Operationalizing 'Self-contained Labour Areas'

as a potential standard geography for disseminating Statistics Canada data

Alessandro Alasia

Centre for Special Business Projects

Canadian Rural Revitalization Foundation, 2016 Conference

Guelph, October 13th, 2016

Statistics Canada   Statistique Canada
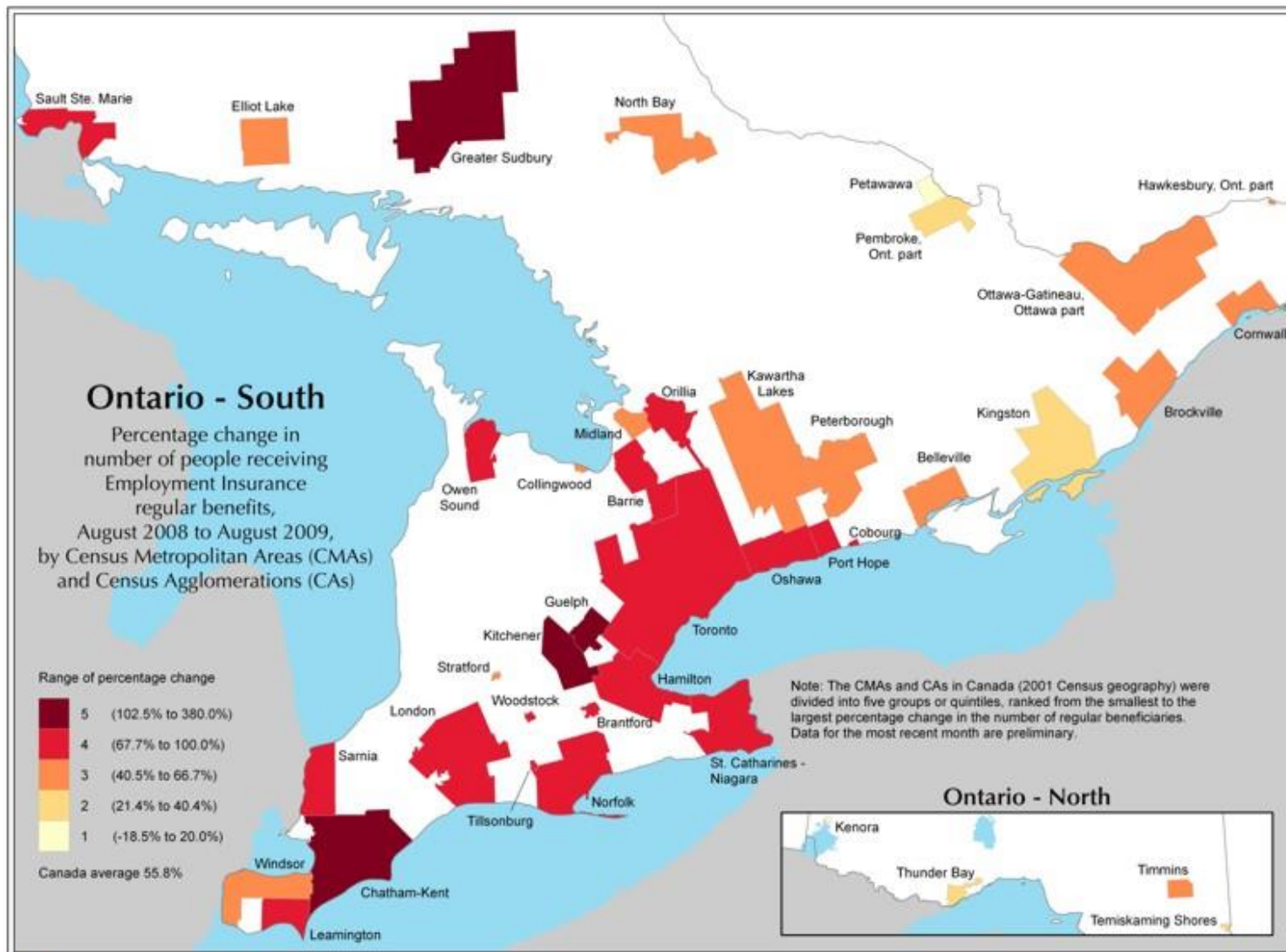
Canada

Context

- The self-contained labour areas (SLA) came from the desire to explore new ways to deal with local data, particularly for non-urban areas.

- In 2011, 19% of Canadian residents lived outside of the Census Metropolitan Areas (CMAs) and Census Agglomerations (CAs) These areas are often:

  - Less clearly delineated than urban areas, due to administrative differences between regions

  - Less accessible for analysis, as populations are often small for survey or sample statistics to be created or tracked reliably

