Federal statistical datasets allowing disaggregation of Asian, Native Hawaiian, and Other Pacific Islander communities

Abstract
This document provides a searchable catalog of federal datasets that include disaggregated data about Asian, Native Hawaiian, and Pacific Islander populations. These datasets cover a variety of topics and are sortable by the disaggregated subgroups included, both of which can be searched below. For each dataset, the search will return information about the contents of relevant datasets and how to gain access to the data.
Contents

I. Annual Business Survey (ABS) .......................................................................................................................... 2
II. American Community Survey (ACS) .............................................................................................................. 4
III. ACS/HUD Administrative Data Linked Files ............................................................................................... 7
IV. American Housing Survey (AHS) .................................................................................................................. 11
V. American Time Use Survey (ATUS) ............................................................................................................. 14
VI. Baccalaureate and Beyond Longitudinal Study (B&B) .............................................................................. 16
VII. 2020 Census ............................................................................................................................................................ 19
VIII. Current Population Survey (CPS) ................................................................................................................. 21
IX. Early Childhood Longitudinal Study, Birth Cohort (ECLS-B) ..................................................................... 24
X. Early Childhood Longitudinal Study, Kindergarten Class of 1998-99 (ECLS-K) ........................................ 26
XI. Early Childhood Longitudinal Study, Kindergarten Class of 2010-11 (ECLS-K:2011) ......................... 29
XII. Early Childhood Longitudinal Study, Kindergarten Class of 2010-11 (ECLS-K:2011) ....................... 32
XIII. Education Longitudinal Study of 2002 (ELS:2002) .................................................................................. 35
XIV. High School Longitudinal Study of 2009 (HSLS:09) ............................................................................... 38
XV. Household Pulse Survey (HPS) .................................................................................................................... 41
XVI. National Postsecondary Student Aid Study (NPSAS) .............................................................................. 43
XVII. National Health and Nutrition Examination Survey (NHANES) .......................................................... 47
XVIII. National Health Interview Survey (NHIS) ............................................................................................... 50
XIX. National Survey of Family Growth ............................................................................................................. 52
XX. National Vital Statistics System (NVSS)–Birth Data .................................................................................. 55
XXI. National Vital Statistics System (NVSS)-Fetal Death Data ................................................................. 57
XXII. National Vital Statistics System (NVSS)-Linked Birth/Infant Death Data ............................................. 60
XXIII. National Vital Statistics System (NVSS)-Mortality Data ....................................................................... 62
XXIV. Survey of Income and Program Participation (SIPP) ............................................................................ 65
I. **Annual Business Survey (ABS)**

Selected economic and demographic characteristics for businesses and business owners by sex, ethnicity, race, and veteran status.

**Sponsoring agency:** Census Bureau and National Center for Science and Engineering Statistics (NCSES) within the National Science Foundation

**Data universe:**
Included are all nonfarm employer businesses filing the 941, 944, or 1120 tax forms.

**Disaggregated Asian and Native Hawaiian or Other Pacific Islander categories collected:**
Chinese, Filipino, Guamanian or Chamorro, Japanese, Korean, Native Hawaiian, Samoan, Vietnamese

**Are disaggregated Asian and Native Hawaiian or Other Pacific Islander data publicly available?** YES

**Measures available:**
Age, Business and Economy, Education, Geography, Migration and Nativity, Sex and Gender, Veterans

**Sample size and unit of analysis:**
850,000 every 5 years (starting with reference year 2017), 300,000 in between years Firm level.

**Types of geography available to researchers:** State, County, Place

**What data are available?**
PUBLIC-USE data? Yes
RESTRICTED-USE data? Yes

**What is the difference between PUBLIC-USE and RESTRICTED-USE data files:**
PUBLIC-USE is tabulated data with suppressions Restricted use is micro dataset.

**How to access the PUBLIC-USE data:**

1. Go to [census.gov](http://census.gov)
2. Select "Surveys/Programs"
3. Select "Annual Business Survey"
4. Select "Data"
5. Select "ABS Tables"
How to obtain permission to use the RESTRICTED-USE data:
Approved project via the Federal Research Data Center

Constraints surrounding access to RESTRICTED-USE data (e.g., citizenship requirements, cost, etc):
The Census Bureau considers proposals from qualified researchers in social science disciplines consistent with the subject matter of the surveys and censuses collected by the Census Bureau. Proposals may be submitted at any time to ResearchDataGov. All researcher access to restricted-use data occurs at secure Federal Statistical Research Data Centers (FSRDC). Proposed projects must: 1) Provide benefits to Census Bureau programs, 2) Demonstrate scientific merit, 3) Require non-public data, 4) Be feasible given the data, 5) Pose no risk of disclosure. https://www.census.gov/about/adrm/fsrdc.html

Agency policy re. sample sizes/publishable estimates/data suppression standards:
Official estimates put forth by the Census Bureau must adhere to Census’ disclosure-avoidance rules and methodology as approved by the Disclosure Review Board.

Information on when the data were collected/what time period could be examined using the data:
Currently under 5-year agreement to collect data for reference years 2017 – 2022.
2017 data collected July - December 2018, results published in May 2020;
2018 data collected July - December 2019, results published in Jan 2021;
2019 data collected July - December 2020, results published in Oct 2021

Where to find disseminated official estimates:
Published as 3 data series, Company Summary, Characteristics of Businesses, and Characteristics of Business Owners

1. Go to Census.gov
2. Select "surveys and programs"
3. Select "Annual Business Survey"
4. Select "Data"
5. Select "ABS Tables"

Type of data (survey, program, admin etc.): Survey

Do the data include measures of citizenship, or are respondents limited by citizenship (i.e., only U.S. citizens are included in the data)? Survey asks respondents if owners are citizens of the U.S.
Analytic Considerations:
Data analysis on race of owner based on historical information. Disaggregated race data for the AANHPI community are only available in the large sample year (850,000) every five years. In the small sample year (300,000), only broad race categories are produced.

II. American Community Survey (ACS)

The American Community Survey (ACS) is a nationwide survey designed to provide communities with reliable and timely social, economic, housing, and demographic data every year. A separate annual survey, called the Puerto Rico Community Survey (PRCS), collects similar data about the population and housing units in Puerto Rico. The Census Bureau uses data collected in the ACS and the PRCS to provide estimates on a broad range of population, housing unit, and household characteristics for states, counties, cities, school districts, congressional districts, census tracts, block groups, and many other geographic areas.

Sponsoring agency: U.S. Census Bureau

Data universe:
A nationally representative sample of housing unit (HU) addresses and residents of group quarters (GQ) facilities.

Disaggregated Asian and Native Hawaiian or Other Pacific Islander categories collected:
Afghan, Asian Indian, Bangladeshi, Bengali, Bhutanese, Bikinian, Bruneian, Burmese, Buryat, Cambodian, Carolinian, Central Asian, Cham, Chinese, Chuukese, Cook Islander, East Asian, Easter Islander, Ebeye, Ejit, Enewetak Islander, Fijian, Filipino, French Polynesian, Guamanian or Chamorro, Hakka, Han, Hawaiian, Hmong, Hong Kong, I-Kiribati, Indo, Indo Fijian, Indo-Chinese, Indonesian, Iwo Jima, Japanese, Kalmyk, Kazakh, Khmer, Kili, Korean, Kosraean, Kuki, Kwajalein Islander, Kyrgyz, Lahu, Laotian, Macanese, Malay, Malaysian, Maldivian, Maori, Marshallese, Melanesian, Micronesian, Mien, Mili, Mizo, Mongolian, Montagnard, Native Hawaiian, Nauruan, Nepalese, New Caledonian, Niuean, Ni-Vanuatu, Northern Mariana Islander, Okinawan, Pacific Islander, Pakistani, Palauan, Papua New Guinean, Part Hawaiian, Pashtun, Pohnpeian, Polynesian, Punjabi, Rotuman, Saipanese, Samoan, Sikh, Sindhi, Singaporean, Solomon Islander, South Asian, Southeast Asian, Sri Lankan, Tahitian, Tai Dam, Taiwanese, Tajik, Thai, Tibetan, Timorese, Tokelauan, Tongan, Turkmen, Tuvaluan, Ujelang, Urdu, Uzbek, Vietnamese, Wallisian and Futunan, Yapese

Are disaggregated Asian and Native Hawaiian or Other Pacific Islander data publicly available? YES

Measures available:
Age, Assistance Programs, Disability, Early Childhood, Education, Elementary and Secondary Education, Employment Characteristics, Families and Living Arrangements, Food and Nutrition,
Sample size and unit of analysis:
The ACS has an annual sample size of about 3.5 million addresses, with survey information collected nearly every day of the year. Data are pooled across a calendar year to produce estimates for that year. As a result, ACS estimates reflect data that have been collected over a period of time rather than for a single point in time.

Types of geography available to researchers: The RESTRICTED-USE data includes geographic information down to the block group for each record. Standard pre-tabulated ACS data products based on the full RESTRICTED-USE file are available for geographies and geographic components down to the block group unless otherwise suppressed due to population thresholds, quality, or disclosure restrictions.
The PUMS files include geographic information for Regions, Divisions, States, Puerto Rico, and Public Use Microdata Areas (PUMAs). PUMAs are non-overlapping, statistical geographic areas that partition each state or equivalent entity into geographic areas containing no fewer than 100,000 people each.

What data are available?
PUBLIC-USE data? Yes
RESTRICTED-USE data? Yes

What is the difference between PUBLIC-USE and RESTRICTED-USE data files:
PUMS files contain additional edits and modifications to protect respondent confidentiality, such as “top-coding” continuous variables with outlying values (e.g., income, transportation time) and the collapsing of categorical variables (e.g., ancestry, race) that have hundreds of detailed categories.

How to access the PUBLIC-USE data:
For standard pre-tabulated data:

1. Go to census.gov
2. Select "Surveys and Programs"
3. Select "American Community Survey"
4. Select "Data"
5. Click on the data.census.gov link in the center of the "Data" page to explore the data.

Public Use Microdata Sample (PUMS) files enable data users to create custom estimates and tables that are not available through pre-tabulated data products. To access the PUMS data:
0. Go to census.gov
1. Select "Surveys and Programs"
2. Select "American Community Survey"
3. Select "Microdata"
4. From the PUMS page, select whether to use the files within data.census.gov or download from the File Transfer Protocol (FTP) site to use in a database or statistical software package.

How to obtain permission to use the RESTRICTED-USE data:

The Census Bureau considers proposals from qualified researchers in social science disciplines consistent with the subject matter of the surveys and censuses collected by the Census Bureau. Proposals may be submitted at any time:

1. Go to census.gov
2. Select "Surveys and Programs"
3. Select "All surveys and programs"
4. Select "Center for Economic Studies"
5. then navigate to "Data" (on the left side of the CES page).
6. Select "RESTRICTED-USE Data" and then "Apply for access": https://www.census.gov/about/adrm/ced/apply-for-access.html

Constraints surrounding access to RESTRICTED-USE data (e.g., citizenship requirements, cost, etc):

All researcher access to RESTRICTED-USE data occurs at secure Federal Statistical Research Data Centers (FSRDC). Some FSRDCs may charge fees for access to their facilities.

Agency policy re. sample sizes/publishable estimates/data suppression standards:

Areas or groups of 65,000 or more are eligible for 1-, 5- year, and Supplemental Estimates. Areas or groups of 20,000 or more are eligible for 5-year and Supplemental Estimates. Areas or groups of 20,000 or fewer are eligible for 5-year estimates only. https://www2.census.gov/programs-surveys/acs/tech_docs/data_suppression/ACSO_Data_Suppression.pdf

Information on when the data were collected/what time period could be examined using the data:

The ACS data are collected continually and have been disseminated as annual estimates since 2006. Both 1-year and 5-year files are available.
Where to find disseminated official estimates:
Official ACS estimates can be found at data.census.gov:

1. Go to census.gov
2. Select "Surveys and Programs"
3. Select "American Community Survey"
4. Select "Data", and select the orange data tool link: data.census.gov

Type of data (survey, program, admin etc.): Survey data.

Do the data include measures of citizenship, or are respondents limited by citizenship (i.e., only U.S. citizens are included in the data)? Yes, ACS data include measures of nativity, and inclusion in the sample is not limited by citizenship status.

Analytic Considerations:
Detailed information for Asian and Native Hawaiian and Other Pacific Islander (NHPI) groups may be limited in the standard pre-tabulated ACS data products or PUMS files. The ACS Selected Population Tables, which is produced every five years, provides detailed tables iterated for a larger set of disaggregated Asian and NHPI detailed groups.

III. ACS/HUD Administrative Data Linked Files

Social and data scientists at HUD and the Census Bureau developed a procedure to link HUD rental assistance administrative records to the American Community Survey (ACS), thereby identifying ACS households as receiving HUD rental assistance. The ACS contains a wealth of household and demographic information that is not currently collected by HUD. Some examples of information in the ACS include type of occupation, commuting, veteran status, health insurance status, expanded racial categories and household relationship types, and internet access. These link files (ACS-HUD linked files) are available to researchers within Federal Research Data Centers. Using the data, information can be used to gain insights into HUD-assisted households that would otherwise not be possible with current administrative data.

Sponsoring agency: Department of Housing and Urban Development and U.S. Census Bureau

Data universe:
ACS respondents who are in a HUD-Assisted Household

Disaggregated Asian and Native Hawaiian or Other Pacific Islander categories collected:
Afghan, Asian Indian, Bangladeshi, Bengali, Bhutanese, Bikinian, Bruneian, Burmese, Buryat, Cambodian, Carolinian, Central Asian, Cham, Chinese, Chuukese, Cook Islander, East Asian, Easter Islander, Ebeye, Ejit, Enewetak Islander, Fijian, Filipino, French Polynesian, Guamanian or Chamorro, Hakka, Han, Hawaiian, Hmong, Hong Kong, I-Kiribati, Indo, Indo Fijian, Indo-Chinese, Indonesian, Iwo Jiman, Japanese, Kalmyk, Kazakh, Khmer, Kili, Korean, Kosraean, Kuki, Kwajalein Islander, Kyrgyz, Lahu, Laotian, Macanese, Malay, Malaysian, Maldivian, Maori, Marshallese, Melanesian, Micronesian, Mien, Mili, Mizo, Mongolian, Montagnard, Native Hawaiian, Nauruan, Nepalese, New Caledonian, Niuean, Ni-Vanuatu, Northern Mariana Islander, Okinawan, Pacific Islander, Pakistani, Palauan, Papua New Guinean, Part Hawaiian, Pashtun, Pohnpeian, Polynesian, Punjabi, Rotuman, Saipanese, Samoan, Sikh, Sindhi, Singaporean, Solomon Islander, South Asian, Southeast Asian, Sri Lankan, Tahitian, Tai Dam, Taiwanese, Tajik, Thai, Tibetan, Timorese, Tokelauan, Tongan, Turkmen, Tuvaluan, Ujelang, Urdu, Uzbek, Vietnamese, Wallisian and Futunan, Yapese

Are disaggregated Asian and Native Hawaiian or Other Pacific Islander data publicly available? NO

Measures available:
Age, Assistance Programs, Disability, Early Childhood, Education, Elementary and Secondary Education, Employment Characteristics, Families and Living Arrangements, Food and Nutrition, Geography, Health Care and Insurance, Health Insurance Coverage and Access, Homeownership, Housing, Income, Migration and nativity, Postsecondary Education and Beyond, Poverty, Reproductive Health, Sex and Gender, Transportation and Commuting, Unemployment, Veterans, Voting and Registration

Sample size and unit of analysis:
Roughly 60,000 to 75,000 ACS households are linked to as HUD administrative record each year.

