



Installation and configuration files

Welcome

How to Install Hadoop on Windows 10

Easy Steps to Install Hadoop



Easy Steps to Install Hadoop on Windows 10



Step 1

- Download Java 8 version

Windows x64

169.46 MB

 [jdk-8u301-windows-x64.exe](#)

- Download latest Hadoop version

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

We suggest the following mirror site for your download:

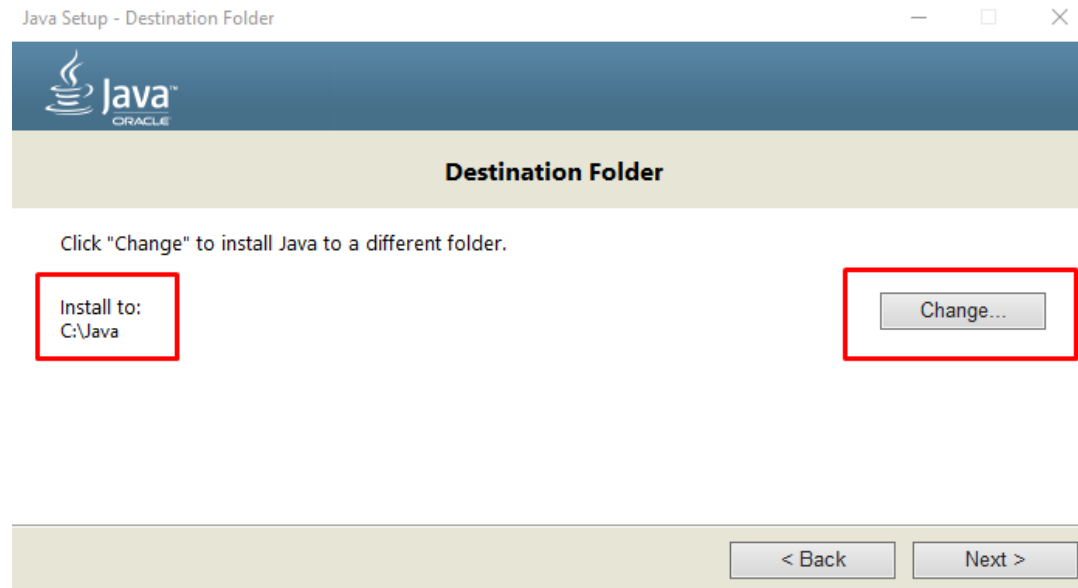
<https://apache.mivzakim.net/hadoop/common/hadoop-3.3.1/hadoop-3.3.1.tar.gz>

Other mirror sites are suggested below.

