Demolinguistic Data for Indigenous Communities in Canada
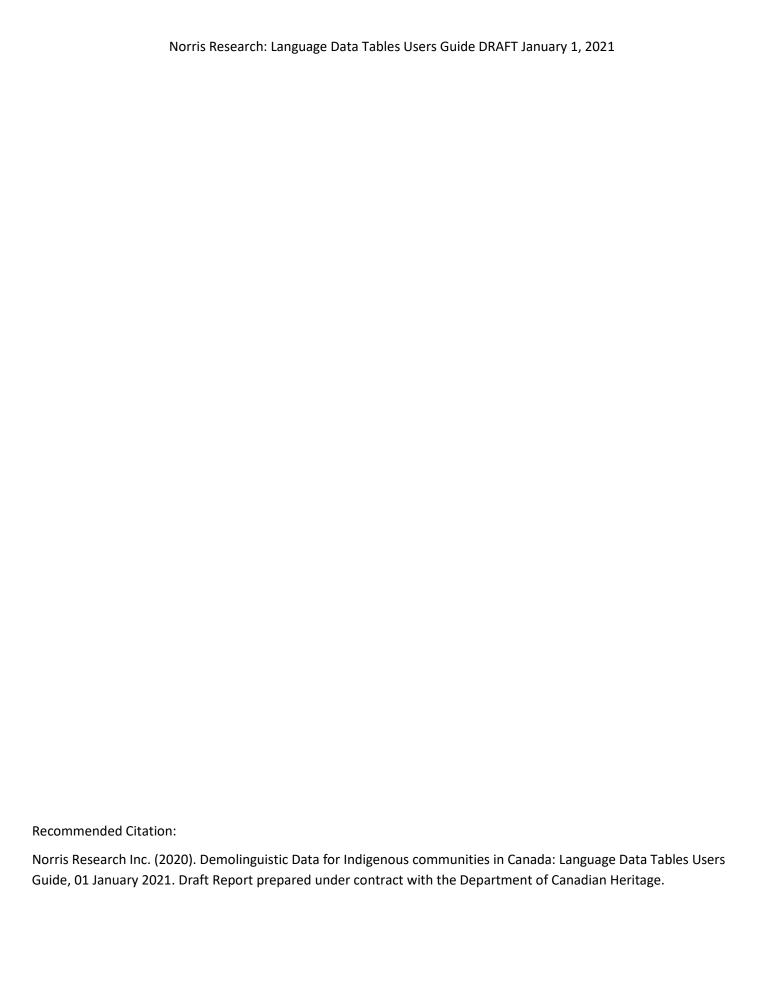
Language Data Tables User Guide

Version 0.7.1

Norris Research Inc.

https://norrisresearch.com/ref_tables.htm

1 January 2021

Table of Contents

!! IMPORTANT !!

A Cautionary Note

This website and the underlying databases constitute a "work in progress". They are built from a disparate collection of data sources, which are full of idiosyncracies. The outputs have not yet been subjected to extensive peer review, and it is probable that some problems have been either:

- not yet documented

    - e.g. problems caused by partial enumeration (see Appendix B)

- not yet detected

    - e.g. the re-assignment of the traditional language of the Dene Tha' First Nation from Dëne Sųłiné [Chipewyan] to Dene-thah [Slavey, South] was only made in 2018, and has had a significant impact on the population counts of those two languages.

- introduced accidentally

    - e.g. StatCan began reporting Ottawa as a separate category only in 2016, and the data manipulation methodology needs to be adjusted in order to correctly assign speakers to the corresponding NRI category.

Please exercise caution when using the data. Any feedback will be gratefully received at norrisinc@videotron.ca

Website Tips and Tricks

The website should be compatible with current versions of all the major web browsers. It is tested with Chrome, Firefox and Microsoft Edge. Note that some features are not compatible with Internet Explorer. The tables and maps are best viewed on a large display (27" or greater).

The website implements three features that are out-of-the-ordinary:

- Tables that can be sorted and filtered
- A tree view of hierarchical data
- KML maps optimised for viewing in the desktop version of Google Earth

The following sectons give some advice on how to take advantage of these capabilities.

Tables

Most of the tables can be sorted and filtered. The best way to demonstrate this capability is by working with a simple example, found here:

> https://norrisresearch.com/ref/csd_types.htm

This table contains counts of the various Census Subdivision (CSD) types for the last four censuses. Each column has a title, and immediately to the right of the title is an icon (⬍). Clicking on this icon will sort the table according to the values in that column. An initial click will sort the data in ascending order. Clicking again will sort the data in descending order. You can test this easily by clicking on the icon for the "Year" column.

Filtering the data in the table is accomplished by typing in one of the boxes in the second row of the table. Test this by typing "ALL" in the box for the "Code" column. The table will be filtered to show just four rows, representing the total number of CSDs for each of the four censuses.

Many of the tables on the website are very large. Judicious use of the sorting and filtering capabilities can give you a more manageable view of the data.

Note that some of the tables contain thousands of rows, and it may take a few seconds (or even minutes with a slow computer or network) for the table to finish loading. You will not be able to sort or filter the table until it has finished loading. You can look at the tab in your browser for a visual clue that the page is still loading. Also, the row containing the filter boxes will not be displayed until the table has been loaded.

Tree View

A tree is a very powerful data structure for displaying hierarchical data sets. It is used here to display the hierarchical classification of indigenous languages in Canada and the USA:

> https://norrisresearch.com/ref/languages_can_usa.tree.htm

The initial view shows just one record for the collapsed tree. Clicking on the "Expand All" button will display the full tree. Branches of the tree can be expanded by clicking on the arrow icon to the left of the label. The icon will change to show that the branch has been expanded, and additional branches or leaves will be shown. Leaves represent the lowest level of each branch. The check box immediately to the left of the label does not currently have any functionality.

Right-clicking on any of the labels will trigger a pop-up box with four rows: "ISO 639-3", "Glottolog", "FPCC" and "Wikipedia". Some of these rows may be greyed-out. The ones which are not greyed-out function as hyperlinks which will take you to the corresponding web page.

The search box allows you to quickly locate any language in the tree. If the branch containing the search term is collapsed, it will be expanded automatically.

KML maps

KML is a data format that is designed for the display of geographic data. There are many software packages that can be used to view KML data. But the KML maps on this website have been optimised for use with the desktop version of Google Earth Pro. This is free software that worls on any ordinary Windows, Mac or Linux desktop computer and can be downloaded from:

https://www.google.com/earth/versions/

If you have no prior experience of Google Earth Pro, you should work through the tutorials provided by Google to familiarise yourself with the basic capabilities of the software.

The KML maps on this website have sophisticated hierarchical structures that can be seen by expanding the tree view in the "Places" pane, in the upper left of the Google Earth window. The search box at the bottom of the "Places" pane is particularly useful for locating geographic features in complex maps. For example, try loading the map of Indian Reserves:

https://norrisresearch.com/kml/ablang/fnp_reserves.kmz

This map contains placemarks for 3326 reserves. Type "pikwakanagan" in the search box at the bottom of the "Places' pane, and the legend will expand and highlight the "Algonquins of Pwakanagan First Nation" row. Double-clicking on this row will cause Google Earth to zoom in on the location of the reserve.

Note that some of the default setting for Google Earth are not ideal for working with these maps. To change the settings, go to Tools|Options, and make the following changes:

- Under the Navigation tab, select "Do not automatically tilt while zooming"
- Under the General tab, check "Allow access to local files and personal data"
- Under the General tab, check "Show web results in external browser"

Additionally, you may wish to toggle off some of the layers in the pane in the lower left of the Google Earth window, to prevent the map from becoming too cluttered.

There are many keyboard shortcuts that one can use for navigating in Google Earth. They can be seen via the Google Earth Help menu. One particularly useful shortcut is "n" which can be used to reorient the map north-south.

About

Purpose

(more…)

The primary goal of this website is to provide high-quality detailed demographic data on the state of Indigenous languages in Canada. The demographic data are summarised in web pages and maps, and can be downloaded as MS Excel spreadsheets and MS Access databases. The maps are presented as KML files that can be downloaded and re-used. All of the data are made available under essentially the same terms as the Government of Canada's Open Data license:

https://open.canada.ca/en/open-government-licence-canada

The primary audience is intended to be policy analysts working in various levels of government and Indigenous organisations. Secondary audiences may include academic researchers and anybody with a general interest in the Indigenous history and language in Canada.

There are numerous other websites that present data that overlap with this one. Some of the more prominent ones are described in a separate section. This one is distinguished from the others in three important aspects:

1. This site includes an abundance of quantitative, well-documented, high-quality, census-derived data.

2. Data are provided at the level of individual communities, and aggregated at a national level.

3. WWW hyperlinks are provided to facilitate comparison with other data sources.

Context

(more…)

In order to interpret the data that are presented here, the user should have a good understanding of:

- Language classification (more…)

- The Canadian Census (more…)

- The historical geography of Indigenous people in Canada (more…)

The website provides some basic material covering these three topics, plus numerous pointers to explore the subjects in more detail.

Terminology

(more…)

In recent years, authors have struggled to use a terminology that is respectful of the Indigenous cultures that they are wrting about. Whilst the general principle of a respectful terminology is almost universally accepted, it can be very hard to apply in practice. Challenges arise for a wide variety of reasons, described in detail elsewhere. In many situations, there is no simple right or wrong terminology. We have attempted to be consistent within this website. We have drawn

attention to some situations where we are uncomfortable or uncertain of our own usage. We expect that our terminology will gradually evolve.

Literature Sources

(more…)

Our classification of Indigenous languages and communities across Canada is based on extensive research in published material. But the amount of material is enormous and frequently contradictory. Much of it is not easily accessible. Our sources are described in more detail elsewhere. Wikipedia serves as our initial starting point for research, but is only used as an authoritative source when other avenues of research have been exhausted.

Language Classification

Introduction

Language classification falls within the general field of "Historical Linguistics", and the narrower field of "Comparative Linguistics". Quoting from Wikipedia:

> Comparative linguistics is a branch of historical linguistics that is concerned with comparing languages in order to establish their historical relatedness. Languages may be related by convergence through borrowing or by genetic descent, thus languages can change and are also able to cross-relate.

> Genetic relatedness implies a common origin or proto-language. Comparative linguistics has the goal of constructing language families, reconstructing proto-languages, and specifying the changes that have resulted in the documented languages.

Amongst historical linguists, there has always been a tension between those who are willing to relate languages based on slim evidence, versus those who require much more compelling evidence. An example of the tension can be found in the writings of Merritt Ruhlen. See especially Chapter ?? in ?? Arguments frequently reduce to disagreements over whether a feature that is common to two languages is due to borrowing or to a common ancestor, or simply coincidental. For non-linguists, the end-result is a confusing set of hierarchical classifications that are similar, but not identical.

Dialects and Languages

The concept of what constitutes a "language" does not have a precise scientific or legal definition. This leads to insurmountable problems when somebody attempts to produce a precise count of the number of languages spoken in Canada. Again quoting from Wikipedia:

> [Dialect] refers to a variety of a language that is a characteristic of a particular group of the language's speakers. Under this definition, the dialects or varieties of a particular language are closely related and, despite their differences, are most often largely mutually intelligible, especially if close to one another on the dialect continuum.

The condition of mutual intelligibility is critically important, but is clearly something that will vary continuously from "perfectly intelligible" to "completely unintelligible". Three languages in Canada – Cree, Ojibwe and Inuktut – are spoken over huge areas, and are subject to significant regional variation, such that mutual intelligibility may be completely lost.

Social and Political Considerations

"A language is a dialect with an army." This oft-quoted phrase neatly sums up the situation for many languages, which are virtually identical from a linguistic perspective, but are considered to be quite distinct by their speakers for various social and political reasons. Norwegian and Swedish make a good example; there are many others. There is little to be gained and much goodwill to be lost by trying to treat the different languages as a single entity.

Language and Genetics

There are strong parallels between language evolution and biological evolution, as summarised in two short papers by Pagel (2017a, 2017b). Combining DNA studies with linguistic studies is likely to reveal a lot about language evolution. These studies have been slow to take off amongst Indigenous populations in North America. Elsewhere in the world, they have resulted in many useful insights. ?? gives a very good high-level overview. ?? gives a recent summary, drawing attention to the lack of work in North America.

Mixed Languages

Sign Languages

Further reading

Wikipedia

https://en.wikipedia.org/wiki/Historical_linguistics#Comparative_linguistics

https://en.wikipedia.org/wiki/Comparative_linguistics

https://en.wikipedia.org/wiki/Dialect

https://en.wikipedia.org/wiki/A_language_is_a_dialect_with_an_army_and_navy#:~:text=%22A%20language%20is%20a%20dialect,of%20a%20language%20or%20dialect.

University-level textbooks

Campbell

Trask/??

Specialised textbooks

Campbell/Poser

Ruhlen

Journal papers

Pagel

The historical geography of Indigenous people in Canada

Introduction

A critical initial step in the data analysis that is presented here is to develop a list of indigenous communtities in Canada. A "community" can be defined in many different ways. It may have a simple correlation with the physical geography. Quoting from www.sociologyguide.com:

> The term community is one of the most elusive and vague in sociology and is by now largely without specific meaning. At the minimum it refers to a collection of people in a geographical area.

This minimal meaning is the one that is implied in most situatiuons within this website. Expanded meanings are sometimes used in relation to Indigenous communities within large urban areas. Our census-based data analysis necessitates a definition of community that is directly linked to Census Subdivisions (CSDs). The precise rules are explained in a separate section. In order to properly understand the current distribution of Indigenous languages in Canada, we need to understand how these communities have evolved.

