Introduction to QGIS (version 3.16) and Geoprocessing using COVID-19 data

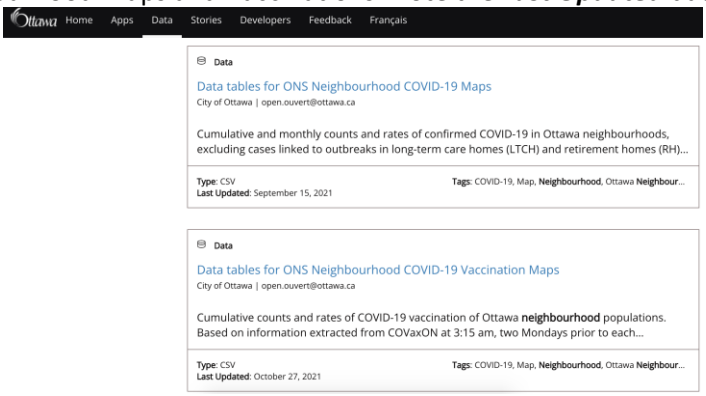
Rebecca Bartlett, GIS Librarian, November 2021

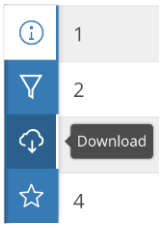
Over the next two weeks, we'll examine possible connections between demographic information (income, walkability, proximity to resources, etc.) and COVID-19 cases and vaccinations in Ottawa neighbourhoods.

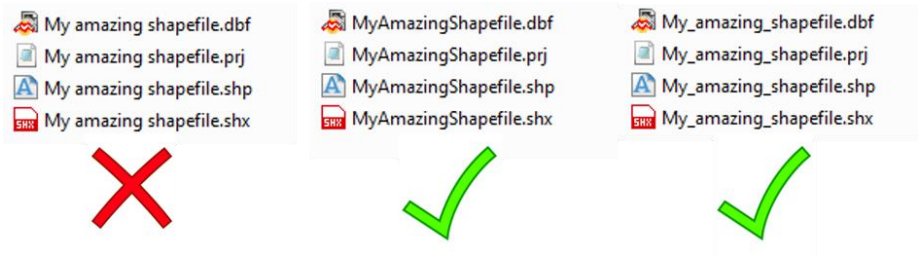
Table of Contents

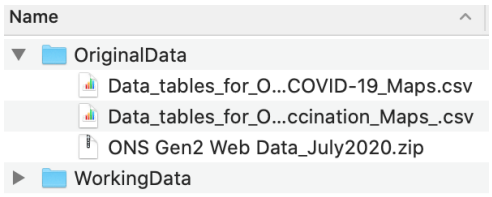
1. Download data for Ottawa neighbourhoods and COVID-19

1) Download data from http://www.davidmckie.com/ONS Gen2 Web Data_July2020.zip.
    a. This includes 2 files that we'll need for Weeks 1 & 2: ONS Gen2 Web Data_July2020.xlsx and ONS_Boundaries_Gen2.zip (it contains the neighbourhood boundary shapefile).

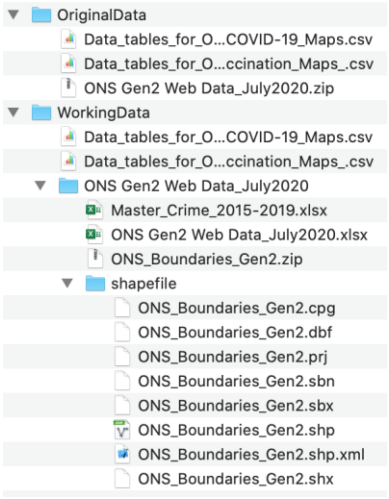
2) Go to https://open.ottawa.ca/
3) Search for 'neighbourhood' and then enter to go to the results page.
4) On the results page for the 'neighbourhood' search, look for Data Tables for COVID19 by Ottawa Neighbourhood Maps and Vaccinations. Note the ***Last Updated*** dates.
    a. 

    b. Click on each dataset and download the CSVs using the little Download button next to the table
    c. 

