**Name:- Samreen Shaikh Roll no:- 64**

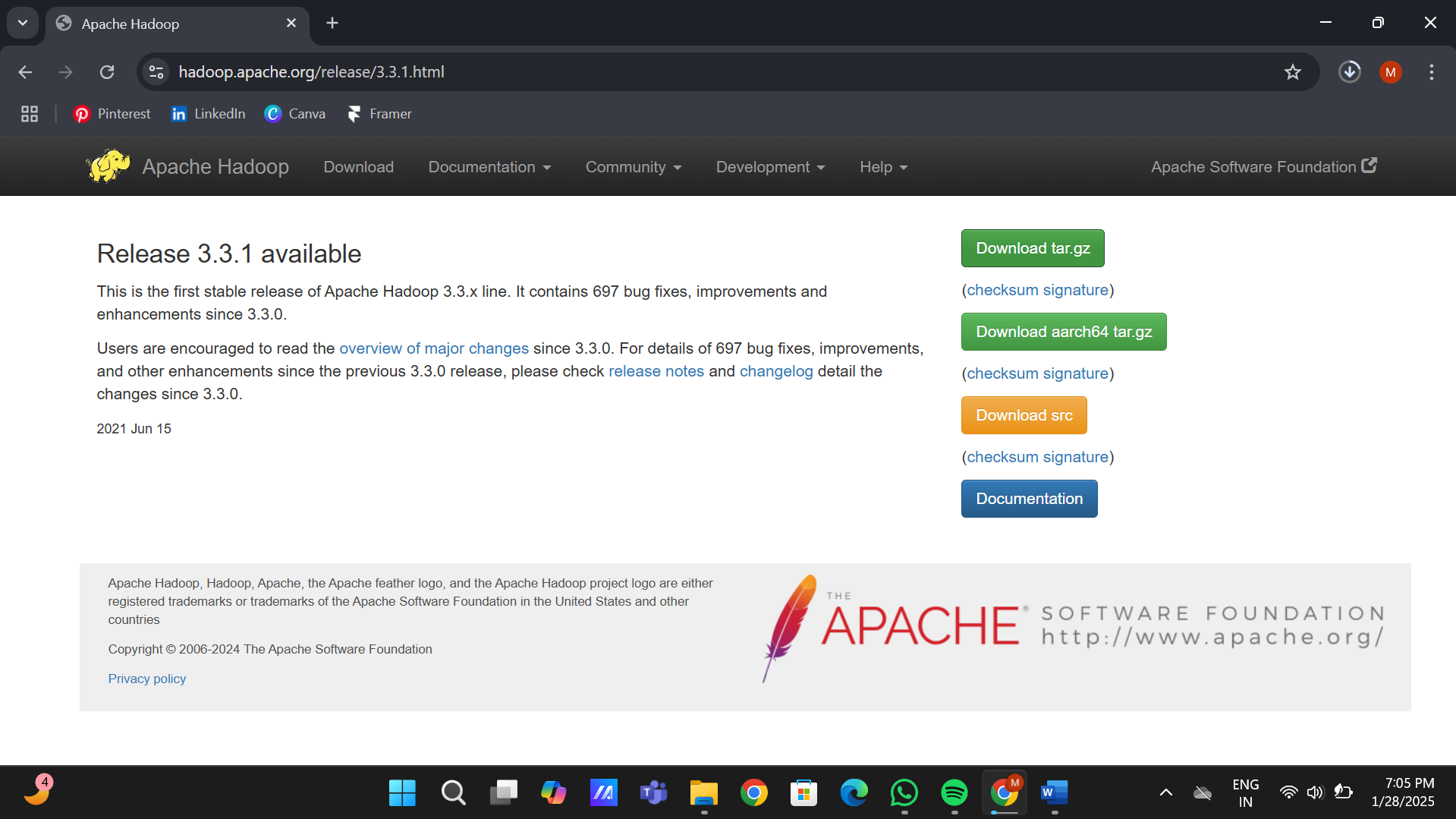
**Subject:- Big Data Analytics Date:- 28/01/2025**

