Data Dictionary Guide | CRA Analytics Data Tables

Board of Governors of the Federal Reserve System

March 6, 2020

Data Dictionary Framework

The CRA Analytics Data Tables group and aggregate loan data variables at the most functional levels for analysis. Notes about naming convention and organization are below:

1. The Retail Loan Table is organized differently than the Performance Evaluation (PE) Table and the Merged Data Table:
   - **Retail Loan Table: year, county, bank**
     - Each row is uniquely identified by a combination of year, county (*state_code_fips + county_code_fips*), and bank (*id_rssd*).
   - **PE and Merged Data Tables: exam period, Assessment Area (AA), bank**
     - Each row is uniquely identified by a combination of exam period (*…eval_period_begin* and *…eval_period_end*), AA (*assessment_area*), and bank (*id_rssd*).
2. There are variables with the same names across tables, but organized at table-specific levels.
   - For example, H_tL1 in Retail Loan and Merged Data Tables are both Home Mortgage Disclosure Act (HMDA) count of loans for low- and moderate- income (LMI) tracts, but H_tL1 in Retail Loan Table is at the year, county, and bank level, while H_tL1 in Merged Data Table is at the exam period, AA, and bank level.
3. All variables are bank level, unless marked aggregate ("A"). The PE Table does not have aggregate variables.
   - **Retail Loan Table**: Aggregate variables are at the *county level*, including all lenders, e.g., AC_1 is Count of CRA loans at the county level for all lenders.
   - **Merged Data Table**: Aggregate variables are at the *AA level*, including all lenders, e.g., AH_bL1 is Count of HMDA loans by LMI borrowers at the AA (identified in the row) level for all lenders.
4. All other variables not marked "A" are at the individual bank level.

Naming Convention

The variable names can be broken into four parts or positions to understand their definitions (figure 1).

- *Position 1* shows whether the variable is an aggregate variable ("A") or collected from PEs ("p"). It also reveals the data source ("H" - HMDA, "C" - CRA (sum of "sb" and "sf"), "sb" - CRA small business, "sf" - CRA small farm, "rm" - residential mortgage).
  - "p" denotes that the loan data were manually collected from the PEs and not reported directly to the FFIEC. The data are from banks that do not report CRA data (i.e., Small Banks), but may or may not report HMDA data. All "p" variables are at the individual bank level.
- *Position 2* shows whether the variable is specific to inside ("in") or outside ("out") AA. If the variable is a count *not of loans*, position 2 shows "count." If the variable is a percentage, position 2 shows "pct."

- *Position 3* contains abbreviations to denote
  - Loan Recipient Type: borrowers (b) or tracts (t)
    - If both "b" and "t" appear in the variable name, the variable contains loan values for borrowers (of a certain income level) within tracts (of a certain income level). For example, H_bLtL_3 stands for the bank's total HMDA loan amount to LMI borrowers in LMI tracts. In other words, the co-appearance of "b" and "t" is *not* additive and does *not* describe values that are "borrower *or* tract".
  - Income Categories: LMI ("L"), MUI ("M"), Other/Unknown ("O"), Low ("W"), Moderate ("R"), Middle Income Distressed ("D")
    - The last three indicate the more granular income category data; these variables are limited and few in number.
- *Position 4* shows the loan value type using numeric values
  1) Origination count
  2) Origination count for affiliates and subsidiaries (A/S)
  3) Origination amount
  4) Origination amount A/S
  5) Purchase count
  6) Purchase count A/S
  7) Purchase amount
  8) Purchase amount A/S
- Additional descriptions of non-loan objects are appended on some variables (e.g., count of families, number of tracts, owner-occupied housing).

Note: The four positions are not always filled, especially with aggregate variables or variables that are overall counts for a loan type. For instance, H_1 represents a bank's total HMDA count of loans (borrower or tract).

