

Data Purchase and Access Working Group meeting

February 2, 2017

1:30 – 3pm Eastern time

<https://cdp.adobeconnect.com/theboardroom>

Teleconference number: 1-888-271-3643

PRESENT

- Amanda Richards – Peel
- Cheryl Hitchen – Kingston
- Christina Maes Nino – Winnipeg
- Heath Priston – Toronto
- Isabelle Lépine – Montréal
- Louisa Wong – Hamilton
- Sian Jones – Calgary
- Natalie Hui – York Region
- Olga Belanovskaya– Simcoe
- Ted Hildebrandt – Halton

Agenda

- 1. Follow up to the January Leads’ Meeting on the Census data purchase..... 1
- 2. Re-working the Census purchase priority list 2
- 3. Discuss/review the Community Poverty Project tables and determine if changes are required 3
- 4. Discuss process for bumping up priority of data tables (requests from Leads) 4
- 5. Discuss methodology for introducing housing data into Target Group Profiles..... 4
- 6. Follow up on Income Inequality tables 5
- 7. Other business 5
- 8. Next meeting 5

ACTIONS

- The CDP team will compile the 2006 tables and its variables. This may take some time, but the team can try to do this task as they prepare table lists and process data requests.
- Question for Statistics Canada - see if they have references that they can give us regarding 2006 and 2016 Census.
- List each target group and suggest a definition that would allow inclusion of housing and family data. These definitions will be sent to the rest of the DPAWG members for suggestions and changes.

1. Follow up to the January Leads' Meeting on the Census data purchase

Nothing to be added.

2. Re-working the Census purchase priority list

The members were reminded about the Census acquisition spreadsheet circulated after the Leads' Meeting on January 19, 2017. The spreadsheet contains product rankings by number of downloads. The team recognizes a few potential flaws, such as that a user may be downloading and disseminating the data or certain products have been in the catalogue longer than others. The CDP team needs to know the table priorities and requests before the annual face-to-face meeting in May, 2017.

One concern regarding the Census tables is the continuity of the 2006 Census and the 2016 Census. The members want to identify the differences between the 2006 and 2016 data variables and see what corresponds. Also, there is interest in requesting 2006 data using 2016 geographies to make accurate comparisons over time.

ACTION: The CDP team will compile the 2006 tables and its variables. This may take some time, but the team can try to do this task as they prepare table lists and process data requests.

Regarding the budget for purchasing tables, the team could probably replicate the entire order, but there should be some change in priorities. There are some tables that were downloaded fewer than 10 times, perhaps some of these tables can be cut from the list. By removing some of these low-demand tables, the team can get more custom geos or get the 2006 data people are interested in. Also, there are other table priorities besides the Census.

The CDP team has been told that the Census standard profiles will be free to the DA level on the Statistics Canada website. Considering this, the team will not get profiles at standard geographies. For those interested in custom geos, the team can focus on that. However, there is no indication that all standard profiles will be free.

Louisa (Hamilton): For those standard tables that would become available that the DA level on the website - we don't need worry about user download numbers for those tables, because those numbers will skew our priorities. Some of those tables are key ones that we use and there's also those bigger special tables like target group profiles at custom geos.

Cheryl (Kingston): The question might be for all of us, which ones are on the list that we would like at our custom geos, and once we have sense of that we can decide how much to take off the list? I think Mike is trying to identify what out of these we want there, or prioritize the ones without custom geos that we might like at custom geos. Would that be useful, pick which one without custom geos that we would like?

Mike (CDP): Yes, it would be helpful.

For other priorities and custom geographies purchases, the team needs to start eliminating the less popular tables. It would be useful to look at the list of less popular tables and the members should identify which ones that are interested in and bump them up the priority list.

The first Census profile will be released on Feb 8th on Population and Dwelling Counts. Statistics Canada will keep updating the profiles as the topic data tables are released. However, the members must decide when to order the profiles at custom geographies - keeping in mind that the team cannot order the profile

at custom geographies with each release. If there is no rush, the most efficient way is to wait until November 29, 2017 when the full profile is available. If there is a rush, the team should pick a midway point when most data is released. That way setting priorities is more effective, and the team isn't ordering too much data twice.

Statistics Canada has told the CDP team that they are in the queue for custom geos. The CDP team will look at the list to see how it relates to the 2006 order.

Sian (Calgary): Do we know if there are any comparability issues between 2006 to 2016? We can use that to prioritize.

Heath (Toronto): Some changes to education and health related questions. More specific than it was before. Also, the income.

Cheryl (Kingston): Maybe looking at 2006 may not be the best idea, as we're not sure if we'll be comparing apples to pears to bananas! Population, age, poor indicators, but things like cultural diversity, there can be some changes in the details, don't want to compare if they don't match. The first two releases will be quite standard. But the later releases will need to be looked at.

Ted (Halton): We will be looking into this, too.

ACTION: Question for Statistics Canada - see if they have references that they can give us regarding 2006 and 2016 Census.

The team sees value in data continuity for the 2006 and 2011 orders, and will link the connections between them. The CDP team will process requests as high priority items as well as use popular downloads as an indication of priority. The team will develop a list to bring to face-to-face meeting in May, 2017 and will have more information before then to help prioritize data and which tables to bump up or down the list.

