Introduction to Big Data

ECAP456

Dr. Rajni Bhalla

Associate Professor

Learning Outcomes

After this lecture, you will be able to

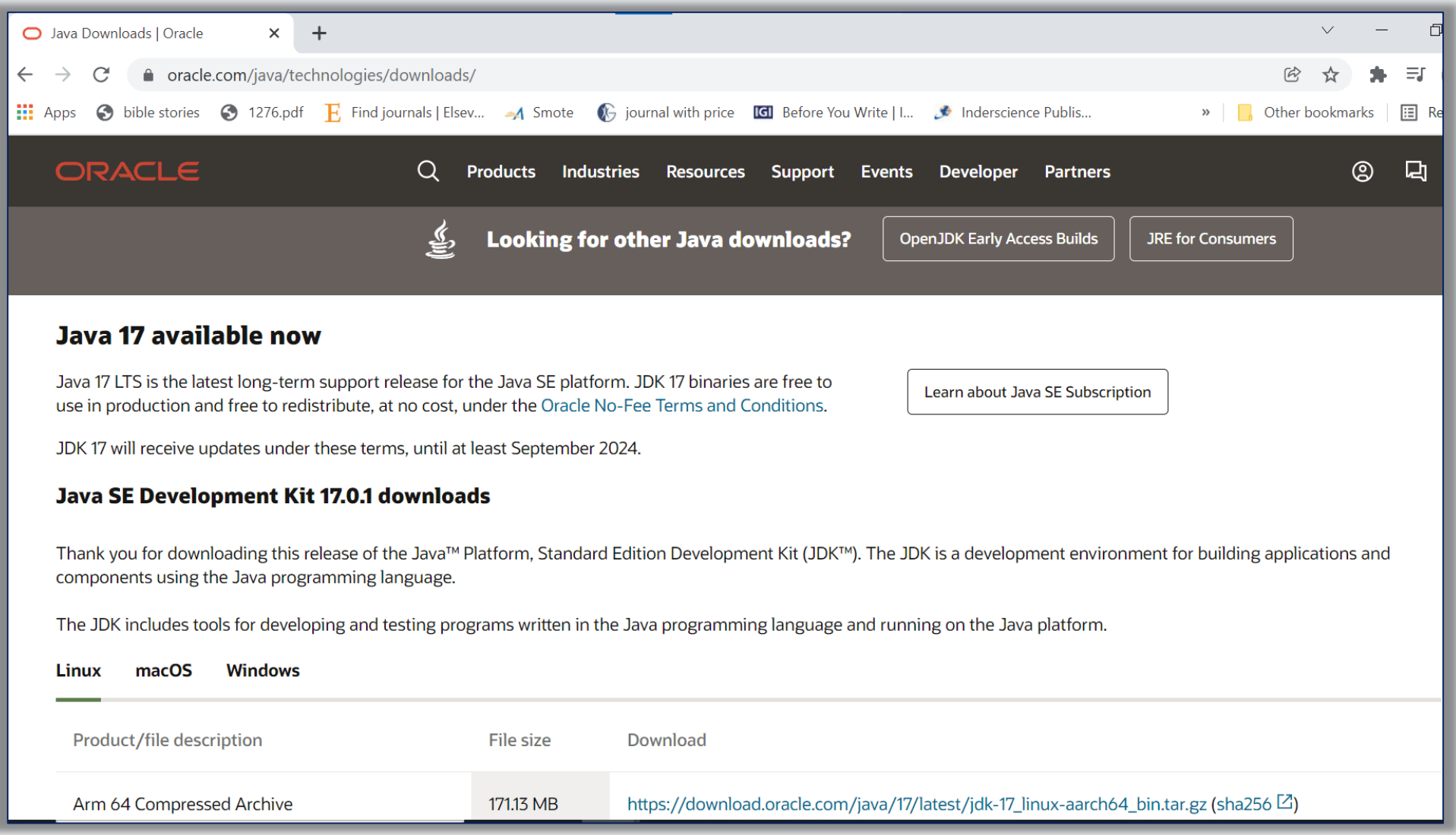
- learn how to install java
- learn how to install hadoop

Prerequisites to install Hadoop

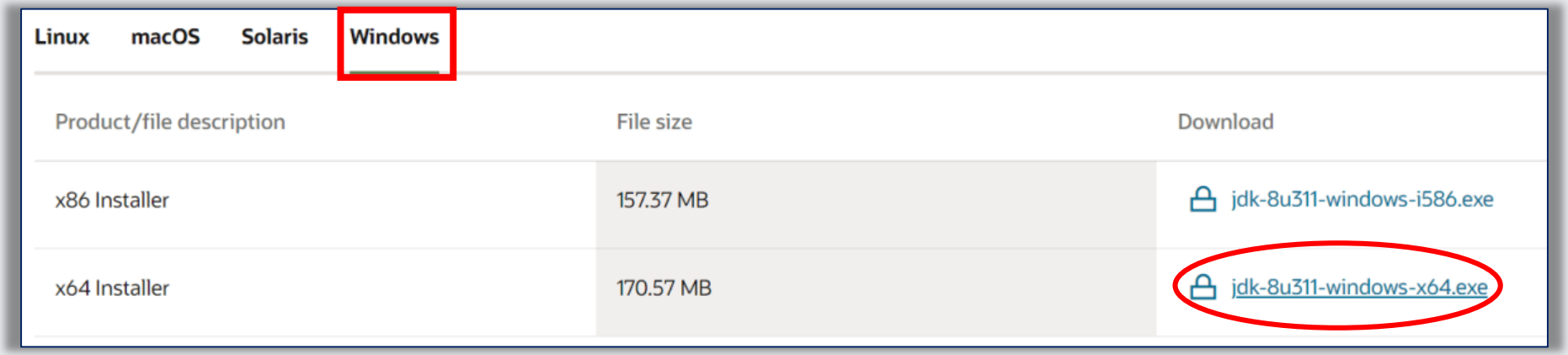
Java™

Prerequisites to install Hadoop

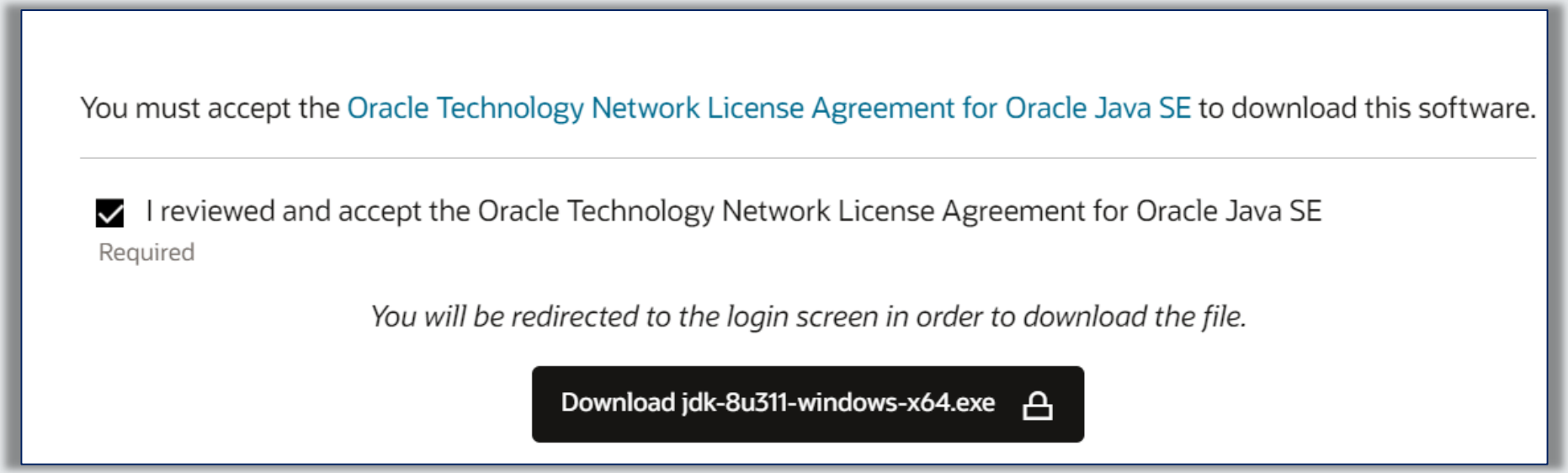
Steps to install Java



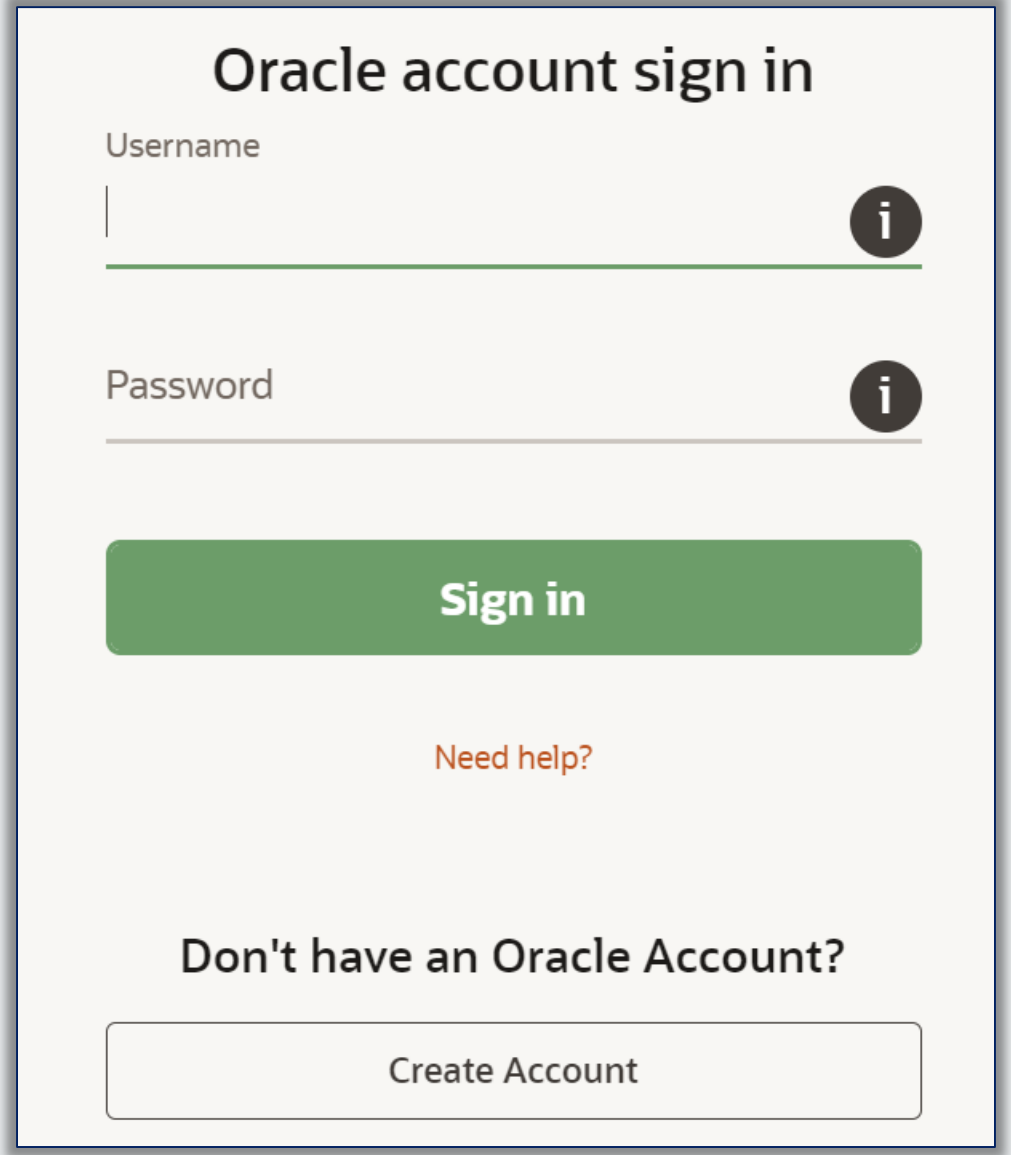
Steps to Install Java



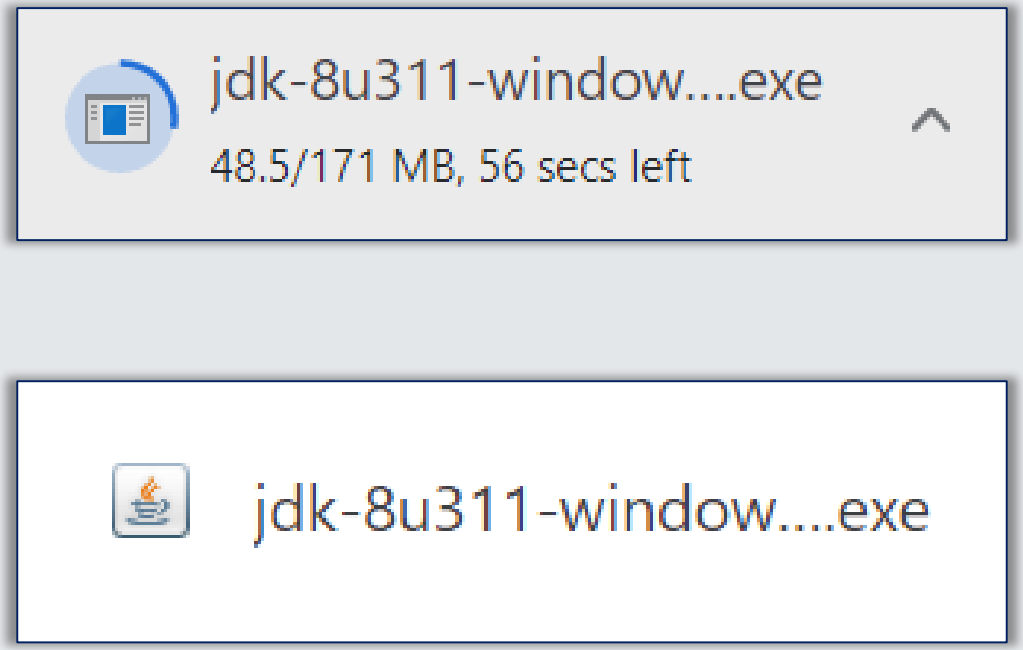
Steps to Install Java



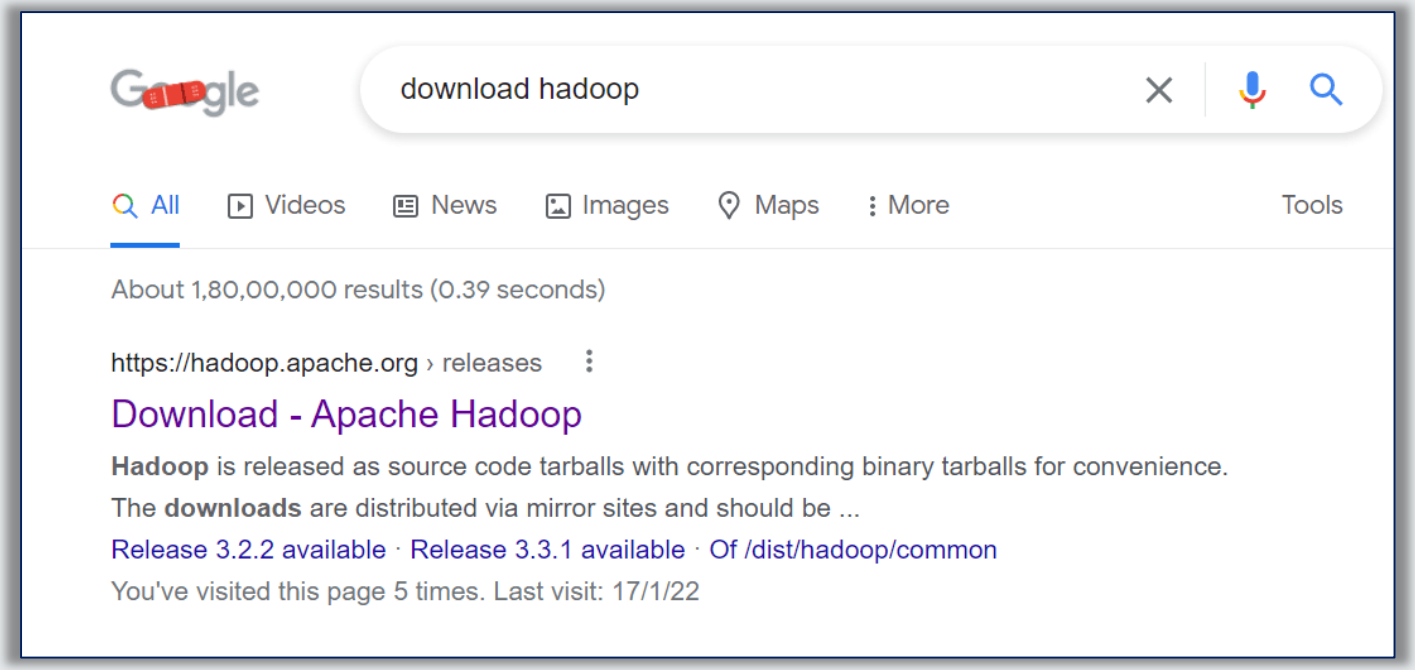
Steps to Install Java



Steps to Install Java

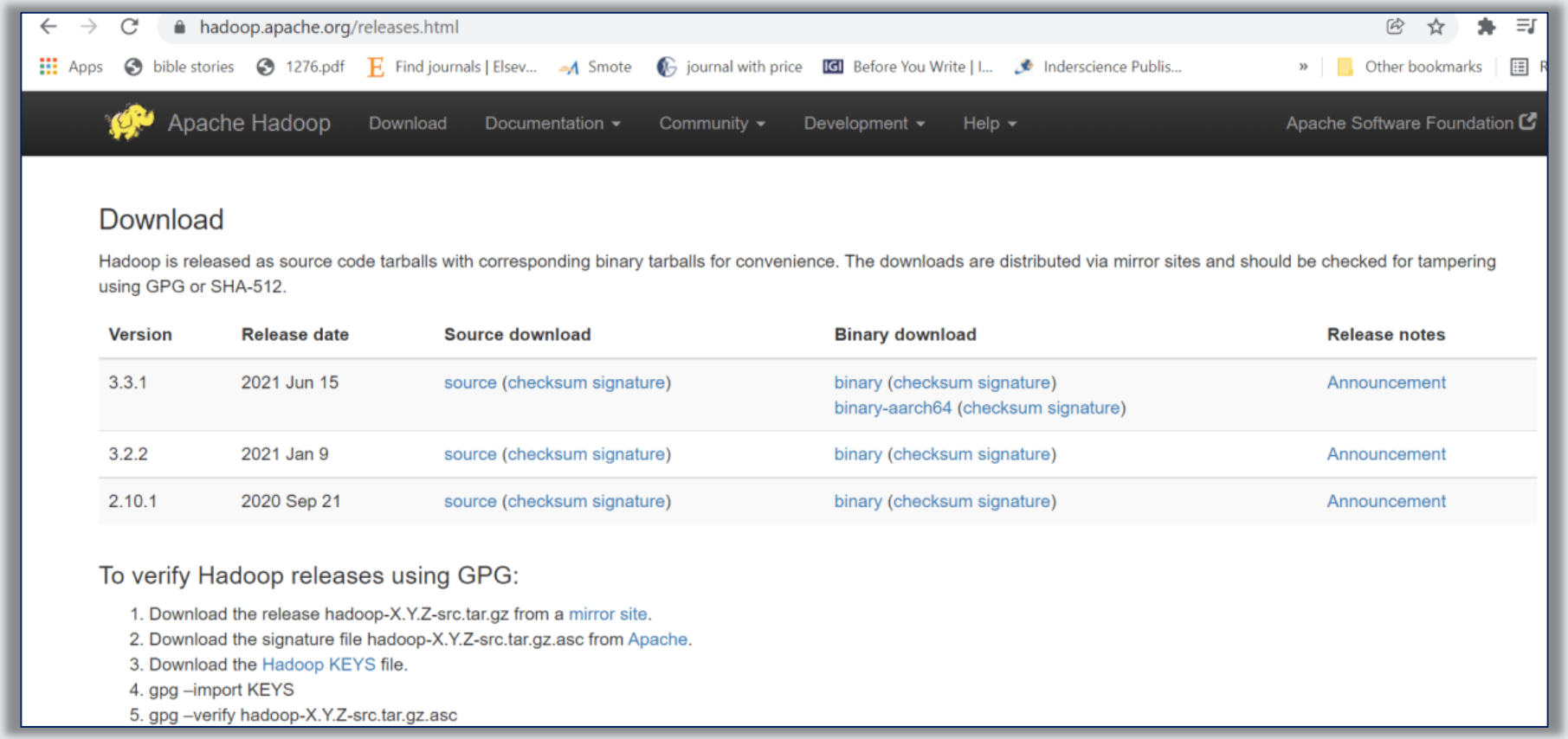


Steps to install hadoop



https://hadoop.apache.org ›

Steps to install hadoop



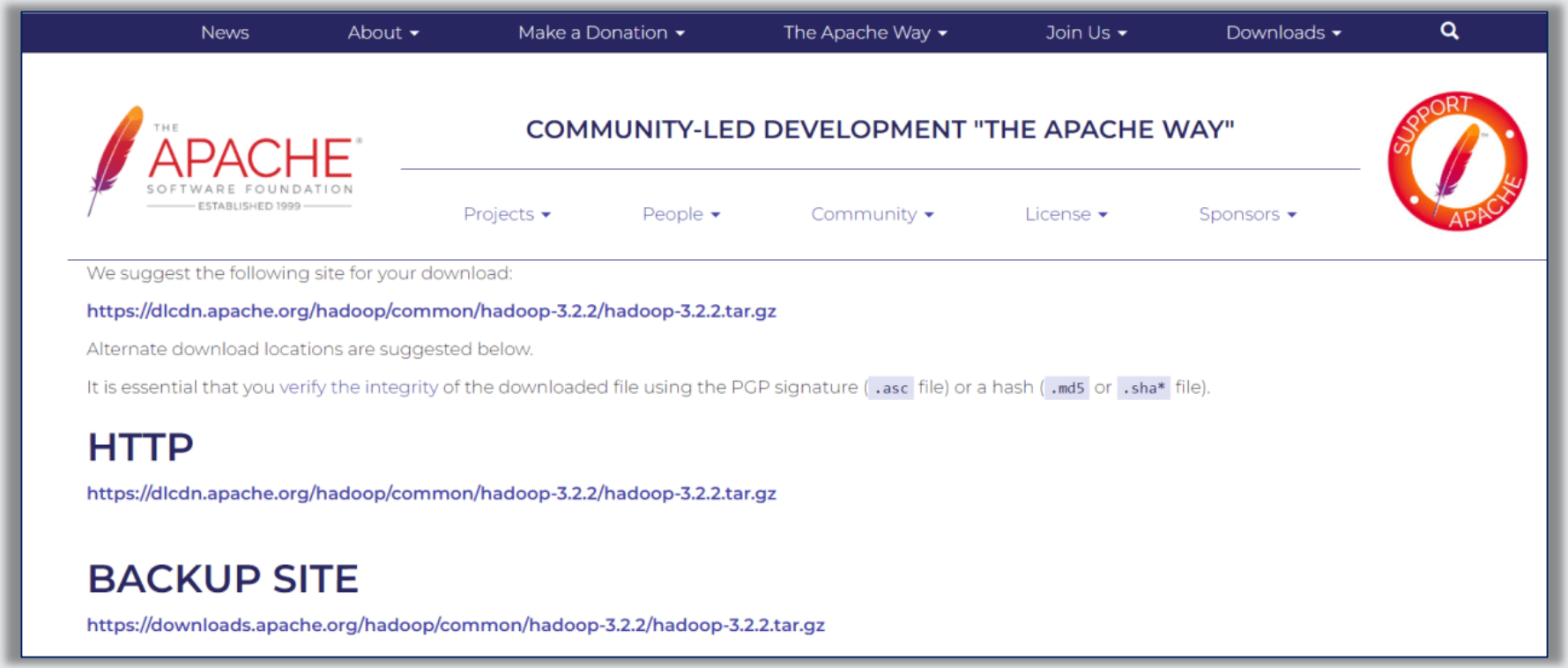
Steps to install hadoop

Download

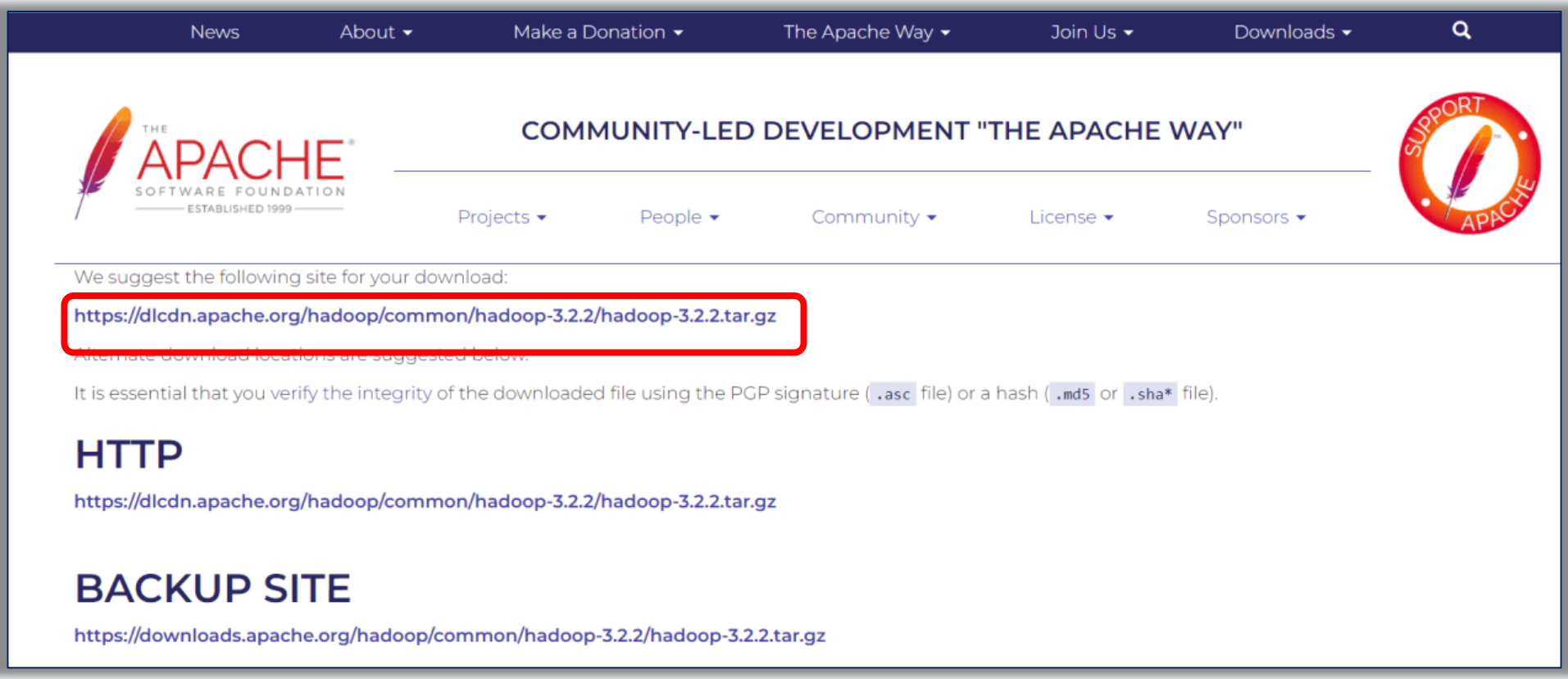
Hadoop is released as source code tarballs with corresponding binary tarballs for convenience. The downloads are distributed via mirror sites and should be checked for tampering using GPG or SHA-512.

