



Canadian Council on Social Development

Community Data Program

Guide to Using Community Data 3.1

Part 2: Microsoft Excel

In this Guide, we explain some basics of Microsoft Excel. The concepts covered are relatively simple to understand, but easy to overlook as a typical user. They will allow you to use Excel more efficiently for any purpose. Feel free to read this Guide all the way through, or simply cherry pick what you need. If you need further assistance, don't hesitate to [get in touch with us](#).

Contents

1. **FAQ**
2. **Filter**
3. **Freeze panes**
4. **Text-to-Columns**
5. **CONCATENATE**
6. **TRIM**
7. **Dollar signs**
8. **Next steps**

1. FAQ

1.1 When should I use Excel?

- Modifying individual cells, rows, and columns
- Using formulas, especially when they depend on values that are expected to change
- Creating a chart (bar, pie, line, etc.)

1.2 When should I not use Microsoft Excel?

- Opening Statistics Canada data only available in the .ivt format (for this you'll need Beyond 20/20 Professional Browser)
- Storing data (well, you can, but we recommend a structured database software like MySQL or Microsoft Access)
- Statistical analysis (again, you can, but we recommend statistical software like SPSS, SAS, STATA, or R)
- Creating a map (we recommend ArcGIS or QGIS)
- Never, never, never use Excel to merge two tables using a unique key

1.3 How can I learn more about Census and Postal geographies?

This Guide will discuss Census and Postal geographies at various times. These geographies can be difficult to understand for beginners.

As the name suggests, Census geographies are used by Statistics Canada for Census data. They're also used for many other data products. Postal geographies are used by Canada Post for its own administrative purposes. Taxfiler data is available at many Postal geographies.

Our [Guide to Using Community Data 2.3 \(Part 1\)](#) (formerly called Product Profile) explains each of the typical Census geographies. You can also view the Statistics Canada page on [Census geographies](#), and use [GeoSearch](#) to find a specific geographic area.

Our [metadata spreadsheets for Taxfiler data](#) explain Postal geographies. For additional information, don't hesitate to [ask us](#).

1.4 What data is available to CDP members?

There's plenty of data available to CDP members in the .xlsx format. All [Taxfiler data](#) and [TransUnion debt data](#) are in .xlsx format. Several other tables are available only in .xlsx format, like the [Census of Agriculture](#), [Building Permits](#), [Insolvency Data](#), [Permanent Residents Rounded Data Cube](#), [Postal Code Conversion File](#), [Seniors Housing Report](#), [Shelter Capacity Report](#), and [Welfare Incomes](#).

More generally, [Schedule B](#) lists the data that we plan to acquire during any given program year (April 1 to March 31).

Finally, you're always welcome to [email us](#) if you need assistance finding data. We respond to all emails within one business day.

1.5 Where can I find other resources on Excel?

Any internet search will provide a list of tutorials, YouTube videos, and official Microsoft how-to's. Unlike Beyond 20/20, Microsoft Excel is one of the most widely-used applications in the world.

[Back to top](#)

2. Filter

Filtering lets you quickly show or hide a subset of rows within a table. It's similar to an SQL query on a database, but much simpler.

For this exercise, let's use [2014 Q1 TransUnion Credit Report Characteristics](#) data. [This spreadsheet](#) explains its column names. Remember that TransUnion data is available at the six-digit Postal Code level for all of Canada.¹

¹ We (the CDP Team) also merge TransUnion data with the same-year Postal Code Conversion File to provide a latitude-longitude coordinate for each six-digit Postal Code. In addition to a coordinate for each Postal Code, the merge provides the community name and province/territory associated with each postal code, according to both the PCCF (commname_pccf, prov_pccf) and the TU (commname_tu, prov_tu). In general, we recommend using the PCCF community name and province/territory.

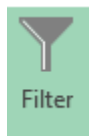
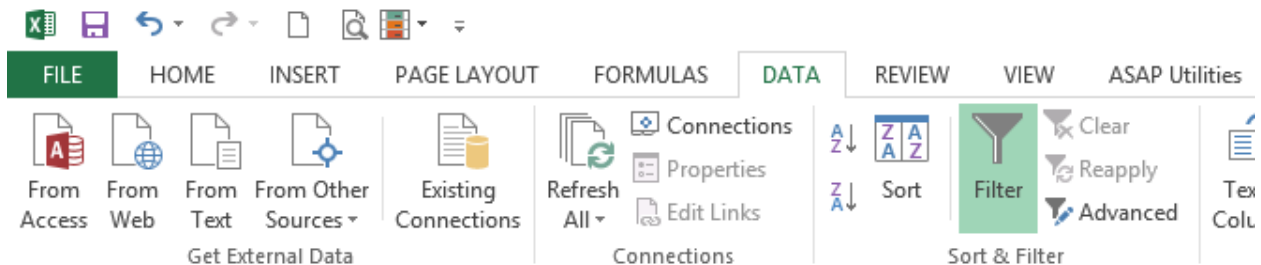
Here's what TransUnion data should look like, more or less.

