Behavioral Risk Factor Surveillance System

2001

Year-to-Date Data Quality Report Handbook

Version 5.0.0
March 27, 2002

Behavioral Risk Factor Surveillance System
2001 Year-to-Date Data Quality Report Handbook

Table of Contents

Documentation for 2001 Year-to-Date Data Quality Report

Introduction

This document discusses the tables found in the 2001 Year-to-Date Quality Control Report. The main target audience are State* BRFSS Coordinators, especially those who do not have day-to-day responsibility for data collection. The data may be collected by a contractor or in the state health department.

The rest of this document is divided into five sections. *General Guidelines for Interpretation* presents some general advice. The following section, *Definitions of Variables*, defines unfamiliar variables. *General Questions The Tables Above Can Help Answer* lists several general questions and identifies the tables that can help answer them. The next section, *Standards of Comparison*, identifies the types of standards that can be used to evaluate the data in the tables. Next is *Specific Guidelines for Interpretation of 2001 BRFSS Year-to-Date Data Quality Report Tables*, which identifies things to look for and, as much as possible, standards for comparison for each table. Finally, the *Appendix* lists the title of each table in the report.

General Guidelines for Interpretation

Many Results Are More Important for What They Point To Rather Than for What They Are In Themselves. Ultimately, you want to make an informed judgment about how good a job the data collector is doing. Any set of results, even with as many tables as are in this report, can reflect only a portion of all that is involved in conducting a survey. The tables are occasions for you to ask questions about and probe more deeply into the data collection process. Many of the results are not, in themselves, very important. But, taken together, they can be suggestive of the quality of the overall process. They can indicate that a closer look at the data collection process may be called for and what aspects of the data collection process merit a closer look.

Nobody's Perfect. With so many tables to look at, every data collector is bound to have some unfavorable results. Also, some months will be better than others. You should keep in mind that one unfavorable result does not ruin a month and that one bad month does not ruin a year.

Data Provide Evidence, Not Proof. Data are ultimately just marks on some medium. They do not, in themselves, tell you how those marks came to be what they are. How those marks came to be what they are is one's interpretation of the data. There is always a possibility that any particular interpretation, which is always an inference from data, may not be true. Thus, your first step should be to show your data collector the data that seem problematic to you and to ask the data collector to provide an interpretation. If the data collector does not give you an interpretation that completely satisfies you, you should investigate further how those findings came to be.

To take a simple example, suppose a particular interviewer shows an excess of interviews with women and with respondents over age 65. There are at least three ways such results could come about: (1) The interviewer may not be pursuing respondents aggressively and is thus disproportionately interviewing more cooperative respondents. (2) The interviewer may be a woman over 65 who is better able to establish a quick rapport with other women over 65 than with other respondents. (3) The interviewer only works on weekdays, when available respondents are disproportionately women over age 65. (4) The

* "State" refers to the 50 states, the District of Columbia, Guam, Puerto Rico, and the Virgin Islands.

interviewer may be miskeying age. Determining which, if any, of these interpretations is or could be true would require looking at additional data or relying on direct observation or other non-statistical sources of knowledge.

Consistency Is a Lot Easier To Measure Than Quality. In many cases, you will end up looking for consistency over time rather than the achievement of some definite standard.

There's Always a Story. If you do have occasion to ask about how some unfavorable results came about, the response will probably be a story. A few stories are OK, but if you start getting the same or a different set of stories for several months, then you might start looking for a systemic problem.

Count Your Money — Monitoring Data Collection Is Part of Your Job, Not a Sign of Distrust. A sound business practice is that any time cash changes hands, the receiving party should count it. This holds no matter how well the receiving party knows the person. In the same way, you cannot just assume that the data collector is doing a good job, no matter how good a relationship you have. You have to look at data, make direct observations, ask questions, and probe.

If the Data Look Too Good To Be True, They Probably Are. Just as teachers can teach to a test, a knowledgeable data collector can make any set of results look good. All of the variables used in these tables can be slanted or manipulated to a greater or lesser extent. For example, you should be sure that you know exactly how final disposition codes are assigned. Many states assign final disposition codes after data collection ends. It is all too easy for a knowledgeable data collector to write a computer program that "fixes" data so that it can pass PCEdits and other data quality checks without reviewing what really happened or correcting the data collection deficiencies that created the original data error.