Most of the Indigenous communities that are identified in this study (?? out of ??) have a one-to-one relationship with Indian Reserves. There are currently over 3,000 Indian Reserves in Canada. Their boundaries are steadily changing. New reserves are being created, and existing ones are changing status. The following sections attempt to place the modern distribution of Indigenous communities in a historical geographic perspective, beginning in the very distant past. The coverage is shallow, but links to further reading are provided to allow a deeper understanding.

The First Americans

The arrival of humans in the Americas remains quite uncertain, and important new discoveries are being steadily reported. DNA analysis is proving to be an extremely powerful new tool for testing and refining our theories. "Modern" humans first appeared in Africa about ?? years ago. Compared to earlier humans, these were probably the first ones to develop sophisticated speech, attributable to a genetic adaptation. Also compared to earlier humans, our ancestors were able to adapt to living in almost all climatic conditions on the planet, and rapidly migrated to places that were previously too inhospitable. One of these locations is the Yana Rhinoceros Horn Site (RHS) in eastern Beringia, where there is very good evidence of human activity from 30,000 years ago.

The Earth has been subjected to several Ice Ages during its history. The most recent one began about 2.5 million years ago, roughly coincident with the initial evolution of the Homo genus (Homo habilis) in Africa about 2 million years ago. The ice has waxed and waned over this time period. The Last Glacial Period (LGP) lasted from about 115,000 to 12,000 years ago. The Last Glacial Maximum (LGM) was the most recent time during the LGP that ice sheets were at their greatest extent. Ice sheets covered most of Canada, reaching as far south as Wisconsin and Massachusetts. The growth of the ice sheets was accompanied by a large drop in sea levels. Growth of ice sheets began about 33,000 years ago and maximum coverage was between 26,500 years and 19,000 years ago. North America and Asia are currently separated from one another by the Bering Strait, which is 80 km wide at its narrowest point. During the LGM, the Bering Strait would have been above sea level. The region around it, extending from the Mackenzie River in the east to the Lena or Kolyma River in the west, is referred to as Beringia. It remained ice-free during the LGM, because of low precipitation.

The "Beringian Standstill Hypothesis" was first proposed in 2007, based on DNA evidence. It proposes that humans migrated into Beringia from other parts of Asia, and became genetically isolated for several thousand years, before

some of them migrated southwards through the Americas. A reasonable scenario that is consistent with the majority of the current archaeological and DNA evidence would be:

1. Humans migrate into Beringia from other parts of Asia about 35,000-30,000 years ago

    a. Evidence from the Yana Rhinoceros Horn Site (RHS)

2. Once in Beringia, they become isolated from other Asian populations

    a. DNA evidence

3. They spread across Beringia, but are prevented from migrating southwards through the Americas by the presence of large ice sheets

    a. Evidence from the Blue Fish Caves in the Yukon

4. As the ice recedes at the end of the LGM, humans rapidly migrate southwards through the Americas

    a. Evidence from multiple sites dated to about 13,000 years ago

Much of the archaeological evidence that could be used to test this scenario was probably destroyed by rising sea levels along the Pacific coast in both Beringia and the Americas.

From a linguistic perspective, this scenario leads to some significant possibilities:

1. If all Indigenous people are descended from a single isolated population in Beringia, they may all share a common ancestral language.

2. Excluding Alaska and the Yukon, the American continents were not populated until about 13,000 years ago. All of the linguistic evolution within the continents has thus taken place over just 13,000 years, although there would have been linguistic evolution within Beringia dating back to about 35,000 years ago.

Whilst the Beringian Standstill Hypothesis, or variants of it, are the most poular explanation of the initial peopling of the Americas, there are other hypotheses. Of these, one of the most prominent postulates arrival from Europe by people traversing the southern edge of the ice sheet in the North Atlantic.

From the LGM to Columbus

The initial populating of the Americas seems to have been a rapid event, roughly 13,000 to 11,000 years ago. Our knowledge of the years between this initial event and the arrival of Europeans in Canada is still sparse, when compared to elsewhere in the world. People migrated into New England from the west at a time when the lands to the north were still ice-covered. As the ice retreated, people migrated northwards, but the migration was complicated by the changing sea levels and rapidly changing drainage networks. The weight of the great ice sheets had pushed the land down. After the ice melted, the land started to slowly rise again, but the melting ice caused sea levels to rise very quickly, resulting in much of the recently deglaciated land to be flooded by sea water. In eastern Canada, the Champlain Sea was formed as an inlet of the Atlantic Ocean. It covered all the land around Montreal and Ottawa to a depth of perhaps 150m. The sea lasted from about 13,000 to 10,000 years ago. Our understanding of the gradual retreat of the ice and the changing drainage patterns is still evolving, but the overall picture is well-known (see the movie at emvc.geol.ucsb.edu). The deglaciation was essentially complete by 5,000 years ago. An important feature is the relatively early emergence of an ice-free corridor from north to south through Alberta.

In the millennia before Columbus's arrival, there was a steady evolution of Indigenous societies in the Americas. Notable societies were the:

- Maya

    o The Mayan civilization was centred on the Yucatan Peninsula in SE Mexico

    o The civilization flourished between 1700 and 1100 years ago

    o The Mayan writing system was highly developed

- Aztec

    o flourished in central Mexico in the 14th and 15th centuries

    o had a simple writing system

- Inca

    o Centred in Peru

    o Flourished between the 12th and 15th centuries

    o Information was recorded on assemblages of knotted strings, known as Quipu

Societies to the north have left far fewer archaeological remains. The Mississippian culture is noteworthy – especially the city of Cahokia, which may have had a population of over 10,000 people. The city was located about 17 km south of the confluence of the Mississippi and Missouri Rivers. The culture flourished between 800 and 600 years ago.

Archaeological evidence in Canada is much more challenging to uncover and interpret than further south. The encyclopaedic work of ?? is perhaps the best summary of our current state of knowledge.

From First European Contact to the Great Peace of 1701

Vikings reached Newfoundland about 1,000 years ago, but had no apparent major impact on the local Indigenous people they encountered. Their presence at L'Anse aux Meadows may have lasted for about 100 years. Christopher Columbus reached the West Indies in 1492, and returned in 1493, 14968 and 1502. John Cabot reached Newfoundland in 1497. These voyages opened the way for European exploration and colonization of the Americas. Basque whalers and fishermen were hunting whales and fishing for cod off Newfoundland within 30 years of Cabot's voyage.

Jacques Cartier sailed to the Gulf of St Lawrence in 1534. He returned in 1535, visiting Stadacona (Quebec City) and reaching Hochelaga on October 2nd. He returned again in 1541, in an attempt to establish a colony. The attempt was unsuccessful and was abandoned in 1543. Pierre de Chauvin de Tonnetuit built a house at Tadoussac in 1600. Samuel de Champlain made his first voyage, of over twenty, across the Atlantic in 159?. In 1603 he travelled to Tadoussac – his first visit to Quebec. In 1608 he founded Quebec City. The colonial population of New France grew very slowly during the 17th century. At the same time, other European countries were establishing colonies all along the Atlantic coast of North America:

- Britain

    o Jamestown, Virginia (1607)

    o Plymouth, Massachusetts (1620)

- o   St John's, Newfoundland (c1620)

- Netherlands

- o   New Amsterdam, New Yoek (1624)

- Spain

- o   St. Augustine, Florida (1565)

Written documentation of the Indigenous people of North America began with these initial explorations and colonisations. The documentation is, of course, entirely one-sided, and needs to be interpreted accordingly.

The fur trade rapidly rose to be of pre-eminent importance to New France and to the English and Dutch colonies further south. This led to competition amongst the European nations and also amongst the Indigenous nations, which in turn led to the so-called Beaver Wars which lasted for several decades until the Great Peace of Montreal was signed in 1701. An important event in these years was the Dutch surrender of its North American territories to Britain in 1674, under the terms of the Treaty of Westminster.

Fur Trade Expansion, 1701-1763

The 18th Century was marked by increasing rivalry and conflict between the British and French in their attemts to control the fur trade. The Hudson's Bay Company was founded in 1670. The British and French fought an endless succession of worldwide wars, which did not finally end until the Battle of Waterloo and the subsequent signing of the 1815 Treaty of Paris on November 20th, 1815. In the context of North America, the most important was the Seven Years War, which ended with the signing on February 10th 1763 of the 1763 Treaty of Paris. New France became a British territory. French colonisation had not proceeded west of Montreal, except for some minor exceptions (??). The British

The Royal Proclamation of 1763 was issued by King George III on October 7, 1763. The Proclamation forbade all settlement west of a line drawn along the Appalachian Mountains. The Royal Proclamation continues to be of legal importance to First Nations in Canada, being the first legal recognition of aboriginal title, rights and freedoms, and is recognized in the Canadian Constitution of 1982. Section 25 of the 1982 Constitution Act reads:

> Aboriginal rights and freedoms not affected by Charter

> 25. The guarantee in this Charter of certain rights and freedoms shall not be construed so as to abrogate or derogate from any aboriginal, treaty or other rights or freedoms that pertain to the aboriginal peoples of Canada including

> (a) any rights or freedoms that have been recognized by the Royal Proclamation of October 7, 1763; and

> (b) any rights or freedoms that now exist by way of land claims agreements or may be so acquired.

Royal Proclamation to Indian Act (1763-1876)

In the century following the Royal Proclamation, the political entity of modern-day Canada began to coalesce. The American Revolutionary War (1775-1783) and the subsequent westward territorial expansion of the USA was a major

controlling factor. The British North America Act of 1867 created the fundamental political structure of the country. The fur trade thrived across Canada until the middle of the 19th century, after which it went into a long, steady decline. The Hudson's Bay Company merged with the North West Company in 1821, and dominated the fur trade business from that point forward.

WWW links

https://en.wikipedia.org/wiki/Community

https://www.sociologygroup.com/community-meaning/

https://www.sociologyguide.com/basic-concepts/Community.php

https://en.wikipedia.org/wiki/Ice_age

https://en.wikipedia.org/wiki/Last_Glacial_Period

https://en.wikipedia.org/wiki/Last_Glacial_Maximum

https://science.sciencemag.org/content/325/5941/710

https://en.wikipedia.org/wiki/Bering_Strait

https://en.wikipedia.org/wiki/Beringia

https://www.ncbi.nlm.nih.gov/pmc/articles/PMC1952074/

https://en.wikipedia.org/wiki/Post-glacial_rebound

https://en.wikipedia.org/wiki/Champlain_Sea

http://emvc.geol.ucsb.edu/2_infopgs/IP2IceAge/aDeglacNoAm.html

https://en.wikipedia.org/wiki/Aztecs

https://en.wikipedia.org/wiki/Maya_civilization

https://en.wikipedia.org/wiki/Inca_Empire

https://en.wikipedia.org/wiki/Mississippian_culture

https://en.wikipedia.org/wiki/L%27Anse_aux_Meadows

https://en.wikipedia.org/wiki/Christopher_Columbus

https://en.wikipedia.org/wiki/John_Cabot

https://en.wikipedia.org/wiki/History_of_Basque_whaling

https://en.wikipedia.org/wiki/Algonquian%E2%80%93Basque_pidgin

https://en.wikipedia.org/wiki/Pierre_de_Chauvin_de_Tonnetuit

http://www.biographi.ca/en/bio/chauvin_de_tonnetuit_pierre_de_1E.html

https://en.wikipedia.org/wiki/Tadoussac

https://en.wikipedia.org/wiki/Samuel_de_Champlain

https://en.wikipedia.org/wiki/Jamestown,_Virginia

https://en.wikipedia.org/wiki/Plymouth_Colony

https://en.wikipedia.org/wiki/New_Amsterdam

https://en.wikipedia.org/wiki/St._Augustine,_Florida

https://en.wikipedia.org/wiki/North_American_fur_trade

https://en.wikipedia.org/wiki/North_American_fur_trade

https://en.wikipedia.org/wiki/Beaver_Wars

https://en.wikipedia.org/wiki/Great_Peace_of_Montreal

https://en.wikipedia.org/wiki/Treaty_of_Westminster_(1674)

https://en.wikipedia.org/wiki/Hudson%27s_Bay_Company

https://en.wikipedia.org/wiki/Treaty_of_Paris_(1815)

https://en.wikipedia.org/wiki/Seven_Years%27_War

https://en.wikipedia.org/wiki/Treaty_of_Paris_(1763)

https://en.wikipedia.org/wiki/Royal_Proclamation_of_1763

https://laws-lois.justice.gc.ca/eng/const/page-15.html

https://en.wikipedia.org/wiki/Indian_Act

https://en.wikipedia.org/wiki/Constitution_Act,_1867

https://en.wikipedia.org/wiki/Hudson%27s_Bay_Company

The Canadian Census

Introduction

The data that are presented here are all derived from the Canadian Census. In order to properly understand and interpret the data, a good general understanding of the Census is required, plus a detailed understanding of a few specific topics. This section provides the basic information. Statistics Canada produces a large number of publications which describe the Census in great detail.

Historical Overview

Governments have been carrying out censuses for millennia. Most famously, the Christmas story is driven by a census of the Roman Empire. Jean Talon was responsible for organising the first Canadian census in 1666. The first detailed modern census of Canada was carried out in 1851-52. National censuses have been carried out continuously since then. From 1851 to 1951, they were held every 10 years. From 1951 to the present, they have been held every 5 years. 2011 was a special case, as described in detail below.

