ASIAN AMERICANS
ADVANCING
JUSTICE
AAJC

VIA EMAIL

October 3, 2022

NSTC Subcommittee on Equitable Data
Office of Science and Technology Policy
Eisenhower Executive Office Building
1650 Pennsylvania Ave. NW
Washington, DC 20504
Email: equitabledata@ostp.eop.gov

Re: Request for Information; Equitable Data Engagement and Accountability
(Document Citation 87 FR 54269)

Asian Americans Advancing Justice – AAJC submits this comment in response to the White House Office of Science and Technology Policy (OSTP) federal register notice regarding the Request for Information; Equitable Data Engagement and Accountability, 87 FR 54269 (September 2, 2022). Asian Americans and Native Hawaiian and Pacific Islanders (NHPIs) comprise vastly diverse racial groups. Without accurate disaggregated data, some of the most disadvantaged in our communities are rendered invisible, leaving their needs unmet. Detailed data are also critical to our ability to break down the stereotype of the "model minority," which has been used to erase the history of exclusion and discrimination against Asian Americans and NHPIs. We write to provide comments to help foster collaboration between all levels of the Federal government, promote engagement of communities that access or participate in Federal programs in data collection and research, and create opportunities for the broader community to publicly access equitable data, including disaggregated and granular data.

Organizational Information

Advancing Justice | AAJC, a nonprofit, nonpartisan 501(c)(3) organization, was incorporated in 1991 and opened its Washington, DC office in 1993. For over thirty years, we have served as the leading Asian American voice on civil rights issues in our nation's capital. Our mission is to advance civil and human rights for Asian Americans and to build and promote a fair and equitable society for all. Our expertise on issues of importance to the Asian American community is widely acknowledged in the media, by the public, and by policymakers at the federal, state, and local levels.

Advancing Justice | AAJC considers data collection and reporting to be the backbone of its mission. We have been working to eliminate the barriers that have historically resulted in the undercounting and underreporting of Asian Americans and NHPIs in federal data collection and analysis efforts, particularly in the decennial census count. Our permanent census program monitors census policy and educates policy makers—including through testifying at Congressional hearings—and conducts community outreach and education on the surveys conducted by the Census Bureau.

GUIDING PRINCIPLES FOR APPROACHING DATA EQUITY

Prioritizing Data Disaggregation

The collection of detailed data are particularly critical for Asian Americans, who are among our nation's fastest growing and most diverse racial groups.[1] Often viewed as homogenous, these communities include more than 30 detailed subgroups that can differ dramatically across key social and economic indicators.[2] While Indian Americans have an average poverty rate of 6%, Mongolian Americans and Burmese Americans have a poverty rate of 25%.[3] Roughly 75% of Taiwanese Americans hold a bachelor's degree, yet only 14% of Laotian Americans do.[4] Another example can be found in health disparities. A study showed that "19.4% of Asian adults compared to 12.9% of whites report[ed] being without a usual source of health care, with Cambodian and Vietnamese [Americans] … three times more likely to skip doctor visits due to cost compared to all Asian [Americans] or U.S. residents."[5] The study further found that U.S.-born Vietnamese American women represent one of the highest risk groups for breast cancer at a rate of being four times more likely to die of breast cancer than any other Asian American groups and that Korean American children are four times more likely to have no health insurance as compared to others. Finally, disaggregating data on the prevalence of smoking in New York City showed that while the prevalence in smoking was lower overall in Asian Americans compared to whites (14.1% vs. 18.6%), that was not the case for some Asian American subgroups, where the actual prevalence of smoking was much higher, such as 35.5% in Korean Americans.[6] And while Japanese Americans, Filipino Americans, and Indian Americans all have English language proficiency at or above 80%, only 36% of Bhutanese Americans speak English proficiently.[7]

We strongly believe in the collection of detailed race and ethnicity data by Federal agencies. Without this requirement, Federal agencies are unlikely to adopt collection of detailed race and ethnicity data. They have long had the option to, but we have not seen significant movement toward detailed reporting when it is not mandatory. Thus, we believe that the requirement should be made of all Federal agencies for data disaggregation, with a process OMB could administer that would allow agencies to apply for an exemption where such collection and reporting of detailed data would create undue hardship or privacy concerns, in which case they must collect and report data based on the minimum categories.