5) **Create a project folder for this assignment and name it something helpful, like JournalismMap**. Also add two sub-folders in your main project folder: OriginalData and WorkingData.
    a. Avoid having spaces in your folder names and file names. GIS software can be finicky and give you errors that make it seem like your data is corrupt…but your data just has a space in the file name. Use CamelCase or underscores for folder and file names instead:
    b. 

6) Copy the zip file into your new OriginalData folder. Your folder should look something like this:
    a. 

    b. We strongly suggest a hierarchy like this because it will really help keep you organized over the next few weeks:
        i. It may seem like overkill right now but trust us on this! You also ***ALWAYS*** want to keep an untouched, unedited copy of your original datasets as back-up.
7) **Copy the COVID CSV files and the ONS Gen 2 ZIP file into your WorkingData folder**
8) Unzip/extract the ZIP file
9) You'll also need to unzip ONS_Boundaries_Gen2.zip
10) Once everything is unzipped, your map project folder should look something like this:



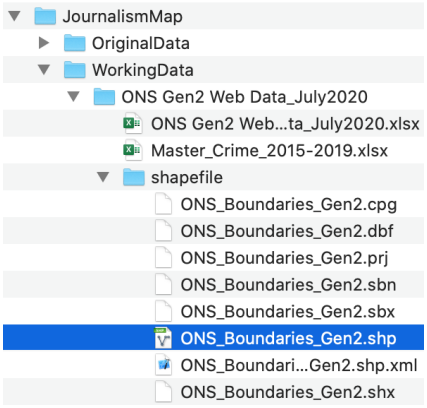
    a. This is as good a time as any to note that shapefiles look like multiple different files when you view them in Finder or Windows Explorer. If you ever copy and paste them somewhere, you need to have **all the parts** (e.g.: all the files that have the same name

before the file extension – there may be up to 8 of them). For more on shape files, please consult pages 121 and 148 of The Data Journalist.

11) You're done the downloading and extracting!

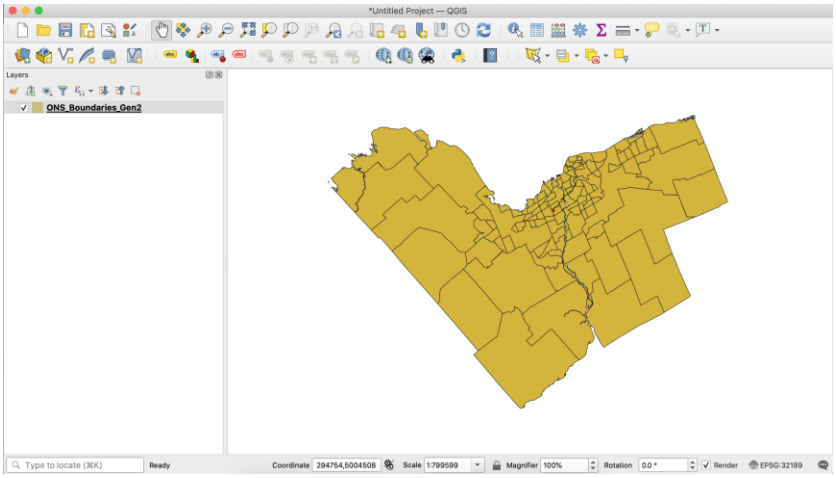
2. Adding the neighbourhood shapefile to QGIS

12) Open QGIS.
13) Select New Empty Project.
14) Add the dataset by clicking on **Layer > Add Layer > Add Vector Layer…** Browse to the project's folder and find the ONS shapefile. Click on the .shp (shown below) and click **Add** to add the shapefile to QGIS. Then close the Add Vector Layer window.



   a. You could also have the project folder open in Finder/Windows Explorer and just drag and drop the .shp file in. Either way works!