Version	Release date	Source download	Binary download	Release notes
3.3.1	2021 Jun 15	source (checksum signature)	binary (checksum signature) binary-aarch64 (checksum signature)	Announcement
3.2.2	2021 Jan 9	source (checksum signature)	binary (checksum signature)	Announcement
2.10.1	2020 Sep 21	source (checksum signature)	binary (checksum signature)	Announcement

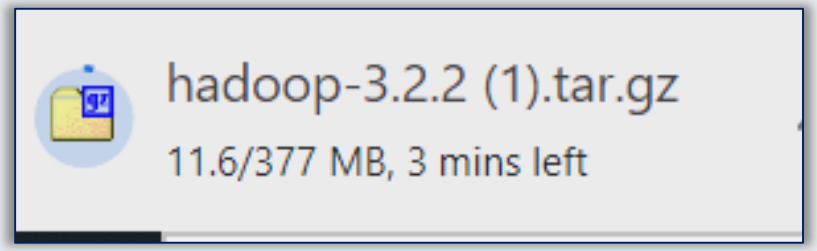
Steps to install hadoop



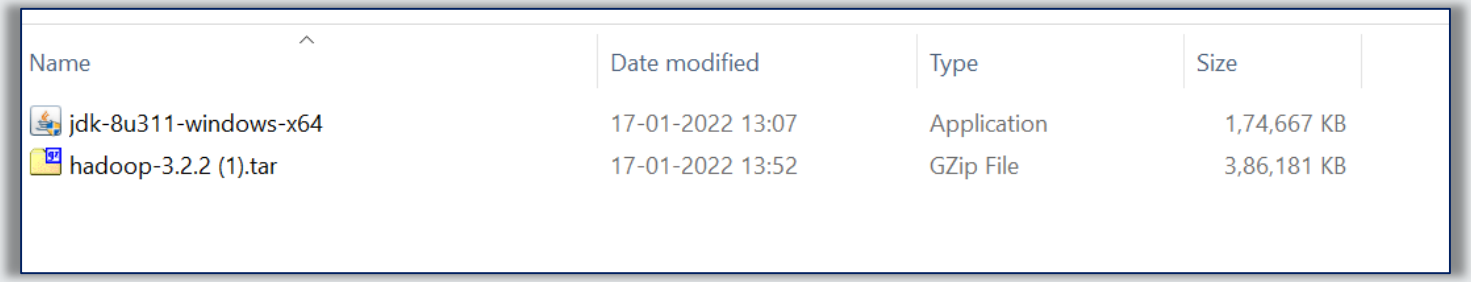
Steps to install hadoop



Steps to install hadoop



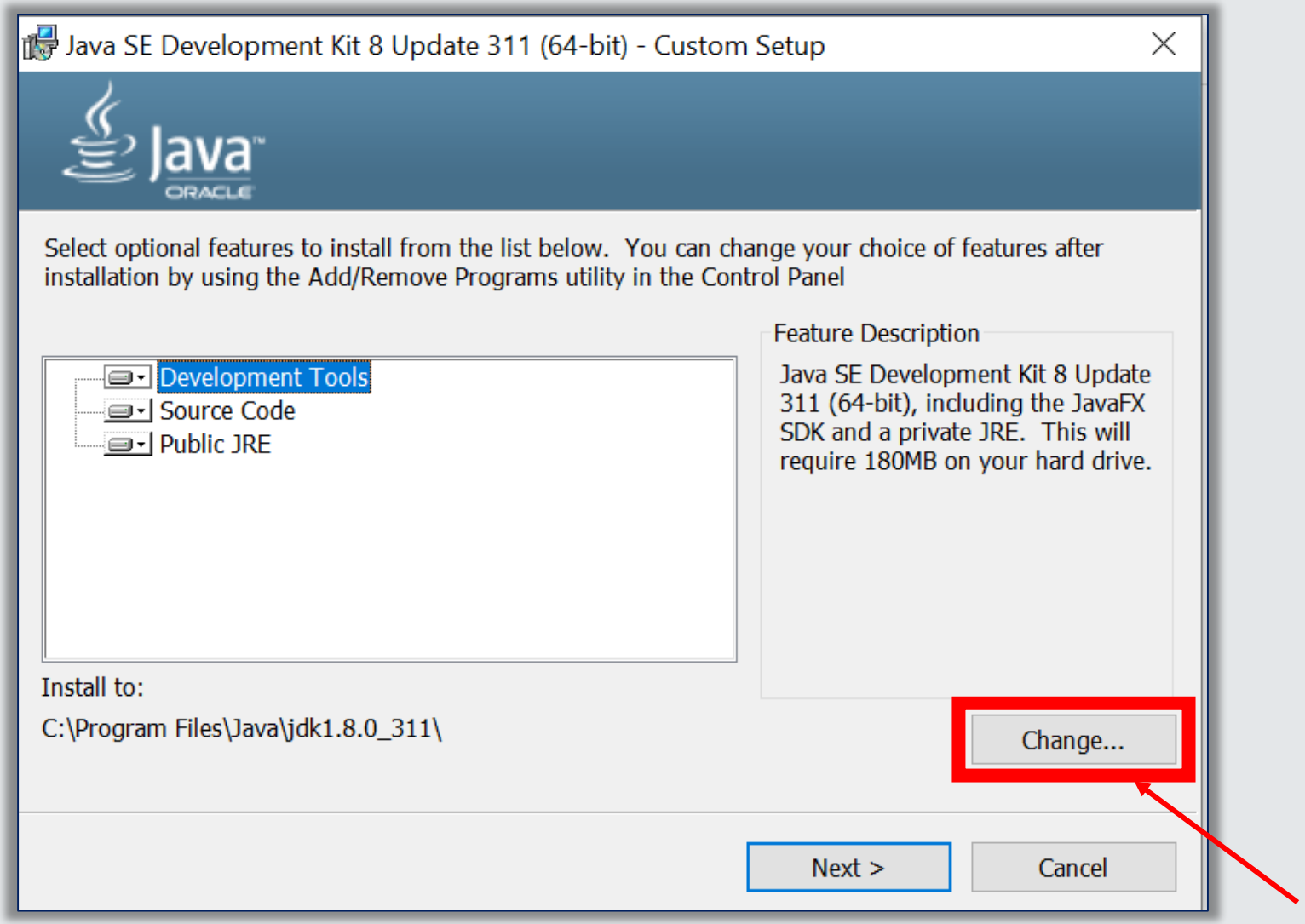
Steps to install hadoop



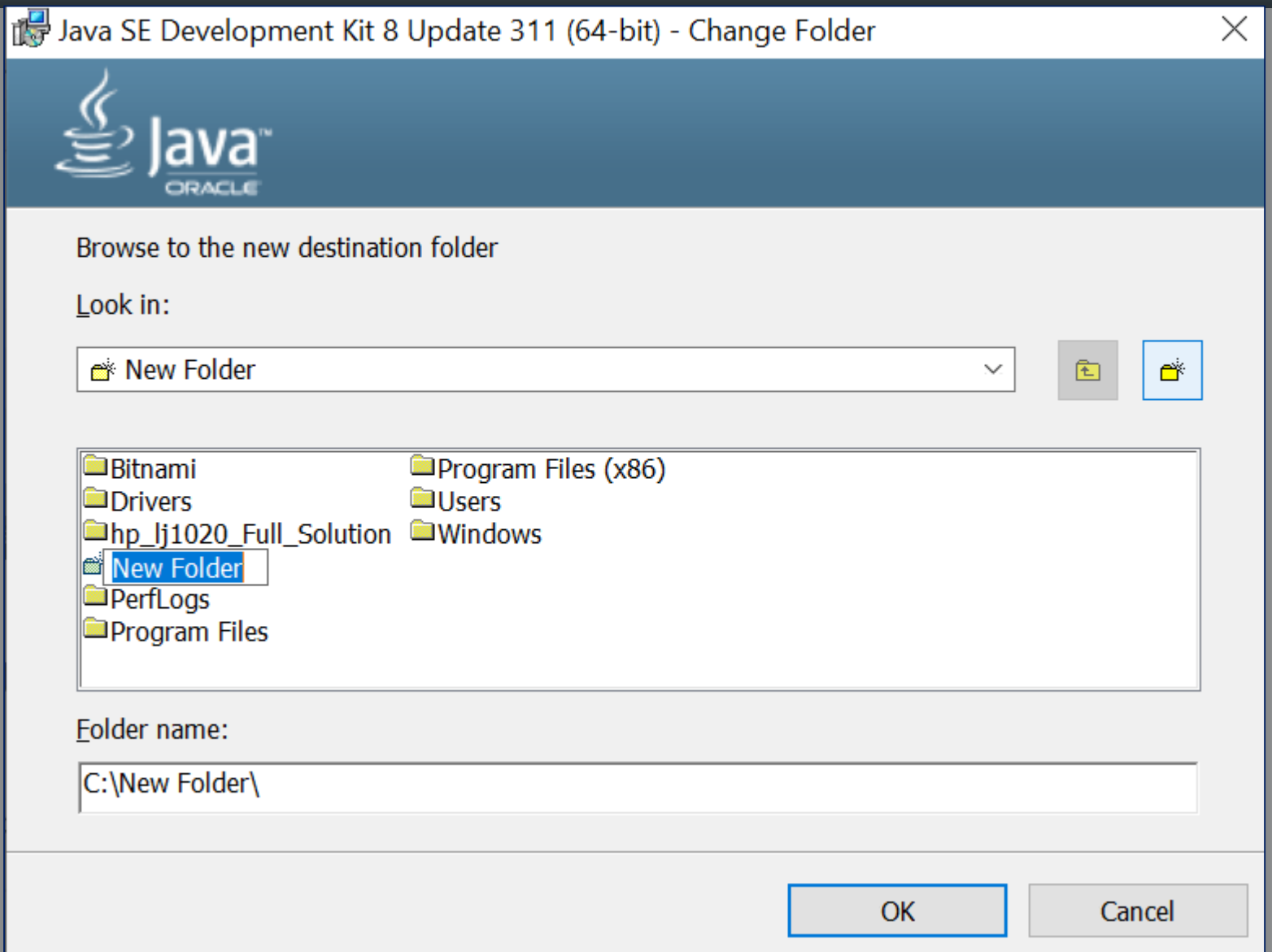
Name	Date modified	Type	Size
jdk-8u311-windows-x64	17-01-2022 13:07	Application	1,74,667 KB
hadoop-3.2.2 (1).tar	17-01-2022 13:52	GZip File	3,86,181 KB

