

American Community Survey Accuracy of the Data (2016)

INTRODUCTION

This document describes the accuracy of the 2016 American Community Survey (ACS) 1-year estimates. The data contained in these data products are based on the sample interviewed from January 1, 2016 through December 31, 2016.

The ACS sample is selected from all counties and county-equivalents in the United States. In 2006, the ACS began collecting data from sampled persons in group quarters (GQs) – for example, military barracks, college dormitories, nursing homes, and correctional facilities. Persons in sample in group quarters and persons in sample in housing units (HUs) are included in all 2016 ACS estimates that are based on the total population. All ACS population estimates from years prior to 2006 include only persons in housing units.

The ACS, like any other statistical activity, is subject to error. The purpose of this document is to provide data users with a basic understanding of the ACS sample design, estimation methodology, and accuracy of the ACS data. The ACS is sponsored by the U.S. Census Bureau, and is part of the Decennial Census Program.

For additional information on the design and methodology of the ACS, including data collection and processing, visit: <https://www.census.gov/programs-surveys/acs/methodology.html>.

To access other accuracy of the data documents, including the **2016 PRCS Accuracy of the Data document** and the **2012-2016 ACS Accuracy of the Data document**¹, visit: <https://www.census.gov/programs-surveys/acs/technical-documentation/code-lists.html>.

¹ The **2012-2016 Accuracy of the Data document** will be available after the release of the 5-year products in December 2017.

Table of Contents

INTRODUCTION.....	1
DATA COLLECTION	3
Housing Units.....	3
Group Quarters.....	3
SAMPLING FRAME	4
Housing Units.....	4
Group Quarters.....	4
SAMPLE DESIGN.....	5
Housing Units.....	5
Group Quarters.....	8
2016 Sample Sizes for Housing Unit Addresses and Group Quarters.....	12
WEIGHTING METHODOLOGY.....	13
Group Quarters Person Weighting	13
Housing Unit and Household Person Weighting	15
CONFIDENTIALITY OF THE DATA	20
Title 13, United States Code	20
Disclosure Avoidance	20
Data Swapping	21
Synthetic Data	21
ERRORS IN THE DATA.....	22
Sampling Error.....	22
Nonsampling Error.....	22
MEASURES OF SAMPLING ERROR.....	23
Confidence Intervals and Margins of Error.....	23
Limitations	24
CALCULATION OF STANDARD ERRORS	25
Sums and Differences of Direct Standard Errors	26
Ratios	27
Proportions/Percents	27
Percent Change.....	28
Products.....	28
TESTING FOR SIGNIFICANT DIFFERENCES	28
EXAMPLES OF STANDARD ERROR CALCULATIONS.....	30
Example 1 – Calculating the Standard Error from the Margin of Error.....	30
Example 2 – Calculating the Standard Error of a Sum or a Difference	30
Example 3 – Calculating the Standard Error of a Proportion/Percent.....	31
Example 4 – Calculating the Standard Error of a Ratio	31
Example 5 – Calculating the Standard Error of a Product	32
CONTROL OF NONSAMPLING ERROR.....	33
Coverage Error.....	33
Nonresponse Error	34
Measurement and Processing Error.....	35
ISSUES WITH APPROXIMATING THE STANDARD ERROR OF LINEAR COMBINATIONS OF MULTIPLE ESTIMATES	37

DATA COLLECTION

Housing Units

The ACS employs four modes of data collection:

1. Internet
2. Mailout/Mailback
3. Computer Assisted Telephone Interview (CATI)
4. Computer Assisted Personal Interview (CAPI)

The general timing of data collection is as follows. Note that mail responses are accepted during all three months of data collection:

- Month 1: Mailable addresses in sample are sent an initial mailing package, which contains information for completing the ACS questionnaire via the internet. If a sample address has not responded online within two weeks of the initial mailing, then it is sent a second mailing package with a paper questionnaire. Sampled addresses then have the option of which mode to use to complete the interview.
- Month 2: All non-responding addresses with an available phone number are sent to CATI.
- Month 3: A sample of mailable non-responses without a good phone number, CATI non-responses, and unmailable addresses are sent to CAPI.

All Remote Alaska addresses that are in sample are sent to CAPI and assigned to one of two data collection periods: January-April or September-December.² Up to four months is allowed to complete the assigned interviews.

Group Quarters

Group Quarters data collection generally spans six weeks. However, for Remote Alaska and Federal prisons, the data collection period lasts up to four months. Group Quarters in Remote Alaska are assigned to one of two data collection periods: January-April or September-December. All Federal prisons are assigned to a September-December data collection period.

Field representatives have several options available to them for data collection. They can complete the questionnaire with the resident either in person or over the telephone, conduct a personal interview with a proxy, such as a relative or guardian, or leave a paper questionnaire for residents to complete. The last option is used for data collection in Federal prisons.

² Prior to the 2011 sample year, all remote Alaska sample cases were subsampled for CAPI at a rate of 2-in-3.

SAMPLING FRAME

Housing Units

The universe for the ACS consists of all valid, residential housing unit addresses in all county and county equivalents in the 50 states, including the District of Columbia. The Master Address File (MAF) is a database maintained by the Census Bureau containing a listing of residential, group quarters, and commercial addresses in the U.S. and Puerto Rico. The MAF is normally updated twice each year with the Delivery Sequence Files (DSF) provided by the U.S. Postal Service; in 2014, however, the MAF was updated only once. The DSF covers only the U.S. These files identify mail drop points and provide the best available source of changes and updates to the housing unit inventory. The MAF is also updated with the results from various Census Bureau field operations, including the ACS.

Group Quarters

As in previous years, due to operational difficulties associated with data collection, the ACS excludes certain types of GQs from the sampling universe and data collection operations. The weighting and estimation accounts for this segment of the population are included in the population controls. The following GQ types are removed from the GQ universe:

- Soup kitchens
- Domestic violence shelters
- Regularly scheduled mobile food vans
- Targeted non-sheltered outdoor locations
- Maritime/merchant vessels
- Living quarters for victims of natural disasters
- Dangerous encampments

The ACS GQ universe file contains both valid and invalid GQs, but only valid GQs are eligible for sampling. This is done in order to maintain an inventory of all GQ records. In this way, any updates to the GQ universe can be applied to the combined valid and invalid file.

SAMPLE DESIGN

Housing Units

The ACS employs a two-phase, two-stage sample design. The first-phase sample consists of two separate address samples: Period 1 and Period 2. These samples are chosen at different points in time. Both samples are selected in two stages of sampling, a first-stage and a second-stage. Subsequent to second-stage sampling, the majority of sample addresses are randomly assigned to one of the twelve months of the sample year (the exception is for addresses in remote Alaska, which are assigned to either January or September). The second-phase of sampling occurs when the CAPI sample is selected.

The Period 1 sample is selected during September and October of the year prior to the sample year (the 2016 Period 1 sample was selected in September and October of 2015). Approximately half of a year's sample is selected at this time. Sample addresses that are not in remote Alaska are randomly assigned to one of the first six months of the sample year; sample addresses in remote Alaska are assigned to January.

Period 2 sampling occurs in January and February of the sample year (the 2016 Period 2 sample was selected during January and February of 2016). This sample accounts for the remaining half of the overall first-phase sample. Period 2 sample addresses that are not in remote Alaska are randomly assigned to one of the last six months of the sample year; Period 2 sample addresses in remote Alaska are assigned to September.³

A sub-sample of non-responding addresses and of any addresses deemed unmailable is selected for the CAPI data collection mode.⁴

The following steps are used to select the first-phase and second-phase samples in both periods.

First-Phase Housing Unit Sample Selection

First-Stage Sampling for Housing Units

First stage sampling defines the universe for the second stage of sampling through three steps. First, all addresses that were in a first-phase sample within the past four years are excluded from eligibility. This ensures that no address is in sample more than once in any five-year period. The second step is to select a 20 percent systematic sample of “new” units, i.e. those units that have never appeared on a previous MAF extract. Each

³ Remote Alaska assignments are made so that the sample addresses are approximately evenly distributed between the two data collection periods.

⁴ Beginning with the August, 2011 CAPI sample all non-mailable and non-responding addresses in the following areas are now sent to CAPI: all Hawaiian Homelands, all Alaska Native Village Statistical areas, American Indian areas with an estimated proportion of American Indian population $\geq 10\%$.

new address is systematically assigned either to the current year or to one of four back-samples. This procedure maintains five equal partitions (samples) of the universe. The third step is to randomly assign all eligible addresses to a period.⁵

Assignment of Blocks to a Second-Stage Sampling Stratum for Housing Units

Second-stage sampling uses 16 sampling strata in the U.S.⁶ The stratum-level rates used in second-stage sampling account for the first-stage selection probabilities. These rates are applied at a block level to addresses in the U.S. by calculating a measure of size for each of the following geographic entities:

- Counties
- Places
- School Districts (elementary, secondary, and unified)
- American Indian Areas
- Tribal Subdivisions
- Alaska Native Village Statistical Areas
- Hawaiian Homelands
- Minor Civil Divisions – in Connecticut, Maine, Massachusetts, Michigan, Minnesota, New Hampshire, New Jersey, New York, Pennsylvania, Rhode Island, Vermont, and Wisconsin⁷
- Census Designated Places – in Hawaii only

The measure of size for all areas except American Indian Areas, Tribal Subdivisions, Alaska Native Village Statistical Areas, and Hawaiian Homelands is an estimate of the number of occupied HUs in the area. This is calculated by multiplying the number of ACS addresses by an estimated occupancy rate at the block level. A measure of size for each Census Tract is also calculated in the same manner.

For American Indian, Tribal Subdivisions, and Alaska Native Village Statistical Areas, the measure of size is the estimated number of occupied HUs multiplied by the proportion of people reporting American Indian or Alaska Native (alone or in combination) in the 2010 Census.

For Hawaiian Homelands, the measure of size is the estimated number of occupied HUs multiplied by the proportion of people reporting Native Hawaiian (alone or in combination) in the 2010 Census.

⁵ Most of the period assignments are made during Period 1 sampling. The only assignments in Period 2 sampling are made for addresses that were not part of the process in Period 1, e.g., new addresses.

