

American Housing Survey

Getting Started with the Public Use File: 1973 - 1996

v1.0

Updated October 19, 2022

U.S. Census Bureau, Department of Commerce
Department of Housing and Urban Development

Contents

Overview	3
Sample Design.....	3
The AHS-N Sample and Sample Changes	3
The AHS-MS Sample and Sample Changes	4
Survey Design/Mode.....	5
Evolution of Dependent Interviewing.....	5
File Structure.....	5
PUF Flat File Structure.....	5
CSV Structure	6
1973 - 1996 Interactive Codebook.....	6
1973 - 1996 Internal Use Files.....	7
Noninterview Types	8
Variable Formatting	8
Not Applicable vs. Not Reported	9
Geography	9
Core and Supplemental Topics	11
Weights	11
Edits and Imputation.....	13
Topcoding and Disclosure	13
Appendix	14
Variable Drop Criteria	14
Variable Type Change Criteria	14
Variable Renames During Revisions.....	14
Metro Panel Years.....	18
Example SAS Code.....	20
Example 1: Generating a Table	20
Example 2: Linking Files	22

Overview

The purpose of this document is to introduce American Housing Survey (AHS) users to the 1973 - 1996 AHS public use file (PUF) microdata. In October 2022, HUD and Census released revisions to all national and metro PUFs from 1973 through 1996 to increase usability for data users. The list below summarizes the revisions made to the files, which are also covered in more detail throughout this guide.

1. Variables without usable information were removed from the files. This is covered in the Variable Formatting section of this guide.
2. Standardized formatting rules were applied to all variables. This is also discussed in more detail in the Variable Formatting section.
3. Noninterview sample, along with variables only relevant for non-interview data, were removed from the files. This is covered in more detail in the Respondent Types section.
4. [Electronic codebook](#) entries for all historic files were reviewed for accuracy, and corrections and updates were applied. This is described in more detail in the Interactive Codebook section.

The remainder of this guide is organized into sections, with each section addressing an important PUF microdata topic.

Sample Design

Throughout its history, the AHS has always consisted of a longitudinal national sample (hereinafter referred to as AHS-N) and independent longitudinal samples of selected metropolitan areas (hereinafter referred to as AHS-MS). Over the length of a longitudinal sample in the AHS, the same housing units are interviewed across multiple survey years, with new housing units added to the sample each year to represent new construction. While this guide focuses on the 1973 - 1996 period, it is important to note that the AHS-N sample is longitudinal (1973-1983 and 1985-2013), and many of the AHS-MS samples are longitudinal from 1974 - 2013. From 1985 - 2013, the AHS-N sample was surveyed in every odd-numbered year. The AHS-MS samples were surveyed periodically throughout the same period, but unlike national, could also exist in even-numbered years. For a full listing of the metro areas sampled in each year, see <https://www.census.gov/content/dam/Census/programs-surveys/ahs/tech-documentation/Quick%20Guide%20to%20Metro%20Area%20Histories%201974%20to%20Current.pdf>.

The AHS-N Sample and Sample Changes

The AHS-N sample was designed to be representative of the nation as well as the four large Census regions and nine Census divisions within the U.S. The national sample was not designed to produce estimates of individual metropolitan area, states, or other areas smaller than the Census division.

The AHS-N sample underwent several changes during the 1973 - 1996 period. The original samples for the AHS-N from 1973 to 1983 were drawn from the 1970 Census. The table below summarizes the changes that occurred after that and their impact on the ability to produce national estimates.

Description of Change	Impact
1985 Sample Redrawn In 1985, AHS-N was administered to a completely new sample drawn from the 1980 Census.	<ul style="list-style-type: none">• This change had no impact on the ability to produce national estimates from the AHS-N PUF.• This change did, however, impact the ability to track changes in particular housing units longitudinally from the prior panel because

	<p>the sample consisted of entirely new households.</p> <ul style="list-style-type: none"> • This new sample panel was surveyed every two years from 1985 to 2013.
<p>Big 6 Oversample These cases were added to the AHS-N sample for 1995 files. These cases are in the Chicago, Detroit, Los Angeles, New York, Northern New Jersey, and Philadelphia Metropolitan Statistical Areas. They continue into later years, as covered in Getting Started with the PUF: 1997-2013.</p>	<ul style="list-style-type: none"> • This change meant that PUF users wishing to produce estimates for the six metropolitan areas would need to use the AHS-N PUF. • This change had no impact on the ability to produce national estimates from the AHS-N PUF.
<p>1987 Sample Drop Some eligible units were missed in the 1987 interview process but recovered in 1989 and later.</p>	<ul style="list-style-type: none"> • This change had no impact on the ability to produce national estimates from the AHS-N PUF.
<p>Neighborhood Sample In 1985, 1989, and 1993 certain AHS-N cases were selected as part of the neighborhood sample. For any case that was in the neighborhood sample, the AHS also collected information regarding up to 10 of its neighbors. These extra cases were not considered to be part of the AHS-N sample.</p>	<ul style="list-style-type: none"> • Although these cases appear on the PUF, they have a zero weight value, meaning they do not contribute to producing national estimates. • For years 1985, 1989, and 1993, neighbor cases can be identified using the variables NEIGH and RURREC.
<p>Rural Oversample These cases were part of a special rural oversample that took place in 1987 and 1991 and only appear on PUFs for those years.</p>	<ul style="list-style-type: none"> • Although these cases appear on the PUF, they have a zero weight value, meaning they do not contribute to producing national estimates.
<p>Extra rural sample These cases were part of a rural oversample that took place from 1974 through 1983 and only appear on the National PUFs for those years.</p>	<ul style="list-style-type: none"> • For years 1974N-83N, rural oversample can be identified using the variable RURREC, which is the flag for the basic national sample. • To make estimates of the 'basic national sample for these years, use cases where RURREC = '1'. • To make estimates of all rural areas (codes 2-5 in URBAN), add cases where RURREC = '2', since these RURREC = '2' cases double-count certain rural units in metro areas whose rural code is suppressed (URBAN=9) in the 'basic national sample.'

For more information about the sample design in each survey year, see Accuracy of the Data at <https://www.census.gov/programs-surveys/ahs/tech-documentation/def-errors-changes.html>.

The AHS-MS Sample and Sample Changes

The data from the independent metropolitan surveys are contained in a single metropolitan dataset for their year. Within that dataset, the cases from each metropolitan area are identified by the SMSA variable. This is important because different metro areas could use the same CONTROL value.

Like with the AHS-N sample, the independent AHS-MS samples are also longitudinal, meaning that they return to the same housing units in each survey year. Data users can use CONTROL and SMSA to link AHS-MS sample longitudinally.

The Appendix of this document displays the range of survey years covered in each panel for each metro sample along with the SMSA code used to identify each metro area. The metro panels included in the table cover all independent metro samples from 1974 through 2013 and exclude metro sample that was embedded into the national sample.