Installation of Java

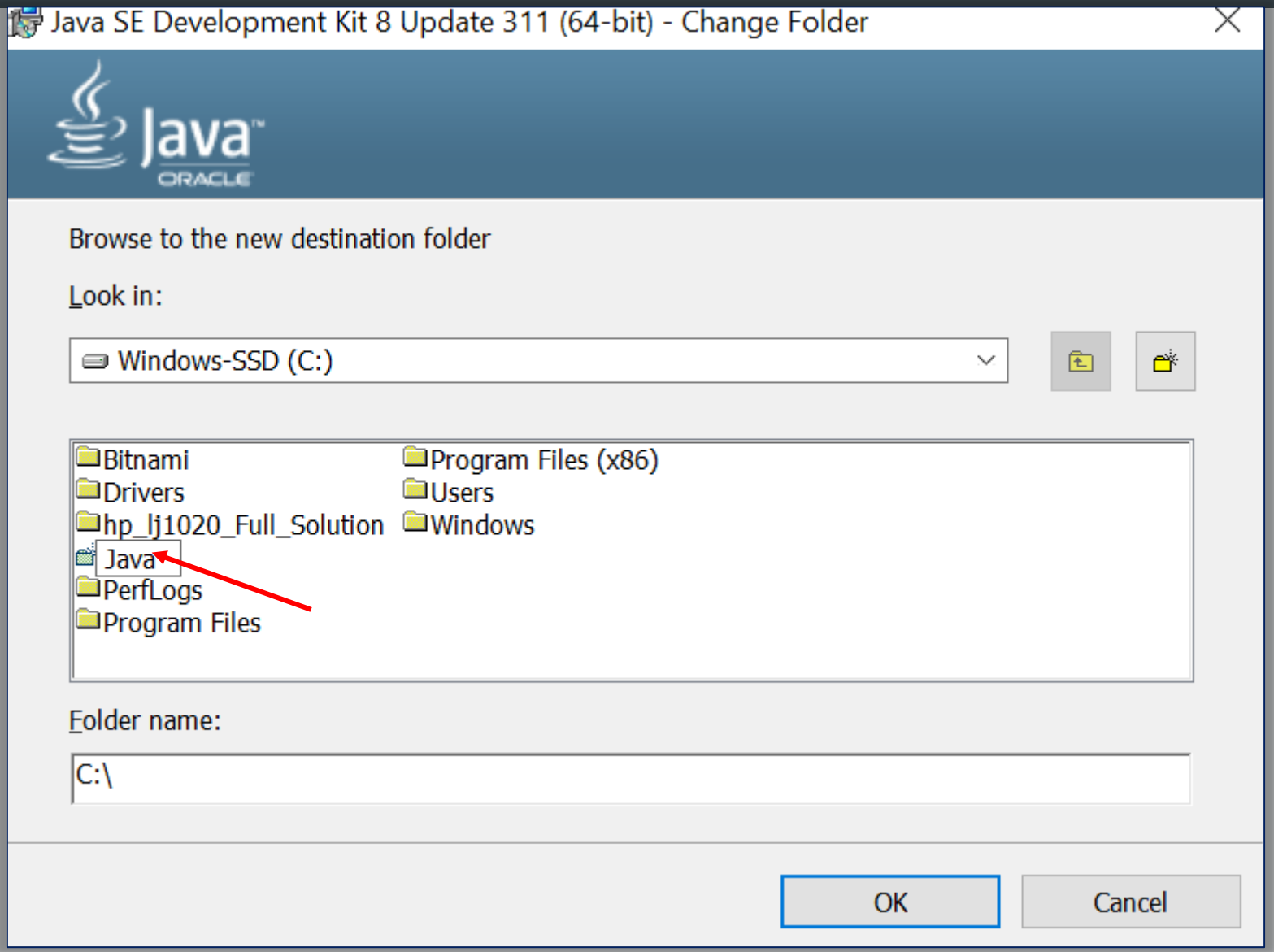
Installation of JAVA



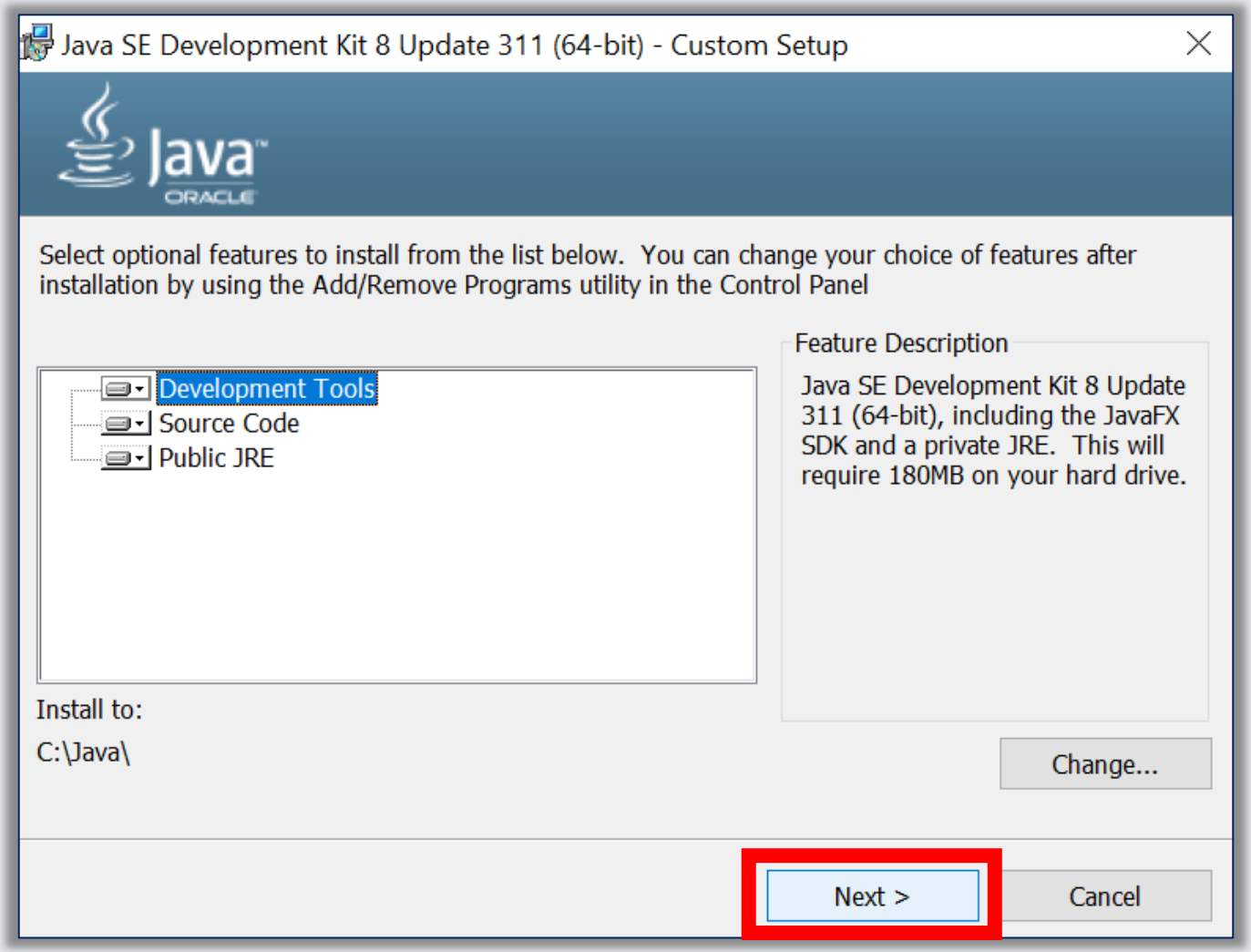
Installation of JAVA



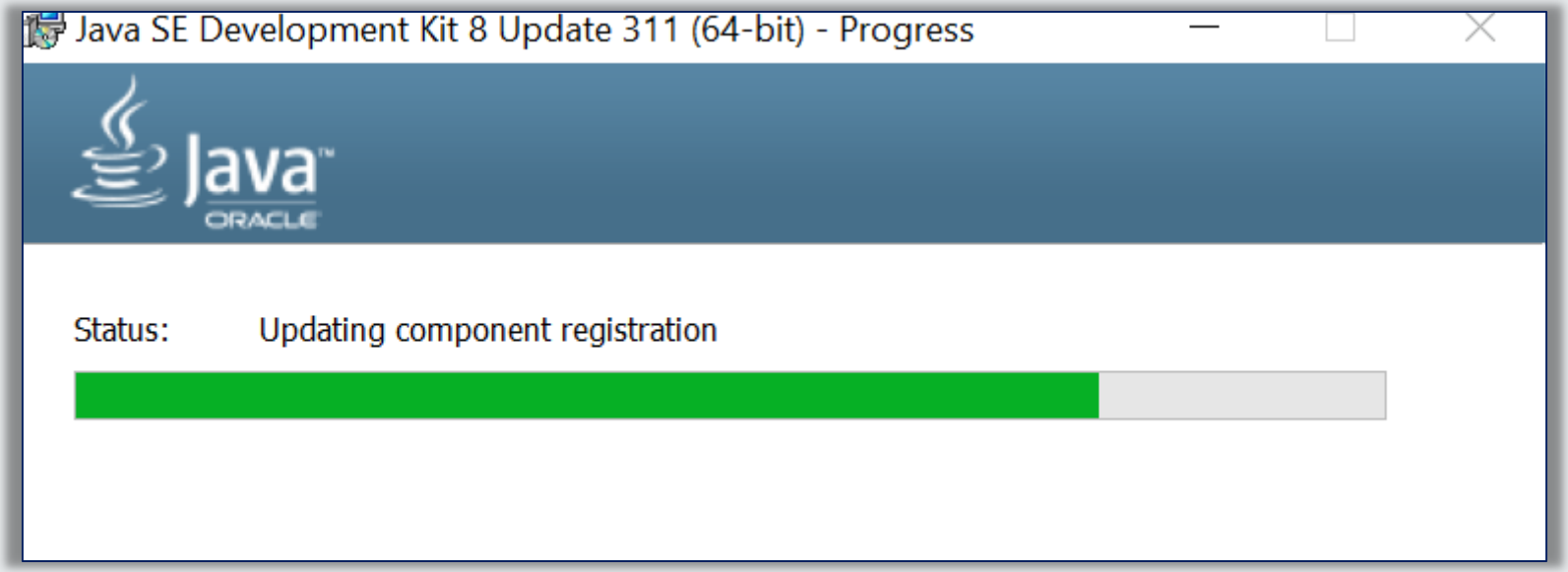
Installation of JAVA



Installation of JAVA

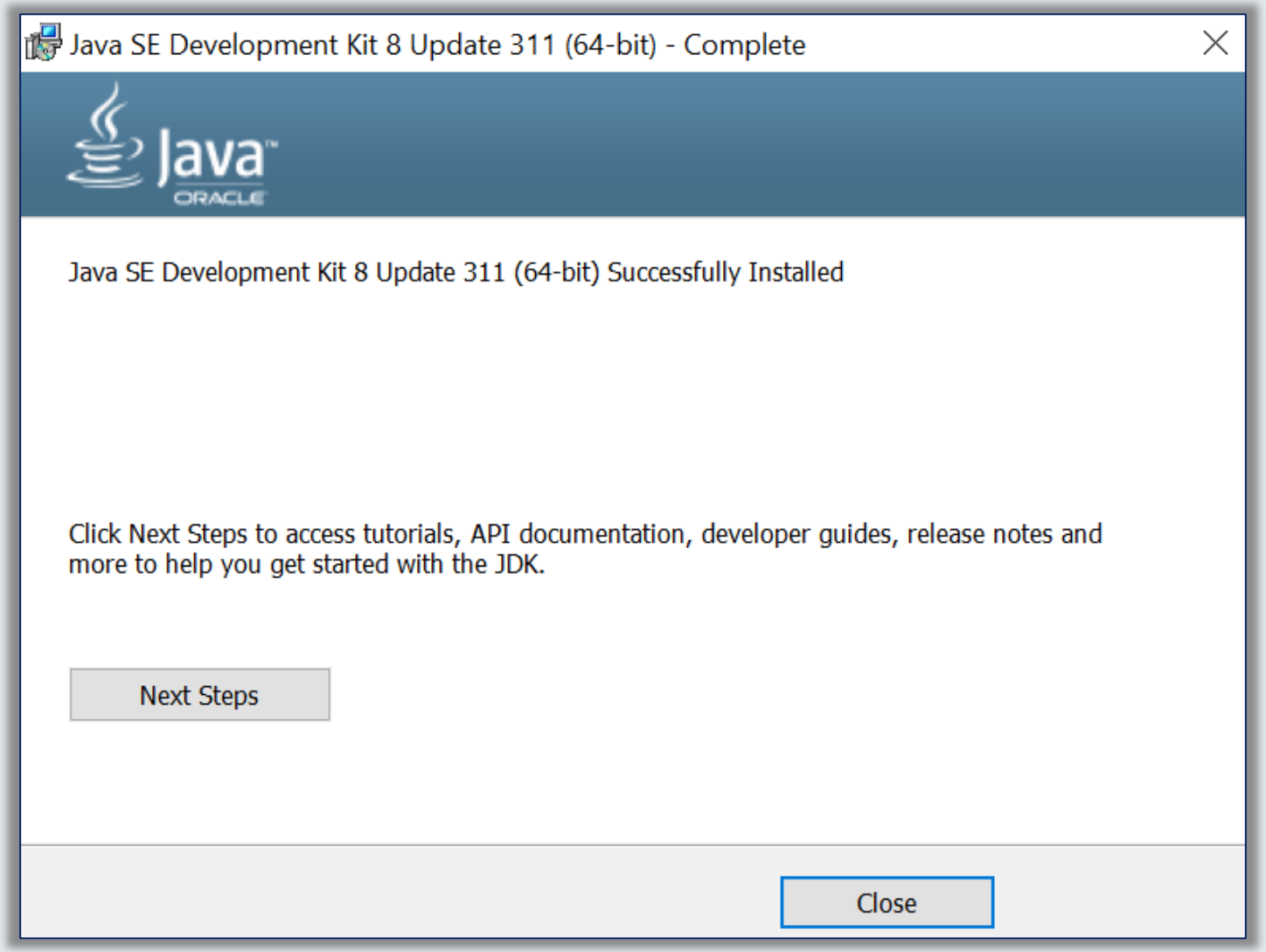


Installation of JAVA

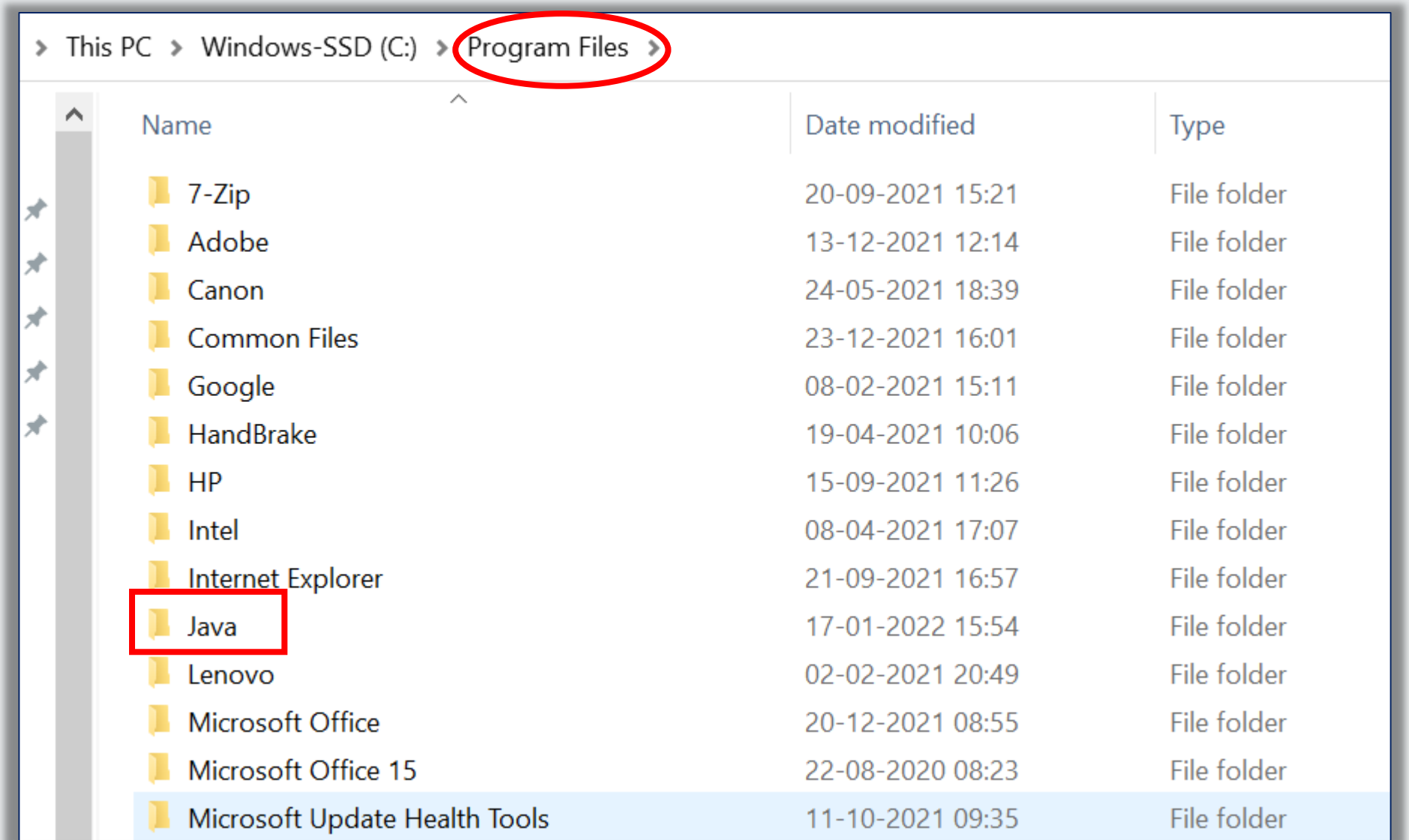


Installation of JAVA

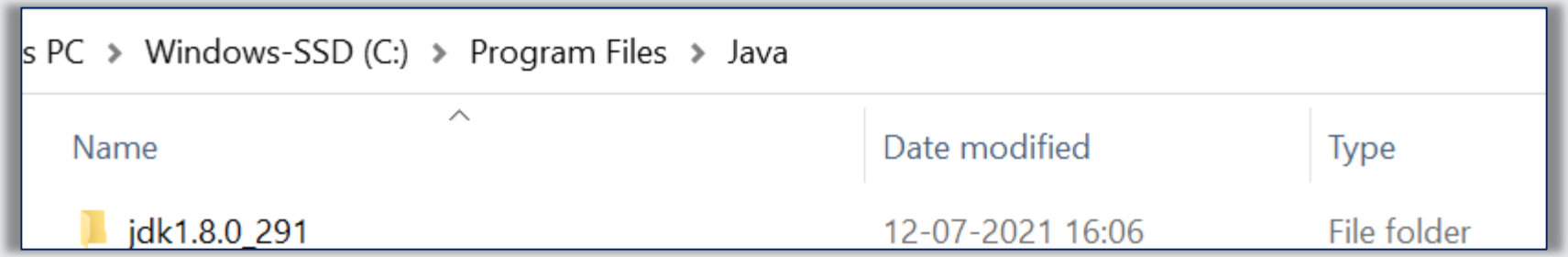
Installation of JAVA

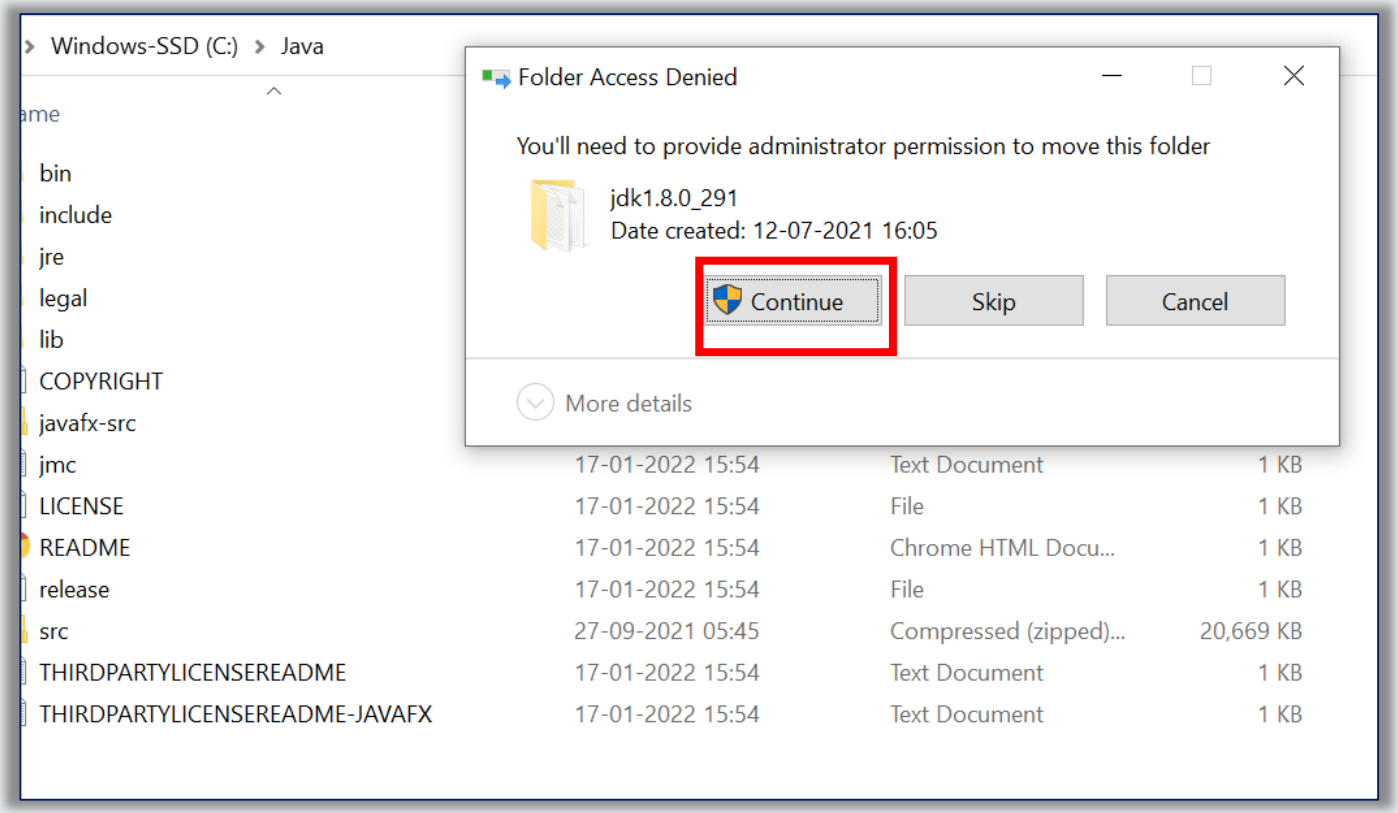


Installation of JAVA

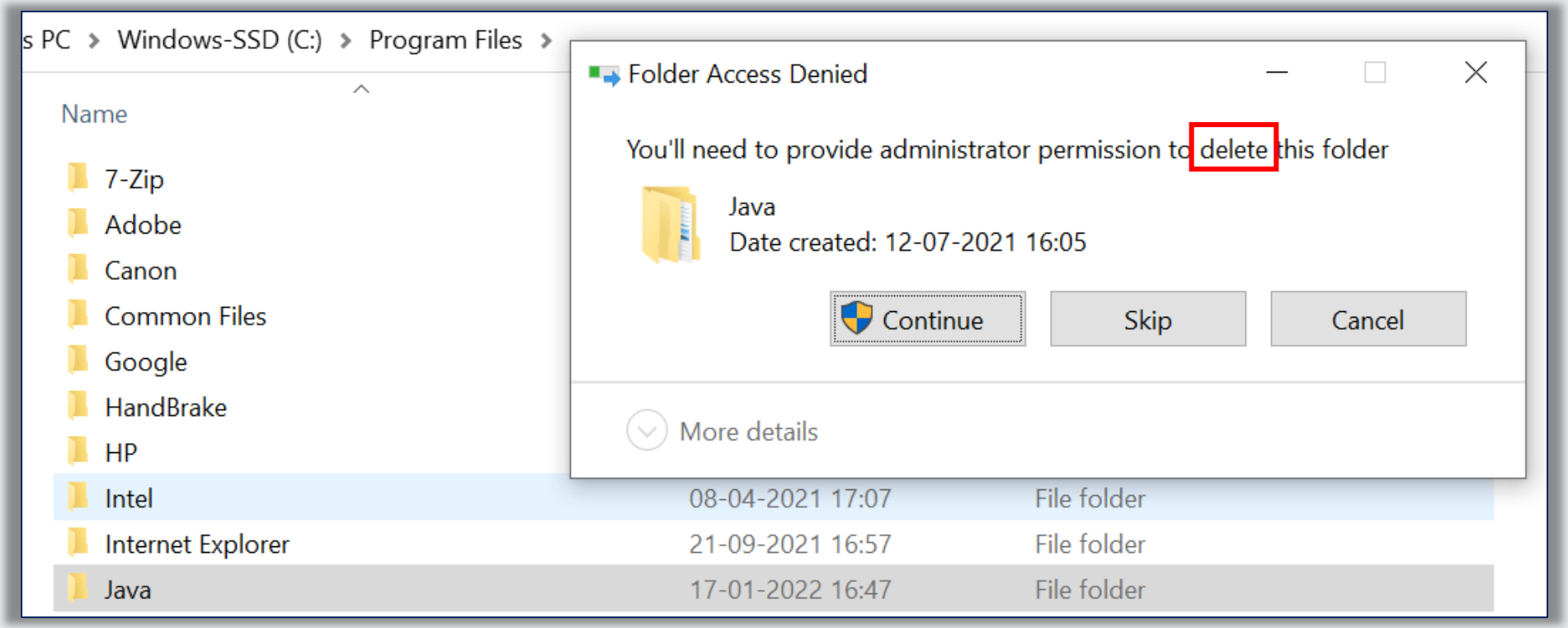


Installation of JAVA





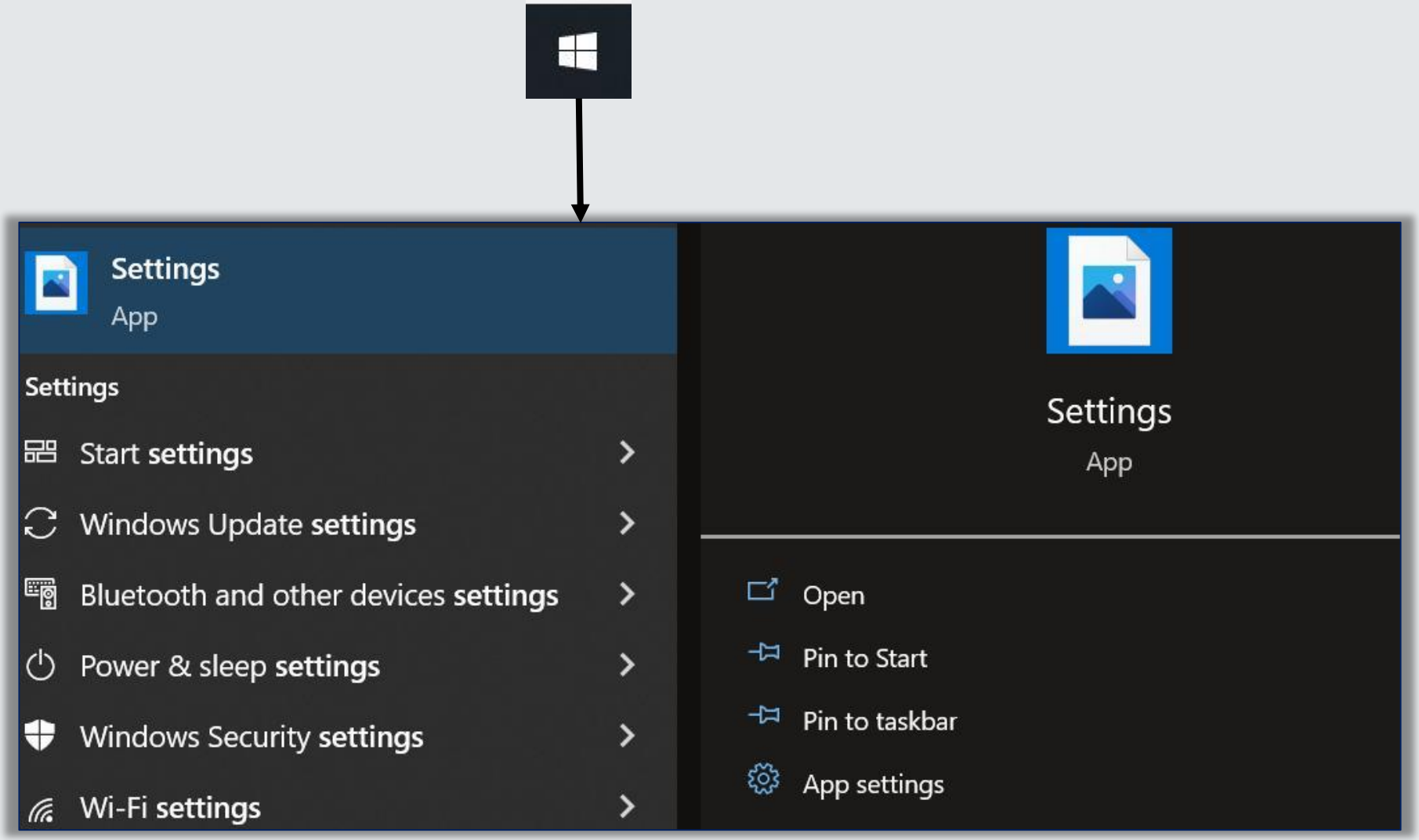
Installation of JAVA



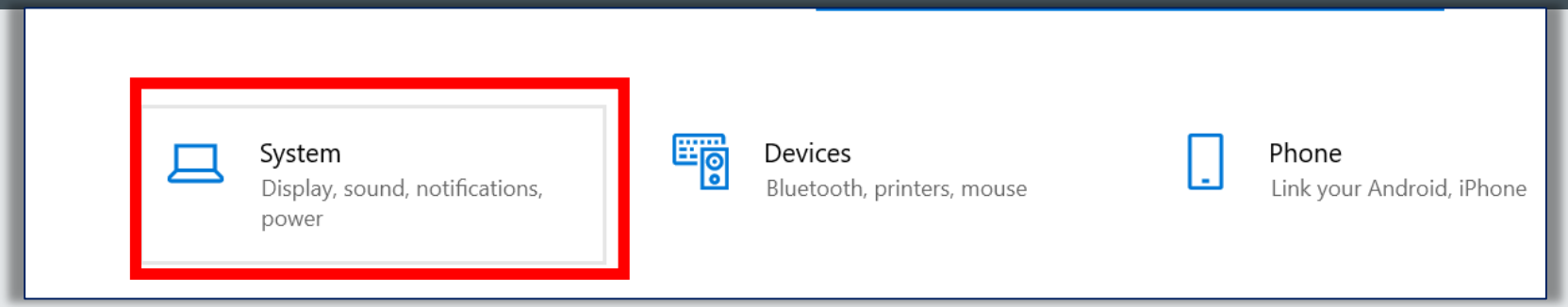
How to set environment variable for JAVA