Easy Steps to Install Hadoop on Windows 10



Step 2 Install Java

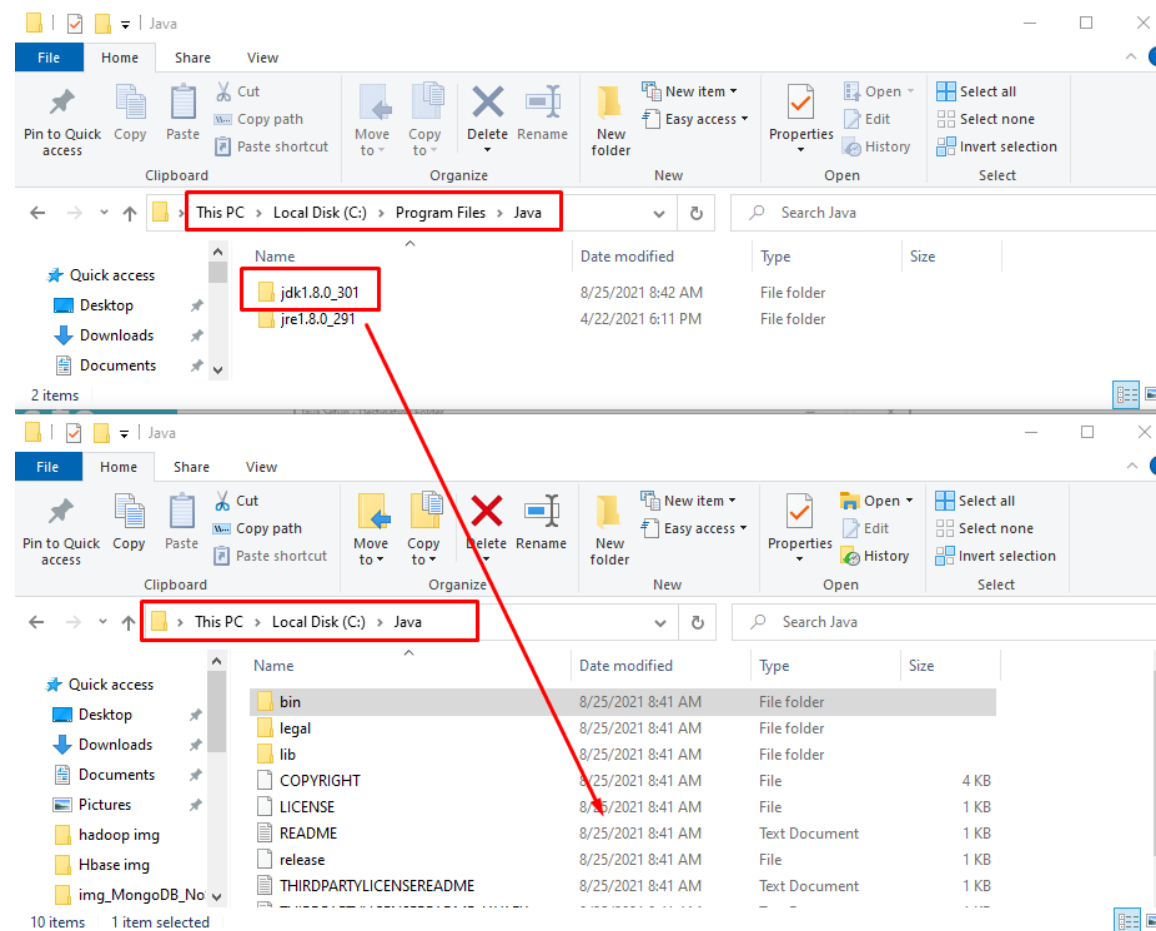


Easy Steps to Install Hadoop on Windows 10



Step 3

Move jdk from C:\Program Files\Java to C:\Java



Easy Steps to Install Hadoop on Windows 10



Step 3 Edit the User Enviroment Variables

Environment Variables

User variables for MarizzaMil

Variable	Value
OneDrive	C:\Users\MarizzaMil\OneDrive
Path	C:\Users\MarizzaMil\AppData\Local\Microsoft\WindowsApps;C:\Pr...
TEMP	C:\Users\MarizzaMil\AppData\Local\Temp
TMP	C:\Users\MarizzaMil\AppData\Local\Temp

New... Edit... Delete

System variables

New User Variable

Variable name: JAVA_HOME

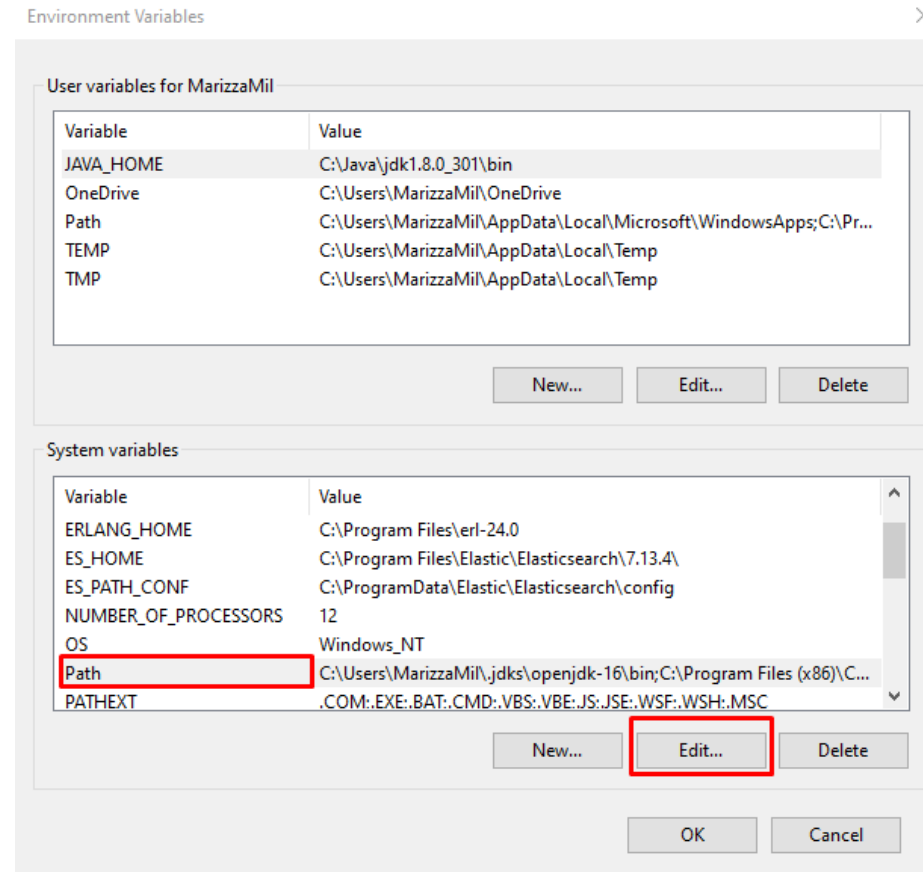
Variable value: C:\Java\jdk1.8.0_301\bin

