LAB 1

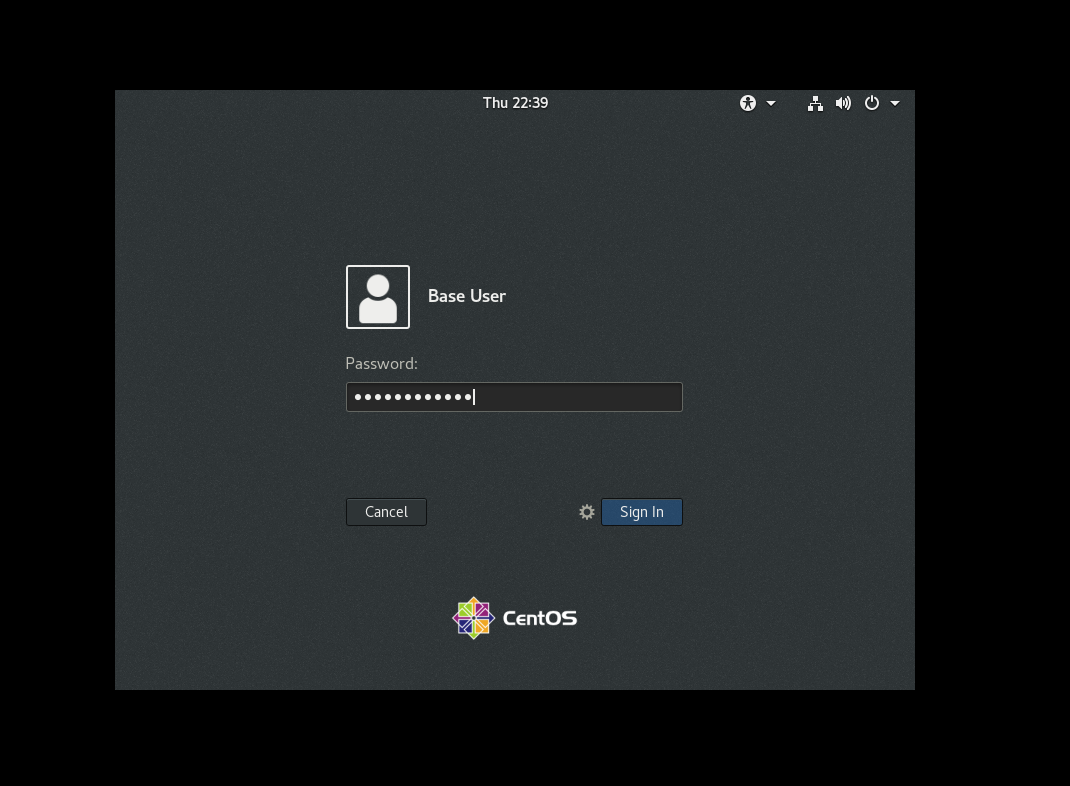
Name: Karim Gamal Mahmoud Mohamed

E-Mail: [21KGMM@queensu.ca](mailto:21KGMM@queensu.ca)

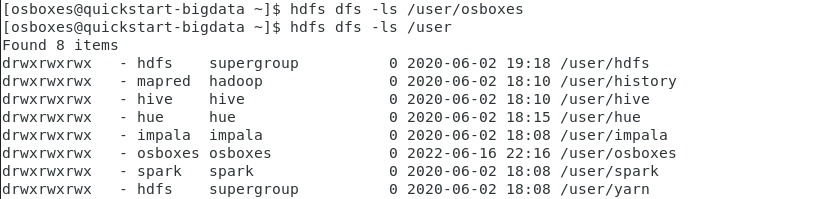
1. The virtual machine login (assuming you have a Cloudera VM installed at your location)

Note: The user : asosboxes

Password : BaseUser@123



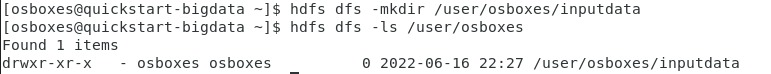
1. List all the files and directories under /user/osboxes in HDFS



3) Create a directory with name 'inputdata' in HDFS (command: hdfs dfs –mkdir directoryname)

&

4) List the folders in /user/osboxes within hdfs



5) Now, from your local machine, download historical stock data (csv format) of at least 5 stock quotes (random) from yahoo financial (finance.yahoo.com) and put them in a local folder of your machine.

Graphical user interface, text, application

Description automatically generated

6) In each of the files put the stock ticker (e.g. TSLA) in the first column of the file. Similarly do the same for others and finally merge them in one file named 'stock.csv’.

A screenshot of a computer

Description automatically generated

7) Install Winscp and connect to VM to transfer the stock.csv file from your local machine to the VM (in a new folder called 'data' under /home/osboxes).



A picture containing text

Description automatically generated

Graphical user interface, application

Description automatically generated

1. Within VM, copy the file stock.csv from /home/osboxes/data (local file system) to 'inputdata' directory in HDFS that are created earlier



9) Show the content of the last few records in the stock.csv file in HDFS.

A screenshot of a computer

Description automatically generated with medium confidence

10)Make a copy of the stock.csv file as stock1.csv from the current hdfs location to another folder named 'testdata' in hdfs.

A picture containing table

Description automatically generated

Text, letter

Description automatically generated