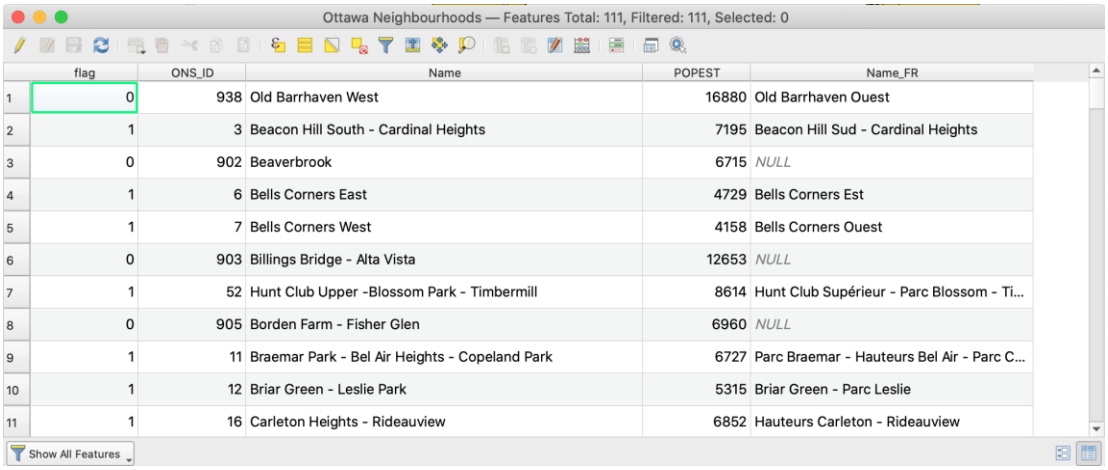
15) Your QGIS window should look something like this. Don't worry if the colour is different – QGIS randomly selects a colour for the data symbology.



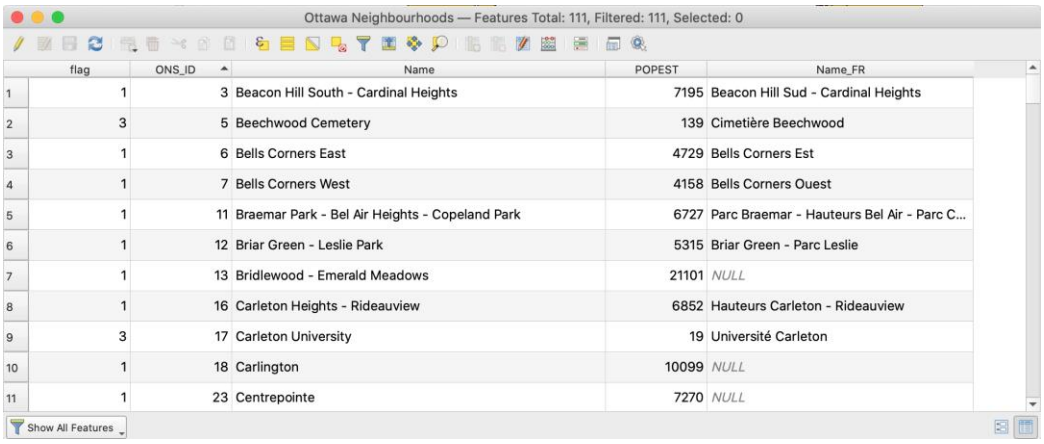
   a.
   b. If you get a "Select Transformation…" pop-up, just click OK

16) First thing to do next is to save the project. QGIS doesn't auto-save so you'll need to manually save frequently.
    a. Save the file in your main project folder
17) We're going to do a little bit of housekeeping since the name of the neighbourhoods shapefile could be more helpful.

a. Right-click on the shapefile layer (ONS_Boundaries_Gen2)
b. Click **Rename Layer** in the menu
c. Give it a helpful name, like Ottawa Neighbourhoods
    i. It is worth noting that renaming the layer only changes the layer name in QGIS. *It does not change the filename of the shapefile*, so if you look in your WorkingData folder the filename is still ONS_Boundaries_Gen2. Only exporting the shapefile with a new name will change the shapefile filename.

