MWSUG 2018 - Paper HS-036

How to Navigate in a Maze of the Raking Macro with Advanced Weight Trimming

Michael P. Battaglia, Battaglia Consulting Group, LLC
David Izrael, Abt Associates
Sarah W. Ball, Abt Associates

ABSTRACT

Raking to population control totals is often the final step in developing survey weights. Raking is an iterative procedure that brings the weighted sample into agreement on socio-demographic variables that are available for the sample and the population. It is primarily used to reduce unit nonresponse bias. Raking can lead to some observations ending up with extreme weights; in other words, weights that are very large or very small compared to the mean weight, resulting in inflated standard errors. In 2009, we enriched a SAS® raking macro implementing weight trimming during the raking iterations, ensuring that the weighted sample agreed with the population. We recently further enhanced the macro to add several options related to weight trimming. Among them, two trimming methods - "AND" or '"OR" – and an option that allows us to set some different convergence criteria for a subset of the raking variables. This paper should help users to navigate among a number of options and parameters to more efficiently use the power of the raking macro with advanced weight trimming.

BACKGROUND ON RAKING TO CONTROL TOTALS AND SURVEY WEIGHTS

Consider a simple random sample of 500 individuals from a population of 100,000. Because each sample individual has an equal probability of selection, the base sampling weight equals 200.0 (100,000/500). In this example, for the survey being conducted, 350 individuals respond requiring the base sampling weight to be adjusted for unit nonresponse to 285.7 (200.0 * (500/350)). This weight is often referred to as the nonresponse-adjusted base sampling weight. Let us assume that for the population of 100,000 individuals we have information on respondent age by gender (for example, 4 by 2 cells) distribution from the most recent decennial census. In most surveys, younger individuals are less likely to respond than older individuals, and males may be somewhat less likely to respond than females. This differential nonresponse can be accounted for by using simple poststratification of the 350 respondents' nonresponse-adjusted base sampling weights to the 8 age by gender population totals (e.g., for females age 18-29, the population count is 20,000 and the sum of the nonresponse-adjusted base sampling weights is 17,000, yielding a simple poststratification weight adjustment multiplier of 1.1765). We therefore end up with eight different simple poststratification nonresponse-adjusted base sampling weight values instead of one.

Now, let us assume that in addition to age by gender we also have the education, marital status, and race/ethnicity distribution of the population, but not the cross-classification of these variables. We refer to this as one two-variable margin and three one-variable margins. We cannot use simple poststratification in this situation because we do not have population totals available for the cross-classification of four variables. Even if the population totals were available the sample size of respondents might not be large enough to have respondents in all cells of the cross-classification.

In such situations, one can often improve the relation between the sample and the population by adjusting the sample design weights of the individuals in the sample so that the marginal totals of the adjusted weights on specified characteristics, referred to as control variables, agree with the corresponding totals for the population. This operation is known as raking ratio estimation, raking, or sample-balancing, and the population totals are usually referred to as control totals. Raking, which is an iterative proportional fitting, is most often used to reduce bias from unit nonresponse in probability sample surveys. Raking is also a useful tool for reducing bias in nonprobability surveys such as Internet panel samples.

Raking usually proceeds one variable at a time, applying a proportional adjustment to the weights of the individuals that belong to the same category of the control variable. The initial sample design weights in the raking process, as discussed above, are often equal to the inverse of the selection probabilities and may have undergone some adjustments for unit nonresponse. The weights from the raking process are used in estimation and analysis. In choosing variables to use in raking one ideally should select variables that are related to nonresponse and are correlated with the key substantive survey variables. Also, in creating a list of raking variables one should place the most important variables at the end since the last variable in the iterative adjustment process will be in exact agreement with the population totals. See Battaglia et al. (2009) for further background on raking. The original SAS raking macro was developed by Izrael et al. (2000).

WHY ARE EXTREME WEIGHTS AN ISSUE?