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

Easy Steps to Install Hadoop on Windows 10



Step 4.1 Edit the System Environment Variables

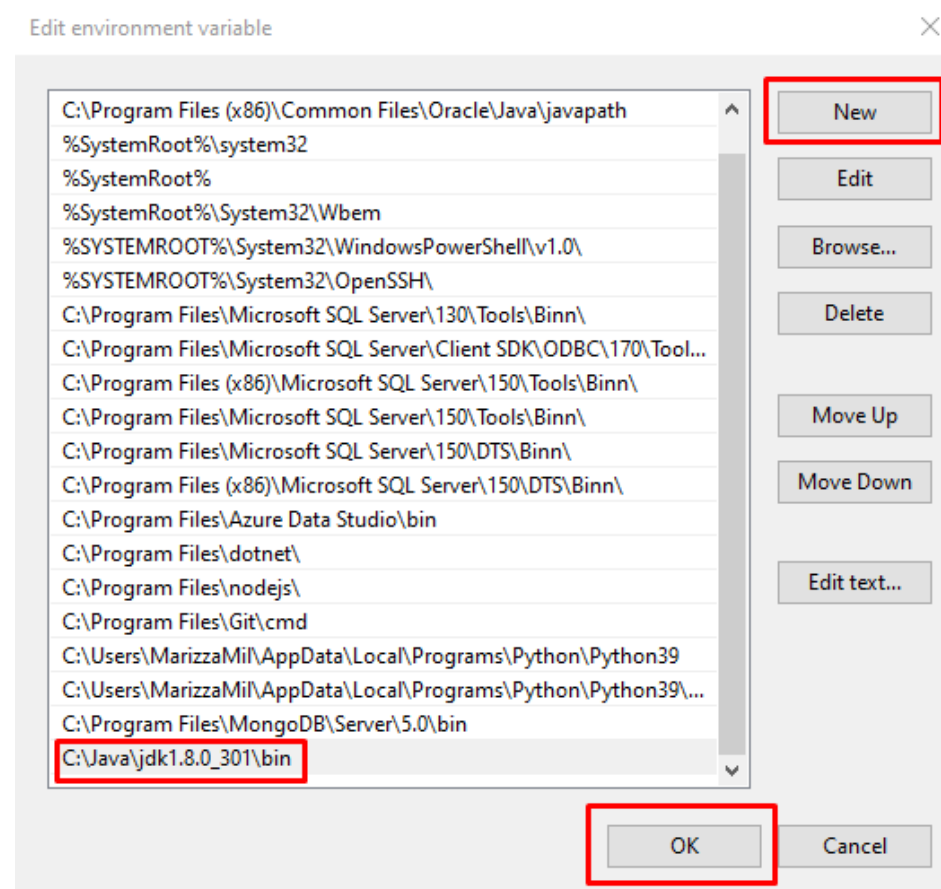


Easy Steps to Install Hadoop on Windows 10



Step 4.2

Edit the System Environment Variables



Easy Steps to Install Hadoop on Windows 10



Step 5

We have successfully installed Java into our local system

Command Prompt

```
Microsoft Windows [Version 10.0.19042.1165]
(c) Microsoft Corporation. All rights reserved.

C:\Users\MarizzaMil>java -version
java version "1.8.0_301"
Java(TM) SE Runtime Environment (build 1.8.0_301-b09)
Java HotSpot(TM) 64-Bit Server VM (build 25.301-b09, mixed mode)

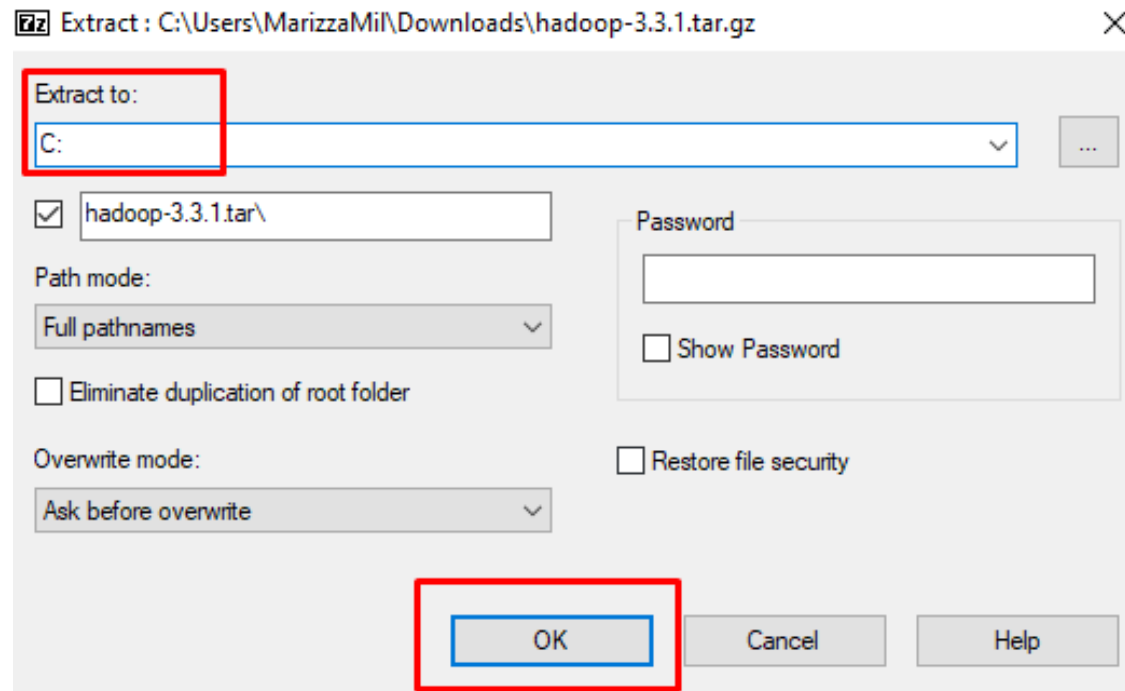
C:\Users\MarizzaMil>
```

