

ASSIGNMENT 1

BUAN 6346 BIG DATA  
DR. JERRY PEREZ

Setu Shah SXS210181



## **TABLE OF CONTENTS**

<b>Description.....</b>	<b>2</b>
<b>Chapter 02 Setup.....</b>	<b>3</b>
Getting Started.....	3
Working With Virtual Machine.....	4
Points to Note During Assignment.....	5
<b>Chapter - 03 - Access HDFS.....</b>	<b>7</b>
Explore the HDFS Command Line Interface.....	7
Uploading Files to HDFS.....	8
View HDFS files.....	25
The Hue File Browser is used browse, view and manage files.....	28
<b>Chapter - 04 - YARN Job.....</b>	<b>35</b>
Explore the YARN cluster.....	35
Submitting an application to the YARN cluster.....	37
View the application in the Hue Job Browser.....	38
View the application in the YARN UI.....	39
<b>Chapter - 05 – SQOOP.....</b>	<b>41</b>
Import the accounts table from MySQL.....	41
View the imported data.....	47
Import incremental updates to accounts.....	49
Import webpage data using an alternate field delimiter.....	51

## **DESCRIPTION**

Apache Hadoop The Apache™ Hadoop® project develops open-source software for reliable, scalable, distributed computing.

A generalized data-flow programming framework, built on Hadoop YARN, which provides a powerful and flexible engine to execute an arbitrary DAG of tasks to process data for both batch and interactive use-cases.

Hadoop Distributed File System (HDFS) – a distributed file-system that stores data on commodity machines, providing very high aggregate bandwidth across the cluster.

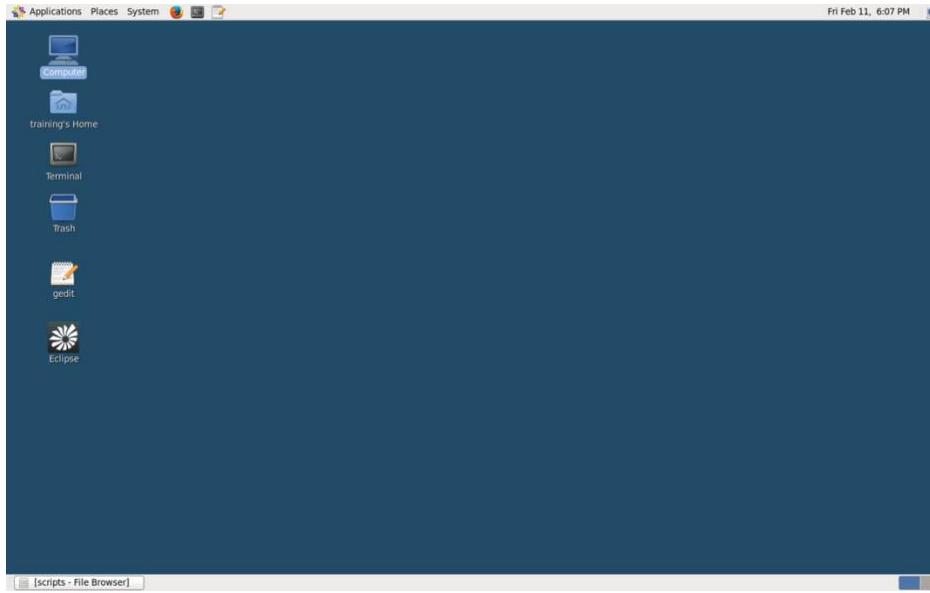
Hadoop YARN – (introduced in 2012) a platform responsible for managing computing resources in clusters and using them for scheduling users' applications.

Hadoop MapReduce – an implementation of the MapReduce programming model for large-scale data processing.

Hue write SQL like a pro Hue brings the best Querying Experience with the most intelligent Autocompletes and Query Editor components. Help users find the correct data among thousands of databases and self-document it. Assist users with their SQL queries and leverage rich preview for links, sharing from Editor etc.

## Chapter - 02 – Setup

The objective of this chapter is to get familiar with the virtual machine and simultaneously setup the virtual machine.

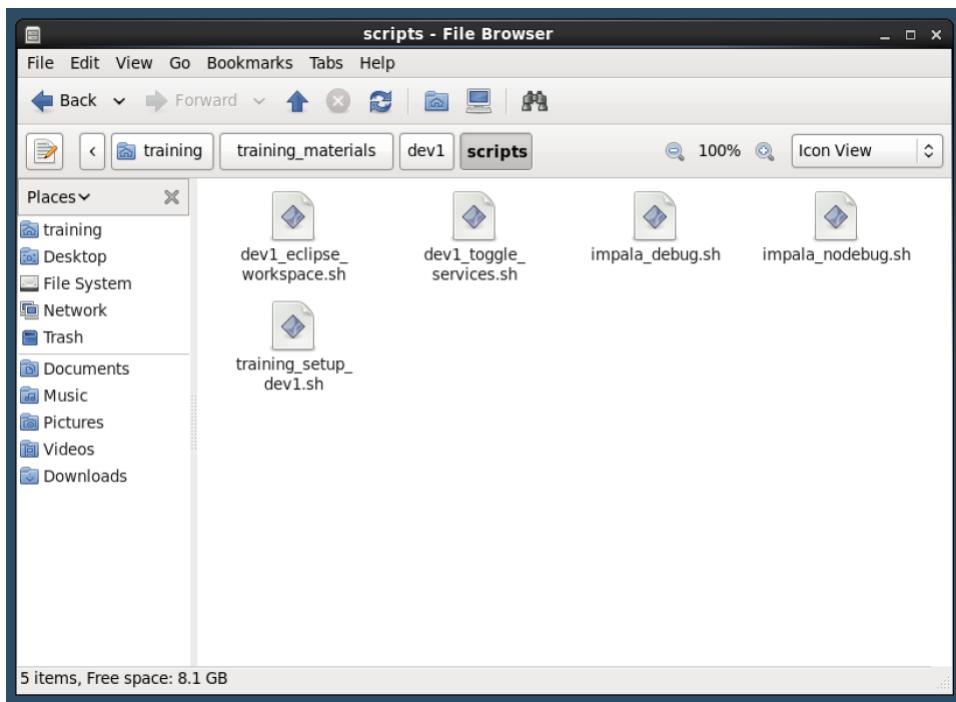


### Working with the Virtual Machine

#### Getting started

```
File Edit View Search Terminal Help
starting Spark history-server (spark-history-server): [ OK ]
JMX enabled by default
Using config: /etc/zookeeper/conf/zoo.cfg
Starting zookeeper ... STARTED
Stopping Apache Archiva...
Apache Archiva was not running.
Stopped Hive Server2: [ OK ]
* Disabling services not required for Dev1 Training
no secondarynamenode to stop
Stopped Hadoop secondarynamenode: [ OK ]
Flume agent is not running [ OK ]
Stopping Hadoop HBase regionserver daemon: no regionserver to stop because no pid file /var/run/hbase/hbase-hbase-regionserver.pid
hbase-regionserver.
no master to stop because no pid file /var/run/hbase/hbase-hbase-master.pid
Stopped HBase master daemon: [ OK ]
no rest to stop because no pid file /var/run/hbase/hbase-hbase-rest.pid
Stopped HBase rest daemon: [ OK ]
no thrift to stop because no pid file /var/run/hbase/hbase-hbase-thrift.pid
Stopped HBase thrift daemon: [ OK ]
Stopped Kafka Server: [ OK ]
Solr server daemon is not running [ OK ]
[Dev] Done setting up for Dev1 Training
[training@localhost ~]$ Setu Shah SXS210181
```

This is the course setup script to enable, and setup all services and data required for the course.

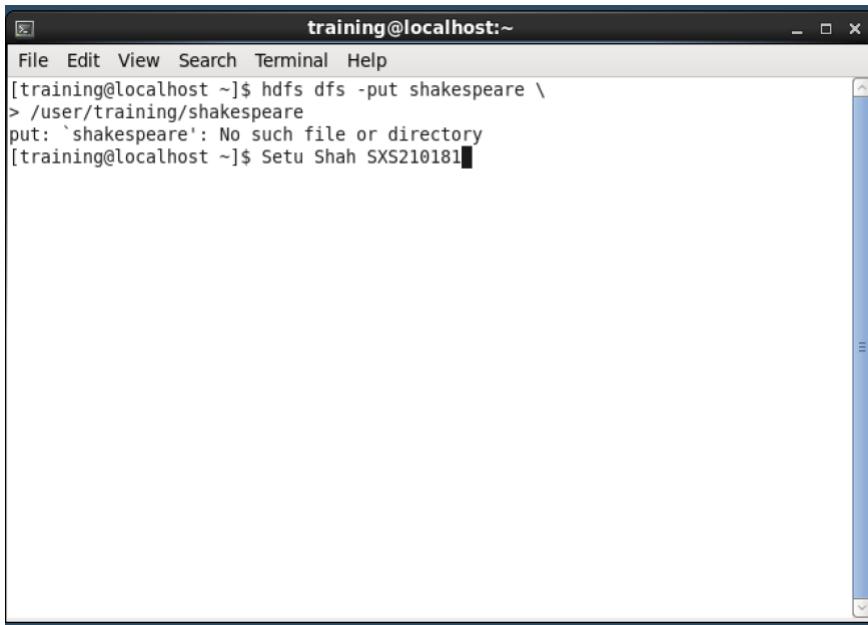


## Working with the Virtual Machine

Virtual Machine is up and running and it is by default logged in as user training. Username and Password are both training.



Below screenshot explains that the dollar sign (\$) toward the start of each line shows the Linux shell prompt. The actual prompt will incorporate extra information. The backslash (\) toward the finish of the mainline means that the order isn't finished and forges ahead with the following line. You can enter the code precisely as displayed (on two lines), or you can enter it on a single line. Assuming you do the last option, you ought not to type in the backslash. I executed the code, and got output "put: 'shakespeare': No such record or index found".

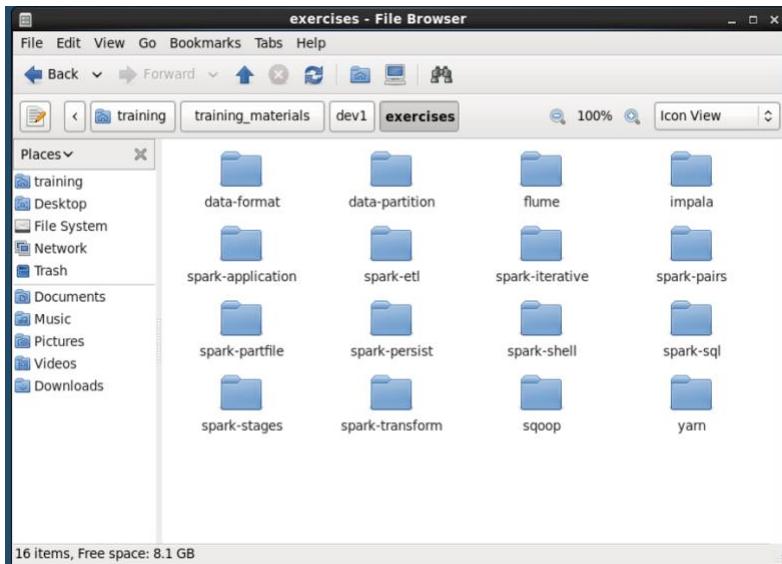


A screenshot of a terminal window titled "training@localhost:~". The window contains the following text:

```
[training@localhost ~]$ hdfs dfs -put shakespeare \
> /user/training/shakespeare
put: 'shakespeare': No such file or directory
[training@localhost ~]$ Setu Shah SXS210181
```

### Points to note during the assignment

Accessing the main directory “\$DEV1/exercises”.

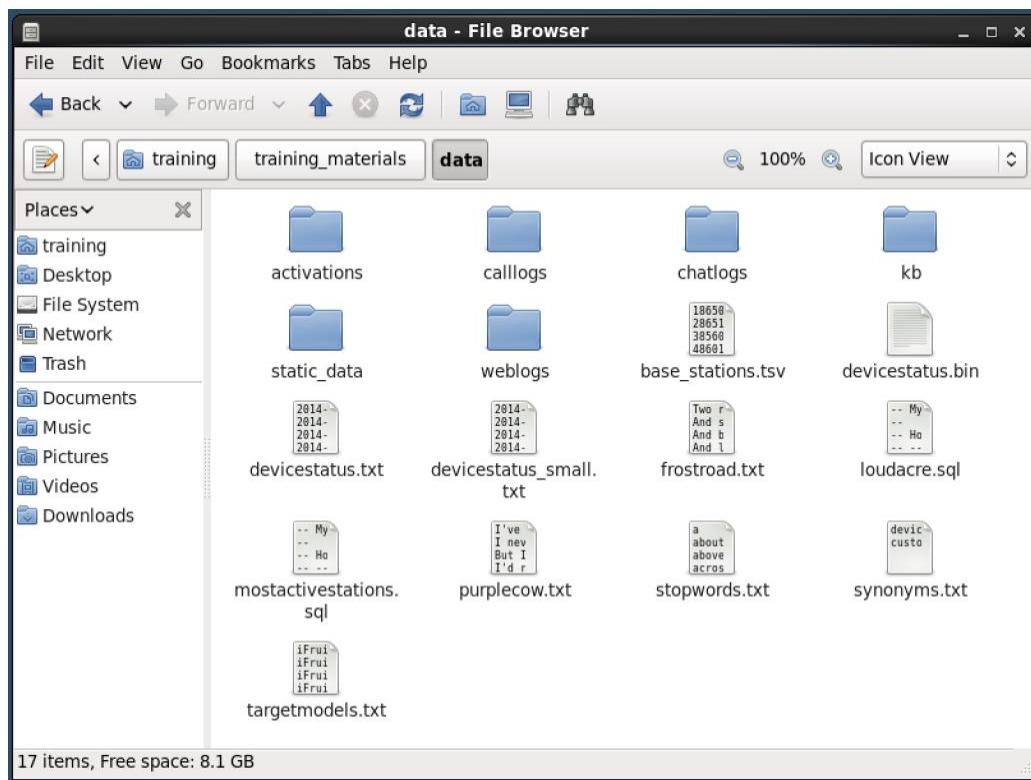


All directories under the above directory are related to the assignments. Within each exercise directory I found the following subdirectories.

<b>solution</b>	Solution code for each exercise
<b>stubs</b>	A few of the exercises depend on provided starter files containing skeleton code
<b>Maven project directories</b>	For exercises for which you must write Scala classes, you have been provided with preconfigured Maven project directories.

---

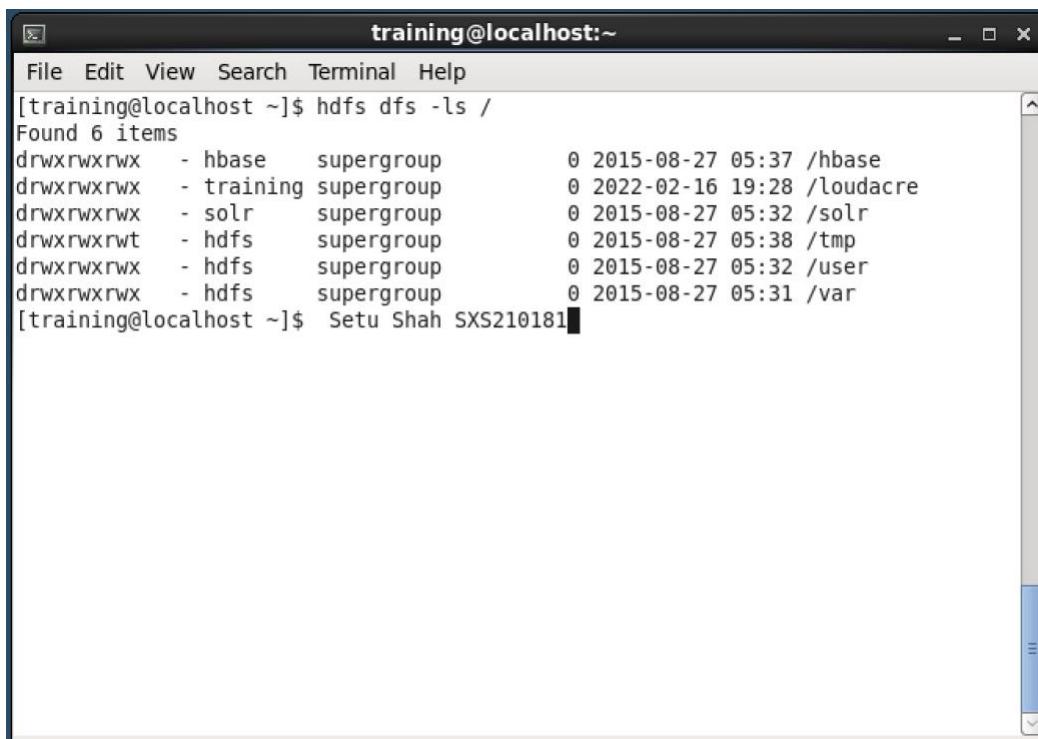
Files used in the exercises are located in “\$DEV1DATA (~/training\_materials/data)”.



## Chapter-3 Access HDFS

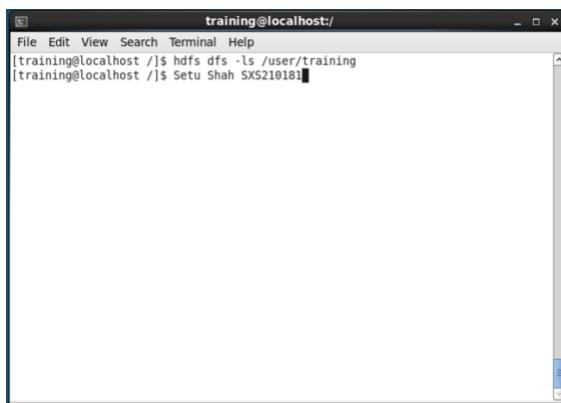
### Explore the HDFS Command line Interface

HDFS has been created, “hdfs dfs -ls /” is used to view the contents of root directory. We can see “loudacre\_Setu\_Shah” which was created when I was setting up the virtual machine.

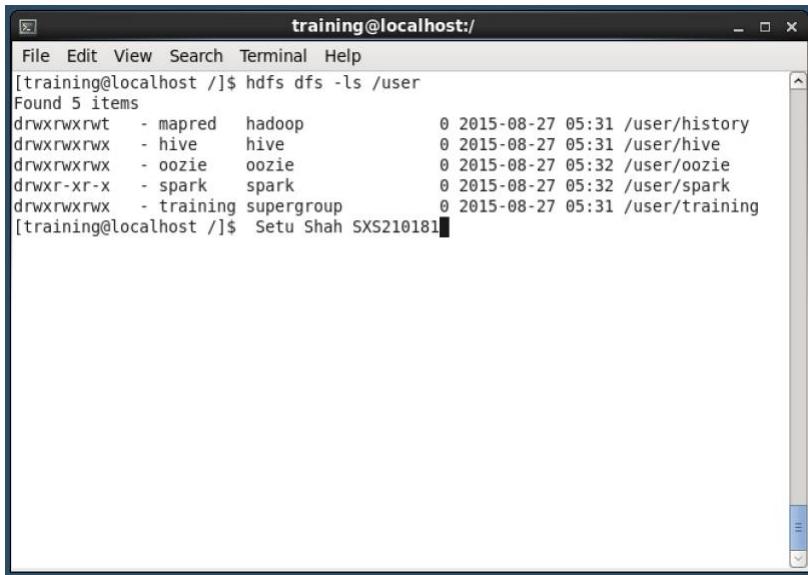


```
File Edit View Search Terminal Help
[training@localhost ~]$ hdfs dfs -ls /
Found 6 items
drwxrwxrwx  - hbase    supergroup      0 2015-08-27 05:37 /hbase
drwxrwxrwx  - training  supergroup      0 2022-02-16 19:28 /loudacre
drwxrwxrwx  - solr     supergroup      0 2015-08-27 05:32 /solr
drwxrwxrwt  - hdfs    supergroup      0 2015-08-27 05:38 /tmp
drwxrwxrwx  - hdfs    supergroup      0 2015-08-27 05:32 /user
drwxrwxrwx  - hdfs    supergroup      0 2015-08-27 05:31 /var
[training@localhost ~]$ Setu Shah SXS210181
```

Using “hdfs dfs -ls/user” to view contents of user directory. We can find home directory “training” under “/user”, which has been named after the username “training”.



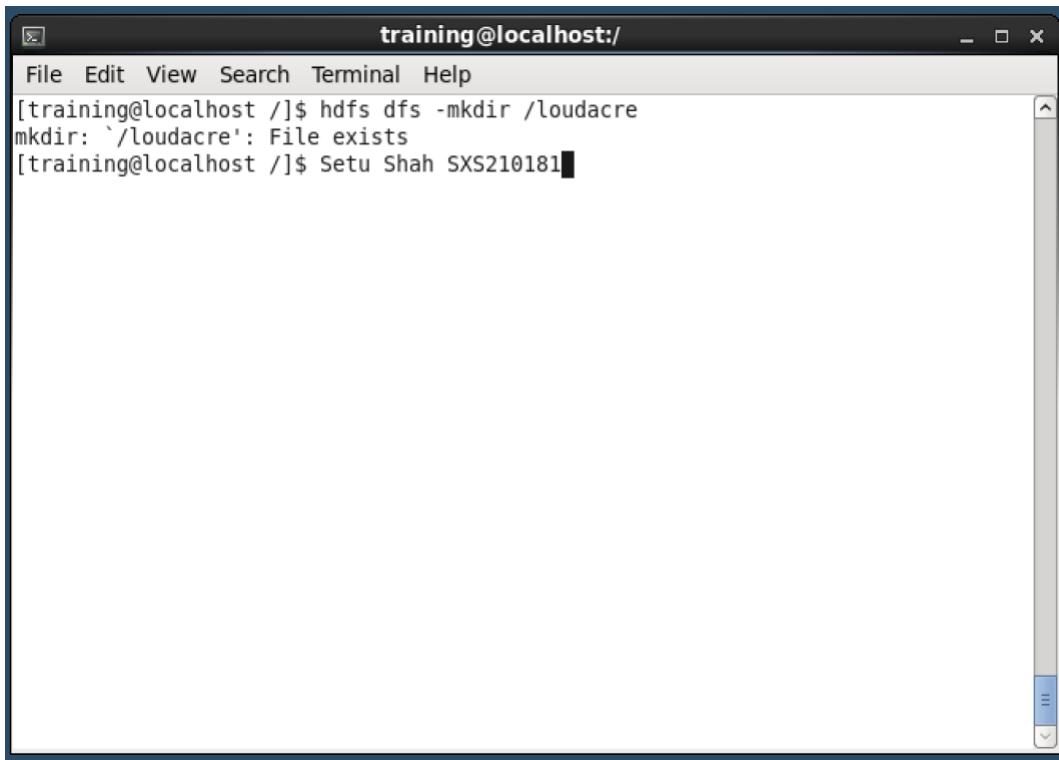
```
File Edit View Search Terminal Help
[training@localhost /]$ hdfs dfs -ls /user/training
[training@localhost /]$ Setu Shah SXS210181
```



```
training@localhost:~$ hdfs dfs -ls /user
Found 5 items
drwxrwxrwt  - mapred   hadoop          0 2015-08-27 05:31 /user/history
drwxrwxrwx  - hive      hive           0 2015-08-27 05:31 /user/hive
drwxrwxrwx  - oozie     oozie          0 2015-08-27 05:32 /user/oozie
drwxr-xr-x  - spark     spark          0 2015-08-27 05:32 /user/spark
drwxrwxrwx  - training  supergroup    0 2015-08-27 05:31 /user/training
[training@localhost ~]$ Setu Shah SXS210181
```

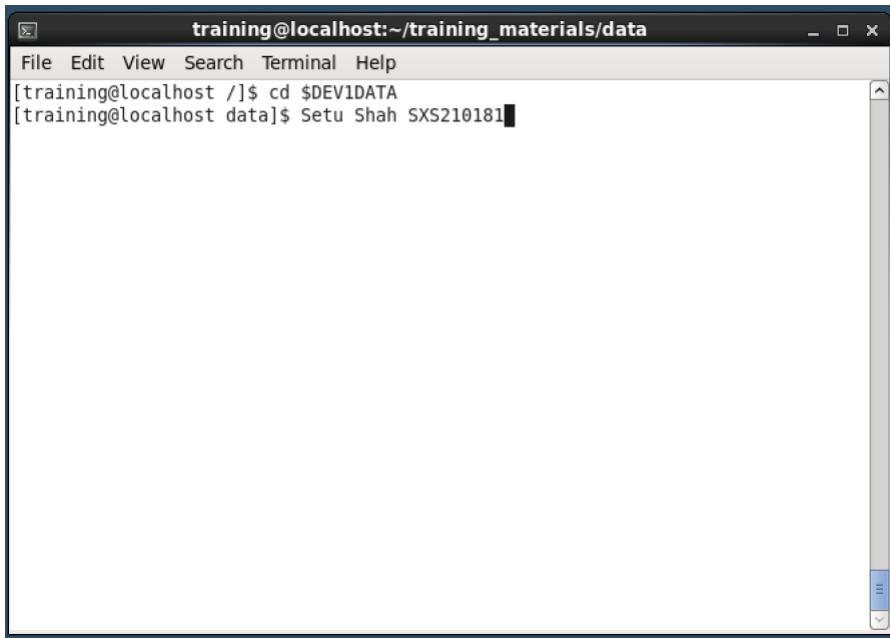
## **UPLOADING FILES TO HDFS**

Using the command “hdfs dfs -mkdir /loudacre” to create a new directory in home directory “training”. For me, after setting up, I used this command and hence it shows that it exists. We will be using this directory for the entire course.



```
training@localhost:~$ hdfs dfs -mkdir /loudacre
mkdir: '/loudacre': File exists
[training@localhost ~]$ Setu Shah SXS210181
```

Changing the directories to local file system directory using “cd \$DEV1DATA”, which contains the sample data.



