**BDA LAB ASSIGNMENT**

**Name-Vivek Pusti**

**Class -Cse IS 3**

**Roll no- 2193283**

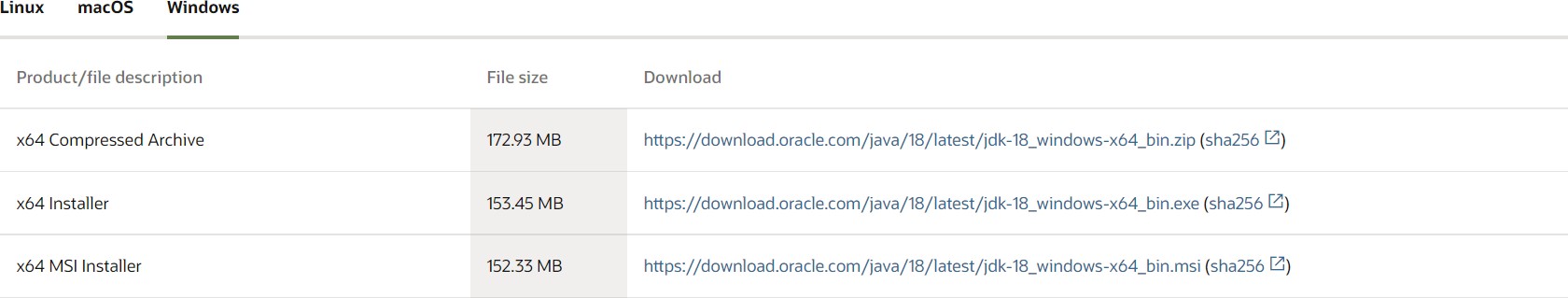
**Enrollno-MITU19BTCS0228**

ASSIGNMENT NO.1: Hadoop Installation

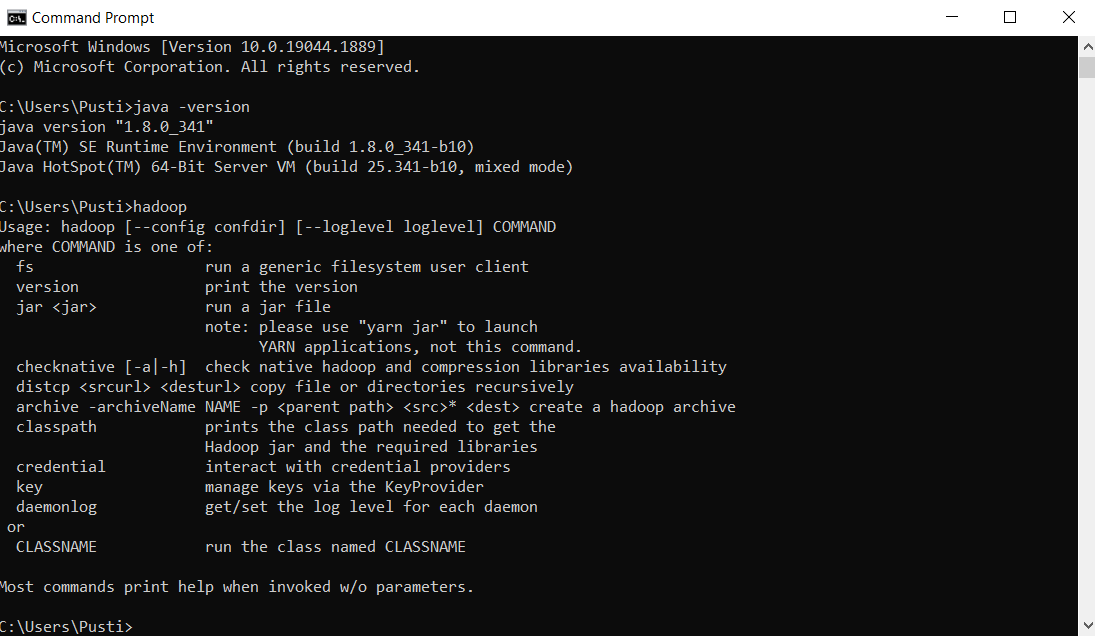
# Aim – Installation and configuration of Hadoop in the windows operating system.

