

# INTRODUCTION TO BIG DATA

ECAP456

**Dr. Rajni Bhalla**  
Associate Professor

# Learning Outcomes



After this lecture, you will be able to

- learn how to install java
- learn how to install hadoop

# Prerequisites to install Hadoop



# Prerequisites to install Hadoop



# Steps to install Java

The screenshot shows the Oracle Java Downloads page for Java 17. The browser's address bar shows the URL `oracle.com/java/technologies/downloads/`. The Oracle logo is in the top left, and navigation links for Products, Industries, Resources, Support, Events, Developer, and Partners are in the top right. A search icon is also present. Below the navigation bar, there is a section titled "Looking for other Java downloads?" with two buttons: "OpenJDK Early Access Builds" and "JRE for Consumers". The main content area features a heading "Java 17 available now" followed by a paragraph about Java 17 LTS and a button "Learn about Java SE Subscription". Below this is a section titled "Java SE Development Kit 17.0.1 downloads" with a paragraph about the JDK and another paragraph about the tools included. At the bottom, there are tabs for "Linux", "macOS", and "Windows", with "Linux" selected. Below the tabs is a table with three columns: "Product/file description", "File size", and "Download".

Java Downloads | Oracle

oracle.com/java/technologies/downloads/

Apps bible stories 1276.pdf Find journals | Elsev... Smote journal with price Before You Write | I... Inderscience Publis... Other bookmarks

ORACLE

Products Industries Resources Support Events Developer Partners

Looking for other Java downloads?

OpenJDK Early Access Builds JRE for Consumers

## Java 17 available now

Java 17 LTS is the latest long-term support release for the Java SE platform. JDK 17 binaries are free to use in production and free to redistribute, at no cost, under the [Oracle No-Fee Terms and Conditions](#).

JDK 17 will receive updates under these terms, until at least September 2024.

[Learn about Java SE Subscription](#)

## Java SE Development Kit 17.0.1 downloads



Thank you for downloading this release of the Java™ Platform, Standard Edition Development Kit (JDK™). The JDK is a development environment for building applications and components using the Java programming language.

The JDK includes tools for developing and testing programs written in the Java programming language and running on the Java platform.

Linux macOS Windows

Product/file description	File size	Download
Arm 64 Compressed Archive	171.13 MB	<a href="https://download.oracle.com/java/17/latest/jdk-17_linux-aarch64_bin.tar.gz">https://download.oracle.com/java/17/latest/jdk-17_linux-aarch64_bin.tar.gz</a> (sha256 <a href="#">[link]</a> )

# Steps to Install Java

Linux	macOS	Solaris	Windows
Product/file description		File size	Download
x86 Installer		157.37 MB	 <a href="#">jdk-8u311-windows-i586.exe</a>
x64 Installer		170.57 MB	 <a href="#">jdk-8u311-windows-x64.exe</a>


# Steps to Install Java

You must accept the [Oracle Technology Network License Agreement for Oracle Java SE](#) to download this software.

---

☒ I reviewed and accept the Oracle Technology Network License Agreement for Oracle Java SE  
Required


*You will be redirected to the login screen in order to download the file.*

Download `jdk-8u311-windows-x64.exe` 


# Steps to Install Java

## Oracle account sign in

Username



Password



**Sign in**

[Need help?](#)

**Don't have an Oracle Account?**

Create Account



# Steps to Install Java



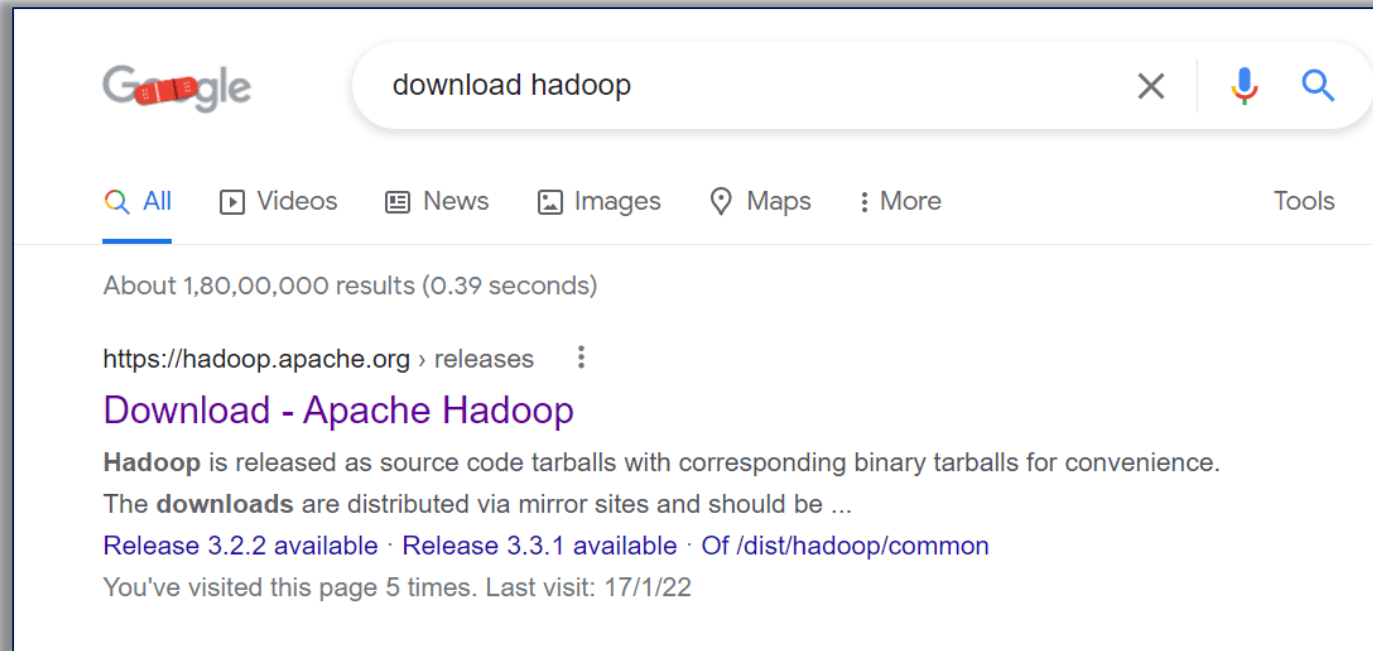
jdk-8u311-window....exe

48.5/171 MB, 56 secs left



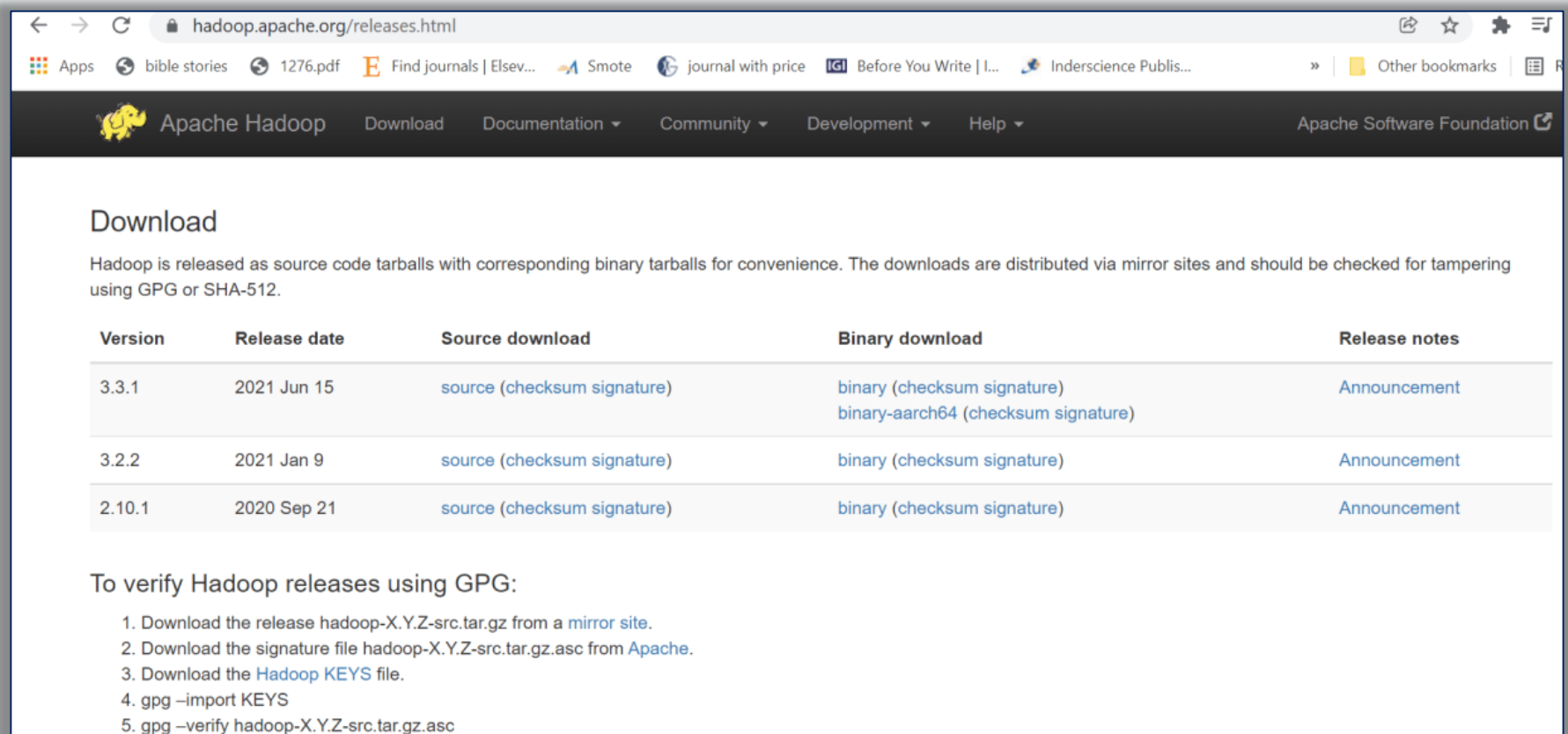
jdk-8u311-window....exe

# Steps to install hadoop





<https://hadoop.apache.org> ›

# Steps to install hadoop

A screenshot of a web browser displaying the Apache Hadoop releases page. The browser's address bar shows 'hadoop.apache.org/releases.html'. The page has a dark navigation bar with the Apache Hadoop logo and links for Download, Documentation, Community, Development, and Help. The main content area is titled 'Download' and contains a paragraph explaining that Hadoop is released as source code tarballs with corresponding binary tarballs. Below this is a table with five columns: Version, Release date, Source download, Binary download, and Release notes. The table lists three versions: 3.3.1 (2021 Jun 15), 3.2.2 (2021 Jan 9), and 2.10.1 (2020 Sep 21). Each version row provides links for source and binary downloads and a link to the release notes. Below the table, a section titled 'To verify Hadoop releases using GPG:' lists five steps for verifying the releases using GPG.

← → ↻ 🔒 hadoop.apache.org/releases.html

Apps bible stories 1276.pdf Find journals | Elsev... Smote journal with price Before You Write | I... Inderscience Publis... » Other bookmarks

 Apache Hadoop Download Documentation ▾ Community ▾ Development ▾ Help ▾ Apache Software Foundation 

## Download

Hadoop is released as source code tarballs with corresponding binary tarballs for convenience. The downloads are distributed via mirror sites and should be checked for tampering using GPG or SHA-512.

Version	Release date	Source download	Binary download	Release notes
3.3.1	2021 Jun 15	<a href="#">source (checksum signature)</a>	<a href="#">binary (checksum signature)</a> <a href="#">binary-aarch64 (checksum signature)</a>	<a href="#">Announcement</a>
3.2.2	2021 Jan 9	<a href="#">source (checksum signature)</a>	<a href="#">binary (checksum signature)</a>	<a href="#">Announcement</a>
2.10.1	2020 Sep 21	<a href="#">source (checksum signature)</a>	<a href="#">binary (checksum signature)</a>	<a href="#">Announcement</a>

### To verify Hadoop releases using GPG:

1. Download the release `hadoop-X.Y.Z-src.tar.gz` from a [mirror site](#).
2. Download the signature file `hadoop-X.Y.Z-src.tar.gz.asc` from [Apache](#).
3. Download the [Hadoop KEYS](#) file.
4. `gpg --import KEYS`
5. `gpg --verify hadoop-X.Y.Z-src.tar.gz.asc`

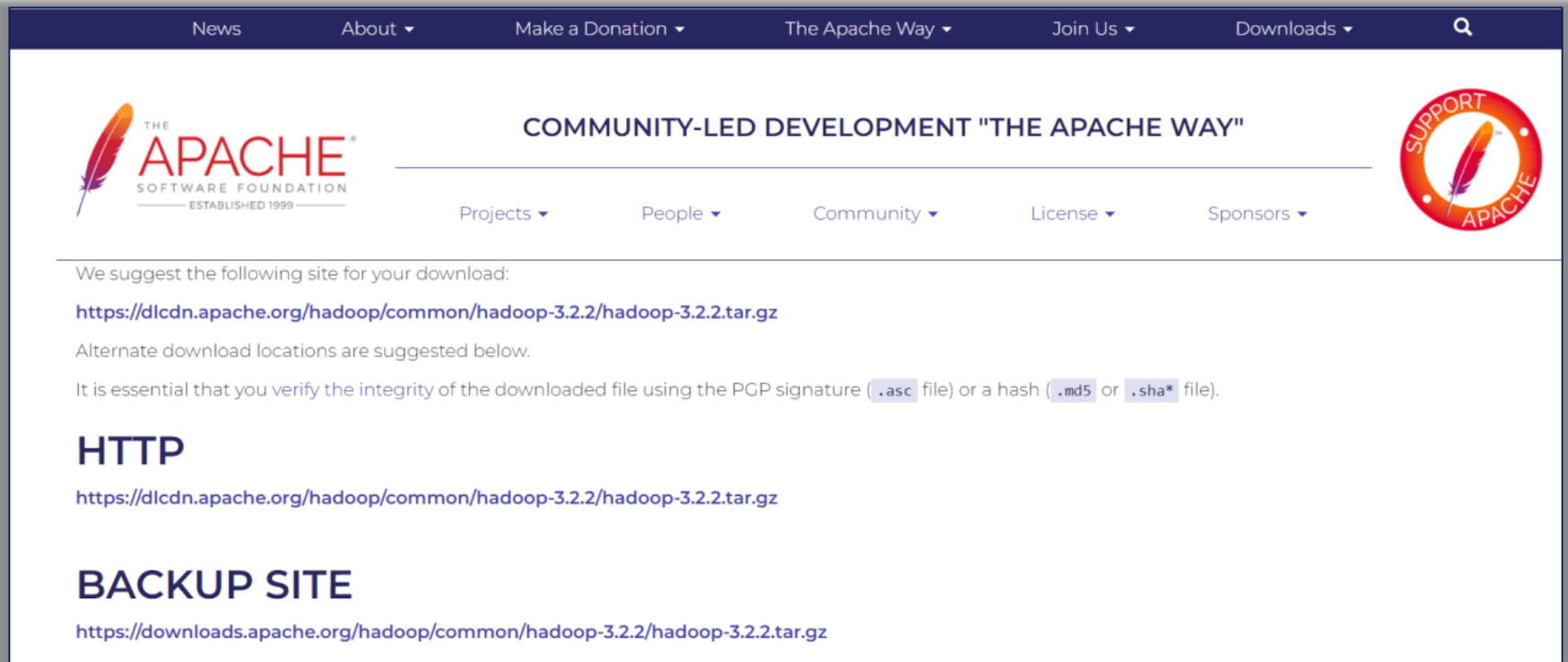
# Steps to install hadoop

## Download

Hadoop is released as source code tarballs with corresponding binary tarballs for convenience. The downloads are distributed via mirror sites and should be checked for tampering using GPG or SHA-512.



Version	Release date	Source download	Binary download	Release notes
3.3.1	2021 Jun 15	<a href="#">source (checksum signature)</a>	<a href="#">binary (checksum signature)</a> <a href="#">binary-aarch64 (checksum signature)</a>	<a href="#">Announcement</a>
✓ 3.2.2	2021 Jan 9	<a href="#">source (checksum signature)</a>	<a href="#">binary (checksum signature)</a>	<a href="#">Announcement</a>
2.10.1	2020 Sep 21	<a href="#">source (checksum signature)</a>	<a href="#">binary (checksum signature)</a>	<a href="#">Announcement</a>

# Steps to install hadoop



The screenshot shows the Apache Software Foundation website. The top navigation bar includes links for News, About, Make a Donation, The Apache Way, Join Us, and Downloads. The main header features the Apache logo on the left, the text "COMMUNITY-LED DEVELOPMENT 'THE APACHE WAY'" in the center, and a circular "SUPPORT APACHE" logo on the right. Below the header, there are links for Projects, People, Community, License, and Sponsors. The main content area provides instructions for downloading Hadoop 3.2.2, including a direct download link and alternate locations. It also mentions the importance of verifying the file's integrity using PGP signatures or hashes. The page is divided into sections for HTTP and BACKUP SITE, each with a corresponding download link.

News About ▾ Make a Donation ▾ The Apache Way ▾ Join Us ▾ Downloads ▾

 **COMMUNITY-LED DEVELOPMENT "THE APACHE WAY"** 

Projects ▾ People ▾ Community ▾ License ▾ Sponsors ▾

We suggest the following site for your download:

<https://dlcdn.apache.org/hadoop/common/hadoop-3.2.2/hadoop-3.2.2.tar.gz>

Alternate download locations are suggested below.

It is essential that you verify the integrity of the downloaded file using the PGP signature (`.asc` file) or a hash (`.md5` or `.sha*` file).

## HTTP


<https://dlcdn.apache.org/hadoop/common/hadoop-3.2.2/hadoop-3.2.2.tar.gz>

## BACKUP SITE

<https://downloads.apache.org/hadoop/common/hadoop-3.2.2/hadoop-3.2.2.tar.gz>


# Steps to install hadoop

[News](#) [About ▾](#) [Make a Donation ▾](#) [The Apache Way ▾](#) [Join Us ▾](#) [Downloads ▾](#) [🔍](#)



COMMUNITY-LED DEVELOPMENT "THE APACHE WAY"

[Projects ▾](#) [People ▾](#) [Community ▾](#) [License ▾](#) [Sponsors ▾](#)



We suggest the following site for your download:

<https://dlcdn.apache.org/hadoop/common/hadoop-3.2.2/hadoop-3.2.2.tar.gz>

Alternate download locations are suggested below.

It is essential that you verify the integrity of the downloaded file using the PGP signature (`.asc` file) or a hash (`.md5` or `.sha*` file).

## HTTP

<https://dlcdn.apache.org/hadoop/common/hadoop-3.2.2/hadoop-3.2.2.tar.gz>

## BACKUP SITE

<https://downloads.apache.org/hadoop/common/hadoop-3.2.2/hadoop-3.2.2.tar.gz>

# Steps to install hadoop



hadoop-3.2.2 (1).tar.gz

11.6/377 MB, 3 mins left





hadoop-3.2.2 (1).tar.gz

215/377 MB, 1 min left



hadoop-3.2.2 (1).tar.gz

# Steps to install hadoop

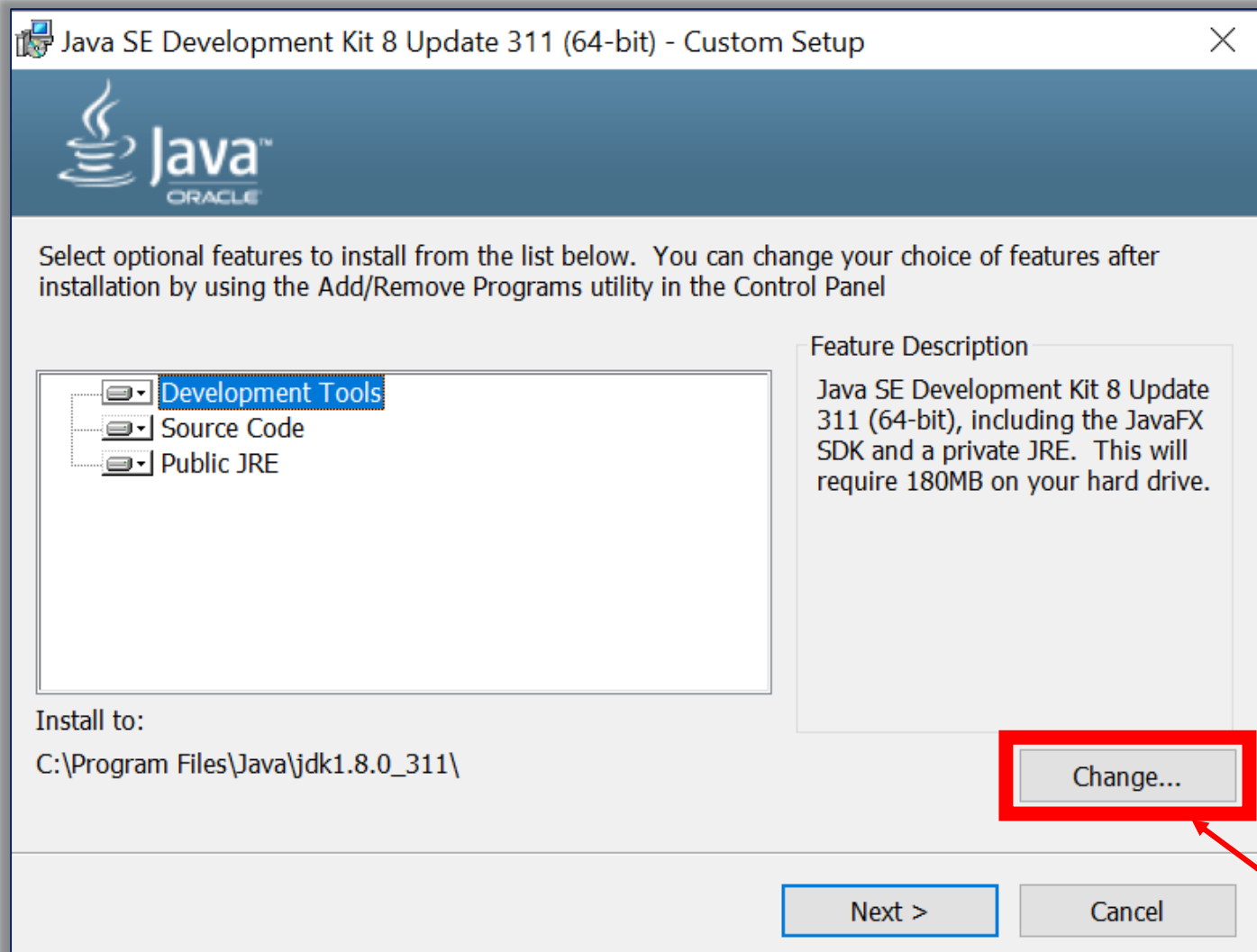
Name	Date modified	Type	Size
 jdk-8u311-windows-x64	17-01-2022 13:07	Application	1,74,667 KB
 hadoop-3.2.2 (1).tar	17-01-2022 13:52	GZip File	3,86,181 KB



