Introduction to QGIS (version 3.10) and Geoprocessing using COVID-19 data

Rebecca Bartlett, GIS & Digital Resources Librarian, November 2020

Over the next two weeks, we're going to take a look at possible connections between demographic information (income, walkability, proximity to resources, etc.) and COVID-19 cases in Ottawa neighbourhoods.

Table of Contents

1. DOWNLOAD DATA FOR OTTAWA NEIGHBOURHOODS AND COVID-19	1
2. Adding the Neighbourhood shapefile to QGIS	2
3. PREPARING THE EXCEL SPREADSHEET	5
4. BRINGING THE CSVS INTO QGIS	7

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 'neighbourhoods' and then enter to go to the results page.
- 4) On the results page for the 'neighbourhoods' search, click the COVID-19 in Ottawa Neighbourhoods result
 - a. It will open a new tab. Scroll down to the bottom of the page, and click the link to the dataset
 - b. The data table for this map is available for download on Open Ottawa 🗷
 - c. A new tab will open again. Click the Download button to download the CSV file.
- 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:



6) Copy the zip file into your new OriginalData folder and **unzip/extract it**. Your folder should look something like this:



- 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 Excel file, the ONS Gen 2 Excel file, and ONS_Boundaries_Gen2.zip into your WorkingData folder
- 8) Unzip the ONS_Boundaries_Gen2.zip file
- 9) Once 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.
- 10) You're done the downloading and extracting!

2. Adding the neighbourhood shapefile to QGIS

- 11) Open QGIS.
- 12) 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!
- 13) 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.



- 14) 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
- 15) We're going to do a little bit of housekeeping since the name of the neighbourhoods shapefile is just a bunch of random numbers and letters not helpful at all!
 - a. Right-click on the shapefile layer (6f216d7_ etc.)
 - b. Click Rename Layer in the menu

а

- c. Give it a helpful name, like Ottawa Neighbourhoods
 - 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 a bunch of letters and numbers. Only
 exporting the shapefile with a new name will change the shapefile filename.
- 16) 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

Į			ا کا 🗧 ک	💊 🍸 🖺 🌺 .	P 🛯 🔚 📗 🗶		R.		
	FID	flag	ONS_ID	Name	POPEST	Name_FR	Shape_Are	Shape_Len	
1	110	0	935	Manor Park	7016	NULL	2506689.093	9336.845392	
2	109	0	947	Riverview	6297	NULL	3663346.257	8424.303921	
3	111	0	914	Chapman Mills	8212	NULL	7334882.691	13761.919186	
4	106	0	941	Parkwood Hill	5234	NULL	1031159.1757	9977.771934	
5	105	0	921	Findlay Creek	4419	NULL	19663736.59	22155.69501	
6	108	0	909	Cardinal Creek	10535	NULL	19374265.96	19214.76014	
7	107	0	920	Fallingbrook	19415	NULL	11601591.68	15376.743114	
8	102	1	86	Pineview	6169	NULL	3689535.734	8298.642861	
9	101	0	926	Greenbelt	2073	Ceinture de V	336185835.1	248571.9411	
10	104	0	937	Old Barrhave	22286	Old Barrhave	12445394.50	20220.34399	
11	103	0	907	Brookside - B	14940	NULL	19117325.757	20399.05450	
	00	1	25	Dunrahin	5560	A.II. II. I	0575040467	00000 55 400	-

- 17) 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.
- 18) Click on the ONS_ID header to sort

c.

a.

a.

•	•		C	ttawa Neighbour	hoods :: Feature	s Total: 111, Filte	red: 111, Selected	i: 0
/			ک 📄 🖻	🔩 🝸 🔳 🔶 !	P i 🛯 🕅 🕅	📓 I 🚍 I 📅 🍳	8.	
	FID	flag	ONS_ID	Name	POPEST	Name_FR	ShapeAre	Shape_Len
1	2	1	3	Beacon Hill S	7195	Beacon Hill S	4533906.765	10362.45230
2	75	3	5	Beechwood	139	Cimetière Be	1351777.0703	5282.712404
3	4	1	6	Bells Corners	4729	Bells Corners	6786857.875	11170.923794
4	5	1	7	Bells Corners	4158	Bells Corners	2772272.183	7467.061208
5	9	1	11	Braemar Park	6727	Parc Braemar	4777807.722	10087.59676
6	10	1	12	Briar Green	5315	Briar Green	3850340.289	11595.39619
7	63	1	13	Bridlewood	21101	NULL	13152825.113	15017.365812
8	11	1	16	Carleton Heig	6852	Hauteurs Carl	4321687.199	12893.05075
9	74	3	17	Carleton Univ	19	Université Ca	1848575.628	9414.831432
10	12	1	18	Carlington	10099	NULL	5499930.957	10693.47955
11	16	1	23	Centrepointe	7270	NULL	4285224.089	8704.279264
	47	4	24	Contratours	24004	Contro Millo	6700261 05 4	10410 40000

- 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.
- 19) 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.