A screenshot of a terminal window titled "training@localhost:~/training\_materials/data". The window has a standard Linux-style title bar with icons for minimize, maximize, and close. Below the title bar is a menu bar with "File", "Edit", "View", "Search", "Terminal", and "Help". The main area of the terminal shows the command history:

```
[training@localhost /]$ cd $DEV1DATA  
[training@localhost data]$ Setu Shah SXS210181
```

The terminal window is set against a white background with a dark blue border. A vertical scroll bar is visible on the right side of the terminal window.

Using the command “hdfs dfs -put kb /loudacre/” to copy local ‘kb’ directory and its content into remote HDFS directory “/loudacre/kb”.

Further I used, “hdfs dfs -ls /loudacre/kb” to list the contents of new HDFS directory. We can see the KB articles that were in the local directory.

As the files are in loudacre, it already exists.

The image shows two terminal windows side-by-side. Both windows have a title bar "training@localhost:~/training\_materials/data".

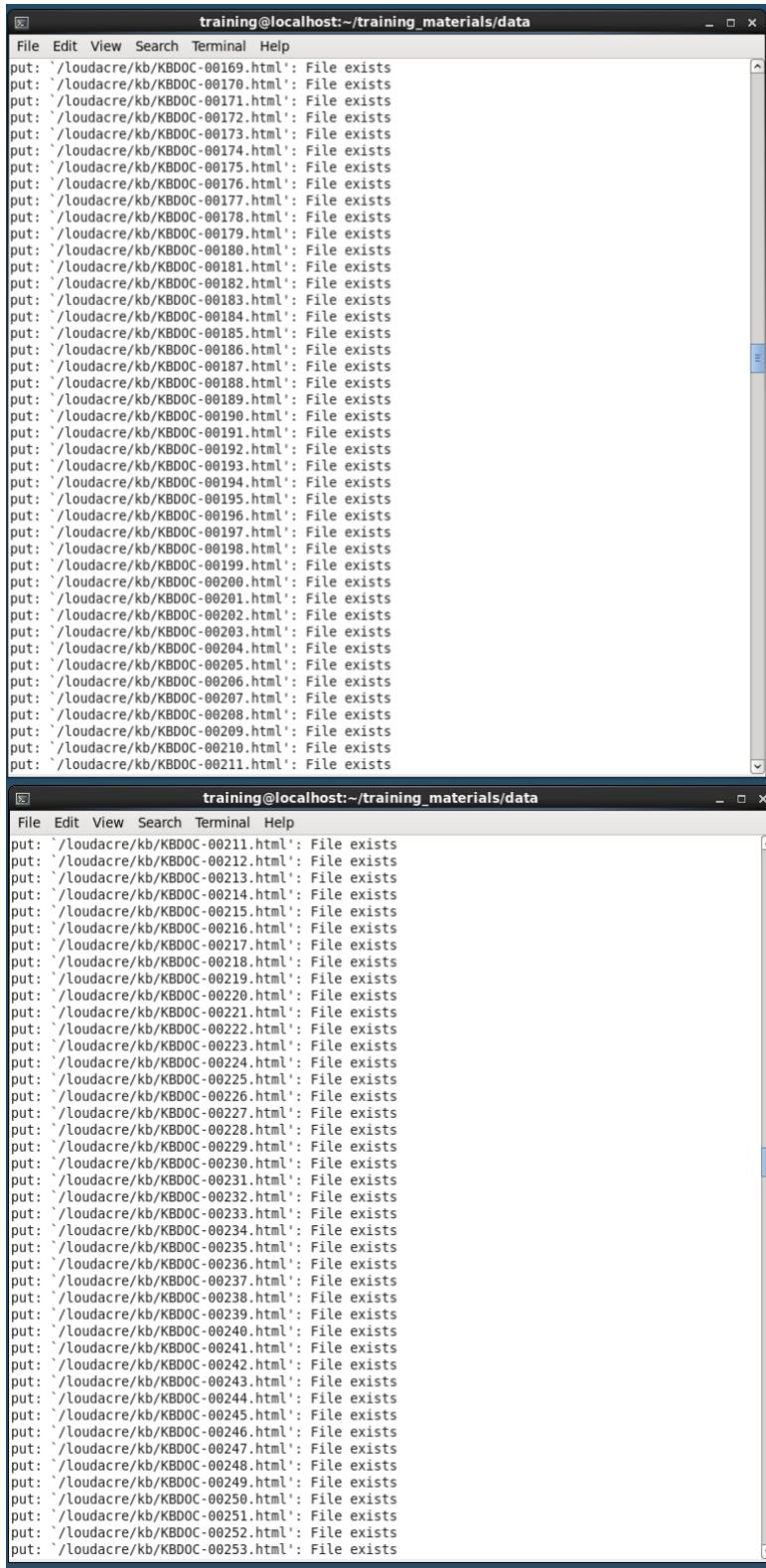
The top window displays the output of the command "hdfs dfs -ls /loudacre/kb". It lists numerous files in the /loudacre/kb directory, all ending in ".html" and starting with "KBDOC-00001" through "KBDOC-00042". Each file is listed as "File exists".

```
File Edit View Search Terminal Help
put: '/loudacre/kb/KBDOC-00001.html': File exists
put: '/loudacre/kb/KBDOC-00002.html': File exists
put: '/loudacre/kb/KBDOC-00003.html': File exists
put: '/loudacre/kb/KBDOC-00004.html': File exists
put: '/loudacre/kb/KBDOC-00005.html': File exists
put: '/loudacre/kb/KBDOC-00006.html': File exists
put: '/loudacre/kb/KBDOC-00007.html': File exists
put: '/loudacre/kb/KBDOC-00008.html': File exists
put: '/loudacre/kb/KBDOC-00009.html': File exists
put: '/loudacre/kb/KBDOC-00010.html': File exists
put: '/loudacre/kb/KBDOC-00011.html': File exists
put: '/loudacre/kb/KBDOC-00012.html': File exists
put: '/loudacre/kb/KBDOC-00013.html': File exists
put: '/loudacre/kb/KBDOC-00014.html': File exists
put: '/loudacre/kb/KBDOC-00015.html': File exists
put: '/loudacre/kb/KBDOC-00016.html': File exists
put: '/loudacre/kb/KBDOC-00017.html': File exists
put: '/loudacre/kb/KBDOC-00018.html': File exists
put: '/loudacre/kb/KBDOC-00019.html': File exists
put: '/loudacre/kb/KBDOC-00020.html': File exists
put: '/loudacre/kb/KBDOC-00021.html': File exists
put: '/loudacre/kb/KBDOC-00022.html': File exists
put: '/loudacre/kb/KBDOC-00023.html': File exists
put: '/loudacre/kb/KBDOC-00024.html': File exists
put: '/loudacre/kb/KBDOC-00025.html': File exists
put: '/loudacre/kb/KBDOC-00026.html': File exists
put: '/loudacre/kb/KBDOC-00027.html': File exists
put: '/loudacre/kb/KBDOC-00028.html': File exists
put: '/loudacre/kb/KBDOC-00029.html': File exists
put: '/loudacre/kb/KBDOC-00030.html': File exists
put: '/loudacre/kb/KBDOC-00031.html': File exists
put: '/loudacre/kb/KBDOC-00032.html': File exists
put: '/loudacre/kb/KBDOC-00033.html': File exists
put: '/loudacre/kb/KBDOC-00034.html': File exists
put: '/loudacre/kb/KBDOC-00035.html': File exists
put: '/loudacre/kb/KBDOC-00036.html': File exists
put: '/loudacre/kb/KBDOC-00037.html': File exists
put: '/loudacre/kb/KBDOC-00038.html': File exists
put: '/loudacre/kb/KBDOC-00039.html': File exists
put: '/loudacre/kb/KBDOC-00040.html': File exists
put: '/loudacre/kb/KBDOC-00041.html': File exists
put: '/loudacre/kb/KBDOC-00042.html': File exists
```

The bottom window displays the output of the command "hdfs dfs -put kb /loudacre". It lists the same set of files as the top window, all ending in ".html" and starting with "KBDOC-00001" through "KBDOC-00042". Each file is listed as "File exists".

```
[training@localhost data]$ hdfs dfs -put kb /loudacre
File Edit View Search Terminal Help
put: '/loudacre/kb/KBDOC-00001.html': File exists
put: '/loudacre/kb/KBDOC-00002.html': File exists
put: '/loudacre/kb/KBDOC-00003.html': File exists
put: '/loudacre/kb/KBDOC-00004.html': File exists
put: '/loudacre/kb/KBDOC-00005.html': File exists
put: '/loudacre/kb/KBDOC-00006.html': File exists
put: '/loudacre/kb/KBDOC-00007.html': File exists
put: '/loudacre/kb/KBDOC-00008.html': File exists
put: '/loudacre/kb/KBDOC-00009.html': File exists
put: '/loudacre/kb/KBDOC-00010.html': File exists
put: '/loudacre/kb/KBDOC-00011.html': File exists
put: '/loudacre/kb/KBDOC-00012.html': File exists
put: '/loudacre/kb/KBDOC-00013.html': File exists
put: '/loudacre/kb/KBDOC-00014.html': File exists
put: '/loudacre/kb/KBDOC-00015.html': File exists
put: '/loudacre/kb/KBDOC-00016.html': File exists
put: '/loudacre/kb/KBDOC-00017.html': File exists
put: '/loudacre/kb/KBDOC-00018.html': File exists
put: '/loudacre/kb/KBDOC-00019.html': File exists
put: '/loudacre/kb/KBDOC-00020.html': File exists
put: '/loudacre/kb/KBDOC-00021.html': File exists
put: '/loudacre/kb/KBDOC-00022.html': File exists
put: '/loudacre/kb/KBDOC-00023.html': File exists
put: '/loudacre/kb/KBDOC-00024.html': File exists
put: '/loudacre/kb/KBDOC-00025.html': File exists
put: '/loudacre/kb/KBDOC-00026.html': File exists
put: '/loudacre/kb/KBDOC-00027.html': File exists
put: '/loudacre/kb/KBDOC-00028.html': File exists
put: '/loudacre/kb/KBDOC-00029.html': File exists
put: '/loudacre/kb/KBDOC-00030.html': File exists
put: '/loudacre/kb/KBDOC-00031.html': File exists
put: '/loudacre/kb/KBDOC-00032.html': File exists
put: '/loudacre/kb/KBDOC-00033.html': File exists
put: '/loudacre/kb/KBDOC-00034.html': File exists
put: '/loudacre/kb/KBDOC-00035.html': File exists
put: '/loudacre/kb/KBDOC-00036.html': File exists
put: '/loudacre/kb/KBDOC-00037.html': File exists
put: '/loudacre/kb/KBDOC-00038.html': File exists
put: '/loudacre/kb/KBDOC-00039.html': File exists
put: '/loudacre/kb/KBDOC-00040.html': File exists
put: '/loudacre/kb/KBDOC-00041.html': File exists
put: '/loudacre/kb/KBDOC-00042.html': File exists
```





The image shows two terminal windows side-by-side, both titled "training@localhost:~/training\_materials/data". The terminals are displaying a series of "put" commands followed by file paths and the message "File exists".

The top terminal window displays the following log:

```
put: '/loudacre/kb/KBDOC-00169.html': File exists
put: '/loudacre/kb/KBDOC-00170.html': File exists
put: '/loudacre/kb/KBDOC-00171.html': File exists
put: '/loudacre/kb/KBDOC-00172.html': File exists
put: '/loudacre/kb/KBDOC-00173.html': File exists
put: '/loudacre/kb/KBDOC-00174.html': File exists
put: '/loudacre/kb/KBDOC-00175.html': File exists
put: '/loudacre/kb/KBDOC-00176.html': File exists
put: '/loudacre/kb/KBDOC-00177.html': File exists
put: '/loudacre/kb/KBDOC-00178.html': File exists
put: '/loudacre/kb/KBDOC-00179.html': File exists
put: '/loudacre/kb/KBDOC-00180.html': File exists
put: '/loudacre/kb/KBDOC-00181.html': File exists
put: '/loudacre/kb/KBDOC-00182.html': File exists
put: '/loudacre/kb/KBDOC-00183.html': File exists
put: '/loudacre/kb/KBDOC-00184.html': File exists
put: '/loudacre/kb/KBDOC-00185.html': File exists
put: '/loudacre/kb/KBDOC-00186.html': File exists
put: '/loudacre/kb/KBDOC-00187.html': File exists
put: '/loudacre/kb/KBDOC-00188.html': File exists
put: '/loudacre/kb/KBDOC-00189.html': File exists
put: '/loudacre/kb/KBDOC-00190.html': File exists
put: '/loudacre/kb/KBDOC-00191.html': File exists
put: '/loudacre/kb/KBDOC-00192.html': File exists
put: '/loudacre/kb/KBDOC-00193.html': File exists
put: '/loudacre/kb/KBDOC-00194.html': File exists
put: '/loudacre/kb/KBDOC-00195.html': File exists
put: '/loudacre/kb/KBDOC-00196.html': File exists
put: '/loudacre/kb/KBDOC-00197.html': File exists
put: '/loudacre/kb/KBDOC-00198.html': File exists
put: '/loudacre/kb/KBDOC-00199.html': File exists
put: '/loudacre/kb/KBDOC-00200.html': File exists
put: '/loudacre/kb/KBDOC-00201.html': File exists
put: '/loudacre/kb/KBDOC-00202.html': File exists
put: '/loudacre/kb/KBDOC-00203.html': File exists
put: '/loudacre/kb/KBDOC-00204.html': File exists
put: '/loudacre/kb/KBDOC-00205.html': File exists
put: '/loudacre/kb/KBDOC-00206.html': File exists
put: '/loudacre/kb/KBDOC-00207.html': File exists
put: '/loudacre/kb/KBDOC-00208.html': File exists
put: '/loudacre/kb/KBDOC-00209.html': File exists
put: '/loudacre/kb/KBDOC-00210.html': File exists
put: '/loudacre/kb/KBDOC-00211.html': File exists
```

The bottom terminal window displays a similar log, starting from file ID 00211:

```
put: '/loudacre/kb/KBDOC-00211.html': File exists
put: '/loudacre/kb/KBDOC-00212.html': File exists
put: '/loudacre/kb/KBDOC-00213.html': File exists
put: '/loudacre/kb/KBDOC-00214.html': File exists
put: '/loudacre/kb/KBDOC-00215.html': File exists
put: '/loudacre/kb/KBDOC-00216.html': File exists
put: '/loudacre/kb/KBDOC-00217.html': File exists
put: '/loudacre/kb/KBDOC-00218.html': File exists
put: '/loudacre/kb/KBDOC-00219.html': File exists
put: '/loudacre/kb/KBDOC-00220.html': File exists
put: '/loudacre/kb/KBDOC-00221.html': File exists
put: '/loudacre/kb/KBDOC-00222.html': File exists
put: '/loudacre/kb/KBDOC-00223.html': File exists
put: '/loudacre/kb/KBDOC-00224.html': File exists
put: '/loudacre/kb/KBDOC-00225.html': File exists
put: '/loudacre/kb/KBDOC-00226.html': File exists
put: '/loudacre/kb/KBDOC-00227.html': File exists
put: '/loudacre/kb/KBDOC-00228.html': File exists
put: '/loudacre/kb/KBDOC-00229.html': File exists
put: '/loudacre/kb/KBDOC-00230.html': File exists
put: '/loudacre/kb/KBDOC-00231.html': File exists
put: '/loudacre/kb/KBDOC-00232.html': File exists
put: '/loudacre/kb/KBDOC-00233.html': File exists
put: '/loudacre/kb/KBDOC-00234.html': File exists
put: '/loudacre/kb/KBDOC-00235.html': File exists
put: '/loudacre/kb/KBDOC-00236.html': File exists
put: '/loudacre/kb/KBDOC-00237.html': File exists
put: '/loudacre/kb/KBDOC-00238.html': File exists
put: '/loudacre/kb/KBDOC-00239.html': File exists
put: '/loudacre/kb/KBDOC-00240.html': File exists
put: '/loudacre/kb/KBDOC-00241.html': File exists
put: '/loudacre/kb/KBDOC-00242.html': File exists
put: '/loudacre/kb/KBDOC-00243.html': File exists
put: '/loudacre/kb/KBDOC-00244.html': File exists
put: '/loudacre/kb/KBDOC-00245.html': File exists
put: '/loudacre/kb/KBDOC-00246.html': File exists
put: '/loudacre/kb/KBDOC-00247.html': File exists
put: '/loudacre/kb/KBDOC-00248.html': File exists
put: '/loudacre/kb/KBDOC-00249.html': File exists
put: '/loudacre/kb/KBDOC-00250.html': File exists
put: '/loudacre/kb/KBDOC-00251.html': File exists
put: '/loudacre/kb/KBDOC-00252.html': File exists
put: '/loudacre/kb/KBDOC-00253.html': File exists
```

```
File Edit View Search Terminal Help
put: '/loudacre/kb/KBDOC-00253.html': File exists
put: '/loudacre/kb/KBDOC-00254.html': File exists
put: '/loudacre/kb/KBDOC-00255.html': File exists
put: '/loudacre/kb/KBDOC-00256.html': File exists
put: '/loudacre/kb/KBDOC-00257.html': File exists
put: '/loudacre/kb/KBDOC-00258.html': File exists
put: '/loudacre/kb/KBDOC-00259.html': File exists
put: '/loudacre/kb/KBDOC-00260.html': File exists
put: '/loudacre/kb/KBDOC-00261.html': File exists
put: '/loudacre/kb/KBDOC-00262.html': File exists
put: '/loudacre/kb/KBDOC-00263.html': File exists
put: '/loudacre/kb/KBDOC-00264.html': File exists
put: '/loudacre/kb/KBDOC-00265.html': File exists
put: '/loudacre/kb/KBDOC-00266.html': File exists
put: '/loudacre/kb/KBDOC-00267.html': File exists
put: '/loudacre/kb/KBDOC-00268.html': File exists
put: '/loudacre/kb/KBDOC-00269.html': File exists
put: '/loudacre/kb/KBDOC-00270.html': File exists
put: '/loudacre/kb/KBDOC-00271.html': File exists
put: '/loudacre/kb/KBDOC-00272.html': File exists
put: '/loudacre/kb/KBDOC-00273.html': File exists
put: '/loudacre/kb/KBDOC-00274.html': File exists
put: '/loudacre/kb/KBDOC-00275.html': File exists
put: '/loudacre/kb/KBDOC-00276.html': File exists
put: '/loudacre/kb/KBDOC-00277.html': File exists
put: '/loudacre/kb/KBDOC-00278.html': File exists
put: '/loudacre/kb/KBDOC-00279.html': File exists
put: '/loudacre/kb/KBDOC-00280.html': File exists
put: '/loudacre/kb/KBDOC-00281.html': File exists
put: '/loudacre/kb/KBDOC-00282.html': File exists
put: '/loudacre/kb/KBDOC-00283.html': File exists
put: '/loudacre/kb/KBDOC-00284.html': File exists
put: '/loudacre/kb/KBDOC-00285.html': File exists
put: '/loudacre/kb/KBDOC-00286.html': File exists
put: '/loudacre/kb/KBDOC-00287.html': File exists
put: '/loudacre/kb/KBDOC-00288.html': File exists
put: '/loudacre/kb/KBDOC-00289.html': File exists
put: '/loudacre/kb/KBDOC-00290.html': File exists
put: '/loudacre/kb/KBDOC-00291.html': File exists
put: '/loudacre/kb/KBDOC-00292.html': File exists
put: '/loudacre/kb/KBDOC-00293.html': File exists
put: '/loudacre/kb/KBDOC-00294.html': File exists
put: '/loudacre/kb/KBDOC-00295.html': File exists
```

The file then starts the list of contents in the directories.  
 These are the list of all the files

```
training@localhost:~/training_materials/data
File Edit View Search Terminal Help
put: '/loudacre/kb/KBDOC-00293.html': File exists
put: '/loudacre/kb/KBDOC-00294.html': File exists
put: '/loudacre/kb/KBDOC-00295.html': File exists
put: '/loudacre/kb/KBDOC-00296.html': File exists
put: '/loudacre/kb/KBDOC-00297.html': File exists
put: '/loudacre/kb/KBDOC-00298.html': File exists
put: '/loudacre/kb/KBDOC-00299.html': File exists
put: '/loudacre/kb/KBDOC-00300.html': File exists
[training@localhost data]$ hdfs dfs -ls /loudacre/kb
Found 300 items
-rw-rw-r- 1 training supergroup 1626 2022-02-16 19:31 /loudacre/kb/KB00
C-00001.html
-rw-rw-r- 1 training supergroup 1298 2022-02-16 19:31 /loudacre/kb/KB00
C-00002.html
-rw-rw-r- 1 training supergroup 1183 2022-02-16 19:31 /loudacre/kb/KB00
C-00003.html
-rw-rw-r- 1 training supergroup 1029 2022-02-16 19:31 /loudacre/kb/KB00
C-00004.html
-rw-rw-r- 1 training supergroup 1155 2022-02-16 19:31 /loudacre/kb/KB00
C-00005.html
-rw-rw-r- 1 training supergroup 1189 2022-02-16 19:31 /loudacre/kb/KB00
C-00006.html
-rw-rw-r- 1 training supergroup 1648 2022-02-16 19:31 /loudacre/kb/KB00
C-00007.html
-rw-rw-r- 1 training supergroup 1303 2022-02-16 19:31 /loudacre/kb/KB00
C-00008.html
-rw-rw-r- 1 training supergroup 1173 2022-02-16 19:31 /loudacre/kb/KB00
C-00009.html
-rw-rw-r- 1 training supergroup 1017 2022-02-16 19:31 /loudacre/kb/KB00
C-00010.html
-rw-rw-r- 1 training supergroup 1149 2022-02-16 19:31 /loudacre/kb/KB00
C-00011.html
-rw-rw-r- 1 training supergroup 1117 2022-02-16 19:31 /loudacre/kb/KB00
C-00012.html
-rw-rw-r- 1 training supergroup 1638 2022-02-16 19:31 /loudacre/kb/KB00
C-00013.html
-rw-rw-r- 1 training supergroup 1282 2022-02-16 19:31 /loudacre/kb/KB00
C-00014.html
-rw-rw-r- 1 training supergroup 1166 2022-02-16 19:31 /loudacre/kb/KB00
C-00015.html
-rw-rw-r- 1 training supergroup 1017 2022-02-16 19:31 /loudacre/kb/KB00
C-00016.html
-rw-rw-r- 1 training supergroup 1160 2022-02-16 19:31 /loudacre/kb/KB00

training@localhost:~/training_materials/data
File Edit View Search Terminal Help
-rw-rw-r- 1 training supergroup 1160 2022-02-16 19:31 /loudacre/kb/KB00
C-00017.html
-rw-rw-r- 1 training supergroup 1105 2022-02-16 19:31 /loudacre/kb/KB00
C-00018.html
-rw-rw-r- 1 training supergroup 1612 2022-02-16 19:31 /loudacre/kb/KB00
C-00019.html
-rw-rw-r- 1 training supergroup 1274 2022-02-16 19:31 /loudacre/kb/KB00
C-00020.html
-rw-rw-r- 1 training supergroup 1176 2022-02-16 19:31 /loudacre/kb/KB00
C-00021.html
-rw-rw-r- 1 training supergroup 1024 2022-02-16 19:31 /loudacre/kb/KB00
C-00022.html
-rw-rw-r- 1 training supergroup 1149 2022-02-16 19:31 /loudacre/kb/KB00
C-00023.html
-rw-rw-r- 1 training supergroup 1103 2022-02-16 19:31 /loudacre/kb/KB00
C-00024.html
-rw-rw-r- 1 training supergroup 1580 2022-02-16 19:31 /loudacre/kb/KB00
C-00025.html
-rw-rw-r- 1 training supergroup 1242 2022-02-16 19:31 /loudacre/kb/KB00
C-00026.html
-rw-rw-r- 1 training supergroup 1143 2022-02-16 19:31 /loudacre/kb/KB00
C-00027.html
-rw-rw-r- 1 training supergroup 1013 2022-02-16 19:31 /loudacre/kb/KB00
C-00028.html
-rw-rw-r- 1 training supergroup 1130 2022-02-16 19:31 /loudacre/kb/KB00
C-00029.html
-rw-rw-r- 1 training supergroup 1101 2022-02-16 19:31 /loudacre/kb/KB00
C-00030.html
-rw-rw-r- 1 training supergroup 1630 2022-02-16 19:31 /loudacre/kb/KB00
C-00031.html
-rw-rw-r- 1 training supergroup 1258 2022-02-16 19:31 /loudacre/kb/KB00
C-00032.html
-rw-rw-r- 1 training supergroup 1143 2022-02-16 19:31 /loudacre/kb/KB00
C-00033.html
-rw-rw-r- 1 training supergroup 1003 2022-02-16 19:31 /loudacre/kb/KB00
C-00034.html
-rw-rw-r- 1 training supergroup 1118 2022-02-16 19:31 /loudacre/kb/KB00
C-00035.html
-rw-rw-r- 1 training supergroup 1096 2022-02-16 19:31 /loudacre/kb/KB00
C-00036.html
-rw-rw-r- 1 training supergroup 1602 2022-02-16 19:31 /loudacre/kb/KB00
C-00037.html
-rw-rw-r- 1 training supergroup 1247 2022-02-16 19:31 /loudacre/kb/KB00
```