⁶ Beginning with the 2011 sample the ACS implemented a change to the stratification, increasing the number of sampling strata and changing how the sampling rates are defined. Prior to 2011 there were seven strata; there are now 16 sampling strata. Table 1 gives a summary of these strata and the rates.

⁷ These are the states where MCDs are active, functioning governmental units.

Each block is then assigned the smallest positive, non-zero measure of size from the set of all entities of which it is a part. The 2016 second-stage sampling strata and the overall first-phase sampling rates are shown in Table 1 below.

Calculation of the Second-Stage Sampling Rates for Housing Units

The overall first-phase sampling rates are calculated using the distribution of ACS valid addresses by second-stage sampling stratum in such a way as to yield an overall target sample size for the year of 3,540,000 (1,770,000 for each period) in the U.S. The first-phase rates are adjusted for the first-stage sample to yield the second-stage selection probabilities.

Table 1. First-phase Sampling Rate Categories for the United States

Sampling Stratum	Type of Area	Rate Definitions	2016 Sampling Rates	
			Period 1	Period 2
1	$0 < \text{MOS}^1 < 200$	15.00%	15.00%	15.00%
2	$200 \leq \text{MOS} < 400$	10.00%	10.00%	10.00%
3	$400 \leq \text{MOS} < 800$	7.00%	7.00%	7.00%
4	$800 \leq \text{MOS} < 1200$	$2.80 \times \text{BR}$	4.34%	4.33%
5	$1200 \leq \text{MOS}$ and $0 < \text{TRACTMOS}^2 < 400$	$3.50 \times \text{BR}$	5.42%	5.41%
6	$1200 \leq \text{MOS}$ and $0 < \text{TRACTMOS} < 400 \text{ HR}^3$	$0.92 \times 3.50 \times \text{BR}$	4.99%	4.97%
7	$1200 \leq \text{MOS}$ and $400 \leq \text{TRACTMOS} < 1000$	$2.80 \times \text{BR}$	4.34%	4.33%
8	$1200 \leq \text{MOS}$ and $400 \leq \text{TRACTMOS} < 1000 \text{ HR}$	$0.92 \times 2.80 \times \text{BR}$	3.99%	3.98%
9	$1200 \leq \text{MOS}$ and $1000 \leq \text{TRACTMOS} < 2000$	$1.70 \times \text{BR}$	2.63%	2.63%
10	$1200 \leq \text{MOS}$ and $1000 \leq \text{TRACTMOS} < 2000 \text{ HR}$	$0.92 \times 1.70 \times \text{BR}$	2.42%	2.42%
11	$1200 \leq \text{MOS}$ and $2000 \leq \text{TRACTMOS} < 4000$	BR^4	1.55%	1.55%
12	$1200 \leq \text{MOS}$ and $2000 \leq \text{TRACTMOS} < 4000 \text{ HR}$	$0.92 \times \text{BR}$	1.43%	1.42%
13	$1200 \leq \text{MOS}$ and $4000 \leq \text{TRACTMOS} < 6000$	$0.60 \times \text{BR}$	0.93%	0.93%
14	$1200 \leq \text{MOS}$ and $4000 \leq \text{TRACTMOS} < 6000 \text{ HR}$	$0.92 \times 0.60 \times \text{BR}$	0.86%	0.85%
15	$1200 \leq \text{MOS}$ and $6000 \leq \text{TRACTMOS}$	$0.35 \times \text{BR}$	0.54%	0.54%
16	$1200 \leq \text{MOS}$ and $6000 \leq \text{TRACTMOS HR}$	$0.92 \times 0.35 \times \text{BR}$	0.50%	0.50%

¹MOS = measure of size (estimated number occupied housing units) of the smallest governmental entity

²TRACTMOS = the measure of size (MOS) at the Census Tract level

³HR = areas where predicted levels of completed mail and CATI interviews are > 60%

⁴BR = base sampling rate

Second-Stage Sample Selection for Housing Units

After each block is assigned to a second-stage sampling stratum in each period, a systematic sample of addresses is selected from the second-stage universe (first-stage sample) within each county and county equivalent.

Sample Month Assignment for Housing Units

After the second stage of sampling, sample addresses selected during Period 1 sampling that are not in remote Alaska are allocated to one of the first six months of the sample year. Sample addresses selected during Period 2 sampling that are not in remote Alaska are assigned to a month between July through December, inclusive. Sample addresses in

remote Alaska are assigned to January or September in Period 1 and Period 2 sampling, respectively.

Second-Phase Housing Unit Sample Selection – CAPI Subsampling

The addresses from which CAPI sub-samples are selected can be divided into two groups. One group includes addresses that are not eligible for any other data collection operation – these consist of unmailable addresses and those in remote Alaska areas. The second group includes addresses that are eligible for the other data collection operations but for which no response was obtained prior to CAPI sub-sampling – these consist of mailable addresses not in a remote Alaska area.

All sample addresses in remote Alaska are sent to the CAPI data collection operation. Most unmailable addresses are selected for CAPI at a rate of 2-in-3 – the exception is when they are in a Hawaiian Homeland area (HH), Alaska Native Village Statistical area (ANVSA), or certain American Indian areas (AI), where all are selected for CAPI.

With one exception, mailable addresses from which a response was not obtained by the time of the CAPI operation are sampled at rates of 1-in-2, 2-in-5, and 1-in-3 – these rates are set at the tract level. The exception is for addresses in HH, ANVSA, and AI areas, where all are selected for CAPI. Table 2 shows the CAPI sub-sampling rates that are associated with each group of addresses.

Table 2. Second-Phase (CAPI) Subsampling Rates for the United States

Address and Tract Characteristics	CAPI Subsampling Rate
Addresses in Remote Alaska*	Take all (100.0%)
Addresses in Hawaiian Homelands, Alaska Native Village Statistical areas and a subset of American Indian areas*	Take all (100.0%)
Unmailable addresses that are not in the previous two categories	66.7%
Mailable addresses in tracts with predicted levels of completed mail and CATI interviews prior to CAPI subsampling between 0% and 35%, inclusive	50.0%
Mailable addresses in tracts with predicted levels of completed mail and CATI interviews prior to CAPI subsampling greater than 35% and less than or equal to 50%	40.0%
Mailable addresses in all other tracts	33.3%

*The full CAPI follow-up procedure for these two categories began in 2011.

Group Quarters

The 2016 group quarters (GQ) sampling frame was divided into two strata: a small GQ stratum and a large GQ stratum. Small GQs have expected populations of fifteen or fewer residents, while large GQs have expected populations of more than fifteen residents.

Samples were selected in two phases within each stratum. In general, GQs were selected in the first phase and then persons/residents were selected in the second phase. Both phases differ between the two strata. Each sampled GQ was randomly assigned to one or more months in 2016 – it was in these months that their person samples were selected.

Small GQ Stratum

First Phase of Sample Selection for Small GQs

There were two stages of selecting small GQs for sample.

1. First stage

The small GQ universe is divided into five groups that are approximately equal in size. All new small GQs are systematically assigned to one of these five groups on a yearly basis, with about the same probability (20 percent) of being assigned to any given group. Each group represents a second-stage sampling frame, from which GQs are selected once every five years. The 2016 second-stage sampling frame was used in 2011 as well, and is currently to be used in 2021, 2026, etc.

2. Second stage

GQs were systematically selected from the 2016 second-stage sampling frame. Each GQ had the same second-stage probability of being selected within a given state, where the probabilities varied between states. Table 3 below shows these probabilities.

Table 3. 2016 Group Quarter State Targeted Sampling Rates for the U.S.

State	Targeted Rate	State	Targeted Rate	State	Targeted Rate
Alabama	2.21%	Kentucky	2.38%	North Dakota	4.51%
Alaska	3.46%	Louisiana	2.76%	Ohio	2.50%
Arizona	2.19%	Maine	3.33%	Oklahoma	2.16%
Arkansas	2.25%	Maryland	2.24%	Oregon	2.58%
California	2.61%	Massachusetts	2.28%	Pennsylvania	2.58%
Colorado	2.33%	Michigan	2.79%	Rhode Island	2.54%
Connecticut	2.39%	Minnesota	2.51%	South Carolina	2.02%
Delaware	5.03%	Mississippi	2.53%	South Dakota	3.99%
District of Columbia	2.89%	Missouri	2.33%	Tennessee	2.35%
Florida	2.34%	Montana	4.29%	Texas	2.24%
Georgia	2.47%	Nebraska	2.48%	Utah	1.82%
Hawaii	3.11%	Nevada	3.80%	Vermont	4.61%
Idaho	4.30%	New Hampshire	2.81%	Virginia	2.42%
Illinois	2.36%	New Jersey	2.64%	Washington	2.45%
Indiana	2.29%	New Mexico	2.68%	West Virginia	2.18%
Iowa	2.45%	New York	2.37%	Wisconsin	2.52%
Kansas	2.46%	North Carolina	2.36%	Wyoming	6.85%

Second Phase of Sample Selection for Small GQs

Persons were selected for sample from each GQ that was selected for sample in the first phase of sample selection. If fifteen or fewer persons were residing at a GQ at the time a field representative (interviewer) visited the GQ, then all persons were selected for sample. Otherwise, if more than fifteen persons were residing at the GQ, then the interviewer selected a systematic sample of ten persons from the GQ's roster.

Targeted Sampling Rate (Probability of Selection) for Small GQs

The targeted state-level sampling rates are the probabilities of selecting any given person in a GQ; it is around these probabilities that the sample design is based. These probabilities reflect both phases of sample selection, and they varied by state. The probabilities for 2016 are shown in Table 3.

The sample was designed so that the second-phase sampling rate would be one-hundred percent for small GQs (i.e., select the entire expected population of fifteen or fewer persons for sample in every small sampled GQ). This means the probability of selecting any person in a small GQ was designed to equal the probability of selecting the small GQ itself.

Large GQ Stratum

First phase of Sample Selection for Large GQs

All large GQs were eligible for being sampled in 2016, as has been the case every year since the inception of the sampling of GQs in 2006. This means there was only a single stage of sampling in this phase. This stage consists of systematically assigning "hits" to GQs independently in each state, where each hit represents ten persons to be sampled.

