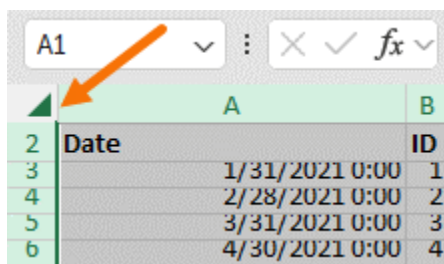# 1: Autofit Columns and Rows

Begin data cleaning by auto-fitting columns and rows to make the data readable. This simple action ensures that all your data is visible, addressing any issues of data being cramped or cut off.

To make your data properly visible, select the whole worksheet by clicking this triangle:

Next, bring your cursor between any 2 column labels within your dataset until it's converted into the double-headed arrow > double-click.

The same can be achieved using the keyboard shortcut CTRL+A to select the data and ALT,H,O,I to auto-fit column width.

Similarly, you can repeat the process to autofit the rows.

|	A	B	C	D
2	Date	ID	Name	Region
3	1/31/2021 0:00	1	John Smith	North
4	2/28/2021 0:00	2	Jane Doe	East
5	3/31/2021 0:00	3	Mike Tyson	West
6	4/30/2021 0:00	4	Anna Belle	South
7	5/31/2021 0:00	5	Chris P. Bacon	East
8	6/30/2021 0:00	6	Peter Parker	
9	7/31/2021 0:00	7	Mary Jane	West
10	8/31/2021 0:00	8	Bruce Wayne	South

The keyboard shortcut for data auto-fitting row height is ALT,H,O,A.

While it's a quick and easy fix, it ensures that no data is overlooked. Once all your data is visible, move on to the next step.

2: Identify and Remove Duplicates

Duplicates are rarely intentional. You can use conditional formatting to highlight duplicates in your data, and then utilize the 'Remove Duplicates' feature under the Data tab to eliminate them, ensuring each row remains unique.

For example, in my dataset, I can quickly identify duplicates in the ID column by going to the Home tab > Conditional Formatting > Highlight Cells Rules > Duplicate values.
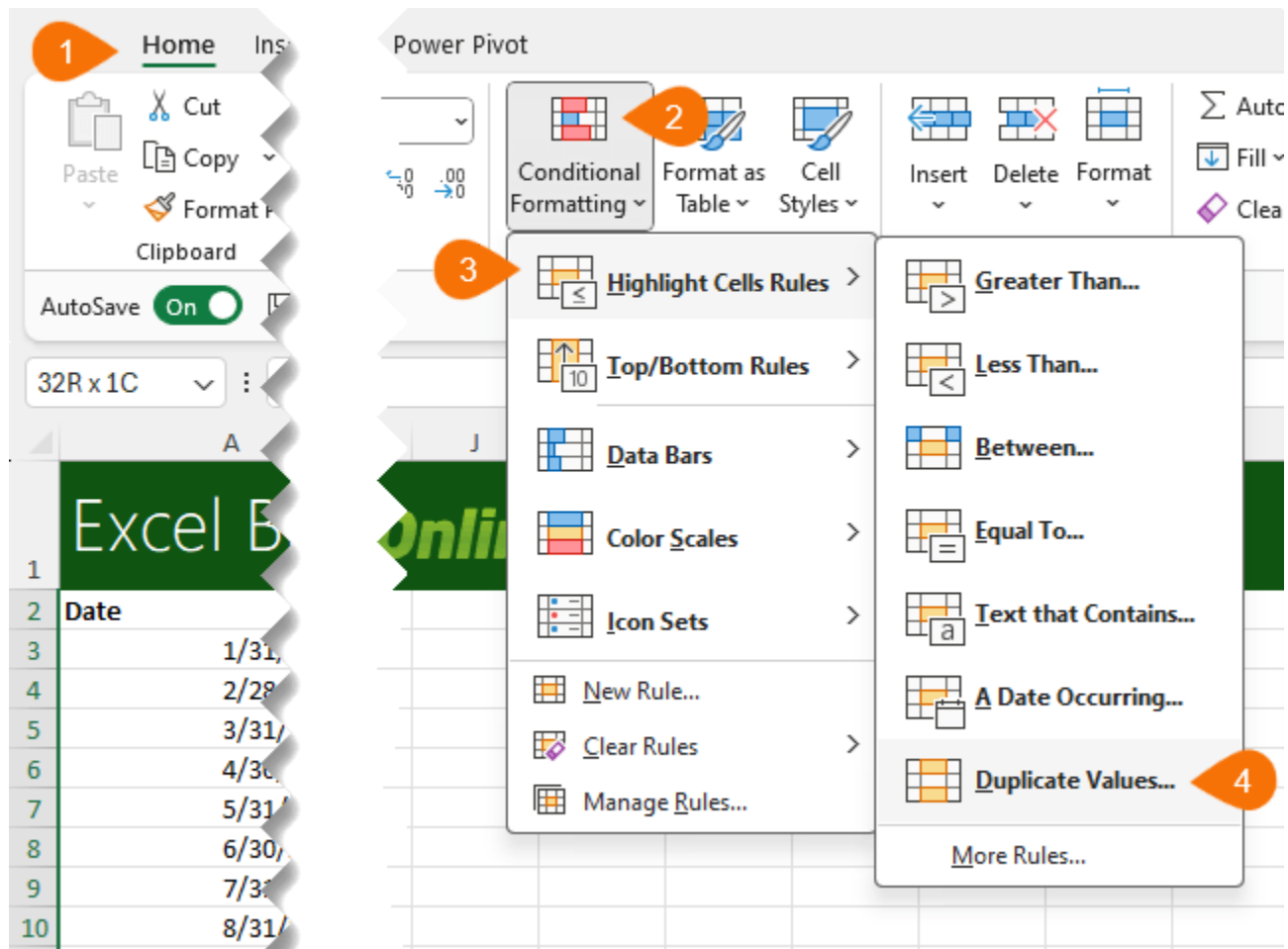
This opens the Duplicate Values dialog box, allowing you to choose the highlight format.
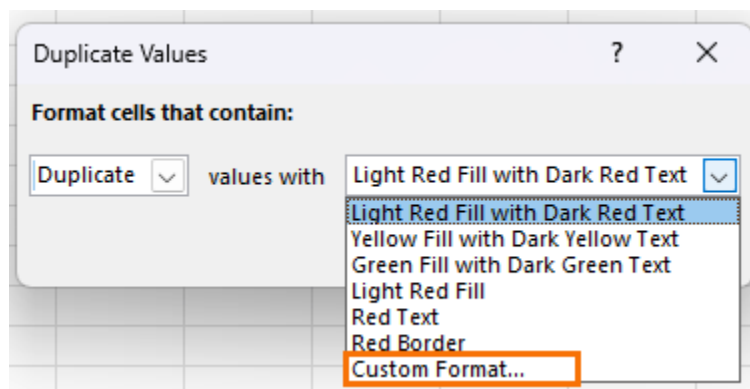

I'll go with the default settings but notice you can choose from these options or create your own custom format too.

Now I can see which rows have duplicate IDs at a glance.

	A	B	C	D	E	F	G	H
2	Date	ID	Name	Region	Rating	Product	Quantity	Price Per Unit
3	1/31/2021 0:00	1	John Smith	North	Good	Magic Wand	10	$20.00
4	2/28/2021 0:00	2	Jane Doe	East	Excelent	Unicorn Horn	15	$10.00
5	3/31/2021 0:00	3	Mike Tyson	West	Poor	Boxing Gloves	0	inf
6	4/30/2021 0:00	4	Anna Belle	South	Average	Fairy Dust	25	$10.00
7	5/31/2021 0:00	5	Chris P. Bacon	East	Good	Bacon Scented Candle	30	$16.67
8	6/30/2021 0:00	6	Peter Parker		Excelent	Web Shooter	0	inf
9	7/31/2021 0:00	7	Mary Jane	West	Poor	Potent Potion	35	$10.00
10	8/31/2021 0:00	8	Bruce Wayne	South	Average	Bat Signal	40	$15.00
11	9/30/2021 0:00	9	Clark Kent	East	Good	Glasses with X-ray Vision	45	$12.22
12	10/31/2021 0:00	10	Diana Prince	North	Excelent	Lasso of Truth	50	$14.00
13	11/30/2021 0:00	11	Tony Stark	West	Poor	Iron Man Suit	5	$160.00
14	12/31/2021 0:00	12	Steve Rogers	South	Average	Captain America Shield	20	$45.00
15	1/31/2022 0:00	13	Natasha Romanoff	East	Good	Black Widow's Bite	0	inf
16	2/28/2022 0:00	14	Bruce Banner		Excelent	Gamma Radiation Serum	30	$36.67
17	4/30/2021 0:00	4	Anna Belle	South	Average	Fairy Dust	25	$10.00
18	5/31/2021 0:00	5	Chris P. Bacon	East	Good	Bacon Scented Candle	30	$16.67
19	6/30/2021 0:00	6	Peter Parker		Excelent	Web Shooter	0	inf
20	3/31/2022 0:00	15	Nick Fury	West	Poor	Eye Patch	35	$34.29

I can also easily remove the duplicates, and eliminate errors from my reports.

