

Second Instruction

Automated Installation of Big Data Tools (Hadoop, Sqoop, Pig, and Python) via Shell Script File

Step 1/6:

Open the terminal, type **sudo su -** and press the **Enter** key to switch to the root user.

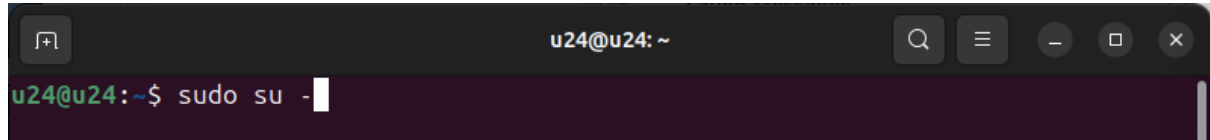


Figure 1. Switching to the root user.

Step 2/6:

Enter the password for the current user and press the **Enter** key.

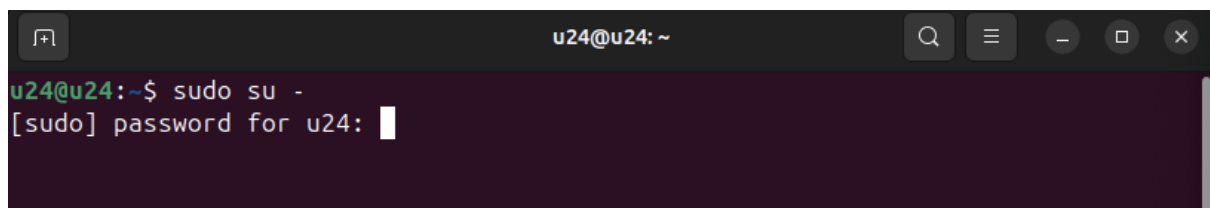


Figure 2. Filling the password for the current user.

Step 3/6:

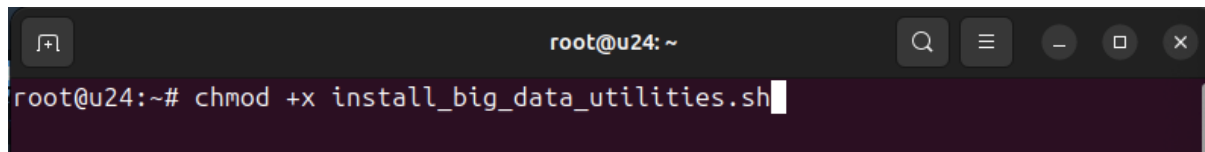
Type **wget**

https://raw.githubusercontent.com/datasciencessource/bash_script/refs/heads/main/install_big_data_utilities.sh

and press **Enter** key to download a script file named `install_big_data_utilities.sh` to store in `/root/` directory.

Step 4/6:

Type **chmod +x install_big_data_utilities.sh** and press **Enter** key to grant execute permission to the script file.

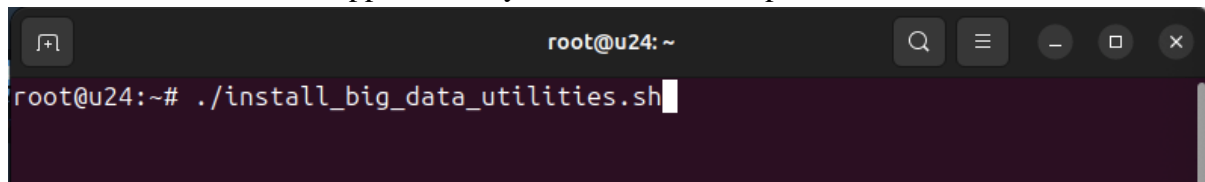
A terminal window with a dark background. The title bar shows 'root@u24: ~'. The command prompt is 'root@u24:~#'. The command entered is 'chmod +x install_big_data_utilities.sh'.

```
root@u24:~# chmod +x install_big_data_utilities.sh
```

Figure 3. Granting execute permission to the script file.

Step 5/6:

Type **`./install_big_data_utilities.sh`** and press **Enter** key to execute the script file and install Big Data utilities such as java, SSH, Hadoop, Sqoop, Pig, and Python. The installation will take approximately 10 minutes to complete.

A terminal window with a dark background. The title bar shows 'root@u24: ~'. The command prompt is 'root@u24:~#'. The command entered is './install_big_data_utilities.sh'.

```
root@u24:~# ./install_big_data_utilities.sh
```

Figure 4. Executing the script file and install big data utilities.

Step 6/6:

After the installation is complete, search the terminal output to verify that all utilities as the following have been installed successfully.

openjdk version "11.0.26" 2025-01-21

Hadoop 2.10.2

Apache Pig version 0.15.0

Sqoop 1.4.7

Python 3.12.3

pip 24.0

```
root@u24: ~  
Processing triggers for desktop-file-utils (0.27-2build1) ...  
  
openjdk version "11.0.26" 2025-01-21  
OpenJDK Runtime Environment (build 11.0.26+4-post-Ubuntu-1ubuntu124.04)  
OpenJDK 64-Bit Server VM (build 11.0.26+4-post-Ubuntu-1ubuntu124.04, mixed mode,  
  sharing)  
Hadoop 2.10.2  
Subversion Unknown -r 965fd380006fa78b2315668fbc7eb432e1d8200f  
Compiled by ubuntu on 2022-05-24T22:35Z  
Compiled with protoc 2.5.0  
From source with checksum d3ab737f7788f05d467784f0a86573fe  
This command was run using /usr/local/hadoop/share/hadoop/common/hadoop-common-2  
.10.2.jar  
4912 NameNode  
7026 Jps  
5170 DataNode  
5570 ResourceManager  
5350 SecondaryNameNode  
5687 NodeManager  
Apache Pig version 0.15.0 (r1682971)  
compiled Jun 01 2015, 11:44:35  
Warning: /usr/local/sqoop/./hbase does not exist! HBase imports will fail.  
Please set $HBASE_HOME to the root of your HBase installation.  
Warning: /usr/local/sqoop/./hcatalog does not exist! HCatalog jobs will fail.  
Sqoop 1.4.7  
git commit id 2328971411f57f0cb683dfb79d19d4d19d185dd8  
Compiled by maugli on Thu Dec 21 15:59:58 STD 2017  
Python 3.12.3  
pip 24.0 from /usr/lib/python3/dist-packages/pip (python 3.12)  
  
root@u24:~#
```

Figure 5. Installation is complete and display in terminal