DATA PURCHASE AND ACCESS WORKING GROUP
July 24, 2019 meeting
2:30pm Eastern time

Connection details:
Adobe Connect: https://cdp.adobeconnect.com/theboardroom/
Teleconference: 1-866-398-2885 ; access code: 908904

<u>Attendance</u>
Michael Ditor – CDP team
Albert Boakye – Winnipeg
Blair Hodgson – WDG
Ted Hildebrandt – Halton
Cheryl Hitchen – KFLA
Graham Good – Muskoka
Louisa Wong – Hamilton
Jasmine Ing – Calgary
Lan Do – Calgary
Sharon Stroick - Calgary
Andrea Dort – Peel
Auburn – Parry Sound-Nipissing
Kholah, Shannon, Bakul , Niket and Danxi – Simcoe
Allison Scott – Federal Census Division 11
Femi Adegeye - Regina

ACTION ITEMS FOR MEMBERS
- Submit your signed Letter Of Consent to Assignment & Assumption if this hasn't already been done.
- If you have priorities for the IMDB order, download the list of variables from the LIPS data working group and indicate your priorities by August 15. This zip file contains data dictionaries with detailed information on the variables.
- T1FF working poor data – how do you use it? The last purchase was for 2014 and 2015 data. We order this data every two years but are considering waiting another year to order 2016, 2017 and 2018 next year. Your feedback on this order is requested.
- Income Inequality data: How are you using these data? Are you using these data?
- Consumer debt and mortgage debt data: Do you use these data? What do you use these data for?

<u>Transition update</u>
CCEDNET is waiting until all agreements have been signed over to them before taking over administration of the program. We urge those who have not yet submitted their signed Letter Of Consent to Assignment & Assumption to do so as soon as possible. Please contact us with any questions or concerns you have about this process.

There are several products that are ready to be purchased but that have been waiting for CCSD to pay for them. We contacted Peggy Taillon, President of CCSD to pay for these tables and **they should be received the week of July 29**. These tables include:

- 12 Target Group Profiles at custom geographies for Group 4 (Winnipeg, Nova Scotia, Immigration Geographies, North Bay Parry Sound Public Health Unit)
- 23 semi-custom tables from the 2016 Census
- the Canadian Survey on Disability custom tabulations for provinces, CDs and CSDs

We will work on getting the data cataloged as soon as possible, starting with the tables at custom geographies.

Data Update – CUSP Energy Poverty and Equity table
The first table on energy poverty and equity from the project with the Canadian Urban Sustainability Practitioners has been cataloged (and paid for by CUSP). Download it here. More tables cross-tabulating these data with demographic profiles of households will be cataloged shortly.

IMDB custom order
We have been in contact with the IMDB group at Statistics Canada about getting permanent resident landings data. IRCC will no longer be providing us with the Permanent Residents data cube but we can get similar data from Statistics Canada in IVT format. As part of the Immigration Sector Network pilot project, we will be acquiring data to assist Local Immigration Partnerships (LIPs) in their reporting to IRCC and making it available through the Immigration Data Portal.

We need to create specifications for a few tables that meet the needs of our members as well as the LIPs. We've just begun diving into the large number of possibilities that can come from the IMDB landing dictionary as well as the T1FF.

The LIPs data working group is working on this as well, starting by prioritizing variables from a spreadsheet that was sent out to them. If you are interested in indicating your priorities for these variables, please download the sheet here. We will use the feedback from the two groups to come up with an order.

This zip file contains data dictionaries indicating the range of variables that can be included in the tables. Also included is a list of variables that were included in the IRCC data cube to which Toronto has access (supplied by Heath Priston), however this list is a subset of the variables that are available. It is also possible to get data on non-permanent residents (NPR) but only for those who have obtained permanent resident (PR) status.

As there is a limit to the number of dimensions/variables that can be included in a table, we will want to acquire a few different tables to cover the range of information needed. These themes were suggested as a starting point for a data order, but your feedback in this regard is also welcome:

- Demographic profile
- Economic outcomes/sources of income
- Retention (+secondary/in-out migration?)
- Non-permanent residents: demographics

Geography: we want to get these data at the CSD level as well as for custom LIP geographies (combinations of CSDs). There are 3 possible geographies with the IMDB:

- Intended destination at admission of the immigrant, at CSD level and up
- Residential geography for each tax year, at CSD level and up
- Service geography for each service year (or tax year), at CD level

Destination geography might be useful for retention information. Can we use first tax year as proxy for destination?

