

**BEST PRACTICES  
FOR THE DISAGGREGATION  
OF FEDERAL DATA ON  
ASIAN AMERICANS  
AND PACIFIC ISLANDERS**



**WHITE HOUSE INITIATIVE  
ON ASIAN AMERICANS  
AND PACIFIC ISLANDERS  
INTERAGENCY WORKING GROUP  
DATA AND RESEARCH SUBCOMMITTEE**

**March 2016**



## INTRODUCTION

In recent years, federal statistical agencies have seen an increase in the demand for data on Asian Americans and Pacific Islander (AAPI) subgroups. As the AAPI population has grown, it has become increasingly clear that the varied AAPI ethnicities differ in socioeconomic status, health status, educational attainment, labor force characteristics, and many others. For example, Asian Indians tend to have higher educational attainment than Samoans, and Japanese have lower unemployment rates than Filipinos.

In order to provide more accurate and meaningful information on the AAPI community to policy makers, the press, and the public, federal statistical agencies should attempt to provide disaggregated AAPI data—that is, data by individual AAPI ethnicities—to the extent possible.

However, providing quality, disaggregated data on the AAPI population can be a challenge for federal agencies. To explore this issue and potential solutions, the White House Initiative on Asian Americans and Pacific Islanders Interagency Working Group established a subcommittee of key federal agencies to discuss data disaggregation. The Data and Research Subcommittee developed this document, which contains an overview of best practices for providing disaggregated AAPI data.



## Three major areas present frequent challenges when providing disaggregated AAPI data from federal surveys and administrative records:

- Data collection
- Data analysis and evaluation
- Data access and dissemination

For each area, best practices are presented for common challenges faced by federal surveys and administrative records.



### Data Collection Best Practices Highlight:

*The American Community Survey Language Assistance Program,  
U.S. Census Bureau*

The language assistance program for the American Community Survey (ACS) includes a set of methods and procedures designed to assist sample households with limited English proficiency in completing the ACS interview. The ACS language assistance program includes the use of several key tools to support each mode of data collection—mail, telephone, and personal visit. For example, the ACS hired bilingual staff that can conduct telephone interviews and personal visit interviews in Chinese, Japanese, Korean, Tagalog, Urdu, and Vietnamese. Additionally, informational material about the survey is available in Chinese, Korean, and Vietnamese, which reflect the ACS's greatest language needs other than Spanish.

Providing language assistance is one of many ways that the ACS has improved nonresponse bias and reduced the introduction of response errors; it ensures that individuals with limited English skills will more fully understand the survey questions. In addition, the ACS collects and provides data on English proficiency and language spoken at home. For more information, visit [www.census.gov/acs](http://www.census.gov/acs).

## DATA COLLECTION CHALLENGES AND BEST PRACTICES

The ability to provide quality, disaggregated AAPI data from federal surveys and administrative records begins with the techniques and methods utilized to collect data. Fundamental challenges involve ensuring that the federal survey sample includes adequate representation of the AAPI community, which is a relatively small population in the United States.

Additionally, a significant proportion of the AAPI community is limited in English proficiency and may not be as accessible via some modes of data collection. A large proportion of the AAPI population is foreign born, and about 1 in 3 have limited English proficiency, which can make it difficult for federal agencies to obtain responses from the AAPI community. Another challenge includes the actual presentation of the question on race. If federal surveys and administrative records only include the minimally required U.S. Office of Management and Budget race categories of “Asian” and “Native Hawaiian and Other Pacific Islander” as response options without collecting any further information, data for the AAPI population cannot be disaggregated.

## Best Practices to Overcome Data Collection Challenges

### *If the challenge is:*

#### **Encouraging AAPI participation—**

- Conduct outreach activities with Asian and Native Hawaiian and Other Pacific Islander community organizations, advocates, ethnic media, and respected leaders ahead of the data collection and at multiple points leading up to the collection point. This engagement with the AAPI community will serve as an opportunity to build awareness of the federal survey or administrative records and to emphasize the importance of participation. Through cooperation and support from the organizations, advocates, and leaders, information can be disseminated to the AAPI population about why the federal survey or administrative records need to collect certain data and how those data can be used to inform and benefit their community. If available, provide specific examples of beneficial uses of the data.

#### **Ensuring adequate AAPI representation—**

- Oversample areas in which there are high concentrations of Asian American and Native Hawaiian and Other Pacific Islander populations. This will increase the number of AAPI respondents included in the federal survey sample and will aid in making the resulting estimates more reliable. This approach is most effective, in terms of resources and disaggregating data, if the survey sampling frame already includes information on detailed AAPI groups. It should be noted that there are trade-offs associated with oversampling one particular group. This could lead to improved estimates for the oversampled group, but also impact estimates for other groups.