**Step 1: Installation of Java**

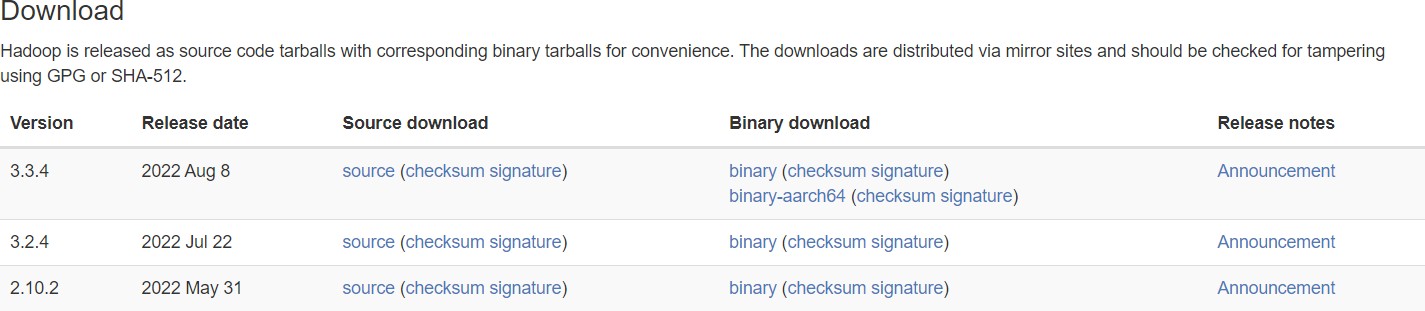
1. Hadoop required java to run it’s services hence java latest version is downloaded from official oracle website <https://www.oracle.com/java/technologies/downloads/>.



1. After successful installation of java jdk install the downloaded x64 installer.
2. To run java into the system we have to set the environment variable into local machine so that java will directly shift too the bin location whenever the java or javac command is encountered.
3. To verify if the path is properly set or not got to command promote and enter command java -version.