Definitions of Variables

Assigned Month. BRFSS protocol calls for states to submit related pre-screened sample records that they never called along with sample records that they did call. Sometimes states fail to do so. For states that receive their sample through BSB, we are able to identify related pre-screened records they may have failed to send and to add them to their data files. Assigned month is the month of the file in which a record was submitted or the month of the file in which a related pre-screened record should have been submitted but was not.

File Month. Each data file submitted to CDC contains the name of a month in the filename. The file month of a record is month of the file in which it was submitted.

Household Roster Status.
None = Number of adults, number of men, number of women are all missing.
Partial = One or two of number of adults, number of men, number of women are missing.
Inconsistent (But Complete) = Number of adults, number of men, number of women are all non-missing but number of adults does not equal the sum of number of men and number of women.
Consistent (And Complete) = Number of adults, number of men, number of women are all non-missing and number of adults equals the sum of number of men and number of women.

Number of Residential Telephone Numbers.
Missing = "Do you have more than one telephone number in your household?" or "How many residential telephone numbers do you have?" was refused.
Otherwise, the reported number of residential telephone numbers.

Density Status. Categorized as Listed, Not listed one-plus block, or Zero block.

Disposition Codes

The final BRFSS disposition codes for 2001 are

01	Completed interview
02	Refused interview after respondent selection
03	Nonworking number
04	Ring no answer
05	Not a private residence
06	No eligible respondent at this number
07	Selected respondent not available during the interviewing period
08	Language barrier
09	Interview terminated within questionnaire
10	Line busy
11	Respondent unable to communicate due to physical or mental impairment
12	Technological barrier
13	On never call list
14	Hang-up or termination before respondent selection

In the tables below, these codes are categorized as follows:

Household records are records with disposition codes of 01, 02, 06, 07, 08, 09, 11, 12, 13, or 14.

Definitions and Labels for Disposition Code Categories

Category	Definition	Format in Tables
Completed interview	Disposition Code = 1	Completed Interview
Eligible Household	(Disposition Code in (7,8,9,14) and Number of Adults not equal to missing) or Disposition Code in (2,9,13)	Elig HH
Household or Probable Household, Eligibility Unknown	(Disposition Code in (7,8,11,14) and Number of Adults equal to missing) or Disposition Code = 12	HH or Prob HH, Elig Unkn
Household, No Eligible Respondent	Disposition Code = 6	HH, No Elig Resp
Non-Contact	Disposition Code in (4,10)	Non-Contact
Non-Household	Disposition Code in (3,5)	Non-HH

General Questions the Tables Above Can Help Answer

The tables are generally organized around questions that a state BRFSS Coordinator, as the ultimate state person responsible for assuring and assessing the quality of BRFSS data should ask.

Is there evidence of significant bias in the data?
The tables in Section II address this question. Specific guidelines for assessing biases are given below.

Is the data collector calling numbers frequently enough and according to the BRFSS callback rules?
Tables III.1 to III.5 address these questions. The discussion for these tables in the *Specific Guidelines for Interpretation* section, below, identifies the patterns to look for.

How consistent is the data collection effort from month to month?
Every table that is run by month can help answer this question. The primary ones to look at are Tables III.6B, III.7B, III.8B, and III.9B. The percentage of records in each disposition code category in Tables III.8B and III.9B should vary by less than two points from month to month. If these show less consistency than you think appropriate, verify that the same patterns exist in Tables III.6B and III.7B**.** These latter tables should show more month to month variability, since their bases are smaller.

Is the data collector dispositioning numbers according to their definitions?
There are a number of different ways, which vary by disposition code, to approach this question. The first place to look is in Section IV, Proper Assignment of Disposition Codes. By looking at the disposition codes of various sets of numbers, you can identify codes that are being assigned to more or fewer records than they should be (Tables IV.1 and IV.2). For example, in Table IV.2, at least 98% of records in zero blocks should be assigned a Non-Contact or Non-Household disposition code. Another way to assess if disposition codes are being appropriately assigned is to determine the consistency between the household roster status and the disposition code (Tables IV.3 to IV.5). A third way is to determine the consistency between the number of attempts and the disposition code (Tables IV.6 and IV.7).

Disposition code 06 No eligible respondent at this number is especially prone to being assigned when another code would be more appropriate. This code should be applied to well under 1% of households. You can check this primarily by looking at Table IV.1. Also, Table IV.4 should show that all records assigned a disposition code of 06 have no household roster.

Are the interviewers adequately trained, supervised, or monitored?
Sections VI and IX can be used to address this question. The standards for comparison in this case are standing relative to other states. Other indications that interviewers are not adequately trained, supervised, or monitored would be: (1) Relatively large percentages of records missing income or weight in Section VI. (2) A relatively large percentage of Hispanics coded Other race and relatively large biases in Section I.