Experiences in other countries

- "Commuting zones (CZ's) and labor market areas (LMA's) were developed because **county boundaries are not adequate confines for an area's economy**. A local economy and its labor market are bounded not by the nearest county line, but by interrelationships between buyers and sellers of labor. If we are to understand the diversity of nonmetro America we need a geographic standard capturing variations in local economic and labor force activities. The central objective of CZ's and LMA's was to develop such a geographic unit that better captures the economic and social diversity of nonmetro areas."
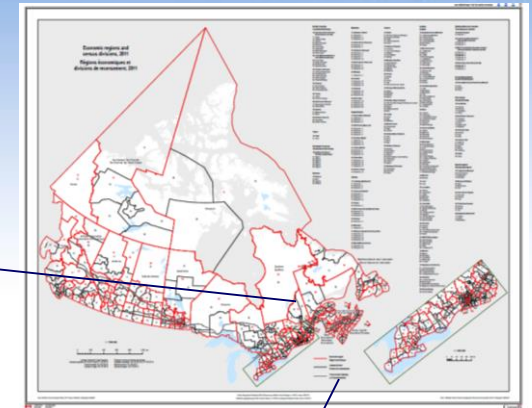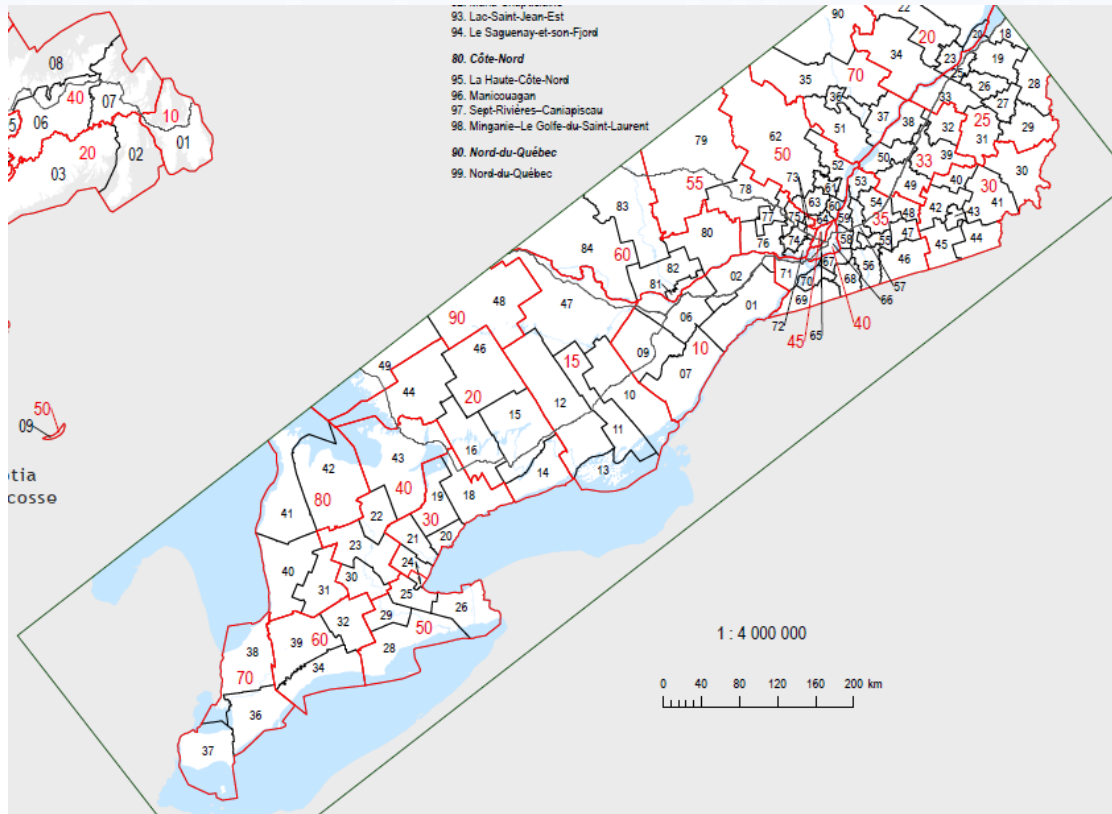
Source: Commuting Zones and Labor Market Areas in the US. Available at: http://www.ers.usda.gov/data-products/commuting-zones-and-labor-market-areas/documentation.aspx

An example

Source: http://www.statcan.gc.ca/pub/73-002-x/2009005/map-carte/ont-eng.htm

Sub-provincial geographies, example

- Economic regions and census divisions

Source: http://www12.statcan.gc.ca/census-recensement/2011/geo/map-carte/ref/sgc-cgt/index-eng.cfm

Project background

- Analysis with 2006 census data. Published in "Self-contained labour areas: A proposed delineation and classification by degree of rurality". Available online at: http://www.statcan.gc.ca/pub/21-006-x/21-006-x2008008-eng.pdf

- Update analysis with 2011 Census/NHS data.

- Currently: cost-recovery project in collaboration with the Statistical Registers and Geography Division (SRGD), and funded by ESDC. Objective: to refine the 2011 SLA delineation, and move it toward a new geographical classification

Definitions

Self-Contained Labour Area (SLA):

- A group of Census Subdivisions (CSDs) that meets or exceeds a threshold of labour self-containment
- Self-containment for any SLA will never be less than 75%

Self-containment has two parts:

- **Self-containment of Area Workers**:
  - % of area jobs that are filled by area residents
  - (How many of the available jobs are filled by local residents?)
- **Self-containment of Area Residents:**
  - % of area residents that have jobs in the area
  - (How many of the local residents stay in the area to work?)

Methodology

Methodology Summary:

- Custom clustering process that grouped areas together based on the reciprocal strength of the commuting ties between them

Features:

- Minimizes urban bias
- Uses a sliding scale that requires higher self-containment for smaller areas
  - From a maximum of 90% for areas under 1,000 workers
  - To a minimum of 75% for areas above 25,000 workers

Methodology: commuting strength

- Equation adapted from OECD studies and that used to define Travel-to-Work Areas for the UK:

$$S_{a,b} = \frac{(F_{a,b}) * (F_{a,b})}{R_a \quad W_b} + \frac{(F_{b,a}) * (F_{b,a})}{R_b \quad W_a}$$

- $F_{a,b}$ is the number of journeys to work from area A to area B
- R_a is the number of workers who live in area A
- W_a is the number of workers who work in area A
- R_b is the number of workers people who live in area B
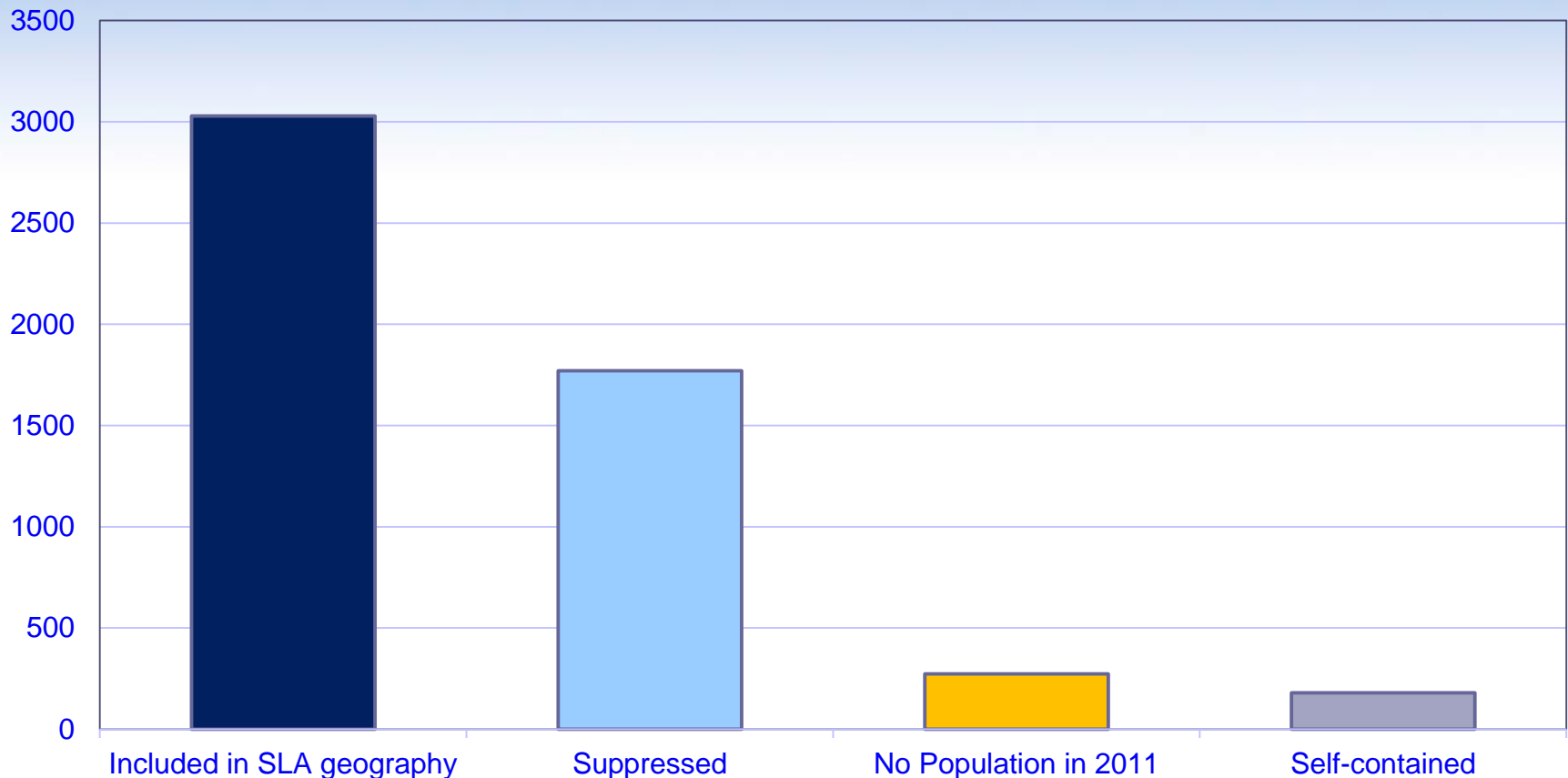- W_b is the number of workers who work in area B