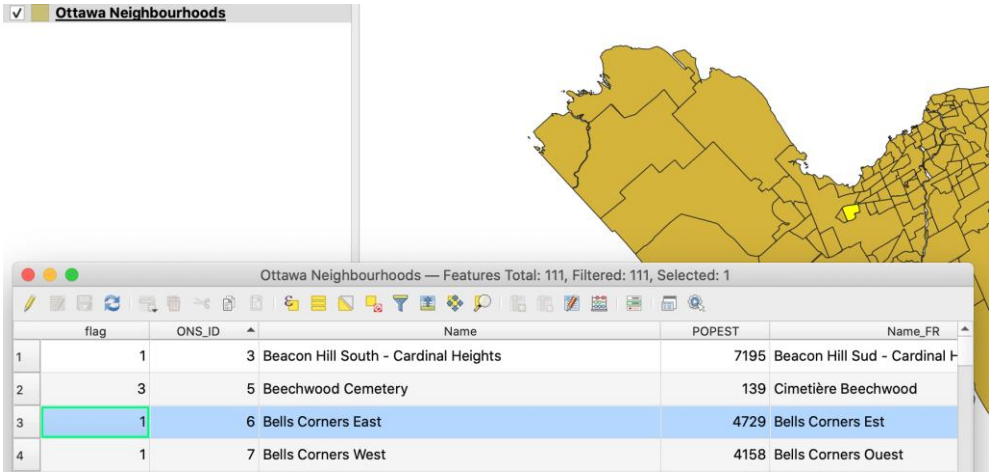
18) GIS files are always built on data, so let's take a look at it.
    a. Right-click on the Ottawa Neighbourhoods layer
    b. **Open Attribute Table**
    c. 

	flag	ONS_ID	Name	POPEST	Name_FR
1	0	938	Old Barrhaven West	16880	Old Barrhaven Ouest
2	1	3	Beacon Hill South - Cardinal Heights	7195	Beacon Hill Sud - Cardinal Heights
3	0	902	Beaverbrook	6715	NULL
4	1	6	Bells Corners East	4729	Bells Corners Est
5	1	7	Bells Corners West	4158	Bells Corners Ouest
6	0	903	Billings Bridge - Alta Vista	12653	NULL
7	1	52	Hunt Club Upper -Blossom Park - Timbermill	8614	Hunt Club Supérieur - Parc Blossom - Ti...
8	0	905	Borden Farm - Fisher Glen	6960	NULL
9	1	11	Braemar Park - Bel Air Heights - Copeland Park	6727	Parc Braemar - Hauteurs Bel Air - Parc C...
10	1	12	Briar Green - Leslie Park	5315	Briar Green - Parc Leslie
11	1	16	Carleton Heights - Rideauview	6852	Hauteurs Carleton - Rideauview

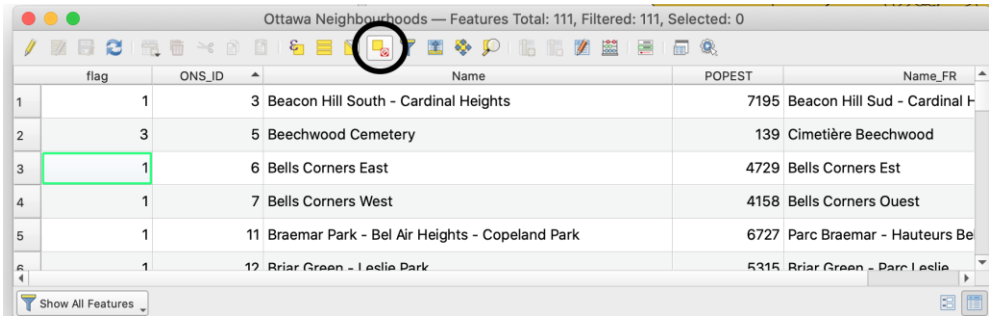
19) Note that there are 111 features. Also note the **ONS_ID** column, as this is a unique ID for each neighbourhood and will come in handy later.
20) Click on the **ONS_ID** header to sort
    a. 