In general, a GQ has either Z or $Z+1$ hits assigned to it. The value for Z is dependent on both the GQ's expected population size and its within-state target sampling rate, shown in Table 3. When this rate is multiplied by a GQ's expected population, the result is a GQ's expected person sample size. If a GQ's expected person sample size is less than ten, then $Z = 0$; if it is at least ten but less than twenty, then $Z = 1$; if it is at least twenty but less than thirty, then $Z = 2$; and so on. See 2.C. below for a detailed example.

If a GQ has an expected person sample size that is less than ten, then this method effectively gives the GQ a probability of selection that is proportional to its size; this probability is the expected person sample size divided by ten. If a GQ has an expected person sample size of ten or more, then it is in sample with certainty and is assigned one or more hits.

Second Phase of Sample Selection for Large GQs

Persons were selected within each GQ to which one or more hits were assigned in the first phase of selection. There were ten persons selected at a GQ for every hit assigned to the GQ. The persons were systematically sampled from a roster of persons residing at the GQ at the time of an interviewer's visit. The exception was if there were far fewer persons residing in a GQ than expected – in these situations, the number of persons to sample at the GQ would be reduced to reflect the GQ's actual population. In cases where fewer than ten persons resided in a GQ at the time of a visit, the interviewer would select all of the persons for sample.

Targeted Sampling Rate (Probability of Selection) for Large GQs

As for small GQs, the targeted state-level sampling rates are the probabilities of selecting any given person in a GQ. The probabilities are shown in Table 3. Note that these rates are the same as for persons in small GQs.

As an example, suppose a GQ in Indiana had an expected population of 250. The target sampling rate in Indiana was 2.29%, meaning any given person in a GQ in Indiana had about a 1-in-43²/₃ chance of being selected. This rate, combined with the GQ's expected population of 250, means that the expected number of persons selected for sample in this GQ would be 5.725 (2.29% × 250). Since this is less than ten, this GQ would have either 0 or 1 hits assigned to it ($Z = 0$). The probability of it being assigned a hit would be the GQ's expected person sample size of 5.725 divided by 10, or 57.25%.

As a second example, suppose a GQ in Idaho had an expected population of 1,000. The target sampling rate in Idaho was 4.30%, meaning any given person in a GQ in Idaho had about a 1-in-23.26 chance of being selected. This rate, combined with the GQ's expected population of 1,000, means that the expected number of persons selected for sample in the GQ would be 43 (4.30% × 1,000); this GQ would be assigned either four or five hits ($Z = 4$).

Group Quarters Sample Month Assignment

All sample GQs were assigned to one or more months (interview months) – these were the months in which interviewers would visit a GQ to select a person sample and conduct interviews. All small GQs, all large GQs that were assigned only one hit, all remote Alaska GQs, all sampled military facilities, and all sampled correctional facilities (regardless of how many hits a military or correctional facility was assigned) were assigned to a single interview month. Remote Alaska GQs were assigned to either January or September; Federal prisons were assigned to September; all of the others were randomly assigned one interview month.

All large GQs that had been assigned multiple hits, but were not in any of the categories above, had each hit randomly assigned to a different interview month. If a GQ had more than twelve hits assigned to it, then multiple hits would be assigned to one or more interview months for the GQ. For example, if a GQ had fifteen hits assigned to it, then there would be three interview months in which two hits were assigned and nine interview months in which

one hit was assigned. There are two restrictions to this process. One restriction was applied to college dormitories, whose hits were randomly assigned to non-summer months only, i.e., January through April and September through December. The other restriction was applied to military ships, whose hits were randomly assigned only to the last ten months of the year, i.e., March through December.

Bureau of Prison Group Quarters

Prior to 2016, all GQs were sampled at the same time for a given year. Starting in 2016, Bureau of Prison GQs (Federal prisons) started to be sampled separately from other GQs. They are sampled using the same procedure shown above, and are all assigned to the September interview month as before.

2016 Sample Sizes for Housing Unit Addresses and Group Quarters.

Counts of sample addresses and GQ persons can be found in two locations on the US Census Bureau website. In American Fact Finder, base tables B98001 and B98002 provide sample size counts for the nation, states, and counties. Sample size counts for the nation, and states are also available in the Sample Size and Data Quality Section of the ACS website, at <https://www.census.gov/acs/www/methodology/sample-size-and-data-quality/>.

WEIGHTING METHODOLOGY

The estimates that appear in this product are obtained from a raking ratio estimation procedure that results in the assignment of two sets of weights: a weight to each sample person record and a weight to each sample housing unit record. Estimates of person characteristics are based on the person weight. Estimates of family, household, and housing unit characteristics are based on the housing unit weight. For any given tabulation area, a characteristic total is estimated by summing the weights assigned to the persons, households, families or housing units possessing the characteristic in the tabulation area. Each sample person or housing unit record is assigned exactly one weight to be used to produce estimates of all characteristics. For example, if the weight given to a sample person or housing unit has a value 40, all characteristics of that person or housing unit are tabulated with the weight of 40.

The weighting is conducted in two main operations: a group quarters person weighting operation which assigns weights to persons in group quarters, and a household person weighting operation which assigns weights both to housing units and to persons within housing units. The group quarters person weighting is conducted first and the household person weighting second. The household person weighting is dependent on the group quarters person weighting because estimates for total population, which include both group quarters and household population, are controlled to the Census Bureau's official 2016 total resident population estimates.

Group Quarters Person Weighting

Starting with the weighting for the 2011 1-year ACS estimates, the group quarters (GQ) person weighting changed in important ways from previous years' weighting. The GQ population sample was supplemented by a large-scale whole person imputation into not-in-sample GQ facilities. For the 2016 ACS GQ data, roughly as many GQ persons were imputed as interviewed. The goal of the imputation methodology was two-fold.

1. The primary objective was to establish representation of county by major GQ type group in the tabulations for each combination that exists on the ACS GQ sample frame. The seven major GQ type groups are defined by the Population Estimates Program and are given in Table 4.
2. A secondary objective was to establish representation of tract by major GQ type group for each combination that exists on the ACS GQ sample frame.

Table 4: Population Estimates Program Major GQ Type Groups

Major GQ Type Group	Definition	Institutional / Non-Institutional
1	Correctional Institutions	Institutional
2	Juvenile Detention Facilities	Institutional
3	Nursing Homes	Institutional
4	Other Long-Term Care Facilities	Institutional
5	College Dormitories	Non-Institutional
6	Military Facilities	Non-Institutional
7	Other Non-Institutional Facilities	Non-Institutional

The GQ sampling frame was modified to create an imputation frame from which all imputed GQs were selected from. The frame was updated with the actual populations and GQ type changes from ACS interviews, as well as any subsequent information gathered in other processes since the sampling frame was initially created. The change in populations for ACS GQ interviews was used to calculate a not-in-sample adjustment factor that was used to update the population for all GQs on the frame not selected for sample. This adjustment factor was calculated at the following level:

GQ Major Type × GQ Size Stratum

There were three size strata used for this process: GQs in sample with certainty, GQs with 16 or more persons, and GQs with less than 16 persons.

For all not-in-sample GQ facilities with an expected population of 16 or more persons (large facilities), we imputed a number of GQ persons equal to 2.5% of the expected population. For those GQ facilities with an expected population of fewer than 16 persons (small facilities), we selected a random sample of GQ facilities as needed to accomplish the two objectives given above. For those selected small GQ facilities, we imputed a number of GQ persons equal to 20% of the facility's expected population.

Interviewed GQ person records were then sampled at random to be donors for the imputed persons of the selected not-in-sample GQ facilities. An expanding search algorithm searched for donors within the same specific type of GQ facility and the same county. If that failed, the search included all GQ facilities of the same major GQ type group. If that still failed, the search expanded to a specific type within state, then a major GQ type group within state. This expanding search continued through division, region, and the entire nation until suitable donors were found.

The weighting procedure made no distinction between sampled and imputed GQ person records. The initial weights of person records in the large GQ facilities equaled the observed or expected population of the GQ facility divided by the number of person records. The initial weights of person records in small GQ facilities equaled the observed or expected population of the GQ facility divided by the number of records, multiplied by the inverse of the fraction of small GQ facilities represented in the weighting to the number on the frame of that tract by major GQ type group combination.

The population totals on the imputation frame are used to ensure that the sub-state distribution of GQ weighting preserves the distribution from the frame. This accomplished through a series of three constraints:

1. *Tract Constraint (TRCON)* – This factor makes the total weight within each tract by major type group equal the total population from the imputation frame.
2. *County Constraint (CYCON)* – This factor makes the total weight within each county by major type group equal the total population from the imputation frame.

3. *State Constraint (STCON)* – This factor makes the total weight within each state by major type group equal the total population from the imputation frame.

As was done in previous years' weighting, we controlled the final weights to an independent set of GQ population estimates produced by the Population Estimates Program for each state by each of the seven major GQ type groups.

Lastly, the final GQ person weight was rounded to an integer. Rounding was performed so that the sum of the rounded weights were within one person of the sum of the unrounded weights for any of the groups listed below:

Major GQ Type Group

Major GQ Type Group × County

Housing Unit and Household Person Weighting

The housing unit and household person weighting uses two types of geographic areas for adjustments: weighting areas and subcounty areas. Weighting areas are county-based and have been used since the first year of the ACS. Subcounty areas are based on incorporated place and minor civil divisions (MCD). Their use was introduced into the ACS in 2010.

Weighting areas were built from collections of whole counties. 2010 Census data and 2007-2011 ACS 5-year data were used to group counties of similar demographic and social characteristics. The characteristics considered in the formation included:

- Percent in poverty (the only characteristic using ACS 5-year data)
- Percent renting
- Density of housing units (a proxy for rural areas)
- Race, ethnicity, age, and sex distribution
- Distance between the centroids of the counties
- Core-based Statistical Area status

Each weighting area was also required to meet a threshold of 400 expected person interviews in the 2011 ACS. The process also tried to preserve as many counties that meet the threshold to form their own weighting areas. In total, there are 2,130 weighting areas formed from the 3,143 counties and county equivalents.