#### **Taking into account limited English proficiency in the AAPI community—**

- Develop a language assistance program. This can include a number of components that will assist AAPIs with limited English proficiency in completing the survey or administrative records. The assistance can range from providing an in-language guide or pamphlet that explains how to complete the survey/administrative records to recruiting and training bilingual interviewers or neighborhood interpreters, providing in-language telephone and Internet assistance, or actually translating survey/administrative record instruments. Efforts could also be made to translate reports, press releases, and blogs into languages such as Chinese, Vietnamese, and Korean.



#### **Eliciting the reporting of specific Asian or Pacific Islander ethnicities—**

- Collect information on detailed ethnicities through the use of example groups which reflect the diversity within the community, and checkbox response categories and/or dedicated areas to enter detailed information in the data collection instrument. The use of checkboxes only may be necessary for federal agencies that lack resources and infrastructure to capture and process written/typed responses to the question on race. In addition, including a short set of AAPI ethnic group examples may help orient those who are not familiar or do not identify with pan-ethnic terms such as “Asian” or “Pacific Islander.”

#### **Increasing AAPI response rates—**

- Collect AAPI data using multiple modes. Higher proportions of certain population groups may be more likely to participate in data collection via particular modes—especially if these groups face language barriers. Segments of the AAPI population that are difficult to obtain responses from via a mail-out/mail-back survey may be more likely to participate in a survey that offers telephone interviewing, or one that can collect data through a face-to-face interview. Include community-based approaches and utilize trusted community partners to endorse, promote, and even conduct research.

## Data Analysis and Evaluation Best Practices Highlight:

*“Asians in the U.S. Labor Force: Profile of a Diverse Population: 2008–2010,”*  
U.S. Bureau of Labor Statistics

In 2011, the Bureau of Labor Statistics published an article in its Monthly Labor Review that contained disaggregated AAPI estimates from the Current Population Survey, a monthly household survey. In order to publish reliable estimates for seven Asian groups—Asian Indian, Chinese, Filipino, Japanese, Korean, Vietnamese, or Other Asian—data were pooled together over a series of months or years.

For example, the article provided 36-month averages (2008–2010) for labor force participation rates by age, nativity, marital status, and educational attainment for each Asian population group. Unemployment rates and alternative measures of labor underutilization for 2003 through 2010 were also presented—providing data users with estimates previously unavailable at that level of detail.

For more information, visit [www.bls.gov/opub/mlr/2011/11/art1full.pdf](http://www.bls.gov/opub/mlr/2011/11/art1full.pdf).

## DATA ANALYSIS AND EVALUATION CHALLENGES AND BEST PRACTICES

After information about the AAPI community has been collected in federal surveys and administrative records, thorough analysis and evaluation of the data are required. A common challenge encountered is that the typically small AAPI sample size can limit data analysis and evaluation. Furthermore, a small AAPI sample size can limit the amount of AAPI data that is suitable for publication. Another challenge is presenting the data clearly to the wide range of interested users.



## Best Practices to Overcome Data Analysis and Evaluation Challenges

### *If the challenge is:*

#### **Working with a dataset that has a small AAPI sample size—**

- Pool multiple months or years of survey data together. This is extremely helpful when attempting to generate more reliable estimates and increases the possibility of analyzing data for relatively small AAPI groups. Pooled data may also require reweighting.

#### **Analyzing AAPI disaggregated data—**

- Produce analytic results for detailed Asian and detailed Native Hawaiian and Other Pacific Islander population groups when reliable estimates are available in order to better show the differences and similarities within these populations. Produce analytic results at state and regional levels when reliable estimates are available to display how trends may differ by geography.

#### **Ensuring that the AAPI data are suitable for publication—**

- Review AAPI data carefully prior to publication. Compare the AAPI estimates with those from other sources (such as the American Community Survey) to see if there is general consistency between the estimates. If the estimates differ substantially, try to identify possible reasons why. AAPI estimates should also be checked to see if they vary dramatically across time periods. Additionally, standard errors should be examined to ensure that the confidence intervals are not too large for quality estimates. Ensuring that the AAPI sample size is adequate for the analysis being conducted and that participant confidentiality has been maintained are also critical factors in determining whether the data are suitable for publication.

#### **Preparing AAPI data analysis for the public—**

- Create a variety of data products that will communicate important findings and trends to different audiences. It is beneficial to design some analytic briefs, reports, tabular data, and other data products that simplify, as much as possible, complex analyses and data trends for novice data users, laypersons, media, and the general AAPI community. In addition, reports, papers, journal articles, and other highly technical documents can be created for savvy data users, academics, and other researchers.