Types of geography available to researchers: The RESTRICTED-USE data includes geographic information down to the block group for each record. Standard pre-tabulated ACS data products based on the full RESTRICTED-USE file are available for geographies and geographic components down to the block group unless otherwise suppressed due to population thresholds, quality, or disclosure restrictions.
The PUMS files include geographic information for Regions, Divisions, States, Puerto Rico, and Public Use Microdata Areas (PUMAs). PUMAs are non-overlapping, statistical geographic areas that partition each state or equivalent entity into geographic areas containing no fewer than 100,000 people each.

What data are available?
PUBLIC-USE data? NO
RESTRICTED-USE data? YES
**What is the difference between PUBLIC-USE and RESTRICTED-USE data files:**

The ACS/HUD PUMS indicator is a model prediction of HUD-subsidy status. It was not built for producing publication-quality estimates of HUD-assisted units.

The PUMS/HUD indicator were created for three purposes:
1. Develop code to produce estimates of interest. Users would then supply their code to HUD so that it can be run on any ACS-HUD linked file from 2011 through 2019 to produce publication-quality estimates.
2. Test sample sizes before moving forward getting Special Sworn Status and submitting a proposal to access the restricted ACS-HUD linked files. If the sample sizes based on the PUMS/HUD indicator need further clarification, users could supply code to HUD to get confirmation of sample sizes. Develop code outside a Federal Statistical Research Data Center to be brought into the Census environment for use on the restricted ACS-HUD linked files.

**How to access the PUBLIC-USE data:**

HUD anticipates that many analysts will want to use the ACS-HUD linked files to produce summary statistics but either cannot get Special Sworn status due to limited resources or will not want to go through the process of obtaining Special Sworn Status without some assurance that their methodology is feasible. To accommodate those users, HUD has developed a “HUD-assisted unit” categorical indicator that can be merged to the 2015 ACS Public Use Microdata Sample (PUMS). This product, called the 2015 PUMS/HUD indicator, classifies each ACS PUMS housing unit as one of four values: public housing (PH), voucher (VO), project-based Section 8 (MF), or not assisted (None). The data can be found on the HUDUser website under the ACS-HUD data linkage page. [https://www.huduser.gov/portal/sites/default/files/pdf/HUDPUMS_Indicator_Howto.pdf](https://www.huduser.gov/portal/sites/default/files/pdf/HUDPUMS_Indicator_Howto.pdf)

**How to obtain permission to use the RESTRICTED-USE data:**

The Census Bureau considers proposals from qualified researchers in social science disciplines consistent with the subject matter of the surveys and censuses collected by the Census Bureau. Proposals may be submitted at any time:

1. Go to [census.gov](https://www.census.gov)
2. Select "Surveys and Programs",
3. Select "All surveys and programs."
4. Select "Center for Economic Studies,"
5. Then navigate to "Data" (on the left side of the CES page).
6. Select "RESTRICTED-USE Data",
7. Then "Apply for access": [https://www.census.gov/about/adrm/ced/apply-for-access.html](https://www.census.gov/about/adrm/ced/apply-for-access.html)

**Constraints surrounding access to RESTRICTED-USE data (e.g., citizenship requirements, cost, etc):**

All researcher access to RESTRICTED-USE data occurs at secure Federal Statistical Research Data Centers (FSRDC). Some FSRDCs may charge fees for access to their facilities.
Agency policy re. sample sizes/publishable estimates/data suppression standards:
Areas or groups of 65,000 or more are eligible for 1-, 5- year, and Supplemental Estimates. Areas or groups of 20,000 or more are eligible for 5-year and Supplemental Estimates. Areas or groups of 20,000 or fewer are eligible for 5-year estimates only. [https://www2.census.gov/programs-surveys/acs/tech_docs/data_suppression/ACSO_Data_Suppression.pdf](https://www2.census.gov/programs-surveys/acs/tech_docs/data_suppression/ACSO_Data_Suppression.pdf)

Information on when the data were collected/what time period could be examined using the data:
ACS/HUD linked files are available for the 2011–2019 one-year ACS microdata.

Where to find disseminated official estimates: NA

Type of data (survey, program, admin etc.): Survey and administrative.

Do the data include measures of citizenship, or are respondents limited by citizenship (i.e., only U.S. citizens are included in the data)? The ACS asks questions about a person’s place of birth, citizenship, and year of entry into the United States to create data about citizens, noncitizens, and the foreign-born population. The ACS/HUD linked files are only suitable for estimates at the housing unit level. Thus, estimates of citizenship of HUD-Subsidy recipients cannot be made at the individual level.

Analytic Considerations:
The linkage process was designed only to identify ACS housing units as receiving HUD assistance. That process means that the linked data set can be used to produce estimates of ACS characteristics for HUD-assisted housing units. HUD made no attempt to determine if the ACS household and the HUD household contained the same members, however.

HUD provides a technical document that details potential sources of error and quality assessment of the linkage. To determine whether the multifaceted linking process performed well, HUD compared the pre-linking known HUD-assisted record count with the post-linking ACS weighted estimate of HUD-assisted housing units. Based on this analysis, HUD concluded that the matching algorithm performed well enough to ensure that the ACS housing unit cases flagged as HUD-assisted units are a representative cross-section of all possible ACS cases that are HUD-assisted units. In statistical terms, although there are false negatives, their omission will not necessarily result in biased estimates of housing or household characteristics of HUD-assisted households.

Because the ACS is a sample-based survey, all estimates are subject to sampling error, including estimates of HUD-assisted housing units.

For more information, see the technical documentation on the HUDUser website under the HUD/ACS Linkage page. [https://www.huduser.gov/portal/datasets/acs-hud-data-linkage.html](https://www.huduser.gov/portal/datasets/acs-hud-data-linkage.html)
IV. **American Housing Survey (AHS)**

The American Housing Survey (AHS) is sponsored by the Department of Housing and Urban Development (HUD) and conducted by the U.S. Census Bureau. The survey has been the most comprehensive national housing survey in the United States since its inception in 1973, providing current information on the size, composition, and quality of the nation’s housing and measuring changes in our housing stock as it ages. The AHS is a longitudinal housing unit survey conducted biennially in odd-numbered years, with samples redrawn in 1985 and 2015. The survey provides up-to-date information about the quality and cost of housing in the United States and major metropolitan areas. The survey also includes questions about: the physical condition of homes and neighborhoods, the costs of financing and maintaining homes, and the characteristics of people who live in these homes.

**Sponsoring agency:** Department of Housing and Urban Development and U.S. Census Bureau

**Data universe:**
Residential housing units in the United States that exist at the time the survey is conducted.

**Disaggregated Asian and Native Hawaiian or Other Pacific Islander categories collected:**
Asian Indian, Chinese, Chuukese, French Polynesian, Hakka, Kalmyk, Kosraean, Nauruan, Sikh, Wallisian and Futunan

**Are disaggregated Asian and Native Hawaiian or Other Pacific Islander data publicly available?** YES

**Measures available:**
Assistance Programs, Community and Economic Development, Families and Living Arrangements, Health, Homeownership, Housing, Income

**Sample size and unit of analysis:**
The 2019 AHS sample, comprised of an integrated national longitudinal sample and an integrated metropolitan longitudinal sample, includes a total of 117,422 housing units. The integrated national longitudinal sample is made up of 86,151 housing units, which includes three parts: (1) 34,876 national cases representative of the U.S. and 9 Census divisions, (2) 46,090 housing units from the Top 15 metropolitan areas (an oversample where National plus the oversample is approximately 3,000 per metro area), and (3) 5,185 housing units that are part of a HUD-assisted renter oversample. The integrated metropolitan longitudinal sample is comprised of 31,271 housing units, which represent 10 additional metropolitan areas (approximately 3,000 per metro area).
Types of geography available to researchers: National data are collected each survey year. From 2015 onward, the same "Top 15" metropolitan areas will be oversampled in every survey cycle, as will an additional 10 metros (referred to as the 'dynamic 10'). The additional 10 metros will be selected from a rotating sample of twenty other Metropolitan areas, which will be in sample every other survey year. A help guide is available on the Census website that summarizes history of metropolitan area oversamples.

What data are available?
PUBLIC-USE data? YES
RESTRICTED-USE data? YES

What is the difference between PUBLIC-USE and RESTRICTED-USE data files:
The AHS internal use file (IUF) contains geographic identifiers (e.g., census tract) that do not appear on the public use file (PUF) and also includes data that has not been topcoded or aggregated for confidentiality reasons. In addition, the IUFs contain variables that can be used to calculate the correct standard errors for a complex random sample. IUF records, however, have been edited and imputed as needed.
The 2015 and later public use files (PUFs) do not include many of the geographic indicators included in prior year PUFs. This is due to the disclosure avoidance procedures required by the U.S. Census Bureau. The 2015 and later integrated national longitudinal sample PUF (national PUF) will include two geographic variables: Census Division and OMB13CBSA, while the independent metropolitan area longitudinal oversample PUF (metropolitan area PUF) will include OMB13CBSA only.

How to access the PUBLIC-USE data:
The U.S. Census Bureau releases public use file (PUF) microdata and summary tables and/or reports approximately 12 months after data collection. The data and summary tables are available on the Census Website.

1. Go to census.gov
2. Select the Surveys/Programs drop down menu, and click on American Housing Survey: https://www.census.gov/programs-surveys/ahs.html

How to obtain permission to use the RESTRICTED-USE data:
Access to the IUF is available through the secure Research Data Center (RDC) network. Approved researchers must obtain a Special Sworn Status (SSS) from the Census Bureau. SSS researchers are sworn for life to protect the confidentiality of the data they access. There are a number of partnering universities and research institutions throughout the United States staffed with a Census Bureau employee. For information on how to apply and the types of proposals that are acceptable see https://www.census.gov/about/adrm/fsrdc.html
Constraints surrounding access to RESTRICTED-USE data (e.g., citizenship requirements, cost, etc):

All researcher access to RESTRICTED-USE data occurs at secure Federal Statistical Research Data Centers (FSRDC). Some FSRDCs may charge fees for access to their facilities.

Agency policy re. sample sizes/publishable estimates/data suppression standards:

Areas or groups of 65,000 or more are eligible for 1-, 5- year, and Supplemental Estimates. Areas or groups of 20,000 or more are eligible for 5-year and Supplemental Estimates. Areas or groups of 20,000 or fewer are eligible for 5-year estimates only. https://www2.census.gov/programs-surveys/acs/tech_docs/data_suppression/ACSO_Data_Suppression.pdf

Information on when the data were collected/what time period could be examined using the data:

The AHS started collecting detailed AANHPI information in 2015.

The AHS is conducted biennially between May and September in odd-numbered years. HUD sometimes adjusts this schedule and/or sample depending on budget constraints. While national data are always collected, typically no more than 30 metropolitan areas are sampled in one survey year, due to budget constraints.

Where to find disseminated official estimates:

AHS publishes official estimates via the AHS Table Creator tool available on the Census website.

Type of data (survey, program, admin etc.): Survey

Do the data include measures of citizenship, or are respondents limited by citizenship (i.e., only U.S. citizens are included in the data)? The AHS collects information on U.S. citizenship of the householder and other housing unit members. The AHS only includes housing unit level weights so estimates of person-level citizenship cannot be made using the AHS.

Analytic Considerations:

For estimates of detailed AANHPI, sample size will be limited. For this reason, many official AANPI estimates on the AHS Table Creator are suppressed. When using the public use file, data users should consult the various help guides and technical documentation available on the AHS webpage.

The AHS only has weight for housing units. Thus, estimates cannot be made at the person level.
V. **American Time Use Survey (ATUS)**

The American Time Use Survey (ATUS) measures the amount of time people spend doing various activities, such as paid work, childcare, volunteering, and socializing.

**Sponsoring agency:** Bureau of Labor Statistics

**Data universe:**
Civilian noninstitutional population age 15 and over.

**Disaggregated Asian and Native Hawaiian or Other Pacific Islander categories collected:**
Asian Indian, Chinese, Filipino, Korean, Japanese, Vietnamese

**Are disaggregated Asian and Native Hawaiian or Other Pacific Islander data publicly available?** YES

**Measures available:**
Age, Business and Economy, Disability, Early Childhood, Education, Employment Characteristics, Families and Living Arrangements, Geography, Income, Migration and Nativity, Postsecondary Education and Beyond, Sex and Gender, Spending and Time Use, Transportation and Commuting, Unemployment, Veterans

**Sample size and unit of analysis:**
Nationally representative sample of roughly 26,000 individuals per year.

**Types of geography available to researchers:** State, county, metropolitan core-based statistical area

**What data are available?**
PUBLIC-USE data? YES
RESTRICTED-USE data? NO

**What is the difference between PUBLIC-USE and RESTRICTED-USE data files:**
N/A

**How to access the PUBLIC-USE data:**
Available on the Bureau of Labor Statistics website:

1. Go to [bls.gov](https://bls.gov)
2. Select "SUBJECTS"
3. Select "How Americans Spend Time"
4. Select "ATUS Data Files": www.bls.gov/tus/data.htm

How to obtain permission to use the RESTRICTED-USE data:

NA

Constraints surrounding access to RESTRICTED-USE data (e.g., citizenship requirements, cost, etc):

NA

Agency policy re. sample sizes/publishable estimates/data suppression standards:

Requirements for official ATUS estimates put forth by the Bureau of Labor Statistics vary depending on the type of estimate being produced. These criteria account for the number of respondents representing the given population, the number of respondents who reported doing a specified activity, and the standard error or coefficient of variation for the estimates. For a detailed description of the statistical reliability criteria necessary for publication, please contact ATUS staff at ATUSinfo@bls.gov.

Information on when the data were collected/what time period could be examined using the data:

Annual data for 2013 – present (2013 does not have this information for all records)

Where to find disseminated official estimates:

NA

Type of data (survey, program, admin etc.): Survey

Do the data include measures of citizenship, or are respondents limited by citizenship (i.e., only U.S. citizens are included in the data)? Includes variables identifying foreign- and native-born status; citizenship status; and country of birth.

Analytic Considerations:

Sample sizes for disaggregated Asian groups are extremely small in the annual ATUS data. For time-use analyses of these groups at the national level, it is necessary to combine multiple years of data. The ATUS occasionally has special modules—that is, sets of special questions that are included on the ATUS, generally for a year or longer. Public use files for these modules are available on the Bureau of Labor Statistics website and can be linked to the ATUS data files. However, the sample sizes for disaggregated Asian groups for modules are very small and are unlikely to provide reliable estimates.
VI. **Baccalaureate and Beyond Longitudinal Study (B&B)**

The Baccalaureate and Beyond Longitudinal Study (B&B) is a nationally representative longitudinal study of students who completed the requirements for a bachelor's degree in a given academic year. Graduating seniors are followed 1, 4, and 10 years after completing their bachelor’s degree. There is a special emphasis on new elementary and secondary teachers. The National Postsecondary Student Aid Study (NPSAS) serves as the base year collection for each B&B cohort. To date, four cohorts have been followed.

Students in the first cohort completed their bachelor’s degree in 1992-93 and were followed up with in 1994, 1997, and 2003. Students in the second cohort completed their bachelor’s degree in 1999–2000 and were only followed up with one time in 2001. Students in the third cohort completed their bachelor’s degree in 2007–08 and were followed up with in 2009, 2012, and 2018. Students in the fourth cohort completed their bachelor’s degree in 2015–16 and were followed up with in 2017 and 2020.

