



CTPP Status Report



U.S. Department of Transportation
Federal Highway Administration
 AASHTO Standing Committee on Planning



TRB Census Subcommittee
Bureau of Transportation Statistics
Federal Transit Administration

Census Transportation Planning Products (CTPP) Highlights

Penelope Weinberger, AASHTO,
pweinberger@ashto.org

Hello All,

The CTPP program oversight board had excellent strategic meeting in February in Fort Lauderdale, FL, but many of the items discussed were placed on hold as the world deals with the current pandemic. However, the CTPP still and will continue to provide an excellent source of data for where your workforce lives; class, industry, and occupation of workers; and other demographic data, including home location. If you are analyzing how many workers are out of work, working from home, and still traveling to work, the CTPP can give you detailed baseline data to quantify these workforce segments and changes. The oversight board are starting to consider and discuss how the pandemic will impact next data tabulation from the US Census Bureau, beyond the obvious of number of workers, travel to work and workplace.

Stay well, stay safe.

Census Data at TRB

Clara Reschovsky, Census Subcommittee
 co-Chair, clara.reschovsky@dot.gov

The Census Data Subcommittee, a subcommittee of TRB's Urban Data Committee (ABJ30), had a visible presence at the TRB Annual Meeting in January 2020 in Washington, DC. The Subcommittee sponsored a poster session, as well as having an in-person meeting. The poster session, Census Transportation Planning Products 2012-2016: New Data Analyses and Uses, was packed with 12 great posters

submitted from a mix of metropolitan planning organization (MPO) and local government staff, consultants, and academics. Topics ranged from autonomous vehicles adoption, transit ridership, work from home trends, and worker characteristics. The poster session was crowded with Annual Meeting attendees wanting to learn more about the CTPP through the posters.

The annual subcommittee meeting was well attended, and the group had a robust discussion of the material presented. We heard an interesting presentation from Kevin Tierney on his poster, "Using Available Data Sources on Work-at-Home Characteristics to Forecast the Potential Effects of Telecommuting Incentives." The poster presented analysis on the demographics of people who report working at home in the ACS in Massachusetts as the State considers incentivizing working at home to reduce congestion. In preparation for the next Commuting in America product, two presentations were given on topics with the intent of getting feedback. The first, "Vehicle Availability Patterns and Trends including Commuting Trends for Zero-Vehicle Households," was presented by Evelyn Blumenberg and Julene Paul, both of UCLA. They presented their preliminary findings, asked for feedback on their findings and directions to take with their research. The other CIA presentation was "The Changing Nature of Work" by Nancy McGuckin. She also presented preliminary results and asked for feedback.

Customized Polygons Now Can Be Used for Selecting Geographies

Penelope Weinberger, AASHTO,
pweinberger@ashto.org

Jiangbo Yu, Cambridge Systematics, Inc., jyu@camsys.com

Have you ever wished that you could upload your own polygons to the CTPP data access website instead of manually selecting zones? Now you can!

This new function allows users to upload custom polygons to assist in their selection for the place-of-residence and place-of-work geographies of their interest. The following article is dedicated to introducing this new function. The April 2019 status report provides general information about the CTPP Data Access Website. You also are encouraged to explore the website itself by going to <http://data5.ctpp.transportation.org/>.

Below are some important details and limitations of this new function:

- Custom polygons functionality is only available to authenticated users of the software. In other words, a user needs to sign in to the CTPP website.

- Custom polygons are only available to the user (owner) who uploaded it. Currently, you cannot share the information about the uploaded polygons with other users through the Data Access Website. GIS file formats including Esri Shapefile or Mapinfo TAB file are supported. The files need to be in a zip format with no other folders included.
- Currently, the file size upload limit is set at 50 MB and a maximum of 500 polygons. Both of these limits may increase in the future.

Where is the new function located?

After you have been authenticated, signed in to the CTPP data access website and chose the dataset of interest, select “Residence” or “Workplace,” and then click “Selection map,” as shown in Figure 1.

