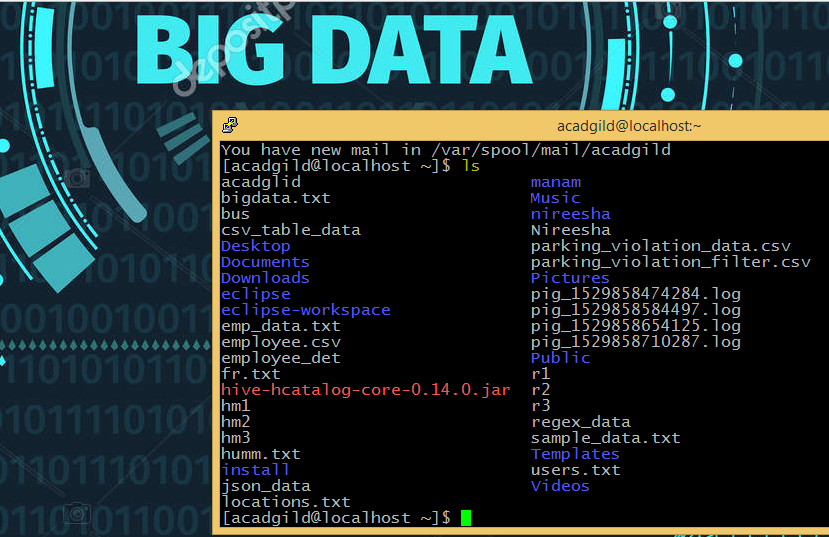
Command: **pwd**

This command displays the path of present working directory.

****

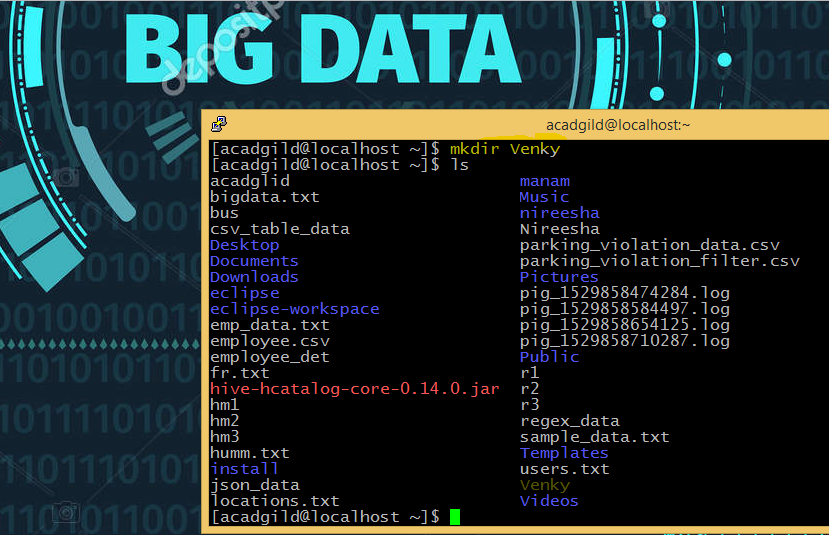
Command: **ls**

This command lists out all the files and directories present in the current directory or the directory of whose path has been given



