

Beyond Matching the Postal Code to Census Geography

Vince Gray

vince@uwo.ca

ACCOLEDS, 2016-11-30

I would like to note that I was honoured to be asked to come out to talk with you – I hope that you find this presentation both interesting and worthwhile. Having said that, I'm rather apprehensive about this presentation.

Why match geography to Postal Code?

A non-exhaustive list of possible reasons:

- Where are my (respondents/clients) from?
- How (close to/distant from) a particular object (e.g., wind turbine, health clinic, proposed pipeline) are my (respondents/clients)?
- What are the socioeconomic characteristics of the region(s) from which my sample is drawn?
- Is the data set that I collected representative of the characteristics of the region(s) from which my sample is drawn?

We now interrupt your regularly scheduled presentation for a disclaimer!

- This isn't the presentation I originally envisioned
- I hit some potholes or encountered detours on the road to the presentation I expected
- I don't know how significant my findings are, but they bothered me, so I thought that I should share them with you (and hopefully receive reassurance from you)

Point 3 – I hope that you will be able to provide me with guidance on them ...

Conventional wisdom - PCCF

- Use PCCF if:

- You don't have SAS and the knowledge to use PCCF+
 - If you have SAS, but don't know how to use PCCF+, see Jeff's 2015 presentation!

OR

- You are *[mainly]* working with urban postal codes

OR

- User doesn't want to be bothered with using PCCF+

Conventional wisdom – PCCF+

- Use PCCF+ if:

- You have SAS and the knowledge to use it AND
- You have *[a lot of]* rural postal codes

Note: zero in the FSA (e.g., A0A) = rural postal code

OR

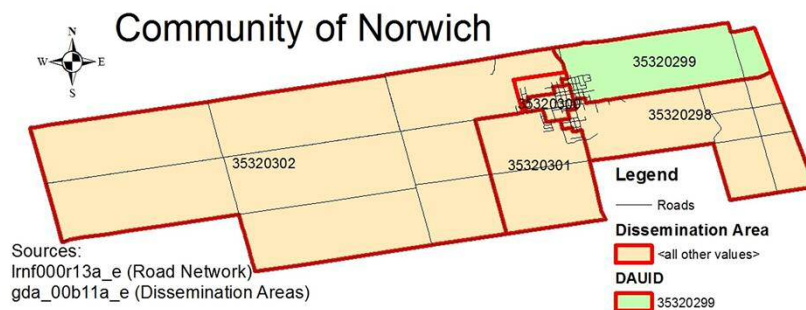
- You have incomplete Postal Codes to match against

OR

- You want to match to multiple census geographies

Why PCCF+ for rural?

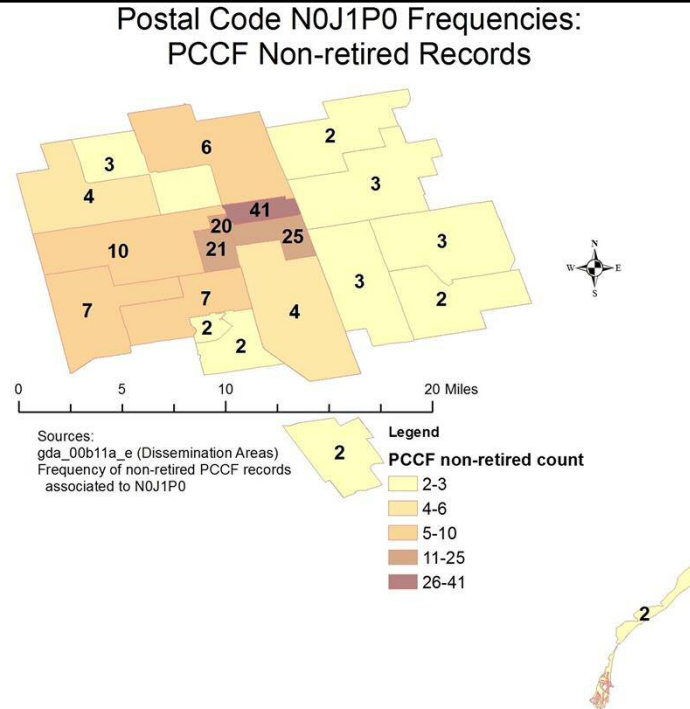
- Best-match PCCF file matches **N0J1P0** (Norwich, Ontario) to DA **35320299**
 - The Norwich post office (where all the mailboxes are) is in 35302300 (not that that's important, just odd);
 - At least 4 other possible DAs will NEVER be selected by PCCF



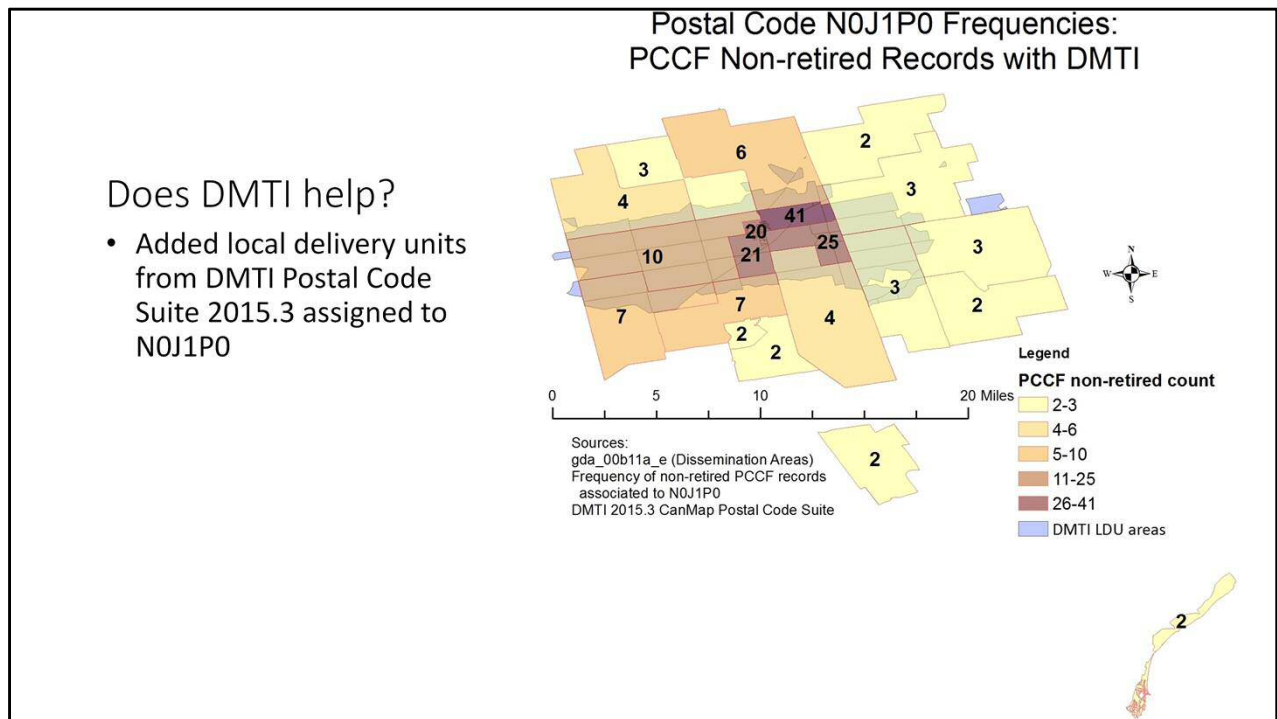
I chose Norwich because my mother-in-law lives there, and my wife is from there. Norwich is a little village of about 2,000-odd people. I've overlaid the village streets on the dissemination area boundaries that include parts of Norwich.

N0J1P0 isn't just the village of Norwich

These are the non-retired dissemination areas linked to postal code N0J1P0 in the May 2013 PCCF: the numbers are the number of times found in the PCCF



A couple of these areas don't pass the smell test to me - Long Point (which juts out into Lake Ontario) is a long way from Norwich, and I'm concerned that there is something wrong. You should be getting a hint now of what went wrong with the presentation ... Note that the "best match" also happens to have the most occurrences in the PCCF.



So, I decided to add the DMTI postal suite boundaries, to figure out where N0J1P0 actually is. I'm left with some concerns about this as well: new western and eastern areas are included, and there are holes and unconnected zones in the area. But, the two unconnected areas to the south are DEFINITELY not part of N0J1P0 according to DMTI. A couple more areas to the south are also excluded, along with two to the north and one to the east.

Problem ...

- What is the real geographic area covered by N0J1P0?
- Do we have related problems?
 - e.g., inaccuracy or conflicting info from DMTI and PCCF?
- How likely is it that N0J1P0 is the only problematic postal code? (i.e., that there really is a problem?)
 - How should we (or should we) test this more rigorously?
 - Something for people who actually know GIS well to work on?
 - Something to ask Statistics Canada about?

So, part of the difficulty with the presentation is revealed ...

Canada Post provides address lookup based on postal code

<https://www.canadapost.ca/cpotools/apps/fpc/personal/findAnAddress?execution=e1s1>

- Looking for N0J1P0 returns 113 records that Canada Post associates with it
- Many records are general delivery, as the mail for the village is all delivered to the boxes in the post office

The screenshot shows the Canada Post website's 'Find an Address' tool. At the top, there is a maintenance notice. Below the navigation bar, the search results for postal code N0J 1P0 are displayed. A message states 'We found more than 100 result(s)' and advises selecting a postal code to continue. A table lists specific addresses, including building numbers, delivery modes, street names, and postal codes. To the right, there is a section for 'Postal Rates' with links to view all rates for Canada, USA, and International destinations.

Building	#	Delivery Mode	Street Name	Suite	City	Prov.	Postal Code
772876-773116	Even	RR 1	59 HWY		NORWICHON		N0J 1P0
772877-773121	odd	RR 1	59 HWY		NORWICHON		N0J 1P0

If anyone know what postal code N0J1P0 means, it should be Canada post – after all, they have copyrighted their work ...

There is also a single entry for the post office

Free tools to geocode
the addresses

More than two programs;
samples are

- <http://hamstermap.com>
- <http://findlatitudeandlongitude.com/batch-geocode>

Principles of both are the
same: type or paste in the
addresses you want to code
and press a button to convert
or geocode

HamsterMap.com

Translate addresses into coordinates and put them on the map!

Paste your addresses into a textbox below and hit the **Geocode!** button. The latitude and longitude will appear in a real-time.

Note that there is no guarantee that geocoding will be 100% correct - best effort has to suffice! Remember, geocoding is not an exact science and you can reasonably expect about 90% precision in towns and slightly less in rural areas.

After you get the decimal coordinates, you can [place these on the MAP](#). You will have to use Tab separated output if you wish to do so.

[Like](#) [Share](#) [334](#)

CONVERT Coordinates **GEOCODE Addresses** **Quick MAP** **Custom MAP**

Type / Copy & Paste your addresses

4 Elmwood Place, London, Ontario, Canada
471 Alston Road, London, Ontario, Canada
37 Florence street, norwich, ontario, canada

Get the Coordinates

42.969557	-81.262536	Exact
42.943178	-81.239091	Exact
42.995369	-80.685095	Next

Filter by ACCURACY:

☒ Exact
☒ Next door
☒ Street
☒ District

☐ Include addresses into output
☐ add found addresses into output
☒ Include accuracy into output

Output type: ☐ Comma separated ☒ Tab separated **Convert** **Reset**

The following diagram demonstrates what exactly different accuracy levels mean

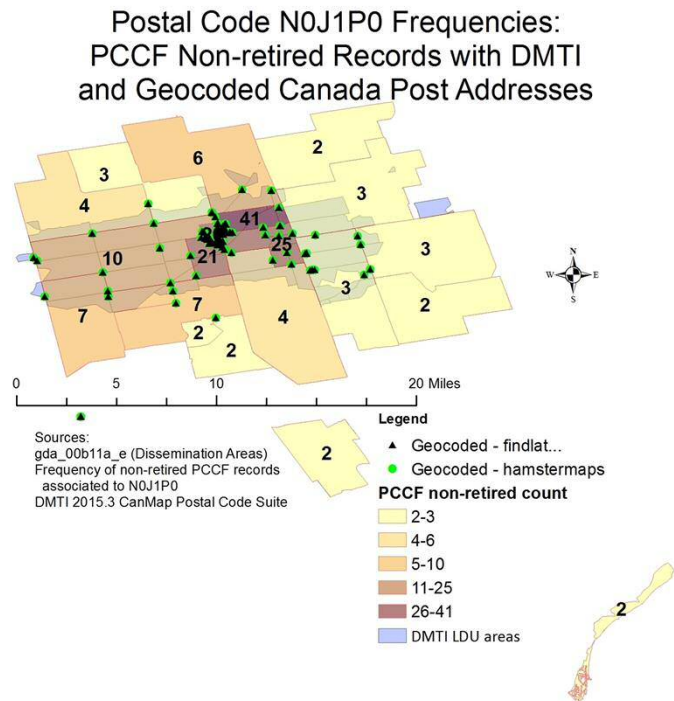
Accuracy: Street Exact Next Door

- The sample records I've got on the screen are some of my relatives' addresses, included to show how it works
- Either will display the level of accuracy coded to; HamsterMap lets you filter records out by accuracy and is clearer on accuracy meaning (uses text rather than a code to represent accuracy)
- I know that there ISN'T a 37 Florence Street – that would be in the cemetery
- HamsterMaps that address Norwich as a "next door" match, while findlat... recorded it as a "street" match
- Findlat ... includes field headers
- For purposes of the Canada Post street segments, I took the first number in the range and associated it with the street, and pasted all 113 records in to geocode them

Overlaying Canada Post's geocoded street segments

Hamstermap and Findlat... appear to use the same algorithm (and database?) since points overlay but use different terminology for accuracy

Have even more concerns about at least the two outliers from non-retired PCCF records: they don't show up with the geocoded addresses



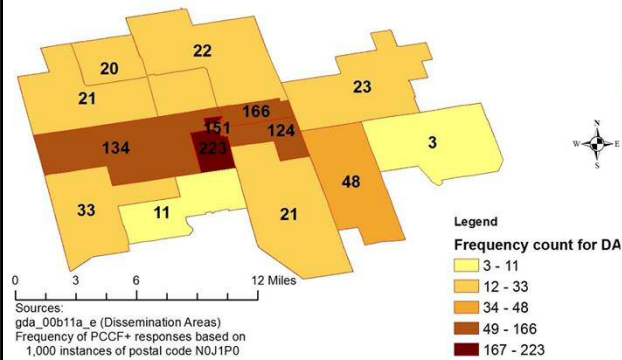
- Note the one outlier

So, how does PCCF+ fare?

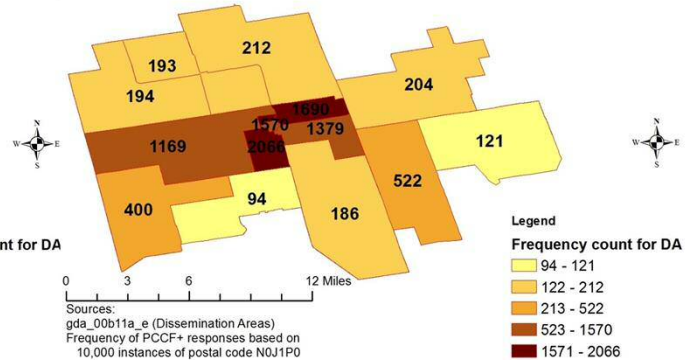
- Problem is how to test it
 - In a sample of London hospital in-patient records, you might get one or a few records with NOJ1P0 addresses
- Created a text file with 1,000 repetitions of NOJ1P0, and a second file with 10,000 repeats
- Ran PCCF+ against both
- Calculated the frequency of dissemination area returned by PCCF+
- Mapped those results

Comparison of PCCF retrieval results: 1,000 and 10,000 occurrences of postal code

Postal Code N0J1P0 Frequencies:
Test of 1,000 PCCF+ Instances



Postal Code N0J1P0 Frequencies:
Test of 10,000 PCCF+ Instances



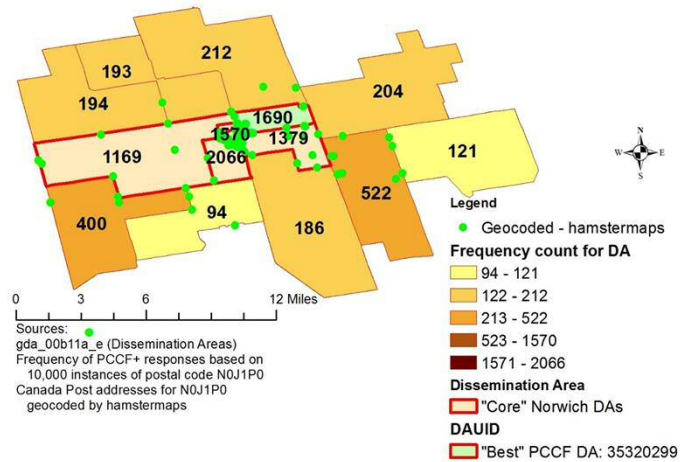
Returning to original map

Bottom line: PCCF+ results coincide with Canada Post's geocoded addresses REASONABLY well – but still don't really know PC borders

The PCCF "best match" record is the second-most frequently selected postal code by PCCF

BUT: how well would a single matched DA actually reflect the postal code?

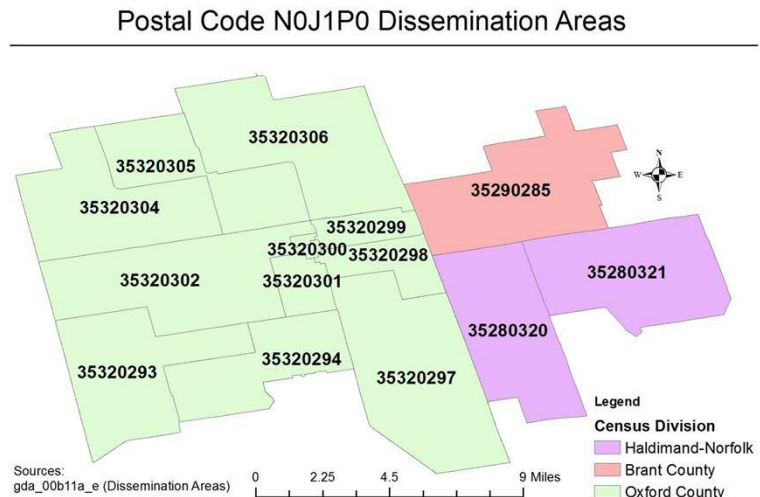
Postal Code N0J1P0 Frequencies:
Test of 10,000 PCCF+ Instances with Canada Post addresses, "core" Norwich records and PCCF Best Match



- I would expect that the Das on the periphery might also be selectable by other postal codes

Looking at the DAs

The whole purpose of going to PCCF or PCCF+ was to obtain census geography: now we can match data to places



If we look more closely, the fourteen dissemination areas represented by this single postal code span three different census divisions, and three census subdivisions – 11 of the 14 are in Oxford County, but 3 are in either Brant County or Haldimand-Norfolk. These boundaries happen to coincide with federal electoral districts

What is the importance of FED?

- If we are studying election results, and know both Postal code and District, may be able to be far more precise in the demographics retrieved

**We now return to your regularly
scheduled presentation**

But beware of more interruptions ...