Response rates in surveys conducted today are typically well below 50% with some response rates below 10%. Differential nonresponse is also an issue in almost all surveys. Therefore, when we rake to several control margins (e.g., age by gender, education, marital status, race/ethnicity, tenure status, etc.) we can end up with weights from the raking that for some respondents are very high and for other respondents are very low. This may result in the weights having a large coefficient of variation, which will lead to increased standard errors. Furthermore, respondents with high weights may carry undue influence on the weighted estimates (e.g., smoking prevalence, prevalence of adults with a rare medical condition, median financial assets, etc.). For example, if some of the respondents with the highest weights report having a rare medical condition, the prevalence estimate for that condition may be too high. On the other hand, respondents with very small weights are essentially not contributing to the estimates. In a sense, the resources used to obtain those interviews are wasted.

BASIC APPROACHES TO DEALING WITH EXTREME WEIGHTS AND WHAT ARE THE ADVANTAGES OF OUR APPROACH

When raking to population control totals to create the weights there are a few options to consider reducing the impact of extreme weights. One option is to implement the raking and then reduce the value of the weights of respondents with high weight values, and increase the value of the weights of respondents with low weight values, then, rescale the weights of all respondents so the weights sum to the size of the population. A drawback of this approach is that the weight trimming may cause the weighted distribution of the respondents to no longer agree with the population distribution on each of the raking variables. Izrael et al. (2009) developed an enhancement to the SAS raking macro that implemented weight trimming during the raking iterations. The weight trimming is implemented for up to 50 cycles. Let h = iteration number, i = raking margin (i.e., control variable), j = category of variable i, k = respondent, l = weight trimming cycle (step) for category j of variable i at iteration h, and m = weight adjustment cycle (step) after trimming for weight trimming cycle l of category j of variable i at iteration h. At cycle l the program indicates how many respondents had low weights that were increased and gives the sum of the weights before and after trimming for those respondents. At cycle l the program also indicates how many respondents had high weights that were decreased and gives the sum of the weights before and after trimming for those respondents.

For example, at iteration h = 1, control variable i = 1, control variable category j = 1, we calculate the sum of WT_{11kl} for the respondents that had their weight trimmed. Call this total X_{1111} (i.e., X_{hijl}). We then calculate $Y_{1111} = POP_{11} - X_{1111}$, where POP_{11} is the control total.

For weight adjustment cycle l = 1, we ratio-adjust the weights of the respondents that did not have their weights trimmed:

$WT_{11k11} = WT_{11kl}$ (Y_{1111} / sum of WT_{11kl} of the respondents who did not have their weights trimmed).

If the respondent had their weight trimmed then $WT_{11k11} = WT_{11kl}$.

This is implemented for each category j of control variable 1. We then go to cycle l = 2 and determine if any respondents that did not have their weights trimmed at cycle l = 1 have weights that now exceed the trimming values. We apply the weighting trimming to those respondents. The cycling is continued until no respondents in each category of control variable 1 had their weights trimmed or a maximum of 50 cycles is reached

Assuming the raking converged, our approach ensures that the weighted distribution of the respondents on each raking variable is in very close agreement with the corresponding population distribution.

WEIGHT TRIMMING FEATURES OF THE NEW RAKING MACRO

<u>To trim or not to trim</u>

The new version of SAS raking macro has an “on/off switch” that allows the user to specify if weight trimming is to take place during the raking iterations. Even if weight trimming is to be used, it is still useful to first run the raking specifying no weight trimming to establish a baseline result that can be compared with the trimmed raking result. This approach also allows for a check on convergence problems due to control total conflicts, etc.

<u>Trimming method</u>

As with the original SAS raking macro that undertakes weight trimming (Izrael et al. 2009), there are up to four trimming values that must be specified by the user of the raking macro. There are two “high end” trimming values – the individual high cap value (Individual high weight cap value (IHCV) factor: Respondent's weight multiplied by user specified value), and the global high cap value (Global high weight cap value factor: Mean input weight times user specified value). On the “low end” there are also two trimming values -- the individual low cap value (Individual low weight cap value (ILCV) factor: Respondent's weight times user specified value), and the global low cap value (Global low weight cap value factor: Mean input weight times user specified value). We have added “switches” to allow the user to turn off one or more of the four trimming methods:

