**HDFS Assignments**

1. Create a text file **userdata.txt** in your home directory having following lines using a text editor:

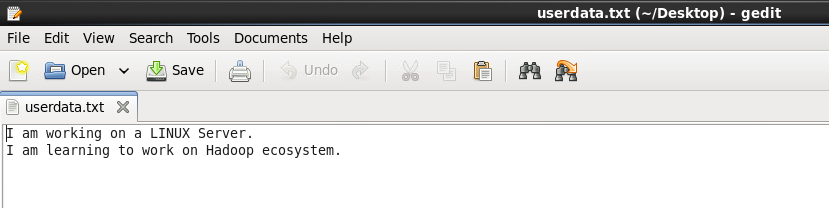
I am working on a Linux Server.

I am learning to work on Hadoop ecosystem.

**hadoop fs –mkdir /COE/userdata.txt**

**echo “I am working on a Linux Server.” /COE/userdata.txt**

**echo** **“I am learning to work on Hadoop ecosystem.” /COE/userdata.txt**



1. Create a folder called **training** under your home directory in HDFS

Write the command used here:

**hdfs dfs –mkdir /COE/ training**



1. Create a folder called **SampleData** under **training**

Write the command used here:

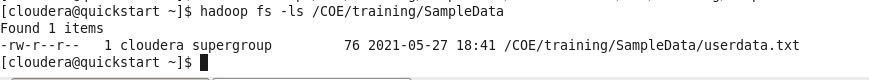
**hdfs dfs –mkdir /COE/ training/SampleData**



1. Copy file **userdata.txt** to **SampleData** folder in HDFS

Write the command used here:

**hadoop fs –copyFromLocal ~/userdata.txt /COE/training/SampleData**



1. Display the content of **userdata.txt** in **hdfs** using **cat** command

Write the command used here:

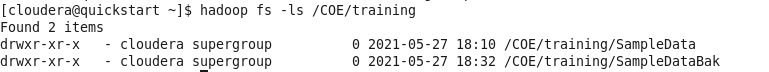
**hadoop fs –cat /COE/training/SampleData/userdata.txt**



1. Create another directory called **SampleDataBak** under training directory in HDFS

Write the command used here:

**hdfs dfs –mkdir /COE/ training/SampleDataBak**



1. Copy the file **userdata.txt** from **SampleData** folder to **SampleDataBak**:

Write the command used here:

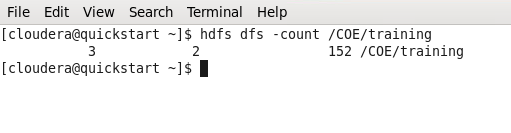
**hdfs dfs –cp /COE/ training /SampleData/userdata.txt /COE/training/SampleDataBak**



1. List all sub-folders and files under **training** folder in **HDFS**:

Write the command used here:

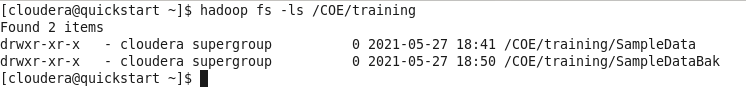
**hdfsdfs –count /COE/training**



**hadoop fs -ls /COE/training**

**hadoop fs -ls /COE/ training /SampleData**

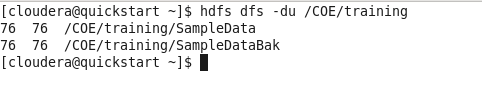
**hadoop fs -ls /COE/training/SampleDataBak**



1. Display the total disk space used by **training** directory

Write the command used here:

**hdfs dfs -du /COE/training**



1. Display the disk space used by each folder under **training** directory

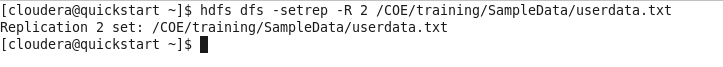
Write the command used here:

**hdfs dfs -du /COE/training**

1. Change the Replication factor of **training/SampleData/userdata.txt** to 2

Write the command used here:

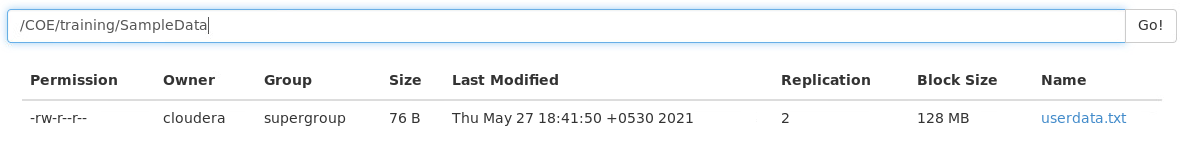
**hdfs dfs -setrep -R 2 /COE/training/SampleData/userdata.txt**

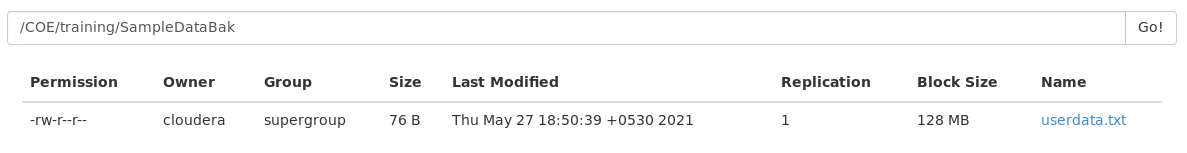


1. Browse the **training** folder in HDFS using Web UI (localhost:50070) and note down the following:

**File / Folder Name Size Replication Block Size**

training





1. Delete training folder along with all sub folders and files:

Write the command used here:

**hadoop -rm -r /COE/training**

