

EX NO:1	DOWNLOADING AND INSTALLING HADOOP, UNDERSTANDING DIFFERENT HADOOP MODES. STARTUP SCRIPTS, CONFIGURATION FILES.
DATE:	

**AIM:**

To Downloading and installing Hadoop; Understanding different Hadoop modes. Startup scripts, Configuration files.

**PREREQUISITES TO INSTALL HADOOP ON WINDOWS**

**VIRTUAL BOX** (For Linux): it is used for installing the operating system on it.

**OPERATING SYSTEM:** You can install Hadoop on Windows or Linux based operating systems. Ubuntu and CentOS are very commonly used.

**JAVA:** You need to install the Java 8 package on your system.

**HADOOP:** You require Hadoop latest version

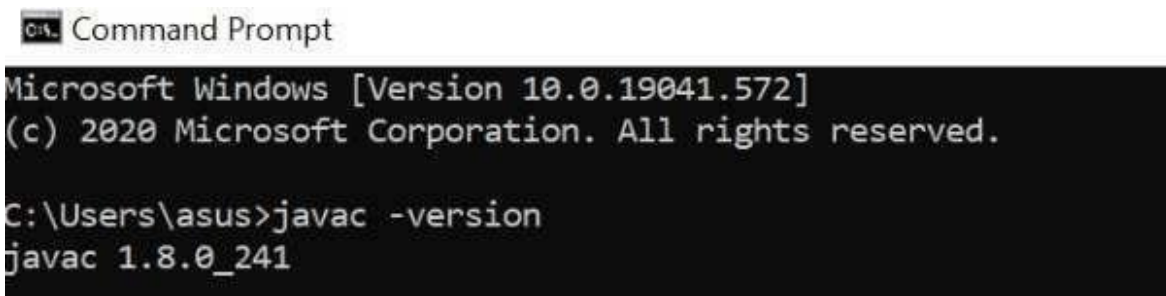
**1. Install Java**

Java JDK Link to download\_

<https://www.oracle.com/java/technologies/javase-jdk8-downloads.html>

Extract and install Java in C:\Java

Open cmd and type -> javac -version



```

C:\> Command Prompt
Microsoft Windows [Version 10.0.19041.572]
(c) 2020 Microsoft Corporation. All rights reserved.

C:\Users\asus>javac -version
javac 1.8.0_241

```

**2. Download Hadoop**

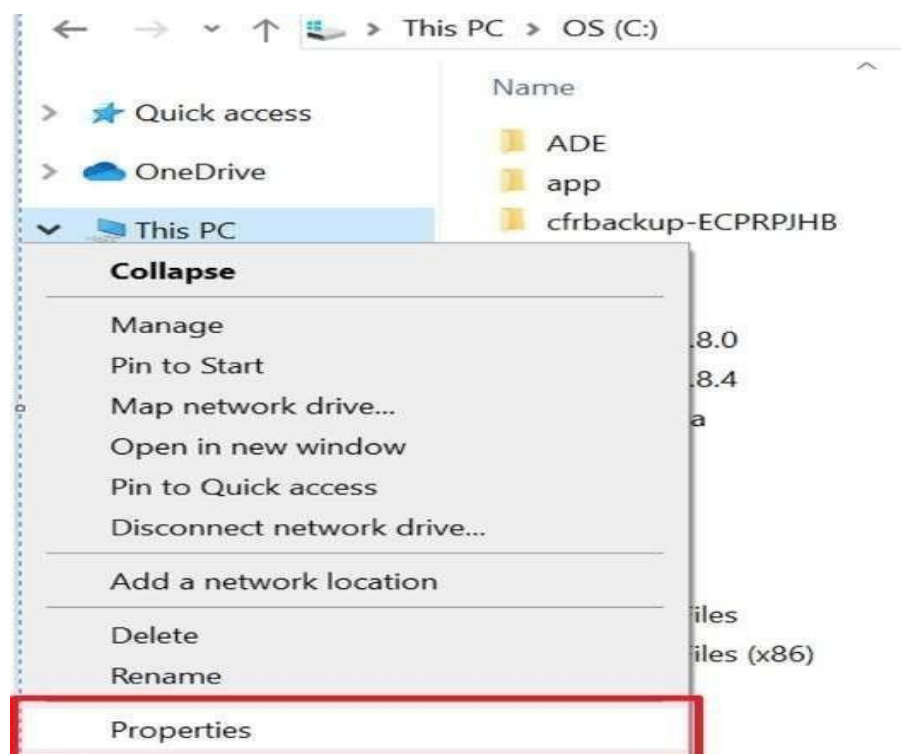
<https://www.apache.org/dyn/closer.cgi/hadoop/common/hadoop-3.3.0/hadoop-3.3.0.tar.gz>

extract to C:\Hadoop

Find your jdk-1.8 folder in C:\program files\java\jdk-1.8 and Put your jdk-1.8 file in a separate folder ;  
C:\Java\jdk-1.8