GL switch: YES/NO

GH switch: YES/NO

IL switch: YES/NO

IH switch: YES/NO

Two weight trimming methods are now offered: the “OR” or the “AND” options. The “OR” method is the original trimming method developed. For “OR” trimming, each of the four trimming switches can be set to YES or NO. A case has its weight reduced if the value of the weight is greater than the individual high cap value or the global high cap value. A case has its weight increased if the value of the weight is less than the individual low cap value or the global low cap value. The “AND“-trimming method is new. The GL and IL switch *pairs* can be set to YES or NO, and the GH and IH switch *pairs* can be set to YES or NO. In the “AND” method a case has its weight reduced if the value of the weight is greater than the individual high cap value and the global high cap value. A case has its weight increased if the value of the weight is less than the individual low cap value and the global low cap value. The “OR” method will trim the weights of more cases than the “AND” method. It therefore allows for more control over extreme weights. However, if one has many raking margins and the weighted distributions prior to raking differ by a large degree from the control total distributions, convergence may be difficult to achieve. Also, keep in mind that the control variables will almost never be statistically independent. The associations between the control variables may cause convergence issues. In this situation, the “AND” method may allow for reasonable control over extreme weight values while achieving convergence.

Izrael et al. (2017) gives full details on the macro calls associated with the various weight trimming options. Below, we demonstrate a part from the raking diagnostics when different methods of trimming are used.

<u>Raking with no weight trimming:</u>

Sample size of completed interviews: 3724

Raking input weight adjusted to population total: COMPOSITE_WT_TRUNC_SCALED_ATPT

Trim weight?: NO

<u>Weight trimming using the OR method</u>:

Sample size of completed interviews: 3724

Raking input weight adjusted to population total: COMPOSITE_WT_TRUNC_SCALED_ATPT

Trim weight?: YES

Trimming method: OR

GL switch: YES

GH switch: YES

IL switch: YES

IH switch: YES

Global low weight cap value factor: Mean input weight times 0.10

Global high weight cap value factor: Mean input weight times 10.0

Individual low weight cap value (ILCV) factor: Respondent's weight times 0.167

Individual high weight cap value (IHCV) factor: Respondent's weight times 6.

<u>Weight trimming using the AND method:</u>

Sample size of completed interviews: 3579

Raking input weight adjusted to population total: COMPOSITE_WT_TRUNC_SCALED_ATPT

Trim weight?: YES

Trimming method: AND

GL and IL switch: YES

GH and IH switch: YES

Global low weight cap value factor: Mean input weight times 0.111

Global high weight cap value factor: Mean input weight times 9.0

Individual low weight cap value (ILCV) factor: Respondent's weight times 0.20

Individual high weight cap value (IHCV) factor: Respondent's weight times 5.0

The trimming diagnostics now give enhanced details on the number of cases with trimmed weights:

Number of Respondents Who Had Their Weights Decreased by the Trimming: 37

Number of Respondents Who Had Their Weights Increased by the Trimming: 315

Number of Respondents Who Had Their Weights Decreased to Global High Cap Value (GHCV): 2

Number of Respondents Who Had Their Weights Increased to Global Low Cap Value (GLCV): 280

Number of Respondents Who Had Their Weights Decreased to Individual High Cap Value (IHCV): 35

Number of Respondents Who Had Their Weights Increased to Individual Low Cap Value (ILCV): 35

GUIDELINES FOR TRIMMING WEIGHTS USING THE NEW RAKING MACRO

The specification of the weight trimming value factors and the use of the "OR versus "AND" option is determined by the user of the weighting macro. We provide guidance on how to proceed: the basic idea is to develop a reasonable approach. Reasonable can be judged relative to not using any weight trimming during raking. Goals might include avoiding cases with extremely high weights, avoiding cases with extremely small weights, and reducing the coefficient of variation of the weights. For some surveys where the average cost per interview is low, we might be less concerned about cases with extremely small weights. Trimming the weights often reduces the coefficient of variation of the raked weights but this is not guaranteed because the trimming of high and low weight values causes the weights of the non-trimmed cases to be adjusted to get agreement with the control totals. Finally, a reasonable approach to weight trimming should not trim the weights of a high proportion of the cases in the sample. Ideally, less than 20% of cases should have their weights trimmed.