Results from clustering algorithm

- Using 2011 Census/NHS data, there were 329 self-contained labour areas resulting from the clustering algorithm

- The average self-contained labour area:
  - Was 96% self-contained
  - Had a resident workforce of 36,000
  - Had a resident population of 89,000
  - Contained 11 CSDs
  - Was geographically contiguous

- Some CSDs were not grouped because they were completely self-contained or had suppressed data

SLA geography exclusions, 2011
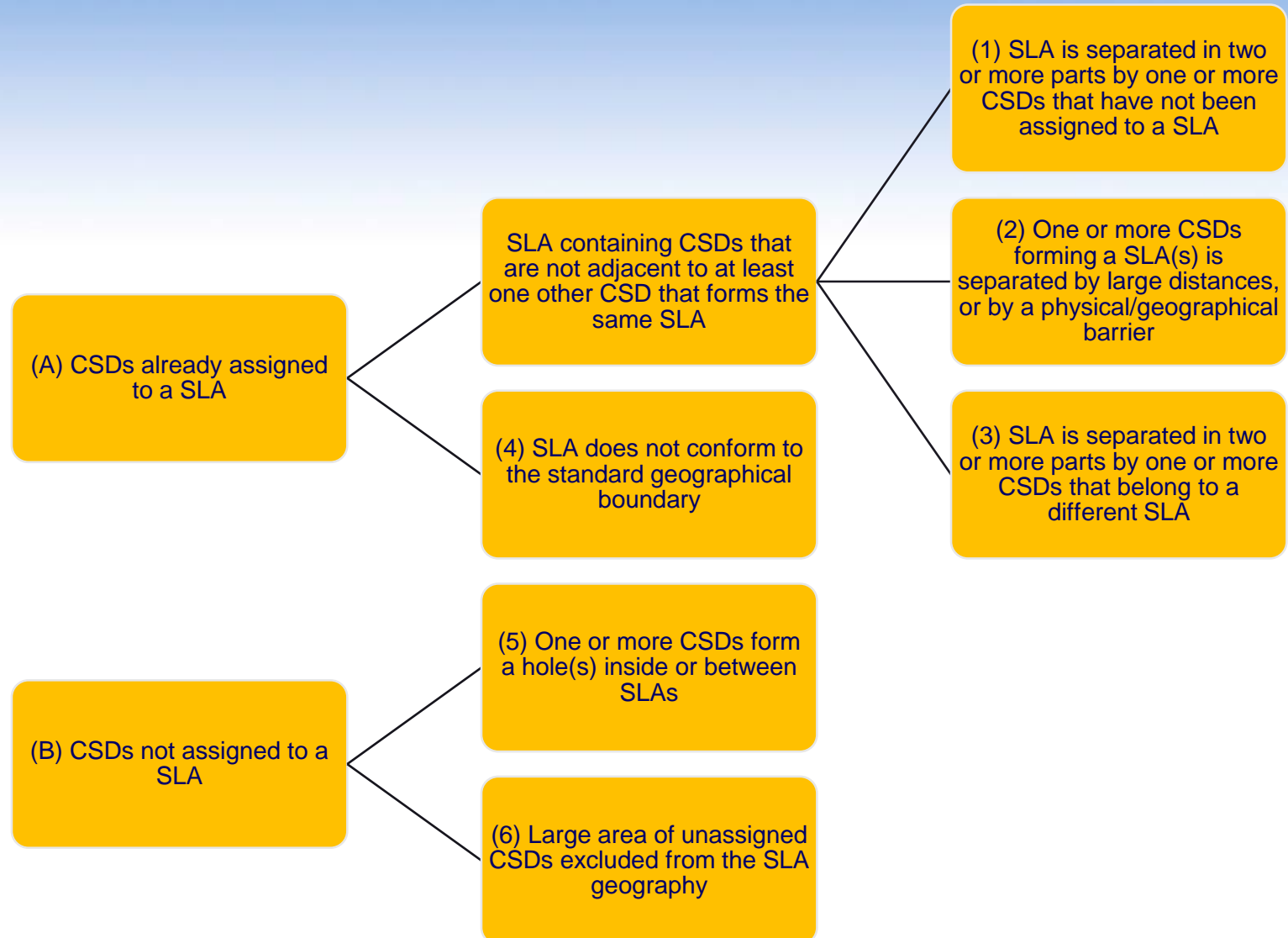

Number of CSDs

Total Number of CSDs = 5253

Bar chart showing Number of CSDs by category:
- Included in SLA geography: ~3025
- Suppressed: ~1775
- No Population in 2011: ~275
- Self-contained: ~180