For more detailed history of the independent metropolitan area longitudinal oversamples, users are strongly encouraged to consult the Metropolitan Oversample Histories: 1973 - 2013 guide at

https://www.census.gov/programs-surveys/ahs/tech-documentation/help-guides/metro_oversamp_hist_2013.html.

Survey Design/Mode

From 1973 through 1996, the AHS was conducted using a paper survey questionnaire administered by Census Bureau Field Representatives (FRs) who visited respondents in their homes to conduct interviews in person. Beginning in 1997, the survey began using computer-assisted personal interviewing (CAPI), where an electronic questionnaire on the FR's laptop was used to conduct the interview. Most interviews were conducted in person, although some were conducted by telephone.

Evolution of Dependent Interviewing

In the paper questionnaire, values for a limited number of variables were carried forward using a paper control card that was kept for each survey year. After development of the electronic questionnaire in 1997, the AHS had the capability to expand the number of values carried forward electronically. A more detailed overview of dependent interviewing can be found at

https://www.census.gov/content/dam/Census/programs-surveys/ahs/tech-documentation/1985-2013/Dependent_Interviewing_1985_to_2013.pdf.

File Structure

Separate PUFs are published for the integrated national longitudinal sample and the metropolitan area samples during the 1973 - 1996 period. These PUFs are referred to as the national PUF and the metropolitan area PUF.

Since 1997, HUD and the Census Bureau have provided the PUF in two structures—relational and flat. From 1973 through 1996, HUD and the Census Bureau provided the PUF in only one format structure called the flat file. The structure of the PUFs is discussed in the next two sections.

All PUFs are published in SAS and CSV formats. Besides the difference in file type, one other important difference between the SAS and CSV formats is that the SAS files now include descriptive column labels.

PUF Flat File Structure

In a flat file, there is one file that contains all the data. Each record represents a housing unit. One-to-many relationships (such as demographic data for persons) are represented by numbered sets of variables. For example, there will be a RACE1, SEX1, GRAD1, etc. for the first person in the household, and another set for the second person (RACE2, SEX2, GRAD2, etc.). If a housing unit has fewer than the maximum number of such entities, the extra sets of variables are filled with missing values.

The following standards were applied to all years:

- For AHS-N PUFs, there will always be exactly one row in the flattened file for each unique CONTROL. For AHS-MS PUFs, there will always be exactly one row for each unique combination of CONTROL and SMSA.
- Flattened variables are numbered consecutively (e.g., AGE1, AGE2, AGE3, AGE4, etc.).
- File names will use the following convention: ahs[4-digit year][n for national or m for metro]. For example, “ahs1974m” or “ahs1973n”.

CSV Structure

All PUFs in 1973 - 1996 are available in both SAS and Comma Separated Value (CSV) formats. The CSV format allows data users to import the data when using statistical packages that are not compatible with SAS file formats.

The October 2022 revisions to the 1973 - 1996 files introduced a standardized CSV format for each PUF. This fixed inconsistencies in CSV formatting with the original files and created CSVs for years where they did not originally exist. In addition, the CSV files are now consistent with the revised PUFs. The following standards were applied to the CSV files:

- Variable names on the CSV datasets match their names on the SAS version exactly.
- Variable values on the CSV datasets match the SAS version exactly.
- Character variables in SAS have single quotes (‘ ’) around their values on the CSV file. Numeric values will not.

1973 - 1996 Interactive Codebook

A codebook is a necessary tool for working with the PUFs. Prior to 2015, HUD and the Census Bureau maintained the AHS Codebook as a living document that spanned surveys conducted from 1997 through 2013. HUD and the Census Bureau made updates to the AHS Codebook each time a new survey was published. Other, static AHS codebooks covered surveys prior to 1997.

Starting with the 2015 PUF, HUD and the Census Bureau implemented an entirely new approach to maintaining the AHS Codebook called the AHS Codebook Interactive Tool. This tool is a web-based utility that includes information previously contained in the AHS Codebook but with enhanced functionality. This approach to presenting codebook-style information enables HUD and the Census Bureau to easily update the AHS Codebook if corrections or additions are required.

The AHS Codebook Interactive Tool now includes both AHS-N and AHS-MS PUFs from 1973 to present, which covers all publicly available AHS data. The data include variables on the AHS PUFs from 1973 - 2013 and internal use file (IUF) variables for 2015 to present. The Codebook was also updated to be consistent with the October 2022 revisions to the 1973 - 1996 files. The current Interactive Codebook is located at <https://www.census.gov/data-tools/demo/codebook/ahs/ahsdict.html>.

Several features are built into the AHS Codebook tool to meet the needs of AHS PUF users. The following table describes various features.

Feature	Description
Filter by survey year	This feature enables users to filter the AHS Codebook by survey year. Users wishing to download an entire codebook for a specific year could use this feature.

Feature	Description
Filter by topic or subtopic	This feature enables users to filter the AHS Codebook by a particular topic and, if desired, a subtopic. This feature is helpful for viewing all the variables in a particular topic.
Download button	This feature enables users to download a selected portion of the AHS Codebook. Format options include PDF and CSV.
Mini-codebook download	This feature enables users to download a PDF of all the variables in any specified survey year. The file includes the following fields: topic, subtopic, variable name, and description. The purpose of the file is to enable users to quickly review all the variables in the selected PUF and IUF.
Historic view	When a variable's metadata has no changes, multiple years are grouped together into one entry. When a variable's metadata has changes for different years, the tool displays a separate entry. This feature allows users to easily identify when a variable's metadata has changed over time.

AHS = American Housing Survey. IUF = internal use file. PUF = public use file.

1973 - 1996 Internal Use Files

HUD and the Census Bureau try to include every possible AHS variable on the PUFs. However, our duty to protect respondents' confidentiality prevents public release of some variables. We maintain another file, the internal use file (IUF), that includes such variables.

If a 1973 - 1996 AHS PUF user cannot complete an analysis because one or more AHS variables are not available on the PUF, HUD and the Census Bureau may be able to help in two ways. First, HUD and Census Bureau staff can run special tabulations, with disclosure methods applied, using IUF-only variables. Second, HUD and Census Bureau staff can assist AHS users with the process of obtaining access to the AHS IUF. Data users can contact an AHS staff member with special table requests at ahsn@census.gov or at 301-763-3235.

AHS users should note this process requires the following steps:

- Submitting a detailed proposal.
- Submitting to a background check and obtaining a non-sensitive federal security clearance.
- Agreeing to become subject to legally binding disclosure restrictions.
- Willingness to travel to a Census Bureau Federal Statistical Research Data Center.¹
- Willingness to pay a fee to access the Research Data Center.

More details on the process are available at <https://www.census.gov/about/adrm/ced/apply-for-access.html>.

Unlike the 2015 and later IUFs, variables unique to the 1973 - 1996 IUFs are not all documented in a publicly available data dictionary. Limited information on historic IUF files is available at https://www.census.gov/content/dam/Census/programs-surveys/ahs/tech-documentation/American_Housing_Survey_AHS_Using_the_Internal_User_File_IUF.pdf. Data users can reach out to the AHS staff contact listed above to verify whether an IUF has the variables needed.