3. Discuss/review the Community Poverty Project tables and determine if changes are required

The team has made a list of all existing CPP tables and their details. It may be too early to discuss now, so the team can have a few email discussions about the CPP tables afterwards. There were a few tables that were not ordered last time (there was a discussion on that and the team was okay with not ordering the data at the time). These tables will be put back into consideration.

Heath (Toronto): Speaking about the housing – it would be nice to have property tax, house value, and related data to understand that part of the household burden. There is interest in housing in general. When looking at poverty in Toronto, we look at issues such as fair equity, transit, mobility... Also, hearing about seniors 65+ is too broad, especially with people living longer and older people immigrating. We will probably need 80+, 85+, etc. This is something we are interested in pursuing, as well as activity limitations.

Cheryl (Kingston): This is something we are also looking at. We need more information on mode of transportation and more age groups. People are drawing conclusions on things that don't align. Looking at CPP tables, we are looking post-November... how long after November can we expect the tables? If we get it early 2018 then I'm happy, but if it's 3 years from now, it's old data.

The CDP team is looking at how much they can frontload the data order. In the last round, this was spread out over 3 years. The team had not planned on putting in any rush orders. Statistics Canada is committed to getting data out faster this time around. Also, the team is working to get these orders ready before the topic release date.

Sian (Calgary): Another topic of interest – getting questions related to carbon tax. We would like to look utility costs around housing and pre-senior age groups (45+ ,55+).

Cheryl (Kingston): It would be good to get these age groups, like 45-55, to compare with TIFF. As well as older, older adults. Would be nice to have property taxes and mortgage. What else?

Heath (Toronto): Property taxes, electricity (separated), municipal services, etc.

The CDP asks that if there are any particular tables that the DPAWG teams wants with certain age groups, let the CDP know. By sending in requests and the CDP team can prioritize them.

4. Discuss process for bumping up priority of data tables (requests from Leads)

The CDP team is committed to treating requests as priority items as well as developing the methodology to do this efficiently.

5. Discuss methodology for introducing housing and family data into Target Group Profiles

Housing and family data – some wanted this incorporated into the TGP. The TGP data was taken from the population of individuals, so the household and family data was not included. For a TGP table, i.e. a minority household or immigrant household, there would be criteria to identify a household that can be classified as part of the target group. For instance, if you have one member of the target group, then the household is identified. Or, if you have 50% of the household with the target group, then the household is identified as part of the target group.

Some members want to have more inclusive definitions, such as needing only one member of the target group to make the profile. However, this definition can change, depending on the situation.

Cheryl (Kingston): Is it 50% of the adults that determine this? Would an Indigenous child count as an Indigenous household? I don't know. What if there were 8 members in the household and one of them was an Indigenous child... would the child count or does it have to be the adults in the household?

Heath (Toronto): If we are looking at low-income population, how is the low-income determined? Presumably, that would be based on the household level measure of income. Every individual that lives in a low-income household is considered low-income?

Cheryl (Kingston): Yes, it's the family's income that they use, identity is a little different. I think we have to be really careful with this! For example, if there is one Francophone in the household with lots of different people in it...

Heath (Toronto): Are we trying to get the number of households that have people that meet the profile? Or are we trying to understand some conditions on how people meet the profile?

Cheryl (Kingston): Trying to remember how this came up – it was our discussion on trying to get household data for the target group profile regarding Francophones, how many individuals versus how many families. We need family dynamic information. So, perhaps just knowing one person is fine for this definition.

The team recognizes that the definition may need to vary. The team will look at each TGP individually and see whether an inclusive definition would be appropriate.

ACTION: List each target group and suggest a definition that would allow inclusion of housing and family data. These definitions will be sent to the rest of the DPAWG members for suggestions and changes.

6. Follow up on Income Inequality tables

About the Income Inequality tables, there was some aggregate income data that was suppressed due to dominance suppression. There were a few places that had the tenth decile income suppressed. Dominance suppression is when one value of the decile is skewing values of cell so much it may compromise confidentiality.

The data were also re-ordered using a ‘purist’ decile methodology in which each decile contains exactly one tenth of the values. Previously, the decile thresholds were determined first and values assigned to each decile, with all values equalling a threshold value put in the same decile. This method resulted in some deciles having more than one tenth of the data values and having a higher aggregate income than ‘higher’ deciles (see notes in the Income Inequality catalogue entries).

The CDP team requested the bottom coding of the first decile to see if dominance suppression could be avoided. The bottom coding did not seem to make a difference, as tables with and without bottom-coding produced unsuppressed data. It appears as though the purist decile methodology fixed the dominance suppression. This methodology is written as notes and are included in the catalogue entries.

The tables are in, with full sets of unsuppressed data in geographies that had previously been suppressed. There are still some questions as to exactly how we got it, so these questions have been put to Statistics Canada. The team will go forward with this purist decile methodology – each decile containing exactly 1/10th of the data values.

7. Other business

A recent report on point-in-time (PiT) counts of homelessness was released by Employment and Social Development Canada (ESDC). The team will try to get this data in Excel format at the community level. However, ESDC had previously indicated that they would not release community level data because of conflicts with communities wanting to release their own data. It would be great to get this data all in one table. The team will need to get the communities’ permission to get this data. If any DPAWG members are involved, please be willing to allow this data to be published.

8. Next meeting

Wednesday April 5th or April 12th, 2017. We will determine the best date via Doodle poll.

Meeting ended 3:00pm ET.