**Data Purchase and Access Working Group Meeting**
**1:30pm EDT -September 16, 2015**
teleconference #: 1-888-271-3643
Adobe Connect meeting: https://meet86536941.adobeconnect.com/dpawg-09-16-2015/

**Agenda**

1. Follow up on orders: Census/NHS, Canadian Survey on Disability, General Social Survey, Labour Force Survey, custom taxfiler data, CIC Permanent Residents Rounded Data Cube
2. Custom geography group #3
3. Information update on Environics Analytics products
4. Core Housing Need variables
5. Longitudinal Administrative Database
6. Other business
7. Next meeting

In attendance:
Isabelle Lepine - Ville de Montréal
Louisa Wong - City of Hamilton
Andrea Dort - Region of Peel
Heath Priston - City of Toronto

**Census/NHS tables**

See the document: "CDP orders update Sept 16, 2015 - Census-NHS tables.docx". We have sent in a request for a cost estimate and are waiting for a response.

**Canadian Survey on Disability**

The following tables have been ordered at the level of CD/CSD, CMA/CA, PT, Canada and Economic Region and Consolidated Census subdivision:

- 115-0001  Prevalence of disability for adults, by sex and age group (all adults, under 65, 65+)
- 115-0002  Adults with and without disabilities, by sex and age group (all adults, under 65, 65+)
- 115-0003  Adults with disabilities, by type (not by sex or age)
- 115-0004  Adults with disabilities, by severity (not by sex or age)

The Economic Regions have lots of good data but the Census Consolidated Subdivisions have anywhere from 75% to 90% suppressed - same as the CSD, low population sizes as expected.

**General Social Survey – Cycle 24**

The following table has been ordered at the level of CD/CSD, CMA/CA, PT, Canada and Economic Region and Consolidated Census subdivision:

Section 4: Unpaid Work Activities
Module: Volunteering (VCG), by sex and (all respondents, under 65, 65+)
- In the past 12 months, did you do unpaid volunteer work for any organization?

**Labour Force Survey**

We repeated the same variables from the last order and requested additional Census Divisions (up to 12) including Simcoe, Saint John, Nipissing, Oxford, Chatham-Kent, Wellington, Dufferin, Alberta Division #8, #15, #16.  We are not particularly hopeful that we will get additional geographies.  The CSD data from previous orders was 'grandfathered in' from the FCM QOLRS project.

**Custom taxfiler data**

Custom taxfiler data for 2011 has already been acquired for this year (working poverty and inequality), and we plan to acquire 2012 and 2013 next year, with the possibility of acquiring at least 2012 this year, budget permitting.

Custom geographies applied to taxfiler data will not be acquired this year. This was confirmed at the annual meeting. We have initiated discussions with STC about how to acquire taxfiler data using custom geos, but don't intend to spend any money until next program year.

**CIC Permanent Residents Rounded Data Cube**

There are no plans to release a 2013 data cube. Data will be made available on the Federal government's Open Data portal, however the date for this release is not known.

**Status and estimated cost for the work on Custom Geography Group #3**

This information comes from Gerald at Statistics Canada who is working on the custom geographies.  He provided an estimate of the time needed, but unfortunately not an estimated delivery date.  We are trying to get this information.

**Toronto:**  Counts are out for verification with TPS.  Still waiting to hear back from the Manager there as to whether he wants any changes.  The Division file we used however was already on our base and used as part of another project so the work involved with that part is minimal.  Estimate: 10-12 hours ($748.50 – $898.20).  Likely on the lower end since I can't really imagine any changes here.

**Durham:**  One layer completed.  Second layer had issues.  Just received revised layer back from Durham and will now finish up the work required.
Estimate:  25 hours ($1,871.25)

**York:**  EDI geography final and paid for by York as part of another project.  Police geographies are approved.  Still waiting on York Schools for approval of their counts.  We took the approach of revising what was previously done as opposed to starting fresh which would have been a lot more time.
Estimate: 15 hours ($1,122.75)

**Vancouver:**  Really a matter of formatting and building the final geography.  Given the number of layers I am estimating  5 – 10 hours (tops out at $748.50).

## Environics Analytics

As a follow up to discussions at the annual meeting we acquired more information about Environics Analytics products.

### Enhanced Postal Code Conversion File

1.  How often is it updated and what are the primary differences between the PCCF produced by EA and STC's?

a.   Updated 2x year
b.   Differences:
> i. Enhanced rural methodology: including a postal code table & a community name table – They publish the community names
> ii. Corrections to STC PCCF
> iii. Verification process to bring in new and changed postal codes using Canada Post data and street files via TomTom

Action: Look at STC's product

**Environics Analytics**
**PRIZM**

2.  What data are available from the PRIZM database? Could the Community Data Program acquire access to parts of the underlying database? What would the cost structure be for this? What licensing restrictions exist?

a.   Every postal code gets a 2-digit code related to a Prizm cluster
b.    Ingredients that go into Prizm will not be divulged (demographic variables and urbanity indicators, Equifax, consumer database etc..)
c.   Associated behavioural, attitudinal and media consumption data

Action: Look at standard promotional material

**Environics Analytics**
**Business Data**

3.  Does EA support any database that would allow us to measure community-level (CD/CSD) performance of local businesses e.g., gross revenue, HST remittances, anything with a dollar figure?

EA acquires two sources of business data. Both include revenue:

a.  InfoCanada
b.  Dunn & Bradstreet

EA would be able to extract specific variables for smaller geographies.

Action: Prepare an order for a specific variable

**Core Housing Need**

We are looking into ordering Core Housing Need data from Statistics Canada.  At the Census Subdivision level we should be able to get all the same cross-tabulations that are available in HiCO:

- household type (total, family households, couples, lone parents, multiple-family households, non-family households)
- Aboriginal status (total, Aboriginal household, non-Aboriginal household)
- age of primary maintainer (total, non-senior households, 15-29, 30-44, 45-64, senior households)
- tenure (total, owners, renters)
- housing standards (total, above all housing standards, below housing standards, below adequacy standard, below affordability standard, below suitability standard)
- household data (household count, avg household income, avg shelter cost, avg shelter-to-income-ratio)
- core housing need status (total, households not in core housing need, households in core housing need)

However at the CT and DA level it's unlikely that we could get all these cross-tabulations without significant suppression.  A request was made to working group members to indicate which dimensions should take priority for these geographic levels.  Tenure and housing standards seemed the most likely to be priority dimensions.  A request was made to see if it was possible to get median household income instead of average household income.  If this variable could be acquired then it would become a priority for the cross-tabulations.

There was also interest expressed in acquiring core housing need data at custom geographies.

**Longitudinal Administrative Database**

This order is on hold, largely because ordering down to CSD is cost-prohibitive and highly suppressed. We can order for CD and higher, but we'll need a stronger set of variables.  Michael

Ditor will investigate what variables could be ordered that would take advantage of the data's longitudinal nature.

**Next Meeting**

The next meeting was set for November 18, 2015 at 1:30pm Eastern Time.


Meeting ended at 2:15pm