•	• •	Ottawa Ne	ighbourhoods ::	Features Total: 1	11, Filtered: 111,	Selected: 1			
1		1 1 ≤ 1 1 1	ا کا 🖹 🗳	🖥 🝸 🗷 🗞 🖇		📓 🗃 👩 (8		
	FID	flag	ONS_ID +	Name	POPEST	Name_FR	Shape_Are	-	
1	2	1	3	Beacon Hill S	7195	Beacon Hill S	4533906.765	10	
2	75	3	5	Beechwood	139	Cimetière Be	1351777.0703	52	Trans
3	4	1	6	Bells Corners	4729	Bells Corners	6786857.875	111	J.
4	5	1	7	Bells Corners	4158	Bells Corners	2772272.183	74	J
5	9	1	11	Braemar Park	6727	Parc Braemar	4777807.722	10	
6	10	1	12	Briar Green	5315	Briar Green	3850340.289	11	
7	63	1	13	Bridlewood	21101	NULL	13152825.113	15(~~
в	11	1	16	Carleton Heig	6852	Hauteurs Carl	4321687.199	12	(
9	74	3	17	Carleton Univ	19	Université Ca	1848575.628	94	

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

	•	Ottawa N	eighbourhood	: Fectures Total:	111, Filtered: 111,	Selected: 1		
/			1 i 🗧 📕 🛛	😼 🕈 🖭 💠 .	P 📙 📙 💋		R.	
	FID	flag	ONS_ID	Name	POPEST	Name_FR	ShapeAre	
1	2	1	3	Beacon Hill S	7195	Beacon Hill S	4533906.765	10
2	75	3	5	Beechwood	139	Cimetière Be	1351777.0703	52
3	4	1	6	Bells Corners	4729	Bells Corners	6786857.875	111

3. Preparing the Excel spreadsheet

i.

i.

i.

a.

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.

- 21) Open **ONS Gen2 Web Data_July2020_transposed.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)

	А	В	С	D	E
1	name	ONS_ID	name_fr		
2	Old Barrhaven West	938	Old Barrha	ven Ouest	
3	Beacon Hill South - Cardinal Heights	3	Beacon Hil	Sud - Cardi	nal Heights
4	Beaverbrook	902	Beaverbroo	ok	

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

Т
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.

	А	В	С	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
		1			

- 22) <u>Take a few minutes to review the Layers sheet and familiarize yourself with what data is</u> <u>available. Write down 5-10 fields you're initially interested in, especially their Polygon</u> <u>Attribute designation (which are often codes such as CDP54b) and if the numbers in that</u> <u>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.

- 23) 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.
- 24) We need to <u>Transpose</u> (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.

a.



25) **Save As** and give the spreadsheet a new name in the WorkingData folder, something like ONS Gen2 Web Data July2020 transposed.xlsx

