New York City Census FactFinder (NYC CFF) User Guide

Link to NYC Census FactFinder Demonstration (Oct. 9, 2015 Workshop)

New York City Census FactFinder Feedback

Feedback concerning the NYC Census FactFinder application is always appreciated, and may be addressed to **CFF_DL@planning.nyc.gov**. Please note that due to the volume of feedback we receive, we may not reply to each inquiry. We kindly request that you supply us with your contact information in case your inquiry requires further clarification.

Table of Contents

About New York City Census FactFinder

New York City Census FactFinder (NYC CFF) is New York City's online map portal for querying Census information. It was developed by the Department of City Planning and Department of Information Technology and Telecommunications. The application provides up-to-date Census data for locations in New York City. To begin a search, choose the type of geographic area you want to profile (census tract or Neighborhood Tabulation Area). Then pick your location. You can search by address, intersection, place of interest, census tract, subway station, or neighborhood. Alternatively, you can simply use the select tool on the interactive map to choose a census tract or Neighborhood Tabulation Area to be profiled. Multiple census tracts can also be selected using the radius tool or by using the select tool and the control key together.

NYC CFF can be accessed from http://maps.nyc.gov/census.

The following is a list of minimum computer requirements to run NYC CFF interactive mapping application.

Browser/OS Support

Windows 7
IE 8
Chrome 8.0
Firefox 3.6.13
Safari 5.0

Windows Vista
IE 8
Safari 5.0
Firefox 3.6
Chrome 8.0

Windows XP
Internet Explorer 7 & 8
Firefox 3.6.12
Google 8.0
Safari 5.0

Mac 10.4
Firefox 3.6.13

Mac 10.6
Firefox 3.6.9
Safari 5.0
Chrome 6.0

Getting Started

New York City Census FactFinder (NYC CFF) enables you to explore the map and view census related information. The web application page has three main parts:

- Along the top, the **Menu** has links to Help (including this User Guide), and the disclaimer.
- The **Map** tab includes the **Toolbar** for interacting with the map. The **2010 Census Profile** tab displays the census profiles from 2010 Census and the **ACS Profile** tab displays profiles from the American Community Survey.
- The **Console**, on the left side of the page, contains the following panels:
    - Geographic Data Source where users select the type of geography that they would like to profile
    - Search for a Location to search by address, intersection, place of interest, census tract, subway station, or Neighborhood Tabulation Area
    - Searched Locations allows users to review and return to location results from previous searches
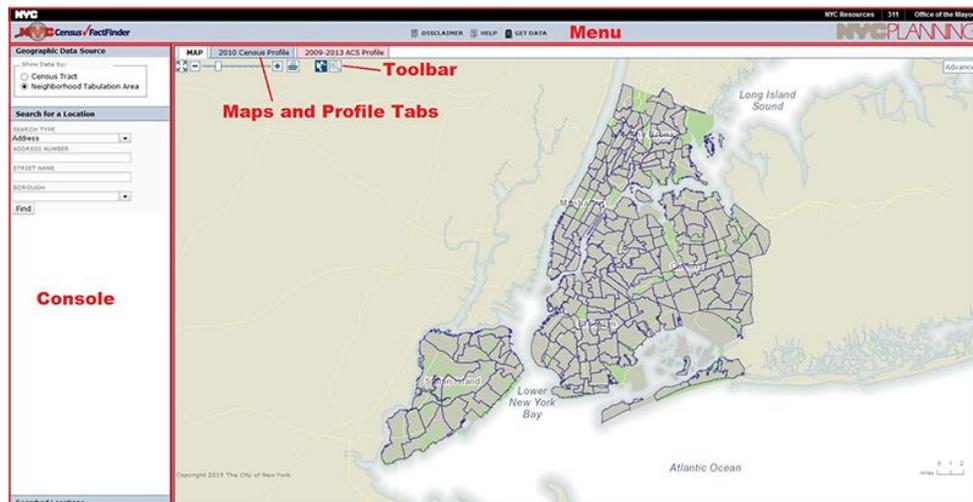
Search for a Location and Searched Locations panels can be expanded and collapsed as needed. By default, "Search for a Location" is expanded. Click on any panel name to expand that panel. The content of the panel is hidden when collapsed, but not deleted.

The console will also display the legend information such as selected location indicated by a star on the map, selected geography, and buffer distance. You can resize the area used by the Console and the Map by clicking and dragging the divider between the two parts.

Using the Toolbar to Explore the Map

This section explains the different ways you can interact with the map. When you first access Census FactFinder, you will see a small scale map of the entire City and surrounding area. From here you can either search for a location in the Console or use tools in the Toolbar, at the top of the map, to begin your search.
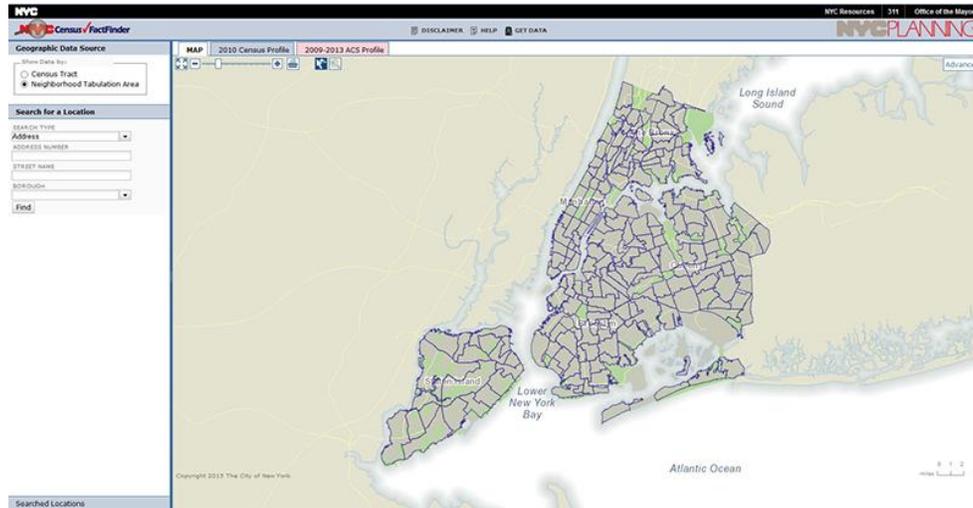
This section explains the toolbar.