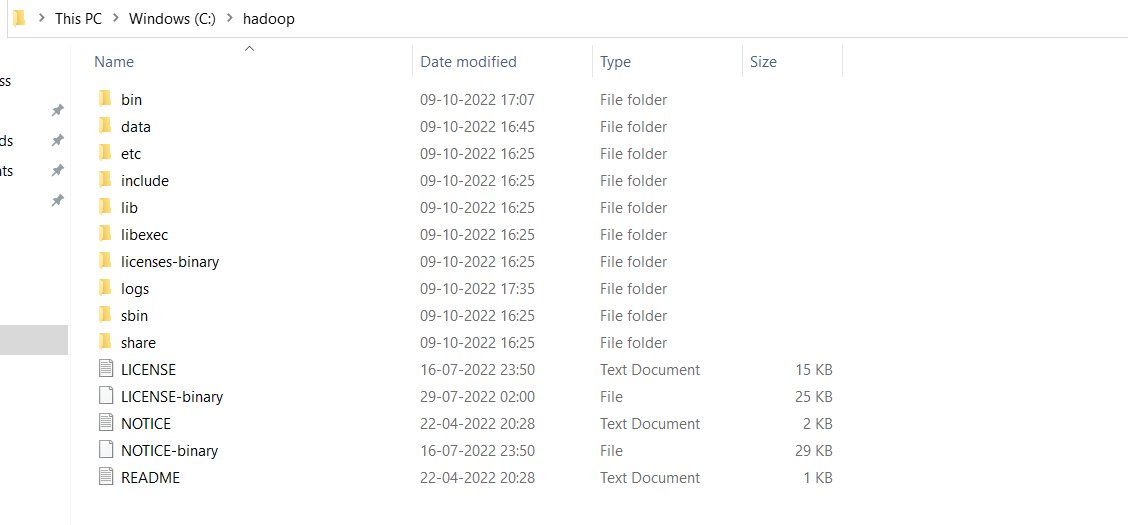


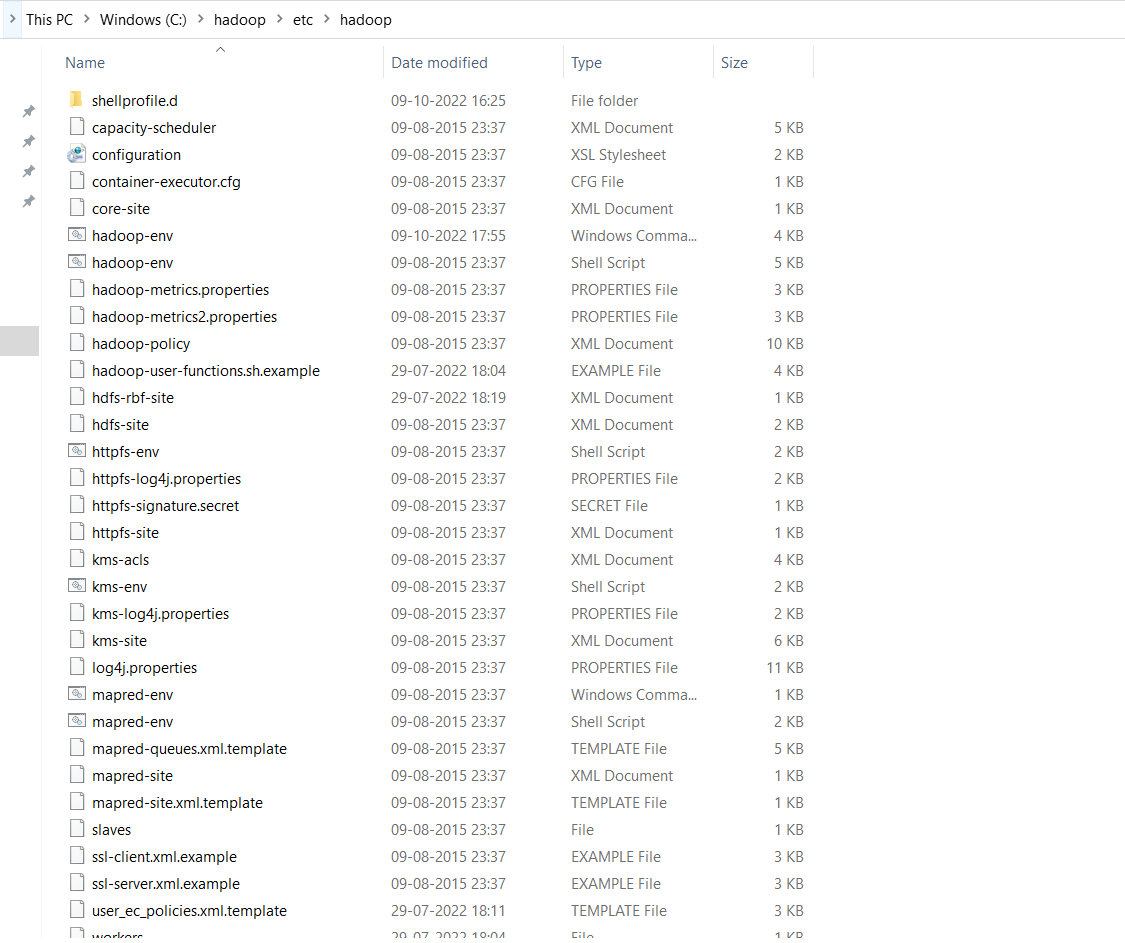
# Step 2: Installation of Hadoop

1. To install Hadoop we have to visit the official Hadoop website in order to downloaded the required files. <https://hadoop.apache.org/releases.html>
2. Download the latest version of Hadoop with binary download which allows us to modify the Hadoop files and it consist of all the required files which is in the compress format.
3. After downloading the compress .tar file extract the file into C drive ( it is not exe file which will automatically extracted)
4. After successful extraction go the the Hadoop 3.3.4 folder
5. Now in order to create datanode and namenode which are important node of Hadoop architecture we have to create a new folder named as data.
6. After creation of data folder we have to create 2 new folders named as datanode and namenode which are empty for now.

