# **Hadoop - Installation** ❗❗ONLY\_FOR\_WINDOWS❗❗

👉 System Requirements

✅ Windows 10 (64 bit)

✅ Minimum 8GB RAM

✅ Internet connection (approx. 2GB)

✅ Virtualization should be enabled 👌

🚫If these conditions are not met exactly then there might be issues in installation🚫

1. Download & install Oracle VirtualBox

🌍[<https://download.virtualbox.org/virtualbox/6.1.6/VirtualBox-6.1.6-137129-Win.exe>]

1. Download Bitnami Hadoop Stack

🌍[<https://bitnami.com/redirect/to/995396/bitnami-hadoop-3.2.1-2-linux-debian-9-x86_64.ova>]

1. Download & install Putty

🌍[<https://www.chiark.greenend.org.uk/~sgtatham/putty/latest.html>]

1. Download & install WinSCP

🌍[<https://winscp.net/download/WinSCP-5.17.3-Setup.exe>]

1. Import the VM
   1. Launch VirtualBox
   2. Select “File➡️ Import Appliance” menu option in VirtualBox and select the *.ova* file downloaded from the Bitnami website. Click “Continue” to proceed.
   3. Click “Import” to proceed.

❗ Ensure that the memory is set to at least 6GB ❗

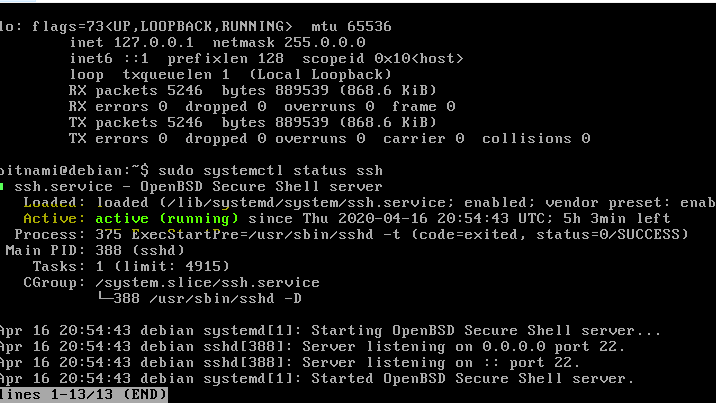
* 1. Settings➡️Network adapter 1:NAT
  2. ➡️Network adapter 2:host-only

1. Start VM
2. Login to the VM using username as ***bitnami*** and password as ***bitnami***

❗ *password is not visible while typing don't be afraid it's there*❗

1. Enable SSH - enter the commands below
   1. sudo rm -f /etc/ssh/sshd\_not\_to\_be\_run
   2. sudo systemctl enable ssh
   3. sudo systemctl start ssh
   4. sudo apt install openssh-server
   5. sudo systemctl status ssh

❗ *After this status should be* ✅*active*❗

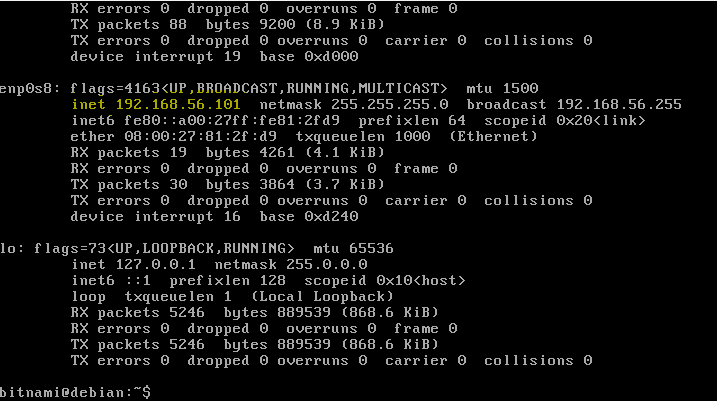


* 1. sudo ufw allow ssh

🌍 Refer this site for further info [<https://linuxize.com/post/how-to-enable-ssh-on-ubuntu-18-04/>]

1. Run the following command to check the IP address of the VM

sudo ifconfig

❗ Note down the **inet** IP. It will be something like **192.168.x.x** ❗

❗ ❗ ❗ If mouse pointer is stuck inside VM press **right ctrl** ❗ ❗ ❗

1. Open Chrome
   1. Type the IP➡️Hit enter
   2. You’ll be prompted to type Username & Password
   3. Go to the VM and type **cat ./bitnami\_credentials**
   4. Note ‘em and type them in the chrome tab [Yes, I know the pass is irritating 😠]
   5. You’ll see a cute 🐘 (Yellow Jumbo) which implies you’re good to proceed to the next step.
2. Create a **.txt** file on your PC ➡️name it accordingly➡️put a ton of 💩 inside it➡️save it.
3. Add the file to the VM
   1. Open WinSCP
   2. Select protocol as **SFTP**
   3. Enter the inet IP address noted above in the Host Name
   4. Enter username and password as **bitnami**
   5. Click Login
   6. The left pane is the VM’s file directory and right one is your file directory
   7. In the VM’s directory navigate to **/home/bitnami** and paste the .txt you created in the above step
   8. Close WinSCP (you can even uninstall it if you want 😂, JK don't uninstall it)
4. SSH to the VM
   1. Open PuTTY
   2. Enter the inet IP address noted above in the Host Name
   3. Connection type: SSH
   4. Click Open
5. Finally 🏆 run these commands in Putty’s terminal
   1. **hadoop fs -rmr /tmp/hdfs-example-output**
   2. **hadoop fs -put -f /home/bitnami/input\_file.txt /tmp/hdfs-example-input**
   3. **hadoop jar /opt/bitnami/hadoop/share/hadoop/mapreduce/hadoop-mapreduce-examples-\*.jar grep /tmp/hdfs-example-input /tmp/hdfs-example-output '[a-zA-Z0-9]+'**
   4. **hadoop fs -cat /tmp/hdfs-example-output/part-r-00000**
6. Take 📸 of the output. Create doc and submit.

**🔥This might look easy for you but it had a lot of hard work and time(almost 2 days & nights)involved,So...cheer up and #workhard.🔥**

#StaySafe!!!...#Go Corona, Corona Go