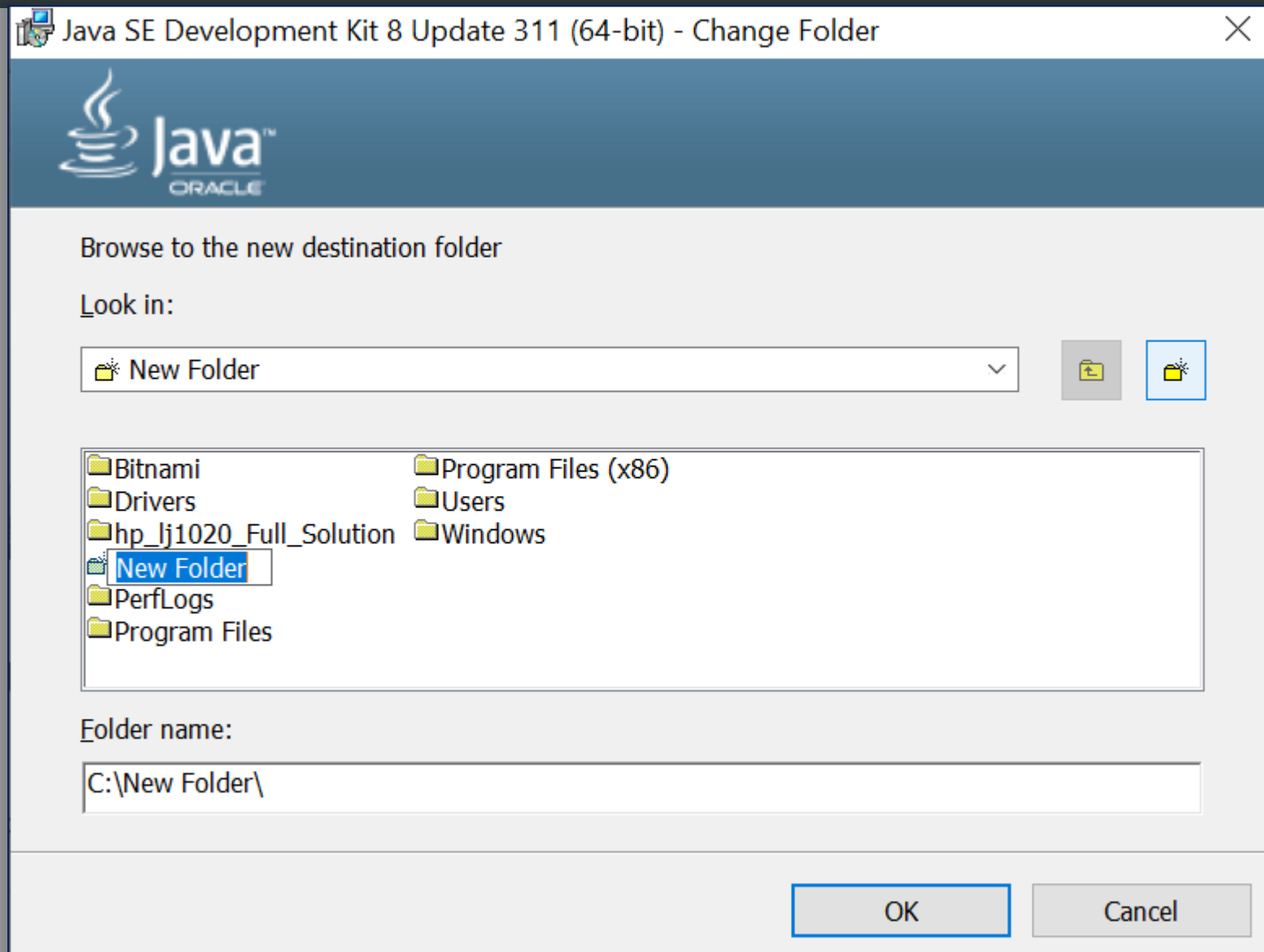
# Installation of Java

---

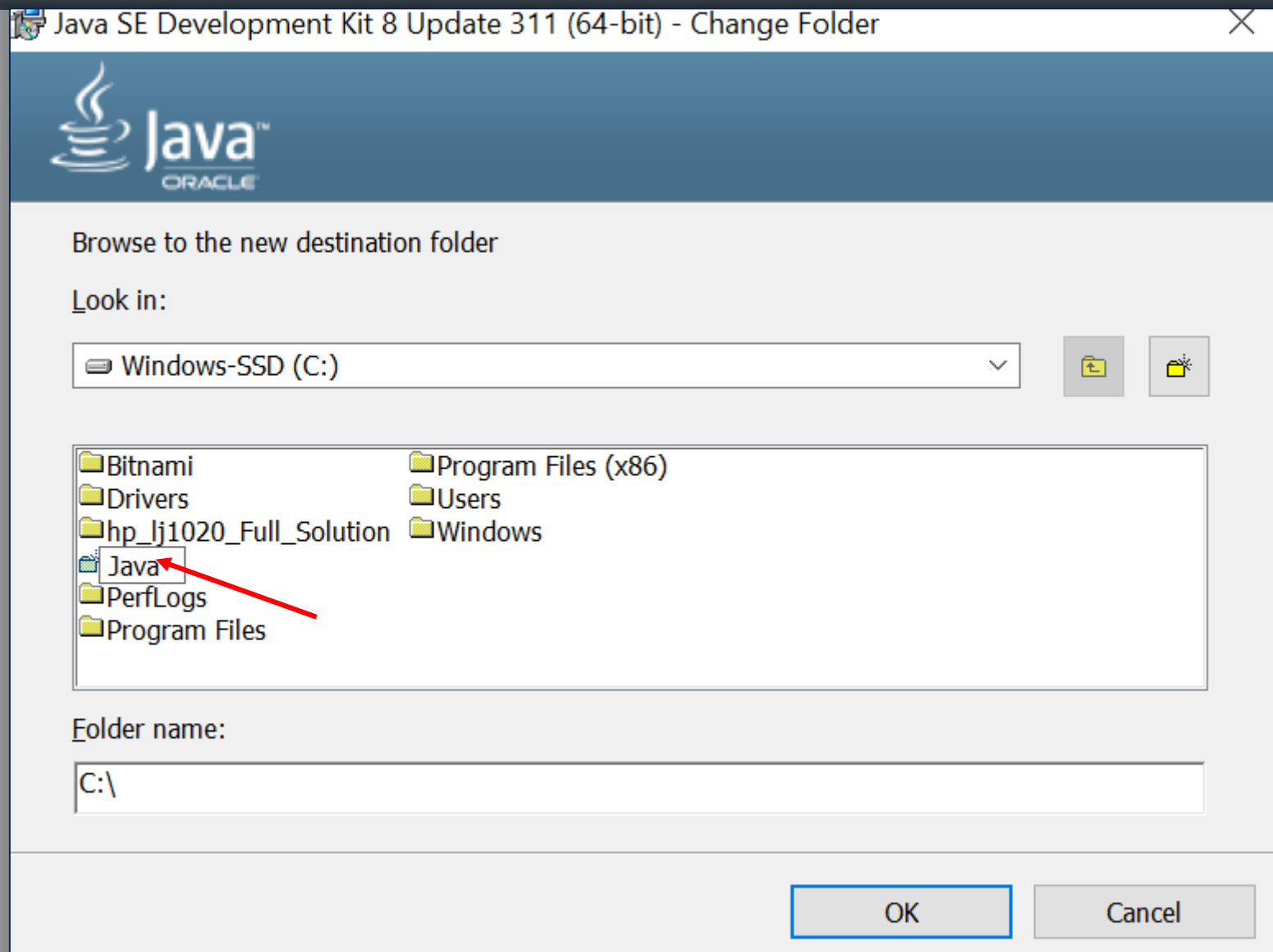
# Installation of JAVA



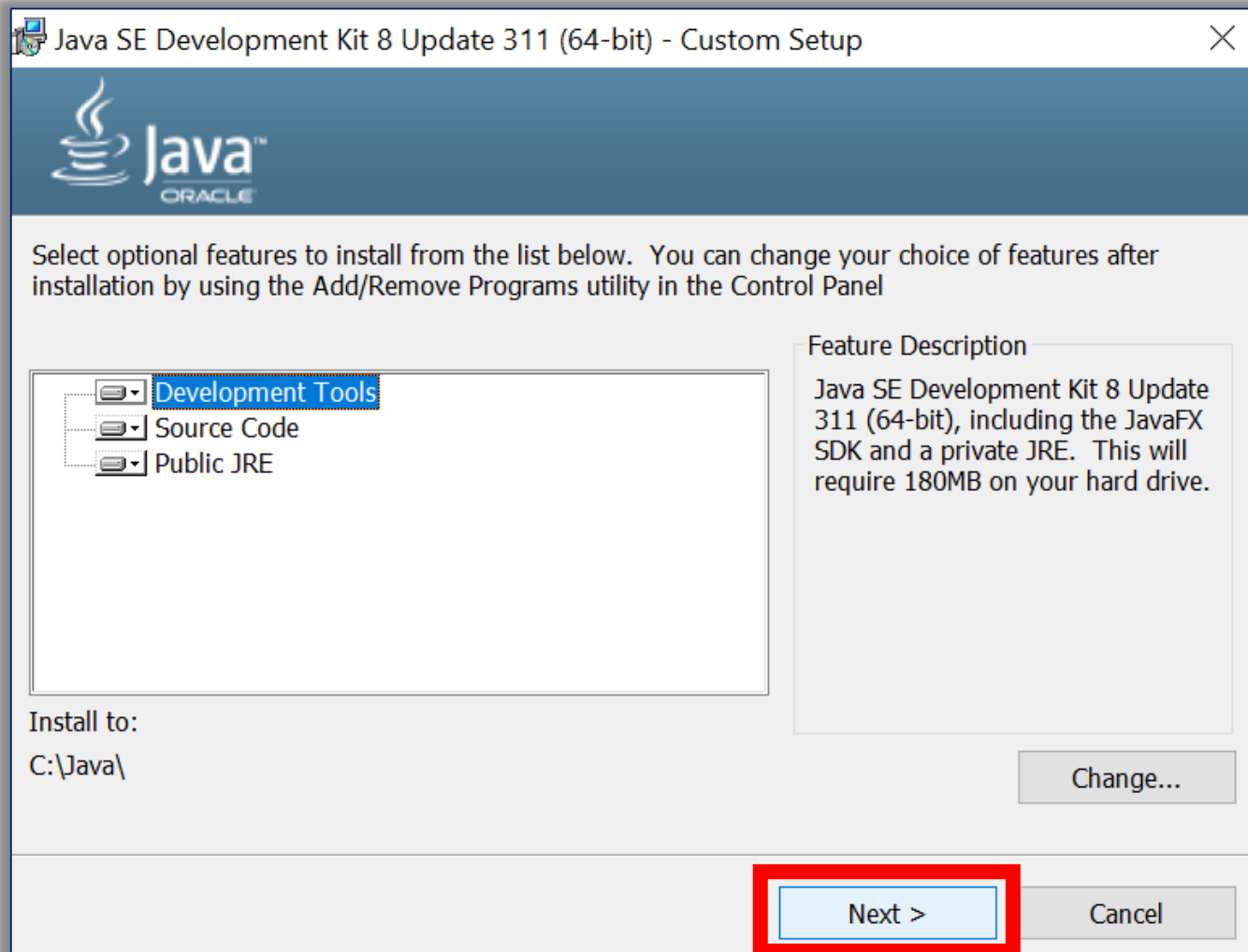
# Installation of JAVA



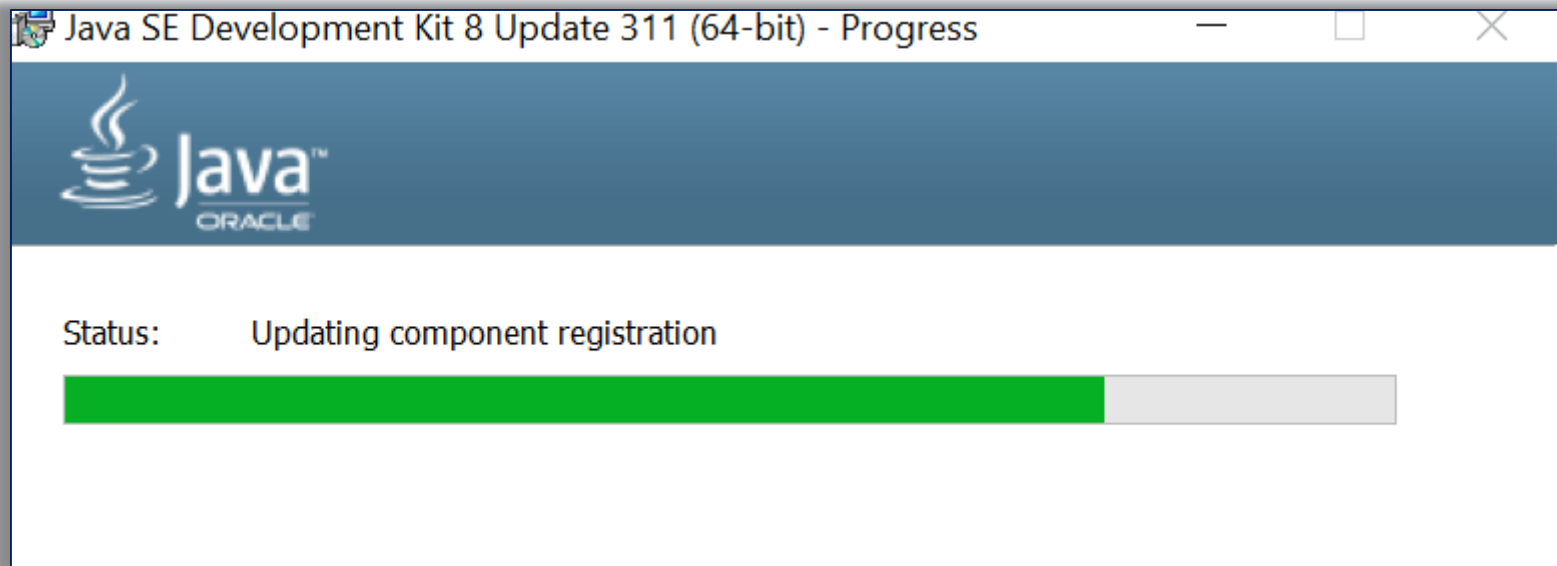
# Installation of JAVA



# Installation of JAVA



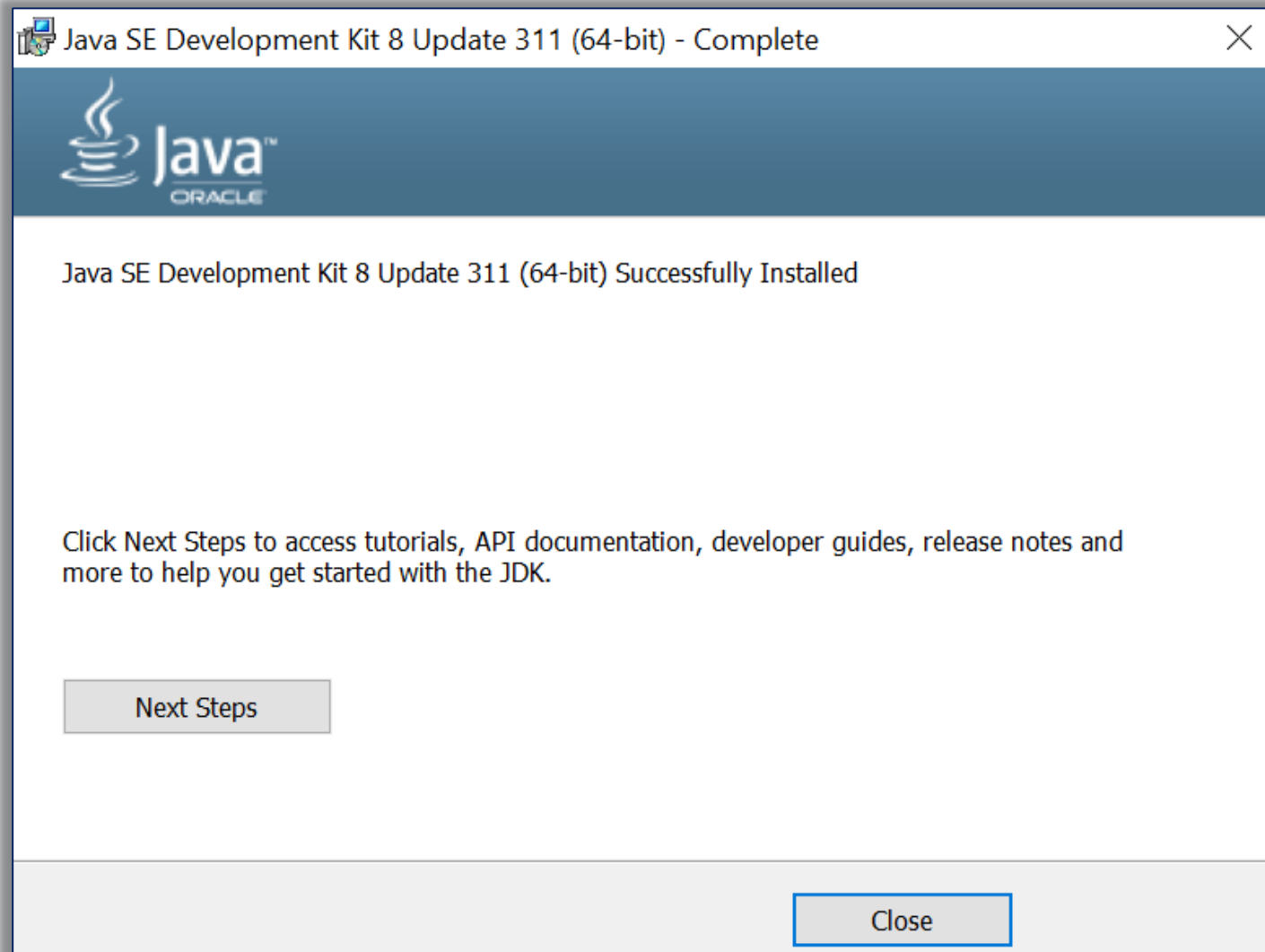
# Installation of JAVA



# Installation of JAVA

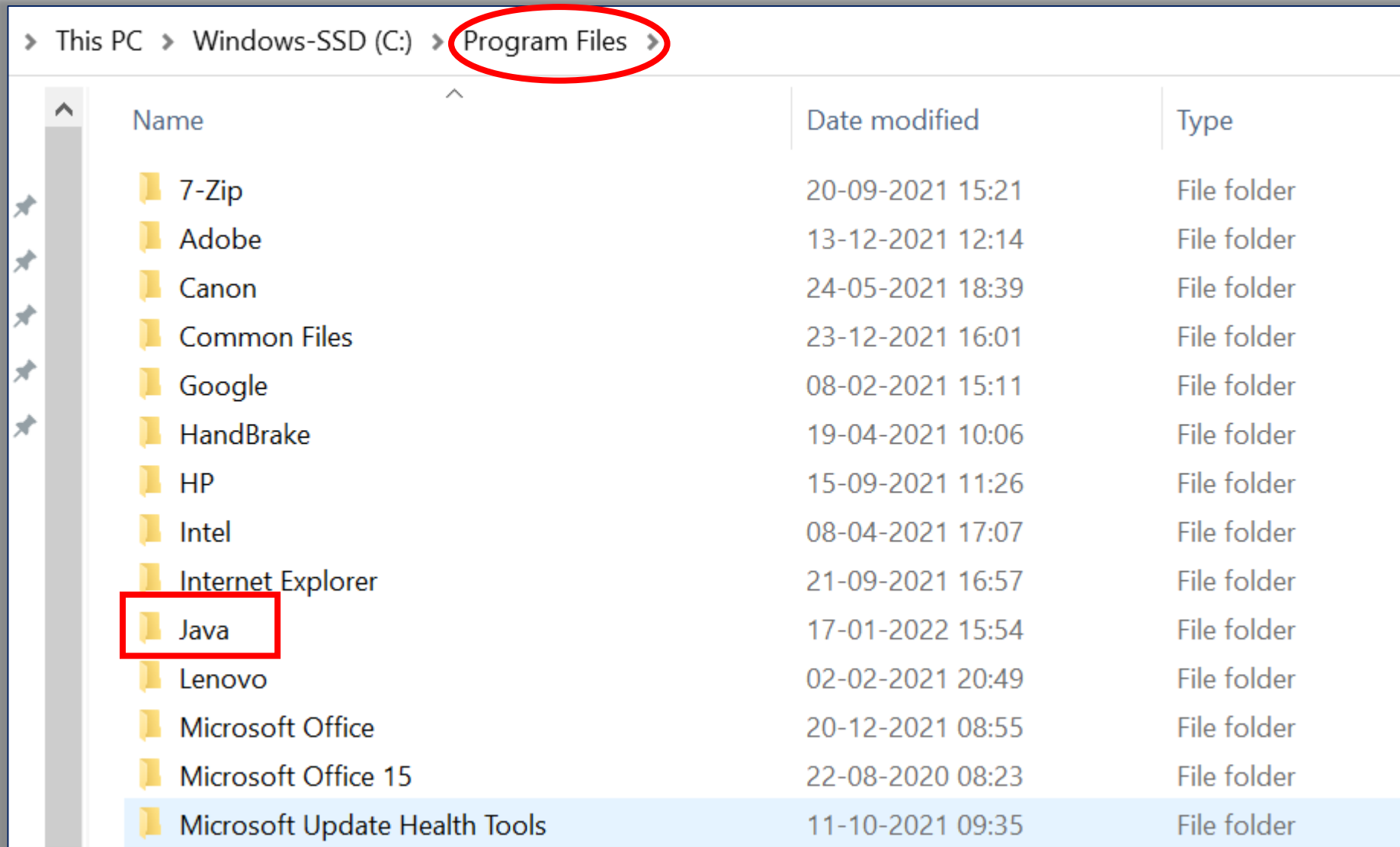


# Installation of JAVA




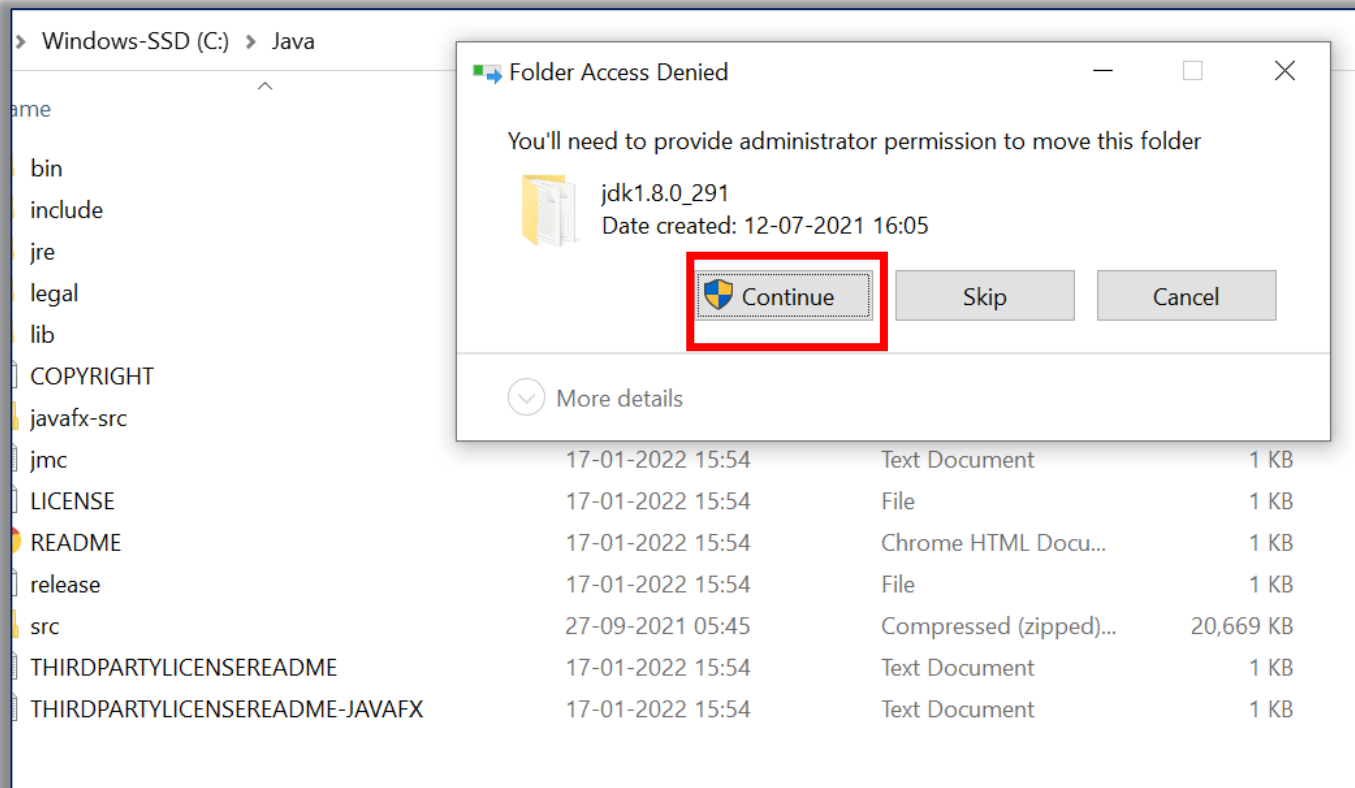


# Installation of JAVA

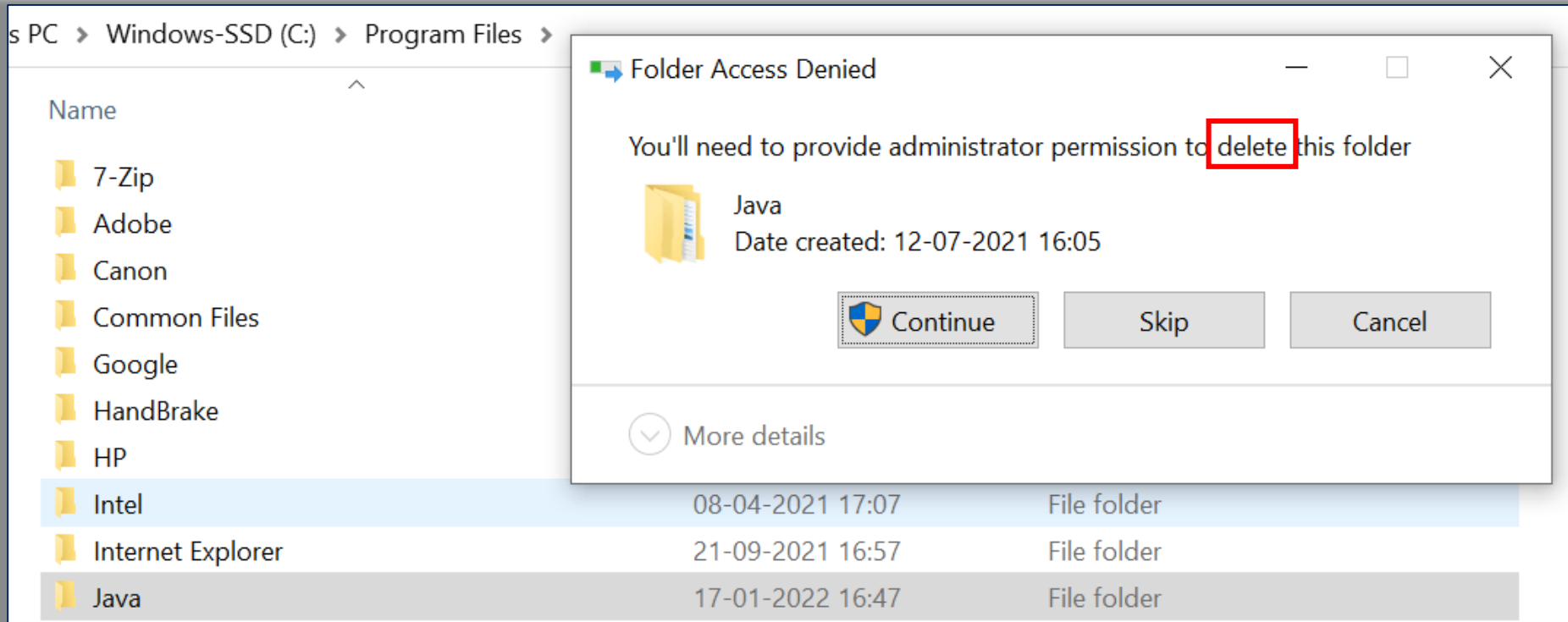


# Installation of JAVA

s PC > Windows-SSD (C:) > Program Files > Java		
Name	Date modified	Type
 jdk1.8.0_291	12-07-2021 16:06	File folder