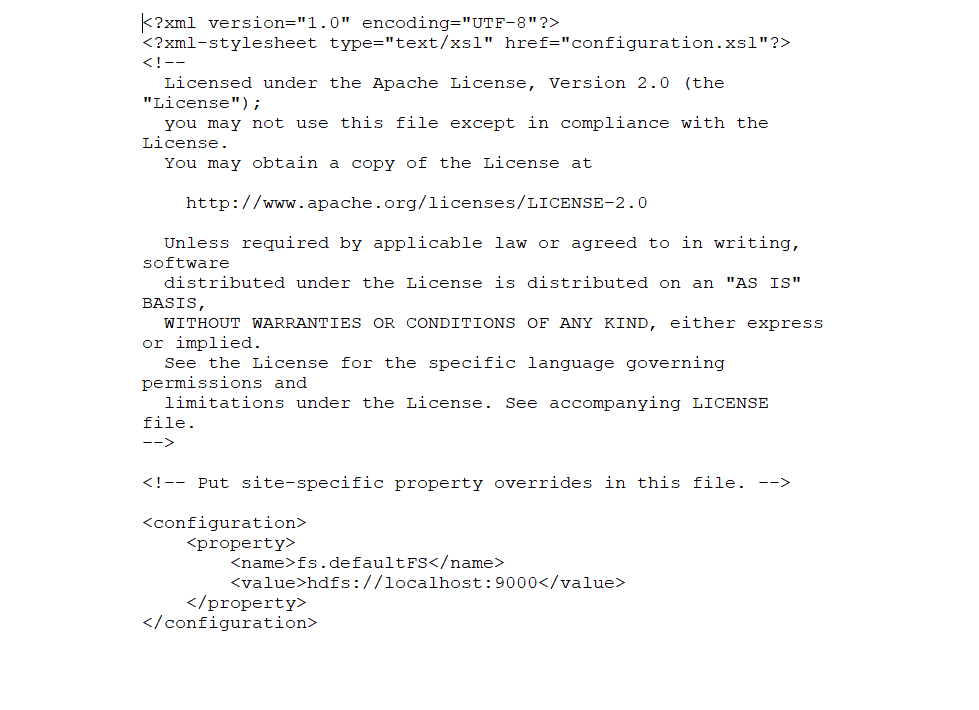
# 3: Setting up the Hadoop files

1. To make your Hadoop run all the services we have to modify the core Hadoop xml files such as hdfs.xml, yarn.xml, and so on.
2. To modify the file go to etc folder of Hadoop installation
3. Go to etc folder and you see all the required file that need to modify.

