

Database Documentation

The 2008 National Survey of Drinking and Driving Attitudes and Behaviors was administered to a general population sample of 6,999 persons 16 and older. Of these, 1607 persons came from cell phone only households and were interviewed using their cell phones. The remaining 5392 respondents were interviewed using landline telephones. The landline sample included persons who also had cell phones (although their interviews were always on landline phones). The cell phone sample was screened to exclude those who also had a landline, thus the cell phone sample had no overlap with landline households.

The survey was administered using a single questionnaire containing approximately 100 questions and assorted skip patterns. The questions included both fixed response and open-ended response items. Users of the databases may wish to first review the questionnaire and its skip patterns to get a sense of how cases are directed through the question series. The questionnaire is included as a document on this Web site.

There are certain characteristics of the database about which database users should be aware when conducting analyses:

1. Questions allowing multiple responses are represented in the database by multiple fields (database variables). The variable name will contain an underscore followed by a number indicating which number field it is (e.g., QN88_1; QN88_2; QN88_3). To get an estimate of the extent to which a particular response was given, the database user will need to collect the information from all fields attached to that questionnaire item. For example, QN88 asks what the respondent did to keep guests at a social event from driving after drinking too much to drive safely. It was asked only of those persons who said in response to QN86 that they had hosted in the past year a social event or party where alcohol was served to adults, a total of 2703 cases. Of those 2703 cases, 2703 were coded as providing 1 response and 728 were coded as having 2 responses. The first field has the variable name QN88_1. It contains responses from all 2703 respondents. The second field has the variable name QN88_2. It has a second response from those 728 respondents recorded as having 2 responses. The third field has the variable name QN88_3. It has a third response from those 234 respondents who gave 3 responses. Thus to obtain the percentage who responded “have someone else drive them home” in QN88, one would have to combine those identified as giving that response in field 1 with those identified as giving it in field 2 and those identified as giving it in field 3 and applying it to the appropriate base (i.e., the total of 2703 cases).

In general, the order in which the responses are provided for any individual case is a function of the order in which the interviewer recorded the responses. That order does not necessarily reflect priority; oftentimes it is the order in which response categories were listed for the interviewers to check-off. For any single case, the number of fields for which there is data equates to the number of responses recorded or coded for that case.

The following questions allowed multiple responses: QN88 (3 fields), QN91 (3 fields), D2 (3 fields), and D6 (5 fields).

2. Some items requiring numeric responses set a ceiling value, whereby any responses above the ceiling value were coded as the ceiling value. Thus the value “13” in QN36 meant 13 months or more than 13 months. A value label was attached here stating “more

than 12 months ago.” In the vast majority of cases, all recorded responses fell below the ceiling value. Exceptions were QN18 (96+), QN36 (13+), QN39 (24+), QN41 (120), and QN66 (97+). In addition, QN20 and QN23 restricted the response to 30 days because those questions asked how many days during a typical month the respondent had consumed specified amounts of alcohol. There also were variables where a value of “0” did not necessarily mean a null value but could be a value less than 1 (e.g., a partial drink): QN31, QN38, QN41, QN65, QN66, and QN126.

3. Items requiring a numeric response from the respondent typically used extreme values for categorizing a “Don’t Know” (DK) response or Refusal (RF). For example, QN139_B, which asked what the minimum drinking age is, used the value 98 to record DK and 99 to record RF. If computing statistics for database variables, database users will need to take into account that some of the recorded numbers may actually be representations for categories rather than numeric values.
4. Review of the databases for statistical disclosure issues led to values for some variables being recoded into floor and/or ceiling ranges. Those recoded were:
 - S1: Set a ceiling value of 6 or more.
 - QND1 Set a ceiling value of age 86 or older
 - QND1B Set a ceiling value of 4 or more children
 - QND9 Set a floor value of less than 100 lbs and ceiling value of more than 325 lbs
 - QND11: Set a ceiling value of 4 or more telephone lines
 -
5. The given variable label may truncate or otherwise shorten the question read to the respondents by the interviewers. Database users would need to check the questionnaire to see all wording read by the interviewers.

Other notes:

S1 variable: The 1607 missing cases who were not asked the number of adults in the household were the cell phone sample. This study treated the cell phone as a single user device attached to the respondent, in contrast to landline phones available to any household member (and affecting the within household probability of selection).

Qn39 series: There are 3 variable fields for Qn 39, which asks the length of time over which respondents consumed alcohol during their most recent instance of drinking and driving. The first field (QN39H) records the number of hours specified by respondents and the second field (QN39M) records the number of minutes specified by respondents. The third field (QN39) combines the two to provide the total response for each respondent. Thus, those respondents who said an hour and a quarter had a “1” entered in the first field, a “15” entered in the second field, and a “1.25” entered in the third field. Across cases, the interviewers recorded into the first 2 fields how the respondents answered the question. For example, some respondents said 90 minutes, while others said an hour and a half. In the first case, the first field would be “time not given in hours” and the second field would be “90.” In the second case, the first field would be “1” and the second field “30.” In both cases, the third field was “1.5.”

Race recode variable: The race question (qnd6) allowed multiple responses. This variable recodes the original responses (qnd6_1 through qnd6_5) and presents a single race category (White only/Black only/Other only/Multi-race) for each respondent.

Normalized weight: This variable contains the survey weights to be used for generating weighted estimates. The weights are normalized in the sense that they add up to the actual sample size of 6999 respondents.

Proj_wt: This is also a variable containing the survey weights. It is essentially the same weight variable (normalized weight) except that it adds up to the estimated population size of about 232 million. So, it is rescaled to a different total. Estimates like population means and proportions will be the same irrespective of what sets of weights (normalized or projected) are used. However, for estimating population aggregates like estimated number of the 16 and older population who drink and drive, the use of projected_wt will be necessary.