Once in the tab “Selection map,” you will see the button “Custom Polygons” in the top bar of the mapping area, as shown in Figure 2.

The custom polygons layer will be available in the drop-down option “Zoom In & Select,” as shown in Figure 3.

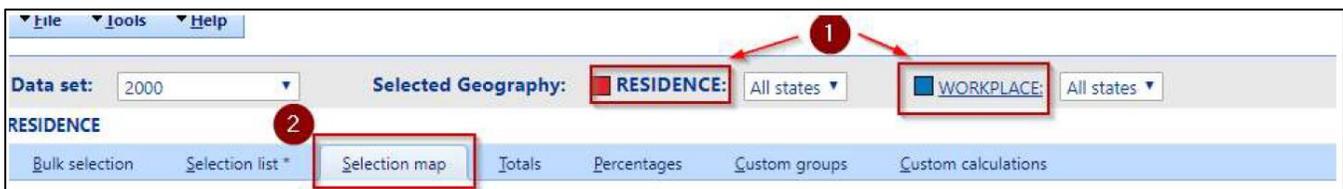


Figure 1. The first step is to select “Residence” or “Workplace” and then click “Selection map.”

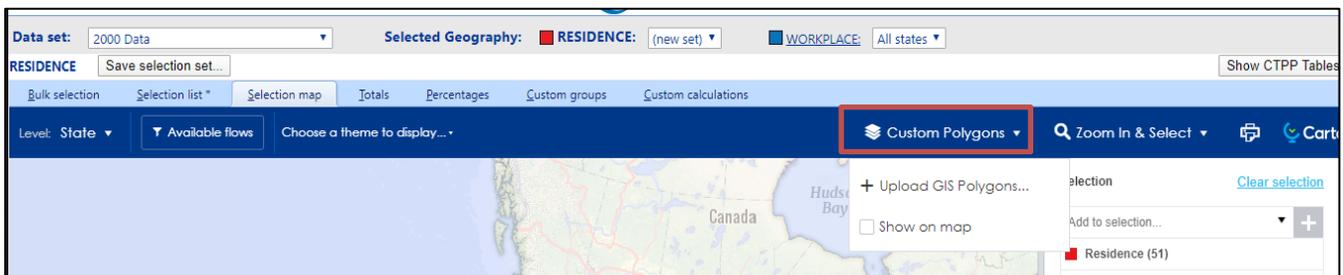


Figure 2. The drop-down item “Custom Polygons” (highlighted with a red rectangle) is in the top bar of the mapping area.

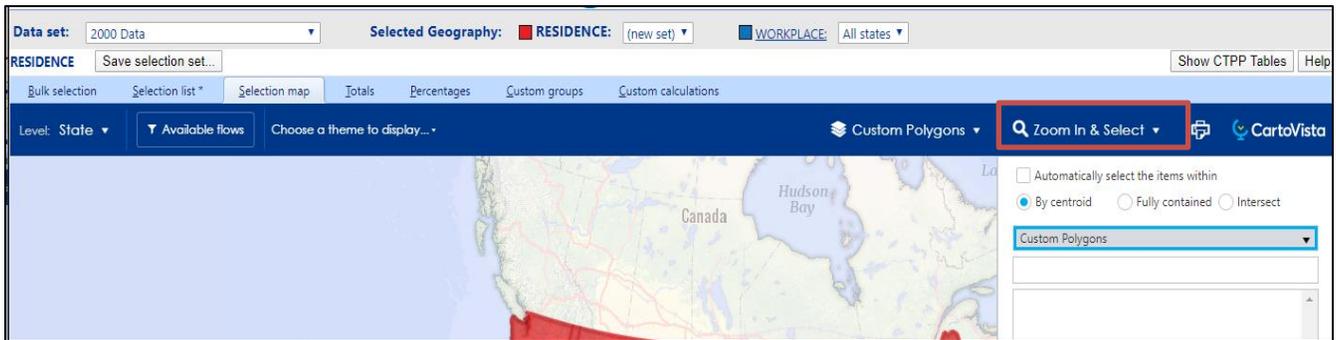


Figure 3: The drop-down item “Zoom In & Select” (highlighted with a red rectangle) is in the top bar of the mapping area.

