```
--Install Jdk1.8 or above
--Create a folder "E:\hadoop-env"
--Download Hadoop binaries
      link: https://www.apache.org/dyn/closer.cgi/hadoop/common/hadoop-
3.2.1/hadoop-3.2.1.tar.gz
-- Extract hadoop-3.2.1 and copy it in E:\hadoop-env

--Download GitHub repository: https://github.com/cdarlint/winutils
--Extract winutils-master.zip and find winutils-master\hadoop-3.2.1\bin
--Copy all files from winutils-master\hadoop-3.2.1\bin to E:\hadoop-
env/hadoop-3.2.1\bin

--Environment Variables
      1. JAVA_HOME: JDK installation folder path
      2. HADOOP_HOME: Hadoop installation folder path
      3. classpath jre/bin
      4. path: as jdk/bin homefolder
      5. path as %JAVA_HOME\bin
      6. path as %HADOOP_HOME%\bin

--Check java path open PowerShell-> hadoop -version
--Replace Progra~1 insted of Program files if JAVA_HOME is incorrectly
set error
--HDFS site configuration (create following folders)
      1. E:\hadoop-env\hadoop-3.2.1\data\dfs\namenode
      2. E:\hadoop-env\hadoop-3.2.1\data\dfs\datanode
--Configuring Hadoop cluster
      1. E:\hadoop-3.2.1\etc\hadoop\hdfs-site.xml
            <property>
            <name>dfs.replication</name>
            <value>1</value>
            </property>
            <property>
            <name>dfs.namenode.name.dir</name>
            <value>file:///E:/hadoop-env/hadoop-
3.2.1/data/dfs/namenode</value>
            </property>
            <property>
            <name>dfs.datanode.data.dir</name>
            <value>file:///E:/hadoop-env/hadoop-
3.2.1/data/dfs/datanode</value>
            </property>

      2. E:\hadoop-3.2.1\etc\hadoop\core-site.xml
            <property>
            <name>fs.default.name</name>
            <value>hdfs://localhost:9820</value>
            </property>

      3. E:\hadoop-3.2.1\etc\hadoop\mapred-site.xml
            <property>
            <name>mapreduce.framework.name</name>
            <value>yarn</value>
            <description>MapReduce framework name</description>
```

```
            </property>

    4. E:\hadoop-3.2.1\etc\hadoop\yarn-site.xml
            <property>
            <name>yarn.nodemanager.aux-services</name>
            <value>mapreduce_shuffle</value>
            <description>Yarn Node Manager Aux Service</description>
            </property>
```

-- Download hadoop-hdfs-3.2.1.jar file from
(https://github.com/FahaoTang/big-data/blob/master/hadoop-hdfs-3.2.1.jar)
-- Make 2 copies and rename one file to hadoop-hdfs-3.2.1.bak
-- Copy the downloaded hadoop-hdfs-3.2.1.jar and hadoop-hdfs-3.2.1.bak to
folder E:\hadoop-env\hadoop-3.2.1\share\hadoop\hdfs

-- Open PowerShell as Admin

-- Formatting Name node
        hdfs namenode -format

-- Open PowerShell open location E:\hadoop-env\hadoop-3.2.1\sbin

-- Starting Hadoop services
        ./start-dfs.cmd

-- Starting Yarn services
        ./start-yarn.cmd

-- use jps command to check running servicess

        then you cand find following servicess is in running
            14560 DataNode
            4960 ResourceManager
            5936 NameNode
            768 NodeManager
            14636 Jps

-- Hadoop Web UI
    1. Name node web page: http://localhost:9870/dfshealth.html
    2. Data node web page: http://localhost:9864/datanode.html
    3. Yarn web page: http://localhost:8088/cluster

--Load Data to HDFS
        hadoop fs -put c:\testdata.txt \input

-- Execute job
        hadoop jar C:\Wordcount_Mapreduce\dist\Wordcount_Mapreduce.jar
/input/testdata.txt /output

-- Stop Hadoop services

```
        ./stop-dfs.cmd

-- Stop Yarn services
        ./stop-yarn.cmd
```