Language Questions

Random rounding and data suppression

Short-form and long-form questionaires

Population universes

Geographical Area Classification

Evolving geography

Evolving language classification

2011 National Household Survey

WWW links

https://www150.statcan.gc.ca/n1/pub/98-187-x/4064814-eng.htm

https://www12.statcan.gc.ca/english/census01/Info/history.cfm

Literature Sources

The website uses a linguistic classification and identification of Indigenous communities that is based on an extensive review of the literature. But the literature is enormous, and it is impossible to comprehensively examine it all. Accordingly, we have had to be selective in what we examined, and this will undoubtedly have led to errors. Additionally, the literature is full of contradictions and inconsistent terminology. It is important, therefore, to understand how we approached the literature review, and how we may have missed important works. The types of sources can be somewhat arbitrarily split into the following categories:

- Wikipedia, and other web-based resources
- Scholarly encyclopaedic works
- Encyclopaedic works aimed at a general audience
- A
- B
- C
- D
- E
- F

We can also classify the literature in terms of the primary focus of each document:

- Linguistic
- Cultural
- Historical / Geographical
- A
- B
- C

Wikipedia, and other web-based resources

These sources differ from the other ones listed, in that they tend to be ephemeral, anonymous and poorly referenced. Notwithstanding these problems, they can be important jumping-off points for further research. For many indigenous communities they are the only readily-available sources of information.

Notes on the hierarchical classification

Background

The first step in deriving the demolinguistic data that are presented on the website was to establish a classification system for indigenous languages in Canada that worked well for our needs and was compatible with other classifications. The rationale behind the classification is explained in detail elsewhere. In this section, we will focus on the hierarchical aspect of the system.

The hierarchy is an organisational tree, **not** an evolutionary tree. Its primary objective is to facilitate correlation between the various classifications that can be found in the literature. All of these classifications are hierarchical to varying degrees and they are all broadly in agreement with one another. But they differ in innumerable details. Correlating one scheme with another can be confusing.

The present classification, which recognises 107 mutually exclusive "languages/dialects", aims to be comprehensive in its coverage of all the Indigenous languages that have been documented in Canada. Every Indigenous person should be able to identify a node amongst the 107 that corresponds to the language they speak, or their ancestors spoke. There will be disagreements related to terminology and dialects, but hopefully no group is unrepresented.

Linguistic hierarchical classifications have a lot in common with biological classifications. Pagel 2017a, 2017b) gives a good introduction to their similarities. In an ideal world, both the linguistic and biological trees would correlate directly with the evolution of languages and species from a common ancestor. In biology, the "last universal common ancestor" is the most recent population of organisms from which all organisms now living on Earth have a common descent. A similar concept can be applied to languages, but is much harder to develop and test. In the real world, there are many languages and species which cannot be assigned to precise locations in an evolutionary tree. In biology, these problems arise most notably with fossils that have no close living relatives, and DNA analysis is not possible. In linguistics, languages that have no close living relatives are generally classed as "isolates".

There is considerable evidence that the Indigenous people of the Americas have a common ancestral population that was already present in Beringia about 30,000 years ago. In an ideal world, it may be possible to relate all of the Indigenous languages of the Americas back to a "proto-American" language that was being spoken in Beringia 30,000 years ago. This ideality is a long way from present-day reality. Nevertheless, Greenberg and Ruhlen (2007, figure 2, page 6) have postulated a phylogenetic structure of the "Amerind" family that encompasses all of the Indigenous languages of the Americas except for the Na-Dene and Eskimo/Aleut families.

Language Families

Despite the difficulties involved in classifying languages, linguists are in general agreement with respect to several large families that collectively encompass almost all of the languages spoken in Canada. These families are:

- Algonquian

- Athapaskan

- Inuit

- Iroquoian

- Salish

- Siouan

- Tsimshian

- Wakashan

The only living languages that cannot be easily placed into one of these families, excluding the special case of Michif, are Haida, Kutenai and Tlingit. This grouping of language families corresponds to the classification used by Statistics Canada for the Canadian Census.

Linguists and archaeologists are in agreement that the Inuit and their language are relatively recent arrivals to Canada – perhaps arriving about 1,000 years ago.

The Athapaskan-speaking people are thought to have arrived in Canada sometime before the Inuit, but well after the initial populating of the Americas.

Tlingit is probably related to the Athapaskan languages. Haida **may** be related to the Athapaskan family, and Kutenai **may** be related to the Algonquian family.

Scope of the hierarchy

Given the above notes and caveats, the hierarchical classification presented at https://norrisresearch.com/ref/languages_can_usa.tree.htm was produced.

The tree is intended to be comprehensive for the eight language families listed above. All of these families except Tsimshian have members in the USA. Hence, the map of the language locations covers much of the USA. The tree is **not** intended to be comprehensive for all of the languages spoken in the USA. The tree contains a total of 232 child nodes, each of which has been mapped to a single precise location. For the 107 languages connected to Canada, all of these locations have been accurately located, not necessarily within Canada. Census data have been used to identify the communities with the largest number of mother tongue speakers. Dormant or extinct languages have been located based on historical or archaeological grounds. For example, Beothuk is located at the "Beothuk Interpretation Centre", which is located a short distance away from an important Beothuk archaeological site.

Domain, Kingdom, Phylum, Class, Order, Family, Genus, Species

Wikipedia Links

https://en.wikipedia.org/wiki/Taxonomy_(biology)

https://en.wikipedia.org/wiki/Most_recent_common_ancestor

https://en.wikipedia.org/wiki/Last_universal_common_ancestor

https://en.wikipedia.org/wiki/Timeline_of_human_evolution

https://en.wikipedia.org/wiki/Settlement_of_the_Americas

Introduction

The web page at https://norrisresearch.com/ref_tables.htm contains links to several tables of data on the indigenous languages of Canada. The data are extracted from a collection of MS Access databases and are presented both as web page tables and MS Excel files. The MS Excel files generally contain more columns than the web page tables. This document explains the content of each of these tables, and also includes notes explaining how the tables were created.

List of tables

N	Table	Row count
1	Indigenous Languages of Canada and the USA	364
2	Indigenous Languages of Canada (NRI data)	
3	Indigenous Languages Cross-Reference : NRI - Statistics Canada	
4	Census Subdivision Types	
5	Census Subdivisions	
6	Census Languages (2001-2016)	
7	"Aboriginal" CSDs with NRI language	
8	CSDs 2001-2016; with mother tongue speakers	
9	CSDs 2001-2016; with mother tongue speakers (extended)	
10	CSD Summary counts by language, Province, and census year	
11	Reserves	
12	First Nations	
13	Tribal Councils	
14	Canada Lands	
15	Friendship Centres	
16	Distinctions	
17	BC Fishing Reserves (in prep)	
18	Historical Atlas Reserves to 1900 (in prep)	
19	BC Reserves, 1943 (in prep)	
20	Specific Claims (in prep)	

Descriptions of the tables

For each table, the following characteristics are given:

1) Data Source:

   a) MS Access table name

   b) MS Access database containing the table

2) Description

   a) A short explanation of the table

3) Columns

   a) A list of all of the columns in the table, with an explanation of each one

Indigenous Languages of Canada and the USA

https://norrisresearch.com/ref/languages_can_usa.htm

Data Source

Database: db_ab_lang.accdb

Table: Languages_Can_USA

Description

This table presents a hierarchical classification of the indigenous languages of Canada and the USA. The traditional territories of many indigenous languages span the Canada/USA border. Many language families are spread across two or more distincts regions. Many speakers were severely displaced in the years following European colonisation, making it difficult to assign their languages to "traditional territories".

The hierarchy follows closely the hierarchies of Glottolog and Ethnologue, but does not match exactly to either one. Deviations were introduced for a variety of reasons, which are discussed in the NRI Language Classification Report (4th edition in preparation).

Columns

N	www	Column Name	Column Type	Description
1		Language_Name	Short Text (60)	This column contains a simple name for the language, using only non-accented Roman letters. It does not necessarily correspond to the preferred name for the language.
2		Language_Name_Pref	Short Text (60)	This column contains the "preferred" name for the language. Most preferred names are autonyms, and often include special non-Roman characters. This has the unfortunate effect of making them hard to type and easy to mis-spell. In many cases, the preferred name is either unknown or it is the same as the simple name; in these cases, the preferred name is arbitrarily set equal to the simple name.
3		Language_Names_Alt	Short Text (100)	This column contains one or more alternate names for the language
4	✓	Language_Name_List	Short Text (120)	This column is populated by either • using the value of Language_Name_Pref • combining the Language_Name_Pref and Language_Names_Alt columns where Language_Names_Alt is not null The resultant value is a single text field which facilitates easy searching for a language, using names that are familiar to most users.
5	✓	Dialects	Long Text	Many languages have been split into dialects. Some of the most commonly-seen dialects are summarised in this column.
6		ISO_639_3_Code	Short Text (3)	The ISO 639-3 code list assigns a three letter code to every language that is recognised by the Ethnologue. The Ethnologue language classification is steadily evolving, and so this code list is also changing. A recent change in Canada saw the retirement of the Nootka (noo) code, and its replacement by codes for Nuu-Chah-Nulth (nuk) and Ditidaht (ddt).
7		Glottolog_Code	Short Text (8)	Similar to the ISO 639-3 code maintained by the Ehnologue, Glottolog assigns an eight character code to each language in its database
8		FPCC_Name	Short Text (20)	This is an ASCII version of the language autonym. It is used to construct a hyperlink to the corresponding page on the FPCC website.
9		href_Wikipedia	Short Text (150)	A link is given to a Wikipedia page, if one is available

10		href_ISO_639_3	Short Text (75)	If the language has an ISO 639-3 code, then this column gives a link to the corresponding page on the SIL (Ethnologue) website.
11		href_Glottolog	Short Text (75)	If the language has a Glottolog code, then this column gives a link to the corresponding page on the Glottolog website.
12		href_FPCC	Short Text (75)	If the language has an FPCC web page, then this column gives a link to the page.
13	✓	Link_Wikipedia	Short Text (180)	A full "<a href ...> html element containing the Wikipedia link
14	✓	Link_ISO_639_3	Short Text (110)	A full "<a href ...> html element containing the ISO 639-3 link
15	✓	Link_Glottolog	Short Text (110)	A full "<a href ...> html element containing the Glottolog link
16	✓	Link_FPCC	Short Text (110)	A full "<a href ...> html element containing the FPCC link
17		Location	Short Text (60)	A place name that is used to provide a point location for the language. Table Place_Names (maintained in MS word document language_data_tables.docx and imported into MS Access database db_ab_lang.accdb) is used to maintain the master copy of lat/long coordinates of these places.
18	✓	Latitude	Decimal	The latitude used to locate the language (based on data in table Language_UNESCO_2016, unless overriden by data in table Place_Names). Generally corresponds to the point location of the CSD with the largest mother tongue population.
19	✓	Longitude	Decimal	The longitude used to locate the language (based on data in table Language_UNESCO_2016, unless overriden by data in table Place_Names). Generally corresponds to the point location of the CSD with the largest mother tongue population.
20	✓	Language_Key	Short Text (43)	A hierarchical code, used to construct the hierarchical tree of languages
21		Lang_Flag	Long Integer	A flag to indicate whether the node in the tree has children (= 0) or not (= 1)
22	✓	Lang_Class	Short Text (1)	A flag to indicate whether the node in the tree has children (= F (family)) or not (= S (single))
23		flag_Glottolog	Long Integer	A flag indicating whether the Glottolog data includes a lat/long location for the language. Note that Glottolog locations are often very imprecise (see table Glottolog_Links (maintained in MS Word document language_data_tables.docx and imported into MS Access database db_ab_lang.accdb)
24		flag_UNESCO	Long Integer	A flag indicating that the UNESCO_2016 table includes a precise location for the data (maintained in MS Word document language_data_tables.docx and imported into MS Access database db_ab_lang.accdb)
25		flag_PlaceName	Long Integer	A flag indicating that the location of the language is taken from table Place_Names (maintained in MS Word document language_data_tables.docx and imported into MS Access database db_ab_lang.accdb).
26	✓	Language_Status	Short Text (25)	A tect field, identifying whether the language is Living, Secondarily Survivng, Dormant, or Extinct. It is populated only for the 107 languages that lie within the NRI "Canadian" subset.
27	✓	flag_Canada	Long Integer	A flag that indicates whether the language is one of the 107 languages that are within the NRI subset.

Indigenous Languages of Canada (NRI data)
https://norrisresearch.com/ref/lang_nri_mt.htm

Data Source

Database: db_csd_lang.accdb

Table: lang_nri_mt

Description

The NRI classification includes 107 "language categories" for Canada, including several that are extinct or dormant. This table contains data for each of these languages for both the 2011 and 2016 censuses, for a total of 214 rows. The population and average age data are obtained using the methodology that is described in the Methodology section of this guide. The status and vitality of each language is derived from the population and average age data.