How to upload my customized polygons?

Clicking the “+ Upload GIS Polygons...”, will open a dialog box “Add Data,” as shown in Figure 4.

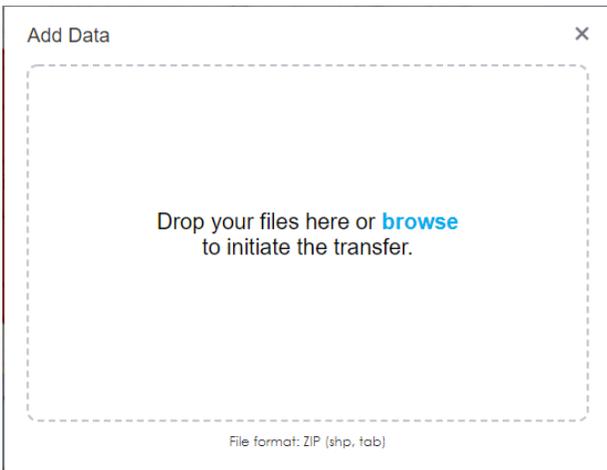


Figure 4. The pop-up dialog box facilitates the polygon upload process.

You can drag the polygon zip file from your local drive to the dialog box area or click “Browse” to choose the polygon zip file. A progress bar will be present while the file is being uploaded. Once the upload is complete, the user *must* select a column to be used for labeling the geographic feature from a dialog box similar to that in Figure 4. Click on the “Add” button to complete the upload, as shown in Figure 5.

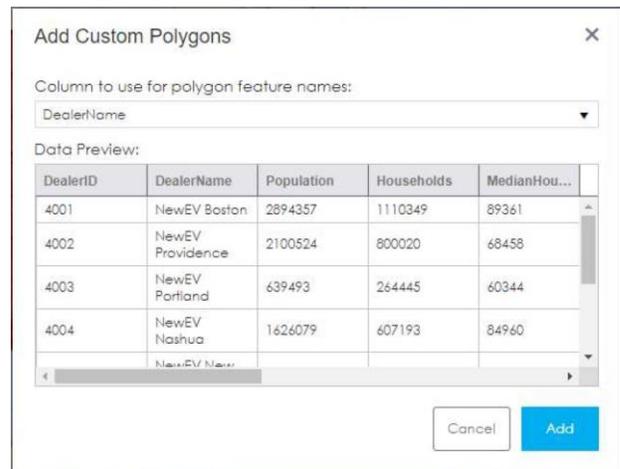


Figure 5. The preview of the attribute table of the uploaded geographic file will be shown for verification. A user should select the column to use for polygon feature names before proceeding.

Once the upload process has finished, the custom polygons will appear in the drop-down manual “Zoom In & Select,” as shown in Figure 6. Users can remove one or all of the uploaded custom polygons.

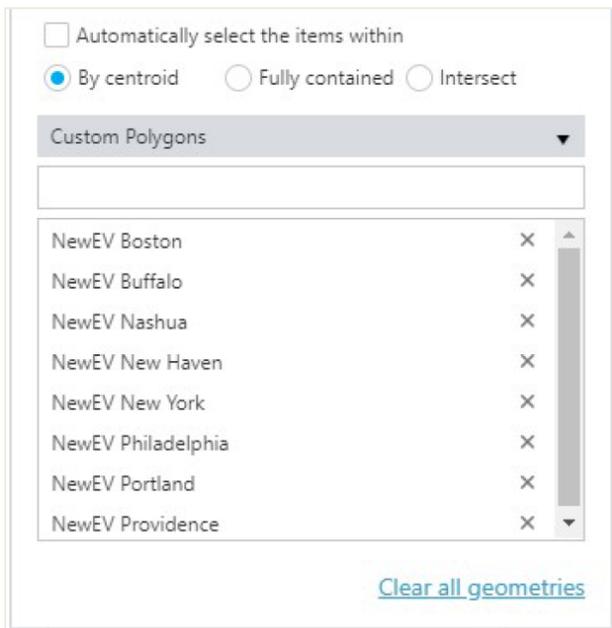


Figure 6. The uploaded custom polygons can be accessed through the drop-down manual “Zoom In & Select.”

