A Deeper Look At:
Census & Differential Privacy

Differential Privacy (DP): a new privacy methodology intended to be applied beginning with the 2020 decennial census.

Background: The U.S. Census must carefully monitor data accuracy and privacy when making census results and data public. In recent testing for privacy vulnerabilities, the census concluded that the privacy protection methods used for decennial census data in the past no longer fulfill Title 13 of the United States Code, which details the obligation of an accurate population count and mandates confidentiality. Although the census data itself may not disclose personal identities, modern data science and computing technologies may now be combined with other public and private data sets to re-identify information specific to an individual. *(Source: Iowa State University Department of Economics, ICIP)*

How does DP work?

DP methodology typically involves adding and subtracting random "noise" numbers to actual values from a data set. "Noise" helps to mask original data enough to avoid analysis that could reveal information at an individual level. This could apply to population counts and other demographic data, such as age, gender, race and income levels. The goal is to increase individual privacy in the census process.

Does adding noise reduce the accuracy of the data?

Because this process involves injecting noise data to census counts, the accuracy of the data will be reduced. The goal of a DP framework is to introduce just enough noise to protect privacy while retaining enough accuracy for reliable data analysis.

To illustrate this, the census has released multiple sets of test data, demonstrating how DP would have impacted 2010 census data as a baseline. Each set of test data appears to improve its accuracy, however, the data still show impacts on data accuracy. *(Source: Iowa State University Department of Economics, ICIP)*

How much will the published Census data differ from actual values?

This is still unknown, but the census recently released demonstration data sets based on 2010 census data that can help us gauge the potential effects of DP for different geographic area types and population groups, including cities and counties in Iowa.

In the latest test data set (9/2020), 144 cities incorrectly showed a count decrease of more than 10% compared to the 2010 census and 84 cities incorrectly showed an increase of more than 10%. Individual cities showed up to a 56% decrease in population counts and up to a 100% increase. In theory, the State total population count should not be changed. However, the mean average percentage change for individual cities was 6.6%. For cities under 500 in population in 2010, that average was 11.2%.
(Raw data source: IPUMS NHGIS, University of Minnesota, www.nhgis.org)

IOWA LEAGUE of CITIES

(515) 244-7282

research@iowaleague.org

www.iowaleague.org

How does this affect cities in Iowa?

While there are many potential impacts, a few include: federal and state funding reductions or increases at the city level that are based on decennial census counts, social programming impacts, city planning, potential state and local redistricting/re-precincting, impacts on housing data, etc. The League is concerned that some individual cities, especially rural communities, will see reduced state and federal funding due to DP.

Can you give me an example of a city revenue that depends at least partly on population?

At the federal level, 316 programs use census-derived data to distribute $1.5 trillion to local governments, businesses, nonprofits and households across the nation (7.8% of GDP in FY17). Iowa's fair share depends on the accuracy of its data count. *(Source: George Washington University Institute of Public Policy)*

Local option sales tax and road use tax fund are two state-shared revenue sources in Iowa that use census numbers as a part of their distribution formula.

What about potential undercounts, generally, or due to DP?

If there were an undercount in Iowa in the 2020 census, Iowa would lose portions of federal funds whose distribution formulas are based on census populations. If there were an undercount of just one-tenth of a percent, this would equate to an approximately $90M loss over 10 years for Iowa.
(Source: League analysis of raw data from George Washington University)

See www.iowaleague.org/_layouts/media/Census2020.pdf