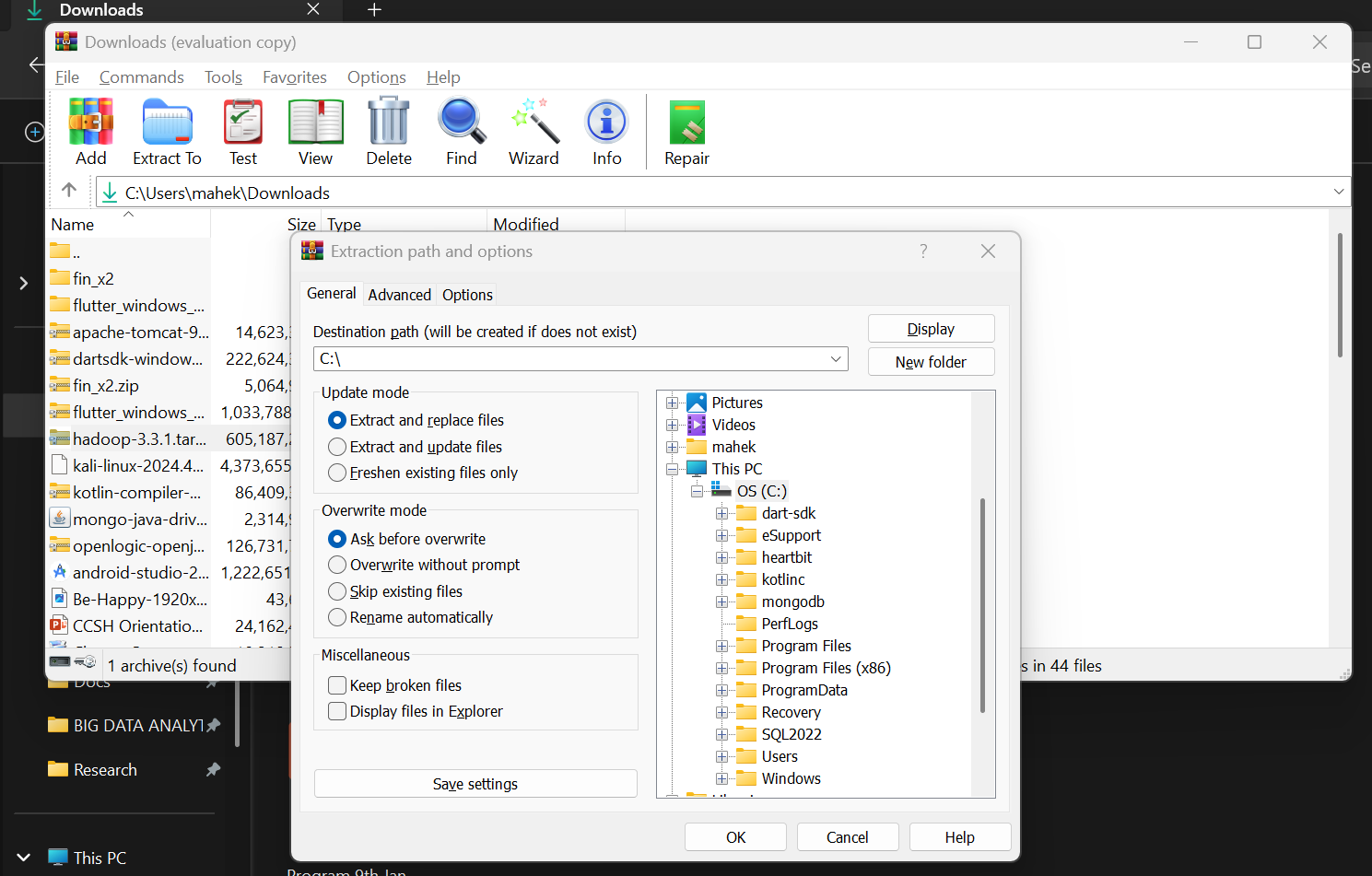
**PRACTICAL - 13**

**HADOOP INSTALLATION AND SETUP**

**INSTALLATION:**

**STEP I: *Install Hadoop 3.3.1 from the official website of Apache and extract in the C drive folder.***

****

****

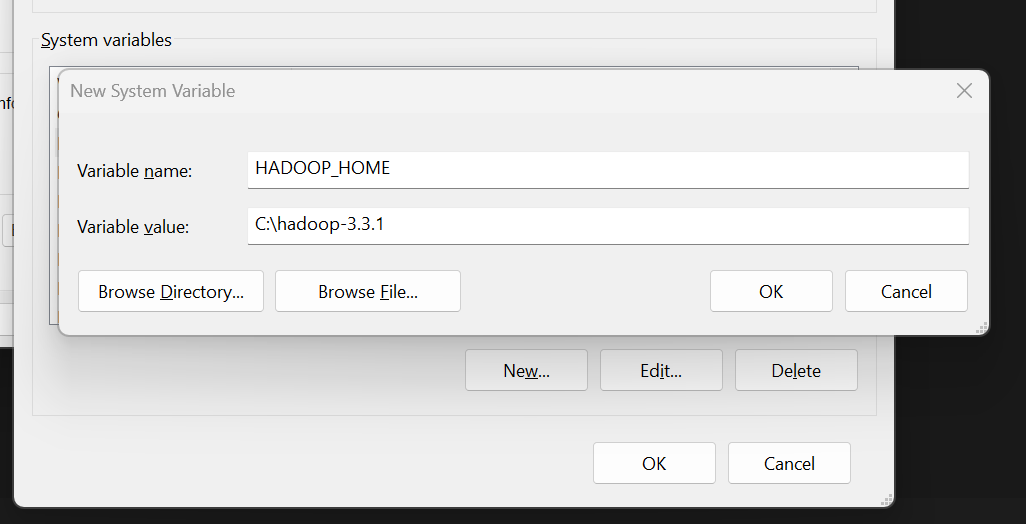
**STEP II: *Set path of HADOOP\_HOME as given below***