Subcounty areas are built from incorporated places and MCDs, with MCDs only being used in the 20 states where MCDs serve as functioning governmental units. Each subcounty area formed has a total population of at least 24,000, as determined by the July 1, 2016 Population Estimates data, which are based on the 2010 Census estimates of the population on April 1, 2010, updated using births, deaths, and migration. The subcounty areas can be incorporated places, MCDs, place/MCD intersections (in counties where places and MCDs are not coexistent), 'balance of MCD,' and 'balance of county.' The latter two types group together unincorporated areas and places/MCDs that do not meet the population threshold. If two or

more subcounty areas cannot be formed within a county, then the entire county is treated as a single area. Thus, all counties whose total population is less than 48,000 will be treated as a single area since it is not possible to form two areas that satisfy the minimum size threshold.

The estimation procedure used to assign the weights is then performed independently within each of the ACS weighting areas.

Initial Housing Unit Weighting Factors

This process produced the following factors:

Base Weight (BW)

This initial weight is assigned to every housing unit as the inverse of its block's sampling rate.

CAPI Subsampling Factor (SSF)

The weights of the CAPI cases are adjusted to reflect the results of CAPI subsampling. This factor is assigned to each record as follows:

Selected in CAPI subsampling: $SSF = 2.0, 2.5, \text{ or } 3.0$ according to Table 2

Not selected in CAPI subsampling: $SSF = 0.0$

Not a CAPI case: $SSF = 1.0$

Some sample addresses are unmailable. A two-thirds sample of these is sent directly to CAPI and for these cases $SSF = 1.5$.

Sample addresses in Remote Alaska, Hawaiian Homelands, Alaska Native Village Statistical areas and a subset of American Indian areas are selected for CAPI at 100% sampling rate and for these cases $SSF = 1.0$.

Variation in Monthly Response by Mode (VMS)

This factor makes the total weight of the Mail, CATI, and CAPI records to be tabulated in a month equal to the total base weight of all cases originally mailed for that month. For all cases, VMS is computed and assigned based on the following groups:

Weighting Area \times Month

Noninterview Factor (NIF)

This factor adjusts the weight of all responding occupied housing units to account for nonresponding housing units. The factor is computed in two stages. The first factor, NIF1, is a ratio adjustment that is computed and assigned to occupied housing units based on the following groups:

Weighting Area \times Building Type (single or multi unit) \times Tract

A second factor, NIF2, is a ratio adjustment that is computed and assigned to occupied housing units based on the following groups:

Weighting Area \times Building Type \times Month

NIF is then computed by applying NIF1 and NIF2 for each occupied housing unit. Vacant housing units are assigned a value of NIF = 1.0. Nonresponding housing units are assigned a weight of 0.0.

Noninterview Factor Mode (NIFM)

This factor adjusts the weight of the responding CAPI occupied housing units to account for CAPI nonrespondents. It is computed as if NIF had not already been assigned to every occupied housing unit record. This factor is not used directly but rather as part of computing the next factor, the Mode Bias Factor.

NIFM is computed and assigned to occupied CAPI housing units based on the following groups:

Weighting Area \times Building Type \times Month

Vacant housing units or non-CAPI (mail and CATI) housing units receive a value of NIFM = 1.0.

Mode Bias Factor (MBF)

This factor makes the total weight of the housing units in the groups below the same as if NIFM had been used instead of NIF. MBF is computed and assigned to occupied housing units based on the following groups:

Weighting Area \times Tenure (owner or renter) \times Month \times Marital Status of the Householder (married/widowed or single)

Vacant housing units receive a value of MBF = 1.0. MBF is applied to the weights computed through NIF.

Housing Unit Post-Stratification Factor (HPF)

This factor makes the total weight of all housing units agree with the 2016 independent housing unit estimates at the subcounty level.

Person Weighting Factors

Initially the person weight of each person in an occupied housing unit is the product of the weighting factors of their associated housing unit ($BW \times \dots \times HPF$). At this point, everyone

in the household has the same weight. The person weighting is done in a series of three steps, which are repeated until a stopping criterion is met. These three steps form a raking ratio or raking process. These person weights are individually adjusted for each person as described below.

The three steps are as follows:

Subcounty Area Controls Raking Factor (SUBEQRF)

This factor is applied to individuals based on their geography. It adjusts the person weights so that the weighted sample counts equal independent population estimates of total population for the subcounty area. Because of later adjustment to the person weights, total population is not assured of agreeing exactly with the official 2016 population estimates at the subcounty level.

Spouse Equalization/Householder Equalization Raking Factor (SPHHEQRF)

This factor is applied to individuals based on the combination of their status of being in a married-couple or unmarried-partner household and whether they are the householder. All persons are assigned to one of four groups:

1. Householder in a married-couple or unmarried-partner household
2. Spouse or unmarried partner in a married-couple or unmarried-partner household (non-householder)
3. Other householder
4. Other non-householder

The weights of persons in the first two groups are adjusted so that their sums are each equal to the total estimate of married-couple or unmarried-partner households using the housing unit weight ($BW \times \dots \times HPF$). At the same time, the weights of persons in the first and third groups are adjusted so that their sum is equal to the total estimate of occupied housing units using the housing unit weight ($BW \times \dots \times HPF$). The goal of this step is to produce more consistent estimates of spouses or unmarried partners and married-couple and unmarried-partner households while simultaneously producing more consistent estimates of householders, occupied housing units, and households.

Demographic Raking Factor (DEMORF)

This factor is applied to individuals based on their age, race, sex and Hispanic origin. It adjusts the person weights so that the weighted sample counts equal independent population estimates by age, race, sex, and Hispanic origin at the weighting area. Because of collapsing of groups in applying this factor, only total population is assured of agreeing with the official 2016 population estimates at the weighting area level.

This uses the following groups (note that there are 13 Age groupings):

Weighting Area \times Race / Ethnicity (non-Hispanic White, non-Hispanic Black, non-Hispanic American Indian or Alaskan Native, non-Hispanic Asian, non-Hispanic Native Hawaiian or Pacific Islander, and Hispanic (any race)) \times Sex \times Age Groups.

These three steps are repeated several times until the estimates at the national level achieve their optimal consistency with regard to the spouse and householder equalization.

The effect Person Post-Stratification Factor (PPSF) is then equal to the product (SUBEQRF \times SPHHEQRF \times DEMORF) from all of iterations of these three adjustments.

The unrounded person weight is then equal to the product of PPSF times the housing unit weight (BW \times ... \times HPF \times PPSF).

Rounding

The final product of all person weights (BW \times ... \times HPF \times PPSF) is rounded to an integer.

Rounding is performed so that the sum of the rounded weights is within one person of the sum of the unrounded weights for any of the groups listed below:

County

County \times Race

County \times Race \times Hispanic Origin

County \times Race \times Hispanic Origin \times Sex

County \times Race \times Hispanic Origin \times Sex \times Age

County \times Race \times Hispanic Origin \times Sex \times Age \times Tract

County \times Race \times Hispanic Origin \times Sex \times Age \times Tract \times Block

For example, the number of White, Hispanic, Males, Age 30 estimated for a county using the rounded weights is within one of the number produced using the unrounded weights.

Final Housing Unit Weighting Factors

This process produces the following factors:

Householder Factor (HHF)

This factor adjusts for differential response depending on the race, Hispanic origin, sex, and age of the householder. The value of HHF for an occupied housing unit is the PPSF of the householder. Since there is no householder for vacant units, the value of HHF = 1.0 for all vacant units.

Rounding

The final product of all housing unit weights (BW \times ... \times HHF) is rounded to an integer. For occupied units, the rounded housing unit weight is the same as the rounded person

weight of the householder. This ensures that both the rounded and unrounded householder weights are equal to the occupied housing unit weight. The rounding for vacant housing units is then performed so that total rounded weight is within one housing unit of the total unrounded weight for any of the groups listed below:

County

County × Tract

County × Tract × Block

CONFIDENTIALITY OF THE DATA

The Census Bureau has modified or suppressed some data on this site to protect confidentiality. Title 13 United States Code, Section 9, prohibits the Census Bureau from publishing results in which an individual's data can be identified.

The Census Bureau's internal Disclosure Review Board sets the confidentiality rules for all data releases. A checklist approach is used to ensure that all potential risks to the confidentiality of the data are considered and addressed.

Title 13, United States Code

Title 13 of the United States Code authorizes the Census Bureau to conduct censuses and surveys. Section 9 of the same Title requires that any information collected from the public under the authority of Title 13 be maintained as confidential. Section 214 of Title 13 and Sections 3559 and 3571 of Title 18 of the United States Code provide for the imposition of penalties of up to five years in prison and up to \$250,000 in fines for wrongful disclosure of confidential census information.

Disclosure Avoidance

Disclosure avoidance is the process for protecting the confidentiality of data. A disclosure of data occurs when someone can use published statistical information to identify an individual who has provided information under a pledge of confidentiality. For data tabulations, the Census Bureau uses disclosure avoidance procedures to modify or remove the characteristics that put confidential information at risk for disclosure. Although it may appear that a table shows information about a specific individual, the Census Bureau has taken steps to disguise or suppress the original data while making sure the results are still useful. The techniques used by the Census Bureau to protect confidentiality in tabulations vary, depending on the type of data. All disclosure avoidance procedures are done prior to the whole person imputation into not-in-sample GQ facilities.

Data Swapping

Data swapping is a method of disclosure avoidance designed to protect confidentiality in tables of frequency data (the number or percent of the population with certain characteristics). Data swapping is done by editing the source data or exchanging records for a sample of cases when creating a table. A sample of households is selected and matched on a set of selected key variables with households in neighboring geographic areas that have similar characteristics (such as the same number of adults and same number of children). Because the swap often occurs within a neighboring area, there is no effect on the marginal totals for the area or for totals that include data from multiple areas. Because of data swapping, users should not assume that tables with cells having a value of one or two reveal information about specific individuals. Data swapping procedures were first used in the 1990 Census, and were used again in Census 2000 and the 2010 Census.