The toolbar allows you to interact with the map. The table below shows each of the tools, with a brief description. Some of the tools perform an immediate action when clicked by the user. Others become the "active tool" after selection and require an interaction with the map with the mouse/cursor. Each tool is explained in more detail below.

By default the Select tool is activated.

The tools available in CFF are:

Name	Icon	Description
Full extent		Immediately zooms map out to the full view of New York City and surrounding areas. The active tool does not change.
Zoom Bar		Click or drag the bar to zoom to a new zoom level. The active tool does not change.
Print		Click on the icon to print the current map. The active tool does not change.
Excel Export		Click on the icon to export the current profile to MS Excel. The active tool does not change.
Select		Select a single or multiple features on the map. To select multiple hold ctrl-key or shift-key and click to select/deselect a census tract
Radius Search		Selects multiple features within a defined distance of a selected location.

Zooming

There are two different tools for changing the map scale.

Zoom Bar
You can jump directly to a certain level of magnification by using the "Zoom" bar in the left portion of the tool bar. The white bar indicates the current level. Click on any other area of the bar to zoom to that level. The very right will zoom in all the way; the very left will zoom out all the way. If you move your mouse over the bar, you will also see some descriptions of the approximate zoom level you would get by clicking there, including "City View," "Borough View," "Neighborhood View," "Block View," and "Building View."

You can zoom in and out one level at a time, without re-centering the map, by clicking on the "Plus" (zoom in) and "Minus" (zoom out) buttons directly to the right and left of the "Zoom" bar. Scrolling your mouse wheel, if you have one, will also zoom in and out.

Full Extent
At any time, you can see the full City coverage map by clicking the "Zoom Out to Full City View" icon at the left corner of the tool bar.

Select

In NYC CFF, there is a single tool to either select or deselect a single or multiple features on the map based on the active geography. In the "Geographic data source" section in the console if Neighborhood Tabulation Area is active, then only a single selection is allowed. If the census tract radio button is active, then the select tool permits single or multiple selections. To select multiple census tracts hold ctrl-key to select/deselect the census tracts (Mac users hold Shift-key).

Radius Search

The radius search tool allows you to buffer your selected (point) location with a predefined distance. This tool is activated only when the active geography is set to census tract in the "Geographic Data Source" section and the selected location is a point based search such as address, intersection, place of interest, or subway stations.

Click on the radius search tool to activate a dropdown box that gives you a list of distances to buffer.
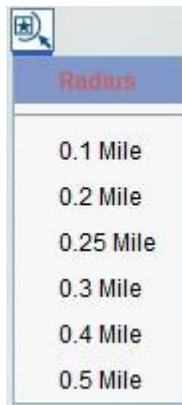
By selecting one of the choices the tool will create a buffer on the selected location and display all census tracts that intersect the buffer area as shown below.
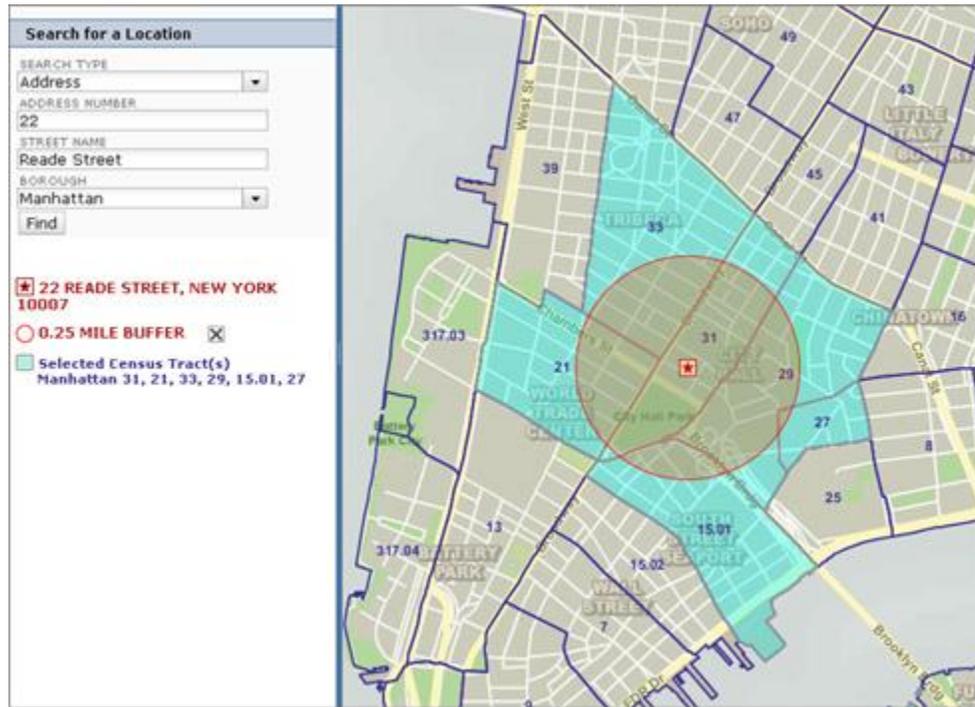
You can also clear the buffer from the map by clicking on the "x" button beside the buffer information in the console panel.

Print from Map View

To print your current display map, click on the Printer icon on the right side of the zooming toolbar. Enter the title of the map to get a version of the page that can be easily printed. If the map has an active selected location, the "Print" tool will print the current extent of the map view. By default the demographic profile is saved. To change the profile type to be printed, see the "2010 Census Profile" or "ACS Profile" section.

