

✓ **Congratulations! You passed!**

Grade received **100%** To pass 100% or higher

[Go to next item](#)

Hands-On Activity: Cleaning data with spreadsheets

Total points 2

1.



1 / 1 point

Activity overview

You've learned about cleaning data and its importance in meeting good data science standards. In this activity, you'll do some data cleaning with spreadsheets, then transpose the data.

By the time you complete this activity, you will be able to perform some basic cleaning methods in spreadsheets. This will enable you to clean and transpose data, which is important for making data more specific and accurate in your career as a data analyst.



What you will need

To get started, first access the data spreadsheet.

To use the spreadsheet for this course item, click the link below and select "Use Template."

Link to data spreadsheet: [Cleaning with spreadsheets](#)

OR

If you don't have a Google account, you can download the template directly from the attachment below.



Data Spreadsheet for Cleaning with Spreadsheets
XLSX File

[Download file](#) ↓

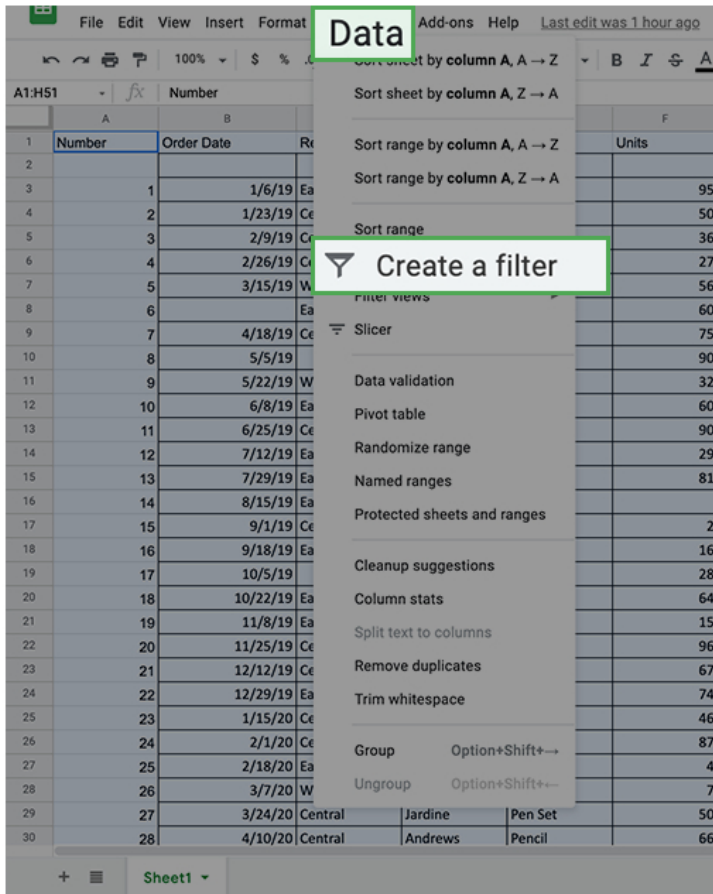


Select and remove blank cells

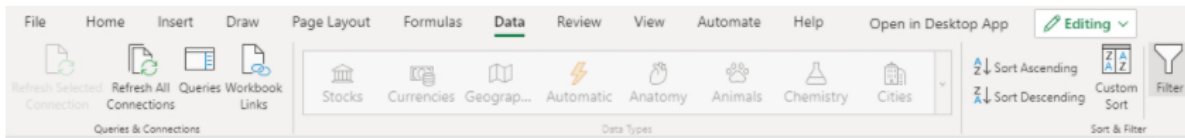
The first technique we'll use is to select and eliminate rows containing blank cells by using filters. To eliminate rows with blank cells:

1. Highlight all cells in the spreadsheet. You can highlight **Columns A-H** by clicking on the header of **Column A**, holding **Shift**, and clicking on the header of **Column H**.