¹ For more information, see <https://www.census.gov/fsrdc>.

Noninterview Types

Though the AHS visits the same housing units each cycle, units may be removed or added to the sample for various reasons throughout the survey process. HUD and the Census Bureau created the National Sample Case History File to track all sampled housing units in the integrated National sample. AHS users who want to know why various sample cases were not surveyed or were nonresponses for 1985 - 2013 can review this file.²

The 1973 - 1996 AHS PUFs include occupied interviews, usual residence elsewhere interviews, and vacant interviews. The variable STATUS indicates which of these three types of interviews was conducted.

While many of the original PUFs from this time period included non-interview sample, the October 2022 revisions removed noninterviews to be consistent with formatting in 1997 and later files.³ As in other years, noninterviews are classified as one of three types indicated in the following table.

Noninterview Type	Reason for Noninterview	Eligibility for Future Surveys
Type A	Respondent refused the interview or could not be contacted.	Housing unit remains in the sample for future surveys.
Type B	Housing unit was currently not habitable but could be habitable in a future year.	Housing unit remains in the sample for future surveys.
Type C	Housing unit was demolished or destroyed.	Housing unit is permanently removed from the sample.

Variable Formatting

The 1973 - 1996 PUFs include more than 1,400 unique variables across all flat files, plus edit flags. There are two types of variables—numeric and categorical.

The October 2022 file revisions applied the following standardization rules to the variables on the PUF data files:

- All variable names were fully capitalized.
- Labels matching the e-codebook descriptions were applied to all files.
- Variables without usable data were removed from the PUFs. This includes variables that had only missing values, only reported one unique value, or were determined to be artifacts of data processing that were not useful for publication. Flattened variables were kept from the PUFs as long as one member of the set had valid data. The criteria for variables being dropped during revisions can be found in the Appendix of this document.
- Truncated values were corrected where appropriate.
- Response codes with duplicate meanings were collapsed.
- Duplicated records were removed.
- Variable types were updated to be consistent across survey years with a few exceptions where the structure of the variable changed from numeric continuous to categorical or vice versa. All data in

² For more information, see https://www.census.gov/programs-surveys/ahs/tech-documentation/help-guides/1985-2013/national-sample-case-history_2013.html.

³ The following variables were also deleted from the files when non-interviews were removed: BBLDG, BOARDU, BOARD, EXPOSE, DISAS, DFIRE, STATUA, and NOINT. These variables were removed because they were only applicable for non-interviewed sample.

the original 1973 - 1996 PUFs were numeric except CONTROL. In the updated PUFs, numerous variables were changed from numeric to character. A variable remained numeric if it had a natural representation as a number, such as a count, dollar, or percent value. The criteria for type changes during revisions can be found in the Appendix of this document.

- Padding zeros were added to the variables where appropriate to standardize variable lengths.
- Variable names were changed to improve consistency across survey years. A full list of variables with updated names can be found in the Appendix of this document.
- Duplicated control variable values were resolved by assigning new control values.
- Control variable values that had only 11-digits on the 1974 - 1978 Metro PUF were updated to add a trailing 0, which allows them to match to the 12-digit CONTROL variables used in other survey years.

Many of the 1973 - 1996 PUF variables represent the same or similar concepts as other years. However, many variable names are different as compared to 2015 and beyond. Because the 2015 AHS included an entirely new sample, HUD and the Census Bureau took the opportunity to change variable names to make them more user friendly and to bring them in line with other surveys. More details on this process are available in the 2015 Getting Started Guide located at https://www.census.gov/programs-surveys/ahs/tech-documentation/help-guides/2015-later/puf_start.html.

Not Applicable vs. Not Reported

The 1973 - 1996 PUFs differentiate between respondents who did not know, refused, or generally gave no response to the question. This differs from more recent PUFs which combine these into one “not reported” category. Most variables have a code for Not Reported (or Not Answered) which is the digit 8, repeated for as many digits as the variable has (e.g., “8”, “98”, “998”, “9998”) and a code for Not Applicable (or Not applicable/Not reported) which is the digit 9, repeated for as many digits as the variable has (e.g., “9”, “99”, “999”, “9999”).

In addition, some numeric variables will have the (.) system missing code, and some character variables will have a blank (‘ ’) missing value code that are not reported, not applicable, or not applicable / not reported. These are most often due to respondents not being asked the question during the interview because they were not eligible at the time or quit the interview early.

However, for some variables it could represent the absence of a positive response, such as a missing parent line number meaning that the person’s parent does not live in the household. In another example, a missing value in a mark all that apply variable (where an ‘X’ means they have the characteristic) means the respondent did not pick that response from the answer categories. In this case, it can usually be treated as a “No” response and is documented as such in the codebook.

Geography

The AHS is limited in the geographic detail it can reveal on the PUF. There are two reasons for this. First, the combination of the survey’s rich data and a localized geographic variable could compromise the confidentiality of the survey’s respondents. Second, the sample design does not allow for statistically sound estimates below the metro level.

The table below summarizes the geographic variables available on the 1973 - 1996 PUFs. For a more comprehensive discussion of AHS geography, including disclosure avoidance methodology and changes in definitions across survey years, see the 1985 - 2013 AHS geography guide at

<https://www.census.gov/programs-surveys/ahs/tech-documentation/help-guides/1985-2013/geography--public-use-file--1985-2013.html>.

Geography Variable	Years	Description
REGION	1973 - 1995 National	The four Census regions. ⁴
STATE	1974 - 1996 Metro	This variable identifies state codes for sample in metropolitan areas. It cannot be used to make state level estimates. Data users are advised to use caution with this variable, because it is masked for some sample to prevent disclosure.
COUNTY	1974 - 1996 Metro	This variable identifies county codes for sample in metropolitan areas. It cannot be used to make county level estimates. In some cases, the variable may refer to a combined set of counties. For more information on the county code, see the codebook. Data users are to use caution with this variable, because it is masked for some sample to prevent disclosure.
DEGREE	1985 - 1995 National	This variable categorizes units into one of six climate categories based on climate data from its county.
CMSA	1985 - 1995 National	The Consolidated Metropolitan Statistical Area (CMSA) is a grouping of two or more metropolitan statistical areas. Users are advised that these are not representative in AHS.
SMSA	1974 - 1996 National and Metro	This code is the primary identifier of a metropolitan area in both the national and metro files. It can be used to make estimates of the indicated area on metro PUFs, but users are to be cautious when using SMSA on the national PUFs. For national, SMSA is only representative for “big 6” sample in years where the oversample is on the file, and in those cases, it sometimes needs to use a collection of SMSA codes.
METRO	1973 - 1996 National and Metro	Identifies the primary and up to five secondary central cities.
PMSA	1984 - 1996 Metro	Based on the 1990 Primary Metropolitan Statistical Area code for the metropolitan areas surveyed in 1998.
ZONE	1984 - 1996 Metro	Area of greater than 100,000 population constructed by HUD and usually based on

⁴ A list of states that make up each Census Region and Division can be found at https://www2.census.gov/geo/pdfs/maps-data/maps/reference/us_regdiv.pdf.