Visualizing the uploaded custom polygons

In the “Custom Polygons” drop-down list, there is a checkbox “Show on Map,” as highlighted in Figure 7. When checked, the uploaded custom polygons will be visible in the map.

The use of the uploaded custom polygons to select the standard geographies

The uploaded custom polygons can be used to select standard geographies stored in the CTPP data access website. As shown in Figure 8, after choosing the appropriate geographic level, select “Automatically select the items within” with a specific intersection type.

When the uploaded polygons are visible, you also can click on them to open a contextual manual allowing you to use the selected polygon for spatial intersection on the current level, as shown in Figure 9. The selection set can be further adjusted through the standard map selection functionality.

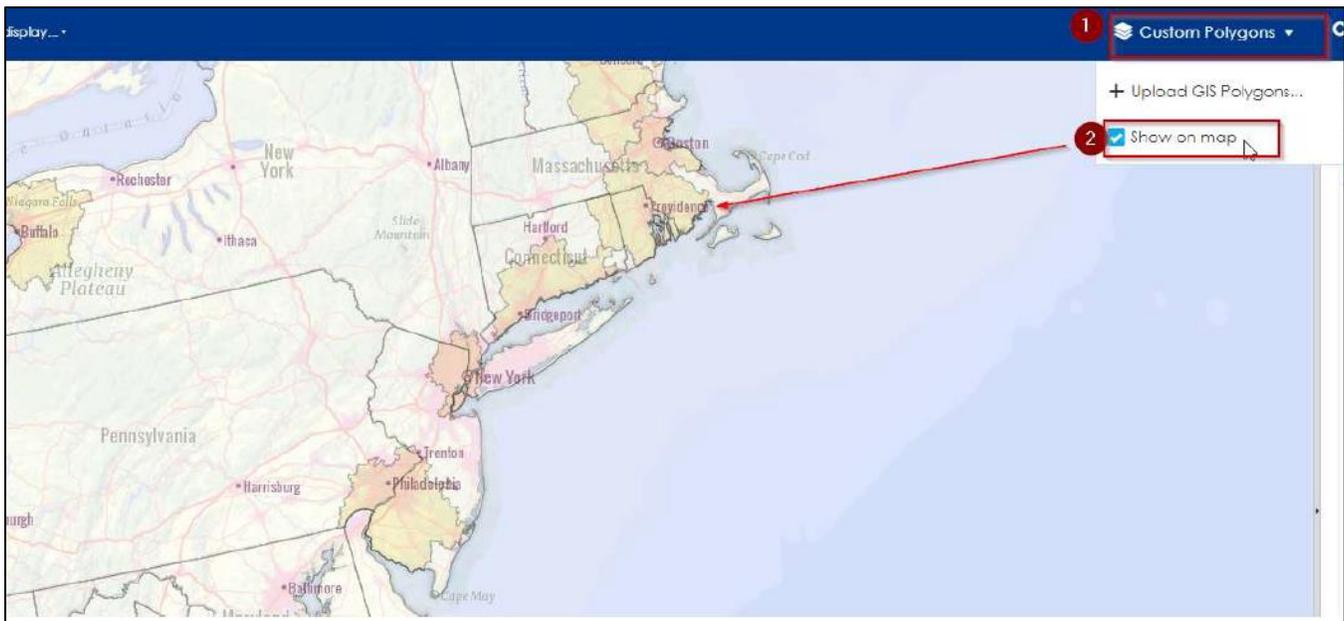


Figure 7. Toggle on the “Show on map” checkbox to visualize the uploaded polygons in the map.

