

**Claritas Update  
Demographics  
Methodology  
Summary**

© 2008 by Claritas Inc. All rights reserved.

**Warning!** The enclosed material is the intellectual property of Claritas Inc. (Claritas is a subsidiary of The Nielsen Company, a global leader in information services for the media and entertainment industries.) This manual and its contents are protected under Federal Copyright Laws, Title 17 of the U.S. Code. Under this copyright law, unauthorized users may be subject to civil liability including an injunction, actual damages, infringer's profits, and statutory damages of up to \$100,000. Criminal penalties may include a fine of up to \$25,000 and/or up to one year's imprisonment. All reproduction of this document without the expressed written permission of Claritas Inc. is strictly prohibited. By reading further in this document, you are implicitly agreeing to these terms and conditions.

*Claritas Update Demographics*<sup>™</sup> is a trademark of Claritas Inc. All rights reserved.

*TotalSource*<sup>™</sup> is a trademark of Equifax.

Other company names and product names are trademarks or registered trademarks of their respective companies and are hereby acknowledged.

# Demographic Estimation Program

This document will summarize the methodology behind the Claritas Update Demographics™. With over 30 years of experience, Claritas has a proven track record as an industry leader producing quality data. Claritas has been able to maintain this position by continually looking for new ways to innovate and improve upon their update methodology, taking advantage of new external resources that can lead to increased accuracy, and continually keeping an eye to the future as well as changes that will impact the demographic estimation process (e.g. the replacement of the Census Long Form with the American Community Survey).

This document provides a summary of the methodology process for the development of the 2008 update. For the full methodology document, please contact your Claritas account representative.

## Experienced Leadership

The Claritas demographics team is headed by Dr Ken Hodges who brings a 25-year history of groundbreaking contributions to the field of applied demography. After earning his doctorate in Sociology and Demography from Cornell University, Dr. Hodges joined Donnelly Marketing Information Services (DMIS) in 1982 before moving to Claritas in 1993.

During his tenure, Dr. Hodges has developed and refined methods for using consumer database information as input to small area estimates, and adapted standard demographic methods for use in nationwide small area estimates.

In addition to his work at Claritas, Dr. Hodge's professional involvements include the following:

- Member, 2010 Census Advisory Committee
- Board Member, Association of Public Data Users
- Board Member, Council of Professional Associations on Federal Statistics
- Member, Population Association of America Committee on Population Statistics
- Member, ICSC North America Research Taskforce

In addition to the extensive experience of Dr Hodges, the Claritas Demographics Estimation Program draws upon the strengths of some of the industry's pioneering programs including: National Planning Data Corporation, Donnelly Marketing Information Services, National Decision Systems, and Market Statistics.

## Evaluation and Support Materials

The Claritas Estimation Program is supported by extensive research and evaluation, with results often documented in professional papers. In addition to this methodology document, papers describing the following topics are available:

- Evaluation of 2000 estimates against 2000 census results.
- Evaluation of consumer database counts against 2000 census results.
- Comparison of 1996 estimates and 2001 projections from alternative suppliers.
- Evaluation of geometric data retrieval methods.

## Annual Demographic Update

The update is a shorthand term for the massive set of demographic estimates and projections produced each year for the Claritas Update Demographics product.

**Estimates** are data prepared for the current year, and **projections** (sometimes called forecasts) are prepared for dates five years in the future.

The update is brought up to date each year for many geographic levels including national, state, county, place (city/town), census tract, and block group. Data is also available for commonly-used areas such as metropolitan areas, ZIP Codes, and media areas such as DMAs. Because it is produced for small areas, the update can be easily aggregated to custom geographic areas specified by the user.

The update begins with the estimation and projection of base counts, such as total population, household population, group quarters population, households, family households, and housing units. Characteristics related to these base counts are then estimated. Population characteristics include age, sex, race, and Hispanic ethnicity. Households are estimated by age of householder and income, family households are estimated by income, and owner-occupied housing units are estimated by value.

## Methodological Changes Implemented in 2008

With over three decades of experience to learn from, Claritas is always looking for ways to innovate and improve on estimation techniques in order to adapt to a changing demographic landscape and to take advantage of new resources. In 2008, the following changes were implemented:

### Enhanced Aggregation of Estimates to ZIP Codes

In some very small, high growth ZIP Codes, rapid growth was successfully estimated for block groups, but was not being fully allocated to the correct ZIP Code. For example, growth that should have been allocated to ZIP Code A was allocated to adjacent ZIP Code B, thus leaving ZIP Code A inconsistent with residential address counts from the United States Postal Service (USPS). This enhanced aggregation provides more up-to-date factors by which to identify the percent of block group growth in ZIP Code A versus ZIP Code B.

**Note:** This enhancement did not reflect a change to the estimation method, but rather to the process used to aggregate block group estimates to ZIP Codes. For a more detailed discussion of how ZIP Code estimates are built, please see the “Zip Code Estimates and Projections” section.