	A	B	C	D	E	F	G	H	I	J	K
	fsaldu	comnname_tu	prov_tu	comnname_pccf	prov_pccf	lat_pccf	lon_pccf	count	nmcd_count	nmcd_sum	nmcd_min
1	fsaldu	comnname_tu	prov_tu	comnname_pccf	prov_pccf	lat_pccf	lon_pccf	count	nmcd_count	nmcd_sum	nmcd_min
2	rtauld	nomdecmm_tu	prov_tu	nomdecmm_fccp	prov_fccp	lat_fccp	lon_fccp	compte	ecnh_compl	ecnh_sommi	ecnh_min
3	A0A1A0	AQUAFORTE	NL	AQUAFORTE	NL	47.0078556	-52.9550234	121	92	4604392	8
4	A0A1B0	AVONDALE	NL	AVONDALE	NL	47.4200549	-53.1927532	619	488	13009761	2
5	A0A1C0	OCEAN POND	NL	BAY BULLS	NL	47.3147876	-52.8173414	1229	1050	33598498	1
6	A0A1E0	BAY DE VERDE	NL	BAY DE VERDE	NL	48.0977797	-52.898312	376	306	7731991	18
7	A0A1G0	BAY ROBERTS	NL	BAY ROBERTS	NL	47.584115	-53.2815241	3112	2546	82651752	2
8	A0A1H0	ST JOHNS	NL	BELL ISLAND FRONT	NL	47.63082	-52.935893	174	133	2910385	5
9	A0A1J0	ST. JOHNS	NL	SHEA HEIGHTS	NL	47.5421336	-52.7132509	1326	1051	22372231	1
10	A0A1K0	BRIGUS	NL	BRIGUS	NL	47.534746	-53.217387	690	566	16312712	3
11	A0A1L0	CARP	NU	BROAD COVE	NL	47.838027	-53.092157	304	245	5651751	2
12	A0A1M0	BURNT POINT B.D.	NL	BURNT POINT	NL	47.9580202	-53.0394918	303	251	6863961	3
13	A0A1N0	CALVERT	NL	CALVERT	NL	47.054913	-52.9117806	215	177	4367671	2
14	A0A1P0	CAPE BROYLE	NL	CAPE BROYLE	NL	47.095503	-52.946778	615	503	14656711	1
15	A0A1R0	CAPLIN COVE BDV	NL	CAPLIN COVE	NL	48.039485	-52.98598	79	67	1540350	9

Let's say that you want to only see data from Saskatchewan. The easiest way to do so would be to filter. Fortunately, filtering is already activated for TransUnion data. You can tell that filtering has been activated if there are boxes with downward-facing triangles at the right side of each heading.

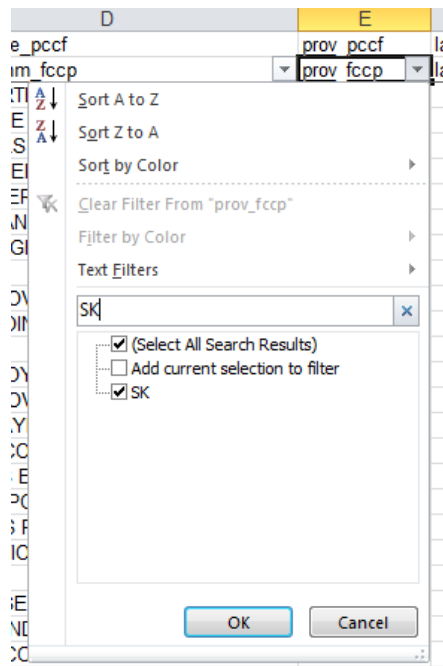


To activate or de-activate filtering, select the **Data** tab and then select **Filter**.



To show only rows in Saskatchewan, you'll need to filter column **prov_pccf** to equal **SK**. To do this, you have to select the downward-facing triangle to the right of cell **E2**. Then in the text box, write **SK**. Alternatively, you can also uncheck (**Select All**) and then check the box next to **SK**.

Then select **OK**.

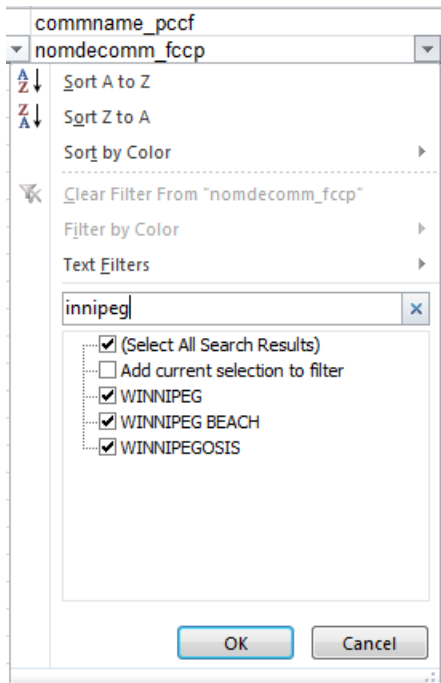


Once you select **OK**, you'll only be able to see all rows—each row a six-digit Postal Code—in Saskatchewan. This happens to be 13,846 out of 463,290 rows, which you can see at the bottom left side of the screen (*13846 of 463290 records found*).