Easy Steps to Install Hadoop on Windows 10



Step 6

Extract the tar version of Hadoop that we've downloaded to drive C:\



Easy Steps to Install Hadoop on Windows 10



Step 7

Rename folder "hadoop-3.3.1.tar" to "hadoop"

> Local Disk (C:) > hadoop >				
Name	Date modified	Type	Size	
bin	6/15/2021 8:52 AM	File folder		
etc	6/15/2021 8:15 AM	File folder		
include	6/15/2021 8:52 AM	File folder		
lib	6/15/2021 8:52 AM	File folder		
libexec	6/15/2021 8:52 AM	File folder		
licenses-binary	6/15/2021 8:52 AM	File folder		
sbin	6/15/2021 8:15 AM	File folder		
share	6/15/2021 9:18 AM	File folder		
LICENSE	6/15/2021 8:02 AM	Text Document	15 KB	
LICENSE-binary	6/15/2021 8:02 AM	File	23 KB	
NOTICE	5/21/2021 7:11 PM	Text Document	2 KB	
NOTICE-binary	6/15/2021 8:02 AM	File	29 KB	
README	5/21/2021 7:11 PM	Text Document	1 KB	

Easy Steps to Install Hadoop on Windows 10

Configuration Files



Configuration Files are the files which are located in the extracted tar.gz file in the etc/hadoop/ directory.

Local Disk (C:) > hadoop > etc > hadoop			
Name	Date modified	Type	Size
shellprofile.d	6/15/2021 8:14 AM	File folder	
capacity-scheduler	6/15/2021 8:38 AM	XML Document	9 KB
configuration	6/15/2021 8:39 AM	XSL Stylesheet	2 KB
container-executor.cfg	6/15/2021 8:38 AM	CFG File	3 KB
core-site	6/15/2021 8:14 AM	XML Document	1 KB
hadoop-env	6/15/2021 8:14 AM	Windows Comma...	4 KB
hadoop-env	6/15/2021 8:52 AM	Shell Script	17 KB
hadoop-metrics2.properties	6/15/2021 8:14 AM	PROPERTIES File	4 KB
hadoop-policy	6/15/2021 8:14 AM	XML Document	12 KB
hadoop-user-functions.sh.example	6/15/2021 8:14 AM	EXAMPLE File	4 KB
hdfs-rbf-site	6/15/2021 8:21 AM	XML Document	1 KB
hdfs-site	6/15/2021 8:18 AM	XML Document	1 KB
httpfs-env	6/15/2021 8:20 AM	Shell Script	2 KB
httpfs-log4j.properties	6/15/2021 8:20 AM	PROPERTIES File	2 KB
httpfs-site	6/15/2021 8:20 AM	XML Document	1 KB
kms-acls	6/15/2021 8:15 AM	XML Document	4 KB
kms-env	6/15/2021 8:15 AM	Shell Script	2 KB
kms-log4j.properties	6/15/2021 8:15 AM	PROPERTIES File	2 KB
kms-site	6/15/2021 8:15 AM	XML Document	1 KB
log4j.properties	6/15/2021 8:14 AM	PROPERTIES File	14 KB
mapred-env	6/15/2021 8:39 AM	Windows Comma...	1 KB
mapred-env	6/15/2021 8:39 AM	Shell Script	2 KB
mapred-queues.xml.template	6/15/2021 8:39 AM	TEMPLATE File	5 KB
mapred-site	6/15/2021 8:39 AM	XML Document	1 KB
ssl-client.xml.example	6/15/2021 8:14 AM	EXAMPLE File	3 KB
ssl-server.xml.example	6/15/2021 8:14 AM	EXAMPLE File	3 KB
user_ec_policies.xml.template	6/15/2021 8:18 AM	TEMPLATE File	3 KB
workers	6/15/2021 8:14 AM	File	1 KB
yarn-env	6/15/2021 8:38 AM	Windows Comma...	3 KB
yarn-env	6/15/2021 8:38 AM	Shell Script	7 KB
yarnservice-log4j.properties	6/15/2021 8:38 AM	PROPERTIES File	3 KB
yarn-site	6/15/2021 8:38 AM	XML Document	1 KB

Easy Steps to Install Hadoop on Windows 10

Configuration Files



Step 8

Let's edit these four files:

