## Demographic Reference Guide

## Introduction to Demographics <br> About Demographics

Demographics is the study of physical characteristics about people. The word comes from the Greek word for people, demos. As an academic discipline, it is generally considered to be part of sociology.

Technically demographics refer to geo-demographics because of the relationship of people to place. At the most basic level, demographics counts people in geographic places. Once the total population is counted, then attributes about those people can be collected and quantified. Attributes such as age, sex, education level, marital status, occupation are all characteristics of people.

Households are also part of demographic study. In this case, it is characteristics of particular households that are captured and reported.

Demographic data is always related to some kind of geographic space.

## Decennial Census

Article 1 of the US Constitution requires that a census (a count) of all citizens be taken every 10 years (thus decennial census) to reapportion seats in the U.S. House of Representatives. A census form is provided (in theory) to every U.S. household where one person representative of the household completes a set of questions about the occupants of the household.

The Census Bureau historically has fielded two kinds of census counts referred to as the "short form" and the "long form." The short form is to go to every household. The long form which is much more extensive, is a sample.

## Standard Census Geographies

The Census Bureau developed and maintains a hierarchical system of geographic areas. Each level aggregates to the next level up in a fairly consistent manner. (There are exceptions, but they are rare.) The Census Bureau graphic illustrates the core hierarchy (running down the middle). It also shows the relationship of other geographic areas such as legislative districts, zip codes and places.

In this Guide, not all Census Bureau geographies are discussed. We only focus on those built into MissionInsite. For more information on other geographies, visit the Census Bureau's

website.

The United States National Boundary

States
Counties

Census Tracts

This boundary includes all geographic areas within the national boundaries of the United States of America.

The first major division is into states.
States are divided into counties or parishes, depending upon the particular state.
In Louisiana, these subdivisions are known as parishes. Alaska has no counties however the county equivalents are boroughs In four states (Maryland, Missouri, Nevada and Virginia), there are one or more cities that are independent of any county and thus constitute primary subdivisions of their states. The District of Columbia has no primary divisions, and the entire area is considered equivalent to a county for statistical purposes.

Each county is subdivided into census tracts.
Census tracts usually contain between 1,500 and 8,000 people with a target size of 4,000. Census tracts are set at each decennial census. Many never change because they exist in established population areas.
On urban edges and country areas, census tracts can be quite large since they are drawn in such a way as to come as close as possible to the targeted population threshold. This can create problems in period between a decennial census if new residential development occurs. While the problem will be corrected at the next census when a census tract is subdivided to reflect the new population reality, spatial queries in the interim
period can fail to accurately capture the real profile of a particular census tract.
The following illustration demonstrates the problem.


Block Groups

Blocks

## Other Geographies

Zip codes

Census Block groups are standard Census Bureau geographies. Typically 4 to 6 block groups aggregate to form a census tract.
A block group is the lowest level of geography for which census data is released-for privacy purposes. Block groups generally contain between 600 and 3,000 people with a targeted optimum size of 1,500. Block groups never cross the boundaries of states, counties, or statistically equivalent entities, except for a block group delineated by American Indian tribal authorities. They never cross the boundaries of a census tract.
The Block Group is the lowest level that the Census Bureau tabulates sample (long form, for example) data.

A block is the smallest geographic unit for which the Census Bureau tabulates 100-percent data. In urban areas, blocks usually follow city blocks, bounded by streets. This rule does not follow in rural areas where a block may encompass multiple square miles and my not be bounded by streets.
Over 11 million blocks are identified for Census 2010. The Census Bureau does not release data at the block level to protect privacy

Technically, zip codes are not geographic areas. Created by the Postal Service, they represent carrier delivery routes. The Census Bureau developed Zip Code Tabulation Areas (ZCTAs) as a way to create geographic areas that approximate the carrier routes reflected in the Postal Service zip codes. This allows geographic information systems to conduct spatial queries on an area that approximates a zip code delivery area. These boundaries are updated a least once per year to reflect the changes of zip codes.


Relative size of zip codes and standard census geographies.

## Cities

School Districts
Cities are incorporated places with municipal boundaries. These are public domain boundaries and may or may not reflect current reality for municipalities. You can query on cities in Missionlnsite.

The Census Bureau maintains a set of boundaries for school districts. These also are public domain boundaries and are not always accurate or up-to-date. The boundaries are available for integration into GIS Systems.
There are three different school district boundary files reflecting the diversity of school district configurations in US. They are: Unified School Districts, Elementary School Districts and High School Districts. These layers can be found in MissionInsite.

## Demographic Analysis

Demographic analysis has two basic steps. First, one must decide determine the geographic area about which information is sought. Second, one looks up that particular data. One can go to the library or increasingly online to find out information on geographies such as counties, zip codes, cities, etc. Several websites allow one to type in their zip code and get some kind of demographic data back. There are limitations to some sites.

Within MissionInsite, you can access geographies such as zip codes, counties, and cities and well as user-defined areas such as an agency boundary or sub-boundaries. Additionally, you can define your own custom geographic areas and derive demographic information about those areas.

## Demographic Data Retrieval

Unless the area you want data on perfectly matches one or more block group boundaries, obtaining demographic data requires allocating some portion of a block group to within the
geographic area to get demographic totals. This process is called a 'spatial query' or demographic data retrieval.

MissionInsite retrieves data within a geographic area by spatially determining all of the full or partial block groups included within the query area. These data are then aggregated to provide totals. When a queried area includes one or more partial block groups, a calculation allocates some portion of the block group data to the aggregated total.

Demographic information systems use three different approaches for data retrieval, and Mission Insite uses the most advanced model for the greatest sensitivity to changes on the ground and in small geographic areas.

## Structure of the System and Resource Guide

Dozens of demographic variables from multiple sources are available in MissionInsite. Some are estimates and projections of the last decennial census but MissionInsite includes data from other sources as well.

Most demographic data falls under one of the following macro categories:

- Population
- Households
- Housing
- Mosaic

This guide's structure follows these macro categories. To find a specific variable, look under its macro category.