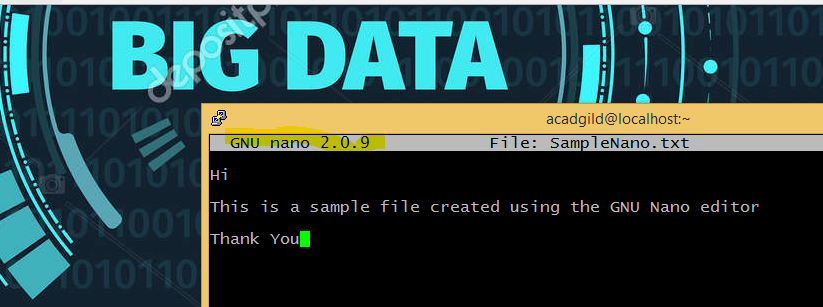
Command: **mkdir**

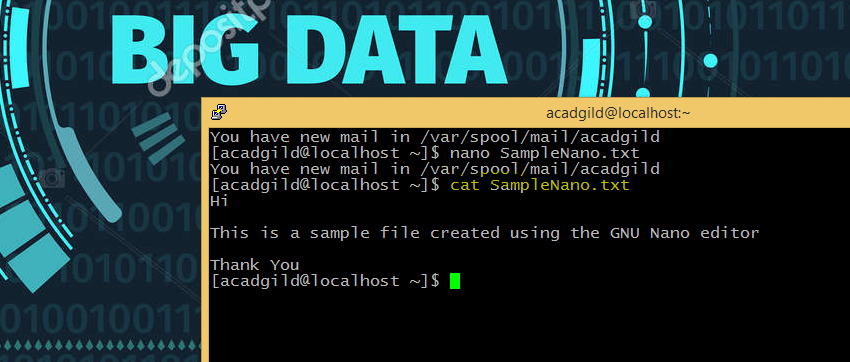
This command is user to create a new directory



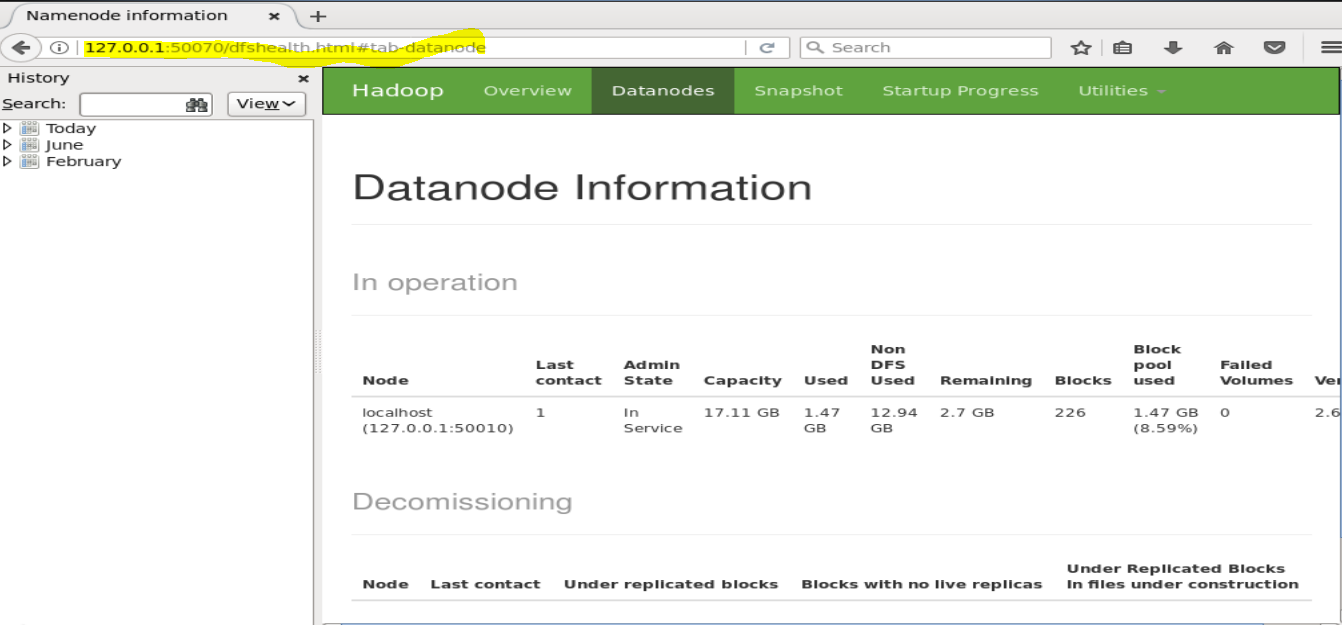
**4**. **Create a file from the terminal using nano editor (for example nano test.txt) and put some content in it. Cat it to see if the content is saved.**

Command: **nano <filename>**





**5. Open the hdfs web page from the browser by typing localhost:50070 in the browser. Check all the details of the HDFS.**