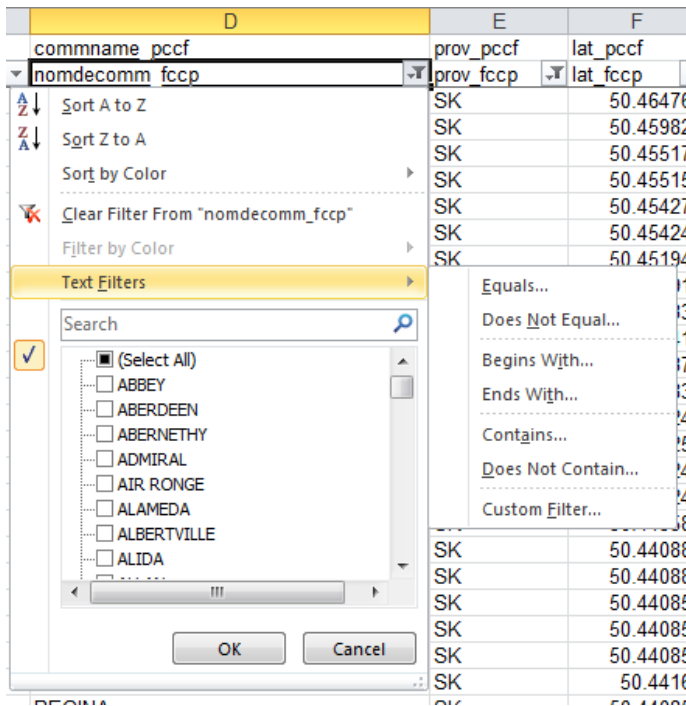
S0A1C0	GOODEVE	SK	GOODEVE	SK	51.060115
S0A1E0	GRAYSON	SK	GRAYSON	SK	50.720856
S0A1G0	HAZEL DELL	SK	HAZEL DELL	SK	51.977916
S0A1J0	HUBBARD	SK	HUBBARD	SK	51.127809
S0A1K0	HYAS	SK	HYAS	SK	51.895121
S0A1L0	INSINGER	SK	INSINGER	SK	51.5226313

13846 of 463290 records found

Note that filtering uses the *contains* operator by default. This means that when you type "innipeg" into the filter text box, you'll see "Winnipeg" in the list of results, because "Winnipeg" contains the characters "innipeg".



If you'd like, you can use other operators, e.g. *does not contain*, *equals*, *starts with*, etc. You can also join multiple filters using Boolean statements (*AND*, *OR*). To view advanced filtering functionality, select *Text Filters*.

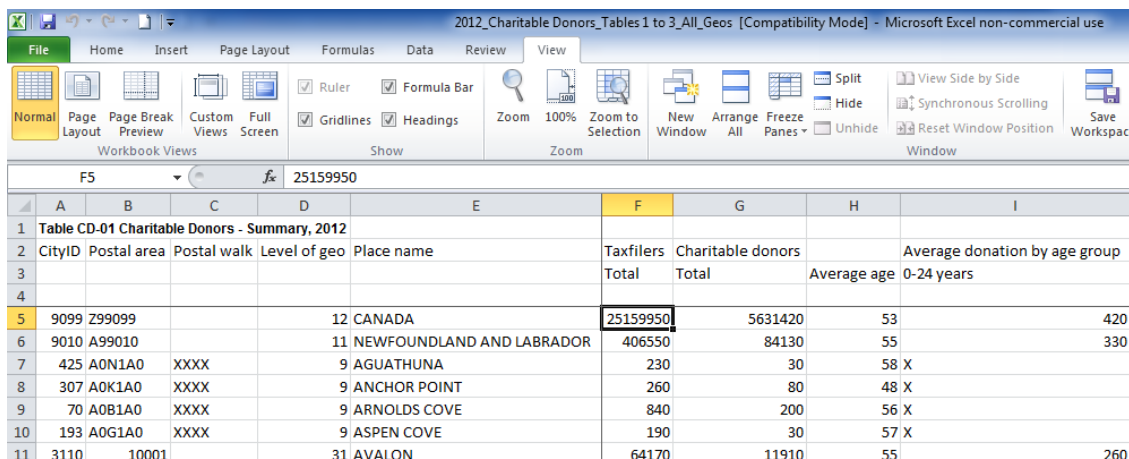
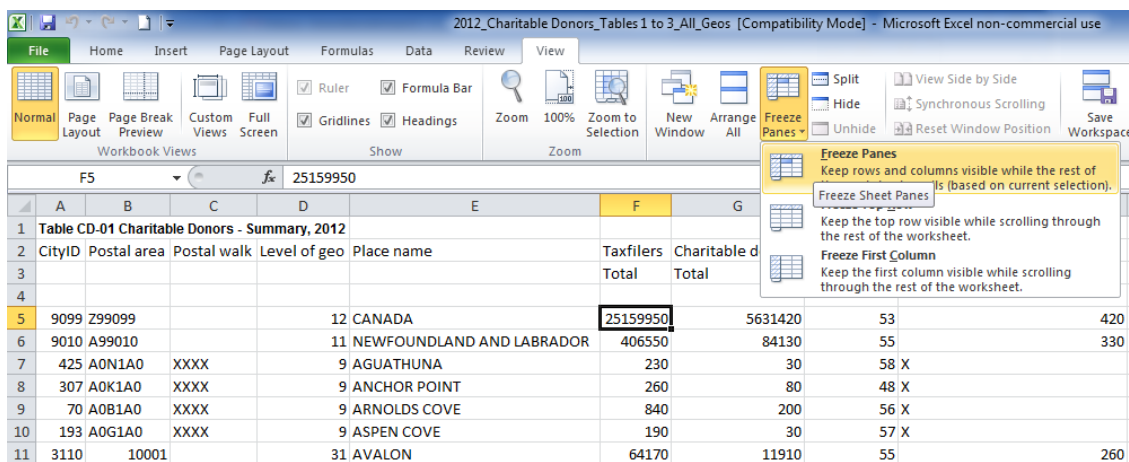


We won't go into those in detail, but hopefully you get the idea with the basic text box and check boxes.