Moreover, we believe Federal agencies should be required to collect detailed race and ethnicity data even when such data could not be responsibly reported due to statistical reliability and confidentiality concerns. This will provide us the option to aggregate the data across time for the same group, (i.e., pool the responses across a period of time), which could address statistical reliability and confidentiality concerns.

Prioritizing Granular Data

In addition to disaggregating data by racial and ethnic group, the federal government must ensure the data it produces illuminate geographical differences by providing Asian American and NHPI data at different geographic levels: local, state, and federal. The more granular the data—for example, drilling down to the county or city level—the more likely it is that problems can be identified in specific localities. While

---

[1] https://www.pewresearch.org/fact-tank/2021/04/29/key-facts-about-asian-origin-groups-in-the-u-s/.

[2] https://www1.nyc.gov/assets/immigrants/downloads/pdf/Fact-Sheet-NYCs-API-Immigrant-Population.pdf.

[3] Id.

[4] https://theconversation.com/asians-could-opt-out-of-naming-a-country-of-origin-on-the-2020-census-a-policymakers-nightmare-92714.

[5] https://www.pfizer.com/news/articles/health_disparities_among_asian_americans_and_pacific_islanders#:~:text=In%20fact%2C%2019.4%25%20of%20Asian,all%20Asians%20or%20U.S.%20residents.

[6] Id.

[7] https://www.pewresearch.org/fact-tank/2021/04/29/key-facts-about-asian-origin-groups-in-the-u-s/.

collecting geographic data should be a priority, granular data on socio-economic status and language ability also provide targeted information that can help address some of the most pressing issues in the Asian American and NHPI community and improve civil rights outcomes. Overlaying geographic data with disaggregated data by Asian American and NHPI subgroups is a first step. The collection of intersectional data can provide evidence of continued discrimination and address problems in order to begin driving solutions.

Not Collapsing Asian Americans or NHPIs into an "Other" Category

We believe it is a necessary step for the federal government, academic institutions, and other organizations that collect data to end the use of an "All Other Races" category. While the practice may have made the presentation of data easier for agencies and others, whereby they could present data on a few racial and ethnic groups, such as white, Black, and Latino, and then presented the rest as "All Other Races" data, the practice does not serve the public at large, particularly those that comprised "All Other Races."

For example, combined, Asian Americans and NHPIs are the "majority" populations in the state of Hawaii.[8] Communities of color are the majority populations in the state of California.[9] Demographics have significantly changed over the last several decades. Asian Americans remain among the fastest growing groups in the United States, with 38.6% growth between the 2010 and 2020 Censuses.[10] The NHPI population also grew rapidly between 2010 and 2020, at 29.5%.[11] These communities are often lumped into "All Other Races," making these fast-growing communities invisible. The practice moving forward should be to have agencies report on the data for, at a minimum, all racial and ethnic categories, with the addition of data on detailed groups as available. We hope that sets a precedent in academia and other industries. Regardless of the reasons why the "All Other Races" category was initially constructed, is no longer acceptable because the diversity within the "All Other Races" category renders this information essentially useless. If "All Other Races" is being used because the study does not have enough data to disaggregate by race, this should be addressed in the methodology. Similarly, if the researcher(s) is/are not motivated to collect disaggregated data at the onset, the results cannot be relied upon. Ultimately, there is no acceptable reason for the use of "All Other Races" in today's changing world.

Creating Accessible Data for the Public

To promote equity and democratic participation and engagement, the federal government should release the maximum amount of data to the public without compromising individual privacy. For individuals who do not have expertise in data manipulation, providing publicly accessible dashboards allows them to see trends easily. Offering trainings about how to best access data could further help educate communities and increase the impact of equitable data collection. And while it is important that data be easily understood by individuals lacking statistical expertise, the data must also be accessible at different levels of statistical competency. Thus, in addition to producing easy-to-understand data products, agencies should provide the underlying data, with proper technical documentation, so that experts can assist communities and advance solutions requiring more complicated statistical analysis. An example of an agency providing both easy-to-understand products and the underlying data can be found with the U.S. Census Bureau's release of its determinations made under Section 203 of the Voting Rights Act. Not only does the Bureau provide the determinations and support materials (such as comparison tables), it also provides a public use dataset that

[8] http://aapidata.com/wp-content/uploads/2018/05/HI-2018.pdf.

