Installing Hadoop on Windows 11 – Step-by-Step Guide

# 1. Introduction

Purpose: This guide walks you through installing Hadoop on Windows 11 in a single-node setup.  
Prerequisites:  
- Windows 11 OS  
- Administrative rights  
- Basic familiarity with command-line environment

# 2. Pre-installation Requirements

## 2.1 Install Java (JDK 8)

- Download JDK 8 (e.g., from Oracle).  
- Install and set JAVA\_HOME environment variable to the JDK installation path.  
- Add %JAVA\_HOME%\bin to the system Path.  
- Verify setup with 'java -version' in Command Prompt.

## 2.2 Install Required Tools

- Download and install 7‑Zip (for file extraction).  
- Install additional dependencies if required (e.g., Visual C++ Redistributables).

# 3. Downloading and Extracting Hadoop

- Download Hadoop (e.g., version 3.2.4 or another stable release).  
- Extract using 7‑Zip into C:\hadoop.  
- Rename folder to simplify path, e.g., C:\hadoop.

# 4. Configuration of Hadoop Environment

## 4.1 Set Environment Variables

- Define HADOOP\_HOME = C:\hadoop.  
- Add C:\hadoop\bin and C:\hadoop\sbin to the system Path.

## 4.2 Edit Configuration Files

Configuration files are located in C:\hadoop\etc\hadoop

core-site.xml:  
<configuration>  
 <property>  
 <name>fs.defaultFS</name>  
 <value>hdfs://localhost:9000</value>  
 </property>  
</configuration>

hdfs-site.xml:  
<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>

mapred-site.xml:  
<configuration>  
 <property>  
 <name>mapreduce.framework.name</name>  
 <value>yarn</value>  
 </property>  
</configuration>

yarn-site.xml:  
<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>

## 4.3 Create Data Directories

Create C:\hadoop\data\namenode and C:\hadoop\data\datanode manually.

# 5. Launching Hadoop

- Open Command Prompt as Administrator.  
- Format the NameNode:  
 hdfs namenode -format  
- Start Hadoop services:  
 cd C:\hadoop\sbin  
 start-dfs.cmd  
 start-yarn.cmd  
- Use 'jps' to confirm Java processes running (NameNode, DataNode, ResourceManager).

# 6. Verifying the Installation

- Open a browser to check Hadoop Web Interfaces:  
 NameNode UI: http://localhost:9870  
 ResourceManager UI: http://localhost:8088  
- In Command Prompt, verify Hadoop commands:  
 hadoop fs -ls /

# 7. Testing with basic HDFS commands

- Create and manage files in HDFS:  
 hadoop fs -mkdir /test  
 hadoop fs -put somefile.txt /test  
 hadoop fs -cat /test/somefile.txt

# 8. Final Checks and Troubleshooting

- Ensure no path conflicts or missing dependencies.  
- Check log files if UI or services don’t come up.  
- Revisit environment variables or .xml configuration files for typos.