Weighting the Data

When data are used without weights, each record counts the same as any other record. Implicit in such use are the assumptions that each record has an equal probability of being selected and that noncoverage and nonresponse are equal among all segments of the population. When deviations from these assumptions are large enough to affect the results obtained from a data set, then weighting each record appropriately can help to adjust for assumption violations. An additional (but conceptually unrelated) reason for weighting 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 is usually a state.

Following is a general description of the process that reflects factors taken into account in weighting the 2012 BRFSS data. Where a factor does not apply, its value is set to one for calculation.

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

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

In 2012 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, a compositing factor was calculated for the mostly cellular telephone dual sampling frame users.. The design weight was multiplied by the compositing factor to generate a composite weight, which is used as the raking input weight.

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 are defined by subsets of area code/prefix combinations, and regions, which are defined 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	The design weight is raked 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 geographic regions are included, four additional margins (region, region by age group, region by gender, region by race/ethnicity) are included.
_LLCPWT	The final weight assigned to each respondent.

Weight trimming is used 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

The design weight for the child weighting is calculated from the stratum weight times the inverse of the number of telephones in the household and then multiplied by the number of children:

Child Design Weight = STRWT * (1/NUMPHON2) * CHILDREN

CHIILDWT = 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

The design weight for the household weighting is calculated from the stratum weight times the inverse of the number of telephones in the household.

Household Design Weight = STRWT * (1/NUMPHON2)

HOUSEWT = STRWT * (1/ NUMPHON2) is raked 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 Formula through 2010

Please note prior to 2011 the data weighting formula used post-stratification as the weighting method. The formulas listed below apply to data weighted through 2010.

FINALWT = STRWT * 1 OVER NPH * NAD * POSTSTRAT

The computational formula above is intended to reflect all the possible factors that could be taken into account in weighting a state's data. Where a factor does not apply its value is set to one.

FINALWT is the final weight assigned 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 are defined by subsets of area code/prefix combinations, and regions, which are defined 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/ 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

CHILDWT = STRWT * 1 OVER NPH * CHILDREN * POSTCH

The computational formula above is intended to reflect all the possible factors that could be taken into account in weighting a state's data. Where a factor does not apply its value is set to one.

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

The design weight for the child weighting is calculated from the stratum weight times the inverse of the number of telephones in the household and then multiplied by the number of children:
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

HOUSEWT = STRWT * 1 OVER NPH * POSTHH

HOUSEWT is the weight assigned to each household.

The design weight for the household weighting is calculated from the stratum weight times the inverse of the number of telephones in the household:
Household Design Weight = STRWT * (1/NPH) * CHILDREN

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.