

Challenges Facing the Disclosure Review Board at Census

August 3, 2016

Billy Wisniewski and Holly Newman

Center for Disclosure Avoidance Research (CDAR)

U.S. Census Bureau

Disclaimer: Any views expressed are those of the authors and not necessarily those of the U.S. Census Bureau

Outline

- Disclosure Review Board (DRB)
- Disclosure Review Board Coordinator (DRBC)
- Disclosure Avoidance Officer (DAO)
- Walkthrough of Challenging Requests

Census Bureau Requirements

- Disseminate high quality data to the public
- Promise to protect the confidentiality of individual respondents
 - Title 13
 - Title 26
 - Confidential Information Protection and Statistical Efficiency Act (CIPSEA)

Disclosure Review Board

- Mission: to ensure that the Census Bureau protects the confidentiality of all Title 13 and Title 26 publicly released data products
- Who is in the DRB?



Disclosure Review Board

- Variety of Roles
 - Establish and review disclosure avoidance policies
 - Review and approve disclosure avoidance procedures
 - Communicate disclosure avoidance policies to others
 - Coordinate research on the disclosure risk of data products
 - Determine how effective the disclosure avoidance techniques are in protecting data products
 - If necessary, revise and update the Census Bureau's disclosure avoidance policies and methodologies

Disclosure Review Board Coordinator

- Plays an important role in the DRB
- Duties:
 - Ensures that all DRB members and Disclosure Avoidance Officers are compliant and up to date with the annual training
 - Escorts visitors who do not have Title 26 training
 - Organizes requests for each meeting
 - Collects meeting minutes
 - Sends out approval/denial memos
 - Uploads minutes and memos to CDAR secure shared drive and CDAR intranet page



Disclosure Avoidance Officer

- Assists researchers and DRB submitters going through the DRB review process, completing (or helping complete):
 - Memo to the Chair of the DRB explaining the request
 - DRB Checklist
 - Questionnaire from the survey or census
 - A list of variables of interest
 - A record layout for requested microdata
 - Table outlines/shells
 - Some cross-tabulations of the variables of interest
- Each program area that releases data appoints a DAO

Disclosure Avoidance Officer

- Performs the final review of data products before sending them to the DRB or into the public domain
- Completes DAO Checklist
 - This helps the DAO go through the review process to make sure that all disclosure avoidance techniques were applied correctly



Examples of Challenging Requests

1.) Releasing All Variables as part of a Public Use File (PUF)

- PUF: microdata data file where each record represents value/characteristics of a particular respondent
- The issue? Re-identification, leading to violation of Census's imperative to confidentiality
- What to do?

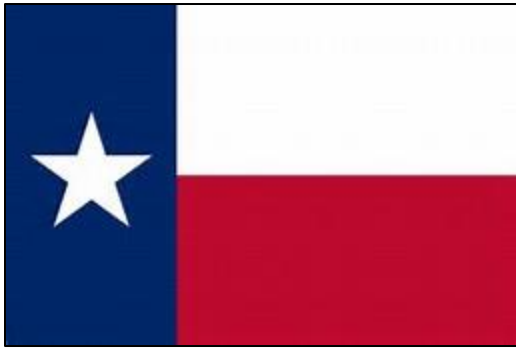
What to do with a PUF request?

- Direct Identifiers
- Geographic Thresholds
- Rounding Content
- Topcoding and Bottomcoding
- Categorical Thresholds
- Data Swapping
- Noise Infusion

2.) Requesting tabular data for every county in the U.S.?

- Tabular data: count/frequency tables
- Almost all cases use American Community Survey or Decennial data
- Standard tables vs. special tabulations
- The issue? Small cells can lead to re-identification

LOVING COUNTY, TEXAS (population 103)



What to do with count/frequency tables?

- Data swapping
- Synthetic data
- Special tabulation rules:
 - Rounding counts
 - Cell thresholds/universe thresholds/categorical thresholds
 - Other techniques

3.) Requests for Economic Magnitude Data

- Economic Magnitude Data: tables containing establishment firms along with an aggregate of some quantity of interest (e.g., sales) over all units in each cell
- Establishment counts (alone) are not considered sensitive
- Magnitude values can be sensitive, must be protected
 - Protection at the firm level
 - Determined by sensitivity rule (e.g., p-percent rule)

Economic Magnitude Data

- The issue?: Re-identifying an attribute, as users can closely approximate the true confidential values of dominating firms
- What to do?
 - Use Cell Suppression

Economic Magnitude Data

- Cell Suppression
 - Use sensitivity rules to determine sensitive cells and suppress from table (primary suppressions)
 - Suppress other cells (complementary suppressions) to protect the primaries
 - Can use noise infusion or synthetic data to create magnitude tables

4.) Requests from the FSRDCs

- Output originating from one of the 23 Federal Statistical Research Data Centers
- Sometimes, subject to DRB review
 - Level of geography below state-level
 - Researcher asks for modification of a prior guideline

Requests from the FSRDCs

- Count rules, concentration ratios
- Other rules to follow:
 - Rounding Ns
 - Interpolating medians
 - Keeping track of all samples, including implicit samples within a project

Conclusion

- DRB, DAOs and DRB Coordinator all play a critical role
- Current disclosure avoidance techniques ensure protection of respondents in data releases
 - DRB facilitates the efficient process of data releases
- DRB will obtain difficult requests from time to time
 - Compromise on content can usually achieve a useful, safe data product

Questions?

- William.Wisniewski@census.gov
- Holly.Newman@census.gov