	flag	ONS_ID	Name	POPEST	Name_FR
1	1	3	Beacon Hill South - Cardinal Heights	7195	Beacon Hill Sud - Cardinal Heights
2	3	5	Beechwood Cemetery	139	Cimetière Beechwood
3	1	6	Bells Corners East	4729	Bells Corners Est
4	1	7	Bells Corners West	4158	Bells Corners Ouest
5	1	11	Braemar Park - Bel Air Heights - Copeland Park	6727	Parc Braemar - Hauteurs Bel Air - Parc C...
6	1	12	Briar Green - Leslie Park	5315	Briar Green - Parc Leslie
7	1	13	Bridlewood - Emerald Meadows	21101	NULL
8	1	16	Carleton Heights - Rideauview	6852	Hauteurs Carleton - Rideauview
9	3	17	Carleton University	19	Université Carleton
10	1	18	Carlington	10099	NULL
11	1	23	Centrepointe	7270	NULL

    b. Sorting columns can be a handy way to get to know the data a bit. For example, are there any NULL values or 0 values? Are there easily identifiable duplicate values? Etc.
21) Click on one of the row numbers, whichever one you'd like. Note that the selected row corresponds to a selected neighbourhood (in yellow) in the polygon dataset.

a. 

22) To deselect, click the **Deselect All** button in the attribute table

a. 

3. Preparing the Excel spreadsheet

In order to connect non-GIS tabular (spreadsheet) data to a GIS file, we need to make sure the data is arranged correctly so we can join the two files.

23) Open **ONS Gen2 Web Data_July2020.xlsx** and have a look at the data in the different sheets.

b. The Neighbourhoods sheet is straightforward with just the neighbourhood names and unique ID numbers (the ONS_ID column)

i.

	A	B	C	D	E
1	name	ONS_ID	name_fr		
2	Old Barrhaven West	938	Old Barrhaven Ouest		
3	Beacon Hill South - Cardinal Heights	3	Beacon Hill Sud - Cardinal Heights		
4	Beaverbrook	902	Beaverbrook		

c. The Layers sheet has definitions for all the different data points for each neighbourhood. The Polygon Attribute column will be particularly helpful!

i.

T
Polygon attribute
D_dent_clinic_ave_dis3
D_dent_clinic_count
D_dent_clinic_countPB
D_dent_clinic_covPop

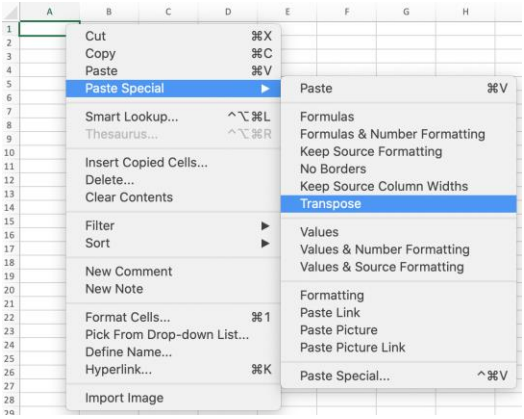
d. The Data sheet has all the data we'll want. Note that the id column matches the Polygon Attribute column in the Layers sheet.

i.

