Analysis Plan for American Indian and Alaska Native Data

Uphold Indigenous Data Sovereignty through Data Practice

Contents

Section I. Special Considerations

Erasure of American Indians and Alaska Natives in data

American Indians and Alaska Natives (AI/AN) are often omitted in data collection and data analysis because they are categorized as "Other Race" or "Multiracial", grouped with other small populations, or misclassified as another race or, more recently, "something else." As a result of these systemic inequities in data, AI/AN are undercounted in local and national data, which leads to an inaccurate representation of social, economic, and health measures and results in a lack of resources being given to Native communities. Ensuring Indigenous people are included in data collection and analysis may require more resources so that agencies are able to collect data that is relevant to Indigenous people, identify datasets that include Indigenous people in the numbers, and use statistical techniques that maximize the sample size since that is usually the largest barrier to presenting AI/AN data without suppression.[1]

Protecting confidentiality of small populations

When analyzing data on small populations, special considerations must be applied to protect the confidentiality of victims. If reportable data contains less than 10 individuals, the data must be suppressed, which means redacting or withholding information to protect individual identities, privacy, and personal information. It is also important to avoid associating or grouping multiple identifying characteristics, including demographic and geographic information, which could, in combination, lead to the identification of victims in small populations.

Indigenous Data Sovereignty

In data collection and analysis, Indigenous Data Sovereignty—Indigenous people's legal authority to govern the collection, ownership, and applications of their own data— must be upheld, as it is derived from the inherent right of Indigenous nations to govern their peoples, lands, and resources as acknowledged in federal and international treaty law.[2–4] This government-to-government relationship between tribes and states means the collection, analysis, and presentation of AI/AN data must be negotiated with tribes and urban Native communities.

Section II. Action Steps to Uphold Indigenous Data Sovereignty

STEP 1	Hold a Tribal Consultation or Urban Confer to request permission to collect tribally specific data elements (e.g. tribal affiliation). Ensure it is understood by all involved parties what information can be released if there is a public records request.
STEP 2	In collaboration with tribes and urban Native communities, develop a Data Use Agreement (DUA) or Memorandum of Understanding (MOU) to establish the permitted uses for tribally specific data, and set protocols for how data will be shared back to tribes and urban Native communities, including frequency.
STEP 3	Request input from tribes and urban Native communities on analysis, interpretation, and presentation of AI/AN data.

Section III. Implementation of Recommended Data Elements

The data elements below are intended to improve the accuracy of data collected on AI/AN people and to strengthen data sharing between tribes and law enforcement agencies with the ultimate goal of bringing justice to victims. It must be known that these recommended data elements may not be comprehensive, as data collection is a continuously evolving process that requires ongoing quality improvement and evaluation. Suggestions for quality improvement (QI) processes are highlighted below.

Collecting race data

The collection of race data should, at a minimum, include the following categories: White/Caucasian, Black/African American, Asian, American Indian/Alaska Native, Native Hawaiian/Pacific Islander, and Other, with an option to specify race (see Table 1). Categories may need to be tailored to local contexts to ensure they adequately capture racial diversity and self-identification of race because, for example, categories in Seattle, WA, won't be the exact same as categories in Los Angeles, CA, due to differences in the population and how they identify racially.

Data tools should allow for the selection of multiple races with the ability to disaggregate—separate into individual categories—the data once collected. When AI/AN is selected, the data tool should allow for the collection of tribal affiliation data. Regardless of if an individual affiliates with a tribe, they should be included in the data for American Indians/Alaska Natives. Some individuals may not know their tribe or may belong to tribes that are not federally recognized.

Terminology key:

Variable: Element you are assessing

Description: Variable description

Data Type: Defines what type of data is allowable

Numeric codes: Values are the allowable numeric codes. Labels are the corresponding descriptors for each numeric code.

Branching Logic: When an answer to one question enables another question to appear

Required: Question cannot be bypassed without a response entered.

Quality Improvement Tip
Review the "Other" category annually. When responses are repeated more than once, consider adding an additional race category. When responses match existing race categories, group with the appropriate categories.