Columns

N	www	Column Name	Column Type	Description
1	✓	Census_Year	Long Integer	
2		Lang_Name_NRI	Short Text (60)	
3		Lang_Name_Pref	Short Text (60)	
4		Lang_Names_Alt	Short Text (100)	
5	✓	Lang_Name_List	Short Text (120)	
6	✓	Lang_Family	Short Text (12)	
7	✓	Pop_Can_NRI	Long Integer	
8		Pop_Can_NRI_txt	Short Text (10)	
9		Pop_Can_NRI_txt2	Short Text (15)	
10		Pop_CSD_NRI	Long Integer	
11	✓	AvAge	Decimal	
12		AvAge_Calc	Decimal	
13		AvAge_StatCan	Decimal	
14		flag_ExactMatch	Long Integer	
15	✓	Language_Status	Short Text (25)	
16	✓	Language_Endangerment	Short Text (30)	
17	✓	Endangerment_Level	Integer	
18	✓	Province_Code	Short Text (2)	
19	✓	Latitude	Decimal	
20	✓	Longitude	Decimal	

Indigenous Languages Cross-Reference: NRI - Statistics Canada

https://norrisresearch.com/ref/lang_correspondences.htm

Data Source

Database: db_csd_lang.accdb

Table: Lang_Correspondences

Description

The classification of Indigenous languages used in the Census has steadily grown more comprehensive over the years. This table tracks the changes from 2001 through to 2016, by correlating each of the four Census classifications with the NRI classification. The table contains 108 records. This number is one more than the number of languages in the NRI classification because of a quirk in the 2001 and 2006 classifications, which counted "Dene" and "Chipewyan" as separate languages.

Columns

N	www	Column Name	Column Type	Description
1	✓	Language_Name	Short Text (60)	
2	✓	SubCode	Short Text (1)	
3		Province_Code	Short Text (2)	
4	✓	StatCan_Family	Short Text (12)	
5	✓	Language_Status	Short Text (25)	
6	✓	Lang_2001_Name	Short Text (30)	
7		Lang_2001_ID	Integer	
8	✓	Lang_2006_Name	Short Text (30)	
9		Lang_2006_ID	Integer	
10	✓	Lang_2011_Name	Short Text (30)	
11		Lang_2011_ID	Integer	
12	✓	Lang_2016_Name	Short Text (60)	
13		Lang_2016_ID	Integer	

Census Subdivision Types

https://norrisresearch.com/ref/csd_types.htm

Data Source

Database: db_statcan_lang.accdb

Table: CSD_types

Description

Each Census Subdivision is assigned to a particular "type", which evolves steadily from one census to the next. Appendix A presents information which describes the different CSD types, and how they have evolved. The DIAND/INAC definition of "on-reserve" is based more-or-less on CSD types. Many CSD types are specific to a particular province or territory. This table lists all of the CSD types broken down by Census Year and Province/Territory. By filtering the *Code* column on the value "ALL", you can get a quick summation of all of the CSDs for each Census Year.

Columns

N	www	Column Name	Column Type	Description
1	✓	Census_Year	Integer	
2	✓	csdtype	Short Text (3)	
3	✓	csdtypeName	Short Text (75)	
4	✓	Canada	Integer	
5	✓	NL	Integer	
6	✓	PE	Integer	
7	✓	NS	Integer	
8	✓	NB	Integer	
9	✓	QC	Integer	
10	✓	ON	Integer	
11	✓	MB	Integer	
12	✓	SK	Integer	
13	✓	AB	Integer	
14	✓	BC	Integer	
15	✓	YT	Integer	
16	✓	NT	Integer	
17	✓	NU	Integer	

Census Subdivisions

https://norrisresearch.com/ref/csds_2011.htm

Data Source

Database: db_statcan_lang.accdb; db_csd_lang.accdb

Table: CSDs_2011

Description

This table lists all 5253 Census Subdivisions for the 2011 Census. It also lists an additional 69 CSDs that were created for the 2016 Census. The table includes links to the corresponding Census Profile page for each CSD – for the 2011 Census in most cases; for the 2016 Census for the 69 additional CSDs. There is also a link to a KML map showing the extent of the CSD.

Columns

N	www	Column Name	Column Type	Description
1	✓	csdUid	Long Integer	
2		prUid	Integer	
3	✓	ProvinceCode	Short Text (2)	
4	✓	csdName	Short Text (60)	
5	✓	csdType	Short Text (3)	
6		cmaUid	Long Integer	
7		cmaname	Short Text (75)	
8	✓	cmatype	Short Text (1)	
9	✓	sactype	Integer	
10	✓	Ab_Flag	Short Text (1)	
11	✓	Flag_2016_CSD	Integer	
12		Flag_Add_CSD	Integer	
13	✓	Pct_Ab_2016	Decimal	
14	✓	Pct_Ab_2011	Decimal	
15	✓	Pct_Ab_2006	Decimal	
16	✓	Pct_Ab_2001	Decimal	
17		Pop_Tot_2016	Long Integer	
18		AvAge_Tot_2016	Double	
19		Pop_Ab_2016	Long Integer	
20		AvAge_Ab_2016	Double	
21		Pop_Tot_2011	Long Integer	
22		AvAge_Tot_2011	Double	
23		Pop_Ab_2011	Long Integer	
24		AvAge_Ab_2011	Double	

25		Pop_Tot_2006	Long Integer	
26		AvAge_Tot_2006	Double	
27		Pop_Ab_2006	Long Integer	
28		AvAge_Ab_2006	Double	
29		Pop_Tot_2001	Long Integer	
30		AvAge_Tot_2001	Double	
31		Pop_Ab_2001	Long Integer	
32		AvAge_Ab_2001	Double	
33		Pop_Ab_MT_2016	Long Integer	
34		Latitude	Decimal	
35		Longitude	Decimal	
36	✓	GE_link	Short Text (7)	

Census Languages (2001-2016)

https://norrisresearch.com/ref/lang_census_mt.htm

Data Source

Database: db_csd_lang.accdb

Table: lang_census_mt

Description

This table shows the classification of Indigenous languages that StatCan has used for the last four Census Years. It revelas how the classification has grown steadily more detailed:

- 2001:   44

- 2006:   46

- 2011   76

- 2016:   86

Also see Appendix C, for a look at the 1971-1981 classifications.

The classification has also evolved in terminology. For example, "Micmac" was replaced by Mi'kmaq in 2006.

Note that the table is not yet fully populated with population counts for the 2001 and 2006 censuses.

Columns

N	www	Column Name	Column Type	Description
1	✓	Census_Year	Integer	
2	✓	Lang_ID	Long Integer	
3	✓	Lang_Name_STC	Short Text (60)	
4		Lang_Name_Group_STC	Short Text (60)	
5		Group_Code	Short Text (4)	
6		Lang_Name_NRI	Short Text (60)	
7		flag_ExactMatch	Long Integer	
8	✓	AvAge	Decimal	
9	✓	Pop	Long Integer	
10	✓	Pop_PR	Long Integer	
11	✓	Pop_CD	Long Integer	
12	✓	Pop_CSD	Long Integer	
13		Pop_CSD_Adj	Long Integer	
14	✓	Pop_AbCSD	Long Integer	
15	✓	Pop_Pct_Diff_1	Decimal	

16	✓	Pop_Pct_Diff_2	Decimal	

"Aboriginal" CSDs with NRI language
https://norrisresearch.com/ref/csd_lang_nri_mt.htm

Data Source

Database: db_csd_lang.accdb

Table: csd_lang_nri_mt

Description

This table summarises the number of mother tongue speakers for the "traditional" languages of each of the 1401 "Indigenous" CSDs for the 2011 and 2016 censuses. It is not yet populated with data for 2001 and 2006.

Columns

N	www	Column Name	Column Type	Description
1	✓	Census_Year	Integer	
2	✓	csdUid	Long Integer	
3		prUid	Integer	
4	✓	ProvinceCode	Short Text (2)	
5	✓	csdName	Short Text (60)	
6	✓	csdType	Short Text (3)	
7		Reserve_Number	Long Integer	
8		Reserve_Name	Short Text (75)	
9		FirstNation_Number	Long Integer	
10		FirstNation_Name	Short Text (60)	
11	✓	Latitude	Decimal	
12	✓	Longitude	Decimal	
13		GE_link	Short Text (7)	
14		hlink_statcan	Long Text (300)	
15		hlink_boundary	Short Text (100)	
16		Lang_Name_NRI	Short Text (60)	
17		Lang_Name_Pref	Short Text (60)	
18		Lang_Name_Alt	Short Text (100)	
19	✓	Lang_Name_List	Short Text (120)	
20		Lang_Name_Idx	Integer	
21		Language_Status	Short Text (25)	
22	✓	Pop_Total	Long Integer	
23	✓	MT_Pop	Long Integer	
24	✓	MT_AvAge	Decimal	
25		flag_TradLang	Integer	
26		txt_TradLang	Short Text (1)	

27		flag_MatchCase	Integer	

CSDs 2001-2016; with mother tongue speakers

https://norrisresearch.com/ref/csd_lang_stc_mt.htm

Data Source

Database: db_csd_lang.accdb

Table: csd_lang_stc_mt

Description

This table lists all of the CSDs that reported Indigenous language mother tongue speakers in the four censuses from 2001 to 2016. The language listed in the table is the one reported by StatCan. Population and average age data are also included. There is a link to the StatCan Census Profile for the corresponding census.

Columns

N	www	Column Name	Column Type	Description
1	✓	Census_Year	Long Integer	
2	✓	csdUid	Long Integer	
3		subID	Long Integer	
4		prUid	Integer	
5	✓	ProvinceCode	Short Text (2)	
6	✓	csdName	Short Text (60)	
7	✓	csdType	Short Text (3)	
8		csdStatus	Short Text (10)	
9		flag_Oddball	Integer	
10	✓	Ab_Flag	Short Text (3)	
11	✓	CSD_2016_Flag	Integer	
12	✓	Latitude	Decimal	
13	✓	Longitude	Decimal	
14		CSD_link	Short Text (7)	
15	✓	hlink_statcan	Long Text (300)	
16		hlink_boundary	Short Text (1000	
17	✓	Language_Name	Short Text (60)	
18		Lang_ID	Integer	
19	✓	Pop	Long Integer	
20	✓	AvAge	Double	

CSDs 2001-2016; with mother tongue speakers (extended)

https://norrisresearch.com/ref/csd_lang_stc_mt_grouped.htm

Data Source

Database: db_csd_lang.accdb

Table: csd_lang_stc_mt_grouped

Description

This table is an extended version of the preceding table. It contains an extra column identifying the CMA that corresponds to a particular CSD, which can be useful for filtering the data. It also includes extra rows corresponding to higher levels in the StatCan language hierarchy. For example it contains records for "Algonquian languages", whereas the previous table contains records for only the lowest levels of the StatCan hierarchy (e.g. Ojibway or Micmac). As a result of adding these extra rows, the html table is very big (13125 records) and may therefore take a long time to load.

Columns

N	www	Column Name	Column Type	Description
1	✓	Census_Year	Long Integer	
2	✓	csdUid	Long Integer	
3		subID	Long Integer	
4		prUid	Integer	
5	✓	ProvinceCode	Short Text (2)	
6	✓	csdName	Short Text (60)	
7	✓	csdType	Short Text (3)	
8	✓	cmaName	Short Text (75)	
9		csdStatus	Short Text (10)	
10		flag_Oddball	Integer	
11	✓	Ab_Flag	Short Text (3)	
12	✓	Latitude	Decimal	
13	✓	Longitude	Decimal	
14		CSD_link	Short Text (7)	
15		hlink_statcan	Long Text (300)	
16		hlink_boundary	Short Text (100)	
17	✓	Language_Name	Short Text (60)	
18		Lang_ID	Integer	
19	✓	Pop	Long Integer	
20	✓	AvAge	Double	

CSD Summary counts by language, Province, and census year
https://norrisresearch.com/ref/csd_lang_summary.htm

Data Source

Database: db_csd_lang.accdb

Table: csd_lang_summary

Description

This table summarises the distribution of languages by Province/Territory using the NRI classification. The table lists the number of CSDs in each jurisdiction where the language is traditionally spoken. For example, for Mi'kmaq [Micmac], the language is traditionally spoken in a total of 79 CSDs, of which 26 are in Nova Scotia. In the 2011 Census, of those 79 CSDs, 23 reported having mother tongue speakers. The Province column identifies thelocation of the CSD with the largest number of mother tongue speakers. For Mi'kmaq [Micmac], this corresponds to Eskasoni, in Nova Scotia.

Columns

N	www	Column Name	Column Type	Description
1	✓	Census_Year	Integer	
2		Summary_Code	Short Text (10)	
3		Lang_Name_NRI	Short Text (60)	
4	✓	Lang_Name_List	Short Text (120)	
5		Language_Status	Short Text (25)	
6	✓	Lang_Family	Short Text (12)	
7	✓	Language_Endangerment	Short Text (30)	
8		Endangerment_Level	Integer	
9	✓	Province_Code	Short Text	
10	✓	Canada	Integer	
11	✓	NL	Integer	
12	✓	PE	Integer	
13	✓	NS	Integer	
14	✓	NB	Integer	
15	✓	QC	Integer	
16	✓	ON	Integer	
17	✓	MB	Integer	
18	✓	SK	Integer	
19	✓	AB	Integer	
20	✓	BC	Integer	
21	✓	YT	Integer	
22	✓	NT	Integer	
23	✓	NU	Integer	

Reserves

https://norrisresearch.com/ref/fnp_reserves.htm

Data Source

Database: db_reserves.accdb

Table: FNP_Placemarks

Description

This table presents a list of all the Indian Reserves found on the INAC "First Nations Profiles" website. The list of Reserves evolves slowly. The data in this table were retrieved from the INAC website in August, 2017.