3. Freeze panes

For this example, we use **Taxfiler Charitable Donations tables from 2012** to illustrate Freeze Panes, but you can try it on any Excel table you wish. Even a blank table will work.

Sometimes you will want to scroll down or across a table while still being able to see the first few rows or columns. To do so, place your cursor such that the row above it and the column to its left are those that you want to freeze in place. Then select the **View** tab, then **Freeze Panes > Freeze Panes**.



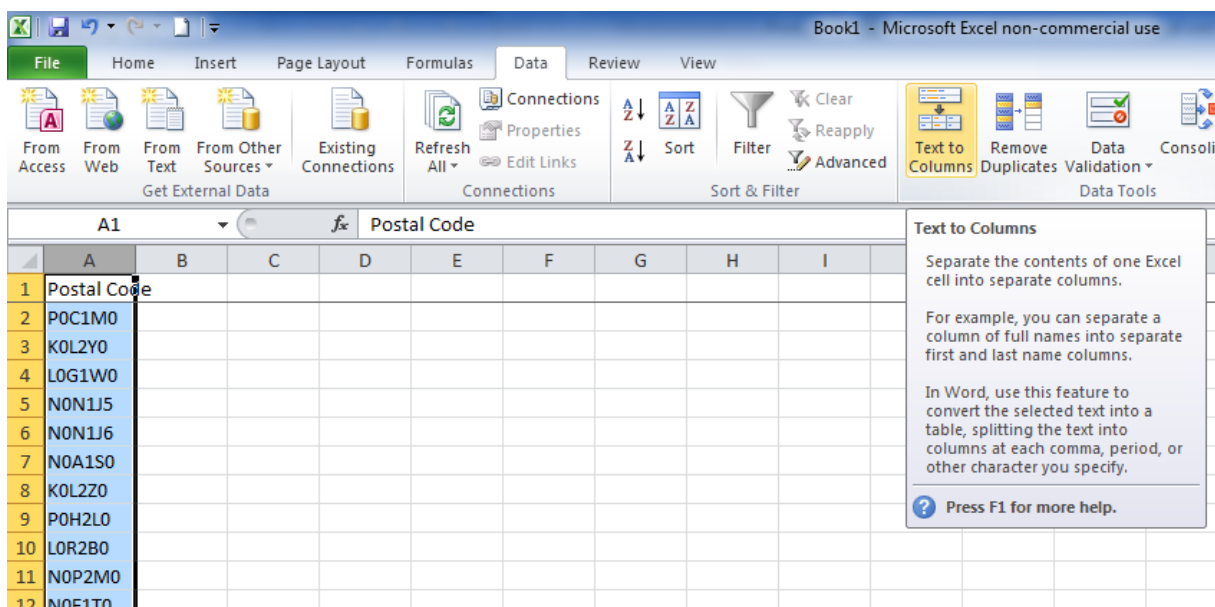
After freezing panes, you can scroll left and right, as well as up and down, without moving the frozen rows and columns. In the screenshot above, we've frozen the first four rows and the first five columns.

4. Text-to-Columns

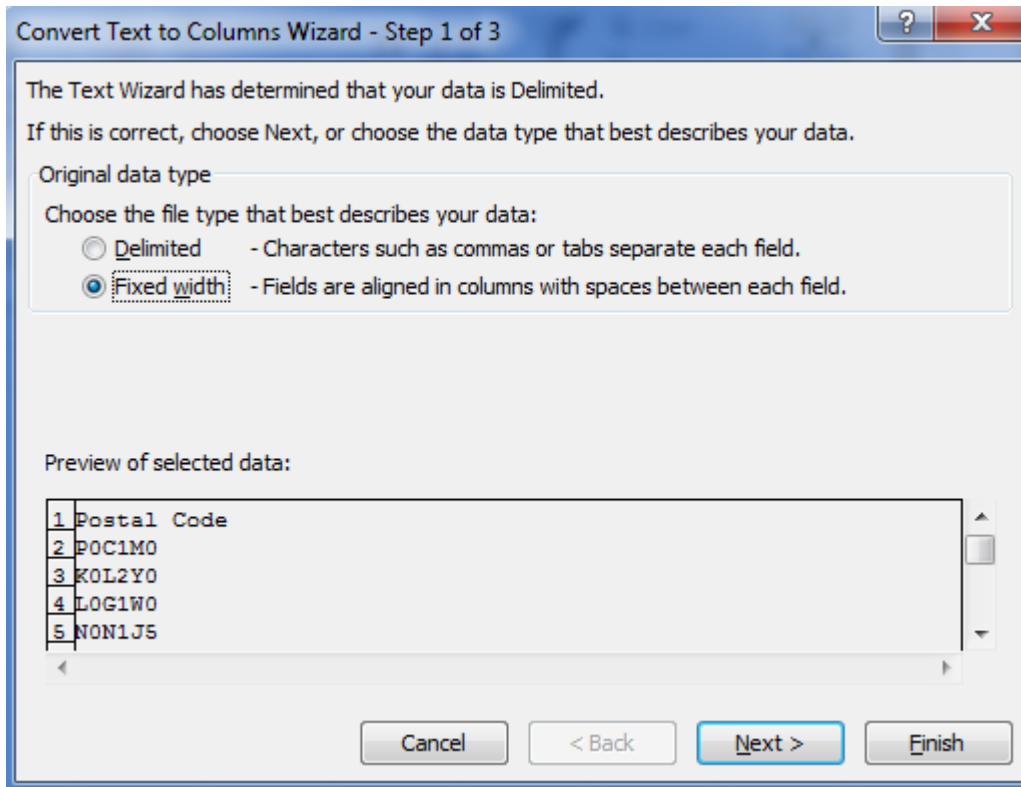
If you have a single column that you want to split into multiple columns, you can do so by using *Text-to-Columns*. There just has to be either (1) a common *break length* (e.g. 3 characters for an area code), or (2) a common symbol—or *delimiter*—separating what you want to show up in multiple columns. A *delimiter* can be any character, as long as it is consistent.

