**Steps to Install Hadoop**

* Install Java JDK 1.8
* Download Hadoop and extract and place under C drive
* Set Path in Environment Variables
* Config files under Hadoop directory
* Create folder datanode and namenode under data directory
* Edit HDFS and YARN files
* Set Java Home environment in Hadoop environment
* Setup Complete. Test by executing start-all.cmd

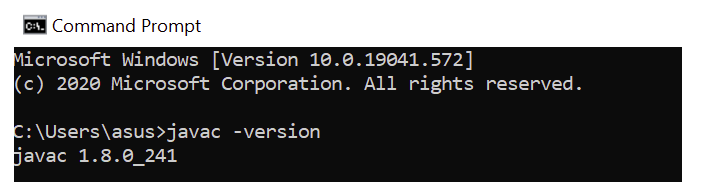
1. **Install Java**

**– Java JDK Link to download**

[**https://www.oracle.com/java/technologies/javase-jdk8-downloads.html**](https://www.oracle.com/java/technologies/javase-jdk8-downloads.html)

**– extract and install Java in C:\Java**

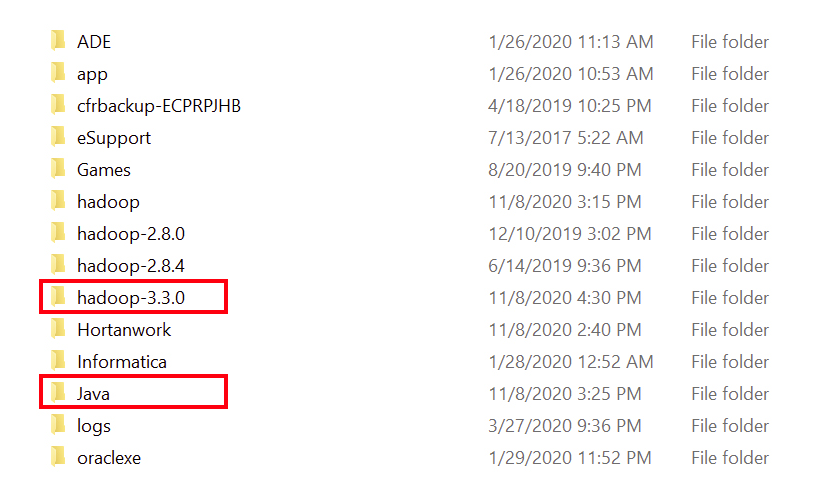
**– open cmd and type -> javac -version**



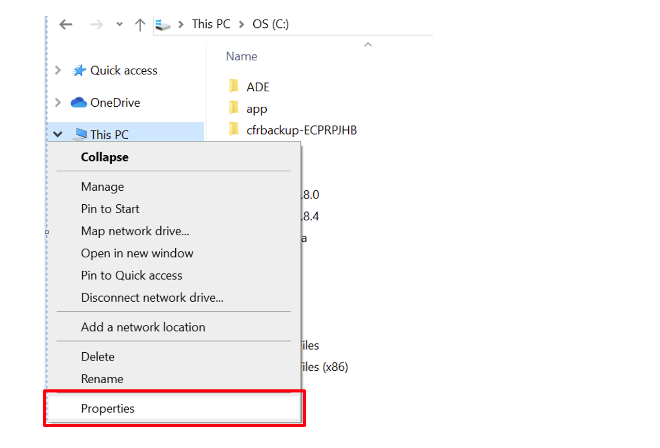
1. **Download Hadoop**

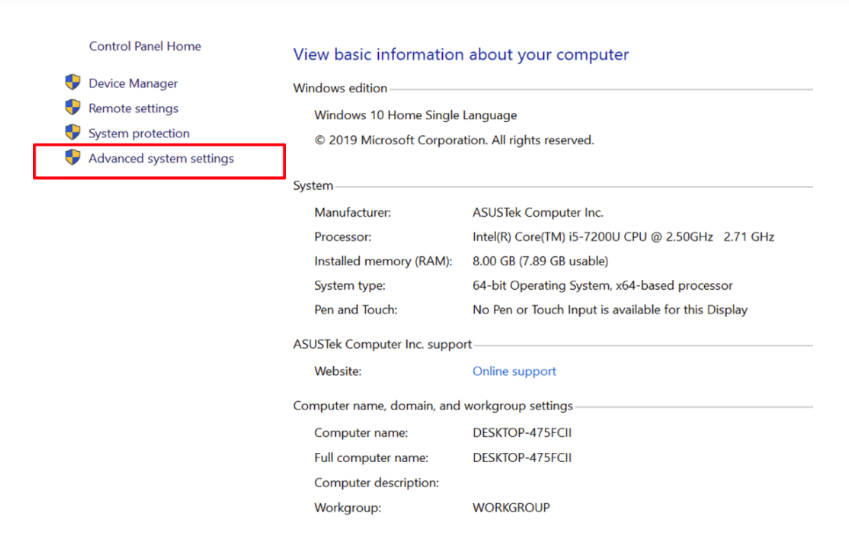
**–**[**https://www.apache.org/dyn/closer.cgi/hadoop/common/hadoop-3.3.0/hadoop-3.3.0.tar.gz**](https://www.apache.org/dyn/closer.cgi/hadoop/common/hadoop-3.3.0/hadoop-3.3.0.tar.gz)