training@localhost:~/training\_materials/data

-rw-rw-rw-	1	training supergroup
C-00037.html	1602	2022-02-16 19:31 /loudacre/kb/KBDO
-rw-rw-rw-	1	training supergroup
C-00038.html	1247	2022-02-16 19:31 /loudacre/kb/KBDO
-rw-rw-rw-	1	training supergroup
C-00039.html	1173	2022-02-16 19:31 /loudacre/kb/KBDO
-rw-rw-rw-	1	training supergroup
C-00040.html	1003	2022-02-16 19:31 /loudacre/kb/KBDO
-rw-rw-rw-	1	training supergroup
C-00041.html	1136	2022-02-16 19:31 /loudacre/kb/KBDO
-rw-rw-rw-	1	training supergroup
C-00042.html	1107	2022-02-16 19:31 /loudacre/kb/KBDO
-rw-rw-rw-	1	training supergroup
C-00043.html	1625	2022-02-16 19:31 /loudacre/kb/KBDO
-rw-rw-rw-	1	training supergroup
C-00044.html	1266	2022-02-16 19:31 /loudacre/kb/KBDO
-rw-rw-rw-	1	training supergroup
C-00045.html	1151	2022-02-16 19:31 /loudacre/kb/KBDO
-rw-rw-rw-	1	training supergroup
C-00046.html	992	2022-02-16 19:31 /loudacre/kb/KBDO
-rw-rw-rw-	1	training supergroup
C-00047.html	1127	2022-02-16 19:31 /loudacre/kb/KBDO
-rw-rw-rw-	1	training supergroup
C-00048.html	1078	2022-02-16 19:31 /loudacre/kb/KBDO
-rw-rw-rw-	1	training supergroup
C-00049.html	1612	2022-02-16 19:31 /loudacre/kb/KBDO
-rw-rw-rw-	1	training supergroup
C-00050.html	1290	2022-02-16 19:31 /loudacre/kb/KBDO
-rw-rw-rw-	1	training supergroup
C-00051.html	1180	2022-02-16 19:31 /loudacre/kb/KBDO
-rw-rw-rw-	1	training supergroup
C-00052.html	1011	2022-02-16 19:31 /loudacre/kb/KBDO
-rw-rw-rw-	1	training supergroup
C-00053.html	1147	2022-02-16 19:31 /loudacre/kb/KBDO
-rw-rw-rw-	1	training supergroup
C-00054.html	1114	2022-02-16 19:31 /loudacre/kb/KBDO
-rw-rw-rw-	1	training supergroup
C-00055.html	1643	2022-02-16 19:31 /loudacre/kb/KBDO
-rw-rw-rw-	1	training supergroup
C-00056.html	1296	2022-02-16 19:31 /loudacre/kb/KBDO
-rw-rw-rw-	1	training supergroup
C-00057.html	1175	2022-02-16 19:31 /loudacre/kb/KBDO
-rw-rw-rw-	1	training supergroup
C-00058.html	1038	2022-02-16 19:31 /loudacre/kb/KBDO

training@localhost:~/training\_materials/data

C-00056.html		
-rw-rw-rw-	1	training supergroup
C-00057.html	1175	2022-02-16 19:31 /loudacre/kb/KBDO
-rw-rw-rw-	1	training supergroup
C-00058.html	1038	2022-02-16 19:31 /loudacre/kb/KBDO
-rw-rw-rw-	1	training supergroup
C-00059.html	1134	2022-02-16 19:31 /loudacre/kb/KBDO
-rw-rw-rw-	1	training supergroup
C-00060.html	1105	2022-02-16 19:31 /loudacre/kb/KBDO
-rw-rw-rw-	1	training supergroup
C-00061.html	1646	2022-02-16 19:31 /loudacre/kb/KBDO
-rw-rw-rw-	1	training supergroup
C-00062.html	1315	2022-02-16 19:31 /loudacre/kb/KBDO
-rw-rw-rw-	1	training supergroup
C-00063.html	1189	2022-02-16 19:31 /loudacre/kb/KBDO
-rw-rw-rw-	1	training supergroup
C-00064.html	1031	2022-02-16 19:31 /loudacre/kb/KBDO
-rw-rw-rw-	1	training supergroup
C-00065.html	1152	2022-02-16 19:31 /loudacre/kb/KBDO
-rw-rw-rw-	1	training supergroup
C-00066.html	1097	2022-02-16 19:31 /loudacre/kb/KBDO
-rw-rw-rw-	1	training supergroup
C-00067.html	1621	2022-02-16 19:31 /loudacre/kb/KBDO
-rw-rw-rw-	1	training supergroup
C-00068.html	1262	2022-02-16 19:31 /loudacre/kb/KBDO
-rw-rw-rw-	1	training supergroup
C-00069.html	1161	2022-02-16 19:31 /loudacre/kb/KBDO
-rw-rw-rw-	1	training supergroup
C-00070.html	1003	2022-02-16 19:31 /loudacre/kb/KBDO
-rw-rw-rw-	1	training supergroup
C-00071.html	1141	2022-02-16 19:31 /loudacre/kb/KBDO
-rw-rw-rw-	1	training supergroup
C-00072.html	1100	2022-02-16 19:31 /loudacre/kb/KBDO
-rw-rw-rw-	1	training supergroup
C-00073.html	1628	2022-02-16 19:31 /loudacre/kb/KBDO
-rw-rw-rw-	1	training supergroup
C-00074.html	1282	2022-02-16 19:31 /loudacre/kb/KBDO
-rw-rw-rw-	1	training supergroup
C-00075.html	1187	2022-02-16 19:31 /loudacre/kb/KBDO
-rw-rw-rw-	1	training supergroup
C-00076.html	1040	2022-02-16 19:31 /loudacre/kb/KBDO
-rw-rw-rw-	1	training supergroup
C-00077.html	1151	2022-02-16 19:31 /loudacre/kb/KBDO



training@localhost:~/training_materials/data			
File	Edit	View	Search Terminal Help
-rw-rw-rw-	1	training supergroup	1135 2022-02-16 19:31 /loudacre/kb/KBDO
C-00119.html			
-rw-rw-rw-	1	training supergroup	1095 2022-02-16 19:31 /loudacre/kb/KBDO
C-00120.html			
-rw-rw-rw-	1	training supergroup	1613 2022-02-16 19:31 /loudacre/kb/KBDO
C-00121.html			
-rw-rw-rw-	1	training supergroup	1283 2022-02-16 19:31 /loudacre/kb/KBDO
C-00122.html			
-rw-rw-rw-	1	training supergroup	1162 2022-02-16 19:31 /loudacre/kb/KBDO
C-00123.html			
-rw-rw-rw-	1	training supergroup	1046 2022-02-16 19:31 /loudacre/kb/KBDO
C-00124.html			
-rw-rw-rw-	1	training supergroup	1151 2022-02-16 19:31 /loudacre/kb/KBDO
C-00125.html			
-rw-rw-rw-	1	training supergroup	1122 2022-02-16 19:31 /loudacre/kb/KBDO
C-00126.html			
-rw-rw-rw-	1	training supergroup	1657 2022-02-16 19:31 /loudacre/kb/KBDO
C-00127.html			
-rw-rw-rw-	1	training supergroup	1298 2022-02-16 19:31 /loudacre/kb/KBDO
C-00128.html			
-rw-rw-rw-	1	training supergroup	1189 2022-02-16 19:31 /loudacre/kb/KBDO
C-00129.html			
-rw-rw-rw-	1	training supergroup	1037 2022-02-16 19:31 /loudacre/kb/KBDO
C-00130.html			
-rw-rw-rw-	1	training supergroup	1151 2022-02-16 19:31 /loudacre/kb/KBDO
C-00131.html			
-rw-rw-rw-	1	training supergroup	1125 2022-02-16 19:31 /loudacre/kb/KBDO
C-00132.html			
-rw-rw-rw-	1	training supergroup	1630 2022-02-16 19:31 /loudacre/kb/KBDO
C-00133.html			
-rw-rw-rw-	1	training supergroup	1307 2022-02-16 19:31 /loudacre/kb/KBDO
C-00134.html			
-rw-rw-rw-	1	training supergroup	1208 2022-02-16 19:31 /loudacre/kb/KBDO
C-00135.html			
-rw-rw-rw-	1	training supergroup	1023 2022-02-16 19:31 /loudacre/kb/KBDO
C-00136.html			
-rw-rw-rw-	1	training supergroup	1165 2022-02-16 19:31 /loudacre/kb/KBDO
C-00137.html			
-rw-rw-rw-	1	training supergroup	1097 2022-02-16 19:31 /loudacre/kb/KBDO
C-00138.html			
-rw-rw-rw-	1	training supergroup	1626 2022-02-16 19:31 /loudacre/kb/KBDO
C-00139.html			
-rw-rw-rw-	1	training supergroup	1319 2022-02-16 19:31 /loudacre/kb/KBDO
C-00140.html			
-rw-rw-rw-	1	training supergroup	1319 2022-02-16 19:31 /loudacre/kb/KBDO
C-00141.html			
-rw-rw-rw-	1	training supergroup	1199 2022-02-16 19:31 /loudacre/kb/KBDO
C-00142.html			
-rw-rw-rw-	1	training supergroup	1043 2022-02-16 19:31 /loudacre/kb/KBDO
C-00143.html			
-rw-rw-rw-	1	training supergroup	1146 2022-02-16 19:31 /loudacre/kb/KBDO
C-00144.html			
-rw-rw-rw-	1	training supergroup	1115 2022-02-16 19:31 /loudacre/kb/KBDO
C-00145.html			
-rw-rw-rw-	1	training supergroup	1643 2022-02-16 19:31 /loudacre/kb/KBDO
C-00146.html			
-rw-rw-rw-	1	training supergroup	1291 2022-02-16 19:31 /loudacre/kb/KBDO
C-00147.html			
-rw-rw-rw-	1	training supergroup	1195 2022-02-16 19:31 /loudacre/kb/KBDO
C-00148.html			
-rw-rw-rw-	1	training supergroup	1046 2022-02-16 19:31 /loudacre/kb/KBDO
C-00149.html			
-rw-rw-rw-	1	training supergroup	1155 2022-02-16 19:31 /loudacre/kb/KBDO
C-00150.html			
-rw-rw-rw-	1	training supergroup	1114 2022-02-16 19:31 /loudacre/kb/KBDO
C-00151.html			
-rw-rw-rw-	1	training supergroup	1644 2022-02-16 19:31 /loudacre/kb/KBDO
C-00152.html			
-rw-rw-rw-	1	training supergroup	1302 2022-02-16 19:31 /loudacre/kb/KBDO
C-00153.html			
-rw-rw-rw-	1	training supergroup	1194 2022-02-16 19:31 /loudacre/kb/KBDO
C-00154.html			
-rw-rw-rw-	1	training supergroup	1016 2022-02-16 19:31 /loudacre/kb/KBDO
C-00155.html			
-rw-rw-rw-	1	training supergroup	1168 2022-02-16 19:31 /loudacre/kb/KBDO
C-00156.html			
-rw-rw-rw-	1	training supergroup	1139 2022-02-16 19:31 /loudacre/kb/KBDO
C-00157.html			
-rw-rw-rw-	1	training supergroup	1619 2022-02-16 19:31 /loudacre/kb/KBDO
C-00158.html			
-rw-rw-rw-	1	training supergroup	1287 2022-02-16 19:31 /loudacre/kb/KBDO
C-00159.html			
-rw-rw-rw-	1	training supergroup	1187 2022-02-16 19:31 /loudacre/kb/KBDO
C-00160.html			

File Edit View Search Terminal Help			
C-00160.html	1	training supergroup	1146 2022-02-16 19:31 /loudacre/kb/KBDO
-rw-rw-rw-			
C-00161.html	1	training supergroup	1135 2022-02-16 19:31 /loudacre/kb/KBDO
-rw-rw-rw-			
C-00162.html	1	training supergroup	1623 2022-02-16 19:31 /loudacre/kb/KBDO
-rw-rw-rw-			
C-00163.html	1	training supergroup	1280 2022-02-16 19:31 /loudacre/kb/KBDO
-rw-rw-rw-			
C-00164.html	1	training supergroup	1175 2022-02-16 19:31 /loudacre/kb/KBDO
-rw-rw-rw-			
C-00165.html	1	training supergroup	1019 2022-02-16 19:31 /loudacre/kb/KBDO
-rw-rw-rw-			
C-00166.html	1	training supergroup	1126 2022-02-16 19:31 /loudacre/kb/KBDO
-rw-rw-rw-			
C-00167.html	1	training supergroup	1099 2022-02-16 19:31 /loudacre/kb/KBDO
-rw-rw-rw-			
C-00168.html	1	training supergroup	1647 2022-02-16 19:31 /loudacre/kb/KBDO
-rw-rw-rw-			
C-00169.html	1	training supergroup	1300 2022-02-16 19:31 /loudacre/kb/KBDO
-rw-rw-rw-			
C-00170.html	1	training supergroup	1210 2022-02-16 19:31 /loudacre/kb/KBDO
-rw-rw-rw-			
C-00171.html	1	training supergroup	1016 2022-02-16 19:31 /loudacre/kb/KBDO
-rw-rw-rw-			
C-00172.html	1	training supergroup	1153 2022-02-16 19:31 /loudacre/kb/KBDO
-rw-rw-rw-			
C-00173.html	1	training supergroup	1120 2022-02-16 19:31 /loudacre/kb/KBDO
-rw-rw-rw-			
C-00174.html	1	training supergroup	1667 2022-02-16 19:31 /loudacre/kb/KBDO
-rw-rw-rw-			
C-00175.html	1	training supergroup	1330 2022-02-16 19:31 /loudacre/kb/KBDO
-rw-rw-rw-			
C-00176.html	1	training supergroup	1254 2022-02-16 19:31 /loudacre/kb/KBDO
-rw-rw-rw-			
C-00177.html	1	training supergroup	1064 2022-02-16 19:31 /loudacre/kb/KBDO
-rw-rw-rw-			
C-00178.html	1	training supergroup	1179 2022-02-16 19:31 /loudacre/kb/KBDO
-rw-rw-rw-			
C-00179.html	1	training supergroup	1125 2022-02-16 19:31 /loudacre/kb/KBDO
-rw-rw-rw-			
C-00180.html	1	training supergroup	1625 2022-02-16 19:31 /loudacre/kb/KBDO
-rw-rw-rw-			
C-00181.html			

File Edit View Search Terminal Help			
C-00180.html	1	training supergroup	1625 2022-02-16 19:31 /loudacre/kb/KBDO
-rw-rw-rw-			
C-00181.html	1	training supergroup	1271 2022-02-16 19:31 /loudacre/kb/KBDO
-rw-rw-rw-			
C-00182.html	1	training supergroup	1165 2022-02-16 19:31 /loudacre/kb/KBDO
-rw-rw-rw-			
C-00183.html	1	training supergroup	1032 2022-02-16 19:31 /loudacre/kb/KBDO
-rw-rw-rw-			
C-00184.html	1	training supergroup	1158 2022-02-16 19:31 /loudacre/kb/KBDO
-rw-rw-rw-			
C-00185.html	1	training supergroup	1099 2022-02-16 19:31 /loudacre/kb/KBDO
-rw-rw-rw-			
C-00186.html	1	training supergroup	1621 2022-02-16 19:31 /loudacre/kb/KBDO
-rw-rw-rw-			
C-00187.html	1	training supergroup	1262 2022-02-16 19:31 /loudacre/kb/KBDO
-rw-rw-rw-			
C-00188.html	1	training supergroup	1167 2022-02-16 19:31 /loudacre/kb/KBDO
-rw-rw-rw-			
C-00189.html	1	training supergroup	1023 2022-02-16 19:31 /loudacre/kb/KBDO
-rw-rw-rw-			
C-00190.html	1	training supergroup	1126 2022-02-16 19:31 /loudacre/kb/KBDO
-rw-rw-rw-			
C-00191.html	1	training supergroup	1126 2022-02-16 19:31 /loudacre/kb/KBDO
-rw-rw-rw-			
C-00192.html	1	training supergroup	1624 2022-02-16 19:31 /loudacre/kb/KBDO
-rw-rw-rw-			
C-00193.html	1	training supergroup	1283 2022-02-16 19:31 /loudacre/kb/KBDO
-rw-rw-rw-			
C-00194.html	1	training supergroup	1178 2022-02-16 19:31 /loudacre/kb/KBDO
-rw-rw-rw-			
C-00195.html	1	training supergroup	1027 2022-02-16 19:31 /loudacre/kb/KBDO
-rw-rw-rw-			
C-00196.html	1	training supergroup	1135 2022-02-16 19:31 /loudacre/kb/KBDO
-rw-rw-rw-			
C-00197.html	1	training supergroup	1091 2022-02-16 19:31 /loudacre/kb/KBDO
-rw-rw-rw-			
C-00198.html	1	training supergroup	1623 2022-02-16 19:31 /loudacre/kb/KBDO
-rw-rw-rw-			
C-00199.html	1	training supergroup	1309 2022-02-16 19:31 /loudacre/kb/KBDO
-rw-rw-rw-			
C-00200.html	1	training supergroup	1191 2022-02-16 19:31 /loudacre/kb/KBDO
-rw-rw-rw-			
C-00201.html			

File Edit View Search Terminal Help					
C-00201.html	-rw-rw-rw-	1	training	supergroup	1030 2022-02-16 19:31 /loudacre/kb/KB00
C-00202.html	-rw-rw-rw-	1	training	supergroup	1152 2022-02-16 19:31 /loudacre/kb/KB00
C-00203.html	-rw-rw-rw-	1	training	supergroup	1100 2022-02-16 19:31 /loudacre/kb/KB00
C-00204.html	-rw-rw-rw-	1	training	supergroup	1604 2022-02-16 19:31 /loudacre/kb/KB00
C-00205.html	-rw-rw-rw-	1	training	supergroup	1271 2022-02-16 19:31 /loudacre/kb/KB00
C-00206.html	-rw-rw-rw-	1	training	supergroup	1160 2022-02-16 19:31 /loudacre/kb/KB00
C-00207.html	-rw-rw-rw-	1	training	supergroup	1029 2022-02-16 19:31 /loudacre/kb/KB00
C-00208.html	-rw-rw-rw-	1	training	supergroup	1128 2022-02-16 19:31 /loudacre/kb/KB00
C-00209.html	-rw-rw-rw-	1	training	supergroup	1091 2022-02-16 19:31 /loudacre/kb/KB00
C-00210.html	-rw-rw-rw-	1	training	supergroup	1614 2022-02-16 19:31 /loudacre/kb/KB00
C-00211.html	-rw-rw-rw-	1	training	supergroup	1268 2022-02-16 19:31 /loudacre/kb/KB00
C-00212.html	-rw-rw-rw-	1	training	supergroup	1160 2022-02-16 19:31 /loudacre/kb/KB00
C-00213.html	-rw-rw-rw-	1	training	supergroup	1020 2022-02-16 19:31 /loudacre/kb/KB00
C-00214.html	-rw-rw-rw-	1	training	supergroup	1142 2022-02-16 19:31 /loudacre/kb/KB00
C-00215.html	-rw-rw-rw-	1	training	supergroup	1090 2022-02-16 19:31 /loudacre/kb/KB00
C-00216.html	-rw-rw-rw-	1	training	supergroup	1607 2022-02-16 19:31 /loudacre/kb/KB00
C-00217.html	-rw-rw-rw-	1	training	supergroup	1275 2022-02-16 19:31 /loudacre/kb/KB00
C-00218.html	-rw-rw-rw-	1	training	supergroup	1178 2022-02-16 19:31 /loudacre/kb/KB00
C-00219.html	-rw-rw-rw-	1	training	supergroup	1036 2022-02-16 19:31 /loudacre/kb/KB00
C-00220.html	-rw-rw-rw-	1	training	supergroup	1144 2022-02-16 19:31 /loudacre/kb/KB00
C-00221.html	-rw-rw-rw-	1	training	supergroup	1112 2022-02-16 19:31 /loudacre/kb/KB00
C-00222.html	-rw-rw-rw-	1	training	supergroup	

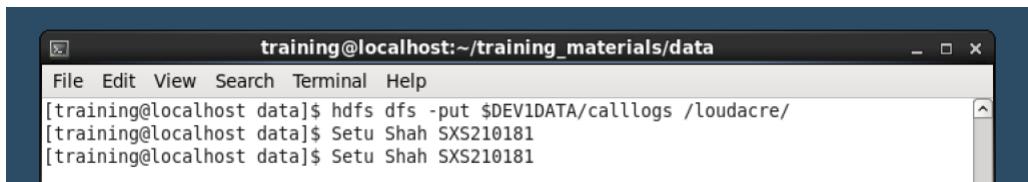
training@localhost:~/training_materials/data					
File	Edit	View	Search	Terminal	Help
C-00220.html					
-rw-rw-rw-	1	training	supergroup	1144	2022-02-16 19:31 /loudacre/kb/KBDO
C-00221.html					
-rw-rw-rw-	1	training	supergroup	1112	2022-02-16 19:31 /loudacre/kb/KBDO
C-00222.html					
-rw-rw-rw-	1	training	supergroup	1640	2022-02-16 19:31 /loudacre/kb/KBDO
C-00223.html					
-rw-rw-rw-	1	training	supergroup	1297	2022-02-16 19:31 /loudacre/kb/KBDO
C-00224.html					
-rw-rw-rw-	1	training	supergroup	1195	2022-02-16 19:31 /loudacre/kb/KBDO
C-00225.html					
-rw-rw-rw-	1	training	supergroup	1030	2022-02-16 19:31 /loudacre/kb/KBDO
C-00226.html					
-rw-rw-rw-	1	training	supergroup	1130	2022-02-16 19:31 /loudacre/kb/KBDO
C-00227.html					
-rw-rw-rw-	1	training	supergroup	1132	2022-02-16 19:31 /loudacre/kb/KBDO
C-00228.html					
-rw-rw-rw-	1	training	supergroup	1645	2022-02-16 19:31 /loudacre/kb/KBDO
C-00229.html					
-rw-rw-rw-	1	training	supergroup	1302	2022-02-16 19:31 /loudacre/kb/KBDO
C-00230.html					
-rw-rw-rw-	1	training	supergroup	1197	2022-02-16 19:31 /loudacre/kb/KBDO
C-00231.html					
-rw-rw-rw-	1	training	supergroup	1049	2022-02-16 19:31 /loudacre/kb/KBDO
C-00232.html					
-rw-rw-rw-	1	training	supergroup	1149	2022-02-16 19:31 /loudacre/kb/KBDO
C-00233.html					
-rw-rw-rw-	1	training	supergroup	1112	2022-02-16 19:31 /loudacre/kb/KBDO
C-00234.html					
-rw-rw-rw-	1	training	supergroup	1616	2022-02-16 19:31 /loudacre/kb/KBDO
C-00235.html					
-rw-rw-rw-	1	training	supergroup	1268	2022-02-16 19:31 /loudacre/kb/KBDO
C-00236.html					
-rw-rw-rw-	1	training	supergroup	1172	2022-02-16 19:31 /loudacre/kb/KBDO
C-00237.html					
-rw-rw-rw-	1	training	supergroup	996	2022-02-16 19:31 /loudacre/kb/KBDO
C-00238.html					
-rw-rw-rw-	1	training	supergroup	1110	2022-02-16 19:31 /loudacre/kb/KBDO
C-00239.html					
-rw-rw-rw-	1	training	supergroup	1093	2022-02-16 19:31 /loudacre/kb/KBDO
C-00240.html					
-rw-rw-rw-	1	training	supergroup	1627	2022-02-16 19:31 /loudacre/kb/KBDO
C-00241.html					