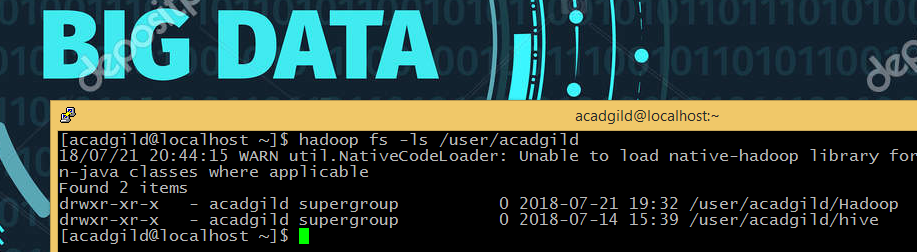
**Task 2**

**1. Check whether /user/acadgild directory exists or not in the HDFS.**

**If it doesn't exist, then create this.**

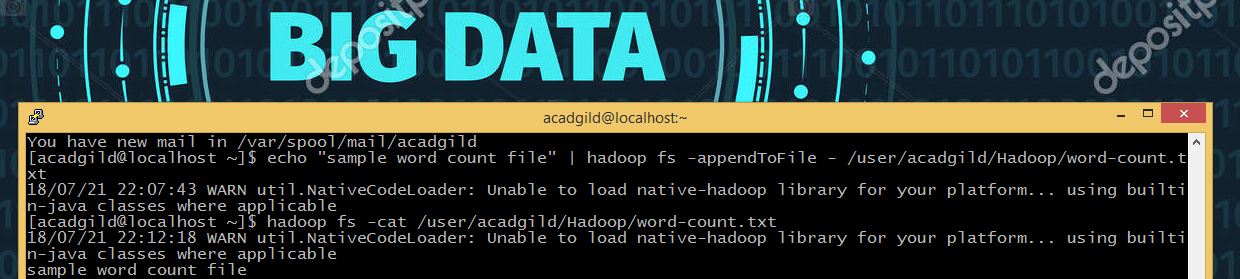
**Create a directory /user/acadgild/hadoop.**

Command: **hadoop fs -mkdir /user/acadgild/Hadoop**

****

**2. Create a file in HDFS under directory /user/acadgild/hadoop, with name word-count.txt.**

**Whatever we type on screen should get appended to the file. Try to type (on screen) few lines from any online article or textbook.**

****

**3. Create a file max-temp.txt in local FS. Put some 10-15 records of date and temperature**

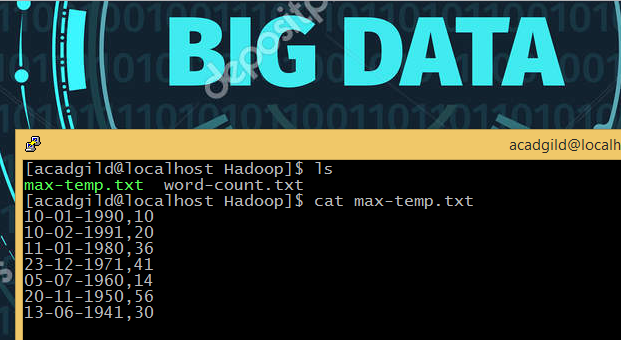
**example: dd-mm-yyyy,temperature**

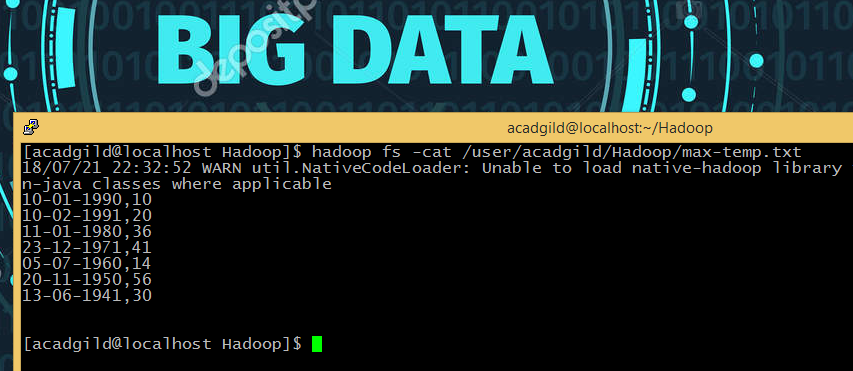
**Example:**

**10-01-1990,10**

**10-02-1991,20**

**Move this file to HDFS at /user/acadgild/hadoop.**

****

****

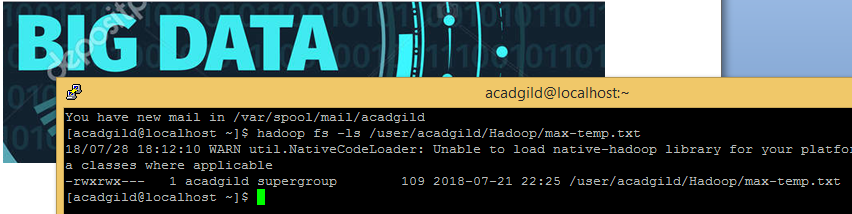
**4. Change the permission of the file /user/acadgild/hadoop/max-temp.txt, such that only the**