Using B&B data, researchers can address questions regarding bachelor’s degree recipients' undergraduate experiences, including their participation in various financial aid programs, student loan debt and repayment of that debt; entrance into and progress through postbaccalaureate education; employment, and family formation. B&B also includes items to help policymakers and researchers better understand the teacher pipeline and experiences and outcomes of teachers.

**Sponsoring agency:** National Center for Education Statistics, Institute of Education Sciences, U.S. Department of Education

**Data universe:**
Undergraduate students who completed the requirements for a bachelor’s degree in a given academic year.

**Disaggregated Asian and Native Hawaiian or Other Pacific Islander categories collected:**
Asian Indian, Chinese, Filipino, Japanese, Korean, Vietnamese

**Are disaggregated Asian and Native Hawaiian or Other Pacific Islander data publicly available?** YES

**Measures available:**
Age, Assistance Programs, Disability, Education, Employment Characteristics, Families and Living Arrangements, Geography, Health, Homeownership, Income, Migration and Nativity, Postsecondary Education and Beyond, Sex and Gender, Unemployment, Veterans, Voting and Registration
**Sample size and unit of analysis:**
B&B:16/17 - Nationally representative sample of approximately 29,000 students who completed the requirements for a bachelor’s degree during the 2015–16 academic year.

B&B:08/18 - Nationally representative sample of approximately 17,200 students who completed the requirements for a bachelor’s degree during the 2007–08 academic year.

B&B:2000/01 - Nationally representative sample of approximately 10,000 students who completed the requirements for a bachelor’s degree during the 1999–2000 academic year.

B&B:1993/2003 - Nationally representative sample of approximately 11,000 students who completed the requirements for a bachelor’s degree during the 1992–93.

**Types of geography available to researchers:** This geographic information is available both in the online data tool and the RESTRICTED-USE file: State for sample member’s residence, baccalaureate institution, job, teaching job, and high school attended; Locale for sample member’s residence, baccalaureate institution, and teaching job; Region for sample member’s residence and baccalaureate institution. The RESTRICTED-USE file also includes ZIP codes for the sample member’s high school, residences, job, and teaching job.

**What data are available?**
PUBLIC-USE data? YES
RESTRICTED-USE data? YES

**What is the difference between PUBLIC-USE and RESTRICTED-USE data files:**
Case-level data are available only in RESTRICTED-USE format. The online data tool allows a researcher to run analyses without accessing the case-level data. The online data tool and the RESTRICTED-USE file generally include the same derived variables, except for ZIP codes. The RESTRICTED-USE file includes the source data files used to create the derived variables.

**How to access the PUBLIC-USE data:**
These data are available through PowerStats, an interactive online interface that enables users to generate custom statistical analyses and data visualizations. To access PowerStats,

1. Go to [nces.ed.gov](http://nces.ed.gov) and type "DataLab" in the Search box.
2. PowerStats will result: [https://nces.ed.gov/datalab](https://nces.ed.gov/datalab).
How to obtain permission to use the RESTRICTED-USE data:
Procedures for applying for an NCES RESTRICTED-USE license can be from the NCES homepage. Go to nces.ed.gov and type "RESTRICTED-USE license" in the Search box: https://nces.ed.gov/statprog/instruct.asp

Constraints surrounding access to RESTRICTED-USE data (e.g., citizenship requirements, cost, etc):
While there are no citizenship requirements for access, currently restricted-use data can only be accessed at institutions within the United States.

Agency policy re. sample sizes/publishable estimates/data suppression standards:
Estimates can be reported if based on a numerator size of at least 3 and a denominator size of at least 30. Unweighted sample sizes based on restricted data must be rounded to the nearest 10.

Information on when the data were collected/what time period could be examined using the data:
B&B:16/17 – data were collected in 2015–16 (base year), and then one year later in 2017. Data were also collected in 2020, four years after graduation.

B&B:08/18 – data were collected in 2007–08 (base year), and then 1, 4, and 10 years later (i.e., 2009, 2012, and 2018, respectively).

B&B:2000/01 – data were collected in 1999–2000 (base year), and then 1 year later in 2001.

B&B:1993/2003 – data were collected in 1992–93 (base year), and then 1, 4, and 10 years later (i.e., 1994, 1997, and 2003, respectively).

Where to find disseminated official estimates:
In published reports on the NCES website: http://nces.ed.gov

B&B data are included in study-specific reports (see https://nces.ed.gov/pubsearch/getpubcats.asp?sid=015), The Condition of Education (https://nces.ed.gov/programs/coe/), and the Digest of Education Statistics (https://nces.ed.gov/programs/digest/). However, estimates disaggregated for the AANHPI subgroups have not been released in any NCES publications.

Estimates can also be found in the DataLab Tables Library: https://nces.ed.gov/datalab/table/library. You can search for tables that include B&B.

Type of data (survey, program, admin etc.): Survey, institution student records (base year only), and administrative data
Do the data include measures of citizenship, or are respondents limited by citizenship (i.e., only U.S. citizens are included in the data)? B&B includes information on students enrolled in postsecondary institutions in the 50 states, Puerto Rico, and the District of Columbia. There was no restriction on citizenship to be included in the sample. The datasets include an indicator of whether the sample member is a U.S. citizen, a permanent resident, or has a visa.

Analytic Considerations:
Though Native Hawaiians and Other Pacific Islanders are included in B&B, information allowing for identification of specific subgroup was not collected.

VII. 2020 Census

The 2020 Census counted every person living in the United States and the five U.S. territories. It marked the 24th census in U.S. history and the first time that households were invited to respond to the census online.

The 2020 Census apportionment results and redistricting data are available. Data yet to be released, however, include detailed racial and ethnic groups. For additional information about current plans for developing and releasing these statistics, please refer to [https://www.census.gov/newsroom/blogs/random-samplings/2022/04/2020-census-detailed-race-ethnic-tribal-data.html](https://www.census.gov/newsroom/blogs/random-samplings/2022/04/2020-census-detailed-race-ethnic-tribal-data.html).

Sponsoring agency: U.S. Census Bureau

Data universe:
Every person living in the United States and the five U.S. territories.

Disaggregated Asian and Native Hawaiian or Other Pacific Islander categories collected:
Afghan, Asian Indian, Bangladeshi, Bengali, Bhutanese, Bikinian, Bruneian, Burmese, Buryat, Cambodian, Carolinian, Central Asian, Cham, Chinese, Chuukese, Cook Islander, East Asian, Easter Islander, Ebeye, Ejit, Enewetak Islander, Fijian, Filipino, French Polynesian, Guamanian or Chamorro, Hakka, Han, Hawaiian, Hmong, Hong Kong, I-Kiribati, Indo, Indo Fijian, Indo-Chinese, Indonesian, Iwo Jiman, Japanese, Kalmyk, Kazakh, Khmer, Kili, Korean, Kosraean, Kuki, Kwajalein Islander, Kyrgyz, Lahu, Laotian, Macanese, Malay, Malaysian, Maldivian, Maori, Marshallese, Melanesian, Micronesian, Mien, Mili, Mizo, Mongolian, Montagnard, Native Hawaiian, Nauruan, Nepalese, New Caledonian, Niuean, Ni-Vanuatu, Northern Mariana Islander, Okinawan, Pacific Islander, Pakistani, Palauan, Papua New Guinean, Part Hawaiian, Pashtun, Pohnpeian, Polynesian, Punjabi, Rotuman, Saipanese, Samoan, Sikh, Sindhi, Singaporean, Solomon Islander, South Asian, Southeast Asian, Sri
Lankan, Tahitian, Tai Dam, Taiwanese, Tajik, Thai, Tibetan, Timorese, Tokelauan, Tongan, Turkmen, Tuvaluan, Ujelang, Urdu, Uzbek, Vietnamese, Wallisian and Futunan, Yapese

**Are disaggregated Asian and Native Hawaiian or Other Pacific Islander data publicly available?** YES

**Measures available:**
Age, Families and Living Arrangements, Geography, Homeownership, Sex and Gender

**Sample size and unit of analysis:**
The decennial census is a count of all persons living in the United States and associated territories as of April 1.

**Types of geography available to researchers:** The lowest level of geography for population counts is still being determined. Counts are currently proposed to be available for the nation, states, counties, and American Indian, Alaska Native, and Hawaiian Homeland Areas.

**What data are available?**
PUBLIC-USE data? YES
RESTRICTED-USE data? YES

**What is the difference between PUBLIC-USE and RESTRICTED-USE data files:**
Still being determined as of April, 2022.

**How to access the PUBLIC-USE data:**
1. Go to census.gov
2. Select "Surveys/Programs", and then select 2020 Census Results from the drop-down menu: https://www.census.gov/programs-surveys/decennial-census/decade/2020/2020-census-results.html

**How to obtain permission to use the RESTRICTED-USE data:**
The Census Bureau considers proposals from qualified researchers in social science disciplines consistent with the subject matter of the surveys and censuses collected by the Census Bureau. Proposals may be submitted at any time.

1. Go to census.gov
2. Select "Surveys and Programs", and "All surveys and programs."
3. Select "Center for Economic Studies," and then navigate to "Data" (on the left side of the CES page).
4. Select "RESTRICTED-USE Data", and then "Apply for access": https://www.census.gov/about/adrm/ced/apply-for-access.html

Constraints surrounding access to RESTRICTED-USE data (e.g., citizenship requirements, cost, etc):

All researcher access to RESTRICTED-USE data occurs at secure Federal Statistical Research Data Centers (FSRDC). Some FSRDCs may charge fees for access to their facilities.

Agency policy re. sample sizes/publishable estimates/data suppression standards:

Confidentiality protections and differential privacy methods are still being determined. https://www.census.gov/newsroom/blogs/random-samplings/2021/04/modernizing_privacy.html

Information on when the data were collected/what time period could be examined using the data:

The decennial census is a count of all persons living in the United States and associated territories as of April 1.

Where to find disseminated official estimates:

1. Data will be released on data.census.gov

Type of data (survey, program, admin etc.): Census

Do the data include measures of citizenship, or are respondents limited by citizenship (i.e., only U.S. citizens are included in the data)? NO

Analytic Considerations:

Population thresholds for detailed Asian and NHPI groups for pre-tabulated 2020 Census data products are still being determined as of April, 2022.

VIII. Current Population Survey (CPS)

The Current Population Survey is a monthly household survey that provides a wide range of information about employment, unemployment, hours of work, earnings, and people not in the labor force. CPS data can be broken down by demographic characteristics, including age, sex, race, Hispanic ethnicity, and educational attainment.

Sponsoring agency: Bureau of Labor Statistics, Census Bureau
Data universe:
Civilian noninstitutional population age 16 and over

Disaggregated Asian and Native Hawaiian or Other Pacific Islander categories collected:
Asian Indian, Chinese, Filipino, Japanese, Korean, Vietnamese

Are disaggregated Asian and Native Hawaiian or Other Pacific Islander data publicly available? YES

Measures available:
Age, Business and Economy, Disability, Education, Employment Characteristics, Families and Living Arrangements, Geography, Income, Migration and Nativity, Postsecondary Education and Beyond, Sex and Gender, Unemployment, Veterans

Sample size and unit of analysis:
Nationally representative sample of roughly 60,000 households (approximately 110,000 people)

Types of geography available to researchers: State, county, metropolitan core-based statistical area

What data are available?
PUBLIC-USE data? YES
RESTRICTED-USE data? NO

What is the difference between PUBLIC-USE and RESTRICTED-USE data files:
NA

How to access the PUBLIC-USE data:
1. Go to census.gov
2. Select "Surveys"
4. Select "Data"

Data are available on the Census Bureau’s website, either as flat downloadable files (at https://www.census.gov/data/datasets/time-series/demo/cps/cps-basic.html) or through the MDAT tool (at https://data.census.gov/mdat/#/).

How to obtain permission to use the RESTRICTED-USE data:
NA
Constraints surrounding access to RESTRICTED-USE data (e.g., citizenship requirements, cost, etc):
NA

Agency policy re. sample sizes/publishable estimates/data suppression standards:
Official estimates put forth by the Bureau of Labor Statistics for the Current Population Survey must adhere to suppression rules. Generally, percentages are suppressed if the monthly base is less than 75,000 persons, the quarterly average base is less than 60,000 persons, or the annual average base is less than 35,000 persons.

Information on when the data were collected/what time period could be examined using the data:
Monthly for 2003 – present (data from 2013 – present appears on public use files)

Where to find disseminated official estimates:
Disaggregated estimates on Asian groups are included in table 2 of the annual report “Labor force characteristics by race and ethnicity.”

1. Go to bls.gov
2. Select "CPS"
3. Select "Publications":
   https://www.bls.gov/cps/demographics.htm#race

Type of data (survey, program, admin etc.): Survey

Do the data include measures of citizenship, or are respondents limited by citizenship (i.e., only U.S. citizens are included in the data)? Includes variables identifying foreign- and native-born status; citizenship status; country of birth; and date of entry.

Analytic Considerations:
Sample sizes for disaggregated Asian groups can be extremely small in the CPS for subnational areas. For reliable estimates for many subnational areas, it is necessary to combine multiple years of data. The CPS often has special supplements—that is, sets of questions that are given to respondents for one month, some of which repeat annually and some of which are conducted rarely. These include the Annual Social and Economic Supplement (sometimes known as the March Supplement), as well as supplements on Unemployment Insurance nonfilers, veterans, people with disabilities,
contingent workers, civic engagement, public participation in the arts, tobacco use, school enrollment, the unbanked or underbanked, volunteering and civic life, voting and registration, computer and internet use, and fertility. Public use files for these supplements are available on the Census Bureau’s website, and the files contain the same variable for disaggregated Asian groups as the monthly CPS files. However, the sample sizes for disaggregated groups for supplements are very small. Multiple years of supplement data would need to be pooled to obtain robust estimates.

IX. Early Childhood Longitudinal Study, Birth Cohort (ECLS-B)

The Early Childhood Longitudinal Study, Birth Cohort (ECLS-B) is a longitudinal study that followed a nationally representative sample of children from birth through kindergarten entry. The purpose of the ECLS-B was to provide a comprehensive and reliable set of data that may be used to describe and to better understand children’s early development; their home learning experiences; their experiences in early care and education programs; their health care, nutrition, and physical well-being; and how their early experiences relate to their later development, learning, and success in school.

The ECLS-B longitudinal 9-month-kindergarten 2007 restricted-use dataset includes data from every round of data collection. It includes information collected from the children, their parents/guardians, and their early care and education providers. The file also includes information collected from teachers and about children’s schools when the children enrolled in kindergarten.

Sponsoring agency: National Center for Education Statistics, Institute of Education Sciences, U.S. Department of Education

Data universe:
Children who were born in the United States in 2001

Disaggregated Asian and Native Hawaiian or Other Pacific Islander categories collected:
Asian Indian, Chinese, Filipino, Guamanian or Chamorro, Hawaiian, Hmong, Japanese, Korean, Pacific Islander, Samoan, Vietnamese

Are disaggregated Asian and Native Hawaiian or Other Pacific Islander data publicly available? NO

Measures available:
Age, Assistance Programs, Behavioral Risk Factors, Disability, Early Childhood, Education, Families and Living Arrangements, Food and Nutrition, Geography, Health, Health Care and Insurance,
Sample size and unit of analysis:

Types of geography available to researchers: State and county FIPS codes for location where the child was born and mother’s residence at the time of the child’s birth, Census region, locale (city, suburb, town, rural) of home and early care and education settings, and ZIP code for home and early care and education settings. All data are in restricted-use files.