Set the Environment Variable for JAVA

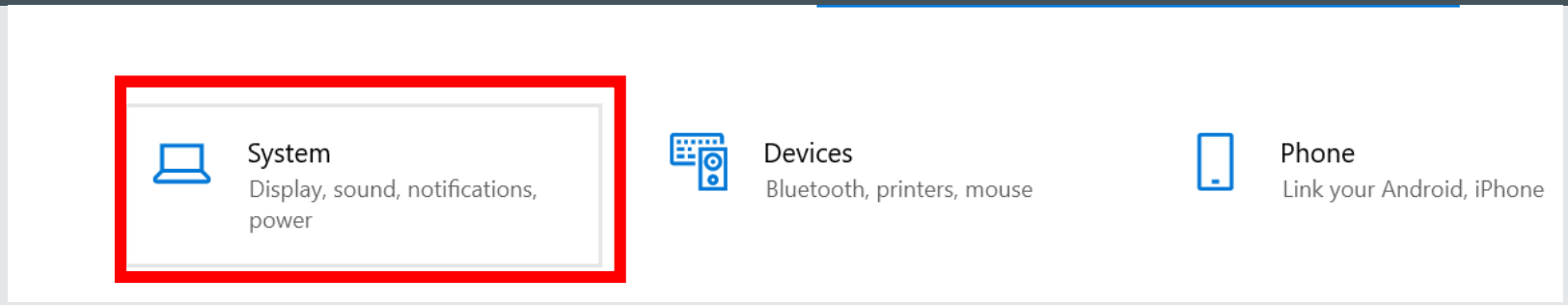
Set the Environment Variable for JAVA

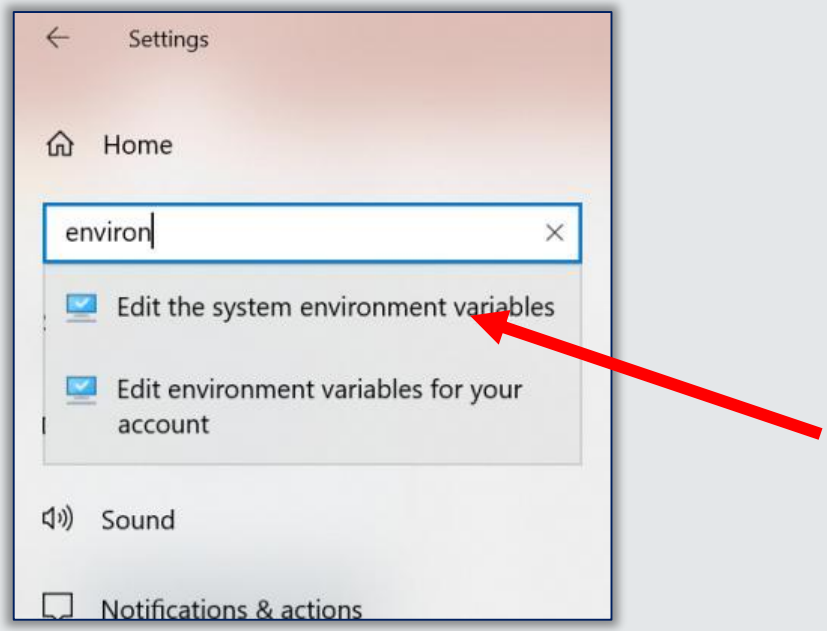


Set the Environment Variable for JAVA

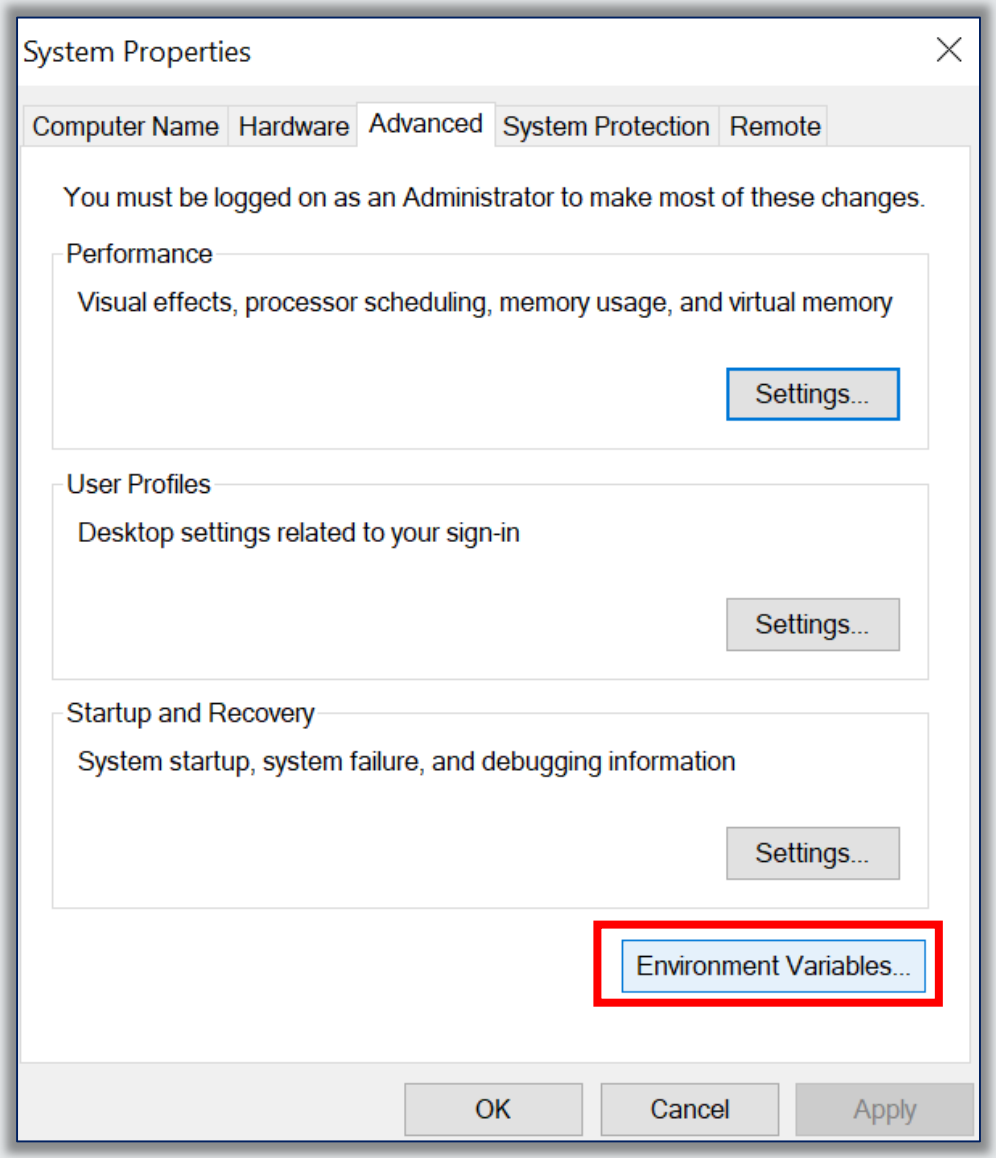


Set the Environment Variable for JAVA

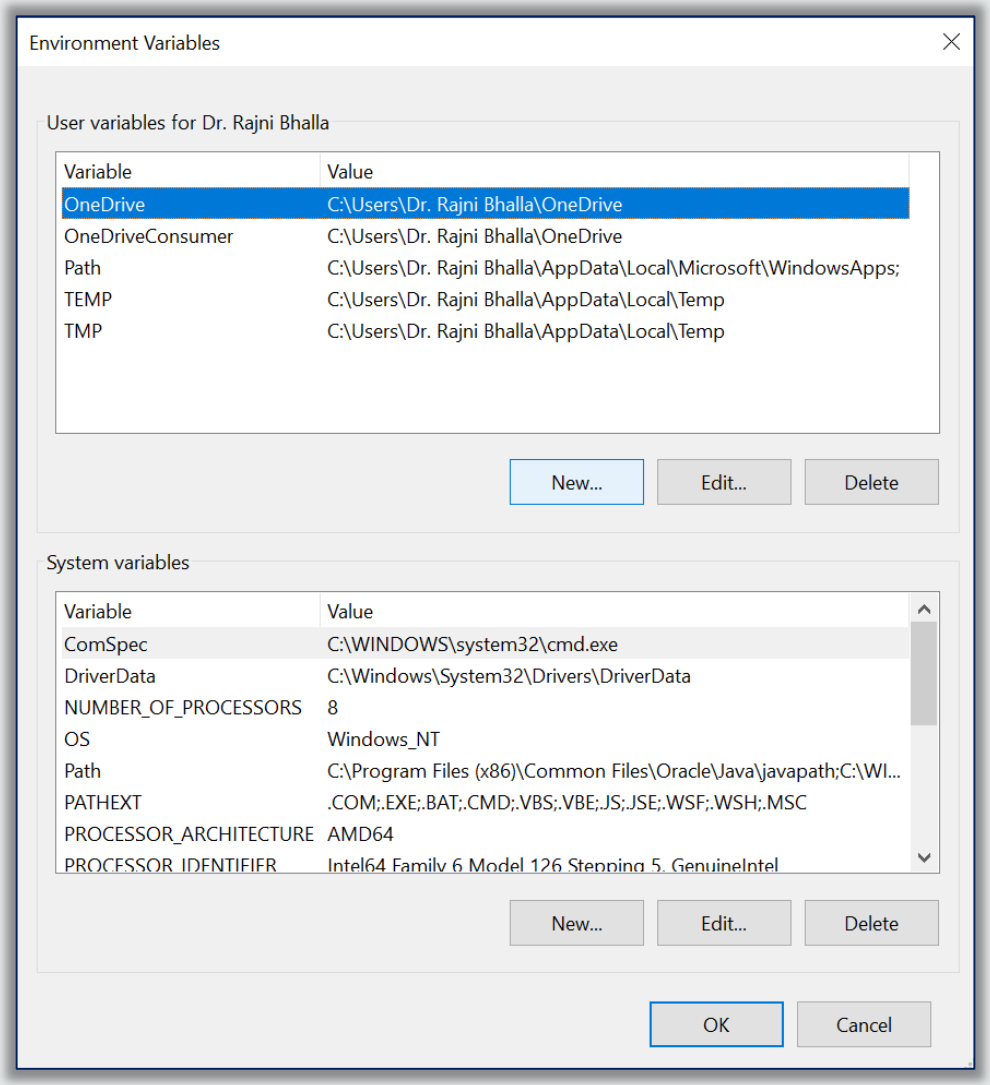




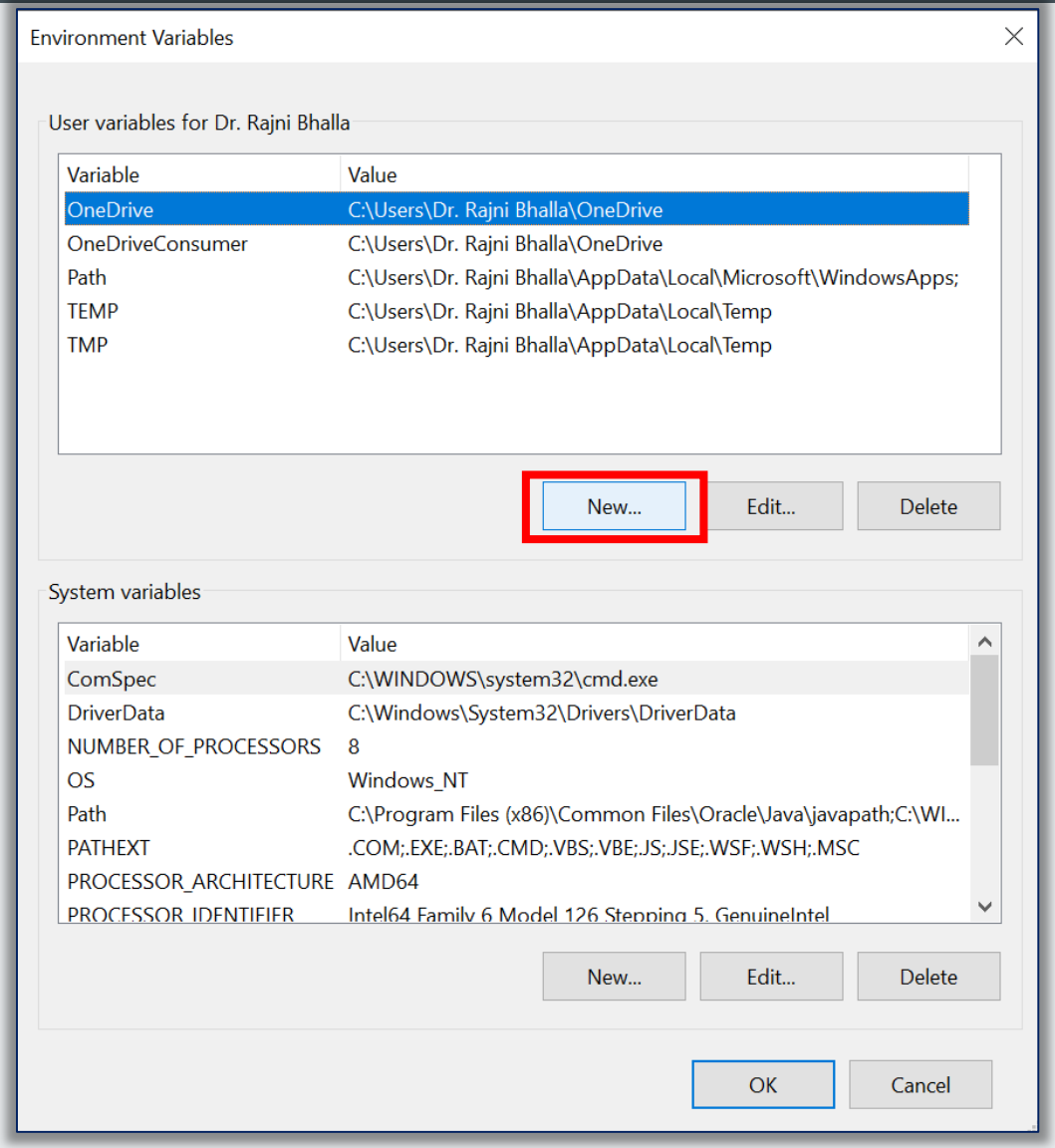
Set the Environment Variable for JAVA



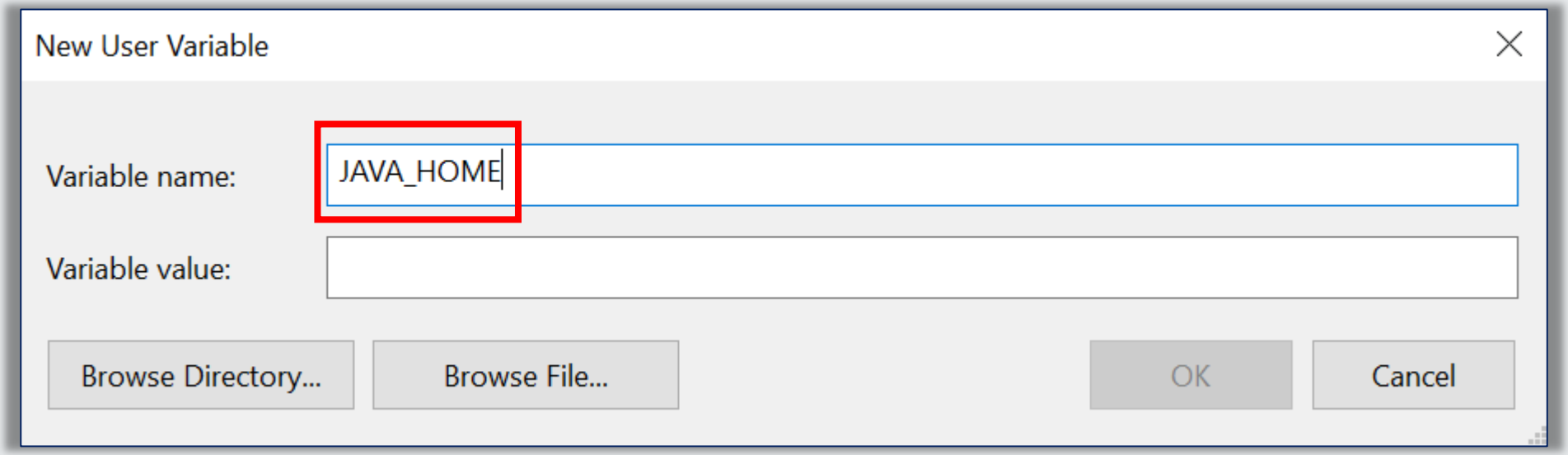
Set the Environment Variable for JAVA



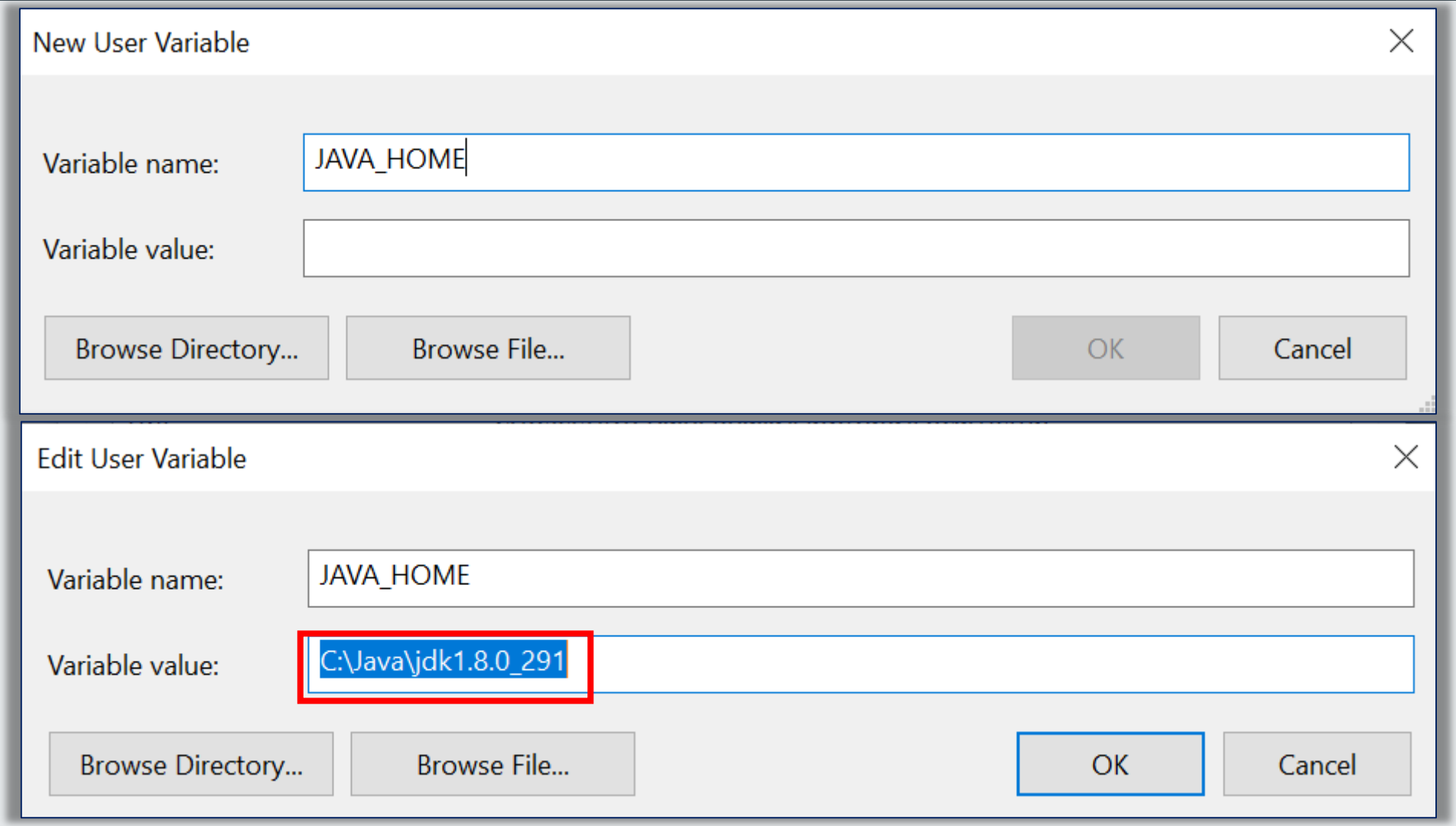
Set the Environment Variable for JAVA



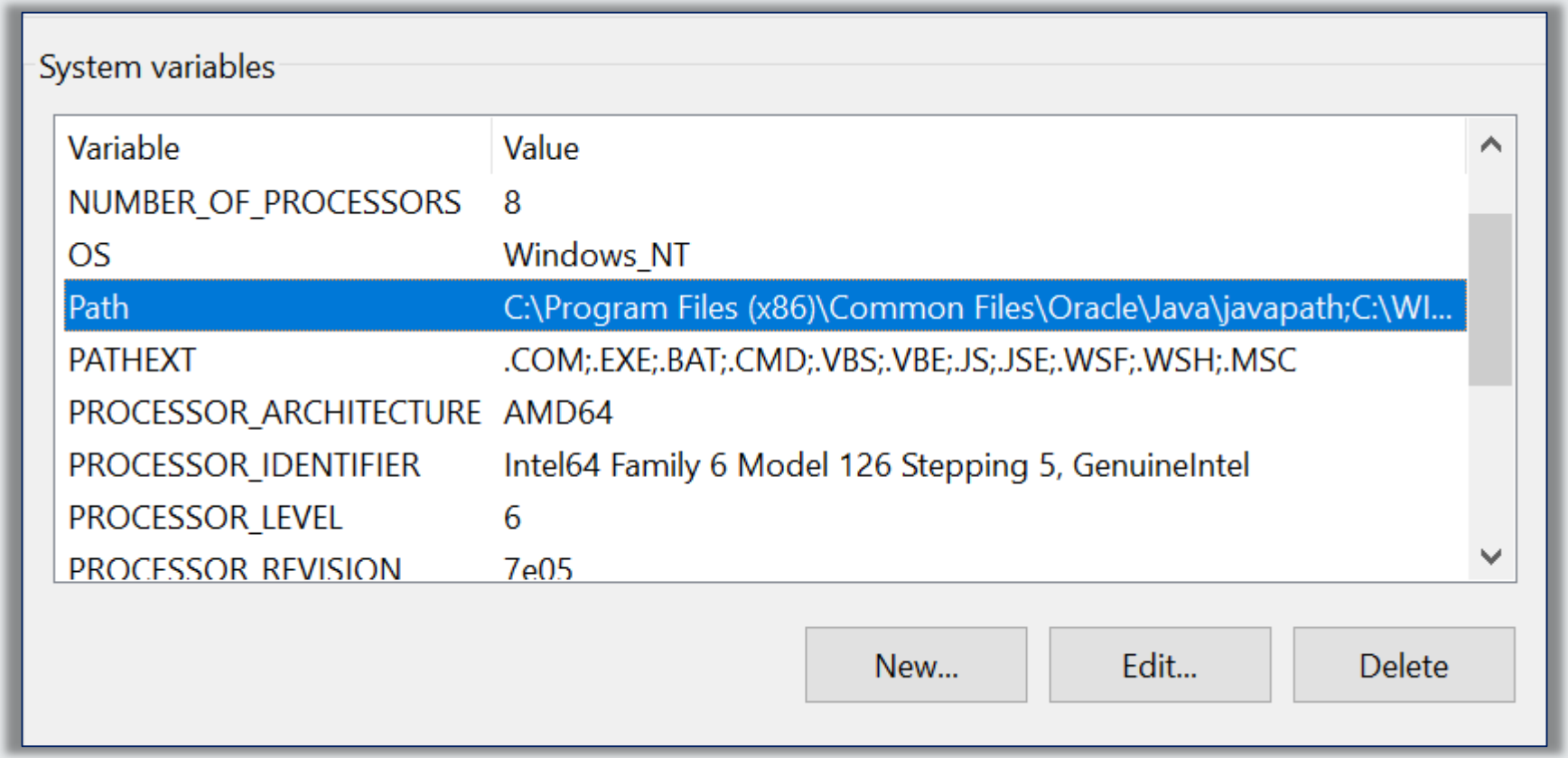
Set the Environment Variable for JAVA



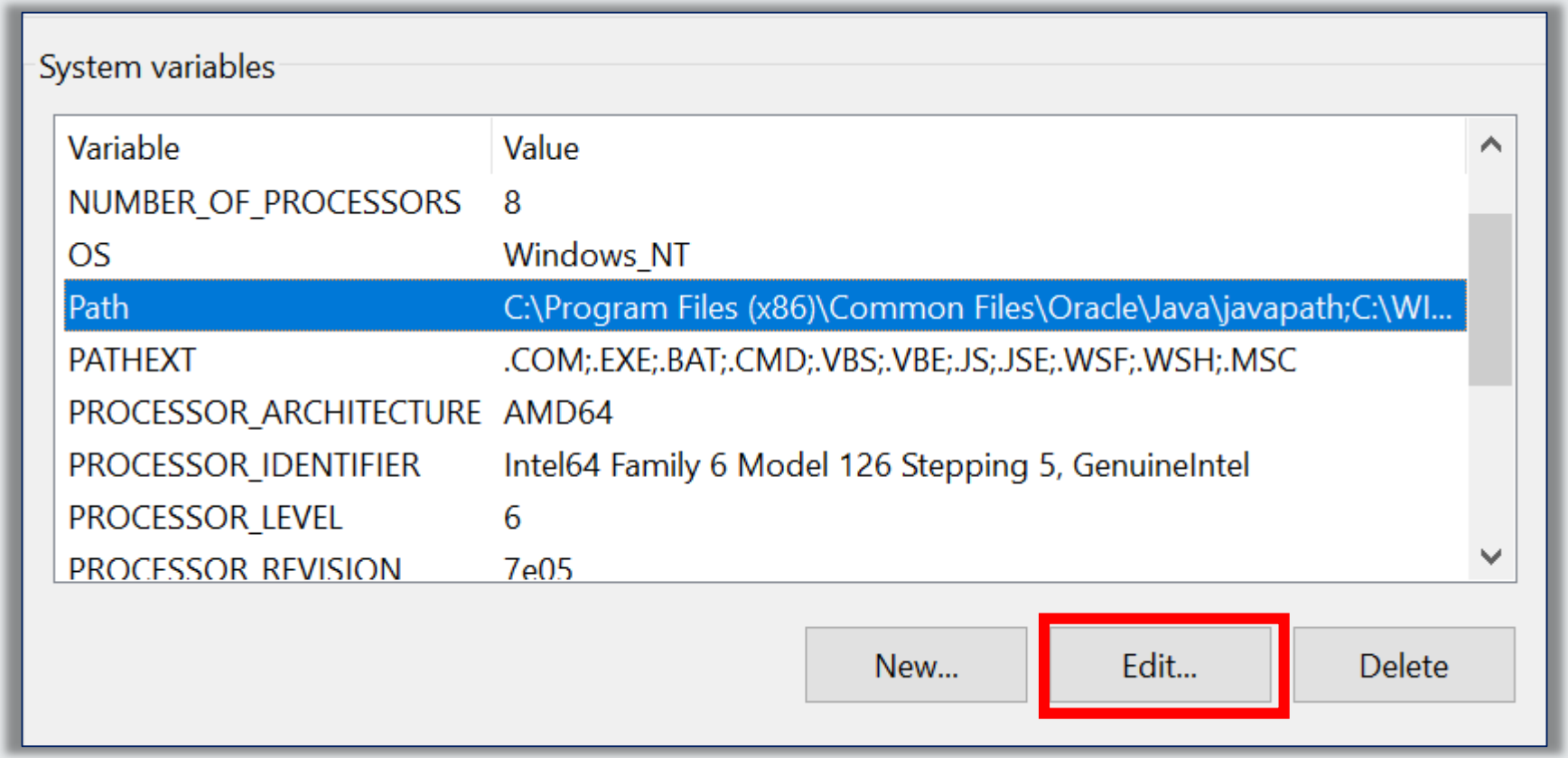
Set the Environment Variable for JAVA



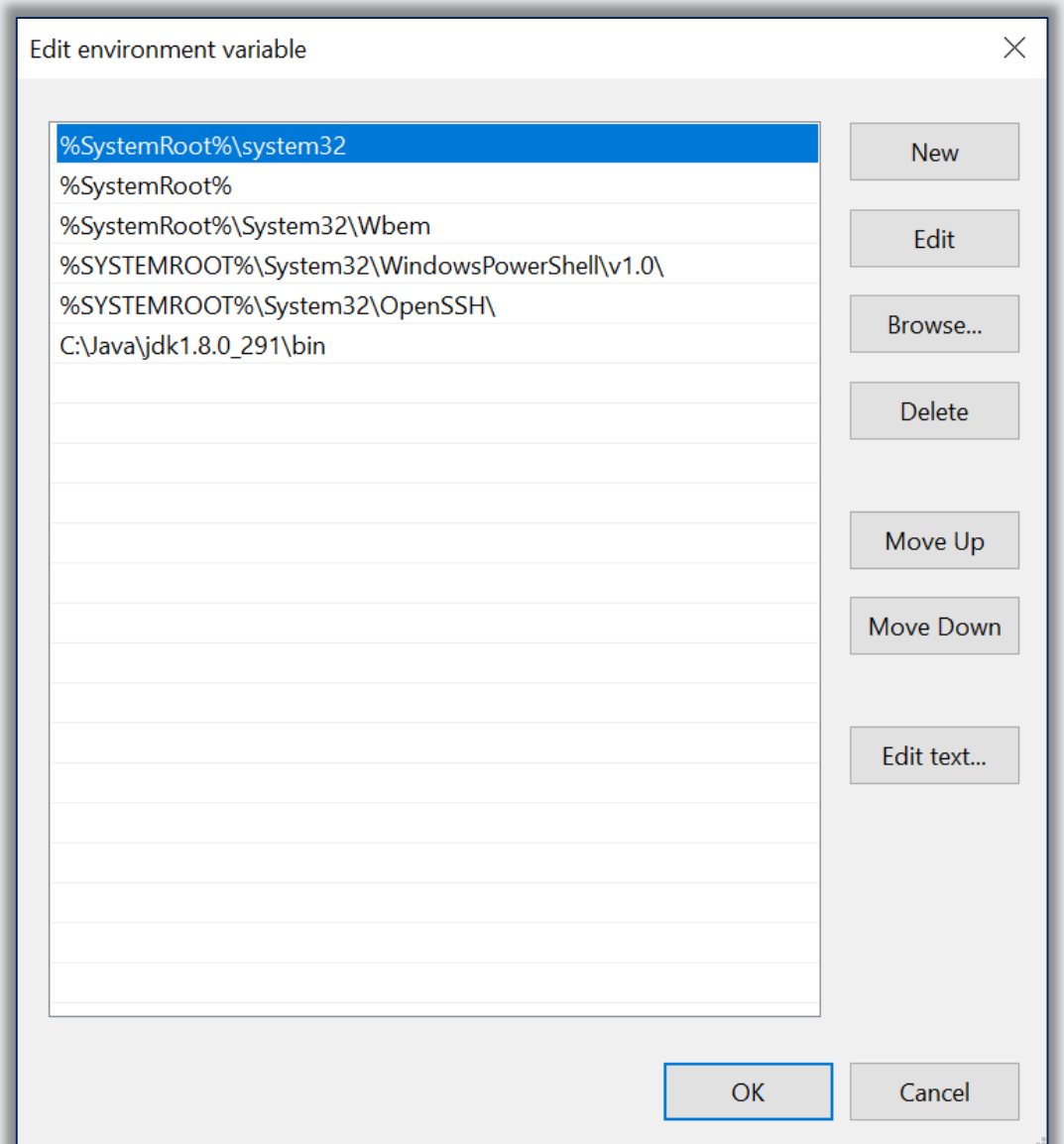
Set the Environment Variable for JAVA



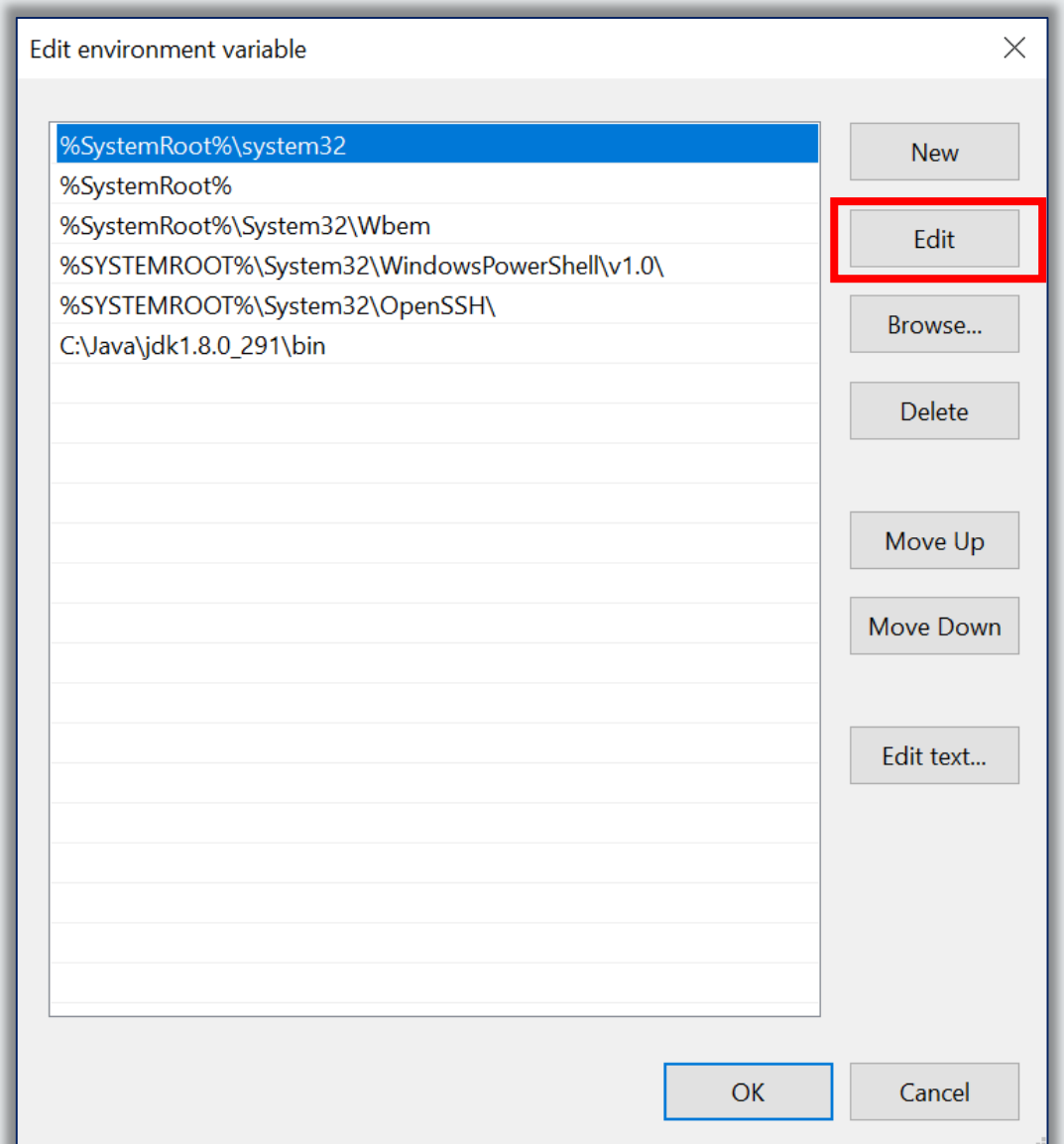
Set the Environment Variable for JAVA



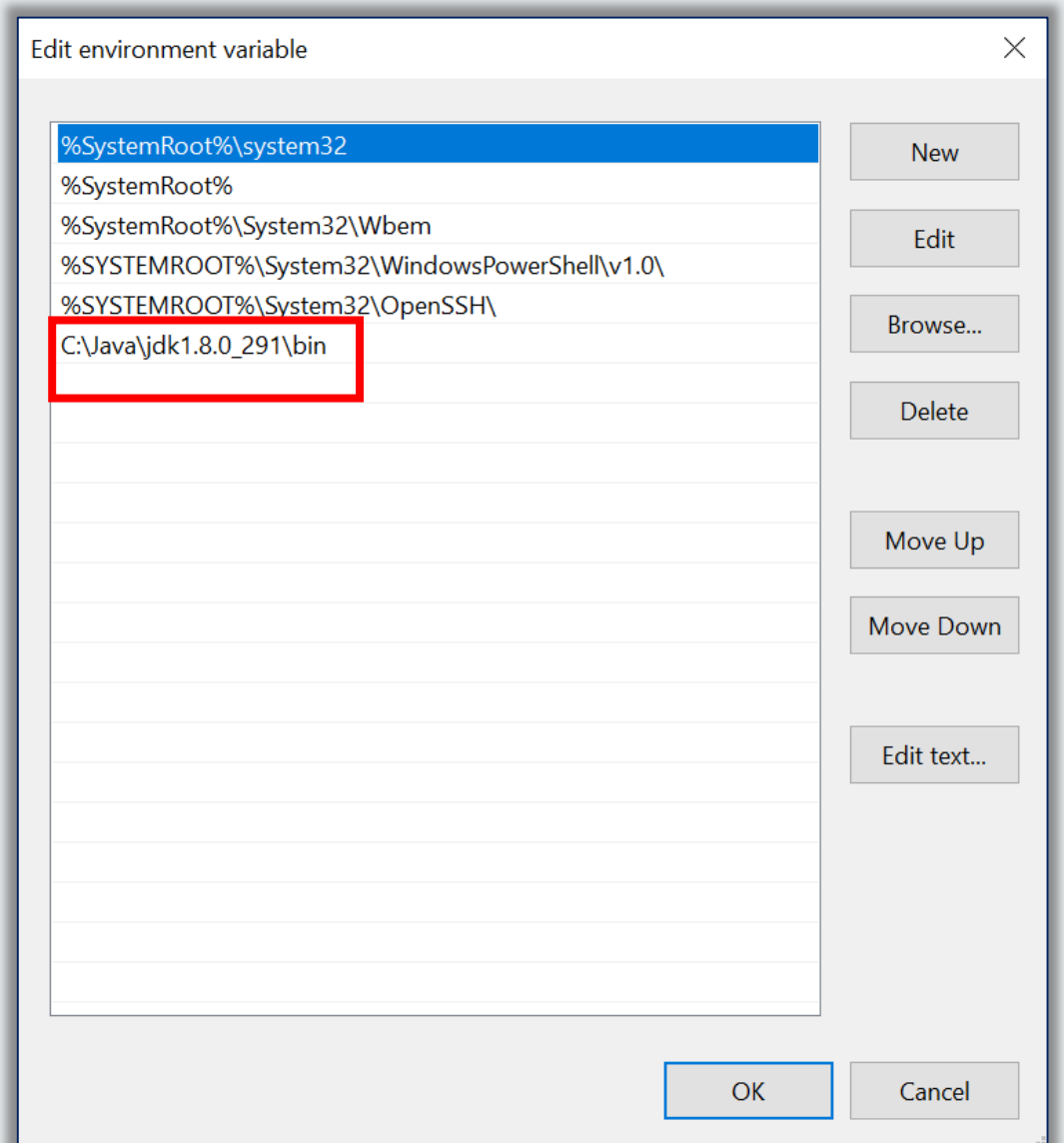
Set the Environment Variable for JAVA



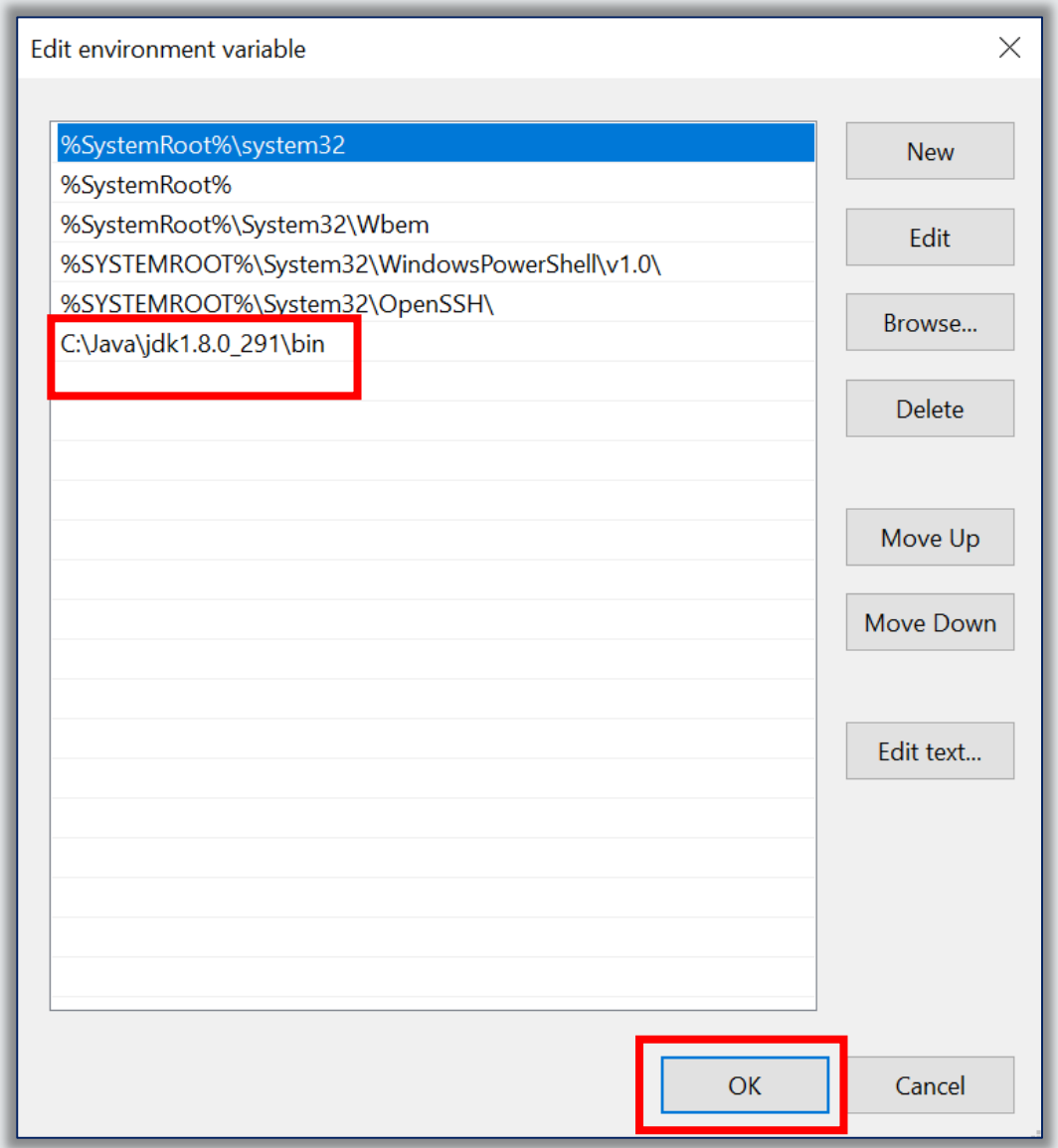
Set the Environment Variable for JAVA



Set the Environment Variable for JAVA



Set the Environment Variable for JAVA



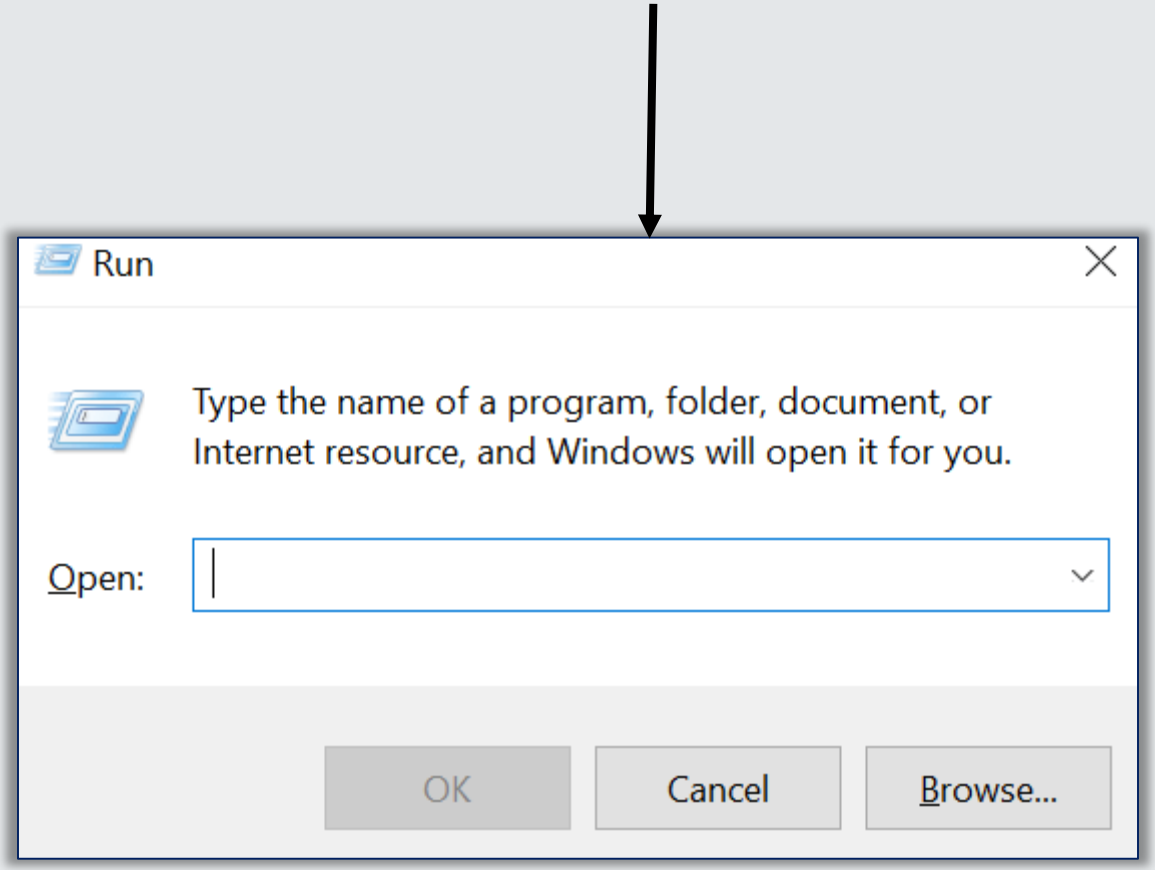
Check whether java is functional or not

Steps to check Java is functional or not

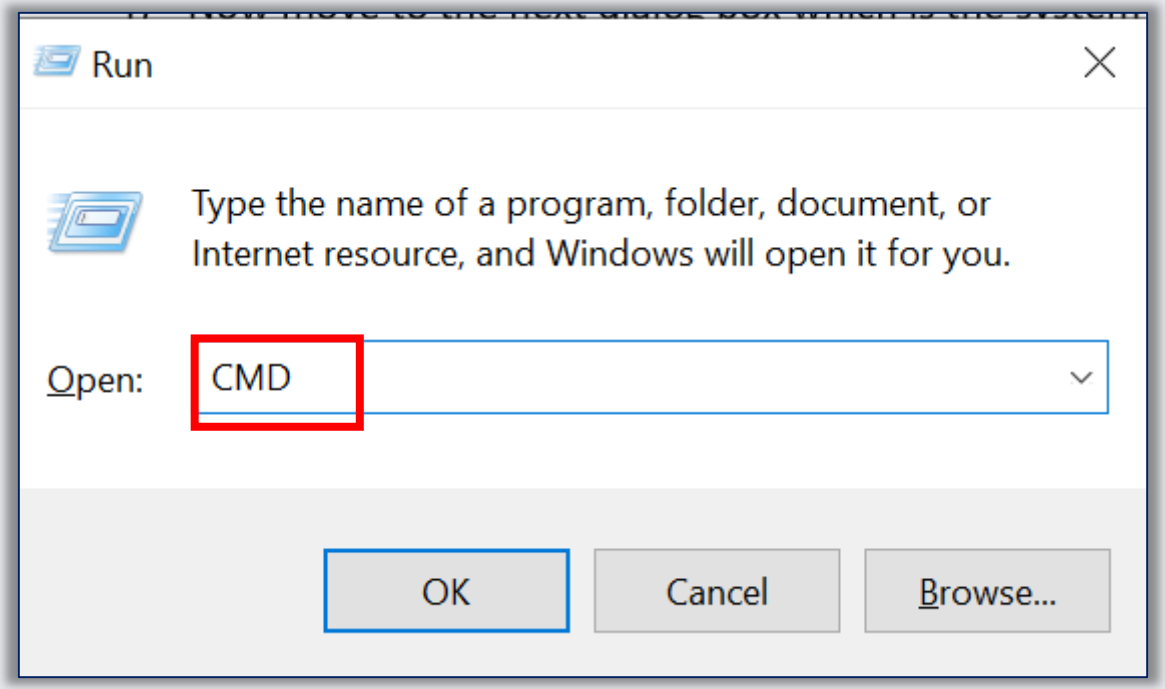
Windows + R

Steps to check Java is functional or not

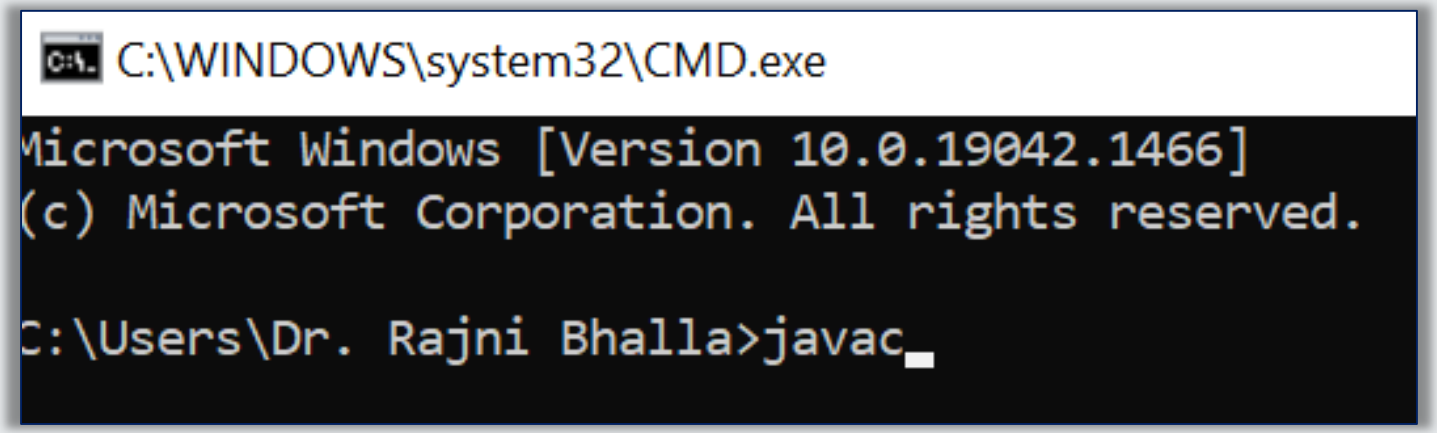
Windows + R



Steps to check Java is functional or not



Steps to check Java is functional or not

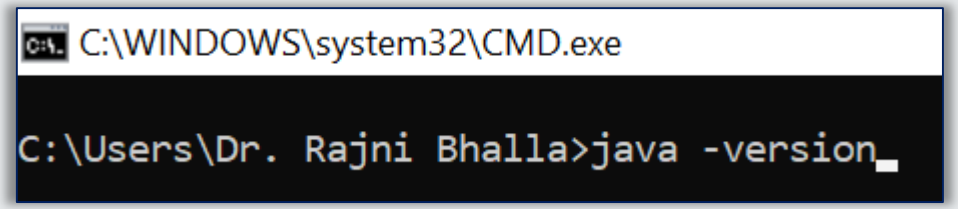


Steps to check Java is functional or not