Legend: ■ Included in SLA geography □ Suppressed ■ No Population in 2011 ■ Self-contained

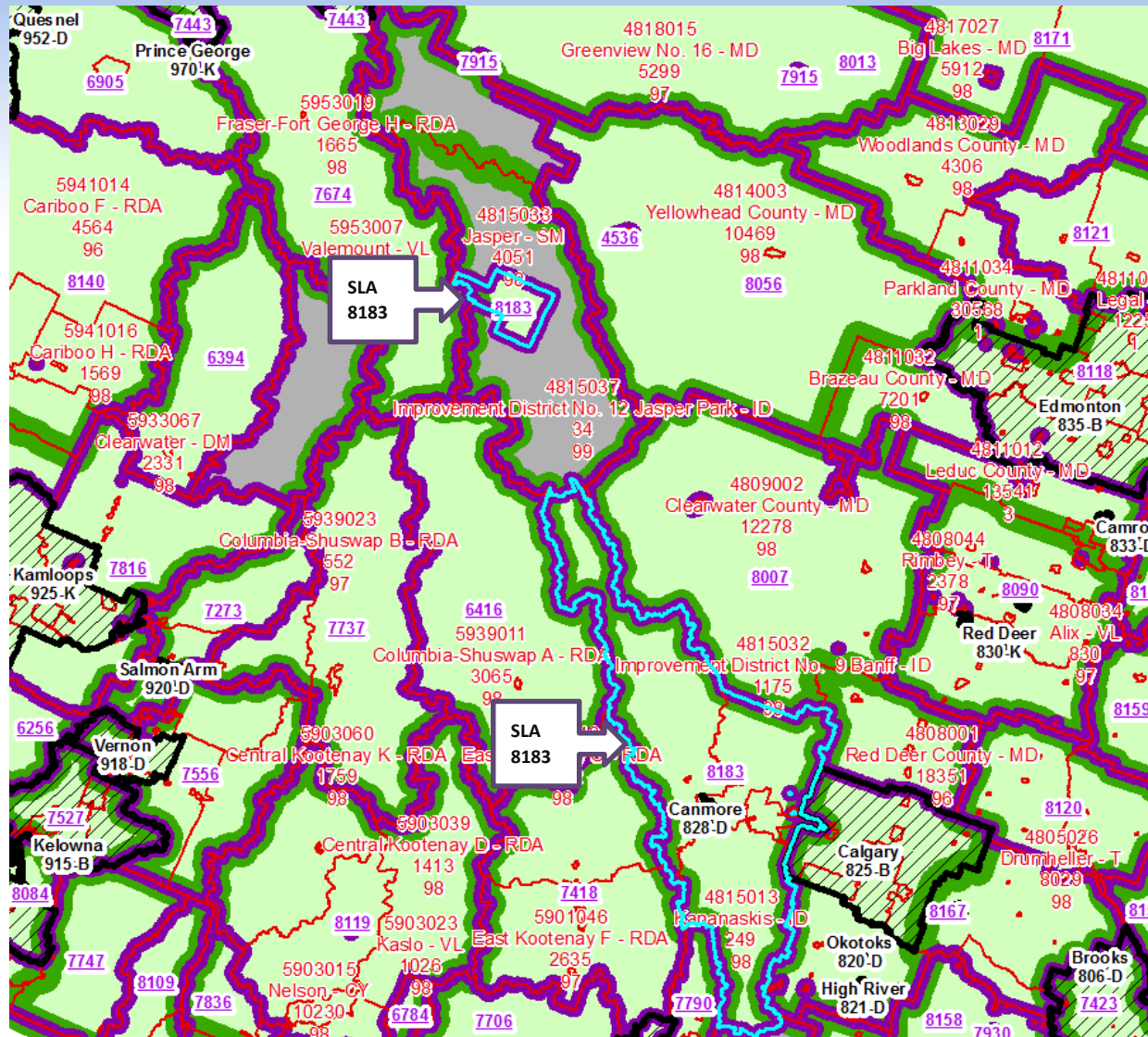# Two key considerations in enhancing the clustering results

- Preserving geographic contiguity versus ensuring minimal impact on the population
  - If too many people are affected by the SLA boundary realignment, then the concept of a "SLA" may not hold true any longer

- Respecting the unique characteristic of a particular labour market versus ensuring an efficient rule to delineate the SLA boundaries
  - If we fail to check the underlying reason that cause a SLA to form in certain way, then realigning the SLA boundary may affect the labour market reality of that particular SLA.

Geographic contiguity

(A) CSDs already assigned to a SLA

SLA containing CSDs that are not adjacent to at least one other CSD that forms the same SLA

- (1) SLA is separated in two or more parts by one or more CSDs that have not been assigned to a SLA
- (2) One or more CSDs forming a SLA(s) is separated by large distances, or by a physical/geographical barrier
- (3) SLA is separated in two or more parts by one or more CSDs that belong to a different SLA