Geography Variable	Years	Description
		central cities of the OMB MSA or PMSA at the time the sample was drawn. However, when possible, the ZONE codes were updated to reflect current central cities and their respective boundaries. Zone definition files are available for 2007 - 2013 on the AHS Metropolitan PUF web pages ⁵ .

Core and Supplemental Topics

The 1973 - 1996 AHS questionnaires are organized into topics, where a topic is a set of related questions about a particular subject matter. HUD and the Census Bureau group all the topics into two types—core and supplemental. Core topics are administered in every survey. In some cases, questions in the core topics may be edited or deleted, or HUD and the Census Bureau may add a question to a core topic. When a question is edited, removed, or added to a core topic, the change is usually intended to be permanent.

Supplemental topics may appear in one survey or a group of surveys and not appear in the next survey but reappear in a future survey. AHS users interested in reviewing the specific questions in the supplemental topics can review the “items booklet” for each survey year located in the Questionnaires section at <https://www.census.gov/programs-surveys/ahs/tech-documentation/def-errors-changes.html>.

In later years, HUD and Census reassessed supplemental topics, with some becoming part of a core topic and others becoming “rotating” topics. They also adopted a different strategy for including supplemental information in the survey by asking some supplemental topics of only half the sample to reduce respondent burden.

The following table lists the supplemental topics and the AHS Interactive Codebook topic (or category) in which they can be found.

Survey Year	Topic	Topic in the Codebook Database
1975N, 1975 - 1977M, 1980 - 1981N, 1982M, 1984 - 1985M	Commuting	Commuting
1985 - 1993N	Second Homes	Second Home
1975 - 1977N, 1980N, 1983N, 1976 - 1981M	Mobile Homes	Home Improvement/Structural
1978N, 1979 - 1982M	Disabilities	Demographics/Accessibility/Healthy Homes

N = National. M = Metro.

Weights

Historically, the PUFs have included multiple weighting options. The table below lists all weighting variables originally published as part of the 1973 - 1996 PUFs, along with a description summarizing how they should be used and whether they were altered in the PUF revisions. As discussed in the AHSN

⁵ Shapefiles can be found at <https://www.census.gov/programs-surveys/ahs/data.html> by selecting the year of interest from the tabs at the top (2007-2013), clicking the link to the AHS Public Use File (PUF), selecting Metro PUF, then scrolling to the Mapping Files section.

Sample Design section above, a special rural oversample took place from 1974 to 1983. This oversample should be removed when making national estimates because those cases are not included in weights.

Variable Name	PUF Availability	Description
WEIGHT ⁶	1973 - 1996 National and Metro	This is the primary weight used to replicate published estimates. Weight is zero on non-interviews, on the neighbor sample, on units accidentally interviewed which are not supposed to be in the sample, and on certain types of vacant and URE units which are not considered part of the housing stock (e.g., vacant and URE units in transient hotels, boats, caves). Before 1984, vacant and URE mobile homes also had zero weights. Weight changes each year to impute Type A non-interviews onto interviews and to match to the AHS estimates with data from other Census Bureau surveys. Beginning in 1981N, Weight is adjusted based on the 1980 Census. Prior to that, it was adjusted based on the 1970 Census.
PWT	1985 - 1995 National and 1983 - 1996 Metro	This is a basic or “pure” weight that reflects the inverse probability of selection without additional adjustments. Advanced users can use this information to create custom weights, including longitudinal weighting. To reduce confusion, this variable was removed from the PUFs during revisions, but remains on the Sample Case History files.
OLDWT	1981, 1991 - 1995 National and 1979 - 1983 Metro	This is a weight of each case in the sample based on the 1970 Census.
WWT1-8	1980 - 1981 National and 1982 Metro	This is a person level weight which is stored for each individual worker as WWT1-8.

⁶ In 1995, weights used to generate national estimates can also be used to generate applicable “big 6” supplemental metro estimates in years they are represented in the national file. However, users are advised that differences caused by disclosure avoidance methods that mask the SMSA value for some cases could contribute to larger differences than what they may see in a national estimate.

Regarding weights, 1973 - 1996 AHS PUF users should note three other pieces of information—

- For additional information about how the weights were constructed, users should consult the “Sample Design, Weighting, and Error Estimation” documents published with each year’s PUFs⁷.
- In general, nearly all STATUS 1-3 housing units have a positive weight on the PUF. However, a small number of housing units are given a positive weight only when regularly occupied, but a weight of 0 if vacant. This occurs with certain types of housing units such as house boats or RVs that may not be considered housing units when vacant.

Edits and Imputation

In some instances, AHS respondents provide answers that HUD and the Census Bureau believe are in error due to conflicts with other responses. Additionally, AHS respondents may not provide a response to certain questions if they do not know or refuse to answer.

In the case of an inconsistent response, for example if a respondent reports the year in which they moved into the unit being before their year of birth, Census may edit their year moved to match their birth year. This is referred to as a consistency edit.

When a respondent does not give a response, some variables will have Not Reported codes (8, 98, etc.) to show missing data, while other variables have data allocated from a similar home, so the data do not appear to be missing. When data in one variable are allocated, a code is put into another variable, the allocation variable, to show that the data really were missing in the original interview. These variables always start with a “ALL”. For surveys from 1984 on, allocation variables always start with an “A” and generally follow the response codes below:

Blank: No edit or allocation

1: Answer was allocated

2: Answer was edited (consistency edits)

Topcoding and Disclosure

To ensure respondent confidentiality, the Census Bureau requires that disclosure avoidance procedures be applied to certain PUF variables. For some variables, this means they cannot be published on the PUFs in any form. This is mentioned in the Internal Use File and Geography sections of this document.

A handful of other disclosure avoidance procedures are applied to 1973 - 1996 PUF variables. These procedures include collapsing into larger categories, perturbation, rounding, and topcoding. To maintain confidentiality, unusually high (and, in a few cases, low) values in the AHS dataset are replaced with maximum (or minimum) values, called top (or bottom) codes. For some variables in the 1973 - 1996 PUFs, the original variable is retained in the IUF.

For additional information about disclosure avoidance methods, users should consult the 1985-2013 disclosure avoidance documentation available at https://www.census.gov/content/dam/Census/programs-surveys/ahs/tech-documentation/1985-2013/Disclosure_Avoidance_1985_to_2013.pdf.

⁷ Sample Design, Weighting and Error Estimation documents can be found in the Accuracy of the Data section at: <https://www.census.gov/programs-surveys/ahs/tech-documentation/def-errors-changes.html>.

Appendix

Variable Drop Criteria

The following criteria were used to determine when variables were dropped:

- All observations for the variable have the same value. For example, if every observation has a '1', the variable will be removed.
- All observations are missing values.
- Variables that had meaningful data only for non-interview cases were dropped.
- Variables had data quality issues based on Census research.
- Census was unable to determine the definition / description for a variable.
- Census was unable to determine what the coding for a variable means, e.g. 1 - Yes, 2 - No.