4.1 Text to columns with a break length

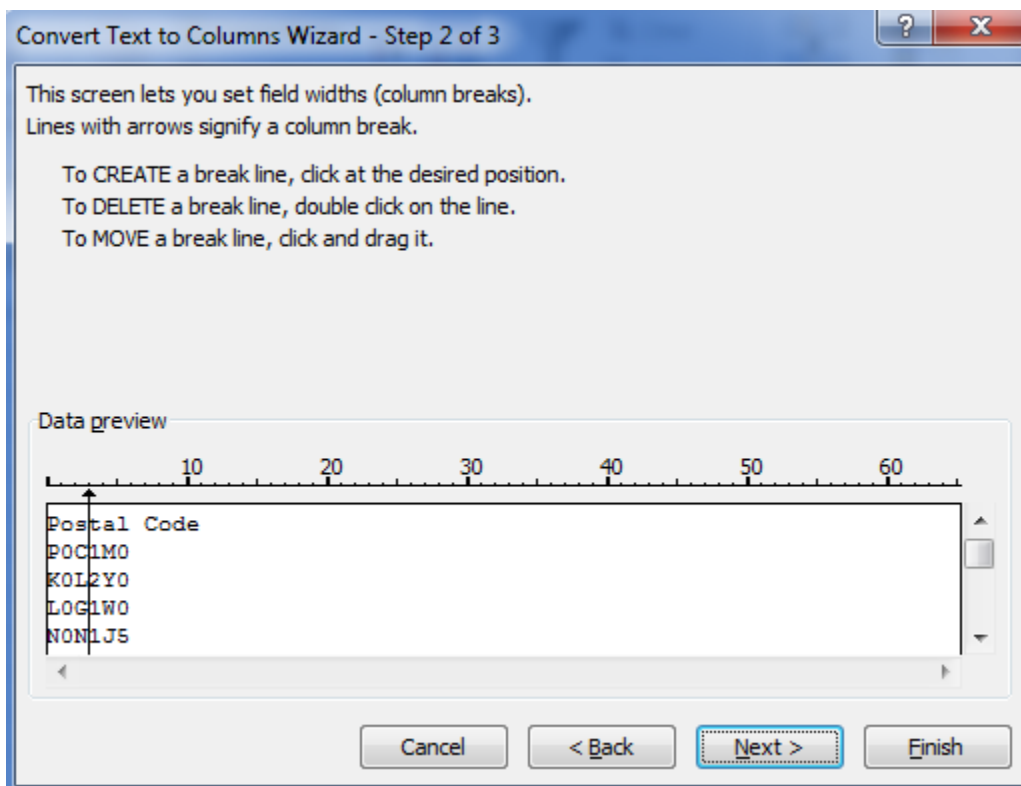
Let's say you have a list of Postal Codes and you want to create column for Forward Sortation Areas (the first-three digits of a Postal Code). There are a few ways you could do this, but we'll use *Text-to-Columns*. To do so, highlight the column you want to split, then select the *Data* tab, then *Text-to-Columns*.