## Data Sources

MissionInsite's demographic provider, Synergos Technologies Inc. (STI), derives information from multiple sources and is updated twice a year.
\(\left.$$
\begin{array}{ll}\text { Bureau of Economic Analysis (BEA) } & \begin{array}{l}\text { The BEA promotes a better understanding of the U.S. } \\
\text { economy by providing timely, relevant, and accurate } \\
\text { economic accounts data. The BEA is an agency of the } \\
\text { Dept. of Commerce. Along with the Census Bureau and } \\
\text { STAT-USA, BEA is part of the Department's Economics } \\
\text { and Statistics Administration }\end{array}
$$ <br>
\& The BLS is an independent national statistical agency <br>
that collects, processes, analyzes, and disseminates <br>

labor economics and statistics data to the public, the\end{array}\right\}\)| U.S. Congress, other federal agencies, state and local |
| :--- |
| governments, business, and labor entities. Among |
| the data used for PopStats is data from the BLS's |
| Local Area Unemployment Statistics (LAUS) program, |
| which produces monthly and annual employment, |
| unemployment, and labor force data for Census regions |
| and divisions, states, counties, metropolitan areas, and |
| many cities, by place of residence. |


| Centers for Disease Control (CDC) | The CDC's mission is to collaborate to create the |
| :--- | :--- |
|  | expertise, information, and tools that people and |
|  | communities need to protect their health - through |
|  | health promotion, prevention of disease, injury and |
| disability, and preparedness for new health threats. |  |
|  | Among the data used for PopStats are the CDC's |
|  | natality and mortality files. |
|  | The DOD is the federal department charged with |
|  | coordinating and supervising all agencies and functions |
| of the government relating directly to national security |  |
| and the military. The DOD has three major components |  |


|  | the U.S. and other nations. NCES is located within the U.S. Dept. of Education and the Institute of Education Sciences. Among the data used for PopStats are the NCES's public and private records. |
| :---: | :---: |
| National Center for Health Statistics (NCHS) | The NCHS provides U.S. public health statistics, including diseases, pregnancies, births, aging, and mortality. It is a division of the CDC. |
| National Parks Service (NPS) | The NPS is the U.S. federal agency that manages all national parks, many national monuments, and other conservation and historical properties. It is an agency of the U.S. Dept. of the Interior, a federal executive department. Among the data used for PopStats are the NPS's park attendance records. |
| Social Security | U.S. Social Security is a social insurance program funded through dedicated payroll taxes called the Federal Insurance Contributions Act (FICA). Tax deposits are formally entrusted to several funds, including primarily the Federal Old-Age and Survivors Insurance Trust Fund. |
| U.S. Census Bureau | As part of the U.S. Dept. of Commerce, the Census Bureau serves as a leading source of data about America's people and economy. The most visible role of the Census Bureau is to perform the official decennial count of people living in the U.S. Public resources from the Census Bureau include population, economic, industry, and geography studies. Along with population data, several reports from the U.S. Census Bureau are used for PopStats, including the American Community Survey (ACS) and the Current Population Survey (CPS). |
| U.S. Postal Service (USPS) | The USPS is an independent agency of the U.S. government responsible for providing postal service. It is one of the few government agencies authorized by the U.S. Constitution. |

## About the Resource Guide

This Guide is a compilation of information from multiple sources. All census variable definitions are from the Census Bureau "Factfinder Glossary." In most cases, the definitions have been captured as provided by the Bureau with some little editing to make the information easier to understand for nontechnical folks.

MOSAIC definitions were acquired from either Tetrad Computer Applications and/or AGS/ Experian. Again, most definitions are provided as written by AGS/Experian. This ensures information accurately reflects the intent of each MOSAIC type and/or group. Finally, several variables are provided by Synergos Technologies, Inc. (STI).

All of this information is easily accessible on either the Census Bureau, STI or AGS websites. In some cases, footnotes to specific web pages have been provided. MissionInsite does not claim original authorship of this information.

## Types of Reports

Single Year Reports

Trend Reports

Forecast Reports

Quarterly Reports

These reports provide a single year data point. In most cases, it is the year of the most recent demographic data update.

Trend reports typically look back at history and forward to show the trends coming from the past and extending into the future. Some may go back as far as the two prior decennial censuses and forward five years into the future.

Forecasts are reports that begin with the current year estimate and look forward 10 years, depending on the data available.

Quarterly reports provide a close up view of recent changes. Most look back eight (8) quarters from the most recent demographic update.

## Demographic Variables and Descriptions

Population is at the most basic level of demographics. This includes total population counts as well as population counts by a multitude of population characteristics and attributes such as racial/ ethnicity, age, educational attainment etc.

## Population Based Variables

Population based variables are all demographic data that are based upon population counts. Data is organized according to the following sub-categories and includes:

- Population
- Age
- Education
- Marital Status
- Housing
- Race and Ethnicity
- Employment


## Population

Population is the number of persons counted at their place of usual residence, the place where the person lives and sleeps most of the time.

Five data points are provided including:

- 2000
- 2010
- Current year estimate
- Five-year projection
- Ten-year forecast when available


## Components of Change (past 12 months w/o Group Quarters)

Components of Change is an analysis of household population change over the past twelve months from the current estimate. This analysis shows changes in population that are attributable to births, deaths, and net migration. This analysis does not include group quarter population changes.

The three elements of the Components of Change are based upon the prior 12 months from the most recent update. They are reported as whole numbers.

Categories include:

- Births (past 12 mths)
- Deaths (past 12 mths$)$
- Migration (past 12 mths)

These three factors determine net changes in population. Births add to the population, and deaths subtract from it. Net migration is added to the result. Net migration is in-migration (people moving into an area) minus out-migration (people moving out of an area). In some cases, net migration can be negative because more people moved out than new people moved in.

Formula: (Births minus deaths) plus (in-migration minus out-migration) = Change in population.

## Population Change Index: Projected to Actual

The Population Change Index: Projected to Actual (PCI) is an indicator that qualifies the type of growth or decline that has occurred in recent quarters. It is a measure of the comparison between expected and actual. This variable is called Expected Value Index (EVI) by STI.

PCI may take on a value ranging from -2 to +2 . The sign represents the direction of the data. A positive means growth and a negative means decline. With respect to a growth situation, a value centering around one indicates that the expected growth and actual growth are essentially equal.

A value between 0 and 1 means actual growth is beginning to taper off (the possible beginnings of a growth plateau). A value between 1 and 2 means growth is accelerating beyond expectations (the possible beginnings of a new growth cycle). A value of zero means no population or NA (Not Applicable). Negative values can be interpreted as the inverse of positive growth.