The table includes links to

- The corresponding page on the INAC website

- the corresponding CanLands map (if available)

- a map of the CSD outline (if applicable)

- a link to the First Nation that corresponds to the reserve, in those cases where there is a 1-to-1 relationship

The table contains data for a total of 3326 reserves. Note that reserves which have been "surrendered" are not included in this list.

Columns

N	www	Column Name	Column Type	Description
1		Reserve_Code	Short Text (12)	
2		CSDuid_2001_1	Long Integer	
3		CSDtype_2001_1	Short Text (3)	
4		CSDuid_2001_2	Long Integer	
5		CSDtype_2001_2	Short Text (3)	
6		CSDuid_2006_1	Long Integer	
7		CSDtype_2006_1	Short Text (3)	
8		CSDuid_2006_2	Long Integer	
9		CSDtype_2006_2	Short Text (3)	
10		CSDuid_2011_1	Long Integer	
11		CSDtype_2011_1	Short Text (3)	
12		CSDuid_2011_2	Long Integer	
13		CSDtype_2011_2	Short Text (3)	
14		CSDuid_2016_1	Long Integer	
15		CSDtype_2016_1	Short Text (3)	

16		CSDuid_2016_2	Long Integer	
17		CSDtype_2016_2	Short Text (3)	
18	✓	Reserve_Number	Long Integer	
19	✓	Reserve_Name	Short Text (75)	
20		CanLands_Name_Orig	Short Text (75)	
21		CanLands_Type	Short Text (75)	
22	✓	Latitude	Decimal	
23	✓	Longitude	Decimal	
24	✓	Province	Short Text (7)	
25		flag_CSD_MetisInuit	Short Text (1)	
26		flag_ILRS_LSS	Long Integer	
27		flag_SubSurface	Long Integer	
28		flag_CanLands	Long Integer	
29		flag_Quirks	Long Integer	
30		FNP2016_Flag	Long Integer	
31		FirstNationName	Short Text (60)	
32		FirstNationNumber	Long Integer	
33		filename_canlands	Short Text (100)	
34	✓	hlink_canlands	Short Text (200)	
35		hlink_csd2001_1	Short Text (100)	
36		hlink_csd2001_2	Short Text (100)	
37		hlink_csd2001	Short Text (200)	
38		hlink_csd2006_1	Short Text (100)	
39		hlink_csd2006_2	Short Text (100)	
40		hlink_csd2006	Short Text (200)	
41		hlink_csd2011_1	Short Text (100)	
42		hlink_csd2011_2	Short Text (100)	
43	✓	hlink_csd2011	Short Text (200)	
44		hlink_csd2016_1	Short Text (100)	
45		hlink_csd2016_2	Short Text (100)	
46		hlink_csd2016	Short Text (200)	
47	✓	hlink_fnp_reserve	Short Text (200)	
48	✓	hlink_fnp_fn	Short Text (200)	
49		flag_CSD_2001	Short Text (2)	
50		flag_CSD_2011	Short Text (2)	
51		flag_CSD_2016	Short Text (2)	

Norris Research: Language Data Tables Users Guide DRAFT January 1, 2021

First Nations
https://norrisresearch.com/ref/first_nations_lang.htm

Data Source

Database: db_fn_lang.accdb

Table: FNP_FirstNations_Lang

Description

This table presents a list of all the First Nations found on the INAC "First Nations Profiles" website. The list of First Nations evolves slowly. The data in this table were retrieved from the INAC website in August, 2017. There is a total of 636 First Nations.

There are hyperlinks to additional information for each First Nation, and also to its associated tribal council. Please note that it is difficult to maintain up-to-date links for the links to websites for the First Nations and Tribal Councils. Some links may be broken, and some may be missing.

The corresponding KML map of First Nations generally maps the location of each First Nation according to its mailing address. For some isolated communities, the mailing address may resolve to a lat/long that is a long way from the actual location of the First Nation. These problems have been corrected when they have been identified, but there are probably still some unrecognised problems.

Columns

N	www	Column Name	Column Type	Description	
1		FileName	Short Text (20)	20	
2		OrgType	Short Text (2)		2
3	✓	OrgName	Short Text (125)	125	
4	✓	OrgNumber	Long Integer		4
5	✓	Province_Code	Short Text (2)		2
6		FNP_Address	Short Text (100)	100	
7	✓	Latitude	Decimal	16	
8	✓	Longitude	Decimal	16	
9		Latitude_GC	Decimal	16	
10		Longitude_GC	Decimal	16	
11		Latitude_KML	Decimal	16	
12		Longitude_KML	Decimal	16	
13	✓	URL_Org	Short Text (250)	250	
14	✓	URL_Wikipedia	Short Text (250)	250	
15		hlinkOrg	Short Text (250)	250	
16		hlinkWikipedia	Short Text (250)	250	

57

17		URL_Type_Wikipedia	Short Text (5)		5
18		URL_Flag	Long Integer		4
19	✓	URL_FNP	Short Text (250)	250	
20		TC_Number	Long Integer		4
21	✓	TC_Name	Short Text (125)	125	
22	✓	URL_TC_Org	Short Text (250)	250	
23	✓	URL_TC_Wikipedia	Short Text (250)	250	
24	✓	URL_TC_FNP	Short Text (250)	250	
25		hlink_Org	Short Text (250)	250	
26		hlinks_Org	Long Text		-
27		hlinks_TC	Long Text		-
28		nTemp	Short Text (6)		6
29		flag_Correction	Long Integer (500)		4
30		Coding_Type_ID	Long Integer (500)		4
31		Language_Name	Short Text (60)	60	
32	✓	Language_Name_List	Short Text (120)	120	

Tribal Councils

https://norrisresearch.com/ref/tribal_councils.htm

Data Source

Database: db_aandc_fnp.accdb

Table: Addresses_tc_kml

Description

This table presents a list of all the Tribal Councils found on the INAC "First Nations Profiles" website. The list of Tribal Councils evolves slowly. The data in this table were retrieved from the INAC website in August, 2017. Please note that it is difficult to maintain up-to-date links for the links to websites for the Tribal Councils. Some links may be broken, and some may be missing.

The corresponding KML map of Tribal Councils maps the location of each Tribal Council according to its mailing address.

Columns

N	www	Column Name	Column Type	Description
1		FileName	Short Text (20)	
2		OrgType	Short Text (2)	
3	✓	OrgName	Short Text (125)	
4	✓	OrgNumber	Long Integer	
5	✓	Province_Code	Short Text (2)	
6		FNP_Address	Short Text (100)	
7		Latitude	Decimal	
8		Longitude	Decimal	
9		Latitude_GC	Decimal	
10		Longitude_GC	Decimal	
11		Latitude_KML	Decimal	
12		Longitude_KML	Decimal	
13	✓	URL_Org	Short Text (250)	
14	✓	URL_Wikipedia	Short Text (250)	
15		hlinkOrg	Short Text (250)	
16		hlinkWikipedia	Short Text (250)	
17		URL_Type_Wikipedia	Short Text (5)	
18		URL_Flag	Long Integer	
19		URL_FNP	Short Text (250)	
20		TC_Number	Long Integer	
21		TC_Name	Short Text 9125)	
22		URL_TC_Org	Short Text (250)	

23		URL_TC_Wikipedia	Short Text (250)	
24		URL_TC_FNP	Short Text (250)	
25		hlink_Org	Short Text (250)	
26		hlinks_Org	Long Text (500)	
27		hlinks_TC	Long Text (500)	
28		nTemp	Short Text (6)	
29		flag_Correction	Long Integer	
30		Coding_Type_ID	Long Integer	

Canada Lands

https://norrisresearch.com/ref/canlands.htm

Data Source

Database: db_canlands.accdb

Table: Placemarks

Description

This table lists all of the "places" retrieved from the KML map of "Canada Lands" that is available on the NRCan website at:

https://www.nrcan.gc.ca/earth-sciences/geomatics/canada-lands-surveys/11092#CLGoogleEarth

The places include

- Indian Reserve   3080
- Municipal Boundary 46
- National Park   46
- Other   55
- Subdivision   89

The 3080 reserves are not a simple subset of the 3326 reserves listed on the INAC FNP website. Reconciling the two lists is not trivial.

The "Other" category encompasses land claim agreements involving different Indigenous groups.

Columns

N	www	Column Name	Column Type	Description
1		ID	Long Integer	
2	✓	Name_Placemark	Short Text (75)	
3		Name_Original	Short Text (75)	
4		Name_Modified_1	Short Text (75)	
5		Name_Modified_2	Short Text (75)	
6		Name_File	Short Text (100)	
7	✓	hlink_canlands	Short Text (200)	
8	✓	Type_Placemark	Short Text (20)	
9	✓	Latitude	Decimal	
10	✓	Longitude	Decimal	
11		Quad	Short Text (4)	

12		UTMZone	Decimal	
13		UTMZone_chr	Short Text (2)	
14	✓	Province	Short Text (2)	

Friendship Centres

https://norrisresearch.com/ref/friendship_centres.htm

Data Source

Database: db_friendship_centres.accdb

Table: Addresses_fc_kml

Description

The list of friendship centres is based on data that were retrieved from the NAFC website:

https://nafc.ca/en/friendship-centres/

The data were extensively edited, to update mailing addresses and websites. The data have not been updated since 2017, and, as with the similar tables of First Nations and Tribal Councils, it is to be expected that some of the data are no longer valid.

Columns

N	www	Column Name	Column Type	Description
1		Postal_Code	Short Text (7)	
2	✓	OrgType	Short Text (10)	
3		OrgTypeID	Long Integer	
4	✓	OrgName	Short Text (125)	
5		OrgNumber	Long Integer	
6	✓	Province_Code	Short Text (2)	
7		FC_Address	Short Text (100)	
8		Address	Short Text (70)	
9	✓	City	Short Text (25)	
10		Telephone	Short Text (30)	
11		Latitude	Decimal	
12		Longitude	Decimal	
13		Latitude_GC	Decimal	
14		Longitude_GC	Decimal	
15		Latitude_KML	Decimal	
16		Longitude_KML	Decimal	
17	✓	URL_Org	Short Text (250)	
18		URL_Wikipedia	Short Text (250)	
19		hlinkOrg	Short Text (250)	
20		hlinkWikipedia	Short Text (250)	
21	✓	Closed_Flag	Short Text (10)	

Distinctions

https://norrisresearch.com/distinctions/distinctions.htm

Data Source

Database: distinctions_01/*.accdb; distinctions_02/*.accdb; distinctions_03/*.accdb

Table: CrossTab_*

Description

The data accessed via this link are quite different from the html tables and KML maps accessed via the other links. There are no simple tables or maps. Instead, the Distinctions tables comprise a set of very large MS Excel spreadsheets. Some of the spreadsheets have row counts approaching 500,000. Working with these data requires a high degree of familiarity with the structure of Census data, in order to reduce the data to ueful subsets.

Norris (2019) gives numerous examples of the possible uses of these datasets.

Columns

N	www	Column Name	Column Type	Description
1				
2				
3				
4				

BC Fishing Reserves

Data Source

Database: db_BC_Fishing_Reserves.accdb

Table: Fishing_Reserves_v04 (in progress)

Description

This table is built from data that was published as an online supplement to *Landing Native Fisheries: Indian Reserves and Fishing Rights in British Columbia, 1849-1925*, by Douglas C. Harris (Vancouver: UBC Press, 2008).

https://open.library.ubc.ca/cIRcle/collections/ubcpress/641/items?year=2008

Columns

N	www	Column Name	Column Type	Description
1				
2				
3				
4				

Historical Atlas Reserves to 1900

Data Source

Database: db_places.accdb

Table: (in progress)

Description

Data are taken from an MS Excel spreadsheet (Native_Reserves_1902.xls) that is available online via the "Historical Atlas of Canada Online Learning Project":

http://www.historicalatlas.ca/website/hacolp/

The data sources used to create the Atlas maps are documented here:

http://www.historicalatlas.ca/website/hacolp/national_perspectives/native_canada/UNIT_15/U15_sources.htm

Columns

N	www	Column Name	Column Type	Description
1				
2				
3				
4				

BC Reserves, 1943

Data Source

Database: BC_Reserves_1943.accdb

Table: BC_Reserves_1943

Description

[data manipulation by script bc_reserves.sql, called from db_sql.bat]

Data are taken from "SCHEDULE OF INDIAN RESERVES IN THE DOMINION OF CANADA PART 2 RESERVES IN THE PROVINCE OF BRITISH COLUMBIA RECOMPILED AND CORRECTED UP TO MARCH 31, 1943." A transcripton of the report is available as a PDF file from Library and Archives Canada:

http://www.bac-lac.gc.ca/eng/discover/aboriginal-heritage/first-nations/indian-affairs-annual-reports/Pages/item.aspx?IdNumber=33382

Further development is on hold until all of the Reserves that are identified in table Ab_Places, database db_places.accdb, have been located (script places.sql). See also Reserves_without_Locations.docx, derived from table Ab_Places (222 records as of March, 2019).

Columns

N	www	Column Name	Column Type	Description
1				
2				
3				
4				

Specific Claims

Data Source

Database: db_places.accdb

Table: Specific_Claims

Description

Data are retrieved from the INAC website:

http://services.aadnc-aandc.gc.ca/SCBRI_E/Main/ReportingCentre/External/externalreporting.aspx

Columns

N	www	Column Name	Column Type	Description
1				
2				
3				
4				

Methodology Overview

This section describes the overall methodology that is used to create the website tables and maps from the source data.