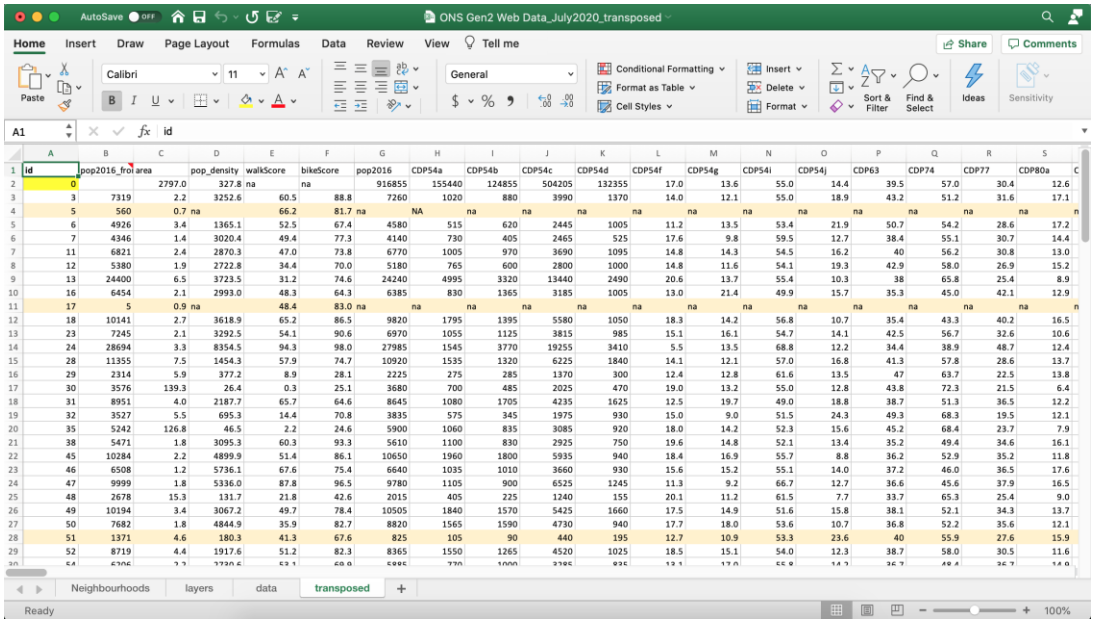
	A	B	C	D	E
1	id	0	3	5	6
2	pop2016_fromDB		7319	560	4926
3	area	2797.0	2.2	0.7	3.4
4	pop_density	327.8	3252.6	na	1365.1
5	walkScore	na	60.5	66.2	52.5
6	bikeScore	na	88.8	81.7	67.4
7	pop2016	916855	7260	na	4580
8	CDP54a	155440	1020	NA	515