What data are available?
PUBLIC-USE data? NO
RESTRICTED-USE data? YES

What is the difference between PUBLIC-USE and RESTRICTED-USE data files:
NA

How to access the PUBLIC-USE data:
NA

How to obtain permission to use the RESTRICTED-USE data:
Procedures for applying for an NCES RESTRICTED-USE license can be found by:

1. Go to nces.ed.gov
2. Type "RESTRICTED-USE license" in the Search box: https://nces.ed.gov/statprog/instruct.asp

Constraints surrounding access to RESTRICTED-USE data (e.g., citizenship requirements, cost, etc):
While there are no citizenship requirements for access, currently RESTRICTED-USE data can only be accessed at institutions within the United States.

Agency policy re. sample sizes/publishable estimates/data suppression standards:
Estimates can be reported if based on a numerator size of at least 3 and a denominator size of at least 30. Unweighted sample sizes for the ECLS-B must be rounded to the nearest 50.

Information on when the data were collected/what time period could be examined using the data:
Data were collected at 5 time points: when the children were about 9 months of age (2001–02), about 2 years of age (2003–04), about 4 years of age (2005–06), and when they were in
kindergarten (ages 5 and 6). In the fall of 2006, data were collected from all participating sample children, approximately 75 percent of whom were in kindergarten or higher. In the fall of 2007, data were collected from the approximately 25 percent of participating sample children who had not yet entered kindergarten or higher in the previous collection, as well as children who were repeating kindergarten in the 2007–08 school year. Information extracted from the children's birth certificate is also included in the data file.

**Where to find disseminated official estimates:**

To find this information from the NCES homepage:

1. Go to [nces.ed.gov](http://nces.ed.gov)
2. Type “early childhood longitudinal study” in the Search box.

ECLS data are included in ECLS-specific reports (see [https://nces.ed.gov/pubsearch/getpubcats.asp?sid=024](https://nces.ed.gov/pubsearch/getpubcats.asp?sid=024)), the *Condition of Education* ([https://nces.ed.gov/programs/coe/](https://nces.ed.gov/programs/coe/)), and the *Digest of Education Statistics* ([https://nces.ed.gov/programs/digest/](https://nces.ed.gov/programs/digest/)). However, estimates disaggregated for the AANHPI subgroups have not been released in any NCES publications.

**Type of data (survey, program, admin etc.):** Surveys, direct assessments starting at 9 months of age, observations of parent-child interactions and childcare settings, and administrative records.

**Do the data include measures of citizenship, or are respondents limited by citizenship (i.e., only U.S. citizens are included in the data)?** All children were born in the United States. For the sampled child’s parents/guardians, information on their county of birth and, if not born in the U.S., the year parents came to the US and citizenship status is collected from parent reports.

**Analytic Considerations:**

The children participating in the study came from diverse socioeconomic and racial/ethnic backgrounds with oversamples of Chinese children, other Asian and Pacific Islander children, American Indian and Alaska Native children, twins, and children born with low and very low birth weight.


The Early Childhood Longitudinal Study, Kindergarten Class of 1998-99 (ECLS-K) is a longitudinal study that followed a nationally representative sample of students from their kindergarten year to the spring of 2007, when most of the students were in eighth grade. The longitudinal nature of the ECLS-K data allows for the examination of how a wide range of family, school, community, and individual factors are associated with educational, socioemotional, and physical development over time.
The ECLS-K longitudinal kindergarten through eighth-grade dataset includes data from every round of data collection. It includes information collected from the students, including direct assessment data in reading, mathematics, science and, in the early grades, fine motor skills and general knowledge; their parents/guardians; their teachers; and their school administrators in each year of the study.

**Sponsoring agency:** National Center for Education Statistics, Institute of Education Sciences, U.S. Department of Education

**Data universe:**
Children who were in kindergarten in the 1998-99 academic year

**Disaggregated Asian and Native Hawaiian or Other Pacific Islander categories collected:**
Asian Indian, Chinese, Filipino, Guamanian or Chamorro, Hawaiian, Hmong, Japanese, Korean, Pacific Islander, Samoan, Vietnamese

**Are disaggregated Asian and Native Hawaiian or Other Pacific Islander data publicly available?** YES

**Measures available:**
Age, Assistance Programs, Behavioral Risk Factors, Disability, Early Childhood, Education, Elementary and Secondary Education, Families and Living Arrangements, Food and Nutrition, Geography, Health, Health Care and Insurance, Housing, Income, Mental Health, Migration and Nativity, Poverty, Sex and Gender, Spending and Time Use, Unemployment

**Sample size and unit of analysis:**
Nationally representative sample of approximately 22,000 participating kindergartners followed through their elementary school years.

**Types of geography available to researchers:** Census region and students' home locale (city, suburb, town, rural) are available in the PUBLIC-USE files. The RESTRICTED-USE files include: school county and state FIPS codes, ZIP code, ZIP code tabulation area (ZCTA), Census tract, and latitude and longitude; and home ZIP code, ZCTA, and Census tract. There is also a supplementary RESTRICTED-USE geocode data file covering the first five rounds of data collection that includes more than 600 variables derived from 2000 Census data that are related to the demographic characteristics for each ZCTA and tract corresponding to schools and children in the ECLS-K sample. Examples of variables included in the file include: the median average income of residents, the percent of residents in different race/ethnicity and age groups, the percent of residents who own or rent their homes, the percent of households that are headed by females, and the percent that are military households.

**What data are available?**
PUBLIC-USE data? YES
RESTRICTED-USE data? YES

What is the difference between PUBLIC-USE and RESTRICTED-USE data files:
RESTRICTED-USE files include data that would be considered sensitive or potentially disclosive in combination with other data due to rarity or small sample sizes. For example, most disability-related data and geographic identifiers are included only in RESTRICTED-USE files. In addition, variables in RESTRICTED-USE files typically contain data at a finer level of detail than the public-use files. For example, specific primary home language is reported in the RESTRICTED-USE files but provided in categories (English, not English) in the PUBLIC-USE file.

How to access the PUBLIC-USE data:
ASCII data files and syntax programs for SAS, Stata, SPSS, SUDAAN and R can be downloaded from the ECLS website: https://nces.ed.gov/ecls/dataproducts.asp. Alternately, to find the data from the NCES homepage:

1. Go to nces.ed.gov
2. Type "ECLS data" in the Search box

How to obtain permission to use the RESTRICTED-USE data:
Procedures for applying for an NCES restricted-use license can be found from the NCES homepage:

1. Go to nces.ed.gov
2. Type "RESTRICTED-USE license" in the Search box: https://nces.ed.gov/statprog/instruct.asp

Constraints surrounding access to RESTRICTED-USE data (e.g., citizenship requirements, cost, etc):
While there are no citizenship requirements for access, currently RESTRICTED-USE data can only be accessed at institutions within the United States.

Agency policy re. sample sizes/publishable estimates/data suppression standards:
Estimates can be reported if based on a numerator size of at least 3 and a denominator size of at least 30. Unweighted sample sizes based on restricted data must be rounded to the nearest 10.

Information on when the data were collected/what time period could be examined using the data:
Data were collected at seven time points: the fall and spring of the 1998-99 school year, when all of the students were in kindergarten; the fall and spring of the 1999-2000 school year, when most of the students were in first grade; the spring of 2002, when most of the students were in third
grade; the spring of 2004, when most of the students were in fifth grade; and the spring of 2007, when most of the students were in eighth grade.

Where to find disseminated official estimates:
To find this information from the NCES homepage:

1. Go to nces.ed.gov
2. Type "early childhood longitudinal study" in the Search box


Type of data (survey, program, admin etc.): Surveys and direct assessments

Do the data include measures of citizenship, or are respondents limited by citizenship (i.e., only U.S. citizens are included in the data)? Children are included regardless of citizenship status. For the sampled child, country of birth, year child came to the US if not born in the US, and citizenship status are ascertained. For the sampled child’s parents/guardians, county of birth, year parent came to the US, and citizenship status if not born in the US are ascertained. This information is collected from parents/guardians in the parent interview.

Analytic Considerations:
Questions asking specific AANHPI subgroup were only asked in the fourth round of data collection, when most study children were in the spring of their first grade year. Therefore, these data are not available for study participants who dropped out of the study before that data collection round.

XI. Early Childhood Longitudinal Study, Kindergarten Class of 2010-11 (ECLS-K:2011)

The Early Childhood Longitudinal Study, Kindergarten Class of 2010-11 (ECLS-K:2011) is a longitudinal study that followed a nationally representative sample of students from their kindergarten year to the spring of 2016, when most of the students were in fifth grade. The longitudinal nature of the ECLS-K:2011 data allows for the examination of how a wide range of family, school, community, and individual factors are associated with educational, socioemotional, and physical development over time.
The ECLS-K:2011 longitudinal kindergarten through fifth-grade dataset includes data from every round of data collection. It includes information collected from the students, including direct assessment data in reading, mathematics, science, and executive function; their parents/guardians; their teachers; and their school administrators in each year of the study. It also includes information collected in the spring of 2011 from their kindergarten-year before- and after-school care providers.

**Sponsoring agency:** National Center for Education Statistics, Institute of Education Sciences, U.S. Department of Education

**Data universe:**
Children who were in kindergarten in the 2010-11 academic year

**Disaggregated Asian and Native Hawaiian or Other Pacific Islander categories collected:**
Asian Indian, Chinese, Filipino, Guamanian or Chamorro, Hmong, Japanese, Korean, Native Hawaiian, Pacific Islander, Samoan, Vietnamese

**Are disaggregated Asian and Native Hawaiian or Other Pacific Islander data publicly available?** YES

**Measures available:**
Age, Assistance Programs, Behavioral Risk Factors, Disability, Early Childhood, Education, Elementary and Secondary Education, Families and Living Arrangements, Food and Nutrition, Geography, Health, Health Care and Insurance, Housing, Income, Mental Health, Migration and Nativity, Poverty, Sex and Gender, Spending and Time Use, Unemployment

**Sample size and unit of analysis:**
Nationally representative sample of approximately 18,170 participating kindergartners followed through their elementary school years.

**Types of geography available to researchers:** Census region, school locale (city, suburb, town, rural), school county and state FIPS codes, school ZIP code and Census tract, home ZIP code and Census tract. Only school locale is available in the PUBLIC-USE files. All other data are in the RESTRICTED-USE files.

**What data are available?**

PUBLIC-USE data? YES
RESTRICTED-USE data? YES

**What is the difference between PUBLIC-USE and RESTRICTED-USE data files?**
RESTRICTED-USE files include data that would be considered sensitive or potentially disclosive in combination with other data due to rarity or small sample sizes. For example, most disability-
related data and geographic identifiers are included only in RESTRICTED-USE files. In addition, variables in RESTRICTED-USE files typically contain data at a finer level of detail than the PUBLIC-USE files. For example, specific primary home language is reported in the RESTRICTED-USE files but provided in categories (English, not English) in the PUBLIC-USE file.

**How to access the PUBLIC-USE data:**

ASCII data files and syntax programs for SAS, Stata, SPSS, SUDAAN and R can be downloaded from the ECLS website: [https://nces.ed.gov/ecls/dataproducts.asp](https://nces.ed.gov/ecls/dataproducts.asp). Alternately, to find the data from the NCES homepage:

1. Go to [nces.ed.gov](http://nces.ed.gov)
2. Type "ECLS data" in the Search box

**How to obtain permission to use the RESTRICTED-USE data:**

Procedures for applying for an NCES restricted-use license can be found from the NCES homepage:

1. Go to [nces.ed.gov](http://nces.ed.gov)
2. Type "RESTRICTED-USE license" in the Search box: [https://nces.ed.gov/statprog/instruct.asp](https://nces.ed.gov/statprog/instruct.asp)

**Constraints surrounding access to RESTRICTED-USE data (e.g., citizenship requirements, cost, etc):**

While there are no citizenship requirements for access, currently RESTRICTED-USE data can only be accessed at institutions within the United States.

**Agency policy re. sample sizes/publishable estimates/data suppression standards:**

Estimates can be reported if based on a numerator size of at least 3 and a denominator size of at least 30. Unweighted sample sizes based on restricted data must be rounded to the nearest 10.

**Information on when the data were collected/what time period could be examined using the data:**

Data were collected at 9 time points in 6 years: the fall and spring of the 2010–11 school year, when all of the students were in kindergarten; the fall and spring of the 2011–12 school year, when most of the students were in first grade; the fall and spring of the 2012-13 school year, when most of the students were in second grade; the spring of 2014, when most of the students were in third grade; the spring of 2015, when most of the students were in fourth grade; and the spring of 2016, when most of the students were in fifth grade.

**Where to find disseminated official estimates:**

To find this information from the NCES homepage:
1. Go to nces.ed.gov
2. Type "early childhood longitudinal study" in the Search box


**Type of data (survey, program, admin etc.):** Surveys and direct assessments

**Do the data include measures of citizenship, or are respondents limited by citizenship (i.e., only U.S. citizens are included in the data):** Children are included regardless of citizenship status. For the sampled child, country of birth, year child came to the US if not born in the US, and citizenship status are ascertained. For the sampled child’s parents/guardians, country of birth and year parent came to the US if not born in the US are ascertained. This information is collected from parents/guardians in the parent interview.

**Analytic Considerations:**
Data were not collected in Hawaii in the first round of data collection (the fall of the children’s kindergarten year), which affects longitudinal analysis of certain AANHPI groups that include base-year data. Questions asking specific AANHPI subgroup were not included until the seventh round of data collection, when most study children were in third grade. Therefore, these data are not available for study participants who dropped out of the study before that data collection round.

---


The Early Childhood Longitudinal Study, Kindergarten Class of 2010-11 (ECLS-K:2011) is a longitudinal study that followed a nationally representative sample of students from their kindergarten year to the spring of 2016, when most of the students were in fifth grade. The longitudinal nature of the ECLS-K:2011 data allows for the examination of how a wide range of family, school, community, and individual factors are associated with educational, socioemotional, and physical development over time.

The ECLS-K:2011 longitudinal kindergarten through fifth-grade dataset includes data from every round of data collection. It includes information collected from the students, including direct assessment data in reading, mathematics, science, and executive function; their parents/guardians; their teachers; and their school administrators in each year of the study. It also includes
information collected in the spring of 2011 from their kindergarten-year before- and after-school care providers.

**Sponsoring agency:** National Center for Education Statistics, Institute of Education Sciences, U.S. Department of Education

**Data universe:**
Children who were in kindergarten in the 2010-11 academic year

**Disaggregated Asian and Native Hawaiian or Other Pacific Islander categories collected:**
Asian Indian, Chinese, Filipino, Guamanian or Chamorro, Hmong, Japanese, Korean, Native Hawaiian, Pacific Islander, Samoan, Vietnamese

**Are disaggregated Asian and Native Hawaiian or Other Pacific Islander data publicly available?** YES

**Measures available:**
Age, Assistance Programs, Behavioral Risk Factors, Disability, Early Childhood, Education, Elementary and Secondary Education, Families and Living Arrangements, Food and Nutrition, Geography, Health, Health Care and Insurance, Housing, Income, Mental Health, Migration and Nativity, Poverty, Sex and Gender, Spending and Time Use, Unemployment

**Sample size and unit of analysis:**
Nationally representative sample of approximately 18,170 participating kindergartners followed through their elementary school years.

