# NYC HANES 2013-14 - WHAT YOU NEED TO KNOW

This document is designed to help data users navigate how best to analyze the NYC HANES 2013-14 dataset, given its population-based and clustered sampling scheme. Below we describe the purpose and content of various resources we have provided.

## Analytic Guidelines

The Analytic Guidelines provide overall guidance on the use of the NYC HANES 2013-14 dataset and statistical weights, as well as other analytic issues pertaining to assessing statistical reliability of estimates.

## White Paper Documenting the Use of Blood Pressure Validation Study to Calibrate NYC HANES 2013-14 Blood Pressure Data

This document explains how NYC HANES 2013-14 blood pressure data were calibrated to be comparable to NYC HANES 2004 blood pressure data. Researchers using blood pressure data should review this document.

## Data Documentation

Data documentation includes the following:

- ***Variable List***

This document lists variables that are included in the public use dataset. The public release data do not include all questions asked in the CAPI or ACASI or all laboratory data about blood or urine testing. Because of confidentiality, quality issues, or other considerations, certain information is not included in the public use dataset. Thus, the dataset does not completely match all the questions asked in different questionnaire sections. Researchers may petition the NYC HANES team for access to additional variables (see Contact Us tab).

- ***Questionnaire***

Questionnaires are provided for CAPI and ACASI. It is important to pay close attention to the skip patterns, targets and eligibility for a particular item.

- ***Codebook***

The codebook contains the text of CAPI and ACASI questions, response codes, and frequency distributions for each variable. Analysts are strongly encouraged to use the codebook in conjunction with the questionnaire. The following table provides a description of the codebook content:

| Codebook Field | Description |
|---|---|
| Section | Describes the component of the study the variable comes from |
| Raw Name | Name of variable as it appears in the dataset |
| Question | Actual question in the questionnaire or brief description of the variable |
| Format | SAS Format Name. User-defined formats are applied for variables, and the format names are provided in the SAS program 'SAS variable formats.sas'. |
| 2004 Study | Name of variable as it appears in the NYC HANES 2004 dataset. Useful for analysis of change over time |
| Notes | Describe the:<br>• NHANES matching variable name, when it applies. Useful for analysis comparing NYC to the National estimates<br>• Preceding skip pattern,<br>• Instructions for the interviewer or examiner during the interview or examination, and<br>• Definitions of recoded variables.<br>The Notes will be useful for identifying the denominator of prevalence estimates and for inclusion criteria. |
| Mean | The mean value of continuous variables. Only available for continuous variables. |
| Value | The possible responses for a variable, and the codes for those responses.<br>**NOTE:** For continuous variables, the value format shows the sample for which the variable has a non-missing value, missing values, Don't Knows and Refusals. |
| Frequency (unweighted) | Frequency of occurrence. |
| Percent (unweighted) | Reflects the occurrence (unweighted) per 100 of the study population. |
| Percent (weighted) | Reflects the weighted occurrence per 100 of the weighted population. |

- ***Formats***

The program (SAS variable formats.sas) includes the SAS formats for the questionnaire and laboratory variables, as well as formats for recoded variables that have been used in previous analyses of the NYC HANES 2013-14 data.

**Note:** Before using the recodes, analysts MUST consider whether the recode is relevant to their analysis by reviewing the literature and clinical guidelines for defining their outcome. The list of 'Previously Used Recodes' is not exhaustive; instead, the recodes are intended to help analysts understand how variables can be used to define outcomes. Users may find that cutoffs or categories should be redefined to suit their specific analysis.