2. Click on the **Data** tab and pick the **Create a filter** option. In Microsoft Excel, this is called **Filter**.



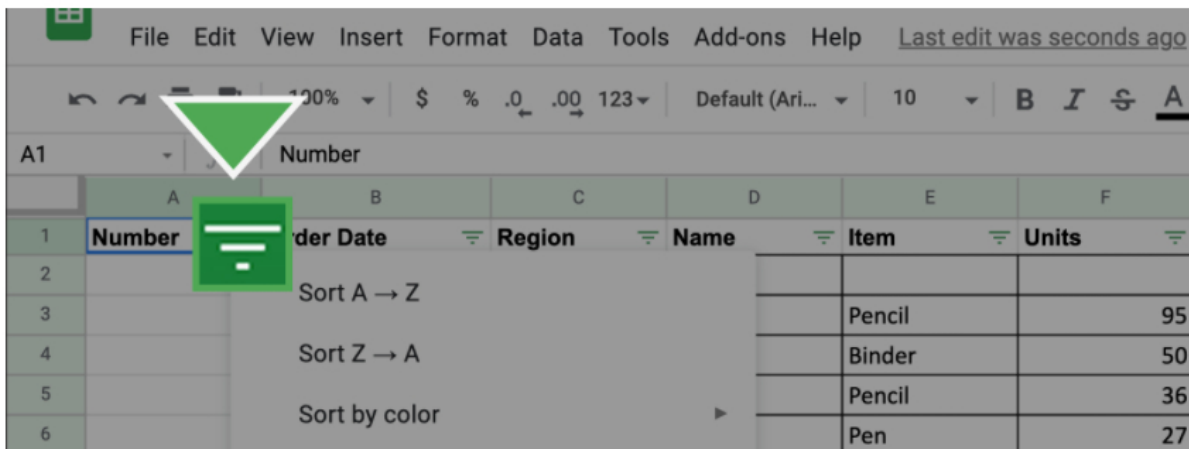
Excel:

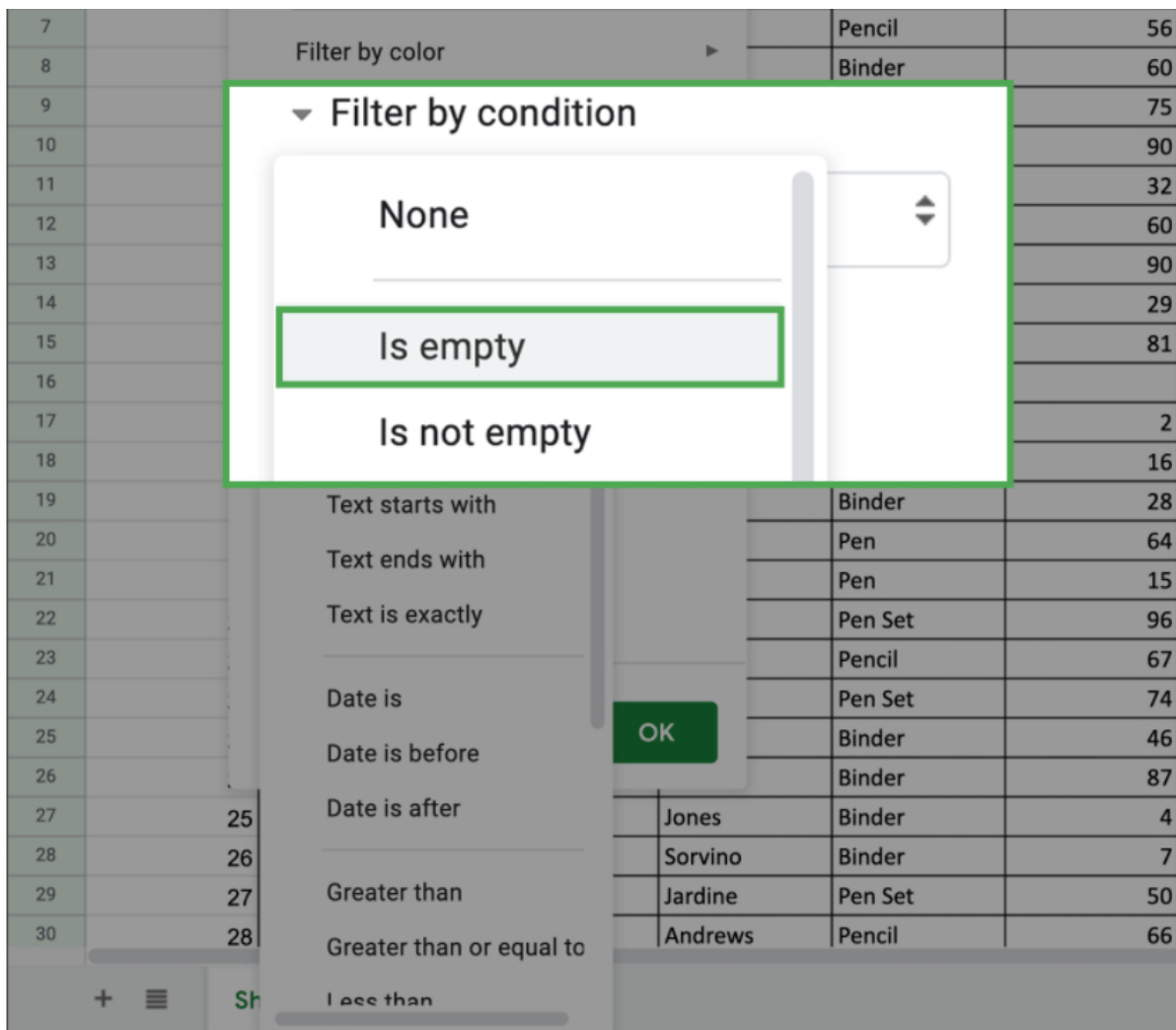


3. Every column now shows a green triangle in the first row next to the column title. Click the green triangle in **Column B** to access a new menu.

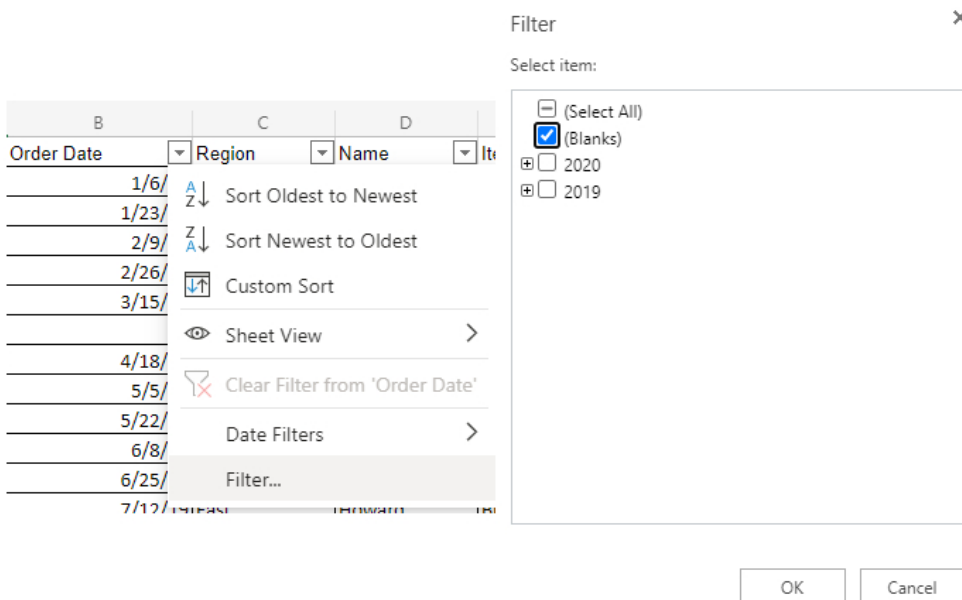
4. On that new menu, click **Filter by condition** and open the dropdown menu to select **Is empty**. Click **OK**.

In Excel, click the dropdown, then **Filter...** then make sure only **(Blanks)** is checked. Click **OK**.





Excel:



You can then review a list of all the rows with blank cells in that column.

5. Select all these cells and delete the rows except the row of column headers.

6. Return to the **Filter by condition** and return it to **None**. In Excel, click **Clear Filter from 'Column'**.

- **Note:** You will now notice that any row that had an empty cell in **Column A** will be removed (including the extra empty rows after the data).

7. Repeat this for **Columns B-H**.

All the rows that had blank cells are now removed from the spreadsheet.