Principal steps - overview

1) Establish a classification system for indigenous languages in Canada
    a) Referred to here as the "NRI" classification
2) Obtain census data for indigenous languages in Canada, at the CSD level
3) Correlate the NRI language classification with StatCan classifications
4) Identify "indigenous" CSDs
5) Identify "traditional" indigenous languages within the indigenous CSDs
6) Assign census populations to NRI languages at the CSD level for "indigenous" CSDs
7) Assign census populations to NRI languages at the CSD level for remaining "non-indigenous" CSDs
8) Reconcile the StaCan population counts with the NRI population counts at the CSD level
9) Assign census populations to NRI languages at the Canada level
10) Reconcile the StaCan population counts with the NRI population counts at the Canada level
11) Resolve anomalies in the data that are revealed in steps 5 to 10, and repeat steps 3 to 10 until all of the anomalies have been dealt with
12) Calculate levels of endangerment for the NRI languages
13) Generate the tables and maps for the website

Principal steps – details

1. Establish a classification system for indigenous languages in Canada

The classification system that is used here has evolved from the classification that was used by Norris Research for the 3rd edition of the UNESCO Atlas of the World's Languages in Danger (Moseley, 2010). The challenges of obtaining a classification are documented in an NRI report that is currently being revised for a 4th edition. The 3rd edition can be found online at:

https://norrisresearch.com/lang.htm

The reader should refer to this report for details. Only a few salient points will be mentioned here:

1. No attempt is made to draw a rigorous distinction between "language" and "dialect". The classification makes no claim to be a definitive list of Indigenous languages in Canada. It is optimised for comparison with other

classifications and for useful geostatistical analysis i.e. populations are not too big and not too small, and similarly for geographic extents.

2. The classification has 107 categories but this includes several languages that are extinct or dormant. There are 87 categories for which we believe there are currently mother tongue speakers living in Canada. There is one language – Wendat [Wendake Huron] – which is classified as secondarily surviving.

3. For languages whose traditional territories span the Canada-USA border, we have taken an "inclusive" approach. Note that some languages which are classed as dormant in Canada may have mother tongue speakers or be secondarily surviving in the USA.

4. Correspondence with the classifications of the Ethnologue, Glottolog and FPCC is generally very good. The major discrepancies are:

    a. Inuktut

        i. We use the classification developed by Dorais (2010), and mapped by Sontag (2007).

    b. Carrier

        i. We use the classification of Poser (2011).

    c. Algonquin

        i. Our bipartite subdivision is based on the work of Drapeau (1996, p142).

2. Obtain census data for indigenous languages in Canada, at the CSD level

Census data were obtained as multi-dimensional arrays in Beyond 2020 format, as custom data retrievals from StatCan. A critical aspect for this project is the provision of average age data for the populations. A second major aspect is that the data for 2001, 2006 and 2016 were mapped to a standard geography based on the 2011 CSD boundaries. Appendix D gives details on some of the characteristics of census data which are particularly important for this project.

3. Correlate the NRI language classification with StatCan classifications

The StatCan classification of Indigenous languages has grown steadily more detailed from one census to the next. In 1971, the classification was simply binary: "Indian" and "Eskimo" (see Appendix C). Unfortunately, increasing detail does not lead to increasingly straightforward correlation. The 2016 StatCan classification is very close to that of NRI, meaning that many languages have a 1-to-1 correlation between the two schemes. The major challenge is with the complex language continua of Cree, Ojibway, and Inuktut. In these cases, many respondents simply identify their language by its broad category e.g. "Cree", rather than "Plains Cree", and the correlation becomes complex.

4. Identify "Indigenous" CSDs

StatCan produces a "census dictionary" to accompany every census. Within each dictionary, there is a section on CSD geography. Within these sections, there is a sub-section on "Census Subdivison Types Associated with 'On-reserve' population", according to criteria established by DIAND/INAC. This list of reserves corresponds to an intial identification of "Indigenous" CSDs. To this list, we can immediately add the approximately 50 Inuit communities in Inuit Nunangat.

This list still does not capture communities with significant populations that self-identify as "Indigenous". For the present exercise, it is important to capture these communities. Accordingly, we add to the list any CSD with at least 25% of the population reporting an Indigenous identity in at least one of the censuses from 2001 to 2016. 25% is a somewhat arbitrary figure, which was selected in order to include some communities with significant Indigenous mother tongue populations (particularly Prince Rupert and Yellowknife).

A complicating factor in this list is that the CSD geography changed between 2011 and 2016. The data for 2001, 2006 and 2011 have all been levelled to the 2011 geography. Some CSDs that were newly created for the 2016 census are included in the list. To view these lists, look at this table:

https://norrisresearch.com/ref/csds_2011.htm

Filter the table for only Indigenous CSDs by typing Y in the *Ab flag* column. Then sort the table on the *2016 Flag* column in descending order. CSDs with a value of 2 are new CSDs for 2016, whilst CSDs with a value of 1 are CSDs that were removed for 2016.

The final list of Indigenous CSDs has 1401 records.

5. Identify "traditional" indigenous languages within the indigenous CSDs

Each of these Indigenous CSDs can be assigned to one or more "expected" Indigenous languages. For example, we would expect to find speakers of Mi'kmaq [Micmac] in Eskasoni, but no speakers of any of the other Indigenous languages. For the Six Nations CSDs, we would "expect" all of the six Iroquoian languages corresponding to the six nations. We are not concerned here with whether any mother tongue speakers actually exist in a CSD – only with what language we might expect them to speak if they do exist.

In identifying traditional languages, we rely heavily on the association of every Reserve with at least one First Nation. The large majority of reserves are associated with only one First Nation. There are a few exceptions to this rule, which are summarised in Appendix E. There are 636 First Nations, and most of them have only one traditional language. Exceptions to this generalisation are listed in Appendix F (although this list is subject to ongoing revision). Knowing the traditional language of a First Nation therefore allows us to establish the traditional language of a particular reserve (i.e. CSD), for the large majority of the 1401 CSDs. We can establish the traditional language for each First Nation by a variety of approaches, relying on previous research, web sites and personal communication. We can also leverage the fact that most tribal councils are organised along cultural similarities, meaning that for many tribal councils, all of the associated First Nations have the same traditional language.

Note that this exercise of assigning traditional languages cannot be carried out rigorously without extensive community consultation, which is beyond the scope of the current project. The assignment is therefore subject to revision as new information is received.

6. Assign census populations to NRI languages at the CSD level for "indigenous" CSDs

Having assigned one or more traditional languages to each of the 1401 "Indigenous" CSDs, we can now correlate the StatCan data on mother tongue speakers with the NRI language. Some examples:

1. Eskasoni 3 (2016): StatCan reported 2475 speakers of Mi'kmaq. The NRI traditional language of this CSD is Mi'kmaq [Micmac], and so all 2475 mother tongue speakers are mapped directly from the StatCan to the NRI classification.

2. Sechelt (Part) (2016): StatCan reported 35 speakers of "Salish languages, n.i.e." The NRI traditional language of this CSD is She shashishalhem [Sechelt]. We established a correlation in step 3 of "Salish languages, n.i.e." with both She shashishalhem [Sechelt] and Nuxálk [Bella Coola]. We assume that all 35 speakers can be assigned to She shashishalhem [Sechelt]. There is, of course, the possibility that some of these 35 people are truly Nuxálk [Bella Coola] speakers. However, the assumption is felt to be reasonable.

3. Cumberland House (2016): StatCan reported 70 speakers of Swampy Cree and 155 speakers of "Cree, n.o.s.". The NRI traditional languages for this community are Dëne Sųliné [Chipewyan] and Cree, Swampy. All of the 225 mother tongue speakers reported by StatCan are assigned to Cree, Swampy.

This methodology is straightforward and allows the large majority of StatCan mother tongue speakers to be mapped to a single NRI language.

The number of exceptions is sufficiently small that they can be handled on a case-by-case basis. Some examples:

1. Pond Inlet (2011): 60 Innu/Montagnais speakers reported by StatCan are all reassigned to Inuktitut, and thence to the NRI dialect of Qikiqtaaluk uannangani [Baffin, North].

2. Natuashish 2 (2016): 605 Naskapi speakers reported by StatCan are re-assigned to Montagnais (Innu), to be consistent with the data reported for 2011, and with the NRI traditional language of Innu-aimun [Montagnais, Eastern]. Note that this is a strange case, and is somewhat problematic because Naskapi and Innu-Aimun are part of the Cree dialect continuum, and Natuashish sits on the boundary of the two dialects. The community has historically self-identified as Innu-Aimun.

Note that the methodology allows for "stray" speakers to be found in unexpected places. This is a fairly common scenario for Cree speakers, who are widely dispersed across many CSDs in small numbers. These speakers are NOT reassigned to the traditional language of the community. There is a very small number of cases of small populations of mother tongue speakers occurring in very unexpected places. For example, there are 10 speakers of Carrier reported in Moosomin 112B in 2011. Whilst this is improbable, it is not impossible. Such cases are flagged as anomalous, but the data are not altered.

7. Assign census populations to NRI languages at the CSD level for remaining "non-indigenous" CSDs

Most mother tongue speakers reside in "Indigenous" CSDS. The validity of this assumption can be seen by examining this table:

https://norrisresearch.com/ref/lang_census_mt.htm

For example, if we filter the columns in this table on Language [SC] = Blackfoot for 2016, we see that the total Blackfoot mother tongue speakers residing on Indigenous CSDs is 2495, whilst the total Blackfoot mother tongue population for all CSDs is 3270. Similar patterns are seen for most other languages, although there are some noteworthy exceptions.

For languages where there is a simple 1-to-1 correspondence between the StatCan and NRI schemes, language assignment is straightforward. The challenge lies in situations where one StatCan language corresponds to multiple NRI categories. For example, there were 160 Ojibway mother tongue speakers in Toronto in 2016. How can we assign these speakers to the different Ojibay dialects? We are forced to make a simplifying assumption: the ratio of the different dialects spoken in Toronto matches exactly to the ratio of the total number of mother tongue speakers in Indigenous CSDs.

As a simple hypothetical example, consider a language X reported by StatCan as having 20 speakers in an urban CSD. Language X is mapped to NRI languages A and B. We have already calculated total populations of languages A and B in Indigenous communities of 60 and 40 respectively. We will therefore asign 60% of the 20 urban speakers to language A and the remaining 40% to language B.

This methodology works well for the large majority of situations. As in the previous step, the number of exceptions is sufficiently small that they can be handled on a case-by-case basis.

As in the previous case, there are examples of mother tongue speakers being found in unlikely places. For example, in 2016 the CSD of Northeastern Manitoulin reported 10 Dene speakers. The data are flagged, but not altered.

8. Reconcile the StatCan population counts with the NRI population counts at the CSD level

At this stage, the StatCan population counts can be compared to the NRI counts. The total counts should be in exact agreement. Any inconsistencies need to be investigated and resolved, before proceeding to the next stage.

9. Assign census populations to NRI languages at the Canada level

Data suppression means that in general, the sum of the populations reported at the CSD level is rather less than the population reported at the Canada level. The extent of the disagreement varies significantly from one language to another, and can be seen in this table:

https://norrisresearch.com/ref/lang_census_mt.htm

Tahltan in 2016 provides an example of a severe disagreement: 75 speakers by summing over all CSDs, versus 110 speakers reported for Canada.

In order for the NRI language counts at the Canada level to be precisely consistent with the StatCan counts at the Canada level, we need to account for this discrepancy. The logic is essentially identical to that used in assigning speakers in non-Indigenous CSDs to NRI languages.

10. Reconcile the StaCan population counts with the NRI population counts at the Canada level

A second reconciliation of speaker counts is now possible. Note that the final NRI population counts are not identical to the StatCan data. There are some data that cannot be sensibly assigned from a StatCan language to an NRI one. A good example is found in the 105 Sarsi mother tongue speakers reported in Toronto in 2011. This number simply cannot be correct. Regardless of how it may have arisen, it must be discarded. There are very few problems of this nature, though, and they can be summarised in a small table (Appendix G, in preparation).

11. Resolve anomalies in the data

If the reconciliation fails, then the logic of the data processing must be invalid, somehow, and the earlier steps need to be revisited, in order to identify the cause of the problem. Note that in practice this resolution of anomalies is an ongoing aspect of the data processing, and not saved until the end of the process.

12. Average Age Calculations

Aggregating population counts for CSDs involves a straightforward summation. Calculating average ages is more complicated. The complication arises because average age data are generally suppressed for small population counts. We need to allow for these suppressed data in the calculation.

13. Calculate levels of endangerment for the NRI languages

The calculation of levels of endangerment is a very straightforward process, based on population count and average age. Note that it needs some refinement, to handle languages with very small populations and also low average ages (less than 35). Such languages (e.g. North Straits Salish in 2016) are anomalous, and lead to a very misleading "unsafe" classification.

14. Generate the tables and maps for the website

Generation of the tables and maps is a highly automated process. Essentially, it involves exporting the tables from the MS Access databases in XML format. The XML is then transformed, using XSLT, into (a) XHTML for delivery as html tables and (b) KML for delivery as maps.

Database Description

In the above 14-step methodology, steps 1 to 5 can be considered as fundamental data acquisition steps, which are required in order to build the database that is the starting point for manipulating the Census data. Steps 6 to 13 are all data processing steps, which are executed by SQL scripts. The "starting point" database consists of several tables, listed below. The primary sources for most of these tables are a set of MS word documents. Tables within the MS Word files can be edited easily, and the revised contents simply sopy-pasted into the MS Access database. Further details on each of the tables can be foud in the corresponding MS Word files.