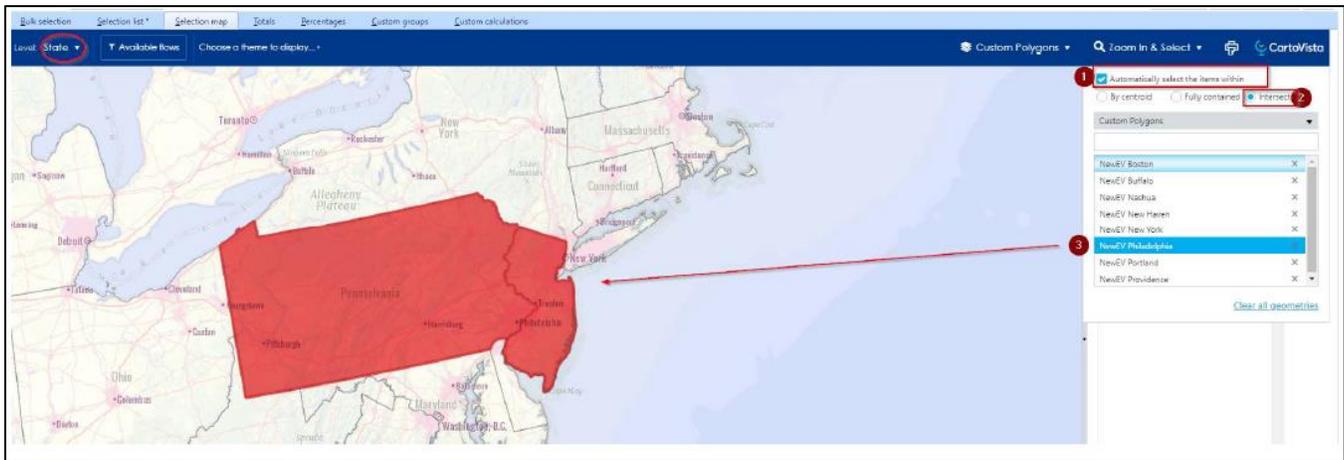


Figure 8. Standard geographies can be selected using the newly uploaded custom polygons, based on “centroid,” “fully contained,” or “Intersect” criteria.



Figure 9. When the uploaded polygons are visible, you also can click them directly on the map allowing you to use the selected polygon for spatial intersection on the current level.

Differential Privacy and Transportation

Krishnan Viswanathan, Cambridge Systematics, Inc., kviswanathan@camsys.com

What is Differential Privacy?

The Census Bureau (CB) is bound by Title 13 of the United States Code. These laws not only provide authority for the work the Bureau does, but also provide strong protection of the information the Bureau collects from individuals and businesses. To allay concerns of

respondents, Census products are reviewed by the Disclosure Review Board (DRB) to ensure that no identifiable information is released. Until and including the 2010 Census, the CB would disclose proof the data using suppression and swapping techniques. The swapping technique involves some households being “swapped,” such that whole households are moved from one block to another to eliminate the possibility of identifying unique households. Table 1 shows the evolution of disclosure avoidance techniques over the decades.

Table 1. Disclosure Avoidance Techniques Over Time¹

		Table Suppression	Swapping	Blank and Impute	Partially Synthetic Data
1970					
	100% Data	X			
	Sample Data	X			
	Households	X			
1980					
	100% Data	X			
	Sample Data	X			
	Households	X			
1990					
	100% Data		X		
	Sample Data			X	
	Households		X	X	
2000					
	100% Data		X		
	Sample Data		X		
	Households		X		
2010					
	100% Data		X		X
	Households		X		
	Group Quarters				X

Data Source: McKenna, L., 2018.

The increasing use and availability of data from different sources such as credit bureaus, big data sources, and smartphone applications, along with the advent of increasing computing power, makes it possible, with some effort, to reconstruct the identity of the respondent. Therefore, the CB has identified differential privacy as a means to enhance disclosure avoidance starting with the 2020 Decennial Census.