Are there particular interviewers who seem to be deviating from sound practices?
The tables in Section IX can address this question. The easiest way to approach it is to look, first, for outliers (interviewers with 0's or *'s next to their records in the boxplots). Since, however, a flat distribution can mask true outliers, you should also look for records with values that are separated from the others, even if they are not marked at outliers. Please remember that there may be valid reasons for deviations from a statistical norm, so these findings should be treated more as indicators of a need for further investigation rather than conclusive proof of inappropriate interviewer practices.

Standards of Comparison

Various standards of comparison are appropriate for different tables.

In many cases, the standard is defined by **BRFSS protocol**. For example, 100% of records dispositioned No answer should have received at least 15 call attempts.

In other cases, a standard cannot be exactly determined but **logic or data may indicate that only values within a fairly narrow range seem reasonable**. In such cases, I have usually set an exact standard as a guideline, knowing that it is arbitrary. For example, the percentage of household or probable household records that should be coded No eligible respondent uses the phone should be well under .1%.

An **outside standard** is appropriate in bias measures. Bias is measured by the difference between a sample value and a population value. For example, a sample with 60% female respondents drawn from a population that is 52% female shows a selection bias of 8 percentage points. The *2001 Year-to-Date*

Data Quality Report contains population data that serve as comparison standards for selection bias measures.

There are cases where the best that can be done is to determine if a state is an **outlier** in a distribution. For example, there is no particular standard that can be determined for the percent of completes among household records by examination of the BRFSS calling rules and protocols. The best that can be done is to see if a state is an outlier compared to other states.

Finally, regardless of how a standard is determined for an individual measure, **consistency** in that measure **is important**. A measure may or may not be consistent at an acceptable level but inconsistency itself can be an indicator of inconsistent or poor quality data **collection** practices.

Standards are indicated by **bold** print.

Specific Guidelines for Interpretation of 2001 BRFSS Year-to-Date Data Quality Report Tables

I. Sample Generation, Release, and Submission

Table I.1. Density Status
This table presents the percentages of listed, not-listed one-plus block, and zero block numbers in the sample. Other things equal, increases in percent listed and decreases in percent zero block numbers should increase efficiency. This table will help in assessing the extent to which changes in sample composition may be responsible for changes in efficiency.

Table I.2. Number of Records in Replicate
Standard: All replicates should contain fifty records.
Telephone sample records should be released by replicate and all records, including those pre-screened as non-working or business, should be sent to BSB. Large numbers of replicates with fewer than fifty records indicate that one of the protocols is not being followed. Small numbers of replicates with fewer than fifty records probably indicate processing problems, either in the CATI software or in post-data collection processing.

Table I.3. Interview Month By File Month
BRFSS protocol states that data collectors should attempt to complete a monthly survey within the prescribed month but that it is more important to call numbers fully according to the BRFSS callback rules than it is to finish within a prescribed month. **The BRFSS standard is that 100% of records in a given file month should be in the identical interview month.** Nevertheless, because of the priority of calling telephone numbers fully according to the BRFSS callback rules, an occasional few percentage points below 100% is not a matter for concern. A substantial deviation from the 100% standard should, however, be an occasion for an inquiry. A chronic deviation from the 100% standard is an indicator that the data collector needs to devote more resources (for example, more hours of calling per month or more interviewers) to the BRFSS in order to complete the survey on time.

Table I.4. File Month By Assigned Month
This table shows whether the data collector is submitting sample records prescreened by GENESYS as non-working or business. BRFSS Policy Memo 98-3 specified that such records should be submitted to BSB.
Standard: 100% of the records in an assigned month should be in the identical file month.

II. Bias

Table II.1, II.2, II.3. Gender, Age, and Race/Ethnicity Biases

These tables show the unweighted percentages among completes of gender, age, and race/ethnicity compared to the 2001 population estimates from Claritas. Large selection biases are a strong indicator of possible biases in the data. They do not, however, indicate anything about the source of the possible biases. The source could be any source of non-sampling error, which causes data to not be representative of the sample. Some sources of non-sampling errors under the control of the data collector are not working the sample hard enough or according to BRFSS protocol, interviewer misconduct (for example, fabrication of interviews, recording one adult in a household in order to be able to interview the person on the phone), a staff of interviewers without the training or skill to induce hard to interview respondents to complete an interview.