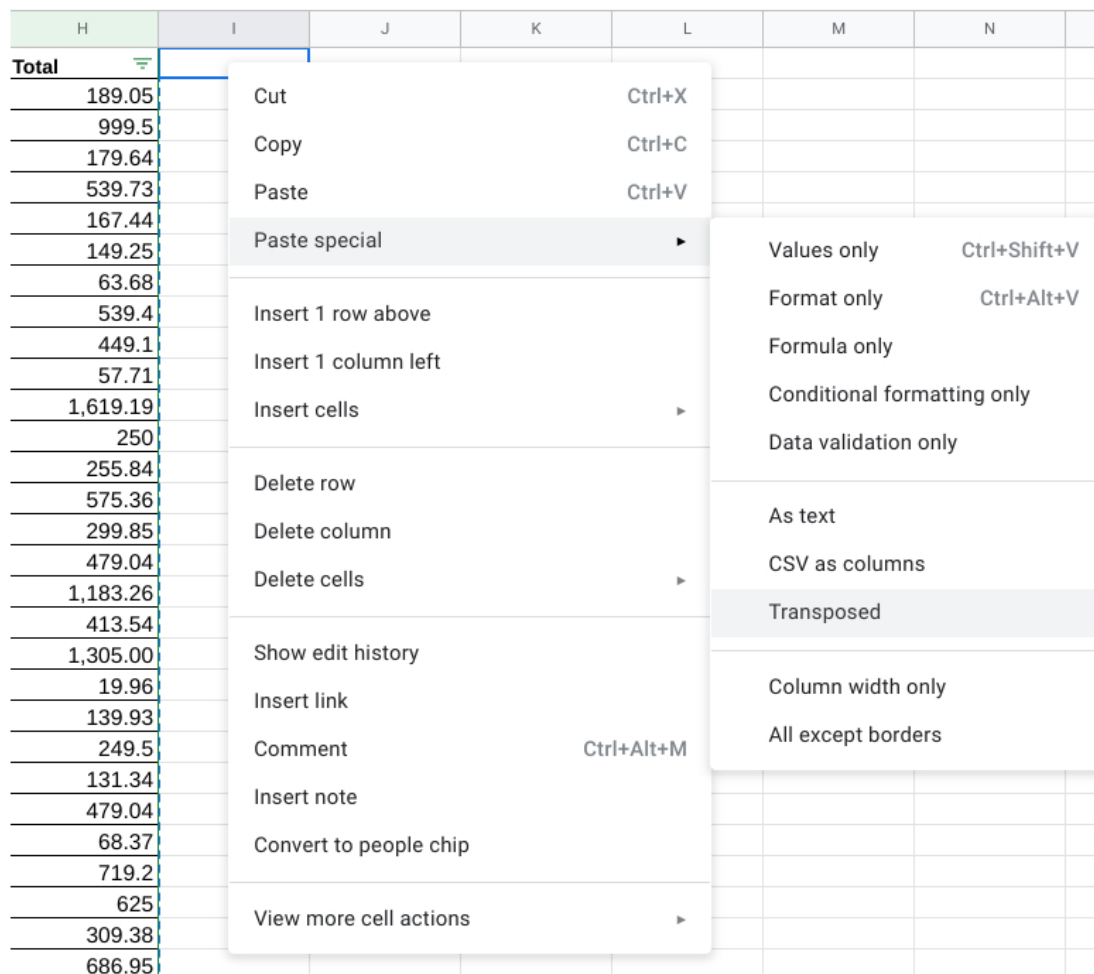
Transpose the data

The second technique you will practice will help you convert the data from the current long format (more rows than columns) to the wide format (more columns than rows). This action is called **transposing**. To transpose your data:

1. Highlight and copy the data that you want to transpose including the column labels. You can do this by highlighting **Columns A-H**. In Excel, highlight only the relevant cells (**A1-H45**) instead of the headers.

2. Right-click on **cell I1**. This is where you want the transposed data to start.

3. Hover over **Paste Special** from the right-click menu. Select the **Transposed** option. In Excel, select the **Transpose** icon under the paste options.



Cut
 Copy

Paste Options

Insert
 Transpose
 Delete
 Clear Contents
 Sort

New Comment
 Number Format...
 Hyperlink...
 Show Changes

You should now find the data transformed into the new wide format. At this point, you should remove the original long data from the spreadsheet.

4. Delete the previous long data. The easiest way to do this is to click on **Column A**, so the entire column is highlighted. Then, hold down the **Shift** key and click on **Column H**. You should find these columns highlighted. Right-click on the highlighted area and select **Delete Columns A - H**.