PCCF+ results

- All levels of current census geography
- Dissemination blocks for 2011/06
- DA or EA for each cycle of the census back to 1981

ID	PCODE	DA2011uid	DA06uid	DB06uid	DA01uid	EA96uid	EA91uid	EA86uid	EA81uid	VERSION
47	900 NOJ1P0	35280320	35280320	3528032011	35280093	35027304	35027304	35026417	35026417	6B1
48	865 NOJ1P0	35280320	35280320	3528032011	35280093	35027304	35027304	35026417	35026417	6B1
49	241 NOJ1P0	35280320	35280320	3528032012	35280093	35027304	35027304	35026417	35026417	6B1
50	845 NOJ1P0	35280321	35280321	3528032101	35280092	35027305	35027305	35026416	35026416	6B1
51	570 NOJ1P0	35280321	35280321	3528032101	35280092	35027305	35027305	35026416	35026416	6B1
52	872 NOJ1P0	35280321	35280321	3528032106	35280093	35027304	35027304	35026416	35026416	6B1
53	336 NOJ1P0	35290285	35290285	3529028501	35290219	35062019	35062002	35004059	35004059	6B1
54	693 NOJ1P0	35290285	35290285	3529028501	35290219	35062019	35062002	35004059	35004059	6B1
55	962 NOJ1P0	35290285	35290285	3529028502	35290218	35062002	35062002	35004058	35004058	6B1
56	875 NOJ1P0	35290285	35290285	3529028502	35290218	35062002	35062002	35004058	35004058	6B1
57	830 NOJ1P0	35290285	35290285	3529028503	35290218	35062002	35062002	35004058	35004060	6B1
58	3 NOJ1P0	35290285	35290285	3529028503	35290218	35062002	35062002	35004058	35004060	6B1
59	64 NOJ1P0	35290285	35290285	3529028504	35290218	35062002	35062002	35004058	35004057	6B1
60	285 NOJ1P0	35290285	35290285	3529028505	35290218	35062002	35062002	35004058	35004058	6B1
61	551 NOJ1P0	35290285	35290285	3529028506	35290218	35062002	35062002	35004058	35004058	6B1
62	713 NOJ1P0	35290285	35290285	3529028507	35290218	35062002	35062002	35004058	35004058	6B1
63	985 NOJ1P0	35290285	35290285	3529028507	35290218	35062002	35062002	35004058	35004058	6B1
64	943 NOJ1P0	35290285	35290285	3529028507	35290218	35062002	35062002	35004058	35004058	6B1

PCCF returns the “best match” or all census geography for specific release

Options for finding data to match to DA/EA

- Beyond 20/20 files
- CHASS Census Analyzer

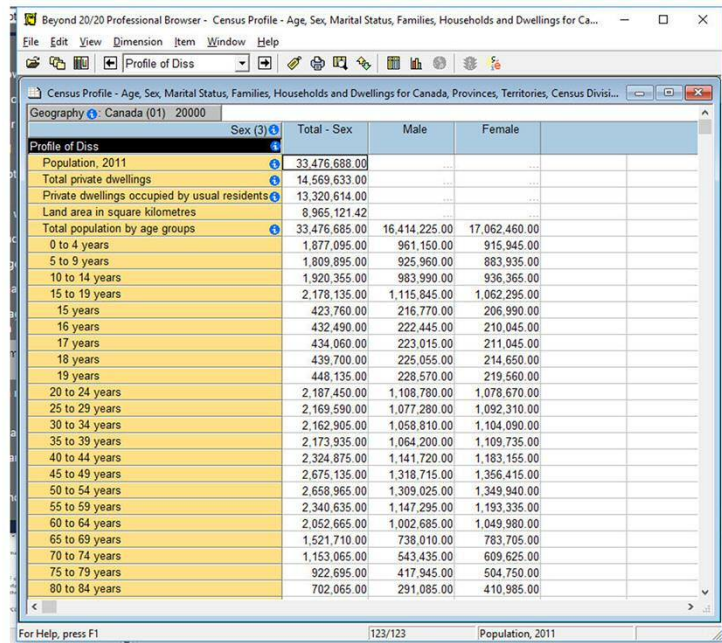
ow, we're going to return to the preconceived presentation ...

Of course, there are population and dwelling counts by dissemination block, but I can't imagine a use of those in conjunction with matching postal codes

Using B20/20
to extract data

Open the file with
needed data

I've selected the 2011
census profile of DAs



The screenshot shows a software window titled "Beyond 20/20 Professional Browser - Census Profile - Age, Sex, Marital Status, Families, Households and Dwellings for Ca...". The main content area displays a table for "Census Profile - Age, Sex, Marital Status, Families, Households and Dwellings for Canada, Provinces, Territories, Census Divi...". The table is filtered for "Geography: Canada (01) 20000". The table has columns for "Sex (3)", "Total - Sex", "Male", and "Female". The rows represent various demographic categories, including "Population, 2011", "Total private dwellings", "Private dwellings occupied by usual residents", "Land area in square kilometres", and "Total population by age groups". The "Total population by age groups" section is expanded, showing data for age groups from "0 to 4 years" to "80 to 84 years".

	Sex (3)	Total - Sex	Male	Female
Population, 2011		33,476,688.00		
Total private dwellings		14,569,633.00		
Private dwellings occupied by usual residents		13,320,614.00		
Land area in square kilometres		8,965,121.42		
Total population by age groups		33,476,688.00	16,414,225.00	17,062,460.00
0 to 4 years		1,877,095.00	961,150.00	915,945.00
5 to 9 years		1,809,895.00	925,960.00	883,935.00
10 to 14 years		1,920,355.00	983,990.00	936,365.00
15 to 19 years		2,178,135.00	1,115,845.00	1,062,290.00
15 years		423,760.00	216,770.00	206,990.00
16 years		432,490.00	222,445.00	210,045.00
17 years		434,060.00	223,015.00	211,045.00
18 years		439,700.00	225,055.00	214,650.00
19 years		448,135.00	228,570.00	219,560.00
20 to 24 years		2,187,450.00	1,108,780.00	1,078,670.00
25 to 29 years		2,169,590.00	1,077,280.00	1,092,310.00
30 to 34 years		2,162,905.00	1,058,810.00	1,104,090.00
35 to 39 years		2,173,935.00	1,064,200.00	1,109,735.00
40 to 44 years		2,324,875.00	1,141,720.00	1,183,155.00
45 to 49 years		2,675,135.00	1,318,715.00	1,356,415.00
50 to 54 years		2,658,965.00	1,309,025.00	1,349,940.00
55 to 59 years		2,340,635.00	1,147,295.00	1,193,335.00
60 to 64 years		2,052,665.00	1,002,685.00	1,049,980.00
65 to 69 years		1,521,710.00	738,010.00	783,705.00
70 to 74 years		1,153,065.00	543,435.00	609,625.00
75 to 79 years		922,695.00	417,945.00	504,750.00
80 to 84 years		702,065.00	291,085.00	410,985.00

Note that the Census profile here does not immediately specify which file was selected – title is there, but not the table identifier (i.e., 98-312-xcb2011007).

Right click on the variable(s) you want – and show it/them

Beyond 20/20 Professional Browser - [Census Profile - Age, Sex, Marital Status, Families, Households and Dwellings for Ca...]

File Edit View Dimension Item Window Help

Sex (3)

Geography Canada (01) 20000

	Sex (3)	Total - Sex	Male	Female
Profile of Diss				
Population, 2011	33,476			
Total private dwellings	14,569			
Private dwellings occupied by usual residents	13,320			
Land area in square kilometres	8,965			
Total population by age groups	33,476			
0 to 4 years	1,877			
5 to 9 years	1,809			
10 to 14 years	1,920			
15 to 19 years	2,178,135.00	1,115,845.00	1,062,295.00	
15 years	423,760.00	216,770.00	206,990.00	
16 years	432,490.00	222,445.00	210,045.00	
17 years	434,060.00	223,015.00	211,045.00	
18 years	439,700.00	225,055.00	214,650.00	
19 years	448,135.00	228,570.00	219,560.00	
20 to 24 years	2,187,450.00	1,108,780.00	1,078,670.00	
25 to 29 years	2,169,590.00	1,077,280.00	1,092,310.00	
30 to 34 years	2,162,905.00	1,058,810.00	1,104,090.00	
35 to 39 years	2,173,935.00	1,064,200.00	1,109,735.00	
40 to 44 years	2,324,875.00	1,141,720.00	1,183,155.00	
45 to 49 years	2,675,135.00	1,318,715.00	1,356,415.00	
50 to 54 years	2,658,965.00	1,309,025.00	1,349,940.00	
55 to 59 years	2,340,635.00	1,147,295.00	1,193,335.00	
60 to 64 years	2,052,665.00	1,002,685.00	1,049,980.00	
65 to 69 years	1,521,710.00	738,010.00	783,705.00	
70 to 74 years	1,153,065.00	543,435.00	609,625.00	
75 to 79 years	922,635.00	417,945.00	504,750.00	
80 to 84 years	702,065.00	291,085.00	410,985.00	
85 years and over	645,515.00	208,300.00	437,215.00	
Median age of the population	40.60	39.60	41.50	

Show only those items which are selected (1) 3/3 Total - Sex

Repeat with other dimensions as needed

CTRL-Click to retain selections before *showing* – ***if you don't, you lose what you've selected!***

	Sex (3)	Total - Sex
Population, 2011		432,490.00
Total private dwellings		434,060.00
Private dwellings occupied by usual residents		439,700.00
Land area in square kilometres		448,135.00
Total population by age groups		2,187,450.00
0 to 4 years		2,169,590.00
5 to 9 years		2,162,905.00
10 to 14 years		2,173,935.00
15 to 19 years		2,324,875.00
15 years		2,675,135.00
16 years		2,658,965.00
17 years		2,340,635.00
18 years		2,052,665.00
19 years		1,521,710.00
20 to 24 years		1,153,065.00
25 to 29 years		922,695.00
30 to 34 years		702,065.00
35 to 39 years		645,515.00
40 to 44 years		
45 to 49 years		
50 to 54 years		
55 to 59 years		
60 to 64 years		
65 to 69 years		
70 to 74 years		
75 to 79 years		
80 to 84 years		
85 years and over		
Median age of the population		40.60
% of the population aged 15 and over		83.20

Here, we've got the variables we want

Geography: Canada (01) 2000

Sex (1)	Total - Sex
Population, 2011	33,476,688.00
Median age of the population	40.60
Average number of children at home per census family	1.10

For Help, press F1 | 1/3 | Total - Sex | ENG

Drag the dimensions around to get the geography in the column

Dimension >> Search for the DA you want

Search data or fields along the active dimension

Geography	Population, 2011	Median age of the popula...	Average number of children at...
Canada (01)	33,476,688.00	40.60	1.10
Newfoundland and Labrador (02)	514,536.00	44.00	0.90
Prince Edward Island (03)	160,000.00	41.60	1.00
Saskatchewan (04)	1,000,000.00	62.00	0.40
Alberta (05)	2,624,410.00	62.00	0.40
British Columbia (06)	4,665,000.00	60.00	0.70
Ontario (07)	13,000,000.00	52.40	0.70
Quebec (08)	7,500,000.00	52.40	0.70
Atlantic Provinces (09)	2,500,000.00	57.60	0.60
Northwest Territories (10)	45,000.00	55.20	0.70
Yukon (11)	35,000.00	57.90	0.30
Nunavut (12)	45,000.00	57.90	0.40
St. John's (1001120) T 00000	81.00	57.90	0.30
St. John's (1001120) U 00000	81.00	57.90	0.40
Division No. 1, Subd. U (1001124) SNO 00000	1,489.00	48.20	1.00
10010737 (10010737) 00000	396.00	51.80	0.90
10010741 (10010741) 00000	649.00	47.80	0.90
10010743 (10010743) 00000	444.00	44.70	1.00
Cape Breton (1001126) T 00000	506.00	47.90	0.90
10010742 (10010742) 00000	506.00	47.90	0.90
Renews-Cappahayden (1001131) T 00000	310.00	53.70	0.60
10010740 (10010740) 00000	310.00	53.70	0.60
Fermeuse (1001136) T 00000	323.00	49.70	1.10
10010738 (10010738) 00000	323.00	49.70	1.10
Port Kirwan (1001140) T 00000	65.00	56.20	0.40
10010739 (10010739) 00000	65.00	56.20	0.80
Aquatorte (1001144) T 00000	83.00	59.50	0.60
10010745 (10010745) 00000	83.00	59.50	0.40
Ferryland (1001149) T 00000	465.00	51.20	0.90

61764/61764 Canada (01) 20000

Specify the code for the DA you want – or for the first DA in the CD – and **find next**

The screenshot shows the 'Census Profile - Age, Sex, Marital Status, Families, Households and Dwellings for Canada, Provinces, Territories, Census Divi...' window. The main data table displays census data for various regions, with columns for 'Profile of Diss', 'Population, 2011', 'Median age of the popula.', and 'Average number of children at...'. A search dialog box is open, allowing the user to search for a specific dimension (Geography) and field (Data). The search results show a list of regions and their corresponding population and age data.

Profile of Diss	Population, 2011	Median age of the popula.	Average number of children at...
Canada (01) 20000	33,476,688.00	40.60	1.10
Newfoundland and Labrador / Terre-Neuve-et-L...	514,536.00	44.00	0.90
Division No. 1 (1001) 00000	262,410.00	41.60	1.00
Division No. 1, Subd. V (1001)			
10010732 (10010732) 01000			
Portugal Cove South (1001105)			
10010734 (10010734) 00000			
Trepassey (1001113) T 00000			
10010735 (10010735) 00000			
10010736 (10010736) 00000			
St. Shott's (1001120) T 00000			
10010733 (10010733) 00000			
Division No. 1, Subd. U (1001)			
10010737 (10010737) 00000			
10010741 (10010741) 00000			
10010743 (10010743) 00000			
Cape Broyle (1001126) T 0000			
10010742 (10010742) 00000			
Renews-Cappahayden (1001131) T 00000	310.00	53.70	0.60
10010740 (10010740) 00000	310.00	53.70	0.60
Fermeuse (1001136) T 00000	323.00	49.70	1.10
10010738 (10010738) 00000	323.00	49.70	1.10
Port Kirwan (1001140) T 00000	65.00	56.20	0.40
10010739 (10010739) 00000	65.00	56.20	0.80
Aquaforte (1001144) T 00000	83.00	59.50	0.60
10010745 (10010745) 00000	83.00	59.50	0.40
Ferryland (1001149) T 00000	465.00	51.20	0.90

This will take you to the first occurrence of that code

Beyond 20/20 Professional Browser - Census Profile - Age, Sex, Marital Status, Families, Households and Dwellings for Ca...

File Edit View Dimension Item Window Help

Geography

Census Profile - Age, Sex, Marital Status, Families, Households and Dwellings for Canada, Provinces, Territories, Census Divi...

Sex (3) Total - Sex

Geography	Profile of Diss	Population, 2011	Median age of the popula...	Average number of children at...
35260833 (35260833) 00001		711.00	37.70	1.20
Haldimand-Norfolk (3528) 00000		109,118.00	44.80	1.00
Haldimand County (3528018) CY 00000		44,876.00	43.10	1.10
35280151 (35280151) 00000		586.00	42.00	1.10
35280152 (35280152) 00000		581.00	41.30	1.10
35280153 (35280153) 00000		556.00	41.50	1.10
35280154 (35280154) 01000		577.00	45.20	0.90
35280155 (35280155) 00000		555.00	36.20	1.10
35280156 (35280156) 00000		543.00	43.80	1.10
35280157 (35280157) 00000		509.00	45.70	1.00
35280158 (35280158) 00000		667.00	37.30	1.10
35280159 (35280159) 00000		475.00	35.70	1.30
35280160 (35280160) 00000		765.00	42.90	1.20
35280161 (35280161) 00000		810.00	40.60	1.00
35280162 (35280162) 00000		779.00	34.70	1.40
35280164 (35280164) 00000		637.00	44.70	1.00
35280166 (35280166) 01000		488.00	47.10	0.90
35280167 (35280167) 00000		375.00	41.90	0.90
35280168 (35280168) 00000		489.00	45.10	1.00
35280169 (35280169) 00000		804.00	41.80	0.90
35280170 (35280170) 00000		709.00	49.70	1.10
35280171 (35280171) 00000		444.00	39.00	1.30
35280172 (35280172) 00000		672.00	38.50	1.20
35280173 (35280173) 00000		512.00	41.30	1.20
35280174 (35280174) 00000		459.00	35.20	1.20
35280175 (35280175) 00000		625.00	52.80	0.90
35280176 (35280176) 00000		449.00	34.40	1.10

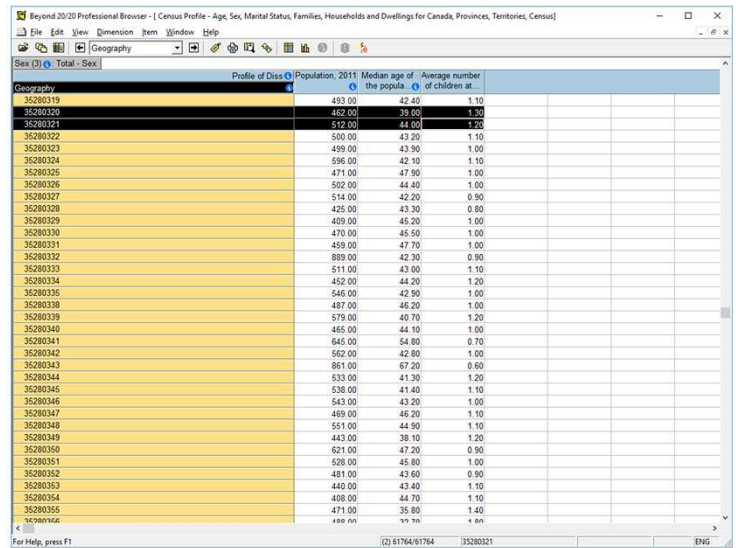
For Help, press F1

61764/61764

Haldimand County (3528018) CY 1

Change the labels to
be only the DA name

Use **Dimension >>**
Change labels as
needed to get only the
DAUID



The screenshot shows a web browser window titled "Beyond 20/20 Professional Browser - [Census Profile - Age, Sex, Marital Status, Families, Households and Dwellings for Canada, Provinces, Territories, Census]". The browser is displaying a table of data under the "Geography" dimension. The table has four columns: "Geography", "Population, 2011", "Median age of the popula...", and "Average number of children at...". The rows are listed by geography code, starting from 35280319 and ending at 35280355. The data is presented in a table format with alternating yellow and white rows.