The distribution of the raking input weights can be examined using a SAS PROC UNIVARIATE. The mean weight and the high and low weight values can be used to develop initial trimming values to start with. In some situations, it may be advisable to do some trimming of the raking input weights before any raking takes place.

We will show an example using CDC's Surveys to Monitor Influenza Vaccination Coverage among Health Care Personnel during the 2016-17 Influenza Season. We chose to illustrate the weight trimming using this survey because it was difficult achieving convergence. All cases in this nonprobability sample have a raking input weight of 6,680.31, so there is no variability in the raking input weights.

From our experience, the first step is to run the raking specifying no weight trimming. We can check to make sure the raking converges and examine statistics on the resulting weight. For the given survey the raking without weight trimming converged after 4 iterations using a convergence criterion of a maximum difference of 0.1 percentage points. The basic statistics on the weights are:

Weight	Mean	Min	Max	CV

Weight	Mean	Min	Max	CV
WEIGHT_ATPT (input weight)	6680.31	6680.31	6680.31	0.000
Final_wgt (raked weight)	6680.31	297.88	78334.24	1.117

We notice that without any weight trimming the maximum weight is 78,334.24, which is 11.7 times greater than the mean weight. The coefficient of variation of the raked weights is 1.117 and if possible, we would like to reduce the CV so there is less variability in the final raked weights. We will typically also examine a PROC UNIVARIATE on the raked weight with no weight trimming to get more details regarding the distribution of the weights (e.g., interquartile range, quantiles, extreme observations, etc.). We also observe that before the raking iterations start (i.e., at iteration 0) there are some very large differences between the weighted percent's and the control total percent's. For example, the occupation control variable has differences as large as 11.7 percentage points:

occupation_rake	Input Weight Sum of Weights	Target Total	Sum of Weights Difference	% of Input Weights	Target % of Weights	Difference in %
Physicians & Dentists	2044175.39	679340	1364835.39	12.551	4.171	8.380
NP/PA/Students	1235857.67	211910	1023947.67	7.588	1.301	6.287
Nurses	1115612.06	3027710	-1912097.94	6.850	18.590	-11.740
Allied Health Professionals	1309341.10	1297480	11861.10	8.039	7.967	0.073
Pharmacists	2050855.70	212610	1838245.70	12.592	1.305	11.287
Technicians/Technologists	1162374.24	1496890	-334515.76	7.137	9.191	-2.054
EMT	982005.82	160050	821955.82	6.030	0.983	5.047
Assistants/Aides	4282079.82	3892620	389459.82	26.292	23.901	2.391
Admin Support Staff/Manager	1850446.35	3716690	-1866243.65	11.362	22.821	-11.459
Non-clinical support staff	253851.85	1591300	-1337448.15	1.559	9.771	-8.212

These large differences are an indication that it may be difficult to achieve convergence when we use weight trimming.

We next used the "OR" trimming method which trims the weights of more cases than the "AND" method. The point of this raking is to ensure that do not have any cases extremely high weights. To do this we specified a global high cap value factor of 10 and an individual high cap value factor of 6, and set the low cap value factors equal to the reciprocals of those values. The following is a part from the respective raking diagnostics:

Trim weight?: YES

Trimming method: OR

GL switch: YES

GH switch: YES

IL switch: YES

IH switch: YES

Global low weight cap value factor: Mean input weight times 0.10

Global high weight cap value factor: Mean input weight times 10.0

Individual low weight cap value (ILCV) factor: Respondent's weight times 0.167

Individual high weight cap value (IHCV) factor: Respondent's weight times 6

This “OR” raking failed to converge. In other words, convergence cannot be achieved because of the trimming of the weights. Therefore, as a next step we kept the same trimming values but switched to the “AND” method which trims the weights of fewer cases. This raking converged but the weights of only 3 cases were reduced by the trimming. This resulted in a slightly reduced highest weight value of 66,803.12 compared to no weight trimming:

Weight	Mean	Min	Max	CV
WEIGHT_ATPT	6680.31	6680.31	6680.31	0.000
Final_wgt	6680.31	667.84	66803.12	1.116

Also note that the coefficient of variation of the weights is almost the same as the CV from no weight trimming.