Standards: Acceptable ranges for gender, age, and race/ethnicity biases are given below. A value above or below the acceptable range should trigger a search for possible causes and remedies. These ranges are based on observed biases in 2000.

Variable	Category	Acceptable Bias Range
Gender	Female	3% to 9%
Age	18-24	-4% to 1%
Age	25-34	-5% to 3%
Age	35-44	-1% to 3%
Age	45-54	-1% to 3%
Age	55-64	-1% to 3%
Age	65+	-3% to 1%
Race/Ethnicity	White/Non-Hispanic	-2% to 3%

Table II.4. Race by Hispanic Origin

Hispanics tend to approach race from a different perspective than non-Hispanics. For this reason, it is difficult to get Hispanics to name one of the standard race categories as their race. As a result, many Hispanics receive a race of Other. In the 1990 Census, about 20% of Hispanics indicated their race as Other.

Standard: More than about 35% of Hispanics with a race of Other may indicate that interviewers have not received appropriate training on probing for the race of Hispanic respondents. Fewer than about 15% of Hispanics with a race of Other may indicate that interviewers are imputing the race of Hispanics.

III. Sample Management: Magnitude and Consistency of Effort

Tables III.1 to III.3. Date and Day of Week of Final Disposition and Minimum, Mean, and Maximum Number of Attempts for Completes, One-Plus Block Numbers, and Zero Block Numbers Respectively.

A good rule-of-thumb is that about 85% of the completes and 75% of one-plus block numbers should received a final disposition in the first half of the interviewing days. For zero block numbers the pattern should be even more pronounced: **About 80% of zero block numbers should receive a final disposition in the first two days of interviewing.** (States that follow the recommendation of calling all the zero block numbers once during the first weekday afternoon of interviewing will disposition almost all of them in a single day.) A bulge in the number of (especially zero-block number) dispositions after the beginning to the interviewing period could be an indicator that additional telephone numbers were released. This should occur early enough in the interviewing period that there is time to fully call all the released numbers according to the BRFSS callback rules. A relatively large number of dispositions per day well into the interviewing period or, worse, an increase in the number of dispositions per day at the end of the interviewing period is probably an indication that the data collector needs to devote more

recources (for example, more hours of calling per month or more interviewers) to the BRFSS in order to fully call all the released numbers according to the BRFSS callback rules. The tables also indicate the day of the week on which numbers receive a final disposition. The day of the week can be used to check on weekend calling. During the last half, or at least the last several days, of the interviewing period, the number of final dispositions by day should be in the single digits.

In Tables III.1 and III.2, after about the third interviewing day, the minimum number of attempts should go above 1 and keep going up until it levels off at close to 15 near the end of the interviewing period. (There may be an occasional number with only a few call attempts, but the large majority of records dispositioned close to the end of the interviewing period should have close to 15 call attempts.) **Similarly, the mean number of attempts should be in the 6-8 range by the fifth interviewing day.**

In Table III.3, the minimum number of attempts should go above 1 after the second interviewing day and the mean number of attempts should be in the 6-8 range by the fourth interviewing day.

The Appendix contains Tables II.1 to III.3 for a state which meets the above standards.

Table III.4. Date, Day of Week, and Final Disposition Code

The frequencies of complete and incomplete households should be heavily concentrated in the first half of the interviewing days. The frequencies of non-working and non-private residences should be heavily concentrated in the first two or three days of the interviewing period. The frequencies of non-contacts should be concentrated toward the end but not at the very end of the interviewing period. You should also look at the extent and pattern of weekend interviewing. The hours for weekend interviewing are more limited than for weekdays and weeknights. At the same time, more people tend to be at home during the weekend. For these reasons, weekend calling should be approached strategically. **In particular, weekend calling should be heavy enough to call all available numbers during a weekend but it should not be wasted on numbers that have not been called before.**

Table III.5. Number of Days in Field for Replicate

Replicates are subsets (in the BRFSS case, of fifty records) of an entire sample that are representative of the entire sample. This table identifies the earliest date on which a sample record from a particular replicate receives a final disposition. As such, it can indicate whether or not one of two events occurred. First, it is common for a CATI package to give priority to sample records that have already been called when sending sample records to interviewers. If that is the case and the data collector does not schedule enough interviewer hours early in the interviewing period, sample records that were not called on the first day or two may not come up until late in the interviewing period. Second, telephone numbers could inappropriately be released in the middle or toward the end of an interviewing period. This table can indicate that one or the other has occurred. It shows only replicates that have been in the field for 14 days or fewer.

