**BIG DATA ENGINEERING FOR HADOOP & SPARK TRAINING**

**ASSIGNMENT 2**

**Task 1 Solution:**

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

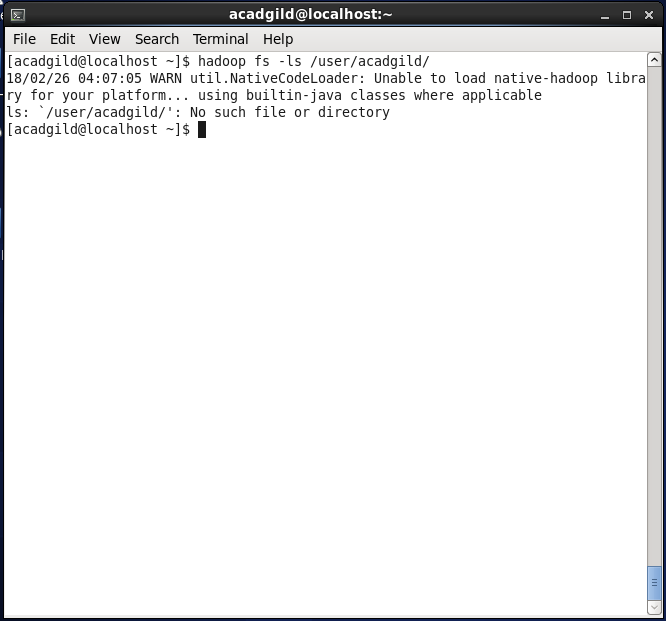
**If it doesn't exist, then create this.**

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

Sol : hadoop fs -ls /user/acadgild/hadoop this command will list the files of the given directory.

If file or the directory doesn’t exists then we get an error “No such file or directory”.

Here we got the error so we have to create the whole directory.

****

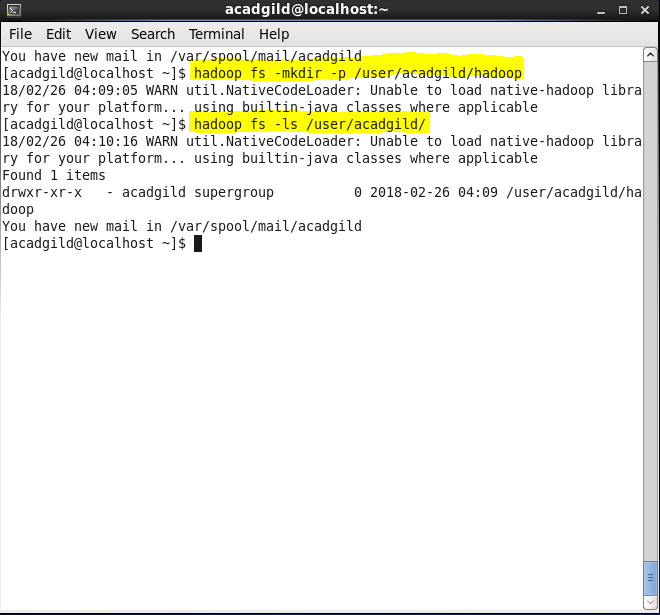
**“**hadoop fs -mkdir –p /user/acadgild/hadoop” command is used to create.

“/user/acadgild/hadoop” directory recursively in hadoop file system as we have used

mkdir with option –p.

Now we can check the directory path created or not by using the command

“hadoop fs -ls /user/acadgild”.

****

**Task 2 Solution:**

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

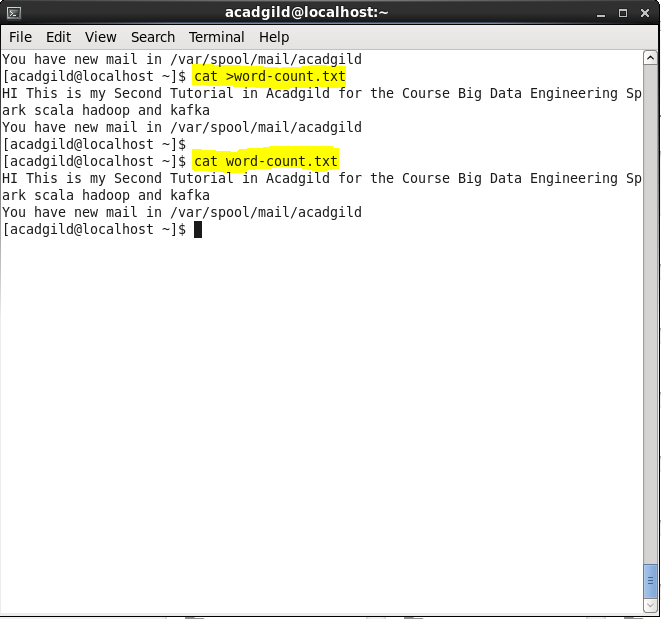
**Whatever we type on screen should get appended to the file.**

