

OpenRefine for Data Cleaning

Facilitation Guide - 60 minutes

Description

In this module, participants will be introduced to OpenRefine, a data cleaning tool. Participants will learn how to import their data and create a project. They will also learn four common features that may optimize their data cleaning experience. Participants will be able to navigate through the user interface well enough to clean and export their project.

Learning Objectives

Participants will...

- Learn to load a file and create a project.
- Learn to use OpenRefine features to optimize their data cleaning.
- Learn how to export their cleaned data.

Materials

Presentation

Pre and Post Assessment

Optional: Virtual Question Parking Lot (i.e. Jamboard)

Facilitator Guide

Welcome and Getting Started (Slides 1-2) - Introduce yourself and your role. Share the purpose of the session and discuss the Learning Objectives.

5 Minutes

Tips & Tricks

- Key Ideas to Highlight
 - This is going to be fun! You'll learn a lot.
- Resources to Share
 - Pre-Assessment

- We will introduce participants to OpenRefine
 - a tool for cleaning data
 - Discuss why we use OpenRefine

Take a few minutes to go through the pre-assessment and see what you know. If you're unsure/have questions about the content, those are great places to start when engaging with the presentation.

Consider – Using a virtual tool like Jamboard to serve as a question parking lot for participants to share questions that come to mind throughout the module.

Purpose (Slides 3-5) - In this section, we'll provide an overview of what OpenRefine can be used for.

5 Minutes

Tips & Tricks

- Key Ideas to Highlight
 - OpenRefine Conventions
- Resources to Share

- We will...
 - Introduce participants to the OpenRefine user interface.
 - Explain what OpenRefine can and cannot do.

Accessing Data (Slides 6-16) - In this section, we'll cover the process of importing data in OpenRefine and how to create a project. We will also detail the installation process.

10 Minutes

Tips & Tricks

- Key Ideas to Highlight
 - App installation.
 - Loading a file.
 - Creating a project.
- Resources to Share

- We will...
 - Guide participants through the process of installing the OpenRefine application.
 - Help participants navigate the process of loading a file.
 - Discuss loadable file formats.
 - Explain common mistakes.
 - Help participants navigate the process of creating a project.
 - Give an example of what can go wrong.

Manipulating Data (Slides 17-39) - In this section, we'll cover four OpenRefine features that assist in data cleaning.

Tips & Tricks

- Key Ideas to Highlight
 - Faceting

<p>20 Minutes</p>	<ul style="list-style-type: none"> ○ Filtering ○ Clustering ○ Reconciliation ● Resources to Share
<ul style="list-style-type: none"> ● We will... <ul style="list-style-type: none"> ○ Discuss how to use Faceting to sort and edit columns ○ Discuss how to use Filtering to search for values. ○ Discuss how to use Clustering to group records from a Faceted column. ○ Discuss how to use Reconciliation to import data from a URL to match against a Faceted column. <p><i>If time allows, run through an example of each function to engage hands-on learners.</i></p>	
<p><i>Exporting Data (Slides 40-44) - In this section, we'll cover how to save a copy of your cleaned data.</i></p> <p>5 Minutes</p>	<p>Tips & Tricks</p> <ul style="list-style-type: none"> ● Key Ideas to Highlight ● Resources to Share
<ul style="list-style-type: none"> ● We will... <ul style="list-style-type: none"> ○ Discuss how to export cleaned data. <ul style="list-style-type: none"> ■ Explore the Custom Tabular Exporter and explain when different options might be utilized. 	
<p><i>Practice and Recap(Slides 45-48) - In this section, we'll offer participants a chance to practice new skills, and revisit the learning objectives. Participants will have a chance to ask questions.</i></p> <p>15 Minutes</p>	<p>Tips & Tricks</p> <ul style="list-style-type: none"> ● Key Ideas to Highlight ● Resources to Share <ul style="list-style-type: none"> ○ Practice links ○ Post Assessment
<ul style="list-style-type: none"> ● Allow time for participants to attempt a practice exercise now, to have access to instructor help if needed. ● Questions? - Check to see if your audience has any questions about the presentation. If used, refer back to the virtual question parking 	

lot (i.e. Jamboard) to answer any last questions.

- **Thank you!** – Give yourself a round of applause! Share link to post-assessment for the final Check for Understanding.