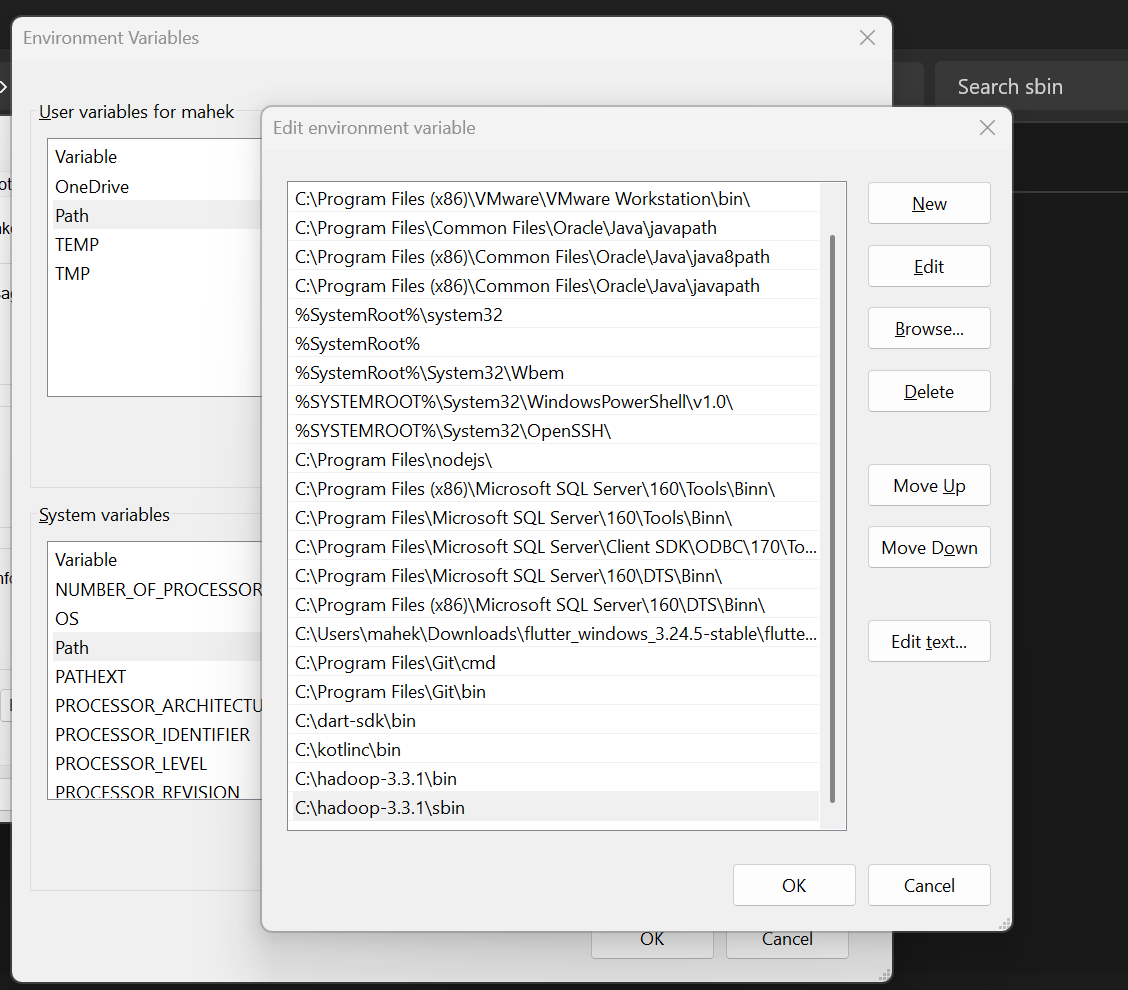
**System Variable : *C:\hadoop-3.3.1***

**Path (Set 2 path for bin and sbin) :**

***C:\hadoop-3.3.1\bin***

***C:\hadoop-3.3.1\sbin***

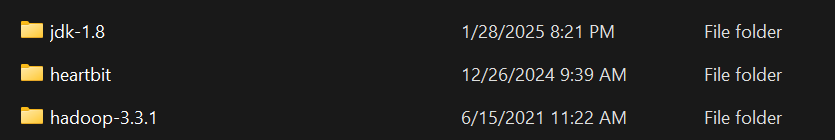
****

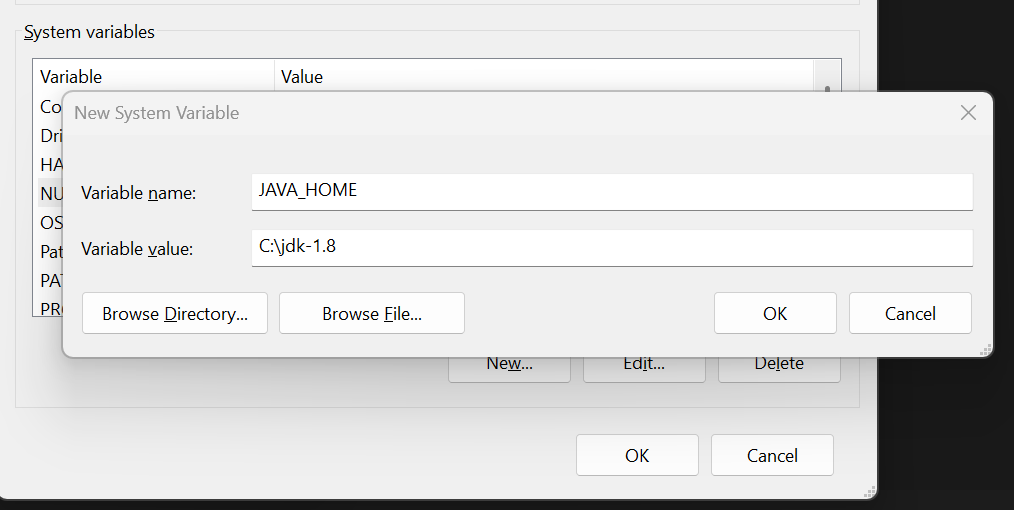
****

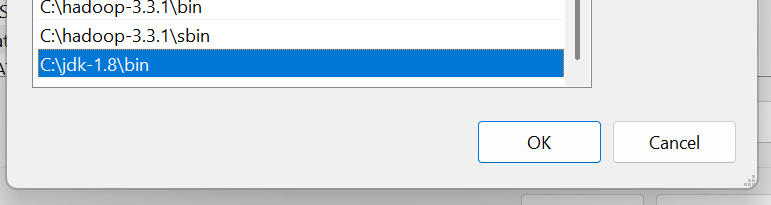
**STEP III: *Setup JAVA\_HOME path in the system environment variable. Make sure jdk is in the C drive exactly where the Hadoop 3.3.1 folder is extracted.***

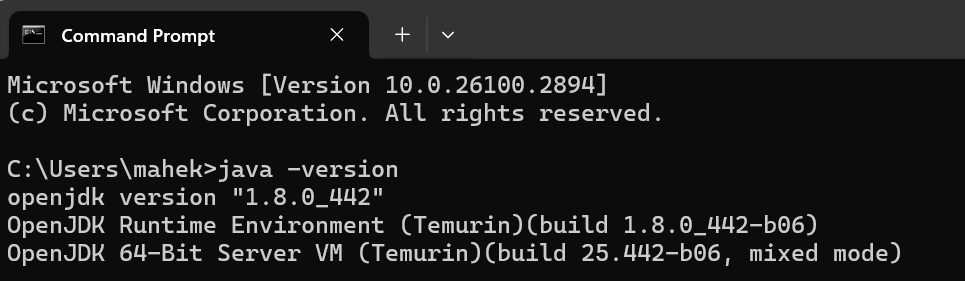
**JAVA\_HOME: *C:\jdk-1.8\***

**Path : *C:\jdk-1.8\bin***

******

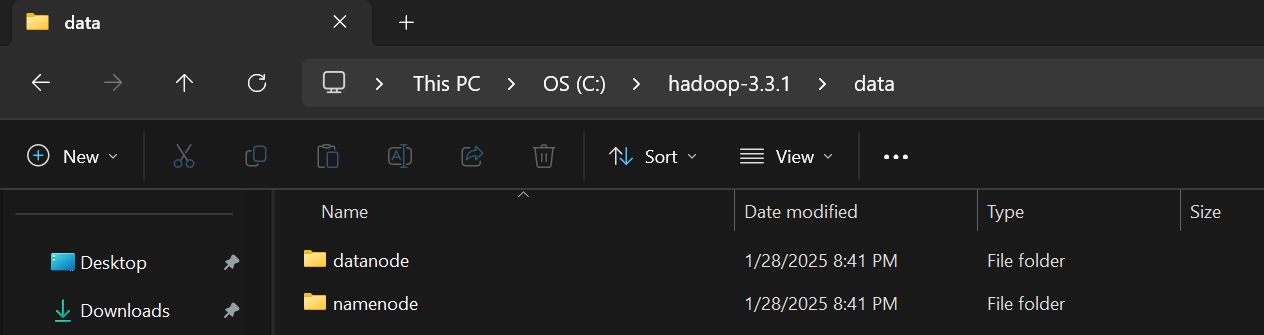
******

******

******

**STEP V: Make folder data in Hadoop 3.3.1 folder and make namenode and datanode folders in the newly created data folder.**