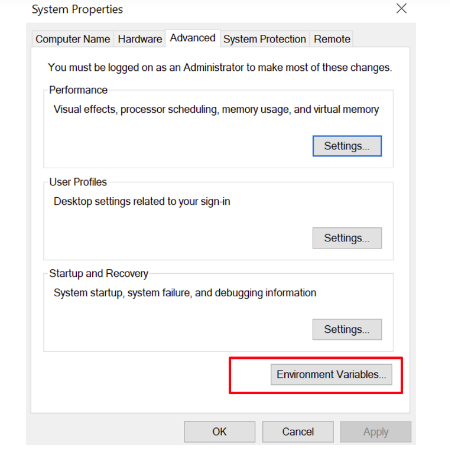
**– extract to C:\Hadoop**

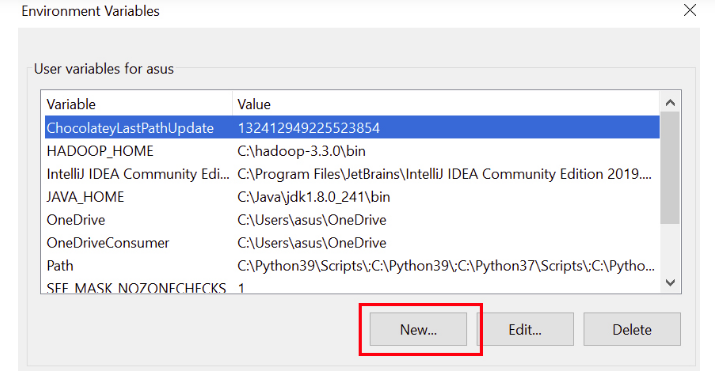


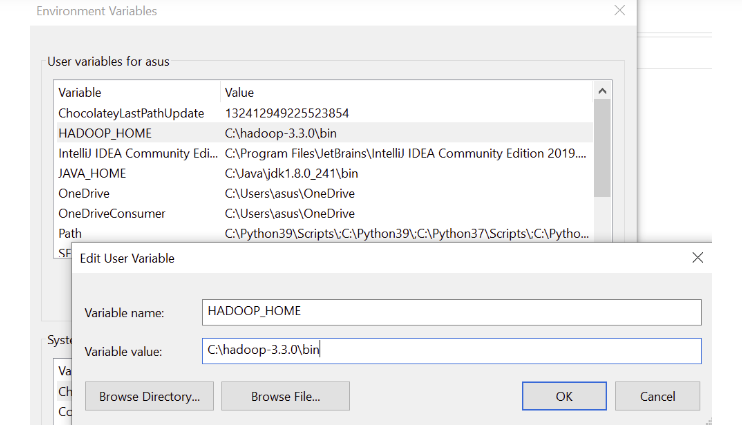
1. **Set the path JAVA\_HOME Environment variable**
2. **Set the path HADOOP\_HOME Environment variable**