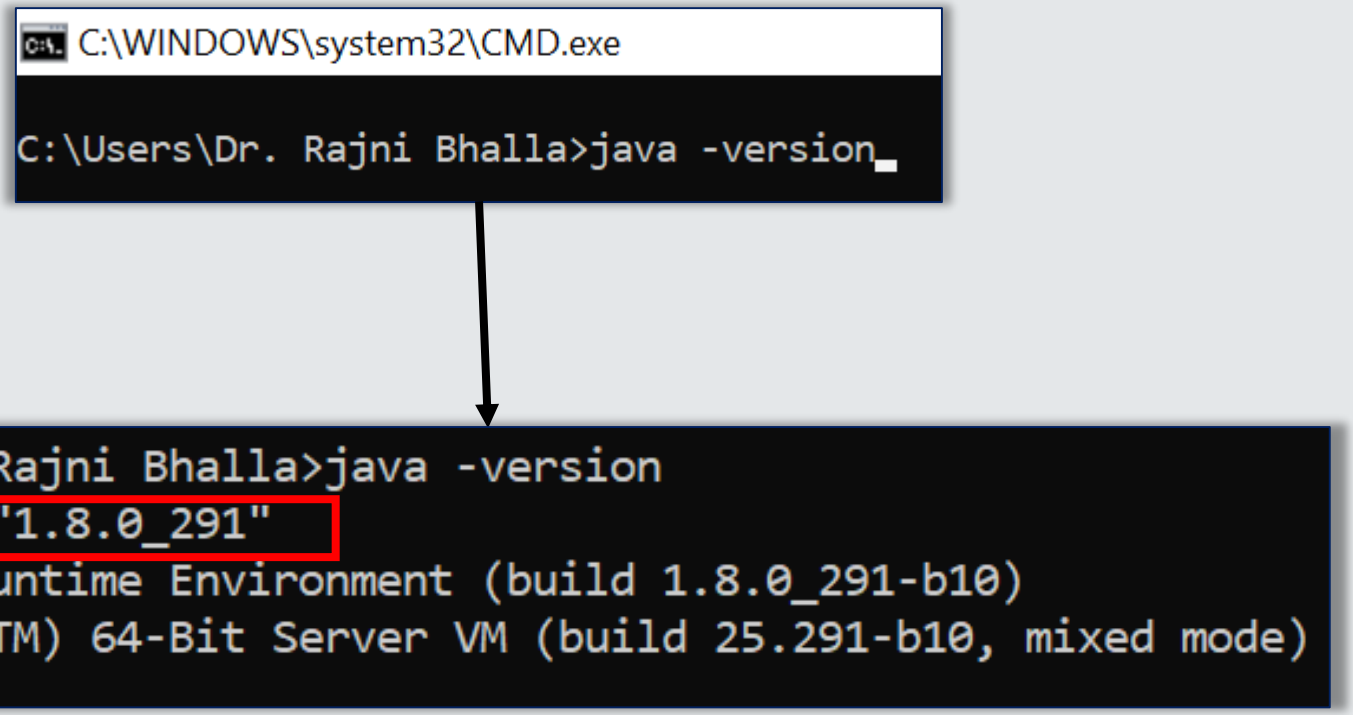
```
C:\Users\Dr. Rajni Bhalla>javac
Usage: javac <options> <source files>
where possible options include:
  -g                         Generate all debugging info
  -g:none                    Generate no debugging info
  -g:{lines,vars,source}     Generate only some debugging info
  -nowarn                    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

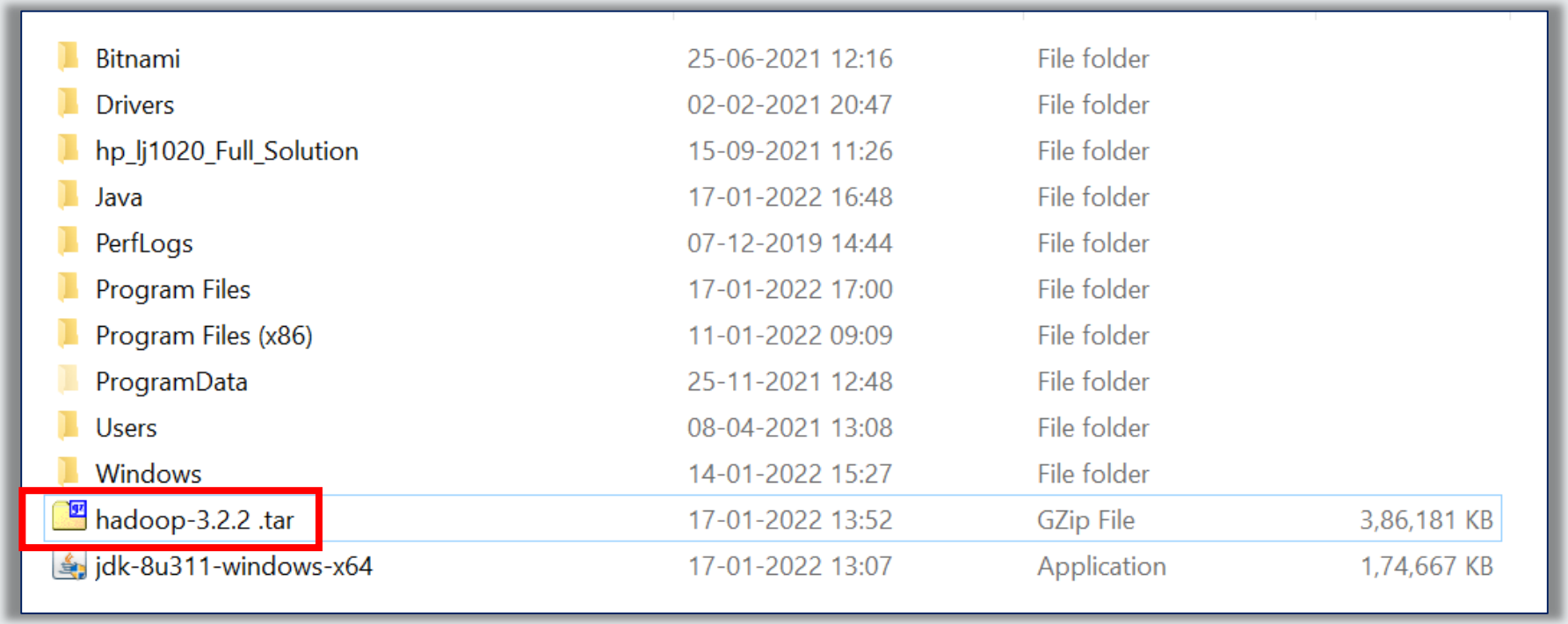

C:\WINDOWS\system32\CMD.exe
C:\Users\Dr. Rajni Bhalla>java -version

CMD

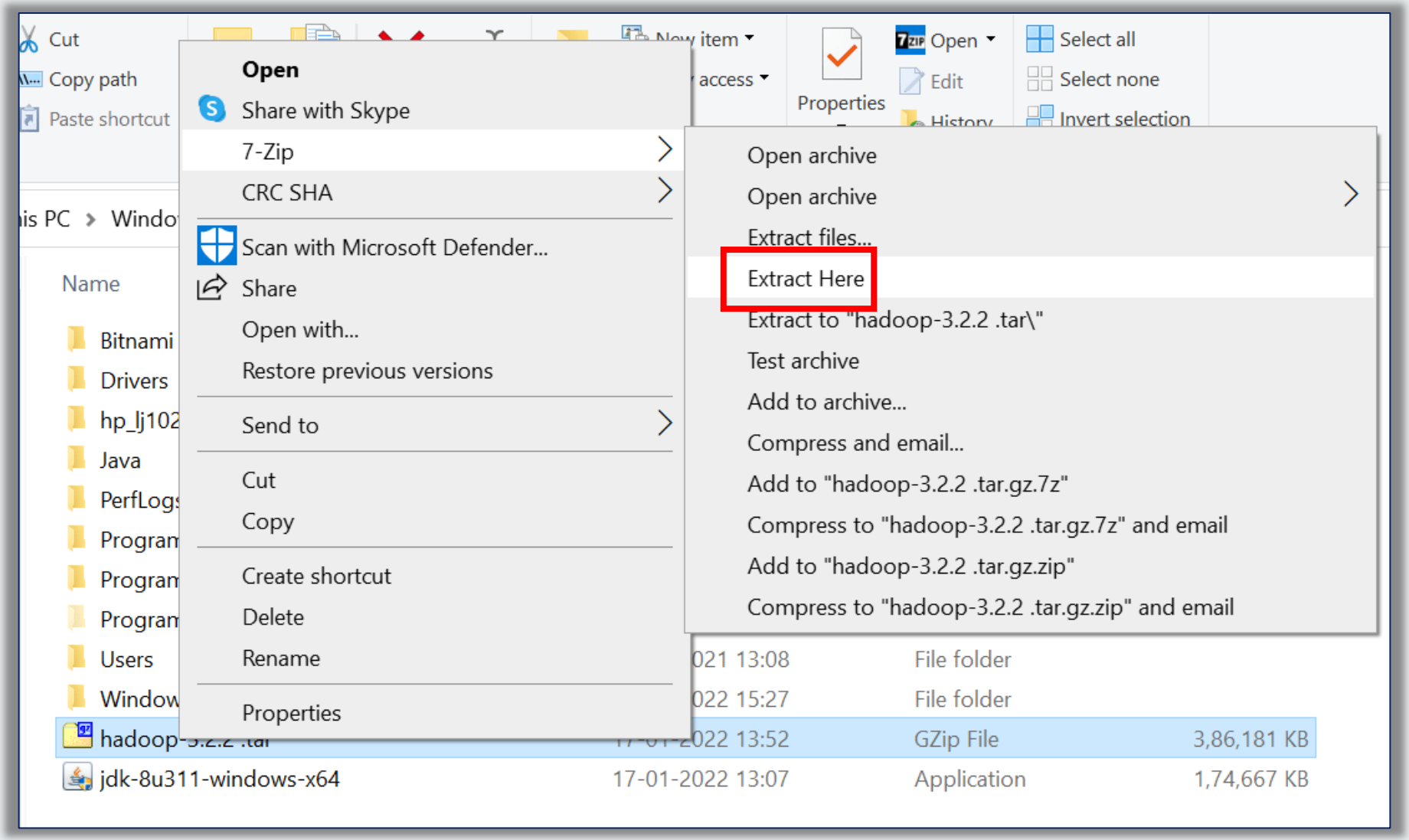

C:\WINDOWS\system32\CMD.exe
C:\Users\Dr. Rajni Bhalla>java -version
C:\Users\Dr. Rajni Bhalla>java -version
java version "1.8.0_291"
Java(TM) SE Runtime Environment (build 1.8.0_291-b10)
Java HotSpot(TM) 64-Bit Server VM (build 25.291-b10, mixed mode)

INSTALL HADOOP INTO OUR LOCAL SYSTEM

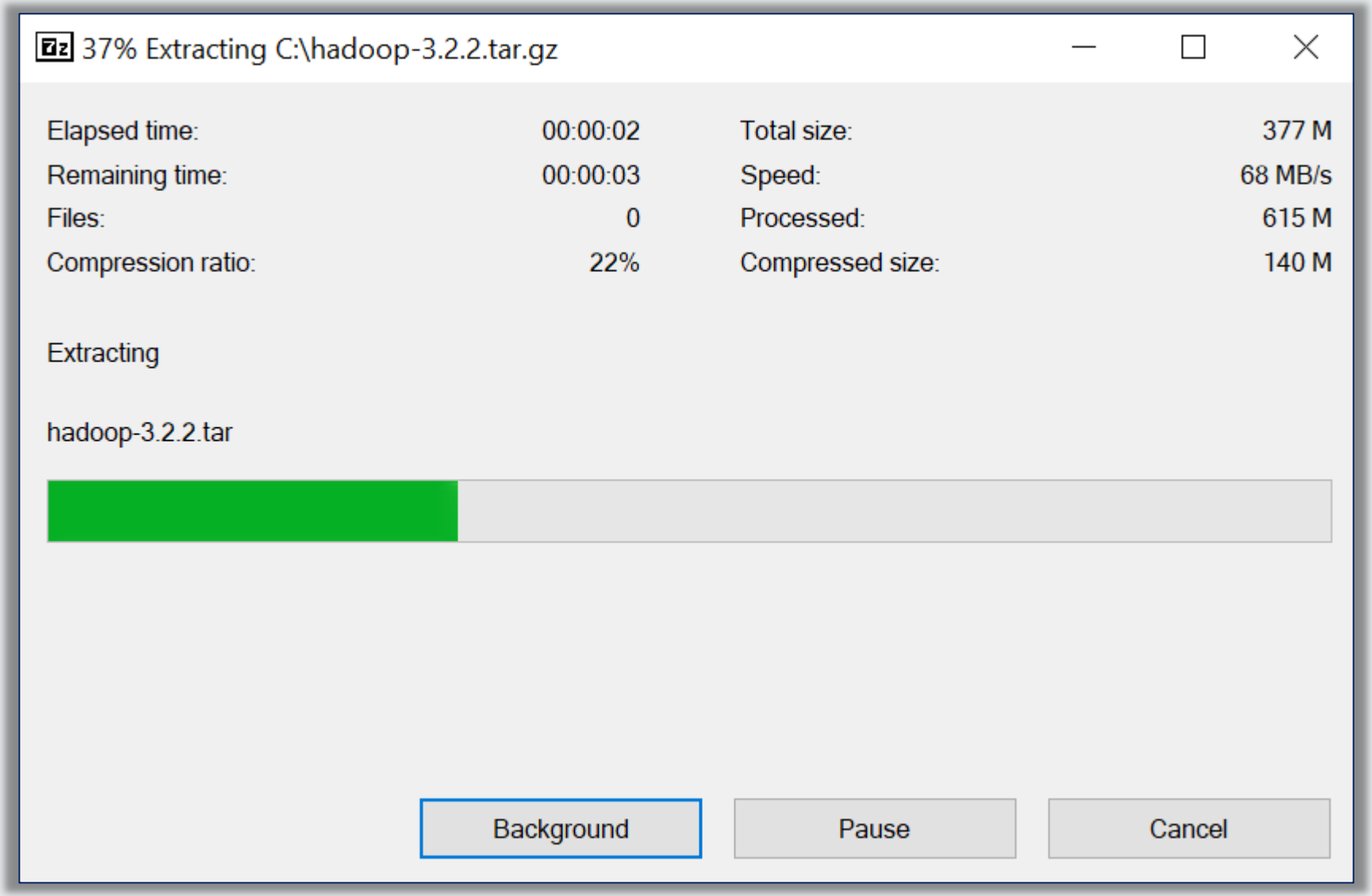
Steps to install Hadoop



Steps to install Hadoop



Steps to install Hadoop

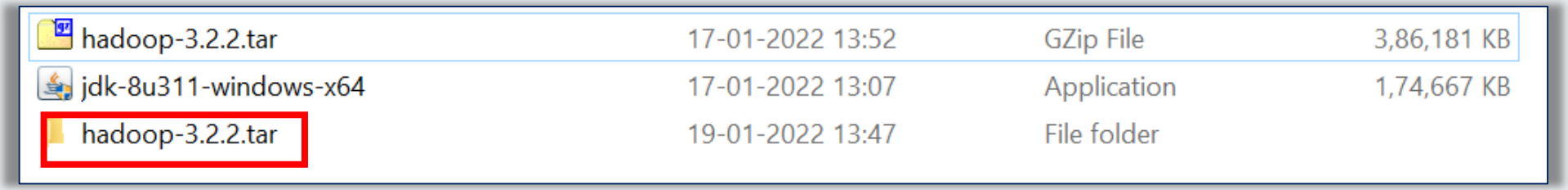


Steps to install Hadoop

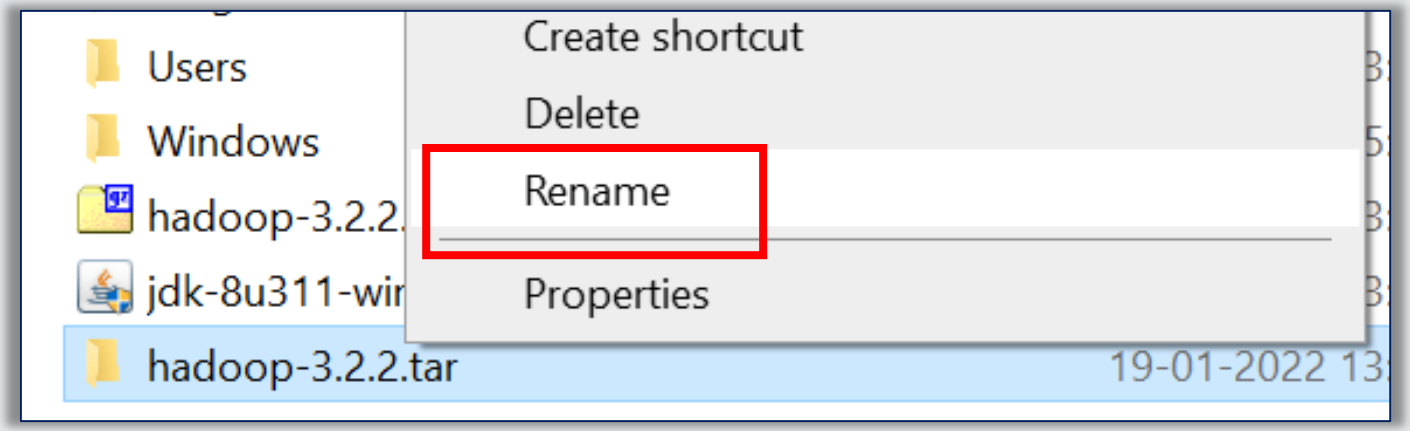
Ignore

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

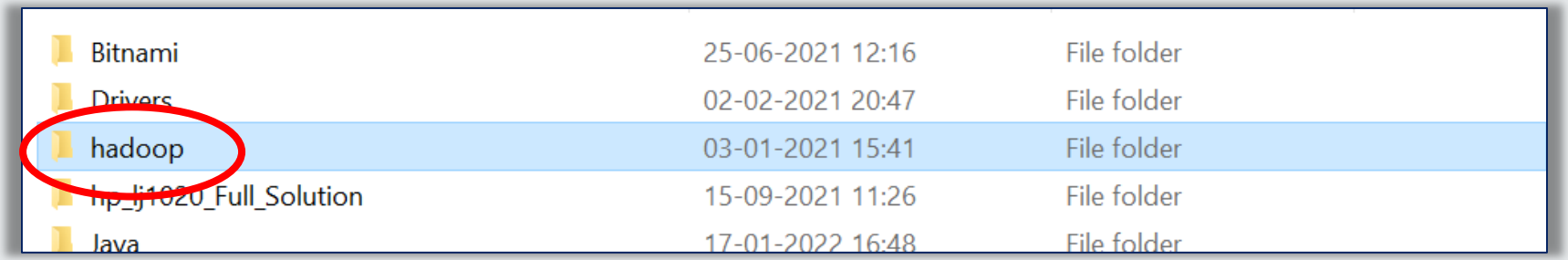
Steps to install Hadoop



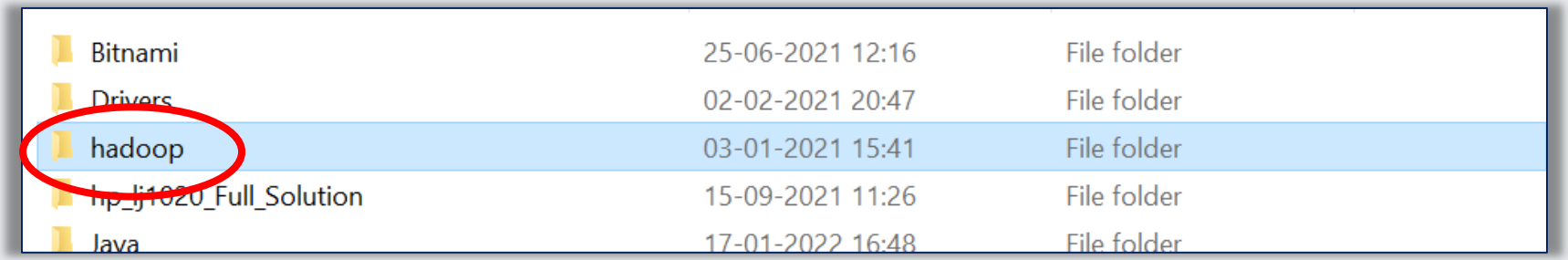
Steps to install Hadoop

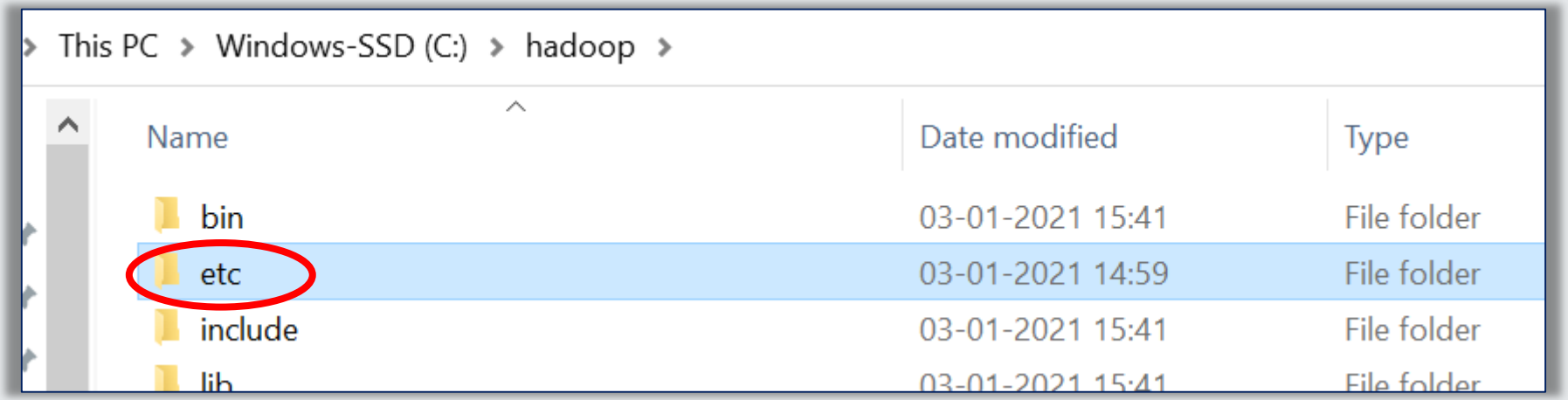


Steps to install Hadoop

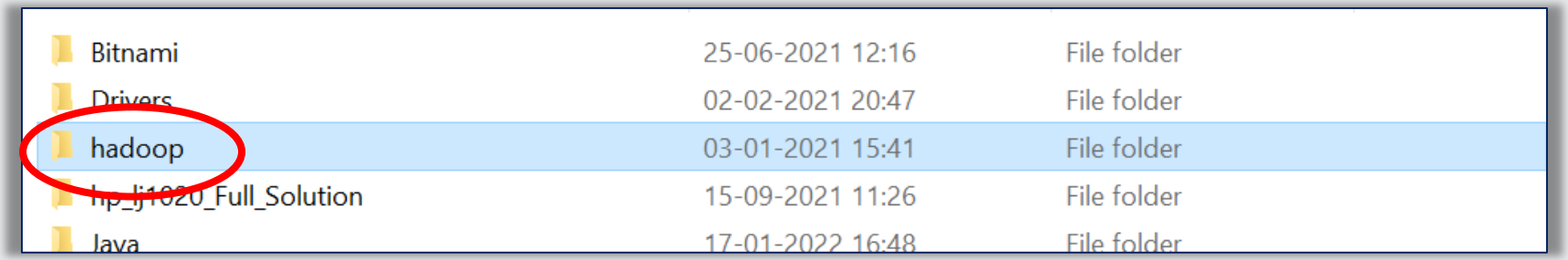


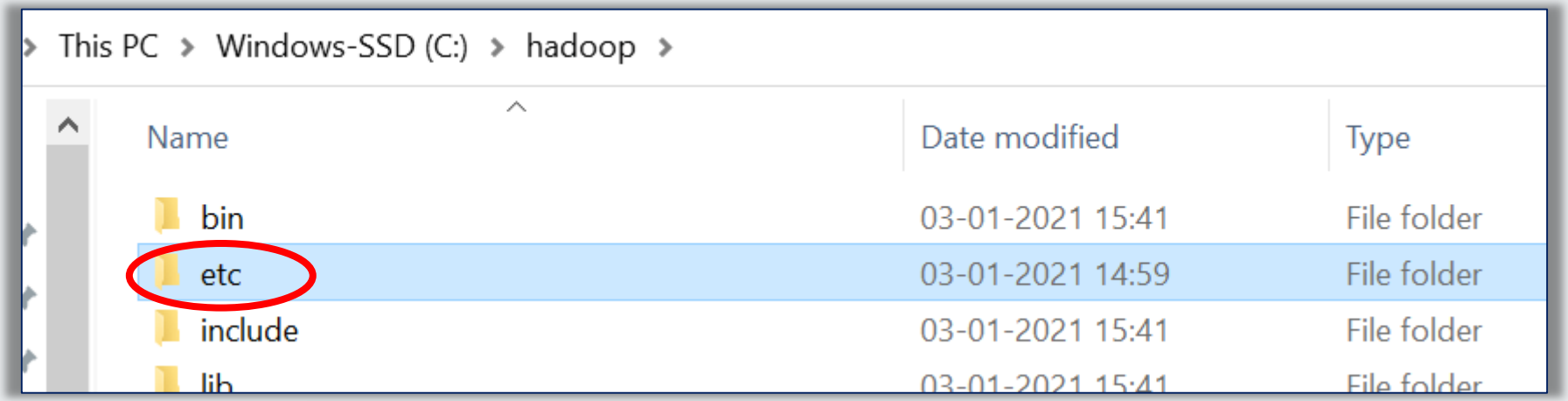
Steps to install Hadoop

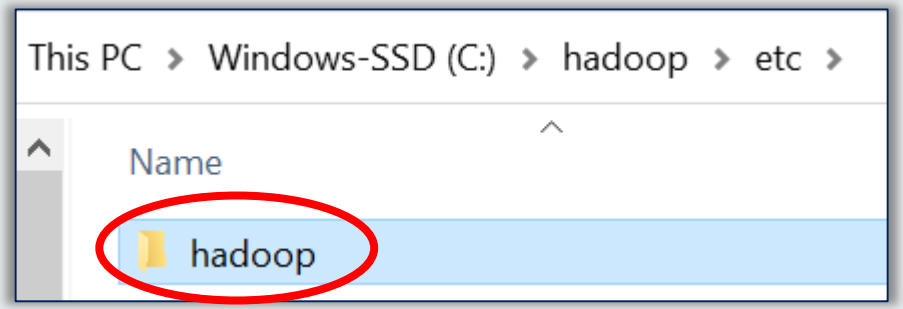




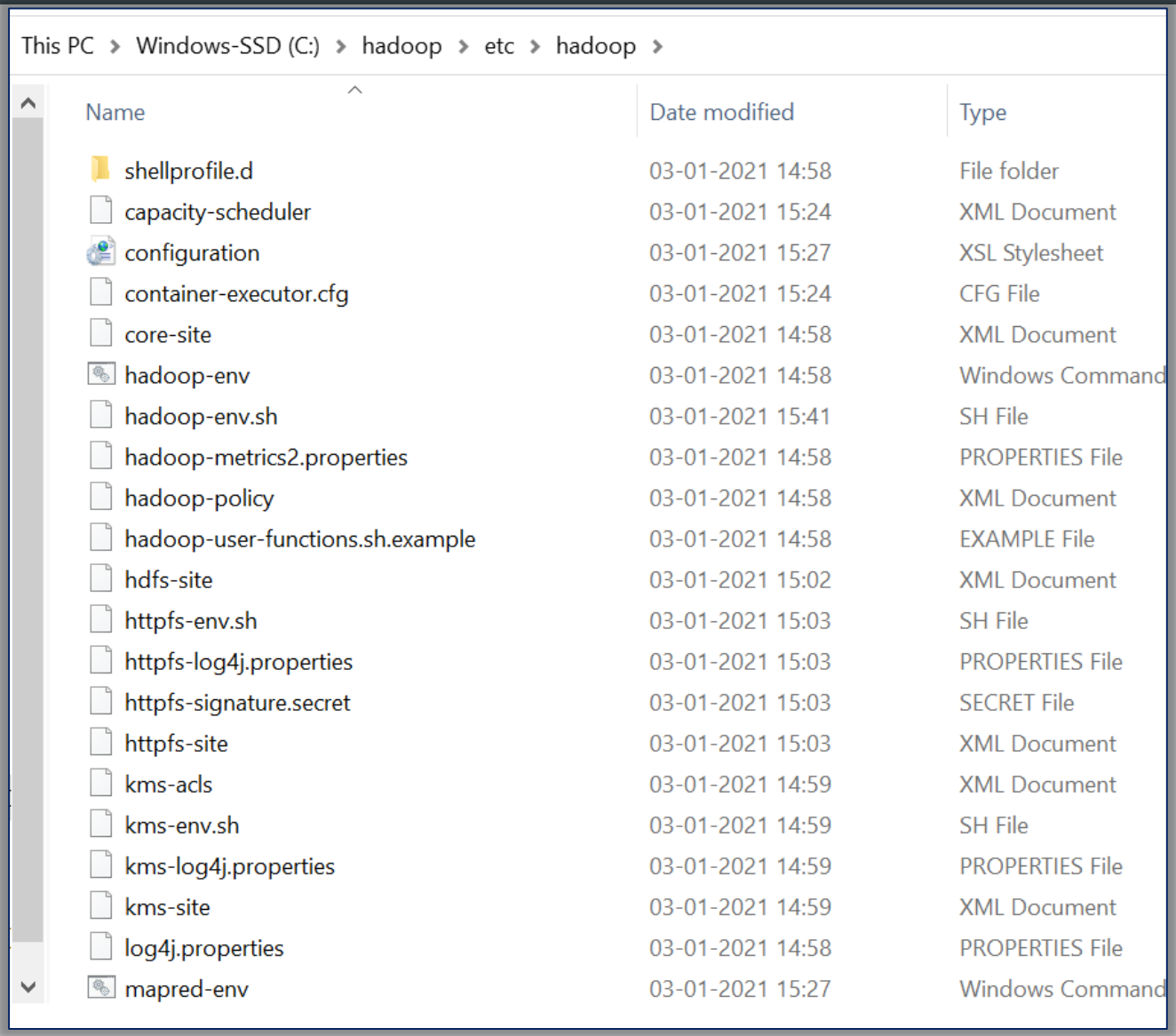
Steps to install Hadoop