****

****

**STEP V: Now make the changes in the code in the files of location C:\hadoop-3.3.1\etc\hadoop as follows**

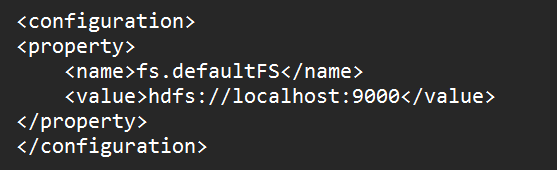
**core-site.xml**

<property>

<name>fs.defaultFS</name>

<value>hdfs://localhost:9000</value>

</property>

****

**hdfs-site.xml**:

<property>

<name>dfs.replication</name>

<value>1</value>

</property>

<property>

<name>dfs.namenode.name.dir</name>

<value>file:///C:/hadoop/tmp/dfs/name</value>

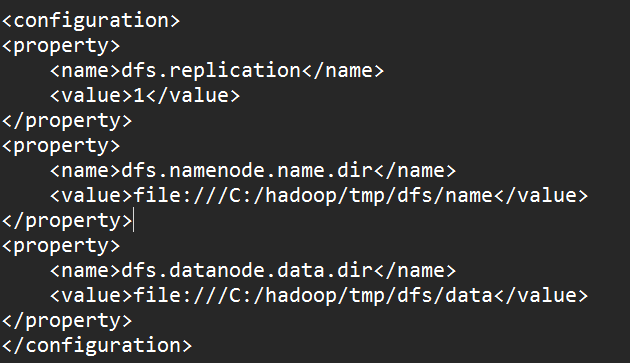
</property>

<property>

<name>dfs.datanode.data.dir</name>

<value>file:///C:/hadoop/tmp/dfs/data</value>

</property>



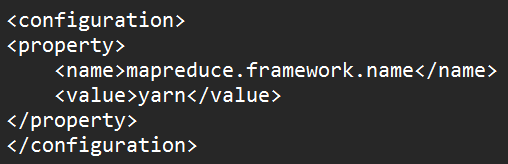
**mapred-site.xml**:

<property>

<name>mapreduce.framework.name</name>

<value>yarn</value>

</property>



**yarn-site.xml**:

<property>

<name>yarn.resourcemanager.address</name>

<value>localhost:8032</value>

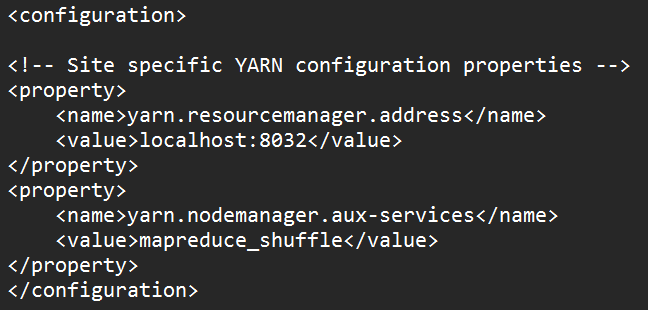
</property>

<property>

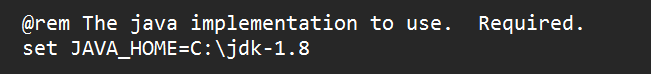
<name>yarn.nodemanager.aux-services</name>