(4) SLA does not conform to the standard geographical boundary

(B) CSDs not assigned to a SLA

- (5) One or more CSDs form a hole(s) inside or between SLAs
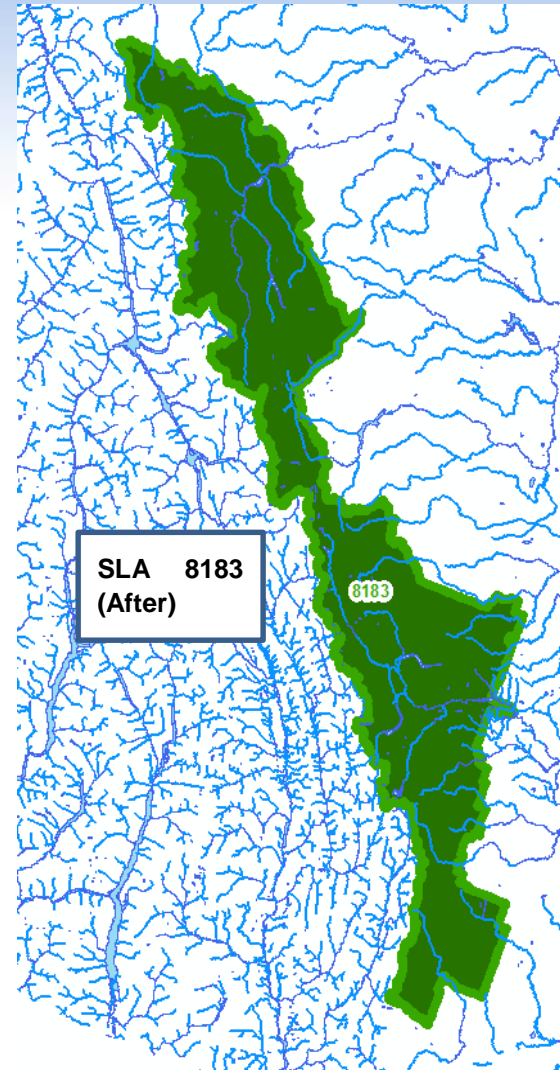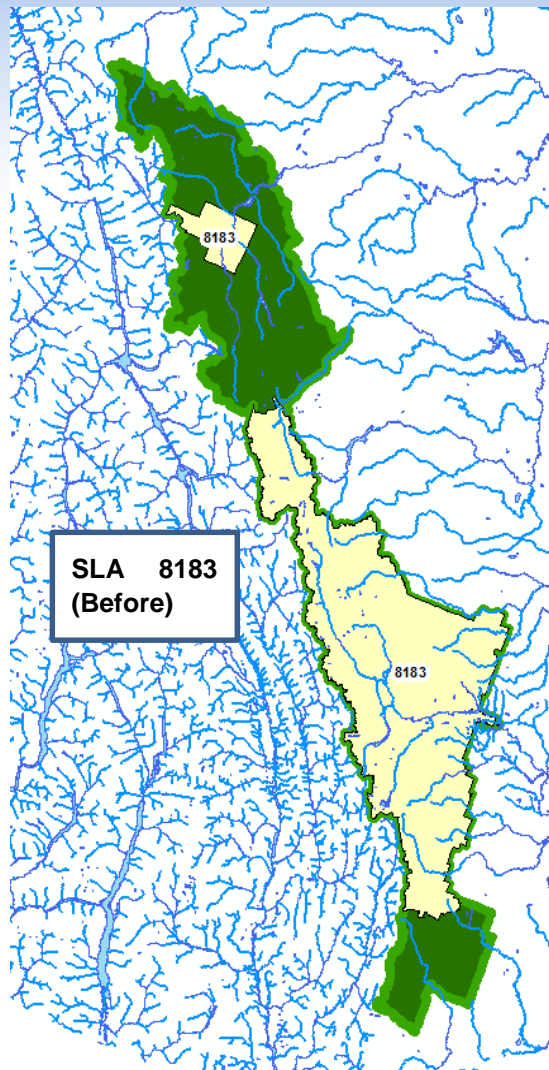- (6) Large area of unassigned CSDs excluded from the SLA geography

Contiguity case example

One or more isolated SLA is surrounded by unassigned CSDs

Contiguity case resolved

One or more isolated SLA is surrounded by unassigned CSDs
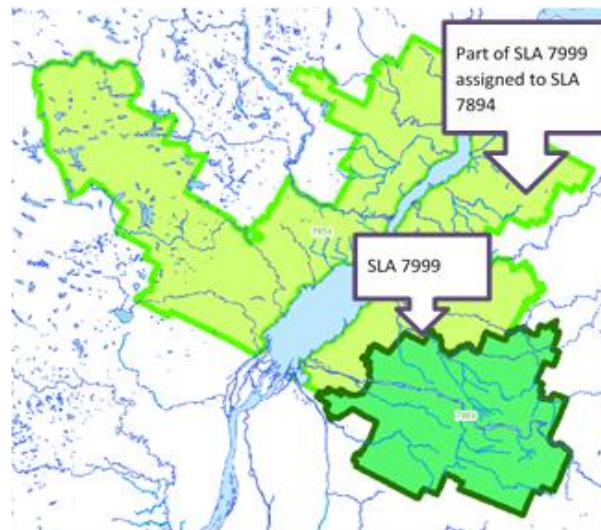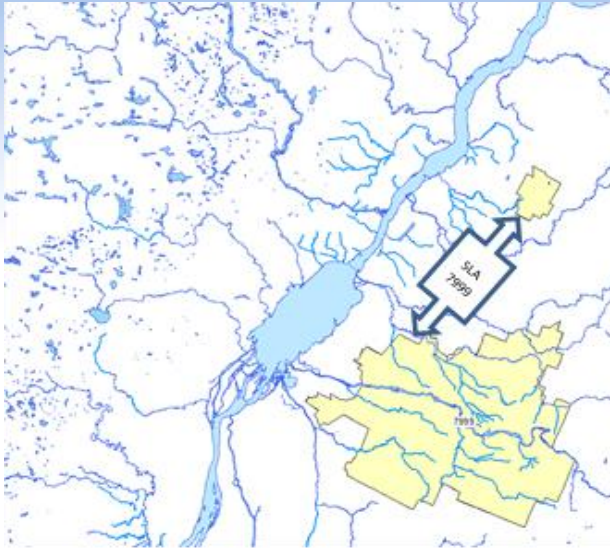

SLA   8183
(Before)

SLA   8183
(After)