# Installation of JAVA



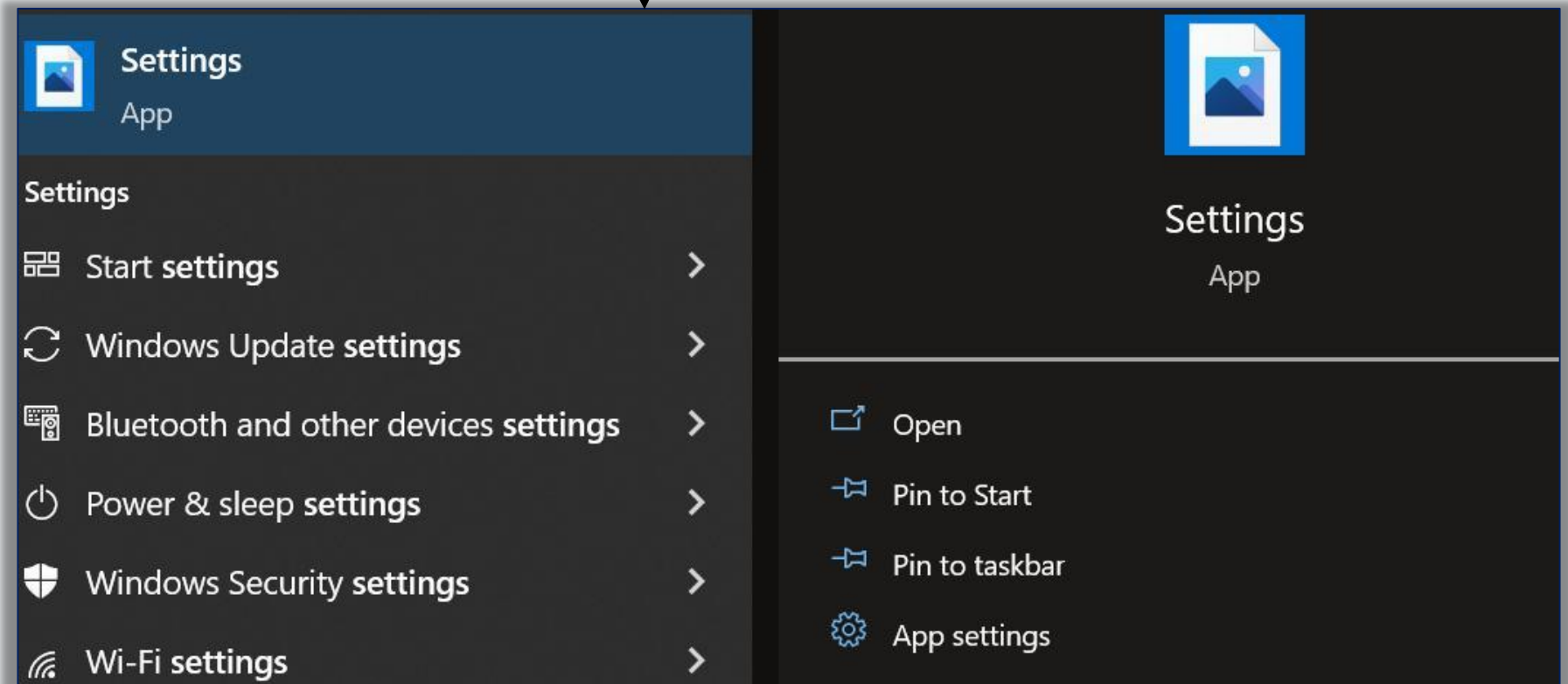
# How to set environment variable for JAVA

---

# Set the Environment Variable for JAVA



# Set the Environment Variable for JAVA



# Set the Environment Variable for JAVA



## System

Display, sound, notifications,  
power



## Devices

Bluetooth, printers, mouse



## Phone

Link your Android, iPhone

# Set the Environment Variable for JAVA



## System

Display, sound, notifications,  
power



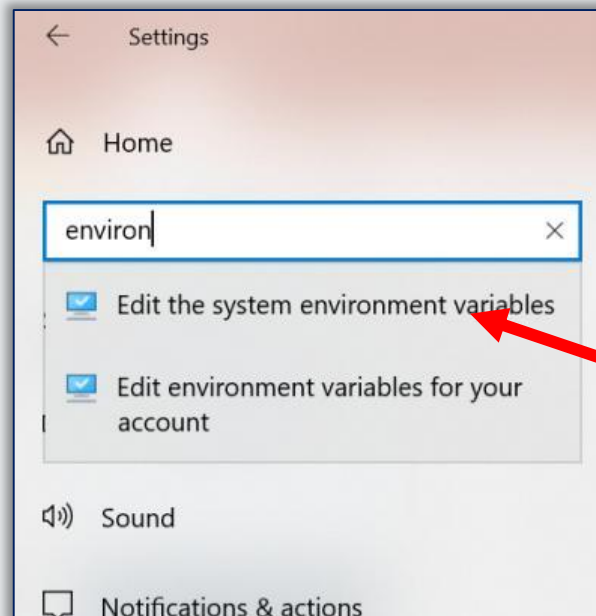
## Devices

Bluetooth, printers, mouse



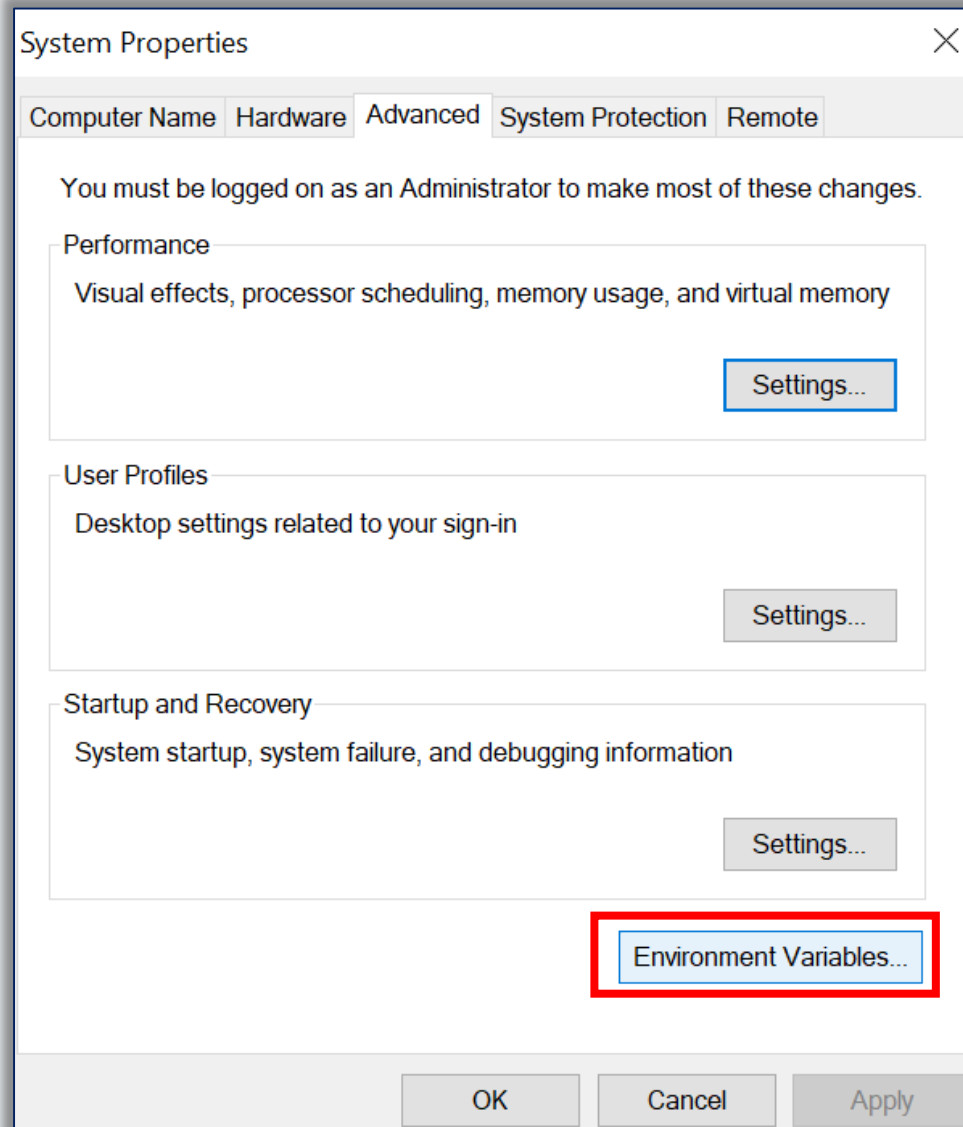
## Phone

Link your Android, iPhone

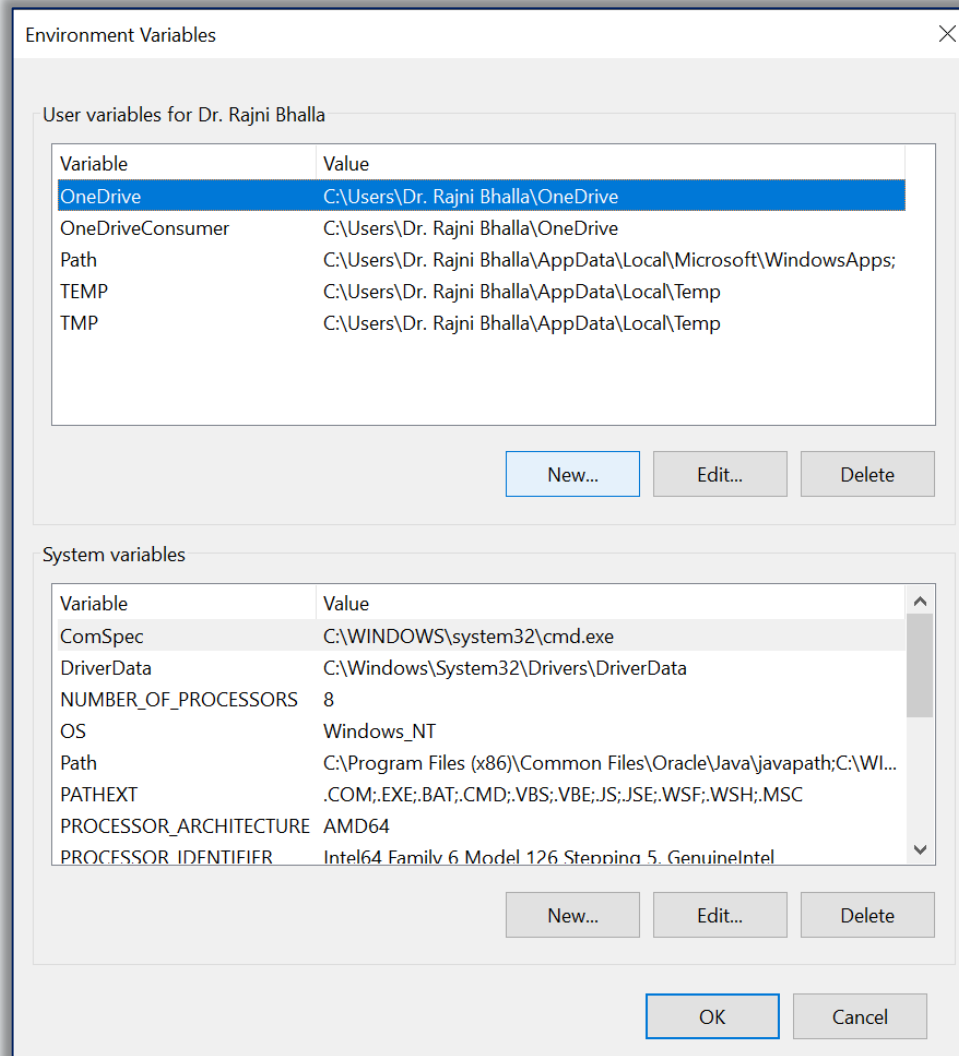




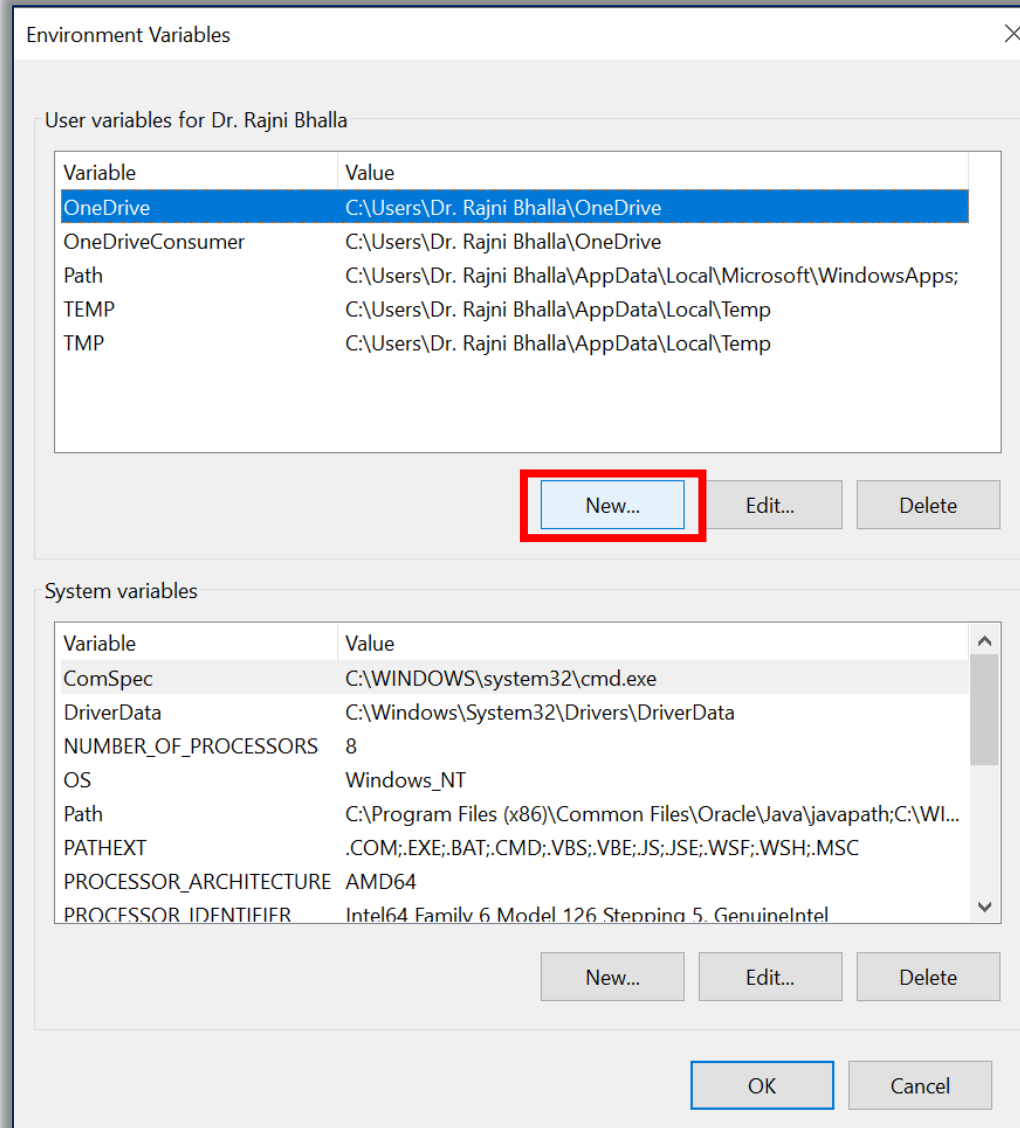
# Set the Environment Variable for JAVA



# Set the Environment Variable for JAVA



# Set the Environment Variable for JAVA



# Set the Environment Variable for JAVA

New User Variable

Variable name: JAVA\_HOME

Variable value:

Browse Directory... Browse File... OK Cancel

# Set the Environment Variable for JAVA

New User Variable ✕

Variable name:

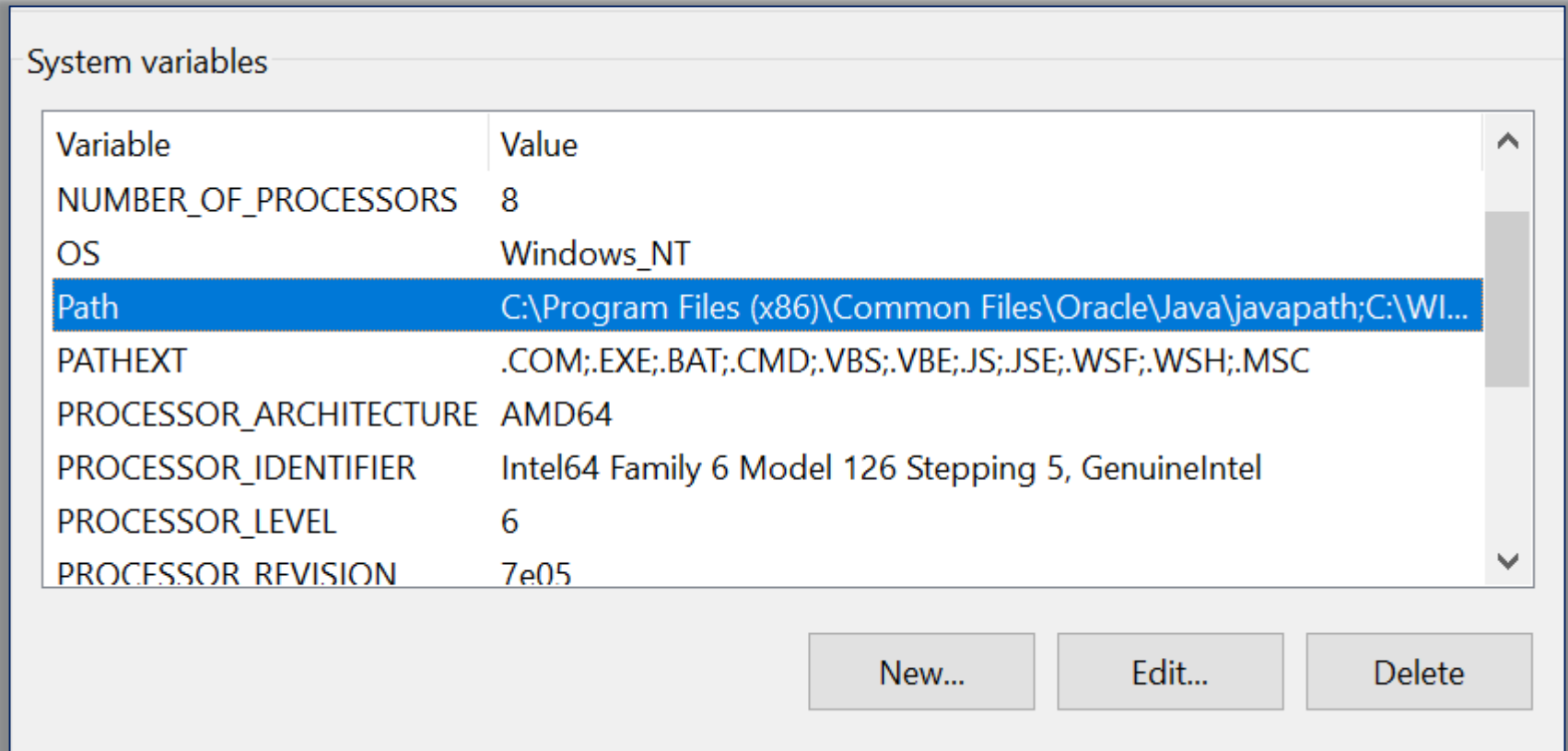
Variable value:

Edit User Variable ✕

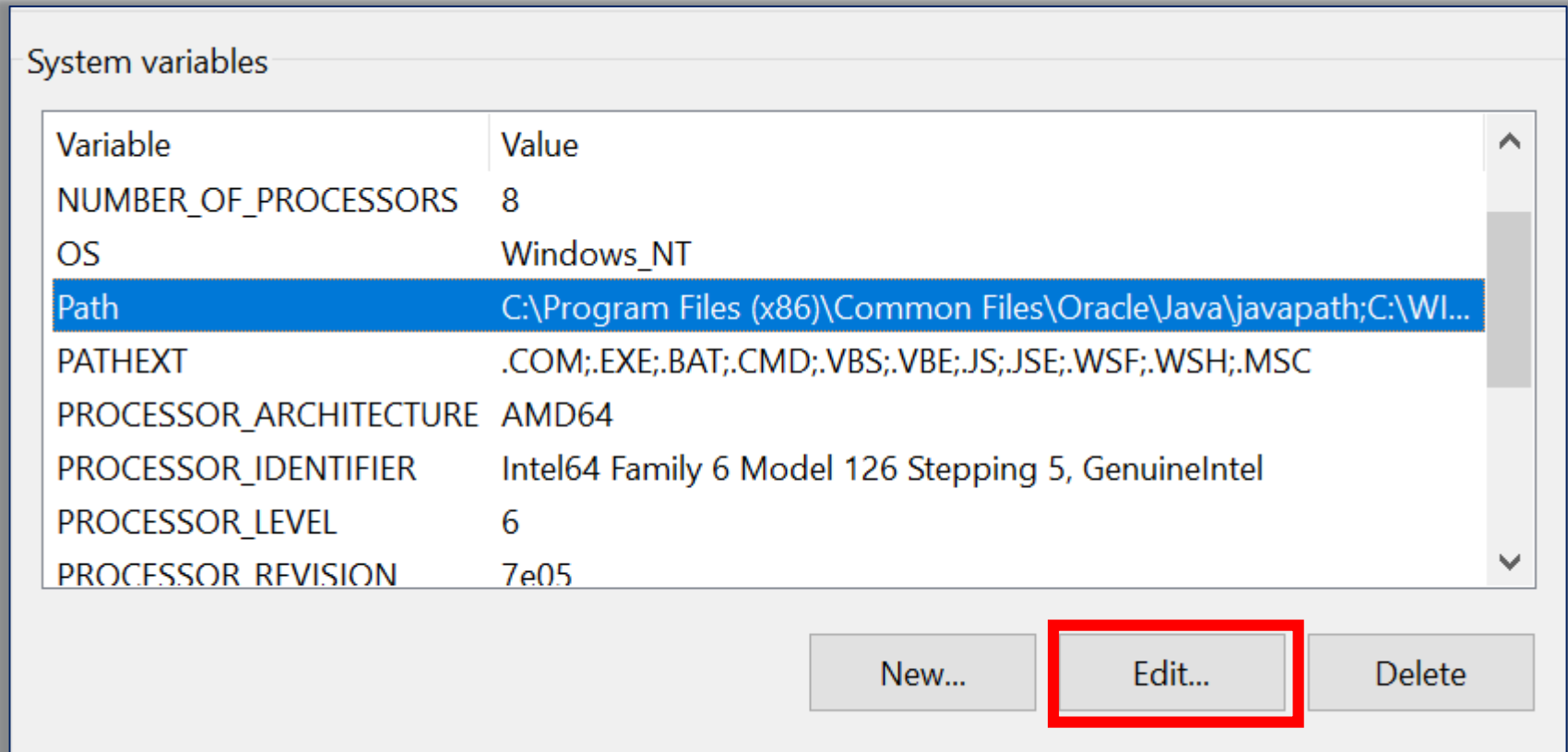
Variable name:

Variable value:

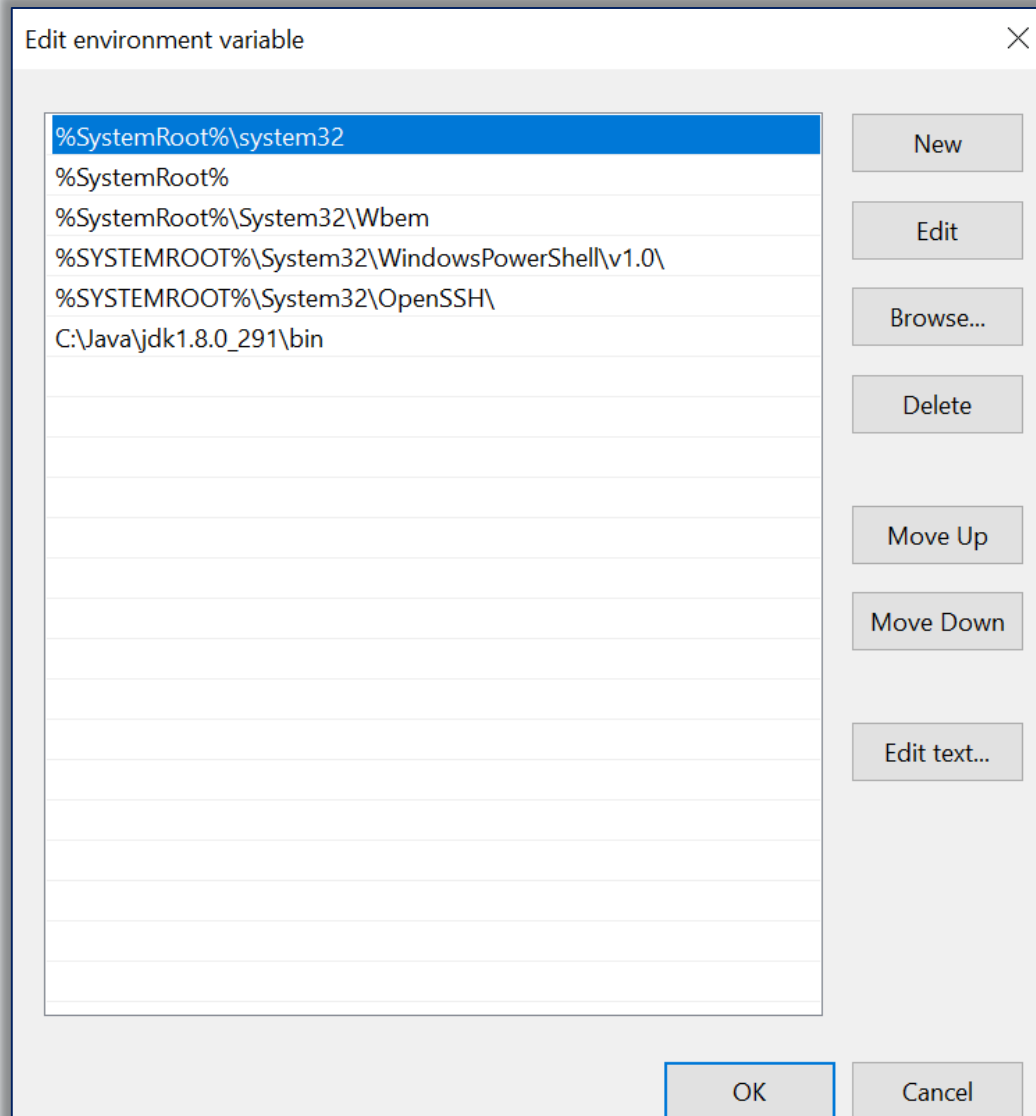
# Set the Environment Variable for JAVA



# Set the Environment Variable for JAVA

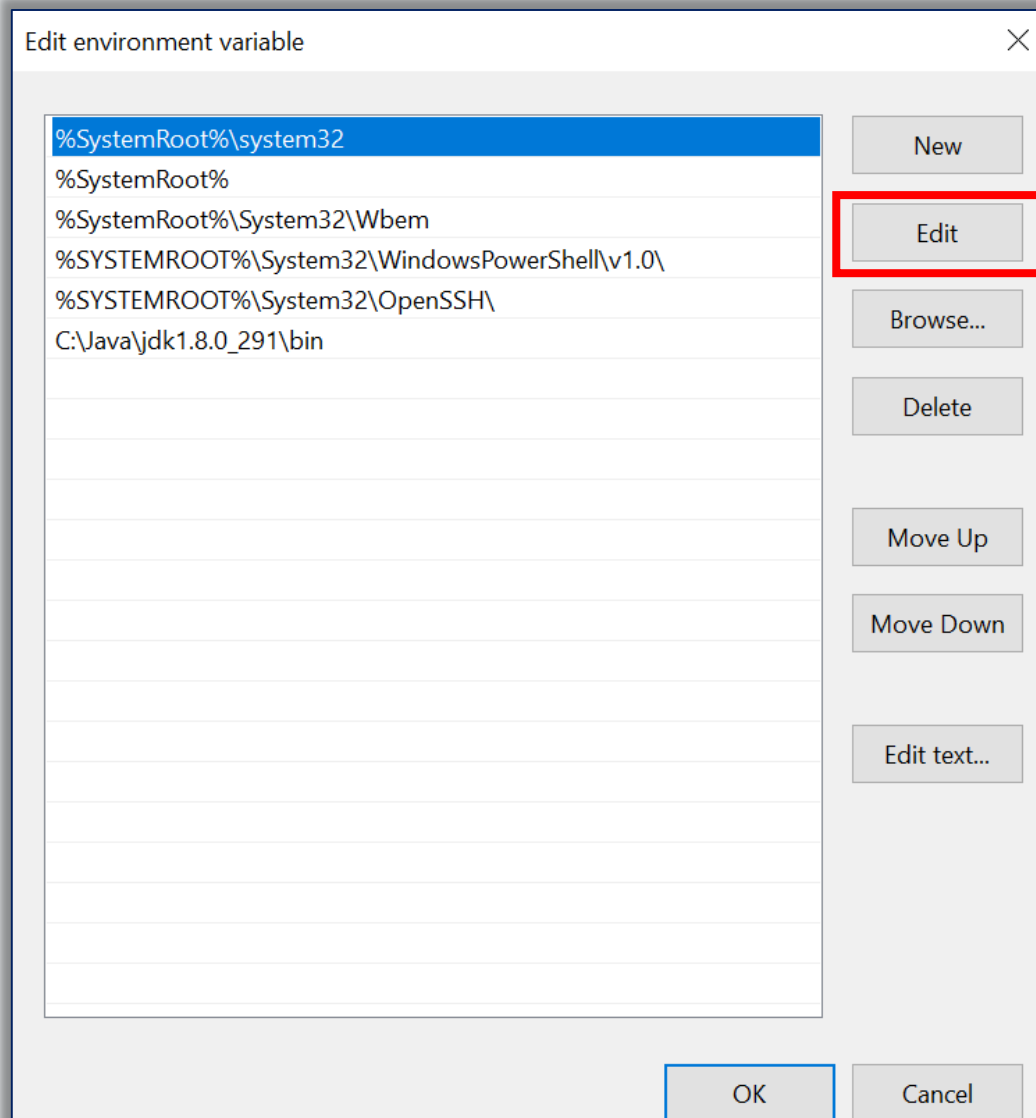


# Set the Environment Variable for JAVA

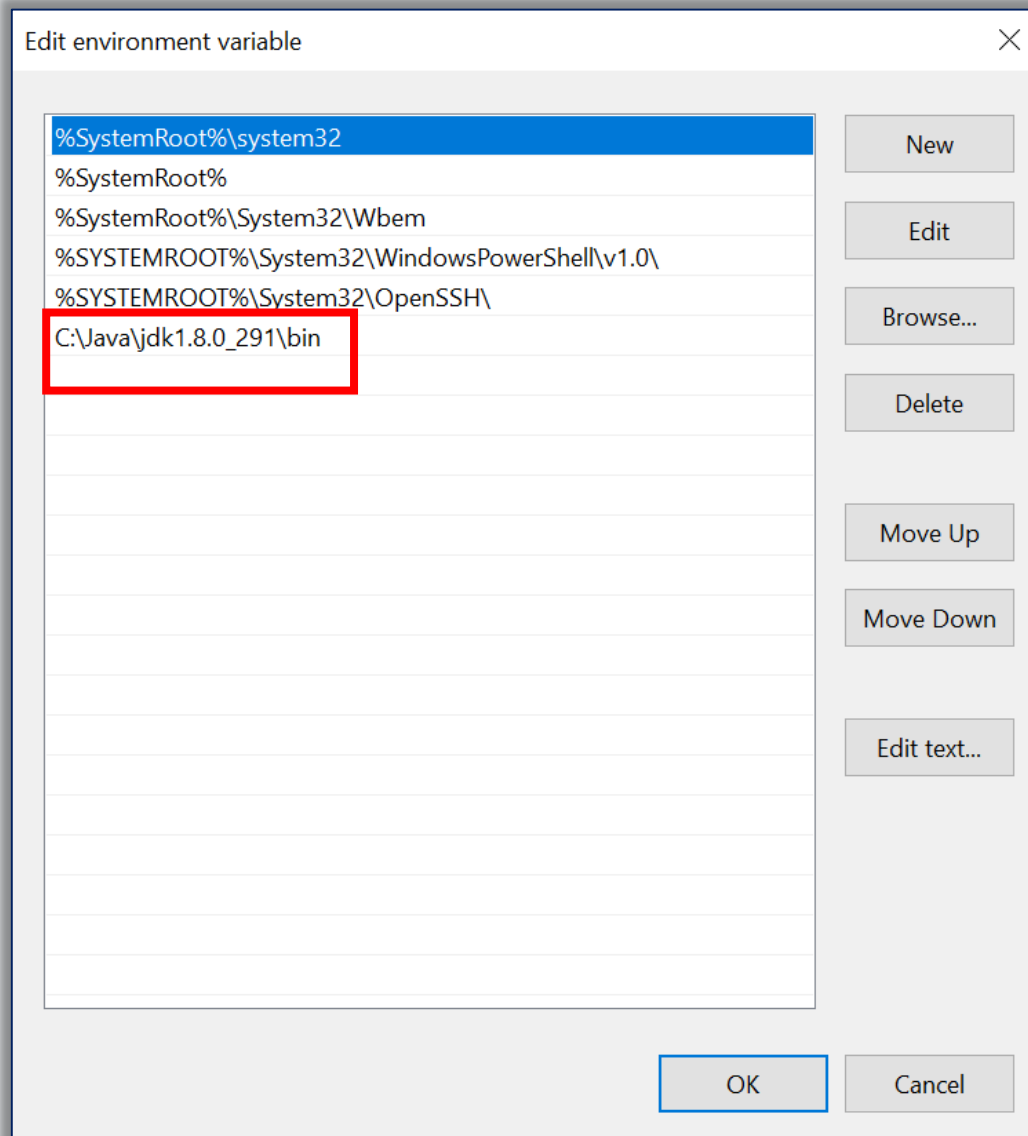




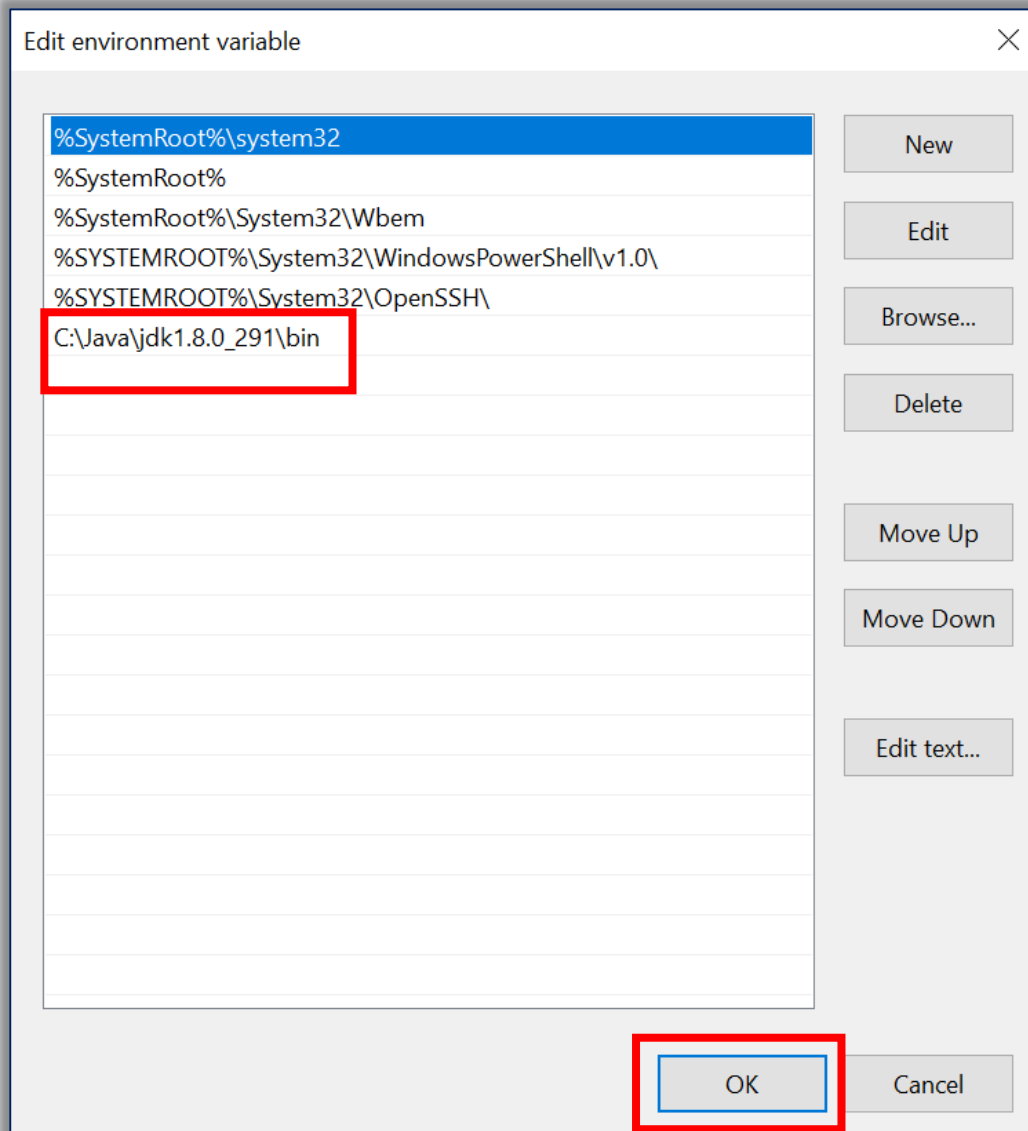
# Set the Environment Variable for JAVA



# Set the Environment Variable for JAVA



# Set the Environment Variable for JAVA



**Check whether java  
is functional or not**

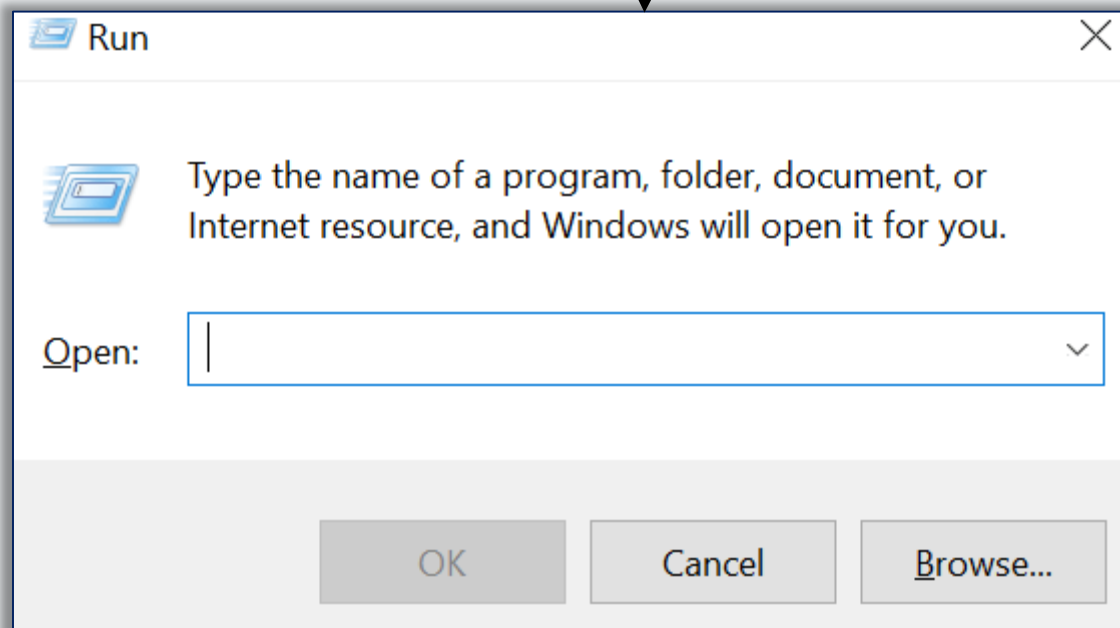
---

# Steps to check Java is functional or not

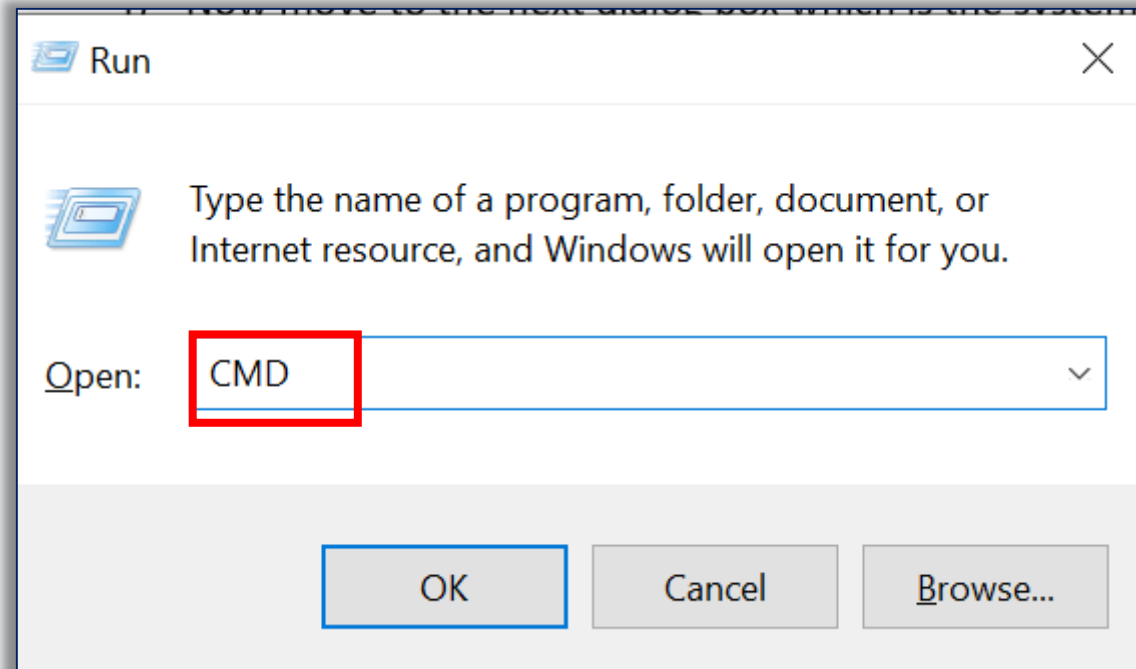
**Windows + R**

# Steps to check Java is functional or not

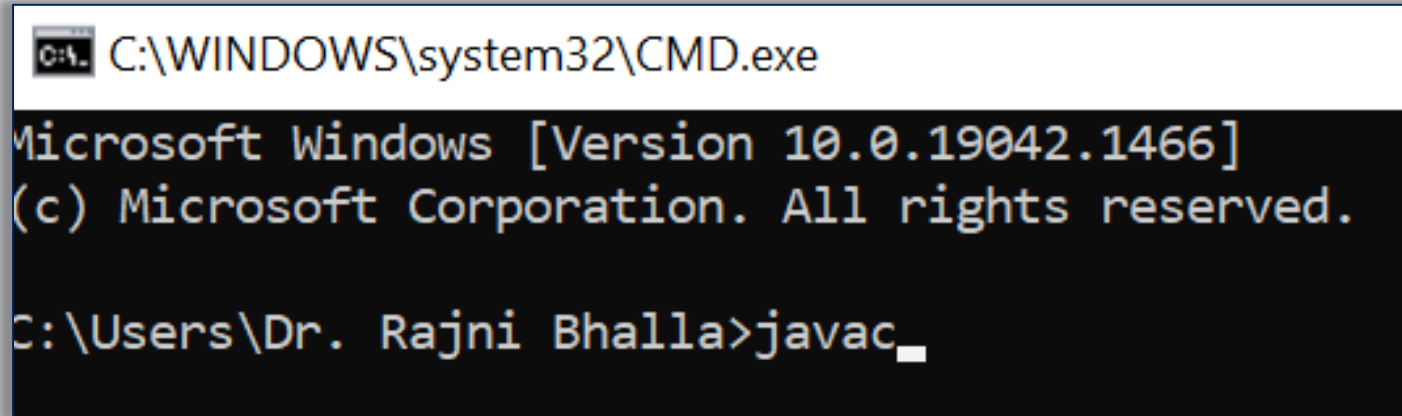
## Windows + R



# Steps to check Java is functional or not



# Steps to check Java is functional or not



```
C:\WINDOWS\system32\CMD.exe
Microsoft Windows [Version 10.0.19042.1466]
(c) Microsoft Corporation. All rights reserved.

C:\Users\Dr. Rajni Bhalla>javac_
```



# Steps to check Java is functional or not

```
C:\Users\Dr. Rajni Bhalla>javac
```

```
Usage: javac <options> <source files>
```

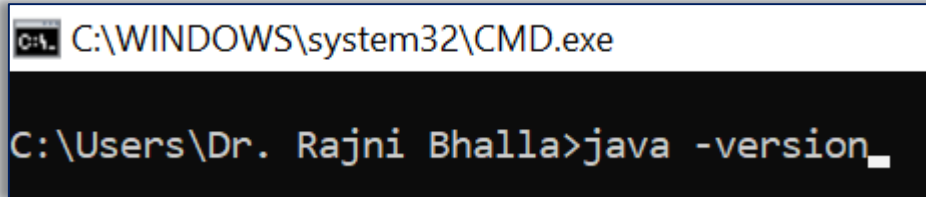
```
where possible options include:
```

-g	Generate all debugging info
-g:none	Generate no debugging info
-g:{lines,vars,source}	Generate only some debugging info
-nowarn	Generate no warnings
-verbose	Output messages about what the compiler is doing
-deprecation	Output source locations where deprecated APIs are used
-classpath <path>	Specify where to find user class files and annotation processors
-cp <path>	Specify where to find user class files and annotation processors
-sourcepath <path>	Specify where to find input source files
-bootclasspath <path>	Override location of bootstrap class files
-extdirs <dirs>	Override location of installed extensions
-endorseddirs <dirs>	Override location of endorsed standards path
-proc:{none,only}	Control whether annotation processing and/or compilation is done.
-processor <class1>[,<class2>,<class3>...]	Names of the annotation processors to run; bypasses default discovery process
-processorpath <path>	Specify where to find annotation processors
-parameters	Generate metadata for reflection on method parameters
-d <directory>	Specify where to place generated class files
-s <directory>	Specify where to place generated source files
-h <directory>	Specify where to place generated native header files
-implicit:{none,class}	Specify whether or not to generate class files for implicitly referenced files
-encoding <encoding>	Specify character encoding used by source files
-source <release>	Provide source compatibility with specified release
-target <release>	Generate class files for specific VM version
-profile <profile>	Check that API used is available in the specified profile
-version	Version information
-help	Print a synopsis of standard options
-Akey[=value]	Options to pass to annotation processors
-X	Print a synopsis of nonstandard options
-J<flag>	Pass <flag> directly to the runtime system
-Werror	Terminate compilation if warnings occur
@<filename>	Read options and filenames from file

# To check version of JAVA

---


# CMD

A screenshot of a Windows Command Prompt window. The title bar at the top shows the icon and the path "C:\WINDOWS\system32\CMD.exe". The main area of the window has a black background with white text. It shows the command prompt "C:\Users\Dr. Rajni Bhalla>" followed by the command "java -version\_" and a cursor at the end of the line.

```
C:\WINDOWS\system32\CMD.exe  
C:\Users\Dr. Rajni Bhalla>java -version_
```

# CMD

```
C:\WINDOWS\system32\CMD.exe  
C:\Users\Dr. Rajni Bhalla>java -version_
```















```
C:\Users\Dr. Rajni Bhalla>java -version  
java version "1.8.0_291"  
Java(TM) SE Runtime Environment (build 1.8.0_291-b10)  
Java HotSpot(TM) 64-Bit Server VM (build 25.291-b10, mixed mode)
```