To do that, go to the 'Data' tab > 'Remove Duplicates':
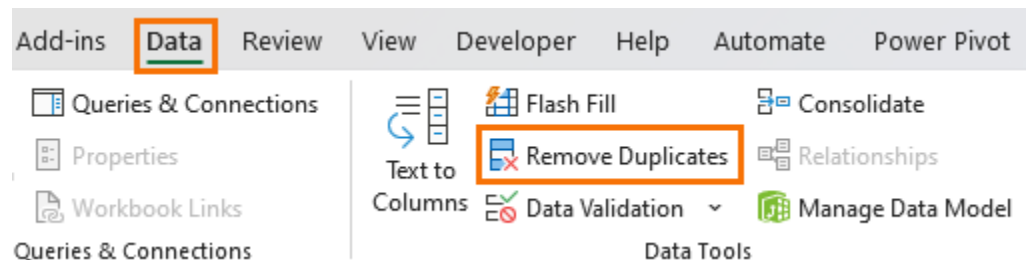
This will open the Remove Duplicates dialog box which lets us choose the columns to compare.

In this case, I want to check for duplicate rows, so I will select all the columns in my table:

Remove Duplicates ? ✕

To delete duplicate values, select one or more columns that contain duplicates.

▤ Select All   ▤ Unselect All   ☑ My data has headers

Columns
☑ Date
☑ ID
☑ Name
☑ Region
☑ Rating
☑ Product
☑ Quantity
☑ Price Per Unit

OK   Cancel

And you can see Excel does the heavy lifting by deleting all the duplicate entries, ensuring each row is unique.

	A	B	C	D	E	F	G	H
2	**Date**	**ID**	**Name**	**Region**	**Rating**	**Product**	**Quantity**	**Price Per Unit**
3	1/31/2021 0:00	1	John Smith	North	Good	Magic Wand	10	$20.00
4	2/28/2021 0:00	2	Jane Doe	East	Excelent	Unicorn Horn	15	$10.00
5	3/31/2021 0:00	3	Mike Tyson	West	Poor	Boxing Gloves	0	inf
6	4/30/2021 0:00	4	Anna Belle	South	Average	Fairy Dust	25	$10.00
7	5/31/2021 0:00	5	Chris P. Bacon	East	Good	Bacon Scented Candle	30	$16.67
8	6/30/2021 0:00	6	Peter Parker		Excelent	Web Shooter	0	inf
9	7/31/2021 0:00	7	Mary Jane	West	Poor	Potent Potion	35	$10.00
10	8/31/2021 0:00	8	Bruce Wayne	South	Average	Bat Signal	40	$15.00
11	9/30/2021 0:00	9	Clark Kent	East	Good	Glasses with X-ray Vision	45	$12.22
12	10/31/2021 0:00	10	Diana Prince	North	Excelent	Lasso of Truth	50	$14.00
13	11/30/2021 0:00	11	Tony Stark	West	Poor	Iron Man Suit	5	$160.00
14	12/31/2021 0:00	12	Steve Rogers	South	Average	Captain America Shield	20	$45.00
15	1/31/2022 0:00	13	Natasha Romanoff	East	Good	Black Widow's Bite	0	inf
16	2/28/2022 0:00	14	Bruce Banner		Excelent	Gamma Radiation Serum	30	$36.67
17	3/31/2022 0:00	15	Nick Fury	West	Poor	Eye Patch	35	$34.29

3: Trim Extra Spaces

Extra spaces can lead to inconsistencies. While some unwanted spaces are obvious and easy to find, others are not always visible.

For example, the extra spaces between Mike and Tyson, Anna and Belle and Peter and Parker are easily visible however, the space in front of Mary Jane is not.

	A	B	C	D	E	F	G	H
2	Date	ID	Name	Region	Rating	Product	Quantity	Price Per Unit
3	1/31/2021 0:00	1	John Smith	North	Good	Magic Wand	10	$20.00
4	2/28/2021 0:00	2	Jane Doe	East	Excelent	Unicorn Horn	15	$10.00
5	3/31/2021 0:00	3	Mike Tyson	West	Poor	Boxing Gloves	0	inf
6	4/30/2021 0:00	4	Anna Belle	South	Average	Fairy Dust	25	$10.00
7	5/31/2021 0:00	5	Chris P. Bacon	East	Good	Bacon Scented Candle	30	$16.67
8	6/30/2021 0:00	6	Peter Parker		Excelent	Web Shooter	0	inf
9	7/31/2021 0:00	7	Mary Jane	West	Poor	Potent Potion	35	$10.00
10	8/31/2021 0:00	8	Bruce Wayne	South	Average	Bat Signal	40	$15.00
11	9/30/2021 0:00	9	Clark Kent	East	Good	Glasses with X-ray Vision	45	$12.22
12	10/31/2021 0:00	10	Diana Prince	North	Excelent	Lasso of Truth	50	$14.00
13	11/30/2021 0:00	11	Tony Stark	West	Poor	Iron Man Suit	5	$160.00
14	12/31/2021 0:00	12	Steve Rogers	South	Average	Captain America Shield	20	$45.00
15	1/31/2022 0:00	13	Natasha Romanoff	East	Good	Black Widow's Bite	0	inf
16	2/28/2022 0:00	14	Bruce Banner		Excelent	Gamma Radiation Serum	30	$36.67
17	3/31/2022 0:00	15	Nick Fury	West	Poor	Eye Patch	35	$34.29

Instead of spotting and removing these one by one, it's best to employ the TRIM function in an empty column beside your data.

For example, I will use the formula TRIM(C3) in cell I3:

	A	B	C	D		I
2	Date	ID	Name	Regio		it
3	1/31/2021 0:00	1	John Smith	North		00 =TRIM(C3)
4	2/28/2021 0:00	2	Jane Doe	East		.00
5	3/31/2021 0:00	3	Mike Tyson	West		
6	4/30/2021 0:00	4	Anna Belle	South		00
7	5/31/2021 0:00	5	Chris P. Bacon	East		.67
8	6/30/2021 0:00	6	Peter Parker			
9	7/31/2021 0:00	7	Mary Jane	West		00

And then, copy it down to fix all the rows:

	A	B	C	R		Trimmed
2	Date	ID	Name			Trimmed
3	1/31/2021 0:00	1	John Smith	No	ᴐ	John Smith
4	2/28/2021 0:00	2	Jane Doe	Eas	ᴑ0	Jane Doe
5	3/31/2021 0:00	3	Mike Tyson	W		Mike Tyson
6	4/30/2021 0:00	4	Anna Belle	Soᵤ	ᴐ	Anna Belle
7	5/31/2021 0:00	5	Chris P. Bacon	Eaₛ	ᴑ7	Chris P. Bacon
8	6/30/2021 0:00	6	Peter Parker			Peter Parker
9	7/31/2021 0:00	7	Mary Jane	Wₑ	ᴐ	Mary Jane
10	8/31/2021 0:00	8	Bruce Wayne	Soᵤ	ᴑ0	Bruce Wayne
11	9/30/2021 0:00	9	Clark Kent	E	ᵎ	Clark Kent
12	10/31/2021 0:00	10	Diana Prince	Nc	ᴐ	Diana Prince
13	11/30/2021 0:00	11	Tony Stark	Wₑ	ᴑ0	Tony Stark
14	12/31/2021 0:00	12	Steve Rogers	Sₜ	ᵎ	Steve Rogers
15	1/31/2022 0:00	13	Natasha Romanoff	Eas		Natasha Romanoff
16	2/28/2022 0:00	14	Bruce Banner		ᴑ7	Bruce Banner
17	3/31/2022 0:00	15	Nick Fury	W	ᵎ	Nick Fury

Once I have trimmed the data, I can copy and paste it to the original column 'as values' via the Paste dropdown > Paste Values > Values:
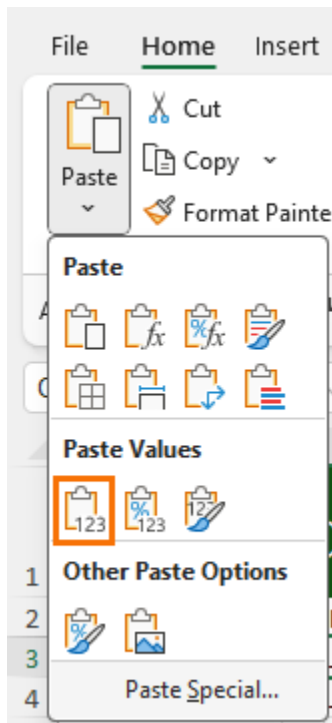
In Microsoft 365, you can also use the new keyboard shortcut CTRL+SHIFT+V to paste as values, while in older versions you can use ALT,E,S,V.

