⇒ Steps to start Hadoop on Ubuntu :

1. Open terminal and type : su - hdoop
   Password : Type the password.

2. cd $HADOOP_HOME

3. cd sbin

4. ./start-dfs.sh

5. ./start-yarn.sh

6. jps.

```
2418 DataNode
2611 SecondaryNameNode
2281 NameNode
3501 Jps
2813 ResourceManager
2943 NodeManager
```

So, All the daemon processes are running.

Date
Page

⇒ wordcount program :

1. create sample text file on the desktop with file name as wc.txt. and enter the contents to the file.

2. create a directory in the hadoop distribute file system, using the following command.

```
hadoop fs -mkdir -p lab6.
```

- check whether the directory has been created successfully in hdfs using the following command.

```
hadoop fs -ls.
```

```
drwxr-xr-x - hdoop supergroup lab6.
```

3. Now let us put the file wc.txt which we have created in step 1 from local filesystem to hdfs use the following command.

```
hadoop dfs -put /home/mdyaseen/Desktop/wc.txt
    lab6/input.txt.
```

- check whether the file has successfully been putted into the hdfs using the following command

```
hadoop fs -ls lab6
```

```
Found 1 items
-rw-r--r--  1 hdoop supergroup  lab6/input.txt
```

4. Now let us execute the jar file. First Download the jar file and place it on to the Desktop and execute the following commands.

```
hadoop jar /home/mdyaseen/Desktop/wordcount.jar
WordCount lab6/  /lab6_output.
```

5. once the job is executed successfully, list the contents of lab6_output Directory with the following command.

```
hadoop dfs -ls /lab6_output.
```

```
Found 2 items
        /lab6_output/_SUCCESS
        /lab6_output/part-r-00000
```

6. Now let us print the output on the terminal

hadoop dfs -cat /lab6_output/part-r-00000

Bear 2
car 3
Deer 2
River 2