- core-site.xml
- hdfs-site.xml
- mapred-site.xml
- yarn-site.xml

Local Disk (C:) > hadoop > etc > hadoop			
Name	Date modified	Type	Size
shellprofile.d	6/15/2021 8:14 AM	File folder	
capacity-scheduler	6/15/2021 8:38 AM	XML Document	9 KB
configuration	6/15/2021 8:39 AM	XSL Stylesheet	2 KB
container-executor.cfg	6/15/2021 8:38 AM	CFG File	3 KB
core-site	6/15/2021 8:14 AM	XML Document	1 KB
hadoop-env	6/15/2021 8:14 AM	Windows Comma...	4 KB
hadoop-env	6/15/2021 8:52 AM	Shell Script	17 KB
hadoop-metrics2.properties	6/15/2021 8:14 AM	PROPERTIES File	4 KB
hadoop-policy	6/15/2021 8:14 AM	XML Document	12 KB
hadoop-user-functions.sh.example	6/15/2021 8:14 AM	EXAMPLE File	4 KB
hdfs-rbf-site	6/15/2021 8:21 AM	XML Document	1 KB
hdfs-site	6/15/2021 8:18 AM	XML Document	1 KB
https-env	6/15/2021 8:20 AM	Shell Script	2 KB
https-log4j.properties	6/15/2021 8:20 AM	PROPERTIES File	2 KB
https-site	6/15/2021 8:20 AM	XML Document	1 KB
kms-acls	6/15/2021 8:15 AM	XML Document	4 KB
kms-env	6/15/2021 8:15 AM	Shell Script	2 KB
kms-log4j.properties	6/15/2021 8:15 AM	PROPERTIES File	2 KB
kms-site	6/15/2021 8:15 AM	XML Document	1 KB
log4j.properties	6/15/2021 8:14 AM	PROPERTIES File	14 KB
mapred-env	6/15/2021 8:39 AM	Windows Comma...	1 KB
mapred-env	6/15/2021 8:39 AM	Shell Script	2 KB
mapred-queues.xml.template	6/15/2021 8:39 AM	TEMPLATE File	5 KB
mapred-site	6/15/2021 8:39 AM	XML Document	1 KB
ssl-client.xml.example	6/15/2021 8:14 AM	EXAMPLE File	3 KB
ssl-server.xml.example	6/15/2021 8:14 AM	EXAMPLE File	3 KB
user_ec_policies.xml.template	6/15/2021 8:18 AM	TEMPLATE File	3 KB
workers	6/15/2021 8:14 AM	File	1 KB
yarn-env	6/15/2021 8:38 AM	Windows Comma...	3 KB
yarn-env	6/15/2021 8:38 AM	Shell Script	7 KB
yarnservice-log4j.properties	6/15/2021 8:38 AM	PROPERTIES File	3 KB
yarn-site	6/15/2021 8:38 AM	XML Document	1 KB

Easy Steps to Install Hadoop on Windows 10

Configuration Files



CORE-SITE.XML

It is one of the important configuration files which is required for runtime environment settings of a Hadoop cluster.

It informs Hadoop daemons where the NAMENODE runs in the cluster.

It also informs the Name Node as to which IP and ports it should bind.

Easy Steps to Install Hadoop on Windows 10

Configuration Files



Step 8.1

Edit CORE-SITE.XML

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

Insert
code



Easy Steps to Install Hadoop on Windows 10

Configuration Files



MAPRED-SITE.XML

It is one of the important configuration files which is required for runtime environment settings of a Hadoop.

It contains the configuration settings for MapReduce.

In this file, we specify a framework name for MapReduce, by setting the `MapReduce.framework.name`.

Easy Steps to Install Hadoop on Windows 10

Configuration Files



Step 8.2

Edit MAPRED-SITE.XML

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

Insert
code

Easy Steps to Install Hadoop on Windows 10

Configuration Files



YARN-SITE.XML

One more important file

This file contains the configuration settings related to YARN.

For example, it contains settings for **Node Manager, Resource Manager, Containers, and Application Master.**