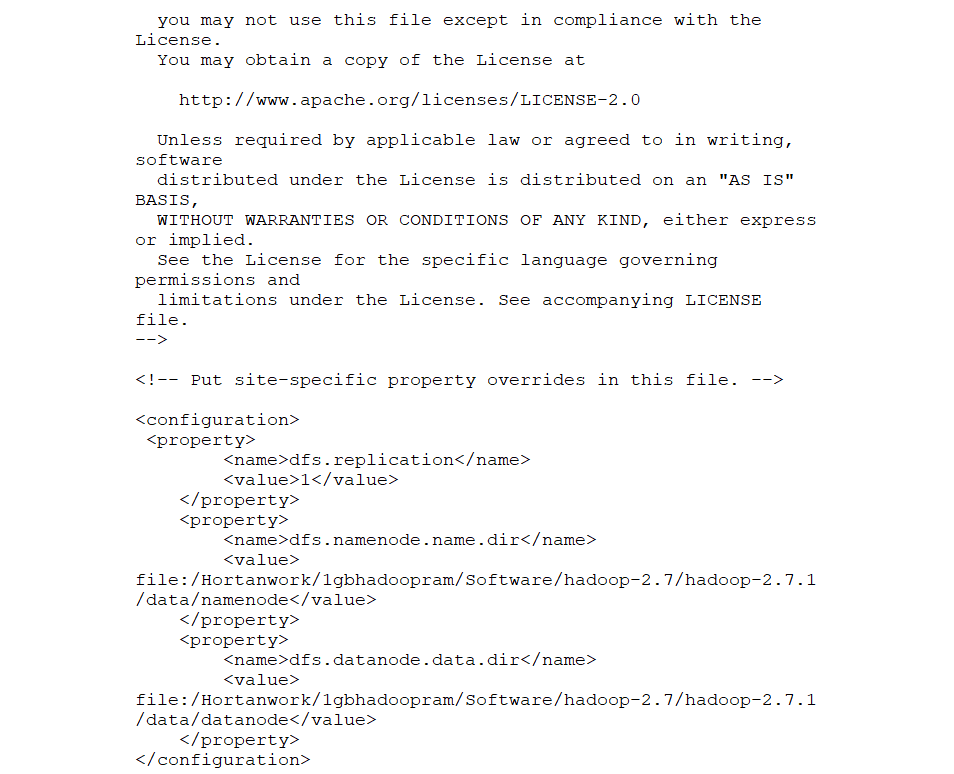




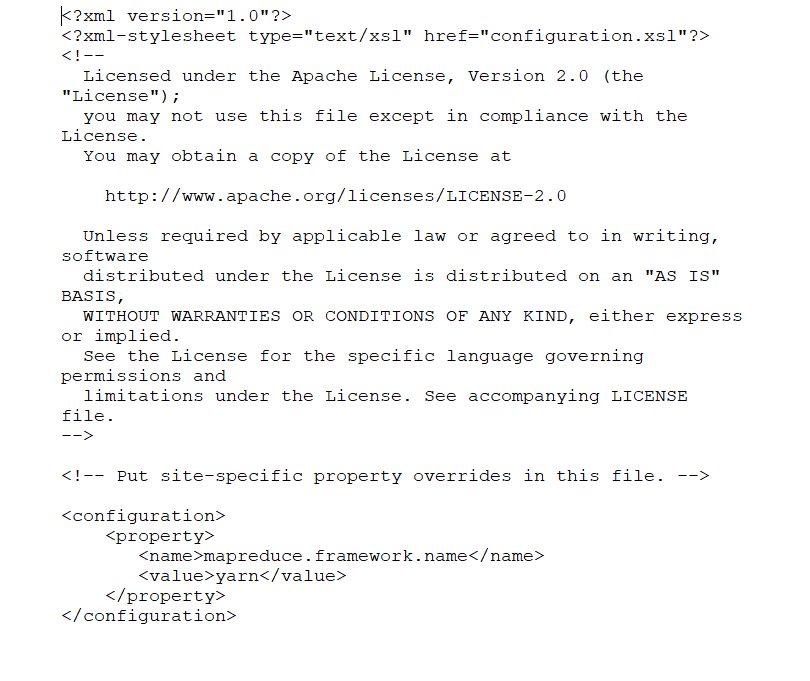
Now to modify or configure the core xml to make Hadoop execute command first open the core-site.xml file and add the below code