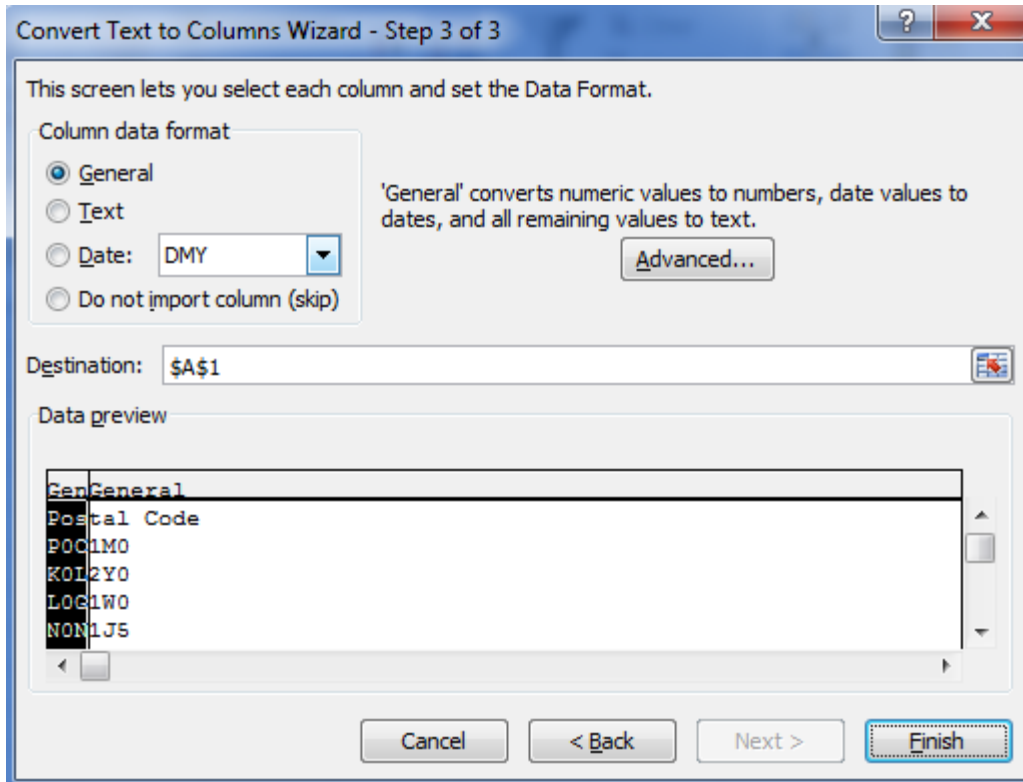


You'll see a prompt. Select *Fixed width* and then *Next*.



If you see a vertical line, reposition it to the three-character location. If you don't see a vertical line, you can add one at the three-character location. Then select *Finish*.





There you have it! An FSA column built from a Postal Code column:

	A	B
1	Pos	tal Code
2	POC	1M0
3	KOL	2Y0
4	LOG	1W0
5	NON	1J5
6	NON	1J6
7	NOA	1S0
8	KOL	2Z0
9	POH	2L0
10	LOR	2B0
11	NOP	2M0
12	NOE	1T0
13	KOK	3J0
14	LOC	1L0
15	NOL	2L0

At this stage, if you want, you can rename the column headers so that they make more sense:

	A	B
1	FSA	LDU
2	POC	1M0
3	KOL	2Y0
4	LOG	1W0
5	NON	1J5
6	NON	1J6
7	NOA	1S0
8	KOL	2Z0
9	POH	2L0
10	LOR	2B0
11	NOP	2M0
12	NOE	1T0
13	KOK	3J0
14	LOC	1L0
15	NOL	2L0

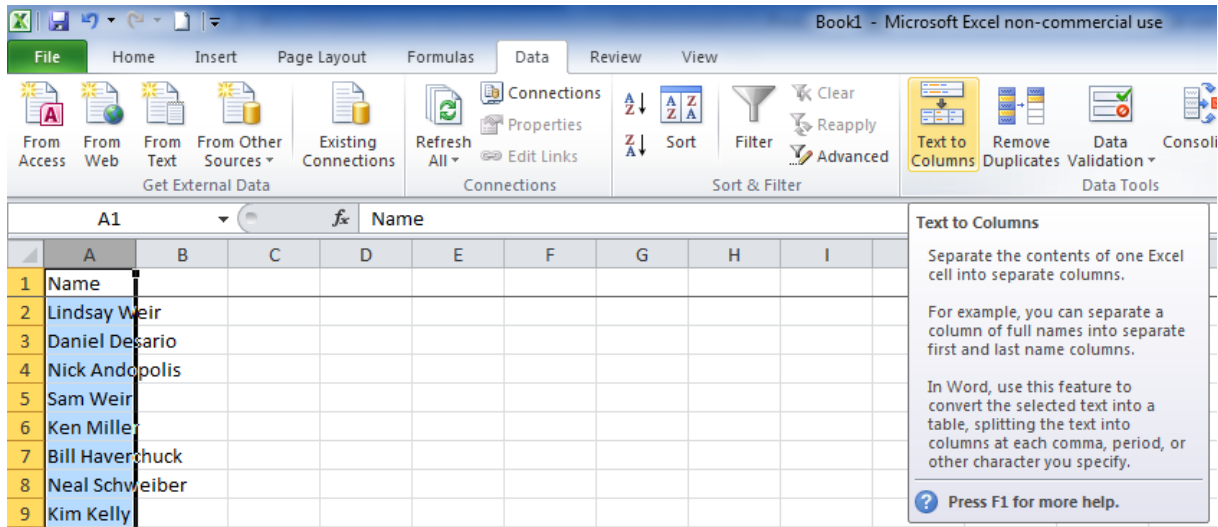
4.2 Text-to-Columns with a delimiter

The *fixed width* option is useful for data that will have a uniform width, but what about words with varying length, e.g. names? This is where *delimiters* are useful.