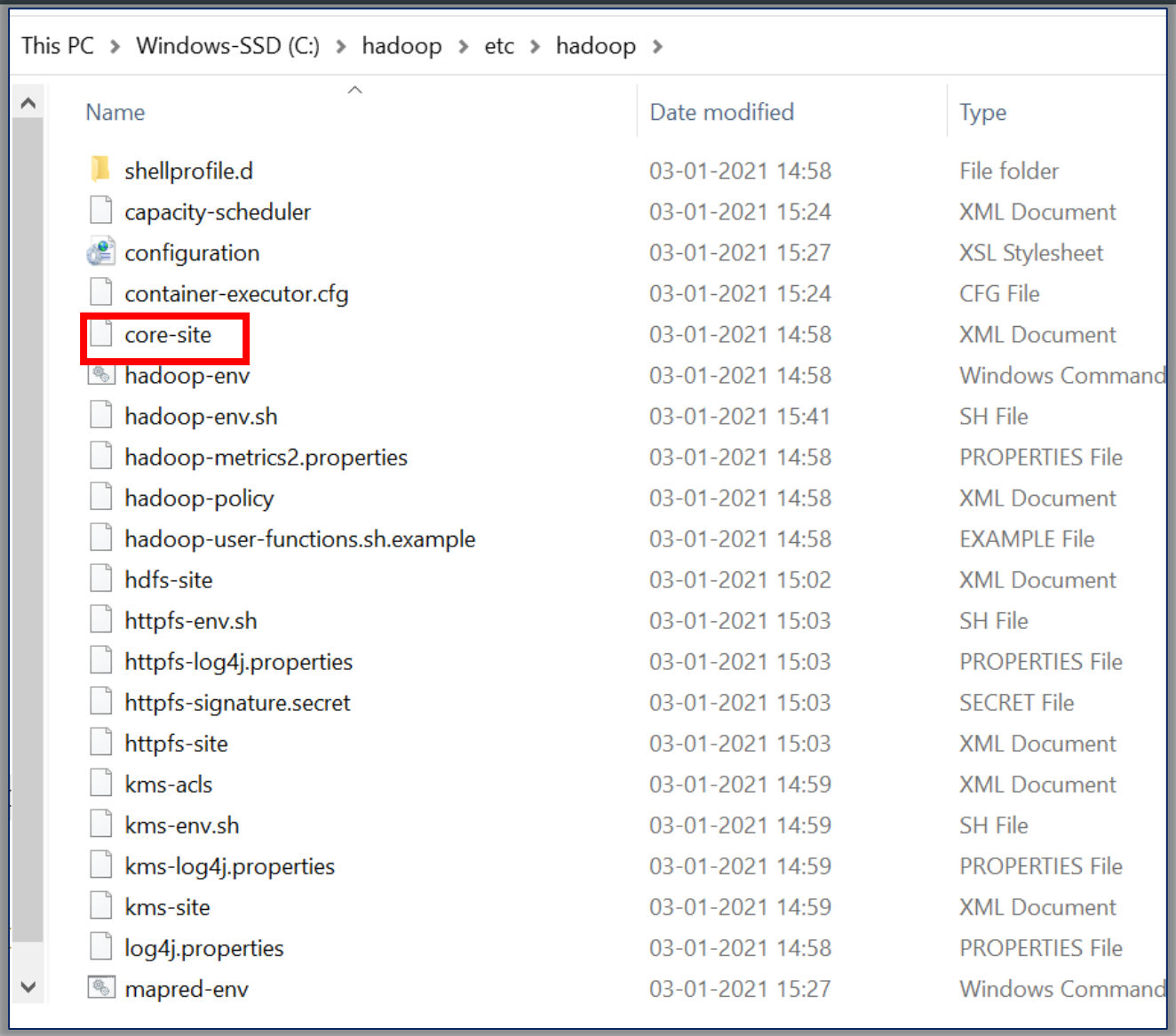




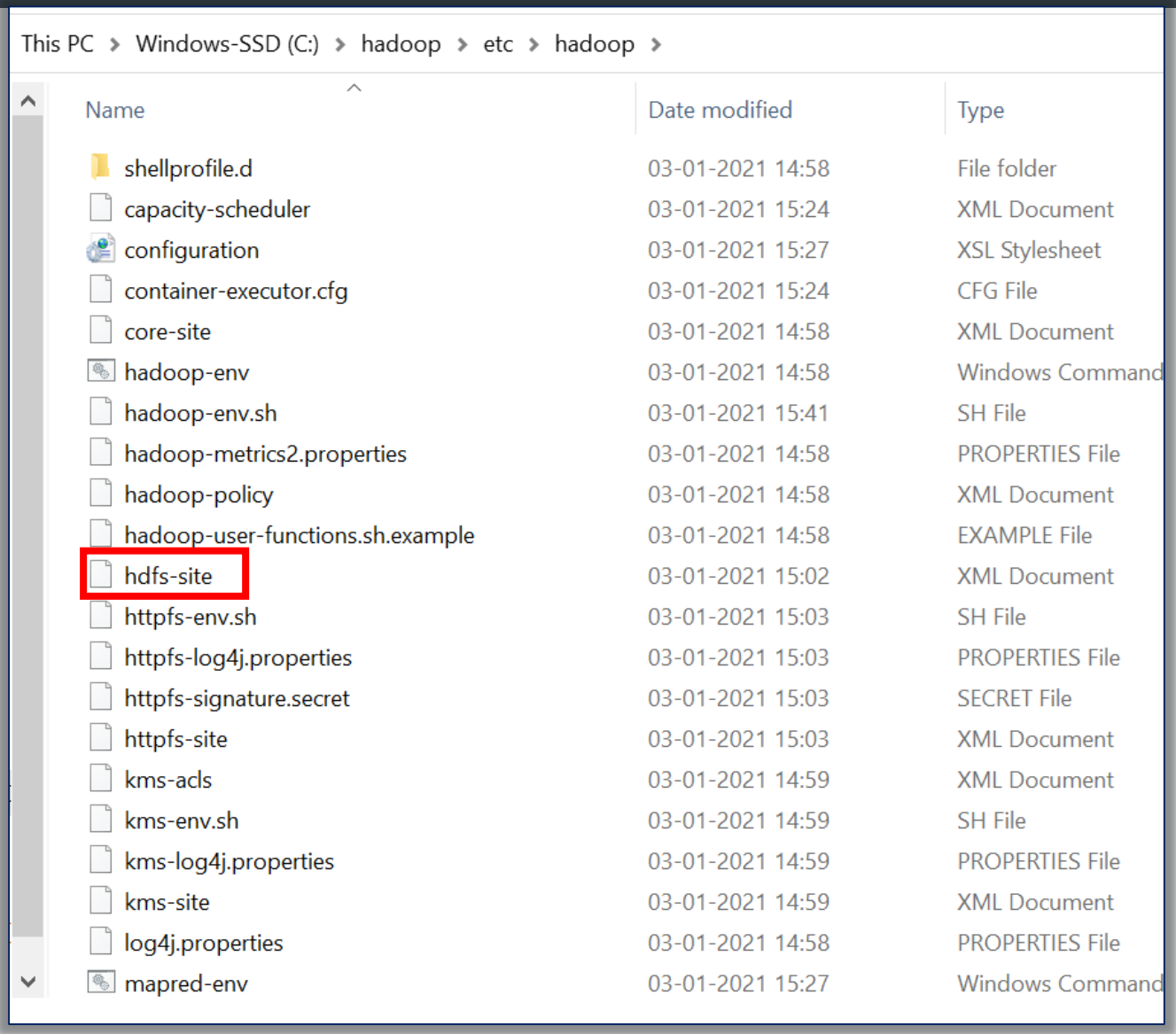
Steps to install Hadoop



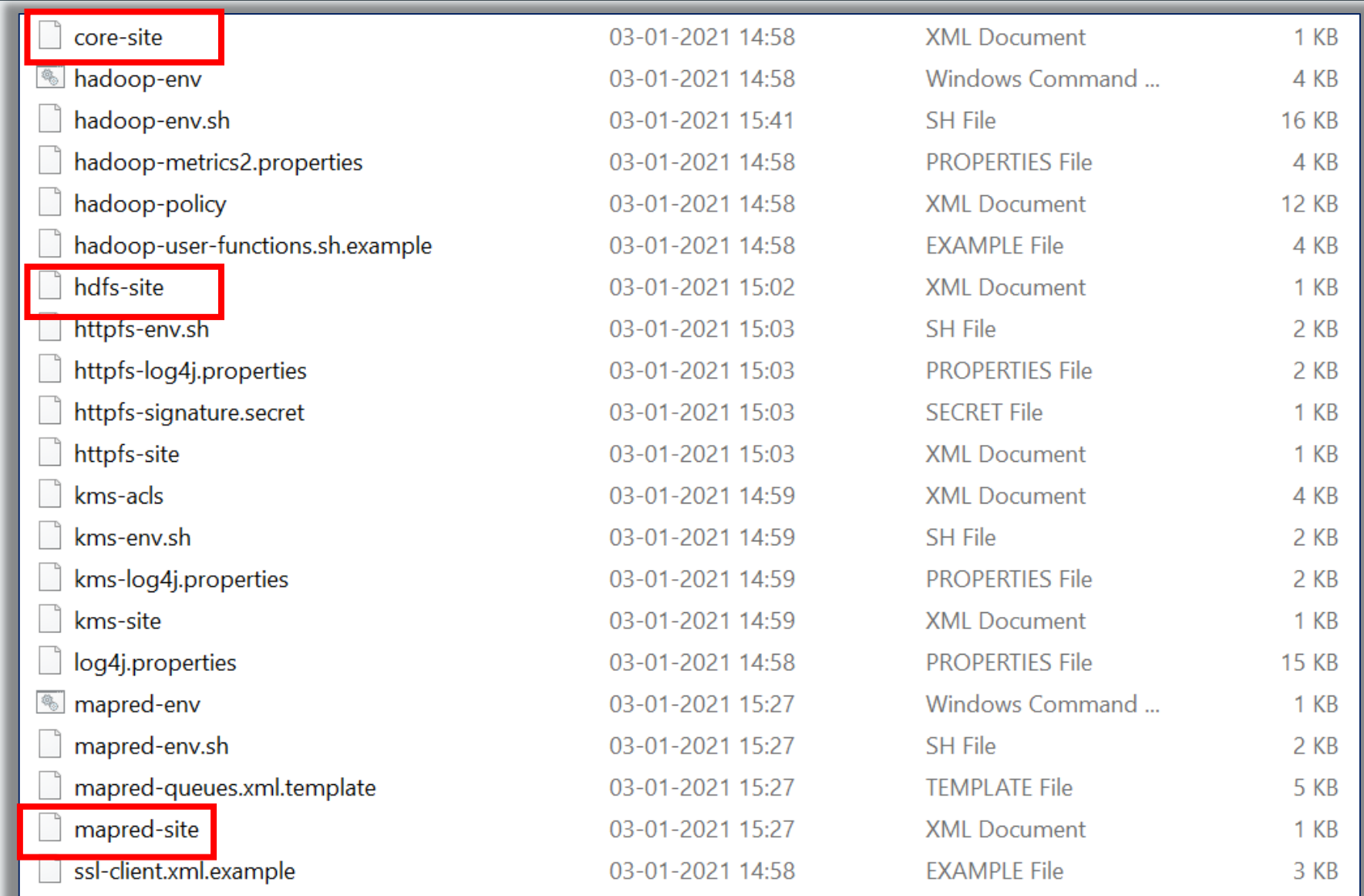
Steps to install Hadoop



Steps to install Hadoop

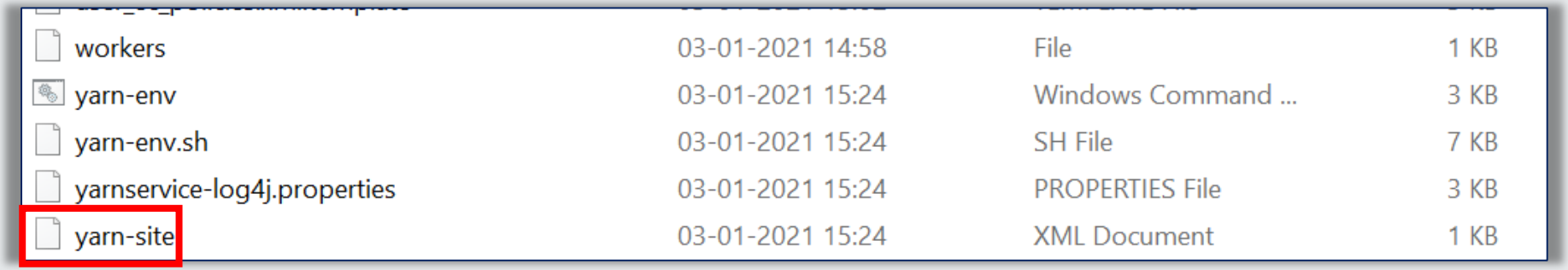


Steps to install Hadoop



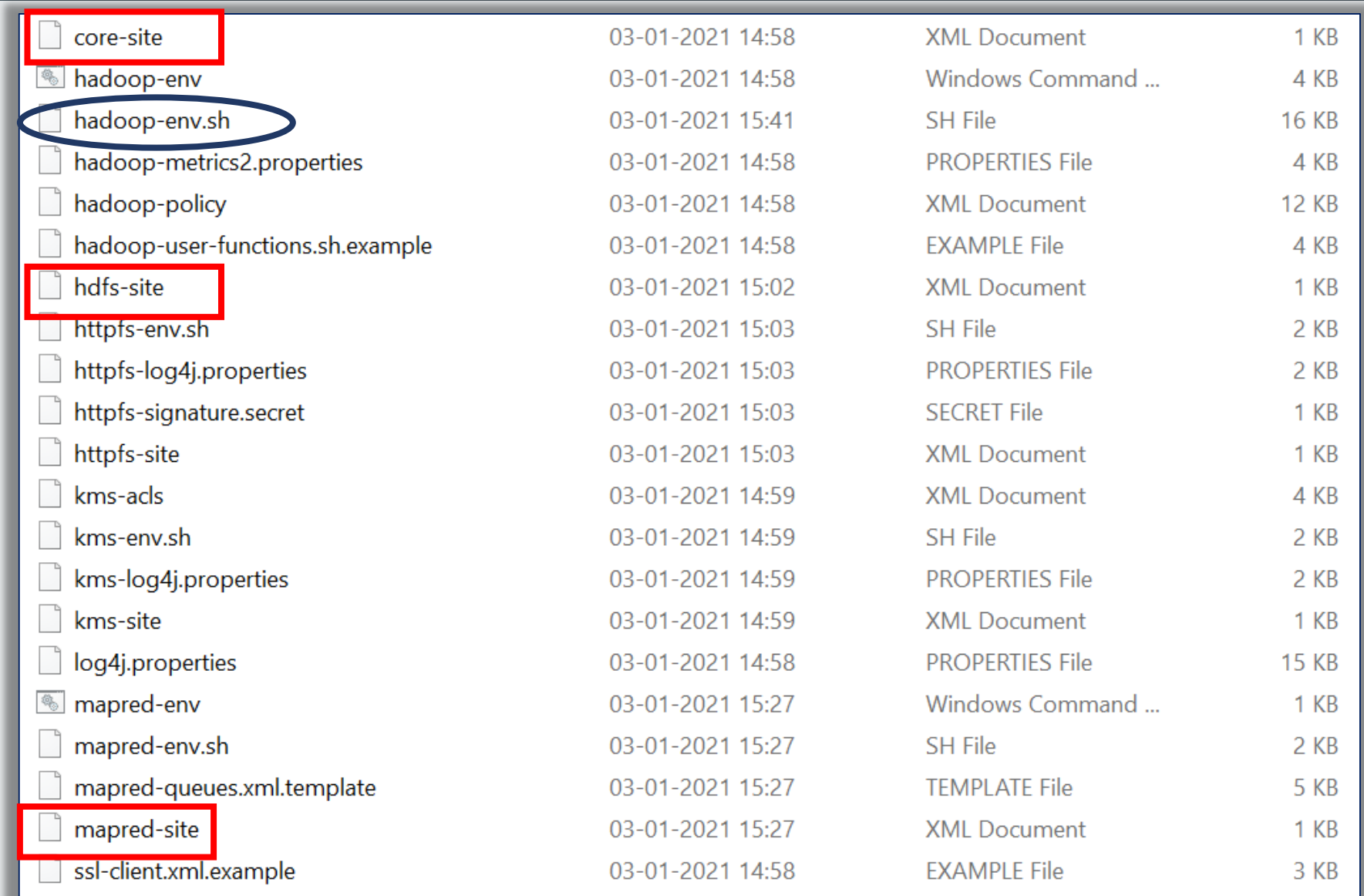
core-site	03-01-2021 14:58	XML Document	1 KB
hadoop-env	03-01-2021 14:58	Windows Command ...	4 KB
hadoop-env.sh	03-01-2021 15:41	SH File	16 KB
hadoop-metrics2.properties	03-01-2021 14:58	PROPERTIES File	4 KB
hadoop-policy	03-01-2021 14:58	XML Document	12 KB
hadoop-user-functions.sh.example	03-01-2021 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



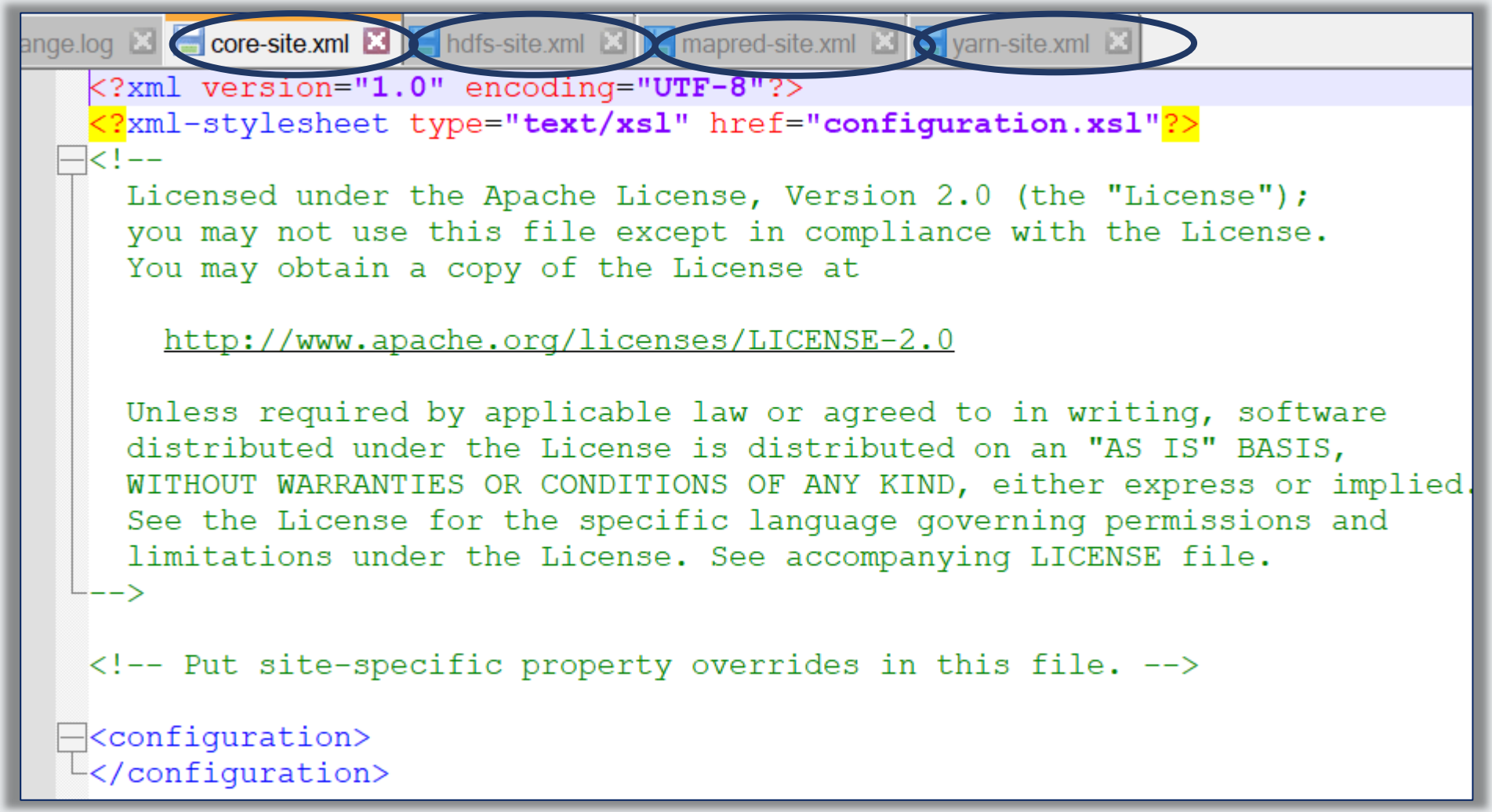
workers	03-01-2021 14:58	File	1 KB
yarn-env	03-01-2021 15:24	Windows Command ...	3 KB
yarn-env.sh	03-01-2021 15:24	SH File	7 KB
yarnservice-log4j.properties	03-01-2021 15:24	PROPERTIES File	3 KB
yarn-site	03-01-2021 15:24	XML Document	1 KB

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

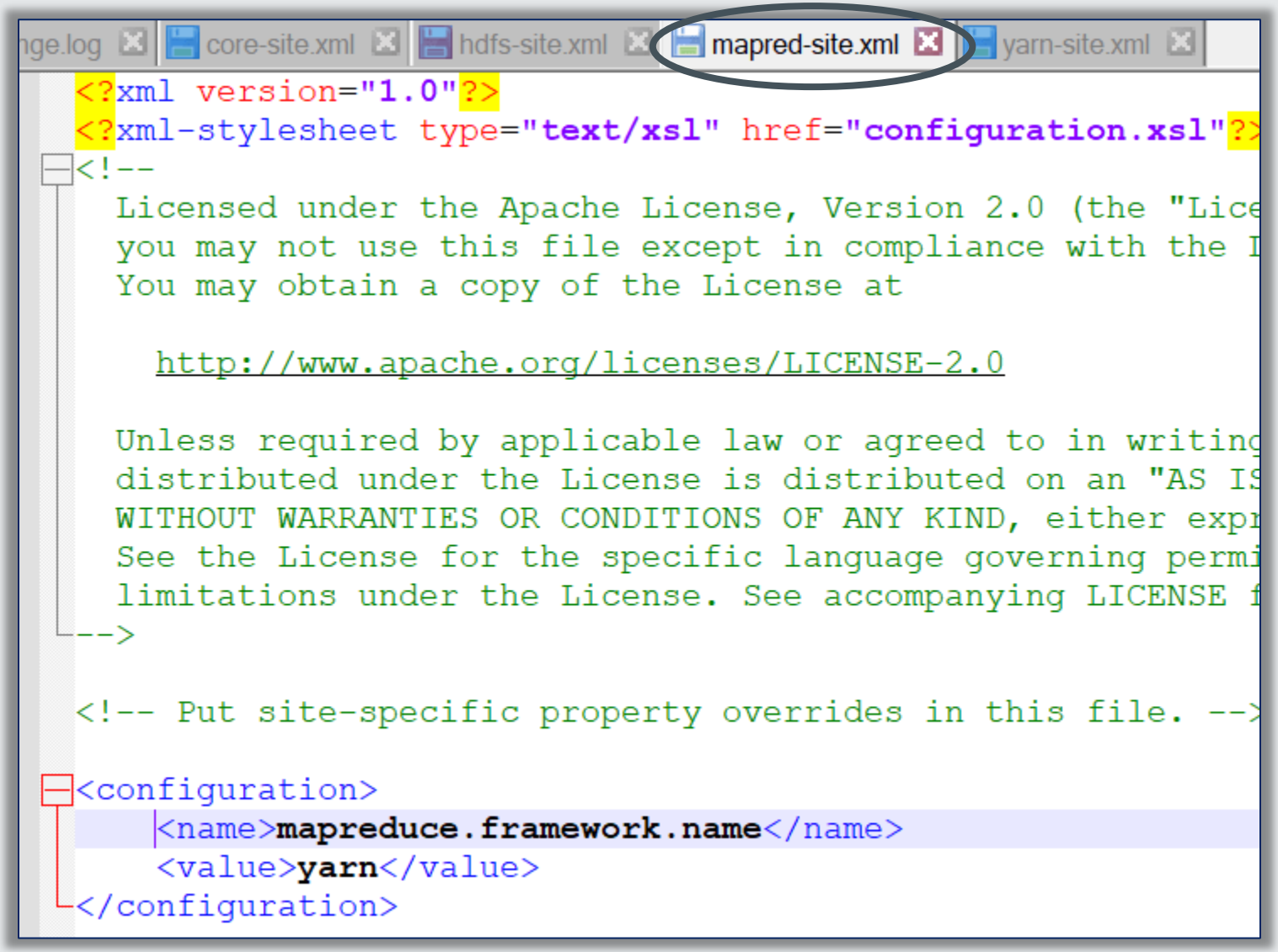


Steps to install Hadoop

core-site.xml

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

Steps to install Hadoop



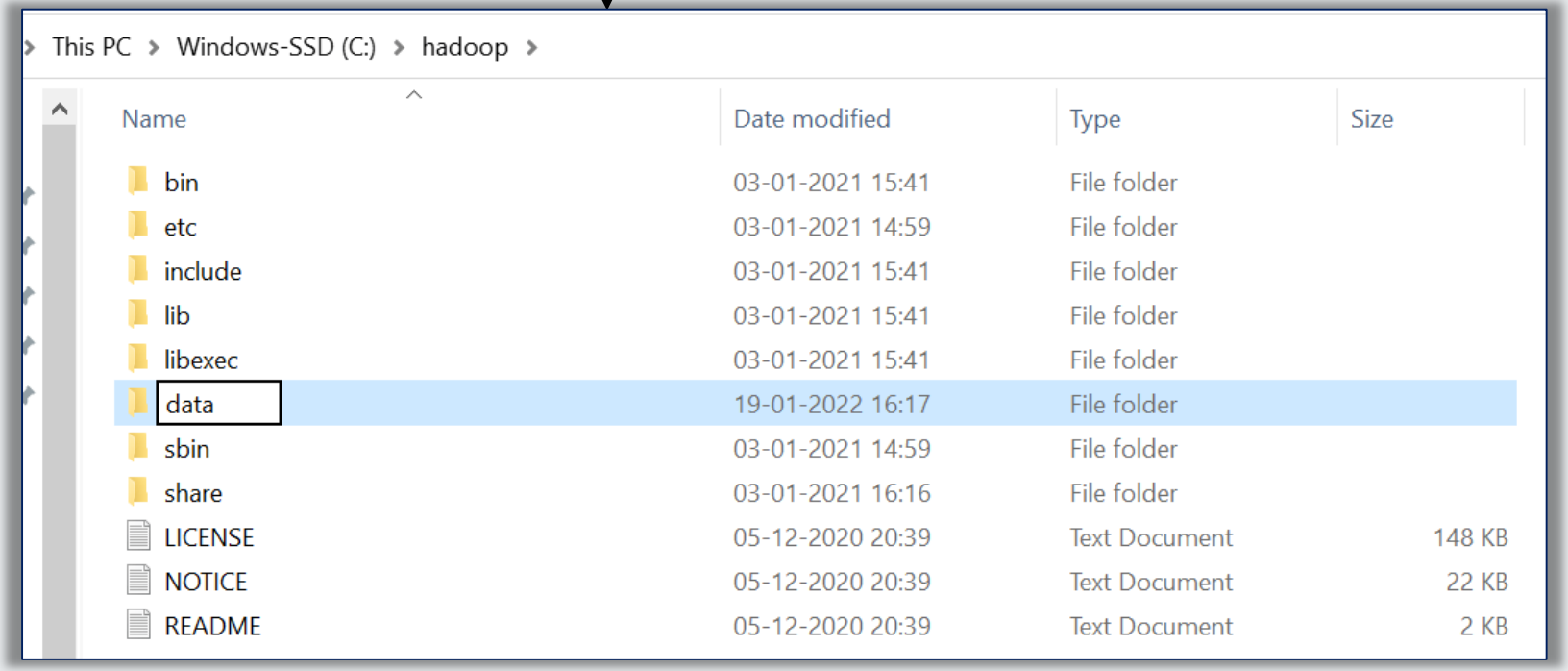
Steps to install Hadoop

e.log | core-site.xml | hdfs-site.xml | mapred-site.xml | yarn-site.xml

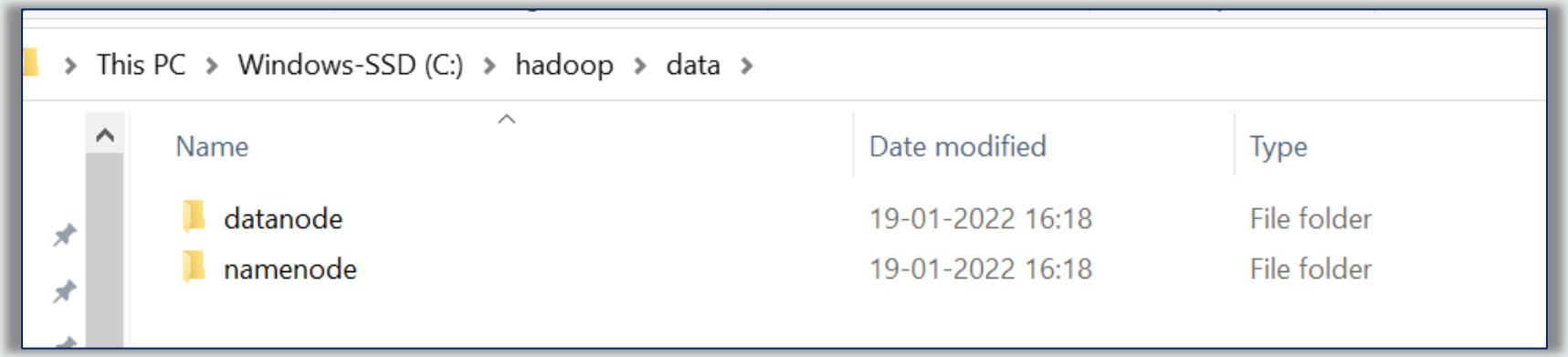
```
<?xml version="1.0"?>
<!--
  Licensed under the Apache License, Version 2.0 (the "License");
  you may not use this file except in compliance with the License.
  You may obtain a copy of the License at