Standard: Replicates should be in the field long enough to ensure that all of the numbers could have been called according to the BRFSS callback rules: an absolute minimum of three days, one of which must be a weekend day.

Tables III.6A to III.9B Disposition Codes

The A versions of these tables are Year-to-Date to give you baseline percentages for the year. The B versions are By Month to give you the monthly variability in the distributions. You can these use disposition codes to measure consistency and, to a lesser extent, the quality of the data collection effort. Because the base of the percentages in Tables III.8A to III.9B is all records, a difference of even 2 percentage points from one month to the next could be meaningful.

IV. Proper Assignment of Disposition Codes

Table IV.1. Disposition Codes for Household or Probable Household Records by Density Status
The treatment of the new disposition codes of Technological Barrier and Hang-Up or Termination Before Respondent Selection as households in 2001 is significantly lowering the percentage of completes among households compared to 2000. As a result, **the following standards should be treated as very rough estimates**: The percent complete should be roughly 30% to 60% among listed households, 25% to 55% among not listed one-plus block numbers, and 20% to 55% among zero block numbers. In addition, the difference in percent complete between listed and zero block numbers should be no greater than about 20 percentage points. Household completion rates below these standards and large differences between listed and zero block rates could indicate that the data collector is not working the sample hard enough or is inappropriately dispositioning non-household numbers as households. Household completion rates above these standards could indicate that the data collector is inappropriately dispositioning household numbers as non-households. Household completion rates are also influenced by characteristics of the population.

Table IV.2. Disposition Codes for All Records by Density Status. **The percent household or probable household should be 65% to 80% among listed numbers, 15% to 30% among not listed one-plus block numbers, and 0.25% to 1.0% among zero block numbers.** A smaller percentage of households, especially for listed numbers, could indicate that sample records are not being called enough or that records which actually ring to households are receiving a non-household disposition. A larger percentage could indicate that records which do not ring to households are receiving a household disposition. A review of one state's 2001 data showed a substantially larger percentage of households for zero block numbers in 2001 compared to 2000. The problem seemed to be due to the inappropriate dispositioning of non-household records as Technological Barrier or Hang-Up or Termination Before Respondent Selection. Keep in mind that household identification rates are influenced by characteristics of the telephone system and of the population. Thus, data outside of these ranges may not be due to any inappropriate practices by the data collector. That, however, can only be determined by a close review of data collection practices.

Tables IV.3 to IV.7. Houshold Rosters or Number of Attempts
The standards for these tables are indicated in the titles.

V. Unit Nonresponse

Table V.1. CASRO Rate
A CASRO Rate below 40% should cause a review of data collector practices that could impact it, especially sample management and interviewer recruitment, retention, training, supervision, and monitoring. Fluctuations of more than about 4 percentage points or a downward trend for three or more months should signal an inquiry into the data collector's practices.

Table V.2. Percent of Eligible Households Among Possibly Eligible Households
This table tells you the percentage of possibly eligible households who answer the Number of Adults question. It indicates the amount of nonresponse before a household roster is begun.

Table V.3. Percent of Eligible Households With Selected Respondent
This table tells you the percentage of households for which a respondent is selected once a household roster is begun. **This percentage should be at or very close to 100%.**

Table V.4. Percent of Completes Among Households With Selected Respondent
This table tells you what happened after a respondent has been selected.

VI. Item Nonresponse

Tables VI.1A to VI.1B. Income Missing Values for Men
The following graph shows the distribution of percent missing values on income by state for men. ("Appendix B: Explanations of Stem-and-Leaf Displays and Boxplots," in this document, explains how to read this graph.)
States should try to stay below 15% missing values on income for men. States with a percentage of missing values above 15% should review their training and monitoring of interviewers with respect to income missing values.

Tables VI.2A to VI.2B. Income Missing Values for Women
States should try to stay below 20% missing values on income for women. States with a percentage of missing values above 20% should review their training and monitoring of interviewers with respect to income missing values.

VII. Household Rosters and Number of Phones

Tables VII.1A and VII.1B. Household Roster Status, Records With Partial or Complete Household Rosters
States should have well under 1% of records with a partial household status and no records with an inconsistent household status.

Tables VII.2A and VII.2B. Number of Adults, Records With Non-Missing Number of Adults
Generally, about one-third of records should indicate one adult and about one-half two adults. Records with six or more adults should be questioned.

Tables VII.3A and VII.3B. Number of Phones
Generally, about 85-90% of records should indicate one phone line. Records with three or more phones should be questioned.