Synthetic Data

The goals of using synthetic data are the same as the goals of data swapping, namely to protect the confidentiality in tables of frequency data. Persons are identified as being at risk for disclosure based on certain characteristics. The synthetic data technique then models the values for another collection of characteristics to protect the confidentiality of that individual.

ERRORS IN THE DATA

Sampling Error

The data in ACS products are estimates of the actual figures that would be obtained by interviewing the entire population. The estimates are a result of the chosen sample, and are subject to sample-to-sample variation. Sampling error in data arises due to the use of probability sampling, which is necessary to ensure the integrity and representativeness of sample survey results. The implementation of statistical sampling procedures provides the basis for the statistical analysis of sample data. Measures used to estimate the sampling error are provided in the next section.

Nonsampling Error

Other types of errors might be introduced during any of the various complex operations used to collect and process survey data. For example, data entry from questionnaires and editing may introduce error into the estimates. Another potential source of error is the use of controls in the weighting. These controls are based on Population Estimates and are designed to reduce variance and mitigate the effects of systematic undercoverage of groups who are difficult to enumerate. However, if the extrapolation methods used in generating the Population Estimates do not properly reflect the population, error can be introduced into the data. This potential risk is offset by the many benefits the controls provide to the ACS estimates, which include the reduction of issues with survey coverage and the reduction of standard errors of ACS estimates. These and other sources of error contribute to the nonsampling error component of the total error of survey estimates.

Nonsampling errors may affect the data in two ways. Errors that are introduced randomly increase the variability of the data. Systematic errors, or errors that consistently skew the data in one direction, introduce bias into the results of a sample survey. The Census Bureau protects against the effect of systematic errors on survey estimates by conducting extensive research and evaluation programs on sampling techniques, questionnaire design, and data collection and processing procedures.

An important goal of the ACS is to minimize the amount of nonsampling error introduced through nonresponse for sample housing units. One way of accomplishing this is by following up on mail nonrespondents during the CATI and CAPI phases. For more information, please see the section entitled “Control of Nonsampling Error”.

MEASURES OF SAMPLING ERROR

Sampling error is the difference between an estimate based on a sample and the corresponding value that would be obtained if the entire population were surveyed (as for a census). Note that sample-based estimates will vary depending on the particular sample selected from the population. Measures of the magnitude of sampling error reflect the variation in the estimates over all possible samples that could have been selected from the population using the same sampling methodology.

Estimates of the magnitude of sampling errors – in the form of margins of error – are provided with all published ACS data. The Census Bureau recommends that data users incorporate margins of error into their analyses, as sampling error in survey estimates could impact the conclusions drawn from the results.

Confidence Intervals and Margins of Error

Confidence Intervals

A sample estimate and its estimated standard error may be used to construct confidence intervals about the estimate. These intervals are ranges that will contain the average value of the estimated characteristic that results over all possible samples, with a known probability.

For example, if all possible samples that could result under the ACS sample design were independently selected and surveyed under the same conditions, and if the estimate and its estimated standard error were calculated for each of these samples, then:

1. Approximately 68 percent of the intervals from one estimated standard error below the estimate to one estimated standard error above the estimate would contain the average result from all possible samples.
2. Approximately 90 percent of the intervals from 1.645 times the estimated standard error below the estimate to 1.645 times the estimated standard error above the estimate would contain the average result from all possible samples.
3. Approximately 95 percent of the intervals from two estimated standard errors below the estimate to two estimated standard errors above the estimate would contain the average result from all possible samples.

The intervals are referred to as 68 percent, 90 percent, and 95 percent confidence intervals, respectively.

Margins of Error

In lieu of providing upper and lower confidence bounds in published ACS tables, the margin of error is listed. All ACS published margins of error are based on a 90 percent confidence level. The margin of error is the difference between an estimate and its upper or lower confidence bound. Both the confidence bounds and the standard error can easily be computed from the margin of error:

$$\begin{aligned}\text{Standard Error} &= \text{Margin of Error} / 1.645 \\ \text{Lower Confidence Bound} &= \text{Estimate} - \text{Margin of Error} \\ \text{Upper Confidence Bound} &= \text{Estimate} + \text{Margin of Error}\end{aligned}$$

Note that for 2005 and earlier estimates, ACS margins of error and confidence bounds were calculated using a 90 percent confidence level multiplier of 1.65. Starting with the 2006 data release, and for every year after 2006, the more accurate multiplier of 1.645 is used. Margins of error and confidence bounds from previously published products will not be updated with the new multiplier. When calculating standard errors from margins of error or confidence bounds using published data for 2005 and earlier, use the 1.65 multiplier.

When constructing confidence bounds from the margin of error, users should be aware of any “natural” limits on the bounds. For example, if a characteristic estimate for the population is near zero, the calculated value of the lower confidence bound may be negative. However, as a negative number of people does not make sense, the lower confidence bound should be reported as zero. For other estimates such as income, negative values can make sense; in these cases, the lower bound should not be adjusted. The context and meaning of the estimate must therefore be kept in mind when creating these bounds. Another example of a natural limit is 100 percent as the upper bound of a percent estimate.

If the margin of error is displayed as ‘*****’ (five asterisks), the estimate has been controlled to be equal to a fixed value and so it has no sampling error. A standard error of zero should be used for these controlled estimates when completing calculations, such as those in the following section.

Limitations

Users should be careful when computing and interpreting confidence intervals.

Nonsampling Error

The estimated standard errors (and thus margins of error) included in these data products do not account for variability due to nonsampling error that may be present in the data. In particular, the standard errors do not reflect the effect of correlated errors introduced by interviewers, coders, or other field or processing personnel or the effect of imputed values

due to missing responses. The standard errors calculated are only lower bounds of the total error. As a result, confidence intervals formed using these estimated standard errors may not meet the stated levels of confidence (i.e., 68, 90, or 95 percent). Some care must be exercised in the interpretation of the data based on the estimated standard errors.

Very Small (Zero) or Very Large Estimates

By definition, the value of almost all ACS characteristics is greater than or equal to zero. The method provided above for calculating confidence intervals relies on large sample theory, and may result in negative values for zero or small estimates for which negative values are not admissible. In this case, the lower limit of the confidence interval should be set to zero by default. A similar caution holds for estimates of totals close to a control total or estimated proportion near one, where the upper limit of the confidence interval is set to its largest admissible value. In these situations, the level of confidence of the adjusted range of values is less than the prescribed confidence level.

CALCULATION OF STANDARD ERRORS

Direct estimates of margin of error were calculated for all estimates reported. The margin of error is derived from the variance. In most cases, the variance is calculated using a replicate-based methodology known as successive difference replication that takes into account the sample design and estimation procedures.

The formula provided below calculates the variance using the ACS estimate (X_0) and the 80 replicate estimates (X_r).

$$\text{Variance} = \frac{4}{80} \sum_{r=1}^{80} (x_r - x_0)^2$$

X_0 is the estimate calculated using the production weight and X_r is the estimate calculated using the r^{th} replicate weight. The standard error is the square root of the variance. The 90th percent margin of error is 1.645 times the standard error.

Additional information on the formation of the replicate weights, can be found in Chapter 12 of the Design and Methodology documentation at:

<https://www.census.gov/programs-surveys/acs/methodology/design-and-methodology.html>.

Beginning with the 2011 ACS 1-year estimates, a new imputation-based methodology was incorporated into processing (see the description in the Group Quarters Person Weighting Section). An adjustment was made to the production replicate weight variance methodology to

account for the non-negligible amount of additional variation being introduced by the new technique.⁸

Excluding the base weights, replicate weights were allowed to be negative in order to avoid underestimating the standard error. Exceptions include:

1. *The estimate of the number or proportion of people, households, families, or housing units in a geographic area with a specific characteristic is zero.* A special procedure is used to estimate the standard error.
2. *There are either no sample observations available to compute an estimate or standard error of a median, an aggregate, a proportion, or some other ratio, or there are too few sample observations to compute a stable estimate of the standard error.* The estimate is represented in the tables by “-” and the margin of error by “**” (two asterisks).
3. *The estimate of a median falls in the lower open-ended interval or upper open-ended interval of a distribution.* If the median occurs in the lowest interval, then a “-” follows the estimate, and if the median occurs in the upper interval, then a “+” follows the estimate. In both cases, the margin of error is represented in the tables by “***” (three asterisks).

Sums and Differences of Direct Standard Errors

Estimates of standard errors displayed in tables are for individual estimates. Additional calculations are required to estimate the standard errors for sums of or the differences between two or more sample estimates.

The standard error of the sum of two sample estimates is the square root of the sum of the two individual standard errors squared plus a covariance term. That is, for standard errors $SE(\widehat{X}_1)$ and $SE(\widehat{X}_2)$ of estimates \widehat{X}_1 and \widehat{X}_2 :

$$SE(\widehat{X}_1 \pm \widehat{X}_2) = \sqrt{[SE(\widehat{X}_1)]^2 + [SE(\widehat{X}_2)]^2 \pm 2\text{cov}(\widehat{X}_1, \widehat{X}_2)} \quad (1)$$

⁸ For more information regarding this issue, see Asiala, M. and Castro, E. 2012. Developing Replicate Weight-Based Methods to Account for Imputation Variance in a Mass Imputation Application. In JSM proceedings, Section on Survey Research Methods, Alexandria, VA: American Statistical Association.

The covariance measures the interactions between two estimates. Currently the covariance terms are not available. Data users should therefore use the following approximation:

$$SE(\widehat{X}_1 \pm \widehat{X}_2) = \sqrt{[SE(\widehat{X}_1)]^2 + [SE(\widehat{X}_2)]^2} \quad (2)$$

However, it should be noted that this approximation will underestimate or overestimate the standard error if the two estimates interact in either a positive or a negative way.

The approximation formula (2) can be expanded to more than two estimates by adding in the individual standard errors squared inside the radical. As the number of estimates involved in the sum or difference increases, the results of the approximation become increasingly different from the standard error derived directly from the ACS microdata. Care should be taken to work with the fewest number of estimates as possible. If there are estimates involved in the sum that are controlled, then the approximate standard error can be increasingly different. Later in this document, examples are provided to demonstrate issues associated with approximating the standard errors when summing large numbers of estimates together.

