

Canadian Century Research Infrastructure Project

Racial Root Background

Root variable: Racial Root

Synopsis of code set construction:

This code set was constructed using the expected responses from the Enumerator Guides and the Dominion Bureau of Statistics (DBS) code books, and by consulting the microdata for imputed values DBS did not expect or did not distinguish. Any imputed value that was a sensible response was given a code.

The Racial Root code set had to account for the fact that responses to the question of “Racial or Tribal Origin” varied greatly from DBS’s expected responses. Many responses are nationalities (such as Canadian, or Mexican, for example), which were not accepted ethnic origins at the time of the censuses. The code set had to be able to separate responses that were tied to specific geographic areas from responses that were not related to a specific place, such as Jewish or Black.

The code set also had to reflect the diversity of Aboriginal origins expressed in the census data. We could not find an existing comprehensive code set of Aboriginal groups in North America. To create a code set that would distinguish between large groups that share similar languages and geographic proximity, and individual First Nations and bands, we used the following online resources:

The First Nations Information Project: <http://www.aboriginalcanada.com/firstnation/>

The Aboriginal Canada Portal:
http://www.aboriginalcanada.gc.ca/acp/community/site.nsf/index_en.html

The Connectivity profiles:
<http://www.aboriginalcanada.gc.ca/acp/site.nsf/en/ao34157.html>

The Department of Indian and Northern Affairs (Federal Government of Canada) search engine:
http://pse2-esd2.ainc-inac.gc.ca/fnprofiles/FNProfiles_home.htm

Understanding the numeric code:

This code is eight digits long.

The first digit indicates large categories of “racial” difference:

- 1 = European
- 2 = American (North, Central, South)
- 3 = African
- 4 = Middle Eastern

Canadian Century Research Infrastructure Project

Racial Root Background

5 = Asian
7 = Oceanian
8 = Jewish
9 = Other

The second digit indicates smaller, usually regions within the larger groups, which vary depending on the area.

Within group 1:

1 = Western European
2 = Central European
3 = Southern European
4 = Eastern European
5 = Northern European
6 = Baltic States
7 = Other European

Within group 2:

1 = North American
2 = Central American
3 = South American

Within group 5:

1 = Northeast Asian
2 = Southeast Asian
3 = Subcontinent

The third digit distinguishes between aboriginal and non-aboriginal groups in North America.

1 = Aboriginal
2 = Non-aboriginal

The fourth and fifth digits indicate aboriginal groups (ex: Cree) in North America.
In all other cases, the fourth and fifth digits indicate ethnic groups within the larger group.

The sixth and seventh digits indicate sub-groups within larger aboriginal groups (Ex: Swampy Cree).
In all other cases, the sixth and seventh digits indicate divisions within ethnic groups.

The eighth digit indicates a band or reserve (the lowest level of detail) in North America.

Code set application notes:

- Dual responses: Many people gave hyphenated or dual responses to the racial origin question (ex: Scottish Canadian, French Irish). We used the following principles to create a code

Canadian Century Research Infrastructure Project

Racial Root Background

set that would reflect this complexity:

- o We used the second term as the primary identity if the first term was deemed to be adjectival (if the first term described the second term). For example, when the response was “African American,” we created an “African American” child code under the American parent. Another example: “Austrian Jewish,” coded to a Jewish parent and an Austrian Jewish child code. In difficult cases, such as “Polish Russian” and “Russian Polish,” this rule takes precedence.
 - o Alternatively, we use the first term as the primary identity if the second term is thought to be employed adjectivally in common speech, as in the case of language group identities (e.g. Swiss German, or Swiss French). Swiss French normally describes a SWISS national origin and a French spoken language or cultural affinity. This would mean such values are coded to a “Swiss” parent code and “Swiss French” and “Swiss German” child codes.
 - o In cases of joint values explicitly separated by a conjunction the first mentioned ethnicity takes precedence over the second. For example, “English and German” would be coded under an English parent code.
 - o Truly uncertain cases should follow rule 3; the first of the two takes precedence as a parent code.
-
- Illegible responses: Data entry operators were instructed to enter a question mark for each illegible character in a response. These values were coded to “Illegible.”
 - A response of “Half Breed” was mapped to “Métis.”
 - American states and, more frequently, Canadian provinces were given their own codes.
 - “Other Breed” was given its own code to account for such values as “o.b.”