Table VII.4. Number of Phones, By Number of Adults
The percentage of records with one phone line should decrease as the number of adults increases. Improbable combinations, for example, one adult and five phones, should be questioned.

VIII. Recruitment, Retention, and Assignment of Interviewers

Tables VIII.1 and VIII.2. Recruitment and Retention of Interviewers
These tables will give you an idea of the stability of the interviewer workforce. **If the turnover seems excessive, you should inquire about the data collectors recruitment base and practices and about their working conditions.**

Table VIII.3. Minimum, Median, Mean, and Maximum Numbers of Completed Interviews per Interviewer
This table will give you an indication of how much BRFSS interviewers are getting with the current year BRFSS survey. **Continuing small median and mean numbers of completed interviews per interviewer as the interviewer year progresses indicates that interviewers are not assigned enough to the survey to gain experience with it.** This could indicate either high turnover or an excessively large number of interviewers assigned to the BRFSS.

IX. Interviewer Outliers

Tables IX.1 to IX.21. These tables should be approached from a systemic and from an individual perspective. From a systemic perspective, the question is whether or not management provides appropriate training, supervision, and monitoring of interviewers. For several variables, one indication that this could be the case would be a poor showing on a global measure from an earlier section coupled with a flat distribution and a large standard deviation among interviewers. For example, **a percent missing income of over 15% in Table VI.1A coupled with a standard deviation of over 5 percentage points in IX.9 should trigger questions about the training, supervision, and monitoring of interviewers in general with respect to asking and probing for income.**

From an individual perspective, the question is whether or not a particular interviewer is following protocol. From this perspective, **the data should be examined with a view toward finding interviewers who are outliers**. Interviewers who are outliers on several measures should be monitored especially carefully, even more especially if they consistently beat the norm. "Appendix B: Explanations of Stem-and-Leaf Displays and Boxplots," in this document, explains how to read these graphs.

Appendix A: List of Tables in the 2001 BRFSS Year-to-Date Data Quality Control Report

List of tables

Note: Each table title has three sections: the name of the dependent variable (with categories in parentheses, if appropriate), either the phrase Year-to-Date or the name of a By-variable, and a description of the records in the table (the base).

I. Sample Generation, Release, and Submission

Table I.1. Density Status, By Assigned Month, Base = All Records

Table I.2. Number of Records in Replicate, By Assigned Month, Base = All Records

Table I.3. Interview Month By File Month, Base = Completes Only

Table I.4. File Month By Assigned Month, Base = All Records

II. Bias

Table II.1. Discrepancy in Gender Between 2001 Claritas Population Estimates and Unweighted BRFSS Data, Year-to-Date, Base = Completes Only

Table II.2. Discrepancy in Age Between 2001 Claritas Population Estimates and Unweighted BRFSS Data, Year-to-Date, Base = Completes Only

Table II.3. Discrepancy in Race/Ethnicity Between 2001 Claritas Population Estimates and Unweighted BRFSS Data, Year-to-Date, Base = Completes Only

Table II.4. Race by Hispanic Origin, Year-to-Date, Base = Completes Only

III. Magnitude and Consistency of Effort

Table III.1 Date and Day of Week of Final Disposition and Minimum, Mean, and Maximum Number of Attempts, By File Month, Base = Completes Only,

Table III.2. Date and Day of Week of Final Disposition and Minimum, Mean, and Maximum Number of Attempts, By File Month, Base = One-Plus Block Numbers With One or More Attempts

Table III.3. Date and Day of Week of Final Disposition and Minimum, Mean, and Maximum Number of Attempts, By File Month, Base = Zero Block Numbers With One or More Attempts

Table III.4. Date, Day of Week, and Final Disposition Code, By File Month, Base = Records With One or More Attempts

Table III.5. Number of Days in Field for Replicate, By File Month, Base = Replicates in Play for Fourteen or Fewer Days

Table III.6A. Disposition Code (Categorized as Completed Interview; Eligible Household; Household or Probable Household, Eligibility Unknown; Household, No Eligible Respondent; Non-Contact; Non-Household), Year-to-Date, Base = Listed Records

Table III.6B. Disposition Code (Categorized as Completed Interview; Eligible Household; Household or Probable Household, Eligibility Unknown; Household, No Eligible Respondent; Non-Contact; Non-Household), By File Month, Base = Listed Records