Easy Steps to Install Hadoop on Windows 10

Configuration Files



Step 8.3

Edit 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>
```

Insert
code

```
<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>
```

Easy Steps to Install Hadoop on Windows 10

Configuration Files



HDFS-SITE.XML

It is one of the important configuration files which is required for runtime environment settings of a Hadoop.

It contains the configuration settings for NAMENODE, DATANODE, SECONDARYNODE.

It is used to specify default block replication.

The actual number of replications can also be specified when the file is created

Easy Steps to Install Hadoop on Windows 10

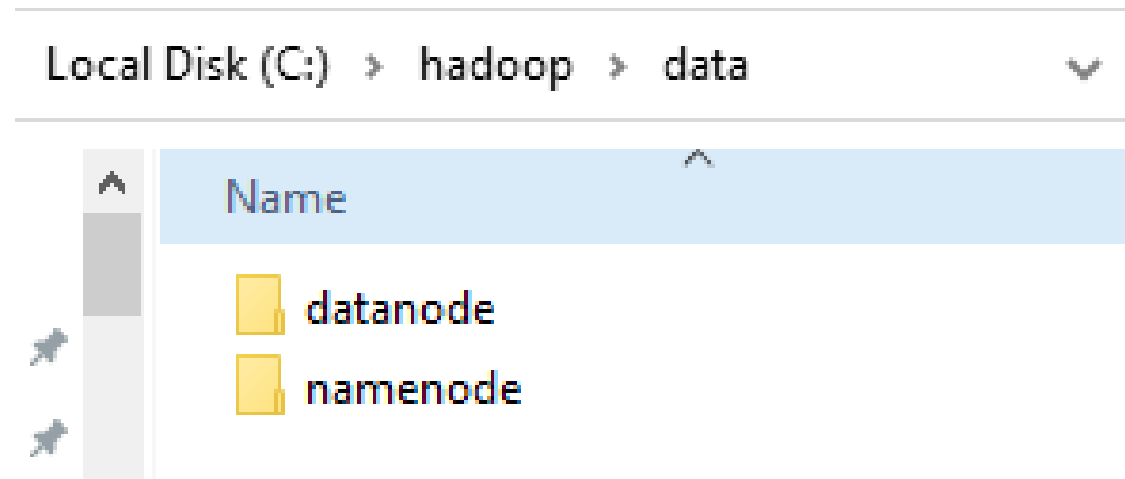
Configuration Files



Step 8.4.1 HDFS-SITE.XML

Inside folder "hadoop" create a new folder with the name "data"

Inside "data" you need to create two more new folders with a "datanode" and "namenode"



Easy Steps to Install Hadoop on Windows 10

Configuration Files



Step 8.4.2 HDFS-SITE.XML

```
<?xml version="1.0" encoding="UTF-8"?>
<?xml-stylesheet type="text/xsl" href="configuration.xsl"?>
<!--
...
-->
```

