Institute on Aging
University of Wisconsin-Madison
2245 Medical Science Center
1300 University Avenue
Madison, Wisconsin 53706-1532
Phone: (608) 262-1818
Fax: (608) 263-6211

MIDUS Geo-coding
README

Linking External Contextual Variables
to MIDUS Datasets
at Different Geographical Levels

This document outlines the process by which MIDUS variables can be merged with contextual data from external sources (such as the Census or Current Population Survey). The MIDUS Admin Core realizes that many geo-coding research questions have scientific merit. The process outlined here attempts to facilitate those endeavors while maintaining established practices regarding data security and the protection of respondent identity.

After reviewing this README, researchers interested in pursuing this process must contact the MIDUS Admin Core (midus_geocoding@aging.wisc.edu) to request a Research Plan and Data Use Agreement.

NOTE: The MIDUS data are NOT geocoded. MIDUS does not provide actual geographic identifiers to researchers, nor are geographic variables of any kind included in publicly available MIDUS datasets. This policy protects participant identities against deductive disclosure. Deductive disclosure is the identification of an individual's identity using known characteristics of that individual. Even when direct identifiers (e.g. name, address, geographic information) are removed from a dataset, it may be possible to identify respondents with unique combinations of characteristics.[1] A number of features of MIDUS make the study susceptible to such risk:

- The sheer number of variables on each MIDUS participant, covering multiple aspects of their lives.
- The comprehensive nature of the variables collected through the different MIDUS projects, which contain personal, cognitive, biological, neurological, and genetic information.
- The potential for extremely small cell counts produced by cross-classifying different domains of data (note, this concern is also monitored by ICPSR when MIDUS submits data for public release).
- Among MIDUS projects with smaller Ns (biomarkers, neuroscience) there can be very few people represented in specific geographic locations (e.g., a given state).
- The sensitive nature of biomarker and clinical data that are protected under HIPPA guidelines.

---

[1] http://www.icpsr.umich.edu/icpsrweb/content/DSDR/disclosure.html

Please note that MIDUS IRB protocols have assured respondents that the information they provide will remain confidential and will in no way be associated with their name. The MIDUS Admin Core cannot emphasize enough how important it is to the respondents, to the study itself, and to the survey research field at large that all precautions be taken to preserve this confidentiality.

The purpose of linking MIDUS data to external contextual variables (e.g., census tract data about neighborhood characteristics) is to allow analyses that examines the characteristics of places of residence *while not revealing actual geographical locations*. The MIDUS geo-coding procedure is performed on a case-by-case basis for individual researchers using a specific process that protects the identity of MIDUS respondents. The process links external contextual data arrayed at different geographic levels (see Example below) but swaps out actual geographic information with noise-added variables that maintain the categorical integrity of the geographic levels while masking their actual location. This allows investigators to distinguish contextual data at different geographic strata without divulging the geographic location or potential identity of the respondents.

In short, ***MIDUS does not provide any geographical identifiers*** for analysis, but authorizes approved users to attach contextual variables at the level of these geographical strata. MIDUS can append contextual data at the following geographic levels:
- FIPS State Code: 2-digit
- FIPS Code: 5-digit, 2-State + 3-County
- Census Tract Code: 11-digit, 2-State + 3-County + 6-Tract
- Census Block Code: 15-digit, 2-State + 3-County + 6-Tract + 4-Block

Example.

Suppose a researcher wishes to examine how the variation in State-level clean air policies affects the reported incidence of asthma among MIDUS respondents who completed the baseline survey. The following is an example of how the process works, in this instance by appending three air-quality variables from the ACS Survey (via the Census Bureau) to the MIDUS data at the State level. ***All data shown are bogus*** and used only to illustrate how contextual data are appended to MIDUS datasets. Table 1 shows the M2ID and State for participants in the MIDUS Core sample.

Table 1. Geographic Data of MIDUS Participants.	
M2ID	STATE
1	CA
2	DE
3	AZ
4	MA
5	TX
6	CT
7	NY
8	PA
9	CA

…	…

Table 2 shows Census data provided to the MIDUS Admin Core by the researcher who is requesting additional contextual information. Note that it is the researcher's responsibility to provide data at the appropriate geographical level, and variables that adhere to Admin Core specifications and formats.

Table 2. Contextual Census Data Provided by Interested Researcher.			
STATE	Census variable 2003	Census variable 2004	Census variable 2005
AK	17.2500	17.9004	6.7908
AL	18.0784	18.1456	6.8509
AR	19.2000	19.1601	6.8800
AZ	19.2281	19.2500	6.9002
CA	19.2761	19.3990	6.9233
CO	19.3866	19.6231	6.9900
CT	19.4321	19.6338	7.0225
DC	19.4837	19.7503	7.0330
DE	19.5443	19.7567	7.0561
…	…	…	…

Table 3 shows the results of the final match of MIDUS data and Census variables – this is the final data product delivered to the researcher. Note that while the Census variables are appended at the individual case level, the actual geographic identifier is coded to limit disclosure of location and potential participant identity.

Table 3. Final Product: Contextual Census Variables Appended at the State Level.				
M2ID	STATE (Noise Added)	Census variable 2003	Census variable 2004	Census variable 2005
1	3	19.2761	19.3990	6.9233
2	15	19.5443	19.7567	7.0561
3	62	19.2281	19.2500	6.9002
4	12	19.6543	19.6542	6.9990
5	90	19.8756	18.9978	7.0651
6	5	19.4321	19.6338	7.0225
7	22	19.5465	19.3344	6.5498
8	88	19.7871	19.5789	6.0321
9	3	19.2761	19.3990	6.9233
…	…	…	…	…