ADE	1/26/2020 11:13 AM	File folder
app	1/26/2020 10:53 AM	File folder
cfrbackup-ECPRPJHB	4/18/2019 10:25 PM	File folder
eSupport	7/13/2017 5:22 AM	File folder
Games	8/20/2019 9:40 PM	File folder
hadoop	11/8/2020 3:15 PM	File folder
hadoop-2.8.0	12/10/2019 3:02 PM	File folder
hadoop-2.8.4	6/14/2019 9:36 PM	File folder
hadoop-3.3.0	11/8/2020 4:30 PM	File folder
Hortanwork	11/8/2020 2:40 PM	File folder
Informatica	1/28/2020 12:52 AM	File folder
Java	11/8/2020 3:25 PM	File folder
logs	3/27/2020 9:36 PM	File folder
oraclexe	1/29/2020 11:52 PM	File folder

3. Set the path JAVA\_HOME Environment variable
4. Set the path HADOOP\_HOME Environment variable



Control Panel Home

Device Manager  
Remote settings  
System protection  
**Advanced system settings**

### View basic information about your computer

Windows edition

Windows 10 Home Single Language

© 2019 Microsoft Corporation. All rights reserved.

#### System

Manufacturer:	ASUSTek Computer Inc.
Processor:	Intel(R) Core(TM) i5-7200U CPU @ 2.50GHz 2.71 GHz
Installed memory (RAM):	8.00 GB (7.89 GB usable)
System type:	64-bit Operating System, x64-based processor
Pen and Touch:	No Pen or Touch Input is available for this Display

ASUSTek Computer Inc. support

Website: [Online support](#)

#### Computer name, domain, and workgroup settings

Computer name:	DESKTOP-475FCII
Full computer name:	DESKTOP-475FCII
Computer description:	
Workgroup:	WORKGROUP

1

System Properties

Computer Name Hardware **Advanced** System Protection Remote

You must be logged on as an Administrator to make most of these changes.

