

String Functions

how to make your data look the way you want it

Where to find the practice data

<http://bit.ly/22xKXfU>

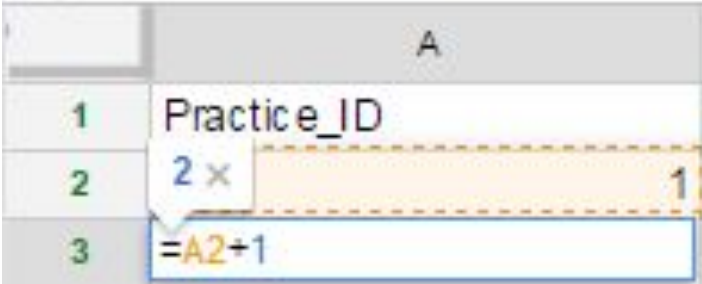
Please download the file, don't open it. If you just open it your edits will apply to everyone!

Creating a unique ID

One handy thing that Sheets will do for you is auto-fill a series. Let's use this to create a unique ID for a table.

This is a number that corresponds to each observation in the data. There should be no repetition of ID's.

- Insert a row to the left of the table
- Name the row something helpful. Suggestion: `unique_ID`
- In the first cell type a 1
- In the next cell type an equals sign, and then click on cell A1. Then type +1. Hit enter.
- A3 should now be equal to 2.
- Then click in the lower right corner of the cell to extend it all the way down the column.



The screenshot shows a Google Sheet with a column labeled 'A'. The first row (row 1) contains the text 'Practice_ID'. The second row (row 2) contains the number '2' in the first cell and '1' in the second cell. The third row (row 3) contains the formula '=A2+1' in the first cell. A dashed orange box highlights the range from row 2 to row 3, indicating the auto-fill operation.

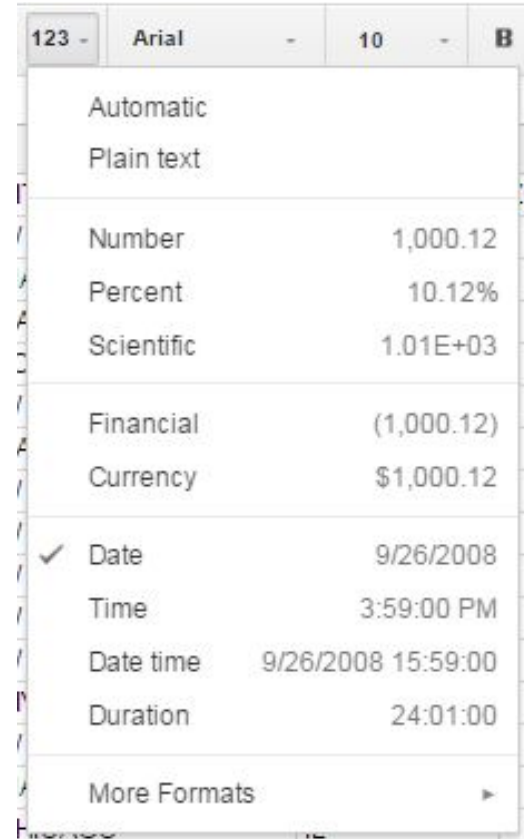
	A
1	Practice_ID
2	2 × 1
3	=A2+1

Changing the format of a field

Sometimes google guesses the wrong format for a field and we want to change it.

In this case take a look at the DOB field. Notice a problem? Let's fix it.

- Click the 'G' above the column to select it.
- Click on the '123' button above to get the dropdown menu.
- Select 'Date'



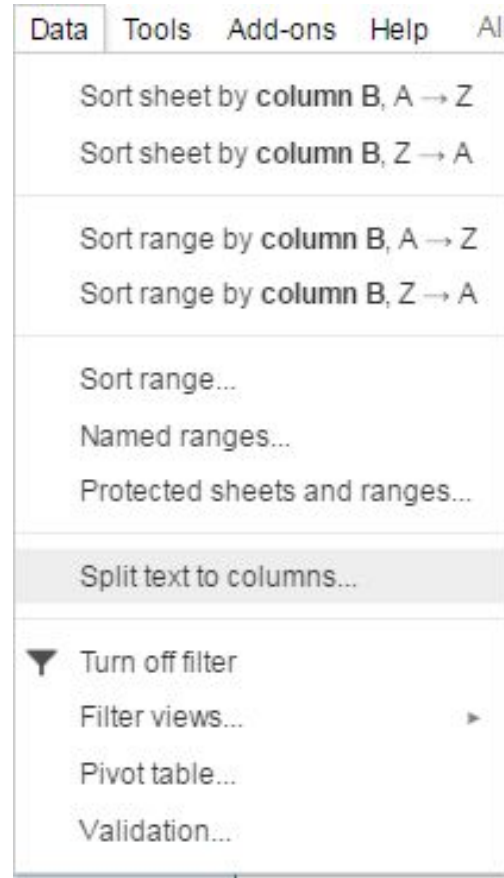
Splitting cells

Sometimes there is too much information in a single column. Sheets has a handy, if imperfect, way of handling this.

We can split the cell into columns using a delimiter. Let's look at the NAME column. What might we want to change here?

Splitting cells

- Let's add 4 empty columns to the right of Column B. (This is one flaw of how Sheets does this - it will overwrite columns if you don't put in empty ones. Better to err with more than less.)
- Select Column B, then do:
Data -> Split text to columns
- Down at the bottom, select 'space' as the separator



Splitting cells

- Scroll through the result.
- What are the pluses and minuses of what we get?
 - William Smith III
 - La Tonya Jones

Sheets can help you out, but you may end up needing to do some cleaning by hand.

Pop quiz!

What do you think the `CONCATENATE` function does? It's okay to use Google to find out.

LEFT/RIGHT/MIDDLE

If we know what number of digits we want to split off from a cell, we can use the LEFT, RIGHT, or MIDDLE functions.

Example: 3-digit ZIP-code

Even in cases when we have a full ZIP-code for each observation in data, it can be useful to use just the first 3-digits. This can be the case when data is sparse, or when we are trying to protect privacy.

LEFT/RIGHT/MIDDLE

Because we want the 3 left-most digits, we will use the 'LEFT' function.

- Select Column G, and insert a row to the left.
- Name it something useful, like ZIP3
- Click in cell F3
- Enter this function, what do you get?

=LEFT(G2,3)

- Double click to extend the function down the entire column.

Transpose

Sometimes we want to switch the columns and rows. Sometimes this is because we want to do a particular calculation, other times it's because of bad table design. Consider the following example. What might we want to change?

	A	B	C	D	E	F	G
1		Person 1	Person 2	Person 3	Person 4	Person 5	Person 6
2	Arrests	2	17	10	19	8	18
3	Convictions	14	2	3	13	7	17

Transpose

- Select a blank cell in the spreadsheet.
- Type this:
 - =TRANSPOSE(
- Then choose the cells you want transposed. If the table is small enough you can do this by holding down the mouse and dragging.
- If the table is really big, you can just enter the cell in the upper left corner : lower right corner.
 - =TRANSPOSE(A2:F9)

Transpose

You should get something like this:

	Arrests	Convictions
Person 1	13	18
Person 2	13	11
Person 3	19	10
Person 4	17	8
Person 5	14	5
Person 6	11	16

Keep exploring!