Working With Messy Data in OpenRefine Workshop

Download the zip file from http://uoft.me/openrefine and save it in a folder on the desktop. Make sure to extract the files from the zip file.

If you are unsure how to do this, you will find detailed instructions in the following *Set Up* section.

Set Up

I recommend you use *Google Chrome* to download all the files for this workshop. To have *Chrome* ask you where to save a file whenever you download something (to avoid having to find it in your Downloads directory and move it to your workshop folder):

1. Start up *Chrome* (if not already running).

2. Click on the customize icon (3 dots arranged vertically) on the top right.

3. Select **Settings**. Scroll down to the bottom and expand the **Advanced** settings. Under **Downloads**, turn on **Ask where to save each file before downloading** and then close the tab.

Next, you should create a folder to store all your files. You are welcome to organize the files as you wish, but if you need some file management assistance, I recommend the following:

4. Right click on the desktop and select **New** and then **Folder**. Name this folder *Workshop*.

Finally, download all the files to this newly created directory.

5. Using *Chrome*, go to http://uoft.me/openrefine (case sensitive)

6. Click on the *WorkshopFiles.zip* file. From the pop-up window, browse to the *Workshop* folder you just created on the Desktop. Then click on **Save**.

7. Go to the *Workshop* folder. Right click on the zip file called *Workshopfiles* and select **7-Zip->Extract Here**

8. This should expand out to a number of files. You should see various data files that we will be using throughout the workshop. You should also see various activity PDF files, which contain activities with steps to follow to learn more about using OpenRefine. Finally, you should see the *Slides* PDF file, which is a copy of the slides for your reference, and the *Handout* PDF file, which are these instructions.