Panning

When no tool is selected in the toolbar, the Planning tool is activated. The map can be panned in all directions. To pan the map, first deselect any active tool (select or radius search). Then, simply left click (the cursor will change to a hand while you hold the left mouse key down) anywhere on the map and drag it in any direction. Once you release the mouse button, the map will shift in the way you just dragged it as if it were a paper map on a table.

Using the Console

Selecting a Geographic Data Source

NYC CFF allows you to show data based on Census Tract or Neighborhood Tabulation Area (NTA). You can make a geographic data source active by clicking on the radio button. This will automatically change the map geography and census profiles. If there is an active selected location, the map view and census profiles will be based on the selected location. By default the NTA geography is activated.
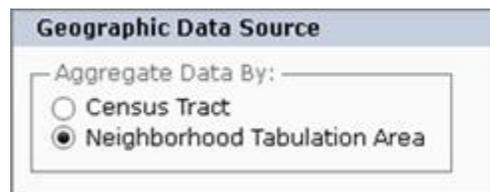
The census tract feature is visible only at a scale below 1:128,000 and if you are above this scale a "Zoom to Layer" hyperlink appears beside the census tract radio button. By clicking this link the user is zoomed to the visible scale of the census tract.
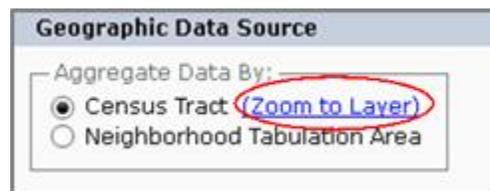
Searching for a Location

You can search for a specific location by using the "Search for a Location" panel in the console to the left of the map. By default the "Search for a Location" panel is open. You can search by address, intersection, place of interest, census tract, subway station, or Neighborhood Tabulation Area.

In CFF "Searching for a Location" is a two step process

1. Select the "Geographic Data Source"
2. Select you search type

Select "Address" in the "SEARCH TYPE" drop-down menu. Enter the building number in the "ADDRESS NUMBER" field, the street name in the "STREET NAME" field, and select the borough from the "BOROUGH" drop-down menu. Then click the "Find" button in the lower-left of the module, or press the "Enter" key. The map will automatically center and zoom in to the selected geography.

The address will either be marked with a red star or by highlighting the building footprint associated with that address and the selected geography is highlighted in a light blue color. The selected location information is displayed below the find button on the "Search for a Location" panel.

Searching by Intersection

Select "Intersection" in the "SEARCH TYPE" drop-down menu. Enter the cross streets in the "CROSS STREET ONE" and "CROSS STREET TWO" fields, and select the borough from the "BOROUGH" drop-down menu. Then click the "Find" button in the lower-left of the module, or press the "Enter" key. The map will automatically center and zoom in on that intersection. The selected intersection will also be marked with a star icon.

Searching by Place of Interest

Select "Place of Interest" in the "SEARCH TYPE" drop-down menu. Enter the name of the place in the "PLACE NAME" field and select the borough from the "BOROUGH" drop-down menu. Then click the "Find" button in the lower-left of the module, or press the "Enter" key. The map will automatically center and zoom in on that place. The selected location will also be marked with a star icon. This feature allows you to search a wide variety of City landmarks and notable destinations.

Searching by Census Tract

Select "Census Tract" in the "SEARCH TYPE" drop-down menu. Select the borough from the "BOROUGH" drop-down menu and choose the census tract number by either typing the value or selecting it from the "CENSUS TRACT" drop down menu. Then click the "Find" button in the lower-left of the module, or press the "Enter" key. The map will automatically center and zoom in on the selected geography. The census tract of the search will be marked with a red boundary and the selected profile area will be highlighted in blue.

Searching by Subway

Select "Subway Station" in the "SEARCH TYPE" drop-down menu. Choose the subway line and the station from the "SUBWAY LINE" and "SUBWAY STATION" drop-down menus. Then click the "Find" button in the lower-left of the module, or press the "Enter" key. The map will automatically center and zoom in on that station. The selected subway station is marked with a red star and the selected profile area is highlighted in blue.

Searching by Neighborhood

Select "Neighborhood" in the "SEARCH TYPE" drop-down menu (Note, these "Neighborhoods" are the Neighborhood Tabulation Areas.) Next, select the borough from the "BOROUGH" drop-down menu and neighborhood (Neighborhood Tabulation Area) from the "NEIGHBORHOOD" drop down menu. Then click the "Find" button in the lower-left of the module, or press the "Enter" key. The map will automatically center and zoom in on that neighborhood. The selected neighborhood is marked with a red boundary and the selected profile area is highlighted in blue.

Map and Profile Tabs

The user can view the Map and 2010 Census Profiles by clicking on the tabs. By default the Map tab is active.
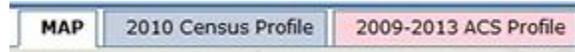
If there is no searched location, or if no census tract or Neighborhood Tabulation Area has been selected, then no profile will be visible under the 2010 Census Profile tab, and the user will receive a message to "Select a location or geography to display profile information.

2010 Census Profile Tab

When you have selected a location, by clicking on the "2010 Census Profile" tab both the "Demographics" and the "Demographic Change" Profile tabs become visible. By default the Demographics tab is selected.

In the Demographics Profile tab users are given a profile from the 2010 Census of their selected area, as well as the profile for New York City. The profiled topics from the 2010 Census include race, sex, age, Asian and Hispanic subgroups, relationship to head of household, household type, housing occupancy,

housing tenure, and household size. Users may change the default contrast geography from New York City to any of the five boroughs by clicking on the drop down menu to the right of the "New York City" heading.

By clicking on the Demographic Change tab users can view the same 2010 Census variables shown in the Demographics tab, along with statistics from the 2000 Census. The numeric change and percent change are also provided.