**Types of geography available to researchers:** Census region, school locale (city, suburb, town, rural), school county and state FIPS codes, school ZIP code and Census tract, home ZIP code and Census tract. Only school locale is available in the PUBLIC-USE files. All other data are in the RESTRICTED-USE files.

**What data are available?**
PUBLIC-USE data? YES
RESTRICTED-USE data? YES

**What is the difference between PUBLIC-USE and RESTRICTED-USE data files:**
RESTRICTED-USE files include data that would be considered sensitive or potentially disclosive in combination with other data due to rarity or small sample sizes. For example, most disability-related data and geographic identifiers are included only in RESTRICTED-USE files. In addition, variables in RESTRICTED-USE files typically contain data at a finer level of detail than the PUBLIC-USE files. For example, specific primary home language is reported in the RESTRICTED-USE files but provided in categories (English, not English) in the PUBLIC-USE file.
How to access the PUBLIC-USE data:
ASCII data files and syntax programs for SAS, Stata, SPSS, SUDAAN and R can be downloaded from the ECLS website: https://nces.ed.gov/ecls/dataproducts.asp. Alternately, to find the data from the NCES homepage:

1. Go to nces.ed.gov
2. Type "ECLS data" in the Search box

How to obtain permission to use the RESTRICTED-USE data:
Procedures for applying for an NCES restricted-use license can be found from the NCES homepage:

1. Go to nces.ed.gov
2. Type "RESTRICTED-USE license" in the Search box: https://nces.ed.gov/statprog/instruct.asp

Constraints surrounding access to RESTRICTED-USE data (e.g., citizenship requirements, cost, etc):
While there are no citizenship requirements for access, currently RESTRICTED-USE data can only be accessed at institutions within the United States.

Agency policy re. sample sizes/publishable estimates/data suppression standards:
Estimates can be reported if based on a numerator size of at least 3 and a denominator size of at least 30. Unweighted sample sizes based on restricted data must be rounded to the nearest 10.

Information on when the data were collected/what time period could be examined using the data:
Data were collected at 9 time points in 6 years: the fall and spring of the 2010–11 school year, when all of the students were in kindergarten; the fall and spring of the 2011–12 school year, when most of the students were in first grade; the fall and spring of the 2012-13 school year, when most of the students were in second grade; the spring of 2014, when most of the students were in third grade; the spring of 2015, when most of the students were in fourth grade; and the spring of 2016, when most of the students were in fifth grade.

Where to find disseminated official estimates:
To find this information from the NCES homepage:

1. Go to nces.ed.gov
2. Type "early childhood longitudinal study" in the Search box

ECLS data are included in ECLS-specific reports (see https://nces.ed.gov/pubsearch/getpubcats.asp?sid=024), a tables library
Do the data include measures of citizenship, or are respondents limited by citizenship (i.e., only U.S. citizens are included in the data): Children are included regardless of citizenship status. For the sampled child, country of birth, year child came to the US if not born in the US, and citizenship status are ascertained. For the sampled child’s parents/guardians, country of birth and year parent came to the US if not born in the US are ascertained. This information is collected from parents/guardians in the parent interview.

Analytic Considerations:
Data were not collected in Hawaii in the first round of data collection (the fall of the children’s kindergarten year), which affects longitudinal analysis of certain AANHPI groups that include base-year data. Questions asking specific AANHPI subgroup were not included until the seventh round of data collection, when most study children were in third grade. Therefore, these data are not available for study participants who dropped out of the study before that data collection round.

XIII. Education Longitudinal Study of 2002 (ELS:2002)

The Education Longitudinal Study of 2002 (ELS:2002) is a nationally representative, longitudinal study of more than 16,000 tenth-graders in public and private high schools who are being followed through high school and into adulthood. The main study students were re-surveyed in 2004 when most were high school seniors. At that time, the study was freshened to be nationally representative of high school seniors in 2004 as well. The study focuses on understanding students’ trajectories from the beginning of high school into postsecondary education or the workforce and beyond.

Sponsoring agency: National Center for Education Statistics, Institute of Education Sciences, U.S. Department of Education

Data universe:
2002 10th-graders, 2004 12th-graders

Disaggregated Asian and Native Hawaiian or Other Pacific Islander categories collected:
Chinese, Filipino, Japanese, Korean, South Asian, Southeast Asian
Are disaggregated Asian and Native Hawaiian or Other Pacific Islander data publicly available? NO

Measures available:
Age, Behavioral Risk Factors, Disability, Education, Elementary and Secondary Education, Employment Characteristics, Families and Living Arrangements, Geography, Homeownership, Housing, Income, Mental Health, Migration and Nativity, Postsecondary Education and Beyond, Poverty, Sex and Gender, Spending and Time Use, Unemployment, Veterans, Voting and Registration

Sample size and unit of analysis:
Nationally representative sample of about 16,000 10th-graders followed through their high school years and beyond.

Types of geography available to researchers: Census region, school locale (city, suburb, town, rural), school county and state FIPS codes, school ZIP code and Census tract, home ZIP code and Census block/tract. Only school locale is available in the public-use files. All other data are in the RESTRICTED-USE files.

What data are available?
PUBLIC-USE data? YES
RESTRICTED-USE data? YES

What is the difference between PUBLIC-USE and RESTRICTED-USE data files:
RESTRICTED-USE files include data that would be considered sensitive or potentially disclosive in combination with other data due to rarity or small sample sizes. For example, most disability-related data and geographic identifiers are included only in RESTRICTED-USE files. In addition, variables in RESTRICTED-USE files typically contain data at a finer level of detail than the PUBLIC-USE files. For example, specific primary home language is reported in the RESTRICTED-USE files but provided in categories (English, not English) in the PUBLIC-USE file.

How to access the PUBLIC-USE data:
ASCII data files and syntax programs for SAS, Stata, SPSS, SUDAAN and R can be downloaded from Online Codebook on the NCES website. https://nces.ed.gov/onlinecodebook. Alternately, to find the data from the NCES homepage:

1. Go to nces.ed.gov
2. Type "Online Codebook" into the Search box.

How to obtain permission to use the RESTRICTED-USE data:
Procedures for applying for an NCES RESTRICTED-USE license can be found on the NCES homepage:

1. Go to nces.ed.gov
2. Type "RESTRICTED-USE" in the Search box:  
   https://nces.ed.gov/statprog/instruct.asp

Constraints surrounding access to RESTRICTED-USE data (e.g., citizenship requirements, cost, etc):
While there are no citizenship requirements for access, currently RESTRICTED-USE data can only be accessed at institutions within the United States.

Agency policy re. sample sizes/publishable estimates/data suppression standards:
Estimates can be reported if based on a numerator size of at least 3 and a denominator size of at least 30. Unweighted sample sizes based on restricted data must be rounded to the nearest 10.

Information on when the data were collected/what time period could be examined using the data:
Data have been collected at 6 time points over 12 years: the spring of the 2001–02 school year, when all of the students were in 10th grade; the spring of the 2003–04 school year, when most of the students were in 12th grade (new students added to “freshen” sample to be representative of students who had been in grades other than 10th in 2002, or outside the US); the spring through fall of 2006, when most of the students were 2 years out of high school; and the summer through fall of 2012 when most students were 8 years out of high school. High school transcripts were collected in 2004 and postsecondary transcripts were collected in 2013–2014.

Where to find disseminated official estimates:
To find this information from the NCES homepage:

1. Go to nces.ed.gov
2. Type “education longitudinal study” into the Search box.

ELS:2002 data are included in ELS-specific reports (see https://nces.ed.gov/pubsearch/getpubcats.asp?sid=107), the Condition of Education (https://nces.ed.gov/programs/coe/), and the Digest of Education Statistics (https://nces.ed.gov/programs/digest/). However, estimates disaggregated for the AANHPI subgroups have not been released in any NCES publications.

Type of data (survey, program, admin etc.): Surveys, direct assessments, admissions test scores, high school transcripts, postsecondary transcripts, financial aid records
Do the data include measures of citizenship, or are respondents limited by citizenship (i.e., only U.S. citizens are included in the data)? No information about citizenship status is collected in this study.

Analytic Considerations:
Though Native Hawaiians and Other Pacific Islanders are included in ELS:2002, information allowing for identification of specific subgroup was not collected. While estimates can be reported for Native Hawaiian and Pacific Islanders as one group in the first round of data collection, in later rounds these data are often combined for reporting with other groups due to study attrition and, therefore, small sample sizes.

XIV. High School Longitudinal Study of 2009 (HSLS:09)
The High School Longitudinal Study of 2009 (HSLS:09) is a nationally representative, longitudinal study of more than 20,000 ninth-graders in about 940 schools who are being followed through their secondary and postsecondary years. The main study students were re-surveyed in 2012 when most were high school eleventh-graders. The study focuses on understanding students’ trajectories from the beginning of high school into postsecondary education or the workforce and beyond. What students decide to pursue when, why, and how are crucial questions for HSLS:09, especially, but not solely, in regard to science, technology, engineering, and math (STEM) courses, majors, and careers.

Sponsoring agency: National Center for Education Statistics, Institute of Education Sciences, U.S. Department of Education

Data universe:
2009 Ninth-graders

Disaggregated Asian and Native Hawaiian or Other Pacific Islander categories collected:
Chinese, Filipino, South Asian, Southeast Asian

Are disaggregated Asian and Native Hawaiian or Other Pacific Islander data publicly available? NO

Measures available:
Age, Behavioral Risk Factors, Disability, Education, Elementary and Secondary Education, Families and Living Arrangements, Geography, Homeownership, Housing, Income, Mental Health, Migration and Nativity, Postsecondary Education and Beyond, Poverty, Sex and Gender, Spending and Time Use, Unemployment, Veterans, Voting and Registration, Employment Characteristics
Sample size and unit of analysis:
Nationally representative sample of about 23,000 ninth-graders followed through their high school years and beyond.

Types of geography available to researchers: Census region, school locale (city, suburb, town, rural), school county and state FIPS codes, school ZIP code and Census tract, home ZIP code and Census block/tract. Only school locale is available in the PUBLIC-USE files. All other data are in the RESTRICTED-USE files.

What data are available?
PUBLIC-USE data? YES
RESTRICTED-USE data? YES

What is the difference between PUBLIC-USE and RESTRICTED-USE data files:
RESTRICTED-USE files include data that would be considered sensitive or potentially disclosive in combination with other data due to rarity or small sample sizes. For example, most disability-related data and geographic identifiers are included only in RESTRICTED-USE files. In addition, variables in RESTRICTED-USE files typically contain data at a finer level of detail than the PUBLIC-USE files. For example, specific primary home language is reported in the RESTRICTED-USE files but provided in categories (English, not English) in the PUBLIC-USE file. Also, the RESTRICTED-USE version of the file includes state-representative data for public school students in 10 states.

How to access the PUBLIC-USE data:
ASCII data files and syntax programs for SAS, Stata, SPSS, SUDAAN and R can be downloaded from the Online Codebook on the NCES website: https://nces.ed.gov/onlinecodebook. Alternately, to find the data from the NCES homepage:

1. Go to nces.ed.gov
2. Type “Online Codebook” into the Search box.

How to obtain permission to use the RESTRICTED-USE data:
Procedures for applying for an NCES RESTRICTED-USE license can be found on the NCES homepage:

1. Go to nces.ed.gov
2. Type “RESTRICTED-USE” in the Search box: https://nces.ed.gov/statprog/instruct.asp

Constraints surrounding access to RESTRICTED-USE data (e.g., citizenship requirements, cost, etc):
While there are no citizenship requirements for access, currently RESTRICTED-USE data can only be accessed at institutions within the United States.

**Agency policy re. sample sizes/publishable estimates/data suppression standards:**
Estimates can be reported if based on a numerator size of at least 3 and a denominator size of at least 30. Unweighted sample sizes based on restricted data must be rounded to the nearest 10.

**Information on when the data were collected/what time period could be examined using the data:**
Data have been collected at 7 time points over 9 years: the fall of the 2009–10 school year, when all of the students were in 9th grade; the spring of the 2011–12 school year, when most of the students were in 11th grade; the summer through fall of 2013, when most of the students had left or completed high school; and the spring through fall of 2016, when most of the students were 3 years out of high school. High school transcripts were collected in 2013 and postsecondary transcripts and student records were collected in 2017–18. A postsecondary data match was undertaken in 2021.

**Where to find disseminated official estimates:**
To find this information from the NCES homepage:

1. Go to [nces.ed.gov](https://nces.ed.gov)
2. Type “high school longitudinal study” into the Search box.


**Type of data (survey, program, admin etc.):** Surveys, direct assessments, admissions test scores, high school transcripts, postsecondary transcripts, financial aid records

**Do the data include measures of citizenship, or are respondents limited by citizenship (i.e., only U.S. citizens are included in the data)?** No information about citizenship status is collected in this study.

**Analytic Considerations:**
Though Native Hawaiians and Other Pacific Islanders are included in HSLS:09, information allowing for identification of specific subgroup was not collected. While estimates can be reported for Native Hawaiian and Pacific Islanders as one group in the first round of data collection, in later rounds these data are often combined for reporting with other groups due to study attrition and, therefore, small sample sizes.
XV. Household Pulse Survey (HPS)

The Household Pulse Survey is a 20-minute online survey studying how the coronavirus pandemic is impacting households across the country from a social and economic perspective. The survey asks questions about childcare, education, employment, energy use, food security, health, housing, household spending, Child Tax Credit payments, and intention to receive a COVID-19 vaccination during the coronavirus pandemic.

Sponsoring agency: US Census Bureau

Data universe:
The HPS sampling frame consisted of all housing units on the Master Address File (MAF) where at least one email address or cell phone number was known.

Disaggregated Asian and Native Hawaiian or Other Pacific Islander categories collected:
Asian Indian, Chinese, Filipino, Guamanian or Chamorro, Hawaiian, Japanese, Korean, Native Hawaiian, Pacific Islander, Samoan, Vietnamese

Are disaggregated Asian and Native Hawaiian or Other Pacific Islander data publicly available? NO

Measures available:
Age, Assistance Programs, Behavioral Risk Factors, Community and Economic Development, Disability, Education, Elementary and Secondary Education, Employment Characteristics, Families and Living Arrangements, Food and Nutrition, Health, Health Care and Insurance, Health Insurance Coverage and Access, Homeownership, Housing, Mental Health, Postsecondary Education and Beyond, Sex and Gender, Spending and Time Use, Transportation and Commuting, Unemployment, Veterans

Sample size and unit of analysis:
The HPS is an experimental data product. Sample sizes were determined such that a three percentage coefficient of variation (CV) for an estimate of 40 percent of the population would be achieved for all sample areas with the exception of the 11 smallest states. In these smaller states, the sample size was reduced to produce a 3.5 percent CV. The overall sample sizes within the sampling areas were adjusted for an anticipated response rate of nine percent. For those counties in one of the top MSAs, the sampling interval was adjusted to select the higher of the sampling rate for either the state or MSA. Approximately 1,043,000 housing units were selected from the sampling frame for the first collection period of Phase 3.2. Of those sampled, approximately 66,000 respondents answered the online questionnaire.
Types of geography available to researchers: MSA, state, national

What data are available?
PUBLIC-USE data? YES
RESTRICTED-USE data? YES

What is the difference between PUBLIC-USE and RESTRICTED-USE data files:
PUBLIC-USE data include top- and bottom-coded values, and some data are suppressed due to small cell size. RESTRICTED-USE data include all collected data.

