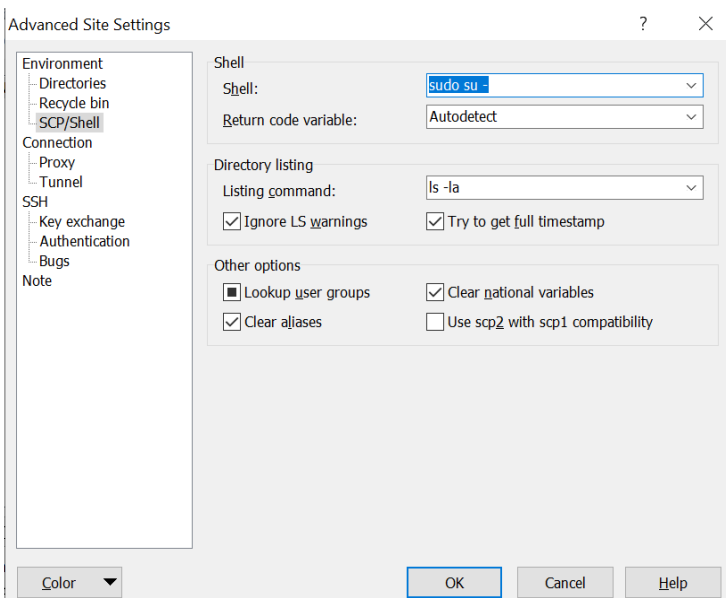# Steps for Downloading Employee and Department Data Set

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. Create two directories separately with the names '**employee**' and '**department'**.

Step 2: Open **WinSCP.** The following 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: Download the employee and department data sets using the following two URLs, respectively:

http://ml-cloud-dataset.s3.amazonaws.com/employee
http://ml-cloud-dataset.s3.amazonaws.com/dept

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

Step 6: Validate that the files are copied successfully to the **employee** and **department** folders by executing a command **ls**.

root@ip-10-0-0-247:~/employee
```
[root@ip-10-0-0-247 ~]# cd employee
[root@ip-10-0-0-247 employee]# ls
employee
[root@ip-10-0-0-247 employee]#
```

root@ip-10-0-0-247:~/department
```
[root@ip-10-0-0-247 ~]# cd department/
[root@ip-10-0-0-247 department]# ls
dept
[root@ip-10-0-0-247 department]#
```