We continued with the “AND” method and reduced the global high weight cap value factor from 10 to 6. This raking failed to converge and we observed that 638 out of 2,438 cases had their weights increased by the trimming.

We therefore modified our “AND” method to reduce both low cap value factors. We also reduced the individual high weight cap value factor to 4.0 to avoid extremely high weights:

Trim weight? YES

Trimming method: AND

GL and IL switch: YES

GH and IH switch: YES

Global low weight cap value factor: Mean input weight times 0.02

Global high weight cap value factor: Mean input weight times 6.0

Individual low weight cap value (ILCV) factor: Respondent's weight times 0.05

Individual high weight cap value (IHCV) factor: Respondent's weight times 4

Although this raking failed to converge, we found that the convergence criterion of 0.1 percentage points was attained for all raking margins except for setting_rake. We also found that reducing the two low cap values resulted in no cases having their weights increased by trimming, and that 39 cases had their weights decreased by the trimming. At this point, we are close to a reasonable solution and to achieve convergence we used a new feature of the raking macro described in the next section.

NOTE: Some users may be new to weight trimming and when we set up the macro we set initial weight trimming values based on our experience with numerous surveys. Those values may be reasonable for many surveys. They are:

Global low weight cap value factor: Mean input weight times 0.091

Global high weight cap value factor: Mean input weight times 11.0

Individual low weight cap value (ILCV) factor: Respondent's weight times 0.2

Individual high weight cap value (IHCV) factor: Respondent's weight times 5

ANOTHER NEW FEATURE -- SPECIFYING A DIFFERENT CONVERGENCE CRITERION FOR ONE OR MORE RAKING MARGINS

In most situations, a convergence criterion is specified and applied to all the raking margins. For example, one might specify as a convergence criterion that the absolute value of the maximum difference between a control total percent and weighted sample percent be less than 0.1 percentage points. This means that convergence is achieved when all the categories of each raking margin have an absolute difference between the control total percent and the weighted sample percent that is less than 0.1 percentage points. As an example, we show a control variable with 5 categories below. The Difference in % column gives the difference between the % of Output Weights and the Target % of Weights. All the absolute value percentage point differences are less than 0.1 Percentage points.

age_rake	Output Weight Sum of Weights	Target Total	Sum of Weights Difference	% of Output Weights	Target % of Weights	Difference in %	Marginal Category Difference in %
<35	5662464.49	5675092	-12627.51	34.768	34.845	-0.078	-0.223
35-44	3586314.00	3587052	-738.00	22.020	22.025	-0.005	-0.021
45-54	3345390.12	3344514	876.12	20.541	20.535	0.005	0.026
55-64	2798558.45	2785226	13332.45	17.183	17.101	0.082	0.479
65+	893872.94	894716	-843.06	5.488	5.494	-0.005	-0.094

We offer an alternative to this standard approach to raking convergence. It is possible that the control totals may come from difference sources and the accuracy of the control totals can differ. For example, one might obtain accurate up-to-date state-level control totals from the American Community Survey for variables such as age, gender, and race/ethnicity. The state-level control totals for a raking margin related to the type of telephone service in the household (cell versus landline telephone) might be derived from the National Health Interview Survey and are less accurate. In that situation, one may want to specify a convergence criterion that applies to age, gender and race/ethnicity and to specify a different convergence criterion for type of telephone service. If we used 0.1 percentage points for age, sex and race/ethnicity and 0.5 percentage points for type of telephone service then convergence would be achieved when all the categories of age, sex and race/ethnicity have an absolute value difference less than 0.1 percentage points, and all the categories of type of telephone service have an absolute value difference less than 0.5 percentage points.