## Data Access and Dissemination Best Practices Highlight:

### *Improving Disaggregated Data in Education*

Disaggregated data will help target resources where they are most needed—in high schools where almost half of Cambodians are not earning a diploma, in colleges where only 14.7 percent of Native Hawaiian or Other Pacific Islanders graduate, and in communities where the average per capita income of Hmong Americans is lower than any racial group nationwide.

As a part of the White House Initiative on Asian Americans and Pacific Islanders' efforts, a Request for Information (RFI) was disseminated to survey state and local educational associations and institutes of higher education about their disaggregation practices and policies. Through the RFI, WHIAAPI was able to collect information about successful practices and policies that schools are using to disaggregate AAPI student data, and the barriers that preclude the collection of disaggregated data.

WHIAAPI, along with CARE, gathered data experts, school administrators, researchers, advocates, and philanthropic organizations for a major convening in June 2013 and September 2015 to discuss these findings and encourage collaboration across all these institutions.

For more information, visit [www.ed.gov/edblogs/aapi/aapi-data-disaggregation/](http://www.ed.gov/edblogs/aapi/aapi-data-disaggregation/).

## DATA ACCESS AND DISSEMINATION CHALLENGES AND BEST PRACTICES

There are a number of issues to consider when preparing AAPI data from federal surveys or administrative records for public access and dissemination. One basic issue to address is how to package/present AAPI data to meet the needs of the AAPI community, as well as others who want to use the data. At times, data products and datasets from federal surveys and administrative records are released in a manner where primarily the seasoned or “power” data users can best navigate and utilize the data. For data users who mainly rely upon published tabulated data, standard data products may not contain the desired AAPI disaggregation. Additionally, getting the word out to the AAPI community about the availability of federal survey and administrative records data on their population can be a challenge. Researchers searching for data on the AAPI population in the United States can access many great data products at [www.data.gov/aapi](http://www.data.gov/aapi).

## Best Practices to Overcome Data Access and Dissemination Challenges

### *If the challenge is:*

#### **Making datasets more accessible to a wider range of data users—**

- Partner with and engage communities to ensure the approaches for access to and dissemination of AAPI data are community relevant and usable.
- Provide materials to assist data users in working with data and datasets, release public use files, and arrange special access for researchers who require analysis of full microdata. Providing online tutorials and/or online courses about how to technically utilize federal survey datasets can greatly help data users who have limited experience with manipulating datasets. These courses can assist intermediate and novice data users in understanding and utilizing the released dataset.
- Include disaggregated AAPI data on public use files (when reliable estimates that maintain respondent confidentiality are produced) so that researchers can conduct their own analyses. This includes ensuring that high quality variables have been added to the public use files and that documentation is clear and complete. Perturbation or topcoding of variables may be necessary to maintain respondent confidentiality. Very advanced data users may want to conduct research using entire microdata files from federal surveys or administrative records.
- Provide restricted on-site access to microdata for approved research projects through agreements with agency research data centers. This could be an option when the public release of these data could otherwise threaten respondent confidentiality. Additionally, federal agencies can consider organizing and joining forces with other agencies that have existing research data centers in order to provide restricted on-site access to data.

#### **Providing data products that meet the needs of a multitude of data users—**

- Make high-quality microdata and analytic results available as widely as practicable as soon as possible after data collection—in as much detail as possible—while maintaining survey participant confidentiality. Federal agencies should make announcements to the public whenever they release new datasets and data products, as well as announce any relevant changes to existing datasets or data products.

- Establish a custom tabulation program. There are data users who may not have the expertise to analyze public use files or microdata, or who find that standard tabulations and other data products created from federal survey and administrative record data do not meet their needs. Therefore, it is advisable for federal agencies to have a process in place for data users to request custom tabulations of particular variables.

### Communicating with the audience—

- Partner with communities to develop data releases to ensure they are culturally and linguistically appropriate and that they can reach the intended audiences for the intended uses. Community-based partners who are trusted by community members can significantly increase the uptake and use of data and research results, as well as be valuable featured speakers and endorsers of the data.
- Produce a research piece to accompany the release of new data items in a news release or report. The research piece can highlight significant findings or trends, as well as provide much more technical detail about the data and present caveats. It can also serve as a valuable resource for agency staff.
- Use a variety of vehicles to share data-related announcements and information (webinars, press releases, blogs, factsheets, etc.). Ensure that the announcements and information are shared with ethnic media, community organizations, advocates, university Asian/Pacific



Islander studies departments, and federal agency advisory committee members. Work to ensure materials are translated into various AAPI languages.