Contiguity case example

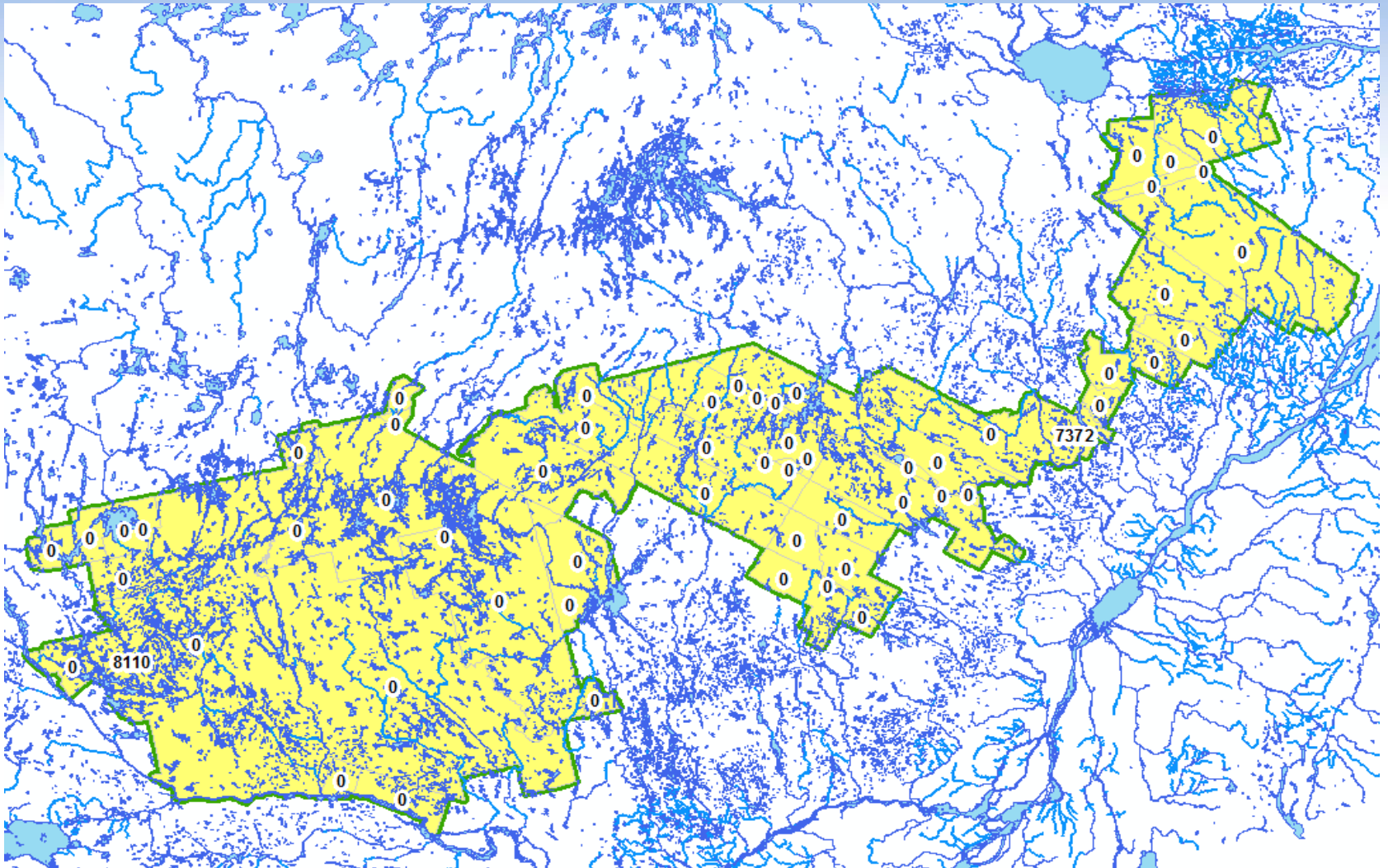Part of a SLA is separated and is surrounded by other SLAs
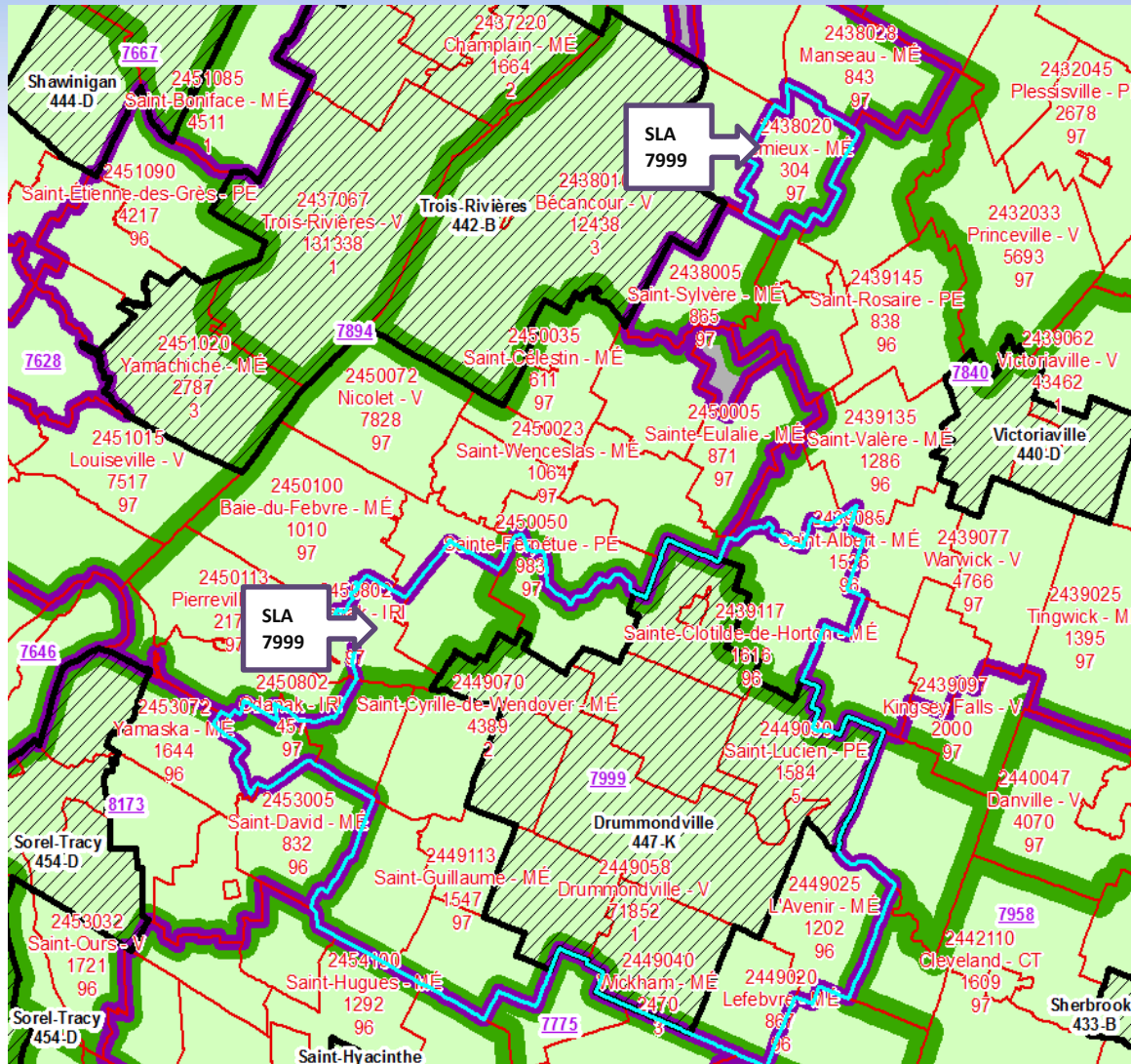
Contiguity case resolved