Ratios

The statistic of interest may be the ratio of two estimates, where the numerator *is not* a subset of the denominator. The standard error of this ratio between two sample estimates is approximated as:

$$SE\left(\frac{\widehat{X}}{\widehat{Y}}\right) = \frac{1}{\widehat{Y}} \sqrt{[SE(\widehat{X})]^2 + \frac{\widehat{X}^2}{\widehat{Y}^2} [SE(\widehat{Y})]^2} \quad (3)$$

Proportions/Percents

For a proportion (or percent), a ratio where the numerator *is* a subset of the denominator, a slightly different estimator is used. If $\widehat{P} = \widehat{X}/\widehat{Y}$ then the standard error of this proportion is approximated as:

$$SE(\widehat{P}) = \frac{1}{\widehat{Y}} \sqrt{[SE(\widehat{X})]^2 - \frac{\widehat{X}^2}{\widehat{Y}^2} [SE(\widehat{Y})]^2} \quad (4)$$

If $\widehat{Q} = 100\% * \widehat{P}$ (P is the proportion and Q is its corresponding percent), then

$SE(\widehat{Q}) = 100\% * SE(\widehat{P})$. Note the difference between the formulas to approximate the standard error for proportions (4) and ratios (3) - the plus sign in the previous formula has been replaced with a minus sign. If the value under the radical is negative, use the ratio standard error formula instead.

Percent Change

The statistic of interest is a percentage change from one time period to another, where the more current estimate is compared to an older estimate, for example, the percent change of a 2016 estimate to a 2015 estimate. If the current estimate = \hat{X} and the earlier estimate = \hat{Y} , then the standard error for the percent change is approximated as:

$$SE\left(\frac{\hat{X} - \hat{Y}}{\hat{Y}}\right) = SE\left(\frac{\hat{X}}{\hat{Y}} - 1\right) = SE\left(\frac{\hat{X}}{\hat{Y}}\right) \quad (5)$$

As a caveat, this formula does not take into account the correlation when calculating overlapping time periods.

Products

For a product of two estimates - for example, deriving a proportion's numerator by multiplying the proportion by its denominator - the standard error can be approximated as:

$$SE(\hat{X} \times \hat{Y}) = \sqrt{\hat{X}^2 \times [SE(\hat{Y})]^2 + \hat{Y}^2 \times [SE(\hat{X})]^2} \quad (6)$$

TESTING FOR SIGNIFICANT DIFFERENCES

Users may conduct a statistical test to see if the difference between an ACS estimate and any other chosen estimate is statistically significant at a given confidence level. "Statistically significant" means that it is not likely that the difference between estimates is due to random chance alone.

To perform statistical significance testing, first calculate a Z statistic from the two estimates (Est_1 and Est_2) and their respective standard errors (SE_1 and SE_2):

$$Z = \frac{Est_1 - Est_2}{\sqrt{(SE_1)^2 + (SE_2)^2}} \quad (7)$$

If $Z > 1.645$ or $Z < -1.645$, then the difference can be said to be statistically significant at the 90 percent confidence level.⁹

Any pair of estimates can be compared using this method, including ACS estimates from the current year, ACS estimates from a previous year, 2010 Census counts, estimates from other Census Bureau surveys, and estimates from other sources. Not all estimates have sampling error (2010 Census counts do not, for example), but when possible, standard errors should be used to produce the most accurate test result.

Users are also cautioned to *not* rely on looking at whether confidence intervals for two estimates overlap in order to determine statistical significance. There are circumstances where comparing confidence intervals will not give the correct test result. If two confidence intervals do not overlap, then the estimates will be significantly different (i.e. the significance test will always agree). However, if two confidence intervals do overlap, then the estimates may or may not be significantly different. The Z calculation shown above is recommended in all cases.

The following example illustrates why using the overlapping confidence bounds rule of thumb as a substitute for a statistical test is not recommended.

Let: $X_1 = 6.0$ with $SE_1 = 0.5$ and $X_2 = 5.0$ with $SE_2 = 0.2$.

The Lower Bound for $X_1 = 6.0 - 0.5 * 1.645 = 5.2$ while the Upper Bound for $X_2 = 5.0 + 0.2 * 1.645 = 5.3$. The confidence bounds overlap, so, the rule of thumb would indicate that the estimates are not significantly different at the 90% level.

However, if we apply the statistical significance test we obtain:

$$Z = \frac{6 - 5}{\sqrt{(0.5)^2 + (0.2)^2}} = 1.857$$

$Z = 1.857 > 1.645$ which means that the difference is significant (at the 90% level).

All statistical testing in ACS data products is based on the 90 percent confidence level. Users should understand that all testing was done using *unrounded* estimates and standard errors, and it may not be possible to replicate test results using the rounded estimates and margins of error as published.

⁹ The ACS Accuracy of the Data document in 2005 used a Z statistic of +/-1.65. Data users should use +/-1.65 for estimates published in 2005 or earlier.

Users completing statistical testing may be interested in using the ACS Statistical Testing Tool. This automated tool allows users to input pairs and groups of estimates for comparison. For more information on the Statistical Testing Tool, visit <https://www.census.gov/programs-surveys/acs/guidance/statistical-testing-tool.html>.

EXAMPLES OF STANDARD ERROR CALCULATIONS

Example 1 – Calculating the Standard Error from the Margin of Error

The estimated number of males, never married is 47,194,876 as found on summary table B12001 (Sex by Marital Status for the Population 15 Years and Over) for the United States for 2016. The margin of error is 89,037. Recall that:

$$\text{Standard Error} = \text{Margin of Error} / 1.645$$

Calculating the standard error using the margin of error, we have:

$$SE(47,194,876) = \frac{89,037}{1.645} = 54,126$$

Example 2 – Calculating the Standard Error of a Sum or a Difference

We are interested in the total number of people who have never been married. From Example 1, we know the number of males, never married is 47,194,876. From summary table B12001 we have the number of females, never married is 41,142,530 with a margin of error of 84,363. Therefore, the estimated number of people who have never been married is $47,194,876 + 41,142,530 = 88,337,406$.

To calculate the approximate standard error of this sum, we need the standard errors of the two estimates in the sum. We calculated the standard error for the number of males never married in Example 1 as 54,126. The standard error for the number of females never married is calculated using the margin of error:

$$SE(41,142,530) = \frac{84,363}{1.645} = 51,284$$

Using formula (2) for the approximate standard error of a sum or difference we have:

$$SE(88,337,406) = \sqrt{54,126^2 + 51,284^2} = 74,563$$

Caution: This method will underestimate or overestimate the standard error if the two estimates interact in either a positive or a negative way.

To calculate the lower and upper bounds of the 90 percent confidence interval around 88,337,406 using the standard error, simply multiply 74,563 by 1.645, then add and subtract the product from 88,337,406. Thus the 90 percent confidence interval for this estimate is [88,337,406 - 1.645(74,563)] to [88,337,406 + 1.645(74,563)] or 88,460,062 to 88,214,750.

Example 3 – Calculating the Standard Error of a Proportion/Percent

We are interested in the percentage of females who have never been married to the number of people who have never been married. The number of females, never married is 41,142,530 and the number of people who have never been married is 88,337,406. To calculate the approximate standard error of this percent, we need the standard errors of the two estimates in the percent.

From Example 2, we know that the approximate standard error for the number of females never married is 51,284 and the approximate standard error for the number of people never married calculated is 74,563.

The estimate is:

$$\left(\frac{41,142,530}{88,337,406}\right) * 100\% = 46.57\%$$

Therefore, using formula (4) for the approximate standard error of a proportion or percent, we have:

$$SE(46.57\%) = 100\% * \left(\frac{1}{88,337,406} \sqrt{51,284^2 - 0.4657^2 * 74,563^2}\right) = 0.04$$

To calculate the lower and upper bounds of the 90 percent confidence interval around 46.57 using the standard error, simply multiply 0.04 by 1.645, then add and subtract the product from 46.57. Thus the 90 percent confidence interval for this estimate is:

$$[46.57 - 1.645(0.04)] \text{ to } [46.57 + 1.645(0.04)], \text{ or } 46.50\% \text{ to } 46.64\%.$$

Example 4 – Calculating the Standard Error of a Ratio

We are interested in the ratio of the number of unmarried males to the number of unmarried females. From Examples 1 and 2, we know that the estimate for the number of unmarried men is 47,194,876 with a standard error of 54,126, and the estimate for the number of unmarried women is 41,142,530 with a standard error of 51,284.

The estimate of the ratio is:

$$47,194,876 / 41,142,530 = 1.147.$$

Using formula (3) for the approximate standard error of this ratio, we have:

$$SE(1.147) = \frac{1}{41,142,530} \sqrt{54,126^2 + 1.147^2 * 51,284^2} = 0.002$$

The 90 percent margin of error for this estimate would be 0.002 multiplied by 1.645, or about 0.003. The 90 percent lower and upper 90 percent confidence bounds would then be [1.147 – 1.645(0.002)] to [1.147 + 1.645(0.002)], or 1.144 and 1.150.

Example 5 – Calculating the Standard Error of a Product

We are interested in the number of single unit detached owner-occupied housing units. The number of owner-occupied housing units is 75,022,569 with a margin of error of 227,992, as found in subject table S2504 (Physical Housing Characteristics for Occupied Housing Units) for 2016, and the percent of single unit detached owner-occupied housing units (called “1, detached” in the subject table) is 82.5% (0.825) with a margin of error of 0.1 (0.001).

Therefore, the number of 1-unit detached owner-occupied housing units is:

$$75,022,569 * 0.825 = 61,893,619.$$

Calculating the standard error for the estimates using the margin of error we have:

$$SE(75,022,569) = \frac{227,992}{1.645} = 138,597$$

and

$$SE(0.825) = \frac{0.001}{1.645} = 0.0006079$$

Using formula (6), the approximate standard error for number of 1-unit detached owner-occupied housing units is:

$$SE(61,893,619) = \sqrt{75,022,569^2 * 0.0006079^2 + 0.825^2 * 138,597^2} = 123,102$$

To calculate the lower and upper bounds of the 90 percent confidence interval around 61,893,619 using the standard error, simply multiply 123,102 by 1.645, then add and subtract the product from 61,893,619. Thus the 90 percent confidence interval for this estimate is [61,893,619 - 1.645(123,102)] to [61,893,619 + 1.645(123,102)] or 61,691,116 to 62,096,122.