**owner and the group members have full control over the file.**

**Others do not have any control over it.**

Command: **chmod**

This command is user to add or remove permissions

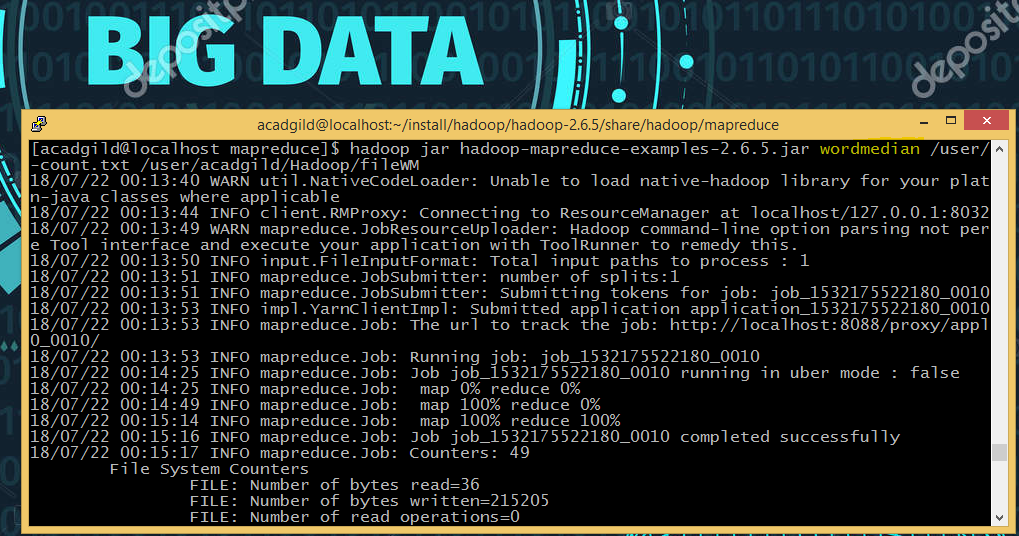
****

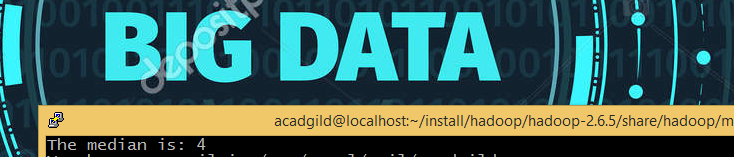
**Task 3:**

**1. Execute WordMedian​, WordMean,​ WordStandardDeviation​ programs using**

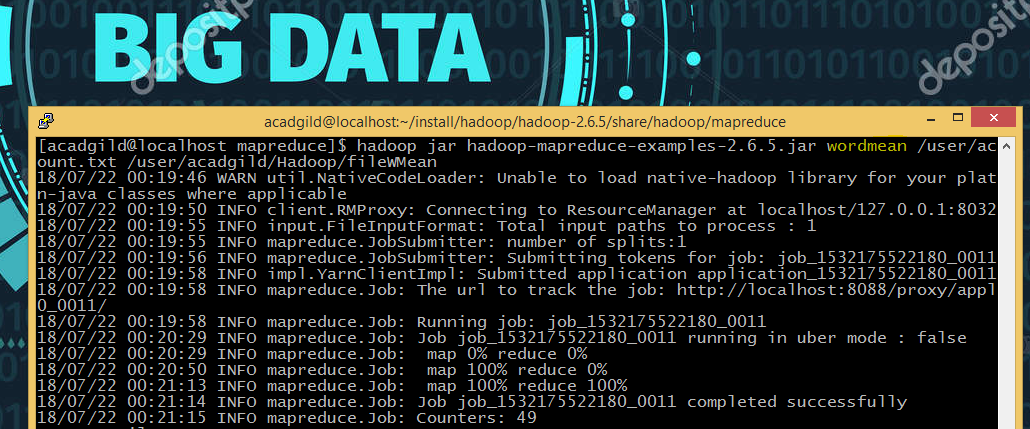
**hadoop-mapreduce-examples-2.9.0.jar file present in your AcadGild VM.**