File Edit View Search Terminal Help					
-rw-rw-rw-	1	training	supergroup	1093	2022-02-16 19:31 /loudacre/kb/KBDO
C-00240.html				1627	2022-02-16 19:31 /loudacre/kb/KBDO
-rw-rw-rw-	1	training	supergroup	1297	2022-02-16 19:31 /loudacre/kb/KBDO
C-00241.html				1194	2022-02-16 19:31 /loudacre/kb/KBDO
-rw-rw-rw-	1	training	supergroup	1048	2022-02-16 19:31 /loudacre/kb/KBDO
C-00242.html				1150	2022-02-16 19:31 /loudacre/kb/KBDO
-rw-rw-rw-	1	training	supergroup	1137	2022-02-16 19:31 /loudacre/kb/KBDO
C-00243.html				1660	2022-02-16 19:31 /loudacre/kb/KBDO
-rw-rw-rw-	1	training	supergroup	1300	2022-02-16 19:31 /loudacre/kb/KBDO
C-00244.html				1187	2022-02-16 19:31 /loudacre/kb/KBDO
-rw-rw-rw-	1	training	supergroup	1030	2022-02-16 19:31 /loudacre/kb/KBDO
C-00245.html				1155	2022-02-16 19:31 /loudacre/kb/KBDO
-rw-rw-rw-	1	training	supergroup	1128	2022-02-16 19:31 /loudacre/kb/KBDO
C-00246.html				1641	2022-02-16 19:31 /loudacre/kb/KBDO
-rw-rw-rw-	1	training	supergroup	1309	2022-02-16 19:31 /loudacre/kb/KBDO
C-00247.html				1176	2022-02-16 19:31 /loudacre/kb/KBDO
-rw-rw-rw-	1	training	supergroup	1026	2022-02-16 19:31 /loudacre/kb/KBDO
C-00248.html				1146	2022-02-16 19:31 /loudacre/kb/KBDO
-rw-rw-rw-	1	training	supergroup	1128	2022-02-16 19:31 /loudacre/kb/KBDO
C-00249.html				1614	2022-02-16 19:31 /loudacre/kb/KBDO
-rw-rw-rw-	1	training	supergroup	1297	2022-02-16 19:31 /loudacre/kb/KBDO
C-00250.html				1212	2022-02-16 19:31 /loudacre/kb/KBDO
-rw-rw-rw-	1	training	supergroup		
C-00251.html					
-rw-rw-rw-	1	training	supergroup		
C-00252.html					
-rw-rw-rw-	1	training	supergroup		
C-00253.html					
-rw-rw-rw-	1	training	supergroup		
C-00254.html					
-rw-rw-rw-	1	training	supergroup		
C-00255.html					
-rw-rw-rw-	1	training	supergroup		
C-00256.html					
-rw-rw-rw-	1	training	supergroup		
C-00257.html					
-rw-rw-rw-	1	training	supergroup		
C-00258.html					
-rw-rw-rw-	1	training	supergroup		
C-00259.html					
-rw-rw-rw-	1	training	supergroup		
C-00260.html					
-rw-rw-rw-	1	training	supergroup		

training@localhost:~/training_materials/data					
File	Edit	View	Search	Terminal	Help
-rw-rw-rw- 1 training supergroup	1297	2022-02-16 19:31	/loudacre/kb/KB00		
C-00260.html					
-rw-rw-rw- 1 training supergroup	1212	2022-02-16 19:31	/loudacre/kb/KB00		
C-00261.html					
-rw-rw-rw- 1 training supergroup	1045	2022-02-16 19:31	/loudacre/kb/KB00		
C-00262.html					
-rw-rw-rw- 1 training supergroup	1165	2022-02-16 19:31	/loudacre/kb/KB00		
C-00263.html					
-rw-rw-rw- 1 training supergroup	1109	2022-02-16 19:31	/loudacre/kb/KB00		
C-00264.html					
-rw-rw-rw- 1 training supergroup	1639	2022-02-16 19:31	/loudacre/kb/KB00		
C-00265.html					
-rw-rw-rw- 1 training supergroup	1302	2022-02-16 19:31	/loudacre/kb/KB00		
C-00266.html					
-rw-rw-rw- 1 training supergroup	1179	2022-02-16 19:31	/loudacre/kb/KB00		
C-00267.html					
-rw-rw-rw- 1 training supergroup	1046	2022-02-16 19:31	/loudacre/kb/KB00		
C-00268.html					
-rw-rw-rw- 1 training supergroup	1132	2022-02-16 19:31	/loudacre/kb/KB00		
C-00269.html					
-rw-rw-rw- 1 training supergroup	1122	2022-02-16 19:31	/loudacre/kb/KB00		
C-00270.html					
-rw-rw-rw- 1 training supergroup	1618	2022-02-16 19:31	/loudacre/kb/KB00		
C-00271.html					
-rw-rw-rw- 1 training supergroup	1268	2022-02-16 19:31	/loudacre/kb/KB00		
C-00272.html					
-rw-rw-rw- 1 training supergroup	1151	2022-02-16 19:31	/loudacre/kb/KB00		
C-00273.html					
-rw-rw-rw- 1 training supergroup	1020	2022-02-16 19:31	/loudacre/kb/KB00		
C-00274.html					
-rw-rw-rw- 1 training supergroup	1105	2022-02-16 19:31	/loudacre/kb/KB00		
C-00275.html					
-rw-rw-rw- 1 training supergroup	1101	2022-02-16 19:31	/loudacre/kb/KB00		
C-00276.html					
-rw-rw-rw- 1 training supergroup	1604	2022-02-16 19:31	/loudacre/kb/KB00		
C-00277.html					
-rw-rw-rw- 1 training supergroup	1253	2022-02-16 19:31	/loudacre/kb/KB00		
C-00278.html					
-rw-rw-rw- 1 training supergroup	1169	2022-02-16 19:31	/loudacre/kb/KB00		
C-00279.html					
-rw-rw-rw- 1 training supergroup	1021	2022-02-16 19:31	/loudacre/kb/KB00		
C-00280.html					
-rw-rw-rw- 1 training supergroup	1135	2022-02-16 19:31	/loudacre/kb/KB00		

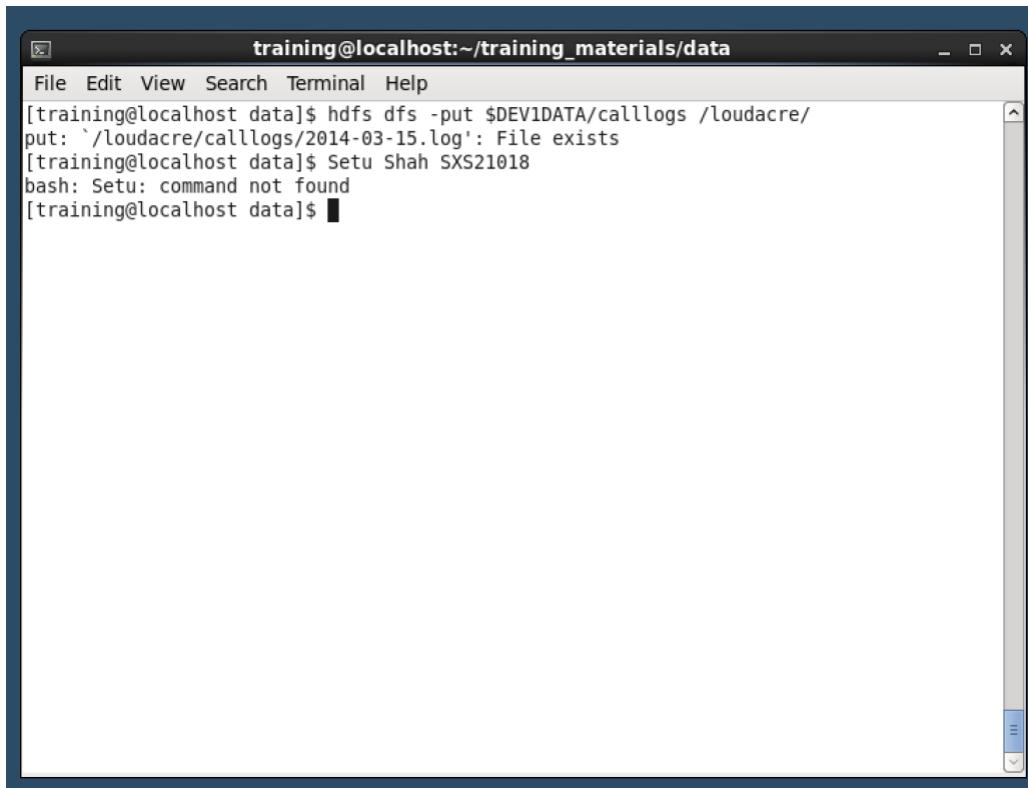
```
training@localhost:~/training_materials/data
File Edit View Search Terminal Help
-rw-rw-rw- 1 training supergroup 1021 2022-02-16 19:31 /loudacre/kb/KB00
C-00280.html
-rw-rw-rw- 1 training supergroup 1135 2022-02-16 19:31 /loudacre/kb/KB00
C-00281.html
-rw-rw-rw- 1 training supergroup 1115 2022-02-16 19:31 /loudacre/kb/KB00
C-00282.html
-rw-rw-rw- 1 training supergroup 1608 2022-02-16 19:31 /loudacre/kb/KB00
C-00283.html
-rw-rw-rw- 1 training supergroup 1263 2022-02-16 19:31 /loudacre/kb/KB00
C-00284.html
-rw-rw-rw- 1 training supergroup 1151 2022-02-16 19:31 /loudacre/kb/KB00
C-00285.html
-rw-rw-rw- 1 training supergroup 1017 2022-02-16 19:31 /loudacre/kb/KB00
C-00286.html
-rw-rw-rw- 1 training supergroup 1129 2022-02-16 19:31 /loudacre/kb/KB00
C-00287.html
-rw-rw-rw- 1 training supergroup 1111 2022-02-16 19:31 /loudacre/kb/KB00
C-00288.html
-rw-rw-rw- 1 training supergroup 1612 2022-02-16 19:31 /loudacre/kb/KB00
C-00289.html
-rw-rw-rw- 1 training supergroup 1253 2022-02-16 19:31 /loudacre/kb/KB00
C-00290.html
-rw-rw-rw- 1 training supergroup 1164 2022-02-16 19:31 /loudacre/kb/KB00
C-00291.html
-rw-rw-rw- 1 training supergroup 1041 2022-02-16 19:31 /loudacre/kb/KB00
C-00292.html
-rw-rw-rw- 1 training supergroup 1110 2022-02-16 19:31 /loudacre/kb/KB00
C-00293.html
-rw-rw-rw- 1 training supergroup 1097 2022-02-16 19:31 /loudacre/kb/KB00
C-00294.html
-rw-rw-rw- 1 training supergroup 1625 2022-02-16 19:31 /loudacre/kb/KB00
C-00295.html
-rw-rw-rw- 1 training supergroup 1270 2022-02-16 19:31 /loudacre/kb/KB00
C-00296.html
-rw-rw-rw- 1 training supergroup 1167 2022-02-16 19:31 /loudacre/kb/KB00
C-00297.html
-rw-rw-rw- 1 training supergroup 1015 2022-02-16 19:31 /loudacre/kb/KB00
C-00298.html
-rw-rw-rw- 1 training supergroup 1126 2022-02-16 19:31 /loudacre/kb/KB00
C-00299.html
-rw-rw-rw- 1 training supergroup 1097 2022-02-16 19:31 /loudacre/kb/KB00
C-00300.html
[training@localhost data]$ Setu Shah SXS210181
```

Using the command “hdfs dfs -put \$DEV1DATA/callogs /loudacre/” to add “callogs” to “loudacre” directory. We can find “callogs” in “loudacre” directory.



A terminal window titled "training@localhost:~/training\_materials/data". The window has a standard OS X style title bar with icons for close, minimize, and zoom. The menu bar includes "File", "Edit", "View", "Search", "Terminal", and "Help". The main pane displays the following command-line session:

```
[training@localhost data]$ hdfs dfs -put $DEV1DATA/callogs /loudacre/
[training@localhost data]$ Setu Shah SXS210181
[training@localhost data]$ Setu Shah SXS210181
```



A terminal window titled "training@localhost:~/training\_materials/data". The window has a standard OS X style title bar with icons for close, minimize, and zoom. The menu bar includes "File", "Edit", "View", "Search", "Terminal", and "Help". The main pane displays the following command-line session:

```
[training@localhost data]$ hdfs dfs -put $DEV1DATA/callogs /loudacre/
put: `/loudacre/callogs/2014-03-15.log': File exists
[training@localhost data]$ Setu Shah SXS21018
bash: Setu: command not found
[training@localhost data]$ █
```

The screenshot shows the Hue File Browser interface. The URL is `localhost:8888/filebrowser/#/loudacre`. The directory structure is as follows:

Name	Size	User	Group	Permissions	Date
hdfs		superuser	superuser	drwxr-xr-x	February 16, 2022 07:28 PM
.		training	superuser	drwxrwxrwx	February 16, 2022 08:16 PM
..		training	superuser	drwxrwxrwx	February 16, 2022 08:16 PM
callogs		training	superuser	drwxrwxrwx	February 16, 2022 07:31 PM
kb		training	superuser	drwxrwxrwx	February 16, 2022 07:31 PM

Show 45 of 2 items

Page 1 of 1

To remove “callogs” from “loudacre” directory using command, “`hdfs dfs -rm -r /loudacre/callogs`”, the directory is then removed.

```
[training@localhost data]$ hdfs dfs -rm -r /loudacre/callogs
22/02/16 20:25:43 INFO fs.TrashPolicyDefault: Namenode trash configuration: Deletion interval = 0 minutes, Emptier interval = 0 minutes.
Deleted /loudacre/callogs
[training@localhost data]$ Setu Shah SXS210181
```

We can see that the “callogs” has been removed from the “loudacre” directory.

The screenshot shows a Firefox browser window titled "Hue - File Browser - Mozilla Firefox". The address bar displays "localhost:8888/filebrowser/#/loudacre". The main content area is a file browser interface with the following details:

**File Browser**

Search for file name: \_\_\_\_\_ Actions: \_\_\_\_\_ Move to trash: \_\_\_\_\_ Upload: \_\_\_\_\_ New: \_\_\_\_\_

Home / loudacre History Trash

Name	Size	User	Group	Permissions	Date
hdbs		supergroup	supergroup	drwxr-xr-x	February 16, 2022 07:28 PM
.		training	supergroup	drwxrwxrwx	February 16, 2022 08:25 PM
kb		training	supergroup	drwxrwxrwx	February 16, 2022 07:31 PM

Show 45 of 1 items Page 1 of 1 Choose What I Share

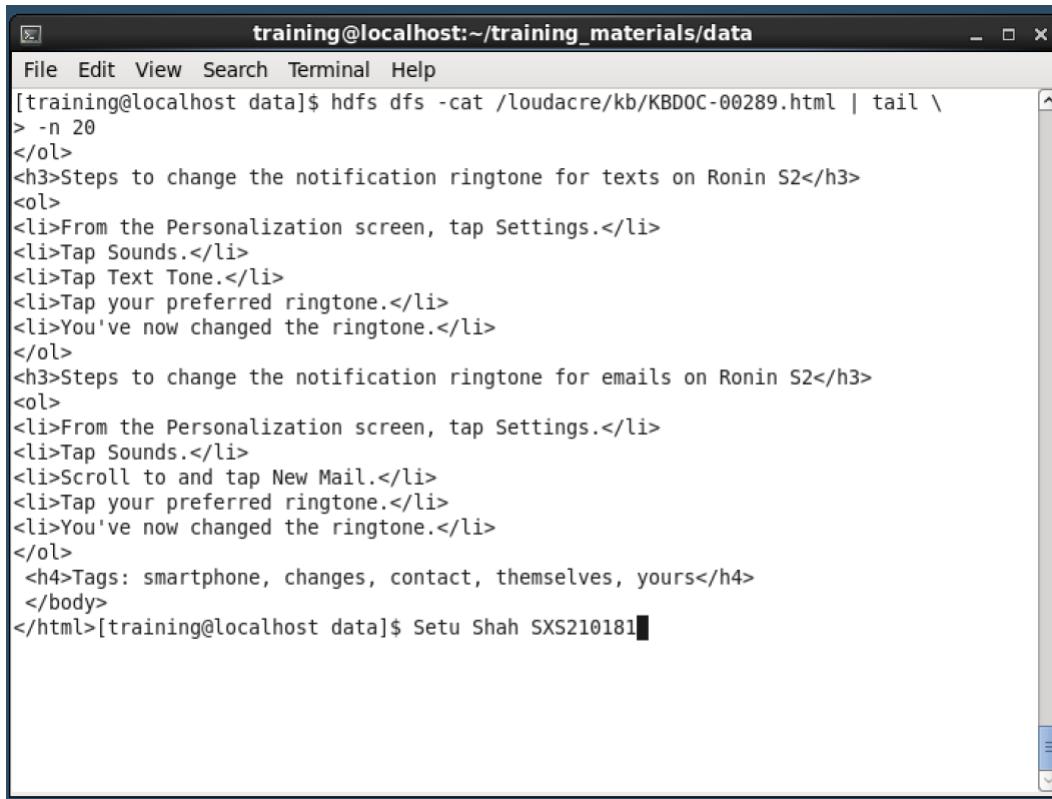
Firefox automatically sends some data to Mozilla so that we can improve your experience.

[data - File Browser] [training@localhost:~] [training@localhost:~/.] [Hue - File Browser - M...]

## **VIEW HDFS FILES**

Because of the huge size, it is inconvenient to see the entire file on terminal.

“hdfs dfs -cat /loudacre/kb/KBDOC-00289.html | tail -n 20”, was used to view the last 20 lines of the article. We can pipe the output of fs -cat command into head, tail, more or less.

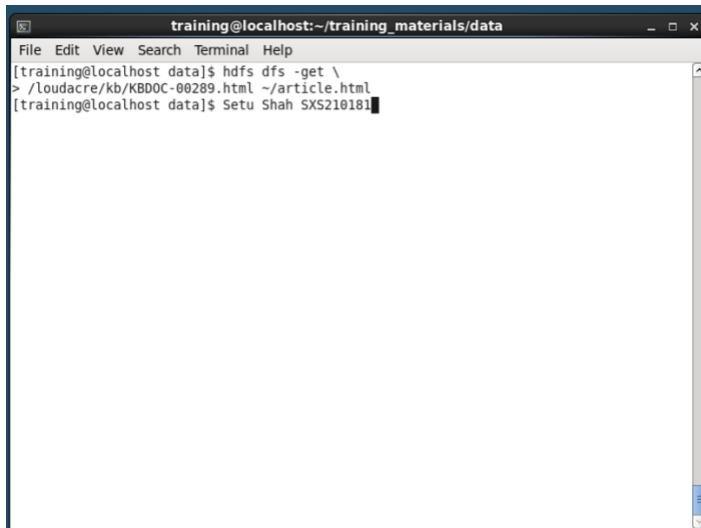


The screenshot shows a terminal window titled "training@localhost:~/training\_materials/data". The window contains the following command and its output:

```
[training@localhost data]$ hdfs dfs -cat /loudacre/kb/KBDOC-00289.html | tail \n> -n 20\n</ol>\n<h3>Steps to change the notification ringtone for texts on Ronin S2</h3>\n<ol>\n<li>From the Personalization screen, tap Settings.</li>\n<li>Tap Sounds.</li>\n<li>Tap Text Tone.</li>\n<li>Tap your preferred ringtone.</li>\n<li>You've now changed the ringtone.</li>\n</ol>\n<h3>Steps to change the notification ringtone for emails on Ronin S2</h3>\n<ol>\n<li>From the Personalization screen, tap Settings.</li>\n<li>Tap Sounds.</li>\n<li>Scroll to and tap New Mail.</li>\n<li>Tap your preferred ringtone.</li>\n<li>You've now changed the ringtone.</li>\n</ol>\n<h4>Tags: smartphone, changes, contact, themselves, yours</h4>\n</body>\n</html>[training@localhost data]$ Setu Shah SXS210181
```

Using the command “hdfs dfs -get” to retrieve a copy of file or directory from HDFS.

Two arguments: HDFS path and local path are taken by the command. “KBDOC-00289.html” will be copied into a local file “article.html”



The screenshot shows a terminal window titled "training@localhost:~/training\_materials/data". The window contains the following command and its output:

```
[training@localhost data]$ hdfs dfs -get \n> /loudacre/kb/KBDOC-00289.html ~article.html\n[training@localhost data]$ Setu Shah SXS210181
```

## Output of the file “less ~article.html”

```
training@localhost:~/training_materials/data
File Edit View Search Terminal Help
[training@localhost data]$ less ~/article.html
[training@localhost data]$ [training@localhost data]$
```

```
Applications Places System
[training@localhost:~/training_materials/data]
File Edit View Search Terminal Help
Wed Feb 16, 8:31 PM
<!DOCTYPE html PUBLIC "-//W3C//DTD XHTML 1.0 Strict//EN"
 "http://www.w3.org/TR/xhtml1/DTD/xhtml1-strict.dtd">
<html xml:lang="en" lang="en">
<head>
<meta http-equiv="Content-Type" content="text/html; charset=utf-8" />
<meta name="device" content="Ronin S2" />
<meta name="docid" content="KBDOC-00289" />
<meta name="createddate" content="2012-07-21" />
<meta name="updateddate" content="2012-08-03" />
<title>Ronin S2 - Change the phone ringtone and notification sound</title>
<link rel="stylesheet" type="text/css" href="theme.css" />
</head>
<body>
<h1>Ronin S2 - Change the phone ringtone and notification sound</h1>
<h2>Document Id: KBDOC-00289</h2>
<h2>Created: 2012-07-21</h2>
<h2>Updated: 2012-08-03</h2>
<h3>Steps to change the phone ringtone for calls on your Ronin S2</h3>
<ol>
<li>From the Personalization screen, tap Settings.</li>
<li>Tap Sounds.</li>
<li>Tap Ringtone.</li>
<li>Tap your preferred ringtone.</li>
<li>You've now changed the ringtone.</li>
</ol>
<h3>Steps to change the notification ringtone for texts on Ronin S2</h3>
<ol>
<li>From the Personalization screen, tap Settings.</li>
<li>Tap Sounds.</li>
<li>Tap Text Tone.</li>
<li>Tap your preferred ringtone.</li>
<li>You've now changed the ringtone.</li>
</ol>
<h3>Steps to change the notification ringtone for emails on Ronin S2</h3>
<ol>
<li>From the Personalization screen, tap Settings.</li>
<li>Tap Sounds.</li>
<li>Scroll to and tap New Mail.</li>
<li>Tap your preferred ringtone.</li>
<li>You've now changed the ringtone.</li>
</ol>
<h4>Tags: smartphone, changes, contact, themselves, yours</h4>
</body>
</html>
~
```

(END)

There are various other commands which can be used for HDFS data.  
 “hdfs dfs” gives the list of these commands such as append, cp, mv, etc.

```

File Edit View Search Terminal Help
[training@localhost data]$ hdfs dfs
Usage: hadoop fs [generic options]
  [-appendToFile <localsrc> ... <dst>]
  [-cat [<ignorecrc>] <src> ...]
  [-copyFromLocal <src> ...]
  [-copyToLocal [-p] [<ignorecrc>] <src> <localdst>]
  [-cp [<src> ...] <path> [<src> ...] <path>]
  [-cp [-f] [<src> ...] <dst>]
  [-createSnapshot <snapshotDir> [<snapshotName>]]
  [-deleteSnapshot <snapshotDir> [<snapshotName>]]
  [-d [<path> ...]]
  [-du [<path> ...]]
  [-get [<path> ...]]
  [-getfacl [-R] <path>]
  [-getfattr [-R] {-n name | -d} [-e en] <path>]
  [-getmerge [-nL] <src> <localdst>]
  [-help [cmd ...]]
  [-ls [<path> ...]]
  [-lsr [<path> ...]]
  [-moveFromLocal <localsrc> ... <dst>]
  [-moveToLocal <src> <localdst>]
  [-mv <src> ... <dst>]
  [-put [-f] [<src> ...] <localsrc> ... <dst>]
  [-renameSnapshot <snapshotDir> <oldName> <newName>]
  [-rm [<path> ...] [-R] [-m] [-p] [<ignorecrc>] <src> ...]
  [-rndr [-R] [<ignorefailonnonempty>] <dir> ...]
  [-setfacl [-R] [<file>|-k] [<acl spec>] <path>|[-set <acl_spec> <path>]
  [-setfattr (-n name [-v value] | -x name) <path>]
  [-setup [-R] [-w] <rep> <path> ...]
  [-stat [format] <path> ...]
  [-tail [-f] <file>]
  [-test <dfs> <path>]
  [-text [<ignorecrc>] <src> ...]
  [-touchz <path> ...]
  [-usage [cmd ...]]]

Generic options supported are
-conf <configuration file> specify an application configuration file
-D <property>value> use value for given property
-fs <local>namenode:port> specify a namenode
-jt <local>resourceManager:port> specify a ResourceManager