```
<!-- 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>
```

Insert
code

Easy Steps to Install Hadoop on Windows 10

Configuration Files



Step 9

Let's edit **hadoop-env.cmd** file

Insert
code

...

@rem The java implementation to use. Required.

set JAVA_HOME=C:\Java\jdk1.8.0_301

@rem The jsvc implementation to use. Jsvc is required to run secure datanodes.

@rem set JSVC_HOME=%JSVC_HOME%

Easy Steps to Install Hadoop on Windows 10



Step 10

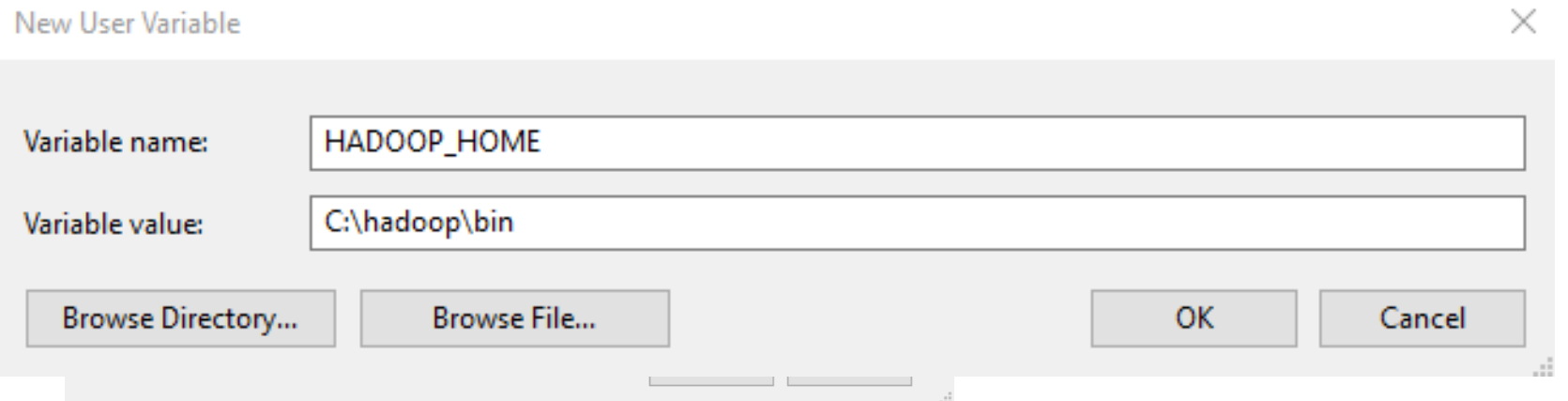
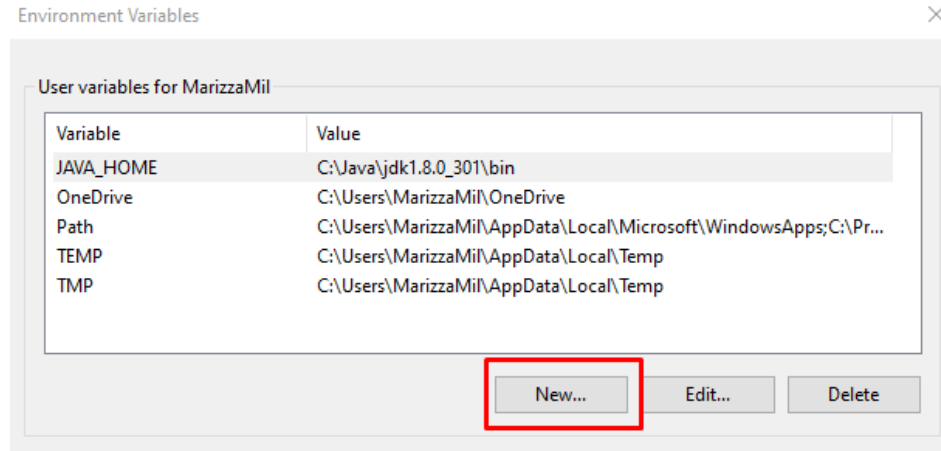
Go to Enviroment Variables and set Home and Path for Hadoop

Easy Steps to Install Hadoop on Windows 10



Step 10.1

Insert to the User Environment Variables
HADOOP_HOME

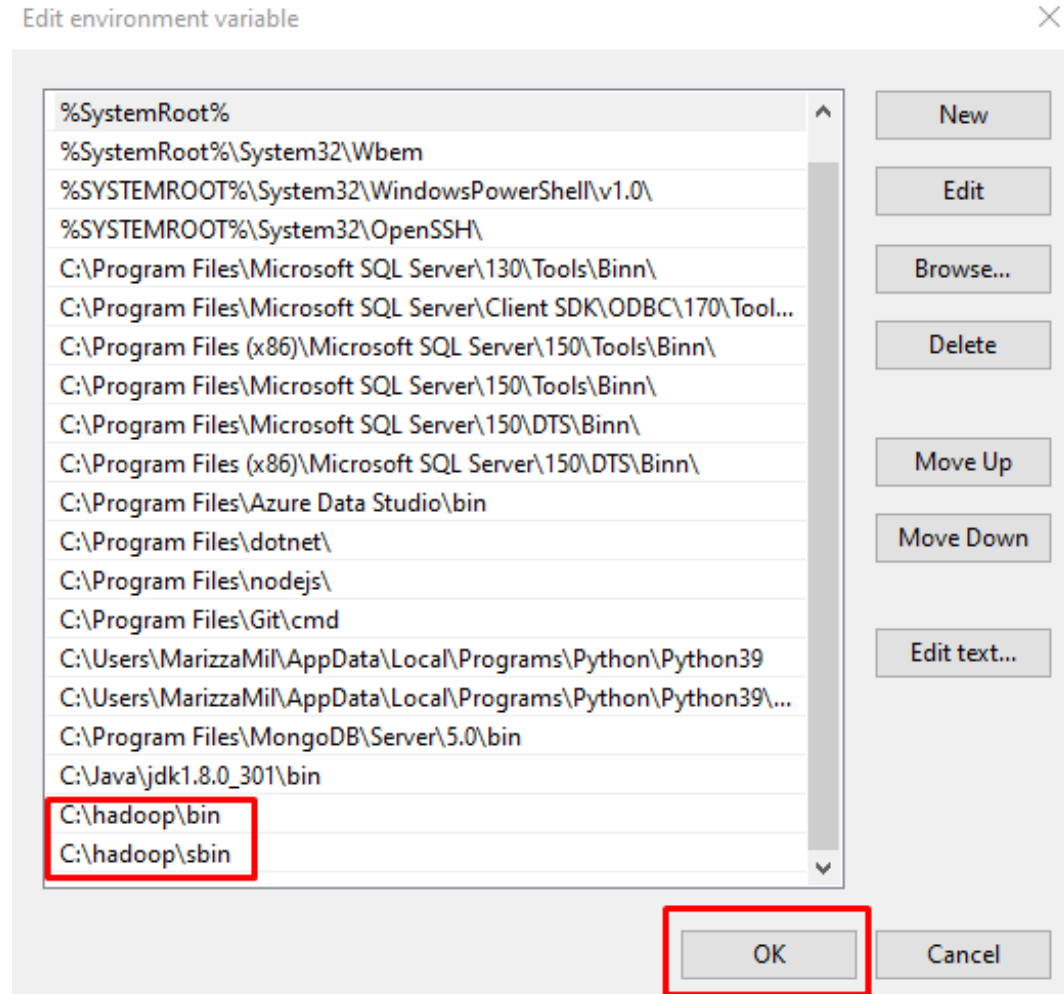


Easy Steps to Install Hadoop on Windows 10



Step 10.2

Edit the System Environment Variables



Easy Steps to Install Hadoop on Windows 10



Step 11

You need download configuration files and replace that particular file with your bin folder in your Hadoop

› Local Disk (C:) › hadoop

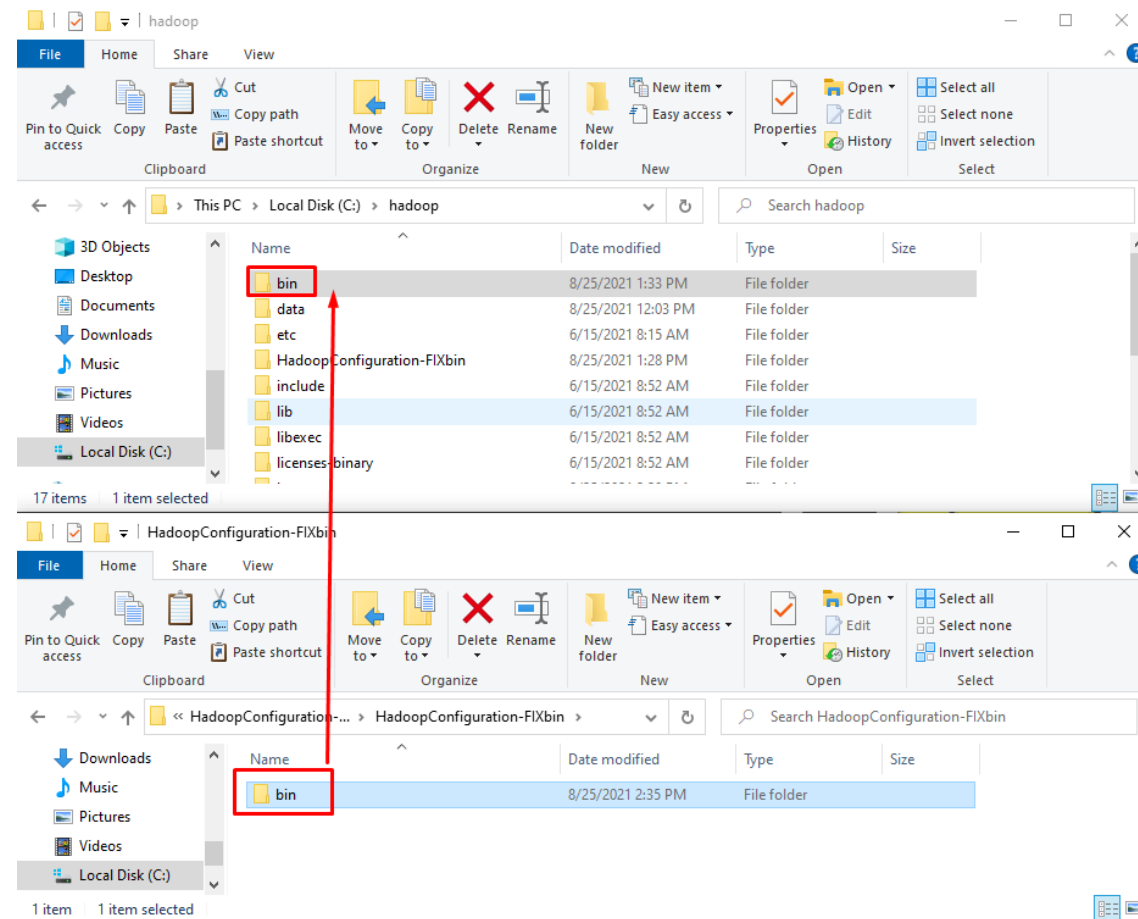
Name	Date modified	Type	Size
etc	6/15/2021 8:15 AM	File folder	
sbin	6/15/2021 8:15 AM	File folder	
lib	6/15/2021 8:52 AM	File folder	
licenses-binary	6/15/2021 8:52 AM	File folder	
bin	6/15/2021 8:52 AM	File folder	
include	6/15/2021 8:52 AM	File folder	