Categories include:

- >1 Growth accelerating
- 1 = Growth as expected
- >0 to 1 Growth tapering off
- <0 to-1 Decline tapering off
- 1 = Decline as expected
- >-1 Decline accelerating

A community expects growth through residential development, but the PCI indicates to what extent the actual growth or decline conforms to what was expected.

There are three possibilities for growth and three for decline.

- An area may be growing faster than expected, so growth is accelerating. An area's growth may slow down and taper off so that the growth is less than expected, or it can be growing as expected.
- An area may be declining faster than expected, so the decline is accelerating. An area's decline may be slow down and taper off so that the decline is less than expected, or it can be declining just as expected.
- If zero (0) is returned, this means no population.


## Population: Recent 8 Quarter History

Quarterly history of population change reflects the estimated changes over the prior eight quarters and as of the most recent demographic update.

Data is reported by quarter as a historical trend. The value of this variable is its sensitivity to immediate changes in population. The standard current year estimate and five and 10 year projections provide the longer view but this variable shows more immediate historical changes at a quarter by quarter level.

## Seasonal Population: Recent 8 Quarter History

Seasonal population is defined as population that resides in a housing unit specifically designated as seasonal housing. Unlike regular housing, seasonal housing is used only for a specific season, like a summer cottage or winter chalet. The unit is typically vacant during the other times of the year.

Seasonal housing must not be confused with second homes. Second homes are typically occupied at various times of the year as opposed to being limited to a specific season. Although all seasonal housing might be considered second homes, all second homes are not necessarily seasonal (and therefore, most importantly, not counted in this estimate). It is reported as the average occupancy for the quarter, not daily as the Transient Population.

Seasonal population resides in a location between six weeks and six months. It can include second homes or some migrant housing, though it is not possible to know aside from knowing an area on the ground. For example, in a community that is known to employ migrant workers, seasonal population may reflect that reality. In an area known for vacation homes, it may tip in that direction.

Seasonal Population is determined by looking at quarterly estimates from the Bureau of Labor Statistics. There is a high correlation between certain industries and their hiring practices with the seasonality of an area. This correlation helps determine seasonal population shifts.

## Transient Population: Recent 8 Quarter History

Transient population is defined as that population which resides in a hotel, campground, or RV park (Recreational Vehicle) for at least one (1) night or up to six weeks but no more.

This figure is exclusive from all other data in MissionInsite. You will not see this number reflected in any of the breakout variables such as age, income, or ethnic makeup. This value is not a part of
the overall population estimate. It is reported as the average daily occupancy for the quarter. Data is reported as a table covering the prior eight (8) quarters.

The estimate is determined based on STI's proprietary model and information from the American Travel Survey (produced by the Bureau of Transportation Statistics), flight statistics from the FAA (Federal Aviation Administration), park attendance from the National Parks Service, and other local/government sources. This is a mathematically determined estimate and not based on direct observation or survey results.

## Population by Gender

Population by gender refers to a person's sex. Individuals were asked to mark either "male" or "female" to indicate their sex.

For most cases in which sex was not reported, it was determined from the person's given (i.e., first) name and household relationship. Otherwise, sex was imputed according to the relationship to the householder and the age of the person. Four data points are reported.

Categories reported:

- Female
- Male


## Age

## Population by Average Age

Average age is calculated by dividing the total aggregate ages of all persons in a defined geography by the total population in that same geography.

## Population by Age

Age is based upon date of birth information as of the most recent decennial census. Updates and projections to age are generated by cohort aging methods.

These data are reported in multiple ways, including:

- Detailed age groupings
- Phase of Life
- School Aged Population
- Outreach Opportunities (Children, Preschool, Youth and Young Adult)


## Population by Phase of Life

Phase of Life re-categorizes the age detail data into seven life stage characteristics. These phase categories attempt to relate age periods to life experiences common to the age ranges.
Categories reported:

- Before Formal Schooling: Ages 0 to 4
- Required Formal Schooling: Ages 5 to 17
- College/Career Starts: Ages 18 to 24
- Singles and Young Families: Ages 25 to 34
- Families and Empty Nesters: Ages 35 to 54
- Enrichment Years Singles/Couples: Ages 55 to 64
- Retirement Opportunities: Age 65 and over


## Education

## Population: Currently Enrolled in Education

The number of persons currently enrolled in public or private school by level. This data is reported by the National Center for Education Statistics and used to set the cohort aging for school-aged children in the age variable. Data reported as a percentage of the total population enrolled in some kind of educational program.

Categories reported:

- Nursery school/preschool
- Kindergarten/Elementary School
- High School
- College/Graduate/Professional school


## Population: 25+ by Educational Attainment

The educational attainment level of persons 25 years of age and greater. Census data on schooling completed reflects self-reported information on the highest level of school completed or the highest degree received.

High school graduates include those who received their diplomas or the equivalent (GED for example) and did not attend college. Graduate/Professional degrees include those in medicine, dentistry, law, pharmacy, and chiropractic.

Degrees from vocational, trade or business schools were not included unless they were collegelevel degrees. Degrees from barber schools, cosmetology schools and the like were specifically excluded from the professional school category. Three data points are reported.

Categories reported:

- Less than 9th Grade
- Some High School, No diploma
- High School Graduate (or GED)
- Some College, No degree
- Associate Degree
- Bachelor Degree
- Graduate or Professional School degree


## Marital Status

## Population by Marital Status 15+

The marital status refers only to the population 15 years old and over. Three data points are reported.

Categories reported:

- Married
- Divorced
- Separated
- Never Married
- Widowed


## Population by Marital Status 15+: Single Female

The marital status single female refers only to the unmarried female population 15 years old and over. Three data points are reported.

Categories reported:

- Divorced
- Never Married
- Widowed


## Population by Marital Status 15+: Single Male

The marital status single male refers only to the unmarried male population 15 years old and over. Three data points are reported.

Categories reported:

- Divorced
- Never Married
- Widowed


## Housing

## Population by Household Type

Population by household type reports the total population by the type of household structure where people dwell. There are three categories and three data points.

Categories reported:

- Population in Family Households
- Population Group Quarters
- Population in Non-Family Households


## Population: Group Quarters

Population in group quarters includes all people not living in households.
Two categories of people are included 1) the institutionalized population which includes people under formally authorized supervised care or custody in institutions (such as correctional institutions, nursing homes, and juvenile institutions) and 2) the non-institutionalized population which includes all people who live in group quarters other than institutions (such as college dormitories, military quarters, and group homes).