```

```

File Edit View Search Terminal Help
[training@localhost data]$ hdfs dfs
Usage: hadoop fs [generic options]
  [-cp [-f] [-p | -p[topax]] <src> ... <dst>]
  [-createSnapshot <snapshotDir> [<snapshotName>]]
  [-deleteSnapshot <snapshotDir> [<snapshotName>]]
  [-du [<path> ...]]
  [-du [<path> ...]]
  [-get [<path> ...]]
  [-getfacl [-R] <path>]
  [-getfattr [-R] {-n name | -d} [-e en] <path>]
  [-getmerge [-nL] <src> <localdst>]
  [-help [cmd ...]]
  [-ls [<path> ...]]
  [-lsr [<path> ...]]
  [-moveFromLocal <localsrc> ... <dst>]
  [-moveToLocal <src> <localdst>]
  [-mv <src> ... <dst>]
  [-put [-f] [<src> ...] <localsrc> ... <dst>]
  [-renameSnapshot <snapshotDir> <oldName> <newName>]
  [-rm [<path> ...] [-R] [-m] [-p] [<ignorecrc>] <src> ...]
  [-rndr [-R] [<ignorefailonnonempty>] <dir> ...]
  [-setfacl [-R] [<file>|-k] [<acl spec>] <path>|[-set <acl_spec> <path>]
  [-setfattr (-n name [-v value] | -x name) <path>]
  [-setup [-R] [-w] <rep> <path> ...]
  [-stat [format] <path> ...]
  [-tail [-f] <file>]
  [-test [<ignorecrc>] <src> ...]
  [-touchz <path> ...]
  [-usage [cmd ...]]]

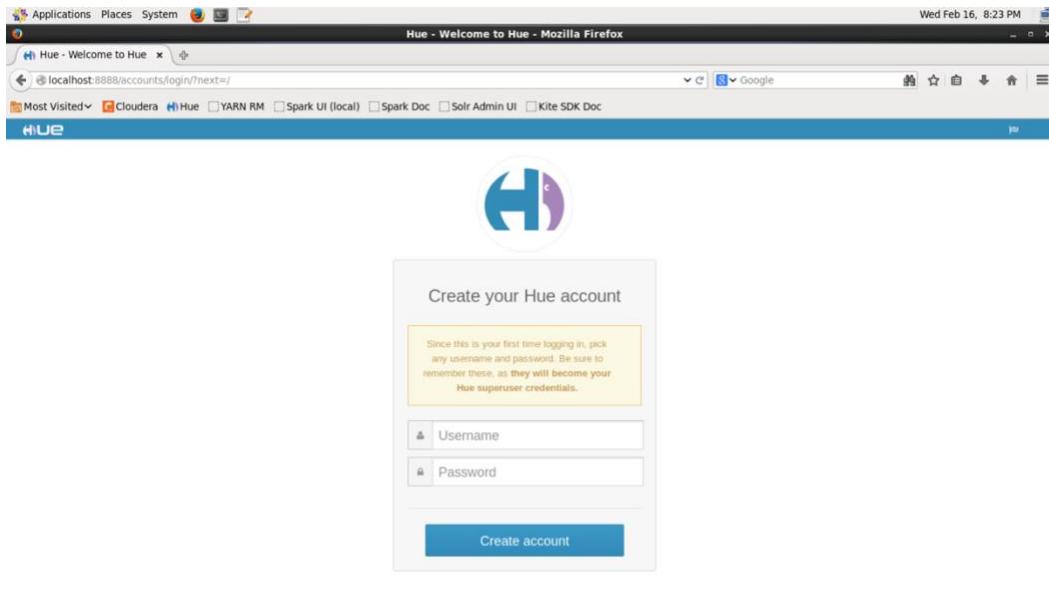
Generic options supported are
-conf <configuration file> specify an application configuration file
-D <property>value> use value for given property
-fs <local>namenode:port> specify a namenode
-jt <local>resourceManager:port> specify a ResourceManager
-files <comma separated list of files> specify comma separated files to be copied to the map reduce cluster
-libjars <comma separated list of jars> specify comma separated jar files to include in the classpath.
-archives <comma separated list of archives> specify comma separated archives to be unarchived on the compute machines.

The general command line syntax is
bin/hadoop command [genericOptions] [commandOptions]

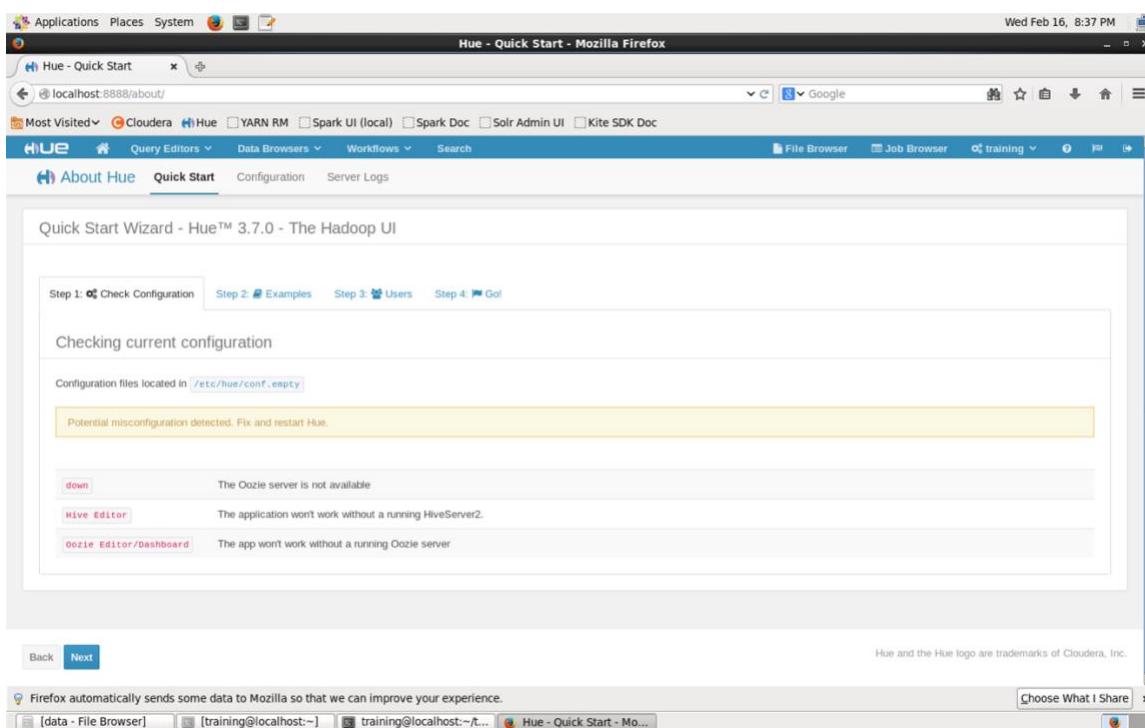
```

## **The Hue File Browser is used to browse, view and manage files**

I circumnavigated HUE page by tapping on bookmark tab and arrived up on the login page. The username and password is 'training'.



Started Mozilla Firefox on the virtual machine vmware to use the Hue File Browser.



Clicking on “File Browser” on menu bar to access the default HDFS home directory, i.e. “/user/training”.

The screenshot shows the Hue File Browser interface in Mozilla Firefox. The URL in the address bar is `localhost:8888/filebrowser/`. The browser title is "Hue - File Browser - Mozilla Firefox". The page displays the contents of the "/user/training" directory. The table has columns: Name, Size, User, Group, Permissions, and Date. There are two entries:

Name	Size	User	Group	Permissions	Date
hdfs		hdfs	supergroup	drwxrwxnwx	February 16, 2022 08:23 PM
training		training	supergroup	drwxrwxnwx	August 27, 2015 05:31 AM

At the bottom, it says "Show 45 of 0 items" and "Page 1 of 1".

Clicking on “/” will let us access the HDFS root directory.  
I can now see all the contents of root directory including “loudacre”.

The screenshot shows the Hue File Browser interface in Mozilla Firefox. The URL in the address bar is `localhost:8888/filebrowser/#/`. The browser title is "Hue - File Browser - Mozilla Firefox". The page displays the contents of the root directory. The table has columns: Name, Size, User, Group, Permissions, and Date. There are several entries:

Name	Size	User	Group	Permissions	Date
.		hdfs	supergroup	drwxr-xr-x	February 16, 2022 07:28 PM
hbase		hbase	supergroup	drwxrwxnwx	August 27, 2015 05:37 AM
loudacre		training	supergroup	drwxrwxnwx	February 16, 2022 08:25 PM
solr		solr	supergroup	drwxrwxnwx	August 27, 2015 05:32 AM
tmp		hdfs	supergroup	drwxrwxnwt	February 16, 2022 08:37 PM
user		hdfs	supergroup	drwxrwxnwx	February 16, 2022 08:23 PM
var		hdfs	supergroup	drwxrwxnwx	August 27, 2015 05:31 AM

At the bottom, it says "Show 45 of 6 items" and "Page 1 of 1".

Further when I click on “loudacre” directory to view it’s contents. We can see “kb” directory that we created earlier.

The screenshot shows the Hue File Browser interface in Mozilla Firefox. The URL is `localhost:8888/filebrowser/#/loudacre`. The browser toolbar includes Applications, Places, System, and a search bar for Google. The main navigation bar has links for Cloudera, Hue, YARN RM, Spark UI (local), Spark Doc, Solr Admin UI, and Kite SDK Doc. Below the navigation is a sub-navigation bar for File Browser, Job Browser, and training. The main content area is titled "File Browser" and shows a list of files and directories under "/loudacre". The "kb" directory is expanded, showing its contents. The table headers are Name, Size, User, Group, Permissions, and Date. The "kb" directory was created by user "training" on February 16, 2022, at 08:25 PM. It contains two files: ".html" and "KBDOC-00001.html". Other files listed are KBDOC-00002.html, KBDOC-00003.html, KBDOC-00004.html, KBDOC-00005.html, KBDOC-00006.html, KBDOC-00007.html, KBDOC-00008.html, KBDOC-00009.html, and KBDOC-00010.html. The permissions for these files are all `-rw-rw-rw-`.

Clicking on “kb” will allow us to access the contents of it. Furthermore, when I clicked on first file or in that case we can click on any random file, we can view the contents of it.

This screenshot shows the Hue File Browser interface in Mozilla Firefox, similar to the previous one but focusing on the "kb" directory under "/loudacre". The URL is `localhost:8888/filebrowser/#/loudacre/kb`. The browser and navigation bars are identical. The main content area shows a list of files under the "kb" directory. The table headers are Name, Size, User, Group, Permissions, and Date. The "kb" directory was created by user "training" on February 16, 2022, at 08:25 PM. It contains 10 files, each named KBDOC-00001.html through KBDOC-00010.html. All files are 1017 bytes in size and were created on February 16, 2022, at 07:31 PM. The permissions for all files are `-rw-rw-rw-`.

I clicked on first file “KBDOC-00001.html” to view contents of it. Below is the screenshot of the clicked file.

The screenshot shows the Mozilla Firefox browser window. The title bar reads "Hue - File Browser - KBDOC-00001.html - File Viewer - Mozilla Firefox". The address bar shows the URL "localhost:8888/filebrowser/view/loudacre/kb/KBDOC-00001.html". The page content is an HTML document with the following structure:

```

<!DOCTYPE html PUBLIC "-//W3C//DTD XHTML 1.0 Strict//EN"
"http://www.w3.org/TR/xhtml1/DTD/xhtml1-strict.dtd">
<html xml:lang="en" lang="en">
<head>
<meta http-equiv="Content-Type" content="text/html; charset=utf-8" />
<meta name="device" content="Sorrento F00L" />
<meta name="docid" content="KBDOC-00001" />
<meta name="createdate" content="2008-10-21" />
<meta name="updatedate" content="2008-10-24" />
<title>Sorrento F00L - Change the phone ringtone and notification sound</title>
<link rel="stylesheet" type="text/css" href="theme.css" />
</head>
<body>
<h1>Sorrento F00L - Change the phone ringtone and notification sound</h1>
<h2>Document Id: KBDOC-00001</h2>
<h2>Created: 2008-10-21</h2>
<h2>Updated: 2008-10-24</h2>
<h3>Steps to change the phone ringtone for calls on your Sorrento F00L</h3>
<ol>
<li>From the Settings screen, tap Personalization.</li>
<li>Tap Sounds.</li>
<li>Tap Ringtone.</li>
<li>Tap your preferred ringtone.</li>
<li>You've now changed the ringtone.</li>
</ol>

```

At the bottom of the page, there is a message from Firefox: "Firefox automatically sends some data to Mozilla so that we can improve your experience." and a "Choose What I Share" button.

The screenshot shows the Mozilla Firefox browser window. The title bar reads "Hue - File Browser - KBDOC-00001.html - File Viewer - Mozilla Firefox". The address bar shows the URL "localhost:8888/filebrowser/view/loudacre/kb/KBDOC-00001.html#p1". The page content is an HTML document with the following structure:

```

<li>Tap Ringtone.</li>
<li>Tap your preferred ringtone.</li>
<li>You've now changed the ringtone.</li>
</ol>
<h3>Steps to change the notification ringtone for texts on Sorrento F00L</h3>
<ol>
<li>From the Settings screen, tap Personalization.</li>
<li>Tap Sounds.</li>
<li>Tap Text Tone.</li>
<li>Tap your preferred ringtone.</li>
<li>You've now changed the ringtone.</li>
</ol>
<h3>Steps to change the notification ringtone for emails on Sorrento F00L</h3>
<ol>
<li>From the Settings screen, tap Personalization.</li>
<li>Tap Sounds.</li>
<li>Scroll to and tap New Mail.</li>
<li>Tap your preferred ringtone.</li>
<li>You've now changed the ringtone.</li>
</ol>
<h4>Tags: tablet, changing, contact, your</h4>

```

At the bottom of the page, there is a message from Firefox: "Firefox automatically sends some data to Mozilla so that we can improve your experience." and a "Choose What I Share" button.

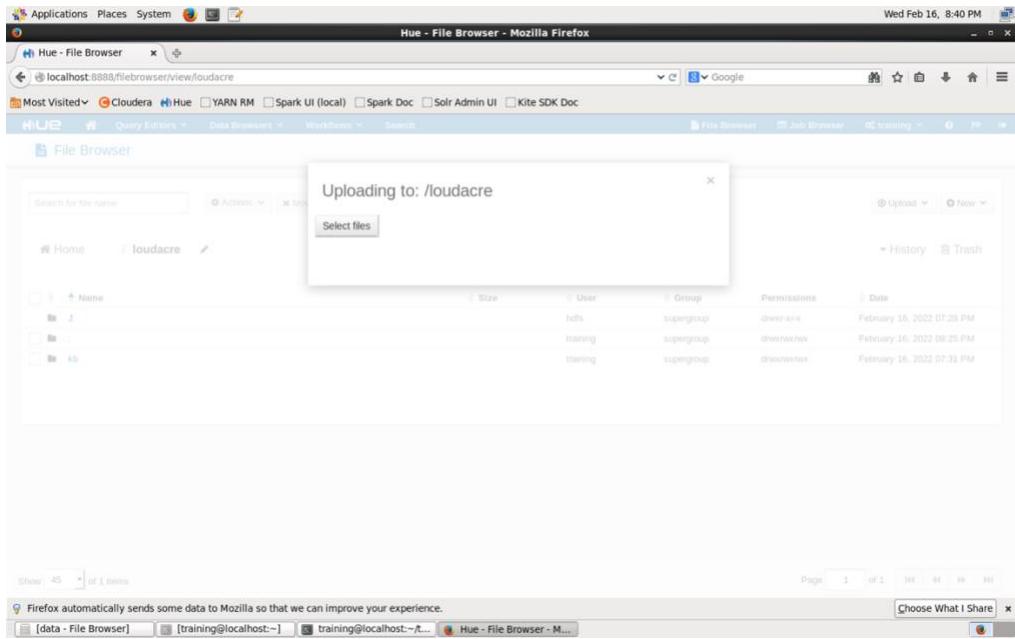
We can click on the “View File Location” on the right Actions panel and we can go back “/loudacre” directory.

Name	Size	User	Group	Permissions	Date
hdfs		hdfs	supergroup	drwxr-xr-x	February 16, 2022 07:28 PM
.		training	supergroup	drwxrwxrwx	February 16, 2022 08:25 PM
kb		training	supergroup	drwxrwxrwx	February 16, 2022 07:31 PM

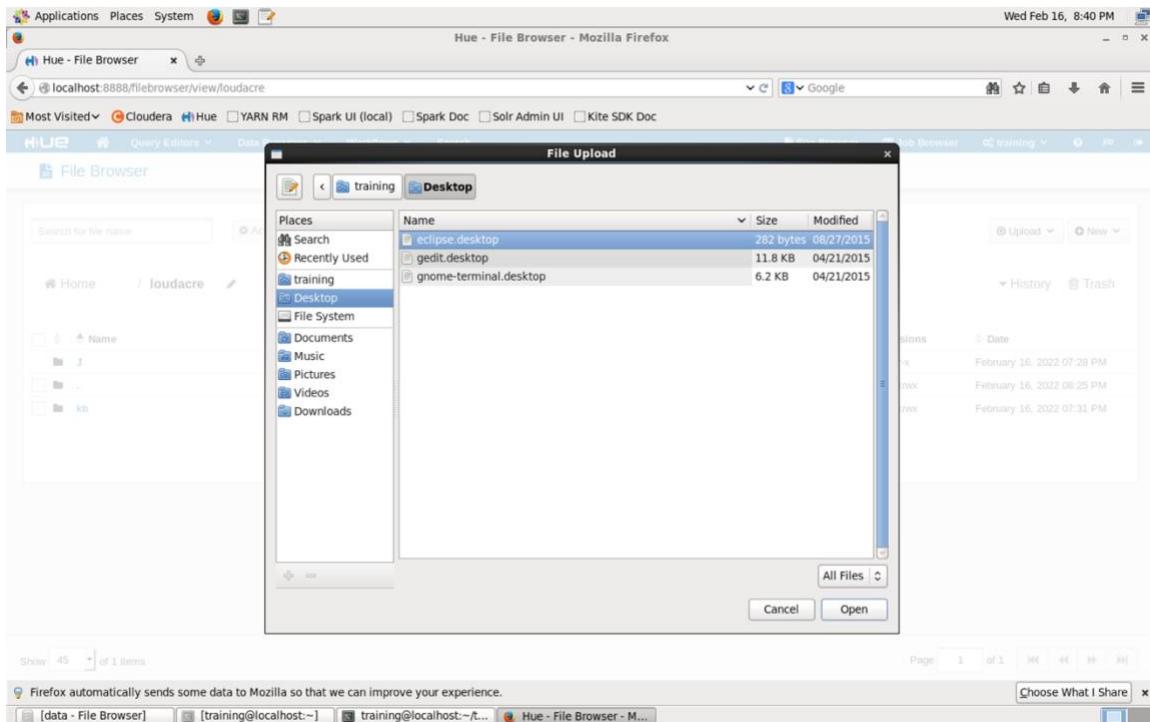
I uploaded a file using “Upload” button and then further clicked on “Files”. There are options to upload a plain file or zipped file.

After clicking on files, selected “Select Files” on the pop up

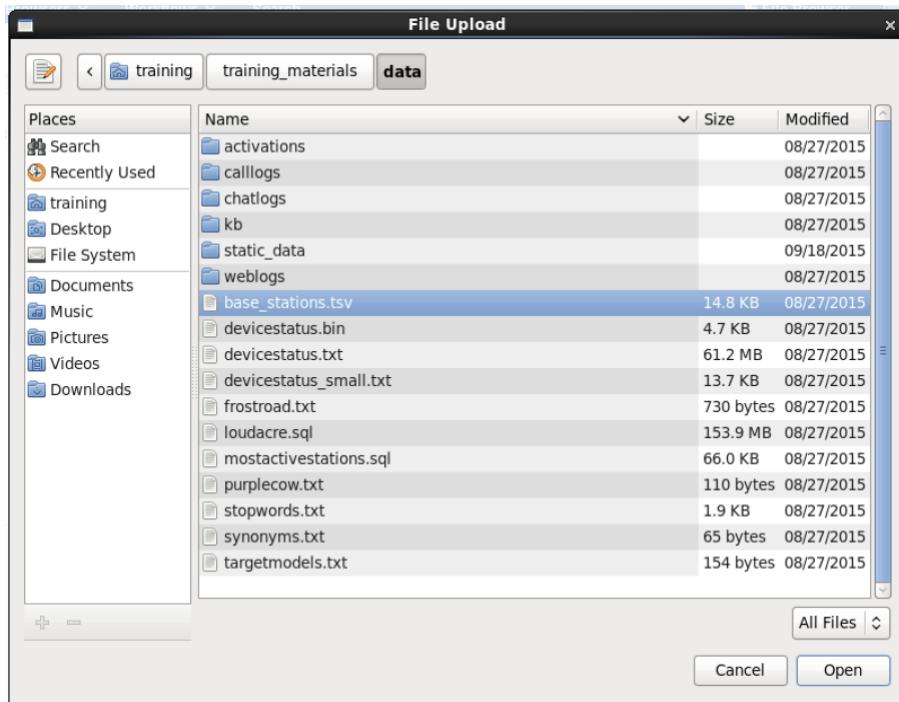
Name	Size	User	Group	Permissions	Date
hdfs		hdfs	supergroup	drwxr-xr-x	February 16, 2022 07:28 PM
.		training	supergroup	drwxrwxrwx	February 16, 2022 08:25 PM
kb		training	supergroup	drwxrwxrwx	February 16, 2022 07:31 PM



After clicking on “Select Files”, linux browser appeared.  
Directed to location “/home/training/Desktop”



Directed to location “/home/training/training\_materials/data” and selected “base\_stations.tsv”. Clicked on “Open” to upload the file.



The uploaded file is now available in “/loudacre”. Then, I clicked on checkbox next to file name and clicked on “Actions” to list the actions that I can perform on the file.

The screenshot shows the Hue - File Browser interface in Mozilla Firefox. The URL is `localhost:8888/filebrowser/view/loudacre?default_to_trash`. The browser toolbar shows the date as Wed Feb 16, 8:42 PM. The main window displays the contents of the 'loudacre' directory:

- Home / loudacre
- base\_stations.tsv (selected)
- kb

A context menu is open over the 'base\_stations.tsv' file, listing the following actions:
 

- Rename
- Move
- Copy
- Download
- Change permissions

The table below lists the files in the directory:

Name	Size	User	Group	Permissions	Date
base_stations.tsv	14.8 KB	training	supergroup	drwxr-xr-x	February 16, 2022 07:28 PM
kb		training	supergroup	-rw-rw-rw-	February 16, 2022 08:40 PM
loudacre.sql		training	supergroup	drwxrwxrwx	February 16, 2022 07:31 PM

At the bottom of the browser window, there is a message from Firefox: "Firefox automatically sends some data to Mozilla so that we can improve your experience." and a "Choose What I Share" button.

## Chapter 4 – YARN Job

### Explore the YARN cluster

Clicking on “YARN RM” on Firefox bookmark tab to visit YARN Resource Manager (RM). The Cluster Metrics displays the status of the application, memory utilization and nodes. When this step was in process there was no applications were running and hence apps running is 0.

The screenshot shows the 'All Applications' page of the YARN Resource Manager. The left sidebar has sections for Cluster (About, Nodes, Applications, Scheduler, Tools), and the main area shows 'Cluster Metrics' and 'User Metrics for dr.who'. The 'Cluster Metrics' table is as follows:

Apps Submitted	Apps Pending	Apps Running	Apps Completed	Containers Running	Memory Used	Memory Total	Memory Reserved	VCores Used	VCores Total	VCores Reserved	Active Nodes	Decommissioned Nodes	Lost Nodes	Unhealthy Nodes
0	0	0	0	0	0 B	2 GB	0 B	0	8	0	1	0	0	0

The 'User Metrics' table is as follows:

Apps Submitted	Apps Pending	Apps Running	Apps Completed	Containers Running	Containers Pending	Containers Reserved	Memory Used	Memory Pending	Memory Reserved	VCores Used	VCores Pending	VCores Reserved	VCores Total	VCores Used	VCores Pending
0	0	0	0	0	0	0	0 B	0 B	0 B	0	0	0	0	0	0

Below these tables, it says "No data available in table".

Next step was to click on the Nodes in Cluster menu.

The bottom section then displayed the list of worker nodes in the cluster. There is only 1 active node.

The screenshot shows the 'Nodes of the cluster' page of the YARN Resource Manager. The left sidebar has sections for Cluster (About, Nodes, Applications, Scheduler, Tools), and the main area shows 'Cluster Metrics' and 'User Metrics for dr.who'. The 'Cluster Metrics' table is identical to the one above. The 'User Metrics' table is as follows:

Nodes Labels	Rack	Node State	Node Address	Node HTTP Address	Last health-update	Health-report	Containers	Mem Used	Mem Avail	VCores Used	VCores Pending	VCores Reserved	VCores Total	VCores Used	VCores Pending
/default-rack	RUNNING	localhost:37932	localhost:8042	Wed Feb 16 20:42:52 -0800 2022			0	0 B	2 GB	0	0	8	8	24	cd

Below these tables, it says "Showing 1 to 1 of 1 entries".

Clicking on the “Node HTTP Address” to open the Node Manager UI of that node. All details about selected node, including amount of available memory, currently running applications, etc. are displayed.  
To return back to home, I clicked on “RM Home”.

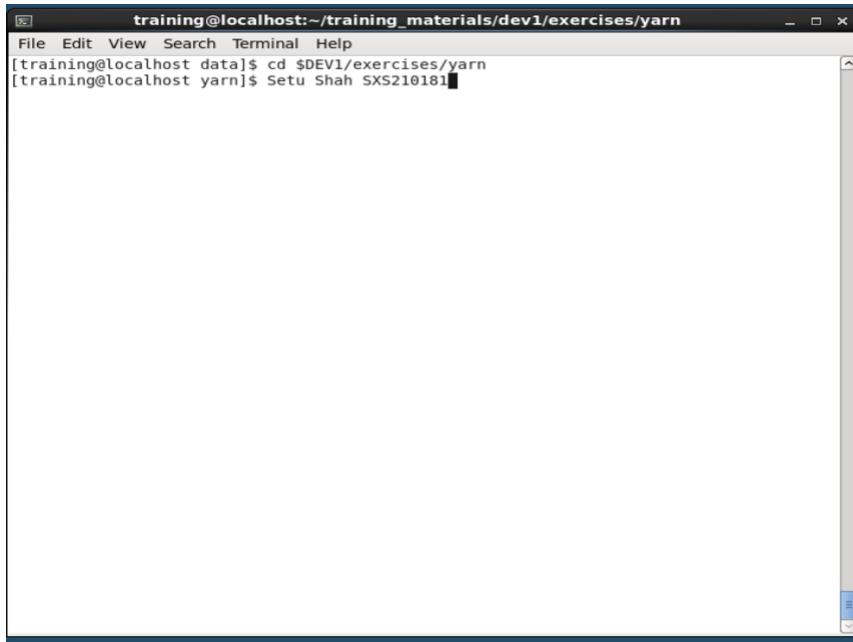
The screenshot shows a Mozilla Firefox browser window with the address bar set to `http://localhost:8042/node`. The title bar indicates it's running on Wednesday, February 16, at 8:46 PM. The main content area displays the Hadoop Node Manager interface. On the left, there's a sidebar with links like 'ResourceManager', 'NodeManager' (which is currently selected), 'Node Information', 'List of Applications', 'List of Containers', and 'Tools'. The right side shows 'NodeManager information' with the following details:

NodeManager information	
Total Vmem allocated for Containers	4.20 GB
Vmem enforcement enabled	false
Total Pmem allocated for Container	2 GB
Pmem enforcement enabled	true
Total VCores allocated for Containers	8
NodeHealthyStatus	true
LastNodeHealthTime	Wed Feb 16 20:44:52 PST 2022
NodeHealthReport	
Node Manager Version:	2.6.0-cdh5.4.3 from 4cd9f51a3f1ef748d45b8d77d0f211ad44296d4b by jenkins source checksum 8bacc650a3181bc0cfbae816dc81eddd on 2015-06-25T02:41Z
Hadoop Version:	2.6.0-cdh5.4.3 from 4cd9f51a3f1ef748d45b8d77d0f211ad44296d4b by jenkins source checksum 4acea6ac185376e0b48b33695e88e7a7 on 2015-06-25T02:34Z

At the bottom of the browser window, there's a message from Firefox: "Firefox automatically sends some data to Mozilla so that we can improve your experience." and a "Choose What I Share" button.