Table 1. Suggested race variable architecture and response categories

Variable	Description	Data Type	Numeric Codes		Branching Logic	Required
			Value	Label		
Race	Self-identified race	Numeric	1	White/Caucasian		Yes
			2	Black/African American		
			3	Asian		
			4	American Indian/Alaska Native	If selected, ask "What is your tribal affiliation?"	
			5	Hawaiian/Pacific Islander		
			6	Other Race	If selected, ask "What is your race?"	
			999	Unknown		

Collecting tribal affiliation data

Tribal affiliation is *the self-identification of association with a particular tribe or tribes and community attachment*.[5] This differs from tribal citizenship or enrollment, as only tribes determine and define tribal citizenship under their rights as sovereign nations.[6] Someone may be affiliated with a tribe by way of descendancy or some other means but not enrolled due to eligibility requirements, missing documentation, or lack of federal recognition of their tribe, among other reasons. If your jurisdiction includes tribes and federally defined tribal lands, however, the addition of tribal affiliation should not be adopted until a Tribal Consultation and/or Urban Confer has taken place.

Selection of tribal affiliation should be prompted when AI/AN is selected as the person's race. Data tools should allow for a drop-down option of all federally recognized tribes with the ability to select more than one tribe (see Table 2). In addition to the list of all federally recognized tribes, there should be an option for "state-recognized tribes" and "Other tribe" with the ability to specify tribe. The "Other tribe" option may be used when a person identifies with a specific subgroup (band, clan, Pueblo) rather than a larger tribe or with other Indigenous groups such as Canadian First Nations.

Quality Improvement Tip
Conduct an annual review of the list of federally recognized tribes to ensure it is up to date. Add in any newly recognized tribes. Updated listings are published each year in the Federal Register.

Table 2. Suggested tribal affiliation variable architecture and response categories

Variable	Description	Data Type	Numeric Codes		Branching Logic	Required
			Value	**Label**		
Tribal affiliation	Self-identified tribal affiliation	Numeric	1-576	See appendix I. for a list of all federally recognized tribes		Yes, if race=4
			777	Other tribe	If selected, ask "What is your tribe?"	
			888	State-recognized tribe	If selected, ask "What is your state-recognized tribe?"	
			999	Unknown		

Collecting ethnicity data

Ethnicity should be collected as a separate data field from race. There should be two categories for ethnicity: Hispanic/Latino and Not Hispanic/Latino (see Table 3).

Table 3. Suggested ethnicity variable architecture and response categories

Variable	Description	Data Type	Numeric Codes		Branching Logic	Required
			Value	**Label**		
Ethnicity	Self-identified ethnicity	Numeric	1	Hispanic/Latino		Yes
			2	Not Hispanic/Latino		
			999	Unknown		

Collecting tribal land information

Prior to collecting tribal land information, a Tribal Consultation must take place so the collecting agency has a clear understanding of how to manage and share tribe-specific information. Agencies with federal trust land (i.e. tribal land held in trust by the federal government for the benefit of a tribe or tribal members)[7] in their jurisdiction should collect information on whether the incident occurred on, the missing person(s) was last seen on, the deceased person(s) was found on, and/or the victim's residence was in Indian Country, an Alaska Native Village, or on tribal corporation lands (See Table 4).[1] Tribal land data fields should be required for all people, not only those who identify as AI/AN. These data fields should include branching logic to allow for the selection of tribal reservation(s) in your jurisdiction.

Quality Improvement Tip
Maintain a record of reservation boundaries and the zip codes in which they intersect for the ease of data entry. Conduct an annual review of the proportion of "Unknown" responses to ensure staff are trained appropriately.

Table 4. Suggested tribal land variables, architecture, and response categories

Variable	Description	Data Type	Numeric Codes		Branching Logic	Required
			Values	**Labels**		
Incident on tribal land	Incident occurred on tribal land	Numeric	1	Yes	If selected, ask "Which tribal land did this occur on?" and choose from drop down.	Yes
			2	No		
			999	Unknown		
			Value	**Labels**		
Missing from tribal land	Missing person last seen on tribal land	Numeric	1	Yes	If selected, ask "Which tribal land did this occur on?" and choose from drop down.	Yes
			2	No		
			999	Unknown		

---

[1] Indian Country, as defined by 18 U.S.C. § 1151, is all land within the limits of any Indian reservation under the jurisdiction of the United States Government, notwithstanding the issuance of any patent, and, including rights-of-way running through the reservation, (b) all dependent Indian communities within the borders of the United States whether within the original or subsequently acquired territory thereof, and whether within or without the limits of a state, and (c) all Indian allotments, the Indian titles to which have not been extinguished, including rights-of-way running through the same (Legal Information Institute).