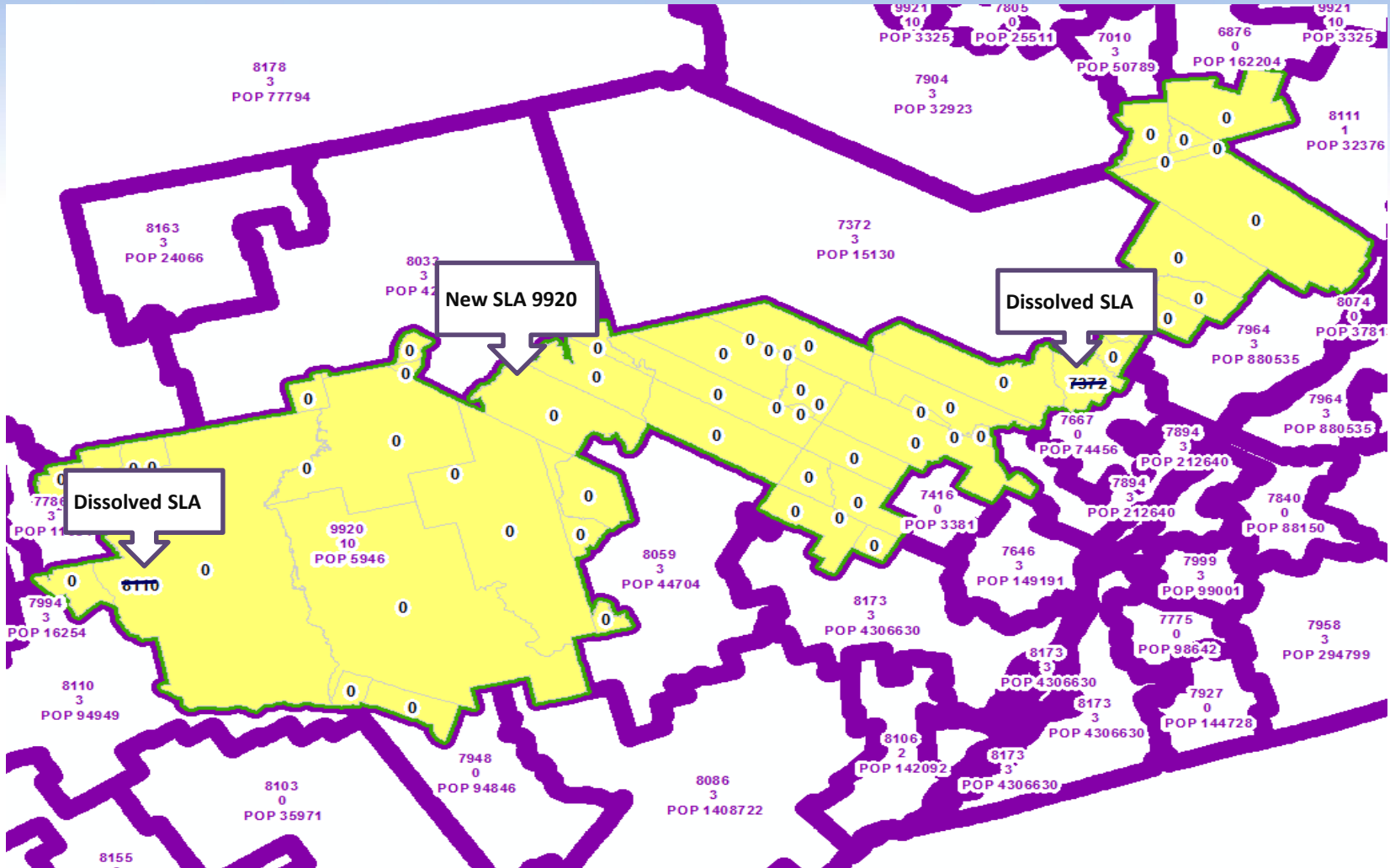Part of a SLA is separated and is surrounded by other SLAs

Contiguity case example

A large area excluded from the SLA geography (CSDs coded as "0")
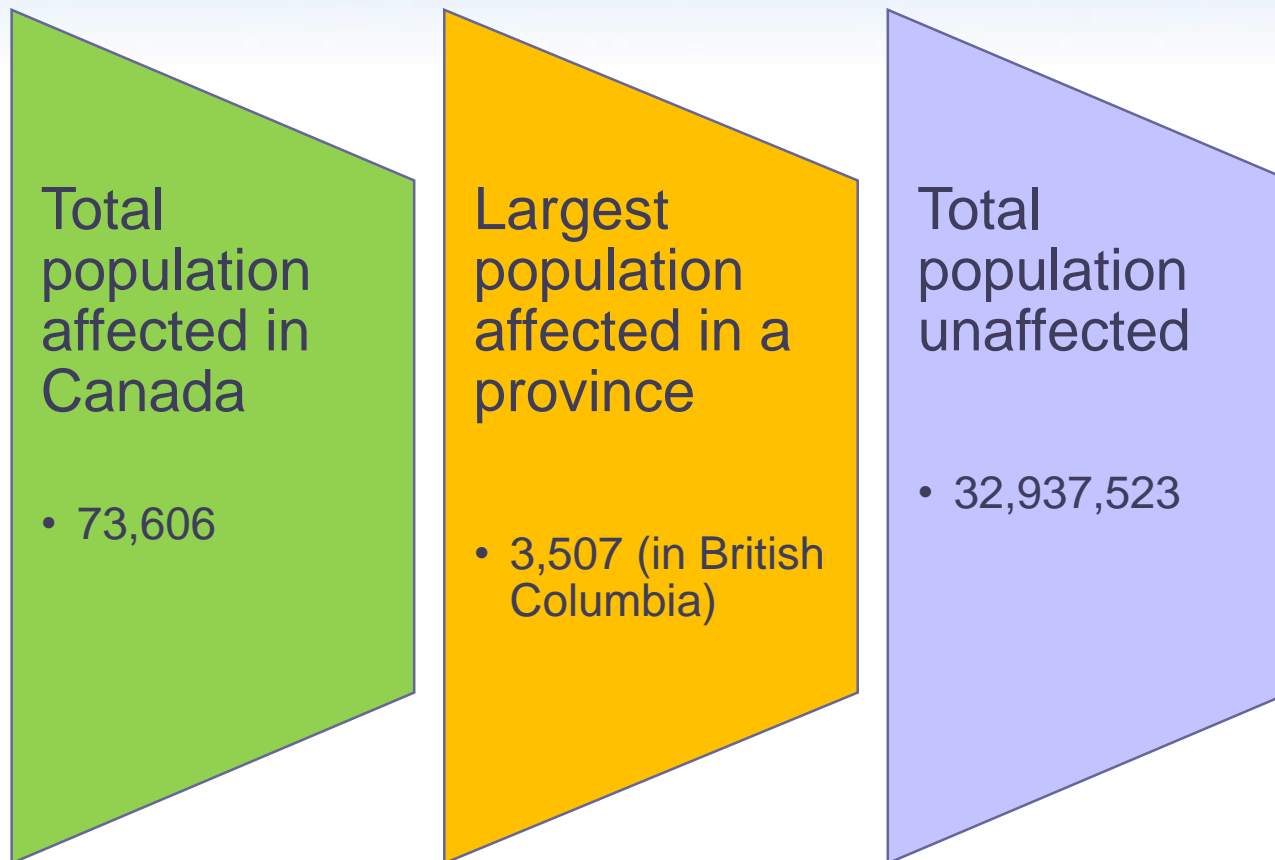
Contiguity case resolved