0				e Layout	Formula	s Data	Review	View	¥ Tell me	2							ß	Share	L Co	mmei
Paste	¥ ⊡ ~ ≪	Calibri B I	<u>U</u> •	•]11 ⊞ • 9	• A^ • • <u>A</u> •	A [×] =	= <u>=</u> \$ = = ⊡ = \≫`	Ge \$	neral × % ୨	€.0 .00 .00 €.00	E Co	nditional For rmat as Table II Styles ~	matting v	Delete	v ∑ v ∎ it v ∳	× A Z V Sort & Filter	Find & Select	4 Ideas	Sensit	ivity
1	÷	$\times \checkmark f_x$	id	D	F	E	c	u			V		м	N	0	P	0	P	_	c
id	р	op2016 froi area	3	pop_density	walkScore	bikeScore	pop2016	CDP54a	CDP54b	CDP54c	CDP54d	CDP54f	CDP54g	CDP54i	CDP54j	CDP63	CDP74	CDP77	CDP	3 80a
	0		2797.0	327.8	na	na	916855	155440	124855	504205	132355	17.0	13.6	55.0	14	.4 39.	5 57.) :	0.4	12
	3	7319	2.2	3252.6	60.5	88.8	7260	1020	880	3990	1370	14.0	12.1	55.0	18	.9 43.	2 51.	2 3	1.6	17
	5	560	0.7	na	66.2	81.7	na	NA	na	na	na	na	na	na	na	na	na	na	na	
	6	4926	3.4	1365.1	52.5	67.4	4580	515	620	2445	1005	11.2	13.5	53.4	21	.9 50.	7 54.	2 2	.8.6	17
	7	4346	1.4	3020.4	49.4	77.3	4140	730	405	2465	525	17.6	9.8	59.5	5 12	.7 38.	4 55.	1 3	.0.7	1
	11	6821	2.4	2870.3	47.0	73.8	6770	1005	970	3690	1095	14.8	14.3	54.5	5 16	.2 4	0 56.	2 3	.0.8	1
	12	5380	1.9	2722.8	34.4	70.0	5180	765	600	2800	1000	14.8	11.6	54.1	1 19	.3 42.	9 58.	2	6.9	1
	13	24400	6.5	3723.5	31.2	74.6	24240	4995	3320	13440	2490	20.6	13.7	55.4	10	.3 3	8 65.	3 1	5.4	
	16	6454	2.1	2993.0	48.3	64.3	6385	830	1365	3185	1005	13.0	21.4	49.9	15	.7 35.	3 45.	4	2.1	1
	1/	5	0.9	na	48.4	83.0	na	na	na	na	na	na	na	na	na	na	na	na	na	
	18	10141	2.7	3618.9	65.2	86.5	9820	1/95	1395	3910	1050	18.3	14.2	50.0	3 10	./ 35.	4 43. E E6	s 4	0.2	- 1
	2.5	28604	2.1	9254 5	04.2	90.0	27995	1033	2770	10255	2410	13.4	12.5	594.	119	2 24	4 29		2.0	
	29	11355	7.5	1454.3	57.0	74.7	10920	1535	1320	6225	1840	14.1	12.1	57 (16	8 41	3 57		8.6	- 1
	29	2314	5.9	377.2	8.9	28.1	2225	275	285	1370	300	12.4	12.8	61.6	13	5 4	7 63	, .	12.5	- 1
	30	3576	139.3	26.4	0.3	25.1	3680	700	485	2025	470	19.0	13.2	55.0	12	.8 43.	8 72.	3 3	1.5	-
	31	8951	4.0	2187.7	65.7	64.6	8645	1080	1705	4235	1625	12.5	19.7	49.0	18	.8 38.	7 51	3 3	6.5	1
	32	3527	5.5	695.3	14.4	70.8	3835	575	345	1975	930	15.0	9.0	51.5	5 24	.3 49.	3 68.	3 1	9.5	1
	35	5242	126.8	46.5	2.2	24.6	5900	1060	835	3085	920	18.0	14.2	52.3	15	.6 45.	2 68.	1 3	3.7	
	38	5471	1.8	3095.3	60.3	93.3	5610	1100	830	2925	750	19.6	14.8	52.1	13	.4 35.	2 49.	t 3	4.6	1
	45	10284	2.2	4899.9	51.4	86.1	10650	1960	1800	5935	940	18.4	16.9	55.7	/ 8	.8 36.	2 52.	9 3	5.2	1
	46	6508	1.2	5736.1	67.6	75.4	6640	1035	1010	3660	930	15.6	15.2	55.1	14	.0 37.	2 46.) i	6.5	1
	47	9999	1.8	5336.0	87.8	96.5	9780	1105	900	6525	1245	11.3	9.2	66.7	12	.7 36.	6 45.	5 3	7.9	1
	48	2678	15.3	131.7	21.8	42.6	2015	405	225	1240	155	20.1	11.2	61.5	5 7	.7 33.	7 65.	3 3	.5.4	
	49	10194	3.4	3067.2	49.7	78.4	10505	1840	1570	5425	1660	17.5	14.9	51.6	5 15	.8 38.	1 52.	1 3	.4.3	1
	50	7682	1.8	4844.9	35.9	82.7	8820	1565	1590	4730	940	17.7	18.0	53.6	5 10	.7 36.	8 52.	2 3	.5.6	1
	51	1371	4.6	180.3	41.3	67.6	825	105	90	440	195	12.7	10.9	53.3	3 23	.6 4	0 55.	9 2	.7.6	1
	52	8719	4.4	1917.6	51.2	82.3	8365	1550	1265	4520	1025	18.5	15.1	54.0	12	.3 38.	7 58.	0 3	0.5	1
-	5.4	6706	2.2	2720 6	62.1	60.0	2002	770	1000	2795	635	121	17.0		1.4	3 36	7 /0		67	1