			Value	Labels		
Deceased on tribal land	Deceased person found on tribal land	Numeric	1	Yes	If selected, ask "Which tribal land did this occur on?" and choose from drop down.	Yes
			2	No		
			999	Unknown		
Primary residence on tribal land	Victim's primary residence is on tribal land	Numeric	**Value**	**Labels**		
			1	Yes	If selected, ask "Which tribal land did this occur on?" and choose from drop down.	Yes
			2	No		
			999	Unknown		

Section IV. Data Analysis and Reporting Guidelines

Conducting analyses with violence data requires special considerations for victim's rights, privacy, and confidentiality. Data should never be used to stigmatize communities intentionally or unintentionally, nor should generalizations be made regarding entire populations. When working with AI/AN data, these same considerations must be applied with additional care for Indigenous Data Sovereignty, or the authority of tribes to govern the collection, ownership, and application of data from their members.[3]

Urban Indian Health Institute (UIHI) recommends providing a written report on tribal-specific data to the identified tribal representatives on a quarterly basis. However, specific reporting procedures—including frequency, format, and the appropriate representatives—should be discussed and determined during a Tribal Consultation.

The following sections outline how to analyze the recommended data elements proposed in Section III and provide analysis practice examples.

Analyzing race and ethnicity data

Apply an inclusive definition of AI/AN.

In data analysis, AI/AN should be defined as anyone marking AI/AN alone or AI/AN in any combination of race or Hispanic ethnicity, given almost half of AI/AN people identify as multiracial.[8]

> **In Practice**
> In 2020, there was a total of 30 victims of sexual violence. Of those, two victims identified as AI/AN. Person X identified as AI/AN, Black, and Hispanic. Person Y identified as AI/AI alone. To calculate the proportion of AI/AN victims of sexual violence, you would count both Person X and Person Y in the numerator (2), and all 30 victims in the denominator.

Protect privacy with data suppression

When reportable data contains less than 10 individuals, use data suppression to protect victims' privacy. Secondary suppression may be required to ensure the suppressed counts cannot be recalculated through subtraction. This occurs when additional counts that *do* exceed 10 individuals are suppressed to avoid inadvertent disclosure. If the value of all information is not the same, suppress counts that contain less meaningful information.[9]

In Practice

Agency Z is producing an annual public report of domestic violence in Jade County, with a focus on the disproportionate impact on communities of color. Among all domestic violence victims, 5 identified as AI/AN, 15 identified as Black, 15 identified as Asian, and 12 identified as White.

Race of the victims would be reported as the following, "In 2020, there was a total of 47 victims of domestic violence. By race, 15 identified as Black and 15 identified as Asian, the number of victims that identified as White and AI/AN has been suppressed to protect the privacy and confidentiality of the victims." In this case, White was chosen for secondary suppression because the focus of the annual report was on communities of color.

Aggregate racial data to limit data suppression

To limit data suppression, increase the sample size by aggregating racial data across years or meaningful geographic regions. This requires uniform data collection procedures across the years or regions of interest.

In Practice

In 2019 there were 7 AI/AN victims of domestic violence. In 2020 there were 8 AI/AN victims of domestic violence. When analyzing one-year intervals, the number of AI/AN victims could not be reported due to suppression guidelines. To overcome this challenge, Agency X aggregated the total number of AI/AN victims across two years to report 15 AI/AN victims from 2019-2020.

Analyzing tribal affiliation data

Uphold Indigenous Data Sovereignty with a Data Use Agreement or Memorandum of Understanding

Self-identified tribal affiliation is intended to capture the vast cultural diversity of AI/AN people. The collection of tribal affiliation should only be implemented if it is approved with the proper Tribal Consultation and/or Urban Confer. Permitted uses of tribally specific data and protocols for sharing data back to tribes should be outlined in a Data Use Agreement (DUA) or Memorandum of Understanding (MOU). Tribal affiliation is highly sensitive information akin to personal identity and should never be used for any purposes other than quality improvement of service delivery. If other purposes are identified, a Tribal Consultation and/or Urban Confer must be required prior to the release of information.

Reinforce the DUA/MOU with an internal agency policy

A DUA/MOU should be reinforced by an internal agency policy to ensure the confidentiality of tribal affiliation information and defer authority to govern this information to tribes. In this policy, it should be stated that tribal affiliation will not be released without the approval of each relevant Tribal Nation. If the request for tribal affiliation is research-related, additional requirements such as a tribal IRB and/or tribal resolution may be necessary. Activities that fall under this category may include informing policy, evaluating systems, or testing hypotheses. If an external agency or subcontractor has access to tribal affiliation data for quality improvement processes, the contract of agreement should recognize tribal affiliation as confidential information subject to the security precautions and protocols outlined in the DUA/MOU.