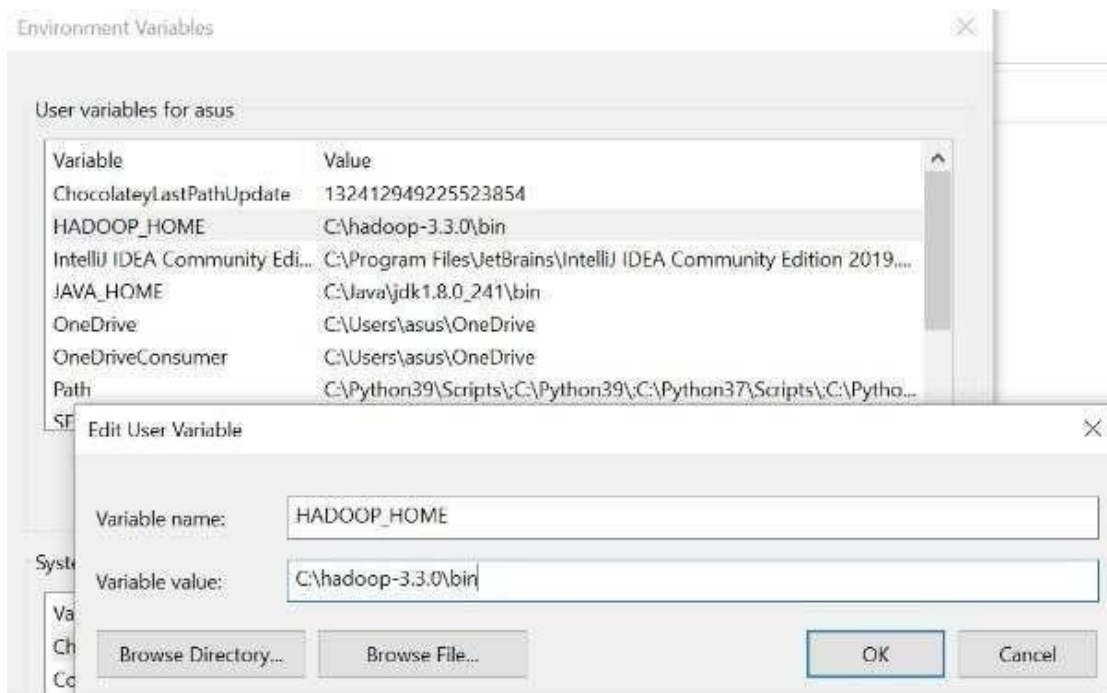
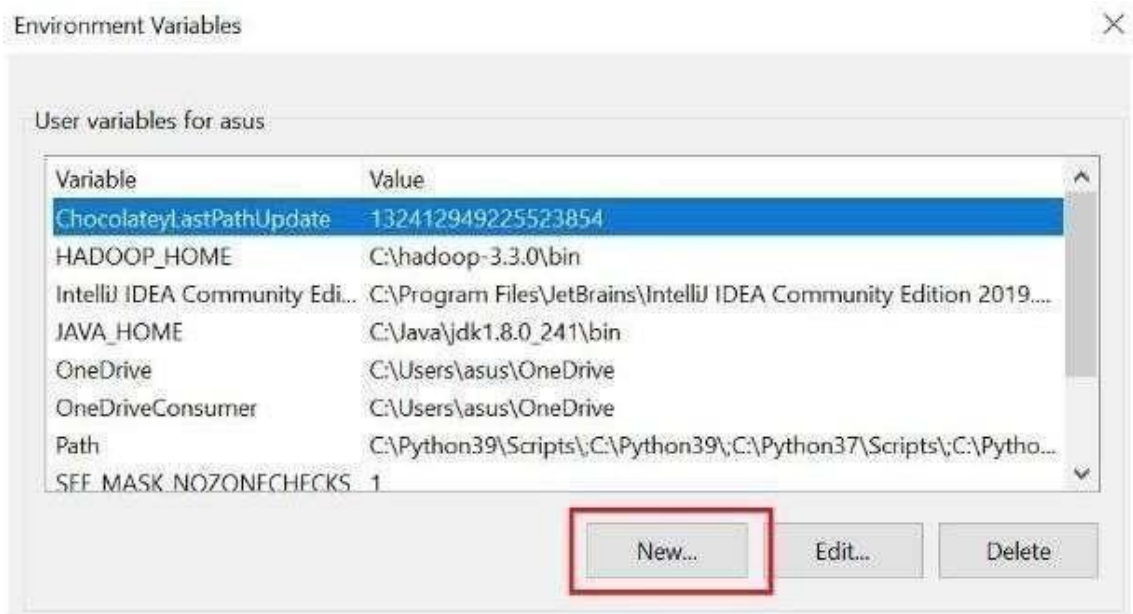
**Performance**  
Visual effects, processor scheduling, memory usage, and virtual memory  
[Settings...](#)

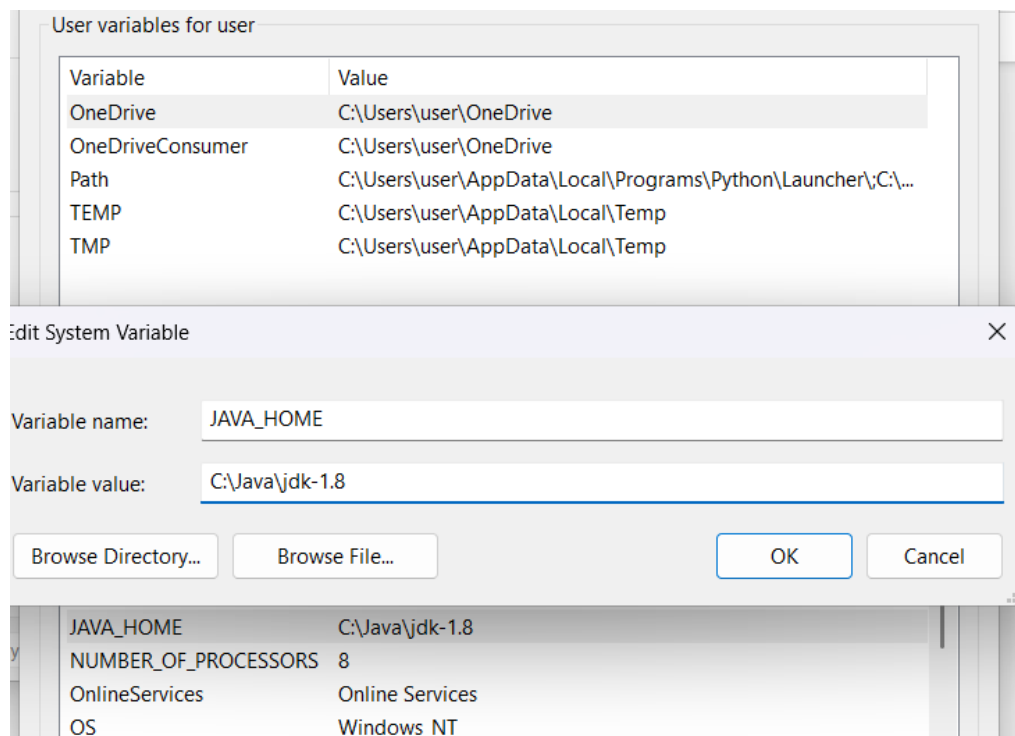
**User Profiles**  
Desktop settings related to your sign-in  
[Settings...](#)