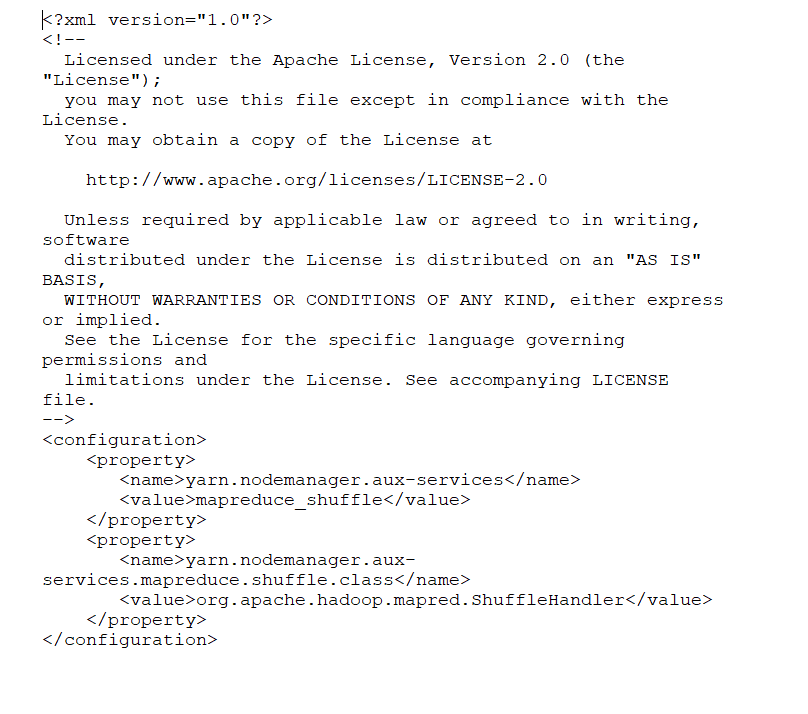
The second file that need to modify is hdfs.xml file add the below code into that file



The third file need to configure is mapred.xml add the below code into that file



The 4th file need to configure is yarn-site.xml which is a framework which keep track of other data clusters which helps in location of datanode and namenode clusters for that use below code.

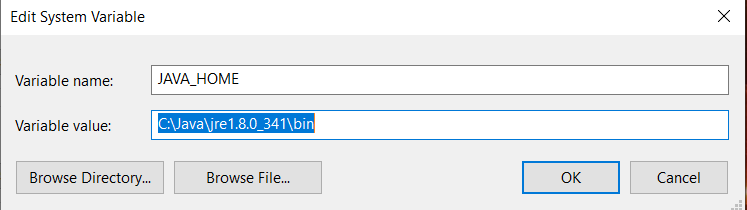


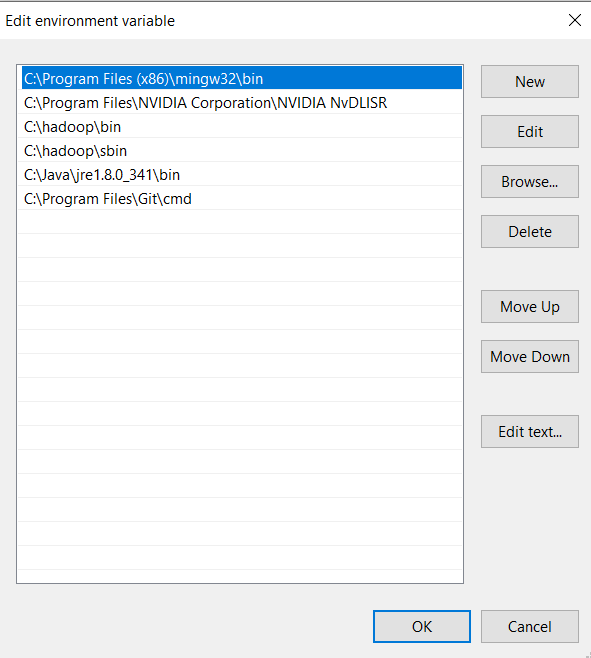
The last file need to configure is Hadoop-env file