Table III.7A. Disposition Code (Categorized as Completed Interview; Eligible Household; Household or Probable Household, Eligibility Unknown; Household, No Eligible Respondent), Year-to-Date, Base = Household or Probable Household Records

Table III.7B. Disposition Code (Categorized as Completed Interview; Eligible Household; Household or Probable Household, Eligibility Unknown; Household, No Eligible Respondent), By File Month, Base = Household or Probable Household Records

Table III.8A. Disposition Code (Categorized as Completed Interview; Eligible Household; Household or Probable Household, Eligibility Unknown; Household, No Eligible Respondent; Non-Contact; Non-Household), Year-to-Date, Base = All Records

Table III.8B. Disposition Code (Categorized as Completed Interview; Eligible Household; Household or Probable Household, Eligibility Unknown; Household, No Eligible Respondent; Non-Contact; Non-Household), By File Month, Base = All Records

Table III.9A. Disposition Code, Year-to-Date, Base = All Records

Table III.9B. Disposition Code, By File Month, Base = All Records

IV. Proper Assignment of Disposition Codes

Table IV.1. Disposition Code (Categorized as Completed Interview; Eligible Household; Household or Probable Household, Eligibility Unknown; Household, No Eligible Respondent; Non-Contact; Non-Household), By Density Status, Base = Household Records

Table IV.2. Disposition Code (Categorized as Completed Interview; Eligible Household; Household or Probable Household, Eligibility Unknown; Household, No Eligible Respondent; Non-Contact; Non-Household), By Density Status, Base = All Records

Table IV.3. Household Roster Status, By Assigned Month, Base = Records With Final Disposition Codes of 01, 02, or 09 (Should Have Only Consistent Household Rosters)

Table IV.4. Household Roster Status, By Assigned Month, Base = Records With Final Disposition Codes of 03, 04, 05, 10, 12, or 13 (Should Have No Household Rosters)

Table IV.5. Household Roster Status, By Assigned Month, Base = Records With Final Disposition Code of 14 (Should Have No or Partial Household Rosters)

Table IV.6. Number of Attempts, By Assigned Month, Base = Records With Final Disposition Codes of 04, 10, or 12 (Should Have 15+ Attempts)

Table IV.7. Number of Attempts, By Assigned Month, Base = Records With Final Disposition Codes of 13 (Should Have 0 Attempts)

V. Unit Nonresponse

Table V.1. CASRO Rate, All Records, Year-to-Date and by Assigned Month, Base= All Records

Table V.2. Percent With Determined Household Eligibility, By File Month, Base = Eligible Households and Households or Probable Households With Undetermined Eligibility (Records with Final Disposition Codes of 01, 02, 07, 08, 09, 11, 12, 13, or 14.)

Table V.3. Percent With Selected Respondent, By File Month, Base = Eligible Households (Records With Final Disposition Codes of 01, 02, 07, 08, 09, 11 and NUMADULT Not Equal to Missing.)

Table V.4. Percent Complete, By File Month, Base = Households With Selected Respondent (Records With Final Disposition Codes of 01, 02, 07, 08, 09, 11 and NUMWOMEN Not Equal To Missing)

VI. Item Nonresponse

Table VI.1A. Income (77 and 99 Collapsed), Year-to-Date, Base = Men Only

Table VI.1B. Income (77 and 99 Collapsed), By File Month, Base = Men Only,

Table VI.2A. Income (77 and 99 Collapsed), Year-to-Date, Base = Women Only

Table VI.2B. Income (77 and 99 Collapsed), By File Month, Base = Women Only

VII. Household Rosters and Number of Phones

Table VII.1A. Household Roster Status, Year-to-Date, Base = Records With Partial or Complete Household Rosters,

Table VII.1B. Household Roster Status, By Assigned Month, Base = Records With Partial or Complete Household Rosters,

Table VII.2A. Number of Adults, Year-to-Date, Base = Records With Non-Missing Number of Adults,

Table VII.2B. Number of Adults, By Assigned Month, Base = Records With Non-Missing Number of Adults,

Table VII.3A. Number of Phones, Year-to-Date, Base = Completes Only

Table VII.3B. Number of Phones, By Assigned Month, Base = Completes Only

TableVII.4. Number of Phones, By Number of Adults, Base = Completes Only

VIII. Recruitment, Retention, and Assignment of Interviewers

Table VIII.1. Number of Interviewers by Number of Months Interviewer Working, Year-to-Date, Base = All Records

Table VIII.2. Recruitment and Retention of Interviewers, By Assigned Month, Base = All Records