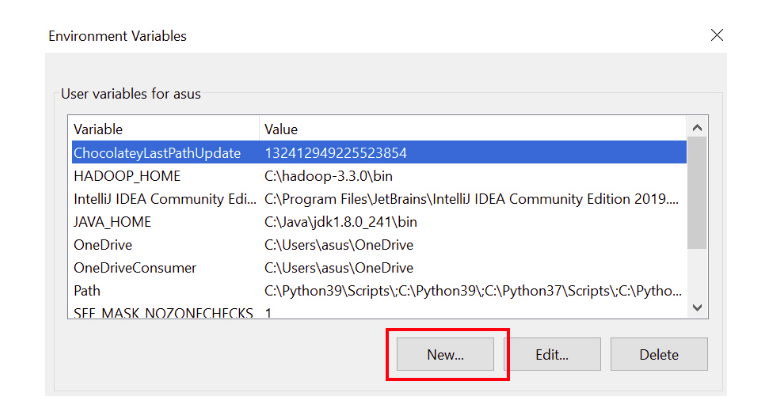


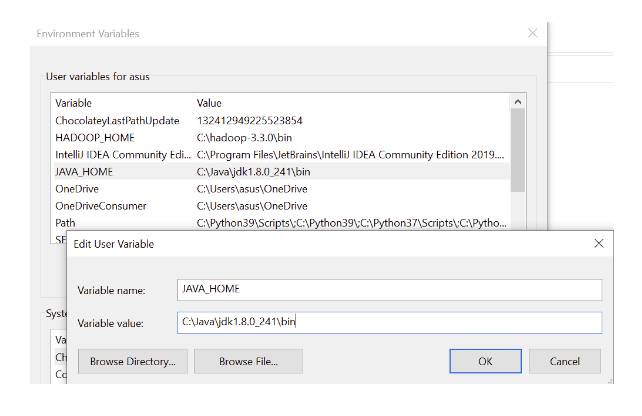


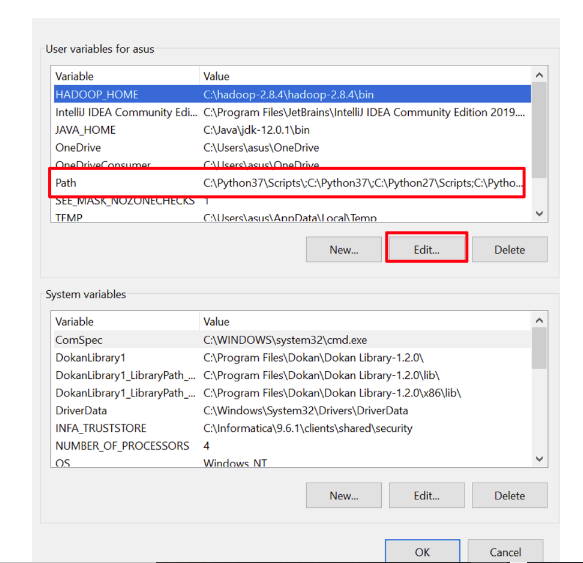


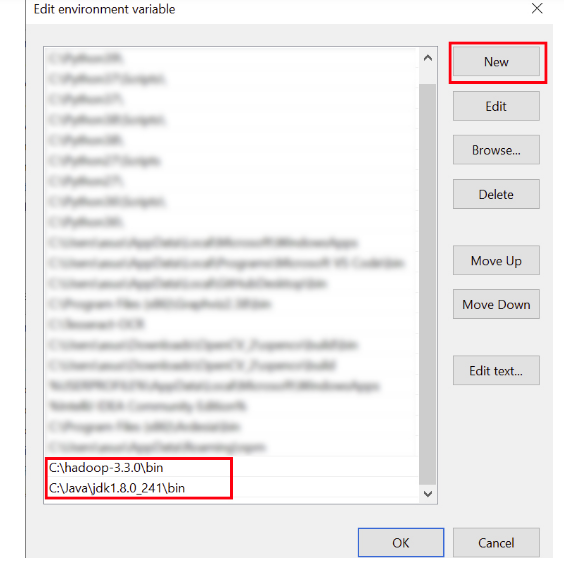












1. **Configurations**

**Edit file C:/Hadoop-3.3.0/etc/hadoop/core-site.xml,**

**paste the xml code in folder and save**

**<configuration>**

**<property>**

**<name>fs.defaultFS</name>**

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

**</property>**

**</configuration>**

**======================================================**

**Rename “mapred-site.xml.template” to “mapred-site.xml” and edit this file C:/Hadoop-3.3.0/etc/hadoop/mapred-site.xml, paste xml code and save this file.**