There are some exceptions to these criteria. When a variable is a part of a set, AGE1 - AGE15, for example, all set members were kept on the files regardless of whether the individual variable members of the set have informative data. In the age example, if AGE1 contains informative data, but AGE9 does not, we still kept AGE9 on the data files. Allocation variables were also kept on the files to show if the variables were allocated or edited.

Variable Type Change Criteria

The following criteria were used to determine when variables were dropped:

- Variables that are categorical in nature (e.g., 1 - Yes, 2 - No) have been converted to SAS character type.
- Variables that are natural representations of numbers (e.g., numeric continuous dollar amounts) are SAS numeric type.
- Variables have consistent types across time.

There are a very small number of variables that have exceptions to these criteria. This happens because the way in which a variable is asked or presented in the data changed from numeric continuous to categorical or vice versa. A good example of this is the variable AMTX - Annual Real Estate Tax Payment. In the very early years on the AHS this variable was reported on the PUF as a numeric continuous variable. However, in 1984 metro the variable AMTX is represented on the PUF as categories in which the code represents a range of dollar amount values. 1 represents \$1 to \$50, 2 represents \$51 to \$100, etc. In variables like AMTX we allow the variable type to fluctuate across time.

Variable Renames During Revisions

Old Name	New Name	Survey Year
AGE	AGE1	1974N, 1975M, 1976M, 1976N, 1977M, 1978M, 1979M, 1980M, 1981M, 1982M, 1983M, 1984M, 1985M, 1985N, 1986M, 1987M, 1987N, 1988M, 1989M, 1989N, 1990M, 1991M, 1991N, 1992M, 1993M, 1993N, 1994M, 1995M, 1995N, 1996M
BADSTE	BADSTEP	1973N, 1974M, 1974N, 1975M, 1976M, 1976N, 1977M, 1978M, 1978N, 1979M, 1979N, 1980M, 1981M, 1981N, 1982M, 1983M, 1983N, 1984M, 1985M, 1986M, 1987M, 1988M, 1989M, 1990M, 1991M, 1992M, 1993M, 1994M, 1995M, 1996M
BDSTEP	BADSTEP	1980N

CPART	CPARK	1983M
CROPS	CROP5	1980N
CSTAD	CSTADD	1991N
FAM	FAM1	1984M, 1985M, 1986M, 1987M, 1988M, 1989M, 1990M, 1991M, 1992M, 1993M, 1994M, 1995M, 1996M
FIRSHO	FRSTHO	1980N
GRADE1	GRAD1	1984M, 1985M, 1985N, 1986M, 1987M, 1987N, 1988M, 1989M, 1989N, 1990M, 1991M, 1991N, 1992M, 1993M, 1993N, 1994M, 1995M, 1995N, 1996M
IFCOL	IFCOLD	1995N
ISTATU	ISTATUS	1973N, 1974M, 1974N, 1976M, 1976N, 1977M, 1978M, 1978N, 1979M, 1979N, 1980M, 1980N, 1981M, 1981N, 1982M, 1983M, 1983N, 1984M, 1985M, 1986M, 1987M, 1987N, 1988M, 1989M, 1990M, 1991M, 1992M, 1993M, 1994M, 1995M, 1996M
KITCHE	KITCHEN	1973N, 1974M, 1974N, 1975M, 1976M, 1976N, 1977M, 1978M, 1978N, 1979M, 1979N, 1980M, 1981M, 1981N, 1982M, 1983M, 1983N, 1984M, 1985M, 1986M, 1987M, 1988M, 1989M, 1990M, 1991M, 1992M, 1993M, 1994M, 1995M, 1996M
KITCHN	KITCHEN	1980N
MAR	MAR1	1974N, 1975M, 1975N, 1976M, 1976N, 1977M, 1977N, 1978M, 1978N, 1979M, 1979N, 1980M, 1980N, 1981M, 1981N, 1982M, 1983M, 1983N, 1984M, 1985M, 1985N, 1986M, 1987M, 1987N, 1988M, 1989M, 1989N, 1990M, 1991M, 1991N, 1992M, 1993M, 1993N, 1994M, 1995M, 1995N, 1996M
MGRHER	MGRHERE	1973N, 1974M, 1974N, 1975M, 1976M, 1976N, 1977M, 1978M, 1983N
MORTNS	MORTINS	1980N
MOVED	MOVE1	1984M, 1985M, 1985N, 1986M, 1987M, 1987N, 1988M, 1989M, 1989N, 1990M, 1991M, 1991N, 1992M, 1993M, 1993N, 1994M, 1995M, 1995N, 1996M
NEWAD	NEWADD	1991N
NEWHUT	NEWSHUT	1980N
NEWICS	NEWICST	1978N, 1979N, 1984M, 1985M, 1986M, 1987M, 1988M, 1989M, 1990M, 1991M, 1992M, 1993M, 1994M
NEWICT	NEWICST	1980N
NEWMOR	NEWMOR	1985N, 1991N, 1993N, 1995N
NUM3BE	NUM3BED	1973N, 1974M, 1974N, 1975M, 1976M, 1976N, 1977M, 1978M
NUMBLO	NUMBLOW	1973N, 1974M, 1974N, 1975M, 1976M, 1976N, 1977M, 1978M, 1978N, 1979M, 1979N, 1980M, 1980N, 1981M, 1981N, 1982M, 1983M, 1983N, 1984M, 1985M, 1986M, 1987M, 1988M, 1989M, 1990M, 1991M, 1992M, 1993M, 1994M, 1995M, 1996M
NUMCLS	NUMCLSD	1973N, 1974M, 1974N, 1975M, 1976M, 1976N, 1977M, 1978M, 1978N, 1979M, 1980M, 1981M, 1981N, 1982M, 1983M, 1983N
NUMCOL	NUMCOLD	1973N, 1974M, 1974N, 1975M, 1976M, 1976N, 1977M, 1978M, 1978N, 1979M, 1980M, 1981M, 1981N, 1982M, 1983M, 1983N, 1984M, 1985M, 1986M, 1987M, 1987N, 1988M, 1989M, 1989N, 1990M, 1991M, 1992M, 1993M, 1994M, 1995M, 1995N, 1996M
OTHCL	OTHCLD	1995N
OTHNH	OTHNHD	1995N
OWNCAR	OWNCARD	1980N, 1983N

