Activity: Setup Database data

This Reading describes the steps to download the datasets and scripts necessary for the Hands On exercises in this course. Before proceeding, make sure have you downloaded and installed VirtualBox and the Cloudera virtual machine. Instructions for this process can be found in the Reading called *Downloading and Installing the Cloudera VM*.

Step 1. **Start the Cloudera VM.** Most of the Hands On exercises in this course use the Cloudera Virtual Machine, so we will download the datasets onto the VM. Start the VM in VirtualBox and perform the remaining steps in the VM.
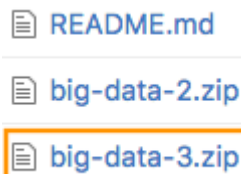
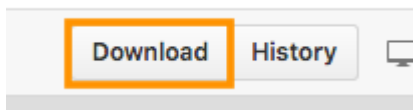Step 2. **Open a web browser.** Open a web browser by clicking on the web browser icon in the top toolbar.

In the web browser, enter the following for the URL:

```
http://github.com/words-sdsc/coursera
```
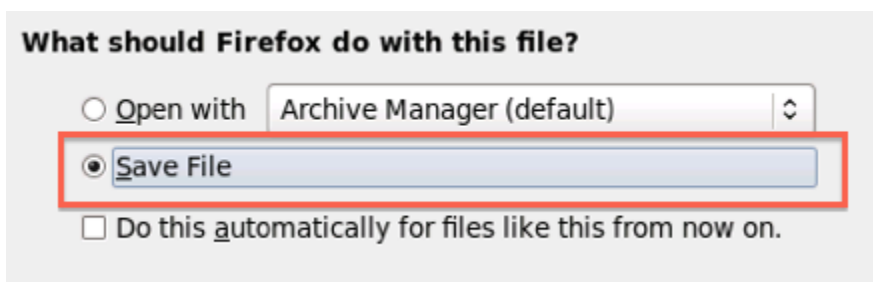
Step 3. **Download the datasets.** Click on *big-data-3.zip*:
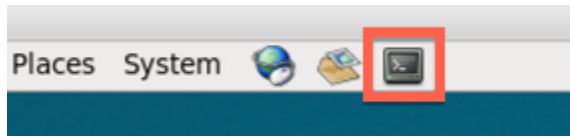

Click on the *Download* button:

In the dialog, select *Save File*:

Click *OK*, and the file big-data-3.zip file will be downloaded to the Downloads directory.

Step 4. **Uncompress the datasets.** Open a terminal shell by clicking on the terminal shell icon in the top toolbar.

In the terminal, run:

```
cd Downloads
unzip -o big-data-3.zip
```

Step 5. **Install tools.** Change directories to *big-data-3* and run *setup.sh* to install tools and libraries.

```
cd big-data-3
./setup.sh
```

During the setup process, Anaconda will give you a series of prompts. First, press *enter* to continue the installation:

```
Welcome to Anaconda3 4.0.0 (by Continuum Analytics, Inc.)

In order to continue the installation process, please review the license
agreement.
Please, press ENTER to continue
>>> ▊
```

Next, read and accept the license:

```
Do you approve the license terms? [yes|no]
>>> ▊
```

Next, press *enter* to accept the default installation location:

```
Anaconda3 will now be installed into this location:
/home/cloudera/anaconda3

  - Press ENTER to confirm the location
  - Press CTRL-C to abort the installation
  - Or specify a different location below

[/home/cloudera/anaconda3] >>> ▊
```

Next, enter *yes* when it asks if you want to prepend the install location to PATH:

```
installation finished.
Do you wish the installer to prepend the Anaconda3 install location
to PATH in your /home/cloudera/.bashrc ? [yes|no]
[no] >>> yes
```

The setup of tools and datasets should continue.

Finally, source $HOME/.bashrc:

```
source $HOME/.bashrc
```