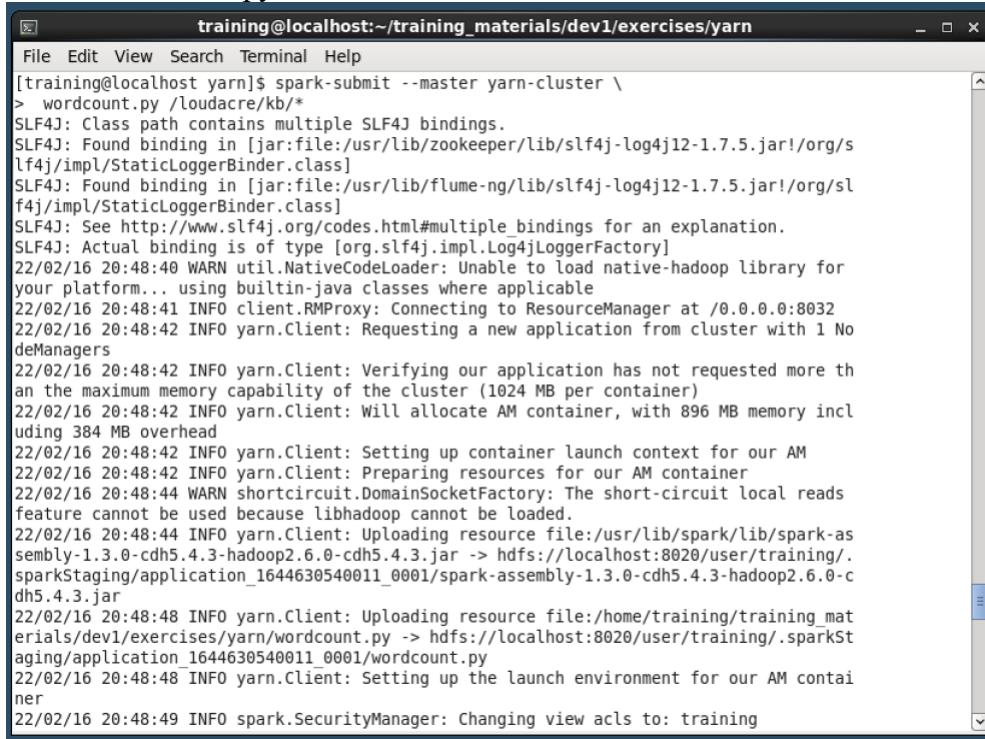
## **Submitting an application to the YARN cluster**

First step: To submit an application is to change the exercise directory. For that, I used command “cd \$DEV1/exercises/yarn” to achieve it.



```
training@localhost:~/training_materials/dev1/exercises/yarn
File Edit View Search Terminal Help
[training@localhost data]$ cd $DEV1/exercises/yarn
[training@localhost yarn]$ Setu Shah SXS210181
```

Second Step: I executed example wordcount.py program on the YARN cluster to count the frequency of words in the knowledge dataset. I used command “spark-submit --master yarn-cluster wordcount.py /loudacre/kb/\*” for execution.



```
training@localhost:~/training_materials/dev1/exercises/yarn
File Edit View Search Terminal Help
[training@localhost yarn]$ spark-submit --master yarn-cluster \
> wordcount.py /loudacre/kb/*
SLF4J: Class path contains multiple SLF4J bindings.
SLF4J: Found binding in [jar:file:/usr/lib/zookeeper/lib/slf4j-log4j12-1.7.5.jar!/org/slf4j/impl/StaticLoggerBinder.class]
SLF4J: Found binding in [jar:file:/usr/lib/flume-ng/lib/slf4j-log4j12-1.7.5.jar!/org/slf4j/impl/StaticLoggerBinder.class]
SLF4J: See http://www.slf4j.org/codes.html#multiple_bindings for an explanation.
SLF4J: Actual binding is of type [org.slf4j.impl.Log4jLoggerFactory]
22/02/16 20:48:40 WARN util.NativeCodeLoader: Unable to load native-hadoop library for your platform... using builtin-java classes where applicable
22/02/16 20:48:41 INFO client.RMProxy: Connecting to ResourceManager at /0.0.0.0:8032
22/02/16 20:48:42 INFO yarn.Client: Requesting a new application from cluster with 1 No deManagers
22/02/16 20:48:42 INFO yarn.Client: Verifying our application has not requested more than the maximum memory capability of the cluster (1024 MB per container)
22/02/16 20:48:42 INFO yarn.Client: Will allocate AM container, with 896 MB memory including 384 MB overhead
22/02/16 20:48:42 INFO yarn.Client: Setting up container launch context for our AM
22/02/16 20:48:42 INFO yarn.Client: Preparing resources for our AM container
22/02/16 20:48:44 WARN shortcircuit.DomainSocketFactory: The short-circuit local reads feature cannot be used because libhadoop cannot be loaded.
22/02/16 20:48:44 INFO yarn.Client: Uploading resource file:/usr/lib/spark/lib/spark-assembly-1.3.0-cdh5.4.3-hadoop2.6.0-cdh5.4.3.jar -> hdfs://localhost:8020/user/training/.sparkStaging/application_1644630540011_0001/spark-assembly-1.3.0-cdh5.4.3-hadoop2.6.0-cdh5.4.3.jar
22/02/16 20:48:48 INFO yarn.Client: Uploading resource file:/home/training/training_materials/dev1/exercises/yarn/wordcount.py -> hdfs://localhost:8020/user/training/.sparkStaging/application_1644630540011_0001/wordcount.py
22/02/16 20:48:48 INFO yarn.Client: Setting up the launch environment for our AM container
22/02/16 20:48:49 INFO spark.SecurityManager: Changing view acls to: training
```

```

File Edit View Search Terminal Help
user: training
22/02/16 20:49:09 INFO yarn.Client: Application report for application_1644630540011_001
01 (state: RUNNING)
22/02/16 20:49:10 INFO yarn.Client: Application report for application_1644630540011_001
01 (state: RUNNING)
22/02/16 20:49:11 INFO yarn.Client: Application report for application_1644630540011_001
01 (state: RUNNING)
22/02/16 20:49:12 INFO yarn.Client: Application report for application_1644630540011_001
01 (state: RUNNING)
22/02/16 20:49:13 INFO yarn.Client: Application report for application_1644630540011_001
01 (state: RUNNING)
22/02/16 20:49:14 INFO yarn.Client: Application report for application_1644630540011_001
01 (state: RUNNING)
22/02/16 20:49:15 INFO yarn.Client: Application report for application_1644630540011_001
01 (state: RUNNING)
22/02/16 20:49:16 INFO yarn.Client: Application report for application_1644630540011_001
01 (state: RUNNING)
22/02/16 20:49:17 INFO yarn.Client: Application report for application_1644630540011_001
01 (state: RUNNING)
22/02/16 20:49:18 INFO yarn.Client: Application report for application_1644630540011_001
01 (state: RUNNING)
22/02/16 20:49:19 INFO yarn.Client: Application report for application_1644630540011_001
01 (state: RUNNING)
22/02/16 20:49:20 INFO yarn.Client: Application report for application_1644630540011_001
01 (state: RUNNING)
22/02/16 20:49:21 INFO yarn.Client: Application report for application_1644630540011_001
01 (state: RUNNING)
22/02/16 20:49:22 INFO yarn.Client: Application report for application_1644630540011_001
01 (state: RUNNING)
22/02/16 20:49:23 INFO yarn.Client: Application report for application_1644630540011_001
01 (state: RUNNING)

```

## View the application in the Hue Job Browser

I clicked on Hue on the bookmark tab and selected, Job Browser from Hue Menu bar. We can view the currently running and recently completed applications by clicking on Job Browser.

Job Browser is displaying that application wordcount.py is currently running.

Logs	ID	Name	Status	User	Maps	Reduces	Queue	Priority	Duration	Submitted
	1644630540011_0001	wordcount.py	RUNNING	training	1	1	root.training	N/A	6m:4s	02/16/22 20:48:50

## View the application in the YARN UI

After reloading YARN RM in Firefox, we can see wordcount.py in the bottom section of RM homepage. We can get more detailed view of the cluster in YARN UI

Apps Submitted	Apps Pending	Apps Running	Apps Completed	Containers Running	Memory Used	Memory Total	Memory Reserved	VCores Used	VCores Total	VCores Reserved	Active Nodes	Decommissioned Nodes
1	0	1	0	2	2 GB	2 GB	1 GB	2	8	1	1	0

Apps Submitted	Apps Pending	Apps Running	Apps Completed	Containers Running	Containers Pending	Containers Reserved	Memory Used	Memory Pending	Memory Reserved	VCores Used	VCores Pending	VCores Reserved
0	0	1	0	0	0	0	0 B	0 B	0 B	0	0	0

ID	User	Name	Application Type	Queue	StartTime	FinishTime	State	FinalStatus	Running Containers	Allocated CPU vcores	Allocated Mem MB
application_1644630540011_0001	training	wordcount.py	SPARK	root.training	Wed Feb 16 20:48:50 -0800 2022	N/A	RUNNING	UNDEFINED	2	2	2048

Next, I clicked on Nodes to see the active nodes and selected nodes HTTP address to view statistics of node in the Node Manager UI.

Apps Submitted	Apps Pending	Apps Running	Apps Completed	Containers Running	Memory Used	Memory Total	Memory Reserved	VCores Used	VCores Total	VCores Reserved	Active Nodes	Decommissioned Nodes	Lost Nodes	Unhealthy Nodes
1	0	1	0	2	2 GB	2 GB	1 GB	2	8	1	1	0	0	0

Node Labels	Rack	Node State	Node Address	Node HTTP Address	Last health-update	Health-report	Containers	Mem Used	Mem Avail	VCores Used	VCores Pending	VCores Available
/default-rack	RUNNING	localhost:37932	localhost:8042	Wed Feb 16 20:56:52 -0800 2022			2	2 GB	0 B	2	6	6

To view, We can view list of applications running in the specified node by clicking on the “List of Applications” on the left side panel.

Applications running on this node

ApplicationId	ApplicationState
application_1644630540011_0001	RUNNING

Showing 1 to 1 of 1 entries

To view what the containers Resource Manager has allocated on the specified selected node for current application, we can do that by clicking on List of Container on the left panel.

Application's information

Container ID	Application ID	User
container_1644630540011_0001_01_000001	application_1644630540011_0001	training
container_1644630540011_0001_01_000002		

Show 20 entries

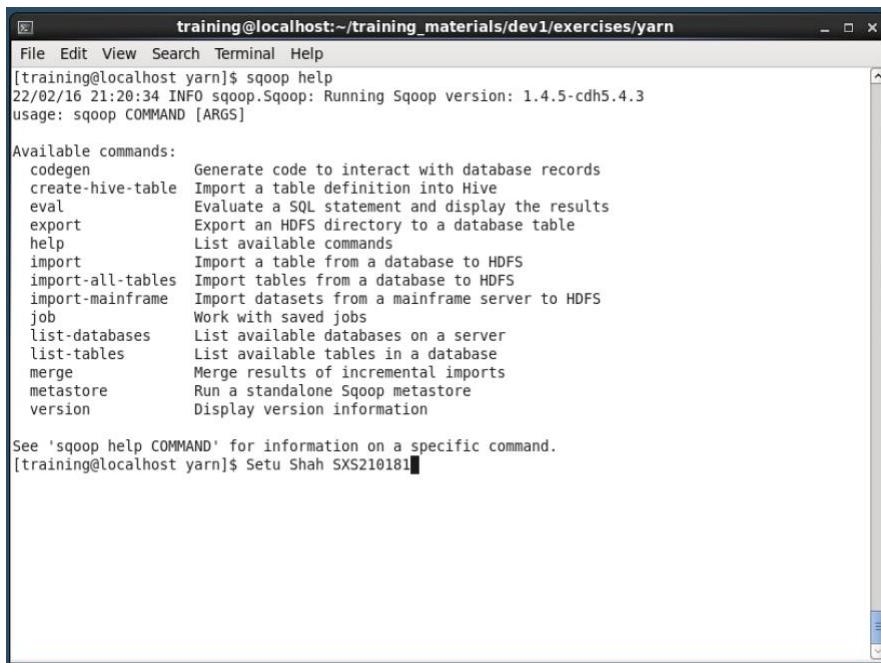
container\_1644630540011\_0001\_01\_000001  
container\_1644630540011\_0001\_01\_000002

Showing 1 to 2 of 2 entries

## Chapter 5 Sqoop

### Import the accounts table from MySQL

To be able to get quickly familiarized with sqoop, I used “sqoop help”. We can see a list of commands in sqoop such as “import”, “merge”, etc.

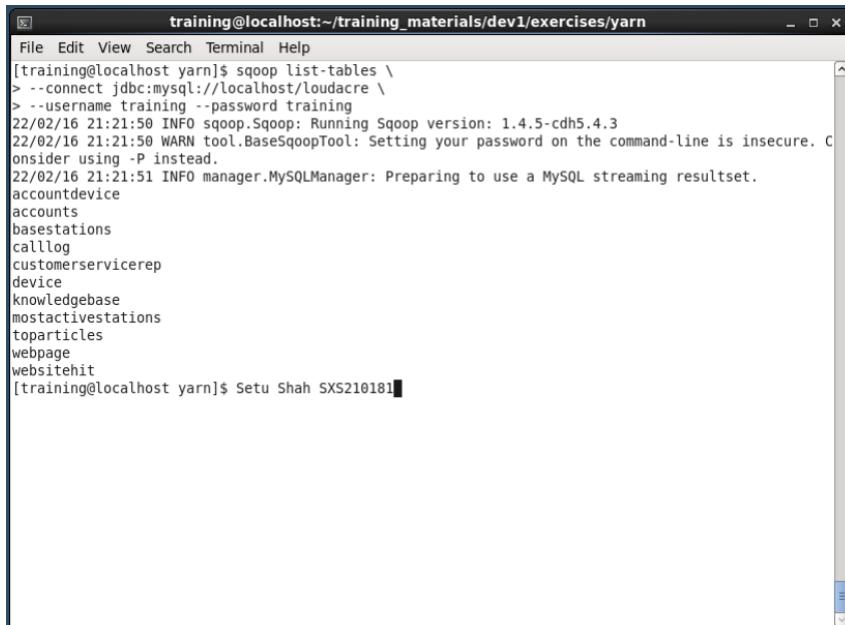


```
training@localhost:~/training_materials/dev1/exercises/yarn
File Edit View Search Terminal Help
[training@localhost yarn]$ sqoop help
22/02/16 21:20:34 INFO sqoop.Sqoop: Running Sqoop version: 1.4.5-cdh5.4.3
usage: sqoop COMMAND [ARGS]

Available commands:
codegen      Generate code to interact with database records
create-hive-table Import a table definition into Hive
eval         Evaluate a SQL statement and display the results
export       Export an HDFS directory to a database table
help         List available commands
import        Import a table from a database to HDFS
import-all-tables Import tables from a database to HDFS
import-mainframe Import datasets from a mainframe server to HDFS
job          Work with saved jobs
list-databases List available databases on a server
list-tables   List available tables in a database
merge        Merge results of incremental imports
metastore    Run a standalone Sqoop metastore
version      Display version information

See 'sqoop help COMMAND' for information on a specific command.
[training@localhost yarn]$ Setu Shah SXS210181
```

Then, I further checked the list of tables in loudacre database.



```
training@localhost:~/training_materials/dev1/exercises/yarn
File Edit View Search Terminal Help
[training@localhost yarn]$ sqoop list-tables \
> --connect jdbc:mysql://localhost/loudacre \
> --username training --password training
22/02/16 21:21:50 INFO sqoop.Sqoop: Running Sqoop version: 1.4.5-cdh5.4.3
22/02/16 21:21:50 WARN tool.BaseSqoopTool: Setting your password on the command-line is insecure. Consider using -P instead.
22/02/16 21:21:51 INFO manager.MySQLManager: Preparing to use a MySQL streaming resultset.
accountdevice
accounts
basestations
calllog
customerservicerep
device
knowledgebase
mostactivesstations
toparticles
webpage
websitehit
[training@localhost yarn]$ Setu Shah SXS210181
```

Executed “sqoop import -help” to view more options in importing the table.

```
training@localhost:~/training_materials/dev1/exercises/yarn
File Edit View Search Terminal Help
[training@localhost yarn]$ sqoop import --help
22/02/16 21:23:17 INFO sqoop.Sqoop: Running Sqoop version: 1.4.5-cdh5.4.3
usage: sqoop import [GENERIC-ARGS] [TOOL-ARGS]

Common arguments:
  --connect <jdbc-uri>          Specify JDBC connect
                                 string
  --connection-manager <class-name>  Specify connection manager
                                 class name
  --connection-param-file <properties-file>  Specify connection
                                 parameters file
  --driver <class-name>          Manually specify JDBC
                                 driver class to use
  --hadoop-home <hdir>          Override
                                 $HADOOP_MAPRED_HOME_ARG
  --hadoop-mapred-home <dir>    Override
                                 $HADOOP_MAPRED_HOME_ARG
  --help                         Print usage instructions
  -p                            Read password from console
  --password <password>          Set authentication
                                 password
  --password-alias <password-alias> Credential provider
                                 password alias
  --password-file <password-file> Set authentication
                                 password file path
  --relaxed-isolation           Use read-uncommitted
                                 isolation for imports
  --skip-dist-cache             Skip copying jars to
                                 distributed cache
  --username <username>          Set authentication
                                 username
  --verbose                      Print more information
                                 while working

Import control arguments:
  --append                       Imports data
                                 in append
                                 mode
  --as-avrodatafile              Imports data
                                 to Avro data
                                 files
  --as-parquetfile               Imports data
```

```
training@localhost:~/training_materials/dev1/exercises/yarn
File Edit View Search Terminal Help
Import control arguments:
  --append                       Imports data
                                 in append
                                 mode
  --as-avrodatafile              Imports data
                                 to Avro data
                                 files
  --as-parquetfile               Imports data
                                 to Parquet
                                 files
  --as-sequencefile              Imports data
                                 to
                                 SequenceFile
                                 s
  --as-textfile                  Imports data
                                 as plain
                                 text
                                 (default)
  --boundary-query <statement>  Set boundary
                                 query for
                                 retrieving
                                 max and min
                                 value of the
                                 primary key
  --columns <col,col,col...>    Columns to
                                 import from
                                 table
  --compression-codec <codec>    Compression
                                 codec to use
                                 for import
  --delete-target-dir            Imports data
                                 in delete
                                 mode
  --direct                        Use direct
                                 import fast
                                 path
  --direct-split-size <n>        Split the
                                 input stream
                                 every 'n'
                                 bytes when
                                 importing in
                                 direct mode
```

```

training@localhost:~/training_materials/dev1/exercises/yarn
File Edit View Search Terminal Help

Incremental import arguments:
--check-column <column>           Source column to check for incremental
                                    change
--incremental <import-type>        Define an incremental import of type
                                    'append' or 'lastmodified'
--last-value <value>               Last imported value in the incremental
                                    check column

Output line formatting arguments:
--enclosed-by <char>              Sets a required field enclosing
                                    character
--escaped-by <char>                Sets the escape character
--fields-separated-by <char>       Sets the field separator character
--lines-separated-by <char>        Sets the end-of-line character
--mysql-delimiters                 Uses MySQL's default delimiter set:
                                    fields: , lines: \n escaped-by: \
                                    optionally-enclosed-by: '
--optionally-enclosed-by <char>   Sets a field enclosing character

Input parsing arguments:
--input-enclosed-by <char>         Sets a required field encloser
--input-escaped-by <char>           Sets the input escape
                                    character
--input-fields-separated-by <char> Sets the input field separator
--input-lines-separated-by <char>  Sets the input end-of-line
                                    char
--input-optionally-enclosed-by <char> Sets a field enclosing
                                    character

Hive arguments:
--create-hive-table                Fail if the target hive
                                    table exists
--hive-database <database-name>   Sets the database name to
                                    use when importing to hive
--hive-delims-replacement <arg>  Replace Hive record \0x01
                                    and row delimiters (\n\r)
                                    from imported string fields
                                    with user-defined string
--hive-drop-import-delims          Drop Hive record \0x01 and
                                    row delimiters (\n\r) from
                                    imported string fields

```

```

training@localhost:~/training_materials/dev1/exercises/yarn
File Edit View Search Terminal Help

objects
--class-name <name>               Sets the generated class name.
                                    This overrides --package-name.
                                    When combined with --jar-file,
                                    sets the input class.
--input-null-non-string <null-str> Input null non-string
                                    representation
--input-null-string <null-str>    Input null string representation
--jar-file <file>                 Disable code generation; use
                                    specified jar
--map-column-java <arg>           Override mapping for specific
                                    columns to java types
--null-non-string <null-str>     Null non-string representation
--null-string <null-str>          Null string representation
--outdir <dir>                   Output directory for generated
                                    code
--package-name <name>             Put auto-generated classes in
                                    this package

Generic Hadoop command-line arguments:
(must precede any tool-specific arguments)
Generic options supported are
-conf <configuration file>      specify an application configuration file
-D <property>=<value>           use value for given property
-fs <local|namenode>:<port>    specify a namenode
-jt <local|resourcemanager>:<port> specify a ResourceManager
-files <comma separated list of files>  specify comma separated files to be copied to the map reduce cluster
-libjars <comma separated list of jars>  specify comma separated jar files to include in the classpath.
-archives <comma separated list of archives>  specify comma separated archives to be unarchived on the compute machines.

The general command line syntax is
bin/hadoop command [genericOptions] [commandOptions]

At minimum, you must specify --connect and --table
Arguments to mysqldump and other subprograms may be supplied
after a '---' on the command line.

[training@localhost yarn]$ Setu Shah SXS210181