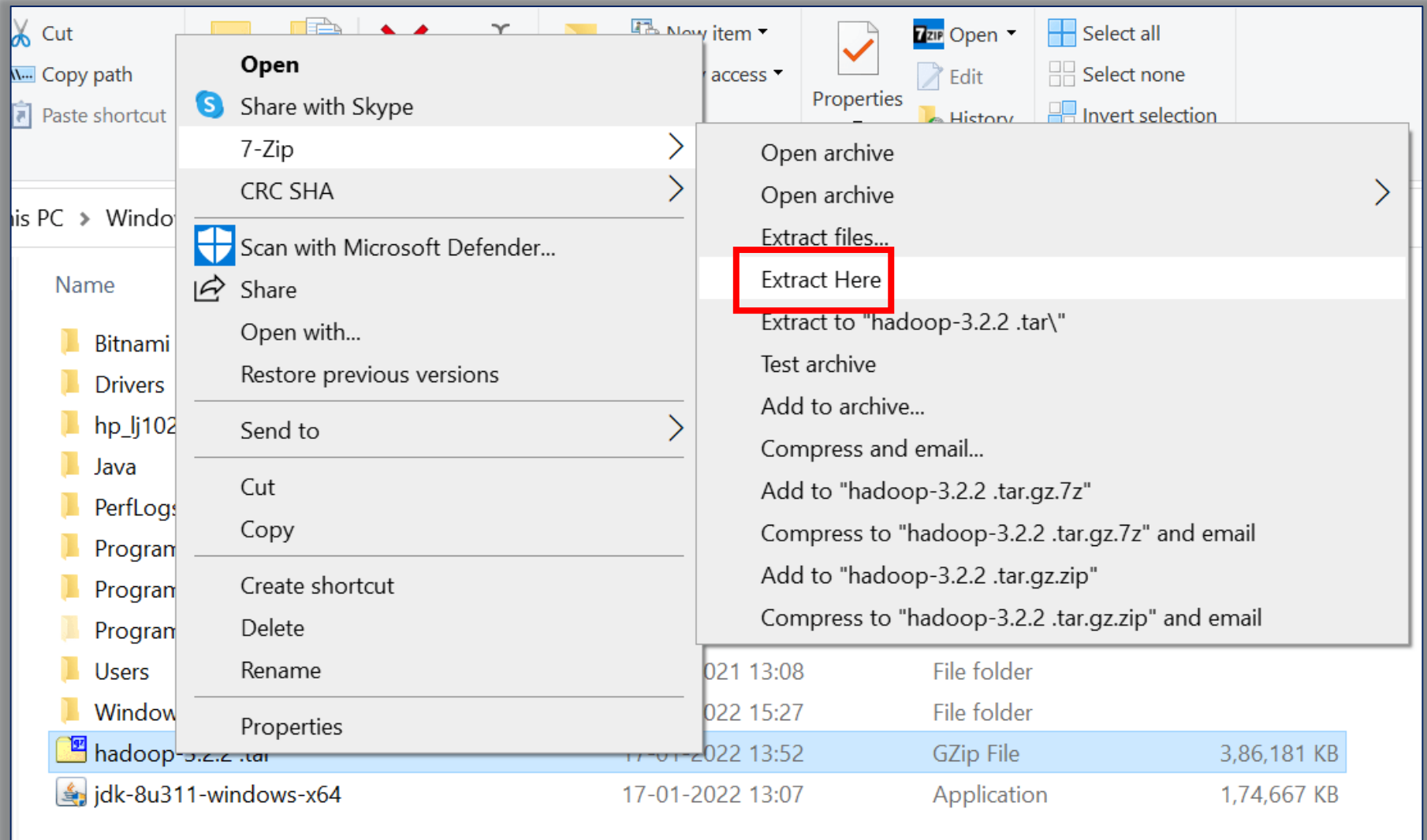
# INSTALL HADOOP INTO OUR LOCAL SYSTEM

---

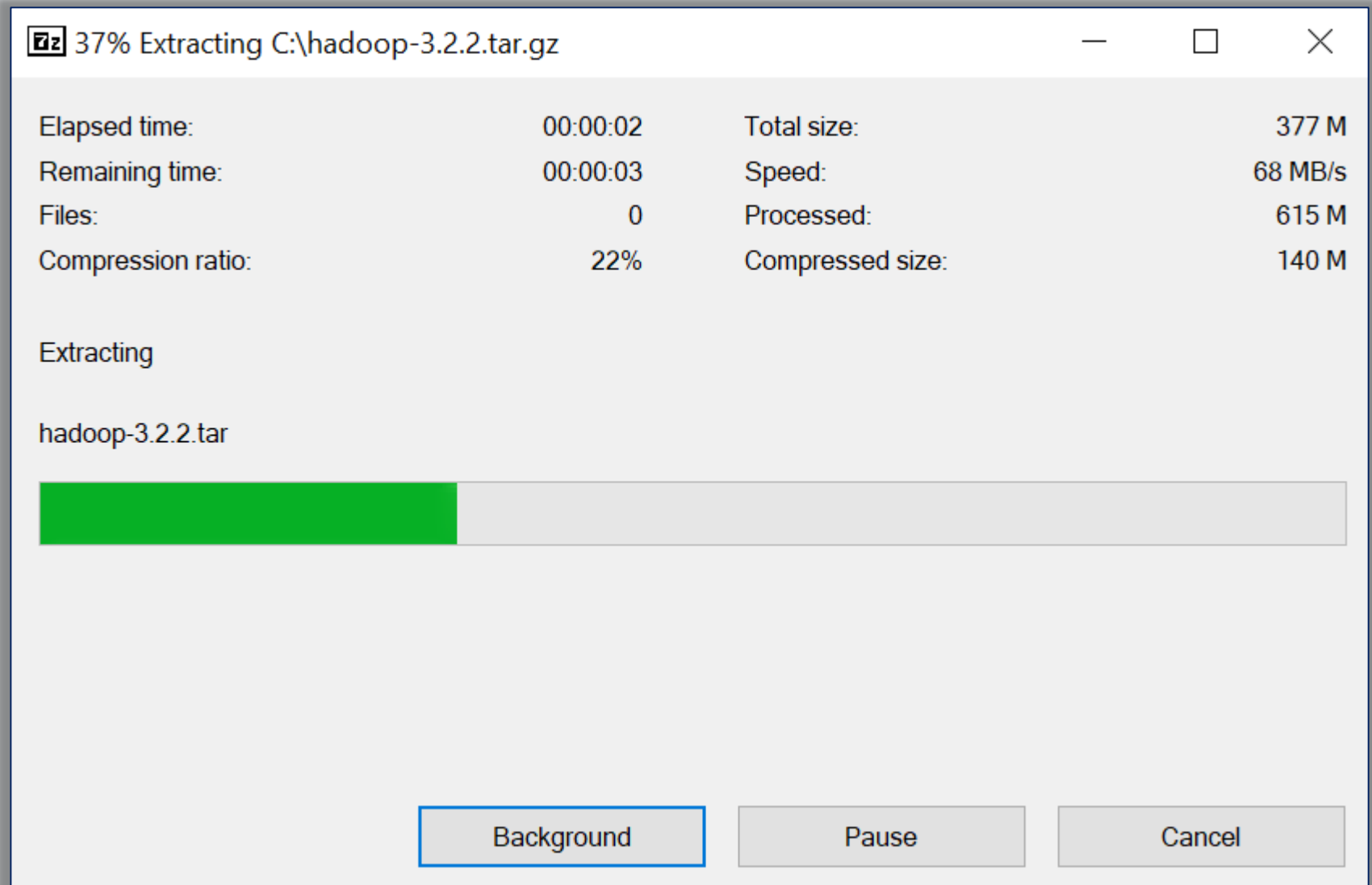
# Steps to install Hadoop

 Bitnami	25-06-2021 12:16	File folder	
 Drivers	02-02-2021 20:47	File folder	
 hp_lj1020_Full_Solution	15-09-2021 11:26	File folder	
 Java	17-01-2022 16:48	File folder	
 PerfLogs	07-12-2019 14:44	File folder	
 Program Files	17-01-2022 17:00	File folder	
 Program Files (x86)	11-01-2022 09:09	File folder	
 ProgramData	25-11-2021 12:48	File folder	
 Users	08-04-2021 13:08	File folder	
 Windows	14-01-2022 15:27	File folder	
 hadoop-3.2.2 .tar	17-01-2022 13:52	GZip File	3,86,181 KB
 jdk-8u311-windows-x64	17-01-2022 13:07	Application	1,74,667 KB

# Steps to install Hadoop



# Steps to install Hadoop








# Steps to install Hadoop

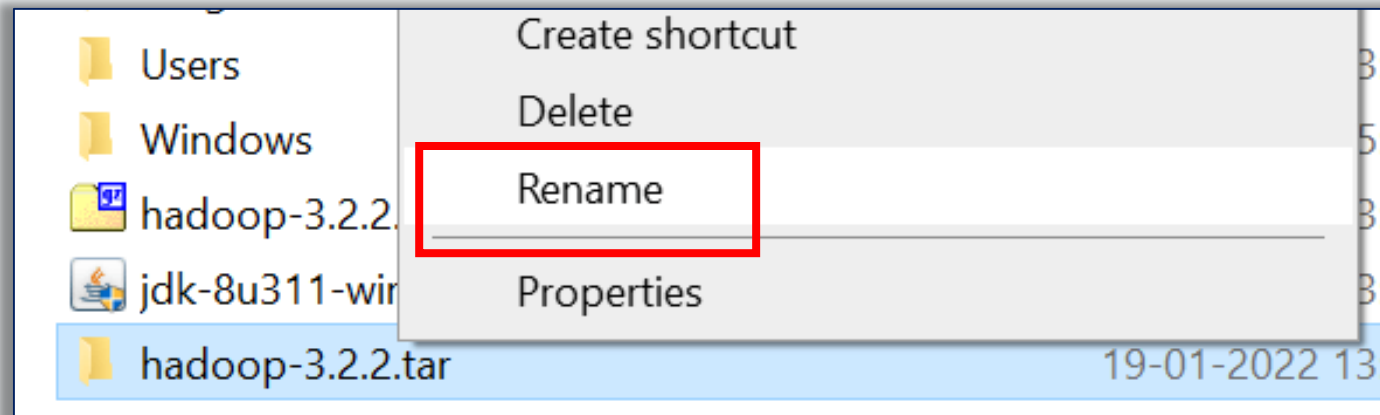
Ignore

1. Cannot create a symbolic link:....
2. Cannot create a symbolic link:....
3. Cannot create a symbolic link:....

# Steps to install Hadoop

 hadoop-3.2.2.tar	17-01-2022 13:52	GZip File	3,86,181 KB
 jdk-8u311-windows-x64	17-01-2022 13:07	Application	1,74,667 KB
 hadoop-3.2.2.tar	19-01-2022 13:47	File folder	

# Steps to install Hadoop



# Steps to install Hadoop

Bitnami	25-06-2021 12:16	File folder
Drivers	02-02-2021 20:47	File folder
hadoop	03-01-2021 15:41	File folder
hp_lj1020_Full_Solution	15-09-2021 11:26	File folder
java	17-01-2022 16:48	File folder

# Steps to install Hadoop

Bitnami	25-06-2021 12:16	File folder
Drivers	02-02-2021 20:47	File folder
hadoop	03-01-2021 15:41	File folder
hp_lj1020_Full_Solution	15-09-2021 11:26	File folder
java	17-01-2022 16:48	File folder

This PC > Windows-SSD (C:) > hadoop >			
Name	Date modified	Type	
bin	03-01-2021 15:41	File folder	
etc	03-01-2021 14:59	File folder	
include	03-01-2021 15:41	File folder	
lib	03-01-2021 15:41	File folder	

# Steps to install Hadoop

Bitnami	25-06-2021 12:16	File folder
Drivers	02-02-2021 20:47	File folder
hadoop	03-01-2021 15:41	File folder
hp_j1020_Full_Solution	15-09-2021 11:26	File folder
java	17-01-2022 16:48	File folder

» This PC » Windows-SSD (C:) » hadoop »

Name	Date modified	Type
bin	03-01-2021 15:41	File folder
etc	03-01-2021 14:59	File folder
include	03-01-2021 15:41	File folder
lib	03-01-2021 15:41	File folder

This PC » Windows-SSD (C:) » hadoop » etc »

Name
hadoop

# Steps to install Hadoop

This PC > Windows-SSD (C:) > hadoop > etc > hadoop >

Name	Date modified	Type
shellprofile.d	03-01-2021 14:58	File folder
capacity-scheduler	03-01-2021 15:24	XML Document
configuration	03-01-2021 15:27	XSL Stylesheet
container-executor.cfg	03-01-2021 15:24	CFG File
core-site	03-01-2021 14:58	XML Document
hadoop-env	03-01-2021 14:58	Windows Command
hadoop-env.sh	03-01-2021 15:41	SH File
hadoop-metrics2.properties	03-01-2021 14:58	PROPERTIES File
hadoop-policy	03-01-2021 14:58	XML Document
hadoop-user-functions.sh.example	03-01-2021 14:58	EXAMPLE File
hdfs-site	03-01-2021 15:02	XML Document
httpfs-env.sh	03-01-2021 15:03	SH File
httpfs-log4j.properties	03-01-2021 15:03	PROPERTIES File
httpfs-signature.secret	03-01-2021 15:03	SECRET File
httpfs-site	03-01-2021 15:03	XML Document
kms-acls	03-01-2021 14:59	XML Document
kms-env.sh	03-01-2021 14:59	SH File
kms-log4j.properties	03-01-2021 14:59	PROPERTIES File
kms-site	03-01-2021 14:59	XML Document
log4j.properties	03-01-2021 14:58	PROPERTIES File
mapred-env	03-01-2021 15:27	Windows Command

# Steps to install Hadoop

This PC > Windows-SSD (C:) > hadoop > etc > hadoop >

Name	Date modified	Type
shellprofile.d	03-01-2021 14:58	File folder
capacity-scheduler	03-01-2021 15:24	XML Document
configuration	03-01-2021 15:27	XSL Stylesheet
container-executor.cfg	03-01-2021 15:24	CFG File
core-site	03-01-2021 14:58	XML Document
hadoop-env	03-01-2021 14:58	Windows Command
hadoop-env.sh	03-01-2021 15:41	SH File
hadoop-metrics2.properties	03-01-2021 14:58	PROPERTIES File
hadoop-policy	03-01-2021 14:58	XML Document
hadoop-user-functions.sh.example	03-01-2021 14:58	EXAMPLE File
hdfs-site	03-01-2021 15:02	XML Document
httpfs-env.sh	03-01-2021 15:03	SH File
httpfs-log4j.properties	03-01-2021 15:03	PROPERTIES File
httpfs-signature.secret	03-01-2021 15:03	SECRET File
httpfs-site	03-01-2021 15:03	XML Document
kms-acls	03-01-2021 14:59	XML Document
kms-env.sh	03-01-2021 14:59	SH File
kms-log4j.properties	03-01-2021 14:59	PROPERTIES File
kms-site	03-01-2021 14:59	XML Document
log4j.properties	03-01-2021 14:58	PROPERTIES File
mapred-env	03-01-2021 15:27	Windows Command
























# Steps to install Hadoop






This PC > Windows-SSD (C:) > hadoop > etc > hadoop >

Name	Date modified	Type
shellprofile.d	03-01-2021 14:58	File folder
capacity-scheduler	03-01-2021 15:24	XML Document
configuration	03-01-2021 15:27	XSL Stylesheet
container-executor.cfg	03-01-2021 15:24	CFG File
core-site	03-01-2021 14:58	XML Document
hadoop-env	03-01-2021 14:58	Windows Command
hadoop-env.sh	03-01-2021 15:41	SH File
hadoop-metrics2.properties	03-01-2021 14:58	PROPERTIES File
hadoop-policy	03-01-2021 14:58	XML Document
hadoop-user-functions.sh.example	03-01-2021 14:58	EXAMPLE File
hdfs-site	03-01-2021 15:02	XML Document
httpfs-env.sh	03-01-2021 15:03	SH File
httpfs-log4j.properties	03-01-2021 15:03	PROPERTIES File
httpfs-signature.secret	03-01-2021 15:03	SECRET File
httpfs-site	03-01-2021 15:03	XML Document
kms-acls	03-01-2021 14:59	XML Document
kms-env.sh	03-01-2021 14:59	SH File
kms-log4j.properties	03-01-2021 14:59	PROPERTIES File
kms-site	03-01-2021 14:59	XML Document
log4j.properties	03-01-2021 14:58	PROPERTIES File
mapred-env	03-01-2021 15:27	Windows Command






















# Steps to install Hadoop

 core-site	03-01-2021 14:58	XML Document	1 KB
 hadoop-env	03-01-2021 14:58	Windows Command ...	4 KB
 hadoop-env.sh	03-01-2021 15:41	SH File	16 KB
 hadoop-metrics2.properties	03-01-2021 14:58	PROPERTIES File	4 KB
 hadoop-policy	03-01-2021 14:58	XML Document	12 KB
 hadoop-user-functions.sh.example	03-01-2021 14:58	EXAMPLE File	4 KB
 hdfs-site	03-01-2021 15:02	XML Document	1 KB
 httpfs-env.sh	03-01-2021 15:03	SH File	2 KB
 httpfs-log4j.properties	03-01-2021 15:03	PROPERTIES File	2 KB
 httpfs-signature.secret	03-01-2021 15:03	SECRET File	1 KB
 httpfs-site	03-01-2021 15:03	XML Document	1 KB
 kms-acls	03-01-2021 14:59	XML Document	4 KB
 kms-env.sh	03-01-2021 14:59	SH File	2 KB
 kms-log4j.properties	03-01-2021 14:59	PROPERTIES File	2 KB
 kms-site	03-01-2021 14:59	XML Document	1 KB
 log4j.properties	03-01-2021 14:58	PROPERTIES File	15 KB
 mapred-env	03-01-2021 15:27	Windows Command ...	1 KB
 mapred-env.sh	03-01-2021 15:27	SH File	2 KB
 mapred-queues.xml.template	03-01-2021 15:27	TEMPLATE File	5 KB
 mapred-site	03-01-2021 15:27	XML Document	1 KB
 ssl-client.xml.example	03-01-2021 14:58	EXAMPLE File	3 KB

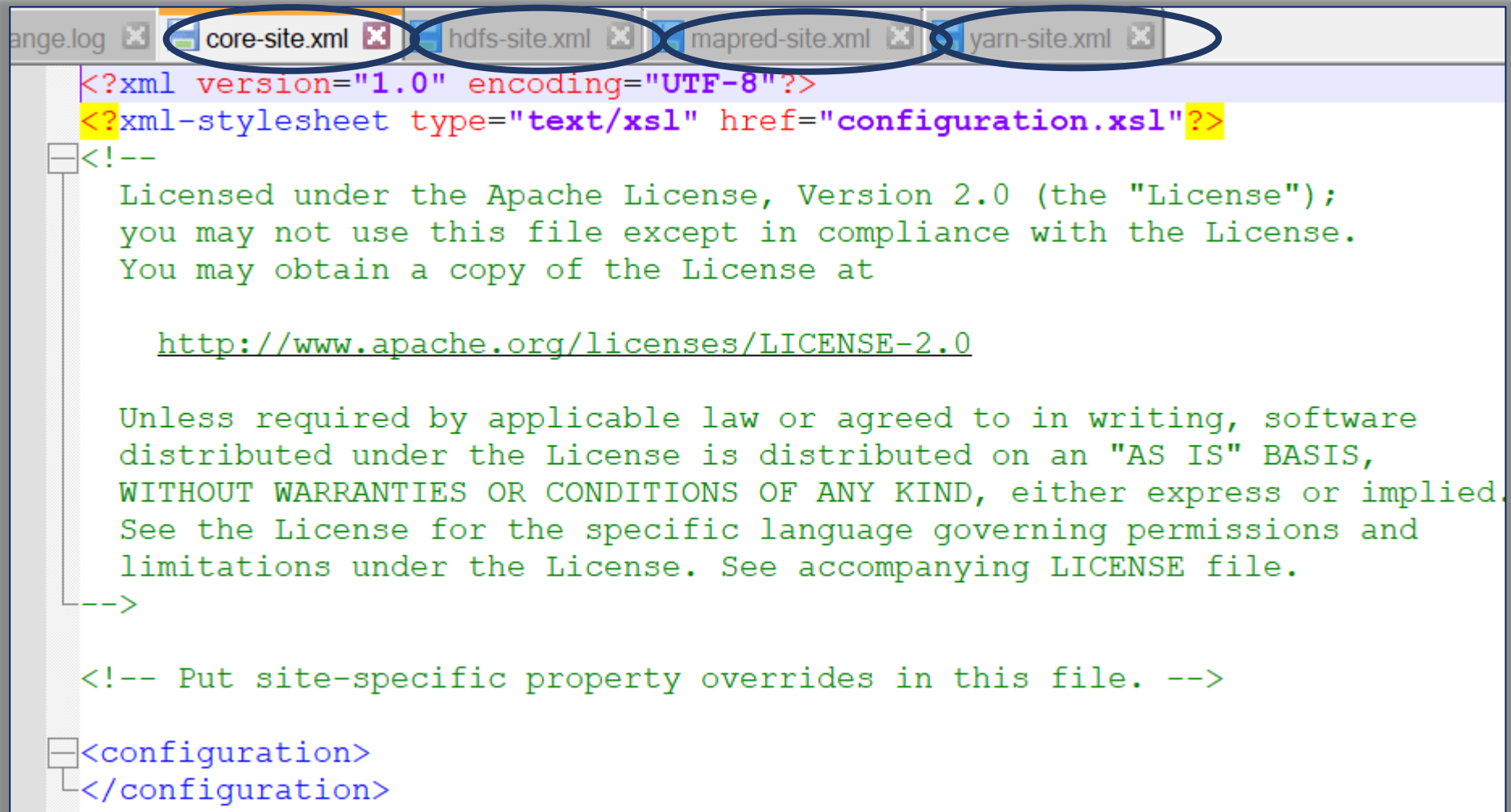
# Steps to install Hadoop

	workers	03-01-2021 14:58	File	1 KB
	yarn-env	03-01-2021 15:24	Windows Command ...	3 KB
	yarn-env.sh	03-01-2021 15:24	SH File	7 KB
	yarnservice-log4j.properties	03-01-2021 15:24	PROPERTIES File	3 KB
	yarn-site	03-01-2021 15:24	XML Document	1 KB

# Steps to install Hadoop

 core-site	03-01-2021 14:58	XML Document	1 KB
 hadoop-env	03-01-2021 14:58	Windows Command ...	4 KB
 hadoop-env.sh	03-01-2021 15:41	SH File	16 KB
 hadoop-metrics2.properties	03-01-2021 14:58	PROPERTIES File	4 KB
 hadoop-policy	03-01-2021 14:58	XML Document	12 KB
 hadoop-user-functions.sh.example	03-01-2021 14:58	EXAMPLE File	4 KB
 hdfs-site	03-01-2021 15:02	XML Document	1 KB
 httpfs-env.sh	03-01-2021 15:03	SH File	2 KB
 httpfs-log4j.properties	03-01-2021 15:03	PROPERTIES File	2 KB
 httpfs-signature.secret	03-01-2021 15:03	SECRET File	1 KB
 httpfs-site	03-01-2021 15:03	XML Document	1 KB
 kms-acls	03-01-2021 14:59	XML Document	4 KB
 kms-env.sh	03-01-2021 14:59	SH File	2 KB
 kms-log4j.properties	03-01-2021 14:59	PROPERTIES File	2 KB
 kms-site	03-01-2021 14:59	XML Document	1 KB
 log4j.properties	03-01-2021 14:58	PROPERTIES File	15 KB
 mapred-env	03-01-2021 15:27	Windows Command ...	1 KB
 mapred-env.sh	03-01-2021 15:27	SH File	2 KB
 mapred-queues.xml.template	03-01-2021 15:27	TEMPLATE File	5 KB
 mapred-site	03-01-2021 15:27	XML Document	1 KB
 ssl-client.xml.example	03-01-2021 14:58	EXAMPLE File	3 KB

# Steps to install Hadoop



```
<?xml version="1.0" encoding="UTF-8"?>
<?xml-stylesheet type="text/xsl" href="configuration.xsl"?>
<!--
Licensed under the Apache License, Version 2.0 (the "License");
you may not use this file except in compliance with the License.
You may obtain a copy of the License at

    http://www.apache.org/licenses/LICENSE-2.0

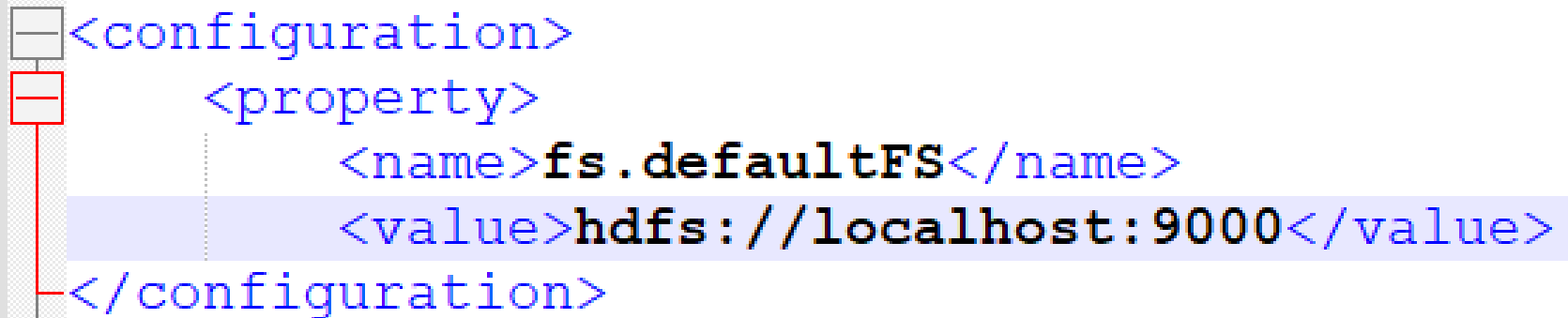
Unless required by applicable law or agreed to in writing, software
distributed under the License is distributed on an "AS IS" BASIS,
WITHOUT WARRANTIES OR CONDITIONS OF ANY KIND, either express or implied.
See the License for the specific language governing permissions and
limitations under the License. See accompanying LICENSE file.
-->

<!-- Put site-specific property overrides in this file. -->

<configuration>
</configuration>
```