How to access the PUBLIC-USE data:
1. Go to census.gov
2. Select Explore Data, and then Experimental Data Products from the drop-down menu.
3. HPS is currently listed on that page. Click the link; the HPS front page includes a link to the PUBLIC-USE data. https://www.census.gov/programs-surveys/household-pulse-survey/datasets.html

How to obtain permission to use the RESTRICTED-USE data:
The Census Bureau considers proposals from qualified researchers in social science disciplines consistent with the subject matter of the surveys and censuses collected by the Census Bureau. Proposals may be submitted at any time.

1. Go to census.gov
2. Select "Surveys and Programs", and "All surveys and programs."
3. Select "Center for Economic Studies," and then navigate to "Data" (on the left side of the CES page).
4. Select “RESTRICTED-USE Data”, and then “Apply for access”: https://www.census.gov/about/adrm/ced/restricted-use-data.html

Constraints surrounding access to RESTRICTED-USE data (e.g., citizenship requirements, cost, etc):
All researcher access to RESTRICTED-USE data occurs at secure Federal Statistical Research Data Centers (FSRDC). Some FSRDCs may charge fees for access to their facilities.

Agency policy re. sample sizes/publishable estimates/data suppression standards:
Throughout all processes associated with estimation, unauthorized release of protected information or administratively restricted information must be prevented by following federal laws (e.g., Title 13, Title 15, and Title 26), Census Bureau policies (e.g., Data Stewardship Policies), and additional provisions governing the use of the data (e.g., as may be specified in a memorandum of
understanding or data-use agreement). (Additional information available in Statistical Quality Standard S1, Protecting Confidentiality.)

**Information on when the data were collected/what time period could be examined using the data:**

The HPS started collecting data in April 2020 in multiple phases of data collection that mirrored real-world events of the COVID-19 pandemic.

**Where to find disseminated official estimates:**

1. Go to census.gov
2. Select Explore Data, and then Experimental Data Products from the drop-down menu.
3. HPS is listed on that page. Click the link; the HPS front page includes links to both data tables and an interactive data tool https://www.census.gov/data/experimental-data-products/household-pulse-survey.html

**Type of data (survey, program, admin etc.):** Experimental survey.

**Do the data include measures of citizenship, or are respondents limited by citizenship (i.e., only U.S. citizens are included in the data)?** NO

**Analytic Considerations:**

The Census Bureau is fielding the Household Pulse Survey as a part of the agency’s Experimental Data Series; as such, data products may not meet some of the Census Bureau’s statistical quality standards. Data are subject to suppression based on overall response and disclosure avoidance thresholds.

---

**XVI. National Postsecondary Student Aid Study (NPSAS)**

The National Postsecondary Student Aid Study (NPSAS) is a cross-sectional, nationally representative sample survey of undergraduate and graduate students enrolled for a given academic year in institutions eligible to participate in federal financial aid programs. The primary purpose of NPSAS is to measure how students and their families pay for postsecondary education at Title IV eligible institutions, with particular emphasis on federal student aid provided through Title IV of the Higher Education Act.

NPSAS has been conducted every 3 to 4 years since 1987. Student interviews and administrative records are used to provide details concerning student financial aid. Data for each NPSAS collection are released in individual data files and made available in an online data analysis tool.
**Sponsoring agency:** National Center for Health Statistics, Centers for Disease Control and Prevention, U.S. Department of Health and Human Services

**Data universe:**
Undergraduate and graduate students enrolled in postsecondary education in the given academic year

**Disaggregated Asian and Native Hawaiian or Other Pacific Islander categories collected:**
Asian Indian, Chinese, Filipino, Japanese, Korean, Vietnamese

**Are disaggregated Asian and Native Hawaiian or Other Pacific Islander data publicly available?** YES

**Measures available:**
Age, Assistance Programs, Disability, Education, Employment Characteristics, Families and Living Arrangements, Geography, Health, Homeownership, Income, Migration and Nativity, Postsecondary Education and Beyond, Sex and Gender, Veterans

**Sample size and unit of analysis:**
NPSAS:16 - Nationally representative sample of approximately 89,000 undergraduate students and 24,000 graduate students enrolled during the 2015–16 academic year.

NPSAS:12 - Nationally representative sample of approximately 95,000 undergraduate students and 16,000 graduate students enrolled during the 2011–12 academic year.

NPSAS:08 - Nationally representative sample of approximately 113,500 undergraduate students and 14,200 graduate students enrolled during the 2007–08 academic year. NPSAS:08 was also designed to provide state-level estimates for undergraduate students within four institutional strata—public 2-year institutions, public 4-year institutions, private not-for-profit 4-year institutions, and private for-profit institutions— for six states: California, Georgia, Illinois, New York, Minnesota, and Texas.

NPSAS:04 - Nationally representative sample of approximately 79,900 undergraduate students and 10,900 graduate students enrolled during the 2003–04 academic year. NPSAS:04 was also designed to provide state-level estimates for undergraduate students within three institutional strata—public 2-year institutions, public 4-year institutions, and private not-for-profit 4-year institutions— for 12 states that were categorized into three groups based on population size—four large, four medium, and four small: California, Connecticut, Delaware, Georgia, Illinois, Indiana, Maine, Nebraska, New York, Oregon, Tennessee, and Texas.

NPSAS:2000 - Nationally representative sample of approximately 50,000 undergraduate students and 12,000 graduate students enrolled during the 1999–2000 academic year.
NPSAS:96 - Nationally representative sample of approximately 41,500 undergraduate students and 7,000 graduate students enrolled during the 1995–96 academic year.

NPSAS:93 - Nationally representative sample of approximately 52,700 undergraduate students and 13,400 graduate students enrolled during the 1992–93 academic year.

NPSAS:90 - Nationally representative sample of approximately 46,800 undergraduate students and 14,300 graduate students enrolled during the 1989–90 academic year.

NPSAS:87 - Nationally representative sample of approximately 34,500 undergraduate students and 8,600 graduate students enrolled during the 1986–87 academic year.

Types of geography available to researchers: This geographic information is available both in the online data tool that provides public access to the data and the case-level restricted-use file: Degree of urbanization of student’s permanent address, Student’s state of legal or permanent residence, Degree of urbanization in which the NPSAS institution is located, Census Division-Geographic region where NPSAS sample institution is located.

The RESTRICTED-USE file also includes the ZIP code of the student’s permanent address in the interview source file.

What data are available?
PUBLIC-USE data? YES
RESTRICTED-USE data? YES

What is the difference between PUBLIC-USE and RESTRICTED-USE data files:
Case-level data are available only in RESTRICTED-USE format. The online data tool allows a researcher to run analyses without accessing the case-level data. The online data tool and the RESTRICTED-USE file include the same derived variables. The RESTRICTED-USE file includes the source data files used to create the derived variables.

How to access the PUBLIC-USE data:
These data are available through PowerStats, an interactive online interface that enables users to generate custom statistical analyses and data visualizations. To access PowerStats,

1. Go to nces.ed.gov and type “DataLab” in the Search box.
2. PowerStats will result: https://nces.ed.gov/datalab.

How to obtain permission to use the RESTRICTED-USE data:
Procedures for applying for an NCES RESTRICTED-USE license can be found on the NCES homepage:
1. Go to nces.ed.gov
2. Type "NCES RESTRICTED-USE license" in the Search box: https://nces.ed.gov/statprog/instruct.asp

Constraints surrounding access to RESTRICTED-USE data (e.g., citizenship requirements, cost, etc):
While there are no citizenship requirements for access, currently RESTRICTED-USE data can only be accessed at institutions within the United States.

Agency policy re. sample sizes/publishable estimates/data suppression standards:
Estimates can be reported if based on a numerator size of at least 3 and a denominator size of at least 30. Unweighted sample sizes based on restricted data must be rounded to the nearest 10.

Information on when the data were collected/what time period could be examined using the data:
For NPSAS:16, data were collected for the 2015–16 academic year.
For NPSAS:12, data were collected for the 2011–12 academic year.
For NPSAS:08, data were collected for the 2007–08 academic year.
For NPSAS:04, data were collected for the 2003–04 academic year.
For NPSAS:2000, data were collected for the 1999–2000 academic year.
For NPSAS:96, data were collected for the 1995–96 academic year.
For NPSAS:93, data were collected for the 1992–93 academic year.
For NPSAS:90, data were collected for the 1989–90 academic year.
For NPSAS:87, data were collected for the 1986–87 academic year.
Because NPSAS is a repeated cross-sectional study, these data can be used to look at population trends from the 1988-87 academic year to the 2015–16 academic year.

Where to find disseminated official estimates:
In published reports on the NCES website: http://nces.ed.gov

NPSAS data are included in NPSAS-specific reports (see https://nces.ed.gov/pubsearch/getpubcats.asp?sid=013 ), the Condition of Education (https://nces.ed.gov/programs/coe/), and the Digest of Education
Statistics (https://nces.ed.gov/programs/digest/). However, estimates disaggregated for the AANHPI subgroups have not been released in any NCES publications.

Estimates can also be found in the DataLab Tables Library: https://nces.ed.gov/datalab/table/library. You can search for tables that include NPSAS.

**Type of data (survey, program, admin etc.):** Survey, institution student records, and administrative data

**Do the data include measures of citizenship, or are respondents limited by citizenship (i.e., only U.S. citizens are included in the data)?** Except for NPSAS:87 and NPSAS:12, NPSAS captures information on students enrolled in postsecondary institutions in the 50 states, Puerto Rico, and the District of Columbia. There was no restriction on citizenship to be included in the sample.
For NPSAS:87 and NPSAS:12, students enrolled in the 50 states and the District of Columbia were included, but students enrolled in postsecondary institutions in Puerto Rico were excluded. There was no restriction on citizenship to be included in the sample.
The datasets include an indicator of whether the sample member is a U.S. citizen. More recent collections also indicate whether noncitizens are a permanent resident or have a visa.

**Analytic Considerations:**
Though Native Hawaiians and Other Pacific Islanders are included in NPSAS, information allowing for identification of specific subgroup was not collected.

### XVII. National Health and Nutrition Examination Survey (NHANES)

NHANES is the principal source of information on the measured health and Nutritional status of the civilian noninstitutionalized population of the United States. The main objective of the NHANES is to monitor the health and Nutritional status of the United States population through the collection and analysis of data on a broad range of health and nutrition topics. A major strength of this survey lies in the collection of physical measurements and laboratory specimens.

**Sponsoring agency:** National Center for Health Statistics, Centers for Disease Control and Prevention, U.S. Department of Health and Human Services

**Data universe:**
The non-institutionalized, civilian U.S. population
Disaggregated Asian and Native Hawaiian or Other Pacific Islander categories collected:
Asian Indian, Chinese, Filipino, Japanese, Korean, Vietnamese

Are disaggregated Asian and Native Hawaiian or Other Pacific Islander data publicly available? NO

Measures available:
Age, Assistance Programs, Behavioral Risk Factors, Disability, Diseases and Conditions, Early Childhood, Education, Elementary and Secondary Education, Employment Characteristics, Families and Living Arrangements, Food and Nutrition, Geography, Health, Health Care and Insurance, Health Care Services, Health Insurance Coverage and Access, Homeownership, Housing, Income, Mental Health, Migration and Nativity, Mortality, Postsecondary Education and Beyond, Poverty, Reproductive Health, Sex and Gender, Veterans

Sample size and unit of analysis:
In the 2017–March 2020 pre-pandemic data set there were 15,560 nationally representative Sample Participant interviews and 14,300 examinations.

Types of geography available to researchers: Geocoded NHANES data are available on the restricted use files via the Research Data Centers.

What data are available?
PUBLIC-USE data? YES
RESTRICTED-USE data? YES

What is the difference between PUBLIC-USE and RESTRICTED-USE data files:
RESTRICTED-USE files include data that would be considered sensitive or potentially pose a disclosure risk in combination with other data due to rarity or small sample sizes. This includes levels of geography, finer levels of race and ethnicity classification, and more detail on other questions and measurements included in the public use file.

How to access the PUBLIC-USE data:
NHANES data files are available on the NHANES website.

1. Go to CDC.gov
2. Select "NCHS"

How to obtain permission to use the RESTRICTED-USE data:
Permission to use NCHS RESTRICTED-USE data is obtain through submission of a research proposal to NCHS’s Research Data Center (RDC). For details, please visit www.cdc.gov/rdc.
Constraints surrounding access to RESTRICTED-USE data (e.g., citizenship requirements, cost, etc):

For access to RESTRICTED-USE data, researchers must have an approved proposal on file, pay user fees, complete training and administrative forms, and submit their output for a disclosure review.

Agency policy re. sample sizes/publishable estimates/data suppression standards:

NCHS staff are required to use the NCHS Data Presentation Standards for Proportions (CDC→NCHS→publications→Vital and Health Statistics→Series 2→No. 175) to publish statistically reliable estimates. Researchers outside of NCHS are highly encouraged to use these standards as well.

Information on when the data were collected/what time period could be examined using the data:

NHANES data have been collected dating back to 1960 as a periodic survey. In 1970 nutrition was added to the survey and the name changed from the National Health Examination Survey to the National Health and Nutrition Examination Survey. Since 1999 it became continuous with data released in 2-year cycles. For measurements included across the surveys, trends can be conducted, in some cases using adjustment equations to adjust for laboratory changes.

Where to find disseminated official estimates:

In numerous published reports on the NCHS website:

1. Go to cdc.gov
2. Type "National Center for Health Statistics" into the Search box
3. Select "NHANES – Publications and Products":
   https://www.cdc.gov/nchs/nhanes/nhanes_products.htm

Type of data (survey, program, admin etc.): Survey

Do the data include measures of citizenship, or are respondents limited by citizenship (i.e., only U.S. citizens are included in the data)? NHANES does not limit participation according to citizenship but has asked about citizenship as part of the survey questionnaire in the past. Currently NHANES does not ask about citizenship.

Analytic Considerations:

Analysis considerations for NHANES data are detailed in survey documentation and analysis guidelines on the CDC website:

1. Go to cdc.gov
2. Select "NCHS"
3. Select "NHANES"
4. Select "Survey Data and Documentation"


XVIII. National Health Interview Survey (NHIS)

The National Health Interview Survey (NHIS) is the principal source of information on the health of the civilian noninstitutionalized population of the United States. The main objective of the NHIS is to monitor the health of the United States population through the collection and analysis of data on a broad range of health topics. A major strength of this survey lies in the ability to categorize these health characteristics by many demographic and socioeconomic characteristics.

**Sponsoring agency:** National Center for Health Statistics, Centers for Disease Control and Prevention, U.S. Department of Health and Human Services

**Data universe:**
The non-institutionalized, civilian U.S. population

**Disaggregated Asian and Native Hawaiian or Other Pacific Islander categories collected:**
Asian Indian, Chinese, Filipino, Guamanian or Chamorro, Japanese, Korean, Vietnamese, Samoan

**Are disaggregated Asian and Native Hawaiian or Other Pacific Islander data publicly available?** NO

**Measures available:**
Age, Assistance Programs, Behavioral Risk Factors, Disability, Diseases and Conditions, Early Childhood, Education, Elementary and Secondary Education, Employment Characteristics, Families and Living Arrangements, Food and Nutrition, Geography, Health, Health Care and Insurance, Health Care Services, Health Insurance Coverage and Access, Homeownership, Income, Injuries, Mental Health, Migration and Nativity, Postsecondary Education and Beyond, Poverty, Reproductive Health, Rural Economy & Population, Sex and Gender, Veterans

**Sample size and unit of analysis:**
In 2020, there were 31,568 nationally representative Sample Adult interviews and 5,790 Sample Child interviews.

