

Steps to Open Hive in Console

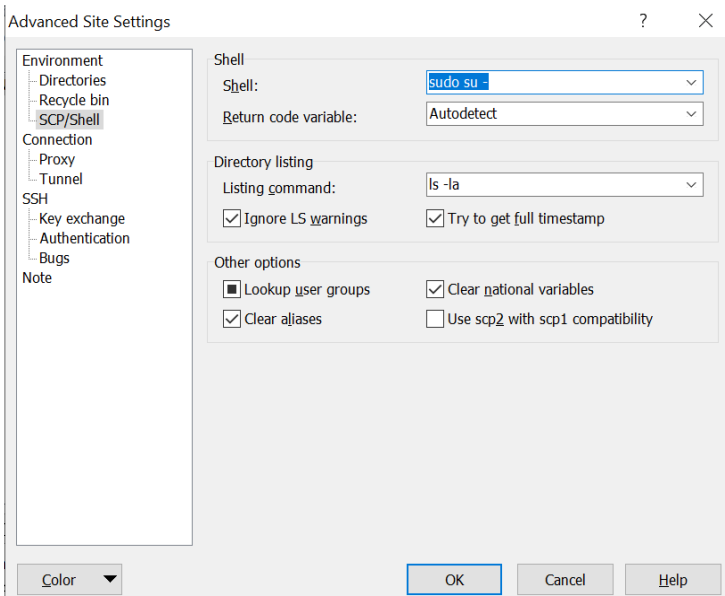
After successful EMR cluster creation and creation of S3 bucket, let's download the data set into the EMR instance.

Step 1: Once, you are in the EMR instance, create a directory called **movie_ratings**. For this example, we are performing these steps in the **root user** but these same steps can be easily followed for the **hadoop user** as well.

```
[root@ip-10-0-0-247 ~]# mkdir movie_ratings
[root@ip-10-0-0-247 ~]# cd movie_ratings
[root@ip-10-0-0-247 movie_ratings]#
```

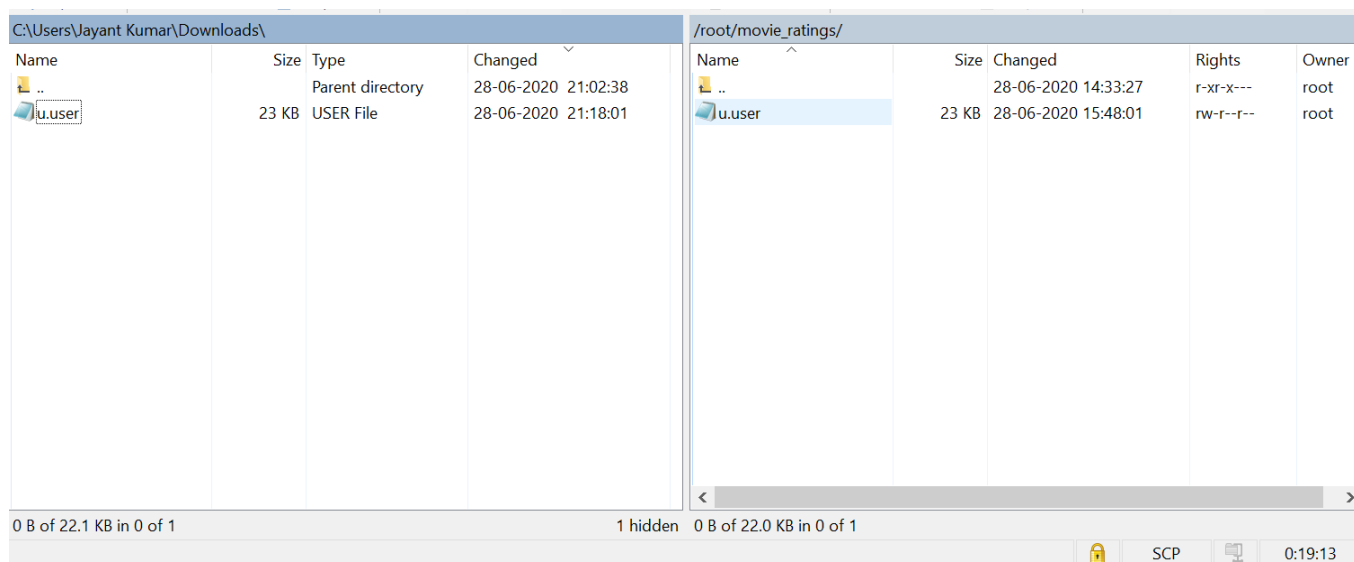
Step 2: Download WinSCP from the following link and install it: <https://winscp.net/eng/download.php>. These steps will be the same as you would have followed in the **Introduction to Cloud computing and AWS Setup module**. Follow those steps to login to WinSCP

Step 3: To get access to the EMR root user, you need to configure the **sudo su -** command, Go to **Environment>SCP/Shell**. Select **sudo su -** from the drop-down list available in the **Shell** field.




Step 4: Click on **Login**. On successful login, you will get root user access to the EMR instance. Before going further, you will also have to download the movie_ratings data set from the following URL:
<http://ml-cloud-dataset.s3.amazonaws.com/u.user>

Step 5: Go to the **Download** folder, and drag the file to the **movie_ratings** folder that you created in Step 1.



Step 6: Validate that the file is copied successfully to the **movie_ratings** folder by executing a command **ls**.

 root@ip-10-0-0-247:~/movie_ratings

```
[root@ip-10-0-0-247 ~]# cd movie_ratings/
[root@ip-10-0-0-247 movie_ratings]# ls
u.user
[root@ip-10-0-0-247 movie_ratings]#
```

Step 7: Now, you can execute a command **hive** to get access to the **HIVE** console.