Geography	Population, 2011	Median age of the popula...	Average number of children at...
35280319	493.00	42.40	1.10
35280320	462.00	39.00	1.30
35280321	512.00	44.00	1.20
35280322	500.00	43.20	1.10
35280323	499.00	43.90	1.00
35280324	596.00	42.10	1.10
35280325	471.00	47.90	1.00
35280326	502.00	44.40	1.00
35280327	514.00	42.20	0.90
35280328	425.00	43.30	0.80
35280329	409.00	45.20	1.00
35280330	470.00	45.50	1.00
35280331	459.00	47.70	1.00
35280332	889.00	42.30	0.90
35280333	511.00	43.00	1.10
35280334	452.00	44.20	1.20
35280335	546.00	42.90	1.00
35280336	487.00	46.20	1.00
35280337	579.00	40.70	1.20
35280340	465.00	44.10	1.00
35280341	645.00	54.80	0.70
35280342	562.00	42.80	1.00
35280343	881.00	57.20	0.50
35280344	533.00	41.30	1.20
35280345	538.00	41.40	1.10
35280346	543.00	43.20	1.00
35280347	499.00	46.20	1.10
35280348	551.00	44.90	1.10
35280349	443.00	38.10	1.20
35280350	621.00	47.20	0.90
35280351	528.00	45.80	1.00
35280352	481.00	43.60	0.90
35280353	440.00	43.40	1.10
35280354	408.00	44.70	1.10
35280355	471.00	35.80	1.40
35280356	408.00	39.70	1.40

Select the DAs that
you need – from the
list obtained in ArcGIS

The screenshot shows a software window titled "Beyond 20/20 Professional Browser - [Census Profile - Age, Sex, Marital Status, Families, Households and Dwellings for Canada, Provinces, Territories, Census]". The window displays a table of census data for various geographical areas (DAs). The table has four main columns: "Geography", "Population, 2011", "Median age of the popula.", and "Average number of children at...". The data is presented in a list format with a yellow background for the first 10 rows. The status bar at the bottom indicates "For Help, press F1", "14/61764", "3520321", and "ENG".

Geography	Population, 2011	Median age of the popula.	Average number of children at...
3520320	462.00	39.00	1.30
3520321	512.00	44.00	1.20
3520325	575.00	41.20	1.40
3520323	502.00	26.90	2.00
3520324	614.00	30.30	1.60
3520327	529.00	37.40	1.30
3520326	700.00	37.80	1.20
3520329	843.00	32.60	1.30
3520300	754.00	37.40	1.20
3520301	910.00	35.10	1.30
3520302	586.00	28.50	1.80
3520304	529.00	31.90	1.60
3520305	619.00	33.60	1.70
3520306	613.00	33.20	1.40

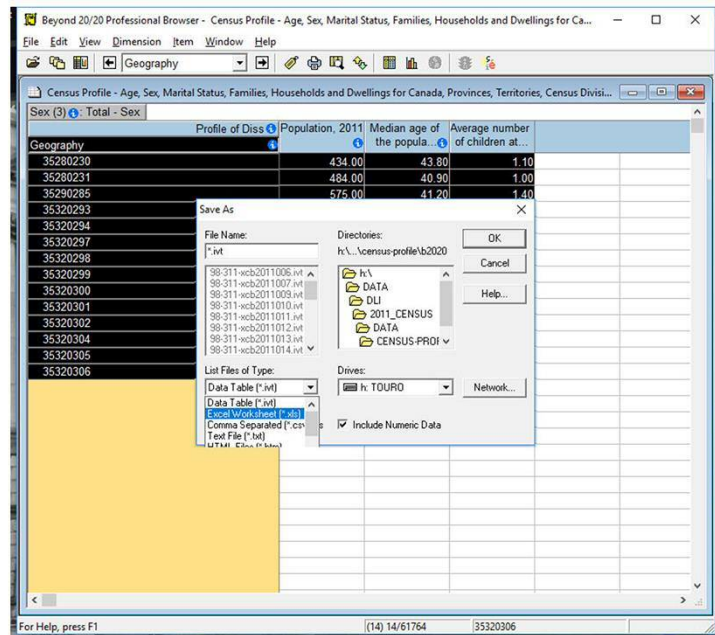
Save the data

Be sure to ITEM
SELECT ALL – *if you
don't, you may only
select and save a
single cell!!!*

The screenshot shows the 'Beyond 20/20 Professional Browser' window. The main table displays census data for 'Dwellings for Canada, Provinces, Territories, Census Divisi...'. The 'Geography' column is highlighted in yellow, and a context menu is open over it, showing options like 'Summary', 'Show', 'Hide', 'Select All', 'Select Group', and 'Cell Information'. The 'Select All' option is highlighted. The table has columns for 'Sex (3) Total - Sex', 'Geography', and numerical data. The status bar at the bottom indicates 'Select all shown items of active dimension', '14/61764', and '35320306'.

Sex (3) Total - Sex	Geography	011 Median age of the popula...	Average number of children at...
35280230		43.80	1.10
35280231		40.90	1.00
35290285		41.20	1.40
35320293		26.90	2.00
35320294		30.30	1.60
35320297		37.40	1.30
35320298		37.80	1.20
35320299		32.60	1.30
35320300		37.40	1.20
35320301		35.10	1.30
35320302		28.50	1.80
35320304		31.90	1.60
35320305		33.60	1.70
35320306		33.20	1.40

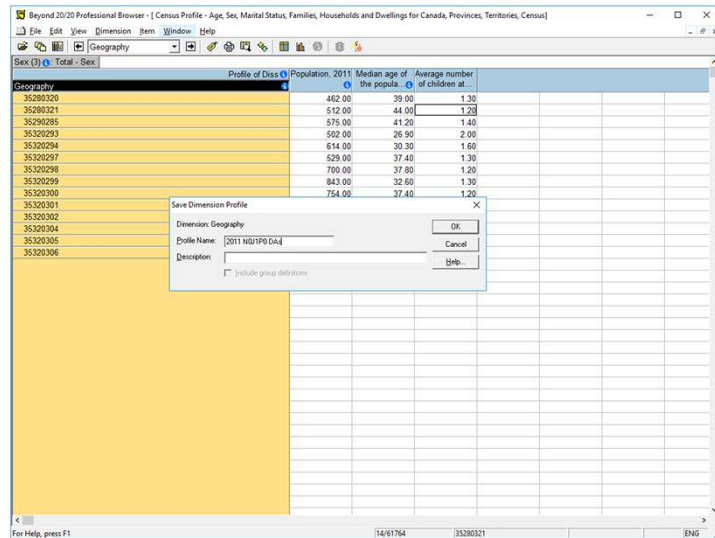
Specify the file name,
directory, and format
to save in



You can stop
here, but ...

***Dimension >> Save
profile***

to allow you to easily
select the same
characteristics in
other B2020 files



If you save the profile, you can reuse it in other similar files – I hope to use it against NHS to avoid having to select all fourteen Dissemination areas again

Rinse & repeat with NHS

Beyond 20/20 Professional Browser - [2011 NHS Semi-custom Profile for Canada, Provinces, Territories, Census Divisions, Census Subdivisions and Dissemination]

File Edit View Dimension Item Window Help

2011 NHS

Geography	Canada (26.1%)	Newfoundland and Labrador...	Division No. 1 (31.1%)	Division No. 1 Subd. V SN.	10010732 (12.0%)	Pop. Sqr.
2011 NHS						
Total population in private households by citizenship	32,852,320	507,270	258,470	60	60	
Canadian citizens	20,895,310	502,750	254,940	55	55	
Canadian citizens aged under 18	6,516,425	53,100	47,595	0	0	
Canadian citizens aged 18 and over	24,318,885	409,650	206,945	55	55	
Not Canadian citizens	1,957,015	4,520	3,530	0	0	
Males in private households by citizenship	16,163,115	247,955	125,565	30	30	
Canadian citizens	15,232,595	245,450	123,525	25	30	
Canadian citizens aged under 18	3,381,280	48,615	25,025	0	0	
Canadian citizens aged 18 and over	11,851,315	196,835	98,500	25	25	
Not Canadian citizens	530,515	2,510	2,035	0	0	
Females in private households by citizenship	16,688,210	259,315	132,965	30	30	
Canadian citizens	15,662,715	257,300	131,415	30	30	
Canadian citizens aged under 18	3,195,140	44,485	22,965	0	0	
Canadian citizens aged 18 and over	12,467,570	212,815	108,450	30	30	
Not Canadian citizens	1,026,500	2,015	1,490	0	0	
Total population in private households by immigrant status and period of immigration	32,852,320	507,270	258,470	55	55	
Non-immigrants	25,720,175	496,405	250,540	60	55	
Immigrants	6,775,765	9,160	6,510	0	0	
Before 1971	1,261,060	2,185	1,415	0	0	
1971 to 1980	870,775	1,520	1,065	0	0	
1981 to 1990	949,895	1,095	780	0	0	
1991 to 2000	1,539,655	1,300	975	0	0	
2001 to 2011	2,154,985	3,065	2,280	0	0	
2001 to 2005	992,675	845	580	0	0	
2006 to 2011	1,162,915	2,220	1,695	0	0	
Non-permanent residents	356,285	1,705	1,420	0	0	
Males in private households by immigrant status and period of immigration	16,163,115	247,960	125,560	30	30	
Non-immigrants	12,753,235	242,175	121,270	25	25	
Immigrants	3,231,365	4,720	3,385	0	0	
Before 1971	605,430	1,140	730	0	0	
1971 to 1980	416,675	735	520	0	0	
1981 to 1990	454,670	640	465	0	0	
1991 to 2000	724,905	630	460	0	0	
2001 to 2011	1,029,790	1,575	1,205	0	0	
2001 to 2005	474,540	435	310	0	0	
2006 to 2011	555,250	1,140	865	0	0	

For Help, press F1

(1) 2620/2620

Total population in private household

ENG

- Select variables
- Reset to only display DAUID code

Beyond 20/20 Professional Browser - 2011 NHS Semi-custom Profile for Canada, Provinces, Territories, Census Divisions, Census Subdivisions and Dissemination Areas

File Edit View Dimension Item Window Help

2011 NHS Semi-custom Profile for Canada, Provinces, Territories, Census Divisions, Census Subdivisions and Dissemination Areas, National Household Survey, 2011 (all GHs) (Read-only)

Geography	2011 NHS Total population in private h.	Canadian citizens	Canadian citizens aged 15 and over	Canadian citizens aged 15 and over, not in private h.	Not Canadian citizens	Males in private household	Canadian citizens
01	32 852 325	30 895 310	6 576 425	24 318 885	1 957 015	16 163 115	15 232 5
10	507 270	502 750	93 100	409 650	4 620	247 555	245 4
1001	258 470	254 940	47 995	206 945	3 530	125 565	123 5
1001101	55	55	0	55	0	30	
10010732	55	55	0	55	0	30	
1001105	160	160	20	140	0	85	
10010734	160	160	15	140	0	90	
1001113	545	545	75	475	0	290	2
10010735	310	310	40	270	0	185	1
10010736	235	240	30	205	0	110	1
1001120	55	55	0	50	0	25	
10010733	55	55	0	45	0	25	
1001124	1 490	1 490	240	1 250	0	675	6

Census Profile - Age, Sex, Marital Status, Families, Households and Dwellings for Canada, Provinces, Territories, Census Divisions, Census Subdivisions and Dissemination Areas, 2011

Sex (3) Total - Sex

Geography	Profile of Dis	Population, 2011	Median age of the population	Average number of children at
3520020	462 00	39 00	1 30	
35200321	512 00	44 00	1 20	
35200285	575 00	41 20	1 40	
35200293	502 00	26 30	2 00	
35200294	614 00	30 30	1 60	
35200297	529 00	37 40	1 30	
35200298	700 00	37 80	1 20	
35200299	643 00	32 60	1 30	
35200300	754 00	37 40	1 20	
35200301	910 00	36 10	1 30	
35200302	586 00	28 50	1 80	
35200304	529 00	31 90	1 60	
35200305	619 00	33 60	1 70	

For Help, press F1

59408/59408 10010732 (12.0%) ENG

Here, I have the census profile and the NHS profile both on the same screen – the census has my preselected DAs, the NHS has all levels of geography

- Apply the profile you saved before ...

Beyond 20/20 Professional Browser - 2011 NHS Semi-custom Profile for Canada, Provinces, Territories, Census Divisions, C...

File Edit View Dimension Item Window Help

Summary

Sort... Search...

Reverse Show All Reset Drill Down Drill Up Drill Back Change Labels Load Profile... Save Profile...

2011 NHS	Total population in private h...	Canadian citizens	Canadian citizens aged...	Canad citizens ag
01	32,852,325	30,895,310	6,676,425	24,31
10	507,270	502,750	93,100	40
1001	258,470	254,940	47,995	20
1001101	55	55	0	
10010732	55	55	0	
1001105	160	160	20	
10010734	160	160	15	
1001113	545	545	75	
10010735	310	310	40	
10010736	235	240	30	

Census Profile - Age, Sex, Marital Status, Families, Households and Dwellings for Canada, Provinces, Territories, Census Divis...

Sex (3) Total - Sex

Profile of Diss	Population, 2011	Median age of the popula...	Average number of children at...
35280230	434.00	43.80	1.10
35280231	484.00	40.90	1.00
35290285	575.00	41.20	1.40
35320293	502.00	26.90	2.00
35320294	614.00	30.30	1.60
35320297	529.00	37.40	1.30
35320298	700.00	37.80	1.20
35320299	843.00	32.60	1.30
35320300	754.00	37.40	1.20
35320301	910.00	35.10	1.30

Load a profile for the active dimension 59408/59408 Canada (26.1%)

- TADA

2011 NHS Semi-custom Profile for Canada, Provinces, Territories, Census Divisions, Census Subdivisions and Dissemination Ar...

Geography	2011 NHS Total population in private h.	Canadian citizens	Canadian citizens aged...	Canadian citizens ag
01	32,852,325	30,895,310	6,676,425	24,31
10	507,270	502,750	93,100	40
1001	258,470	254,940	47,995	20
1001101				0
10010732				0
1001105				20
10010734				15
1001113				75
10010735				40
10010736				30
1001111				0

Load Profile

Dimension: Geography

Profile Name Description

N01120-2011DAs

Show Select Cancel Help... Delete Profile

Include group definitions

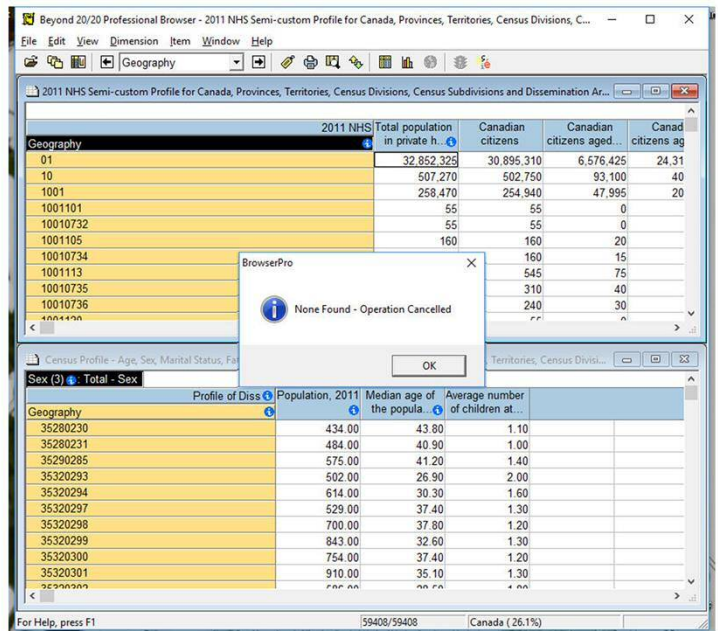
Geography	35280230	35280231	35290285	35320293	35320294	35320297	35320298	35320299	35320300	35320301
	484.00	40.90	1.00							
	575.00	41.20	1.40							
	502.00	26.90	2.00							
	614.00	30.30	1.60							
	529.00	37.40	1.30							
	700.00	37.80	1.20							
	843.00	32.60	1.30							
	754.00	37.40	1.20							
	910.00	35.10	1.30							

For Help, press F1

59408/59408 Canada (26.1%)

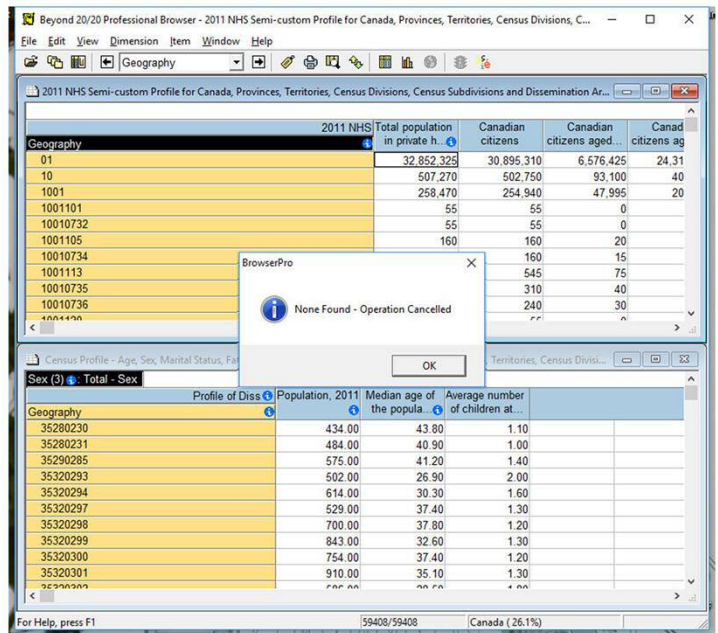
It's all so simple ...

- Well, not so much ... or should that be #@\$%@#!



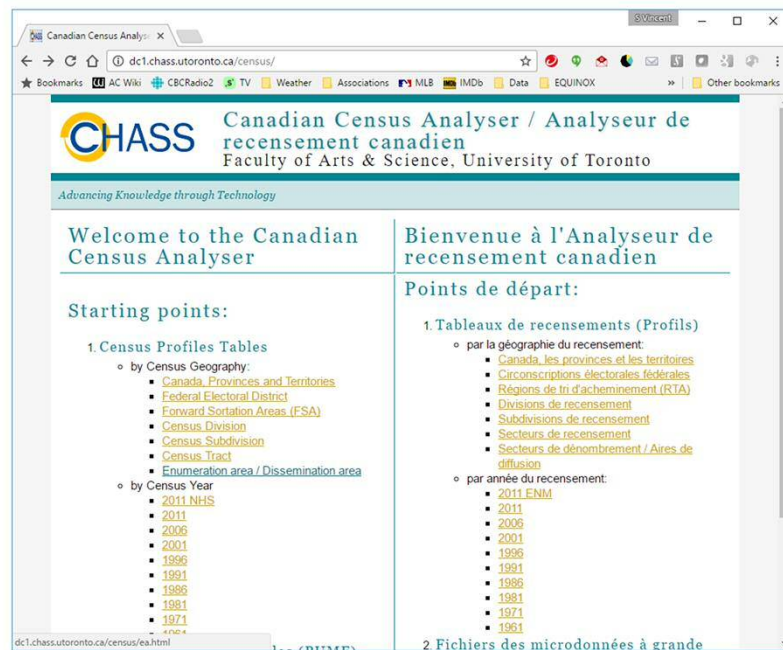
For whatever reason, NHS B20/20 Geography isn't compatible with Census B20/20 Geography

- So, you have to manually select the DAs all over again

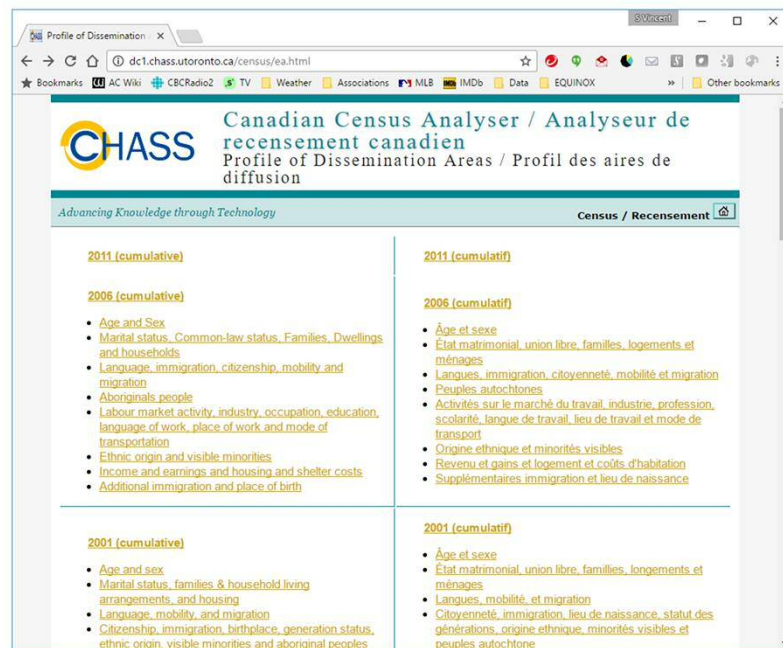


So, how does CHASS' Census Analyzer fare?