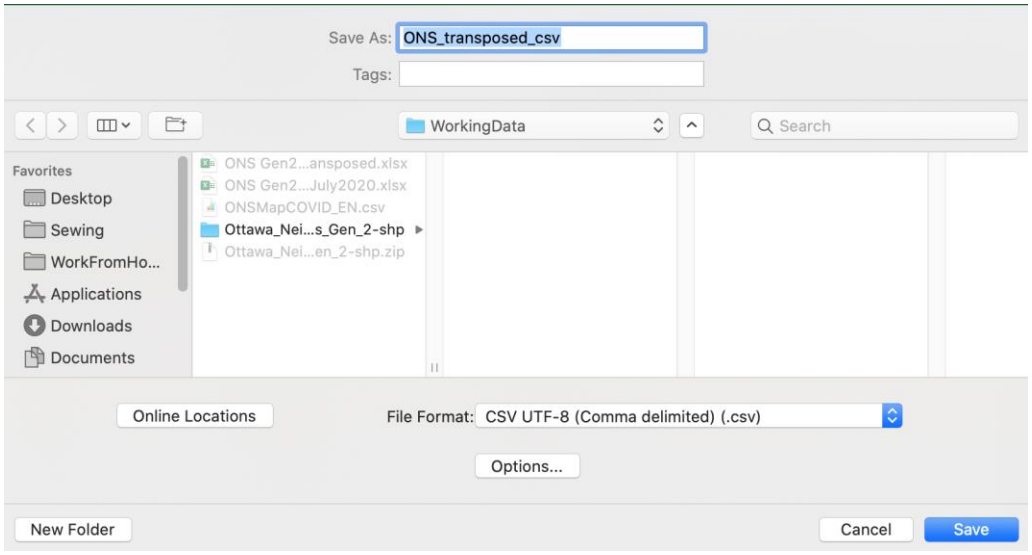
24) **<u>Take a few minutes to review the Layers sheet and familiarize yourself with what data is available. Write down 5-10 fields you're initially interested in, especially their Polygon Attribute designation (which are often codes such as CDP54b) and if the numbers in that column are integers or have decimals.</u>**
    a. You'll need to select 5-10 fields before you can map them, so might as well do it now while you have the field definitions right there.
    b. It's also going to be helpful to know if they're integer or decimal numbers later on.
25) After looking at the Data sheet, it's evident that each neighbourhood is a column and each characteristic (e.g. area, population, walkScore, etc.) is a row.
    e. This is an issue because, in the attribute table for the neighbourhoods shapefile, each neighbourhood is a row and each characteristic is a column.
26) We need to **Transpose** (turn the columns into rows and the rows into columns) for the join to the shapefile to work correctly.
    a. Create a new sheet in the Excel file and give it a helpful name, like Transposed
    b. Copy all the data in the Data sheet
        i. I had to click and drag to select all the rows and columns since using the Select All button (the square between the row numbers and the column letters in the upper left) didn't work when I went to paste into a new sheet.
    c. **Right-click on cell A1** in the empty Transposed sheet
    d. Select **Paste Special… > Transpose**

i.



27) **Save As** and give the spreadsheet a new name in the WorkingData folder, something like ONS Gen2 Web Data_July2020_transposed.xlsx

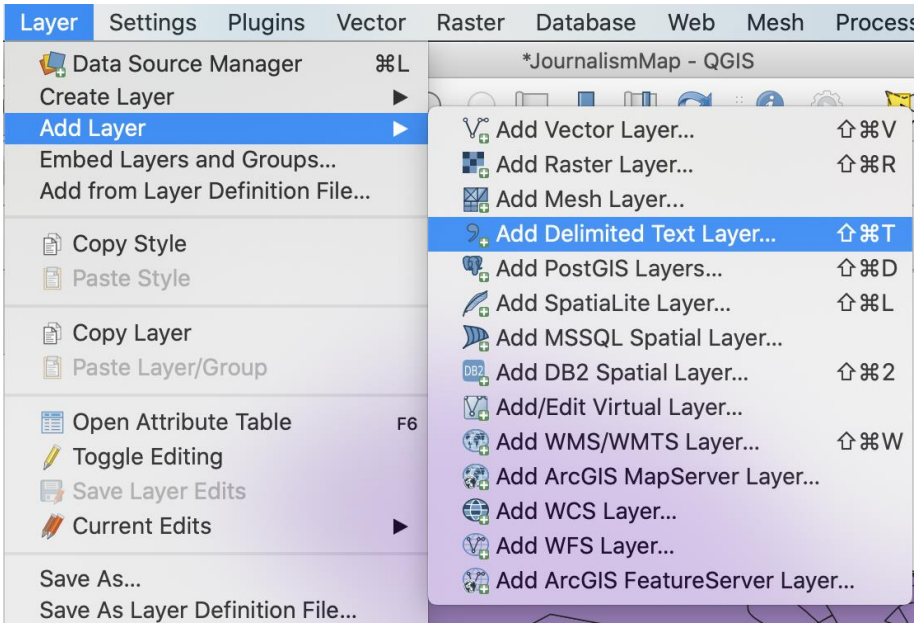
a. 