Categories reported:

- Institutionalized
- Non-Institutionalized: College
- Non-Institutionalized: Military
- Non-Institutionalized: Other


## Race and Ethnicity

## Diversity Index

This index measures the homogeneity of race in a Census Block Group. The measure does not tell which race is the dominate group, and you need to rely on the actual race data to make that determination. This index includes the ethnic class Hispanic as a separate class in its determination.

Diversity is a function of both the percentage represented of each group and the number of groups represented. The Index can take on a value of 0 to 1 , where 1 indicates there is only one race (or ethnic group) represented in that block group. The closer to zero the score is, the more diverse the geographic area is. The Diversity Index is a quick way of determining the level of racial/ ethnic diversity of a defined geography.

The data is reported by the following categories:

- Very diverse
- Somewhat diverse
- Somewhat homogeneous
- Very homogeneous


## Population: Racial/Ethnic Trends

Population by Racial/Ethnic Trends presents five different racial or ethnic categories. It includes three data points. It is a fairly typical practice to provide a high level variable called "Race/ Ethnicity" that rolls up some of the smaller ethnic groups and includes Hispanic or Latino.

People can indicate any race and also Hispanic or Latino. To create this specific report, the racial groups (White, Black/African American, Asian, Pacific Islander/American Indian/Alaska Native/ Other/Some other race and two or more races) are reported as "alone" allowing those persons to be reported separate from those who did indicate Hispanic or Latino.

You can generate a Hispanic/Latino category within the Race/Ethnicity report without double counting.

Categories reported:

- Asian (Non-Hisp)
- Hispanic or Latino
- Black/African American (Non-Hisp)
- Pac Is/Am Ind/Alaska Nat/Oth (Non-Hisp)
- White (Non-Hisp)


## Population: Asian Alone

The Asian Alone variable reflects the self-identification among people of Asian descent.
Categories reported:

- Asian Indian
- Cambodian
- Chinese, except Taiwanese
- Filipino
- Hmong
- Japanese
- Korean
- Laotian
- Other Asian
- Thai
- Vietnamese


## Population by Ancestry

Ancestry data represents self-classification by people according to the ancestry group or groups with which they most closely identify. Ancestry refers to a person's ethnic origin or descent, "roots," heritage, or the place of birth of the person, the person's parents, or their ancestors before their arrival in the United States.

Not all respondents completed this field. As a result, most Ancestry reports create a large group on "Ancestry Unclassified."

Categories reported:

- American
- Scotch Irish
- Central American
- British
- Chinese
- Dutch
- Cuban
- Italian
- Dominican
- Mexican
- French
- Native American (Indian/Eskimo)
- German
- Other
- Hawaiian/Pacific Islander
- Other European (e.g. Greek/Russian)
- Japanese
- Polish
- Korean
- Puerto Rican
- Middle Eastern
- South American
- Other Asian
- South Central Asian (e.g. Indian)
- Other Hispanic
- South East Asian (e.g. Vietnamese)
- Scandinavian
- Unclassified

Note: Puerto Rico is not included.

## Population Hispanic or Latino by Origin

The terms "Spanish," "Hispanic origin,'" and "Latino" are used interchangeably. Some respondents identify with all three terms, while others may identify with only one of these three specific terms.
Hispanic or Latino Origin can be viewed as the heritage, nationality group, lineage, or country of birth of the person or the person's parents or ancestors before their arrival in the United States. People who identify their origin as Spanish, Hispanic, or Latino may be of any race.

Categories reported:

- Cuban
- Mexican
- Puerto Rican
- Other Hispanic or Latino


## Population 5+ by Language Spoken at Home

Language spoken at home is for the population five (5) years of age and older and reflects the language of the home environment. Two reports are provided, a summary of the major groups and a detail report with the sub-groups. A single data year is given.

Categories include:

- Speak only English
- Spanish or Spanish Creole
- European/Indo\#European
- Asian/Pacific
- Other Languages


## Employment

Employed includes all civilians 16 years old and over who were either (1) "at work" -- those who did any work at all during the reference week as paid employees, worked in their own business or profession, worked on their own farm, or worked 15 hours or more as unpaid workers on a family farm or in a family business; or (2) were "with a job but not at work" -- those who did not work during the reference week but had jobs or businesses from which they were temporarily absent due to illness, bad weather, industrial dispute, vacation, or other personal reasons.

Excluded from the employed are people whose only activity consisted of work around the house or unpaid volunteer work for religious, charitable, and similar organizations. People on active duty in the United States Armed Forces are also excluded.

## Employed Civilian Population 16+ by Occupation

Civilian (i.e. excluding military) population 16+ by occupational categories. A single data point is reported.

Categories reported:

- Bldg Maintenance \& Cleaning
- Construction
- Farming, Fishing, \& Forestry
- Food Preparation Serving
- Healthcare support
- Managerial executive
- Office Admin
- Personal Care
- Production Transportation
- Prof specialty
- Protective
- Sales


## Employed Civilian Population 16+ Blue/White Collar

Civilian population $16+$ segmented between traditionally consider blue and white collar occupations. A single data points is reported.

Categories reported:

- Blue Collar - typically labor intensive in either the production or service sectors.
- White Collar - typically professional, managerial, sales and administrative support.


## Employed Population 16+ by Sector Current Year

Estimated current year population 16+ by sector.
Categories reported:

- For-Profit Private Workers
- Non-Profit Private Workers
- Self-Employed Worker in Own Business
- Federal Government Workers
- State Government Workers
- Unpaid Family Workers
- Local Government Workers


## Population: 16+ by Employment Status Current Year

Employment status by two data points.
Categories reported:

- In labor force: Civilian: Employed
- Not in labor force
- In labor force: Civilian: Unemployed
- In labor force: In Armed Forces


## Unemployment

This variable presents historical data of unemployment rates. As reported, they are based on the census bureau's definition of unemployment, which differs significantly from the Bureau of Labor Statistics' (which is what is commonly reported by the media). The rates are not seasonally adjusted and are reported by the eight most recent quarters.

By providing the rates by the eight most recent quarters, one is able to discern the pattern of unemployment in a defined geographic area. If the unemployment rate has been increasing in the historical quarters, this provides an insight into the health of the economy in that area.