ACTIONS FOR MIKE – CHECK WITH STC/IMDB DICTIONARIES:
Check for smaller geographies (CT?) for larger municipalities.
Can we use first tax year as proxy for destination?
Information about the year of initial arrival in Canada (vs the landing year)?

Data at the CSD level will be subject to suppression rules, which we should keep in mind when designing the tables. Note that having a year of admission and a tax year attached to each record effectively decreases the population size (if we want to use any longitudinal aspect of the data). It's possible to group multiple admission years as long as income variables are not included. For example, we could create a demographic profile table for recent landed immigrants (2011-2016) but not include any tax file information.

Halton: Community Development Halton is working with the IMDB data for their LIP. Age groups ordered from IMDB data: Under 20, 20 to 44, 45 to 64, Over 64.

Mike (CDP) will continue to coordinate with Lan Do and Sharon Stroick (City of Calgary) of the LIPs data working group.

ACTION FOR WG MEMBERS: If you have priorities for the IMDB order, download the list of variables from the LIPS data working group and indicate your priorities by August 15. This zip file contains data dictionaries with detailed information on the variables. Priorities can be expressed as ranked variables or simply concepts that you would like to see integrated into the tables.

2016 Census Working Poor data order
There are issues with the Working Poor data order that will make it incomparable with the data we get from the T1FF.

Universe:
We use the education deduction from the T1FF to determine if someone is a student. There is no access with Census file as to whether they claimed an education deduction. We have access to whether they attended school, which is a better concept but not comparable.

Taxfiler data includes residents of collective dwellings while the long-form Census is only for private households.

LIM definition:
Taxfiler uses the Census family aggregation level to determine LIM status. Census uses household level aggregation – might be a bit more expensive to get Census family aggregation.

ACTION (CDP): use the household concept to keep in line with other Census tables.

T1FF Working Poor data order
ACTION FOR WG MEMBERS: T1FF working poor data – how do you use it? The last purchase was for 2014 and 2015 data. We order this data every two years but are considering waiting another year to order 2016, 2017 and 2018 next year. Your feedback on this order is requested.

Calgary: order twice every five years, taking into account coverage by Census.

Income Inequality data order
There are issues with this order that will make it extremely expensive to recreate this table. We would like to hear how this table is used by our members.

Decile groups: geographic-specific decile groups will be very expensive. Canada-wide deciles will be much cheaper

If we get these data, we'd have to get them down to the CSD level – not the CT or ADA level (too expensive).

The data will also not be completely comparable with taxfiler income inequality tables (different universes). Also, the top 1% data will be relatively useless – lots of suppression in this category.

ACTION FOR WG MEMBERS: How are you using these data? Are you using these data?

Custom geography updates
We are still waiting for CCSD to pay for Group 4's 12 TGPs. The next production round for tables at custom geographies, including Vancouver and Halton, will begin shortly. It's not clear yet when the next set of tables (for all custom geos) will be produced. Note that we will be acquiring three BC province-wide custom health geography sets in the next custom geography run.

Modest income table
We will try to get a discussion going on the specs for a modest income table (CPP-10). The continuation of this discussion will need to happen via email.

Child care costs table
Custom child care costs table requested by Simcoe County. CDP asking Simcoe County to start putting together specifications for a custom table.

Table request: household income groups crossed with tenure
Toronto request for table that crosses household income groups against housing tenure.

99-014-x2011028 had this but has been demoted in priority because we already have a number of the stubs covered in other orders (it's mostly about Core Housing Need elements).

Consider getting 2016 Census equivalent table: 98-400-X2016225.
Toronto: acquire at custom geographies. CSD/CD may also be useful for others.

Calgary: HH TGP should cover this data gap.

Production capacity
In response to our request to have more analysts assigned to our order, STC confirmed that they would be distributing the work among several analysts.

TransUnion debt data
Do we need to order this? Please let me know whether you are interested in this data set. If you have used these data, what did you use it for? Was anything produced?

ACTION FOR WG MEMBERS: Do you use these data? What do you use these data for?

ACTION FOR CDP: Delay order of these data until next program year (unless strong pushback occurs).

Data modeling project (liquid assets)
We have $30K in the budget to proceed with this project to model liquid assets data at the CSD level. We've spoken with a potential partner who may want to contribute towards the cost of the contract. We can't go forward with this project until after the transition.

T1FF custom data (low priority for now - could be cut from agenda)
The custom data sets are available in the catalogue. If we were to buy this same order in 2020-21, it would cost in the range of $30K (double what we paid this time). If you are interested in these data, please take a look between now and next year's annual meeting (yes, you've got time but the data aren't getting any younger).

Next meeting
Will send out Doodle poll for week of Sept 16