Report tribally specific information as an aggregated total

The release of tribal affiliation by individual tribe is not only a violation of Indigenous Data Sovereignty but could also expose the identity of victims, endangering their safety. In public reports, never identify individual tribes by name. Tribal affiliation should <u>only</u> be reported as the total number of tribes served. This requires tribal affiliation data to be disaggregated to calculate how many unique tribes are represented. Individuals with multiple tribal affiliations may appear more than once in the total number of tribes.

> **In Practice**
> In 2020, there were 7 assault victims affiliated with Tribe X, 12 assault victims affiliated with Tribe Y, and 1 assault victim affiliated with both Tribe X and Tribe Y. When reporting this information publicly, it should be stated as, "In 2020, Agency Z served 20 AI/AN victims across 2 different tribes". If the number of victims were less than 10, then only the number of tribes served would be reported and the number of victims would be suppressed.

Report back to tribes the total number of victims affiliated with their tribe

In sharing data on tribal affiliation back to tribes, include the total number of victims affiliated with their tribe and the total number of victims affiliated with all other deidentified tribes.

>>> **In Practice**
In Agency X's DUA, it states they will report the number of victims of sexual violence to Tribe Y each quarter. In Quarter 4 of 2020, Agency X reports the following to Tribe Y, "There was a total of 40 sexual violence victims this quarter, 25 identified with Tribe Y and 10 identified with other tribes."

Notify tribes when suppression must be exercised

If suppression is required, the Tribe should still be notified when victim(s) are affiliated with their Tribe, with the statement that the number of victims was less than 10, thus requiring data suppression to protect victim privacy.

>>> **In Practice**
In 2020, there were 5 victims of sexual violence affiliated with Tribe Y. Agency X reports the following to Tribe Y, "In 2020, the number of sexual violence victims affiliated with Tribe Y was less than 10. The exact number of victims has been suppressed to protect the privacy of the victims."

Analyzing tribal land data

Report tribal land data as an aggregated total stratified by race

Tribal land data fields should be required for all persons, not only those who identify as AI/AN. In public reports of incidents, missing persons, deceased persons, or victim's residence, specific tribal lands should never be identified as this could violate the privacy of the victim. Instead, only report the total number of incidents, missing persons, deceased persons, or victim's residences on tribal lands stratified by race and ethnicity. Stratifying or creating a group of data based on a variable, in this case, race, allows for a comparative analysis of violence on tribal lands across races, as many non-Native persons live on tribal lands.

>>> **In Practice**
In a public report of missing persons in Washington State, Agency X reports the following, "There were 100 missing persons in Washington State this year. Among those, 30 persons were last seen on tribal lands. By race, 18 identified AI/AN and 12 identified as Black." If the total number of missing persons last seen on tribal lands was less than 10, the data would be suppressed.

Suppress stratified tribal land data to protect privacy

As with other reportable data, suppression should be exercised when counts are less than 10. For tribal land data, secondary suppression may be required after it is stratified by race.

> **In Practice**
>
> In 2020, there were 18 incidents on tribal lands. Among those, 5 victims identified as AI/AN, 2 victims identified as Black, and 11 victims identified as Asian. In public reports of incidents on tribal lands, Agency X would state, "There was a total of 18 incidents on tribal lands. Among those, 11 identified as Asian. The total number of victims that identified as Black and AI/AN has been suppressed to protect the privacy and confidentiality of the victims."

Report data related to tribal lands back to tribes

When reporting data related to tribal lands back to tribes, include the total number of incidents, missing persons, deceased persons, and victim's residences on their tribal land and the total number on all other deidentified tribal lands. All data should be stratified by race and ethnicity when suppression is not necessary. To provide tribes with the most granular data possible, continue to stratify tribal land data fields by demographic and violence-related categories, including, but not limited to, age, gender, sexual orientation, homelessness, and violence type.

> **In Practice**
>
> Agency X is preparing to send data on missing persons last seen on tribal lands to Tribe Y. They first calculate the total number of missing persons last seen on tribal lands. Then they stratify by reservations in their jurisdiction.
>
> Agency X reports the following to Tribe Y, "A total of 43 missing persons were last seen on tribal lands. Of those, 21 missing persons were last seen on Tribe Y's lands and 12 persons were last seen on other tribal lands. Among those that were last seen on Tribe Y's tribal lands, 11 identified as AI/AN and 10 identified as Black. The race of the missing victims last seen on other tribal lands is suppressed due to counts less than 10."

