



Census Bureau, Department of Commerce

Re: Suggestions on 2030 Census Preliminary Research

On behalf of the Center for Data Innovation ([datainnovation.org](http://datainnovation.org)), I am pleased to submit this letter in response to the Census Bureau's request for comment on preliminary research for the 2030 Census.<sup>1</sup>

The Center for Data Innovation is the leading think tank studying the intersection of data, technology, and public policy. With staff in Washington, London, and Brussels, the Center formulates and promotes pragmatic public policies designed to maximize the benefits of data-driven innovation in the public and private sectors. It educates policymakers and the public about the opportunities and challenges associated with data, as well as important data-related technology trends. The Center is a part of the nonprofit, nonpartisan Information Technology and Innovation Foundation.

In preparation for the 2030 Census, the Center would encourage the Census Bureau to consider its role in closing the data divide. The data divide refers to the social and economic inequalities that result from a lack of data collection and use of data by certain individuals and communities.<sup>2</sup> In a world where opportunities and services are intricately linked to data, the way people are reflected in datasets and their ability to use datasets about themselves significantly impacts their ability to participate in the data economy.

Census data is important for data-driven decision making and understanding quality of life among different demographic groups and geographies in the United States. The Census Bureau should pay particular attention to how it can improve data representativeness in the 2030 Census because incomplete data impacts the data divide.

Data representativeness refers to the extent to which a given dataset sufficiently describes the characteristics of a larger population. Sometimes, datasets exclude a population outright or exclude certain details about a population. For example, the Census Bureau acknowledges that it has historically undercounted the American Indian and Alaska Native (AI/AN) community, particularly those groups living on reservations.<sup>3</sup> The Census Bureau also acknowledges that it has struggled to count individuals in group housing (e.g., college dorms, nursing facilities, homeless shelters, etc.).<sup>4</sup> The Census Bureau should pursue innovative methods to increase data representativeness, including leveraging new data sources (e.g., third-party commercial

---

<sup>1</sup> "Soliciting Input or Suggestions on 2030 Census Preliminary Research," Census Bureau, August 2022,.

<sup>2</sup> Gillian Diebold, "Closing the Data Divide for a More Equitable U.S. Digital Economy," (Center for Data Innovation, August 2022), <https://datainnovation.org/2022/08/closing-the-data-divide-for-a-more-equitable-u-s-digital-economy/>.

<sup>3</sup> "American Indians and Alaska Natives Living on Reservations Have the Highest 2020 Census Undercount," National Congress of American Indians, March 10, 2022, <https://www.ncai.org/news/articles/2022/03/10/american-indians-and-alaska-natives-living-on-reservations-have-the-highest-2020-censusundercount>.

<sup>4</sup> Robert Santos, "It's not too late to submit your ideas for the 2030 Census!" October 17, 2022, Census Bureau, <https://www.census.gov/newsroom/blogs/director/2022/10/submit-your-ideas-2030-census.html>.



data and federal, state, and local administrative records) as well as greater use of technology, including Internet-based data collection.

To improve data representativeness for the 2030 Census, the Census Bureau will need to better understand data quality in the 2020 Census. Unfortunately, external reviewers have not had access to the detailed data necessary to evaluate data quality metrics for the 2020 Census. The American Statistical Association (ASA) task force that evaluated the 2020 Census noted that it did not get access to data below the state level or data with characteristics of household occupants that it needed to assess data quality.<sup>5</sup> Likewise, an interim report from the National Academies of Sciences, Engineering, and Medicine on the quality of the 2020 Census also noted that characterizing the quality of the 2020 Census would be impossible without data at substate levels.<sup>6</sup> The Census Bureau should make this information available to these external reviewers and update its plans for the 2030 Census to provide information necessary to evaluate data quality in a more timely manner.

Prioritizing closing the data divide will ensure that all communities have the opportunity to benefit from the 2030 Census and subsequent data-driven policymaking. Data can transform society for the better, but only if data collection is inclusive of all Americans.

Sincerely,

Gillian Diebold  
Policy Analyst  
Center for Data Innovation  
gdiebold@datainnovation.org

---

<sup>5</sup> "ASA Releases Task Force Report on 2020 Census Analysis and Replicability Report," American Statistical Association, 2021, <https://www.amstat.org/news-listing/2021/09/14/asa-releases-task-force-report-on-2020-census-analysis-and-replicability-report>.

<sup>6</sup> National Academies of Sciences, Engineering, and Medicine, *Understanding the Quality of the 2020 Census: Interim Report*, (National Academies Press, 2022), <https://nap.nationalacademies.org/download/26529>.