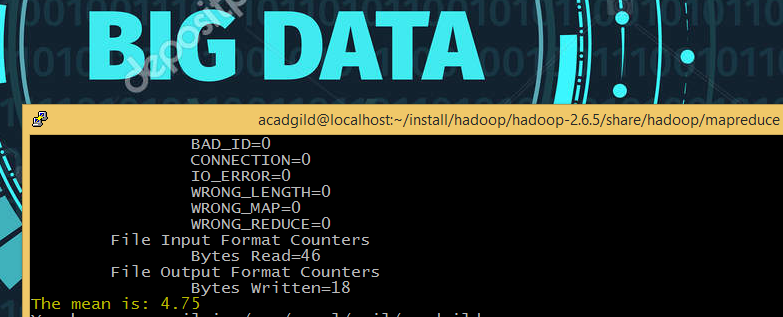
**WordMedian**

****

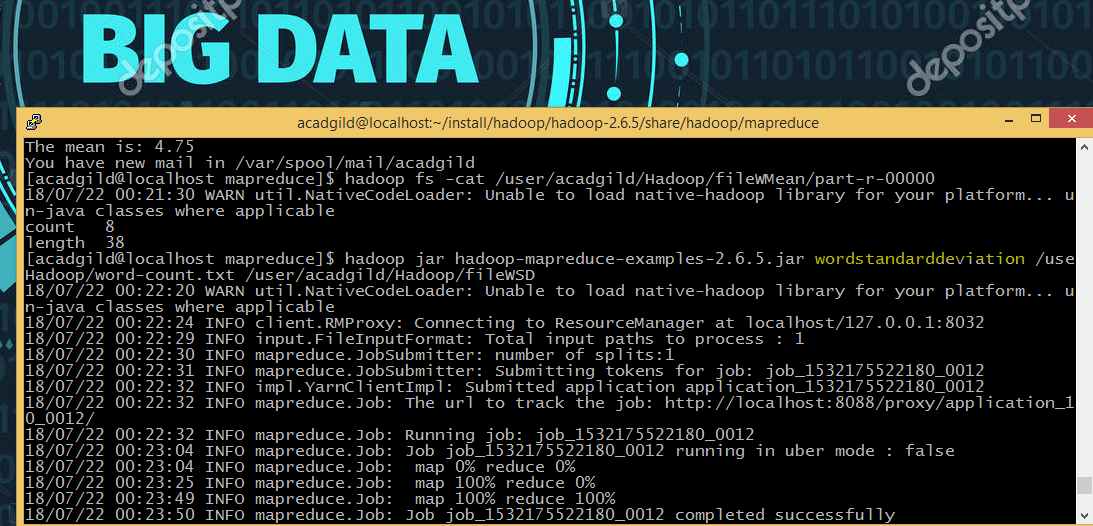
****

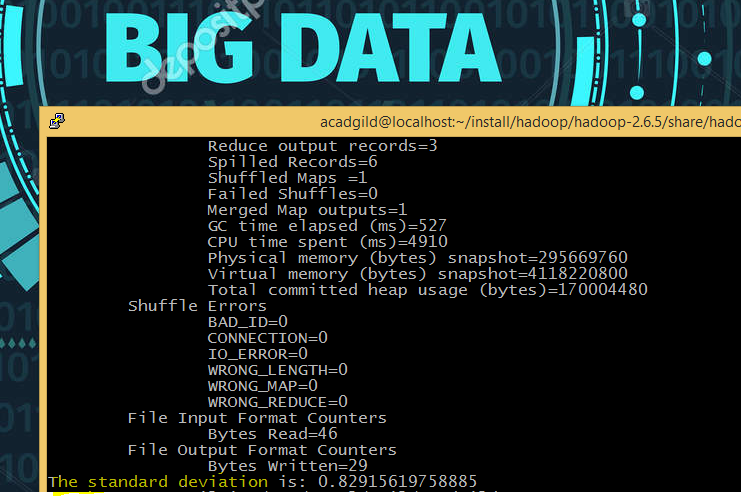
**WordMean**

****

****

**WordStandardDeviation**

****

****