100%
 \$ % .0 .00 123
 Default (Ari...
 10
 B I A

	A	B	C	D	E	F	G	
1	Number	Order Date	Region	Name	Item	Units	Unit Cost	Total
2	1	1/6/19	East	Jones	Pencil	95	1.99	
3	2	1/23/19	Central	Kivell	Binder	50	19.99	
4	3	2/9/19	Central	Jardine	Pencil	36	4.99	
5	4	2/26/19	Central	Gill	Pen			
6	5	3/15/19	West	Sorvino	Pencil			
7	7	4/18/19	Central	Andrews	Pencil			
8	9	5/22/19	West	Thompson	Pencil	32	1.99	
9	10	6/8/19	East	Jones	Binder	60	8.99	
10	11	6/25/19	Central	Morgan	Pencil	90	4.99	
11	12	7/12/19	East	Howard	Binder	29	1.99	
12	13	7/29/19	East	Parent	Binder	81	19.99	
13	15	9/1/19	Central	Smith	Desk	2	125	
14	16	9/18/19	East	Jones	Pen Set	16	15.99	
15	18	10/22/19	East	Jones	Pen	64	8.99	
16	19	11/8/19	East	Parent	Pen	15	19.99	
17	20	11/25/19	Central	Kivell	Pen Set	96	4.99	
18	22	12/29/19	East	Parent	Pen Set	74	15.99	
19	23	1/15/20	Central	Gill	Binder	46	8.99	
20	24	2/1/20	Central	Smith	Binder	87	15	

Cut
 Copy
 Paste
 Paste special
 Insert 8 left
 Delete columns A - H
 Hide columns A - H
 Resize columns A - H
 Group columns A - H
 Ungroup columns A - H
 Sort sheet A → Z
 Sort sheet Z → A
 Randomize range

Your screen should now appear like this:

	A	B	C	D	E	F	G	H	I	J
1	Number	1	2	3	4	5	7	9	10	11
2	Order Date	1/6/19	1/23/19	2/9/19	2/26/19	3/15/19	4/18/19	5/22/19	6/8/19	6/25/19
3	Region	East	Central	Central	Central	West	Central	West	East	Central
4	Name	Jones	Kivell	Jardine	Gill	Sorvino	Andrews	Thompson	Jones	Morgan

5	Item	Pencil	Binder	Pencil	Pen	Pencil	Pencil	Pencil	Binder	Pencil	
6	Units	95	50	36	27	56	75	32	60	90	
7	Unit Cost	1.99	19.99	4.99	19.99	2.99	1.99	1.99	8.99	4.99	
8	Total	189.05	999.5	179.64	539.73	167.44	149.25	63.68	539.4	449.1	

Get rid of extra spaces in cells with string data

Now that you have transposed the data, eliminate the extra spaces in the values of the cells.

1. Highlight the data in the spreadsheet.
2. Click on the **Data** tab, then hover over **Data cleanup** and select **Trim whitespace**.

The screenshot shows the Google Sheets interface with the **Data** tab selected. The **Data cleanup** menu is open, showing options: **Cleanup suggestions**, **Remove duplicates**, and **Trim whitespace**. The background spreadsheet shows a table with columns F, G, and H, containing various data points including dates, names, and numbers.

In Excel, you can use the TRIM command to get rid of white spaces. In any space beneath your data (such as cell **A10**), type =TRIM(A1). Then, drag the bottom right corner of the cell to the bottom right to call the data without the white spaces.

Now all the extra spaces in the cells have been removed.

Change Text Lower/Uppercase/Proper Case

Next, you'll process string data. The easiest way to clean up string data will depend on the spreadsheet program you are using. If you are using Excel, you'll use a simple formula. If you are using Google Sheets, you can use an Add-On to do this with a few clicks. Follow the steps in the relevant section below.