## SUMMARY

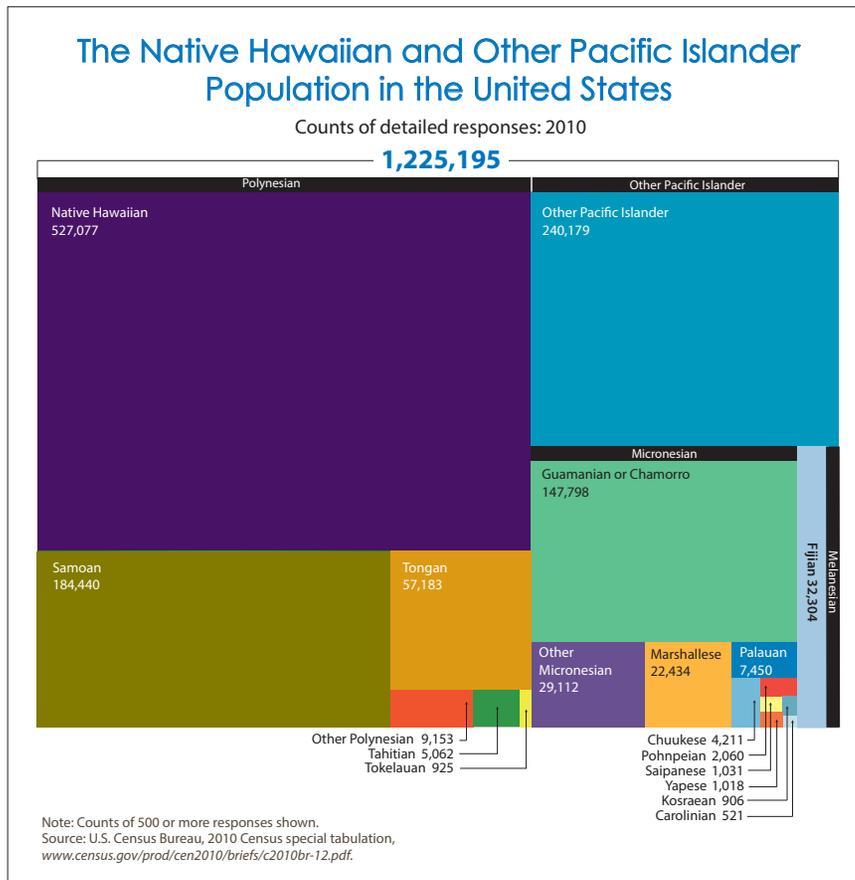
While there are many challenges facing federal surveys in their efforts to provide high quality disaggregated information on the fast-growing, diverse AAPI population, there are a number of ways these difficulties can be addressed. Federal statistical agencies should persist in finding and exploring new methods and techniques that assist in making disaggregated AAPI data more widely available. This will allow data users to examine the diversity within the AAPI community and the differences that exist in socioeconomic status, educational attainment, health and other areas. This document, created by the Data and Research Subcommittee, contains the following set of best practices for providing disaggregated AAPI data from federal surveys.

## DATA COLLECTION

- Conduct outreach activities with AAPI community organizations, advocates, and respected leaders to encourage AAPI participation.
- Oversample the AAPI population to ensure adequate AAPI representation.
- Develop a language assistance program to account for limited English proficiency in the AAPI community.
- Collect information on detailed ethnicities with an adequate set of separate ethnic checkbox response categories that reflect the diversity within the community and/or write-in/type-in areas in the survey instrument.
- Collect AAPI data using multiple modes, considering that the AAPI community is diverse and may not be as responsive to surveys in one particular mode of data collection and may require culturally and linguistically appropriate approaches.

## DATA ANALYSIS AND EVALUATION

- Pool multiple months or years of survey data together and combine samples when working with a dataset that has a small AAPI sample.
- Produce analytic results for disaggregated AAIs when reliable estimates are available in order to better show the similarities and differences within the population.
- Review AAPI data carefully to ensure that the data are suitable for publication.
- Create a variety of data products that will communicate important findings and trends to different audiences.



## DATA ACCESS AND DISSEMINATION

- Partner with AAPI communities in access to and dissemination of data to ensure these approaches are relevant and usable for target audiences.
- Provide materials to assist data users in working with data and datasets, release public use files, and arrange special access for researchers who require analysis of full microdata to make datasets more accessible to a wider range of data users.
- Make high-quality microdata and analytic results available as widely as practicable as soon as possible after data collection—in as much detail as possible—while maintaining survey participant confidentiality.
- Use a variety of vehicles to share data-related announcements and information (webinars, press releases, blogs, factsheets, etc.) and ensure that these are distributed to ethnic media, community organizations, Asian/Pacific Islander studies departments, and federal statistical advisory groups.
- Partner with communities to develop data releases that are culturally and linguistically appropriate, so that they reach the intended audiences.