4: Eliminate Blank Cells

Blank cells can interrupt data analysis. To avoid that, you can utilize the 'Go To Special' option under the Find & Select menu to quickly select and fill blank cells, maintaining data uniformity.
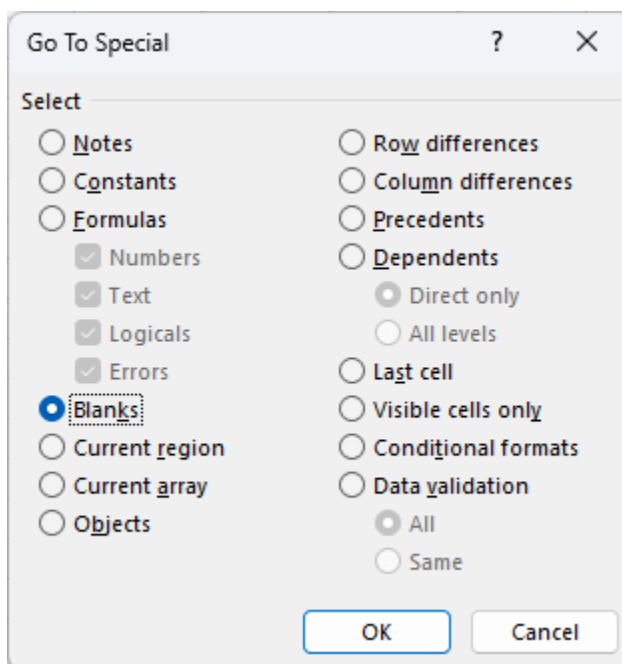
For example, some cells in my Region column are empty:

	A	B	C	D	E	F	G	H
2	Date	ID	Name	Region	Rating	Product	Quantity	Price Per Unit
3	1/31/2021 0:00	1	John Smith	North	Good	Magic Wand	10	$20.00
4	2/28/2021 0:00	2	Jane Doe	East	Excelent	Unicorn Horn	15	$10.00
5	3/31/2021 0:00	3	Mike Tyson	West	Poor	Boxing Gloves	0	inf
6	4/30/2021 0:00	4	Anna Belle	South	Average	Fairy Dust	25	$10.00
7	5/31/2021 0:00	5	Chris P. Bacon	East	Good	Bacon Scented Candle	30	$16.67
8	6/30/2021 0:00	6	Peter Parker		Excelent	Web Shooter	0	inf
9	7/31/2021 0:00	7	Mary Jane	West	Poor	Potent Potion	35	$10.00
10	8/31/2021 0:00	8	Bruce Wayne	South	Average	Bat Signal	40	$15.00
11	9/30/2021 0:00	9	Clark Kent	East	Good	Glasses with X-ray Vision	45	$12.22
12	10/31/2021 0:00	10	Diana Prince	North	Excelent	Lasso of Truth	50	$14.00
13	11/30/2021 0:00	11	Tony Stark	West	Poor	Iron Man Suit	5	$160.00
14	12/31/2021 0:00	12	Steve Rogers	South	Average	Captain America Shield	20	$45.00
15	1/31/2022 0:00	13	Natasha Romanoff	East	Good	Black Widow's Bite	0	inf
16	2/28/2022 0:00	14	Bruce Banner		Excelent	Gamma Radiation Serum	30	$36.67
17	3/31/2022 0:00	15	Nick Fury	West	Poor	Eye Patch	35	$34.29

I will first select my data > Go to the 'Find & Select' menu > Select 'Go To Special':
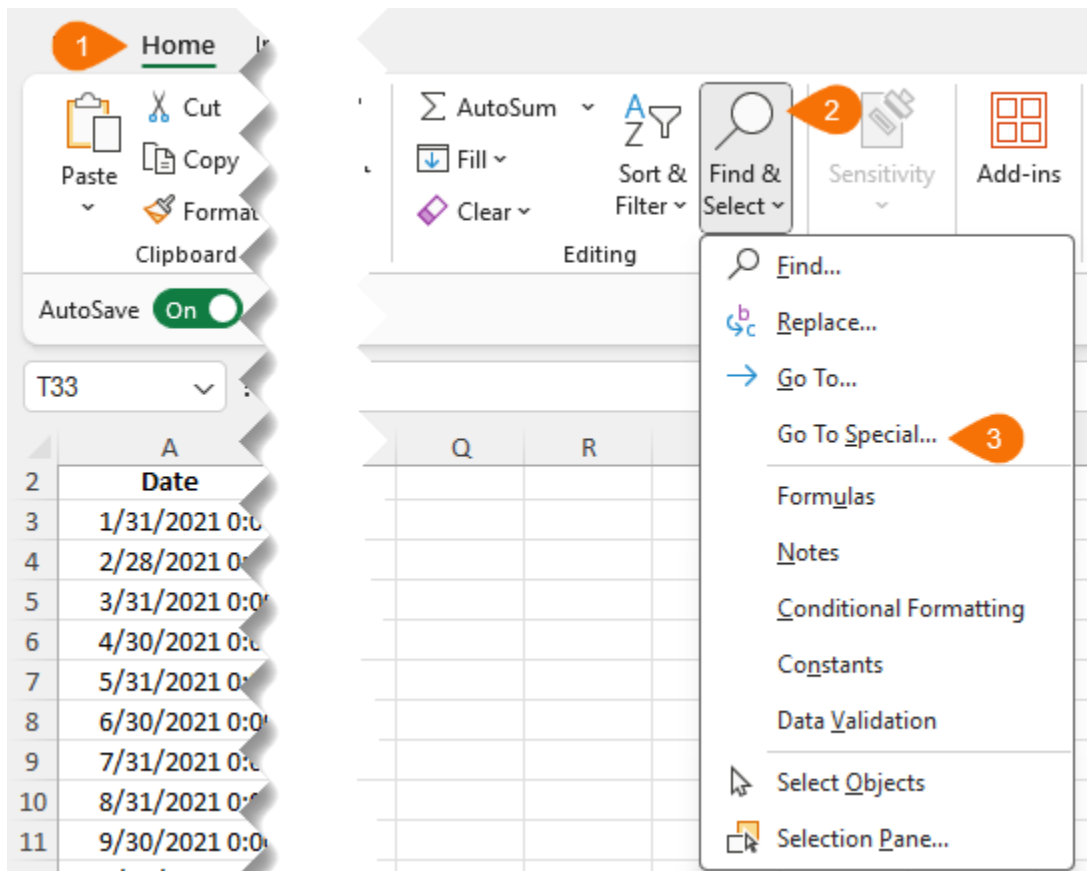
This will open the 'Go To Special' dialog box in which I need to choose 'Blanks':

And within seconds, all the blank cells in the dataset get selected:

	A	B	C	D	E	F	G	H
2	Date	ID	Name	Region	Rating	Product	Quantity	Price Per Unit
3	1/31/2021 0:00	1	John Smith	North	Good	Magic Wand	10	$20.00
4	2/28/2021 0:00	2	Jane Doe	East	Excelent	Unicorn Horn	15	$10.00
5	3/31/2021 0:00	3	Mike Tyson	West	Poor	Boxing Gloves	0	inf
6	4/30/2021 0:00	4	Anna Belle	South	Average	Fairy Dust	25	$10.00
7	5/31/2021 0:00	5	Chris P. Bacon	East	Good	Bacon Scented Candle	30	$16.67
8	6/30/2021 0:00	6	Peter Parker		Excelent	Web Shooter	0	inf
9	7/31/2021 0:00	7	Mary Jane	West	Poor	Potent Potion	35	$10.00
10	8/31/2021 0:00	8	Bruce Wayne	South	Average	Bat Signal	40	$15.00
11	9/30/2021 0:00	9	Clark Kent	East	Good	Glasses with X-ray Vision	45	$12.22
12	10/31/2021 0:00	10	Diana Prince	North	Excelent	Lasso of Truth	50	$14.00
13	11/30/2021 0:00	11	Tony Stark	West	Poor	Iron Man Suit	5	$160.00
14	12/31/2021 0:00	12	Steve Rogers	South	Average	Captain America Shield	20	$45.00
15	1/31/2022 0:00	13	Natasha Romanoff	East	Good	Black Widow's Bite	0	inf
16	2/28/2022 0:00	14	Bruce Banner		Excelent	Gamma Radiation Serum	30	$36.67
17	3/31/2022 0:00	15	Nick Fury	West	Poor	Eye Patch	35	$34.29
18	4/30/2022 0:00	16	Phil Coulson		Average	Agent ID Card	0	inf

I can simply fill a placeholder value as 'TBA' in all these cells.

To do that, I can type TBA in cell D8 and then press CTRL+ENTER to enter it in all selected cells in one go.

