Installing mrjob on HDP 2.65

First you need to connect to your running Hortonworks Sandbox HDP virtual machine from a terminal.

On Windows, you can use [PuTTY](https://putty.org/). Enter **maria\_dev@127.0.0.1** for the hostname, and **2222**for the port.

On Mac or Linux, open a terminal window and run ssh -p 2222 127.0.0.1

To log in, the password for maria\_dev is also **maria\_dev**.

You will now need to escalate your privileges in order to set things up. Enter:

su root

The default root password is **hadoop**. You will be prompted to change this.

Once you are successfully at the root level (you should see a # prompt instead of $), enter the following commands:

1. cd /etc/yum.repos.d
2. mkdir /etc/yum.repos.d/old
3. mv /etc/yum.repos.d/CentOS\*.repo /etc/yum.repos.d/old/
4. mv /etc/yum.repos.d/epel\*.repo /etc/yum.repos.d/old/
5. mv /etc/yum.repos.d/ambari\*.repo /etc/yum.repos.d/old/
6. mv /etc/yum.repos.d/hdp\*.repo /etc/yum.repos.d/old/
8. wget http://media.sundog-soft.com/hadoop/CentOS.repo
9. wget http://media.sundog-soft.com/hadoop/epel.repo
10. yum clean all
12. #NOTE! That's a lowercase L in ius-release-el7.rpm below, not the number 1
13. yum install -y https://repo.ius.io/ius-release-el7.rpm
15. rm ambari\*
16. yum check-update
18. yum install -y python-pip
19. pip install --trusted-host pypi.python.org pathlib
20. pip install --trusted-host pypi.python.org mrjob==0.7.4
21. pip install --trusted-host pypi.python.org PyYAML==5.4.1
22. yum install -y nano
24. cd /home/maria\_dev
25. wget http://media.sundog-soft.com/hadoop/RatingsBreakdown.py
26. wget http://media.sundog-soft.com/hadoop/ml-100k/u.data