

# Data Wrangling (or Cleaning) with OpenRefine

Pre-class activities: 15 min

Face-to-face presentation: 15 min

Face-to-face activities: 45 min

**Presenter: Shahira Khair (skhair@uvic.ca)**

## Learning outcomes

By the end of this session, participants will gain:

1. Understanding of the importance of data cleaning
2. Experience conducting key data cleaning practices using OpenRefine:
  - A. Analyzing the occurrence of values throughout a data set
  - B. Clustering and standardizing values
  - C. Separating multiple values contained in the same field
  - D. Joining multiple values contained in separate fields

## Step 1: Pre-Class Videos, Readings, & Exercises

- [Install](#) OpenRefine (make sure to follow the [instructions](#) for your operating system) (~ 5 min)
- Download the [Samuel Maclure dataset](#) and save it locally to your laptop (< 1 min)
- Download the [McTaggart Cowan dataset](#) and save it locally to your laptop (< 1 min)
- Watch this introductory [YouTube video](#) about OpenRefine (previously called Google Refine) (7 min)
- Resize Handouts for Your Laptop Screen (2 min) - <https://youtu.be/lqk5hZUfzN0>
- Enabling Desktop Sharing in Zoom & Privately Request Assistance during the workshop (2 min) - <https://youtu.be/TWb-bms3R2Y>

## Step 2: Download In-Class Active Learning Activities & Instruction

- [Presenter slides](#) (<http://bit.ly/2LggTrd>)
- [Activity instructions](#) (<http://bit.ly/2WiS6no>)

## Step 3: More Resources, Post Class Assessment & Course Feedback

- [Using OpenRefine guidebook](#) (E-book available from UVic Libraries)
- [Online Guide to OpenRefine](#) (University of Illinois Urbana Champaign Library)
- [GREL Recipes](#) (OpenRefine Github webpage)