### New Data Sources

Effective with the 2008 Update, ADVO data on new housing units were introduced. This source provides a powerful indication of which small areas have experienced significant housing growth in recent years, and is making valuable contributions to the Rapid Change Review process.

Also effective with the 2008 update, Acxiom data on housing values (based on parcel-level records) are contributing to the Claritas estimates of home value in selected markets.

### American Community Survey (ACS) Data

Although the Census Bureau’s American Community Survey (ACS) will not provide block group data until 2010, it is already providing valuable estimates for states and

large counties (with populations of 65,000 or more). And with new ACS estimates released every year, the Claritas Update continues to make use of ACS income and home value estimates to guide our estimates at these levels.

## **Top-Down vs. Bottom-up Methodologies**

Over the years, some descriptions of small area estimation methods have emphasized a distinction between top-down and bottom-up methodologies. The distinction is overdone, as estimates produced with pure top-down or bottom-up methods are rare. A pure top-down method might produce estimates at the county level, and then distribute to tracts and block groups without reference to any post-census tract and block group data. This approach has obvious and severe limitations for small areas. On the other hand, a pure bottom-up method would use post-census block group data to produce estimates at that level, and then sum to all larger areas, without reference to independent estimates for larger areas such as cities, counties, and states. This approach has severe limitations for larger areas.

Claritas estimates combine the strengths of top-down and bottom-up methods. Post-census data for tracts and block groups drive the estimates at those levels, while Post-census data for cities, counties, and states drive the estimates at those levels. The small area estimates are then adjusted to conform with the independent estimates for larger areas. This approach is common because it is designed to achieve maximum accuracy at all levels. Thus it cannot be described as either top-down or bottom-up.

In a bottom-up approach Claritas identifies growth trends using small data sources such as USPS, local government agencies (where available), ADVO new construction data, and the Claritas Master Address File. The Master Address File contains over 135 million records and is informed by various household level sources such as Equifax, ADVO, InfoUSA, Targus, and TeleAtlas North America.

Once the trends have been identified and estimates have been generated at the block group level, a top-down approach is taken in applying controls that will prevent the generation of unrealistic estimates at larger geographic levels such as county or state. Claritas utilizes county level data provided annually by the Census Bureau and other federal agencies as control totals for the small area estimates.

## **Trending**

To take full advantage of methodological refinements and new data resources, each set of updates begins not with the previous year's estimates, but with data from the most recent decennial census. For this reason, the difference between estimates for consecutive years is not an estimate of change from one year to the next. Change is estimated with reference to the previous census numbers.

## **Estimation Date**

The target date for estimates and projections is January 1 of the relevant year.

## Variable Categories

Claritas Update Demographics includes the categories and their respective data items listed below.

<b>Base Counts</b>	Population Households (occupied housing units) Family households (households with two or more related persons) Group quarters population (e.g., dormitories, military barracks, prisons) Housing units (house, apartment, or group of rooms intended as separate living quarters)
<b>Population Characteristics</b>	Population by age Population by sex Population by race Population by Hispanic ethnicity Population by age by sex by race by Hispanic ethnicity
<b>Household characteristics</b>	Households by income Households by size (number of persons) Age of householder Income by age of householder Households by Effective Buying Income Householders by race and Hispanic ethnicity Households by income by race of householder Households by income by ethnicity of householder Households by year householder moved into unit
<b>Housing Characteristics</b>	Total owner-occupied units Owner-occupied units by value Housing units by year structure built
<b>Means and Medians</b>	Mean and median household income Mean and median family household income Mean and median Effective Buying Income Per capita income Median age of population Median age of householders Median home value

# Geography

The Claritas Demographic Update is prepared for a wide range of geographic areas including the following:

Standard Macro Geographies	Standard Micro Geographies	Industry Geographies
Country [1]	Place [25K+]	Cable [9K+]
Divison* [9]	Tract* [60K+]	Gas Utility [1K+]
State* [51]	Block Group* [200K+]	Electric Utility [2K+]
County* [3K+]	Block* [6M+]	Wire Centers [20K+]
Designated Market Area (DMA) [200+]	Zip Code [30K+]	Yellow Pages Directory (YPD) [7K+]
Core Statistical Area (CSA) [100+]	Minor Civil Division (MCD) [Includes Census Civil Divisions (CCD)*] [35K+]	Major & Basic Trading Area Wireless (MTA & BTA) [500 +]
Core Based Statistical Area (CBSA) [900+]		Metro/Rural Service Area for Wireless (RSA) [700+]
Three-Digit Zip (TDZ) [800+]		Federal Congressional Districts [400+]

\* Census geographies

## Data Item Summary

### Base Counts

Base counts are the basic totals for population, households, family households, group quarters population, and housing units. At the national, state, county, and place geography levels, base count updates are based on estimates from the Census Bureau and, in some cases, state demographers. At the census tract and block group levels, base count information is based on sources including local estimates, trends in United States Postal Service (USPS) deliverable address counts, counts from the Claritas MAF, trends in consumer counts from the Equifax TotalSource™ database, and ADVOCOUNTS counts of new housing units.