In order to view a contrasting geography in the Demographic Change tab, users can select New York City or any of the boroughs from the drop down menu to the right of the selected area heading.

At any time users can return to the map view by clicking the map tab, or operate any of the elements in the console section.

2009-2013 ACS Profile Tab

When you have selected a location, by clicking on the "2009-2013 ACS Profile" tab, you will see four profile tabs (Demographics, Social, Economic, and Housing). By default the Demographics tab is selected, and the reliability data is in Hide mode.

When you switch the profile from Hide to Show view for the reliability data, you will see Margin of Error (MOE) and Coefficient of Variation (CV) fields. For more information regarding these fields, please refer to the About the Data section.

In Hide view, users may change the default contrasting geography from New York City to any of the five boroughs by clicking on the drop down menu to the right of the "New York City" heading.

In Show view, users may New York City and any of the five boroughs by clicking on the drop down menu to the right of the "Selected Tract(s)" or "Selected Neighborhood" heading.

In the **Demographics** Profile tab, users are given a profile from the 2009-2013 American Community Survey (ACS) of their selected area, as well as the profile for New York City. The profile topics from the 2009-2013 ACS include sex, age, race, and Asian subgroup.

In the **Social** tab, users are given a profile from the 2009-2013 ACS of their selected area, as well as the profile for New York City. The profiled topics from the 2009-2013 ACS include household type, relationship to head of household, marital status, grandparents, school enrollment, educational attainment, veteran status, disability status, residence from a year ago, place of birth, U.S. citizenship status, year of entry, foreign born region of birth, language spoken at home, and ancestry.

In the **Economic** tab, users are given a profile from the 2009-2013 ACS of their selected area, as well as the profile for New York City. The profiled topics from the 2009-2013 ACS include employment status, commuting to work, occupation, industry, class of worker, income and benefits, health insurance coverage, and families and people whose income in the past 12 months is below the poverty level.

In the **Housing** tab, users are given a profile from the 2009-2013 ACS of their selected area, as well as the profile for New York City. The profiled topics from the 2009-2013 ACS include housing occupancy, units in structure, the year structure was built, rooms, housing tenure, year householder moved into unit, vehicles available, occupants per room, value of unit, mortgage status, selected monthly owner costs as a percentage of household income (SMOCAPI), gross rent, and gross rent as a percentage of household income (GRAPI).

Accessing the Help

At any time you can access the "Help" section by clicking "Help" link at the top of the CFF application.
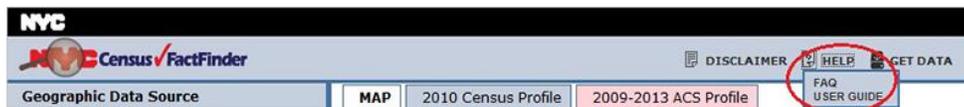
It consists of FAQs (Frequently Asked Questions) and a User Guide that can be accessed at anytime.

About the Data

Census FactFinder is built upon three major data components: the two geographic layers of census tracts and Neighborhood Tabulation Areas, decennial census data from the United States Census Bureau, and American Community Survey data, also from the Census Bureau. The following sections provide background on these three subject areas.

Geographic Data Sources

Census Tracts

These are census tracts developed for the 2010 US Census. The boundary files are derived from the US Census Bureau's TIGER project and have been geographically modified to fit the New York City base map. In New York City, there are 2,168 census tracts, which typically have a population of about 3,000-4,000 and an average land area of 90 acres.

Neighborhood Tabulation Areas (NTAs)

Neighborhood Tabulation Areas, or NTAs, are aggregations of census tracts from the 2010 Census that are subsets of New York City's 55 Public Use Microdata Areas (PUMAs), which roughly represent the city's Community Districts. NTA boundaries and their associated names do not definitively represent neighborhoods. NTAs were created to project populations at a small area level, from 2000 to 2030 for PlaNYC, the long-term sustainability plan for New York City. Since population size affects the error associated with population projections, these geographic units needed to have a minimum population, which was determined to be 15,000. This criterion resulted in combinations of neighborhoods into NTAs.

New York's 189 NTAs (this count includes Rikers Island, but excludes the airports and NTAs designated as "parks, cemeteries, etc.", just as they are excluded in the CFF application) are a valuable summary level for data from both the 2010 Census and the American Community Survey (ACS). Regarding the decennial census, these geographic areas offer a good compromise between the very detailed data for census tracts (2,168) and the broad strokes provided by PUMAs (55). For the ACS, NTAs offer a statistically reliable alternative to the high sampling error that renders data for most individual census tracts unusable. For more on **the limitations of tract level ACS data** please refer to the section entitled "Usability of Tract Data from the 2008-2012 ACS", in Chapter 2, from the National Academy of Sciences book *Realizing the Potential of the American Community Survey*.

2010 Census Profile Data

Data for the 2010 Census Profile shown on NYC Census FactFinder are derived from two Census Bureau sources: Summary File 1 (SF1) of the 2000 Census and Summary File 1 of the 2010 Census. The following is a description of the data compiled from these two datasets, important facts to be aware of, and the processing that was involved in preparing the data for use in NYC CFF.

Data from 2010

The 2010 census provides a demographic portrait of New York City's population. This includes data on age, sex, race and Hispanic origin, household size and type, living arrangements and relationships of household members, and housing units by occupancy status (occupied, vacant) and tenure (owner, renter); all included in NYC CFF. All 2010 data in NYC CFF originate from Summary File 1 (SF1) of the 2010 decennial census. When aggregating data from individual census tracts NYC CFF typically sums up component variables. However, there are certain statistics that require more complex formulas than basic summations including the calculations of rates, averages, and medians.