[9] https://www.ppic.org/publication/californias-population/.

[10] https://www.census.gov/library/visualizations/interactive/race-and-ethnicity-in-the-united-state-2010-and-2020-census.html.

[11] Id.

allows organizations, such as ours, to further analyze the data to determine, for example, jurisdictions that just missed coverage.[12]

How We Collect Data Equitably

While the content of the data that are collected matters tremendously, the methods used to collect data should also reflect our commitment to equity and civil rights. For example, surveys should be available in-language. Before these are widely distributed, they should be reviewed by community members to ensure translations are accurate and understandable. And we must guarantee they are distributed widely, including to households lacking internet access and other hard-to-reach individuals.

SPECIFIC USE CASES FOR ASIAN AMERICAN AND NHPI COMMUNITIES

As an organization committed to advancing the civil and human rights of Asian Americans, we offer use cases on how these guiding principles play out in our own work on immigration, education, telecommunications, technology, and media, and anti-Asian hate.

Immigration

Some overarching principles undergird the equitable collection of immigration data. It is important to ensure immigration data are accessible to the community and can be used by community members and experts. More dashboards like those available from Customs and Border Patrol (CBP) would increase opportunities for community-based organizations (CBO) to use equitable data to hold government accountable to the American public. Regular updates of the data are far more useful than end of year reports. Quarterly reports allow researchers to review the rate at which family-based immigration visas are being processed.

There are specific actions federal agencies can take to improve data equity in the immigration space. Immigration and Customs Enforcement (ICE) presents the most problems in terms of access to equitable data. ICE's data are often late, incomplete, and not publicly available except through outside organizations like Transactional Records Access Clearinghouse (TRAC), which analyzes ICE data obtained through Freedom of Information Act (FOIA) requests. This makes it difficult to get a sense of whether deportations of Southeast Asian Americans have decreased. It also poses challenges in studying how ICE implements its enforcement priorities.

While the yearly annual report for ERO (Enforcement and Removal Operations) usually includes data on removals by country of citizenship, two issues exist.[13] First, the data are often late. The Fiscal Year 2021 report has yet to be released. The long delay between the end of the fiscal year and the release of the ICE ERO data make it hard to identify any issues with ICE operations in a relevant time frame. Second, the data are presented in pdf format, and are not cross tabulated. This poses challenges to analyzing trends in terms of enforcement across different categories (like how many individuals from a particular country of citizenship have criminal convictions when apprehended by ICE).

---

[12] https://www.census.gov/programs-surveys/decennial-census/about/voting-rights/voting-rights-determination-file.html. For an example of how our organization used that dataset to produce additional analysis, see https://www.advancingjustice-aajc.org/publication/jurisdictions-and-languages-just-missed-coverage-2021-section-203-determinations.

[13] https://www.ice.gov/doclib/news/library/reports/annual-report/eroReportFY2020.pdf.

CBP is better than ICE because it provides publicly available dashboards[14] that allow for deeper analysis, including a dashboard[15] that includes some enforcement statistics for the current fiscal year to date. CBP also has datasets in CSV format available for download that allow for some independent analysis/manipulation. For example, CBP offers a data tool to explore data around southwest land border encounters.[16] The data can be filtered by fiscal year, CBP component, certain characteristics related to family, country of citizenship, and legal authority. It is useful for CBP to have all its datasets available on one website, with CSV files for download and data dictionaries available. Nevertheless, the datasets are not as useful as they could be, especially because they lose crucial information about country of citizenship. The datasets vary in terms of how many countries are included in the "other" category. For southwest border encounters, for example, the citizenship variable only includes El Salvador, Guatemala, Honduras, Mexico, and Other, despite Other accounting for almost half of all encounters.

United States Citizenship and Immigration Services (USCIS) has some data available, but mostly in pdf format.[17] Expanding the types of files available for download will improve data accessibility and impact. Department of Homeland Security (DHS) publishes data about legal immigration and migration flows online. These data are also better than ICE. It includes perhaps the most comprehensive breakdown of legal immigration and adjustment of status information.[18] Immigration statistics are centralized on one page.[19] However, there are still some issues. First, the cross tabulations are limited. For example, there is no raw dataset with information about the age of different refugees by country of origin. Second, unlike CBP, there is a dearth of dashboards that allow for easy manipulation of the data. Third, in the case of family-immigration statistics, reports are often delayed – for example, it is unclear how close USCIS and the State Department are to reaching the cap on family-based immigrant visas for this fiscal year, because data are only available for the first two quarters of this fiscal year.