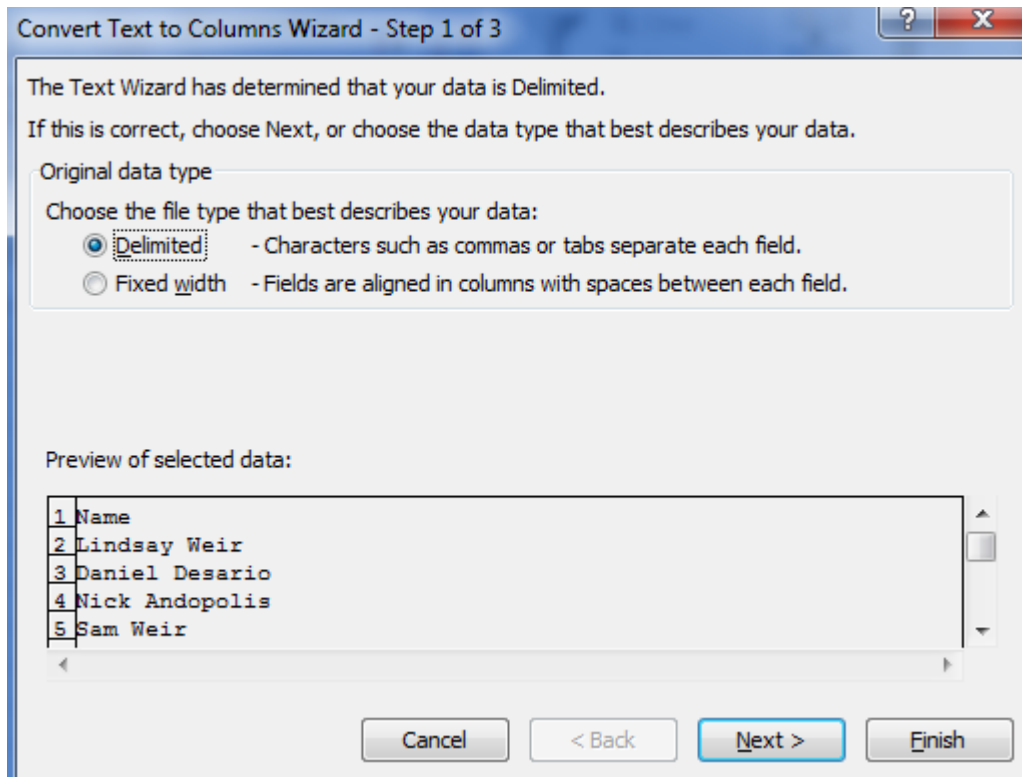
Say for example you had a column *Name* showing full (first and last) names.

	A	B
1	Name	
2	Lindsay Weir	
3	Daniel Desario	
4	Nick Andopolis	
5	Sam Weir	
6	Ken Miller	
7	Bill Haverchuck	
8	Neal Schweiber	
9	Kim Kelly	
10	Jeff Rosso	
11	Millie Kentner	

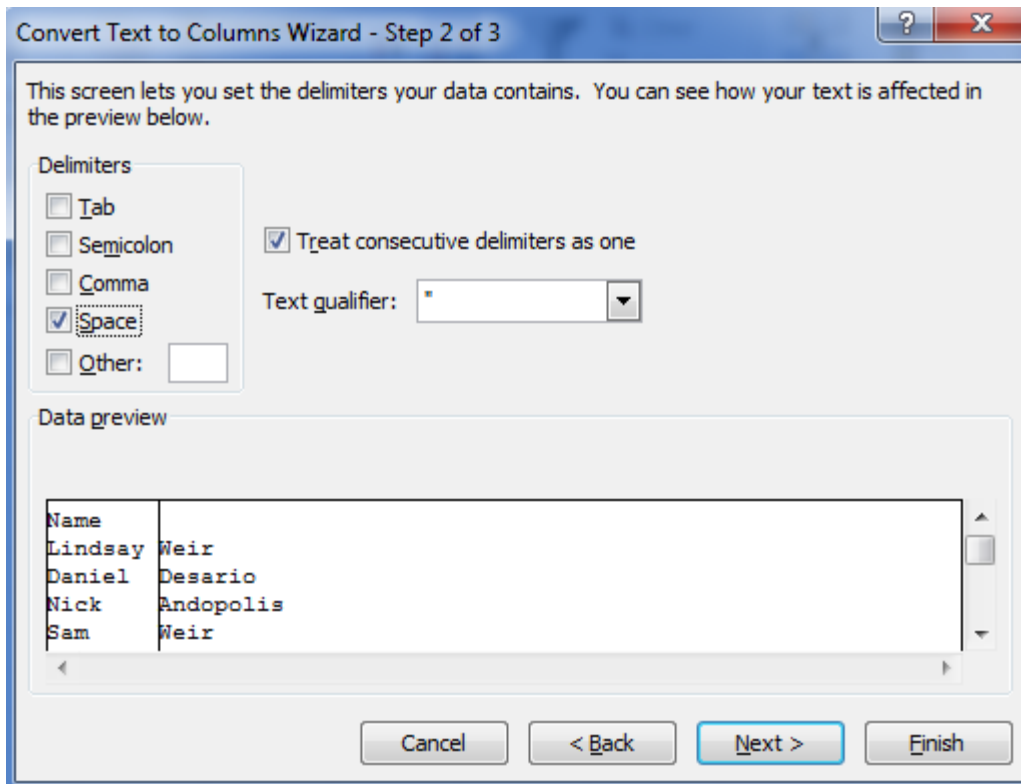
Now say that you want to separate this column into two columns: *First Name* and *Last Name*. You could use *Text-to-Columns* with a *space* as the *delimiter*. To do so, highlight the column in question, then select the *Data* tab, and then *Text-to-Columns*.



You'll see this prompt.



Select *Delimited* and then *Next*.



Uncheck *Tab* and check *Space*. Then select *Finish*.

