

Step-by-Step: Even more data-cleaning techniques

This reading outlines the steps the instructor performs in the next video, [Even more data-cleaning techniques](#). This video teaches you different methods data analysts use in data mapping. Data mapping is the process of matching fields from one database to another. It's critical to the success of data migration, data integration, and many other data-management activities. This video contains one activity for you to practice.

Keep this step-by-step guide open as you watch the video. It can serve as a helpful reference if you need additional context or clarification while following the video steps. This is not a graded activity, but you can complete these steps to practice the skill demonstrated in the video.

What you'll need

If you'd like to follow along with the example in this video, choose a spreadsheet tool, such as Google Sheets or Excel.

To access the spreadsheet the instructor uses in this video, click the link to the template to create a copy of the dataset. If you don't have a Google account, download the data directly from the attachments below.

Link to templates:

[International Logistics Association memberships](#)

[Global Logistics Association](#)

[Logistics Association Merger](#)

Downloads:

[International Logistics Association memberships](#)

[XLSX File](#)

[Global Logistics Association](#)

[XLSX File](#)

[Logistics Association Merger](#)

[XLSX File](#)

Example: **CONCATENATE**

CONCATENATE is a function that joins together two or more text strings. In the video, you'll learn how to use **CONCATENATE** to clean data after two datasets have been combined.

1. Open the dataset spreadsheet titled **Global Logistics Association**. When prompted, select **USE TEMPLATE**.
2. **Insert** a new column to the right of column **E**. Label it **New Address** in cell **F1**.
3. In the second row of the new column (cell **F2**), enter **=CONCATENATE (D2 ,E2)** and press **Enter**.
 1. You will notice that some results need a space between the street address and the unit or suite number, such as: 25 Dyas RdSte. 101.
 2. You could manually clean the data later to add a space between Rd and Ste., but **CONCATENATE** can actually do it for you.
 3. The **CONCATENATE** formula can help you format the data as it is merged by entering an additional string to insert a space between Rd and Ste.
 4. Enter **=CONCATENATE (D2 , " ", E2)** and you will have an address that is formatted like this: 25 Dyas Rd Ste. 101. Much better!
4. Ensure the new data in the cell accurately reflects the merging of the two previous columns.
5. **Select cell F2** and **drag down** to apply the formula to all rows in the column.