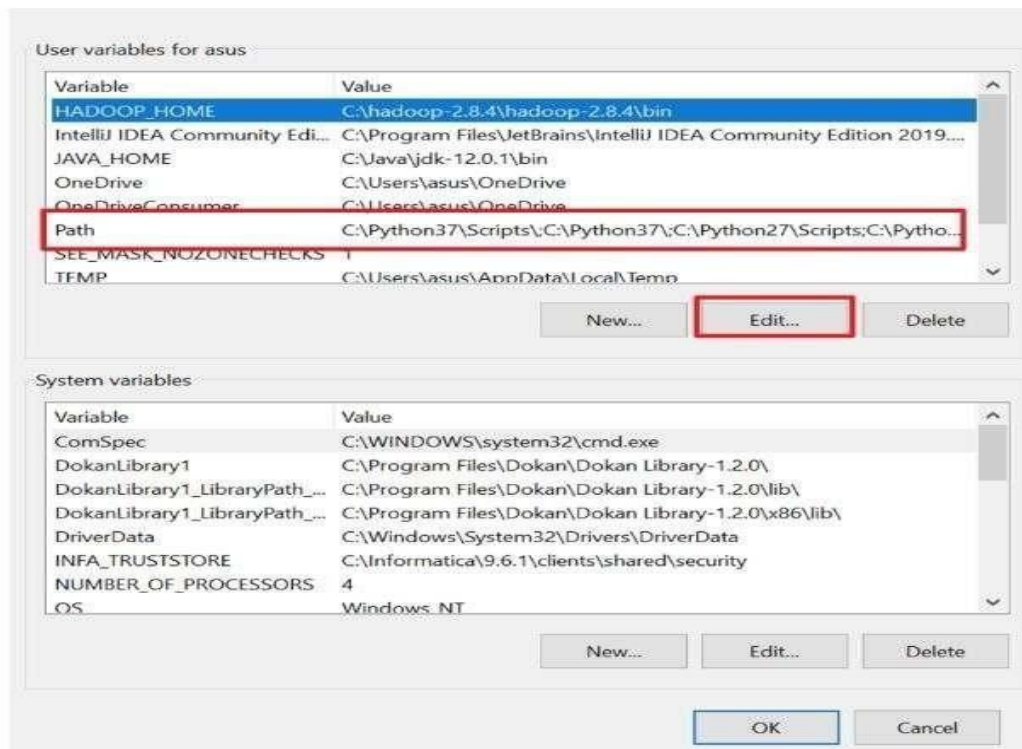
**Startup and Recovery**  
System startup, system failure, and debugging information  
[Settings...](#)

[Environment Variables...](#)