Notice that your table now has two columns where it initially had one. As before, you can rename the column names at this stage, as necessary.

	A	B
1	Name	
2	Lindsay	Weir
3	Daniel	Desario
4	Nick	Andopolis
5	Sam	Weir
6	Ken	Miller
7	Bill	Haverchuck
8	Neal	Schweiber
9	Kim	Kelly
10	Jeff	Rosso
11	Millie	Kentner

Now let's change gears and focus on the opposite of *Text-to-Columns*: the *CONCATENATE* function.

[Back to top](#)

5. CONCATENATE

CONCATENATE allows you to append the contents of two or more cells. For example, if you have separate **FSA** and **LDU** columns and you want to create one new column that is called Postal Code, you could use the **CONCATENATE** function. To do so, enter into the new column **=CONCATENATE()**. In the brackets, include the cells you want to append, separated by a comma. In the example below, we're concatenating cells **A2** and **B2**. We've created a space between the **FSA** and **LDU** by placing a space, between quotation marks, between commas (, " ").

	A	B	C	D	E	F
1	FSA	LDU				
2	POC	1M0	=CONCATENATE(A2," ",B2)			
3	KOL	2Y0				
4	LOG	1W0				
5	N0N	1J5				
6	N0N	1J6				
7	N0A	1S0				
8	KOL	2Z0				
9	POH	2L0				
10	LOR	2B0				

Here's the result:

	A	B	C
1	FSA	LDU	
2	POC	1M0	POC 1M0
3	KOL	2Y0	
4	LOG	1W0	
5	N0N	1J5	
6	N0N	1J6	
7	N0A	1S0	
8	KOL	2Z0	
9	POH	2L0	
10	LOR	2B0	
11	N0P	2M0	
12	N0E	1T0	
13	K0K	3J0	
14	L0C	1L0	
15	N0L	2L0	

You can do the same for each row by copy-pasting cell **C2** to the cells below, or by double-clicking on the tiny black square at the bottom right corner of cell **C2**. In this example, we also rename cell **C1** to “Postal Code” Here’s the result:

	A	B	C
1	FSA	LDU	Postal Code
2	POC	1M0	POC 1M0
3	KOL	2Y0	KOL 2Y0
4	LOG	1W0	LOG 1W0
5	NON	1J5	NON 1J5
6	NON	1J6	NON 1J6
7	NOA	1S0	NOA 1S0
8	KOL	2Z0	KOL 2Z0
9	POH	2L0	POH 2L0
10	LOR	2B0	LOR 2B0
11	NOP	2M0	NOP 2M0
12	NOE	1T0	NOE 1T0
13	KOK	3J0	KOK 3J0
14	LOC	1L0	LOC 1L0
15	NOL	2L0	NOL 2L0

[Back to top](#)

6. TRIM

If you’re dealing with a spreadsheet that has several trailing spaces that you want to clean up, then you can use the *TRIM* function. No need to go through each cell and delete excess spaces, or use *Find-and-Replace*. *TRIM* removes all leading and trailing spaces, and converts multiple spaces into single spaces.

To *TRIM* a cell, select an adjacent cell, and enter `=TRIM()` with the target cell in the brackets.

Here’s an example where we are trimming the contents of cell **A1** from “ Community Data Program ” to “Community Data Program”.

	A	B	C	D	E
1	Community Data Program				
2	=TRIM(A1)				
3	TRIM(text)				

	A	B	C
1	Community Data	Program	
2	Community Data	Program	
3			

This function is especially useful for cleaning up messy data, e.g. survey responses.

[Back to top](#)

7. Dollar signs

Dollar signs allow you to specify a cell in an equation, and not have that cell change when you use the equation on other cells. For example, if you have a price column, and you want to multiply the cells in that column by a given tax rate that will not change depending on the item, then you can enter the tax rate in a different cell and use dollar signs to multiply each cell in the price column by the tax rate cell.

SUM		=B2*\$F\$1				
	A	B	C	D	E	F
1	Item	Price	Price incl. tax		Tax rate:	1.15
2	Desk	100	=B2*\$F\$1			
3	Chair	50				
4	Bed	200				
5	Mirror	80				
6	Table	150				
7	Shelf	30				
8	Dresser	125				

Notice the *dollar signs* in the screenshot above. They ensure that the reference to cell *F1* stays unchanged when copy-pasting that cell into the cells below.

C2		=B2*\$F\$1				
	A	B	C	D	E	F
1	Item	Price	Price incl. tax		Tax rate:	1.15
2	Desk	100	115			
3	Chair	50				
4	Bed	200				
5	Mirror	80				
6	Table	150				
7	Shelf	30				
8	Dresser	125				

	A	B	C	D	E	F
1	Item	Price	Price incl. tax		Tax rate:	1.15
2	Desk	100	115			
3	Chair	50	57.5			
4	Bed	200	230			
5	Mirror	80	92			
6	Table	150	172.5			
7	Shelf	30	34.5			
8	Dresser	125	143.75			

[Back to top](#)

8. Next steps

This Guide is just the tip of the iceberg. There are plenty of other interesting things Excel can do (*IF*, *PivotTables*, *VLOOKUP*, *ASAP Utilities*, etc.). If there's something not covered in this Guide that you want to know more about, don't hesitate to [get in touch with us](#). We're always happy to help out.