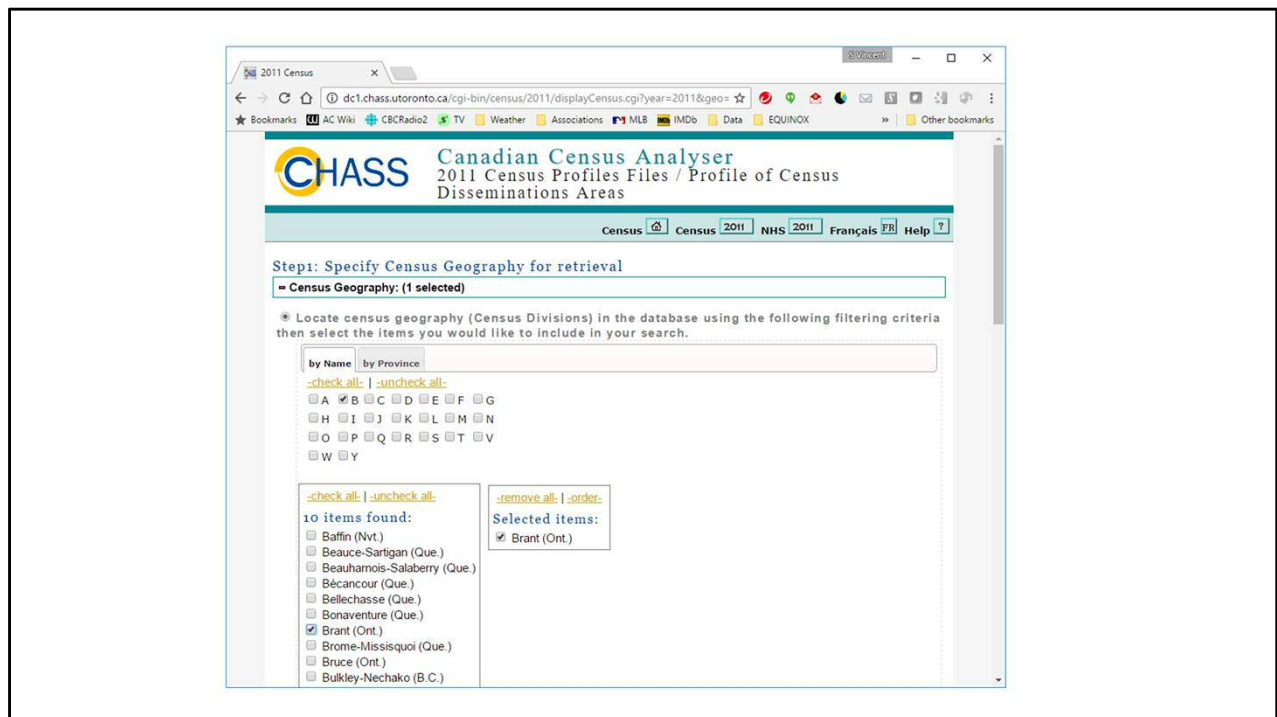
Remember, it's a subscription-based service, so you may NOT have access to it



I want to retrieve data by DA (or EA)

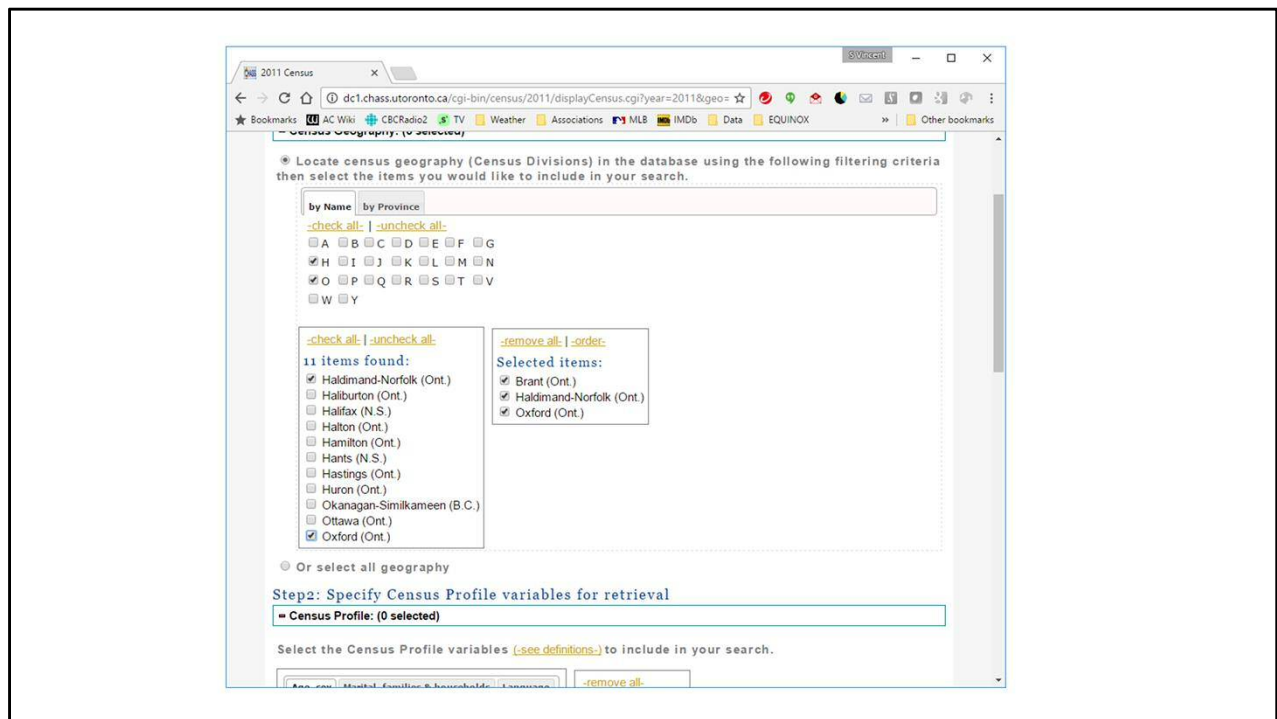


So, I'll pick 2011

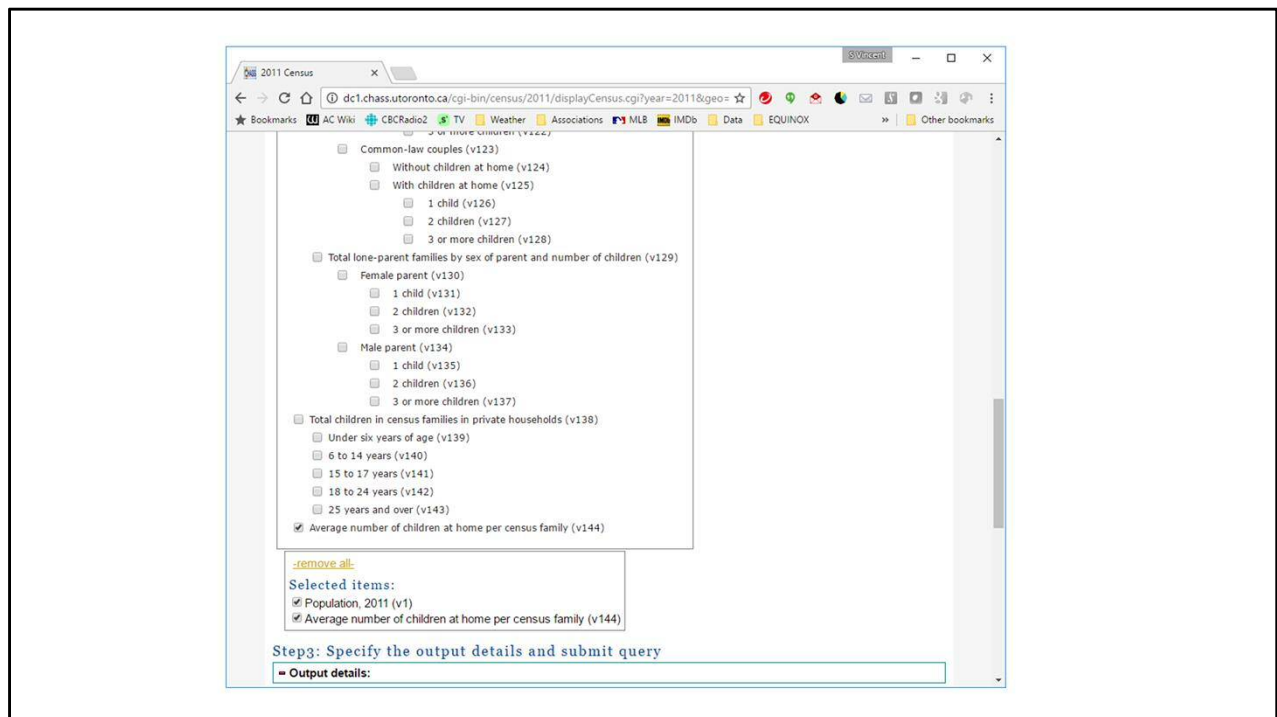


Notice at the top I have the choice of Census or NHS – both are loaded into CHASS

I can pick by Census Division – here, I’m not going to bother to go down to specific DAs, but grab all DAs within the area



I know I want Brant, Haldimand-Norfolk and Oxford.



Now, to pick my variables – there are different tabs that I can select which list specific variable groups. Bottom line, here I'll pull 2 variables (ignoring median age, which I grabbed in B2020).

A screenshot of a web browser window showing the '2011 Census' form at the URL `dc1.chass.utoronto.ca/cgi-bin/census/2011/displayCensus.cgi?year=2011&geo=`. The form includes several sections: a 'DA name' section with a text input field and a note about the format; a 'Census variables to be listed as' section with radio buttons for 'columns' and 'rows'; an email address input field for notifications; a 'zip' file compression option; and an output format selection dropdown menu. The dropdown menu is open, showing options under 'Screen output' (Text, HTML, Comma-Separated Values (CSV) for spreadsheet, MS Excel ready, SAS, SPSS) and 'Download to a file' (Comma-Separated Values (CSV) file for spreadsheet, dBase (DBF) file). A 'Submit Query' button is at the bottom, with a note that it will open a new tab. Footer links include 'CENSUS Home', 'Data Centre', 'CHASS Home', and 'Email Us', with a copyright notice for 2014 CHASS.

2011 Census

dc1.chass.utoronto.ca/cgi-bin/census/2011/displayCensus.cgi?year=2011&geo=

DA name

Optionally filter Dissemination area names (the format of DA name is xxxx);
(e.g. 0 1-4 7-9 0059 302)

Census variables to be listed as: (apply only to Screen output format)

columns: rows:

Optionally provide an email address for larger downloads notifications: (apply only to Download to a file output format)

Optionally enable zip file compression: (apply only to Download to a file output format)

none: zip:

Select the output format:

Screen output

- Text
- HTML
- Comma-Separated Values (CSV) for spreadsheet
- MS Excel ready
- SAS
- SPSS

Download to a file

- Comma-Separated Values (CSV) file for spreadsheet
- dBase (DBF) file

Submit Query. This will open a new tab or window in the browser.

Submit Query

[CENSUS Home](#) | [Data Centre](#) | [CHASS Home](#) | [Email Us](#)

Copyright © 2014 CHASS. All rights reserved.

Specify my output format – all I want is the DA code, so I won't add province, DA name, etc. I can have either a CSV or Excel ready file

```
"COL0","GEO UID"
"COL1","Population and dwelling counts / Population, 2011"
"COL2","Family characteristics / Average number of children at home per census family"
"COL0","COL1","COL2"
3528,109118.00,1.00
35280151,586.00,1.10
35280152,581.00,1.10
35280153,556.00,1.10
35280154,577.00,0.90
35280155,555.00,1.10
35280156,543.00,1.10
35280157,509.00,1.00
35280158,667.00,1.10
35280159,475.00,1.30
35280160,765.00,1.20
35280161,510.00,1.00
35280162,779.00,1.40
35280163,946.00,
35280164,637.00,1.00
35280165,121.00,
35280166,488.00,0.90
35280167,375.00,0.90
35280168,489.00,1.00
35280169,804.00,0.90
35280170,709.00,1.10
35280171,444.00,1.30
35280172,672.00,1.20
35280173,512.00,1.20
35280174,459.00,1.20
35280175,625.00,0.90
35280176,449.00,1.10
35280177,633.00,0.90
35280178,461.00,1.10
35280179,482.00,1.00
35280180,547.00,0.80
35280181,495.00,0.80
35280182,502.00,1.00
35280183,575.00,0.90
35280184,531.00,0.70
35280185,526.00,1.00
35280186,449.00,1.00
35280187,541.00,0.90
35280188,551.00,0.90
```

CSV for spreadsheet – it's ok, but I'm going to have to do more work with it, so I'll grab the excel-ready format

	GEO UID	COL1	COL2
3528	109118	1	
35280151	586	1.1	
35280152	581	1.1	
35280153	556	1.1	
35280154	577	0.9	
35280155	555	1.1	
35280156	543	1.1	
35280157	509	1	
35280158	667	1.1	
35280159	475	1.3	
35280160	765	1.2	
35280161	810	1	
35280162	779	1.4	
35280163	946		
35280164	637	1	
35280165	121		
35280166	488	0.9	
35280167	375	0.9	
35280168	489	1	
35280169	804	0.9	
35280170	709	1.1	
35280171	444	1.3	
35280172	672	1.2	
35280173	512	1.2	

So, I have my info, but I have to change it to read into ArcGIS

The image displays two side-by-side screenshots of a Microsoft Excel spreadsheet, illustrating the process of cleaning and reformatting census data for use in ArcGIS.

Left Screenshot (Original Data):

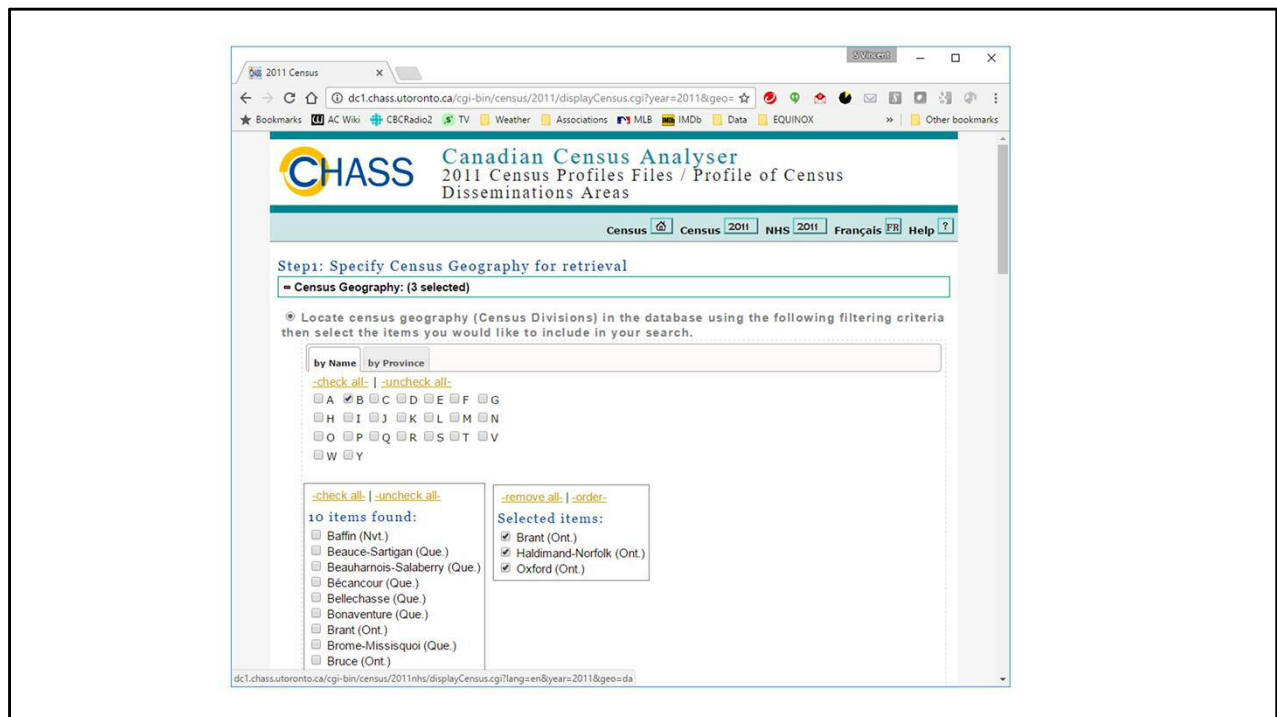
- Worksheet Name:** 2011 Census Profiles Files / Profile of Census Disseminations Areas
- Columns:** A1, B1, C1, D1, E1, F1, G1, H1, I1, J1, K1, L1
- Row 1:** 2011 Census Profiles Files / Profile of Census Disseminations Areas
- Row 2:** COLD GEO UID
- Row 3:** COL1 Population and dwelling counts / Population, 2011
- Row 4:** COL2 Family characteristics / Average number of children at home per census family
- Row 5:** COLD COL1 COL2
- Row 6:** 3528 109118 1
- Row 7:** 35280151 586 1.1
- Row 8:** 35280152 581 1.1
- Row 9:** 35280153 556 1.1
- Row 10:** 35280154 577 0.9
- Row 11:** 35280155 555 1.1
- Row 12:** 35280156 543 1.1
- Row 13:** 35280157 509 1
- Row 14:** 35280158 667 1.1
- Row 15:** 35280159 475 1.3
- Row 16:** 35280160 765 1.2
- Row 17:** 35280161 810 1
- Row 18:** 35280162 779 1.4
- Row 19:** 35280163 946
- Row 20:** 35280164 637 1
- Row 21:** 35280165 121
- Row 22:** 35280166 488 0.9
- Row 23:** 35280167 375 0.9
- Row 24:** 35280168 405 1
- Row 25:** 35280169 804 0.9
- Row 26:** 35280170 709 1.1
- Row 27:** 35280171 444 1.3
- Row 28:** 35280172 672 1.2
- Row 29:** 35280173 512 1.2
- Row 30:** 35280174 459 1.2
- Row 31:** 35280175 625 0.9
- Row 32:** 35280176 449 1.1
- Row 33:** 35280177 633 0.9
- Row 34:** 35280178 461 1.1
- Row 35:** 35280179 482 1
- Row 36:** 35280180 547 0.8

