

Linux and Introduction to Big Data Hadoop

Homework

Sujatha K
211720104148
3rd year CSE C

edureka!

edureka!

Q1. Open the file */var/log/messages* in the vi editor and delete line number 150.

To open the file: `$ vi [filename]`
Delete line numbers: `$ 150d`

Q2. Write a shell script to add two numbers?

```
$ sum=$(( $a + $b ))
```

Q3. User root wants to copy */etc*, including all subdirectories and files to */tmp*. How will you achieve this task?

```
$ copy /etc /root  
$ cp -r /etc/* /root  
$ cp -v /etc/* /root  
$ rcp /etc/* /root  
$ cp -R /etc/*.* /root
```

Q4. Create a file that contains only the username and the user id of all the users present on the server.

```
$ touch file1.txt  
$ sudo useradd -u username
```

Q5. How will you provide a count of all users on the system except for adm user?

```
$ ps -eaho user
```

Q6. How will you list all files in */tmp* in increasing order of their size?

```
ls -lS /path/to/folder/
```

Q7. What command is used to clear history on the Linux server?

```
$ history -d
```

Q8. Explain “Big Data” and what are five V’s of Big Data?

The process of examining large data sets to uncover hidden parts, market trends.
Five V's: Volume, Velocity, Validation, Veracity, Variety

Q9. What is Hadoop and its components?

Hadoop is an open source framework from Apache and is used to store process and analyze data which are very huge in volume.
Components: Hadoop kernel, Map reduce, Hadoop distributed file system, Apache hive, Hbase, Zookeeper.