A large area excluded from the SLA geography (CSDs coded as "0")

Results of reclassification

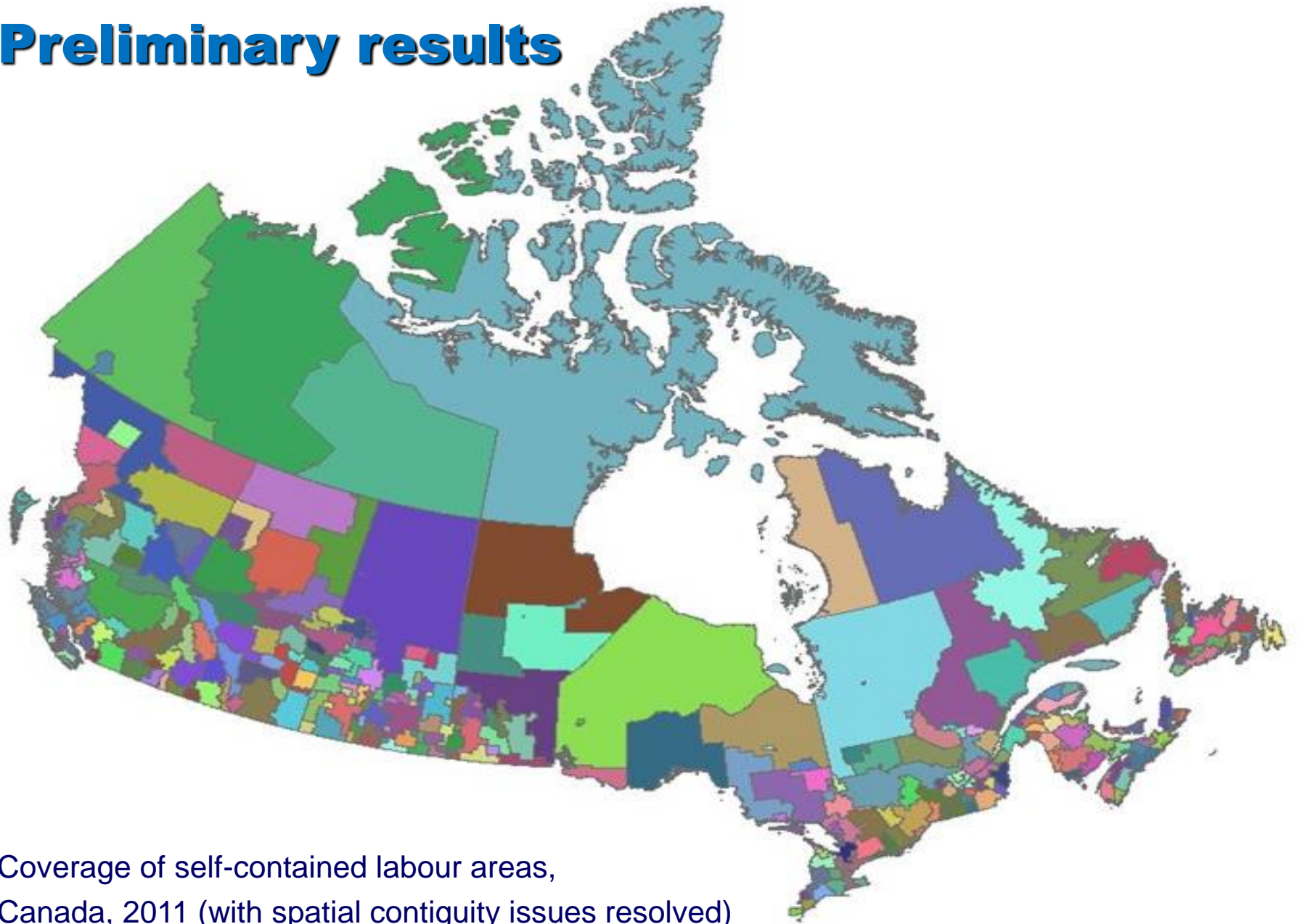- The latest iteration to resolve the contiguity issue in the SLA geography has had little impact on the population

Total population affected in Canada

- 73,606

Largest population affected in a province

- 3,507 (in British Columbia)

Total population unaffected

- 32,937,523

Preliminary results

Coverage of self-contained labour areas,
Canada, 2011 (with spatial contiguity issues resolved)

Preliminary results

	Number of SLAs		SLA Population 2011	
	Count	%	Count	%
Newfoundland and Labrador	36	10%	514,536	2%
Prince Edward Island	4	1%	140,204	<1%
Nova Scotia	11	3%	921,727	3%
New Brunswick	14	4%	751,171	2%
Quebec	61	17%	7,903,001	24%
Ontario	40	11%	12,851,821	38%
Manitoba	24	7%	1,208,268	4%
Saskatchewan	67	19%	1,033,381	3%
Alberta	31	9%	3,645,257	11%
British Columbia	55	16%	4,400,057	13%
Yukon	2	1%	33,897	<1%
Northwest Territories	3	1%	41,462	<1%
Nunavut	1	<1%	31,906	<1%
Total SLAs in all provinces	**349**	**100%**	**33,476,688**	100%
SLAs that span provincial boundaries	13		2,021,529	
Total number of unique SLAs	**336**			

Summary

- Currently: 336 unique self-contained labour areas, which maximize self-containment and provide increased delineation for geographies outside of CMA/CA areas:
    - 204 SLAs achieved self-containment status without any CSDs within a CMA/CA.
    - 108 SLAs contain some CSDs in a CMA/CA and some that fall outside of a CMA/CA
    - 24 SLAs are entirely self-contained within a CMA/CA boundary

- SLA geography was run using both 2006 and 2011 data; the geographies produced were largely coherent between the two time frames

Future improvements

- Merging adjacent SLAs (previously established SLAs)
  - Economic characteristics (industry composition)
  - Connectivity with road network file

- Users feedback/consultations

Questions/Comments