## Workers 16+: Home or Away

Location of workers 16+. A single data point is reported.
Categories reported:

- Worked at Home
- Worked away from Home


## Workers 16+: Transport to Work

Means of transportation to work refers to the principal mode of travel or type of conveyance that the worker usually used to get from home to work. Data were tabulated for workers 16 years old and over. A single data point is reported.

Categories reported:

- Bicycle
- Car, truck, or van: Carpooled
- Car, truck, or van: Drove alone
- Motorcycle
- Other Means
- Public Transportation
- Walked


## Workers 16+: Travel Time to Work

Amount of time workers 16+ travel to work and/or home. The elapsed time includes time spent waiting for public transportation, picking up passengers in carpools, and time spent in other activities related to getting to work. A single data point is reported.

- Less than 15 minutes
- 15 to 29 Minutes
- 30 to 44 Minutes
- 45 to 59 Minutes
- 60 or more minutes


## Household Based Variables

Household based variables are all demographic data that are based upon household counts. Households are further classified into family households and non-family households.

Family households consist of a householder and one or more other persons living in the same household who are related to the householder by birth, marriage, or adoption.

Non-family households can consist of one person living alone or unrelated individuals living together.

Householder: The person, or one of the people, in whose name the home is owned, being bought, or rented. If there is no such person present, any household member 15 years old and over can serve as the householder for the purposes of the census.

Two types of householders are distinguished: a family householder and a nonfamily householder. A family householder is a householder living with one or more people related to him or her by birth, marriage, or adoption. The householder and all people in the household related to him are family members. A nonfamily householder is a householder living alone or with non-relatives only.

These data are organized according to the following sub-categories and include:

- Households
- Size and Type
- Income
- Families
- Non-families
- Vehicles


## Households

A household includes all people who occupy a housing unit. A housing unit is a house, an apartment, a mobile home, a group of rooms, or a single room occupied (or if vacant, intended for occupancy) as separate living quarters. Separate living quarters are those in which the occupants live separately from any other people in the building and that have direct access from the outside of the building or through a common hall.

The occupants may be a single family, one person living alone, two or more families living together, or any other group of related or unrelated people who share living quarters.

Five data points are provided including

- 2000
- 2010
- Current year estimate
- Five-year projection
- Ten-year forecast

People not living in households are classified as living in group quarters. Group quarters are addressed under population.

## Household Size and Type

Household Type Trends
Household Type Trends presents three household types over three time points: the last census, a current year estimate, and a five year projection.

The types reported include:

- One person households: a single person residing alone
- Family households: a householder and one or more other persons living in the same household who are related to the householder by birth, marriage, or adoption
- Non-family households: unrelated individuals living together


## Households by Size

Number of persons per household. This variable indicates the size of households in a selected geography. Households include family and non-family members living in the same house. Three data points are reported.

Categories reported:

- 1-person household
- 2-person household
- 3-person household
- 4-person household
- 5-person household
- 6-person household
- 7-or-more person household


## Householder by Gender

Householders by gender can be either female or male. Householders can include children under 18 or not. Two data points are reported.

Categories reported:

- Female Householder
- Male Householder


## Income

## Households: Average Household Income

Average income is obtained by dividing the aggregate income of a particular geographic area by the number of household units in that geographic area.

## Disposable Household Income

A measure of a household's disposable income from salary and wages, dividends, interest, profits less all government taxes. Two data points are reported; current year estimate and five year projection.

Categories reported:

- Less than $\$ 15,000$
- \$15,000 to \$24,999
- \$25,000 to \$34,999
- \$35,000 to \$49,999
- \$50,000 to \$74,999
- \$75,000 to \$99,999
- \$100,000 to \$124,999
- \$125,000 to \$149,999
- \$150,000 to \$199,999
- \$200,000 or more


## Households: Per capita income

Per capita income is the mean income computed for every man, woman, and child in a particular group. It is derived by dividing the total income of a particular group by the total population in that group.

Two data points are provided, current year and five year forecast.

## Household Income Forecast

Household Income Forecast is the total money received in the stated calendar year by all household members 15 -years-old and over.

Household income differs from family household income by including income from all persons age 15 years and older in all households, including persons living alone and other non-family households. The income is presented in terms of current dollars for the particular year in question.
'Total income is the sum of the amounts reported separately for wage or salary income; net selfemployment income; interest, dividends, or net rental or royalty income or income from estates and trusts; social security or railroad retirement income; Supplemental Security Income (SSI); public assistance or welfare payments; retirement, survivor, or disability pensions; and all other income.

Receipts from the following sources are not included as income: capital gains, money received from the sale of property (unless the recipient was engaged in the business of selling such property); the value of income "in kind" from food stamps, public housing subsidies, medical care, employer contributions for individuals, etc.; withdrawal of bank deposits; money borrowed; tax refunds; exchange of money between relatives living in the same household; and gifts and lumpsum inheritances, insurance payments, and other types of lump-sum receipts.

Median income is that point where there are as many households with incomes greater than the median as there are households with incomes less than the median. Two data points are provided, current year and five year projection.

Categories reported:

- Less than \$15,000
- \$15,000 to \$24,999
- \$25,000 to \$34,999
- \$35,000 to \$49,999
- \$50,000 to \$74,999
- \$75,000 to \$99,999
- \$100,000 to \$124,999
- \$125,000 to \$149,999
- \$150,000 to \$199,999
- \$200,000 or more


## Family Household Income Forecast

Total money received in a stated calendar year for all family household members 15 years old and over. The income is presented in terms of current dollars for the particular year in question. Two date points are provided.

Categories reported:

- Less than \$15,000
- \$15,000 to \$24,999
- \$25,000 to \$34,999
- \$35,000 to \$49,999
- \$50,000 to \$74,999
- \$75,000 to \$99,999
- \$100,000 to \$124,999
- \$125,000 to \$149,999
- \$150,000 to \$199,999
- \$200,000 or more


## Nonfamily Household Income Forecast

Total money received in a stated calendar year for nonfamily household members 15 years old and over. The income is presented in terms of current dollars for the particular year in question. Two date points are provided.

Categories reported:

- Less than \$15,000
- \$15,000 to \$24,999
- $\$ 25,000$ to $\$ 34,999$
- \$35,000 to \$49,999
- \$50,000 to \$74,999
- \$75,000 to \$99,999
- \$100,000 to \$124,999
- \$125,000 to \$149,999
- \$150,000 to \$199,999
- \$200,000 or more


## Asset to Debt

Asset to Debt reflects the number of households with the following characteristics. This variable provides insight into the level of capital available in a defined area vis-à-vis debt loads. A single data year is reported.