```

Then, I imported the accounts table in the loudacre database and saved it in HDFS under /loudacre.

```
training@localhost:~/training_materials/dev1/exercises/yarn
File Edit View Search Terminal Help
[training@localhost yarn]$ sqoop import \
> --connect jdbc:mysql://localhost/loudacre \
> --username training --password training \
> --table accounts \
> --target-dir /loudacre/accounts \
> --null-non-string '\\N'
22/02/16 21:27:05 INFO sqoop.Sqoop: Running Sqoop version: 1.4.5-cdh5.4.3
22/02/16 21:27:05 WARN tool.BaseSqoopTool: Setting your password on the command-line is insecure. Consider using -P instead.
22/02/16 21:27:06 INFO manager.MySQLManager: Preparing to use a MySQL streaming resultset.
22/02/16 21:27:06 INFO tool.CodeGenTool: Beginning code generation
22/02/16 21:27:07 INFO manager.SqlManager: Executing SQL statement: SELECT t.* FROM `accounts` AS t
  LIMIT 1
22/02/16 21:27:07 INFO manager.SqlManager: Executing SQL statement: SELECT t.* FROM `accounts` AS t
  LIMIT 1
22/02/16 21:27:07 INFO orm.CompilationManager: HADOOP_MAPRED_HOME is /usr/lib/hadoop-mapreduce
Note: /tmp/sqoop-training/compile/2bc4af65bbb82bc2519c6f2cade55766/accounts.java uses or overrides a deprecated API.
Note: Recompile with -Xlint:deprecation for details.
22/02/16 21:27:11 INFO orm.CompilationManager: Writing jar file: /tmp/sqoop-training/compile/2bc4af65bbb82bc2519c6f2cade55766/accounts.jar
22/02/16 21:27:12 WARN manager.MySQLManager: It looks like you are importing from mysql.
22/02/16 21:27:12 WARN manager.MySQLManager: This transfer can be faster! Use the --direct
22/02/16 21:27:12 WARN manager.MySQLManager: option to exercise a MySQL-specific fast path.
22/02/16 21:27:12 INFO manager.MySQLManager: Setting zero DATETIME behavior to convertToNull (mysql)
22/02/16 21:27:12 INFO mapreduce.ImportJobBase: Beginning import of accounts
22/02/16 21:27:12 INFO Configuration.deprecation: mapred.job.tracker is deprecated. Instead, use mapreduce.jobtracker.address
22/02/16 21:27:14 INFO Configuration.deprecation: mapred.jar is deprecated. Instead, use mapreduce.job.jar
22/02/16 21:27:15 INFO Configuration.deprecation: mapred.map.tasks is deprecated. Instead, use mapreduce.job.maps
22/02/16 21:27:16 INFO client.RMProxy: Connecting to ResourceManager at /0.0.0.0:8032
22/02/16 21:27:19 INFO db.DBInputFormat: Using read committed transaction isolation
22/02/16 21:27:19 INFO db.DataDrivenDBInputFormat: BoundingValsQuery: SELECT MIN(`acct_num`), MAX(`acct_num`) FROM `accounts`
22/02/16 21:27:19 INFO mapreduce.JobSubmitter: number of splits:4
22/02/16 21:27:20 INFO mapreduce.JobSubmitter: Submitting tokens for job: job_1644630540011_0002
22/02/16 21:27:20 INFO impl.YarnClientImpl: Submitted application application_1644630540011_0002
22/02/16 21:27:20 INFO mapreduce.Job: The url to track the job: http://localhost:8088/proxy/application_1644630540011_0002/
```

```
training@localhost:~/training_materials/dev1/exercises/yarn
File Edit View Search Terminal Help
22/02/16 21:27:20 INFO impl.YarnClientImpl: Submitted application application_1644630540011_0002
22/02/16 21:27:20 INFO mapreduce.Job: The url to track the job: http://localhost:8088/proxy/application_1644630540011_0002/
22/02/16 21:27:20 INFO mapreduce.Job: Running job: job_1644630540011_0002
22/02/16 21:27:35 INFO mapreduce.Job: Job job_1644630540011_0002 running in uber mode : false
22/02/16 21:27:35 INFO mapreduce.Job: map 0% reduce 0%
22/02/16 21:27:55 INFO mapreduce.Job: map 25% reduce 0%
22/02/16 21:28:09 INFO mapreduce.Job: map 50% reduce 0%
22/02/16 21:28:18 INFO mapreduce.Job: map 75% reduce 0%
22/02/16 21:28:33 INFO mapreduce.Job: map 100% reduce 0%
22/02/16 21:28:34 INFO mapreduce.Job: Job job_1644630540011_0002 completed successfully
22/02/16 21:28:35 INFO mapreduce.Job: Counters: 30
  File System Counters
    FILE: Number of bytes read=0
    FILE: Number of bytes written=544560
    FILE: Number of read operations=0
    FILE: Number of large read operations=0
    FILE: Number of write operations=0
    HDFS: Number of bytes read=470
    HDFS: Number of bytes written=18737322
    HDFS: Number of read operations=16
    HDFS: Number of large read operations=0
    HDFS: Number of write operations=8
  Job Counters
    Launched map tasks=4
    Other local map tasks=4
    Total time spent by all maps in occupied slots (ms)=0
    Total time spent by all reduces in occupied slots (ms)=0
    Total time spent by all map tasks (ms)=52661
    Total vcore-seconds taken by all map tasks=52661
    Total megabyte-seconds taken by all map tasks=13481216
  Map-Reduce Framework
    Map input records=129761
    Map output records=129761
    Input split bytes=470
    Spilled Records=0
    Failed Shuffles=0
    Merged Map outputs=0
    GC time elapsed (ms)=2288
    CPU time spent (ms)=18090
    Physical memory (bytes) snapshot=384864256
    Virtual memory (bytes) snapshot=3382542336
```

```

training@localhost:~/training_materials/dev1/exercises/yarn
File Edit View Search Terminal Help
22/02/16 21:28:33 INFO mapreduce.Job: map 100% reduce 0%
22/02/16 21:28:34 INFO mapreduce.Job: Job job_1644630540011_0002 completed successfully
22/02/16 21:28:35 INFO mapreduce.Job: Counters: 30
    File System Counters
        FILE: Number of bytes read=0
        FILE: Number of bytes written=544560
        FILE: Number of read operations=0
        FILE: Number of large read operations=0
        FILE: Number of write operations=0
        HDFS: Number of bytes read=470
        HDFS: Number of bytes written=18737322
        HDFS: Number of read operations=16
        HDFS: Number of large read operations=0
        HDFS: Number of write operations=8
    Job Counters
        Launched map tasks=4
        Other local map tasks=4
        Total time spent by all maps in occupied slots (ms)=0
        Total time spent by all reduces in occupied slots (ms)=0
        Total time spent by all map tasks (ms)=52661
        Total vcore-seconds taken by all map tasks=52661
        Total megabyte-seconds taken by all map tasks=13481216
    Map-Reduce Framework
        Map input records=129761
        Map output records=129761
        Input split bytes=470
        Spilled Records=0
        Failed Shuffles=0
        Merged Map outputs=0
        GC time elapsed (ms)=2288
        CPU time spent (ms)=18090
        Physical memory (bytes) snapshot=384864256
        Virtual memory (bytes) snapshot=3382542336
        Total committed heap usage (bytes)=74940416
    File Input Format Counters
        Bytes Read=0
    File Output Format Counters
        Bytes Written=18737322
22/02/16 21:28:35 INFO mapreduce.ImportJobBase: Transferred 17.8693 MB in 79.6724 seconds (229.6677 KB/sec)
22/02/16 21:28:35 INFO mapreduce.ImportJobBase: Retrieved 129761 records.
[training@localhost yarn]$ Setu Shah SXS210181

```

We can view the sqoop job in Hue Job Browser and YARN Web UI. The below screenshot displays sqoop job in Hue Job Browser.

The screenshot shows the Hue Job Browser interface. At the top, there's a navigation bar with links like 'Hue - Job Browser', 'localhost:8888/jobbrowser', and 'Most Visited'. Below the navigation is a search bar with 'Username: training' and a 'Search for text' input field. A toolbar above the main table includes buttons for 'Successful', 'Running', 'Failed', and 'Killed'. The main area displays a table of job entries:

Log#	ID	Name	Status	User	Maps	Reduces	Queue	Priority	Duration	Submitted
	1644630540011_0002	Accounts.jar	SUCCEEDED	training	100%	100%	root.training	N/A	1m:12s	02/16/22 21:27:20
	1644630540011_0003	wontcount.py	SUCCEEDED	training	100%	100%	root.training	N/A	15m:32s	02/16/22 20:48:50

At the bottom of the table, it says 'Showing 1 to 2 of 2 entries'. The browser's status bar at the bottom indicates 'Firefox automatically sends some data to Mozilla so that we can improve your experience.'

The below screenshot is of YARN UI browser.  
The completed job is displayed in the last section.

**All Applications - Mozilla Firefox**

Wed Feb 16, 9:34 PM

localhost:8088/cluster

Most Visited ▾ Cloudera Hue YARN RM Spark UI (local) Spark Doc Solr Admin UI Kite SDK Doc

## All Applications

**Cluster Metrics**

Apps Submitted	Apps Pending	Apps Running	Apps Completed	Containers Running	Memory Used	Memory Total	Memory Reserved	VCores Used	VCores Total	VCores Reserved	Active Nodes	Decommissioned Nodes	Lost Nodes
2	0	0	2	0	0 B	2 GB	0 B	0	8	0	1	0	0

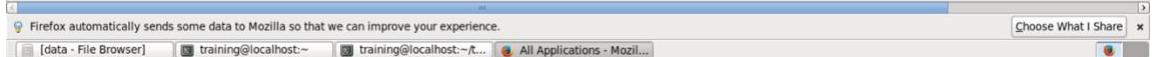
**User Metrics for dr.who**

Apps Submitted	Apps Pending	Apps Running	Apps Completed	Containers Running	Containers Pending	Containers Reserved	Memory Used	Memory Pending	Memory Reserved	VCores Used
0	0	0	2	0	0	0	0 B	0 B	0 B	0

Show 20 entries

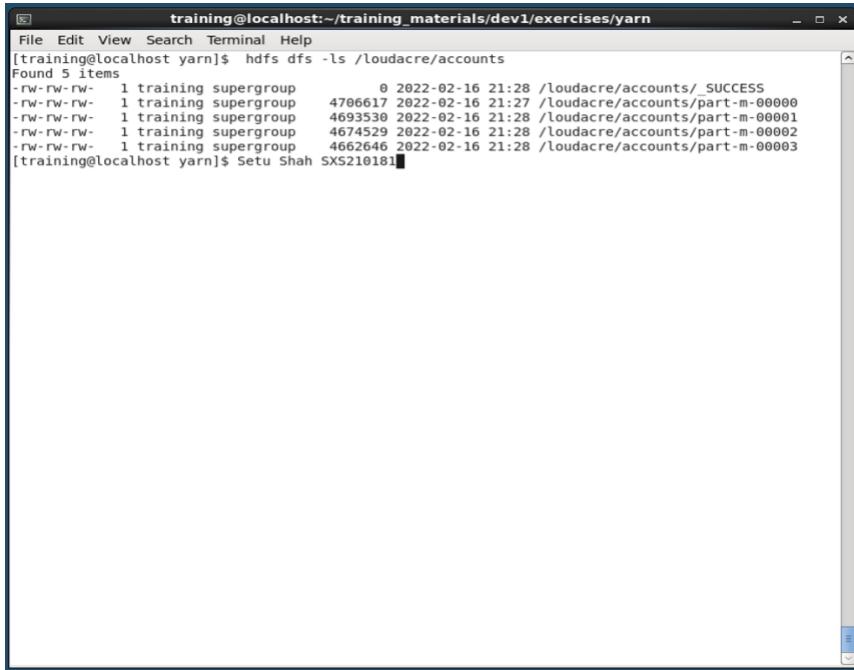
ID	User	Name	Application Type	Queue	StartTime	FinishTime	State	FinalStatus	Running Containers	Allocated CPU	Allocated VM
application_1644630540011_0002	training	accounts.jar	MAPREDUCE	root.training	Wed Feb 16 21:27:20 -0800 2022	Wed Feb 16 21:28:33 -0800 2022	FINISHED	SUCCEEDED	N/A	N/A	N/A
application_1644630540011_0001	training	wordcount.py	SPARK	root.training	Wed Feb 16 20:48:50 -0800 2022	Wed Feb 16 21:04:22 -0800 2022	FINISHED	SUCCEEDED	N/A	N/A	N/A

Showing 1 to 2 of 2 entries



## **View the imported data**

Using command “hdfs dfs -ls /loudacre/accounts” to view the contents of account directory.



```
training@localhost:~/training_materials/dev1/exercises/yarn
File Edit View Search Terminal Help
[training@localhost yarn]$ hdfs dfs -ls /loudacre/accounts
Found 5 items
-rw-rw-rw- 1 training supergroup 0 2022-02-16 21:28 /loudacre/accounts/_SUCCESS
-rw-rw-rw- 1 training supergroup 4706617 2022-02-16 21:27 /loudacre/accounts/part-m-00000
-rw-rw-rw- 1 training supergroup 4693530 2022-02-16 21:28 /loudacre/accounts/part-m-00001
-rw-rw-rw- 1 training supergroup 4674529 2022-02-16 21:28 /loudacre/accounts/part-m-00002
-rw-rw-rw- 1 training supergroup 4662646 2022-02-16 21:28 /loudacre/accounts/part-m-00003
[training@localhost yarn]$ Setu Shah SXS210181
```

Next step : I used tail command to view the last part of each file, for each of the Mapreduce partition files.



```
training@localhost:~/training_materials/dev1/exercises/yarn
File Edit View Search Terminal Help
[training@localhost yarn]$ hdfs dfs -tail /loudacre/accounts/part-m-00000
14-03-18 13:30:46.0
32434,2012-01-08 12:31:49.0,2013-12-25 04:55:12.0,Anne,Boyle,4215 McVaney Road,Reno,NV,89535,775054
6952,2014-03-18 13:30:46.0,2014-03-18 13:30:46.0
32435,2012-10-10 07:00:37.0,\N,Marjorie,Morrison,4417 Coffman Alley,Los Angeles,CA,90042,2132809416
,2014-03-18 13:30:46.0,2014-03-18 13:30:46.0
32436,2012-08-09 03:58:02.0,\N,Helen,Perez,1093 Columbia Road,Santa Barbara,CA,93490,8054121301,201
4-03-18 13:30:46.0,2014-03-18 13:30:46.0
32437,2012-01-14 10:27:37.0,\N,Melissa,Hayes,3434 Parkway Drive,Salem,OR,97394,5031584725,2014-03-1
8 13:30:46.0,2014-03-18 13:30:46.0
32438,2012-08-22 20:40:31.0,\N,Violet,Searcy,2601 Twin Oaks Drive,San Francisco,CA,94060,4152327199
,2014-03-18 13:30:46.0,2014-03-18 13:30:46.0
32439,2012-12-15 08:45:17.0,2013-06-19 00:08:39.0,Eunice,Myers,4867 Reeves Street,Long Beach,CA,907
23,5628246641,2014-03-18 13:30:46.0,2014-03-18 13:30:46.0
32440,2012-09-19 09:11:15.0,\N,Robert,Huskey,953 Neville Street,San Jose,CA,95180,4086716161,2014-0
3-18 13:30:46.0,2014-03-18 13:30:46.0
[training@localhost yarn]$ hdfs dfs -tail /loudacre/accounts/part-m-00001
:31:52.0
64874,2012-03-19 16:11:35.0,\N,Clara,Safford,4614 Spirit Drive,Glendale,CA,91217,7476693718,2014-03
-18 13:31:52.0,2014-03-18 13:31:52.0
64875,2012-10-10 20:07:25.0,\N,Erica,Oleary,3185 Fairway Drive,Long Beach,CA,90805,5624213548,2014-
03-18 13:31:52.0,2014-03-18 13:31:52.0
64876,2012-03-24 08:57:43.0,2014-01-27 09:14:34.0,Leslie,Jeanbaptiste,2573 Loving Acres Road,Kingma
n,AZ,86486,9286393819,2014-03-18 13:31:52.0,2014-03-18 13:31:52.0
64877,2012-04-24 05:52:44.0,\N,Jermaine,Harmon,4489 Shearwood Forest Drive,Burbank,CA,91503,8186973
079,2014-03-18 13:31:52.0,2014-03-18 13:31:52.0
64878,2012-08-04 01:49:37.0,\N,Mary,Hanson,522 Massachusetts Avenue,Oakland,CA,94676,5108035413,201
4-03-18 13:31:52.0,2014-03-18 13:31:52.0
64879,2012-10-21 23:36:04.0,\N,Richard,Alligood,4359 Agriculture Lane,San Diego,CA,92087,6196073804
,2014-03-18 13:31:52.0,2014-03-18 13:31:52.0
64880,2012-10-07 15:51:15.0,\N,Florence,Guevara,3507 School Street,San Francisco,CA,94183,415202013
0,2014-03-18 13:31:52.0,2014-03-18 13:31:52.0
[training@localhost yarn]$ hdfs dfs -tail /loudacre/accounts/part-m-00002
3:32:51.0
97314,2013-01-18 14:05:59.0,2013-04-14 03:44:13.0,Eloy,Murphy,58 College View,San Mateo,CA,94493,65
06652828,2014-03-18 13:32:51.0,2014-03-18 13:32:51.0
97315,2013-04-10 03:29:30.0,\N,Beatrice,Dodson,252 Farm Meadow Drive,Sacramento,CA,95762,9165704077
,2014-03-18 13:32:51.0,2014-03-18 13:32:51.0
97316,2013-03-26 14:13:25.0,\N,Gayle,Montelongo,364 Jail Drive,Kingman,AZ,86482,9285524623,2014-03-
18 13:32:51.0,2014-03-18 13:32:51.0
97317,2013-05-23 08:26:18.0,\N,Marc,Stock,1262 Ocala Street,Van Nuys,CA,91424,8184747448,2014-03-18
13:32:51.0,2014-03-18 13:32:51.0
```

The first six digit in the screenshot is the Account ID.  
The highest Account ID listed from all is 129761.

```
training@localhost:~/training_materials/dev1/exercises/yarn
File Edit View Search Terminal Help
n,AZ,86486,9286393819,2014-03-18 13:31:52.0,2014-03-18 13:31:52.0
64877,2012-04-24 05:52:44.0,\N,Jermaine,Harmon,4489 Shearwood Forest Drive,Burbank,CA,91503,8186973
079,2014-03-18 13:31:52.0,2014-03-18 13:31:52.0
64878,2012-08-04 01:49:37.0,\N,Mary,Hanson,522 Massachusetts Avenue,Oakland,CA,94676,5108035413,201
4-03-18 13:31:52.0,2014-03-18 13:31:52.0
64879,2012-10-21 23:36:04.0,\N,Richard,Alligood,4359 Agriculture Lane,San Diego,CA,92087,6196073804
,2014-03-18 13:31:52.0,2014-03-18 13:31:52.0
64880,2012-10-07 15:51:15.0,\N,Florence,Guevara,3507 School Street,San Francisco,CA,94183,415202013
0,2014-03-18 13:31:52.0,2014-03-18 13:31:52.0
[training@localhost yarn]$ hdfs dfs -tail /loudacre/accounts/part-m-00002
3:32:51.0
97314,2013-01-18 14:05:59.0,2013-04-14 03:44:13.0,Eloy,Murphy,58 College View,San Mateo,CA,94493,65
06652828,2014-03-18 13:32:51.0,2014-03-18 13:32:51.0
97315,2013-04-10 03:29:30.0,\N,Beatrice,Dodson,252 Farm Meadow Drive,Sacramento,CA,95762,9165704077
,2014-03-18 13:32:51.0,2014-03-18 13:32:51.0
97316,2013-03-26 14:13:25.0,\N,Gayle,Montelongo,364 Jail Drive,Kingman,AZ,86482,9285524623,2014-03-
18 13:32:51.0,2014-03-18 13:32:51.0
97317,2013-05-23 08:26:18.0,\N,Marc,Stock,1262 Ocala Street,Van Nuys,CA,91424,8184747448,2014-03-18
13:32:51.0,2014-03-18 13:32:51.0
97318,2013-10-11 15:20:07.0,\N,Roger,Weyand,349 Concord Street,Tucson,AZ,85794,9285391567,2014-03-1
8 13:32:51.0,2014-03-18 13:32:51.0
97319,2013-04-20 00:00:32.0,2014-01-31 00:41:43.0,Leona,Stannard,783 Stonecoal Road,Palm Springs,CA
,92240,7609402808,2014-03-18 13:32:51.0,2014-03-18 13:32:51.0
97320,2013-09-22 13:57:58.0,2014-03-14 15:55:49.0,Bryant,Nemec,4007 West Fork Drive,Eugene,OR,97440
,5411715739,2014-03-18 13:32:51.0,2014-03-18 13:32:51.0
[training@localhost yarn]$ hdfs dfs -tail /loudacre/accounts/part-m-00003
13:33:49.0
129755,2014-01-06 16:11:39.0,2014-02-21 11:22:56.0,John,Hale,1718 Irish Lane,San Bernardino,CA,9238
8,9097978236,2014-03-18 13:33:49.0,2014-03-18 13:33:49.0
129756,2014-01-26 19:34:06.0,\N,Danny,Graf,1104 Golden Ridge Road,San Jose,CA,95000,4088662329,2014
-03-18 13:33:49.0,2014-03-18 13:33:49.0
129757,2014-01-27 02:19:44.0,\N,John,McCall,4117 Conaway Street,Long Beach,CA,90609,5624227213,2014
-03-18 13:33:49.0,2014-03-18 13:33:49.0
129758,2014-03-14 20:25:15.0,\N,Robert,Pitt,2502 Heather Sees Way,Oakland,CA,94635,5100736336,2014-
03-18 13:33:49.0,2014-03-18 13:33:49.0
129759,2014-01-16 15:43:50.0,2014-02-19 07:31:53.0,Deborah,Hutchings,2071 Rodney Street,Globe,AZ,85
517,9288081544,2014-03-18 13:33:49.0,2014-03-18 13:33:49.0
129760,2014-01-06 04:24:56.0,\N,Zola,Tedder,1575 Linden Avenue,San Francisco,CA,94097,4151620942,20
14-03-18 13:33:49.0,2014-03-18 13:33:49.0
129761,2014-01-05 16:23:40.0,\N,Ruth,Ebersole,4684 Irving Place,Klamath Falls,OR,97646,5418511185,2
014-03-18 13:33:49.0,2014-03-18 13:33:49.0
[training@localhost yarn]$ Setu Shah SXS210181
```

## Importing incremental updates to accounts

I ran “add\_new\_accounts.py”, to add the latest accounts to MySQL.

I executed the code in below screenshot to import and append the newly added accounts to the account’s directory.

```
training@localhost:~/training_materials/dev1/exercises/yarn
File Edit View Search Terminal Help
[training@localhost yarn]$ sqoop import \
> --connect jdbc:mysql://localhost/loudacre \
> --username training --password training \
> --incremental append \
> --null-non-string '\\N' \
> --table accounts \
> --target-dir /loudacre/accounts \
> --check-column acct_num \
> --last-value 129761
22/02/16 21:43:19 INFO sqoop.Sqoop: Running Sqoop version: 1.4.5-cdh5.4.3
22/02/16 21:43:19 WARN tool.BaseSqoopTool: Setting your password on the command-line is insecure. Consider using - instead.
22/02/16 21:43:19 INFO manager.MySQLManager: Preparing to use a MySQL streaming resultset.
22/02/16 21:43:19 INFO tool.CodeGenTool: Beginning code generation
22/02/16 21:43:20 INFO manager.SqlManager: Executing SQL statement: SELECT t.* FROM `accounts` AS t
LIMIT 1
22/02/16 21:43:21 INFO manager.SqlManager: Executing SQL statement: SELECT t.* FROM `accounts` AS t
LIMIT 1
22/02/16 21:43:21 INFO orm.CompilationManager: HADOOP MAPRED HOME is /usr/lib/hadoop-mapreduce
Note: /tmp/sqoop-training/compile/c02e1c41a7af481e9bf2f00238c732a/accounts.java uses or overrides a deprecated API.
Note: Recompile with -Xlint:deprecation for details.
22/02/16 21:43:27 INFO orm.CompilationManager: Writing jar file: /tmp/sqoop-training/compile/c02e1c41a7af481e9bf2f00238c732a/accounts.jar
22/02/16 21:43:30 INFO tool.ImportTool: Maximal id query for free form incremental import: SELECT MAX(`acct_num`) FROM `accounts`
22/02/16 21:43:30 INFO tool.ImportTool: Incremental import based on column `acct_num`
22/02/16 21:43:30 INFO tool.ImportTool: Lower bound value: 129761
22/02/16 21:43:30 INFO tool.ImportTool: Upper bound value: 129764
22/02/16 21:43:30 WARN manager.MySQLManager: It looks like you are importing from mysql.
22/02/16 21:43:30 WARN manager.MySQLManager: This transfer can be faster! Use the --direct
22/02/16 21:43:30 WARN manager.MySQLManager: option to exercise a MySQL-specific fast path.
22/02/16 21:43:30 INFO manager.MySQLManager: Setting zero DATETIME behavior to convertToNull (mysql)
22/02/16 21:43:31 INFO mapreduce.ImportJobBase: Beginning import of accounts
22/02/16 21:43:31 INFO Configuration.deprecation: mapred.job.tracker is deprecated. Instead, use mapreduce.jobtracker.address
22/02/16 21:43:31 INFO Configuration.deprecation: mapred.jar is deprecated. Instead, use mapreduce.jar
22/02/16 21:43:31 INFO Configuration.deprecation: mapred.map.tasks is deprecated. Instead, use mapreduce.job.maps
22/02/16 21:43:31 INFO client.RMProxy: Connecting to ResourceManager at /0.0.0.0:8032
22/02/16 21:43:36 INFO db.DBInputFormat: Using read committed transaction isolation
```

```
training@localhost:~/training_materials/dev1/exercises/yarn
File Edit View Search Terminal Help
22/02/16 21:43:36 INFO db.DBInputFormat: Using read committed transaction isolation
22/02/16 21:43:36 INFO db.DataDrivenDBInputFormat: BoundingValsQuery: SELECT MIN(`acct_num`), MAX(`acct_num`) FROM `accounts` WHERE ( `acct_num` > 129761 AND `acct_num` <= 129764 )
22/02/16 21:43:37 INFO mapreduce.JobSubmitter: number of splits:3
22/02/16 21:43:37 INFO impl.YarnClientImpl: Submitted application application_1644630540011_0003
22/02/16 21:43:37 INFO mapreduce.Job: The url to track the job: http://localhost:8088/proxy/application_1644630540011_0003/
22/02/16 21:43:37 INFO mapreduce.Job: Running job: job_1644630540011_0003
22/02/16 21:43:52 INFO mapreduce.Job: Job job_1644630540011_0003 running in uber mode : false
22/02/16 21:43:52 INFO mapreduce.Job: map 0% reduce 0%
22/02/16 21:44:07 INFO mapreduce.Job: map 33% reduce 0%
22/02/16 21:44:19 INFO mapreduce.Job: map 67% reduce 0%
22/02/16 21:44:27 INFO mapreduce.Job: map 100% reduce 0%
22/02/16 21:44:27 INFO mapreduce.Job: Job job_1644630540011_0003 completed successfully
22/02/16 21:44:28 INFO mapreduce.Job: Counters: 30
    File System Counters
        FILE: Number of bytes read=0
        FILE: Number of bytes written=410586
        FILE: Number of read operations=0
        FILE: Number of large read operations=0
        FILE: Number of write operations=0
        HDFS: Number of bytes read=361
        HDFS: Number of bytes written=391
        HDFS: Number of read operations=12
        HDFS: Number of large read operations=0
        HDFS: Number of write operations=6
    Job Counters
        Launched map tasks=3
        Other local map tasks=3
        Total time spent by all maps in occupied slots (ms)=0
        Total time spent by all reduces in occupied slots (ms)=0
        Total time spent by all map tasks (ms)=27945
        Total vcore-seconds taken by all map tasks=27945
        Total megabyte-seconds taken by all map tasks=7153920
    Map-Reduce Framework
        Map input records=3
        Map output records=3
        Input split bytes=361
        Spilled Records=0
        Failed Shuffles=0
        Merged Map outputs=0
```

```

training@localhost:~/training_materials/dev1/exercises/yarn
File Edit View Search Terminal Help
FILE: Number of write operations=0
HDFS: Number of bytes read=361
HDFS: Number of bytes written=391
HDFS: Number of read operations=12
HDFS: Number of large read operations=0
HDFS: Number of write operations=6
Job Counters
Launched map tasks=3
Other local map tasks=3
Total time spent by all maps in occupied slots (ms)=0
Total time spent by all reduces in occupied slots (ms)=0
Total time spent by all map tasks (ms)=27945
Total vcore-seconds taken by all map tasks=27945
Total megabyte-seconds taken by all map tasks=7153920
Map-Reduce Framework
  Map input records=3
  Map output records=3
  Input split bytes=361
  Spilled Records=0
  Failed Shuffles=0
  Merged Map outputs=0
  GC time elapsed (ms)=558
  CPU time spent (ms)=3750
  Physical memory (bytes) snapshot=283234304
  Virtual memory (bytes) snapshot=2533376000
  Total committed heap usage (bytes)=55877632
File Input Format Counters
  Bytes Read=0
File Output Format Counters
  Bytes Written=391
22/02/16 21:44:28 INFO mapreduce.ImportJobBase: Transferred 391 bytes in 57.0195 seconds (6.8573 bytes/sec)
22/02/16 21:44:28 INFO mapreduce.ImportJobBase: Retrieved 3 records.
22/02/16 21:44:28 INFO util.AppendUtils: Appending to directory accounts
22/02/16 21:44:28 INFO util.AppendUtils: Using found partition 4
22/02/16 21:44:28 INFO tool.ImportTool: Incremental import complete! To run another incremental import of all data following this import, supply the following arguments:
22/02/16 21:44:28 INFO tool.ImportTool: --incremental append
22/02/16 21:44:28 INFO tool.ImportTool: --check-column acct_num
22/02/16 21:44:28 INFO tool.ImportTool: --last-value 129764
22/02/16 21:44:28 INFO tool.ImportTool: (Consider saving this with 'sqoop job --create')
[training@localhost yarn]$ Setu Shah SX5210181