    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

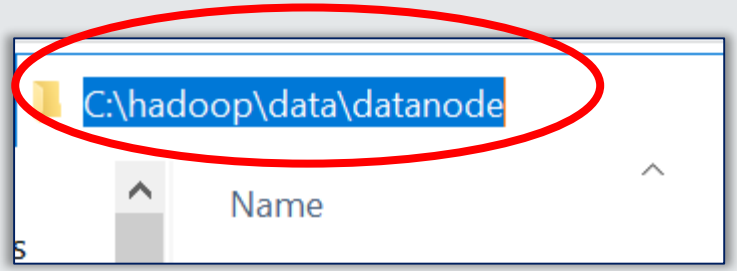
C:\hadoop

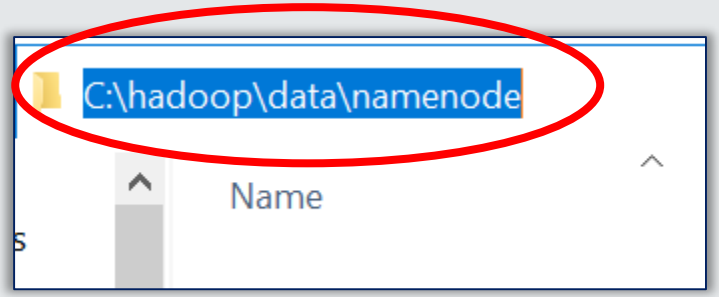


Steps to install Hadoop



Steps to install Hadoop



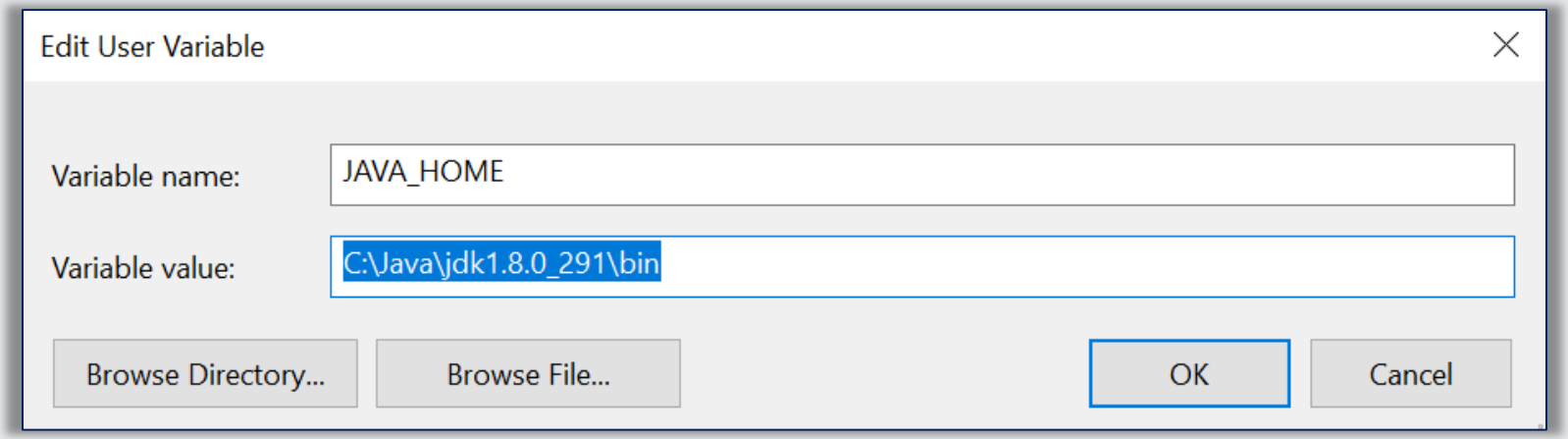


Steps to install hadoop

e.log | core-site.xml | hdfs-site.xml | mapred-site.xml | yarn-site.xml | hadoop-env.cmd

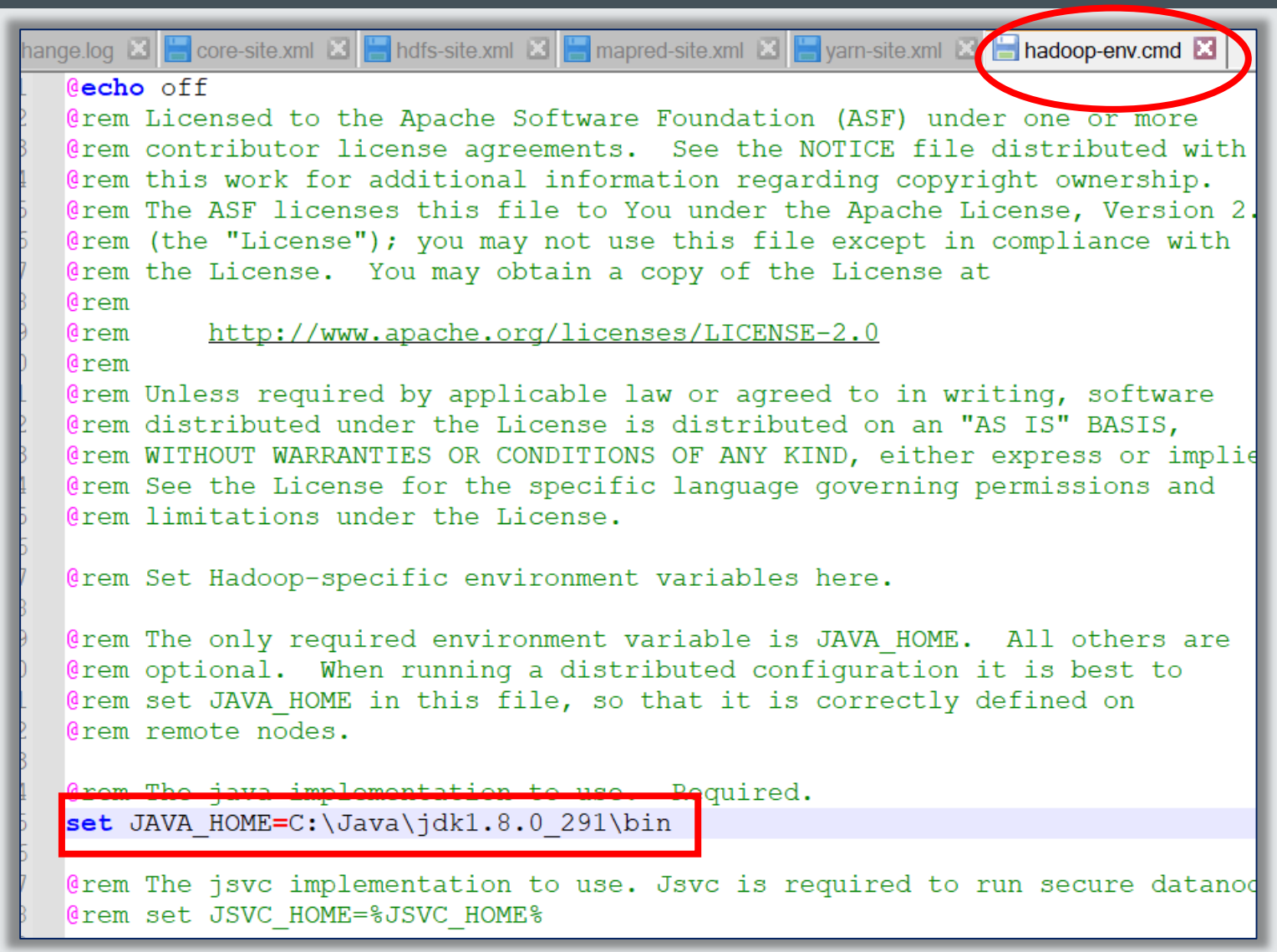
```
    http://www.apache.org/licenses/LICENSE-2.0

  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>
    <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

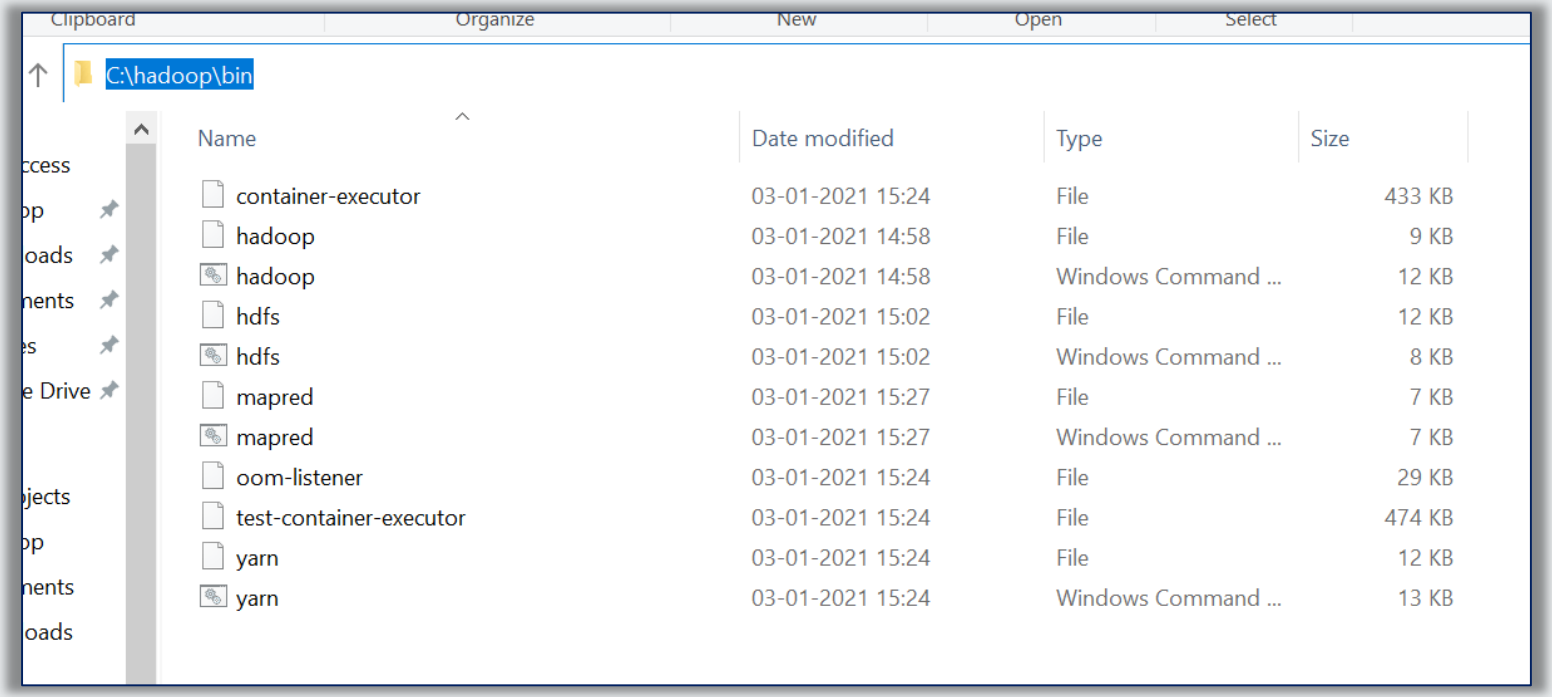


To get the correct location, copy from path

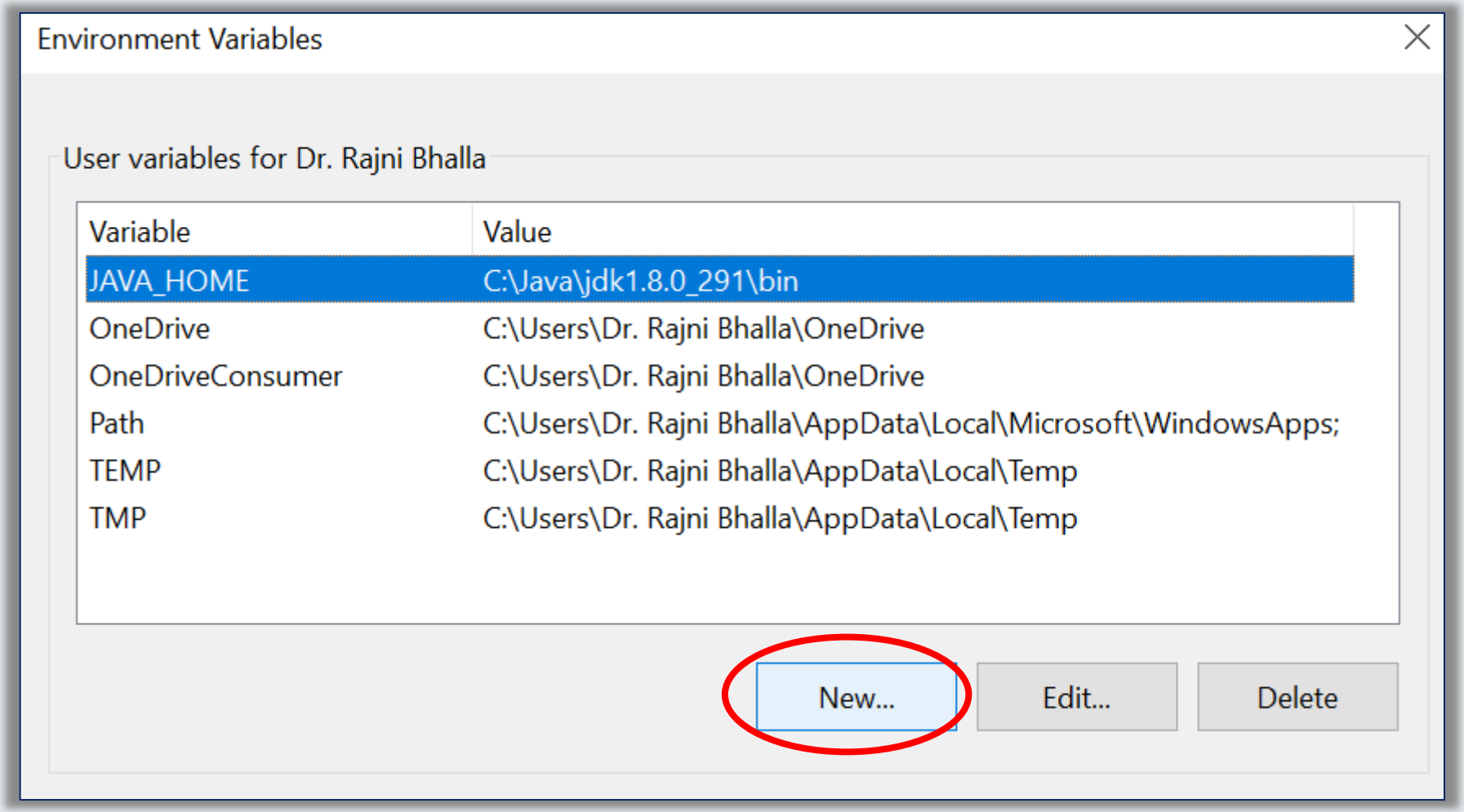
Steps to install hadoop



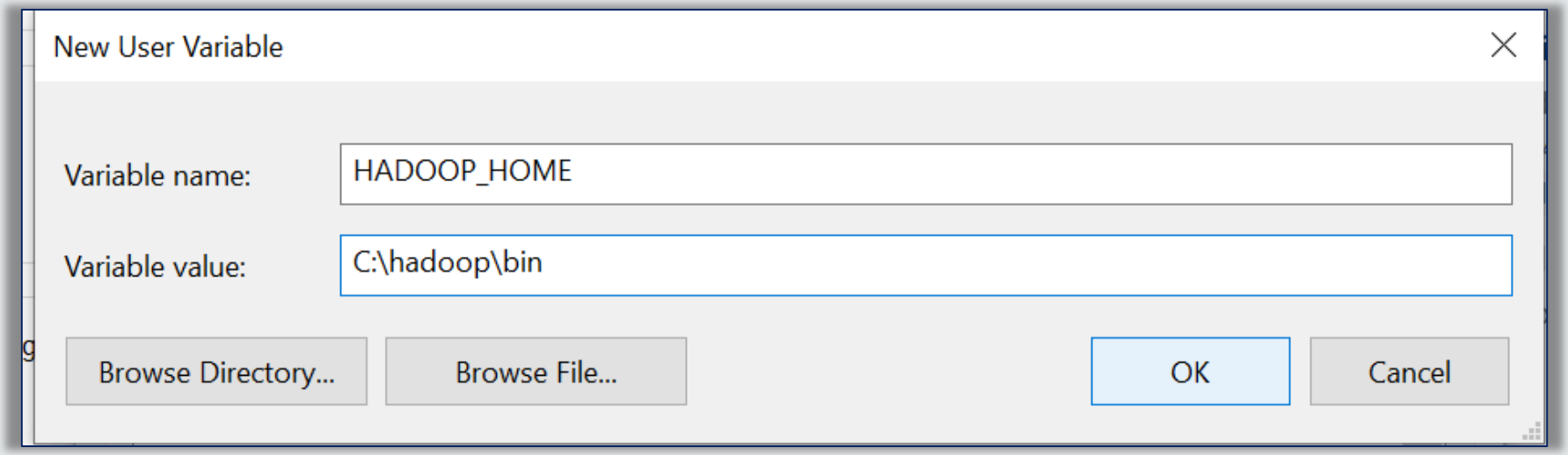
Steps to install hadoop



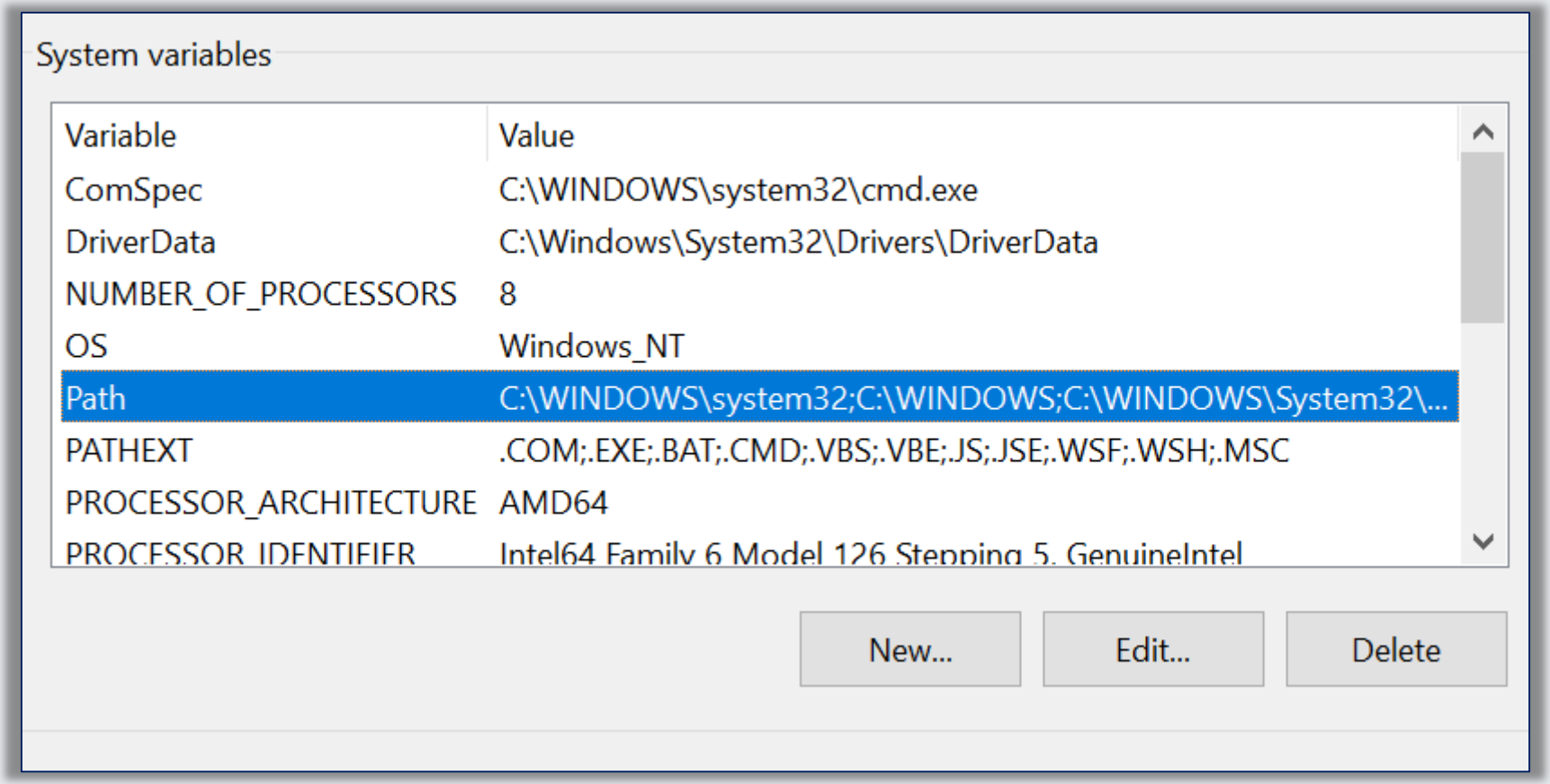
Steps to install hadoop



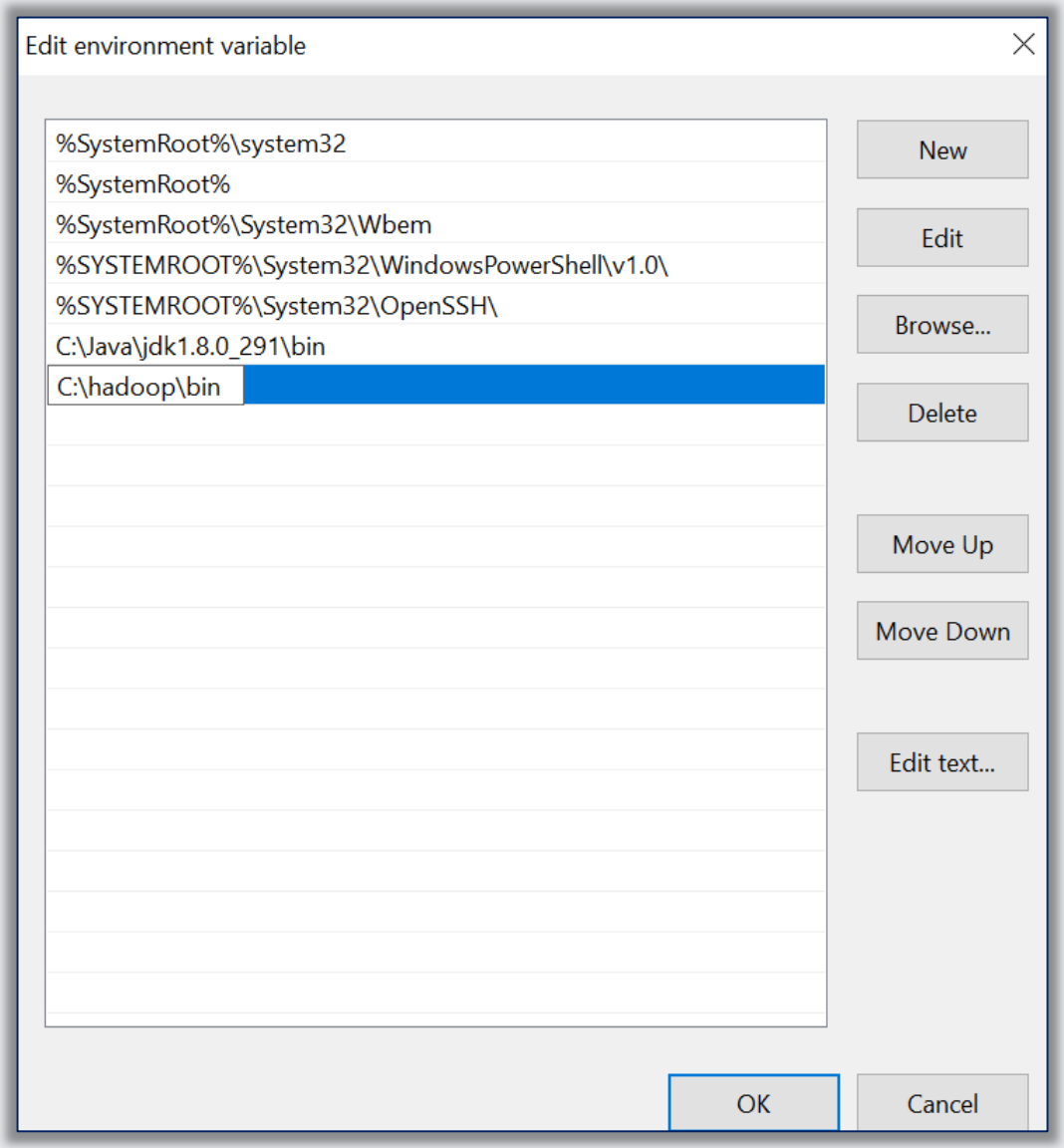
Steps to install hadoop



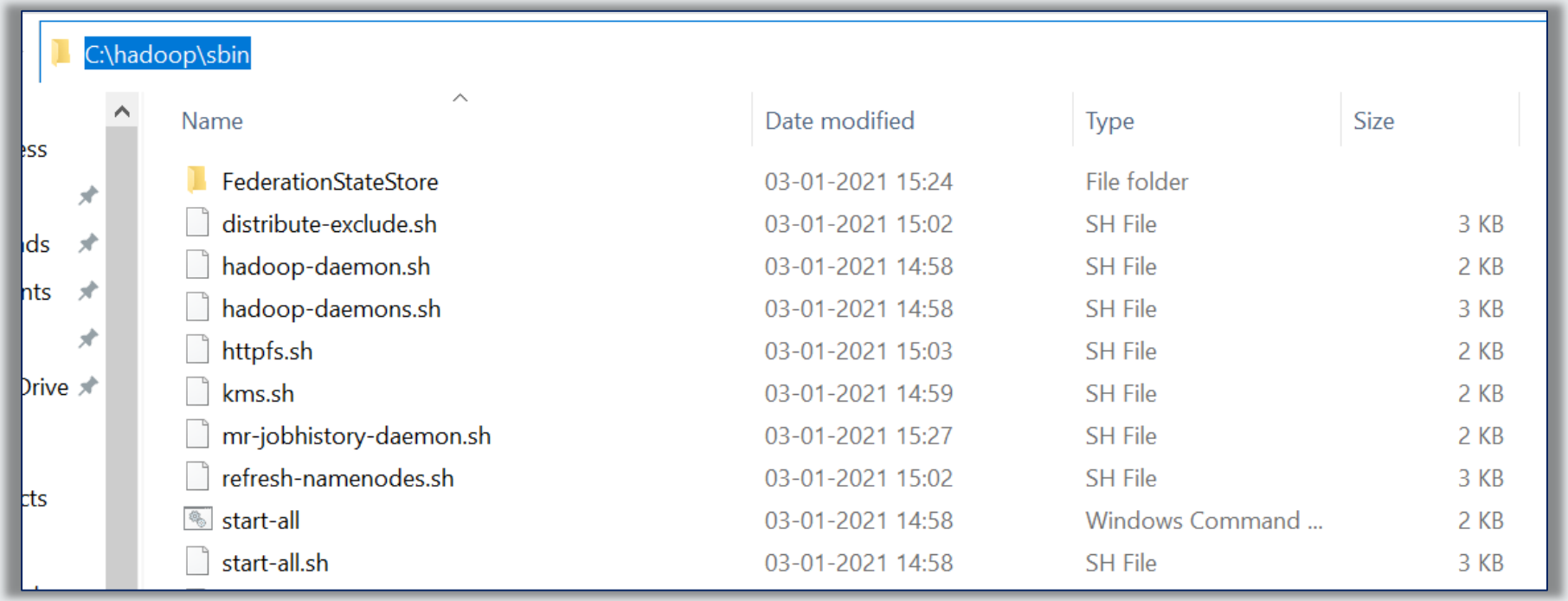
Steps to install hadoop



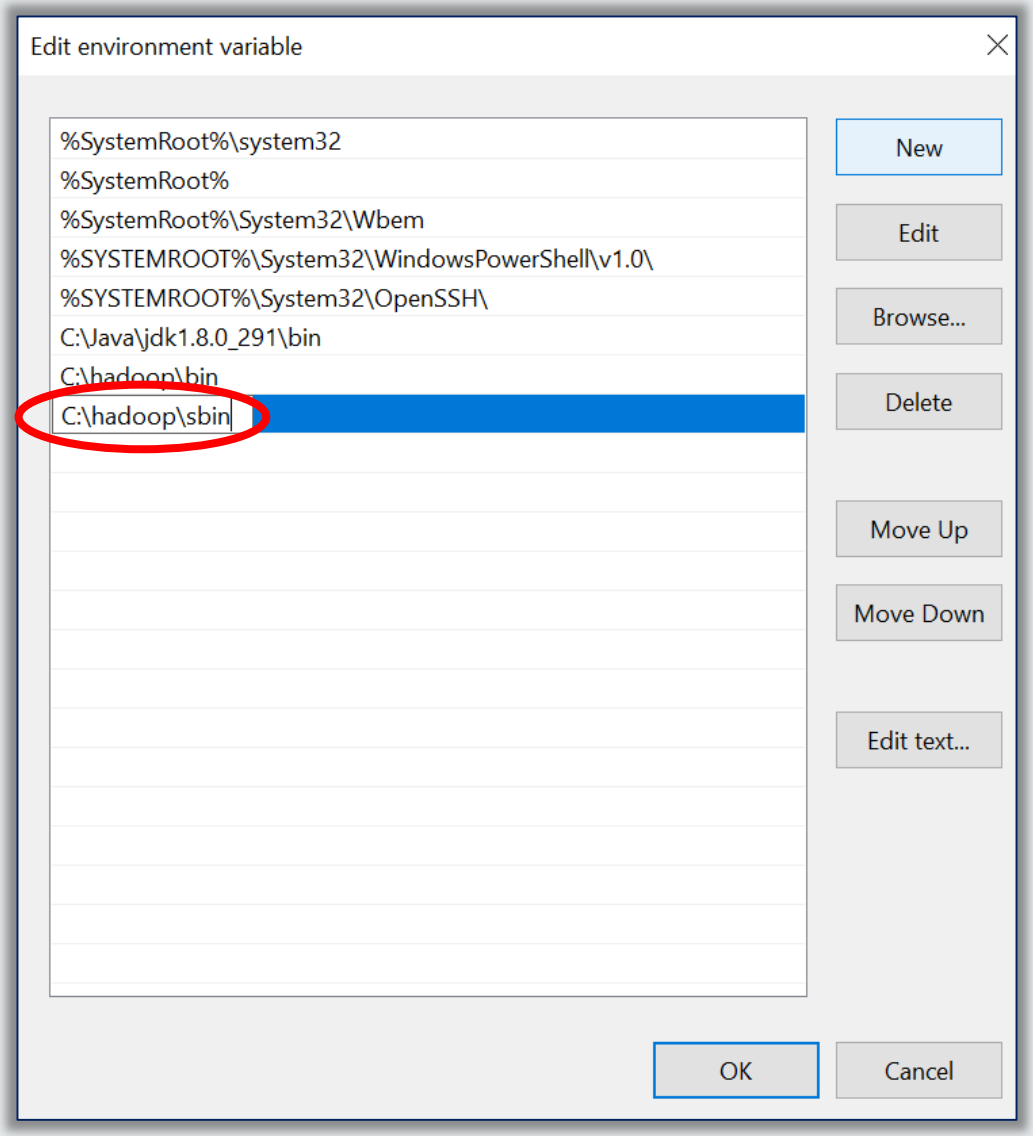
Steps to install Hadoop



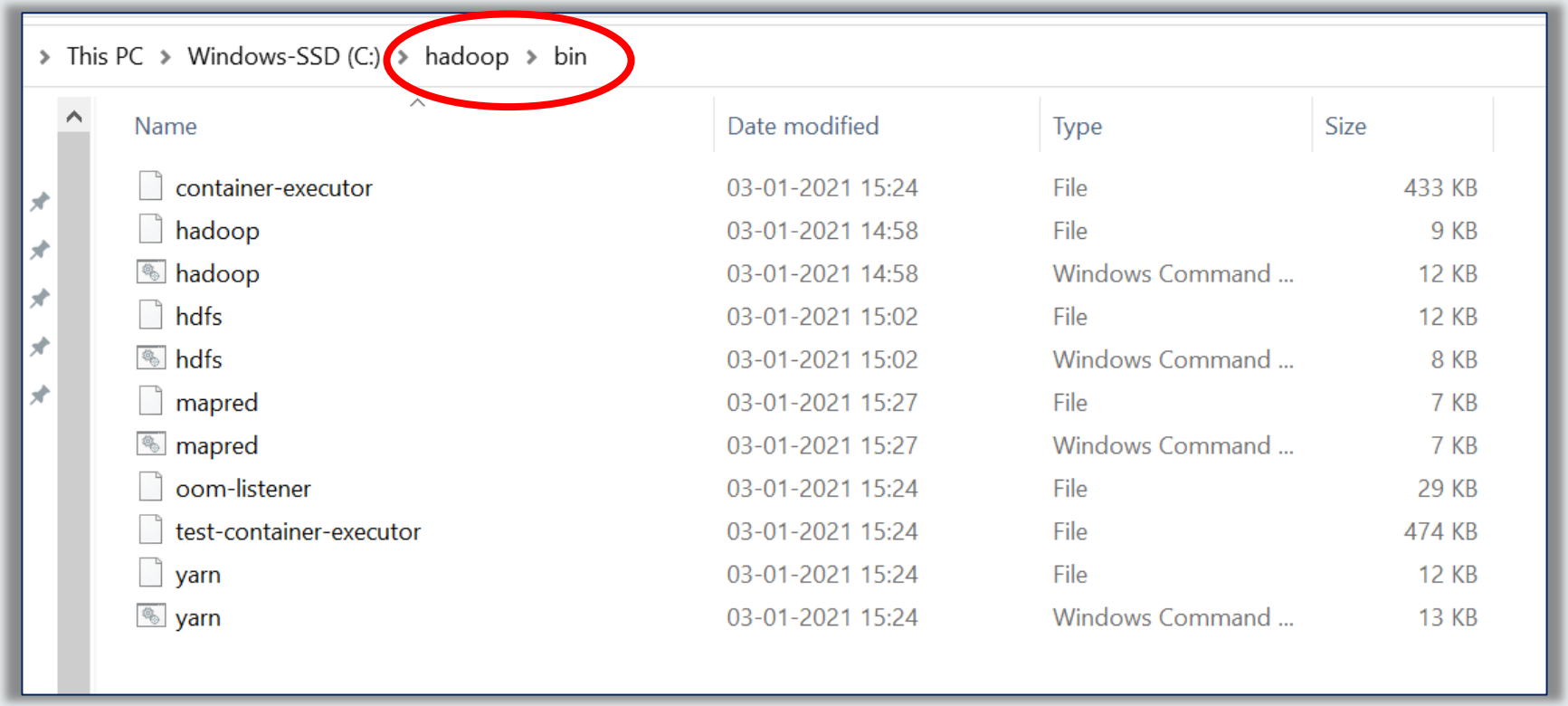
Steps to install Hadoop



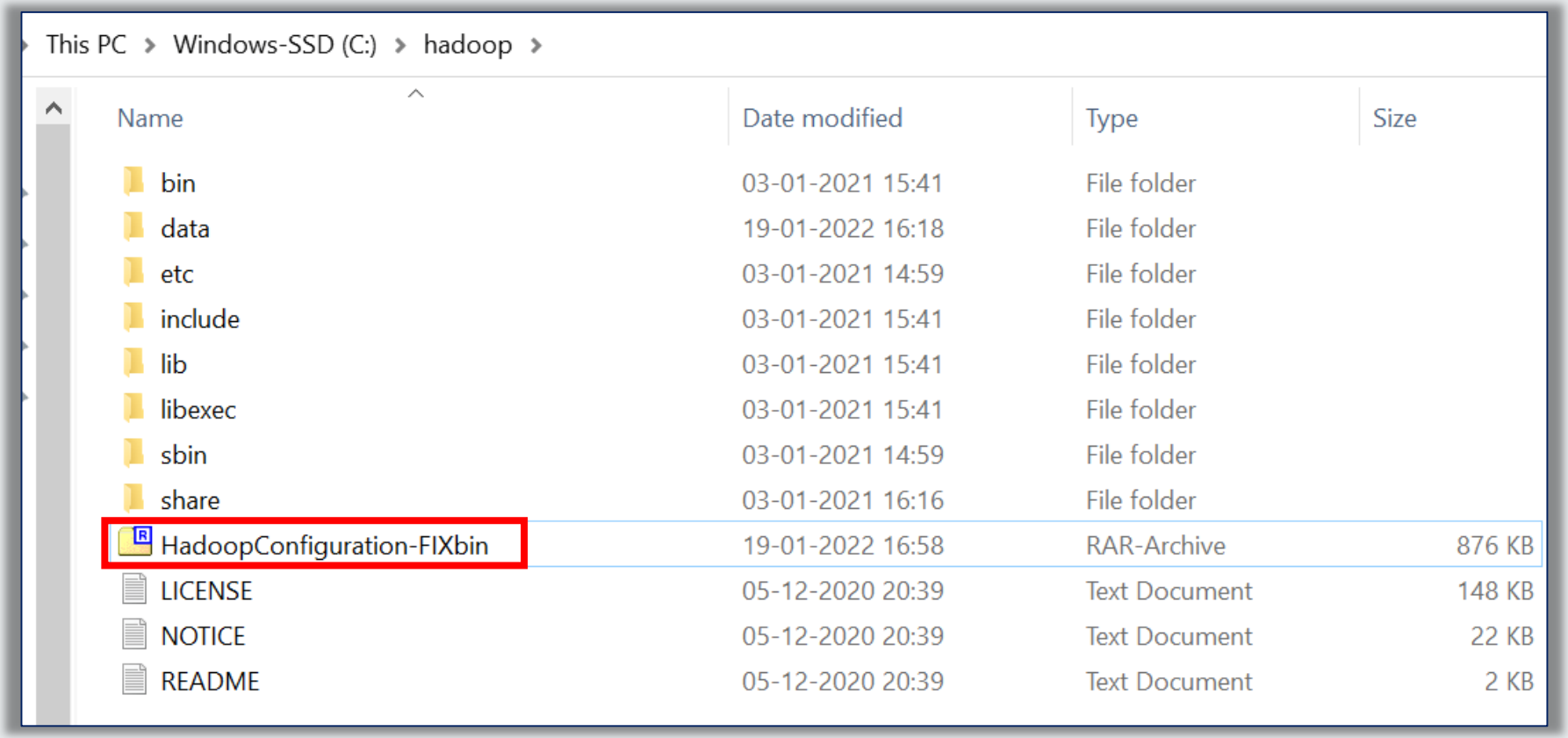
Steps to install Hadoop



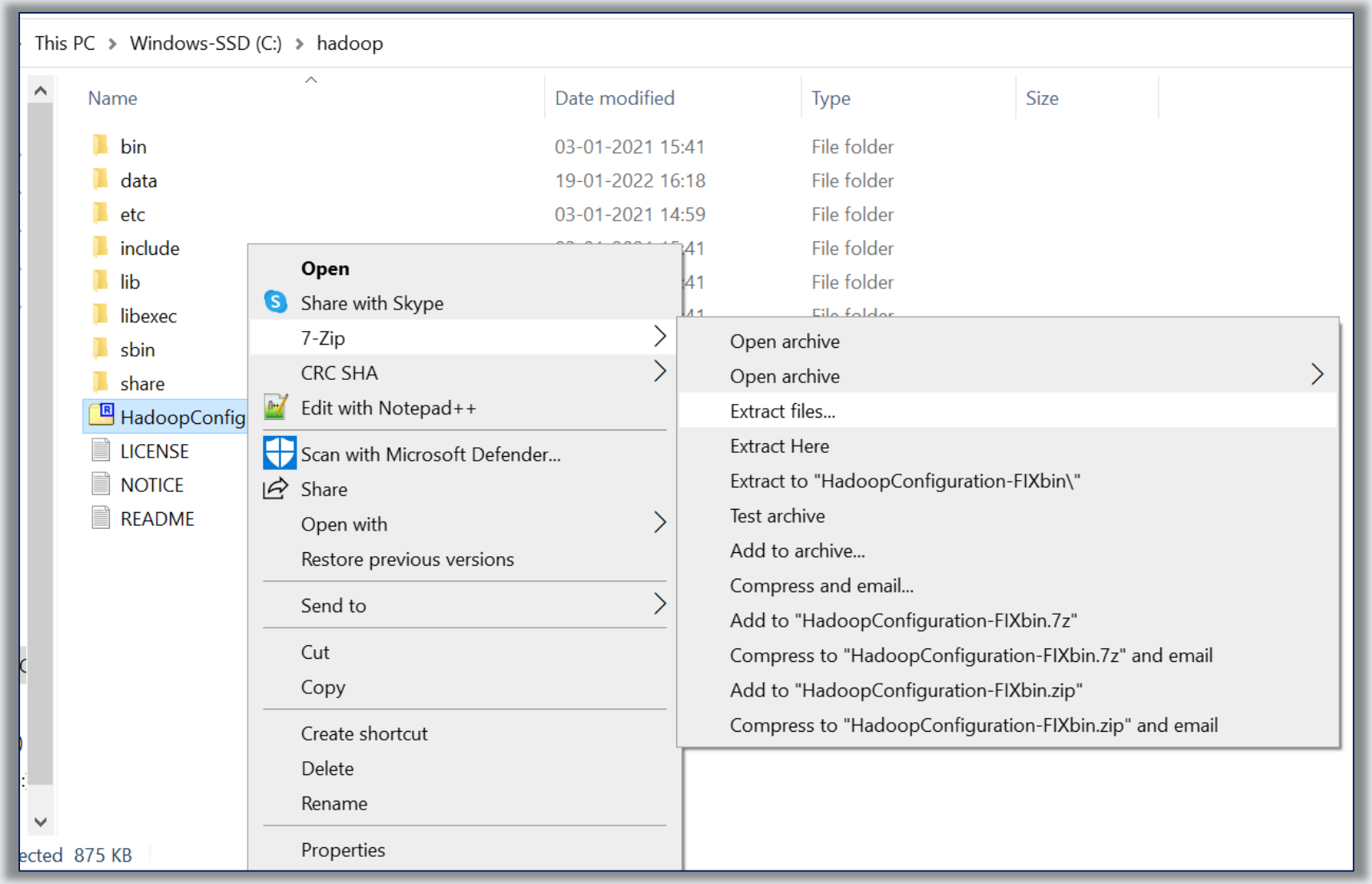
Steps to install Hadoop



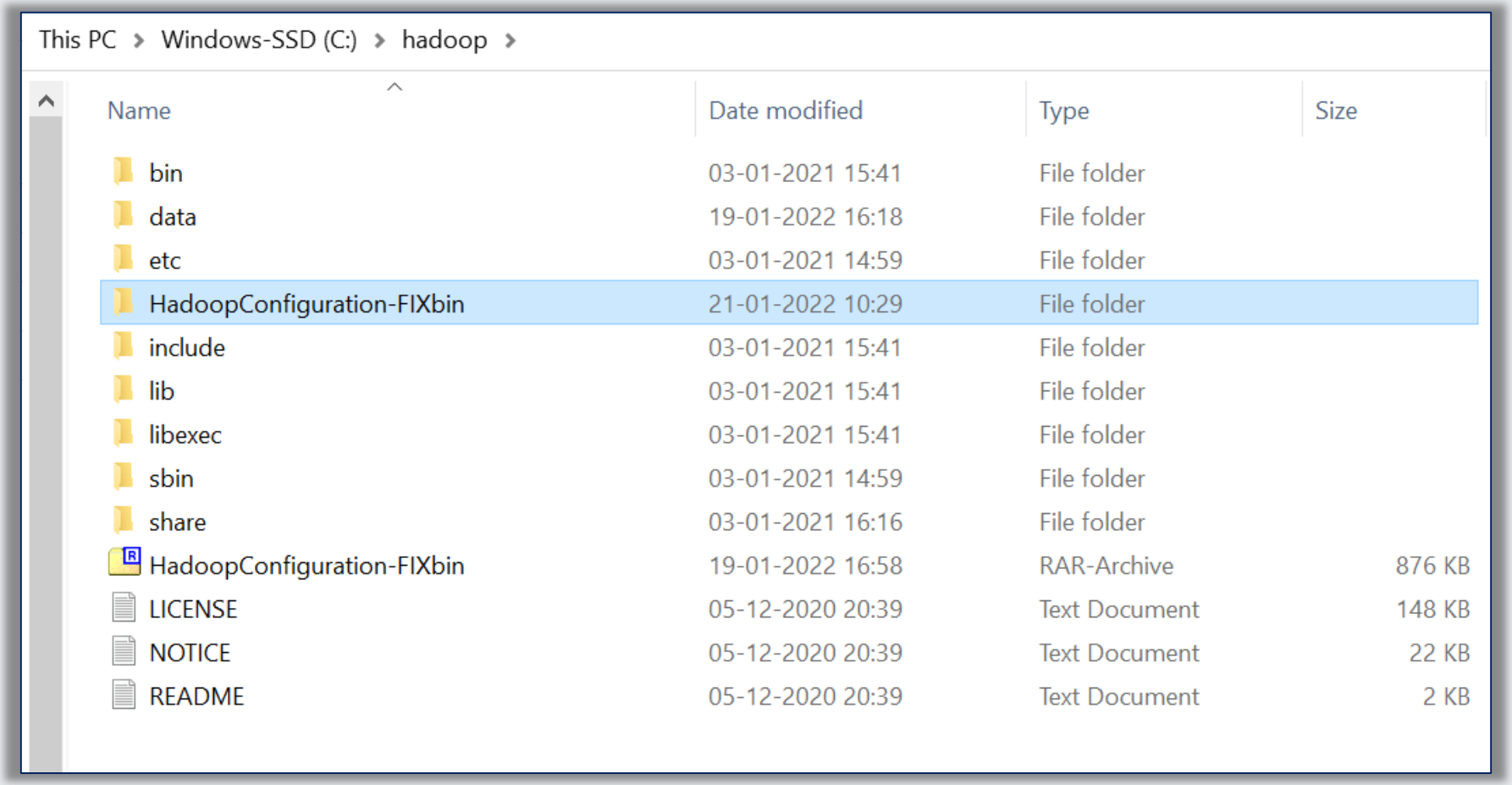
Steps to install Hadoop



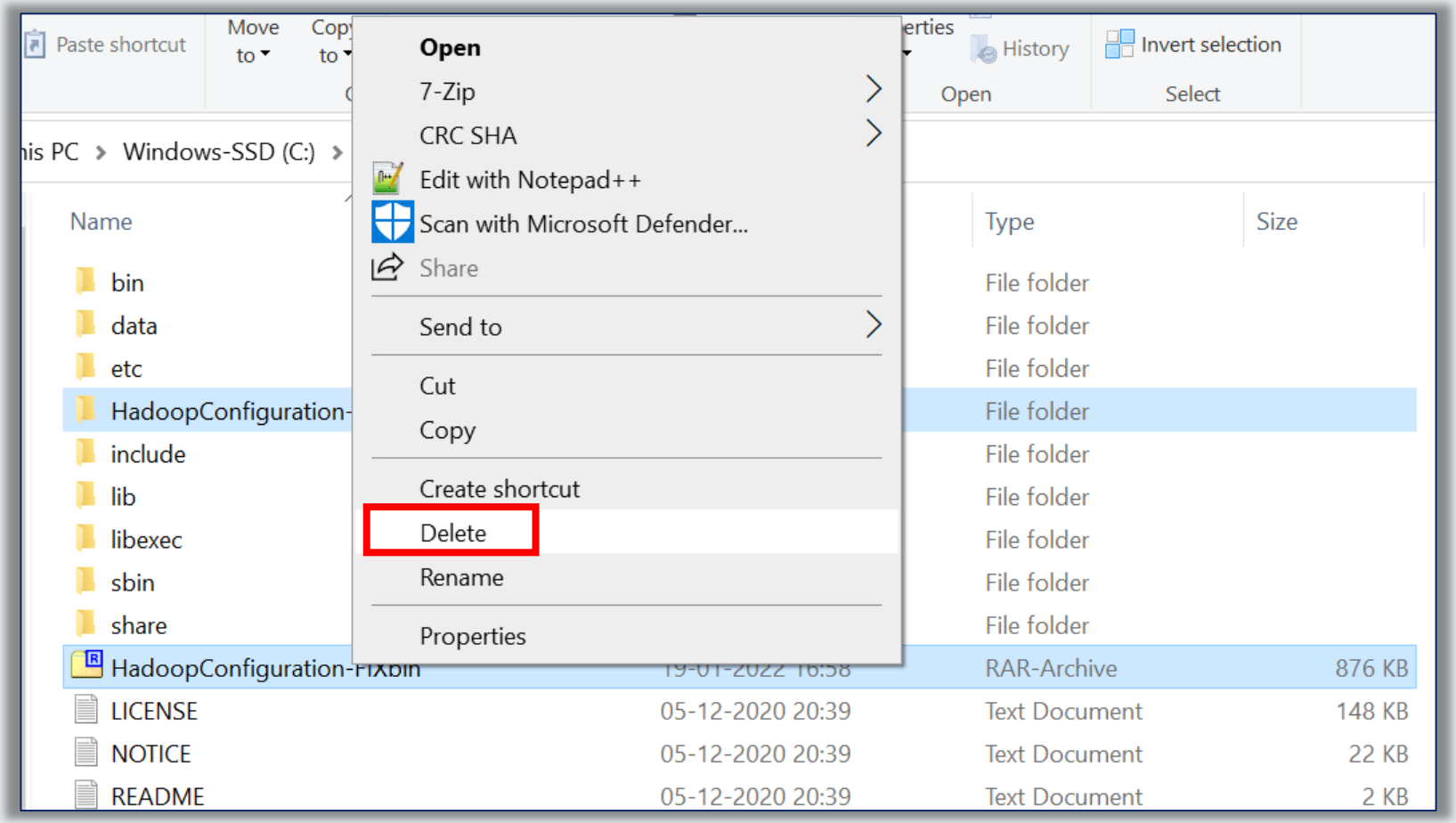
Steps to install Hadoop



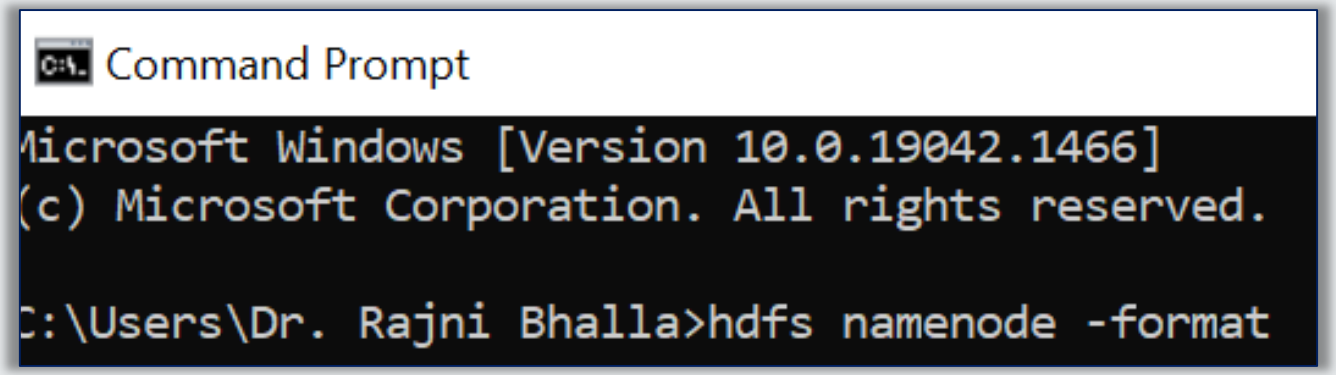
Steps to install Hadoop



Steps to install Hadoop



Steps to install Hadoop



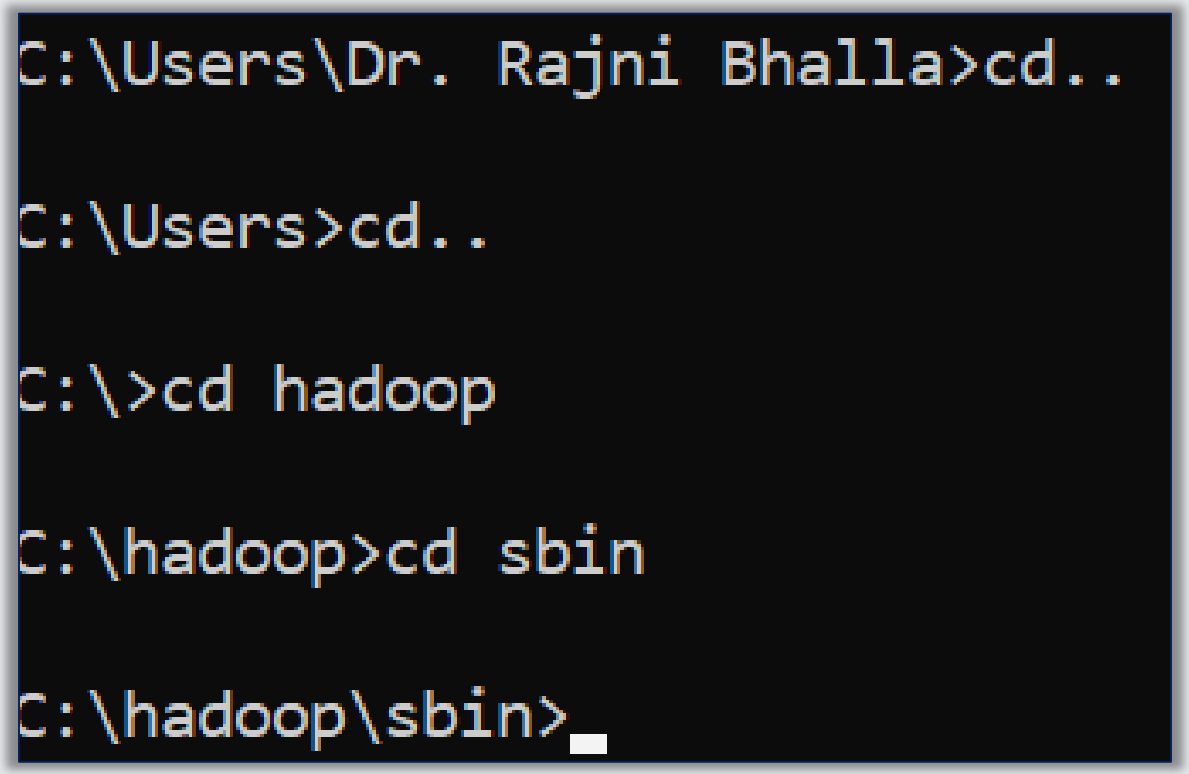
Steps to install Hadoop