Table VIII.3. Minimum, Median, Mean, and Maximum Numbers of Completed Interviews per Interviewer, By Assigned Month, Base = Completes Only

Table VIII.4. Number of Completed Interviews by Interviewer ID, Year-to-Date, Base = Completes Only (Not currently available)

IX. Interviewer Outliers

A. Bias

Table IX.1. Distribution of Percent Female, By Interviewer ID

Table IX.2. Distribution of Percent Age 18-24, By Interviewer ID, Base = Completes Only

Table IX.3. Distribution of Percent Age 65 Plus, By Interviewer ID, Base = Completes Only

Table IX.4. Distribution of Percent White, By Interviewer ID, Base = Completes Only

Table IX.5. Distribution of Percent Hispanic, By Interviewer ID, Base = Completes Only

Table IX.6. Distribution of Percent White, Non-Hispanic, By Interviewer ID, Base = Completes Only

Table IX.7. Distribution of Percent Other Race, By Interviewer ID, Base = Hispanics Only

Table IX.8. Distribution of Percent White, By Interviewer ID, Base = Hispanics With Reported Race Only

B. Item Nonresponse

Table IX.9. Distribution of Percent Income Missing (77 and 99 Collapsed), By Interviewer ID, Base = Men Only

Table IX.10. Distribution of Percent Income Less Than $10,000, By Interviewer ID, Base = Men Only

Table IX.11. Distribution of Percent Income Greater Than $75,000, By Interviewer ID, Base = Men Only

Table IX.12. Distribution of Percent Income Missing (77 and 99 Collapsed), By Interviewer ID, Base = Women Only

Table IX.13. Distribution of Percent Income Less Than $10,000, By Interviewer ID, Base = Women Only

Table IX.14. Distribution of Percent Income Greater Than $75,000, By Interviewer ID, Base = Women Only

Table IX.15. Distribution of Percent Income Don't Know/Not Sure (77), By Interviewer ID, Base = Men Only

Table IX.16. Distribution of Percent Income Refused (99), By Interviewer ID, Base = Men Only

Table IX.17. Distribution of Percent Income Don't Know/Not Sure (77), By Interviewer ID, Base = Women Only

Table IX.18. Distribution of Percent Income Refused (99), By Interviewer ID, Base = Women Only

Table IX.19. Distribution of Percent Weight Missing (777 and 999 Collapsed), By Interviewer ID Women Only,

C. Number of Adults

Table IX.20. Distribution of Percent One Adult, By Interviewer ID, Base = Records With Non-Missing Number of Adults

Table IX.21. Distribution of Percent Three or More Adults, By Interviewer ID, Base = Records With Non-Missing Number of Adults

Appendix B: Explanations of Stem-and-Leaf Displays and Boxplots

Stem-and-leaf displays and boxplots allow the examination of a distribution for the purpose of identifying extreme values, or outliers. Stem-and-leaf displays and boxplots are found in the output of Section IX. They also appear in the discussions of Tables VI.1A to VI.2B on pages 11 and 12 of this document.

In a stem-and-leaf display, the first column of numbers is the stem. In these tables, the stem usually represents a whole percentage number. The second sets of numbers are the leaves. Each observation is represented by one digit (leaf). In these tables, the leaves usually represent rounded tenths of a percent. As a whole, a stem-and-leaf can be thought of as a vertical histogram in that the lengths of the leaves are proportional to the relative frequencies in an interval. Immediately to the right of the stem-and-leaf display proper is a column showing the number of observations in each interval. Observations toward the ends of a distribution that are separated from other observations by one or more blank intervals are candidates to be considered as outliers.

A box plot provides a more formal statistical approach to identifying outliers. "The bottom and top edges of the box correspond to the sample 25th (Q1) and 75th (Q3) percentiles. The box length is one interquartile range (Q3-Q1). The center horizontal line with asterisk endpoints corresponds to the sample median. The central plus sign (+) corresponds to the sample mean. If the mean and median are equal, the plus sign falls on the line inside the box. The vertical lines that project out from the box, are called whiskers; they extend as far as the data extend, up to a distance of 1.5 interquartile ranges. Values farther away are potential outliers. The procedure identifies the extreme values with a zero or an asterisk (*). If zero appears, the value is between 1.5 and 3 interquartile ranges from the top or bottom edge of the box. If an asterisk appears, the value is more extreme." (SAS Institute Inc., SAS Procedures Guide, Version 8, Cary, NC: SAS Institute Inc., 1999. 1643 pp. Page 1389.)