



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](#).

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:		
	<u>Data cleaning checklist</u>	<u>Preferred cleaning methods</u>	<u>Data cleaning motto</u>
	<ul style="list-style-type: none">• <i>Type your response here</i>• Size of the data set (e.g., length of post code)• Data types (Specify string, integer)• Duplicated records• Misspelled values• Missing values• Unformatted data (date, price)	<ul style="list-style-type: none">• <i>Type your response here</i>• Conditional Formatting• COUNTIF()• LEFT(), RIGHT(), MID()• Chart (show something unusual than others)• Split to columns• Data Cleanup (Remove duplicates, Trim whitespace)• Data Formatting	<ul style="list-style-type: none">• <i>Type your response here</i><i>"Avoid applying complicated solutions to simple problems."</i>



Reflection:	Write 1-2 sentences (20-40 words) answering each of the following questions:
Questions and responses:	<ul style="list-style-type: none">• What items did you add to your data cleaning checklist? Why did you decide these were important to check for? <i>You can directly look on the above table. Most of my cleaning checklist contains a general thing that should be done to make the clean data, and things that should be included on the data collection stage (to make the data clean beforehand, in case that you can collect the data yourself by using survey)</i>• 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? <i>The more you face the problem, the more you know how to deal with the problem. I have a little change to see a lot of data (apart from self-learning), however I can practice cleaning external data that I got from various sources when doing practice.</i>• How did you decide on your data cleaning motto? <i>Out of three mottos, I love all of them, however after my considerations, this one matches with me the most =)</i>