**Learning Log: Develop your approach to cleaning data**

**Instructions**You can use this document as a template for the learning log activity: Consider how data analysts approach tasks. Type your answers in this document, and save it on your computer or Google Drive.

We recommend that you save every learning log in one folder and include a date in the file name to help you stay organized. Important information like course number, title, and activity name are already included. After you finish your learning log entry, you can come back and reread your responses later to understand how your opinions on different topics may have changed throughout the courses.

To review detailed instructions on how to complete this activity, please return to Coursera: [Learning Log: Develop your approach to cleaning data](https://www.coursera.org/learn/process-data/supplement/rKHUX/learning-log-develop-your-approach-to-cleaning-data).

| **Date:** <enter date> | **Course/topic:** Course 4: Process Data from Dirty to Clean | | |
| --- | --- | --- | --- |
| **Learning Log:** Develop your approach to cleaning data | | |
| **Think about “clean” and “dirty” data** | Use the instructions from the learning log to complete the Data Cleaning Approach Table below:   | **Data cleaning checklist** | **Preferred cleaning methods** | **Data cleaning motto** | | --- | --- | --- | | * Clear formatting * Convert data type * Standardize capitalization * Remove duplicates * Remove irrelevant data * Handle missing values * Fix errors | * Listed below for better readability | * Data cleaning, One Piece at a time. | | | |
| **Reflection:** | Write 1-2 sentences (20-40 words) answering each of the following questions: | | |
| **Questions and responses:** | * What items did you add to your data cleaning checklist? Why did you decide these were important to check for?   *Type your response here*   * How have your own experiences with data cleaning affected your preferred cleaning methods? Can you think of an example where you needed to perform one of these cleaning tasks?   *Type your response here*   * How did you decide on your data cleaning motto?   *Type your response here* | | |

**Preferred cleaning methods**

Clear formatting:

* Use Excel's "Clear Formatting" button to remove any unwanted formatting, such as font size or color, from selected cells.
* Use the "Format Painter" tool to copy formatting from one cell to another.

Convert data type:

* Use Excel's "Change Data Type" option to convert a column of data from one data type to another (e.g. from text to number).
* Use formulas such as =TEXT() or =VALUE() to convert data types.

Standardize capitalization:

* Use Excel's "Upper" or "Lower" functions to convert all text to upper or lower case.
* Use the "Proper" function to capitalize the first letter of each word.

Remove duplicates:

* Use Excel's "Remove Duplicates" option to remove any duplicate rows in a selected range of data.
* Use the "Advanced Filter" option to filter out duplicates based on a specific column or columns.

Remove irrelevant data:

* Use Excel's "Delete" option to delete rows or columns that are not needed for the analysis.
* Use the "Filter" option to hide rows or columns that are not needed for the analysis.

Handle missing values:

* Use Excel's "IF" function to replace missing values with a specific value or text.
* Use the "IFERROR" function to replace missing values with a specific value or text when the formula encounters an error.

Fix errors:

* Use Excel's "Find and Replace" option to find and replace any errors in the data.
* Use the "IF" function in combination with logical operators to check for errors and replace them with correct values.