	A	B	C	D	E	F	G	H
2	**Date**	**ID**	**Name**	**Region**	**Rating**	**Product**	**Quantity**	**Price Per Unit**
3	1/31/2021 0:00	1	John Smith	North	Good	Magic Wand	10	$20.00
4	2/28/2021 0:00	2	Jane Doe	East	Excelent	Unicorn Horn	15	$10.00
5	3/31/2021 0:00	3	Mike Tyson	West	Poor	Boxing Gloves	0	inf
6	4/30/2021 0:00	4	Anna Belle	South	Average	Fairy Dust	25	$10.00
7	5/31/2021 0:00	5	Chris P. Bacon	East	Good	Bacon Scented Candle	30	$16.67
8	6/30/2021 0:00	6	Peter Parker	TBA	Excelent	Web Shooter	0	inf
9	7/31/2021 0:00	7	Mary Jane	West	Poor	Potent Potion	35	$10.00
10	8/31/2021 0:00	8	Bruce Wayne	South	Average	Bat Signal	40	$15.00
11	9/30/2021 0:00	9	Clark Kent	East	Good	Glasses with X-ray Vision	45	$12.22
12	10/31/2021 0:00	10	Diana Prince	North	Excelent	Lasso of Truth	50	$14.00
13	11/30/2021 0:00	11	Tony Stark	West	Poor	Iron Man Suit	5	$160.00
14	12/31/2021 0:00	12	Steve Rogers	South	Average	Captain America Shield	20	$45.00
15	1/31/2022 0:00	13	Natasha Romanoff	East	Good	Black Widow's Bite	0	inf
16	2/28/2022 0:00	14	Bruce Banner	TBA	Excelent	Gamma Radiation Serum	30	$36.67
17	3/31/2022 0:00	15	Nick Fury	West	Poor	Eye Patch	35	$34.29
18	4/30/2022 0:00	16	Phil Coulson	TBA	Average	Agent ID Card	0	inf
19	5/31/2022 0:00	17	Peggy Carter	East	Good	Vintage Pistol	40	$35.00
20	6/30/2022 0:00	18	Howard Stark	North	Excelent	Arc Reactor	45	$33.33
21	7/31/2022 0:00	19	Hank Pym	West	Poor	Ant-Man Suit	50	$32.00

Alternatively, if I want, I can also copy the value from the cell above.

To do that, with the three empty cells selected, Type = in cell D8 > Press the up arrow > Press CTRL+ENTER.

This will fill the region of the cell above each of these blank cells, respectively.

	A	B	C	D	E	F	G	H
2	Date	ID	Name	Region	Rating	Product	Quantity	Price Per Unit
3	1/31/2021 0:00	1	John Smith	North	Good	Magic Wand	10	$20.00
4	2/28/2021 0:00	2	Jane Doe	East	Excelent	Unicorn Horn	15	$10.00
5	3/31/2021 0:00	3	Mike Tyson	West	Poor	Boxing Gloves	0	inf
6	4/30/2021 0:00	4	Anna Belle	South	Average	Fairy Dust	25	$10.00
7	5/31/2021 0:00	5	Chris P. Bacon	East	Good	Bacon Scented Candle	30	$16.67
8	6/30/2021 0:00	6	Peter Parker	East	Excelent	Web Shooter	0	inf
9	7/31/2021 0:00	7	Mary Jane	West	Poor	Potent Potion	35	$10.00
10	8/31/2021 0:00	8	Bruce Wayne	South	Average	Bat Signal	40	$15.00
11	9/30/2021 0:00	9	Clark Kent	East	Good	Glasses with X-ray Vision	45	$12.22
12	10/31/2021 0:00	10	Diana Prince	North	Excelent	Lasso of Truth	50	$14.00
13	11/30/2021 0:00	11	Tony Stark	West	Poor	Iron Man Suit	5	$160.00
14	12/31/2021 0:00	12	Steve Rogers	South	Average	Captain America Shield	20	$45.00
15	1/31/2022 0:00	13	Natasha Romanoff	East	Good	Black Widow's Bite	0	inf
16	2/28/2022 0:00	14	Bruce Banner	East	Excelent	Gamma Radiation Serum	30	$36.67
17	3/31/2022 0:00	15	Nick Fury	West	Poor	Eye Patch	35	$34.29
18	4/30/2022 0:00	16	Phil Coulson	West	Average	Agent ID Card	0	inf
19	5/31/2022 0:00	17	Peggy Carter	East	Good	Vintage Pistol	40	$35.00

5: Spell Check

Always run a spell check to correct typos and misspellings, a crucial step for maintaining professionalism in your reports.

For example, in my dataset, I should run a spellcheck on columns D through F, excluding the name and numeric columns.

To do that, I will select these columns > Go to the 'Review' tab > Click 'Spelling'. Alternatively, I can also use the keyboard shortcut F7.

This will open the Spelling dialog box which goes through each word in the selected cells to find spelling errors.
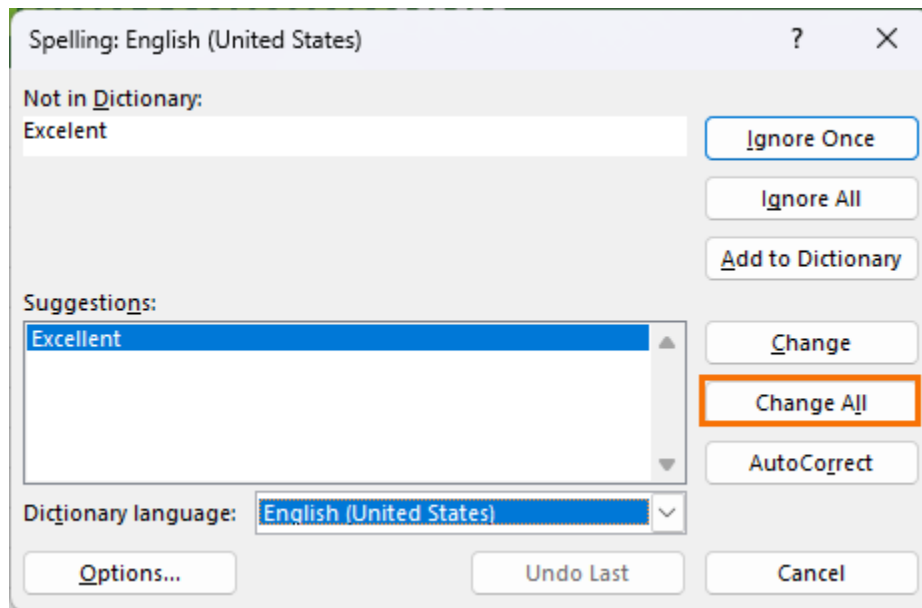
Here, it's found a spelling error in the word 'Excelent', and suggests the correct spelling 'Excellent'.

We can either ignore it or change it. It's better to change it and by clicking Change All, we ensure that any other occurrence of this misspelling is also rectified.

We can also select the 'Add to Dictionary' option for proper nouns such as Asgard that are not recognized English words if this is a common name we'll be using in other files.
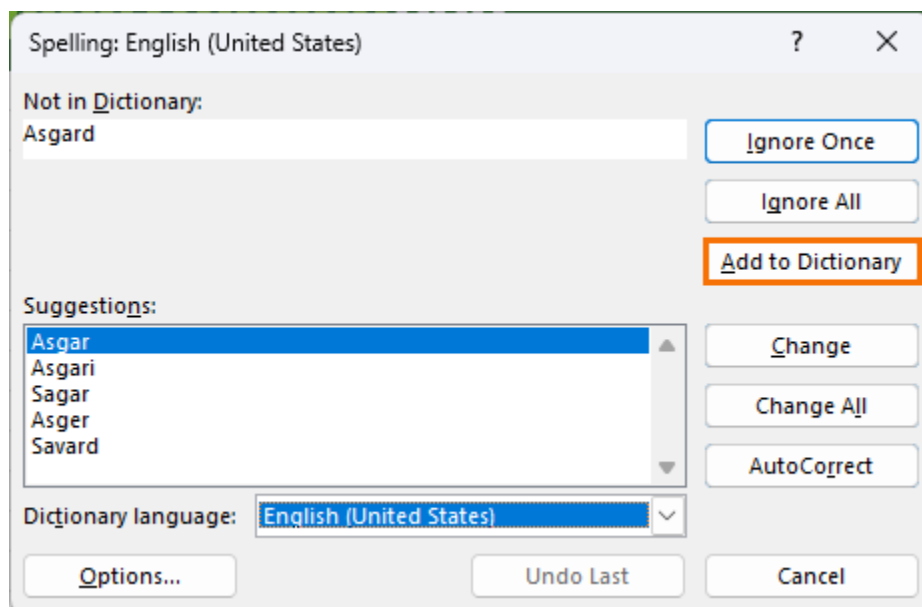
And we're good to go!
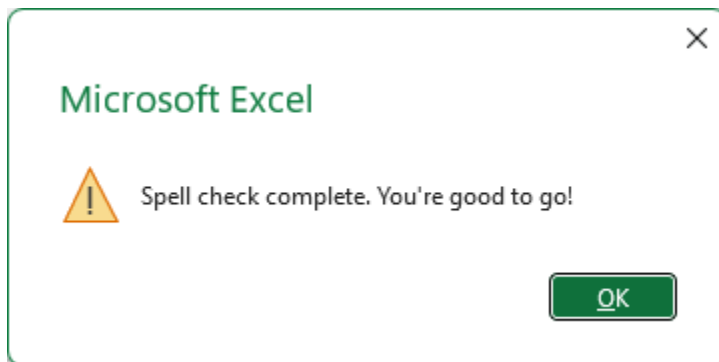
6: Data Validation