<value>mapreduce\_shuffle</value>

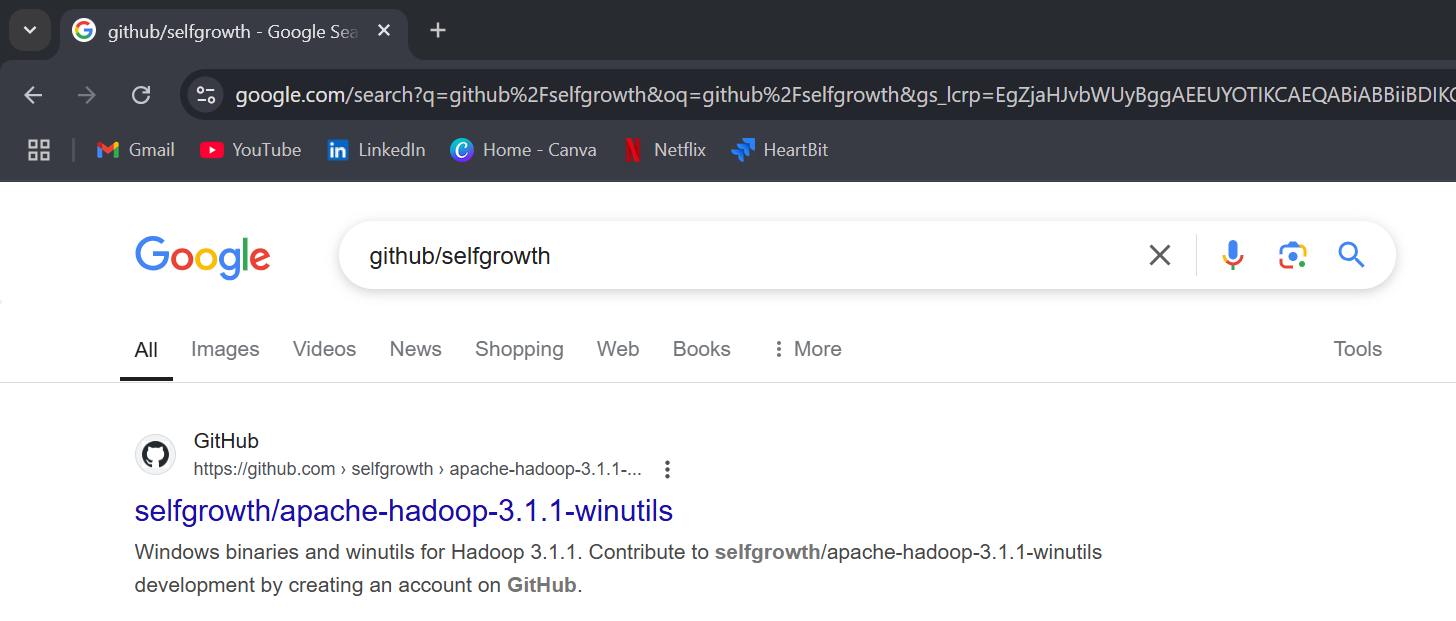
</property>

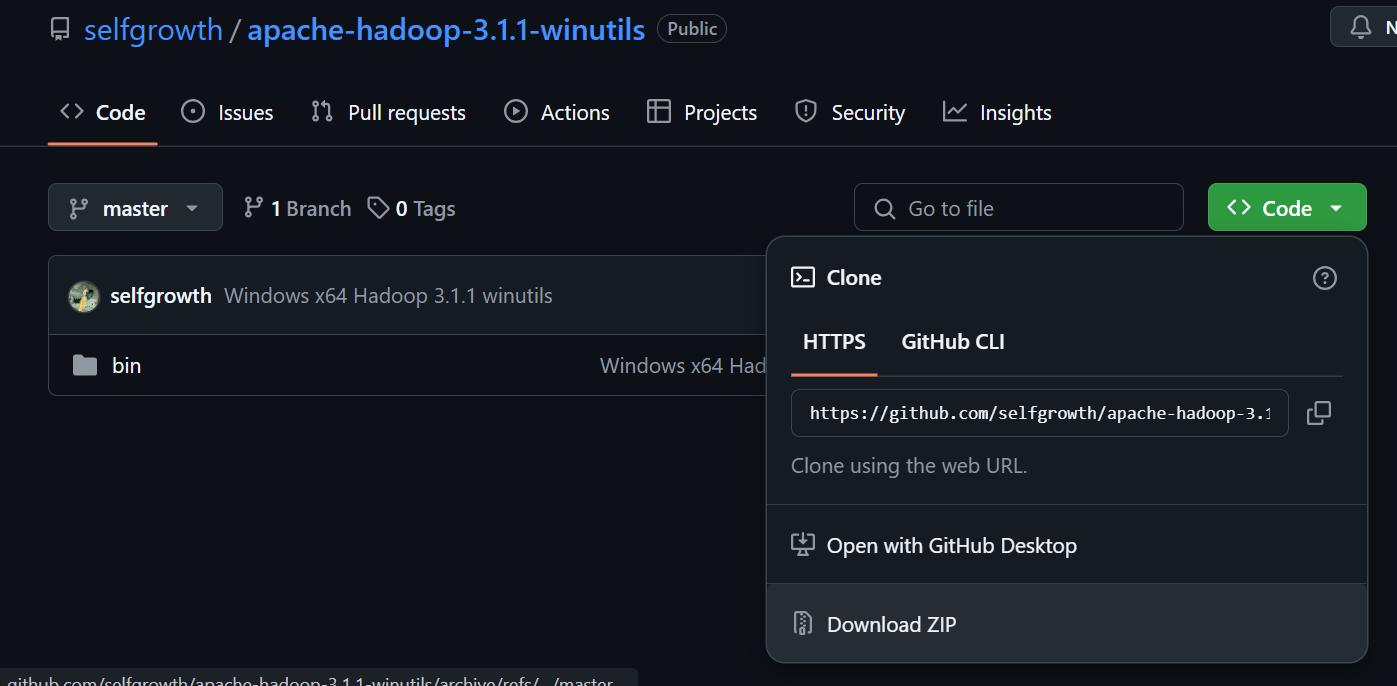


**hadoop-env.cmd : *Uncomment and add the JAVA\_HOME Path* in this file.**

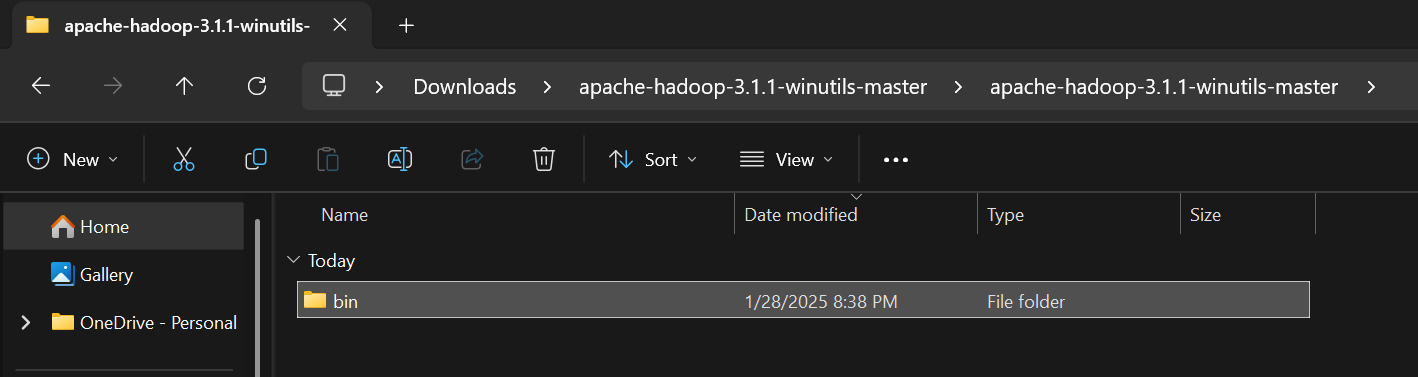


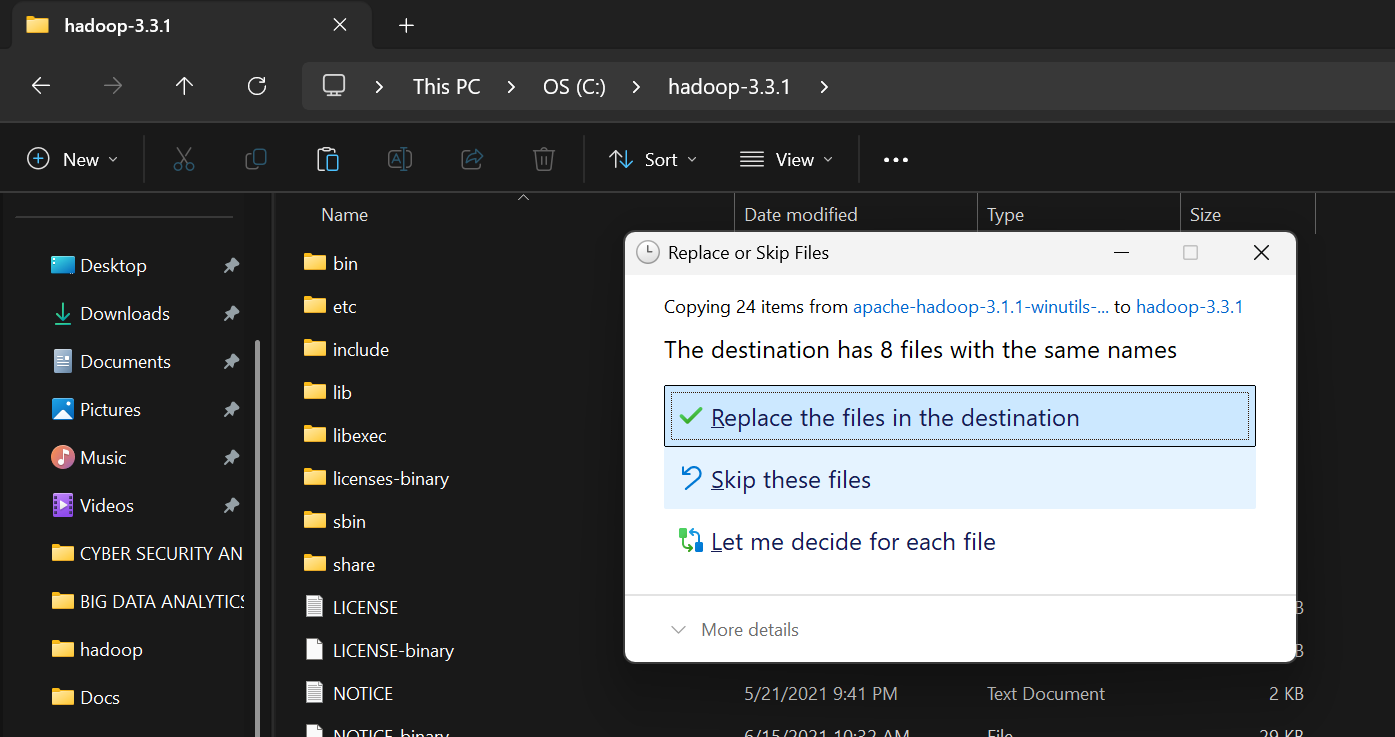
**STEP VI: *Download Hadoop utils for Hadoop 3.3.1 and replace the original bin with the new bin folder.***