Thus, many organizations rely on TRAC immigration data to have access to some of those tools.[20] But those data are based on federal data recovered through FOIA requests – it would be more efficient if ICE presented the same kind of data without TRAC having to issue FOIA requests to ICE and others.

Telecommunications, Technology, and Media

For Asian Americans and NHPIs, one of the first and primary challenges to broadband access is accurate data. Few studies on broadband access specifically focus on Asian Americans, fewer still disaggregate their data, and even fewer data exist for Pacific Islanders.[21] Below, we outline some specific areas where additional data and disaggregation would benefit the Asian American and NHPI community, while also detailing opportunities for additional interagency collaboration.

The Universal Administration Co. (USAC) website states that they publish total households enrolled and state enrollment data weekly, zip code data monthly, and county data quarterly.[22] However, the data do not appear to be published on a regular or timely schedule. Having access to real-time enrollment information would be helpful to know how effective outreach efforts are in targeting unserved and underserved

---

[14] https://www.cbp.gov/newsroom/stats/cbp-public-data-portal.

[15] https://www.cbp.gov/newsroom/stats/cbp-enforcement-statistics.

[16] https://www.cbp.gov/newsroom/stats/southwest-land-border-encounters.

[17] https://www.uscis.gov/tools/reports-and-studies.

[18] https://www.dhs.gov/immigration-statistics/special-reports/legal-immigration.

[19] https://www.dhs.gov/immigration-statistics.

[20] https://trac.syr.edu/immigration/.

[21] https://www.advancingjustice-aajc.org/sites/default/files/2021-05/2021.05.06%20E%26C%20Broadband%20Disparities%20Testimony.pdf.

[22] https://www.usac.org/about/affordable-connectivity-program/acp-enrollment-and-claims-tracker/.

communities. Moreover, having broadband mapping data that overlay census data would illustrate what communities are unserved or underserved, and would increase efforts to promote broadband equity. The Federal Communications Commission (FCC) is currently going through a process to update the agency's broadband data and mapping tools. We encourage the FCC to engage in this as a regular process as the current maps are inadequate. Finally, we have concerns over how the government is collecting and using biometrics data. While Government Accountability Office reports offer some information, there are still too many unknowns. Information about the data agencies have access to, what biometrics information they collect, and what they use it for are essential. For example, we know that DHS has access to multiple databases – including state driver's license data – and that they use biometrics, but it is unclear how they are using these data, if they are purchasing data from other databases, who has access to these data, and whether they share data with other government agencies.

Education

Disaggregating data, providing more granular data, eliminating the "Other" category, and collecting data equitably are especially important when it comes to improving educational outcomes for Asian Americans and NHPIs. Ideally, data should be disaggregated by Chinese, Asian Indian, Filipino, Japanese, Korean, Vietnamese, Pakistani, Cambodian, Hmong, Laotian, Thai, Taiwanese, Burmese, Bangladeshi, and Nepalese sub-populations, "Other Asian" sub-populations (to capture less-populated communities), and Native Hawaiian, Samoan, Chamorro, Tongan, iTaukei, Marshallese, and "Other Pacific Islander" sub-populations. If possible, subgroups such as the lu Mien and Montagnards under Asian American should also be disaggregated. This is not only true for students; these data must also be collected about staff and teachers. When available, disaggregated data provides insight into educational disparities within the Asian American and NHPI community.

More granular data will provide further insight into educational disparities as they apply to Asian Americans and NHPIs. Specifically, data on English learner status, sex, disability, college and career readiness, staffing and finance (e.g., teacher experience and per-pupil expenditures), household income, and parental education should be collected. More data on curriculum, specifically the number of Asian American Studies classes and data on how many credit hours are spent teaching different subjects, would help address curricular inequalities. Data should also be collected on anti-Asian hate incidents in schools. Moreover, data on the English language ability/language preference of parents should be tracked.