**Types of geography available to researchers:** On the public data file, data are available by region and by a 4-level measure of urban/rural status, based on the [2013 NCHS Urban-Rural](https://www.cdc.gov/nchs/products Accessed 30 August 2022) classification system.
**Classification Scheme for Counties.** Geocoded NHIS data are available on the restricted use files via the Research Data Centers.

**What data are available?**
PUBLIC-USE data? YES  
RESTRICTED-USE data? YES

**What is the difference between PUBLIC-USE and RESTRICTED-USE data files:**
RESTRICTED-USE files include data that would be considered sensitive or potentially pose a disclosure risk in combination with other data due to rarity or small sample sizes. This includes finer levels of geography, race and ethnicity classification, and more detail on other questions included in the public use file. More information is available in our RESTRICTED-USE codebooks: Sample Adult ([Restricted use no-frequency codebook (cdc.gov)](https://www.cdc.gov/)) and Sample Child ([Restricted use no-frequency codebook (cdc.gov)](https://www.cdc.gov/)).

**How to access the PUBLIC-USE data:**
Annual PUBLIC-USE data is available on the NCHS/NHIS website:

1. Go to [cdc.gov](https://www.cdc.gov)
2. Search "National Health Interview Study"

**How to obtain permission to use the RESTRICTED-USE data:**
Permission to use NCHS RESTRICTED-USE data is obtain through submission of a research proposal to NCHS’s Research Data Center (RDC). For details, please visit [www.cdc.gov/rdc](https://www.cdc.gov/rdc).

**Constraints surrounding access to RESTRICTED-USE data (e.g., citizenship requirements, cost, etc):**
For access to RESTRICTED-USE data, researchers must have an approved proposal on file, pay user fees, complete training and administrative forms, and submit their output for a disclosure review.

**Agency policy re. sample sizes/publishable estimates/data suppression standards:**
NCHS staff are required to use the [NCHS Data Presentation Standards for Proportions](https://www.cdc.gov/) to publish statistically reliable estimates. Researchers outside of NCHS are highly encouraged to use these standards as well.

**Information on when the data were collected/what time period could be examined using the data:**
NHIS data have been collected annually dating back to 1957. However, the survey undergoes periodic redesigns which prevents comparable evaluation of very long-term trends. Currently data
from 1997–2018 are considered comparable for trending. Starting in 2019, the NHIS underwent a questionnaire redesign and changes in the approach to sampling. Therefore, it is not currently recommended that NHIS data pre-2019 be combined with, or trended with, NHIS data starting in 2019.

**Where to find disseminated official estimates:**
In the NHIS Interactive Data Query System ([https://www.cdc.gov/nchs/nhis/shs.htm](https://www.cdc.gov/nchs/nhis/shs.htm)) and numerous published reports on the NCHS website:

1. Go to [cdc.gov](https://www.cdc.gov)
2. Type "National Health Interview Survey" into the Search box

**Type of data (survey, program, admin etc.):** Survey

**Do the data include measures of citizenship, or are respondents limited by citizenship (i.e., only U.S. citizens are included in the data)?** NHIS does not limit participation according to citizenship but does ask about citizenship as part of the survey questionnaire.

**Analytic Considerations:**
Analysis using NHIS data should be aware of the periodic changes to the survey which impact comparability of trends over time which is detailed in annual survey documentation and on the NHIS website:

1. Go to [cdc.gov](https://www.cdc.gov)
2. Type "National Health Interview Survey" into the Search box

---

**XIX. National Survey of Family Growth**

Includes nationally representative data on family life, marriage/divorce/cohabitation, pregnancy, infertility, use of contraception, and men’s and women’s health.

**Sponsoring agency:** National Center for Health Statistics, Centers for Disease Control and Prevention, U.S. Department of Health and Human Services

**Data universe:**
Reproductive age (15-49 years; 15-44 prior to 2015) population (men and women) in US households.

**Disaggregated Asian and Native Hawaiian or Other Pacific Islander categories collected:**
Asian Indian, Chinese, Filipino, Guamanian or Chamorro, Japanese, Korean, Vietnamese, Samoan, Native Hawaiian

**Are disaggregated Asian and Native Hawaiian or Other Pacific Islander data publicly available?** NO

**Measures available:**
Age, Behavioral Risk Factors, Diseases and Conditions, Early Childhood, Education, Elementary and Secondary Education, Families and Living Arrangements, Health, Health Care and Insurance, Health Care Services, Health Insurance Coverage and Access, Housing, Income, Postsecondary Education and Beyond, Poverty, Reproductive Health, Religion, Sex and Gender, Veterans

**Sample size and unit of analysis:**
Approximately 5,000 men and women aged 15-49 have been interviewed each year during continuous data collection 2006–2010 and 2011–2019. Similar numbers are expected to participate annually in NSFG when data collection resumes in 2022.

**Types of geography available to researchers:** A data file consisting of the case ID, State and County FIPS codes, and State FIPS code of the respondent’s address at the time of the previous Census may only be accessed through an RDC. Block group and tract are also available through an RDC. A list of geographic variables is available here [https://www.cdc.gov/rdc/geocodes/geowt_nsfg.htm](https://www.cdc.gov/rdc/geocodes/geowt_nsfg.htm) These variables may be used to merge external information about place of residence onto the files. Estimates cannot be produced for specific states or counties.

**What data are available?**
PUBLIC-USE data? YES

RESTRICTED-USE data? YES

**What is the difference between PUBLIC-USE and RESTRICTED-USE data files:**
Geographic variables are only available on restricted use data files. A number of NSFG variables have been suppressed or modified on public use files to minimize the risk of disclosure. See Appendix 7 for detailed information ([https://www.cdc.gov/nchs/nsfg/nsfg_2017_2019_puf.htm](https://www.cdc.gov/nchs/nsfg/nsfg_2017_2019_puf.htm)).

**How to access the PUBLIC-USE data:**
Links to data files and related documentation for 1973–2019 are available on the NCHS website:
1. Go to cdc.gov
2. Type "National Survey of Family Growth" into the Search box
3. Select "Questionnaires, Datasets, and Related Documentation":
   https://www.cdc.gov/nchs/nsfg/nsfg_questionnaires.htm

**How to obtain permission to use the RESTRICTED-USE data:**
Permission to use NCHS restricted-use data is obtain through submission of a research proposal to NCHS’s Research Data Center (RDC). For details, please visit www.cdc.gov/rdc

**Constraints surrounding access to RESTRICTED-USE data (e.g., citizenship requirements, cost, etc):**
For access to RESTRICTED-USE data, researchers must have an approved proposal on file, pay user fees, complete training and administrative forms, and submit their output for a disclosure review.

**Agency policy re. sample sizes/publishable estimates/data suppression standards:**
Products must be reviewed by the RDC before submission for publication to ensure there is no disclosure risk for study participants. Publication cannot include information that could be used to identify an individual or establishment, or exact dates.

**Information on when the data were collected/what time period could be examined using the data:**
Six cycles of the NSFG were fielded periodically from 1973 to 2002. A continuous fieldwork design was adopted beginning in 2006. Data collection was paused in 2019 and will resume in 2022. Data files are available for 1973–2019. Data files were released every two years during the continuous fieldwork design period, and this release plan is expected going forward in 2022+ also.

**Where to find disseminated official estimates:**
To find this information from the NCES homepage:

1. Go to cdc.gov
2. Type "National Survey of Family Growth" into the Search box

**Type of data (survey, program, admin etc.):** Survey

**Do the data include measures of citizenship, or are respondents limited by citizenship (i.e., only U.S. citizens are included in the data)?** No citizenship information is obtained in the NSFG.

**Analytic Considerations:**
NA
XX. National Vital Statistics System (NVSS)—Birth Data

Data on all births occurring in the U.S. are collected and compiled through the cooperative activities of state vital records offices and NCHS. Standard forms and procedures are used for the uniform collection of data.

**Sponsoring agency:** National Center for Health Statistics, Centers for Disease Control and Prevention, U.S. Department of Health and Human Services

**Data universe:**
All births occurring in the U.S.

**Disaggregated Asian and Native Hawaiian or Other Pacific Islander categories collected:**
Asian Indian, Chinese, Filipino, Guamanian or Chamorro, Japanese, Korean, Vietnamese, Native Hawaiian, Samoan

**Are disaggregated Asian and Native Hawaiian or Other Pacific Islander data publicly available?** YES

**Measures available:**
Age, Behavioral Risk Factors, Diseases and Conditions, Education, Elementary and Secondary Education, Food and Nutrition, Geography, Health, Health Care and Insurance, Health Care Services, Health Insurance Coverage and Access, Postsecondary Education and Beyond, Reproductive Health, Sex and Gender

**Sample size and unit of analysis:**
Approximately 3.6 million births/year in recent years; one record per birth.

**Types of geography available to researchers:** No geographic information is available on public use data files. Restricted use data files include state and county information.

**What data are available?**
PUBLIC-USE data? YES
RESTRICTED-USE data? YES

**What is the difference between PUBLIC-USE and RESTRICTED-USE data files:**
No geographic information is available on public use data files. Restricted use data files include state and county information.

**How to access the PUBLIC-USE data:**
Links to data files and related documentation for 1968 forward are available on the NCHS website:
1. Go to cdc.gov
2. Type "National Center for Health Statistics" into the Search box
3. Select "Data and Tools"
4. Select "Vital Stats Online - Access to downloadable datasets"; the resulting page includes data guides: https://www.cdc.gov/nchs/data_access/vitalstatsonline.htm
5. Select "Births" for NVSS births data: https://www.cdc.gov/nchs/data_access/Vitalstatsonline.htm#Births

**How to obtain permission to use the RESTRICTED-USE data:**

Access to restricted use data files requires submission of an application and review and approval by NCHS. Users interested in finding out more about the NVSS restricted use access policy can go here: RESTRICTED-USE Vital Statistics Data (cdc.gov)

**Constraints surrounding access to RESTRICTED-USE data (e.g., citizenship requirements, cost, etc):**

Request must demonstrate an intended use that focuses on studies of health and the factors affecting health, and a reasonable likelihood of contributing to knowledge on a health-related issue. Data files cannot be sent to or accessed remotely from outside the US. Data may not be linked to other data sets which may re-identify cases, or be used for commercial or resale purposes. Data files must be stored on, or accessed from, the secure computer system of the researcher's affiliated organization or institution and may not be stored on or accessed from a cloud-based system. Studies that require exact dates may only be conducted at a Federal Statistical Research Data Center (FSRDC). For these requests researchers must have an approved proposal on file, pay user fees, complete training and administrative forms, and submit their output for a disclosures review.; there is a fee involved. Researchers requiring data for only a single state should first request data from that state’s vital records office.


**Agency policy re. sample sizes/publishable estimates/data suppression standards:**

Counts, rates, and percentages for sub-national geographic areas must be suppressed and not displayed in any manner if based on <10 observations in the numerator, denominator, or total, regardless of the number of years combined. This data suppression rule applies to all text, tables, and figures (including maps) contained in main and supplemental files. No data on an identifiable case should be derivable through subtraction or other calculation from the combination of tables in a given publication. No data should permit disclosure when used in combination with other known data.

**Information on when the data were collected/what time period could be examined using the data:**
Data files are available for 1968 to the present. Data files for a calendar year become available approximately 8 months after the close of the year.

**Where to find disseminated official estimates:**
Links to numerous recent reports are available on the NCHS's National Vital Statistics System webpage:

1. Go to [cdc.gov](http://cdc.gov)
2. Type "National Vital Statistics" into the Search box

**Type of data (survey, program, admin etc.):** Complete census of births occurring in the U.S.

**Do the data include measures of citizenship, or are respondents limited by citizenship (i.e., only U.S. citizens are included in the data)?** No citizenship information is available.

**Analytic Considerations:**

**XXI. National Vital Statistics System (NVSS)-Fetal Death Data**

Data on all fetal deaths occurring in the U.S. at 20 weeks gestation or later are collected and compiled through the cooperative activities of state vital records offices and NCHS. Standard forms and procedures are used for the uniform collection of data.

**Sponsoring agency:** National Center for Health Statistics, Centers for Disease Control and Prevention, U.S. Department of Health and Human Services

**Data universe:**
All fetal deaths at 20+ weeks gestation occurring in the U.S.

**Disaggregated Asian and Native Hawaiian or Other Pacific Islander categories collected:**
Asian Indian, Chinese, Filipino, Guamanian or Chamorro, Japanese, Korean, Vietnamese, Native Hawaiian, Samoan

**Are disaggregated Asian and Native Hawaiian or Other Pacific Islander data publicly available?** YES

**Measures available:**
Age, Behavioral Risk Factors, Diseases and Conditions, Education, Elementary and Secondary Education, Geography, Health, Injuries, Mortality, Postsecondary Education and Beyond, Reproductive Health, Sex and Gender

**Sample size and unit of analysis:**
21,000+ fetal deaths per year in recent years; single record for each death.

**Types of geography available to researchers:** No geographic information is available on public use data files. Restricted use data files include state and county information.

**What data are available?**
PUBLIC-USE data? YES
RESTRICTED-USE data? YES

**What is the difference between PUBLIC-USE and RESTRICTED-USE data files:**
No geographic information is available on public use data files. Restricted use data files include state and county information.

**How to access the PUBLIC-USE data:**
Links to data files and related documentation for 1982 forward are available on the NCHS website:

1. Go to [cdc.gov](http://cdc.gov)
2. Type "National Center for Health Statistics" into the Search box
3. Select "Data and Tools"
4. Select "Vital Stats Online – Access to downloadable datasets"; the resulting page includes data guides: [https://www.cdc.gov/nchs/data_access/vitalstatsonline.htm](https://www.cdc.gov/nchs/data_access/vitalstatsonline.htm)
5. Click on "Fetal Death" for NVSS fetal death data: [https://www.cdc.gov/nchs/data_access/vitalstatsonline.htm#Fetal_Death](https://www.cdc.gov/nchs/data_access/vitalstatsonline.htm#Fetal_Death)

**How to obtain permission to use the RESTRICTED-USE data:**
Access to restricted use data files requires submission of an application and review and approval by NCHS. Users interested in finding out more about the NVSS restricted use access policy can go here: [RESTRICTED-USE Vital Statistics Data (cdc.gov)](http://cdc.gov)
Constraints surrounding access to RESTRICTED-USE data (e.g., citizenship requirements, cost, etc):

Request must demonstrate an intended use that focuses on studies of health and the factors affecting health, and a reasonable likelihood of contributing to knowledge on a health-related issue. Data files cannot be sent to, or accessed remotely from outside the U.S. Data may not be linked to other data sets which may re-identify cases, or be used for commercial or resale purposes. Data files must be stored on, or accessed from, the secure computer system of the researcher's affiliated organization or institution and may not be stored on or accessed from a cloud-based system. Studies that require exact dates may only be conducted at a Federal Statistical Research Data Center (FSRDC). For these requests researchers must have an approved proposal on file, pay user fees, complete training and administrative forms, and submit their output for a disclosures review; there is a fee involved. Researchers requiring data for only a single state should first request data from that state's vital records office. See: https://www.cdc.gov/nchs/nvss/nvss-restricted-data.htm and www.cdc.gov/rdc for more information.