- 26) 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

	1	WorkingDa	ita 🗘	<u> </u>	Q Search	
Favorites Desktop Sewing WorkFromHo Applications Downloads Documents	 ONS Gen2ansposed.xl ONS Gen2July2020.xl ONSMapCOVID_EN.csv Ottawa_Neien_2-shp.z 	sx sx ip II				
Online	e Locations	ile Format: CS	V UTF-8 (Comma delir	mited) (.cs	sv)	0

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

4. Bringing the CSVs into QGIS

28) In QGIS, go to Layer > Add Layer > Add Delimited Text Layer...

	Layer	Settings	Plugins	Vector	Raster	Database	Web	Mesh	Process
	🐙 Da	ata Source	Manager	ℋL		*JournalismN	lap - QG	SIS	
	Creat	te Layer		•				::	
	Add I	Layer			V _a Ac	d Vector Lay	/er		<mark>ዮ</mark> ፝፞፝ <mark>ដ</mark> ែ
	Embe	ed Layers a	nd Groups		🛃 Ac	d Raster Lay	/er		ΰ₩R
	Add	from Layer I	Definition I	File	Ac	d Mesh Laye	er		
	C	opy Style			ି <mark>କ</mark> Ac	d Delimited	Text La	yer	℃₩T
	Pa Pa	aste Style			🗣 Ac	d PostGIS La	ayers		℃¥D
					R Ac	d SpatiaLite	Layer		<mark>ፚ</mark> ፞፝፞፝ቘL
	C 🗊	opy Layer			De Ac	d MSSQL Sp	batial La	ayer	
	Pa Pa	aste Layer/G	Foup		DB2 Ac	d DB2 Spati	al Layer	·	企業2
		nen Attribut	to Table	EG	🔀 Ac	dd/Edit Virtua	al Layer.		
		aalo Editin		FO	😨 Ac	d WMS/WM	TS Laye	er	<mark>ዮ</mark> ₩₩
			J		🔐 Ac	d ArcGIS Ma	pServe	r Layer.	
		ave Layer Edito	JILS		🕄 Ac	d WCS Laye	r		
		urrent Edits			M Ac	d WFS Laye	r		
	Save	As			🔐 Ac	d ArcGIS Fe	atureSe	erver Lay	/er
a.	Save	As Layer D	efinition Fi	le		\sim	-	()	A A

- 29) 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)

	Data Source Manager Delimited Text
Browser	File name rkFromHome/Presentations/David McKie/JournalismMap/WorkingData/ONS_transposed_csv.csv
Vector	Layer name ONS_transposed_csv Encoding UTF-8
Raster	▼ File Format
Mesh	CSV (comma separated values)
Delimited Text	Regular expression delimiter
GeoPackage	Custom delimiters
🖉 SpatiaLite	Record and Fields Options
PostgreSQL	▼ Geometry Definition
MSSQL	Point coordinates Mell knows to MM(T)
22 DB2	No geometry (attribute only table)
Virtual Layer	Layer Settings
имѕ/wмтs	Sample Data
异 wcs	id pop2016_fromDB area pop_density walkScore bikeScore pop2016 CDP54a CDP54b CDP54c CD ↑ 0 0000000000000000000000000000000000
WFS	1 0 2/9/0 32/8 na na 916855 155440 124855 504205 13. 2 3 7319 2.2 3252.6 60.5 88.8 7260 1020 880 3990 13.
ArcGIS Map Server	
ArcGIS Feature Server	
GeoNode	Add Clo

- 30) Click Add
- 31) Before closing the Delimited Text window, fill it in the same way but for the **ONSMapCOVID_EN CSV** file.
 - a. Remember, no geometry!
- 32) Add, then Close.
 - a. Your Layers pane should look like this:



33) Save your QGIS project.

34) Congratulations, you're now ready to work with all three datasets in QGIS!