# Steps to install Hadoop

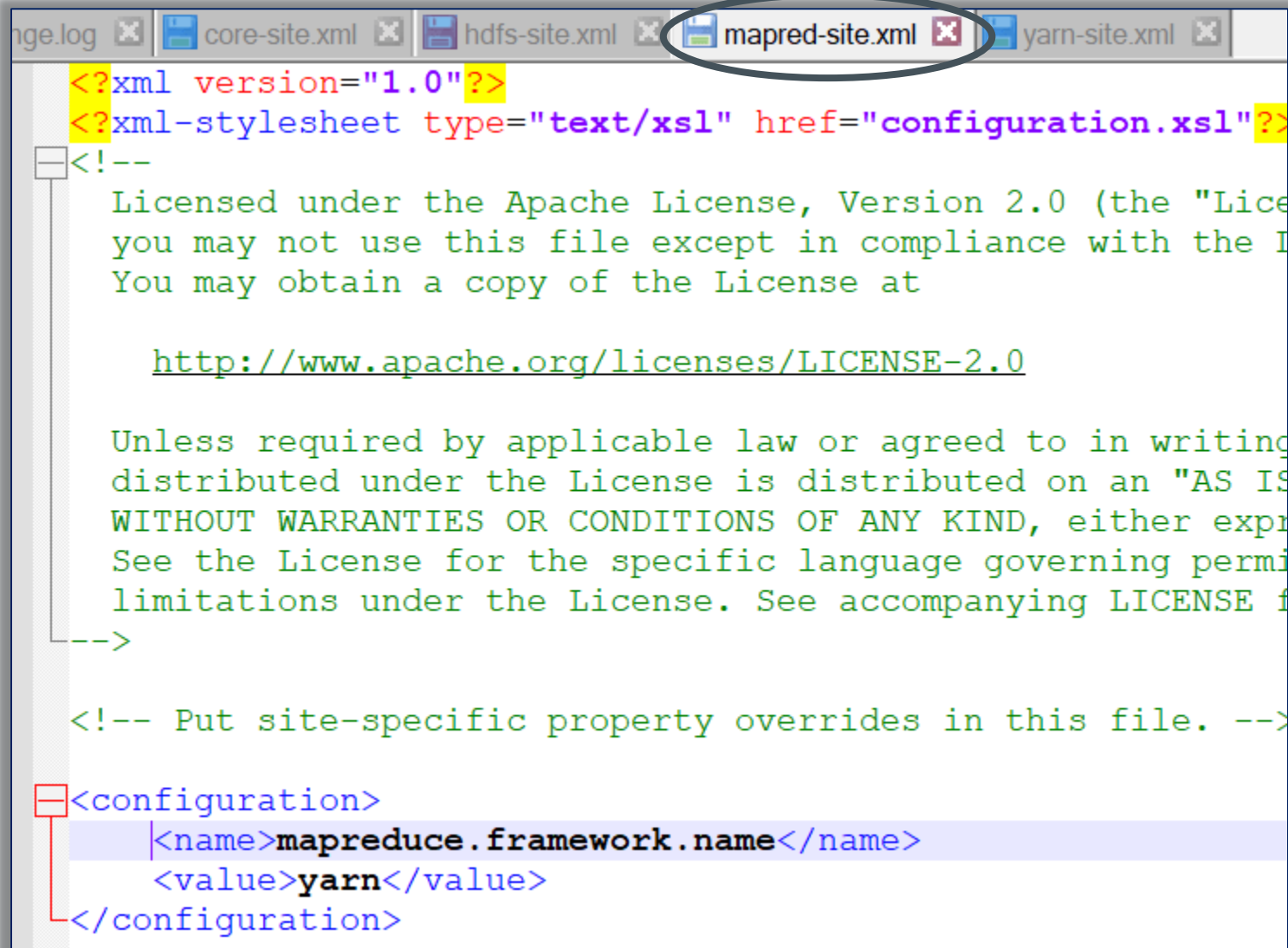
core-site.xml



The image shows a snippet of XML code within an editor. On the left, there is a tree view with a root node 'configuration' and a child node 'property'. The 'property' node is selected, highlighted with a red border, and a red line connects it to the corresponding XML element in the code. The XML code is as follows:

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

# Steps to install Hadoop



```
<?xml version="1.0"?>
<?xml-stylesheet type="text/xsl" href="configuration.xsl"?>
<!--
Licensed under the Apache License, Version 2.0 (the "License");
you may not use this file except in compliance with the License.
You may obtain a copy of the License at

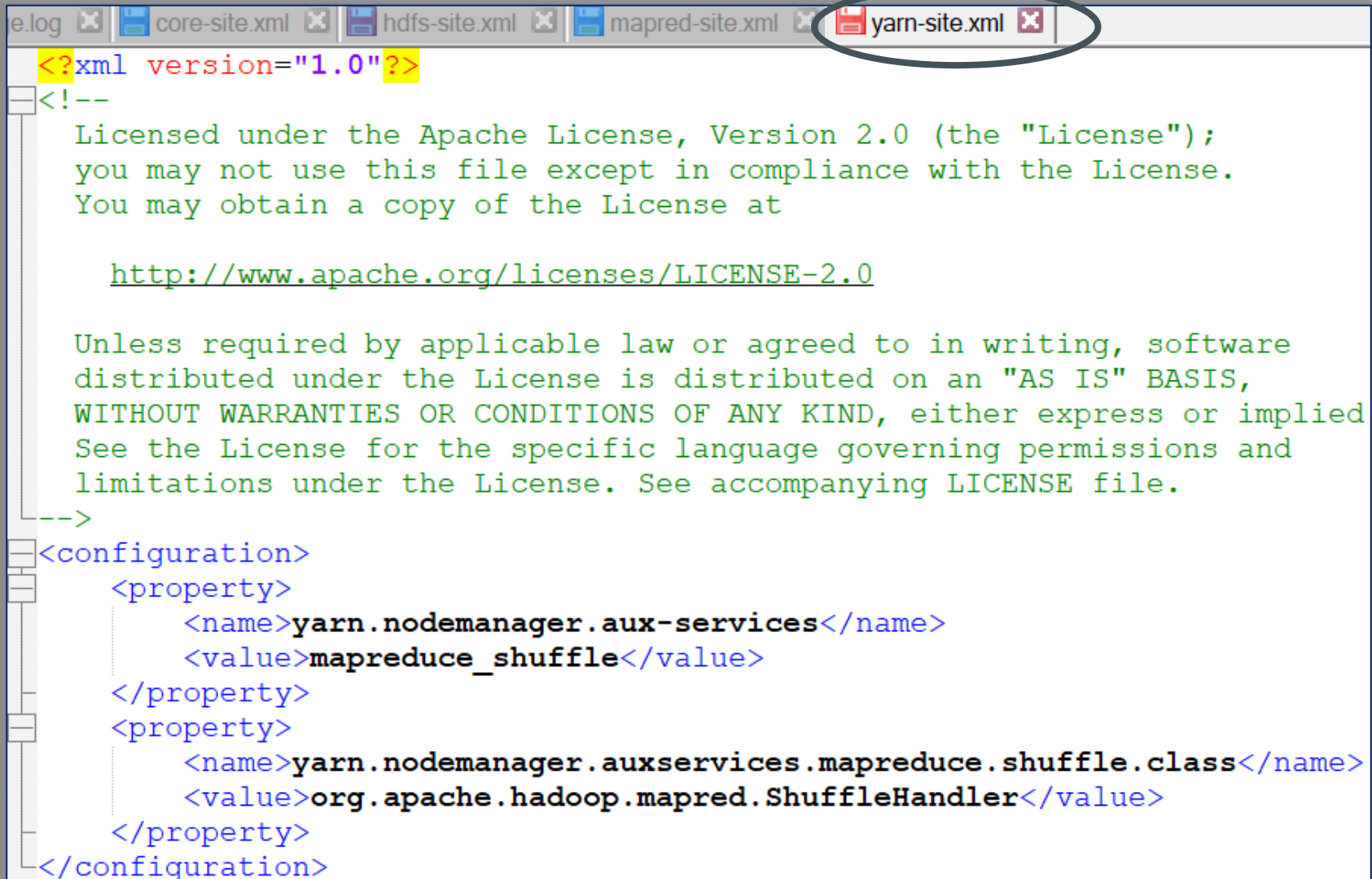
    http://www.apache.org/licenses/LICENSE-2.0

Unless required by applicable law or agreed to in writing, software
distributed under the License is distributed on an "AS IS" BASIS,
WITHOUT WARRANTIES OR CONDITIONS OF ANY KIND, either expressed or
implied. See the License for the specific language governing permissions
and limitations under the License. See accompanying LICENSE file for
details.
-->

<!-- Put site-specific property overrides in this file. -->

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

# Steps to install Hadoop



```
<?xml version="1.0"?>
<!--
Licensed under the Apache License, Version 2.0 (the "License");
you may not use this file except in compliance with the License.
You may obtain a copy of the License at

    http://www.apache.org/licenses/LICENSE-2.0

Unless required by applicable law or agreed to in writing, software
distributed under the License is distributed on an "AS IS" BASIS,
WITHOUT WARRANTIES OR CONDITIONS OF ANY KIND, either express or implied.
See the License for the specific language governing permissions and
limitations under the License. See accompanying LICENSE 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>
```



# Steps to install Hadoop

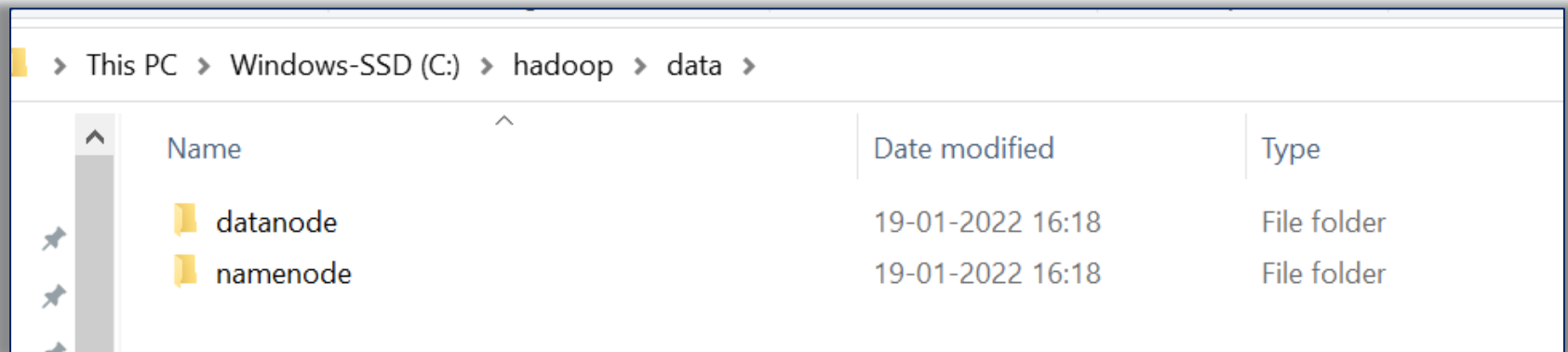
C:\hadoop



This PC > Windows-SSD (C:) > hadoop >

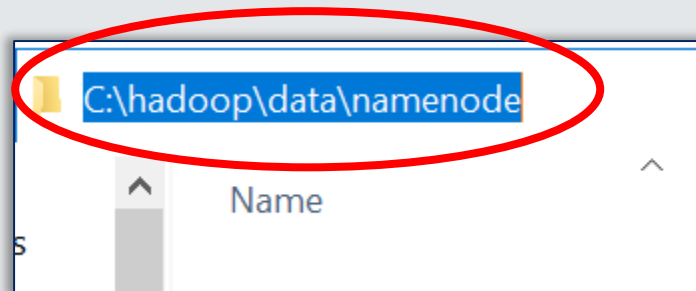
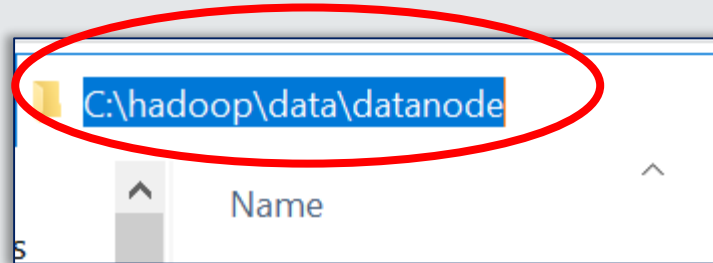
Name	Date modified	Type	Size
bin	03-01-2021 15:41	File folder	
etc	03-01-2021 14:59	File folder	
include	03-01-2021 15:41	File folder	
lib	03-01-2021 15:41	File folder	
libexec	03-01-2021 15:41	File folder	
data	19-01-2022 16:17	File folder	
sbin	03-01-2021 14:59	File folder	
share	03-01-2021 16:16	File folder	
LICENSE	05-12-2020 20:39	Text Document	148 KB
NOTICE	05-12-2020 20:39	Text Document	22 KB
README	05-12-2020 20:39	Text Document	2 KB

# Steps to install Hadoop

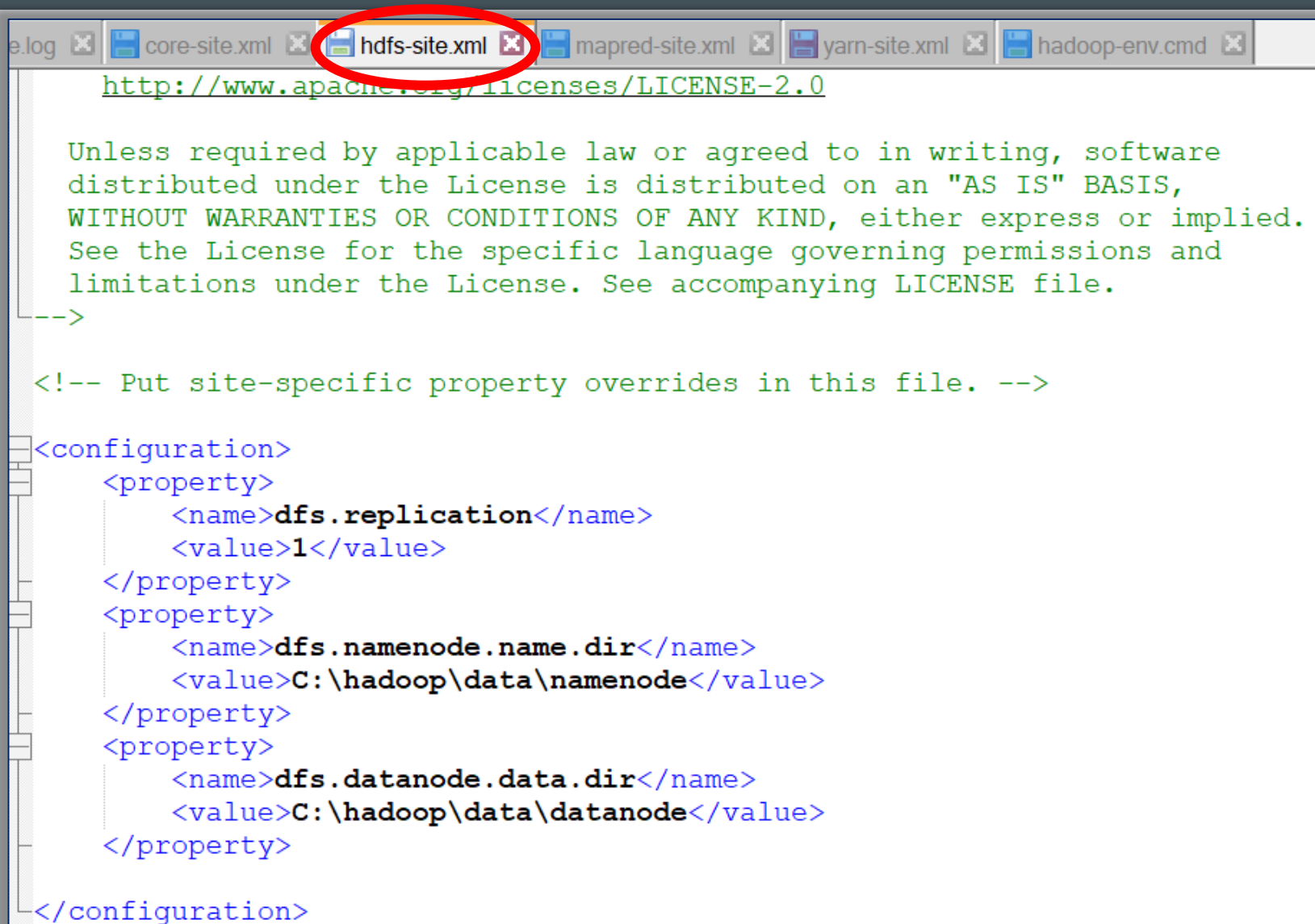


> This PC > Windows-SSD (C:) > hadoop > data >			
	Name	Date modified	Type
	datanode	19-01-2022 16:18	File folder
	namenode	19-01-2022 16:18	File folder

# Steps to install Hadoop



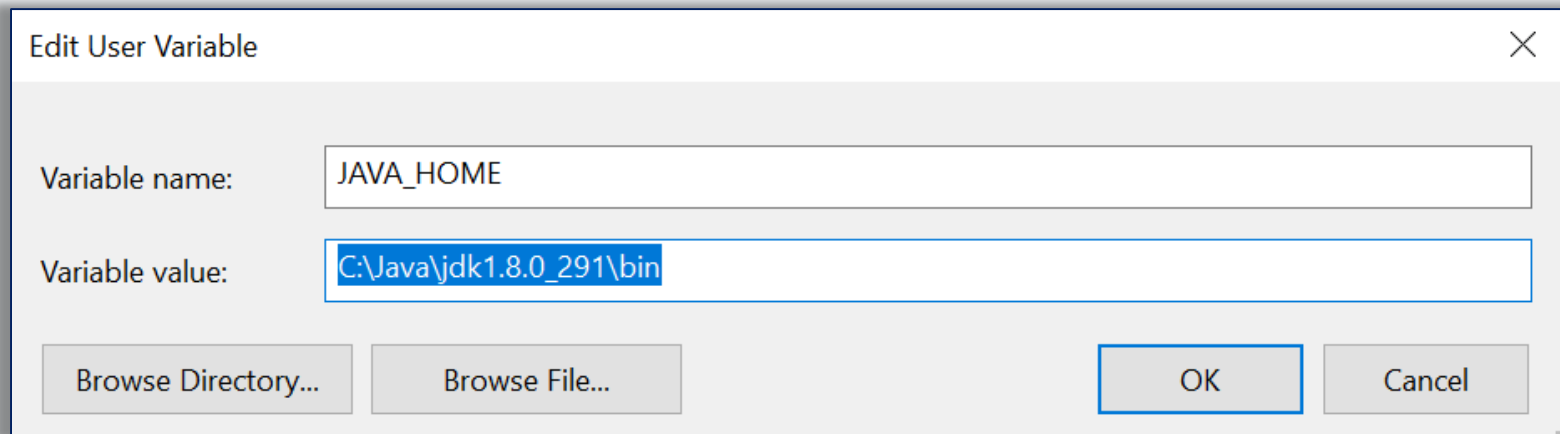
# Steps to install hadoop



The screenshot shows a web browser window with multiple tabs. The active tab is titled 'hdfs-site.xml' and is circled in red. The address bar shows the URL 'http://www.apache.org/licenses/LICENSE-2.0'. The main content area displays the text of the license, followed by a series of XML configuration tags for HDFS. The tags are as follows:

```
-->  
  
<!-- Put site-specific property overrides in this file. -->  
  
<configuration>  
  <property>  
    <name>dfs.replication</name>  
    <value>1</value>  
  </property>  
  <property>  
    <name>dfs.namenode.name.dir</name>  
    <value>C:\hadoop\data\namenode</value>  
  </property>  
  <property>  
    <name>dfs.datanode.data.dir</name>  
    <value>C:\hadoop\data\datanode</value>  
  </property>  
</configuration>
```

# Steps to install hadoop



Edit User Variable

Variable name: JAVA\_HOME

Variable value: C:\Java\jdk1.8.0\_291\bin

Browse Directory... Browse File... OK Cancel

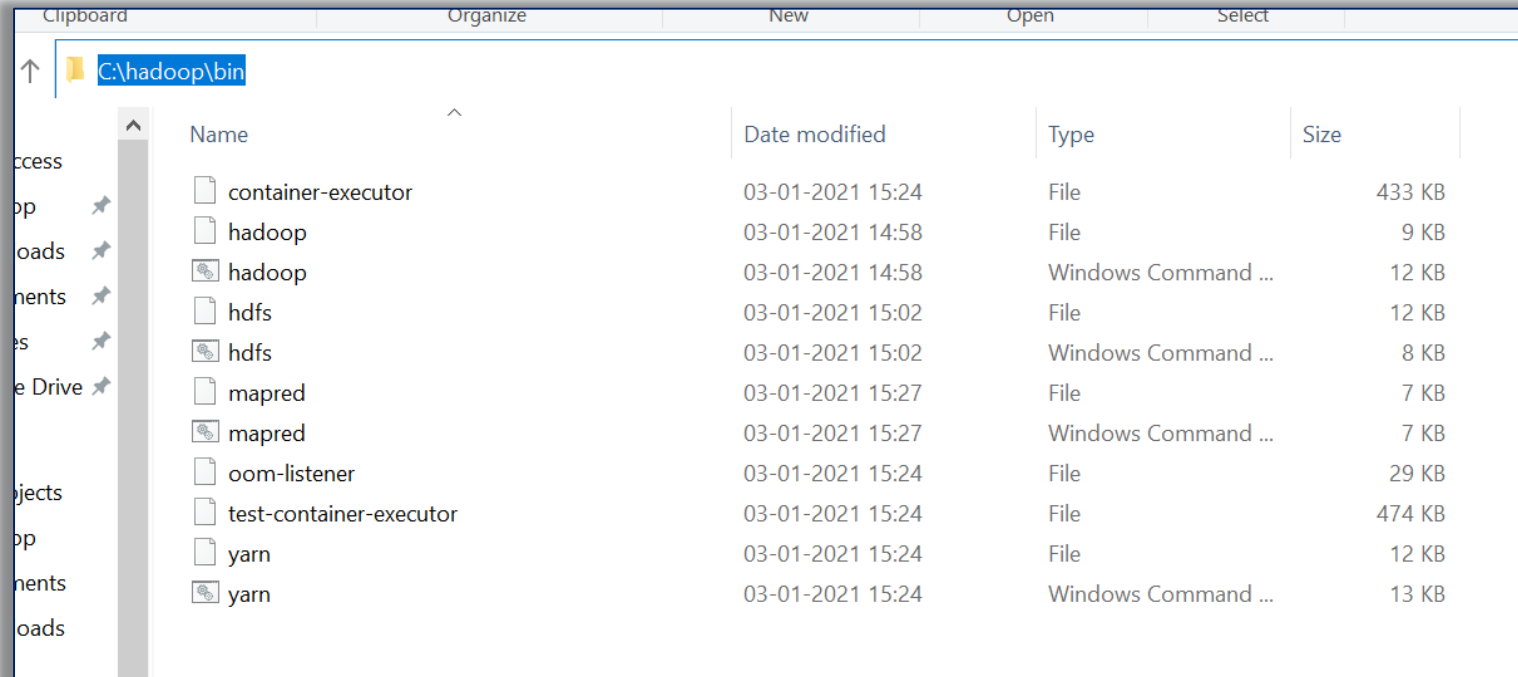
To get the correct location, copy from  
path

# Steps to install hadoop



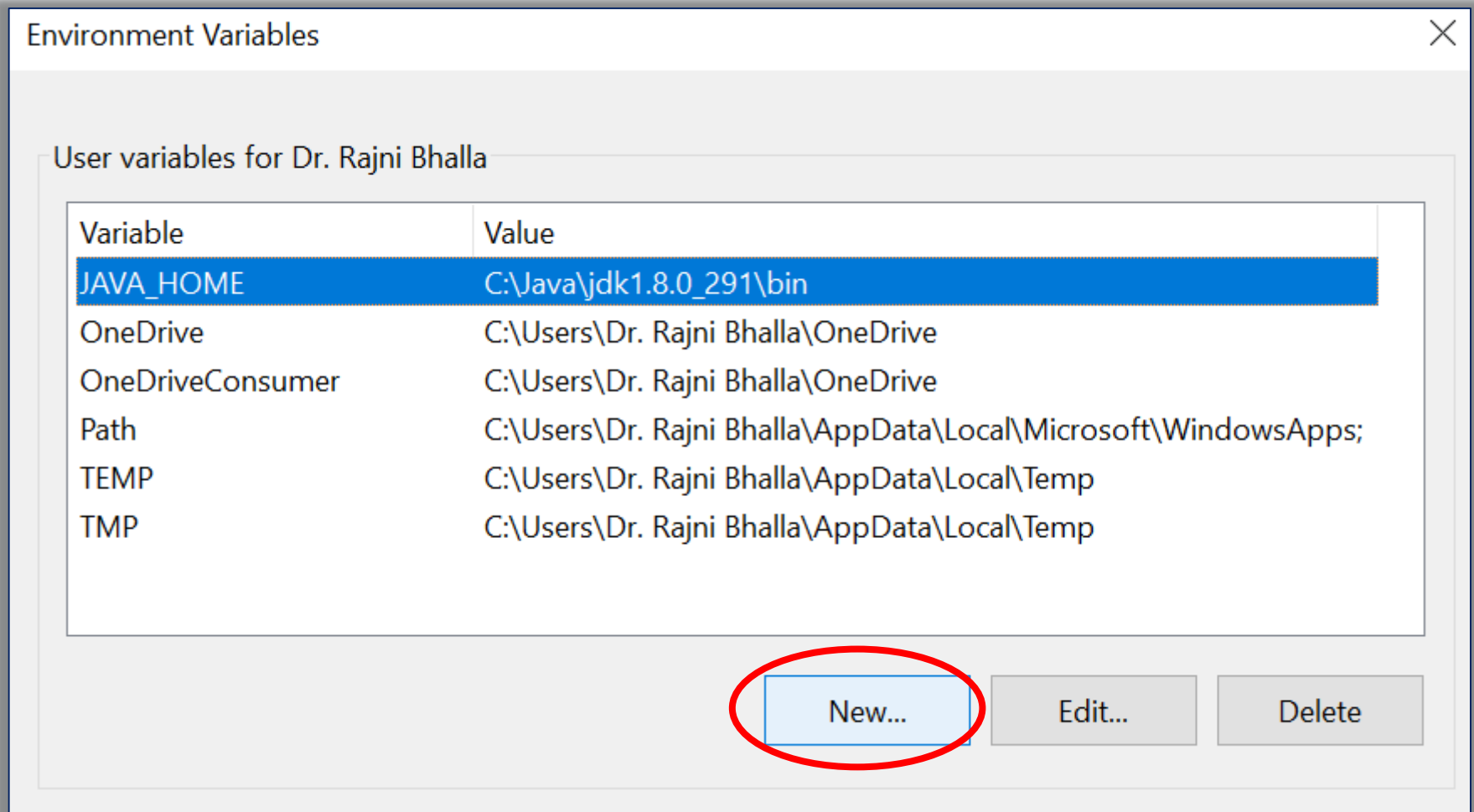
```
change.log x core-site.xml x hdfs-site.xml x mapred-site.xml x yarn-site.xml x hadoop-env.cmd x
1 @echo off
2 @rem Licensed to the Apache Software Foundation (ASF) under one or more
3 @rem contributor license agreements. See the NOTICE file distributed with
4 @rem this work for additional information regarding copyright ownership.
5 @rem The ASF licenses this file to You under the Apache License, Version 2.
6 @rem (the "License"); you may not use this file except in compliance with
7 @rem the License. You may obtain a copy of the License at
8 @rem
9 @rem    http://www.apache.org/licenses/LICENSE-2.0
10 @rem
11 @rem Unless required by applicable law or agreed to in writing, software
12 @rem distributed under the License is distributed on an "AS IS" BASIS,
13 @rem WITHOUT WARRANTIES OR CONDITIONS OF ANY KIND, either express or implied.
14 @rem See the License for the specific language governing permissions and
15 @rem limitations under the License.
16
17 @rem Set Hadoop-specific environment variables here.
18
19 @rem The only required environment variable is JAVA_HOME. All others are
20 @rem optional. When running a distributed configuration it is best to
21 @rem set JAVA_HOME in this file, so that it is correctly defined on
22 @rem remote nodes.
23
24 @rem The java implementation to use. Required.
25 set JAVA_HOME=C:\Java\jdk1.8.0_291\bin
26
27 @rem The jsvc implementation to use. Jsvc is required to run secure datanode
28 @rem set JSVC_HOME=%JSVC_HOME%
```

# Steps to install hadoop



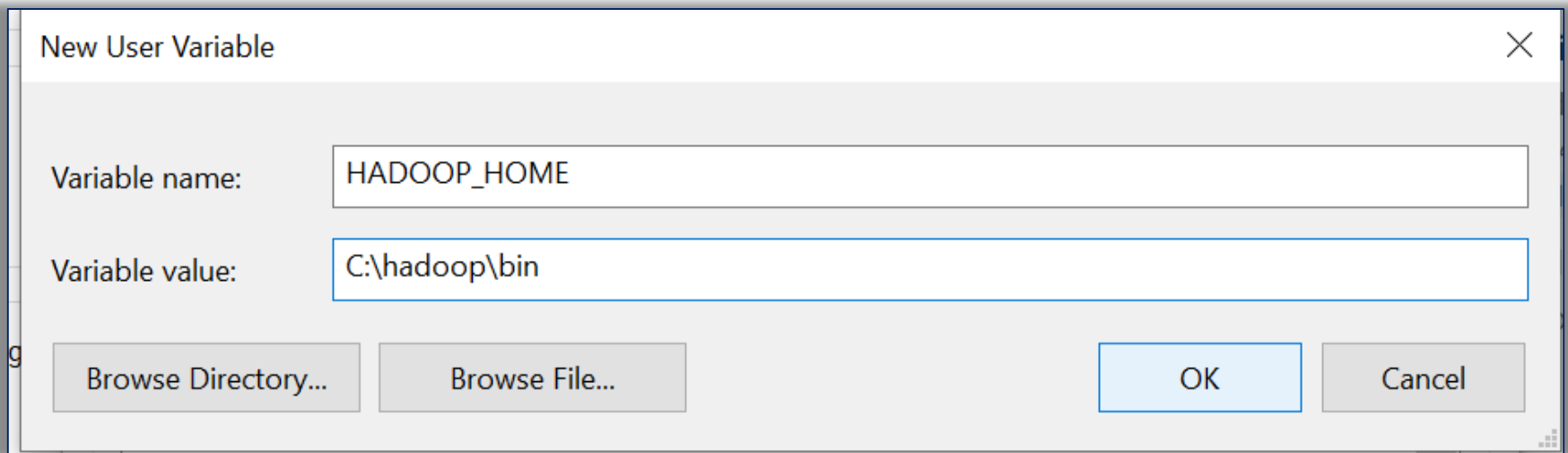
Clipboard					Organize					New					Open					Select				
↑					C:\hadoop\bin																			
Name					Date modified					Type					Size									
container-executor					03-01-2021 15:24					File					433 KB									
hadoop					03-01-2021 14:58					File					9 KB									
hadoop					03-01-2021 14:58					Windows Command ...					12 KB									
hdfs					03-01-2021 15:02					File					12 KB									
hdfs					03-01-2021 15:02					Windows Command ...					8 KB									
mapred					03-01-2021 15:27					File					7 KB									
mapred					03-01-2021 15:27					Windows Command ...					7 KB									
oom-listener					03-01-2021 15:24					File					29 KB									
test-container-executor					03-01-2021 15:24					File					474 KB									
yarn					03-01-2021 15:24					File					12 KB									
yarn					03-01-2021 15:24					Windows Command ...					13 KB									

# Steps to install hadoop





# Steps to install hadoop



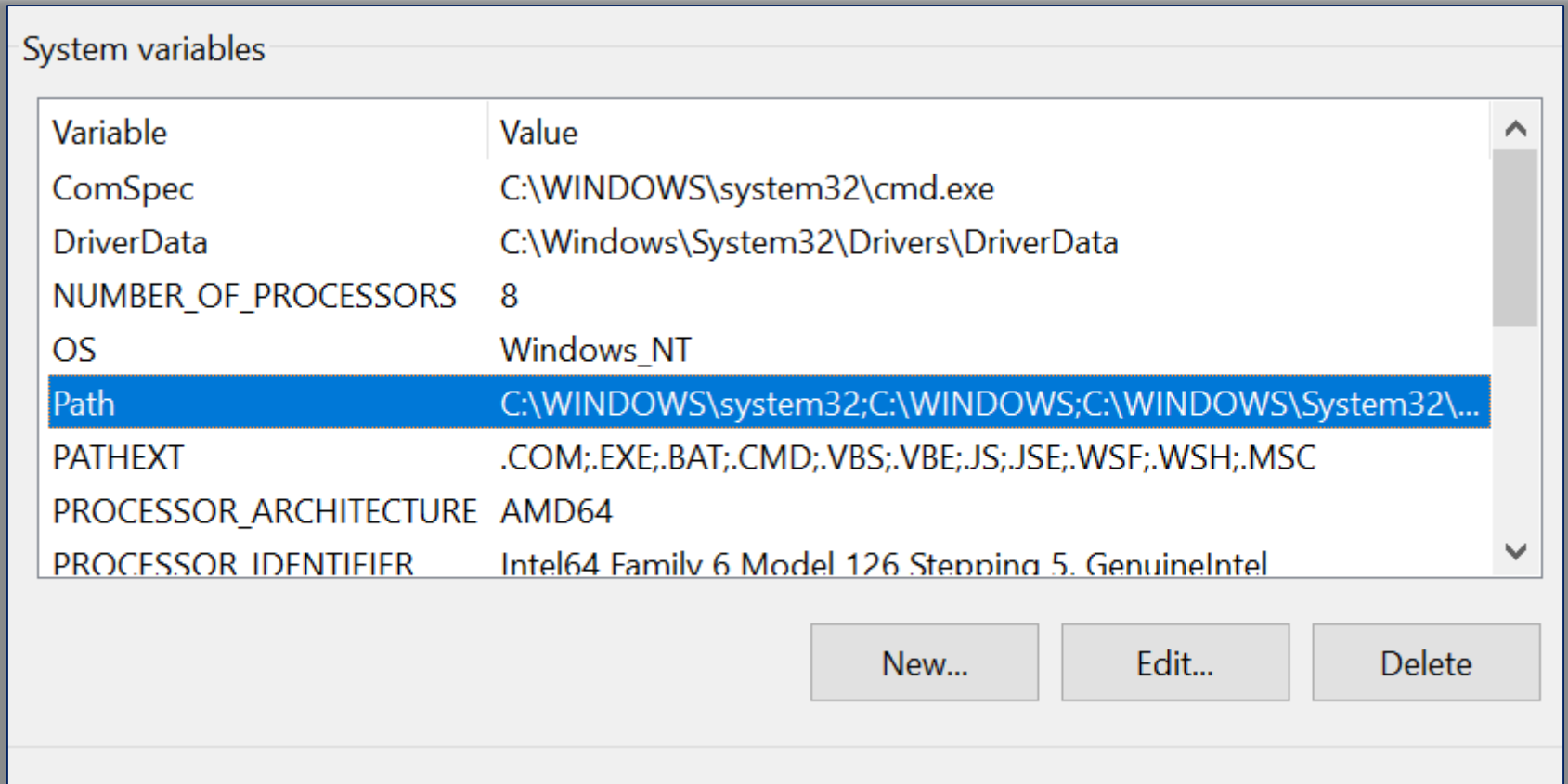
New User Variable

Variable name: HADOOP\_HOME

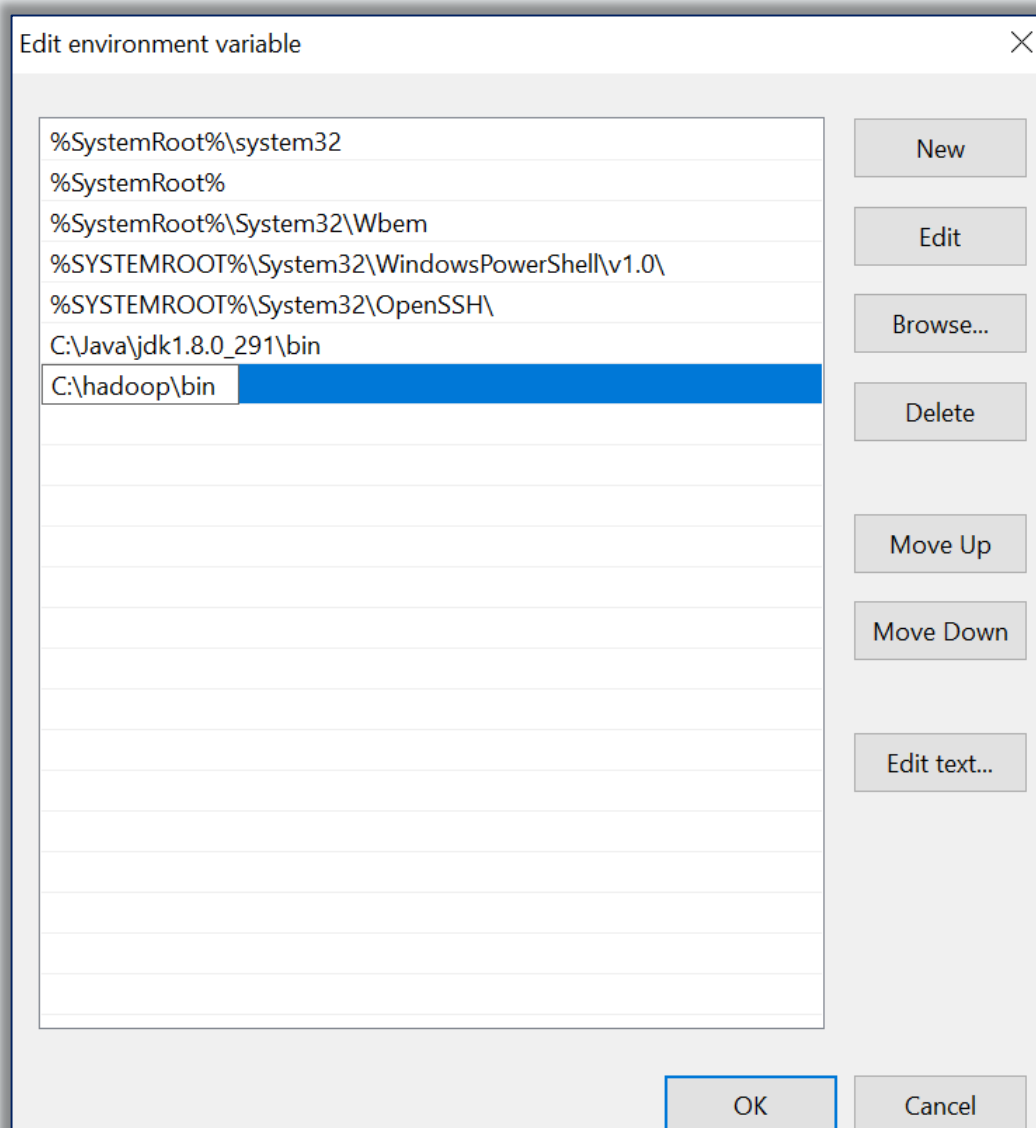
Variable value: C:\hadoop\bin

Browse Directory... Browse File... OK Cancel

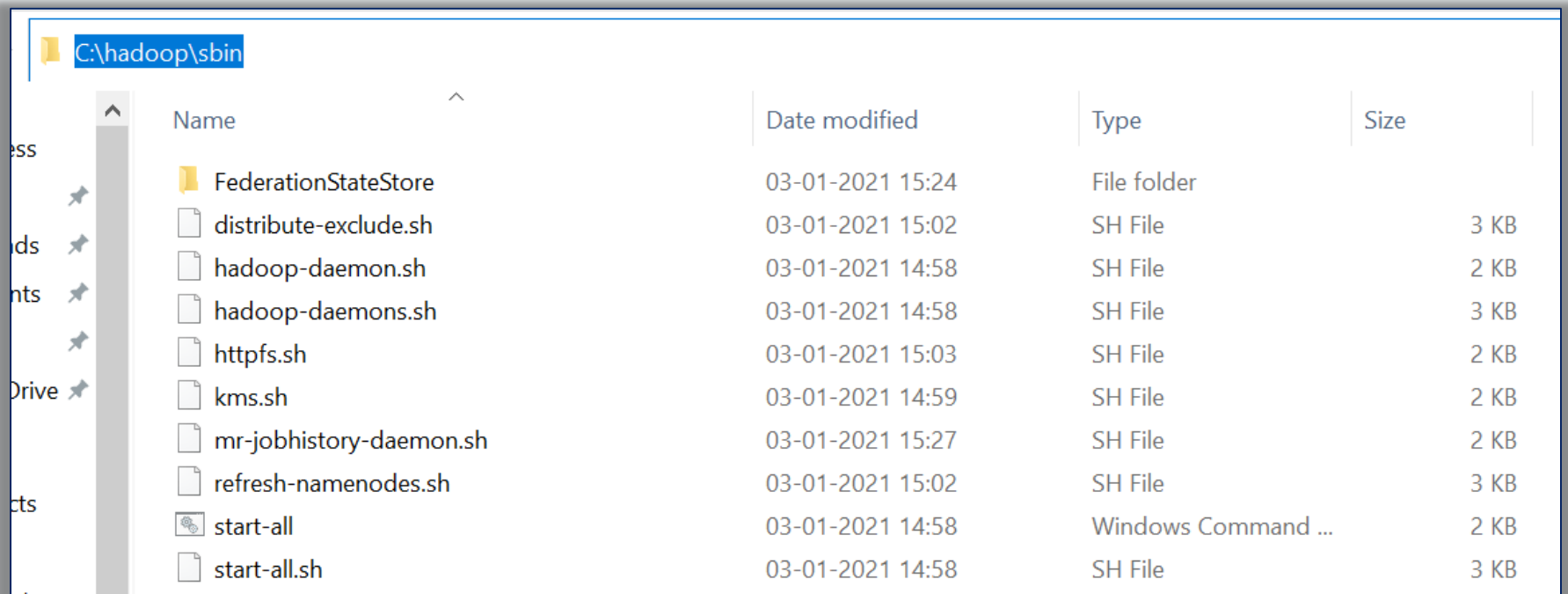
# Steps to install hadoop



# Steps to install Hadoop

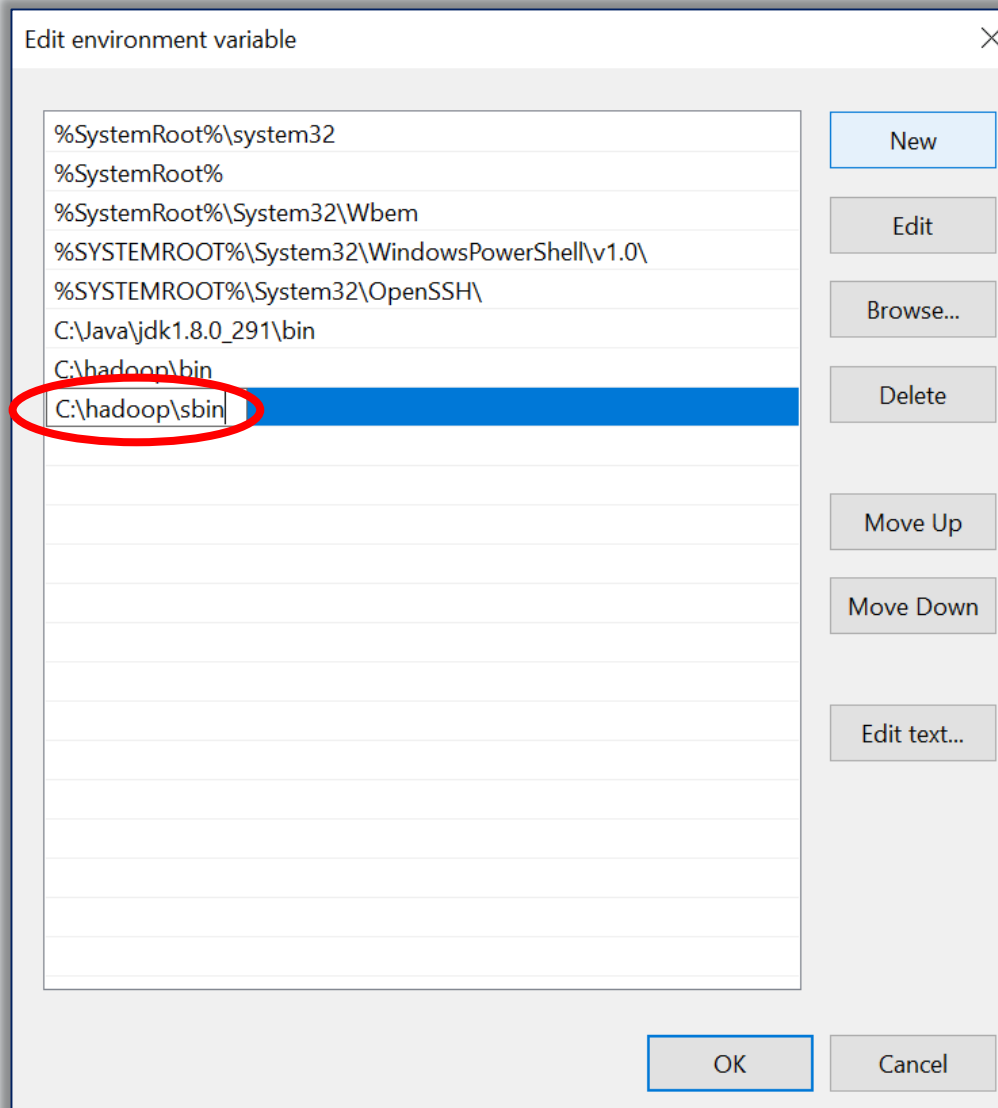


# Steps to install Hadoop

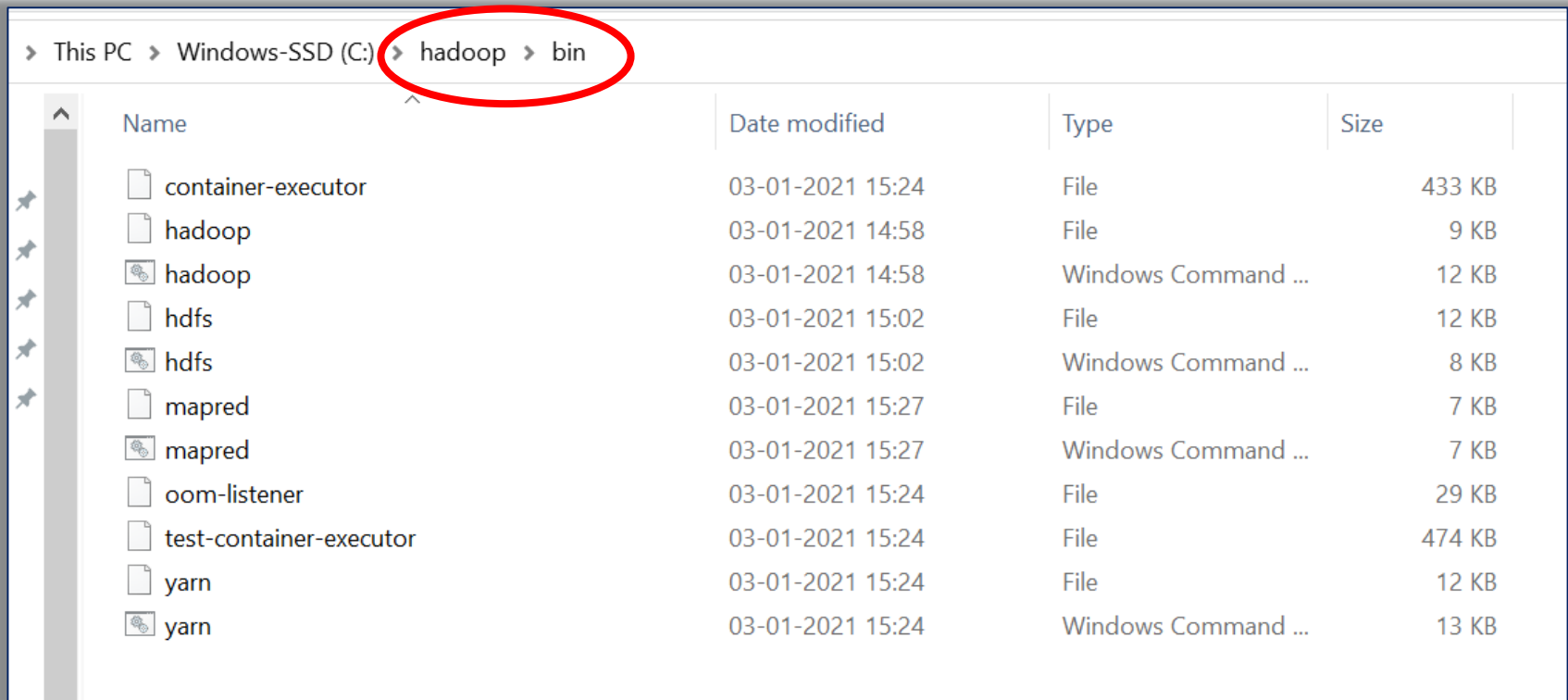


Name	Date modified	Type	Size
FederationStateStore	03-01-2021 15:24	File folder	
distribute-exclude.sh	03-01-2021 15:02	SH File	3 KB
hadoop-daemon.sh	03-01-2021 14:58	SH File	2 KB
hadoop-daemons.sh	03-01-2021 14:58	SH File	3 KB
httpfs.sh	03-01-2021 15:03	SH File	2 KB
kms.sh	03-01-2021 14:59	SH File	2 KB
mr-jobhistory-daemon.sh	03-01-2021 15:27	SH File	2 KB
refresh-namenodes.sh	03-01-2021 15:02	SH File	3 KB
start-all	03-01-2021 14:58	Windows Command ...	2 KB
start-all.sh	03-01-2021 14:58	SH File	3 KB

# Steps to install Hadoop







# Steps to install Hadoop

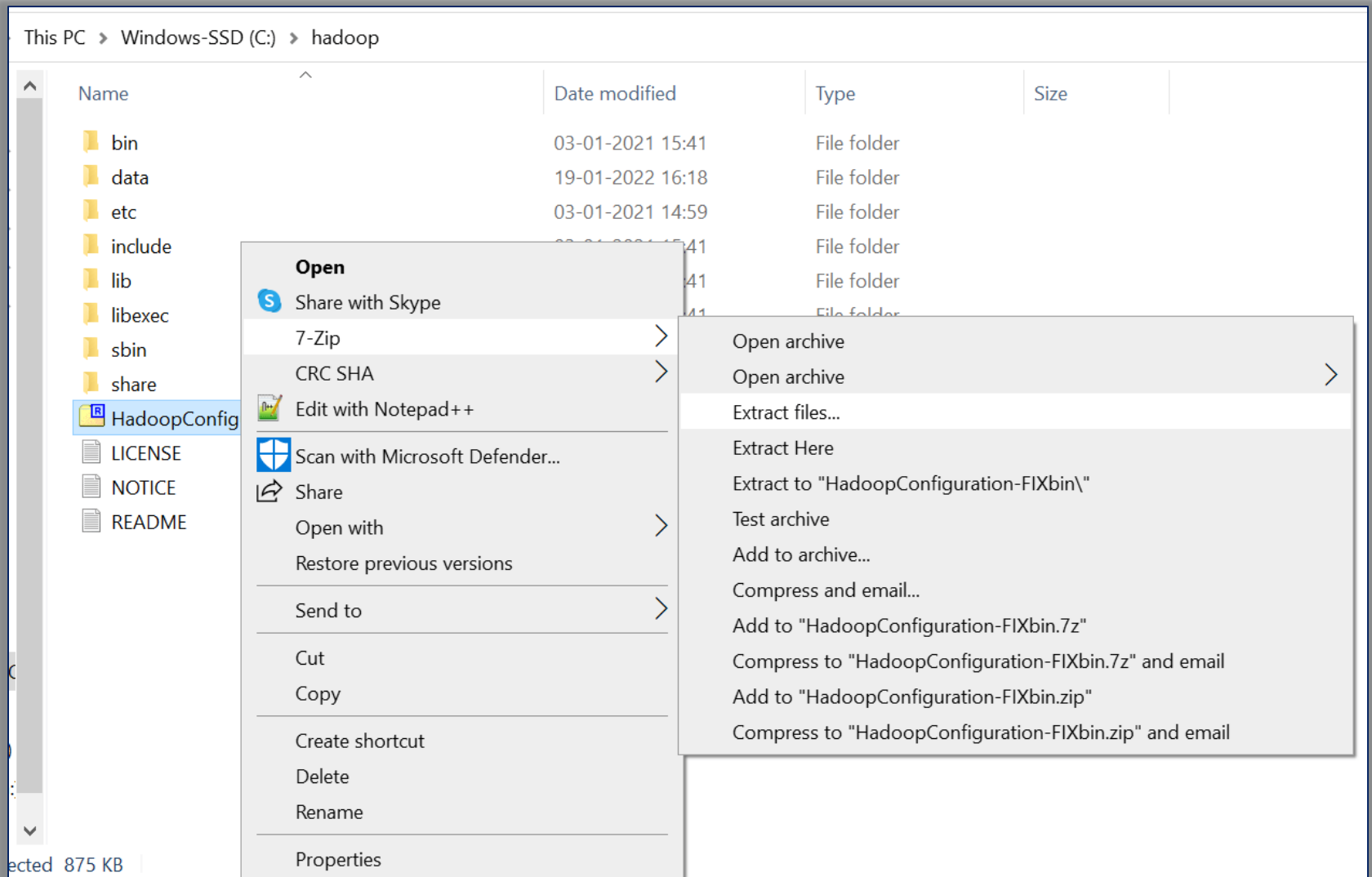


# Steps to install Hadoop

This PC > Windows-SSD (C:) > hadoop >

Name	Date modified	Type	Size
bin	03-01-2021 15:41	File folder	
data	19-01-2022 16:18	File folder	
etc	03-01-2021 14:59	File folder	
include	03-01-2021 15:41	File folder	
lib	03-01-2021 15:41	File folder	
libexec	03-01-2021 15:41	File folder	
sbin	03-01-2021 14:59	File folder	
share	03-01-2021 16:16	File folder	
 HadoopConfiguration-FIXbin	19-01-2022 16:58	RAR-Archive	876 KB
 LICENSE	05-12-2020 20:39	Text Document	148 KB
 NOTICE	05-12-2020 20:39	Text Document	22 KB
 README	05-12-2020 20:39	Text Document	2 KB





# Steps to install Hadoop



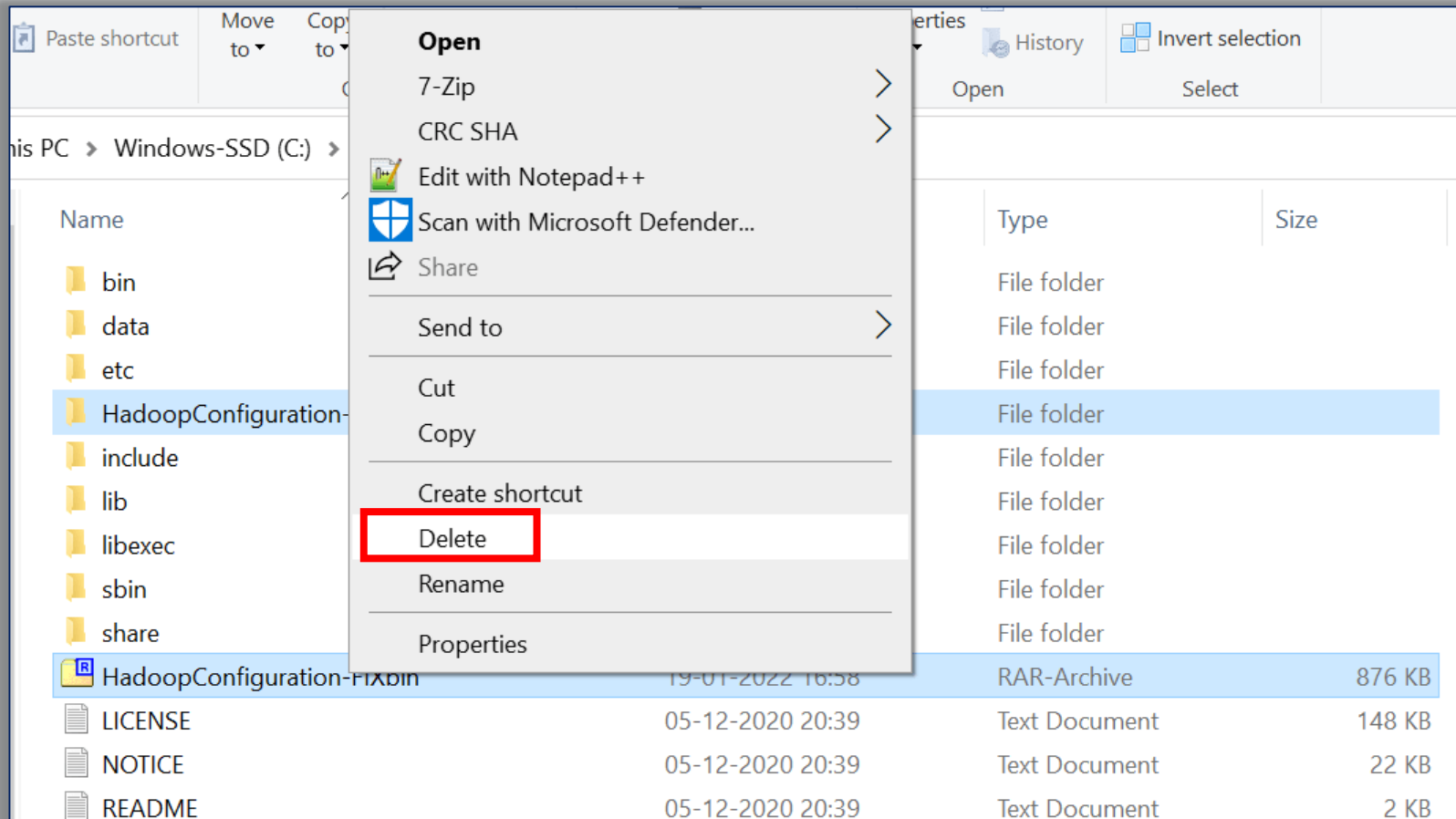


# Steps to install Hadoop

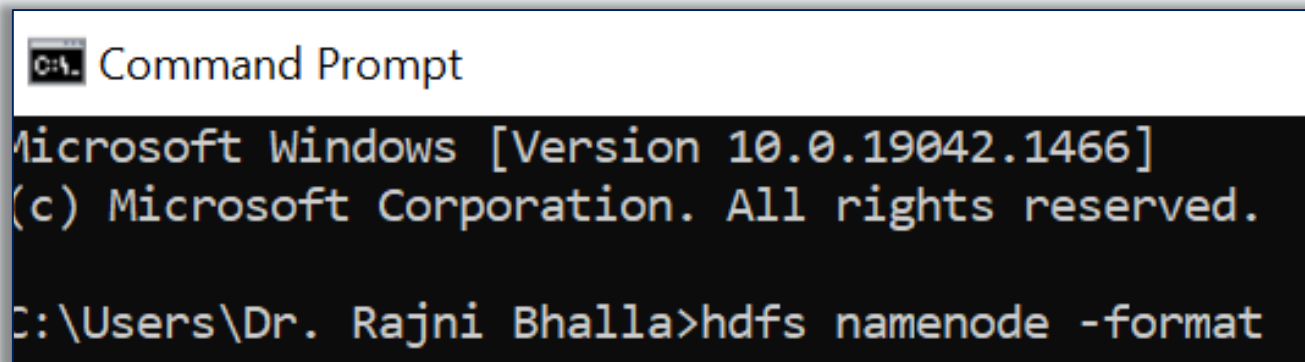
This PC > Windows-SSD (C:) > hadoop >

Name	Date modified	Type	Size
bin	03-01-2021 15:41	File folder	
data	19-01-2022 16:18	File folder	
etc	03-01-2021 14:59	File folder	
HadoopConfiguration-FIXbin	21-01-2022 10:29	File folder	
include	03-01-2021 15:41	File folder	
lib	03-01-2021 15:41	File folder	
libexec	03-01-2021 15:41	File folder	
sbin	03-01-2021 14:59	File folder	
share	03-01-2021 16:16	File folder	
 HadoopConfiguration-FIXbin	19-01-2022 16:58	RAR-Archive	876 KB
 LICENSE	05-12-2020 20:39	Text Document	148 KB
 NOTICE	05-12-2020 20:39	Text Document	22 KB
 README	05-12-2020 20:39	Text Document	2 KB

# Steps to install Hadoop



# Steps to install Hadoop



