

Behavioral Risk Factor Surveillance System

Weighting BRFSS Data
BRFSS 2013





Weighting BRFSS Data

Introduction: Weighting Rationale

When data are used without weights, each record counts the same as any other record—and this implies the following assumptions:

- 1) that each record has equal probabilities of being selected, and
- 2) noncoverage and nonresponse are equal among all segments of the population.

During data collection, the conditions that actually happen may be different than the above assumptions, and the differences can become large enough to affect the results of the data set. Weighting each record appropriately can help to adjust for these differences (or “assumption violations”). Another reason for weighting—one that is conceptually unrelated to the first reason—is to make the total number of cases equal to some desired number which, for state BRFSS data, is the number of people in the state who are aged 18 years and older. In the BRFSS, such raking serves as a blanket adjustment for noncoverage and nonresponse and forces the total number of cases to equal population estimates for each geographic region, which for the BRFSS sums to the state population.

This document provides a general description of the process that reflects factors taken into account in weighting the 2013 BRFSS data. Where a factor does not apply, its value is set to one for calculation. Regardless of state sample design, use of the final weight in analysis is necessary if users are to make generalizations from the sample to the population.

Weighting Methods and BRFSS

From 1984 to 2010, BRFSS used a technique called post stratification to weight BRFSS survey data to account for known proportions of age, race and ethnicity, gender, geographic region, and other known

characteristics of a population. This type of weighting was important because it helped the sample represent the population accurately and adjusted for nonresponse bias. In 2011 a new statistical method

called raking replaced the older method because it could help record additional population characteristics such as education level, marital status, and home ownership status of respondents. It also helps the data fit into an analysis-friendly format called dual-frame survey design, so organizers could account for the type of telephones the respondents used while completing the survey. Because of the changes in the methodology, researchers are advised to avoid comparing data collected before the changes (up to 2010) with data collected from 2011 and onward. Details of these methodology changes can be found with the BRFSS 2011 annual data page:

http://wwwdev.cdc.gov/brfss/annual_data/annual_2011.htm

More About Raking

The raking weighting methodology is comprised of two sections: **design weight** and **raking**.

Design Weight = $STRWT * (1/NUMPHON2) * NUMADULT$

In 2013, the inclusion of cellular telephone respondents who received between 90 and 99 percent of their telephone calls on their cellular telephone required an adjustment to the design weights to account for the overlapping sample frames. From each of the two sample frames, BRFSS calculated a compositing factor for the mostly cellular telephone dual sampling frame users. BRFSS did this calculation by multiplying the design weight by the compositing factor to generate a composite weight for the records in the overlapping sample frames. BRFSS then truncated the design weight based on quartiles within geographic region, which BRFSS uses as the raking input weight. To increase the value of extremely low weights and decrease the value of extremely high weights (and to reduce errors in the outcome estimates that could result), BRFSS uses weight trimming. The objective of weight trimming is to reduce errors in the outcome estimates caused by unusually high or low weights in some categories.

The stratum weight accounts for differences in the basic probability of selection among strata (subsets of area code/prefix combinations). It is the inverse of the sampling fraction of each stratum. There is rarely a complete correspondence between strata, which BRFSS defines by subsets of area code/prefix combinations, and regions, which BRFSS defines by the boundaries of government entities.

The stratum weight (**STRWT**) is calculated using the following items:

- **Number of available records** (**NRECSTR**) and the number of records selected (**NRECSEL**) within each geographic strata and density strata.
- **Geographic strata** (**GEOSTR**), which may be the entire state or a geographic subset such as counties, census tracts, etc.
- **Density strata** (**DENSTR**) indicating the density of the phone numbers for a given block of numbers as listed or not listed.

Within each **_GEOSTR*_DENSTR** combination, the stratum weight (**_STRWT**) is calculated from the average of the **NRECSTR** and the sum of all sample records used to produce the **NRECSEL**. The stratum weight is equal to **NRECSTR / NRECSEL**.

1/ NUMPHON2	The inverse of the number of residential telephone numbers in the respondent's household.
NUMADULT	The number of adults 18 years and older in the respondent's household.
FINAL WEIGHT	BRFSS rakes the design weight to 8 margins (age group by gender, race/ethnicity, education, marital status, tenure, gender by race/ethnicity, age group by race/ethnicity, phone ownership). If BRFSS includes geographic regions, four additional margins (region, region by age group, region by gender, region by race/ethnicity) are included. If at least one county has 500 or more respondents, BRFSS includes four additional margins (county, county by age group, county by gender, county by race/ethnicity).
_LLCPWT	The final weight assigned to each respondent.