When calculating homeowner and rental vacancy rates, average household size, average family size, average household size of owner-occupied units, average household size of renter-occupied units, and median age NTA and tract aggregation values are calculated using specialized formulas. In order to calculate homeowner vacancy rates the number of vacant units for sale only is divided by the sum of owner-occupied units and vacant units that are for sale only, and then multiplying by 100. In order to calculate rental vacancy rates the number of vacant units for rent is divided by the sum of the renter-occupied units and the number of vacant units for rent, and then multiplying by 100. Average household size is obtained by dividing the number of people in households by the total number of households (or householders). Average family size is obtained by dividing the number of members of families by the total number of families (or family householders). Average household size of owner-occupied units is obtained by dividing the number of people living in owner-occupied housing units by the number of owner-occupied housing units. Average household size of renter-occupied units is obtained by dividing the number of people living in renter-occupied housing units by the number of renter-occupied housing units. Finally, the median age divides the age distribution in a stated area into two equal parts: one-half of the population falling below the median value and one-half above the median value. To calculate this statistic first the median instance is obtained by dividing the total population into halves. Then, the single year age group that includes the median instance is identified by examining a cumulative frequency distribution. Then a median age is calculated by inflating the year of single year age group (for example age 62) in the same proportion that the median instance falls in the frequency distribution for that age group. (For example 62 could be inflated to 62.25 if the median instance is one quarter the way through the frequency distribution of 62 year olds.)

A few more notes about the 2010 data:

- The count for Chinese, under Asian subgroups, includes individuals that identified themselves as Taiwanese.
- The 2010 Census enumerated New York City's population at 8,175,133. The Census Bureau, in cooperation with the Department of City Planning, had estimated the city's population to be around 8.4 million as of July 2010. It is the City of New York's belief that there was a significant undercount of the population. This was partly due to the 2010 census finding an increase of 82,000 vacant units in New York City, or a 46 percent rise since 2000. A disproportionate share of this increase was found in two local census offices covering southern Brooklyn and northwest Queens, two vibrant sections of the city. The huge concentration of vacant units in these two areas cannot be explained by new construction or foreclosures; nor is it consistent with other survey and administrative data. Although New York City formally challenged the 2010 Census count, this challenge was rejected by the Census Bureau on the grounds that no errors were identified that were permissible under the Bureau's Count Question Resolution process. While there is now no formal mechanism to correct any undercount, users should still be cognizant of reliability issues when examining data in southern Brooklyn and northwest Queens. While percent distributions in these areas may be accurate, overall counts may be lower than in reality. For more information on this topic visit the section of the Department of City Planning's website that addresses the challenge of the 2010 Census.

Data from 2000

Data from the 2000 Census were included in the NYC CFF application in order to show change between 2000 and 2010, on the Demographic Change tab. The 2000 census provided a demographic portrait of New York City's population at the turn of the last century. This includes data on age, sex, race and Hispanic origin, household size and type, living arrangements and relationships of household members, and housing units by occupancy status (occupied, vacant) and tenure (owner, renter); all included in NYC CFF. All 2000 data in NYC CFF originate from Summary File 1 (SF1) of the 2000 decennial census.

All 2000 Census data used in NYC CFF were put in contrast to 2010 data in order to show change between 2000 and 2010. Because of this, and because all geographic data used in NYC CFF mapping were from the 2010 Census, it was necessary to put all 2000 Census data into 2010 geographic terms. This conversion was done, for the most part, by taking 2000 Census data at a census block level and aggregating the data up to 2010 census tracts. This method of putting 2000 Census data into 2010 geographies yielded very little error, save for a few blocks in 2000 that were split by 2010 census tract boundaries. Many of these split blocks did not involve areas with any population in 2000. However, when split census blocks did involve population, the whole block was allocated to one individual 2010 census tract, based on housing unit counts in either side of the split, using the PLUTO database. A select number of variables from the 2000 Census were not available at a census block level for use with NYC CFF, including: single year age groups used in median age calculations, all Asian and Hispanic subgroup data, and all information on unmarried partners. These variables were converted into 2010 census tracts, from 2000 census tracts, by splitting the variables in the same proportions that the overall population in 2000 was split, based on block level analysis. Because of two different methods of converting 2000 data into 2010 census tracts were necessary, it is possible that certain data incongruencies will arise. For example, if one were to compare the Hispanic total represented in the Hispanic subgroup section (which was derived by converting 2000 census tract data into 2010 census

tract geographies) of the 2000 data with the Hispanic total under the Race and Hispanic Origin section (which was derived by converting 2000 census block data into 2010 census tract geographies) there are times that the numbers will not match due to the different conversion methods.

When aggregating data from individual census tracts NYC CFF typically sums up component variables from the 2000 Census. However, as with the 2010 data, there are certain statistics that require more complex formulas than basic summations including the calculations of rates, averages, and medians. Specifically, when calculating homeowner and rental vacancy rates, average household size, average family size, average household size of owner-occupied units, average household size of renter-occupied units, and median age NTA and tract aggregation values are calculated using specialized formulas. For a detailed explanation of how these specialized statistics are calculated, see section on *Data from 2010*.

A few more notes about the 2000 data:

- As with the 2010 data, the count for Chinese, under Asian subgroups, includes individuals that identified themselves as Taiwanese.
- Users should also be cognizant of the limitations of the Hispanic subgroup data from 2000. Although the number of most specific Hispanic groups increased from 1990 to 2000, the percent of the total Hispanic population for these groups decreased. It is widely believed that a change in the wording of the Census questionnaire in 2000 resulted in an inflated number of Hispanics reporting their Hispanic subgroup as "Other", thus reducing the percent share of Hispanics in specific subgroup categories. This may result in certain Hispanic subgroup counts to be understated in 2000 (noteably in New York City, Dominicans). To correct this, the Census Bureau released the Simulated Hispanic File to show users a more likely count of Hispanic subgroups in 2000.