CONTROL OF NONSAMPLING ERROR

As mentioned earlier, sample data are subject to nonsampling error. Nonsampling error can introduce serious bias into the data, increasing the total error dramatically over that which would result purely from sampling. While it is impossible to completely eliminate nonsampling error from a survey operation, the Census Bureau attempts to control the sources of such error during the collection and processing operations. Described below are the primary sources of nonsampling error and the programs instituted to control for this error.¹⁰

Coverage Error

It is possible for some sample housing units or persons to be missed entirely by the survey (undercoverage). It is also possible for some sample housing units and persons to be counted more than once (overcoverage). Both undercoverage and overcoverage of persons and housing units can introduce bias into the data. Coverage error can also increase both respondent burden and survey costs.

To avoid coverage error in a survey, the frame must be as complete and accurate as possible. For the ACS, the frame is an address list in each state, the source of which is the Master Address File (MAF). An attempt is made to assign each MAF address to the appropriate geographic codes via an automated procedure using the Census Bureau TIGER (Topologically Integrated Geographic Encoding and Referencing) files. A manual coding operation based in the appropriate regional offices is attempted for addresses that could not be automatically coded.

In 2016, the MAF was used as the source of addresses for selecting sample housing units and mailing questionnaires. TIGER produced the location maps for CAPI assignments. Sometimes the MAF contains duplicates of addresses. This could occur when there is a slight difference in the address such as 123 Main Street versus 123 Maine Street, and can introduce overcoverage.

In the CATI and CAPI nonresponse follow-up phases, efforts were made to minimize the chances that housing units that were not part of the sample were mistakenly interviewed instead of units in sample. If a CATI interviewer called a mail nonresponse case and was not able to reach the exact address, no interview was conducted and the case became eligible for CAPI. During the CAPI follow-up, the interviewer had to locate the exact address for each sample housing unit. If the interviewer could not locate the exact sample unit in a multi-unit structure, or found a different number of units than expected, the interviewers were instructed to list the units in the building and follow a specific procedure to select a replacement sample

¹⁰ The success of these programs is contingent upon how well the instructions were carried out during the survey.

unit. Person overcoverage can occur when an individual is included as a member of a housing unit but does not meet ACS residency rules.

Coverage rates give a measure of undercoverage or overcoverage of persons or housing units in a given geographic area. Rates below 100 percent indicate undercoverage, while rates above 100 percent indicate overcoverage. Coverage rates are released concurrent with the release of estimates on American FactFinder in the B98 series of detailed tables (Table IDs B98011, B98012, B98013, and B980014). Coverage rate definitions and coverage rates for total population for nation and states are also available in the Sample Size and Data Quality Section of the ACS website, at <https://www.census.gov/acs/www/methodology/sample-size-and-data-quality/>.

Nonresponse Error

Survey nonresponse is a well-known source of nonsampling error. There are two types of nonresponse error – unit nonresponse and item nonresponse. Nonresponse errors affect survey estimates to varying levels depending on amount of nonresponse and the extent to which the characteristics of nonrespondents differ from those of respondents. The exact amount of nonresponse error or bias on an estimate is almost never known. Therefore, survey researchers generally rely on proxy measures, such as the nonresponse rate, to indicate the potential for nonresponse error.

Unit Nonresponse

Unit nonresponse is the failure to obtain data from housing units in the sample. Unit nonresponse may occur because households are unwilling or unable to participate, or because an interviewer is unable to make contact with a housing unit. Unit nonresponse is problematic when there are systematic or variable differences in the characteristics of interviewed and non-interviewed housing units. Nonresponse bias is introduced into an estimate when differences are systematic; the nonresponse error of an estimate evolves from variable differences between interviewed and non-interviewed households.

The ACS made every effort to minimize unit nonresponse, and thus, the potential for nonresponse error. First, the ACS used a combination of mail, CATI, and CAPI data collection modes to maximize response. The mail phase included a series of three to four mailings to encourage housing units to return the questionnaire. Subsequently, mail nonrespondents (for which phone numbers are available) were contacted by CATI for an interview. Finally, a subsample of the mail and telephone nonrespondents were contacted by personal visit to attempt an interview. Combined, these three efforts resulted in a very high overall response rate for the ACS.

ACS response rates measure the percent of units with a completed interview. The higher the response rate (and, consequently, the lower the nonresponse rate), the lower the chance that

estimates are affected by nonresponse bias. Response and nonresponse rates, as well as rates for specific types of nonresponse, are released concurrent with the release of estimates on American FactFinder in the B98 series of detailed tables (Table IDs B98021 and B98022). Unit response rate definitions and unit response rates by type for the nation and states are also available in the Sample Size and Data Quality Section of the ACS website, at <https://www.census.gov/acs/www/methodology/sample-size-and-data-quality/>.

Item Nonresponse

Nonresponse to particular questions on the survey can introduce error or bias into the data, as the unknown characteristics of nonrespondents may differ from those of respondents. As a result, any imputation procedure using respondent data may not completely reflect difference either at the elemental level (individual person or housing unit) or on average.

Some protection against the introduction of large errors or biases is afforded by minimizing nonresponse. In the ACS, item nonresponse for the CATI and CAPI operations was minimized by requiring that the automated instrument receive a response to each question before the next question could be asked. Questionnaires returned by mail were reviewed by computer for content omissions and population coverage and edited for completeness and acceptability. If necessary, a telephone follow-up was made to obtain missing information. Potential coverage errors were included in this follow-up.

Allocation tables provide the weighted estimate of persons or housing units for which a value was imputed, as well as the total estimate of persons or housing units that were eligible to answer the question. The smaller the number of imputed responses, the lower the chance that the item nonresponse is contributing a bias to the estimates. Allocation tables are released concurrent with the release of estimates on American Factfinder in the B99 series of detailed tables with the overall allocation rates across all person and housing unit characteristics in the B98 series of detailed tables (Table IDs B98031 and B98032). Allocation rate definitions and allocation rates by characteristic at the nation, and states are also available in the Sample Size and Data Quality Section of the ACS website, at <https://www.census.gov/acs/www/methodology/sample-size-and-data-quality/>.

Measurement and Processing Error

Measurement error can arise if the person completing the questionnaire or responding an interviewer's questions responds incorrectly. However, to mitigate this risk, the phrasing survey questions underwent cognitive testing and households were provided detailed instructions on how to complete the questionnaire.

Processing error can be introduced in numerous areas during data collection and capture, including during interviews, during data processing and during content editing.

Interviewer monitoring

An interviewer could introduce error by:

1. Misinterpreting or otherwise incorrectly entering information given by a respondent.
2. Failing to collect some of the information for a person or household.
3. Collecting data for households that were not designated as part of the sample.

To control for these problems, the work of interviewers was monitored carefully. Field staff was prepared for their tasks by using specially developed training packages that included hands-on experience in using survey materials. A sample of the households interviewed by CAPI interviewers was also reinterviewed to control for the possibility that interviewers may have fabricated data.

Processing Error

The many phases involved in processing the survey data represent potential sources for the introduction of nonsampling error. The processing of the survey questionnaires includes the keying of data from completed questionnaires, automated clerical review, follow-up by telephone, manual coding of write-in responses, and automated data processing. The various field, coding and computer operations undergo a number of quality control checks to insure their accurate application.

Content Editing

After data collection was completed, any remaining incomplete or inconsistent information was imputed during the final content edit of the collected data. Imputations, or computer assignments of acceptable codes in place of unacceptable entries or blanks, were most often needed either when an entry for a given item was missing or when information reported for a person or housing unit was inconsistent with other information for the same person or housing unit. As in other surveys and previous censuses, unacceptable entries were to allocated entries for persons or housing units with similar characteristics. Imputing acceptable values in place of blanks or unacceptable entries enhances the usefulness of the data.

ISSUES WITH APPROXIMATING THE STANDARD ERROR OF LINEAR COMBINATIONS OF MULTIPLE ESTIMATES

The following examples demonstrate how approximated standard errors of sums can differ from those derived and published with ACS microdata.¹¹

Example A

Suppose we are interested in the total number of males with income below the poverty level in the past 12 months for the state of Wyoming. We want to find the estimate using both state and PUMA level estimates. Part of the collapsed table C17001 is displayed in Table A below.

Table A: Excerpt from C17001: Poverty Status in the Past 12 Months by Sex by Age (2009)

Characteristic	Wyoming		PUMA 00100		PUMA 00200		PUMA 00300		PUMA 00400	
	Estimate	MOE	Estimate	MOE	Estimate	MOE	Estimate	MOE	Estimate	MOE
Male	23,001	3,309	5,264	1,624	6,508	1,395	4,364	1,026	6,865	1,909
Under 18 Years Old	8,479	1,874	2,041	920	2,222	778	1,999	750	2,217	1,192
18 to 64 Years Old	12,976	2,076	3,004	1,049	3,725	935	2,050	635	4,197	1,134
65 Years and Older	1,546	500	219	237	561	286	315	173	451	302

Source: 2009 American FactFinder

First, sum the three state-level male age group estimates for Wyoming:

$$\text{Estimate(Male)} = 8,479 + 12,976 + 1,546 = 23,001.$$

The approximation for the standard error for the summed state-level age groups is:

$$\text{SE(Male)} = \sqrt{\left(\frac{1,874}{1.645}\right)^2 + \left(\frac{2,076}{1.645}\right)^2 + \left(\frac{500}{1.645}\right)^2} = 1,727.1$$

Next, sum the four PUMA estimates for males:

$$\text{Estimate(Male)} = 5,264 + 6,508 + 4,364 + 6,865 = 23,001.$$

¹¹ Due to differences in the definition, in rare instances summing PUMA estimates within a state may not equal the state estimate.

