

Setting Up Mahout for Shell Jobs

1) Login as "hadoop" user to your existing AWS VM / navigate to your home directory:

```
cd /usr/local/lib
```

1.1) Download Mahout Source Code:

```
sudo git clone https://github.com/apache/mahout.git mahout
```

2.0) Edit variables to access Mahout:

```
export MAHOUT_HOME=/usr/local/lib/mahout
```

```
export MAHOUT_LOCAL=true # for running standalone on your dev machine,
```

2.1) Download new version of mvn (3.3.+) and build Maven

```
sudo wget http://www-eu.apache.org/dist/maven/maven-3/3.3.9/binaries/apache-maven-3.3.9-bin.zip
```

```
sudo unzip apache-maven-3.3.9-bin.zip
```

```
cd apache-maven-3.3.9-bin.zip
```

```
sudo vi ~/.bashrc
```

Add the following path of Ubuntu (may be different where you put maven on OS X)

SHIFT + I

```
export PATH=$PATH:/usr/local/lib/apache-maven-3.3.9/bin
```

CTRL + C then :wq!

```
source ~/.bashrc
```

Check Maven is installed correctly:

```
mvn -v
```

```
Apache Maven 3.3.9 (bb52d8502b132ec0a5a3f4c09453c07478323dc5; 2015-11-10T16:41:47+00:00)
Maven home: /usr/local/lib/apache-maven-3.3.9
Java version: 1.7.0_111, vendor: Oracle Corporation
Java home: /usr/lib/jvm/java-7-openjdk-amd64/jre
Default locale: en_US, platform encoding: UTF-8
OS name: "linux", version: "3.13.0-91-generic", arch: "amd64", family: "unix"
```

```
sudo wget http://mirror.symnds.com/software/Apache/mahout/0.12.2/apache-mahout-distribution-0.12.2.tar.gz
```

```
sudo tar -xvf apache-mahout-distribution-0.12.2.tar.gz apache-mahout-distribution-0.12.2
```

Setting Up Mahout for Eclipse IDE Use (Optional)

If you are using a different machine with a GUI then please make sure you have the following perquisites installed:

- Eclipse IDE
- Maven
- Mahout

If you are using an AWS Instance or a local Ubuntu image you should be able to run Eclipse IDE or create a VNC to your AWS Box and run an Eclipse IDE. Here is a link to download Eclipse:

<http://www.eclipse.org/downloads/packages/eclipse-ide-java-developers/marsr> (Typically Mars)

If not sure download the installation utility and it will help you choose once it is installed. Important step is to pick Java EE.

Just a quick recap if you are using OS X or Ubuntu and have brew installed:

```
brew install maven
```

```
cd Desktop
```

```
git clone https://github.com/apache/mahout
```

```
cd mahout
```

```
mvn eclipse:eclipse
```

This will build a Java compatible version of Mahout that you can open in your Java IDE. If you would like a guide with sample code for your HW there are multiple Mahout related guides, you can find here:

<https://mahout.apache.org/users/>

Mahout Mvn (Java IDE) Recommendation Tutorial

<https://www.youtube.com/watch?v=63k560Livmg>