American Community Survey (ACS) Profile Data

Data for the American Community Survey (ACS) Profile shown on NYC Census FactFinder (CFF) are derived from the 2009-2013 Summary File. The following is a description of the data compiled from this dataset, important facts to be aware of, especially concerning data reliability, and the processing that was involved in preparing the data for use in NYC CFF.

Data from 2009-2013 ACS Summary File

The American Community Survey (ACS) is the most extensive nationwide survey currently available. From its annual releases we are able to examine the city's detailed demographic, socioeconomic, and housing characteristics. Each month, questionnaires are sent to a sample of approximately 295,000 addresses across the country, so households are continuously receiving and responding to the ACS. In order to have a large enough sample from which to create estimates of characteristics, the ACS "rolls-up" the sample for 12-, 36-, or 60-month periods, depending on the size of the geographic area. Estimates are prepared using 1 year of sample for places of at least 65,000 residents, 3 years of sample for places of at least 20,000 residents, and 5 years of sample for all places regardless of size. This means that estimates can be obtained for a single year (e.g. 2012) or for multi-year periods (e.g. 2010-2012 or 2008-2012). In the New York City Census FactFinder application, 5-year period estimates were used, from 2009-2013. Users should note that all dollar values are inflated by the Census Bureau to the final year of any period estimate. For example, dollar values shown from the 2009-2013 release are all in

2013 dollars. For more information on the inflation factors employed in the ACS, see the Bureau of Labor Statistics webpage on the Consumer Price Index Research Series.

Reliability of ACS Data

As mentioned earlier, the ACS is a nationwide survey about a host of household characteristics. Because the ACS is a survey, the data derived from it are considered to be *estimates*, or values obtained from a statistical sample. These estimates are subject to *sampling variability* – the variability that occurs by chance because a value is derived from a sample rather than a census of the entire population. This issue of sampling variability associated with survey estimates is generally referred to as data reliability – *reliability* refers to the ability of a measurement instrument to produce consistent results, and in this case the measurement instrument is the ACS survey. The reliability of ACS data is highly variable and differs depending on a number of factors. To give users a sense of ACS estimate reliability the Census Bureau, and the NYC Census FactFinder application, publish **Margins of Error (MOEs)** along with every estimate. These MOEs describe the precision of an estimate within a 90-percent confidence interval. (The following is drawn from the ACS Handbook that our NYC City Planning Population Division wrote for the Census Bureau.) For example, if given an ACS estimate of 43,527 kindergarteners in Utah, with a MOE of plus or minus 2,834. The MOE tells us if we had the time and money to create the same estimate of kindergartners several thousand times, from several thousand samples, that 90 percent of the estimates of kindergarteners in Utah would be between 40,693 (43,527 – 2,834) and 46,361 (43,527 + 2,834), a fairly precise estimate. In more technical terms, MOEs provide an idea of how much variability (i.e., sampling error) is associated with the estimate. As an MOE gets larger, relative to the size of an estimate, the estimate becomes less reliable. A measure called the **Coefficient of Variation (CV)** can also be used to discern the level of reliability of an estimate. (CVs are published in the NYC Census Factinder application alongside Margins of Error.) This measure is constructed in two steps:

a) Calculate the standard error: SE = MOE/1.645

b) Calculate the Coefficient of Variation: CV = SE/Estimate * 100

With respect to the above example, the standard error associated with the estimate for kindergarteners in Utah is 2,834/1.645 or 1,723. The CV equals 1,723/43,527 * 100, or 4.0 percent. This means that one standard error is 4 percent of the estimate, a fairly low level of variability that indicates that the estimate is reliable.

The issue of **what constitutes a reliable CV** is the subject of much discussion. A recent National Academy of Sciences publication, *Using the American Community Survey: Benefits and Challenges*, suggests "CVs of 10-12 percent or less are often accepted as a reasonable standard of precision for an estimate." However, acceptable levels of reliability really depend on how estimates are being used. If one only wants to get a general sense of whether an area, for example, has a concentration of poverty, then even CVs in excess of 20% could be acceptable. However, most applications will require higher levels of precision (CVs of under 20%). That is why in New York City Census FactFinder all data (both estimates and percents) associated with CVs of 20% or more are grayed out, alerting users about the questionable reliability. This was important because CVs of 20% mean that MOEs are nearly one third the size of the estimate. For example, a median household income of $100,000 with a CV of 20% would have a MOE of $32,900. So, in this example, the estimate could be as high as $132,900 or as low as $67,100. Of course,

many users will wish to go beyond this level of data precision. For example, if one is producing a thematic map with four or five categories of data, then typically CVs will be about 5% to 15%. An even greater level of precision is required if one is creating a ranking or targeting areas for outreach, in which case CVs of under 10% are preferable. Unfortunately, there are no hard and fast rules regarding CVs and acceptable levels of reliability and, ultimately, all applications need to be judged on a case by case basis.

Processing of ACS Data

For the New York City Census FactFinder (NYC CFF) application, Summary File profile data (and other selected Summary File tables), from 2009-2013 for New York City, all constituent boroughs, and all constituent census tracts were downloaded from the Census Bureau's American FactFinder website and processed for use. This processing included:

- The removal of all MOEs and CVs associated with estimates of zero (these cells will appear blank in corresponding NYC CFF output).
- The removal of all non-numeric cell values, except for top- and bottom-coded median estimates (these cells will appear blank in corresponding NYC CFF output).
- All top-coded median estimates had non-numeric elements removed (ie, plus signs). All bottom-coded median estimates had non-numeric elements removed (ie, minus signs) and were numerically reduced by one (eg, bottom-coded *median non-family income* of "2,500-" was represented as "2,499" in the CFF database).
- Percent estimates of 0% and 100% had their MOEs removed
- Where estimates associated with percent totals (percents of 100%) were zero, the percent estimate was removed along with all percent estimates that employed this base.
- The calculation of CVs for all estimates that did not equal zero. For the formula, see section on CVs.
- In certain instances NYC CFF employs different bases for percent totals than does the Census Bureau, and in these cases percent estimates and percent MOEs were re-calculated. For methodology, see section below on the approach used for NTA calculations.
- The same approach used for NTA calculations was used to calculate the estimate and MOE associated with the *high school graduate or higher* and the *bachelor's degree or higher* variables. For methodology, see section below on the approach used for NTA calculations.
- The same approach used for NTA calculations was used to calculate the percent estimate and percent MOE associated with the second instance of the *unemployed* variable – the *Unemployed* variable that uses *civilian labor force* as its base. For methodology, see section below on the approach used for NTA calculations.