### Population Characteristics

Population is estimated for the following demographic characteristics:

- Age
- Sex
- Race
- Hispanic ethnicity
- Age by sex by race by Hispanic ethnicity

#### **Population by Age/Sex**

Age/sex distribution is estimated using a modified cohort survival method, which ages population based on age/sex specific survival probabilities, and estimates births

over the estimation period. Group quarters and other populations that do not age in place are not aged. The method is applied first at county level, using the United States Census Bureau's most recent estimates of county population by age/sex as a starting point. Tract age/sex estimates are produced next, and [controlled](#) to the county estimates, then block group age/sex estimates are produced and controlled to tract level.

### ***Population by Race/Ethnicity***

Race by Hispanic ethnicity is estimated for 14 categories reflecting [single classification](#) race. County estimates are produced first, based on the Census Bureau's most recent county race/Hispanic estimates. Tract estimates are produced next based on 1990-2000 census trends, and are controlled to county level. Block group race/Hispanic estimates are produced next based on projected 1990-2000 census trends, and are controlled to tract level. The 1990-2000 census trends are identified through Claritas' bridging of 1990 census race data to the 2000 census race definitions. Estimates of [all-inclusive](#) race are derived from the single classification estimates through the use of Census 2000 ratios of race counts and tallies.

## **Household Characteristics**

Households are estimated for the following characteristics:

- Household income
- Household size
- Age of householder
- Race and ethnicity
- Year householder moved into unit

### ***Household Income***

Income estimates and projections reflect the census money income definition, and are produced for current dollar values. Rates of change in median income are estimated first, then the 2000 census income distributions are advanced to reflect the estimated rate of change. Income estimates at the county level and above reflect income change indicated by the Bureau of Economic Analysis (BEA) income estimates, income statistics from the Internal Revenue Service (IRS), and in large counties, income estimates from the Census Bureau's new American Community Survey (ACS).

Income change at the tract and block group levels is estimated based on a combination of:

- Change in consumer financial information from the Equifax Consumer Marketing database
- Change in income summarized from the TotalSource consumer household database
- Projections of inter-censal trends

Distributions of 2000 census income are advanced to the estimated and projected years through a process that estimates the movement of households from one income category to the next based on the specific area's estimated rate of income growth.

### ***Household Size***

The distribution of households by size starts with the 2000 census distributions, and advances them to current year based on estimated change in persons per household (average household size). Iterative proportional fitting (IPF) is then used to ensure

consistency with previously estimated household totals and average household size. For more information about [IPF](#), see the “Additional Terminology” section.

### ***Income by Age of Householder***

The income-by-age estimates are produced after those for population by age and households by income. The household-by-income estimates serve as totals for the income dimension, but persons-by-age estimates are converted to householders-by-age through the use of headship rates reflecting 2000 census householder-by-age data. The households-by-income and householders-by-age estimates serve as income and age row and column totals for the estimated income by age table. Cell values (specific income-by-age categories) are estimated through iterative proportional fitting of Census 2000 income-by-age data to the estimated income and age totals. This process yields income-by-age values that not only sum to the income and age estimates, but also preserve the statistical relationship between income and age for each area as measured by the census.

## **Housing Unit Characteristics**

Housing units are estimated for the following characteristics:

- Total count of owner-occupied units
- Value of owner-occupied units
- Year structure built

### ***Housing Value***

Housing value is estimated for all owner occupied housing units. As with income, the method begins with the estimation of a rate of change, which is then used to advance the 2000 census distribution to current and projection year.

At the state and national levels, target rates of change in value are based on change in value estimated by the 2006 ACS, as well as change in the House Price Index from the Office of Federal Housing Enterprise Oversight (OFHEO).

At county level, the OFHEO data is combined with change in median sales price data from the National Association of Realtors to estimate change. An additional data source contributing to the estimates in counties with large populations, is the annual data from the ACS that is currently available for counties with populations in excess of 65,000. Tract rates of change are based on a combination of projected inter-censal trends and post-2000 change in average mortgage amounts from the Equifax Consumer Marketing database. And in some markets, home value data from the Axiom files contribute to the small area estimates of home value.

As with income, estimated rates of change are used to advance the 2000 census distributions to current year. The national and state rates serve only as targets (not control totals) for the county estimates, while the tract and block group estimates are both controlled to the next higher level.

## **Smoothed Data**

In addition to the annual demographic estimates and projections, the update provides a series of detailed census tables that have been ratio adjusted, or smoothed, to relevant current-year totals. These tables purport only to show the effect of applying decennial census distributions to estimated base count totals at the block group level.