**Steps to install Hadoop**

Open terminal

**Step 1.**

**First install java**

$ sudo apt install openjdk-8-jdk

To check jdk

$cd /usr/lib/jvm/

It enters into folder /lib/usr/jvm/

$ls

It shows the folders in that jdk is shown means jdk installed successfully.

Give cd to return to home

**Step 2.**

**Configure bashrc file**

To enter into bashrc file

$sudo nano .bashrc

The codings to add in bashrc file is available in website

Goto browser search the website codewitharjun.medium.com

export JAVA\_HOME=/usr/lib/jvm/java-8-openjdk-amd64   
export PATH=$PATH:/usr/lib/jvm/java-8-openjdk-amd64/bin   
export HADOOP\_HOME=~/hadoop-3.2.3/   
export PATH=$PATH:$HADOOP\_HOME/bin   
export PATH=$PATH:$HADOOP\_HOME/sbin   
export HADOOP\_MAPRED\_HOME=$HADOOP\_HOME   
export YARN\_HOME=$HADOOP\_HOME   
export HADOOP\_CONF\_DIR=$HADOOP\_HOME/etc/hadoop   
export HADOOP\_COMMON\_LIB\_NATIVE\_DIR=$HADOOP\_HOME/lib/native   
export HADOOP\_OPTS="-Djava.library.path=$HADOOP\_HOME/lib/native"   
export HADOOP\_STREAMING=$HADOOP\_HOME/share/hadoop/tools/lib/hadoop-streaming-3.2.3.jar  
export HADOOP\_LOG\_DIR=$HADOOP\_HOME/logs   
export PDSH\_RCMD\_TYPE=ssh

Copy the .bashrc content and past at the bottom of .bashrc file

To write give ctrl+o

Give enter

To exit ctrl+x

$sudo apt-get install ssh

**Step 3.**

Download apache hadoop

Download the apache hadoop binary file from the website. Its download as zip file.

To extract the zip file use the following command

$tar -zxvf ~/downloads/hadoop

$cd hadoop3.2.3 #here I given hadoop3.2.3 its depend upon the version your download.

Then move to the configuration file as follows.

$cd etc/hadoop/

$ls

This will list all the configuration file.

$sudu nano hadoop-env.sh

Paste java home in that file

To get java home $echo java\_home

JAVA\_HOME=/usr/lib/jvm/java-8-openjdk-amd64

$sudo nana core-site.xml

Copy core site code from the website and past in the file

<configuration>   
<property>   
<name>fs.defaultFS</name>   
<value>hdfs://localhost:9000</value> </property>   
<property>   
<name>hadoop.proxyuser.dataflair.groups</name> <value>\*</value>   
</property>   
<property>   
<name>hadoop.proxyuser.dataflair.hosts</name> <value>\*</value>   
</property>   
<property>   
<name>hadoop.proxyuser.server.hosts</name> <value>\*</value>   
</property>   
<property>   
<name>hadoop.proxyuser.server.groups</name> <value>\*</value>   
</property>   
</configuration>

$sudo nano hdfs-site.xml

Copy hdfs code from the website and past in the configuration

<configuration>   
<property>   
<name>dfs.replication</name>   
<value>1</value>   
</property>   
</configuration>

$sudo nano mapred-site.xml

Copy mapred code from the website and past in the configuration

<configuration>   
<property>   
<name>mapreduce.framework.name</name> <value>yarn</value>   
</property>   
<property>  
<name>mapreduce.application.classpath</name>   
  
<value>$HADOOP\_MAPRED\_HOME/share/hadoop/mapreduce/\*:$HADOOP\_MAPRED\_HOME/share/hadoop/mapreduce/lib/\*</value>   
</property>   
</configuration>

$sudo nano yarn-site.xml

Copy yarn code from the website and past in the configuration

<configuration>   
<property>   
<name>yarn.nodemanager.aux-services</name>   
<value>mapreduce\_shuffle</value>   
</property>   
<property>   
<name>yarn.nodemanager.env-whitelist</name>   
  
<value>JAVA\_HOME,HADOOP\_COMMON\_HOME,HADOOP\_HDFS\_HOME,HADOOP\_CONF\_DIR,CLASSPATH\_PREP END\_DISTCACHE,HADOOP\_YARN\_HOME,HADOOP\_MAPRED\_HOME</value>   
</property>   
</configuration>

**Step 4.**

$ssh localhost

Copy ssh commands and execute

ssh localhost

Key geteration

ssh-keygen -t rsa -P '' -f ~/.ssh/id\_rsa

Copy file authrization

cat ~/.ssh/id\_rsa.pub >> ~/.ssh/authorized\_keys

Mode configuration

chmod 0600 ~/.ssh/authorized\_keys

Format the name node

hadoop-3.2.3/bin/hdfs namenode -format

Export ssh(format the file system)

export PDSH\_RCMD\_TYPE=ssh

Then start all services

$start-all.sh

**Step 5.**

Goto browser and type(Start NameNode daemon and DataNode daemon)   
localhost:9870

<http://localhost:9870/>

Hadoop will start

Creare the directory

hadoop fs -mkdir /user

Create file in directory

hadoop fs -mkdir /user/example.txt

Create csv file

touch demo.csv # touch command is used to create empty file

Move one file to another file

hadoop fs -put demo.csv /user/example.txt