libexec	6/15/2021 8:52 AM	File folder	
share	6/15/2021 9:18 AM	File folder	
data	8/25/2021 12:03 PM	File folder	
NOTICE	5/21/2021 7:11 PM	Text Document	2 KB
README	5/21/2021 7:11 PM	Text Document	1 KB
LICENSE	6/15/2021 8:02 AM	Text Document	15 KB
LICENSE-binary	6/15/2021 8:02 AM	File	23 KB
NOTICE-binary	6/15/2021 8:02 AM	File	29 KB
HadoopConfiguration-FIXbin.rar	8/25/2021 1:09 PM	RAR File	876 KB

Easy Steps to Install Hadoop on Windows 10



Step 12

Replace folder bin from C:\hadoop to folder bin with configuration that we've downloaded



Hadoop
Installation and
Configuration
Completed



Testing if
installation
was successful

Run Cmd as administrator

```
hdfs namenode -format
```

Open new Cmd as administrator

```
cd\  
cd hadoop\sbin  
start-dfs  
start-yarn
```



Testing if
installation
was successful



Open in your browser <http://localhost:9870/>

Hadoop	Overview	Datanodes	Datanode Volume Failures	Snapshot	Startup Progress	Utilities ▾
Overview 'localhost:9000' (✓active)						
Started:	Wed Aug 25 14:54:36 +0300 2021					
Version:	3.3.1, ra3b9c37a397ad4188041dd80621bdeefc46885f2					
Compiled:	Tue Jun 15 08:13:00 +0300 2021 by ubuntu from (HEAD detached at release-3.3.1-RC3)					
Cluster ID:	CID-752265e6-3f27-41e2-8afd-4d7f8a2001e4					
Block Pool ID:	BP-415388049-192.168.56.1-1629888967293					