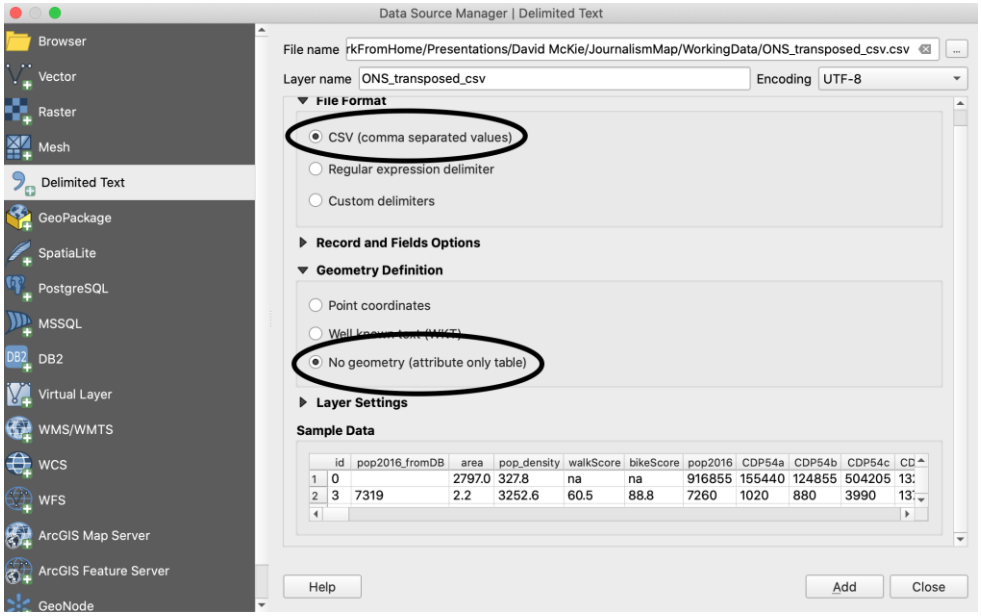
28) We're going to save the Transposed sheet as a CSV file for import into QGIS (Excel files can get a bit messy: no column names, numeric fields are text fields, etc.).
    a. **Save As a new CSV file** with a helpful name, and save in the WorkingData folder
    b. 

    c. You'll get a warning but just click OK.
29) Close Excel.

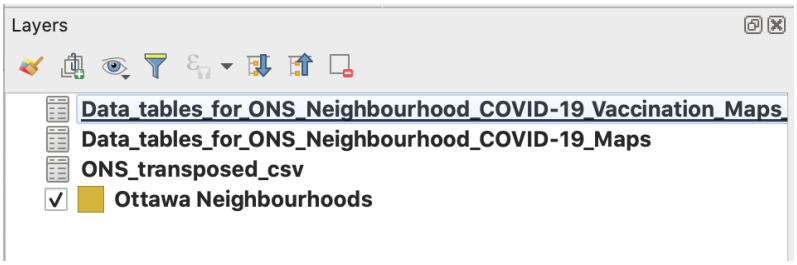
4. Bringing the CSVs into QGIS

30) In QGIS, go to Layer > Add Layer > Add Delimited Text Layer…

a. 

31) In the new window, browse to the **ONS_transposed_csv** you just created (it should be in your WorkingData folder!) and fill out the window as follows:
    a. File format: CSV
    b. **Geometry definition: No Geometry (attribute only table)**
    c. 

32) Click Add
33) Before closing the Delimited Text window, fill it in the same way but for the **Data_tables_for_ONS_Neighbourhood_COVID-19_Maps.csv** and **Data_tables_for_ONS_Neighbourhood_COVID-19_Vaccination_Maps_.csv** files.
    a. Remember, no geometry!
34) Add both, then Close.
    a. Your Layers pane should look like this:



35) **Save your QGIS project.**

36) Congratulations, you're now ready to work with all four datasets in QGIS next week!