While knowing data cleaning techniques is recommended, knowing how to prevent future errors is another skill to have in your Excel quiver.
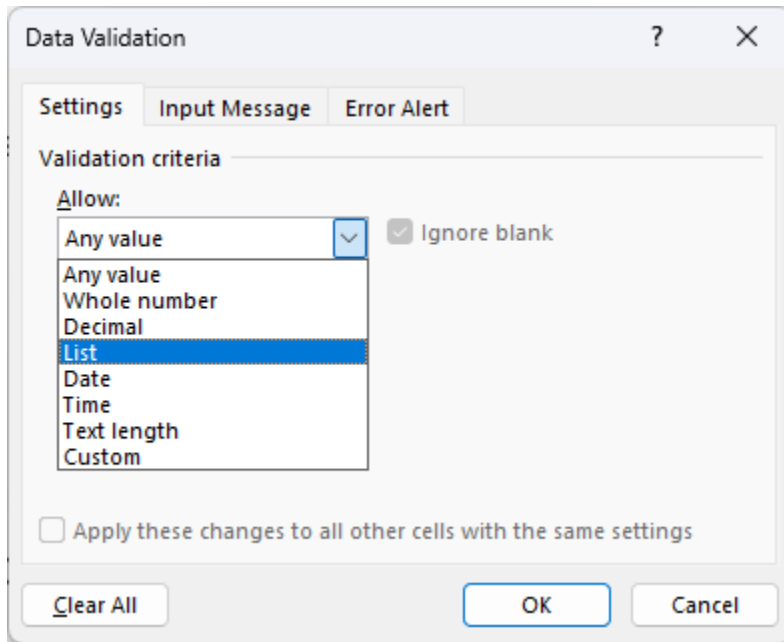
Setting up data validation rules, like creating drop-down lists, is one of the top ways of preventing future data cleaning, as it enforces data integrity from the start.

For example, in my dataset, I can set up a dropdown list for the region column.

To do that, I'll select the region column > Go to the Data tab > Select Data Validation:
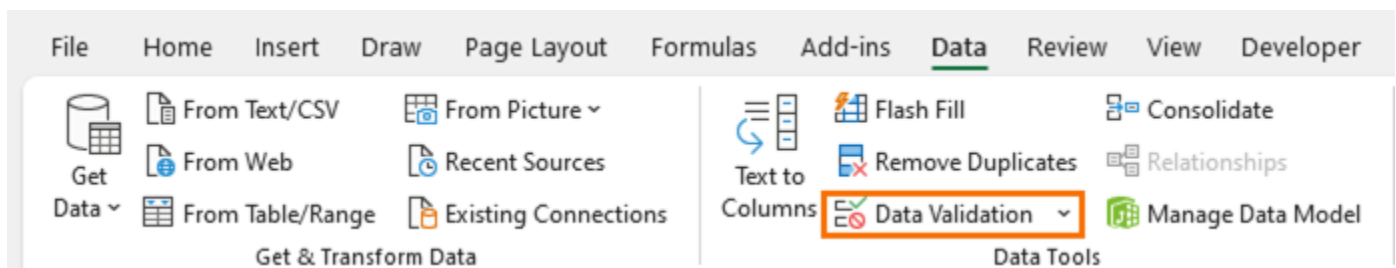
From the 'Allow' dropdown, I will select 'List':
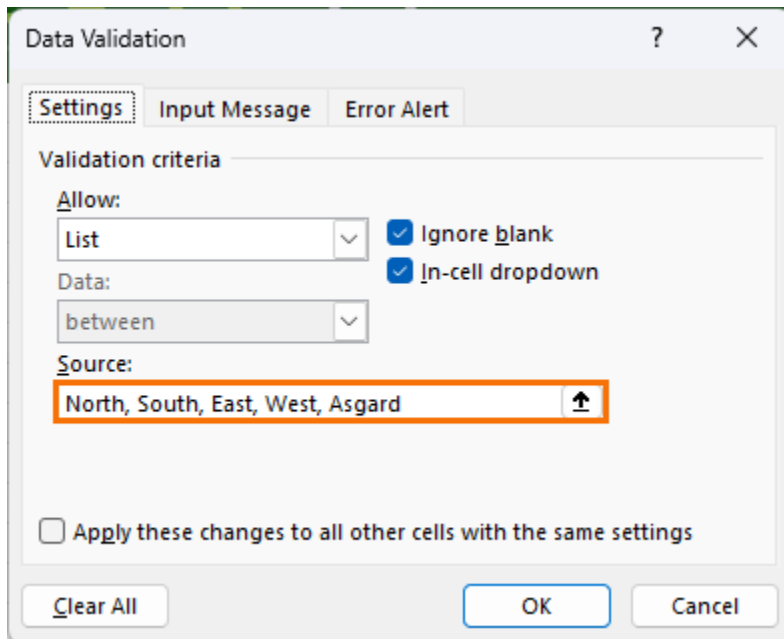
For the source, I can either reference a range of cells that contain the values, or I can just type them in with a comma separator.

I'll type in North, South, East, West, Asgard:

Now, whenever I will add new data, I don't need to type in the region, I can simply choose from the dropdown list:

	Date	ID	Name	Region	Rating	Product	Quantity	Price Per Unit
3	1/31/2021 0:00	1	John Smith	North ▾	ood	Magic Wand	10	$20.00
4	2/28/2021 0:00	2	Jane Doe	North	xcellent	Unicorn Horn	15	$10.00
5	3/31/2021 0:00	3	Mike Tyson	South	oor	Boxing Gloves	0	inf
6	4/30/2021 0:00	4	Anna Belle	East	verage	Fairy Dust	25	$10.00
7	5/31/2021 0:00	5	Chris P. Bacon	West	ood	Bacon Scented Candle	30	$16.67
8	6/30/2021 0:00	6	Peter Parker	Asgard	Excellent	Web Shooter	0	inf
9	7/31/2021 0:00	7	Mary Jane		Poor	Potent Potion	35	$10.00
10	8/31/2021 0:00	8	Bruce Wayne		Average	Bat Signal	40	$15.00
11	9/30/2021 0:00	9	Clark Kent		Good	Glasses with X-ray Vision	45	$12.22
12	10/31/2021 0:00	10	Diana Prince		Excellent	Lasso of Truth	50	$14.00
13	11/30/2021 0:00	11	Tony Stark		Poor	Iron Man Suit	5	$160.00
14	12/31/2021 0:00	12	Steve Rogers		Average	Captain America Shield	20	$45.00
15	1/31/2022 0:00	13	Natasha Romanoff		Good	Black Widow's Bite	0	inf
16	2/28/2022 0:00	14	Bruce Banner		Excellent	Gamma Radiation Serum	30	$36.67
17	3/31/2022 0:00	15	Nick Fury		Poor	Eye Patch	35	$34.29

And if I enter something not in the list, I get an error, preventing unwanted values and future data cleaning:

Name	Region	Rating	Product	Quantity	Price Per Unit
John Smith	Earth ▾				? ✕
Jane Doe	East	Exc			
Mike Tyson	West	Po	**Microsoft Excel**		
Anna Belle	South	Av			
Chris P. Bacon	East	Go	⊗ This value doesn't match the data validation		
Peter Parker		Exc	restrictions defined for this cell.		
Mary Jane	West	Po			
Bruce Wayne	South	Av	Retry Cancel		
Clark Kent	East	Go			
Diana Prince	North	Excellent	Lasso of Truth	50	$14.00

7: Use Tables

Another preventive data cleaning technique is using Excel tables

Storing data in an Excel Table format can make data easier to manage, format, and analyse.

But before you can insert a table, make sure your data is in a tabular layout, and each column contains the same type of data.

In our example,

1. We have columns for the date, ID, name, region, and so on
2. Each row represents a unique record
3. Column headers are also in a single row, i.e. they're not split across multiple rows leading to merged cells.