```
    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)
eption in thread "Thread-0" java.lang.RuntimeException: com.ctc.wstx.exc.WstxParsingException: Unexpected close tag </configuratic
 [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$1.run(ShutdownHookManager.java:102)
sed by: com.ctc.wstx.exc.WstxParsingException: Unexpected close tag </configuration>; expected </property>.
 [row,col,system-id]: [23,15,"file:/C:/hadoop/etc/hadoop/core-site.xml"]
    at com.ctc.wstx.sr.StreamScanner.constructWfcException(StreamScanner.java:621)
    at com.ctc.wstx.sr.StreamScanner.throwParseError(StreamScanner.java:491)
    at com.ctc.wstx.sr.StreamScanner.throwParseError(StreamScanner.java:475)
    at com.ctc.wstx.sr.BasicStreamReader.reportWrongEndElem(BasicStreamReader.java:3365)
    at com.ctc.wstx.sr.BasicStreamReader.readEndElem(BasicStreamReader.java:3292)
    at com.ctc.wstx.sr.BasicStreamReader.nextFromTree(BasicStreamReader.java:2911)
    at com.ctc.wstx.sr.BasicStreamReader.next(BasicStreamReader.java:1123)
    at org.apache.hadoop.conf.Configuration$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
```

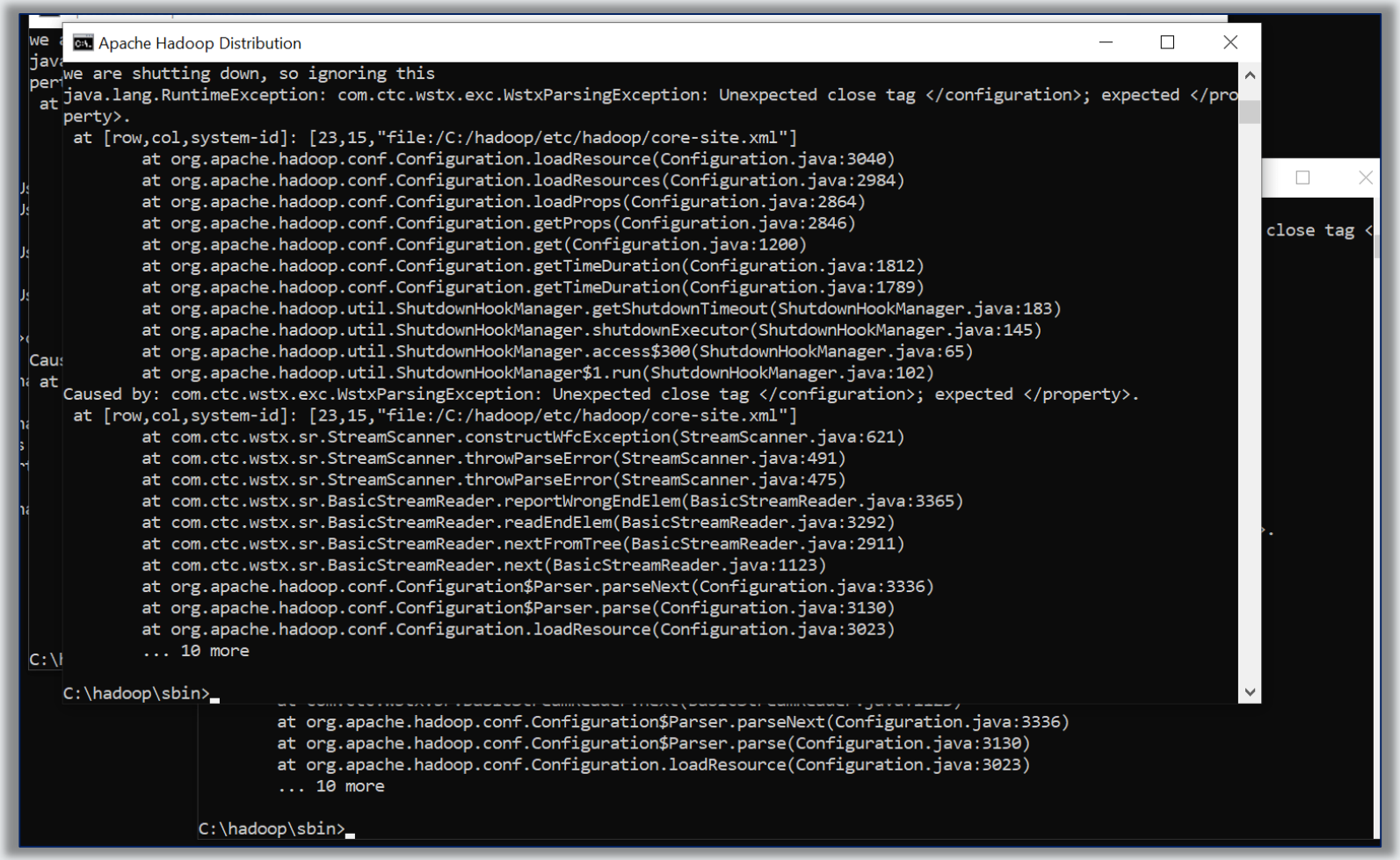
Steps to install Hadoop

```
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-A1IFEVVJ/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-9.4.2
```

Steps to install Hadoop



Steps to install Hadoop



That's all for now...

CMD