**Try to type (on screen) few lines from any online article or textbook.**

**Sol:** Create a file in the local file system and add some data to it using “cat >word-count.txt”

Then check the file content using “cat word-count.txt” this will display the content of the

File if the file exists.



Now copy the file “word-count.txt” to hadoop file system using the command.

**“hadoop fs -put /home/user/word-count.txt /user/acadgild/hadoop/word-count.txt”.**

This command will copy the local file to hadoop file system.

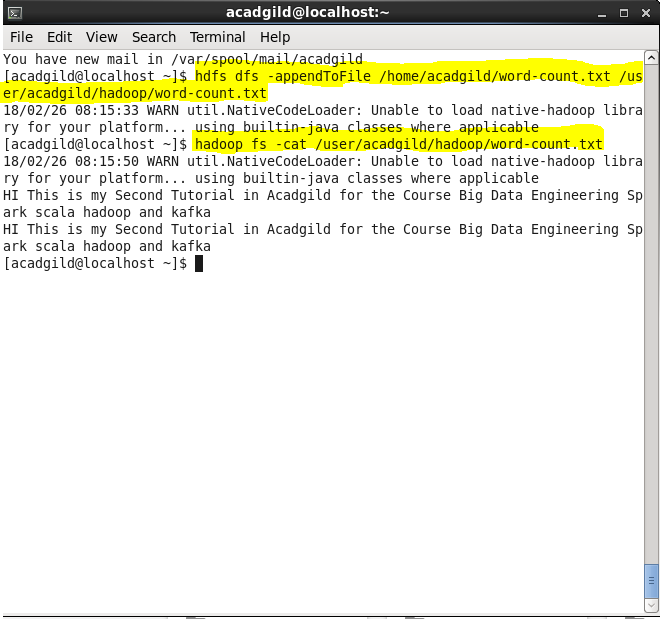
To Append the file with the existing file which is available in the local directory use the

Following command

**“hdfs dfs –appendToFile /home/acadgild/word-count.txt /user/acadgild/hadoop/**

**word-count.txt”**.

Now check the file in hdfs is updated with the contents of local file .



**Task 3 Solution :**

**Q) Create a file max-temp.txt in local FS.**

**Put some 10-15 records of date and temperature example:**

**dd-mm-yyyy,temperature**

**Example:**

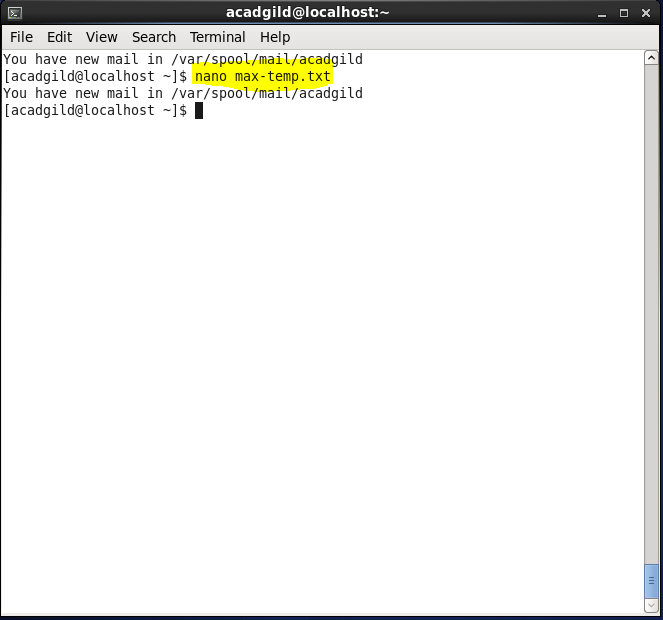
**10-01-1990,10**

**10-02-1991,20**

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

**Sol )** Use nano editor to create a file and update the file with data specified in the above format

**“nano max-temp.txt”** command will create a file in the present working directory.



Update the file with the data of 15 records in the given format and save the file.