OWNHER	OWNHERE	1973N, 1974M, 1974N, 1975M, 1976M, 1976N, 1977M, 1978M, 1983N, 1984M, 1985M, 1986M, 1987M, 1988M, 1989M, 1990M, 1991M, 1992M, 1993M, 1994M, 1995M, 1996M
PAR	PAR1	1984M, 1985M, 1985N, 1986M, 1987M, 1987N, 1988M, 1989M, 1989N, 1990M, 1991M, 1991N, 1992M, 1993M, 1993N, 1994M, 1995M, 1995N, 1996M
PERSIN	PERSINT	1981N, 1983N, 1984M, 1985M, 1985N, 1986M, 1987M, 1987N, 1988M, 1989M, 1989N, 1990M, 1991M, 1991N, 1992M, 1993M, 1993N, 1994M, 1995M, 1995N, 1996M
PLASTE	PLASTER	1974N, 1976M, 1976N, 1977M, 1978M, 1978N, 1979M, 1979N, 1980M, 1980N, 1981M, 1981N, 1982M, 1983M, 1983N
PLIN10	PLINE10	1975M, 1976M, 1978N, 1979M, 1979N, 1980M, 1980N, 1981M, 1981N, 1982M, 1983M, 1983N
PLIN11	PLINE11	1974N, 1975M, 1976M, 1976N, 1977M, 1978M, 1978N, 1979M, 1979N, 1980M, 1980N, 1981M, 1981N, 1982M, 1983M, 1983N
PLIN12	PLINE12	1974N, 1975M, 1976M, 1976N, 1977M, 1978M, 1978N, 1979M, 1979N, 1980M, 1980N, 1981M, 1981N, 1982M, 1983M, 1983N
PLIN13	PLINE13	1974N, 1975M, 1976M, 1976N, 1977M, 1978M, 1978N, 1979M, 1979N, 1980M, 1980N, 1981M, 1981N, 1982M, 1983M, 1983N
PLIN14	PLINE14	1974N, 1975M, 1976M, 1976N, 1977M, 1978M, 1978N, 1979M, 1979N, 1980M, 1980N, 1981M, 1981N, 1982M, 1983M, 1983N
PLIN15	PLINE15	1974N, 1975M, 1976M, 1976N, 1977M, 1978M, 1978N, 1979M, 1979N, 1980M, 1980N, 1981M, 1981N, 1982M, 1983M, 1983N
PLINE	PLINE1	1974N, 1975M, 1975N, 1976M, 1976N, 1977M, 1977N, 1978M, 1978N, 1979M, 1979N, 1980M, 1980N, 1981M, 1981N, 1982M, 1983M, 1983N, 1984M, 1985M, 1985N, 1986M, 1987M, 1987N, 1988M, 1989M, 1989N, 1990M, 1991M, 1991N, 1992M, 1993M, 1993N, 1994M, 1995M, 1995N, 1996M
PLINE1	PLINE10	1974N, 1976N, 1977M, 1978M
PLINEA	PLINE10	1984M, 1985M, 1985N, 1986M, 1987M, 1987N, 1988M, 1989M, 1989N, 1990M, 1991M, 1991N, 1992M, 1993M, 1993N, 1994M, 1995M, 1995N, 1996M
PLINEB	PLINE11	1984M, 1985M, 1985N, 1986M, 1987M, 1987N, 1988M, 1989M, 1989N, 1990M, 1991M, 1991N, 1992M, 1993M, 1993N, 1994M, 1995M, 1995N, 1996M
PLINEC	PLINE12	1984M, 1985M, 1985N, 1986M, 1987M, 1987N, 1988M, 1989M, 1989N, 1990M, 1991M, 1991N, 1992M, 1993M, 1993N, 1994M, 1995M, 1995N, 1996M
PLINED	PLINE13	1984M, 1985M, 1985N, 1986M, 1987M, 1987N, 1988M, 1989M, 1989N, 1990M, 1991M, 1991N, 1992M, 1993M, 1993N, 1994M, 1995M, 1995N, 1996M
PLINEE	PLINE14	1984M, 1985M, 1985N, 1986M, 1987M, 1987N, 1988M, 1989M, 1989N, 1990M, 1991M, 1991N, 1992M, 1993M, 1993N, 1994M, 1995M, 1995N, 1996M
PLINEF	PLINE15	1984M, 1985M, 1985N, 1986M, 1987M, 1987N, 1988M, 1989M, 1989N, 1990M, 1991M, 1991N, 1992M, 1993M, 1993N, 1994M, 1995M, 1995N, 1996M
RACE	RACE1	1974N, 1975M, 1975N, 1976M, 1976N, 1977M, 1977N, 1978M, 1978N, 1979M, 1979N, 1980M, 1980N, 1981M, 1981N, 1982M, 1983M, 1983N, 1984M, 1985M, 1985N, 1986M, 1987M, 1987N, 1988M, 1989M, 1989N, 1990M, 1991M, 1991N, 1992M, 1993M, 1993N, 1994M, 1995M, 1995N, 1996M

REL	REL1	1974N, 1975M, 1975N, 1976M, 1976N, 1977M, 1977N, 1978M, 1978N, 1979M, 1979N, 1980M, 1980N, 1981M, 1981N, 1982M, 1983M, 1983N, 1984M, 1985M, 1985N, 1986M, 1987M, 1987N, 1988M, 1989M, 1989N, 1990M, 1991M, 1991N, 1992M, 1993M, 1993N, 1994M, 1995M, 1995N, 1996M
S15OM1	S15OM1	1985N, 1987N, 1989N
S15OM2	S15OM2	1985N, 1987N, 1989N
S15OM3	S15OM3	1985N, 1987N, 1989N
S15OM4	S15OM4	1985N, 1987N, 1989N
S15OM5	S15OM5	1985N, 1987N, 1989N
S15OM6	S15OM6	1985N, 1987N, 1989N
SEX	SEX1	1974N, 1975M, 1975N, 1976M, 1976N, 1977M, 1977N, 1978M, 1978N, 1979M, 1979N, 1980M, 1980N, 1981M, 1981N, 1982M, 1983M, 1983N, 1984M, 1985M, 1985N, 1986M, 1987M, 1987N, 1988M, 1989M, 1989N, 1990M, 1991M, 1991N, 1992M, 1993M, 1993N, 1994M, 1995M, 1995N, 1996M
SPAN	SPAN1	1984M, 1985M, 1985N, 1986M, 1987M, 1987N, 1988M, 1989M, 1989N, 1990M, 1991M, 1991N, 1992M, 1993M, 1993N, 1994M, 1995M, 1995N, 1996M
SPOS	SPOS1	1984M, 1985M, 1985N, 1986M, 1987M, 1987N, 1988M, 1989M, 1989N, 1990M, 1991M, 1991N, 1992M, 1993M, 1993N, 1994M, 1995M, 1995N, 1996M
STPLAC	STPLACE	1979N, 1980N
SUBF	SUBF1	1984M, 1985M, 1985N, 1986M, 1987M, 1987N, 1988M, 1989M, 1989N, 1990M, 1991M, 1991N, 1992M, 1993M, 1993N, 1994M, 1995M, 1995N, 1996M
SWOO	SWOOD	1995N
TAXPMT	TXPMT2	1985N, 1991N, 1993N, 1995N
TEN	TEN1	1984M, 1985M, 1985N, 1986M, 1987M, 1987N, 1988M, 1989M, 1989N, 1990M, 1991M, 1991N, 1992M, 1993M, 1993N, 1994M, 1995M, 1995N, 1996M
THOSP	THOSP1	1991N, 1993N, 1995N
VACAC	VACVAC	1987M, 1988M, 1989M, 1990M, 1991M, 1992M, 1993M, 1994M, 1995M, 1996M
VACANC	VACANCY	1973N, 1974M, 1975M, 1976M, 1977M, 1978M, 1978N, 1979M, 1979N, 1980M, 1981M, 1981N, 1982M, 1983M, 1983N, 1984M, 1985M, 1986M, 1987M, 1988M, 1989M, 1990M, 1991M, 1992M, 1993M, 1994M, 1995M, 1996M
VACANT	VACANCY	1974N, 1976N, 1980N
WHONE	WHAVL	1985N, 1987N, 1995N
WHYCL	WHYCLD	1995N
WHYMOV	WHYMOVE	1973N, 1974M, 1974N, 1975M, 1976M, 1976N, 1977M, 1978M, 1978N, 1979M, 1979N, 1980M, 1980N, 1981M, 1981N, 1982M, 1983M, 1983N, 1984M, 1985M, 1986M, 1987M, 1987N, 1988M, 1989M, 1989N, 1990M, 1991M, 1992M, 1993M, 1994M, 1995M, 1996M
WHYR	WHYRD	1995N
XCOMDC	XCOMDOC	1980N
XCOMDO	XCOMDOC	1973N, 1974M, 1974N, 1975M, 1976M, 1976N, 1977M, 1978M, 1978N, 1979M, 1979N, 1980M, 1981M, 1981N, 1982M, 1983M, 1983N
XNUITS	XNUNITS	1980N