Right Screenshot (Cleaned Data):

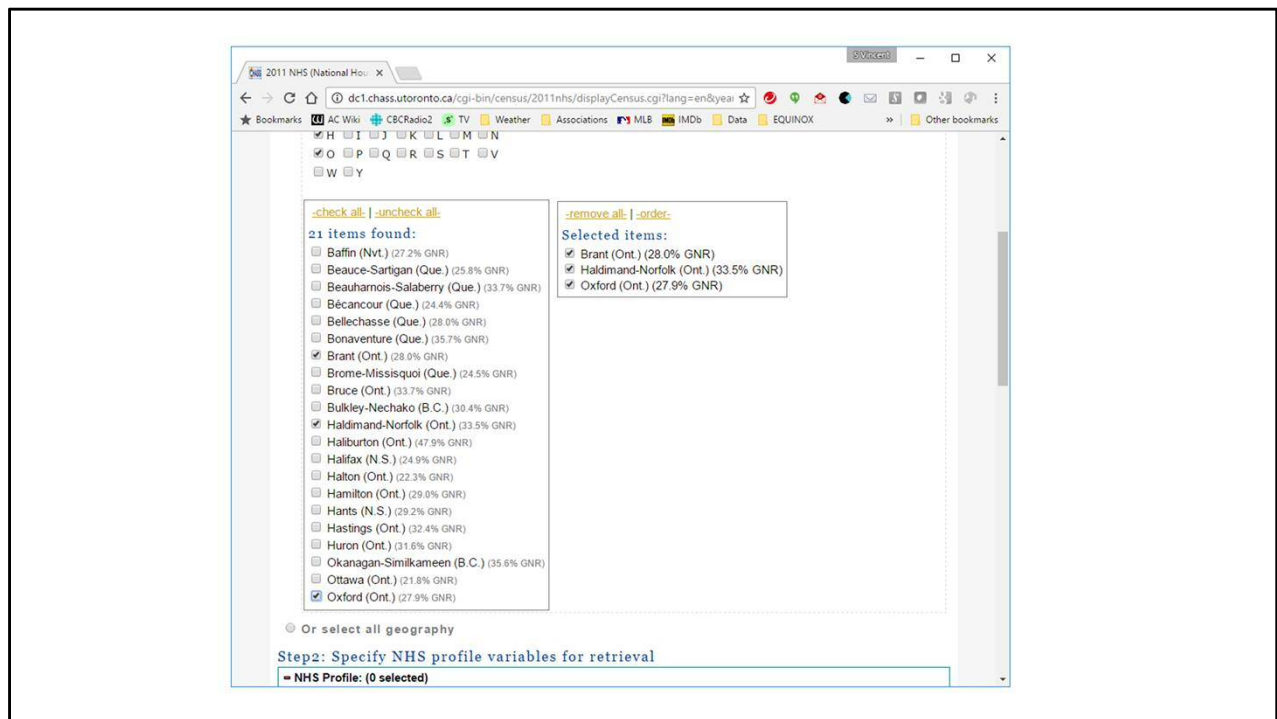
- Worksheet Name:** DAUID
- Columns:** A1, B1, C1, D1, E1, F1, G1, H1, I1, J1, K1, L1
- Row 1:** DAUID Popn2011 KidsprFam
- Row 2:** 3528 109118 1
- Row 3:** 35280151 586 1.1
- Row 4:** 35280152 581 1.1
- Row 5:** 35280153 556 1.1
- Row 6:** 35280154 577 0.9
- Row 7:** 35280155 555 1.1
- Row 8:** 35280156 543 1.1
- Row 9:** 35280157 509 1
- Row 10:** 35280158 667 1.1
- Row 11:** 35280159 475 1.3
- Row 12:** 35280160 765 1.2
- Row 13:** 35280161 810 1
- Row 14:** 35280162 779 1.4
- Row 15:** 35280163 946
- Row 16:** 35280164 637 1
- Row 17:** 35280165 121
- Row 18:** 35280166 488 0.9
- Row 19:** 35280167 375 0.9
- Row 20:** 35280168 405 1
- Row 21:** 35280169 804 0.9
- Row 22:** 35280170 709 1.1
- Row 23:** 35280171 444 1.3
- Row 24:** 35280172 672 1.2
- Row 25:** 35280173 512 1.2
- Row 26:** 35280174 459 1.2
- Row 27:** 35280175 625 0.9
- Row 28:** 35280176 449 1.1
- Row 29:** 35280177 633 0.9
- Row 30:** 35280178 461 1.1
- Row 31:** 35280179 482 1
- Row 32:** 35280180 547 0.8

Got to delete the first few rows, then rename the columns, and strip the final blank row. Finally, to work with ArcGIS, the DAUID must be reformatted as text. (and Excel WILL complain that you are treating numbers as text).

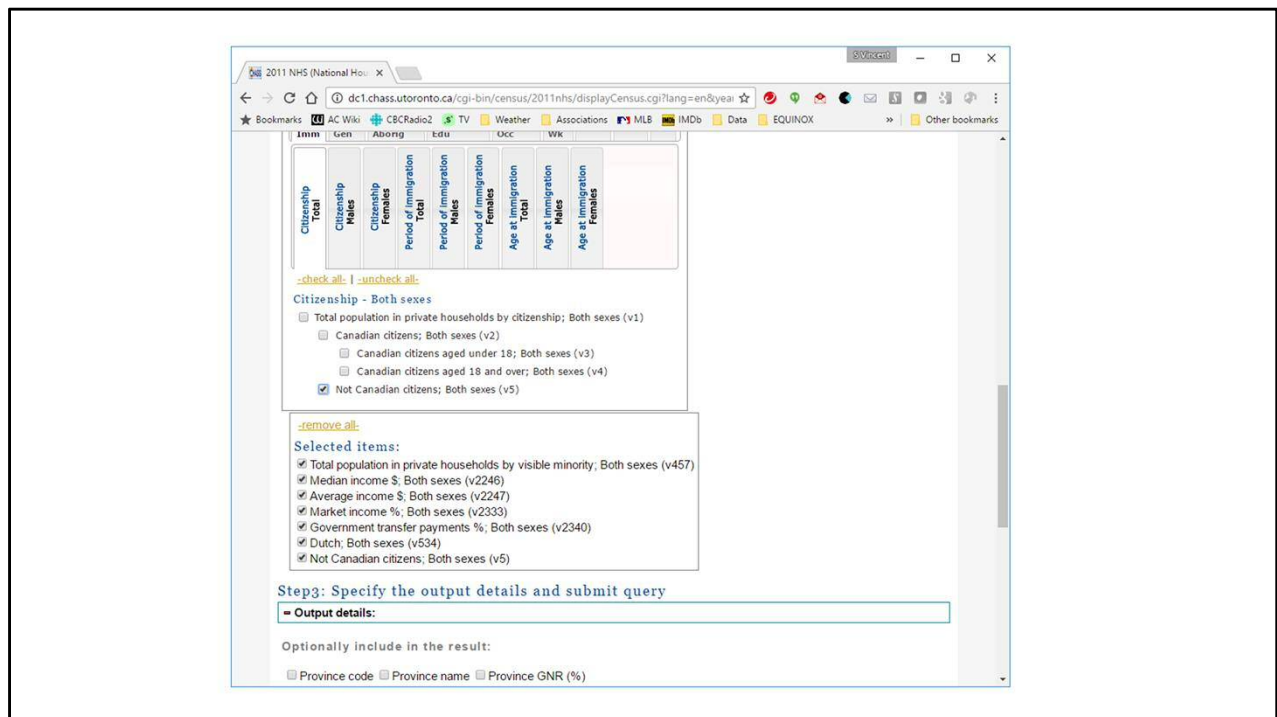
Repeat for NHS data



Now, I've selected NHS, though it's not apparent



Now it becomes evident I have NHS – down at the bottom!



Note the tabs with variables listed under them ... I've gone through my variable selections

2011 NHS (National Household Survey) - displayCensus.cgi?lang=en&year=2011

☐ DA name ☐ DA GNR (%)

Optionally filter Dissemination area names (the format of DA name is xxxx):
(e.g. 0 1-4 7-9 0059 302)

NHS variables to be listed as: (apply only to Screen output format)

columns: ☒ rows: ☐

Optionally provide an email address for larger downloads notifications: (apply only to Download to a file output format)

Optionally enable zip file compression: (apply only to Download to a file output format)

none: ☒ zip: ☐

Select the output format:

Screen output

- Text
- HTML
- Comma-Separated Values (CSV) for spreadsheet
- MS Excel ready**
- SAS
- SPSS

Download to a file

- Comma-Separated Values (CSV) file for spreadsheet
- dBase (DBF) file

Submit Query. This will open a new tab or window in the browser.

census (2).csv Show all

Once again, I've chosen the Excel ready file

[illegible]

53

So, which do you prefer?

- Selecting variables at the DA level is an absolute pain in B2020 (*CTRL-Click can be a shortcut to frustration*)
- It's somewhat less painful in CHASS Census Analyzer ... especially if you cheat and grab an entire census division at a time
 - ArcGIS won't care if you feed it records it doesn't have a match for
 - SPSS (or other stats packages) don't care either if records aren't used to match to postal codes

Finally, we have data to match to
postal geography

Look at the possible data that could be selected

	A	B	C	D	E	F	G	H	I	J	K	L	M	N	O	P	Q	R	S
	DAUID	NHSNonR	TtlPopn	PctNonCd	ImmPctDt	VisMin	PopPctDu	PctOthChi	Participati	EmplRate	UnempRt	LFPctAgri	IncIndMex	IncIndAvg	LowIncPre	IncHldM	IncHldAv	IncFamMe	IncFamAv
1	35280320	28.0	415	0.0	0.0	0	31.3	36.0	77.2	68.4	0.0	11.4	40745	50412	2.4	86862	90965	123022	115898
2	35280321	37.5	495	0.0	0.0	0	10.1	6.8	70.1	60.9	13.1	29.5	30691	37285	13.1	70404	92408	76252	92355
3	35290285	25.6	560	26.8	18.2	145	12.5	9.3	81.0	78.0	3.7	50.6	18376	29929	7.1	62947	88036	54615	73168
4	35320293	27.6	435	10.3	76.9	0	28.7	58.6	81.0	81.0	0.0	23.1	18345	27889	6.9	46865	62831	46964	65201
5	35320294	11.2	615	4.9	55.6	0	26.8	37.1	61.6	59.3	3.8	9.4	33171	39468	5.7	69891	80679	79234	87474
6	35320297	54.5	530	0.0	0.0	0	5.7	15.4	74.7	69.9	0.0	27.4	33332	37686	10.4	77595	82301	77398	80343
7	35320298	49.0	695	0.0	53.3	0	35.3	60.0	74.1	61.6	15.7	7.2	22048	31053	11.4	61800	62609	64456	72702
8	35320299	37.3	825	3.6	71.4	0	37.0	32.4	78.1	76.6	0.0	10.0	29938	36385	6.1	79953	72925	91903	91669
9	35320300	37.6	675	3.0	0.0	0	17.0	16.9	53.8	47.9	12.7	6.3	24589	30342	20.0	51819	58609	51966	63229
10	35320301	36.0	915	0.0	38.9	0	32.8	40.7	69.2	68.5	0.0	6.7	29647	37709	6.6	64957	76689	71204	81558
11	35320302	44.0	585	0.0	0.0	0	33.3	68.1	78.0	78.0	0.0	44.4	18981	23383	28.2	59160	63912	67278	68535
12	35320304	35.2	530	7.5	70.6	0	24.5	21.1	65.3	58.7	0.0	40.8	23180	35853	45.3	46931	84126	51442	91665
13	35320305	38.7	615	0.0	81.8	0	56.1	50.0	76.8	76.8	0.0	24.2	24127	32493	8.9	64993	69235	80745	79485
14	35320306	42.3	610	5.7	52.4	0	45.1	53.9	69.9	67.7	3.1	35.4	29713	36019	5.7	65656	83737	65964	86464
15	Minimum	11.2	415	0.0	0.0	0	5.7	6.8	53.8	47.9	0.0	6.3	18345	23383	2.0	46865	58609	46964	63229
16	Maximum	54.5	915	26.8	81.8	145.0	56.1	68.1	81.0	81.0	15.7	50.6	40745	50412	45.0	86862	92408	123022	115898
17	Median	37.4	598	1.5	45.7	0	30.0	36.6	74.4	68.5	0.0	23.7	27118	35936	8.0	64975	78684	69241	80951
18	Average	36.0	607	4.4	37.1	10.0	28.3	36.2	72.2	68.1	3.7	23.3	26920	34708	13.0	64988	76362	71603	82125
19	Stan.Dev	10.7	138	7.3	32.8	39.0	13.8	20.1	7.8	9.5	5.7	15.3	6624	6412	12.0	11739	11370	19695	13819
20	95%low	14.6	331	0.0	0.0	0	0.0	0.0	56.5	49.0	0.0	0.0	13672	21883	0.0	41510	53621	32213	54487
21	95%high	57.4	884	19.0	100.0	88.0	55.9	76.4	87.9	87.2	15.1	53.9	40168	47532	37.0	88466	99102	110993	109762

After all, it's possible that we don't actually need to do GIS analysis – we want to test our sample to see if it's representative. Let's look at the data – how representative would ANY specific DA be of this area?

On NHS Non-response, we have values ranging from 11.2% to 54.5%. Visible minority – 13 of the 14 DAs have none, but one DA has 25% visible minority population. Unemployment rate – 0% in 8 DAs, up to 5% in 3, and over 12% in 3. Labour force engaged in agriculture ranges as a percentage from 6.3% to 50.6%. Median family income from 47,000 to 123,000, and average from 63,000 to 116,000.

I've indicated, using blue fill, the values that fall outside two standard deviations of the observed values – in other words that would be expected less than 5% of the time. Only six of the 14 possible DAs have no values that fall outside 2 standard deviations on these 18 variables.

Now, we have the second problem I have with the whole process - even when we find data, should we tell users to use it?

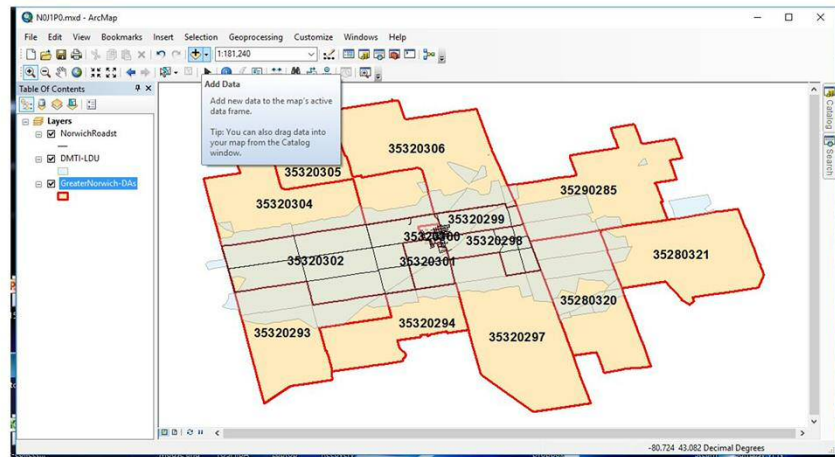
So, let's combine data with geography

- In ArcGIS
- In SPSS (or stats packages generally)

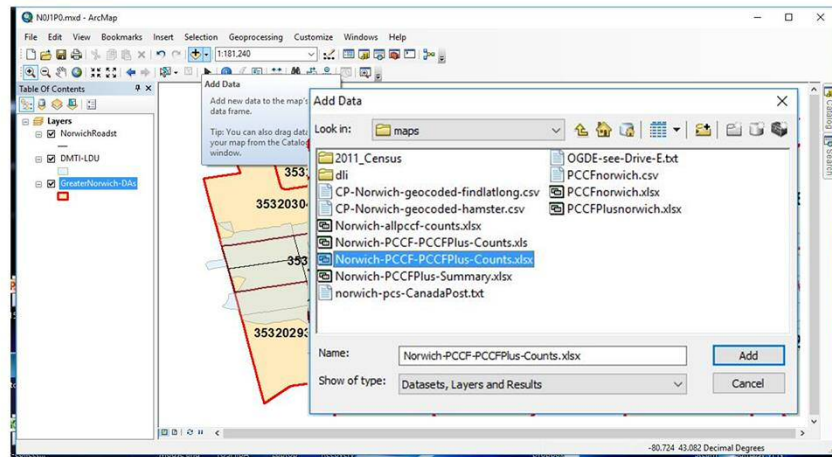
Review: merging data into ArcGIS

Many (most?) of you know far more about this than I do

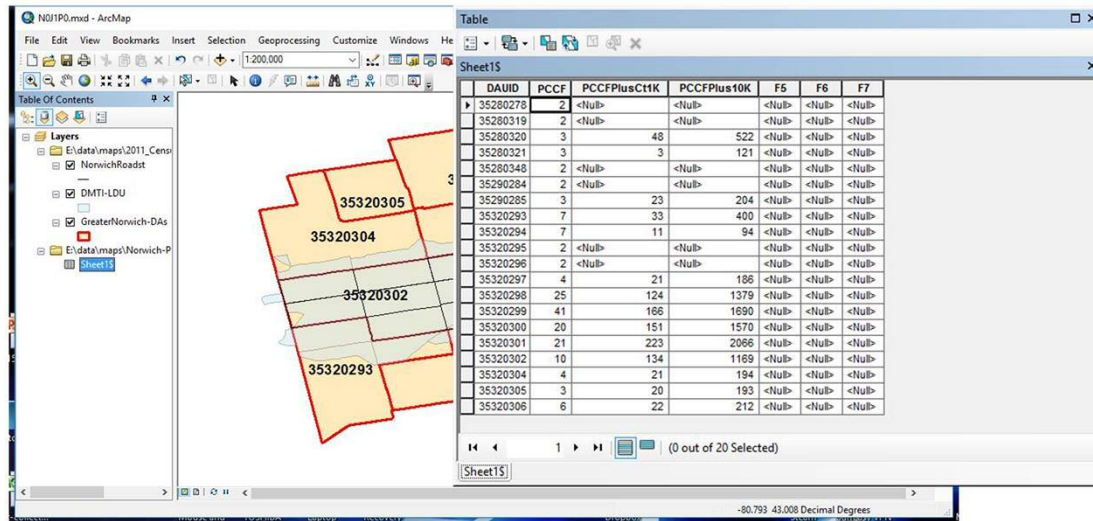
Load the map layers; add data



Load the map layers; add data



Data (counts of frequencies of DA selection) have been added



The screenshot shows the ArcMap interface with a map of DA selection counts and a data table. The map displays several DA selection areas with labels 35320305, 35320304, 35320302, and 35320293. The data table, titled 'Table', shows the following data:

DAUID	PCCF	PCCFPlusCt1K	PCCFPlus10K	F5	F6	F7
35280278	2	<Nul>	<Nul>	<Nul>	<Nul>	<Nul>
35280319	2	<Nul>	<Nul>	<Nul>	<Nul>	<Nul>
35280320	3	<Nul>	48	522	<Nul>	<Nul>
35280321	3	<Nul>	3	121	<Nul>	<Nul>
35280348	2	<Nul>	<Nul>	<Nul>	<Nul>	<Nul>
35290284	2	<Nul>	<Nul>	<Nul>	<Nul>	<Nul>
35290285	3	<Nul>	23	204	<Nul>	<Nul>
35320293	7	<Nul>	33	400	<Nul>	<Nul>
35320294	7	<Nul>	11	94	<Nul>	<Nul>
35320295	2	<Nul>	<Nul>	<Nul>	<Nul>	<Nul>
35320296	2	<Nul>	<Nul>	<Nul>	<Nul>	<Nul>
35320297	4	<Nul>	21	186	<Nul>	<Nul>
35320298	25	<Nul>	124	1379	<Nul>	<Nul>
35320299	41	<Nul>	166	1690	<Nul>	<Nul>
35320300	20	<Nul>	151	1570	<Nul>	<Nul>
35320301	21	<Nul>	223	2066	<Nul>	<Nul>
35320302	10	<Nul>	134	1169	<Nul>	<Nul>
35320304	4	<Nul>	21	194	<Nul>	<Nul>
35320305	3	<Nul>	20	193	<Nul>	<Nul>
35320306	6	<Nul>	22	212	<Nul>	<Nul>

Note in the dataset that there are DAs that do not have records in the PCCF+ - these are the six DAs that were found in the PCCF that weren't found in the PCCF+

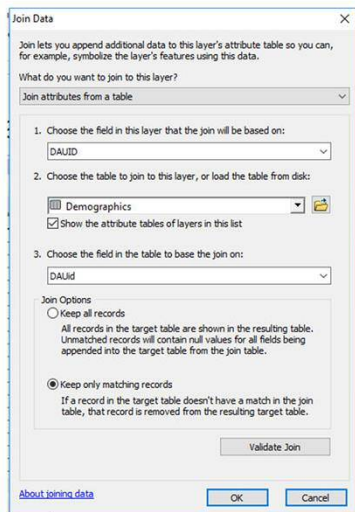
Repeat to add the demographic data

Table									
Demographics									
DAUID	NHSNonResp	TtlPopn	CitCanada	CitNonCan	ImmigTtl	ImmigDtch	VisMin	EthncDtch	Chr
35280320	28	415	415	0	40	0	0	130	
35280321	37.5	495	485	0	25	0	0	50	
35290285	25.6	560	410	150	55	10	145	70	
35320293	27.6	435	390	45	65	50	0	125	
35320294	11.2	615	580	30	90	50	0	165	
35320297	54.5	530	515	0	45	0	0	30	
35320298	49	695	690	0	75	40	0	245	
35320299	37.3	825	790	30	70	50	0	305	
35320300	37.6	675	655	20	20	0	0	115	
35320301	36	915	900	0	90	35	0	300	
35320302	44	585	570	0	80	0	0	195	
35320304	35.2	530	490	40	85	60	0	130	
35320305	38.7	615	605	0	55	45	0	345	
35320306	42.3	610	575	35	105	55	0	275	
Minimum	11.2	415	390	0	20	0	0.0	30	
Maximum	54.5	915	900	150	105	60	145.0	345	
Median	37.4	598	573	10	68	38	0.0	148	
Average	36	607	576	25	64	28	10.0	177	
Stan.Dev	10.7	138	146	40	25	25	39.0	102	
05561000	14.6	234	284	0	12	0	0.0	0	

1 (0 out of 21 Selected)

Demographics

Join data to geographic layer



The image shows a 'Join Data' dialog box with a title bar and a close button. The main text explains that joining appends data to a layer's attribute table. Below this, a dropdown menu is set to 'Join attributes from a table'. The dialog is divided into three numbered steps: 1. 'Choose the field in this layer that the join will be based on:' with a dropdown showing 'DAUID'. 2. 'Choose the table to join to this layer, or load the table from disk:' with a dropdown showing 'Demographics' and a checkbox for 'Show the attribute tables of layers in this list' which is checked. 3. 'Choose the field in the table to base the join on:' with a dropdown showing 'DAUID'. Below these steps is a 'Join Options' section with two radio buttons: 'Keep all records' (unselected) and 'Keep only matching records' (selected). Explanatory text for each option is provided. At the bottom, there is a 'Validate Join' button, a link to 'About joining data', and 'OK' and 'Cancel' buttons.

Join Data

Join lets you append additional data to this layer's attribute table so you can, for example, symbolize the layer's features using this data.

What do you want to join to this layer?

Join attributes from a table

1. Choose the field in this layer that the join will be based on:

DAUID

2. Choose the table to join to this layer, or load the table from disk:

Demographics

☒ Show the attribute tables of layers in this list

3. Choose the field in the table to base the join on:

DAUID

Join Options

☐ Keep all records

All records in the target table are shown in the resulting table. Unmatched records will contain null values for all fields being appended into the target table from the join table.

☒ Keep only matching records

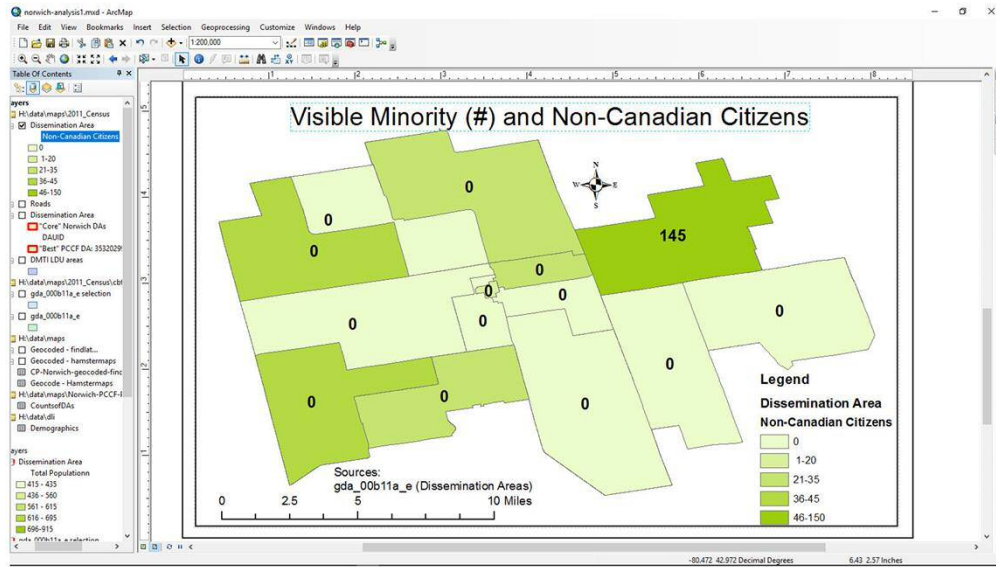
If a record in the target table doesn't have a match in the join table, that record is removed from the resulting target table.

Validate Join

[About joining data](#)

OK Cancel

Map data



Note that ArcGIS uses the labels it sees

Generated labels

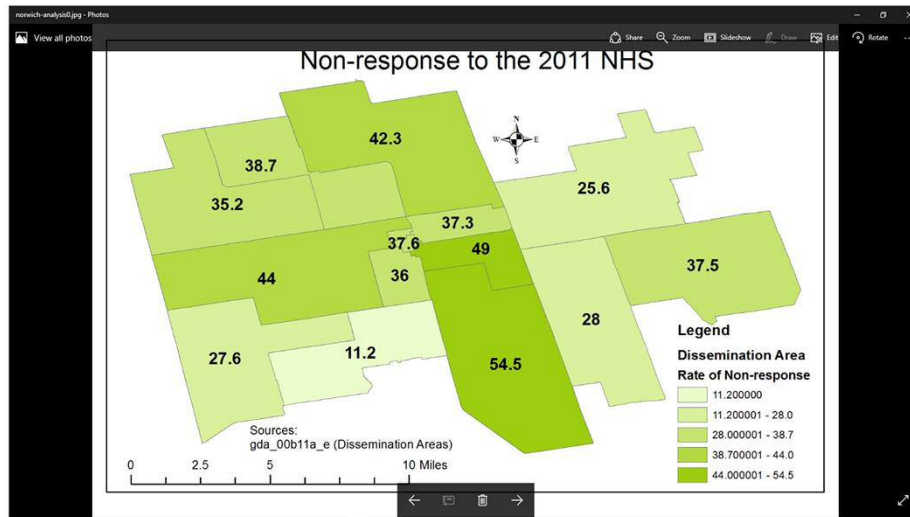
- 0
- 1 – 20
- 21 – 35
- 36 – 45
- 46 – 150

What the labels SHOULD be due to StatCan's rounding policies

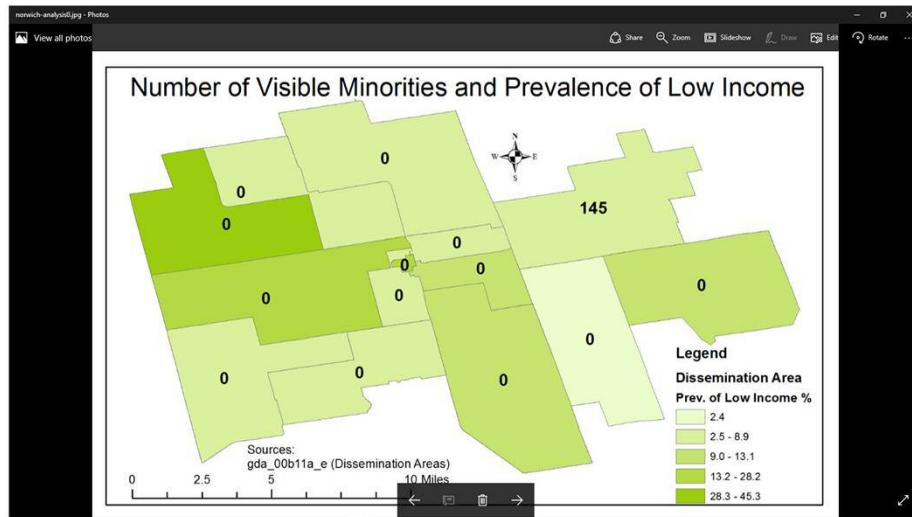
- 0 – 4
- 1 – 24
- 21 – 39
- 36 – 49
- 46 – 149

Do I have a volunteer to be the first person to explain this to a user?

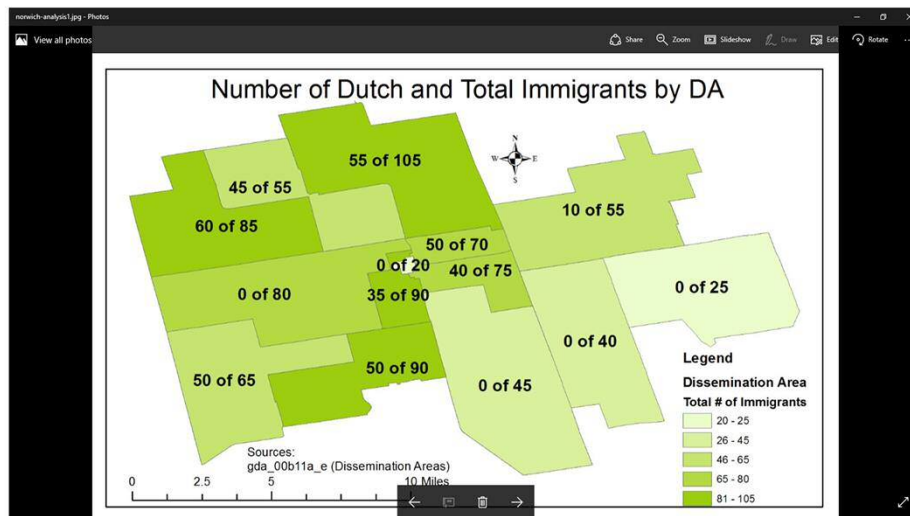
Does non-response matter at DA level?



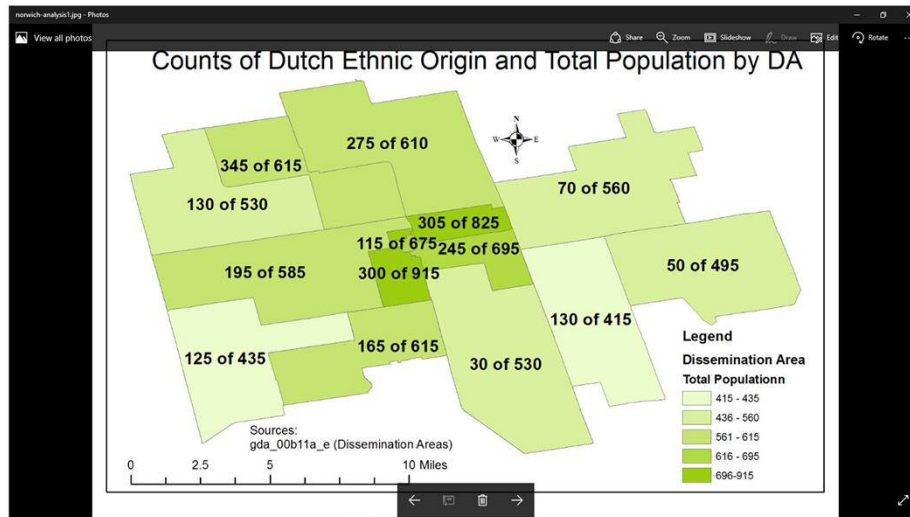
Is low income related to VisMin status?



Dutch immigrants and total immigrants



Location and preponderance of Dutch ethnic origin residents

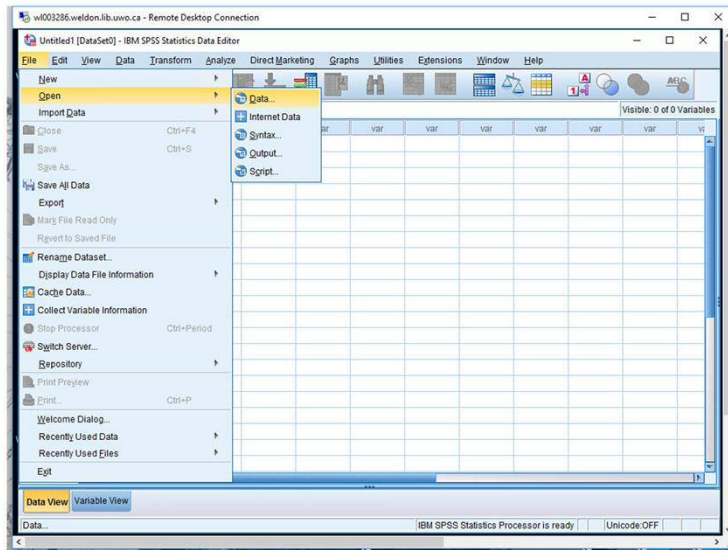


Let's look at SPSS

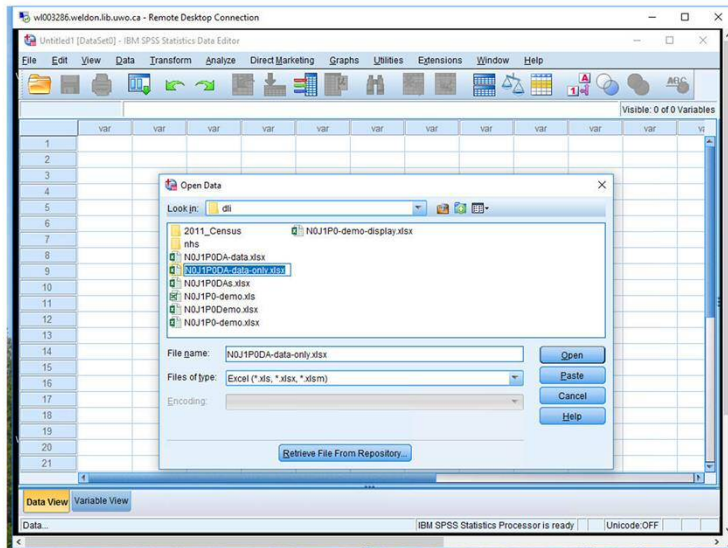
Again, I anticipate that merging data into or combining files in SPSS is well-known to you

Here, the interesting bit (well, interesting if you're as twisted as I am) is that we're going to assume that the PCCF Plus results are representative of the frequency with which we should retrieve the data – and therefore, we can use it to establish better benchmarks for the data.

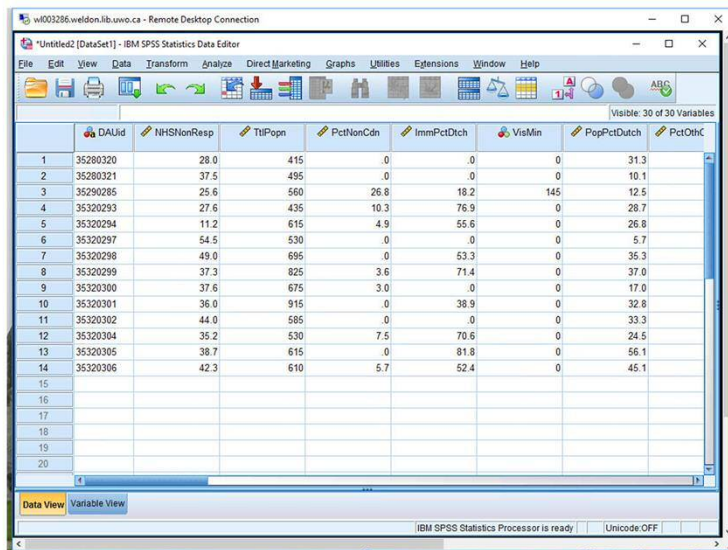
After opening SPSS, open either data or geography



I choose data



I choose data

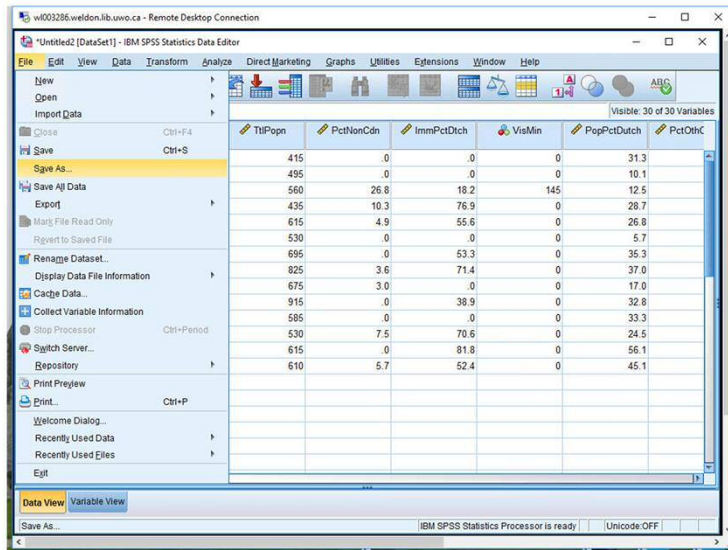


The screenshot shows the IBM SPSS Statistics Data Editor window. The menu bar includes File, Edit, View, Data, Transform, Analyze, Direct Marketing, Graphs, Utilities, Extensions, Window, and Help. The toolbar contains icons for file operations, data manipulation, and analysis. The variable list at the top shows 30 variables: DAUID, NHSNonResp, TstPopn, PctNonCdn, ImmPctDich, VisMin, PopPctDutch, and PctOthC. The data view shows 20 rows of data. The status bar at the bottom indicates 'IBM SPSS Statistics Processor is ready' and 'Unicode OFF'.