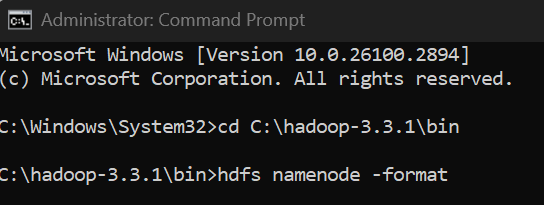


***Copy this bin and replace with original bin in Hadoop.3.3.1 folder.***

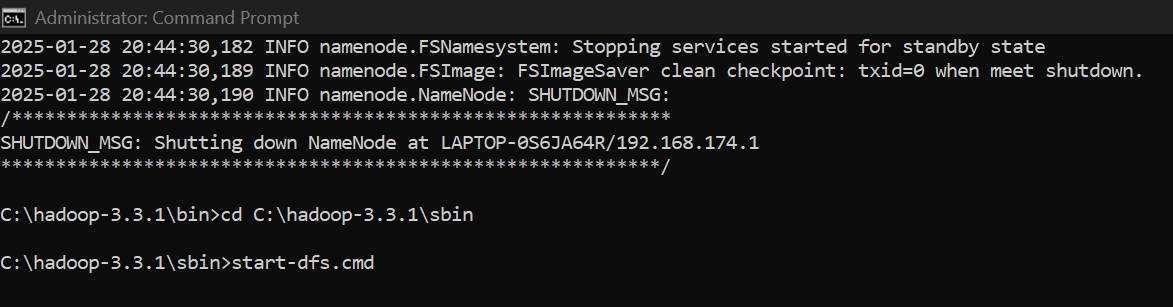




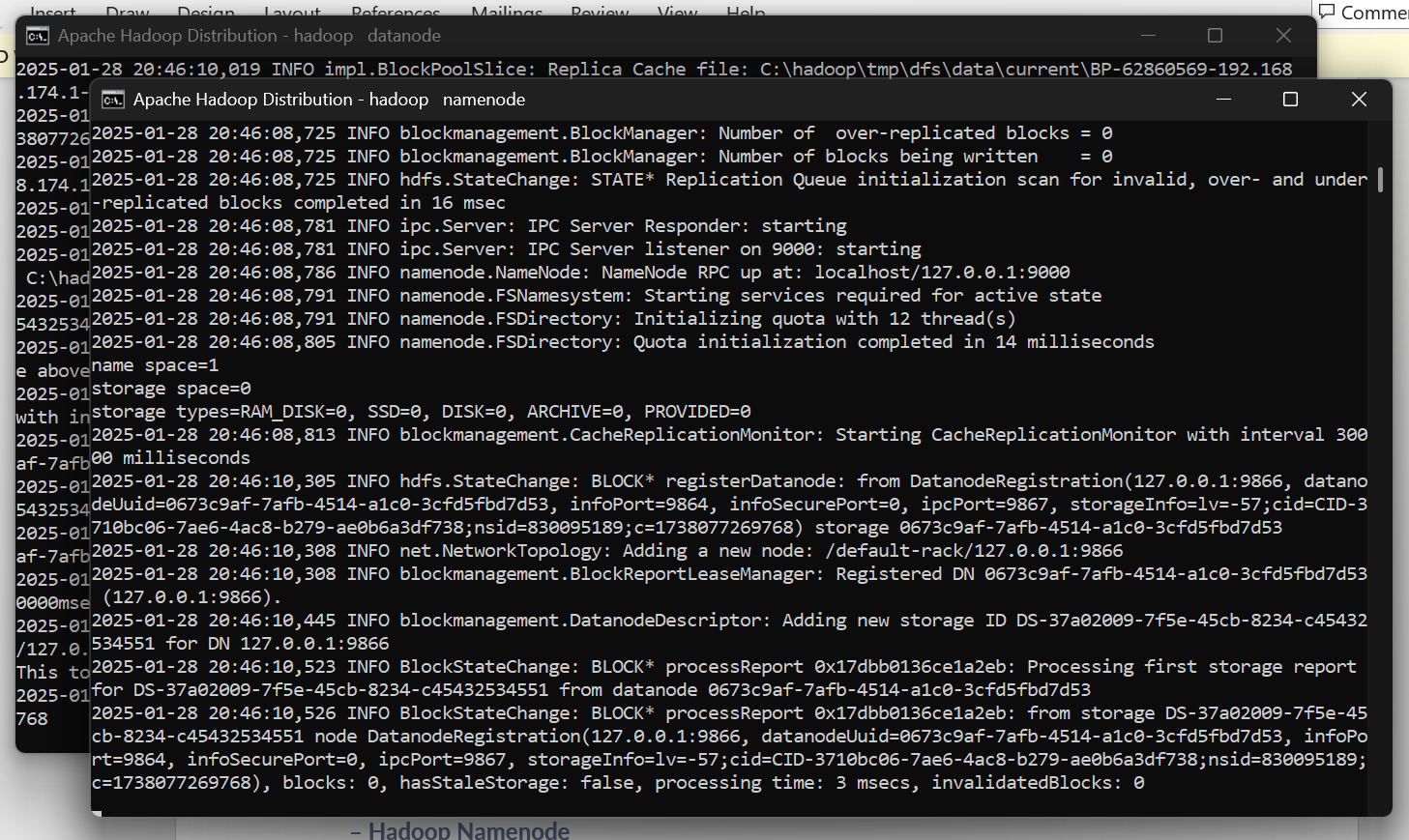
**STEP VII: *Open Command prompt as administrator and cd to the the bin path in Hadoop and type this command to format the name node.***