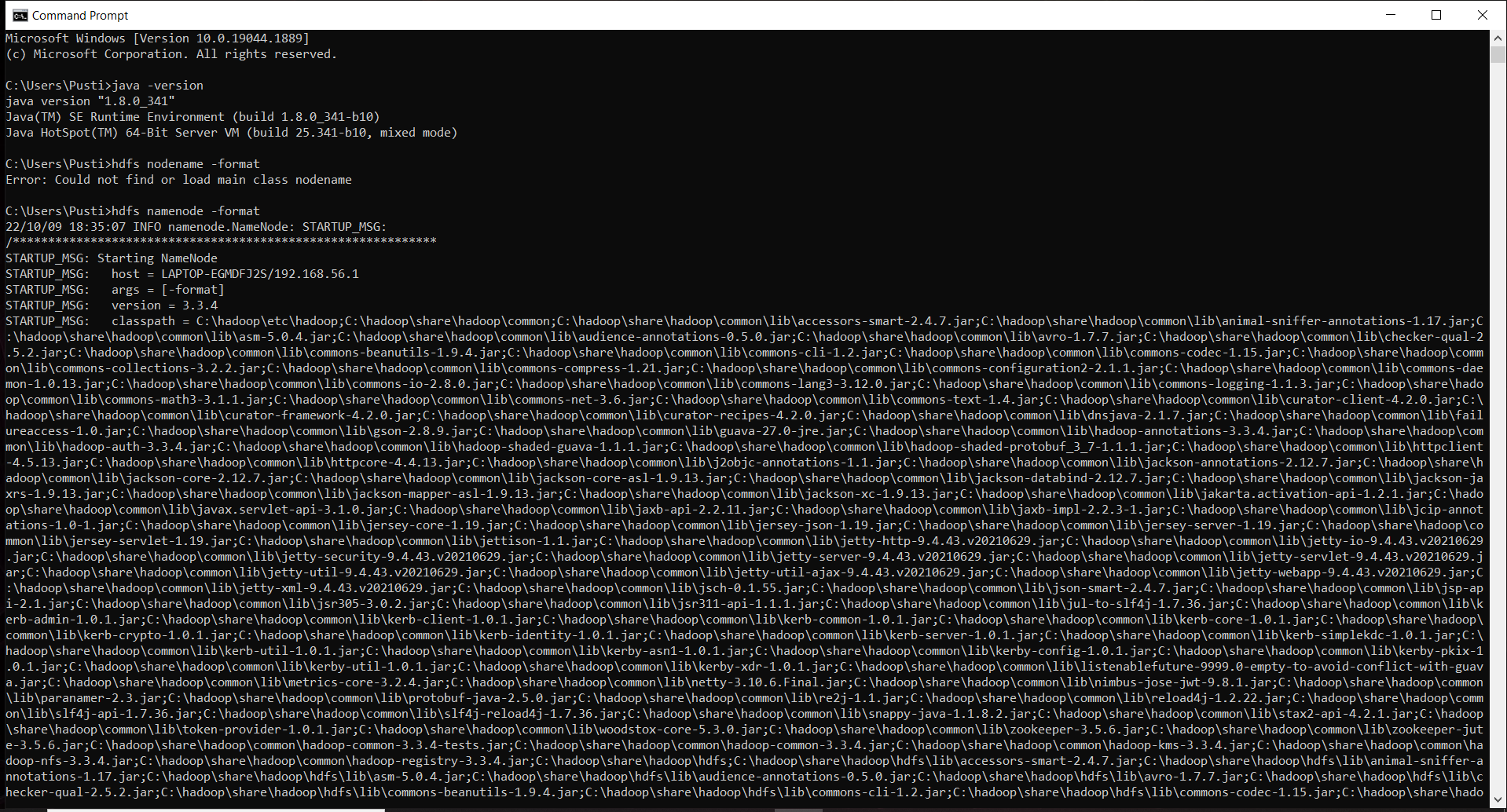


# Step 4: Setting up the environment path

1. In order to execute the Hadoop code we have to set the path of bin and sbin folder into the environmental variable for that search for environmental variable.
2. On the first section of user variable create new variable and HADOOP\_HOME and set the path till bin directory.
3. Similarly for system variable click on path and click on edit you will se the new window.







**Conclusion** – Thus we have studied about the installation of Hadoop and required configuration of Hadoop.