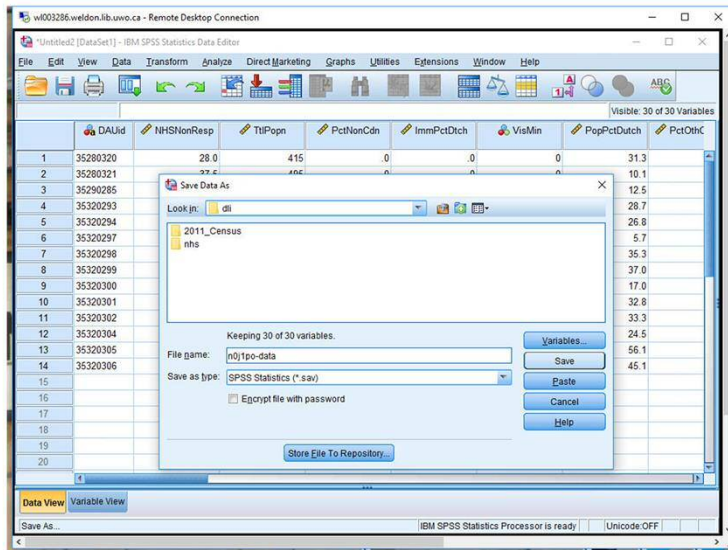
	DAUID	NHSNonResp	TstPopn	PctNonCdn	ImmPctDich	VisMin	PopPctDutch	PctOthC
1	35280320	28.0	415	0	.0	0	31.3	
2	35280321	37.5	495	0	.0	0	10.1	
3	35290285	25.6	560	26.8	18.2	145	12.5	
4	35320293	27.6	435	10.3	76.9	0	28.7	
5	35320294	11.2	615	4.9	55.6	0	26.8	
6	35320297	54.5	530	0	.0	0	5.7	
7	35320298	49.0	695	0	53.3	0	35.3	
8	35320299	37.3	825	3.6	71.4	0	37.0	
9	35320300	37.6	675	3.0	.0	0	17.0	
10	35320301	36.0	915	0	38.9	0	32.8	
11	35320302	44.0	585	0	.0	0	33.3	
12	35320304	35.2	530	7.5	70.6	0	24.5	
13	35320305	38.7	615	0	81.8	0	56.1	
14	35320306	42.3	610	5.7	52.4	0	45.1	
15								
16								
17								
18								
19								
20								

Notice, though – that the DA variable name is DAUID ... for SPSS, variable names you want to match on must be the same format and have the same name.

Save the file as a SPSS system file



Save the file as a SPSS system file

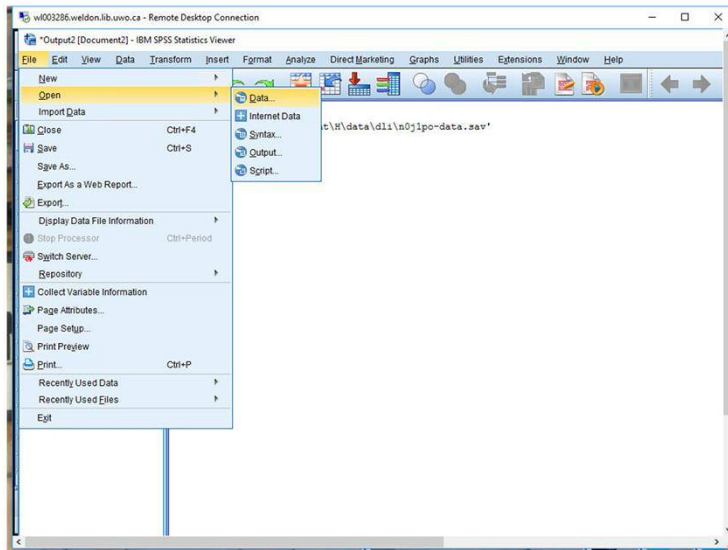


And, obviously, if you save the file name with an accurate name, it will help no 1po, but 1p0 ☹

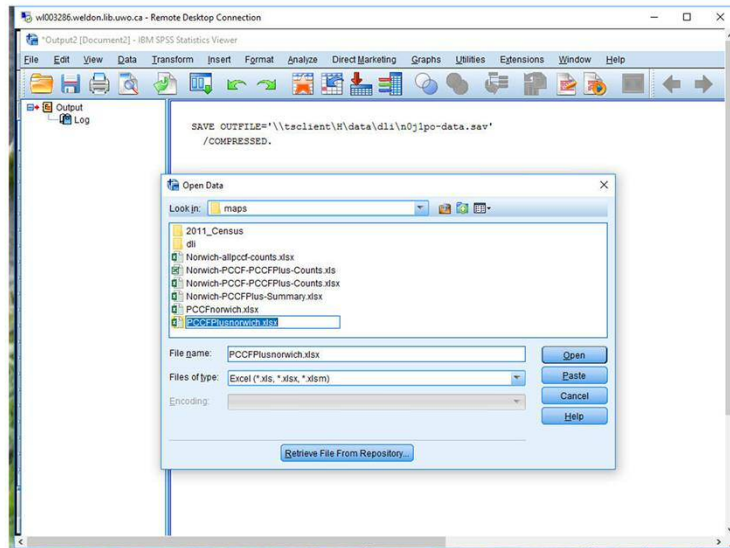
The 1000-record PCCF+ file (geography)

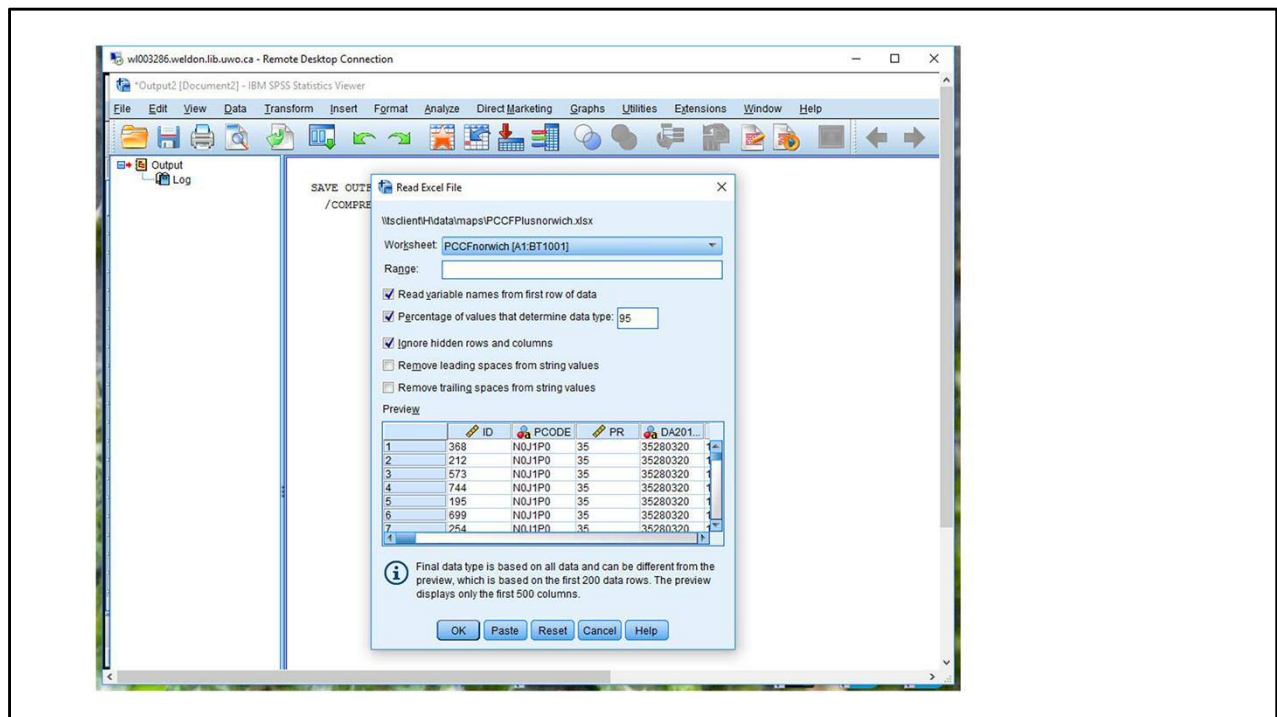
ID	PCODE	PR	DA2011uid	DB	DB_ir2011	CSuid	CSName	CMA	CMAtype	CMAName	CTName	Tracted	SACcode	SACType	CCSuid	FEDuid	FEDName	DPLuid
2	368	N011P0	35	35280320	1	F	3528052	Norfolk Ci	547 D	Norfolk	0	0	547	3	3528052	35028	Haldimani	35
3	212	N011P0	35	35280320	1	F	3528052	Norfolk Ci	547 D	Norfolk	0	0	547	3	3528052	35028	Haldimani	35
4	573	N011P0	35	35280320	1	F	3528052	Norfolk Ci	547 D	Norfolk	0	0	547	3	3528052	35028	Haldimani	35
5	744	N011P0	35	35280320	1	F	3528052	Norfolk Ci	547 D	Norfolk	0	0	547	3	3528052	35028	Haldimani	35
6	195	N011P0	35	35280320	1	F	3528052	Norfolk Ci	547 D	Norfolk	0	0	547	3	3528052	35028	Haldimani	35
7	699	N011P0	35	35280320	1	F	3528052	Norfolk Ci	547 D	Norfolk	0	0	547	3	3528052	35028	Haldimani	35
8	254	N011P0	35	35280320	1	F	3528052	Norfolk Ci	547 D	Norfolk	0	0	547	3	3528052	35028	Haldimani	35
9	500	N011P0	35	35280320	2	F	3528052	Norfolk Ci	547 D	Norfolk	0	0	547	3	3528052	35028	Haldimani	35
10	278	N011P0	35	35280320	2	F	3528052	Norfolk Ci	547 D	Norfolk	0	0	547	3	3528052	35028	Haldimani	35
11	496	N011P0	35	35280320	2	F	3528052	Norfolk Ci	547 D	Norfolk	0	0	547	3	3528052	35028	Haldimani	35
12	633	N011P0	35	35280320	2	F	3528052	Norfolk Ci	547 D	Norfolk	0	0	547	3	3528052	35028	Haldimani	35
13	80	N011P0	35	35280320	2	F	3528052	Norfolk Ci	547 D	Norfolk	0	0	547	3	3528052	35028	Haldimani	35
14	266	N011P0	35	35280320	3	F	3528052	Norfolk Ci	547 D	Norfolk	0	0	547	3	3528052	35028	Haldimani	35
15	244	N011P0	35	35280320	3	F	3528052	Norfolk Ci	547 D	Norfolk	0	0	547	3	3528052	35028	Haldimani	35
16	103	N011P0	35	35280320	4	F	3528052	Norfolk Ci	547 D	Norfolk	0	0	547	3	3528052	35028	Haldimani	35
17	624	N011P0	35	35280320	4	F	3528052	Norfolk Ci	547 D	Norfolk	0	0	547	3	3528052	35028	Haldimani	35
18	37	N011P0	35	35280320	4	F	3528052	Norfolk Ci	547 D	Norfolk	0	0	547	3	3528052	35028	Haldimani	35
19	334	N011P0	35	35280320	4	F	3528052	Norfolk Ci	547 D	Norfolk	0	0	547	3	3528052	35028	Haldimani	35
20	773	N011P0	35	35280320	4	F	3528052	Norfolk Ci	547 D	Norfolk	0	0	547	3	3528052	35028	Haldimani	35
21	701	N011P0	35	35280320	4	F	3528052	Norfolk Ci	547 D	Norfolk	0	0	547	3	3528052	35028	Haldimani	35
22	107	N011P0	35	35280320	4	F	3528052	Norfolk Ci	547 D	Norfolk	0	0	547	3	3528052	35028	Haldimani	35
23	568	N011P0	35	35280320	6	F	3528052	Norfolk Ci	547 D	Norfolk	0	0	547	3	3528052	35028	Haldimani	35

Open it



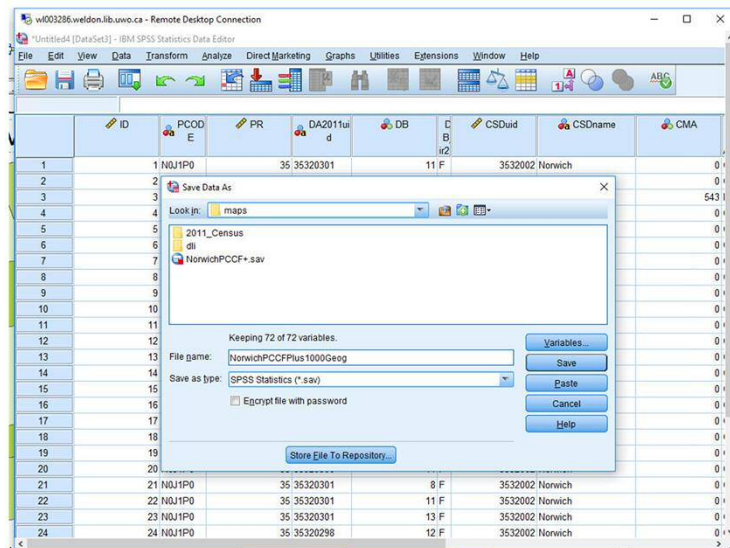
Locate the file



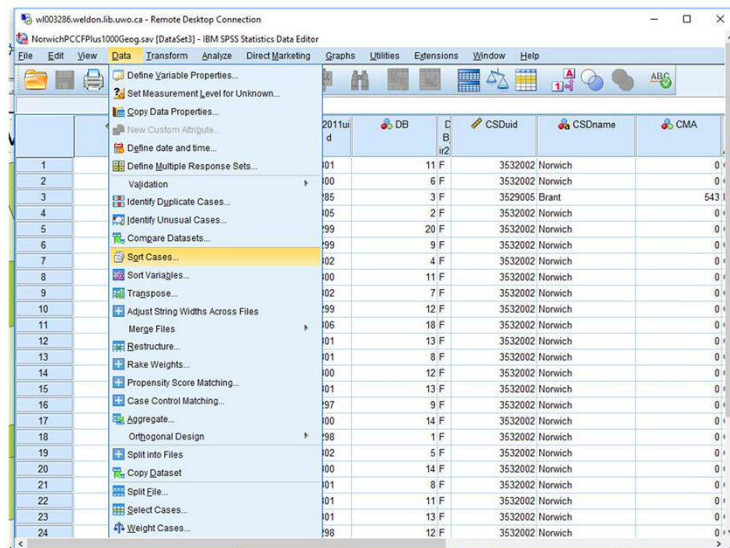


This is the intermediate step I forgot to show when I loaded the data ...

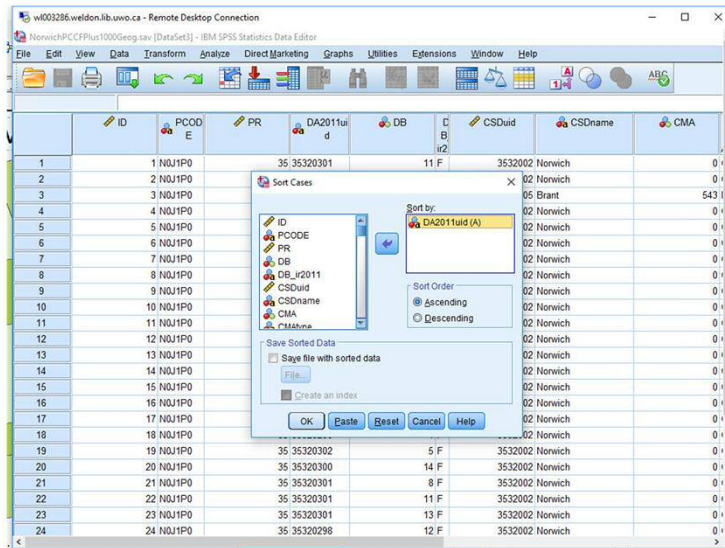
And then save the geography file



Be sure to sort both files by the DAUid

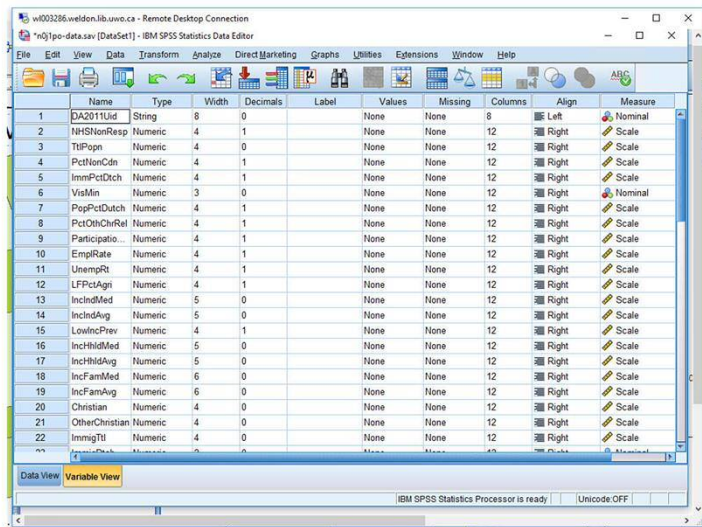


Here called DA2011Uid



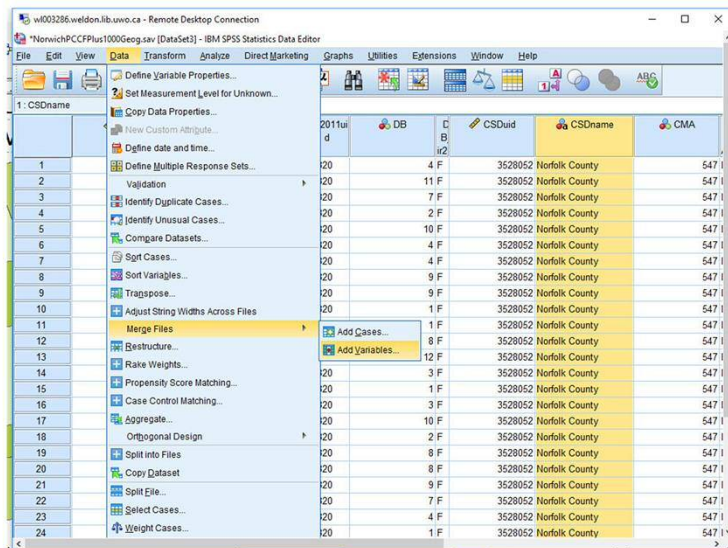
I'll skip some tedium – you also sort the data file, and have to rename either DAUid or DA2011UID so that they match. Once you've done that, you can merge the file and add the data variables to the geographic.