****

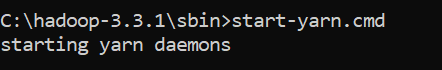
**STEP VIII: *Change directory to sbin and Start the dfs and yarn cmd.***

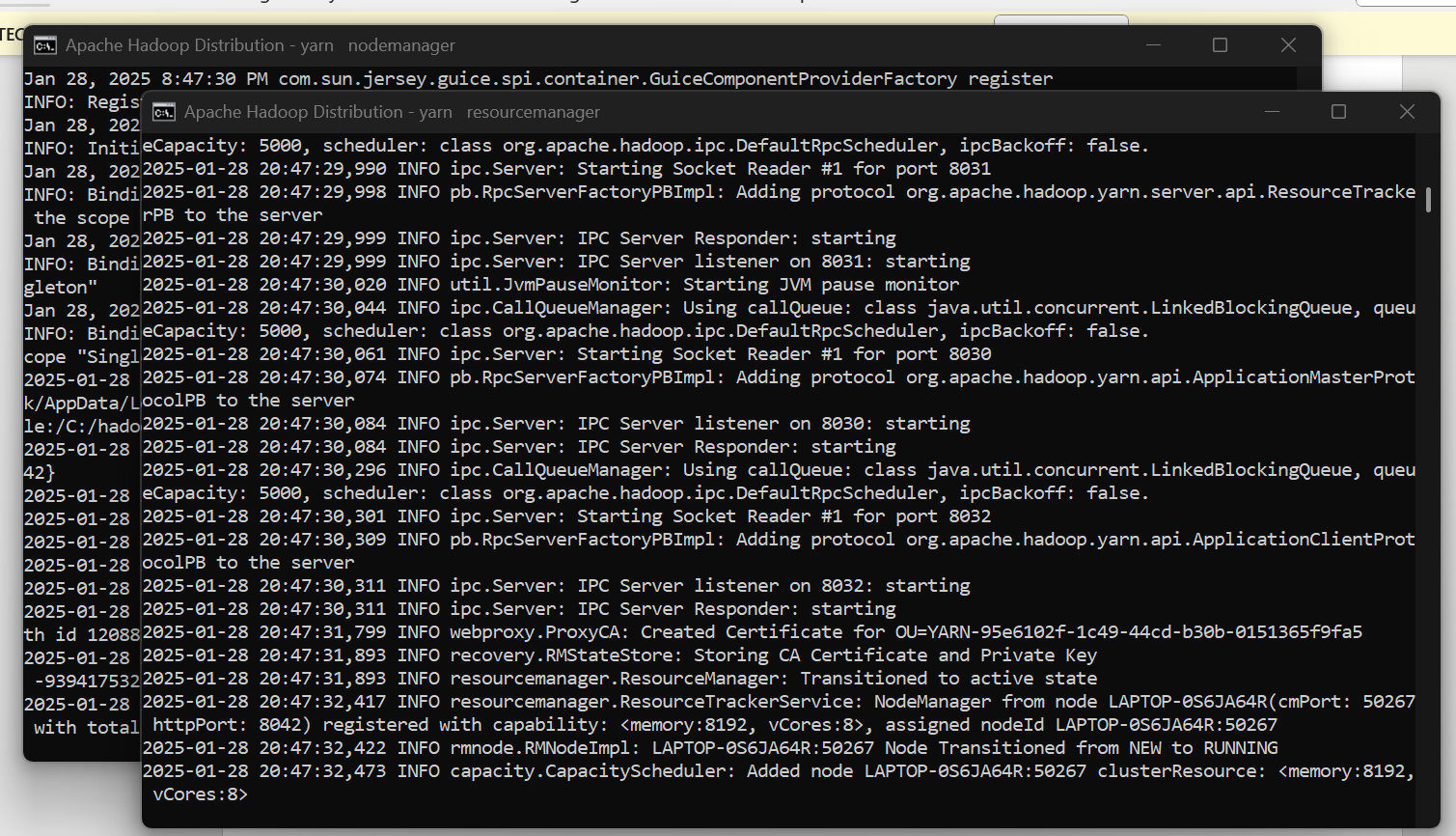
****

**These 2 windows will open**

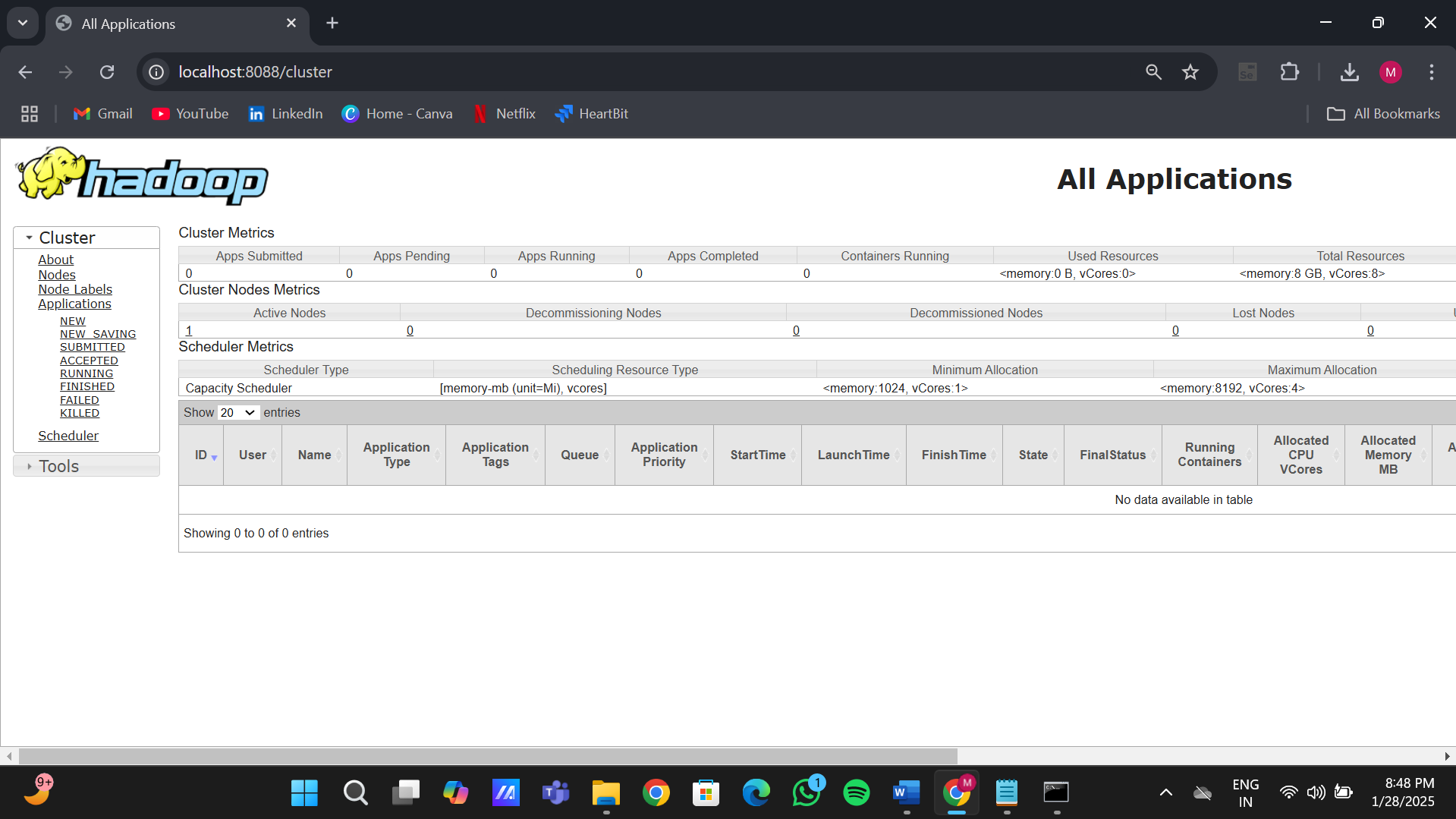