SRI LANKAN  
INDONESIAN  
CAMBODIAN JAPANESE CHUUKESE  
TONGAN **INDIAN** MALAYSIAN **KOREAN** PALAUAN  
THAI **CHINESE** TAIWANESE  
BURMESE **CHINESE** HMONG  
FIJIAN BANGLADESHI **FILIPINO**  
**SAMOAN** BHUTANESE  
LAOTIAN **VIETNAMESE** NEPALESE  
**PAKISTANI CHAMORRO**  
NATIVE HAWAIIAN  
MARSHALLESE

*This word cloud is a visual representation of the diversity of the AAPI community.*



Asian Americans and Pacific Islanders (AAPIs) are now the fastest growing racial group in the United States, increasing over four times as rapidly as the total U.S. population and expected to double to more than 47 million by 2060. Recognizing this tremendous growth and the unique needs within AAPI communities, President Barack Obama signed Executive Order 13515 reestablishing the White House Initiative on Asian Americans and Pacific Islanders (Initiative) on October 14, 2009. The Initiative, chaired by Acting Secretary of Education John King and U.S. Surgeon General Dr. Vivek Murthy, and led by Executive Director Doua Thor, works to improve the quality of life and opportunities for AAPIs by facilitating increased access to and participation in federal programs, where AAPIs remain underserved.

The Executive Order also established the President's Advisory Commission on AAPIs and the Initiative's Interagency Working Group. The Commission comprises community and business leaders who serve as the eyes and ears of the Administration. The Interagency Working Group, representing 24 federal agencies and offices, is charged with creating and implementing strategic agency plans to help increase the AAPI community's access to federal resources. In 2013, the Initiative launched its Regional Network—more than 200 federal officials from regional offices across the country working to connect local AAPI communities to federal resources. In 2014, the Initiative started its E3! Ambassadors Program, enlisting young leaders to “educate, engage, and empower” their campuses and communities on key issues.

Through its innovative approaches ranging from interagency collaboration to on-the-ground community outreach, the Initiative seeks to highlight both the tremendous unmet needs in the AAPI community as well as the dynamic community assets that can be leveraged to meet many of those needs.

For more information, visit [www.whitehouse.gov/aapi](http://www.whitehouse.gov/aapi).

This document provides high-level information on common challenges and best practices for disaggregating AAPI data in federal surveys and administrative records. More detailed information on how federal agencies can implement these best practices may require consultation. Federal agency representatives in the Data and Research Subcommittee can direct agencies to subject matter experts for in-depth consultation, discussion, and guidance.



U.S. Department of Labor  
U.S. Bureau of Labor Statistics  
Dori Allard [Allard.Dorinda@dol.gov](mailto:Allard.Dorinda@dol.gov)



U.S. Department of Housing  
and Urban Development  
Alejandro Aviles [Alejandro.Aviles@hud.gov](mailto:Alejandro.Aviles@hud.gov)  
Shawn Bucholtz [Shawn.J.Bucholtz@hud.gov](mailto:Shawn.J.Bucholtz@hud.gov)



U.S. Department of Commerce  
U.S. Census Bureau  
Enrique Lamas [enrique.lamas@census.gov](mailto:enrique.lamas@census.gov)  
Victoria Velkoff [victoria.a.velkoff@census.gov](mailto:victoria.a.velkoff@census.gov)  
Karen Humes [karen.humes@census.gov](mailto:karen.humes@census.gov)  
Nicholas Jones [nicholas.a.jones@census.gov](mailto:nicholas.a.jones@census.gov)



U.S. Department of Education  
Tom Snyder [Tom.Snyder@ed.gov](mailto:Tom.Snyder@ed.gov)  
Peggy Carr [Peggy.Carr@ed.gov](mailto:Peggy.Carr@ed.gov)



U.S. Department of Health  
and Human Services  
Sam Wu [samuel.wu@hhs.gov](mailto:samuel.wu@hhs.gov)  
Virginia Cain [Vxc6@cdc.gov](mailto:Vxc6@cdc.gov)



U.S. Office of  
Management and Budget  
Jennifer Park [Jennifer\\_E\\_Park@omb.eop.gov](mailto:Jennifer_E_Park@omb.eop.gov)