Differential privacy works to prevent accurate reconstruction attacks while making certain that the data are still useful for statistical analyses. It does this by injecting a controlled amount of uncertainty in the form of mathematical randomness, also called noise, into the calculations that are used to produce data products. The range of noise can be shared publicly because an attacker cannot know exactly how much noise was introduced into any particular table. With differential privacy, it is still possible to reconstruct a database, but the database that is reconstructed will include privacy-ensuring noise. In other words, the

individual records become synthetic byproducts of the statistical system.

Multiple ways to build a system to protect the confidentiality of census data under differential privacy exist. The Census Bureau opted to focus on a top-down approach to maximize utility for redistricting and funding allocation. By understanding how data users might cross tabulate variables or otherwise conduct statistical analyses, noise could be added to cells in the tables, such that it would be difficult to discern individuals. To do this well requires understanding how different statistical tables interact with each other. A top-down approach can produce more tailored statistical tables because noise is injected in ways that preserve the most important statistical efforts. If race is the most important variable, more noise can be introduced to the age variable to preserve the accuracy of the race variable. If one table is more important than another, more noise can be provided to the less-frequently-used table than the one that is more essential. While this control

is an asset, it also introduces significant governance and policy implications.

Noise is introduced to statistical tables to reduce information leakage. Within the construction of a differentially private system, it is necessary to make choices about how much noise is tolerable and how much risk, or total privacy loss, is acceptable. The creators of a differentially private system set a maximum cumulative privacy-loss budget. This is defined by the variable *epsilon* (between ~0 and infinity). Details of how Epsilon works, and its implications are from Boyd (2019). Epsilon can be conceptually understood as a set of knobs. Dial the knobs one way to create higher levels of noise, but lower levels of confidentiality risk.

Dial them the other way and the data are more accurate, but the risk to confidentiality (and, therefore, “privacy loss”) is higher. The system has a global privacy-loss budget, but each geographic level and table also has a local privacy-loss budget that must be managed, such that the interaction of all tables does not result in leakage that exceeds the global amount. The actual local budget is a combination of the particular table and geographic level. Because of these locally defined epsilons, it is possible to introduce more noise to a specific variable (e.g., sex) than others (e.g., race).

The differences between the 2000/2010 Census and the 2020 Census are shown in Figure 10 and Figure 11.