****

**STEP IX: *Start the yarn cmd* .**

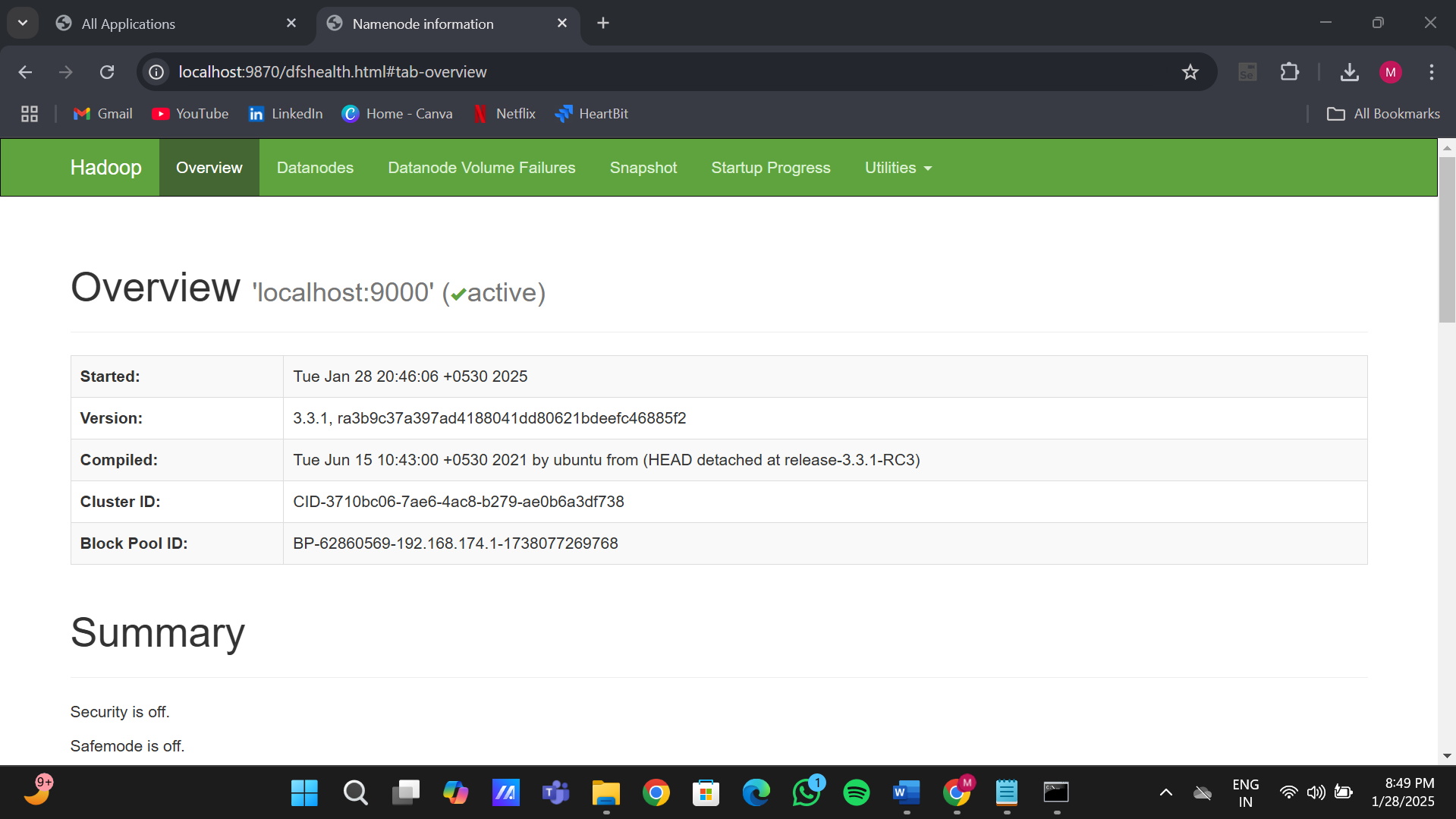
****

****

**STEP X: *Type localhost:8088 on web browser.***

****

**STEP XI: *Type localhost:9870 on web browser.***

****