Since the Census Bureau does not publish data at an NTA level, estimates, MOEs, CVs, percent estimates, and percent MOEs associated with NTAs were all calculated by the Department of City Planning's Population Division. Since NTAs are census tract aggregates, the same approaches used to calculate NTA values were employed by DoITT programmers to calculate ACS values for tract aggregations performed by users through the CFF application interface.

The following is a **brief explanation of the calculations** behind the processing of ACS data at an NTA level and for any tract aggregation performed through the CFF application:

Calculations associated with basic count estimates

Basic count estimates, like *total population* or *workers 16 and over that walked to work*, were relatively easy to produce. Estimates were produced through summations. MOEs were produced using the Census Bureau's *General Approximation Formula (*the square root of the sum of the squared MOEs). CVs were produced using the standard CV formula (for formula, see section on CVs). Percents were calculated by summing the component parts (for both numerator and denominator), dividing the numerator by the denominator, and multiplying the quotient by 100. The formula for the Percent MOE was the most complicated calculation associated with basic count estimates. It involved squaring the MOE of the base, multiplying it by the percent estimate divided by 100 squared, then subtracting this from the MOE of the numerator squared (if this result was negative, then the subtraction performed earlier is changed to addition to avoid having a negative value under a radical), the square root of the result is calculated, before finally multiplying the answer by one divided by the base estimate (an example using this formula is written out below).

Example - Given two census tracts (X and Y) that need to be aggregated to produce results for the cumulative area of both tracts X and Y, then the following methods were followed to produce output:

Tract X	Estimate	MOE	CV	Percent	Percent MOE
Population	100	30	18.2	100.0	
Male	50	20	24.3	50.0	13.2

Tract Y	Estimate	MOE	CV	Percent	Percent MOE
Population	200	40	12.2	100.0	
Male	90	30	20.3	45.0	12.0

Tract X and Y	Estimate	MOE	CV	Percent	Percent MOE
Population	300	50	10.1	100.0	
Male	**140**	**36**	**15.7**	**46.7**	**9.1**
	↑ 1	↑ 2	↑ 3	↑ 4	↑ 5

1. Estimate of the *male population* for Tract X and Y: 50 + 90 = **140**
2. MOE for the *male population* for Tract X and Y: $\sqrt{20^2 + 30^2}$ = **36**
3. CV for the *male population* for Tract X and Y: ((36/1.645)/140)*100 = **15.7**
4. Percent for the *male population* for Tract X and Y: (140/300)*100 = **46.7**
5. Percent MOE for the *male population* for Tract X and Y:
   $(1/300) * \sqrt{36^2 - ((46.7/100)^2 * 50^2)} * 100$ = **9.1**

For more information on calculating the MOEs, CVs, and percent MOEs for derived estimates, see Appendix 3 in the ACS Handbook entitled What State and Local Governments Need to Know.

Calculations associated with medians

10 medians were included in the Census FactFinder application:

1. Median age (years)
2. Median household income (dollars)
3. Median family income (dollars)
4. Median nonfamily income (dollars)
5. Median earnings for workers (dollars)
6. Median earnings for male full-time, year-round workers (dollars)
7. Median earnings for female full-time, year-round workers (dollars)
8. Median number of rooms in housing units
9. Median value of owner-occupied units (dollars)
10. Median gross rent of occupied units paying rent (dollars)

Linear interpolation was used to calculate the estimates for these medians. Although the Census Bureau also employs linear interpolation to calculate their published median estimates, it was not possible to replicate the intervals that the Census Bureau uses (the Bureau uses much finer intervals producing more precise estimates of medians) because we had to rely on published, rather than raw, data. For more on the linear interpolation that the Census Bureau uses, see the ACS Subject Definitions. Note that the calculation of *median value of owner-occupied units* excludes values of zero. Also note that in calculating *median number of rooms in housing units*, the whole number is used as the midpoint of the interval; thus the category "3 rooms" is treated as an interval ranging from 2½ to 3½ rooms.

When calculating medians from grouped data it is common to use the practice of top-coding and, less often, bottom-coding. This top- or bottom-coding occurs when the median instance falls in an open ended interval/grouping. For example, the Census Bureau's highest grouping for *median household income* is *$250,000 or more*, and if the median were to be in this interval, then the Bureau would "top-code" the median estimate as "$250,000+". However, as was mentioned earlier, without access to the Census Bureau's raw data, we had to rely on published data, and sometimes the top- and bottom-coded interval differs from the Bureau's interval. For example, the highest grouping for *median household income* in Census FactFinder calculations (estimates for NTAs and tract aggregates) is *$200,000 or more.* Therefore, it is possible to encounter two different top-coded values in Census FactFinder profiles. If a top-coded value is shown for New York City, a borough, or a census tract, then it came from the Census Bureau, and will consequently use the Bureau's top-coded value (Also note that non-numeric elements, like plus or minus signs, used by the Census Bureau have been removed in CFF output.) If, however, a top-coded value is shown for a NTA or a grouping of census tracts, then it is a derived estimate calculated by the Census FactFinder application and will use CFF top-coded values. The following is a summary of the top- and bottom-coding used in the Census FactFinder application. Note that the differences between Census Bureau and CFF coding are shown, since both sets are possible in CFF output.

