



Nielsen Demographic  
Update  
Summary  
Methodology  
Document

May 2010



© 2010 The Nielsen Company. All rights reserved.

**Warning!** The enclosed material is the intellectual property of The Nielsen Company (US), LLC. This manual and its contents are protected under Federal Copyright Laws, Title 17 of the U.S. Code. Unauthorized users may be subject to civil and criminal liability, including statutory damages of up to \$100,000, criminal fines of up to \$25,000 and/or imprisonment. Reproduction of this document without the express written permission of The Nielsen Company (US), LLC is strictly prohibited.

The Nielsen logo is a trademark or registered trademark of CZT/ACN Trademarks, L.L.C.

*Pop-Facts*<sup>™</sup> is a trademark of The Nielsen Company (US), LLC. All rights reserved.

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

# Nielsen Demographic Update

This document will outline the methodology behind the 2010 Nielsen Demographic Update. With over 30 years of experience, Nielsen has a proven track record as an industry leader producing quality data. Nielsen 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. One example of this was the replacement of the Census Long Form with the American Community Survey (ACS).

This document will outline the methodology process for the development of the 2010 update that includes household and population data covering over 4,000 demographic variables from Census 1990 and Census 2000 as well as current year estimates and five-year projections.

## Experienced Leadership

The Nielsen 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, now part of The Nielsen Company, 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 Nielsen, Dr. Hodges' professional involvements include the following:

- Member, 2010 Census Advisory Committee
- Co-Author, *A Compass for Understanding and Using American Community Survey Data: What the Business Community Needs to Know*
- 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 Nielsen 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 Nielsen Demographic Update 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 Census 2000 results.
- Evaluation of consumer database counts against Census 2000 results.
- Comparison of 1996 estimates and 2001 projections from alternative suppliers.
- Evaluation of geometric data retrieval methods.

## Demographic Update

The update is a shorthand term for the massive set of demographic estimates and projections produced multiple times per year for the Nielsen Pop-Facts products. Estimates consist of 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 multiple times per 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. Owner-occupied housing units are estimated by value.

## Recent Methodological Changes and New Data Sources

With over three decades of experience to learn from, Nielsen 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 recent updates, the changes below were implemented.

### Enhanced Aggregation of Estimates to ZIP Codes

In 2008, 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.

### New Data Sources

The 2009 update introduced Hanley Wood data on new residential development. Identifying the construction and sale of new residential housing units, this source is a

valuable addition to the resources for estimating population and housing growth at the neighborhood level.

Also introduced in 2009 was Hispanic population data from Ethnic Technologies (E-Tech). E-Tech Hispanic counts provides a powerful new tool for estimating changes in Hispanic composition at the neighborhood level. Data from E-Tech is used to enhance the estimation of Hispanic ethnicity in areas that have experienced the most rapid growth of this population group. The E-Tech Hispanic counts—based on analyses of first names, surnames, prefixes, and suffixes against name/address database resources—are not used as is, but as a way to identify tracts where preliminary Hispanic estimates need to be increased.

## **American Community Survey Data**

Although the ACS will not provide block group data until 2011, it was already providing valuable estimates for states and large counties (those with populations of 20,000 or more) as early as 2009. And with new ACS estimates released every year, the Nielsen Demographic Update continues to make use of ACS estimates of median income and home value to guide estimates at these levels.

Effective with the 2010 update, the ratio-adjusted data was significantly upgraded through the use of control totals based on the most recent one-year and three-year estimates from the ACS. The 2010 Nielsen Demographic Update represents the most extensive implementation of one-year and three-year ACS data inputs to date. A list of ratio-adjusted data items is included later in this document.

## **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.

Nielsen 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 Nielsen identifies growth trends using small data sources such as USPS, local government agencies (where available), new construction data, Hanley Wood residential development counts, and the Nielsen Master Address File (MAF). The Master Address File contains over 135 million records and is informed by various household level sources such as Equifax, Valassis, and InfoUSA.

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. Nielsen 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

The 2010 Nielsen Demographic Update includes the categories and data items listed below.

Population & Race	<ul style="list-style-type: none"> <li>--Base Count Population, Households, Housing Units, Families, Group Quarters</li> <li>--Population by Age, Sex, Race, Ethnicity</li> <li>--Population by Ancestry/Origin</li> <li>--Population by Language Spoken</li> <li>--Population by Marital Status</li> <li>--Means and Medians</li> </ul>
Housing & Household Composition	<ul style="list-style-type: none"> <li>--Households by Age, Race, Ethnicity, and Tenure</li> <li>--Housing Units by Value</li> <li>--Households by Size (number of persons) and Presence of Children</li> <li>--Housing Units by Year Built</li> <li>--Housing Units by Number of Units</li> <li>--Households by Year Householder Moved In</li> <li>--Households by Vehicles Available</li> <li>--Means and Medians</li> </ul>
Affluence & Education	<ul style="list-style-type: none"> <li>--Households by Income, Age, Race and Ethnicity of Householder</li> <li>--Effective Buying Income (EBI)</li> <li>--Household Income by Type</li> <li>--Families by Poverty Status</li> <li>--Population Educational Attainment by Sex and Ethnicity</li> <li>--Means and Medians</li> </ul>
Employment & Occupation	<ul style="list-style-type: none"> <li>--Population by Employment Status and Sex</li> <li>--Population by Transportation to Work</li> <li>--Population by Travel Time to Work</li> <li>--Population by Class of Worker</li> <li>--Population by Industry and Occupation</li> <li>--Means and Medians</li> </ul>

# Geography

The 2010 Nielsen Demographic Update is prepared for a wide range of geographic areas.

Standard Macro Geographies	Standard Micro Geographies	Industry Geographies
Country* [1]	Place [25K+]	Cable [9K+]
State* [51]	Tract* [60K+]	Wire Centers [20K+]
County* [3K+]	Block Group* [200K+]	Yellow Pages Directory (YPD) [7K+]
Designated Market Area (DMA) [200+]	Block* [8M+]	Major & Basic Trading Area Wireless (MTA & BTA) [500 +]
Combined Statistical Area (CSA) [100+]	Zip Code [30K+]	Metro/Rural Service Area for Wireless (RSA) [700+]
Core Based Statistical Area – Metropolitan/Micropolitan (CBSA) [900+]	Minor Civil Division (MCD) [Includes Census Civil Divisions (CCD)*] [35K+]	
Three-Digit Zip (TDZ) [800+]		

\* 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 Nielsen Master Address File, Hanley Wood new residential development data, and Valassis 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 E-Tech Hispanic population counts, 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 Nielsen bridging of Census 1990 race data to the Census 2000 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 Census 2000 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 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 Census 2000 income are advanced to the estimated and projected to target dates 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 Census 2000 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 Census 2000 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 number 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 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 2008 ACS, as well as change in the House Price Index from the Federal Housing Finance Agency (FHFA). FHFA is the federal agency formerly known as the Office of Federal Housing Enterprise Oversight (OFHEO).

At the metropolitan area level, the FHFA data is combined with change in median sales price data from the National Association of Realtors (NAR) 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 20,000. Tract rates of change are estimated based on a combination of projected inter-censal trends, post-2000 change in average mortgage amounts from the Equifax Consumer Marketing database, and home value data from Acxiom files.

As with income, estimated rates of change are used to advance the Census 2000 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.

## **Ratio-Adjusted**

In addition to the demographic estimates and projections, past updates have provided a series of detailed Census 2000 tables that have been ratio adjusted to relevant current-year totals. Effective with the 2010 update, most of these tables have been enhanced with the most recent ACS estimates for counties with populations of 20,000 or more.