	A	B	C	D	E	F	G	H
2	Date	ID	Name	Region	Rating	Product	Quantity	Price Per Unit
3	31/01/2021	1	John Smith	North	Good	Magic Wand	10	$20.00
4	28/02/2021	2	Jane Doe	East	Excellent	Unicorn Ho	15	$10.00
5	31/03/2021	3	Mike Tyson	West				
6	30/04/2021	4	Anna Belle	South				0.00
7	31/05/2021	5	Chris P. Bacon	East				6.67
8	30/06/2021	6	Peter Parker	TBA				
9	31/07/2021	7	Mary Jane	West	Poor	Potent Potion	35	$10.00
10	31/08/2021	8	Bruce Wayne	South	Average	Bat Signal	40	$15.00
11	30/09/2021	9	Clark Kent	East	Good	Glasses with X-ray Vision	45	$12.22
12	31/10/2021	10	Diana Prince	North	Excellent	Lasso of Truth	50	$14.00
13	30/11/2021	11	Tony Stark	West	Poor	Iron Man Suit	5	$160.00
14	31/12/2021	12	Steve Rogers		Average	Captain America Shield	20	$45.00
15	31/01/2022	13	Natasha Ro			's Bite	0	inf
16	28/02/2022	14	Bruce Bann			ation Serum	30	$36.67
17	31/03/2022	15	Nick Fury	West	Poor	Eye Patch	35	$34.29
18	30/04/2022	16	Phil Coulson	TBA	Average	Agent ID Card	0	inf
19	31/05/2022	17	Peggy Carter	East	Good	Vintage Pistol	40	$35.00
20	30/06/2022	18	Howard Stark	North	Excellent	Arc Reactor	45	$33.33
21	31/07/2022	19	Hank Pym	West	Poor	Ant-Man Suit	50	$32.00

- Single row of headers
- No merged cells
- Same 'type' of data in each column

Each row represents a unique record

To insert a table, use the keyboard shortcut CTRL+T. This will open the Create Table dialog box:

Create Table ? ✕

Where is the data for your table?

$A$2:$H$30

☑ My table has headers

OK   Cancel

As my table has headers, I will check the box and click OK.

On inserting a table, you can instantly see the data is a lot easier to work with.

	A	B	C	D	E
2	Date	ID	Name	Region	Rating
3	1/31/2021 0:00	1	John Smith	North	Good
4	2/28/2021 0:00	2	Jane Doe	East	Excellent
5	3/31/2021 0:00	3	Mike Tyson	West	Poor
6	4/30/2021 0:00	4	Anna Belle	South	Average
7	5/31/2021 0:00	5	Chris P. Bacon	East	Good
8	6/30/2021 0:00	6	Peter Parker	East	Excellent
9	7/31/2021 0:00	7	Mary Jane	West	Poor
10	8/31/2021 0:00	8	Bruce Wayne	South	Average
11	9/30/2021 0:00	9	Clark Kent	East	Good
12	10/31/2021 0:00	10	Diana Prince	North	Excellent
13	11/30/2021 0:00	11	Tony Stark	West	Poor
14	12/31/2021 0:00	12	Steve Rogers	South	Average
15	1/31/2022 0:00	13	Natasha Romanoff	East	Good

1. The headers are clearly formatted to differentiate them from the data
2. The rows are banded to allow you to easily glance across a row, which is super important in wide tables.
3. On scrolling down, the headers automatically pin to the top of the sheet:

A2			fx	Date			
	Date	ID	Name	Region	Rating	Product	Quantity
5	31/03/2021 0:00	3	Mike Tyson	West	Poor	Boxing Gloves	0
6	30/04/2021 0:00	4	Anna Belle	South	Average	Fairy Dust	25
7	31/05/2021 0:00	5	C. P. Bacon	East	Good	Bacon Scented Candle	30
8	30/06/2021 0:00	6			Excellent	Web Shooter	0
9	31/07/2021 0:00	7			Poor	Potent Potion	35
10	31/08/2021 0:00	8			Average	Bat Signal	40
11	30/09/2021 0:00	9	Clark Kent	East	Good	Glasses with X-ray Vision	45
12	31/10/2021 0:00	10	Diana Prince	North	Excellent	Lasso of Truth	50
13	30/11/2021 0:00	11	Tony Stark	West	Poor	Iron Man Suit	5
14	31/12/2021 0:00	12	Steve Rogers	South	Average	Captain America Shield	20

Headers auto pinned at the top

4. Each column has a filter button, allowing you to easily sort or filter your data:

Date	ID	Name	Region	Rating	Product	Quantity
4	2/28/2021 0:00	2	Ja			Unicorn Horn	15
5	3/31/2021 0:00	3	M			Boxing Gloves	0
6	4/30/2021 0:00	4	Ar			Fairy Dust	25
7	5/31/2021 0:00	5	Ch			Bacon Scented Candle	30
8	6/30/2021 0:00	6	Pe			Web Shooter	0
9	7/31/2021 0:00	7	M			Potent Potion	35
10	8/31/2021 0:00	8	Br			Bat Signal	40
11	9/30/2021 0:00	9	Cl			Glasses with X-ray Vision	45
12	10/31/2021 0:00	10	Di			Lasso of Truth	50
13	11/30/2021 0:00	11	To			Iron Man Suit	5
14	12/31/2021 0:00	12	St			Captain America Shield	20
15	1/31/2022 0:00	13	Na			Black Widow's Bite	0
16	2/28/2022 0:00	14	Br			Gamma Radiation Serum	30
17	3/31/2022 0:00	15	Ni			Eye Patch	35
18	4/30/2022 0:00	16	Ph			Agent ID Card	0
19	5/31/2022 0:00	17	Pe			Vintage Pistol	40
20	6/30/2022 0:00	18	Ho			Arc Reactor	45
21	7/31/2022 0:00	19	Ha			Ant-Man Suit	50
22	8/31/2022 0:00	20	Ja			Wasp's Wings	55
23	9/30/2022 0:00	21	Ku			Comic Book	60
24	10/31/2022 0:00	22	Ge			Drawing Pad	0
25	11/30/2022 0:00	23	Ro			Notepads	65
26	12/31/2022 0:00	24	To			Pen Set	70
27	1/31/2023 0:00	25	Loki Laufeyson	Asgard	Mischief	Trickster's Hat	75
28	2/28/2023 0:00	26	Thor Odinson	Asgard	Worthy	Mjolnir	80

Filter menu overlay:
- A↓ Sort A to Z
- Z↓ Sort Z to A
- Sort by Color >
- Sheet View >
- Clear Filter From "Rating"
- Filter by Color >
- Text Filters >
- Search
  - ☑ (Select All)
  - ☑ Average
  - ☑ Excellent
  - ☑ Good
  - ☑ Leader
  - ☑ Mischief
  - ☑ Poor
  - ☑ Spy
  - ☑ Worthy
- OK Cancel

5. You can also choose from different styles on the Table Design contextual tab > Table Styles:

Table Design
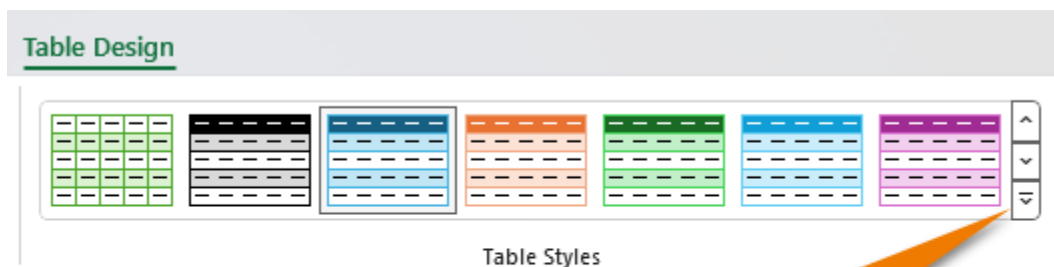
Table Styles

click for more styles or create your own.

6. And, you can add a total row and change the aggregation method as required:

8: Handle Errors with IFERROR

Another great data cleaning technique involves using the IFERROR function to neatly manage and display errors in your dataset, preventing unsightly error messages from cluttering your reports.

For example, let's add a column in our dataset to calculate the Sales amount by multiplying Quantity and Price per Unit.

	G	H	I
2	Quantity ▼	Price Per Unit ▼	Sales ▼
3	10	$20.00	=[@Quantity]*[@[Price Per Unit]]
4	15	$10.00	$150.00
5	0	inf	#VALUE!
6	25	$10.00	$250.00
7	30	$16.67	$500.10
8	0	inf	#VALUE!
9	35	$10.00	$350.00

Notice the formula refers to the columns by name instead of G2*H2? These are called table structured references and Excel inserts them automatically when your data is formatted in an Excel table, and they make writing formulas super easy.

Also, when I press enter, the table copies the formula down the column for me, I don't even need to drag it down.

But wherever the price per unit is missing, the formula gives the #VALUE error.

Instead of displaying the error, we can return something else by wrapping the formula in the following IFERROR function:

=IFERROR([@Quantity]*[@[Price Per Unit]], "Missing")

G	H	I
Quantity ▼	Price Per Unit ▼	Sales ▼
10	$20.00	=IFERROR([@Quantity]*[@[Price Per Unit]], "Missing")
15	$10.00	$150.00
0	inf	Missing
25	$10.00	$250.00
30	$16.67	$500.10
0	inf	Missing
35	$10.00	$350.00

This will return 'Missing' in the rows with missing prices, instead of those ugly #VALUE errors, making your data look error-free.

Alternatively, if you don't want to mix data types in a column (which is never a good idea in your source data), you can use the below formula that returns a blank value instead of 'Missing.'

=IFERROR([@Quantity]*[@[Price Per Unit]] , "")

| G	H	I
Quantity ▼	Price Per Unit ▼	Sales ▼
10	$20.00	=IFERROR([@Quantity]*[@[Price Per Unit]], "")
15	$10.00	$150.00
0	inf	
25	$10.00	$250.00
30	$16.67	$500.10
0	inf	

Keep in mind that you can return anything in place of an error, even another formula.