```
C:\> Command Prompt

Microsoft Windows [Version 10.0.19042.1466]
(c) Microsoft Corporation. All rights reserved.

C:\Users\Dr. Rajni Bhalla>hdfs namenode -format
```

# Steps to install Hadoop

```
at org.apache.hadoop.conf.Configuration$Parser.parseNext(Configuration.java:3336)
at org.apache.hadoop.conf.Configuration$Parser.parse(Configuration.java:3130)
at org.apache.hadoop.conf.Configuration.loadResource(Configuration.java:3023)
at org.apache.hadoop.conf.Configuration.loadResources(Configuration.java:2984)
at org.apache.hadoop.conf.Configuration.loadProps(Configuration.java:2864)
at org.apache.hadoop.conf.Configuration.getProps(Configuration.java:2846)
at org.apache.hadoop.conf.Configuration.get(Configuration.java:1200)
at org.apache.hadoop.conf.Configuration.getTimeDuration(Configuration.java:1812)
at org.apache.hadoop.conf.Configuration.getTimeDuration(Configuration.java:1789)
at org.apache.hadoop.util.ShutdownHookManager.getShutdownTimeout(ShutdownHookManager.java:183)
at org.apache.hadoop.util.ShutdownHookManager.shutdownExecutor(ShutdownHookManager.java:145)
at org.apache.hadoop.util.ShutdownHookManager.access$300(ShutdownHookManager.java:65)
at org.apache.hadoop.util.ShutdownHookManager$1.run(ShutdownHookManager.java:102)
Exception in thread "Thread-0" java.lang.RuntimeException: com.ctc.wstx.exc.WstxParsingException: Unexpected close tag </configuration>
[row,col,system-id]: [23,15,"file:/C:/hadoop/etc/hadoop/core-site.xml"]
at org.apache.hadoop.conf.Configuration.loadResource(Configuration.java:3040)
at org.apache.hadoop.conf.Configuration.loadResources(Configuration.java:2984)
at org.apache.hadoop.conf.Configuration.loadProps(Configuration.java:2864)
at org.apache.hadoop.conf.Configuration.getProps(Configuration.java:2846)
at org.apache.hadoop.conf.Configuration.get(Configuration.java:1200)
at org.apache.hadoop.conf.Configuration.getTimeDuration(Configuration.java:1812)
at org.apache.hadoop.conf.Configuration.getTimeDuration(Configuration.java:1789)
at org.apache.hadoop.util.ShutdownHookManager.getShutdownTimeout(ShutdownHookManager.java:183)
at org.apache.hadoop.util.ShutdownHookManager.shutdownExecutor(ShutdownHookManager.java:145)
at org.apache.hadoop.util.ShutdownHookManager.access$300(ShutdownHookManager.java:65)
at org.apache.hadoop.util.ShutdownHookManager$1.run(ShutdownHookManager.java:102)
Caused by: com.ctc.wstx.exc.WstxParsingException: Unexpected close tag </configuration>; expected </property>.
[row,col,system-id]: [23,15,"file:/C:/hadoop/etc/hadoop/core-site.xml"]
at com.ctc.wstx.sr.StreamScanner.constructWfcException(StreamScanner.java:621)
at com.ctc.wstx.sr.StreamScanner.throwParseError(StreamScanner.java:491)
at com.ctc.wstx.sr.StreamScanner.throwParseError(StreamScanner.java:475)
at com.ctc.wstx.sr.BasicStreamReader.reportWrongEndElem(BasicStreamReader.java:3365)
at com.ctc.wstx.sr.BasicStreamReader.readEndElem(BasicStreamReader.java:3292)
at com.ctc.wstx.sr.BasicStreamReader.nextFromTree(BasicStreamReader.java:2911)
at com.ctc.wstx.sr.BasicStreamReader.next(BasicStreamReader.java:1123)
at org.apache.hadoop.conf.Configuration$Parser.parseNext(Configuration.java:3336)
at org.apache.hadoop.conf.Configuration$Parser.parse(Configuration.java:3130)
at org.apache.hadoop.conf.Configuration.loadResource(Configuration.java:3023)
... 10 more
```

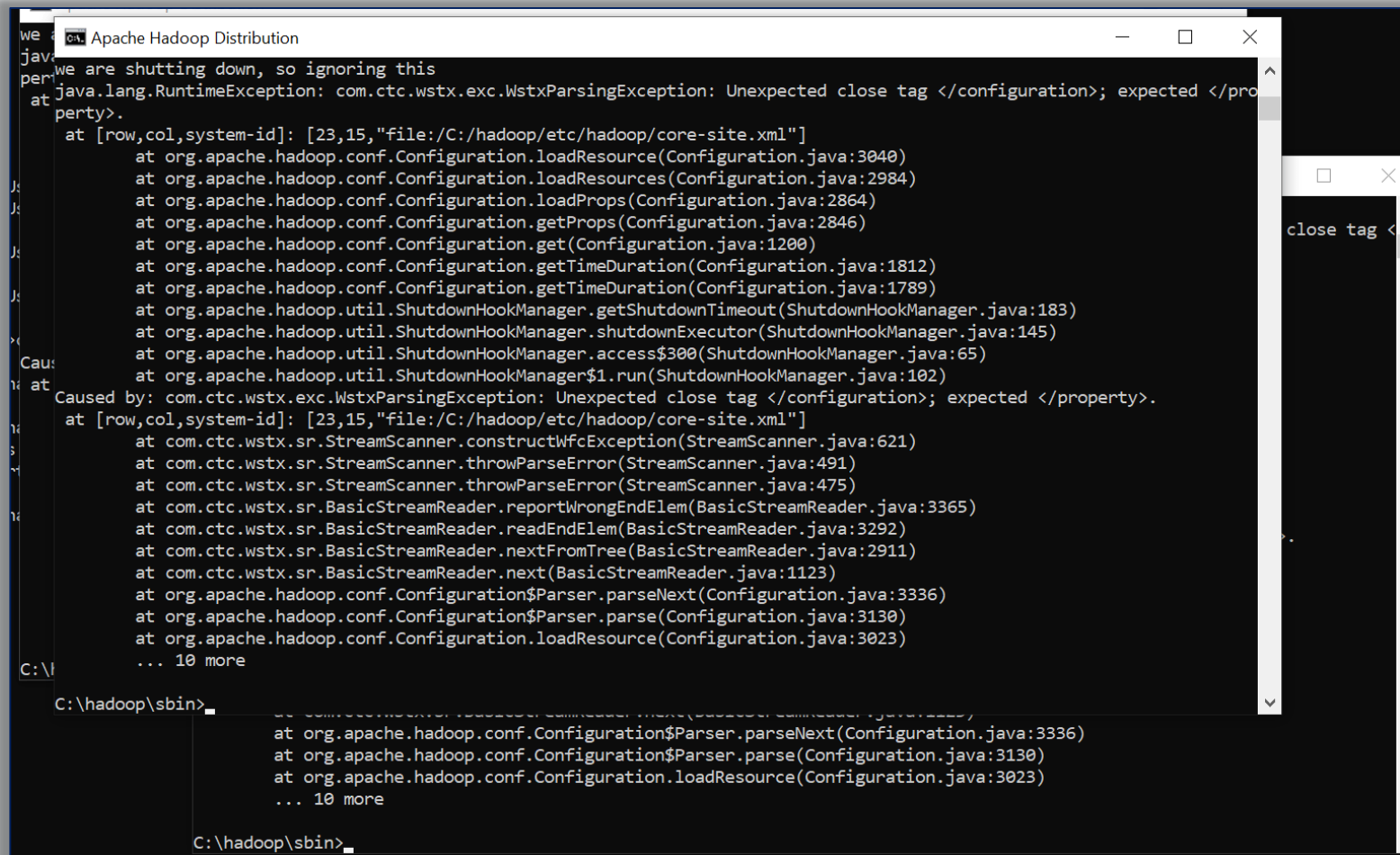
# Steps to install Hadoop

```
C:\Users\Dr. Rajni Bhalla>hdfs namenode -format
2022-01-21 11:37:41,968 INFO namenode.NameNode: STARTUP_MSG:
*****
STARTUP_MSG: Starting NameNode
STARTUP_MSG: host = LAPTOP-A1FEVWJ/10.33.160.168
STARTUP_MSG: args = [-format]
STARTUP_MSG: version = 3.2.2
STARTUP_MSG: classpath = C:\hadoop\etc\hadoop;C:\hadoop\share\hadoop\common;C:\hadoop\share\hadoop\common\lib\accessors-smart-1.2.jar;C:\hadoop\share\hadoop\common\lib\animal-sniffer-annotations-1.17.jar;C:\hadoop\share\hadoop\common\lib\asm-5.0.4.jar;C:\hadoop\share\hadoop\common\lib\audience-annotations-0.5.0.jar;C:\hadoop\share\hadoop\common\lib\avro-1.7.7.jar;C:\hadoop\share\hadoop\common\lib\checker-qual-2.5.2.jar;C:\hadoop\share\hadoop\common\lib\commons-beanutils-1.9.4.jar;C:\hadoop\share\hadoop\common\lib\commons-cli-1.2.jar;C:\hadoop\share\hadoop\common\lib\commons-codec-1.11.jar;C:\hadoop\share\hadoop\common\lib\commons-collections-3.2.2.jar;C:\hadoop\share\hadoop\common\lib\commons-compress-1.19.jar;C:\hadoop\share\hadoop\common\lib\commons-configuration2-2.1.1.jar;C:\hadoop\share\hadoop\common\lib\commons-io-2.5.jar;C:\hadoop\share\hadoop\common\lib\commons-lang3-3.7.jar;C:\hadoop\share\hadoop\common\lib\commons-logging-1.1.3.jar;C:\hadoop\share\hadoop\common\lib\commons-math3-3.1.1.jar;C:\hadoop\share\hadoop\common\lib\commons-net-3.6.jar;C:\hadoop\share\hadoop\common\lib\commons-text-1.4.jar;C:\hadoop\share\hadoop\common\lib\curator-client-2.13.0.jar;C:\hadoop\share\hadoop\common\lib\curator-framework-2.13.0.jar;C:\hadoop\share\hadoop\common\lib\curator-recipes-2.13.0.jar;C:\hadoop\share\hadoop\common\lib\dnsjava-2.1.7.jar;C:\hadoop\share\hadoop\common\lib\error_prone_annotations-2.2.0.jar;C:\hadoop\share\hadoop\common\lib\failureaccess-1.0.jar;C:\hadoop\share\hadoop\common\lib\gson-2.2.4.jar;C:\hadoop\share\hadoop\common\lib\guava-27.0-jre.jar;C:\hadoop\share\hadoop\common\lib\hadoop-annotations-3.2.2.jar;C:\hadoop\share\hadoop\common\lib\hadoop-auth-3.2.2.jar;C:\hadoop\share\hadoop\common\lib\htrace-core4-4.1.0-incubating.jar;C:\hadoop\share\hadoop\common\lib\httpclient-4.5.13.jar;C:\hadoop\share\hadoop\common\lib\httpcore-4.4.13.jar;C:\hadoop\share\hadoop\common\lib\j2objc-annotations-1.1.jar;C:\hadoop\share\hadoop\common\lib\jackson-annotations-2.9.10.jar;C:\hadoop\share\hadoop\common\lib\jackson-core-2.9.10.jar;C:\hadoop\share\hadoop\common\lib\jackson-core-asl-1.9.13.jar;C:\hadoop\share\hadoop\common\lib\jackson-databind-2.9.10.4.jar;C:\hadoop\share\hadoop\common\lib\jackson-jaxrs-1.9.13.jar;C:\hadoop\share\hadoop\common\lib\jackson-mapper-asl-1.9.13.jar;C:\hadoop\share\hadoop\common\lib\jackson-xc-1.9.13.jar;C:\hadoop\share\hadoop\common\lib\javax.activation-api-1.2.0.jar;C:\hadoop\share\hadoop\common\lib\javax.servlet-api-3.1.0.jar;C:\hadoop\share\hadoop\common\lib\jaxb-api-2.2.11.jar;C:\hadoop\share\hadoop\common\lib\jaxb-impl-2.2.3-1.jar;C:\hadoop\share\hadoop\common\lib\jcip-annotations-1.0-1.jar;C:\hadoop\share\hadoop\common\lib\jersey-core-1.19.jar;C:\hadoop\share\hadoop\common\lib\jersey-json-1.19.jar;C:\hadoop\share\hadoop\common\lib\jersey-server-1.19.jar;C:\hadoop\share\hadoop\common\lib\jersey-servlet-1.19.jar;C:\hadoop\share\hadoop\common\lib\jettison-1.1.jar;C:\hadoop\share\hadoop\common\lib\jetty-http-9.4.20.v20190813.jar;C:\hadoop\share\hadoop\common\lib\jetty-io-9.4.20.v20190813.jar;C:\hadoop\share\hadoop\common\lib\jetty-security-9.4.20.v20190813.jar;C:\hadoop\share\hadoop\common\lib\jetty-server-9.4.20.v20190813.jar;C:\hadoop\share\hadoop\common\lib\jetty-servlet-9.4.20.v20190813.jar
```

# Steps to install Hadoop

```
C:\Users\Dr. Rajni Bhalla>cd..  
  
C:\Users>cd..  
  
C:\>cd hadoop  
  
C:\hadoop>cd sbin  
  
C:\hadoop\sbin>_
```

# Steps to install Hadoop



```
we are shutting down, so ignoring this
java.lang.RuntimeException: com.ctc.wstx.exc.WstxParsingException: Unexpected close tag </configuration>; expected </property>.
    at [row,col,system-id]: [23,15,"file:/C:/hadoop/etc/hadoop/core-site.xml"]
    at org.apache.hadoop.conf.Configuration.loadResource(Configuration.java:3040)
    at org.apache.hadoop.conf.Configuration.loadResources(Configuration.java:2984)
    at org.apache.hadoop.conf.Configuration.loadProps(Configuration.java:2864)
    at org.apache.hadoop.conf.Configuration.getProps(Configuration.java:2846)
    at org.apache.hadoop.conf.Configuration.get(Configuration.java:1200)
    at org.apache.hadoop.conf.Configuration.getTimeDuration(Configuration.java:1812)
    at org.apache.hadoop.conf.Configuration.getTimeDuration(Configuration.java:1789)
    at org.apache.hadoop.util.ShutdownHookManager.getShutdownTimeout(ShutdownHookManager.java:183)
    at org.apache.hadoop.util.ShutdownHookManager.shutdownExecutor(ShutdownHookManager.java:145)
    at org.apache.hadoop.util.ShutdownHookManager.access$300(ShutdownHookManager.java:65)
    at org.apache.hadoop.util.ShutdownHookManager$1.run(ShutdownHookManager.java:102)
Caused by: com.ctc.wstx.exc.WstxParsingException: Unexpected close tag </configuration>; expected </property>.
    at [row,col,system-id]: [23,15,"file:/C:/hadoop/etc/hadoop/core-site.xml"]
    at com.ctc.wstx.sr.StreamScanner.constructWfcException(StreamScanner.java:621)
    at com.ctc.wstx.sr.StreamScanner.throwParseError(StreamScanner.java:491)
    at com.ctc.wstx.sr.StreamScanner.throwParseError(StreamScanner.java:475)
    at com.ctc.wstx.sr.BasicStreamReader.reportWrongEndElem(BasicStreamReader.java:3365)
    at com.ctc.wstx.sr.BasicStreamReader.readEndElem(BasicStreamReader.java:3292)
    at com.ctc.wstx.sr.BasicStreamReader.nextFromTree(BasicStreamReader.java:2911)
    at com.ctc.wstx.sr.BasicStreamReader.next(BasicStreamReader.java:1123)
    at org.apache.hadoop.conf.Configuration$Parser.parseNext(Configuration.java:3336)
    at org.apache.hadoop.conf.Configuration$Parser.parse(Configuration.java:3130)
    at org.apache.hadoop.conf.Configuration.loadResource(Configuration.java:3023)
    ... 10 more

C:\hadoop\sbin>
```



CMD  
**That's all for now...**