Now copy “max-temp.txt” file to hdfs file system using the command

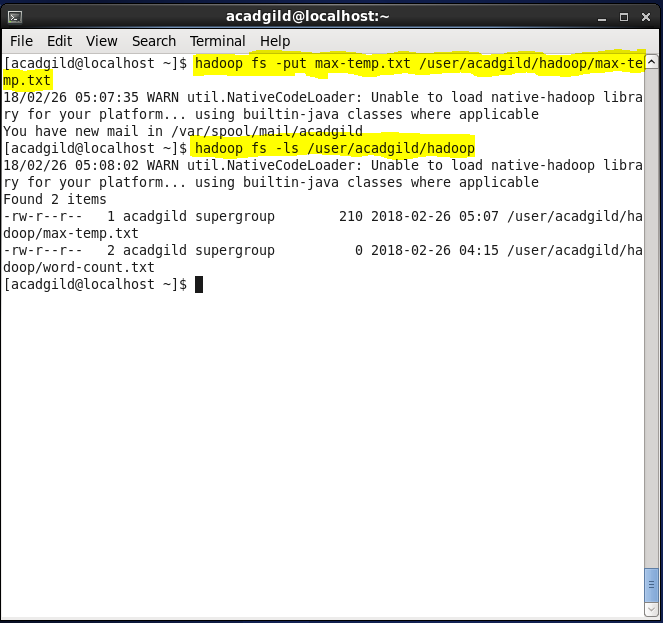
**“hadoop fs -put max-temp.txt /user/acadgild/hadoop/max-temp.txt”**

The above command will copy the max-temp.txt file to hdfs directory

“/user/acadgild/hadoop/”

Now to confirm whether the file is copied or not list the files in hdfs directory

Using the command **“hadoop fs -ls /user/acadgild/hadoop”**



**Task 4 Solution :**

**Q)** **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.**

Sol) we have chmod command to change the read write and execute permissions for the

User,group user and others.

Ex : chmod u=rwx,g=rw,o=r <filename>

The letters **u**, **g**, and **o** stand for "**user**", "**group**", and "**other**". The equals sign ("**=**") means "set the permissions exactly like this," and the letters "**r**", "**w**", and "**x**" stand for "read", "write", and "execute", respectively. The commas separate the different classes of permissions, and there are no spaces in between them.

This is equivalent to “chmod 754 <filename>

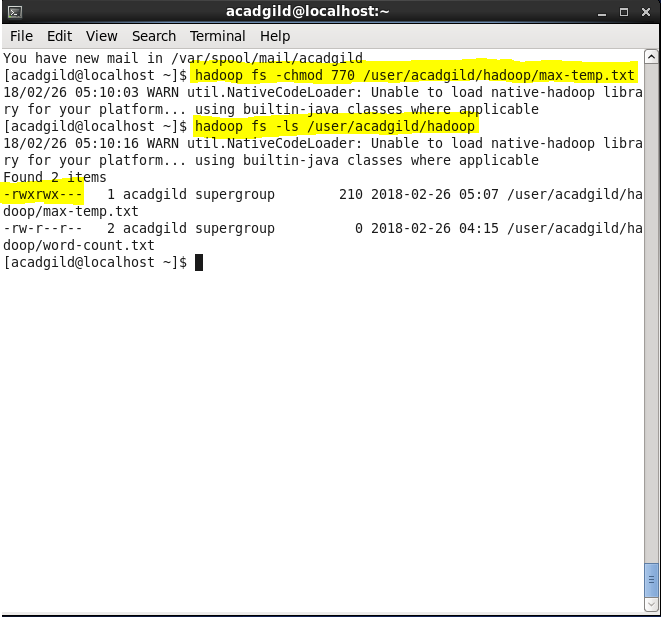
Here the digits **7**, **5**, and **4** each individually represent the permissions for the user, group, and others, in that order. Each digit is a combination of the numbers **4**, **2**, **1**, and **0**:

Now here we have to provide full permissions for user and group and no permissions to

Others.

We have to use the command “hadoop fs –chmod 770 /user/acadgild/hadoop/

max-temp.txt”.

****