The approximation for the standard error of the summed PUMA level estimates is:

$$SE(\text{Male}) = \sqrt{\left(\frac{1,624}{1.645}\right)^2 + \left(\frac{1,395}{1.645}\right)^2 + \left(\frac{1,026}{1.645}\right)^2 + \left(\frac{1,909}{1.645}\right)^2} = 1,851.9$$

Finally, we will sum up all three age groups for all four PUMAs to obtain a third estimate of males:

$$\text{Estimate}(\text{Male}) = 2,041 + 2,222 + \dots + 451 = 23,001$$

The approximated standard error for the summed age-group PUMA level estimates:

$$SE(\text{Male}) = \sqrt{\left(\frac{920}{1.645}\right)^2 + \left(\frac{778}{1.645}\right)^2 + \dots + \left(\frac{302}{1.645}\right)^2} = 1,649.0$$

We also know that the standard error using the published MOE is $3,309 / 1.645 = 2,011.6$.

In this instance, all of the approximations under-estimate the published standard error and should be used with caution.

Example B

Suppose we wish to estimate the total number of males at the national level using age and citizenship status. The relevant data from table B05003 is displayed in Table B below.

Table B: Excerpt from B05003: Sex by Age by Citizenship Status (2009)

Characteristic	Estimate	MOE
Male	151,375,321	27,279
Under 18 Years	38,146,514	24,365
Native	36,747,407	31,397
Foreign Born	1,399,107	20,177
Naturalized U.S. Citizen	268,445	10,289
Not a U.S. Citizen	1,130,662	20,228
18 Years and Older	113,228,807	23,525
Native	95,384,433	70,210
Foreign Born	17,844,374	59,750
Naturalized U.S. Citizen	7,507,308	39,658
Not a U.S. Citizen	10,337,066	65,533

Source: 2009 American FactFinder

The estimate and its MOE that we are interested in are actually published. However, if they were not available in the tables, we could find calculate them. To find the estimate for the number of males, we would sum the number of males under 18 and over 18:

$$\text{Estimate(Male)} = 38,146,514 + 113,228,807 = 151,375,321$$

The approximated standard error is:

$$\text{SE(Male)} = \text{SE}(151,375,321) = \sqrt{\left(\frac{24,365}{1.645}\right)^2 + \left(\frac{23,525}{1.645}\right)^2} = 20,588.8$$

Another method would be to add up the estimates for the three subcategories (Native, and the two subcategories for Foreign Born: Naturalized U.S. Citizen, and Not a U.S. Citizen), for males under and over 18 years of age.

From these six estimates we find:

$$\begin{aligned} \text{Estimate(Male)} &= 36,747,407 + 268,445 + 1,130,662 \\ &\quad + 95,384,433 + 7,507,308 + 10,337,066 \\ &= 151,375,321 \end{aligned}$$

With an approximated standard error of:

$$\begin{aligned} \text{SE(Male)} &= \text{SE}(151,375,321) \\ &= \sqrt{\left(\frac{31,397}{1.645}\right)^2 + \left(\frac{10,289}{1.645}\right)^2 + \left(\frac{20,228}{1.645}\right)^2 + \left(\frac{70,210}{1.645}\right)^2 + \left(\frac{39,658}{1.645}\right)^2 + \left(\frac{65,533}{1.645}\right)^2} \\ &= 67,413.0 \end{aligned}$$

We know that the standard error using the published margin of error is $27,279 / 1.645 = 16,583.0$.

With a quick glance, we can see that the ratio of the standard error of the first method to the published-based standard error yields 1.24, an over-estimate of roughly 24%, whereas the second method yields a ratio of 4.07 or an over-estimate of 307%. This is an example of what could happen to the approximate SE when the sum involves a controlled estimate. In this case, the controlled estimate is sex by age.

Example C

Suppose we are interested in the total number of people aged 65 or older. Table C shows some of the estimates at the national level from table B01001 (the estimates in gray were derived for the purpose of this example only).

Table C: Estimates from AFF Table B01001: Sex by Age (2009)

Age Category	Estimate, Male	MOE, Male	Estimate, Female	MOE, Female	Total	Estimated MOE, Total
65 and 66 years old	2,492,871	20,194	2,803,516	23,327	5,296,387	30,854
67 to 69 years old	3,029,709	18,280	3,483,447	24,287	6,513,156	30,398
70 to 74 years old	4,088,428	21,588	4,927,666	26,867	9,016,094	34,466
75 to 79 years old	3,168,175	19,097	4,204,401	23,024	7,372,576	29,913
80 to 84 years old	2,258,021	17,716	3,538,869	25,423	5,796,890	30,987
85 years and older	1,743,971	17,991	3,767,574	19,294	5,511,545	26,381
Total	16,781,175	NA	22,725,473	NA	39,506,648	74,932

Source: 2009 American FactFinder

To begin we find the total number of people aged 65 and over by adding the totals for males and females:

$$16,781,175 + 22,725,542 = 39,506,717$$

An alternate method would be to sum males and female for each age category. We could then use the MOEs for the age category estimates to approximate the standard error for the total number of people over 65.

$$\text{MOE}(\text{Age 65 and 66}) = \text{MOE}(5,296,387) = \sqrt{20,194^2 * 23,327^2} = 30,854$$

$$\text{MOE}(\text{Age 67 and 69}) = \text{MOE}(6,513,225) = \sqrt{18,280^2 * 24,287^2} = 30,398$$

... etc. ...

With this method, we calculate for the number of people aged 65 or older to be 39,506,648.

We approximate the standard error as:

$$\begin{aligned} \text{SE}(39,506,648) &= \sqrt{\left(\frac{30,854}{1.645}\right)^2 + \left(\frac{30,398}{1.645}\right)^2 + \left(\frac{34,466}{1.645}\right)^2 + \left(\frac{29,913}{1.645}\right)^2 + \left(\frac{30,987}{1.645}\right)^2 + \left(\frac{26,381}{1.645}\right)^2} \\ &= 45,552 \end{aligned}$$

For this example, the estimate and its MOE are published in table B09017. As such, we know that the total number of people aged 65 or older is 39,506,648 with a margin of error of 20,689.

Therefore the published-based standard error is:

$$SE(39,506,648) = \frac{20,689}{1.645} = 12,577$$

The approximated standard error, calculated using six derived age group estimates, yields an approximated standard error roughly 3.6 times larger than the published-based standard error.

As a note, there are two additional ways to approximate the standard error of people aged 65 and over in addition to the way used above. The first is to find the published MOEs for the males age 65 and older and of females aged 65 and older separately and combine them to find the approximate standard error of the total. The second is to use all twelve of the published estimates together (all estimates from the male age categories and female age categories) to create the SE for people aged 65 and older. In this particular example, the results from all three ways are the same; the same approximation for the SE is obtained regardless of the method. This result differs from that found in Example A.

Example D

This example gives an alternative to the methodology of Example C. Here, we derive the estimate and its corresponding SE by summing the estimates for the ages less than 65 years old and subtracting them from the estimate for the total population. Due to the large number of estimates, Table D does not show all of the age groups. Again, the estimates in shaded part of the table were derived for the purposes of this example and cannot be found in table B01001.

Table D: Estimates from AFF Table B01001: Sex by Age (2009)

Age Category	Estimate, Male	MOE, Male	Estimate, Female	MOE, Female	Total	Estimated MOE, Total
Total Population	151,375,321	27,279	155,631,235	27,280	307,006,556	38,579
Under 5 years	10,853,263	15,661	10,355,944	14,707	21,209,207	21,484
5 to 9 years old	10,273,948	43,555	9,850,065	42,194	20,124,013	60,641
10 to 14 years old	10,532,166	40,051	9,985,327	39,921	20,517,493	56,549
...		
62 to 64 years old	4,282,178	25,636	4,669,376	28,769	8,951,554	38,534
Total for Age 0 to 64 years old	134,594,146	117,166	132,905,762	117,637	267,499,908	166,031
Total for Age 65 years and older	16,781,175	120,300	22,725,473	120,758	39,506,648	170,454

Source: 2009 American FactFinder

To find an estimate for the number of people age 65 and older, subtract the population between the ages of zero and 64 years old from the total population:

Number of people aged 65 and older:

$$307,006,556 - 267,499,908 = 39,506,648.$$

The SE approximation uses the same methodology as in part C. First, sum male and female estimates across each age category, then approximate the MOEs:

$$\text{MOE}(\text{Total Population}) = \text{MOE}(307,006,556) = \sqrt{27,279^2 * 27,280^2} = 38,579$$

$$\text{MOE}(\text{Under 5 years}) = \text{MOE}(21,209,207) = \sqrt{15,661^2 * 14,707^2} = 21,484$$

... etc. ...

The SE for the total number of people aged 65 and older is:

$$\text{SE}(\text{Age 65 and older}) = \text{SE}(39,506,648)$$

$$= \sqrt{\left(\frac{38,579}{1.645}\right)^2 + \left(\frac{21,484}{1.645}\right)^2 + \left(\frac{60,641}{1.645}\right)^2 + \left(\frac{56,549}{1.645}\right)^2 + \dots + \left(\frac{38,534}{1.645}\right)^2}$$

$$= 103,620$$

Again, as in Example C, the estimate and its MOE are we published in B09017. The total number of people aged 65 or older is 39,506,648 with a margin of error of 20,689. Therefore the standard error is:

$$\text{SE}(39,506,648) = \frac{20,689}{1.645} = 12,577$$

The approximated standard error using the thirteen derived age group estimates yields a standard error roughly 8.2 times larger than the actual SE.

Additional Resources

Data users can mitigate the problems shown in examples A through D by utilizing a collapsed version of a detailed table or the less detailed annual Supplemental Tables. Using these tables, if available, may reduce the number of estimates used in the approximation. These issues may also be avoided by creating estimates and SEs using the Public Use Microdata Sample (PUMS) or by requesting a custom tabulation, a fee-based service offered under certain conditions by the Census Bureau. For more information regarding custom tabulations, visit: <https://www.census.gov/programs-surveys/acs/data/custom-tables.html>.