XNUNIT	XNUNITS	1973N, 1974M, 1974N, 1975M, 1976M, 1976N, 1977M, 1978M, 1978N, 1979M, 1979N, 1980M, 1981M, 1981N, 1982M, 1983M, 1983N
XTENUR	XTENURE	1973N, 1974M, 1974N, 1975M, 1976M, 1976N, 1977M, 1978M, 1978N, 1979M, 1979N, 1980M, 1980N, 1981M, 1981N, 1982M, 1983M, 1983N
ZPRICM	ZPRICEM	1983N
ZSRVM2	ZSERVM2	1974N, 1976N
ZXCROW	ZXCROWD	1973N, 1974M, 1974N, 1975M, 1976M, 1976N, 1977M, 1978M, 1978N, 1979M, 1979N, 1980M, 1980N, 1981M, 1981N, 1982M, 1983M, 1983N

Metro Panel Years

Metropolitan Area	SMSA	First Panel Years	Second Panel Years	Third Panel Years
Albany, NY	0160	1974, 1977, 1980	None	None
Allentown, PA	0240	1976, 1980	None	None
Anaheim-Santa Ana, CA	0360	1974, 1977, 1981, 1986, 1990, 1994, 2002, 2011	None	None
Atlanta, GA	0520	1975, 1978, 1982, 1987, 1991	1996, 2004, 2011	None
Austin, TX	0640	2013	None	None
Baltimore, MD	0720	1976, 1979, 1983, 1987, 1991	1998, 2007, 2013	None
Birmingham, AL	1000	1976, 1980, 1984, 1988, 1992	1998, 2011	None
Boston, MA	1120	1974, 1977, 1981, 1985, 1989, 1993	1998, 2007, 2013	None
Buffalo, NY	1280	1976, 1979, 1984, 1988, 1994, 2002, 2011	None	None
Charlotte, NC	1520	1995, 2002, 2011	None	None
Chicago, IL	1600	1975, 1979, 1983, 1987, 1991	1995, 1999, 2003 (national)	2009, 2013 (national)
Cincinnati, OH	1640	1975, 1978, 1982, 1986, 1990	1998, 2011	None
Cleveland, OH	1680	1976, 1979, 1984, 1988, 1992	1996, 2004, 2011	None
Colorado Springs, CO	1720	1975, 1978	None	None
Columbus, OH	1840	1975, 1978, 1982, 1987, 1991	1995, 2002, 2011	None
Dallas, TX	1920	1974, 1977, 1981, 1985, 1989, 1994, 2002, 2011	None	None
Denver, CO	2080	1976, 1979, 1983, 1986, 1990	1995, 2004, 2011	None
Detroit, MI	2160	1974, 1977, 1981, 1985, 1989, 1993	1995, 1999, 2003 (national)	2009, 2013 (national)

Fort Worth, TX	2800	1974, 1977, 1981, 1985, 1989, 1994, 2002, 2011	None	None
Grand Rapids, MI	3000	1976, 1980	None	None
Hartford, CT	3280	1975, 1979, 1983, 1987, 1991	1996, 2004, 2013	None
Honolulu, HI	3320	1976, 1979, 1983	None	None
Houston, TX	3360	1976, 1979, 1983	1987, 1991	1998, 2007, 2013
Indianapolis, IN	3480	1976, 1980, 1984, 1988, 1992	1996, 2004, 2011	None
Jacksonville, FL	3600	2013	None	None
Kansas City, MO	3760	1975, 1978, 1982, 1986, 1990	1995, 2002, 2011	None
Las Vegas, NV	4120	1976, 1979	2013	None
Louisville, KY	4520	1976, 1980, 1983	2013	None
Los Angeles, CA	4480	1974, 1977, 1980, 1985, 1989	1995, 1999, 2003 (national)	2011 (national)
Madison, WI	4720	1975, 1977, 1981	None	None
Memphis, TN	4920	1974, 1977, 1980, 1984, 1988, 1992	1996, 2004, 2011	None
Miami-Ft Lauderdale, FL	5000	1975, 1979, 1983, 1986, 1990	1995, 2002, 2007, 2013	None
Milwaukee, WI	5080	1975, 1979, 1984, 1988, 1994, 2002, 2011	None	None
Minneapolis-St Paul, MN	5120	1974, 1977, 1981, 1985, 1989, 1993	1998, 2007, 2013	None
Nashville, TN	5360	2013	None	None
New Orleans, LA	5560	1975, 1978, 1982, 1986, 1990	1995, 2004, 2009, 2011	None
New York, NY	5600	1976, 1980, 1983, 1987, 1991	1995, 1999, 2003 (national)	2009, 2013 (national)
Newark, NJ	5640	1974, 1977, 1981	None	None
Northern New Jersey, NJ		1987, 1991	1995, 1999, 2003 (national)	2009, 2013 (national)
Patterson, NJ	6040	1975, 1978, 1982	None	None
Norfolk, VA	5680	1975, 1978, 1984, 1988, 1992	1998	None
	5720		2011	None
Oakland, CA	5775	1998, 2011	None	None
Oklahoma City, OK	5880	1976, 1980, 1984, 1988, 1992	1996, 2004, 2013	None
Omaha, NE	5920	1976, 1979	None	None
Orlando, FL	5960	1974, 1977, 1981	2013	None
Philadelphia, PA	6160	1975, 1978, 1982, 1985, 1989	1995, 1999, 2003 (national)	2009, 2013 (national)
Phoenix, AZ	6200	1974, 1977, 1981, 1985, 1989, 1994, 2002, 2011	None	None