This enhancement can also be used to solve convergence problems and we demonstrate below how it was applied to the example survey. Recall that all raking margins except *setting_rake* met the convergence criterion of 0.1 percentage points. In that situation convergence can be achieved by increasing the convergence criterion for the one margin that failed to converge to the maximum difference observed for the categories of that control variable. For *setting_rake* the *Other* category has a difference of -0.15 percentage points. The respective raking diagnostics follow:

Trim weight?: YES

Trimming method: AND

GL and IL switch: YES

GH and IH switch: YES

Global low weight cap value factor: Mean input weight times 0.02

Global high weight cap value factor: Mean input weight times 6.0

Individual low weight cap value (ILCV) factor: Respondent's weight times 0.05

Individual high weight cap value (IHCV) factor: Respondent's weight times 4

General tolerance (percentage points): 0.10

Raking variable with different tolerance: setting_rake

Respective different tolerances: 0.15

The raking converged in 5 iterations. We observe (see Attachment for complete raking results) that the high weight value has been reduced from 78,334.24 with no weight trimming to 41,876.32 (a 47 percent reduction). We were also able to achieve a small reduction in the coefficient of variation of the weights. Given that our main objective for this raking was to avoid extremely high weights we deem this to be a reasonable stopping point and these weights were used for analysis.

Izrael et al. (2017) gives details on the macro calls associated with changing the convergence criterion for one or more of the raking control variables.

CONCLUSION

Many surveys have low response rates causing the socio-demographic characteristics of the sample to differ from the population characteristics. Surveys that rely on nonprobability sampling may also exhibit large differences based on the source of the sample. In these situations, raking survey data without weight trimming can lead to extreme weights and a high degree of variability in the weights. Our raking macro offers many features that can be used to avoid extreme weights and in many situations also reduce variability in the weights. The ability of try different trimming methods, different trimming values, and different convergence criteria increases the likelihood that convergence will be achieved and extreme weights will be avoided.

REFERENCES

David Izrael, Michael P. Battaglia, Annabella A. Battaglia, Sarah W. Ball. You Do Not Have To Step On The Same Rake: SAS Raking Macro – Generation IV. *Proceedings of the SAS Global Forum 2017, Orlando, FL.*

Michael P. Battaglia, David Izrael, David C. Hoaglin, and Martin R. Frankel. Practical Considerations in Raking Survey Data. *Survey Practice, Online Journal, June 2009.*

Izrael D, Hoaglin DC, and Battaglia MP. (2000). A SAS Macro for Balancing a Weighted Sample. *Proceedings of the Twenty-Fifth Annual SAS Users Group International Conference*, Cary, NC: SAS Institute Inc., pp. 1350-1355.

David Izrael, Michael P. Battaglia, Martin R. Frankel. Extreme Survey Weight Adjustment as a Component of Sample Balancing (a.k.a. Raking). *Proceedings of SAS Global Forum 2009,* March 2009, Washington, D.C.

David Izrael, Michael P. Battaglia, Annabella A. Battaglia, Sarah W. Ball. You Do Not Have To Step On The Same Rake: SAS Raking Macro – Generation IV. *Proceedings of SAS Global Forum 2017,* April 2017, Orlando, FL

CONTACT

David Izrael,
David_Izrael@abtassoc.com
617.610.4382

ATTACHMENT

RAKING WITH TRIMMING WEIGHT DIAGNOSTOCS

Sample size of completed interviews: **2438**
Raking input weight adjusted to population total: **WEIGHT_ATPT**
Mean value of raking input weight adjusted to population total: **6680.31**
Minimum value of raking input weight: **6680.31**
Maximum value of raking input weight: **6680.31**
Coefficient of variation of raking input weight: **0.00**

Trim weight?: **YES**
Trimming method: **AND**
GL and IL switch: **YES**
GH and IH switch: **YES**
Global low weight cap value (GLCV): **133.61**
Global low weight cap value factor: Mean input weight times **0.02**
Global high weight cap value (GHCV): **40081.87**
Global high weight cap value factor: Mean input weight times **6.0**
Individual low weight cap value (ILCV) factor: Respondent's weight times **0.05**
Individual high weight cap value (IHCV) factor: Respondent's weight times **4**
Number of respondents who have an individual high weight cap value less than the global low weight cap value (GLCV used in weight trimming): **0**
Number of respondents who have an individual low weight cap value greater than the global high weight cap value (GHCV used in weight trimming): **0**

General tolerance (percentage points): **0.10**
Raking variable with different tolerance: **setting_rake**
Respective different tolerances: **0.15**