Position 1	
A	Aggregate
p	PE
H	HMDA
C	CRA (sb + sf)
sb	Small Business
sf	Small Farm
rm	Residential Mortgage

Position 2	
In	Inside Assessment Area
Out	Outside Assessment Area
Count	Count (not loans)
Pct	Percent

Position 3	
b	Borrower
t	Tract
L	LMI (Low to Moderate Income)
M	MUI (Middle to Upper Income)
O	Other/Unknown Income
W	Low Income
R	Moderate Income
D	Middle Income Distressed

Position 4	
1	Count of Loans
2	Count of Loans A/S
3	Amount of Loans
4	Amount of Loans A/S
5	Purchase Count of Loans
6	Purchase Count of Loans A/S
7	Purchase Amount of Loans
8	Purchase Amount of Loans A/S

A/S – Affiliates / Subsidiaries

Figure 1. Naming Convention for CRA Analytics Data Tables - *This figure shows a tabular, colored form of the four positions mentioned earlier.*

Figures 2 and 3 are graphical examples of variable names that explain how to understand a variable's definition from its name.

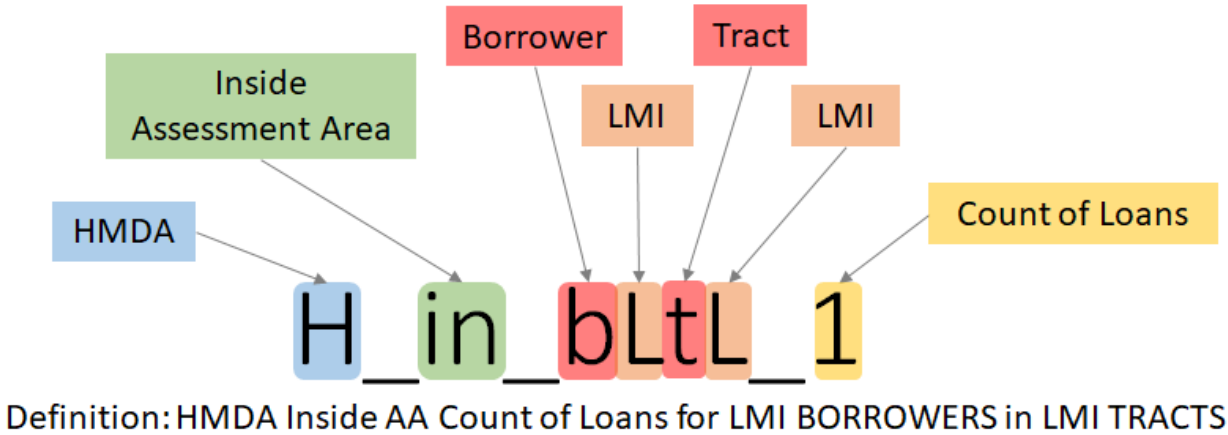

Figure 2. Bank level HMDA loan variable - *This variable has "H" to represent HMDA in position 1, "in" to represent inside AA in position 2, "bLtL" to represent borrower LMI and tract LMI in position 3, and "1" to represent that the variable values show count of loans.*

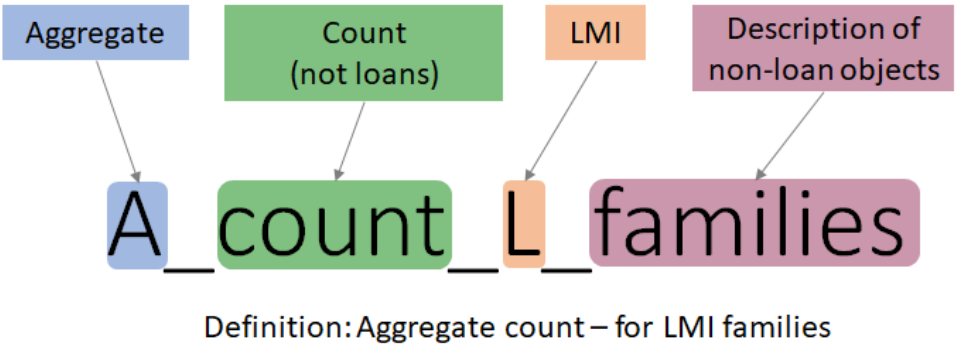

Figure 3. Aggregate count variable from census for LMI families - *This variable has "A" to represent Aggregate in position 1, "count" to represent that the value is a pure count (not of loans) in position 2, "L" to represent LMI in position 3, and "families" as additional description for non-loan objects.*