- Owning any financial assets
- Owning any non-financial assets
- With any debt


## Families

A family includes a householder and one or more other people living in the same household who are related to the householder by birth, marriage, or adoption.

All people in a household who are related to the householder are regarded as members of his or her family. A family household may contain people not related to the householder, but those people are not included as part of the householder's family in census tabulations. The number of family households is equal to the number of families, but family households may include more members than do families.

A household can contain only one family for purposes of census tabulations. Not all households contain families since a household may be comprised of a group of unrelated people or of one person living alone.

Categories reported:

- Families (Total)
- Families Change
- Percent Change


## Families: Forecast by Size

Number of persons per family household. This variable indicates the size of family households in a selected geography. Two data points are reported.

Categories reported:

- 2-person household
- 3-person household
- 4-person household
- 5-person household
- 6-person household
- 7-or-more person household


## Families/Non families: With Children

Households in which at least one child under 18 resides. The residents may or may not be family. Two data points are reported.

Categories reported:

- Family: Married-couple
- Other Family: Female Householder, no husband present
- Other Family: Male Householder, no wife present
- Nonfamily: Female Householder
- Nonfamily: Male Householder


## Families/Non families: With No Children

Households in which no children under 18 reside. The residents may or may not be family. Two data points are reported.

Categories reported:

- Family: Married-couple
- Other Family: Female Householder, no husband present
- Other Family: Male Householder, no wife present
- Nonfamily: Female Householder
- Nonfamily: Male Householder


## Families: Poverty

Families are classified below the poverty level when the total income of the family or of the nonfamily householder is below the appropriate poverty threshold. The poverty thresholds vary depending upon three criteria: size of family, number of children, and age of the family householder or unrelated individual for one and two-person households.

The poverty thresholds are revised annually to allow for changes in the cost of living as reflected in the Consumer Price Index. Poverty thresholds are applied on a national basis and not adjusted for regional, state or local variations in the cost of living.

This variable summarizes the number of families above and below the poverty level, Two data points are reported; 2010 census and an estimate for the current year.

Categories reported in summary form include:

- Above poverty level
- Below poverty level

Categories reported in detail form include:

- Above Poverty Level
- Married-couple family: With related children under 18 years
- Male Householder: no wife present: With related children under 18 years
- Male Householder: no wife present: No related children under 18 years
- Female Householder: no husband present: With related children under 18 years
- Female Householder: no husband present: No related children under 18 years
- Below Poverty Level
- Married-couple family: With related children under 18 years
- Male Householder: no wife present: With related children under 18 years
- Male Householder: no wife present: No related children under 18 years
- Female Householder: no husband present: With related children under 18 years
- Female Householder: no husband present: No related children under 18 years


## Nonfamily

A non-family household consists of a person living alone or a householder living with persons not related.

## Nonfamilies: Forecast by Size

A household consists of a persons living alone or a householder living with persons not related. This variable reports the persons per household in nonfamily households.

Categories reported:

- 1-person household
- 2-person household
- 3-person household
- 4-person household
- 5-person household
- 6-person household
- 7-or-more person household


## Nonfamily Householder by Gender

Nonfamily householders by gender either female or male. Householders can include children under 18. Two data points are reported.

Categories reported:

- Female Householder
- Male Householder


## Vehicles

## Households by Number of Vehicles

The presence or absence of vehicles by household. Two data points are provided.

Three categories are reported:

- No vehicle available
- One vehicle available
- Two or more vehicles available


## Owner households by Number of Vehicles

The presence or absence of vehicles by owner occupied households. Two data points are provided.

Six categories are reported:

- No vehicle available
- 1 vehicle available
- 2 or vehicles available
- 3 or vehicles available
- 4 or vehicles available
- 5 or more vehicles available


## Renter households by number of vehicles

The presence or absence of vehicles by renter occupied households. Two data points are provided.

## Owner households by Number of Vehicles

The presence or absence of vehicles by owner occupied households. Two data points are provided.

Six categories are reported:

- No vehicle available
- 1 vehicle available
- 2 or vehicles available
- 3 or vehicles available
- 4 or vehicles available
- 5 or more vehicles available


## Housing Related Variables

Housing variables present characteristics of housing stock in a community.
Housing includes a house, an apartment, a mobile home or trailer, a group of rooms, or a single room occupied as separate living quarters, or if vacant, intended for occupancy as separate living quarters. Separate living quarters are those in which the occupants live separately from any other individuals in the building and which have direct access from outside the building or through a common hall.

For vacant units, the criteria of separateness and direct access are applied to the intended occupants whenever possible.

## Housing

Housing is the baseline variable establishing the total number of housing units of any type within a geographic area. Three data points are reported.

## Housing Availability Score

Housing Availability Score (HAS) or Demand Saturation. The HAS Score is similar to an occupancy rate in that it measures the percent of existing housing that is currently occupied. It differs from a true occupancy rate in that it does not distinguish between permanently vacant housing and housing available but currently vacant.

The Housing Availability Score is presented as a percentage. A score of $100 \%$ means there is saturation-no open housing available.

- $97 \%$ to $100 \%$ No open housing available
- $90 \%$ to $96 \%$ Limited open housing available
- $70 \%$ to $89 \%$ Some open housing available
- Less than 70\% Open housing available

This score may be used as an indicator of potential population growth without new construction taking place. For a school district, this score could indicate an area where there is potential for growth in student enrollment. If there is open housing and families with children move in, enrollment may increase. Conversely, if the percentage begins to drop in a neighborhood, this indicates the loss of households and can contribute to a decline in enrollment in local schools.

The Housing Availability Score can act as an indicator of housing saturation (all existing housing is occupied). A score of $97 \%$ or greater should be considered a fully saturated neighborhood, this is due to the time elapse between one household moving out and the other moving in.

The indicator is purposely set to zero in very rural areas, due to insufficient information to make a determination

## Housing Units by Occupancy

The Housing Units by Occupancy presents both the number of units occupied and vacant. A total of all housing units is the sum of these two.

Three data points provided and the following categories.

- Occupied
- Vacant


## Housing Trends: By Dwelling Type

Housing units are segmented into one of three types; single unit attached, single unit detached and two or more units (or multiple unit dwellings).