We strongly support the collection of school climate and safety data as these are especially important for our communities. At a national level, a 2016 study showed U.S.-born Laotian and Cambodian men, ages 18-39 years, were reported to have the highest incarceration levels of all Asian Americans and NHPIs at 7.26% of their total population, and many Southeast Asian American males report being stereotyped by their teachers as troublemakers, gang members, delinquents, and failures. [23] Students with disabilities are disproportionately disciplined, further highlighting the importance of this particular data collection.[24] We are hopeful that such data will continue to be gathered and recommend that further analyses be conducted to investigate disproportionate discipline of other disadvantaged groups, specifically English learners and recent arrivals to the United States.

Finally, it is essential that these data are made available to the public, as parents and students can use this information to make informed decisions about their educations. Recently, the availability of user-friendly data has allowed students, parents, and advocates to access data regarding their own schools as well as

[23] http://store.samhsa.gov/product/A-Snapshot-of-Behavioral-Health-Issues-for-Asian-American-Native-Hawaiian-Pacific-Islander-Boys-and-Men-Jumpstarting-an-Overdue-Conversation/SMA16-4959.

[24] https://www2.ed.gov/programs/osepidea/618-data/LEA-racial-ethnic-disparities-tables/disproportionality-analysis-by-state-analysis-category.pdf.

schools attended by others, which has helped shape advocacy and build greater understanding of unequal educational opportunity. All data should be publicly reported for transparency, civil rights enforcement, monitoring, and research purposes.[25] These data are especially important in the wake of the pandemic, when access to educational opportunity was not equitably available. Data can help make those inequalities transparent and provide the impetus for remedies.

Anti-Asian Hate

The flaws with hate crimes data are systemic – largely arising out of the fact that reporting by law enforcement agencies is not mandatory but voluntary. The anti-Asian hate crimes and hate incidents that make the news—as well as those reported to Advancing Justice's Stand Against Hatred site, Stop AAPI Hate, and other community groups—are deeply troubling. And yet we know that they very likely are only a fraction of what our communities are facing since hate crimes are chronically un- and under-reported.[26] As detailed in a recently-released report from the Movement Advancement Project, only an estimated 3.6% of hate crimes are reported to the FBI each year.[27] There is a significant gap between the FBI's Uniform Crime Reporting Program's database of reported hate crimes and overall hate crime experiences of people living in the United States, as collected by the Department of Justice's National Crime Victimization Survey.[28]

Even with concerns about under-reporting, we saw a significant increase in hate crimes against Asian Americans in 2020. The FBI's 2020 Hate Crimes Report documents a significant increase in criminal incidents over 2019, including a 76% increase in hate crime incidents motivated by anti-Asian bias: 279 hate crime incidents against individuals of Asian descent were reported in 2020, compared to 158 incidents reported in 2019.[29] Passage of the COVID-19 Hate Crimes Act, which incorporates the Jabara-Heyer NO HATE Act, was an important step forward to promote better hates crimes reporting and data collection.[30] But it is not enough. To combat racism effectively, we must have data to make informed decisions about prevention. In addition to following the guiding principles provided here (e.g., not collapsing Asian American data into an "other" category, disaggregated by subgroups, etc.), we urge the Department of Justice and other agencies involved in collecting hate crime data to address the systemic issues that prohibit accurate data reporting on anti-Asian hate.

Conclusion

We appreciate the opportunity to provide comments on Equitable Data Engagement and Accountability. We look forward to working with the Administration on its ambitious equitable data efforts described in the Executive Order on Advancing Racial Equity and Support for Underserved Communities Through the Federal Government (E.O. 13985), including the *Vision for Equitable Data* issued to the President in April 2022. Please contact Terry Ao Minnis, Senior Director of Census and Voting Programs, at tminnis@advancingjustice-aajc.org if you have any further questions.

---

[25] http://www.seattleschools.org/modules/groups/homepagefiles/cms/1583136/File/Departmental%20Content/siso/disprof/2011/disprof_2011.

[26] https://www.npr.org/2021/03/11/975592502/asian-americans-experience-far-more-hate-incidents-than-numbers-indicate.

[27] https://www.lgbtmap.org/file/2021-report-hate-crime-laws.pdf.

[28] Id.

[29] https://abcnews.go.com/US/hate-crimes-asians-rose-76-2020-amid-pandemic/story?id=80746198, *see also,* https://crime-data-explorer.fr.cloud.gov/pages/explorer/crime/hate-crime.

[30] Public Law No: 117-13.