BRFSS uses weight trimming to increase the value of extremely low weights and decrease the value of extremely high weights. The objective of weight trimming is to reduce errors in the outcome estimates caused by unusually high or low weights in some categories.

Calculation of a Child Weight

BRFSS calculates the design weight for the child weighting from the stratum weight times the inverse of the number of telephones in the household and then multiplies it by the number of children:

$$\text{Child Design Weight} = \text{STRWT} * (1/\text{NUMPHON2}) * \text{CHILDREN}$$

CHILDWT = BRFSS rakes the child design weight is raked to 5 margins including age by gender, race/ethnicity, gender by race/ethnicity, age by race/ethnicity, and phone ownership.

_CLLCPWT is the weight assigned for each child interview.

Calculation of a Household Weight

BRFSS calculates the design weight for the household weighting from the stratum weight times the inverse of the number of telephones in the household.

$$\text{Household Design Weight} = \text{STRWT} * (1/\text{NUMPHON2})$$

BRFSS rakes **HOUSEWT** = $\text{STRWT} * (1/\text{NUMPHON2})$ to the Number of Persons in a Household, Number of Adults in a Household, and Number of Children in a Household

_HHOLDWT is the weight assigned to each household.

BRFSS Weighting prior to 2011

Please note prior to 2011, the data weighting formula used post-stratification as the weighting method. The formulas listed below apply to data BRFSS weighted prior to 2011.

$$\text{FINALWT} = \text{STRWT} * 1 \text{ OVER NPH} * \text{NAD} * \text{POSTSTRAT}$$

BRFSS intended the computational formula above to reflect all the possible factors that it could take into account in weighting a state's data. Where a factor does not apply, BRFSS set its value to one.

FINALWT is the final weight BRFSS assigns to each respondent.

The stratum weight accounts for differences in the basic probability of selection among strata (subsets of area code/prefix combinations). It is the inverse of the sampling fraction of each stratum. There is rarely a complete correspondence between strata, which BRFSS defines by subsets of area code/prefix combinations, and regions, which are defined by the boundaries of government entities.

BRFSS calculates the stratum weight (**STRWT**) using the following items:

- *Number of available records* (**NRECSTR**) and the number of records selected (**NRECSEL**) within each geographic strata and density strata.
- *Geographic strata* (**GEOSTR**), which may be the entire state or a geographic subset such as counties, census tracts, etc.
- *Density strata* (**DENSTR**) indicating the density of the phone numbers for a given block of numbers as listed or not listed.

Within each **_GEOSTR*_DENSTR** combination, BRFSS calculates the stratum weight (**_STRWT**) from the average of the **NRECSTR** and the sum of all sample records used to produce the **NRECSEL**. The stratum weight is equal to **NRECSTR / NRECSEL**.

1/NPH The inverse of the number of residential telephone numbers in the respondent's household.

NAD The number of adults 18 years and older in the respondent's household.

POSTSTRAT is the number of people in an age-by-gender or age-by-race-by-gender category in the population of a region or a state divided by the sum of the products of the preceding weights for the respondents in that same age-by-gender or age-by-race-by-gender category. It adjusts for non-coverage and non-response and, before 1995, also adjusts for different probabilities of selection by region, where applicable.

Calculation of a Child Weight prior to 2011

$$\mathbf{CHILDWT = STRWT * 1 \text{ OVER } NPH * CHILDREN * POSTCH}$$

BRFSS intends the computational formula above to reflect all the possible factors that could be taken into account in weighting a state's data. Where a factor does not apply, BRFSS sets its value to one.

CHILDWT is the final weight BRFSS assigns to each randomly selected child as identified by the adult respondent.

BRFSS calculates the design weight for the child weighting from the stratum weight times the inverse of the number of telephones in the household and then multiplies it by the number of children:

$$\mathbf{Child \ Design \ Weight = STRWT * (1/NPH) * CHILDREN}$$

POSTCH is the number of children in an age-by-gender or age-by-race-by-gender category in the population of a region or a state divided by the sum of the products of the preceding weights for the children in that same age-by-gender or age-by-race-by-gender category. It adjusts for non-coverage and non-response.

Calculation of a Household Weight prior to 2011

$$\mathbf{HOUSEWT = STRWT * 1 \text{ OVER } NPH * POSTHH}$$

HOUSEWT is the weight BRFSS assigns to each household.

BRFSS calculates the design weight for the household weighting from the stratum weight times the inverse of the number of telephones in the household:

$$\mathbf{Household \ Design \ Weight = STRWT * (1/NPH)}$$

POSTHH is the number of households in the population of a region or a state divided by the sum of the products of the preceding weights for the households in that same category. It adjusts for non-coverage and non-response.