Single Unit attached: A structure that has one or more walls extending from ground to roof separating it from adjoining structures. In row houses (sometimes called townhouses), double houses, or houses attached to nonresidential structures, each house is a separate, attached structure if the dividing or common wall goes from ground to roof.

Single-Unit, detached: A structure detached from any other house; that is, with open space on all four sides. Such structures are considered detached even if they have an adjoining shed or garage. Mobile homes or trailers to which one or more permanent rooms have been added or built also are included.

Two or more units (Multi-unit dwellings): Units in structures containing 2 or more housing units, further categorized as units in structures with 2, 3 or 4,5 to 9,10 to 19,20 to 49, and 50 or more units.

Categories reported:

- 1 Unit Attached
- 1 Unit Detached
- 2 Units
- 3 to 19 Units
- 20 to 49 Units
- 50 or More Units
- Mobile Home or Trailer
- Other


## Housing Trends: Units by Year Built

Units by Year Built refers to when the building was first constructed, not when it was remodeled, added to, or converted. A single data point is provided.

Categories include:

- 1939 or Earlier
- 1940 to 1949
- 1950 to 1959
- 1960 to 1969
- 1970 to 1979
- Built 1980 to 1989
- Built 1990 to 1994
- Built 1995 to 1998
- Built 1999 to March 2003


## Housing by Occupancy Type

This variable segments the housing stock by owner versus renter occupied.

Owner occupied: A housing unit is owner occupied if the owner or co-owner lives in the unit even if it is mortgaged or not fully paid for.

Renter occupied: All occupied housing units that are not owner occupied, whether they are rented for cash rent or occupied without payment of cash rent, are classified as renter occupied.

Three data points are provided. Categories include:

- Owner Occupied
- Renter Occupied


## Housing Trends: Value Owner-Occupied

Value reflects an estimate of how much a property (house and lot, mobile home and lot, or condominium unit) would sell for if it were for sale. It includes only owner-occupied dwellings. A single data point is provided.

- Less than $\$ 20,000$
- \$20,000 to \$39,999
- \$40,000 to \$59,999
- \$60,000 to \$79,999
- \$80,000 to \$99,999
- \$100,000 to \$149,999
- \$150,000 to \$199,999
- \$200,000 to \$299,999
- \$300,000 to \$399,999
- \$400,000 to \$499,999
- \$500,000 to \$749,999
- \$750,000 to \$999,999
- \$1,000,000 or more


## Mortgage Risk: Filings by Risk Ratio

Mortgage Risk Exposure expresses the number of mortgage filings over a three year moving average and categorizes them by the risk exposure ratio given to each one. The total number of mortgages within the period are noted and then spread across 7 aggregated ratio ranges.

- 1.2 or less - Very high debt to income/very high risk
- 1.3 to 1.8 - High debt to income/high risk
- 1.9 to 2.4 - Somewhat high debt to income/ somewhat high risk
- 2.5 to 3.0 - Acceptable debt to income/Reasonable risk
- 3.1 to 3.6 - Low debt to income/Low risk
- 3.7 to 4.2 - Very low debt to income/Very low risk
- 4.3 or greater - Extremely low debt to income/Extremely low risk

Traditionally, a debt to income ratio considered acceptable to banks to make mortgages was 2.5 to 1. This means a person's income should be two and a half times their monthly debt load.

When the ratios are greater than 2.5 it suggests the person has disposable income. When the ratio drops below 2.5 it suggests that the person's debt load is getting to heavy and thus the risk for a mortgage is greater.

A separate score that is an average of the level of risk for a defined geographic area is also available.

## MOSAIC

MOSAIC is a demographic segmentation system created by Experian. It seeks to provide a multidimensional view of a community taking into account multiple socioeconomic and life stage factors.

MOSAIC is truly unique as a demographic segmentation system. It classifies US consumers into one of 71 types and 19 groups. For more information on each of the groups and types, see the MOSAIC Resource documents.

## ViewPoint Variables

ViewPoint variables are derived from the Simmons Market Research Bureau (SMRB) surveys. The SMRB surveys consumer behavior patterns across a broad array of attitudes and practices. There are over 2,600 discrete data items in this SMRB.

After careful review, 120 were selected for inclusion as ViewPoint variables. These data reach beyond demographics to provide a view of people's attitudes and behaviors on an array of topics. The variables chosen are intended to provide planners with 'insite' into the points of view likely to be held by persons in selected geographic areas.
The data are structured around the following thematic categories.

- Life Concerns and Well Being
- Lifestyle Preferences
- Personal Perspectives
- Religious Practices
- Social Values
- Charitable Contributions
- Charitable Contributions: \$200 or more


## Appendix

## About Demographic Retrieval and Spatial Queries

Researchers, planners and marketers all use demographic information to understand the characteristics of a targeted geographic area. Their targeted geographic area may be a particular zip code(s), counties, cities, school districts, neighborhoods or some other defined area. To obtain this data, it must be retrieved from a demographic database. The process of gathering these data out of a demographic database is called 'demographic retrieval' and it is accomplished through a 'spatial query'.

For example, suppose a researcher wanted to know the characteristics of zip code 92714. She would need to "ask" a demographic database for the characteristics of that zip code. Such characteristics might include its current and projected future population, racial-ethnic composition, average household income, and education level, etc. The process of accomplishing this "ask" requires querying a demographic database and retrieving from it the requested information. Because the "ask" is about a geographic area-the zip code-it is called a 'spatial query' because the researcher is looking for data about that piece of "space".

In summary, demographic retrieval is the process of retrieving data via a spatial query resulting in a demographic data report of the population characteristics of the target geographic area.

## About Census Geographic Areas

There are two kinds of geographies to query in sophisticated demographic research systems; a) predefined geographies and b) custom or user defined geographies. Predefined geographies are geographic areas such as zip codes, counties, states, census tracts, school district boundaries, school attendance area boundaries, city boundaries, etc. These geographies already exist.

Custom defined geographies are geographic areas defined by a user at the time of doing the demographic query. Such geographies can be custom polygons or radius studies. For custom defined geographies, the user must use a tool to define the area of inquiry prior to requesting the data.

## About Census Demographic Data

How does demographic retrieval work? To understand this question, one must first understand two concepts about demographic data; the census bureau's hierarchical geographic system and how census data is provided to the public.