And be sure that both the ID names match

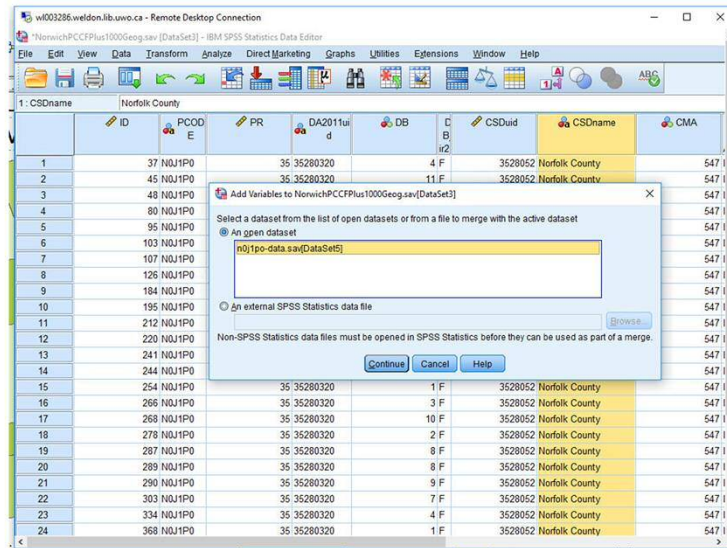


	Name	Type	Width	Decimals	Label	Values	Missing	Columns	Align	Measure
1	DA2011UId	String	8	0		None	None	8	Left	Nominal
2	NHSHNonResp	Numeric	4	1		None	None	12	Right	Scale
3	TtlPopn	Numeric	4	0		None	None	12	Right	Scale
4	PctNonCdn	Numeric	4	1		None	None	12	Right	Scale
5	ImmPctDch	Numeric	4	1		None	None	12	Right	Scale
6	VisMin	Numeric	3	0		None	None	12	Right	Nominal
7	PopPctDutch	Numeric	4	1		None	None	12	Right	Scale
8	PctOthChrRel	Numeric	4	1		None	None	12	Right	Scale
9	Participatio...	Numeric	4	1		None	None	12	Right	Scale
10	EnglRate	Numeric	4	1		None	None	12	Right	Scale
11	UnempRl	Numeric	4	1		None	None	12	Right	Scale
12	LFFctAgri	Numeric	4	1		None	None	12	Right	Scale
13	InclndMed	Numeric	5	0		None	None	12	Right	Scale
14	InclndAvg	Numeric	5	0		None	None	12	Right	Scale
15	LowIncPrev	Numeric	4	1		None	None	12	Right	Scale
16	InclHldMed	Numeric	5	0		None	None	12	Right	Scale
17	InclHldAvg	Numeric	5	0		None	None	12	Right	Scale
18	IncFamMed	Numeric	6	0		None	None	12	Right	Scale
19	IncFamAvg	Numeric	6	0		None	None	12	Right	Scale
20	Christian	Numeric	4	0		None	None	12	Right	Scale
21	OtherChristian	Numeric	4	0		None	None	12	Right	Scale
22	ImmigTtl	Numeric	4	0		None	None	12	Right	Scale

Then, you can merge the files

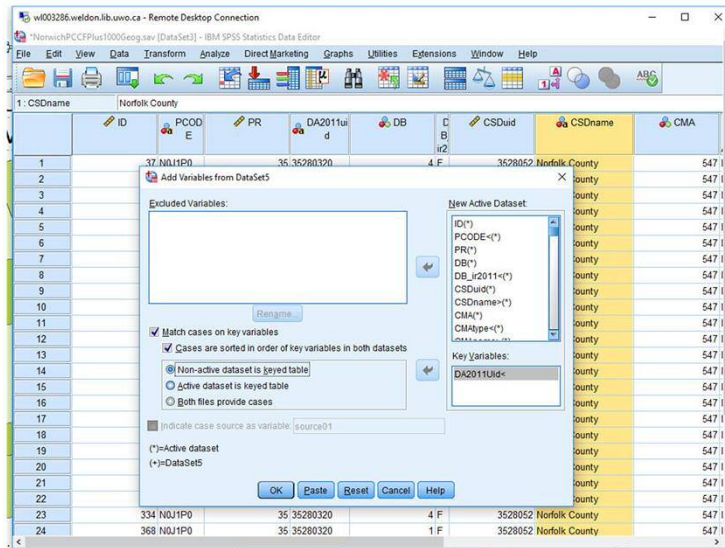


I'm going to merge the data with the geographic records



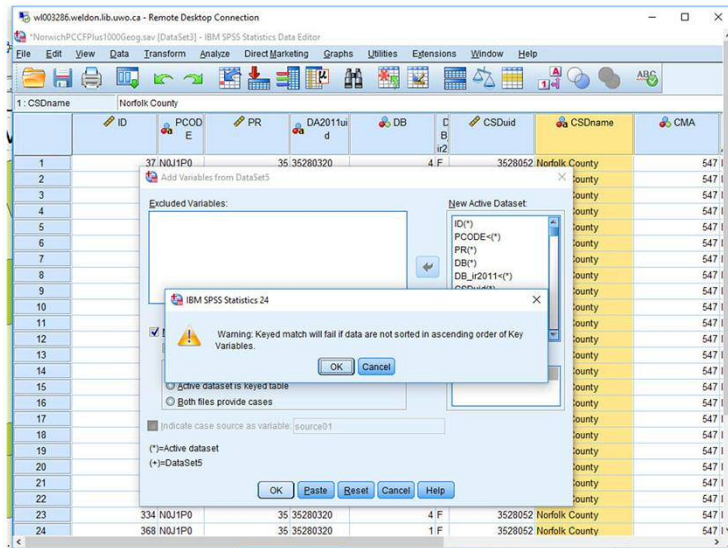
The files don't both have to be open – one can be on disk.

KEYED Table: can be used multiple times in merge

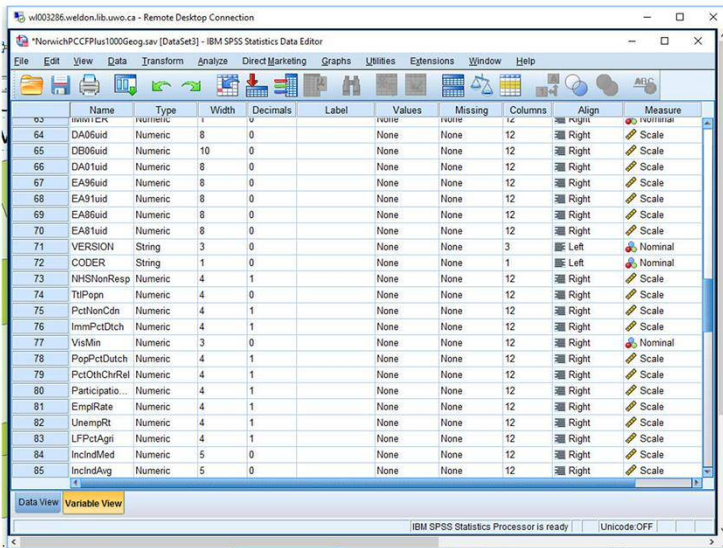


Note that the Key variable must match on variable name (and format)

And a reminder to sort files



TADA ...



IBM SPSS Statistics Data Editor - w003286.weldon.lib.uwo.ca - Remote Desktop Connection

NonichPCCFPlus1000Geog.sav [DataSet1] - IBM SPSS Statistics Data Editor

	Name	Type	Width	Decimals	Label	Values	Missing	Columns	Align	Measure
62	resid1CR	numeric	1	0		resid1	resid1	12	Right	Nominal
64	DA06uid	Numeric	8	0		None	None	12	Right	Scale
65	DB06uid	Numeric	10	0		None	None	12	Right	Scale
66	DA01uid	Numeric	8	0		None	None	12	Right	Scale
67	EA96uid	Numeric	8	0		None	None	12	Right	Scale
68	EA91uid	Numeric	8	0		None	None	12	Right	Scale
69	EA86uid	Numeric	8	0		None	None	12	Right	Scale
70	EA81uid	Numeric	8	0		None	None	12	Right	Scale
71	VERSION	String	3	0		None	None	3	Left	Nominal
72	CODER	String	1	0		None	None	1	Left	Nominal
73	NHSNonResp	Numeric	4	1		None	None	12	Right	Scale
74	TiPopn	Numeric	4	0		None	None	12	Right	Scale
75	PctNonCdn	Numeric	4	1		None	None	12	Right	Scale
76	ImmPotDich	Numeric	4	1		None	None	12	Right	Scale
77	VisMin	Numeric	3	0		None	None	12	Right	Nominal
78	PopPctDutch	Numeric	4	1		None	None	12	Right	Scale
79	PctOthChrRel	Numeric	4	1		None	None	12	Right	Scale
80	Participatio...	Numeric	4	1		None	None	12	Right	Scale
81	EmprRate	Numeric	4	1		None	None	12	Right	Scale
82	UnempRt	Numeric	4	1		None	None	12	Right	Scale
83	LFPctAgri	Numeric	4	1		None	None	12	Right	Scale
84	InclndMed	Numeric	5	0		None	None	12	Right	Scale
85	InclndAug	Numeric	5	0		None	None	12	Right	Scale

Data View Variable View

IBM SPSS Statistics Processor is ready | Unicode OFF

Look at descriptive statistics based on the 1,000 PCCF+ records – and now we can reexamine the data we deliver

The screenshot shows the SPSS 'Descriptives' window. The left sidebar lists 'Output', 'Log', 'Descriptives', 'Notes', and 'Descriptive Stats'. The main area is titled 'Descriptives' and contains a table of 'Descriptive Statistics' for various variables. Each variable has a sample size (N) of 1000, except for 'Valid N (listwise)' which is 1000. The table includes columns for N, Minimum, Maximum, Mean, and Std. Deviation.

	N	Minimum	Maximum	Mean	Std. Deviation
NHISNonResp	1000	11.2	54.5	38.537	6.8330
TbPopn	1000	415	915	710.88	151.939
PctNonCdn	1000	.0	26.8	2.344	4.4965
ImmPctDich	1000	.0	81.8	34.976	29.1412
VisMin	1000	0	145	3.34	21.747
PopPctDutch	1000	5.7	56.1	30.673	8.9514
PctWhChrRel	1000	6.8	68.1	40.828	17.5791
ParticipationRate	1000	53.8	81.0	71.304	8.4610
EmpRRate	1000	47.9	81.0	67.626	10.2026
UnempRt	1000	.0	15.7	4.099	6.2736
LFPctAgn	1000	6.3	50.6	16.308	14.6212
IncInMed	1000	18345	40745	26335.06	5504.998
IncInAvg	1000	23383	50412	33575.63	6121.086
LowIncPrev	1000	2.4	45.3	12.778	9.3445
IncHhldMed	1000	46865	86862	64675.97	10660.160
IncHhldAvg	1000	58609	92408	70736.03	9318.994
IncFamMed	1000	46964	123022	71572.14	17369.733
IncFamAvg	1000	63229	115898	78891.76	12797.795
Christian	1000	270	750	570.20	148.001
OtherChristian	1000	25	395	241.70	116.955
ImmigTtl	1000	20	105	67.27	24.193
ImmigDich	1000	0	60	26.87	21.454
TbLbrFrc15P	1000	220	500	387.91	82.473
AgrIfOfiHu	1000	20	205	59.53	50.884
MtIncM	1000	77.40000000	88.60000000	83.97090000	3.170137688
GovTms	1000	12.40000000	21.70000000	16.03450000	2.960659598
EthncDich	1000	30	345	221.66	84.192
CRCCanada	1000	390	900	688.59	155.602
CbNonCan	1000	0	150	14.88	25.376
Valid N (listwise)	1000				

N01P00A-data-and-1000Avg.xlsx - Microsoft Excel																				
File Home Insert Page Layout Formulas Data Colectica Review View Corel PDF Fusion																				
Clipboard Font Alignment Number Styles Cells Editing																				
A27																				
	A	B	C	D	E	F	G	H	I	J	K	L	M	N	O	P	Q	R	S	T
1		Measure	NHSNonR	TLIPopn	PctNonCd	ImmPctDt	VisMin	PopPctDu	PctOthCh	Participa	EmplrRate	UnempRt	LFpctAgri	IncndMe	IncndAvg	LowIncPr	IncHldM	IncHldAv	IncFamM	IncFamAv
2	Based on averaging the 14 Das	Minimum	11.2	415	0.0	0.0	0.0	5.7	6.8	53.8	47.9	0.0	6.3	18345	23383	2.0	46865	58609	46964	63229
3		Maximum	54.5	915	26.8	81.8	145.0	56.1	68.1	81.0	81.0	15.7	50.6	40745	50412	45.0	86862	92408	123022	115898
4		Average	36.0	607	4.4	37.1	10.0	28.3	36.2	72.2	68.1	3.7	23.3	26920	34708	13.0	64988	76362	71603	82125
5		Stan.Dev	10.7	138	7.3	32.8	39.0	13.8	20.1	7.8	9.5	5.7	15.3	6624	6412	12.0	11739	11370	19695	13819
6		95%low	14.6	331	0.0	0.0	0.0	0.0	0.0	56.5	49.0	0.0	0.0	13672	21883	0.0	41510	53621	32213	54487
7		95%high	57.4	884	19.0	100.0	88.0	55.9	76.4	87.9	87.2	15.1	53.9	40168	47532	37.0	84866	99102	110993	109762
8	Based on averaging the 1,000 test matches from PCCF+	Minimum	11.2	415	0.0	0.0	0.0	5.7	6.8	53.8	47.9	0.0	6.3	18345	23383	2.0	46865	58609	46964	63229
9		Maximum	54.5	915	26.8	81.8	145.0	56.1	68.1	81.0	81.0	15.7	50.6	40745	50412	45.0	86862	92408	123022	115898
10		Average	38.5	711	2.3	35.0	3.3	30.7	40.8	71.3	67.6	4.1	16.3	26335	33576	12.8	64676	70736	71572	78892
11		Stan.Dev	6.8	152	4.5	29.1	21.7	9.0	17.6	8.5	10.2	6.3	14.6	5505	6121	9.3	10660	9319	13730	12798
12		95%low	24.9	407	0	0	0	12.7	5.6	54.3	47.2	0	0	15325	21334	0	43356	52098	36832	53296
13		95%high	52.1	1015.0	11.3	93.2	46.7	48.7	76.0	88.3	88.0	16.7	45.5	37345	45818	31.4	85996	89374	106312	104488
14		DAUID																		
15		35280320	28.0	415	0.0	0.0	0	31.3	36.0	77.2	68.4	0.0	11.4	40745	50412	2.4	86862	90965	123022	115898
16		35280321	37.5	495	0.0	0.0	0	10.1	6.8	70.1	60.9	13.1	29.5	30691	37285	13.1	70404	92408	76252	92355
17	ACTUAL Data	35290285	25.6	560	26.8	18.2	145	12.5	9.3	81.0	78.0	3.7	50.6	18376	29929	7.1	62947	88036	54615	73168
18		35320293	27.6	435	10.3	76.9	0	28.7	58.6	81.0	81.0	0.0	23.1	18345	27889	6.9	46865	62831	46964	65201
19		35320294	11.2	615	4.9	55.6	0	26.8	37.1	61.6	59.3	3.8	9.4	33171	39468	5.7	69891	80679	79234	87474
20		35320297	54.5	530	0.0	0.0	0	5.7	15.4	74.7	69.9	0.0	27.4	33332	37686	10.4	77595	82301	77398	80343
21		35320298	49.0	695	0.0	53.3	0	35.3	60.0	74.1	61.6	15.7	7.2	22048	31053	11.4	61800	62609	64456	72702
22		35320299	37.3	825	3.6	71.4	0	37.0	32.4	78.1	76.6	0.0	10.0	29938	36385	6.1	79953	72925	91903	91669
23		35320300	37.6	675	3.0	0.0	0	17.0	16.9	53.8	47.9	12.7	6.3	24589	30342	20.0	51819	58609	51966	63229
24	2 SDs out	35320301	36.0	915	0.0	38.9	0	32.8	40.7	69.2	68.5	0.0	6.7	29647	37709	6.6	64957	76689	71204	81558
25	Newly out	35320302	44.0	585	0.0	0.0	0	33.3	68.1	78.0	78.0	0.0	44.4	18981	23383	28.2	59160	63912	67278	68535
26	Newly in	35320304	35.2	530	7.5	70.6	0	24.5	21.1	65.3	58.7	0.0	40.8	23180	35853	45.3	46931	84126	51442	91665
27		35320305	38.7	615	0.0	81.8	0	56.1	50.0	76.8	76.8	0.0	24.2	24127	32493	8.9	64993	69235	80745	79485
28		35320306	42.3	610	5.7	52.4	0	45.1	53.9	69.9	67.7	3.1	35.4	29713	36019	5.7	65656	83737	65964	86464

So, using the 1000 iterations of the PCCF match, we get slightly different values – and some values which weren't outside 2 standard deviations are added, and one is removed. So, the question still remains – to what extent are these data representative?

Have I discovered that ice is cold, water is wet, and steam is hot?

- Have you encountered problems with either DMTI or the PCCF pointing to areas they shouldn't?
- Have you looked closely at the data returned by a single run of PCCF+ and compared it to what might have been returned?
- Could PCCF+ return the probability of having selected the particular area?

To think about / discuss - FSAs

- What if you only have FSAs to match against?
 - What do you currently advise?
- Do you need to use PCCF+ (PCCF isn't really an option)?
 - Boundary files for FSAs are already available
 - *[Some]* summary data tables at the FSA level are also available
- Is there value added or removed by using PCCF+?

To think about / discuss – Rural Postal Codes

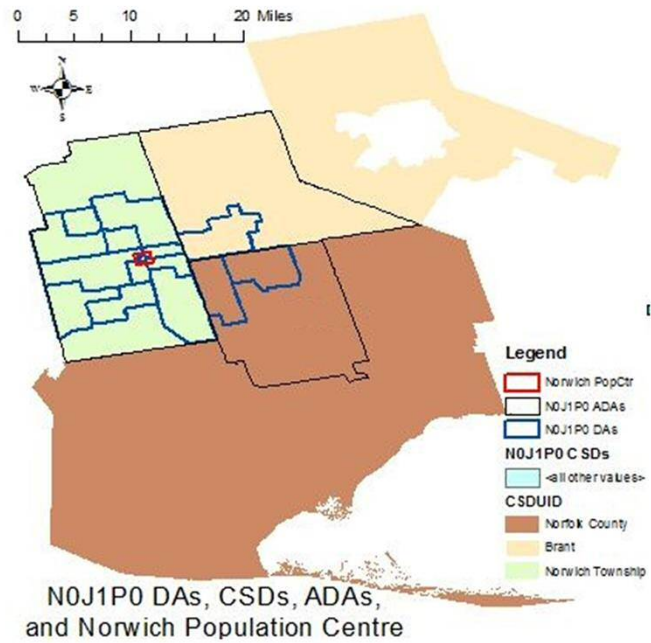
- What do you currently advise?
- Should we recommend users do more work with rural postal codes to get more representative data?
- Should what we recommend depend on what the user intends to do with the data?

CSDs (e.g., township) don't cross county boundaries – and this one postal code contains records for three counties. So, if we used Norwich township, we'd exclude Das that were matched over 8% of the time. Does it matter, or do we care, or should we? What if the CSD boundary includes OTHER non-NOJ1P0 dissemination areas?

Could we suggest a different geographic level?

Aren't a lot of possible candidates

- Have mapped the Population Centre, the Dissemination Areas, the Census Subdivision, and the new Aggregate Dissemination Area



Thanks for your attention

I hope this leads to some discussion around the use of PCCF+ for rural areas – and maybe formulation of best practices for use or informing patrons of the ***potential*** pitfalls of use

I'll leave you with this thought:

Inquiring minds may want to know, but they can also be an absolute pain ...

If it helps (which it shouldn't) I've been fretting about this for over a month now!