Pittsburgh, PA	6280	1974, 1977, 1981, 1986, 1990	1995, 2004, 2011	None
Portland, OR	6440	1975, 1979, 1983, 1986, 1990	1995, 2002, 2011	None
Providence, RI	6480	1976, 1980, 1984, 1988, 1992	1998, 2011	None
Raleigh, NC	6640	1976, 1979	None	None
Richmond, VA	6760	2013	None	None
Riverside, CA	6780	2011	None	None
	7280	1975, 1978, 1982, 1986, 1990, 1994, 2002	None	None
Rochester, NY	6840	1975, 1978, 1982, 1986, 1990	1998, 2013	None
Sacramento, CA	6920	1976, 1980, 1983	1996, 2004, 2011	None
Saginaw, MI	6960	1974, 1977, 1980	None	None
Saint Louis, MO	7040	1976, 1980, 1983, 1987, 1991	1996, 2004, 2011	None
Salt Lake City, UT	7160	1974, 1977, 1980, 1984, 1988, 1992	1998	None
San Antonio, TX	7240	1975, 1978, 1982, 1986, 1990	1995, 2004, 2013	None
San Diego, CA	7320	1975, 1978, 1982, 1987, 1991, 1994, 2002, 2011	None	None
San Francisco, CA	7360	1975, 1978, 1982, 1985, 1989, 1993	1998, 2011	None
San Jose, CA	7400	1984, 1988, 1993	1998, 2011	None
Seattle-Everett, WA	7600	1976, 1979, 1983, 1987, 1991	1996, 2004, 2009, 2013	None
Seattle-Tacoma, WA	8200	1974, 1977, 1981	None	None
Spokane, WA	7840	1974, 1977, 1981	None	None
Springfield, MA	8000	1975, 1978	None	None
Tampa, FL	8280	1985, 1989, 1993	1998, 2007, 2013	None
Tucson, AZ	8520	2013	None	None
Washington, DC	8840	1974, 1977, 1981, 1985, 1989, 1993	1998, 2007, 2013	None
Wichita, KS	9040	1974, 1977, 1981	None	None

Example SAS Code

This section provides some examples of using AHS data. Each example begins with a description of the general procedure, followed by the SAS code for users of that software.

Example 1: Generating a Table

To build a simple crosstabulation of AHS data, you need to specify the variables to tabulate and the weighting variable. ISTATUS is the interview status of a sample case. NUNIT2 is the type of structure. They take these values (as you can find in the Interactive Electronic Codebook):

ISTATUS

Interview status

- 1: Occupied interview
- 2: URE (Usual Residence Elsewhere) interview
- 3: Vacant interview
- 4: Noninterview

NUNIT2

Are these living quarters in a...

- 1: One-unit building, detached from any other building
- 2: One-unit building, attached to one or more buildings
- 3: Building with two or more apartments
- 4: Mobile home, 1 unit
- 5: Mobile home, 2 or more units
- . : Not applicable / Not reported

An example of SAS code that can be used to generate the table is below:

```
PROC FREQ DATA=AHS95N.AHS1995N; /* This line specifies which dataset to use as input. */  
TABLE ISTATUS*NUNIT2; /* This section specifies which variables to use in the tabulation. */  
WEIGHT WEIGHT; /* This section specifies the weighting variable. */  
RUN;
```

Running the code above gives the following output.

The FREQ Procedure

Table of ISTATUS by NUNIT2

ISTATUS(Interview status)		NUNIT2(Structure type)					
Frequency							
Percent							
Row Pct							
Col Pct	1	2	3	4	5	Total	
1	6.083E7	5545054	2.516E7	6150607	13449.1	9.769E7	
	55.57	5.07	22.98	5.62	0.01	89.25	
	62.26	5.68	25.75	6.30	0.01		
	91.93	89.25	85.49	80.59	86.36		
2	441692	76712.9	615343	120427	0	1254175	
	0.40	0.07	0.56	0.11	0.00	1.15	
	35.22	6.12	49.06	9.60	0.00		
	0.67	1.23	2.09	1.58	0.00		
3	4900895	591203	3654258	1360486	2124.9	1.051E7	
	4.48	0.54	3.34	1.24	0.00	9.60	
	46.64	5.63	34.77	12.95	0.02		
	7.41	9.52	12.42	17.83	13.64		
Total	6.617E7	6212970	2.943E7	7631520	15574	1.095E8	
	60.45	5.68	26.89	6.97	0.01	100.00	

Example 2: Linking Files

Another common procedure when using the AHS PUFs is to link variables referring to the same housing unit that exist on two different files. In this example, we create a file called MATCH that has both the 1993 and 1995 interview status variables for any housing unit that exists on either the 1993 or 1995 national flat file.

The code below uses an SQL statement to merge the files, then creates a list tabulation of the 1993 and 1995 ISTATUS variables.

PROC SQL;

```
CREATE TABLE MATCH AS /* Specify the name of the output dataset */
SELECT AHS93.ISTATUS AS ISTATUS1993,
AHS95.ISTATUS AS ISTATUS1995 /* Rename the ISTATUS variable on the 1993 and 1995 files. */
FROM AHS93N.AHS1993N AHS93
FULL JOIN AHS95N.AHS1995N AHS95 /* Specify files to link. */
ON AHS93.CONTROL = AHS95.CONTROL; /* Files are linked by CONTROL. */
```

QUIT;

PROC FREQ DATA=MATCH;

```
TABLES ISTATUS1993*ISTATUS1995 / LIST MISSING; /* Create list table of the variables. */
```

RUN;

After running the code above to link the 1993 and 1995 national flat files by CONTROL, the FREQ procedure generates the following output.

The FREQ Procedure

ISTATUS1993	ISTATUS1995	Frequency	Percent	Cumulative Frequency	Cumulative Percent
	1	7557	11.71	7557	11.71
	2	79	0.12	7636	11.83
	3	939	1.45	8575	13.28
1		11227	17.39	19802	30.67
1	1	35826	55.50	55628	86.17
1	2	181	0.28	55809	86.45
1	3	2092	3.24	57901	89.69
2		122	0.19	58023	89.88
2	1	166	0.26	58189	90.14
2	2	172	0.27	58361	90.40
2	3	181	0.28	58542	90.68
3		1473	2.28	60015	92.97
3	1	2126	3.29	62141	96.26
3	2	188	0.29	62329	96.55
3	3	2227	3.45	64556	100.00

Note that the above output is unweighted and includes missing values. ISTATUS does not have any missing values in any survey year, so we know that the missing values in the above frequency mean that the sample unit does not exist in that year. For example, the first line can be interpreted as there being 7,557 sample units that exist on the national 1995 flat file with a ISTATUS of 1 which don't have a record on the 1993 national flat file.

Files can be linked across years when they are longitudinal, or between datasets in the same year.