Census Bureau Hierarchical System: The Census Bureau maintains a fairly consistent hierarchical system of geographies. At the bottom is the individual house address. These are clustered into census blocks. Census blocks cluster into census block groups, block groups into census tracts, census tracts into counties, counties into states and finally, all states together, the nation. Each geography lower on the hierarchy, consistently aggregates with others of the same unit into the geography above it. This is important to keep in mind when we get to the question of how demographic retrieval works.


Storage of Demographic Data: The wealth of demographic data released to the public is stored at a census block group level. This is to protect privacy. While a block group in populated areas is a fairly small geographic area, the population within it is large enough to blur out any ability to access private household data. The census bureau's optimal target for block groups is roughly 1500 people at the point of the decennial census. Over the course of 10 years that population may grow or decline. Since this is the smallest census geography for which the census bureau releases data, it is at this level that all retrieval engines must work. It is also the level at which most demographic updating vendors will supply data. Retrieval engines then aggregate up whole or partial block groups to provide demographic reports.

## Three Approaches to Demographic Retrieval

Unless the area a person wants data on perfectly matches one or more block group boundaries, obtaining demographic data requires a calculation that allocates some portion of a block group to the spatial query to get demographic totals. There are three approaches to demographic retrieval through spatial queries to consider within the purview of this paper and which are the approaches used in most demographic retrieval systems.

The first model of retrieval requires the system spatial query to identify all of the block groups totally or partially included in a defined geographic area, such as a polygon. For those block groups fully enclosed by the geographic area, $100 \%$ of its data is included in the query results. For block groups partially included, the system calculates what percentage of the block group is included in the geographic area and includes only that percentage of the data for the block group in the query totals. For example, suppose a polygon includes $25 \%$ of the geographic area of a block group. The current year
population estimate for the block group is 100 people. The query would return $25 \%$ of that or 25 people for that block group and add it to the total of the query. This model works fairly well in well established areas. But for areas not fully developed or that will only be partially developed into housing, this model potentially creates a problem. This is easy to illustrate. Suppose again our block group with an estimate of 100 people and suppose that the population is all clustered in one corner so that 90\% of the population for entire block group is really perhaps in $20 \%$ of the total geographic area of the block group. The resulting query will under project population for that block group. Instead of returning 90 out of 100 representing the $90 \%$, the query would only return 20 out of the 100. Now when multiple block groups are aggregated together, undercounts and over counts tend to average out. But in queries that are for smaller areas, the returning projection can be a problem.

Allocation of Block group Population by Census Block Percentage

A second method for retrieval is based upon census blocks. A census block is roughly between 0 and 600 people, though most would be significantly less than 100. While specific data at the block level is not released, the total population and households by block are released. Since all blocks are a part of a block group, adding up the population of all blocks in a block group will equal the population for the block group. Because of this, each block represents a percentage of the total population of the block group. In the illustration below, you will see that the four blocks added together equal the total population of the block group of 1,000 , or $100 \%$. But each block represents a different percentage of the whole because the population of each at the last census varies.
How is this used for retrieval? First, the current year estimate and five year projection for the block group (where the estimates and projections are built) is applied back to the blocks based upon the most recent decennial census percentages for each block. This is done by multiplying the percentage of each block times the population of the current year estimate (or 5 year projection) for the block group.
So in the next example, we see the same block group only the total population for the block group has grown from 1,000 in census year 2000ce to 1,500 for the the current year estimate (which in the example is 2010ce). Using the formula above, the estimated population for the block group is allocated based upon the percentage of the block group each block included in the last census. So now, whereas Block 1's population was 600 in census year 2000, it is now 900 for the current year estimate (2010).
The results of the same calculation are presented for all of the blocks in the block group in this next illustration.
Having calculated the estimated current year population for each block, the second step is for the system to determine which blocks are included in the
user defined geographic area and only includes the data for those blocks captured (by the zip code, radius or polygon). This is the fundamental difference with the first approach. Recapping again the first approach; it calculates what percentage of a block group is captured in the user defined geographic area. If $20 \%$ of the block group is included, then $20 \%$ of the population of the block group gets added to the total.
However, in this second approach instead of calculating a percentage to include, the system determines which blocks have been included. It then aggregates the population data for every block captured. In the next example, assume the geographic area defined by the user includes only Blocks 1 and 2. Since Block 1 was $60 \%$ of the total population in the last census, it receives $60 \%$ of the current year estimate and Block 2 was $5 \%$ in the last census so it receives $5 \%$ of the current year estimate. Together, the system retrieves $65 \%$ or 975 of the total block group's current year estimate and includes that in the total spatial query output.

The following example illustrates the process.
This allocation method is better for the most part than proration approach of the first method in that it is more sensitive to where population actually resided as of the last census. So if a large portion of a block group had no population in the last census and still doesn't, the method of taking a percentage of the total area of the block group and multiplying it times the current year estimate and five year projection will not provide satisfactory results. This second method, tied to where population resided at the last census is more likely to allocate the block group estimate consistent with what is the case. The proration method, for example might conclude that only $10 \%$ of the block group falls with geographic area to be queried. On that basis, instead estimating the population at 975 , it would estimate it at 150.

Allocation of Block group Population by Census Block Percentage after Adjustments for Change

While the second method is a clear improvement over the first, neither takes into account significant changes in the residential information that may have occurred.

The third method of conducting the query, like the second is based upon the allocation of a block's percentage of the population of the block group. But it is different in that it will adjust the block percentages based upon a constant monitoring of two factors: 1) the addition of new zip +4 s and 2 ) indication that first class mail is now being delivered to addresses within the new zip+4 zones.
Let us explain. New zip +4 s are created when the postal service anticipates new residences to which they must deliver mail. These are being created constantly and the data source we use monitors these changes. But in addition, to further confirm that people actually live within the new zip+4s, evidence that first class mail is actually being delivered within the new zip +4 s is monitored. A zip+4 usually includes between 10 and 20 houses so the level of detail is pretty fine grained.

With these data, since the demographers know in which block the zip+ 4 is located, they can adjust the percentages each block reflects of the total estimated and projected population of the block group.
This sensitivity is reflected in the following illustration. Notice again that Block 1 in the last census reflected 60\% of the total population in the block group. It is now estimated that it only represents 20\%. Most likely this shift among the blocks reflects changes in housing development.
This method allows for a greater level of precision when studying small geographic areas and is especially helpful in areas on the edge of residential development.