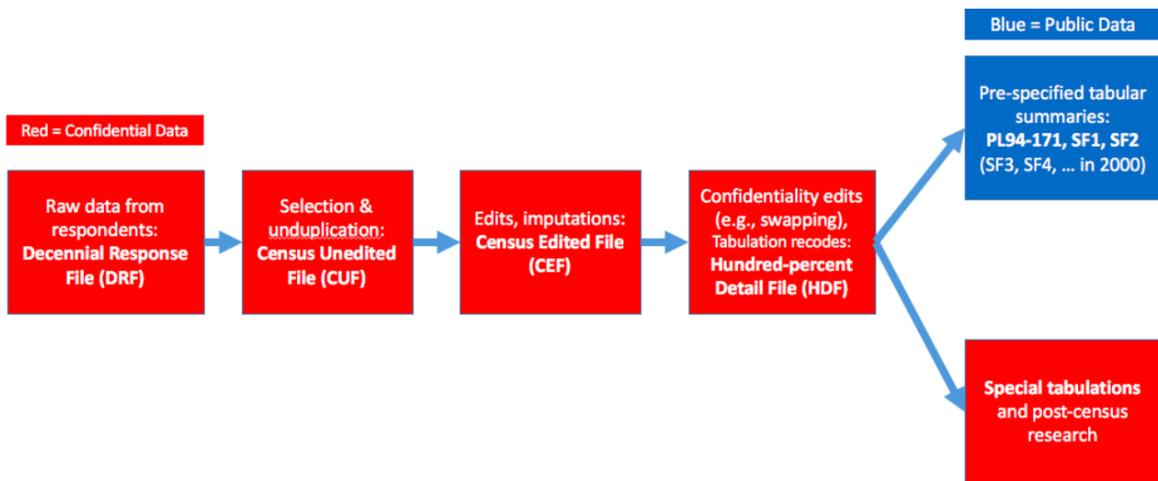
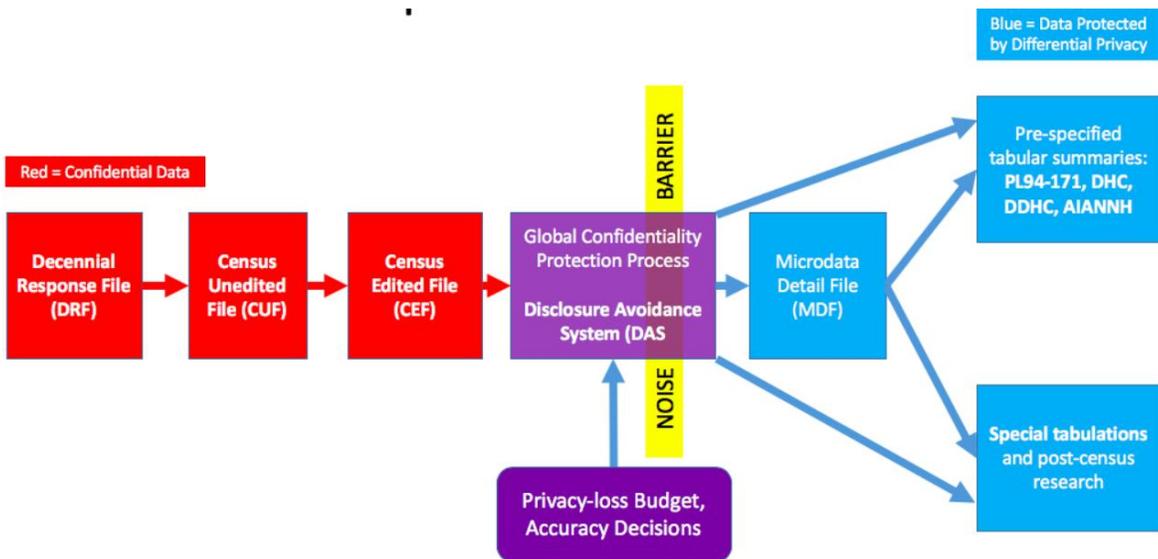


Figure 10. Data Collection and Processing Pipeline (2000 and 2010)



Data Source: Boyd, D., 2019.

Figure 11. Data Collection and Processing Pipeline (2020).

To implement differential privacy, two major questions must be determined before the disclosure avoidance system (DAS) is run: the global privacy-loss budget (or global epsilon) and the allocation of the budget across all of the tables and features (local privacy-loss budget allocation). This means determining all of the data products that will be produced in advance. Further, certain things have to be kept **invariant**—data that cannot be altered, such as State population, which further constrains the data that will be released. In addition, to prevent confusion, the DAS team has decided that all cells must contain positive whole numbers. These decisions by the CB have implications for the transportation community, which is dealt with in the next section.

Transportation Implications

The transportation community relies on small geography data (traffic analysis zones, block groups, parcels etc.) in order to do the analysis expected by transportation modelers and planners. If moving forward, the CB decides to release data that is suppressed using differential privacy, this can lead to loss of confidence by transportation modelers and planners in the data produced by the CB. Below we consider the implications of the decisions made by the CB:

- **Determine data products in advance**—Because of the need for a global privacy loss budget and its allocation across all tables and features, the CB will determine what products will be produced and the transportation community might not be able to weigh in and ultimately get the products it needs for analytical and planning purposes.
- **Invariants**—While the CB has not yet determined the 2020 invariants, there is hope that they will consider small geographies. If they decide to keep invariants at the state level, small geography data will be suppressed and we will end up with

mathematical impossibilities like households greater than household population as demonstrated by Jarosz (2019) using the differential privacy (DP) protected demonstration data released by the CB. Using the DP demonstration data in California, Jarosz (2019) found that in California for one county (out of 58) and 63 places (out of 1,523), the households were greater than the household population.