**<configuration>**

**<property>**

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

**<value>yarn</value>**

**</property>**

**</configuration>**

**======================================================**

**Create folder “data” under “C:\Hadoop-3.3.0”**

**Create folder “datanode” under “C:\Hadoop-3.3.0\data”**

**Create folder “namenode” under “C:\Hadoop-3.3.0\data”**

**======================================================**

**Edit file C:\Hadoop-3.3.0/etc/hadoop/hdfs-site.xml,**

**paste xml code and save this file.**

**<configuration>**

**<property>**

**<name>dfs.replication</name>**

**<value>1</value>**

**</property>**

**<property>**

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

**<value>/hadoop-3.3.0/data/namenode</value>**

**</property>**

**<property>**

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

**<value>/hadoop-3.3.0/data/datanode</value>**

**</property>**

**</configuration>**

**======================================================**

**Edit file C:/Hadoop-3.3.0/etc/hadoop/yarn-site.xml,**

**paste xml code and save this file.**

**<configuration>**

**<property>**

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

**<value>mapreduce\_shuffle</value>**

**</property>**

**<property>**

**<name>yarn.nodemanager.auxservices.mapreduce.shuffle.class</name>**

**<value>org.apache.hadoop.mapred.ShuffleHandler</value>**

**</property>**

**</configuration>**

**======================================================**

**Edit file C:/Hadoop-3.3.0/etc/hadoop/hadoop-env.cmd**

**by closing the command line**

**“JAVA\_HOME=%JAVA\_HOME%” instead of set “JAVA\_HOME=C:\Java”**

**======================================================**

1. **Hadoop Configurations**

**Download**

[**https://github.com/brainmentorspvtltd/BigData\_RDE/blob/master/Hadoop%20Configuration.zip**](https://github.com/brainmentorspvtltd/BigData_RDE/blob/master/Hadoop%20Configuration.zip)

**or (for hadoop 3)**

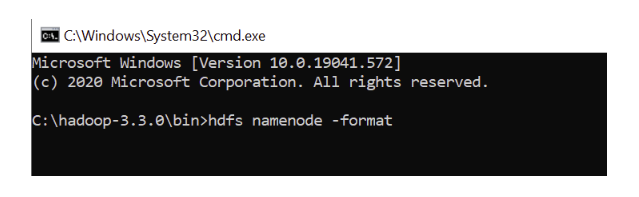
**https://github.com/s911415/apache-hadoop-3.1.0-winutils**

**– Copy folder bin and replace existing bin folder in**

**C:\Hadoop-3.3.0\bin**

**– Format the NameNode**

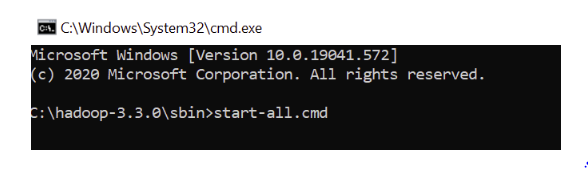
**– Open cmd and type command “hdfs namenode –format”**



1. **Testing**

**– Open cmd and change directory to C:\Hadoop-3.3.0\sbin**

**– type start-all.cmd**



**(Or you can start like this)**

**– Start namenode and datanode with this command**

**– type start-dfs.cmd**

**– Start yarn through this command**

**– type start-yarn.cmd**

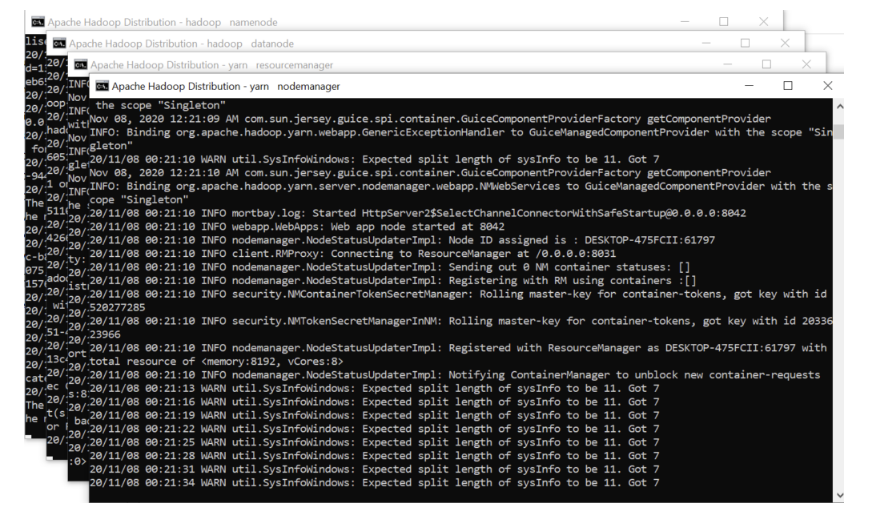
**Make sure these apps are running**

**– Hadoop Namenode**

**– Hadoop datanode**

**– YARN Resource Manager**

**– YARN Node Manager**



**Open:**[**http://localhost:8088**](http://localhost:8088/)