Agency policy re. sample sizes/publishable estimates/data suppression standards:

Counts, rates, and percentages for sub-national geographic areas must be suppressed and not displayed in any manner if based on <10 observations in the numerator, denominator, or total, regardless of the number of years combined. This data suppression rule applies to all text, tables, and figures (including maps) contained in main and supplemental files. No data on an identifiable case should be derivable through subtraction or other calculation from the combination of tables in a given publication. No data should permit disclosure when used in combination with other known data.

Information on when the data were collected/what time period could be examined using the data:

Data files are available for 1982 to the present. Data files for a calendar year become available approximately 12 months after the close of the year.

Where to find disseminated official estimates:

Links to numerous recent reports are available on the NCHS's National Vital Statistics System webpage:

- Go to cdc.gov

Type of data (survey, program, admin etc.): Complete census of fetal deaths occurring in the U.S.
Do the data include measures of citizenship, or are respondents limited by citizenship (i.e., only U.S. citizens are included in the data)? No citizenship information is available.

Analytic Considerations:
Data quality varies among data items.

XXII. National Vital Statistics System (NVSS)-Linked Birth/Infant Death Data

Data on all infant deaths occurring in the U.S. are linked with corresponding birth records through the cooperative activities of state vital records offices and NCHS. Standard forms and procedures are used for the uniform collection of data.

Sponsoring agency: National Center for Health Statistics, Centers for Disease Control and Prevention, U.S. Department of Health and Human Services

Data universe:
All infant deaths occurring in the U.S. linked with corresponding birth data.

Disaggregated Asian and Native Hawaiian or Other Pacific Islander categories collected:
Asian Indian, Chinese, Filipino, Guamanian or Chamorro, Japanese, Korean, Vietnamese, Native Hawaiian, Samoan

Are disaggregated Asian and Native Hawaiian or Other Pacific Islander data publicly available? YES

Measures available:
Age, Behavioral Risk Factors, Diseases and Conditions, Education, Elementary and Secondary Education, Food and Nutrition, Geography, Health, Health Care and Insurance, Health Care Services, Health Insurance Coverage and Access, Injuries, Mortality, Postsecondary Education and Beyond, Reproductive Health, Sex and Gender

Sample size and unit of analysis:
Approximately 21,000 infant deaths/year in recent years; one record per infant death.

Types of geography available to researchers: No geographic information is available on public use data files. Restricted use data files include state and county information.

What data are available?
PUBLIC-USE data? YES
RESTRICTED-USE data? YES

What is the difference between PUBLIC-USE and RESTRICTED-USE data files:
No geographic information is available on public use data files. Restricted use data files include state and county information.

How to access the PUBLIC-USE data:
Links to data files and related documentation for period linked birth-infant death files are available for 1995 forward, and data and user guides for birth cohort linked birth-death files for 1983–2015 are available on the NCHS website:

1. Go to cdc.gov
2. Type "National Center for Health Statistics" into the Search box
3. Select "Data and Tools"
4. Select "Vital Stats Online - Access to downloadable datasets"; the resulting page includes data guides:
   https://www.cdc.gov/nchs/data_access/vitalstatsonline.htm
5. Click on "Period Linked Birth - Infant Death" for NVSS birth-infant death data:
   https://www.cdc.gov/nchs/data_access/vitalstatsonline.htm#Period_Linked

How to obtain permission to use the RESTRICTED-USE data:
Access to restricted use data files requires submission of an application and review and approval by NCHS. Users interested in finding out more about the NVSS restricted use access policy can go here: RESTRICTED-USE Vital Statistics Data (cdc.gov)

Constraints surrounding access to RESTRICTED-USE data (e.g., citizenship requirements, cost, etc):
Request must demonstrate an intended use that focuses on studies of health and the factors affecting health, and a reasonable likelihood of contributing to knowledge on a health-related issue. Data files cannot be sent to, or accessed remotely from outside the U.S. Data may not be linked to other data sets which may re-identify cases, or be used for commercial or resale purposes. Data files must be stored on, or accessed from the secure computer system of the researcher’s affiliated organization or institution and may not be stored on or accessed from a cloud-based system. Studies that require exact dates may only be conducted at a Federal Statistical Research Data Center (FSRDC). For these requests researchers must have an approved proposal on file, pay user fees, complete training and administrative forms, and submit their output for a disclosures review; there is a fee involved. Researchers requiring data for only a single state should first request data from that state’s vital records office.
**Agency policy re. sample sizes/publishable estimates/data suppression standards:**
Counts, rates, and percentages for sub-national geographic areas must be suppressed and not displayed in any manner if based on <10 observations in the numerator, denominator, or total, regardless of the number of years combined. This data suppression rule applies to all text, tables, and figures (including maps) contained in main and supplemental files. No data on an identifiable case should be derivable through subtraction or other calculation from the combination of tables in a given publication. No data should permit disclosure when used in combination with other known data.

**Information on when the data were collected/what time period could be examined using the data:**
Data files are available for 1968 to the present. Data files for a calendar year become available approximately 8 months after the close of the year.

**Where to find disseminated official estimates:**
Links to numerous recent reports are available on the NCHS’s National Vital Statistics System webpage:

- Go to [cdc.gov](https://www.cdc.gov)

**Type of data (survey, program, admin etc.):** Complete census infant deaths occurring in the U.S.

**Do the data include measures of citizenship, or are respondents limited by citizenship (i.e., only U.S. citizens are included in the data)?** No citizenship information is available.

**Analytic Considerations:**
Data quality varies among data items.

---

**XXIII. National Vital Statistics System (NVSS)-Mortality Data**

Data on all deaths occurring in the U.S. are collected and compiled through the cooperative activities of state vital records offices and NCHS. Standard forms and procedures are used for the uniform collection of data.

**Sponsoring agency:** National Center for Health Statistics, Centers for Disease Control and Prevention, U.S. Department of Health and Human Services

**Data universe:**
All deaths occurring in the U.S.

**Disaggregated Asian and Native Hawaiian or Other Pacific Islander categories collected:**
Asian Indian, Chinese, Filipino, Guamanian or Chamorro, Japanese, Korean, Vietnamese, Native Hawaiian, Samoan

**Are disaggregated Asian and Native Hawaiian or Other Pacific Islander data publicly available?** YES

**Measures available:**
Age, Diseases and Conditions, Education, Elementary and Secondary Education, Employment Characteristics, Geography, Health, Injuries, Mortality, Postsecondary Education and Beyond, Sex and Gender, Veterans

**Sample size and unit of analysis:**
Approximately 2.9 million deaths/year in recent years; one record per death.

**Types of geography available to researchers:** No geographic information is available on public use data files. Restricted use data files include state and county information.

**What data are available?**
PUBLIC-USE data? YES

RESTRICTED-USE data? YES

**What is the difference between PUBLIC-USE and RESTRICTED-USE data files:**
No geographic information is available on public use data files. Restricted use data files include state and county information.

**How to access the PUBLIC-USE data:**
Links to data files and related documentation for 1968 forward are available on the NCHS website:

1. Go to [cdc.gov](http://cdc.gov)
2. Type "National Center for Health Statistics" into the Search box
3. Select "Data and Tools"
4. Select "Vital Stats Online - Access to downloadable datasets"; the resulting page includes data guides: [https://www.cdc.gov/nchs/data_access/vitalstatsonline.htm](https://www.cdc.gov/nchs/data_access/vitalstatsonline.htm)
5. Click "Mortality Multiple Cause" for NVSS mortality data: [https://www.cdc.gov/nchs/data_access/vitalstatsonline.htm#Mortality_Multiple](https://www.cdc.gov/nchs/data_access/vitalstatsonline.htm#Mortality_Multiple)

**How to obtain permission to use the RESTRICTED-USE data:**
Access to restricted use data files requires submission of an application and review and approval by NCHS. Users interested in finding out more about the NVSS restricted use access policy can go here: RESTRICTED-USE Vital Statistics Data (cdc.gov)

**Constraints surrounding access to RESTRICTED-USE data (e.g., citizenship requirements, cost, etc):**

Request must demonstrate an intended use that focuses on studies of health and the factors affecting health, and a reasonable likelihood of contributing to knowledge on a health-related issue. Data files cannot be sent to, or accessed remotely from outside the U.S. Data may not be linked to other data sets which may re-identify cases, or be used for commercial or resale purposes. Data files must be stored on, or accessed from the secure computer system of the researcher's affiliated organization or institution and may not be stored on or accessed from a cloud-based system. Studies that require exact dates may only be conducted at a Federal Statistical Research Data Center (FSRDC). For these requests researchers must have an approved proposal on file, pay user fees, complete training and administrative forms, and submit their output for a disclosures review.; there is a fee involved. Researchers requiring data for only a single state should first request data from that state’s vital records office.

See: [https://www.cdc.gov/nchs/nvss/nvss-restricted-data.htm](https://www.cdc.gov/nchs/nvss/nvss-restricted-data.htm) and [www.cdc.gov/rdc](http://www.cdc.gov/rdc) for more information.

**Agency policy re. sample sizes/publishable estimates/data suppression standards:**

Counts, rates, and percentages for sub-national geographic areas must be suppressed and not displayed in any manner if based on <10 observations in the numerator, denominator, or total, regardless of the number of years combined. This data suppression rule applies to all text, tables, and figures (including maps) contained in main and supplemental files. No data on an identifiable case should be derivable through subtraction or other calculation from the combination of tables in a given publication. No data should permit disclosure when used in combination with other known data.

**Information on when the data were collected/what time period could be examined using the data:**

Data files are available for 1968 to the present. Data files for a calendar year become available approximately 12 months after the close of the year.

**Where to find disseminated official estimates:**

Links to numerous recent reports are available on the NCHS’s National Vital Statistics System webpage:

1. Go to [cdc.gov](http://cdc.gov)
Type of data (survey, program, admin etc.): Complete census of deaths occurring in the U.S.

Do the data include measures of citizenship, or are respondents limited by citizenship (i.e., only U.S. citizens are included in the data)? No citizenship information is available.

Analytic Considerations:
Data quality varies among data items.

XXIV. Survey of Income and Program Participation (SIPP)

The Survey of Income and Program Participation (SIPP) is a household-based survey designed as a continuous series of national panels. Each panel generally features a large sample of households that are interviewed multiple times over a four-year period. SIPP data provide the most comprehensive information available on how the nation’s economic well-being changes over time, which has been SIPP’s defining characteristic since its inception in 1983.

Sponsoring agency: US Census Bureau

Data universe:
The SIPP sample is a multistage-stratified sample of the U.S. civilian non-institutionalized population.

Disaggregated Asian and Native Hawaiian or Other Pacific Islander categories collected:
Asian Indian, Chinese, Filipino, Guamanian or Chamorro, Hawaiian, Japanese, Korean, Native Hawaiian, Samoan, Vietnamese

Are disaggregated Asian and Native Hawaiian or Other Pacific Islander data publicly available? NO

Measures available:
Age, Assistance Programs, Business and Economy, Disability, Diseases and Conditions, Early Childhood, Education, Elementary and Secondary Education, Employment Characteristics, Families and Living Arrangements, Food and Nutrition, Health, Health Care and Insurance, Health Insurance Coverage and Access, Homeownership, Housing, Income, Migration and Nativity, Postsecondary Education and Beyond, Poverty, Sex and Gender, Transportation and Commuting, Unemployment, Veterans

Sample size and unit of analysis:
The SIPP survey design is a continuous series of national panels, with sample size ranging from approximately 14,000 to 52,000 interviewed households. The duration of each panel ranges from 2 ½ years to 4 years. The SIPP sample is a multistage-stratified sample of the U.S. civilian non-institutionalized population.

**Types of geography available to researchers:** National, region, state, metro/nonmetro

**What data are available?**

PUBLIC-USE data? YES

RESTRICTED-USE data? YES

**What is the difference between PUBLIC-USE and RESTRICTED-USE data files:**

PUBLIC-USE data include top- and bottom-coded values, and some data are suppressed due to small cell size. RESTRICTED-USE data include all collected data.

**How to access the PUBLIC-USE data:**

1. Go to census.gov
2. Select the "Surveys & Programs" tab, and select "Survey of Income and Program Participation (SIPP)" from the drop-down.
3. Next select "Data" from the left-hand navigation bar: [https://www.census.gov/programs-surveys/sipp/data.html](https://www.census.gov/programs-surveys/sipp/data.html)

**How to obtain permission to use the RESTRICTED-USE data:**

The Census Bureau considers proposals from qualified researchers in social science disciplines consistent with the subject matter of the surveys and censuses collected by the Census Bureau. Proposals may be submitted at any time.

1. Go to census.gov
2. Select "Surveys and Programs", and "All surveys and programs."
3. Select "Center for Economic Studies," and then navigate to "Data" (on the left side of the CES page).
4. Select "RESTRICTED-USE Data", and then "Apply for access": [https://www.census.gov/about/adrm/ced/apply-for-access.html](https://www.census.gov/about/adrm/ced/apply-for-access.html)

**Constraints surrounding access to RESTRICTED-USE data (e.g., citizenship requirements, cost, etc):**

All researcher access to RESTRICTED-USE data occurs at secure Federal Statistical Research Data Centers (FSRDC). Some FSRDCs may charge fees for access to their facilities.

**Agency policy re. sample sizes/publishable estimates/data suppression standards:**
The Census Bureau recommends a minimum weighted estimate - i.e., the weighted totals of the sample cases included in the analysis - of 200,000 for computing cross-sectional estimates and 300,000 for computing longitudinal estimates. Nonsampling error in one or more of the small number of cases providing the estimation can also cause large relative error in that particular estimate. Care must be taken in the interpretation of small differences since even a small amount of nonsampling error can cause a borderline difference to appear significant or not, thus distorting a seemingly valid hypothesis test. https://www2.census.gov/programs-surveys/sipp/tech-documentation/source-accuracy-statements/2019/sipp-2019-SA-24-SEP21.pdf

Information on when the data were collected/what time period could be examined using the data:
The first SIPP panel, the 1984 Panel, began interviews in October 1983. The 1985 Panel began in February 1985. Subsequent panels began in February of each calendar year, resulting in concurrent administration of the survey in multiple panels. In several instances, due to budget and other issues, panels were terminated after seven interviews (28 months). Three panels were terminated even earlier: 1988 (six interviews), 1989 (three interviews), and 2000 (2 interviews). With certain exceptions, each panel overlapped part of the previous panel, with the result that there were two or three active panels at any given time. Overlapping panels allow data users to pool records from different panels, thus having larger samples (and lower standard errors) for cross-sectional analyses. The overlapping feature of the SIPP design was dropped for the 1996 through 2014 Panels, but standard errors for non-overlapping panels remained small since at least 29,000 households were interviewed (considerably more than in the previous panels). Similar to the 1990 through 1993 Panels, the 2018 Panels and beyond overlap with new panels that begin each subsequent year (2019, 2020, and so on).

Where to find disseminated official estimates:

1. Go to census.gov
2. Select the "Surveys & Programs" tab, and select "Survey of Income and Program Participation (SIPP)" from the drop-down.
3. Recent publications are currently listed at the bottom of the page. https://www.census.gov/programs-surveys/sipp.html

For older publications and table packages, select the "Library" tab on the left-hand side of the page. https://www.census.gov/programs-surveys/sipp/library.html

Type of data (survey, program, admin etc.): Survey

Do the data include measures of citizenship, or are respondents limited by citizenship (i.e., only U.S. citizens are included in the data)?
Nativity and citizenship information were collected in the Migration topical module prior to the 2014 panel, and in the core survey in 2014 and beyond. There is no citizenship limitation on survey respondents residing in sampled households.

**Analytic Considerations:** NONE