9: Number Formats

While number formats are great for presenting your data, using plain number formats during the analysis phase avoids unnecessary clutter and complexity in your dataset.

Here are the data cleaning steps for using the right number formats:

1. Use 'General' or 'Number' formats, instead of 'Currency' or 'Accounting' formats during the analysis phase
2. Always use comma separators while working with large numbers
3. Don't use the comma icon from the ribbon as it converts the number into the 'Accounting' format
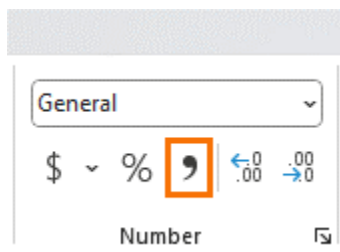
4. Instead, Open the Format Cells dialog box using the shortcut key CTRL+1 > Select Number Format > Check the box for 'Use 1000 Separator'.

Format Cells dialog box showing the Number tab with Category list (General, Number highlighted, Currency, Accounting, Date, Time, Percentage, Fraction, Scientific, Text, Special, Custom), Decimal places set to 2, "Use 1000 Separator (,)" checkbox checked and highlighted, and Negative numbers options (-1,234.10 highlighted, 1,234.10, (1,234.10), (1,234.10)).

5. Also, if you have dates, format them to 'Short Date' if you don't need the time component.

10: Find & Replace

One of the most useful data cleaning techniques leverages the Find & Replace tool for bulk corrections across your dataset to maintain data consistency.

For example, in our dataset, we can replace 'inf', because as we've seen with the IFERROR example, text in a numeric column can cause problems.

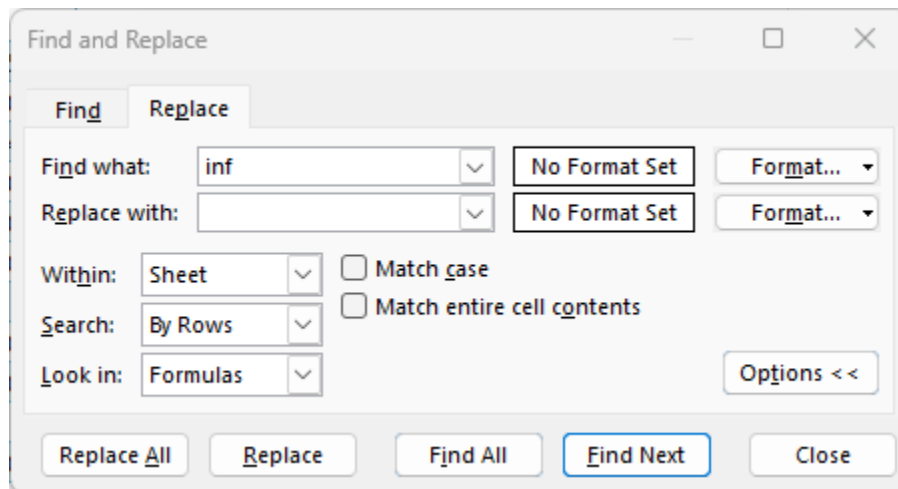To bulk replace 'inf', go to the 'Home' tab > Find and Select > Replace:
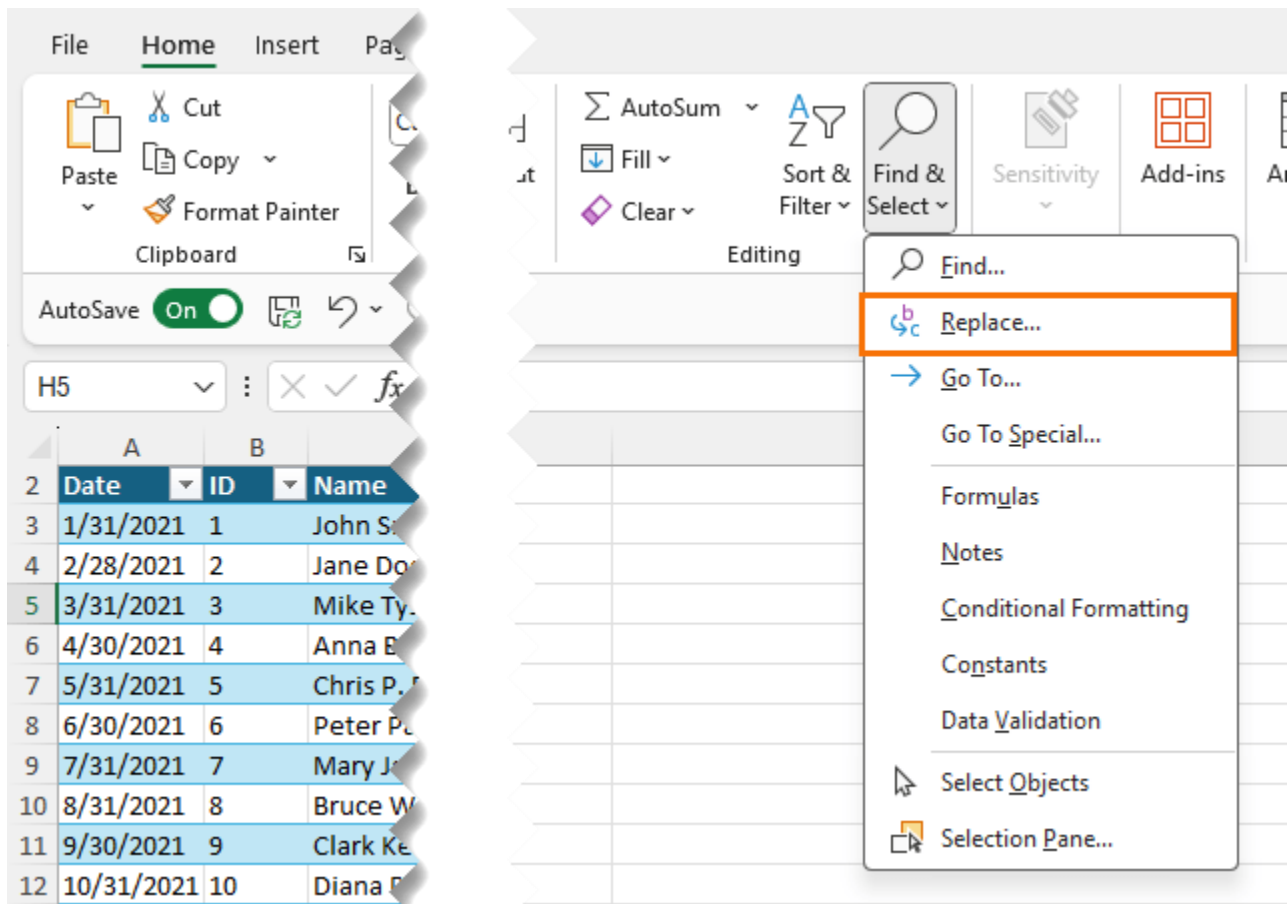
Or use the keyboard shortcut, CTRL+H.

This will open the Find and Replace Dialog box.

Here, enter 'inf' in the 'Find What' field and the value you want to replace it with in the 'Replace With' field.

I'll leave 'Replace With' blank as I want to have an empty cell.

Find and Replace

| | | | | | — | □ | ✕ |

Find | **Replace**

Find what: `inf` ▾ | No Format Set | Format... ▾

Replace with: ▾ | No Format Set | Format... ▾

Within: Sheet ▾ | ☐ Match case

Search: By Rows ▾ | ☐ Match entire cell contents

Look in: Formulas ▾ | Options <<

[Replace All] [Replace] [Find All] [Find Next] [Close]

With Options expanded, notice you can also replace formats and customize where and how the Find and Replace searches the value you want to replace.