Variable	Top-coded Values		Bottom-coded Values	
	Census B.	CFF	Census B.	CFF
Median age	115	85	NA	NA
Median Household Income	250,000	200,000	2,499*	9,999
Median Family Income	250,000	200,000	2,499*	9,999
Median Nonfamily Income	250,000	200,000	2,499*	9,999
Median Earnings for Workers	250,000	100,000	2,499*	2,499
Median Earnings, Male Full-time, Yr.-round Workers	250,000	100,000	2,499*	2,499
Median Earnings, Fem. Full-time, Yr.-round Workers	250,000	100,000	2,499*	2,499
Median Rooms	9	9	NA	NA
Median Value Owner Occupied Units	1,000,000	1,000,000	NA	NA
Median Gross Rent	2,000	2,000	NA	NA

*Recoded in CFF from 2,500 to 2,499 since non-numeric negative sign, marking 2,500 as bottom-coded, was removed.

It should be noted that MOEs and CVs associated with top- or bottom-coded values have been removed and the cells will appear blank. This was done because it is not possible to gauge reliability for open-ended intervals. Additionally, top- and bottom-coded values have been grayed out because the vast majority of these estimates are based on very small sample sizes and consequently, lack statistical reliability.

A slight modification of the Census Bureau's ACS PUMS formula for median MOEs was used to calculate the MOEs associated with median estimates. For the Census Bureau's formula, see the section on Standard Errors for Medians, in the ACS Accuracy of the PUMS documentation. Step two of the calculation, regarding the calculation of the standard error associated with the 50th percentile, was changed from [Design Factor * $\sqrt{(95/(5*Base)) * 502}$] to [Design Factor * $\sqrt{(93/(7*Base)) * 502}$]. This was done to simulate the increased sample size going from the 5-year ACS PUMS to the 5-year ACS estimates typical of New York City. Note that where median estimates were top- or bottom-coded, no MOEs were shown.

CVs were produced using the standard CV formula (for formula, see section on CVs). Note that where median estimates were top- or bottom-coded, no CVs were shown.

Calculations associated with means

7 means were included in the Census FactFinder application:

1. Average household size
2. Average family size
3. Mean travel time to work (minutes)
4. Mean household income (dollars)
5. Per capita income (dollars)
6. Average household size of owner-occupied unit
7. Average household size of renter-occupied unit

The standard formula for arithmetic means (sum of the values divided by the sum of the number of values) was used to calculate the estimates for these means. Most of the calculations required the use of ACS Summary File tables beyond the profiles. For the specific formulas used for each of the seven means included in CFF, refer to the ACS Subject Definitions.

Because there was no Census Bureau endorsed methodology for the calculation of MOEs associated with derived means from ACS Summary File data, we turned to the strong relationship between sample size and variance to model Coefficients of Variation (CVs) for each of the seven means shown in CFF and then estimated the MOE using this modeled CV. The following table summarizes the formulas that resulted from this modeling and the calculation used to backtrack into an estimated MOE using the modeled CV along with the estimated mean.

Mean Estimate	Input X	Formula for CV	Formula for MOE
Average household size	Hsg. Sample Size	$CV=62x^{-.503}$	$(CV/100)*1.645)$ *AvgHHSz
Average family size	Hsg. Sample Size	$CV=51x^{-.473}$	$(CV/100)*1.645)$ *AvgFmSz
Mean travel time to work	Pop. Sample Size	$CV=316x^{-.679}$	$(CV/100)*1.645)$ *MnTrvTm
Mean household income (dollars)	Hsg. Sample Size	$CV=462x^{-.754}$	$(CV/100)*1.645)$ *MnHHInc
Per capita income (dollars)	Pop. Sample Size	$CV=1122x^{-.778}$	$(CV/100)*1.645)$ *PerCapInc
Average household size, owner-occupied unit	Hsg. Sample Size	$CV=564x^{-.762}$	$(CV/100)*1.645)$ *AvgHHSOOc
Average household size, renter-occupied unit	Hsg. Sample Size	$CV=257x^{-.699}$	$(CV/100)*1.645)$ *AvgHHSROc

Calculations associated with rates

2 rates were included in the Census FactFinder application:

1. Homeowner Vacancy Rate
2. Rental Vacancy Rate

These rates were essentially equivalent to the calculation of percentages. For general methodology in the calculation of the estimates, MOEs, and CVs associated with rates, see the calculations for percents, percent MOEs, and CVs, respectively, in the *Calculations associated with basic count estimates* section. For the specific formulas used to calculate both of the rate estimates included in CFF, refer to the ACS Subject Definitions.

Get the Data

Census FactFinder utilizes data from various sources and then format and/or aggregate them for display. The following sections provide instructions on accessing the raw data that are being used in this application.

Geographic Data Sources

Census Tracts

The 2010 Census Tracts and metadata can be accessed from http://www1.nyc.gov/site/planning/data-maps/open-data/districts_download_metadata.page.

Neighborhood Tabulation Areas (NTAs)

The latest NTAs and metadata can be accessed from http://www1.nyc.gov/site/planning/data-maps/open-data.page#other.

2010 Census Data

Census Tracts Level

The 2010 Census Tract data can be accessed from the Census Bureau's American FactFinder.

Neighborhood Tabulation Areas (NTAs) Level

The 2010 Census NTA data can be accessed from City Planning's Population Division website, in the Demographic Tables section of the 2010 Census page.

New York City Census FactFinder Feedback

Feedback concerning the NYC Census FactFinder application is always appreciated, and may be addressed to **CFF_DL@planning.nyc.gov**. Please note that due to the volume of feedback we receive, we may not reply to each inquiry. We kindly request that you supply us with your contact information in case your inquiry requires further clarification.

Copyright Information

All rights to the content of this website are reserved by the New York City Department of City Planning to the extent permitted by law. You may print copies of this information for your own personal use. No part of these pages may otherwise be reproduced or distributed in any form or medium without express permission of the New York City Department of City Planning.