Section V. Advanced Data Analysis

Advanced statistical analyses can be used to identify trends in the data and compare the occurrence of violence between populations. When conducting statistical tests or population comparisons, it is recommended to use a p-value threshold of 0.05. A p-value less than 0.05 signifies statistical significance.

The relative standard error (RSE) can be used to determine unreliable estimates. It is recommended to use an RSE threshold of 30%. When estimates exceed 30% or 0.30, rate instability should be noted.

Key suggested thresholds

P-value = 0.05
For statistical significance

Relative standard error (RSE) = 30%
For reliability of rates

Generating rates of violence

A rate is the occurrence of events over a specific time period. Rates can be used to compare the proportion of events between population groups of different sizes. To calculate the rate of violence, first calculate the total number of incidents of violence. Then, divide the total incidents by the specified population total. To calculate the rate per 100,000, multiply the rate by 100,000 (see Example 1).

As noted previously, it is recommended to apply an inclusive definition of AI/AN. Total population counts of AI/AN alone or AI/AN in any combination of race and Hispanic ethnicity can be obtained from data.census.gov using the American Community Survey. Data are available down to the county level. Additional data on AI/AN population counts may be obtained from state, territorial, and tribal population forecasting organizations. Prior to using any dataset first, determine how AI/AN is defined (e.g. AI/AN alone or AI/AN in any combination of race or Hispanic ethnicity), as data sources may define AI/AN differently.

Example 1

$$\frac{\text{Number of AI/AN victims of sexual violence in Washington*}}{\text{Total AI/AN population in Washington*}} \times 100{,}000$$

*AI/AN defined as AI/AN alone or in any combination of race and ethnicity

Best practices for small populations data analysis

- Recognize that statistical insignificance or a p-value greater than 0.05 does not mean disparities do not exist. Statistical significance indicates whether the relationship between variables is likely not caused by chance.

- Apply weighted sampling to increase statistical power and representation in the data. Weighted sampling is used when your sample population over-or under-estimates your target population. To calculate the statistical weight, divide the target population by the sample population and use the result as a multiplier in your calculations.

- Acknowledge the limitations of the data, including the degree of missingness and racial misclassification,[10] data sources utilized, and statistical assumptions.

- Indicate when descriptions of AI/AN are not inclusive of multiracial individuals and when methodologies for collection of race/ethnicity are not uniform.

- Perform regular evaluations of methodologies to identify areas for improvement.

References

1. Urban Indian Health Institute (UIHI). (2020). *Best Practices for American Indian and Alaska Native Data Collection.* Retrieved Oct 23, 2020 from https://www.uihi.org/resources/best-practices-for-american-indian-and-alaska-native-data-collection/
2. Rainie, Stephanie Carroll, Desi Rodriguez-Lonebear, and Andrew Martinez. 2017. Policy Brief: Indigenous Data Sovereignty in the United States. Tucson: Native Nations Institute, University of Arizona
3. Kukutai, T., & Taylor, J. (2016). *Indigenous data sovereignty: Toward an agenda.* Canberra, Australia: ANU Press.
4. Walter & Andersen. (2013). *Indigenous Statistics: A Quantitative Research Methodology.* Routledge Taylor & Francis Group.
5. U.S. Census Bureau, Understanding and Using American Community Survey Data: What Users of Data for American Indians and Alaska Natives Need to Know, U.S. Government Printing Office, Washington, DC, 2019.
6. Santa Clara Pueblo v. Martinez, 436 U.S. 49 (1978).
7. National Conference of State Legislatures (2012). *Trust Land.* Retrieved February 10, 2021 from https://www.ncsl.org/research/state-tribal-institute/trust-land-overview.aspx
8. Census. (2010). *2010 Census Shows Nearly Half of American Indians and Alaska Natives Report Multiple Races*. Retrieved Oct 23, 2020 from https://www.census.gov/newsroom/releases/archives/2010_census/cb12-cn06.html
9. Washington Department of Health. (2018). Department of Health Agency Standards for Reporting Data with Small Numbers. Retrieved Oct 23, 2020 from https://www.doh.wa.gov/Portals/1/Documents/1500/SmallNumbers.pdf
10. Joshi, S & Warren-Mears, V. (2019). Identification of American Indians and Alaska Natives in Public Health Data Sets: A Comparison Using Linkage-Corrected Washington State Death Certificates. *Journal of Public Health Management & Practice*, Volume 25, page S48-S53.