Once you've selected the desired options, click 'Replace All', and you're good to go!

|	F	G	H	I
2	Product ▾	Quantity ▾	Price Per Unit ▾	Sales ▾
3	Magic Wand	10	20.00	200.00
4	Unicorn Horn	15	10.00	150.00
5	Boxing Gloves	0		0.00
6	Fairy Dust	25	10.00	250.00
7	Bacon Scented Candle	30	16.67	500.10
8	Web Shooter	0		0.00
9	Potent Potion	35	10.00	350.00
10	Bat Signal	40	15.00	600.00
11	Glasses with X-ray Vision	45	12.22	549.90
12	Lasso of Truth	50	14.00	700.00
13	Iron Man Suit	5	160.00	800.00
14	Captain America Shield	20	45.00	900.00
15	Black Widow's Bite	0		0.00
16	Gamma Radiation Serum	30	36.67	1,100.10
17	Eye Patch	35	34.29	1,200.15
18	Agent ID Card	0		0.00

11: Remove Gridlines

Finally, if you want to present your data in a report, removing gridlines can enhance the visual clarity, making your data stand out. The report below still has the gridlines detracting from the data:
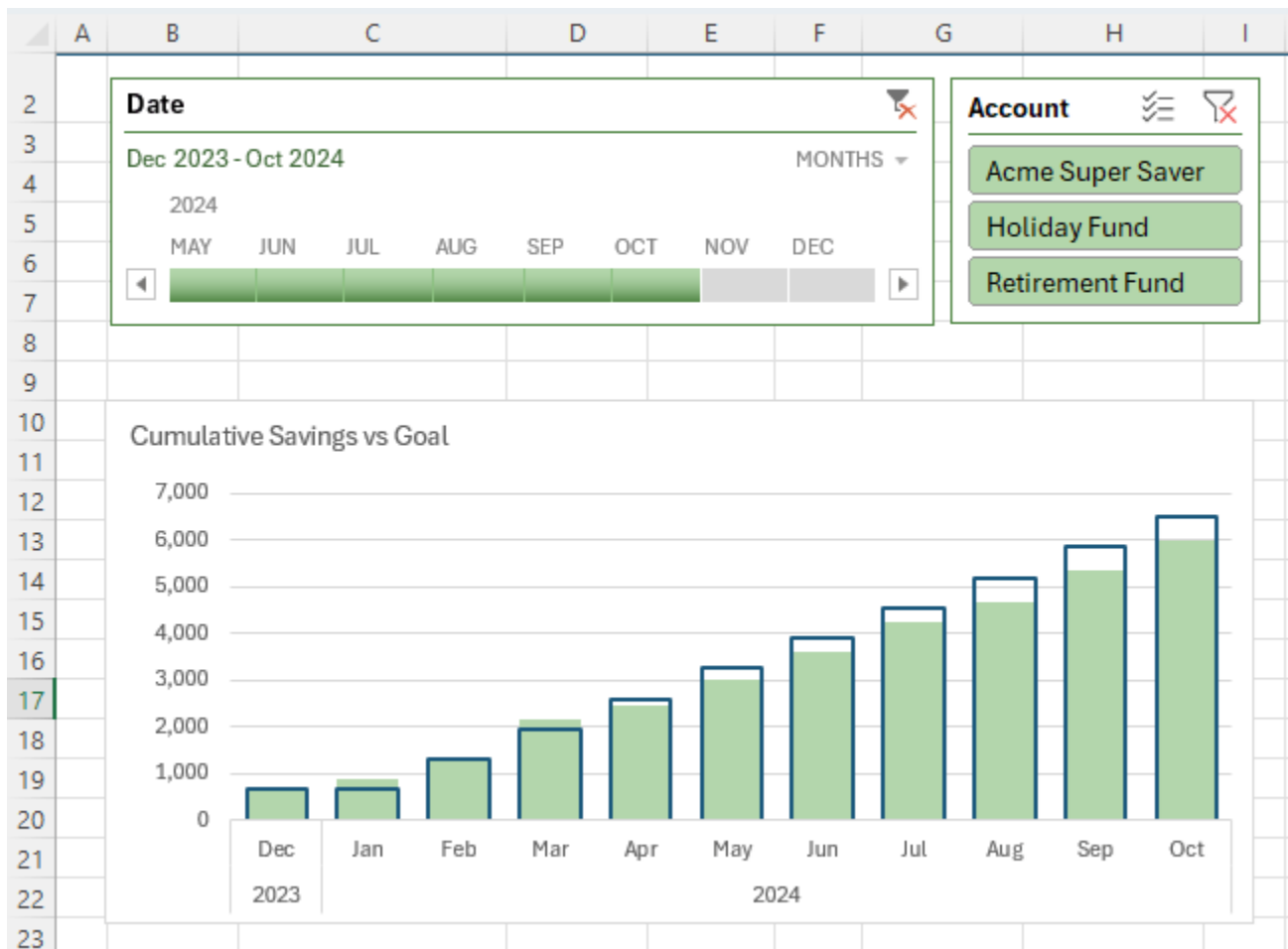
We can remove gridlines easily via the View tab > Uncheck 'Gridlines':
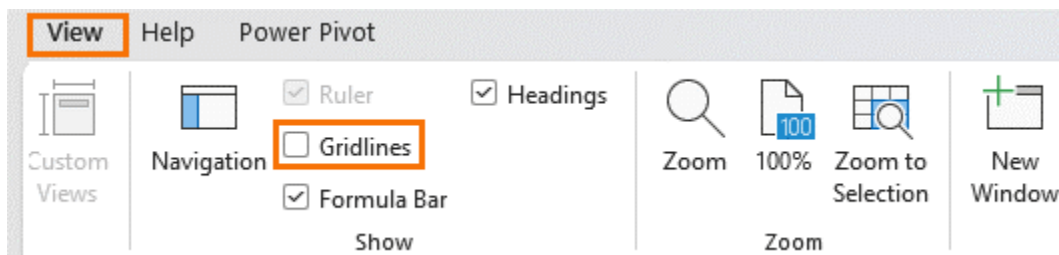
And now our report looks clean and professional:
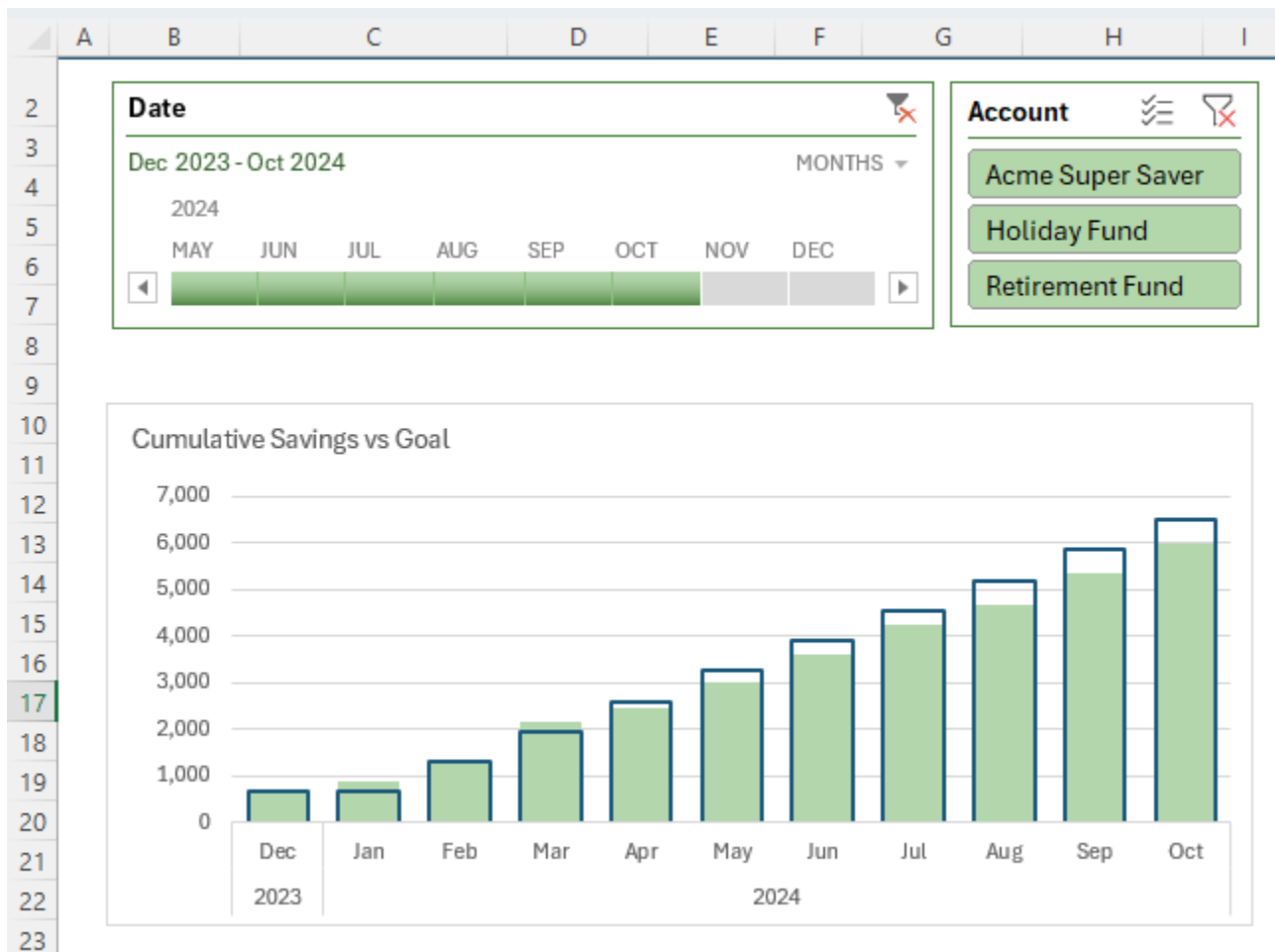
By following these steps, you can transform a dataset into a clean, efficient, and analysis-ready format. Remember, data cleaning is not just about aesthetics, it's about ensuring the accuracy and reliability of your data analysis processes.