- **Only positive whole numbers**—Adding noise without going negative implies that small blocks will get bigger and in turn, because the state level number is kept invariant, the larger blocks will get smaller. This can distort the TAZs used by transportation planners and modelers and has implications for data quality and analysis results.

Given these impacts, is necessary to devise tests with the demonstration data to determine implications for the transportation community.

References

- Boyd, D., 2019, “Balancing Data Utility and Confidentiality in the 2020 US Census”.
https://datasociety.net/wp-content/uploads/2019/12/Differential_Privacy_12_172019.pdf.
- Jarosz, B., 2019, “Importance of Decennial Census for Regional Planning in California”. Presented at the Workshop on 2020 Census Data Products: Data Needs and Privacy Considerations.
https://sites.nationalacademies.org/cs/groups/dbasse/documents/webpage/dbasse_197498.pdf.
- McKenna, L., 2018, “Disclosure Avoidance Techniques Use for the 1970 through 2010 Decennial Censuses of Population and Housing,” Census Research and Methodology Directorate, US Census Bureau, Washington, DC:
<https://www2.census.gov/ces/wp/2018/CES-WP-18-47.pdf>.

CTPP Contact List

Email: CTPPSupport@camsys.com

2012-2016 CTPP Data: <https://ctpp.transportation.org/2012-2016-5-year-ctpp/>

CTPP website: https://www.fhwa.dot.gov/planning/census_issues/ctpp/

FHWA website for Census issues: https://www.fhwa.dot.gov/planning/census_issues

AASHTO website for CTPP: <https://ctpp.transportation.org>

2006-2010 CTPP Data: <https://ctpp.transportation.org/ctpp-data-set-information/5-year-data/>

1990 and 2000 CTPP data downloadable via Transtats: <https://transtats.bts.gov/>

TRB Subcommittee on census data: <http://www.trbcensus.com>

AASHTO

Penelope Weinberger

Phone: (202) 624-3556

Email: pweinberger@aaashto.org

Jessie Jones, ARDOT

Chair, CTPP Oversight Board

Phone: (501) 569-2201

Email: Jessie.Jones@ahtd.ar.gov

Guy Rousseau, Atlanta Regional
Commission

Vice Chair, CTPP Oversight Board

Phone: (404) 463-3274

Email: GRousseau@atlantaregional.com

U.S. Census Bureau: Social, Economic and Housing Statistics Division

Brian McKenzie

Phone: (301) 763-6532

Email: brian.mckenzie@census.gov

Federal Transit Administration (FTA)

Ken Cervenka

Phone: (202) 493-0512

Email: ken.cervenka@dot.gov

Bureau of Transportation Statistics (BTS)

Clara Reschovsky

TRB Census Subcommittee co-Chair

Phone: (202) 366-2857

Email: clara.reschovsky@dot.gov

Federal Highway Administration (FHWA)

Joseph Hausman

Phone: (202) 366-9629

Email: Joseph.Hausman@dot.gov

TRB Committees

Stacey Bricka

Consultant

Chair, TRB Urban Data Committee

Email: Sgbricka@gmail.com

Mara Kaminowitz

TRB Census Subcommittee Co-Chair

Phone: (410) 732-0500

Email: mkaminowitz@baltometro.org

CTPP Technical Support

Jingjing Zang

Phone: (301) 347-9100

Email: CTPPSupport@camsys.com

CTPP Listserv

The CTPP Listserv serves as a web-forum for posting questions and sharing information on Census and ACS. Currently, more than 700 users are subscribed to the listserv. To subscribe, please register by completing a form posted at: <http://www.chrispy.net/mailman/listinfo/ctpp-news>.

On the form, you can indicate if you want emails to be batched in a daily digest. The website also includes an archive of past emails posted to the listserv.