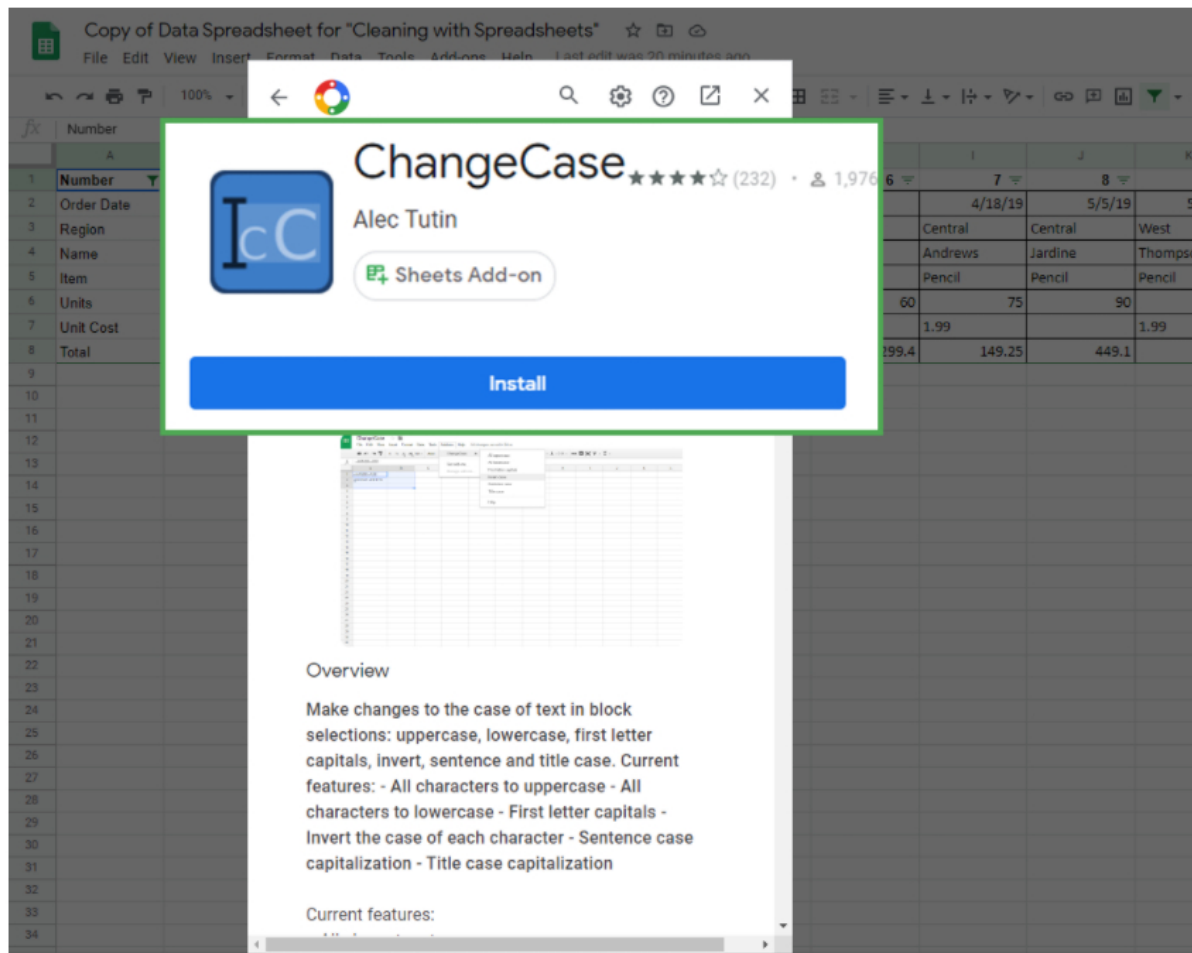
If you are using Microsoft Excel, [this documentation](#) explains how to use a formula to change the case of a text string. Follow these instructions to clean the string text and then move on to the confirmation and reflection section of this activity.

Google sheets

If you're completing this exercise using Google Sheets, you'll need to install an add-in that will give you the functionality needed to easily clean string data and change cases.