Other source tables are secondary, not primary, and are imported from other MS Access databases, after executing SQL scripts to transform original primary source data.

Several of the tables hold distinct sets of data for each of the four language variables, MT, KN, HLA and HLB.

N	Table	MS Word source document	Database	Lang var
1	Aboriginal_CSDs_NRI_Lang	language_data_tables_aboriginal_csds_nri_lang	db_csd_lang	
2	Census_Corrections_Canada	language_data_tables.census_data	db_csd_lang	Yes
3	CSD_Changes	language_data_tables.miscel	db_csd_lang	
4	CSD_Changes_Lang_NRI_Ab	language_data_tables.census_data	db_csd_lang	Yes
5	CSD_Changes_Lang_NRI_nonAb_miscel	language_data_tables.census_data	db_csd_lang	Yes
6	CSD_Changes_Lang_NRI_nonAb_nie	language_data_tables.census_data	db_csd_lang	Yes
7	CSD_Changes_Lang_STC	language_data_tables.census_data	db_csd_lang	Yes
8	CSD_stray_speakers	language_data_tables.stray_speakers	db_csd_lang	Yes
9	CSDs_2011			
10	CSDs_Lang_Pop		db_csd_lang	Yes
11	Endangerment_Levels	language_data_tables.miscel	db_csd_lang	

12	FNP_Placemarks			
13	Lang_Correspondences	language_data_tables.miscel	db_csd_lang	
14	Lang_Xref_NRI_stc	language_data_tables.miscel	db_csd_lang	
15	Language_Groupings_NRI	language_data_tables.miscel	db_csd_lang	
16	Language_Status_Changes	language_data_tables.miscel	db_csd_lang	
17	Language_Variables	language_data_tables.miscel	db_csd_lang	
18	Languages_Can_USA			
19	Languages_Map			
20	Lookup_Lang	language_data_tables.miscel	db_csd_lang	
21	nie_Speakers_Canada	language_data_tables.census_data	db_csd_lang	Yes
22	Population_Factor_Flags	language_data_tables.census_data	db_csd_lang	Yes
23	PR_list	language_data_tables.miscel	db_csd_lang	

Table CSDs_Lang_Pop

Lang_Var_Code	Census_Year	N
HLA	2001	260674
HLA	2006	268408
HLA	2011	235657
HLA	2016	322759
HLB	2001	236936
HLB	2006	189441
HLB	2011	226153
HLB	2016	296902
KN	2001	286471
KN	2006	301376
KN	2011	249067
KN	2016	297333
MT	2001	239167
MT	2006	249051
MT	2011	211051
MT	2016	295385

References

Aboriginal Affairs and Northern Development Canada (AANDC) 2013. "Aboriginal Demographics from the 2011 National Household Survey" Planning, Research and Statistics Branch, May 2013.

Dorais, L.-J. (2010). The Language of the Inuit: Syntax, Semantics, and Society in the Arctic. McGill-Queen's University Press. 396 pp.

Drapeau, L. (1996). The state of the art in linguistic research, standardization and modernization in Quebec Aboriginal languages. *In* J. Maurais (Ed.), Quebec's Aboriginal Languages: History, Planning, Development. Multilingual Matters Ltd., pp. 129-158.

Maurais, J. (1996). Quebec's Aboriginal Languages: History, Planning, Development. Mulitilingual Matters Limited. 334 pp.

Norris, M.J. (2019). Trends in the State of Indigenous Languages in Canada, 1996/2001 to 2011: Distinctions by First Nation, Métis and Inuit Identities and Registered Indian Status, by Place of Residence. Report prepared under contract with the Department of Canadian Heritage. 355 pp.

Moseley, C. (2010). Atlas of the World's Languages in Danger. UNESCO Publishing. 154 pp.

Poser, W.J. (2011). The Carrier Language – a Brief Introduction. College of New Caledonia Press. 60 pp.

Ray, A. J. (1999). Treaty 8: A British Columbian Anomaly. B.C. Studies, no 123, p 5-58.

Sontag, N. (2007). The Inuit Language in Inuit Communities in Canada. University of Alaska Press. Map

Appendix A CSD Geography, 2001-2016

This appendix contains excerpts describing the CSD geography, taken from the census dictionaries for 2001-2016. The source documents can be found online in PDF format, by searching for their StatCan catalogue numbers:

Year	Catalogue Number
2001	92-378-XIE02002
2006	92-566-eng
2011	98-301-X2011001-eng
2016	98-301-x2016001-eng

2001 CSD Geography

2006 CSD Geography

2011 CSD Geography

2016 CSD Geography

Appendix B Incompletely Enumerated Reserves, 2001-2016

This appendix contains excerpts describing the incompletely enumerated CSDs, taken from the census documents for 2001-2016. For 2006-2016, the source documents are the Aboriginal Peoples Technical Reports. There was no equivalent document produced for the 2001 census. The relevant data were found in the Coverage Technical Report. For 2016, additional information was also taken from the *Guide to the Census of Population*. The source documents can be found online in PDF format, by searching for their StatCan catalogue numbers:

Year	Catalogue Number
2001	92-394-XIE
2006	92-569-X2006001-eng
2011	99-011-x2011002-eng
2016	98-307-x2016001-eng
2016	98-304-x2016001-eng

2001 Partially Enumerated CSDs

2006 Partially Enumerated CSDs

2011 Partially Enumerated CSDs

2016 Partially Enumerated CSDs

Tabulations of incompletely enumerated reserves, 2001-2016

The tabulations are taken from the following web pages, which contain additional information:

Year	Hyperlink
2001	https://www12.statcan.gc.ca/english/census01/Products/Reference/tech_rep/coverage/tab14-3-1.cfm
2006	https://www12.statcan.gc.ca/census-recensement/2006/ref/notes/aboriginal-autochtones-eng.cfm
2011	https://www12.statcan.gc.ca/nhs-enm/2011/ref/aboriginal-autochtones-eng.cfm
2016	https://www12.statcan.gc.ca/census-recensement/2016/ref/98-304/app-ann1-2-eng.cfm

2001

Province / Reserve or settlement name	Status in 1996 Census	Census count or estimate in 1996	Estimated census count in 2001	Under coverage rate in 2001 (%)	Estimated missed in 2001	Population estimate in 2001
Quebec		**10,717**	**11,717**	**7.4**	**931**	**12,648**
Akwesasne (Partie)	Incomplete	2,265	2,471		196	2,667
Doncaster 17	Complete	0	4		0	4
Kahnawake 14	Incomplete	7,087	7,757		617	8,374
Kanesatake	Incomplete	1,137	1,235		98	1,333
Lac-Rapide	Complete	228	250		20	270
Ontario		**13,176**	**14,345**	**10.1**	**1,615**	**15,960**
Akwesasne (Part) 59	Incomplete	1,229	1,335		150	1,485
Bear Island 1	Complete	153	169		19	188
Chippewas of the Thames First Nation 42	Incomplete	777	839		94	933
Goulais Bay 15A	Incomplete	46	54		6	60
Marten Falls 65	Complete	204	228		26	254
Moose Factory 68	Complete	0	4		0	4
Munsee-Delaware Nation 1	Incomplete	214	235		26	261
Ojibway Nation of Saugeen (Savant Lake)	Incomplete	203	224		25	249
Oneida 41	Incomplete	1,226	1,332		150	1,482
Pikangikum 14	Complete	1,17	1,261		142	1,403
Rankin Location 15D	Incomplete	670	723		81	804

Six Nations (Part) 40[5]	Incomplete	632	681		77	758
Six Nations (Part) 40	Incomplete	4,731	5,174		583	5,757
Tyendinaga Mohawk Territory	Incomplete	1,15	1,249		141	1,390
Wahta Mohawk Territory	Incomplete	154	171		19	190
Whitefish Bay 32A	Incomplete	502	538		61	599
Whitesand	Complete	115	128		14	142
Manitoba		**85**	**97**	**12.1**	**13**	**110**
Dakota Tipi 1	Incomplete	85	97		13	110
Saskatchewan		**451**	**491**	**15.6**	**90**	**581**
Big Head 124	Incomplete	451	491		90	581
Alberta		**3,775**	**4,115**	**17.3**	**862**	**4,977**
Ermineskin 138	Incomplete	1,300	1,413		296	1,709
Little Buffalo	Incomplete	221	243		51	294
Saddle Lake 125	Incomplete	2,254	2,459		515	2,974
British Columbia		**225**	**255**	**2.9**	**8**	**263**
Esquimalt	Incomplete	82	93		3	96
Marble Canyon 3	Complete	67	76		2	78
Pavilion 1	Complete	76	86		3	89
Provincial Total		**28,429**	**31,020**	**10.4**	**3,521**	**34,541**

2006

Province	Incompletely enumerated Indian reserves and Indian settlements, 2006	Population, 2001	Population, 1996
Quebec	Gesgapegiag (formerly Gesgapegiag 2)	488	442
	Doncaster (formerly Doncaster 17)	¶	0
	Kanesatake	¶	¶
	Kahnawake (formerly Kahnawake 14)	¶	¶
	Akwesasne (formerly Akwesasne (Partie))	¶	¶
	Lac-Rapide	¶	228
	Wendake	1,555	¶
Ontario	Fort Severn 89	401	362
	Attawapiskat 91A	1,293	1,258
	Factory Island 1	1,430	1,286
	Bear Island 1	¶	153
	Tyendinaga Mohawk Territory	¶	¶
	Wahta Mohawk Territory	¶	¶
	Six Nations (Part) 40	¶	¶
	Six Nations (Part) 40	¶	¶
	Oneida 41	¶	¶
	Akwesasne (Part) 59	¶	¶
Saskatchewan	Big Island Lake Cree Territory (formerly Big Head 124)	¶	¶
Alberta	Little Buffalo	¶	¶
	Saddle Lake 125	¶	¶
	Tsuu T'ina Nation 145 (Sarcee 145)	1,982	1,509
British Columbia	Esquimalt	¶	¶

2011

Province	Incompletely enumerated Indian reserves and Indian settlements, 2011	Enumeration status (reasons for absence of data)	Population, 2006	Population, 2001
Quebec	Kanesatake	Permission not given	¶	¶
	Wendake	Permission not given	¶	1,555
	Doncaster	Permission not given	¶	¶
	Kahnawake	Permission not given	¶	¶
	Akwesasne	Permission not given	¶	¶
	Lac-Rapide	Enumeration not completed – other	¶	¶
Ontario	Walpole Island 46	Permission not given	1,878	1,843
	Akwesasne (Part) 59	Permission not given	¶	¶
	Rankin Location 15D	Permission not given	566	¶
	Goulais Bay 15A	Permission not given	82	¶
	Bear Island 1	Permission not given	¶	¶
	Attawapiskat 91A	Enumeration not completed – other	¶	1,293
	Wahta Mohawk Territory	Permission not given	¶	¶
	Fort Hope 64	Natural event	1,144	1,001
	Weagamow Lake 87	Natural event	700	697
	Poplar Hill	Natural event	457	373
	Sachigo Lake 1	Natural event	450	443
	Pikangikum 14	Natural event	2100	¶
	Webequie	Natural event	614	600
	Wunnumin 1	Natural event	487	407
	Summer Beaver	Natural event	362	276
	Neskantaga	Natural event	265	0
	Bearskin Lake	Natural event	459	363
	Kasabonika Lake	Natural event	681	740
	Kingfisher Lake 1	Natural event	415	368
	Wawakapewin (Long Dog Lake)	Natural event	21	31
Manitoba	Brokenhead 4	Enumeration not completed – other	467	372
	Dauphin River 48A	Enumeration not completed – other	84	89
Saskatchewan	Big Island Lake Cree Territory	Permission not given	¶	¶
Alberta	Saddle Lake 125	Permission not given	¶	¶

Province	Incompletely enumerated Indian reserves and Indian settlements, 2011	Enumeration status (reasons for absence of data)	Population, 2006	Population, 2001
British Columbia	**Esquimalt**	Permission not given	¶	¶

2016

Province	Incompletely enumerated Indian reserves and Indian settlements, 2016	Enumeration status for the 2016 Census (reasons for absence of data)	Population, 2011	Population, 2006
Quebec	Kanesatake	Permission not given		
	Doncaster	Permission not given		
	Kahnawake	Permission not given		
	Lac-Rapide	Permission not given		
Ontario	Six Nations (Part) 40	Permission not given	946	
	Six Nations (Part) 40	Permission not given	6,213	
	Chippewas of the Thames First Nation 42	Permission not given	762	747
	Oneida 41	Permission not given	1,282	
	Wahta Mohawk Territory	Permission not given		
	Rankin Location 15D	Permission not given		566
	Goulais Bay 15A	Permission not given		82
	Pikangikum 14	Dwelling enumeration not completed – other		2,100
Alberta	Saddle Lake 125	Permission not given		
British Columbia	Esquimalt	Permission not given		