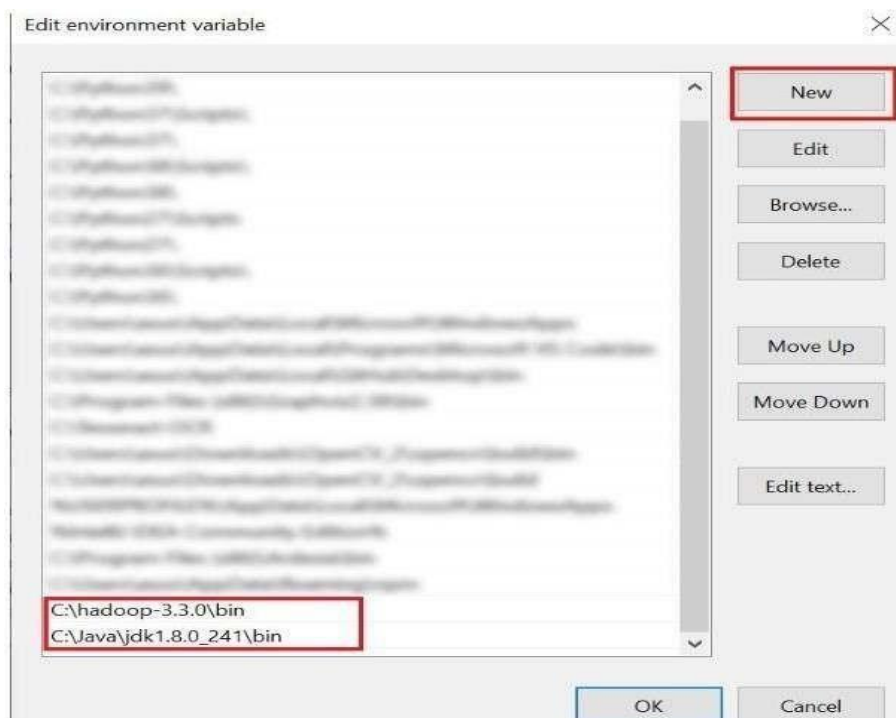
OK Cancel Apply







Also add the C:\hadoop-3.3.0\sbin in path



## 5. Configurations

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

```
<configuration>
<property>
<name>fs.defaultFS</name>
<value>hdfs://localhost:9000</value>
</property>
</configuration>
```

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

```
<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/mapred-site.xml,

```
<configuration>
<property>
<name>mapreduce.framework.name</name>
<value>yarn</value>
</property>
</configuration>
```

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

```
<configuration>
  <property>
    <name>yarn.nodemanager.aux-services</name>
    <value>mapreduce_shuffle</value>
  </property>
  <property>
    <name>yarn.timeline-service.enabled</name>
    <value>>false</value>
  </property>
</configuration>
```

6.Edit file hadoop-env.cmd:

C:/Hadoop-3.3.0/etc/hadoop/hadoop-env.cmd by closing the command line “JAVA HOME=%JAVA HOME%” instead of set ;

```
@rem The java implementation to use. Required.  
set JAVA_HOME=C:\Java\jdk-1.8
```

Also add these line ad the end of the file:

```
set HADOOP_CLASSPATH=%HADOOP_CLASSPATH%;%HADOOP_HOME%\share  
\hadoop\yarn\timelineservice\*
```

#### 6.1.Hadoop Configurations:

Download <https://github.com/s911415/apache-hadoop-3.1.0-winutil> Copy folder bin and replace existing bin folder in  
C:\Hadoop-3.3.0\bin

Now open command prompt as a administrator:

Change to the bin folder in your Hadoop: C:/Hadoop-3.3.0/bin

Type the command:

*Hdfs namenode -format*



```
C:\Windows\System32\cmd.exe
Microsoft Windows [Version 10.0.19041.572]
(c) 2020 Microsoft Corporation. All rights reserved.

C:\hadoop-3.3.0\bin>hdfs namenode -format
```

## 7. Testing

- Open cmd and change directory to C:\Hadoop-3.3.0\sbin
- type start-all.cmd

```
C:\Windows\System32\cmd.exe
Microsoft Windows [Version 10.0.19041.572]
(c) 2020 Microsoft Corporation. All rights reserved.

C:\hadoop-3.3.0\sbin>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

## 45

[illegible]

**Open:** <http://localhost:8088>



**Open:** <http://localhost:9870>

The screenshot shows a web browser window with the address bar displaying 'localhost:9870/dfshealth.html#tab-overview'. The page has a green navigation bar with tabs: 'Hadoop', 'Overview', 'Datanodes', 'Datanode Volume Failures', 'Snapshot', 'Startup Progress', and 'Utilities'. The 'Overview' tab is selected. Below the navigation bar, the title 'Overview 'localhost:9000' (✓active)' is displayed. A table provides details about the Hadoop installation:

Started:	Sun Nov 08 16:53:46 +0530 2020
Version:	3.3.0 [REDACTED] 9af
Compiled:	Tue Jul 07 00:14:00 +0530 2020 by brahma from branch-3.3.0
Cluster ID:	C [REDACTED]
Block Pool ID:	B [REDACTED] 14

Below the table, the heading 'Summary' is followed by the text 'Hadoop installed Successfully.....'.

**RESULT:**

Downloaded and installed Hadoop and also understand different Hadoop modes. Startup scripts, Configuration files are successfully implemented.