Google Sheets Add-on Instructions:

1. Click on the **Add-Ons** option at the top of Google Sheets.
2. Click on **Get add-ons**.
3. Search for **ChangeCase**. It should appear like this:



4. Click on **Install** to install the add-on. It may ask you to login or verify the installation permissions.

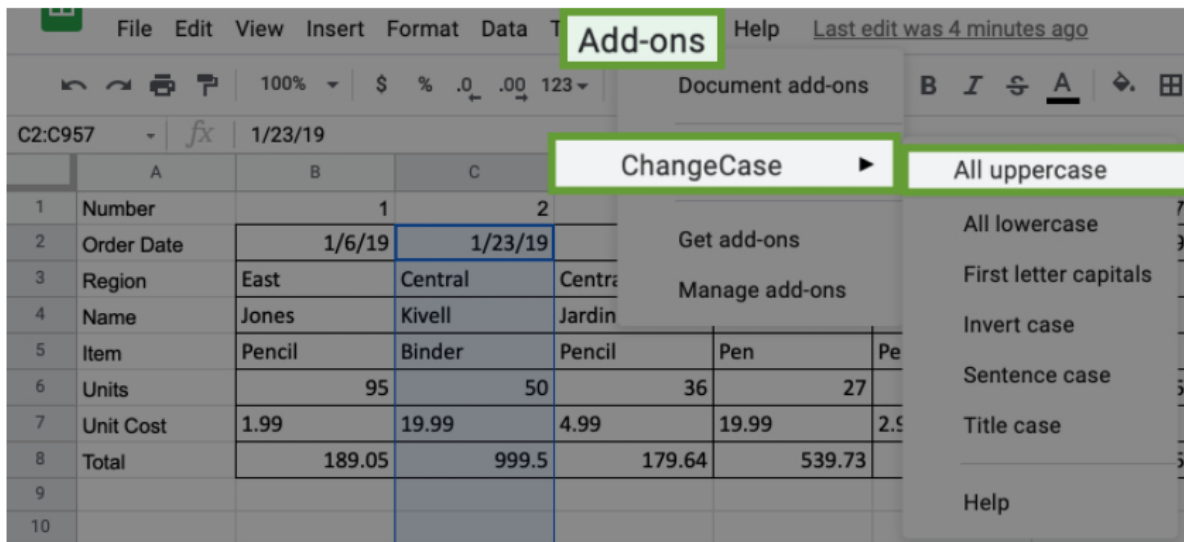
Once you have installed the add-on successfully, you can access it by clicking on the **Add-ons** menu again.

Now, you can change the case of text data that shows up. To change the text in Column C to all uppercase:

1. Click on **Column C**. Be sure to deselect the column header, unless you want to change the case of that as well (which you

don't).

2. Click on the **Add-Ons** tab and select **ChangeCase**. Select the option **All uppercase**. Notice the other options that you could have chosen if needed.



Delete all formatting

If you want to clear the formatting for any or all cells, you can find the command in the **Format** tab. To clear formatting:

1. Select the data for which you want to delete the formatting. In this case, highlight all the data in the spreadsheet by clicking and dragging over **Rows 1-8**.

2. Click the **Format** tab and select the **Clear Formatting** option.

In Excel, go to the **Home** tab, then hover over **Clear** and select **Clear Formats**.

You will notice that all the cells have had their formatting removed.

Confirmation and reflection

Review the final product of the spreadsheet you cleaned during this activity. Which of the following is the rightmost column?

- ☐ Column AZ
- ☐ Column Z
- ☐ Column AA
- ☒ Column AS

✓ **Correct**

In the final product of this activity, the rightmost column is Column AS. You are able to find this information after you

properly transpose the data. Going forward, you can apply what you learned about data cleaning and transposing to work with data in the future.

2. In this activity, you practiced cleaning and transposing data. In the text box below, write 2-3 sentences (40-60 words) in response to each of the following questions:

1 / 1 point

- What was the most challenging part of cleaning the data?
- Why is cleaning and transposing data important for data analysis?
- If you had to clean this data again, what would you do differently? Why?

What was the most challenging part of cleaning the data?

Some tools are too slow.

Why is cleaning and transposing data important for data analysis?

You can work better with the data depending on the format.

If you had to clean this data again, what would you do differently? Why?

I would search a function to change my blanks automatically to 'None'.



Correct

Congratulations on completing this hands-on activity! In this activity, you cleaned and transposed data on a spreadsheet. A good response would include that cleaning is a fundamental step in data science as it greatly increases the integrity of the data.

Good data science results rely heavily on the reliability of the data. Data analysts clean data to make it more accurate and reliable. This is important for making sure that the projects you will work on as a data analyst are completed properly.