Appendix C Languages reported in Census documents, 1971-1981

Appendix D Aspects of Census Data releavant to Indigenous Languages

1) Census takes place every five years
   a) 2011 was an anomaly – technically known as the National Household Survey (NHS). Completion of the questionnaire was voluntary, which led to major concerns about data quality
   b) Present study looks at data from 2001, 2006, 2011 and 2016
2) Data suppression is used to ensure that confidentiality is maintained. Suppression of CSD data happens at two levels:
   a) All data for the CSD is suppressed if the population is less than 40
   b) Data for a particular variable is suppressed if the population count is less than 10
3) Data rounding is also used to ensure that confidentiality is maintained.
   a) All populations are reported in multiples of 5.
   b) The minimum reported population size is 10.
   c) A sophisticated algorithm is used to perform the random rounding
      i) For example, a reported population of 135 could reflect an actual population of anywhere between 131 and 139, whilst a population of 140 could reflect an actual population of anywhere between 136 and 144.
4) Country is divided into ~5,000 census subdivisions (CSDs)
   a) CSDs evolve slowly from one census to the next, primarily in response to growth of urban centres and restructuring of municipal government
   b) Big change in Indian Reserve CSDs from 2006 to 2011
      i) 2006 and earlier: a 1-to-1 relationship between reserves and CSDs
         (1) Many reserves have very small populations, leading to data suppression, especially in BC
      ii) 2011: small reserves assigned to the same First Nation were amalgamated to form larger CSDs
         (1) Data suppression was greatly reduced
   c) The census question for identifying indigenous language allows a free-form response. StatCan tries to respect these free-form responses as much as possible, but there is some translation of responses into the final categories that StaCan reports. StatCan does not describe the challenges of translating responses, but the following observations can be made:
      i) StatCan records Kaska as "Kaska (Nahani)", despite the fact that Nahani was discarded by anthropologists as a useful label many years ago (see footnote in Ray (1999, p47), quoting Beryl Gillespie). Presumably, StatCan records Nahani because elders in the Kaska communities continue to use this label when identifying their language to outsiders.
      ii) It is probable that languages such as Navajo and Cherokee, which have large speaker populations in the USA, also have a few speakers in Canada. How are they classified? Are they grouped in with "Athapaskan languages, n.i.e." and "Iroquoian languages, n.i.e.", respectively?
      iii) Many Ojibway speakers can be expected to report their language as Anishinaabemowin. There are numerous other cases where the StatCan language name is different from the name that the speakers themselves use. The diversity of alternative names and the possibility of mis-spellings can be expected to cause problems.

iv) Inuktut causes particular problems. There is considerable confusion over terminology even within the Inuit community. "Inuktut" is currently being promoted as an all-encompassing name, replacing "Inuit language". "Inuvialuktun" speakers are reported by StatCan, but this label does not have a linguistic basis. It is a term of convenience that encompasses the various dialects that are spoken within the Northwest Territories. The existence of two very distinct written forms of the language may have an impact on how people report their language. The syllabic form dominates in Nunavut and Nunavik (northern Quebec), where most people report their language as "Inuktitut", whilst the Roman alphabet form dominates further west, where it corresponds roughly to the Inuinnaqtun dialect, and also in Labrador.

v) It is not clear what languages are included in the various "n.i.e." (not included elsewhere) groups. For example, "Salish n.i.e." as reported in 2016 will include Nuxálk [Bella Coola] and She shashishalhem [Sechelt], but will it also include Salish languages whose traditional territories lie in the USA?

vi) Do some people identify themselves as speaking a language family, rather than a specific language e.g. Salish, Sioux, Iroquois? If so, then logically they should be identified as "n.o.s.", e.g. "Salish, n.o.s." But if the numbers are very small, then perhaps it is not practical to have both "n.i.e." and "n.o.s." categories for the same family.

Appendix E Reserves assigned to more than one First Nation

The following table is taken from:

https://norrisresearch.com/ref/fnp_reserves.htm

Filtering for "multiple" on the First Nation column

Prov	Reserve Number	Reserve Name
BC	17	Lillooet 1a
NB	6027	The Brothers 18
NS	6033	Franklin Manor No.22 (Part)
NS	6042	Malagawatch 4
QC	6096	Doncaster 17
ON	6143	Saugeen and Cape Croker Fishing Isl. 1
ON	6172	Abitibi 70
ON	6197	Islands In The Trent Waters 36a
ON	6207	Indian River
ON	6224	Glebe Farm 40b
ON	6225	Six Nations 40
ON	6240	Agency 30
ON	6285	Shoal Lake 34b2
ON	6336	Agency 1
ON	6338	Chippewa Island
ON	6356	Summer Beaver Settlement
MB	6400	Fishing Station 62a
SK	6622	Last Mountain Lake 80a
AB	6640	Big Horn 144a
AB	6641	Eden Valley 216
AB	6642	Stoney 142-143-144
AB	6643	Stoney 142b
AB	6660	Pigeon Lake 138a
BC	6835	Saturna Island 7
BC	6836	Pender Island 8
BC	6837	Bare Island 9
BC	6838	Goldstream 13
BC	7130	Grassy Island 17
BC	7170	Hihium Lake 6a
BC	7171	Hihium Lake 6b
BC	7172	Hihium Lake 6

Prov	Reserve Number	Reserve Name
BC	7437	St. Mary's 01a
BC	7463	Uncha Lake 13a
BC	7649	Port Essington
BC	7701	Stony Point (Fishery Bay) 10
BC	7718	Kshwan 27a
BC	7731	Lakata 41
BC	7732	Dogfish Bay 42
BC	7733	Pearse Island 43
BC	7757	Khyex 8
BC	7758	Khtahda 10
BC	7759	Scuttsap 11
BC	7761	Dashken 22
BC	7762	Kshaoom 23
BC	7763	Meanlaw 24
BC	7764	Lakelse 25
BC	7841	Red Bluff 88
BC	7856	Gander Island 14
BC	8113	Grass 15
BC	8114	Skumalasph 16
YT	8400	Haines Junction
YT	8405	Champagne Landing 10
YT	8443	Kloo Lake Settlement
SK	9063	Willow Cree
YT	9116	Klukshu River Settlement
AB	9138	Blue Quills First Nation I.R.
SK	9329	Treaty Four Reserve Grounds 77
BC	9657	Pekw'xe:Yles (Peckquaylis)
ON	9670	Assabaska
BC	10024	Tsimpsean 2a
BC	10025	Point Veitch 7
BC	10026	Willaclough 6

Appendix F First Nations with multiple traditional languages

These records are take from table FN_Lang_rev, in database db_fn_lang.accdb

Prov	Band_Number	SubID	Band_Name	Language_Name
ON	121	1	Six Nations of the Grand River	Cayuga
ON	121	2	Six Nations of the Grand River	Mohawk
ON	121	3	Six Nations of the Grand River	Oneida
ON	121	4	Six Nations of the Grand River	Seneca
ON	121	5	Six Nations of the Grand River	Onondaga
ON	121	6	Six Nations of the Grand River	Tuscarora
ON	121	7	Six Nations of the Grand River	Munsee
ON	182	1	Constance Lake	Cree, Moose
ON	182	2	Constance Lake	Oji-Cree
ON	182	3	Constance Lake	Ojibwa, Central
AB	441	1	Paul	Stoney
AB	441	2	Paul	Cree, Plains
YT	497	1	Ross River	Kaska
YT	497	2	Ross River	Slavey, South
YT	506	1	White River First Nation	Tanana, Upper
YT	506	2	White River First Nation	Tutchone, Northern

Appendix G "Hard-to-interpret" Census Data on Mother Tongue speakers

In Progress

Appendix H SQL Scripts

Combining census data into a single table

MT

INSERT INTO CSDs_Lang_Pop (Lang_Var_Code, Census_Year, csdUid, Lang_ID,

  Lang_Var, AbIdent, Sex, Pop, AvAge)

 SELECT 'MT', 2001, csdUid,

  Lang_2001_ID, Lang_Var, AbIdent, Sex, Pop_2001_uA, AvAge_2001_uA

 FROM MT_CSDs_2001_uA

INSERT INTO CSDs_Lang_Pop (Lang_Var_Code, Census_Year, csdUid, Lang_ID,

  Lang_Var, AbIdent, Sex, Pop, AvAge)

 SELECT 'MT', 2006, csdUid,

  Lang_2006_ID, Lang_Var, AbIdent, Sex, Pop_2006_uA, AvAge_2006_uA

 FROM MT_CSDs_2006_uA

INSERT INTO CSDs_Lang_Pop (Lang_Var_Code, Census_Year, csdUid, Lang_ID,

  Lang_Var, AbIdent, Sex, Pop, AvAge)

 SELECT 'MT', 2011, csdUid,

  Lang_2011_ID, Lang_Var, AbIdent, Sex, Pop_2011, AvAge_2011

 FROM MT_CSDs_2011

 INSERT INTO CSDs_Lang_Pop (Lang_Var_Code, Census_Year, csdUid, Lang_ID,

  Lang_Var, AbIdent, Sex, Pop, AvAge)

 SELECT 'MT', 2016, csdUid,

Lang_2016_ID, Lang_Var, AbIdent, Sex, Pop_2016, AvAge_2016

FROM MT_CSDs_2016

KN

INSERT INTO CSDs_Lang_Pop (Lang_Var_Code, Census_Year, csdUid, Lang_ID,

Lang_Var, AbIdent, Sex, Pop, AvAge)

SELECT 'KN', 2001, csdUid,

Lang_2001_ID, Lang_Var, AbIdent, Sex, Pop_2001_uA, AvAge_2001_uA

FROM KN_CSDs_2001_uA

INSERT INTO CSDs_Lang_Pop (Lang_Var_Code, Census_Year, csdUid, Lang_ID,

Lang_Var, AbIdent, Sex, Pop, AvAge)

SELECT 'KN', 2006, csdUid,

Lang_2006_ID, Lang_Var, AbIdent, Sex, Pop_2006_uA, AvAge_2006_uA

FROM KN_CSDs_2006_uA

INSERT INTO CSDs_Lang_Pop (Lang_Var_Code, Census_Year, csdUid, Lang_ID,

Lang_Var, AbIdent, Sex, Pop, AvAge)

SELECT 'KN', 2011, csdUid,

Lang_2011_ID, Lang_Var, AbIdent, Sex, Pop_2011, AvAge_2011

FROM KN_CSDs_2011

INSERT INTO CSDs_Lang_Pop (Lang_Var_Code, Census_Year, csdUid, Lang_ID,

Lang_Var, AbIdent, Sex, Pop, AvAge)

SELECT 'KN', 2016, csdUid,

Lang_2016_ID, Lang_Var, AbIdent, Sex, Pop_2016, AvAge_2016

FROM KN_CSDs_2016

HLA

INSERT INTO CSDs_Lang_Pop (Lang_Var_Code, Census_Year, csdUid, Lang_ID,

  Lang_Var, AbIdent, Sex, Pop, AvAge)

 SELECT 'HLA', 2001, csdUid,

  Lang_2001_ID, Lang_Var, AbIdent, Sex, Pop_2001_uA, AvAge_2001_uA

 FROM HL_A_CSDs_2001_uA

INSERT INTO CSDs_Lang_Pop (Lang_Var_Code, Census_Year, csdUid, Lang_ID,

  Lang_Var, AbIdent, Sex, Pop, AvAge)

 SELECT 'HLA', 2006, csdUid,

  Lang_2006_ID, Lang_Var, AbIdent, Sex, Pop_2006_uA, AvAge_2006_uA

 FROM HL_A_CSDs_2006_uA

INSERT INTO CSDs_Lang_Pop (Lang_Var_Code, Census_Year, csdUid, Lang_ID,

  Lang_Var, AbIdent, Sex, Pop, AvAge)

 SELECT 'HLA', 2011, csdUid,

  Lang_2011_ID, Lang_Var, AbIdent, Sex, Pop_2011, AvAge_2011

 FROM HL_A_CSDs_2011

 INSERT INTO CSDs_Lang_Pop (Lang_Var_Code, Census_Year, csdUid, Lang_ID,

  Lang_Var, AbIdent, Sex, Pop, AvAge)

  SELECT 'HLA', 2016, csdUid,

  Lang_2016_ID, Lang_Var, AbIdent, Sex, Pop_2016, AvAge_2016

  FROM HL_A_CSDs_2016

HLB

INSERT INTO CSDs_Lang_Pop (Lang_Var_Code, Census_Year, csdUid, Lang_ID,

  Lang_Var, AbIdent, Sex, Pop, AvAge)

```
  SELECT 'HLB', 2001, csdUid,

    Lang_2001_ID, Lang_Var, AbIdent, Sex, Pop_2001_uA, AvAge_2001_uA

  FROM HL_B_CSDs_2001_uA


INSERT INTO CSDs_Lang_Pop ( Lang_Var_Code, Census_Year, csdUid, Lang_ID,

    Lang_Var, AbIdent, Sex, Pop, AvAge )

  SELECT 'HLB', 2006, csdUid,

    Lang_2006_ID, Lang_Var, AbIdent, Sex, Pop_2006_uA, AvAge_2006_uA

  FROM HL_B_CSDs_2006_uA


INSERT INTO CSDs_Lang_Pop ( Lang_Var_Code, Census_Year, csdUid, Lang_ID,

    Lang_Var, AbIdent, Sex, Pop, AvAge )

  SELECT 'HLB', 2011, csdUid,

    Lang_2011_ID, Lang_Var, AbIdent, Sex, Pop_2011, AvAge_2011

  FROM HL_B_CSDs_2011


  INSERT INTO CSDs_Lang_Pop ( Lang_Var_Code, Census_Year, csdUid, Lang_ID,

    Lang_Var, AbIdent, Sex, Pop, AvAge )

   SELECT 'HLB', 2016, csdUid,

    Lang_2016_ID, Lang_Var, AbIdent, Sex, Pop_2016, AvAge_2016

   FROM HL_B_CSDs_2016
```