```

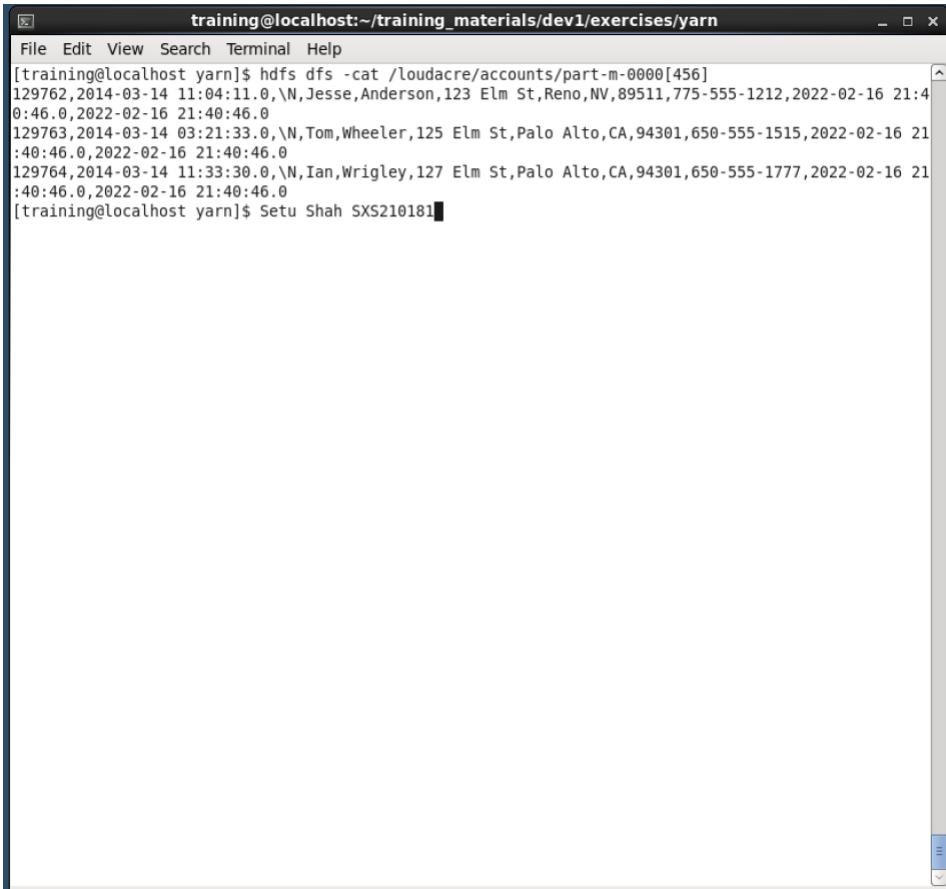
Verifying the sqoop import, I executed the code “hdfs dfs -ls /loudacre/accounts” and viewed the contents of the accounts directory. We can see new added files.

```

training@localhost:~/training_materials/dev1/exercises/yarn
File Edit View Search Terminal Help
[training@localhost yarn]$ hdfs dfs -ls /loudacre/accounts
Found 8 items
-rw-rw-rw- 1 training supergroup          0 2022-02-16 21:28 /loudacre/accounts/_SUCCESS
-rw-rw-rw- 1 training supergroup 4706617 2022-02-16 21:27 /loudacre/accounts/part-m-00000
-rw-rw-rw- 1 training supergroup 4693530 2022-02-16 21:28 /loudacre/accounts/part-m-00001
-rw-rw-rw- 1 training supergroup 4674529 2022-02-16 21:28 /loudacre/accounts/part-m-00002
-rw-rw-rw- 1 training supergroup 4662646 2022-02-16 21:28 /loudacre/accounts/part-m-00003
-rw-rw-rw- 1 training supergroup 129 2022-02-16 21:44 /loudacre/accounts/part-m-00004
-rw-rw-rw- 1 training supergroup 131 2022-02-16 21:44 /loudacre/accounts/part-m-00005
-rw-rw-rw- 1 training supergroup 131 2022-02-16 21:44 /loudacre/accounts/part-m-00006
[training@localhost yarn]$ Setu Shah SX5210181

```

In the below screenshot ‘cat’ command is used to view the entire contents of these files



The screenshot shows a terminal window titled "training@localhost:~/training\_materials/dev1/exercises/yarn". The window contains the following text:

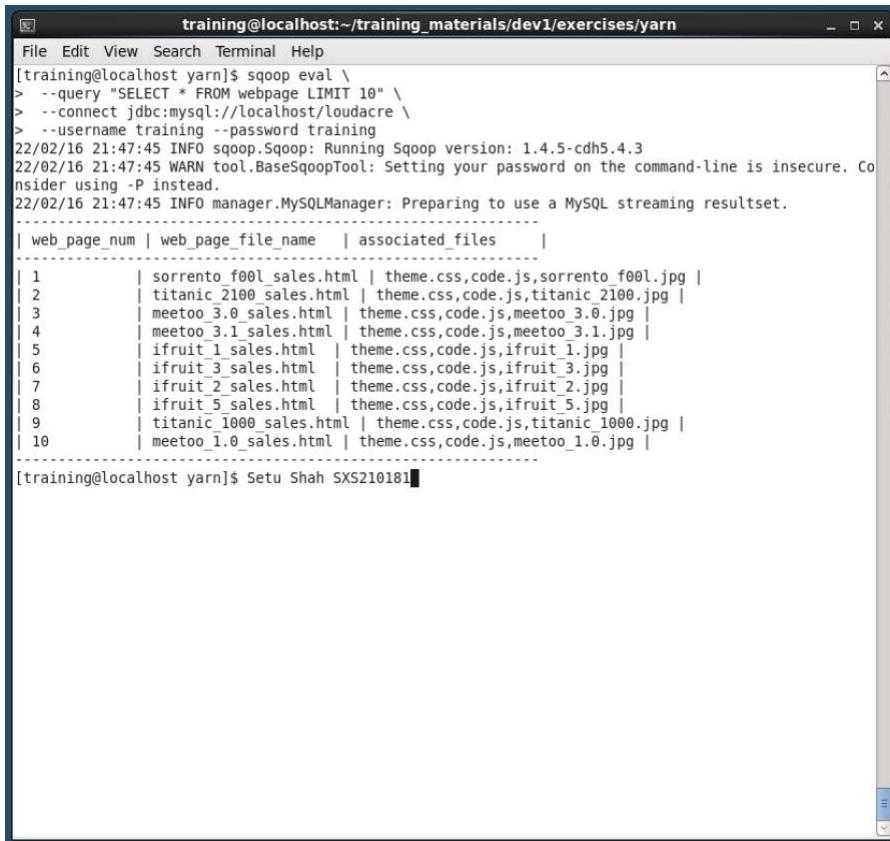
```
[training@localhost yarn]$ hdfs dfs -cat /loudacre/accounts/part-m-0000[456]
129762,2014-03-14 11:04:11.0,\N,Jesse,Anderson,123 Elm St,Reno,NV,89511,775-555-1212,2022-02-16 21:4
0:46.0,2022-02-16 21:40:46.0
129763,2014-03-14 03:21:33.0,\N,Tom,Wheeler,125 Elm St,Palo Alto,CA,94301,650-555-1515,2022-02-16 21
:40:46.0,2022-02-16 21:40:46.0
129764,2014-03-14 11:33:30.0,\N,Ian,Wrigley,127 Elm St,Palo Alto,CA,94301,650-555-1777,2022-02-16 21
:40:46.0,2022-02-16 21:40:46.0
[training@localhost yarn]$ Setu Shah SXS210181
```

## **Import webpage data using an alternate field delimiter**

In the below screenshot, I imported the webpage table to HDFS.

I used command “sqoop eval” to view the records in the table.

I can see that the values in the last column contains comma.



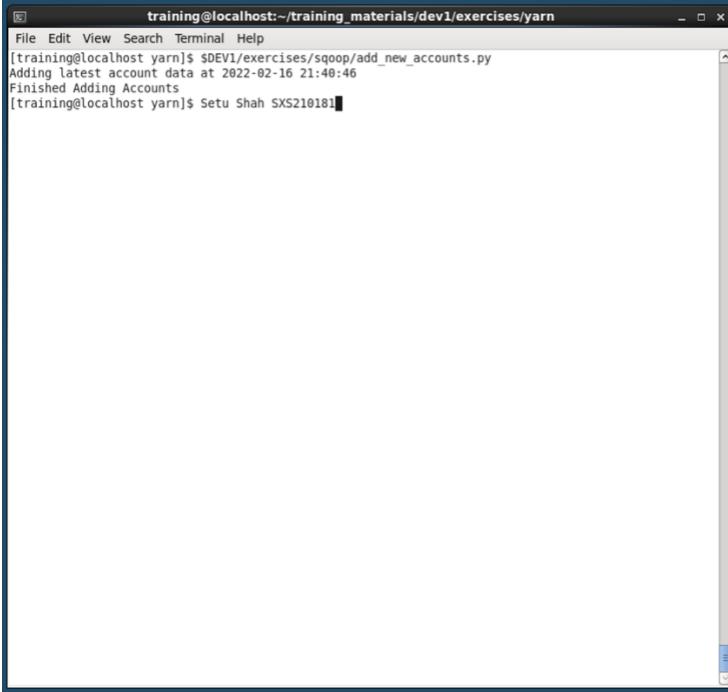
The screenshot shows a terminal window titled "training@localhost:~/training\_materials/dev1/exercises/yarn". The window displays the output of a Sqoop eval command. The command used was:

```
[training@localhost yarn]$ sqoop eval \
> --query "SELECT * FROM webpage LIMIT 10" \
> --connect jdbc:mysql://localhost/loudacre \
> --username training --password training
```

The output shows 10 rows of data from the webpage table, with columns labeled "web\_page\_num", "web\_page\_file\_name", and "associated\_files". The "associated\_files" column contains multiple file paths separated by commas.

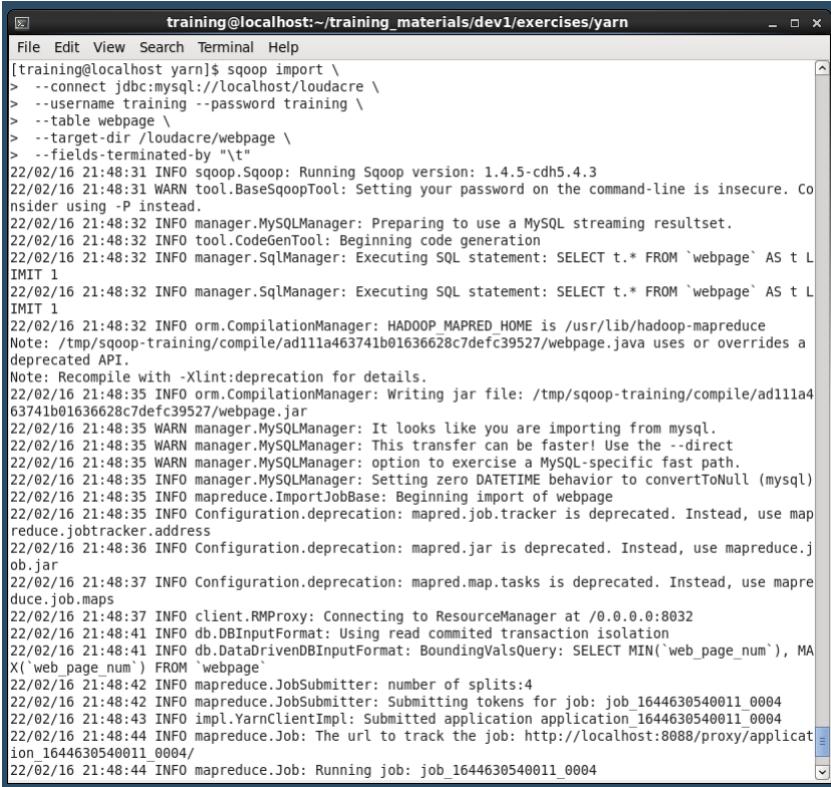
web_page_num	web_page_file_name	associated_files
1	sorrento_f00l_sales.html	theme.css,code.js,sorrento_f00l.jpg
2	titanic_2100_sales.html	theme.css,code.js,titanic_2100.jpg
3	meetoo_3.0_sales.html	theme.css,code.js,meetoo_3.0.jpg
4	meetoo_3.1_sales.html	theme.css,code.js,meetoo_3.1.jpg
5	ifruit1_sales.html	theme.css,code.js,ifruit1.jpg
6	ifruit3_sales.html	theme.css,code.js,ifruit3.jpg
7	ifruit2_sales.html	theme.css,code.js,ifruit2.jpg
8	ifruit5_sales.html	theme.css,code.js,ifruit5.jpg
9	titanic_1000_sales.html	theme.css,code.js,titanic_1000.jpg
10	meetoo_1.0_sales.html	theme.css,code.js,meetoo_1.0.jpg

[training@localhost yarn]\$ Setu Shah SXS21018!



```
training@localhost:~/training_materials/dev1/exercises/yarn
File Edit View Search Terminal Help
[training@localhost yarn]$ $DEV1/exercises/sqoop/add_new_accounts.py
[training@localhost yarn]$ Adding latest account data at 2022-02-16 21:40:46
Finished Adding Accounts
[training@localhost yarn]$ Setu Shah SXS21018:
```

In the below screenshot, I imported the webpage table to HDFS but used the tab character “\t” instead of default (comma) as the field terminator.



```
training@localhost:~/training_materials/dev1/exercises/yarn
File Edit View Search Terminal Help
[training@localhost yarn]$ sqoop import \
> --connect jdbc:mysql://localhost/loudacre \
> --username training --password training \
> --table webpage \
> --target-dir /loudacre/webpage \
> --fields-terminated-by "\t"
22/02/16 21:48:31 INFO sqoop.Sqoop: Running Sqoop version: 1.4.5-cdh5.4.3
22/02/16 21:48:31 WARN tool.BaseSqoopTool: Setting your password on the command-line is insecure. Consider using -P instead.
22/02/16 21:48:32 INFO manager.MySQLManager: Preparing to use a MySQL streaming resultset.
22/02/16 21:48:32 INFO tool.CodeGenTool: Beginning code generation
22/02/16 21:48:32 INFO manager.SqlManager: Executing SQL statement: SELECT t.* FROM `webpage` AS t L
IMIT 1
22/02/16 21:48:32 INFO manager.SqlManager: Executing SQL statement: SELECT t.* FROM `webpage` AS t L
IMIT 1
22/02/16 21:48:32 INFO orm.CompilationManager: HADOOP_MAPRED_HOME is /usr/lib/hadoop-mapreduce
Note: /tmp/sqoop-training/compile/ad111a463741b01636628c7defc39527/webpage.java uses or overrides a
deprecated API.
Note: Recompile with -Xlint:deprecation for details.
22/02/16 21:48:35 INFO orm.CompilationManager: Writing jar file: /tmp/sqoop-training/compile/ad111a4
63741b01636628c7defc39527/webpage.jar
22/02/16 21:48:35 WARN manager.MySQLManager: It looks like you are importing from mysql.
22/02/16 21:48:35 WARN manager.MySQLManager: This transfer can be faster! Use the --direct
22/02/16 21:48:35 WARN manager.MySQLManager: option to exercise MySQL-specific fast path.
22/02/16 21:48:35 INFO manager.MySQLManager: Setting zero DATETIME behavior to convertToNull (mysql)
22/02/16 21:48:35 INFO mapreduce.ImportJobBase: Beginning import of webpage
22/02/16 21:48:35 INFO Configuration.deprecation: mapred.job.tracker is deprecated. Instead, use map
reduce.jobtracker.address
22/02/16 21:48:36 INFO Configuration.deprecation: mapred.jar is deprecated. Instead, use mapreduce.j
ob.jar
22/02/16 21:48:37 INFO Configuration.deprecation: mapred.map.tasks is deprecated. Instead, use mapre
duce.job.maps
22/02/16 21:48:37 INFO client.RMProxy: Connecting to ResourceManager at /0.0.0.0:8032
22/02/16 21:48:41 INFO db.DBInputFormat: Using read committed transaction isolation
22/02/16 21:48:41 INFO db.DataDrivenDBInputFormat: BoundingValsQuery: SELECT MIN(`web_page_num`), MA
X(`web_page_num`) FROM `webpage`
22/02/16 21:48:42 INFO mapreduce.JobSubmitter: number of splits:4
22/02/16 21:48:42 INFO mapreduce.JobSubmitter: Submitting tokens for job: job_1644630540011_0004
22/02/16 21:48:43 INFO impl.YarnClientImpl: Submitted application application_1644630540011_0004
22/02/16 21:48:44 INFO mapreduce.Job: The url to track the job: http://localhost:8088/proxy/applicat
ion_1644630540011_0004/
22/02/16 21:48:44 INFO mapreduce.Job: Running job: job_1644630540011_0004
```

```
training@localhost:~/training_materials/dev1/exercises/yarn
File Edit View Search Terminal Help
22/02/16 21:48:44 INFO mapreduce.Job: Running job: job_1644630540011_0004
22/02/16 21:49:05 INFO mapreduce.Job: Job job_1644630540011_0004 running in uber mode : false
22/02/16 21:49:17 INFO mapreduce.Job: map 0% reduce 0%
22/02/16 21:49:28 INFO mapreduce.Job: map 25% reduce 0%
22/02/16 21:49:38 INFO mapreduce.Job: map 50% reduce 0%
22/02/16 21:49:45 INFO mapreduce.Job: map 75% reduce 0%
22/02/16 21:49:46 INFO mapreduce.Job: Job job_1644630540011_0004 completed successfully
22/02/16 21:49:46 INFO mapreduce.Job: Counters: 30
  File System Counters
    FILE: Number of bytes read=0
    FILE: Number of bytes written=543672
    FILE: Number of read operations=0
    FILE: Number of large read operations=0
    FILE: Number of write operations=0
    HDFS: Number of bytes read=480
    HDFS: Number of bytes written=3203
    HDFS: Number of read operations=16
    HDFS: Number of large read operations=0
    HDFS: Number of write operations=8
  Job Counters
    Launched map tasks=4
    Other local map tasks=4
    Total time spent by all maps in occupied slots (ms)=0
    Total time spent by all reduces in occupied slots (ms)=0
    Total time spent by all map tasks (ms)=31461
    Total vcore-seconds taken by all map tasks=31461
    Total megabyte-seconds taken by all map tasks=8054016
  Map-Reduce Framework
    Map input records=53
    Map output records=53
    Input split bytes=480
    Spilled Records=0
    Failed Shuffles=0
    Merged Map outputs=0
    GC time elapsed (ms)=550
    CPU time spent (ms)=3780
    Physical memory (bytes) snapshot=379711488
    Virtual memory (bytes) snapshot=3377889280
    Total committed heap usage (bytes)=75026432
  File Input Format Counters
    Bytes Read=0
```

```
training@localhost:~/training_materials/dev1/exercises/yarn
File Edit View Search Terminal Help
22/02/16 21:49:45 INFO mapreduce.Job: map 100% reduce 0%
22/02/16 21:49:45 INFO mapreduce.Job: Job job_1644630540011_0004 completed successfully
22/02/16 21:49:46 INFO mapreduce.Job: Counters: 30
  File System Counters
    FILE: Number of bytes read=0
    FILE: Number of bytes written=543672
    FILE: Number of read operations=0
    FILE: Number of large read operations=0
    FILE: Number of write operations=0
    HDFS: Number of bytes read=480
    HDFS: Number of bytes written=3203
    HDFS: Number of read operations=16
    HDFS: Number of large read operations=0
    HDFS: Number of write operations=8
  Job Counters
    Launched map tasks=4
    Other local map tasks=4
    Total time spent by all maps in occupied slots (ms)=0
    Total time spent by all reduces in occupied slots (ms)=0
    Total time spent by all map tasks (ms)=31461
    Total vcore-seconds taken by all map tasks=31461
    Total megabyte-seconds taken by all map tasks=8054016
  Map-Reduce Framework
    Map input records=53
    Map output records=53
    Input split bytes=480
    Spilled Records=0
    Failed Shuffles=0
    Merged Map outputs=0
    GC time elapsed (ms)=550
    CPU time spent (ms)=3780
    Physical memory (bytes) snapshot=379711488
    Virtual memory (bytes) snapshot=3377889280
    Total committed heap usage (bytes)=75026432
  File Input Format Counters
    Bytes Read=0
  File Output Format Counters
    Bytes Written=3203
22/02/16 21:49:46 INFO mapreduce.ImportJobBase: Transferred 3.1279 KB in 68.2397 seconds (46.9375 bytes/sec)
22/02/16 21:49:46 INFO mapreduce.ImportJobBase: Retrieved 53 records.
[training@localhost yarn]$ Setu Shah SXS210181
```

Now, I can view the imported data files in HDFS command line as well as in Hue browser.

```
training@localhost:~/training_materials/dev1/exercises/yarn
File Edit View Search Terminal Help
[training@localhost yarn]$ hdfs dfs -ls /loudacre/webpage
Found 5 items
-rw-rw-rw- 1 training supergroup          0 2022-02-16 21:49 /loudacre/webpage/_SUCCESS
-rw-rw-rw- 1 training supergroup      751 2022-02-16 21:49 /loudacre/webpage/part-m-00001
-rw-rw-rw- 1 training supergroup      874 2022-02-16 21:49 /loudacre/webpage/part-m-00001
-rw-rw-rw- 1 training supergroup      884 2022-02-16 21:49 /loudacre/webpage/part-m-00002
-rw-rw-rw- 1 training supergroup      774 2022-02-16 21:49 /loudacre/webpage/part-m-00003
[training@localhost yarn]$ Setu Shah SXS210181
```

Hue - File Browser - Mozilla Firefox  
Wed Feb 16, 9:53 PM

File Browser - Mozilla Firefox

localhost:8888/filebrowser/#/loudacre

File Browser

Name	Size	User	Group	Permissions	Date
_SUCCESS		hdfs	supergroup	drwxr-xr-x	February 16, 2022 07:28 PM
.		training	supergroup	drwxrwxrwx	February 16, 2022 09:49 PM
accounts		training	supergroup	drwxrwxrwx	February 16, 2022 09:44 PM
base_stations.tsv	14.8 KB	training	supergroup	-rw-rw-rw-	February 16, 2022 08:40 PM
kb		training	supergroup	drwxrwxrwx	February 16, 2022 07:31 PM
webpage		training	supergroup	drwxrwxrwx	February 16, 2022 09:49 PM

Show 45 of 4 items

Page 1 of 1

Choose What I Share