Weighted Distribution Prior To Raking. Iteration 0

setting_rake	Input Weight Sum of Weights	Target Total	Sum of Weights Difference	% of Input Weights	Target % of Weights	Difference in %
Hospital	5090397.54	5675260	-584862.46	31.255	34.846	-3.591
Long term	3400278.67	4505510	-1105231.33	20.878	27.664	-6.786
Other	7795923.79	6105830	1690093.79	47.867	37.490	10.377

race_rake	Input Weight Sum of Weights	Target Total	Sum of Weights Difference	% of Input Weights	Target % of Weights	Difference in %
Black non-Hispanic	1970691.96	2247281	276689.04	12.100	14.412	2.312

GEND - What is your gender?	Input Weight Sum of Weights	Target Total	Sum of Weights Difference	% of Input Weights	Target % of Weights	Difference in %
Male	5097077.85	5093782	3295.85	31.296	31.276	0.020
Female	11189522.15	11192818	-3295.85	68.704	68.724	-0.020

age_rake	Input Weight Sum of Weights	Target Total	Sum of Weights Difference	% of Input Weights	Target % of Weights	Difference in %
<35	4415686.05	5675092	-1259405.95	27.112	34.845	-7.733
35-44	4155153.90	3587052	568101.90	25.513	22.025	3.488
45-54	3660810.83	3344514	316296.83	22.477	20.535	1.942
55-64	3353516.49	2785226	568290.49	20.591	17.101	3.489
65+	701432.73	894716	-193283.27	4.307	5.494	-1.187

Region	Input Weight Sum of Weights	Target Total	Sum of Weights Difference	% of Input Weights	Target % of Weights	Difference in %
Northeast	3360196.80	3141863	218333.80	20.632	19.291	1.341
Midwest	2932656.85	3591906	-659249.15	18.007	22.054	-4.048
South	5892034.95	5930911	-38876.05	36.177	36.416	-0.239
West	4101711.40	3621920	479791.40	25.185	22.239	2.946

occupation_rake	Input Weight Sum of Weights	Target Total	Sum of Weights Difference	% of Input Weights	Target % of Weights	Difference in %
Physicians & Dentists	2044175.39	679340	1364835.39	12.551	4.171	8.380
NP/PA/Students	1235857.67	211910	1023947.67	7.588	1.301	6.287
Nurses	1115612.06	3027710	-1912097.94	6.850	18.590	-11.740
Allied Health Professionals	1309341.10	1297480	11861.10	8.039	7.967	0.073
Pharmacists	2050855.70	212610	1838245.70	12.592	1.305	11.287
Technicians/Technologists	1162374.24	1496890	-334515.76	7.137	9.191	-2.054
EMT	982005.82	160050	821955.82	6.030	0.983	5.047
Assistants/Aides	4282079.82	3892620	389459.82	26.292	23.901	2.391
Admin Support Staff/Manager	1850446.35	3716690	-1866243.65	11.362	22.821	-11.459
Non-clinical support staff	253851.85	1591300	-1337448.15	1.559	9.771	-8.212

****** Program terminated at iteration 5 because raking converged ******

Weighted Distribution After Raking

setting_rake	Output Weight Sum of Weights	Target Total	Sum of Weights Difference	% of Output Weights	Target % of Weights	Difference in %	Marginal Category Difference in %
Hospital	5690671.02	5675260	15411.02	34.941	34.846	0.095	0.272
Long term	4511614.30	4505510	6104.30	27.701	27.664	0.037	0.135
Other	6084314.68	6105830	-21515.32	37.358	37.490	-0.132	-0.352

race_rake	Output Weight Sum of Weights	Target Total	Sum of Weights Difference	% of Output Weights	Target % of Weights	Difference in %	Marginal Category Difference in %
Black,non-Hispanic	2340792.89	2347381	-6588.11	14.373	14.413	-0.040	-0.281
Hispanic	2155512.69	2154803	709.69	13.235	13.231	0.004	0.033
Other,non-Hispanic	11790294.42	11784416	5878.42	72.393	72.357	0.036	0.050

GEND - What is your gender?	Output Weight Sum of Weights	Target Total	Sum of Weights Difference	% of Output Weights	Target % of Weights	Difference in %	Marginal Category Difference in %
Male	5109803.97	5093782	16021.97	31.374	31.276	0.098	0.315
Female	11176796.03	11192818	-16021.97	68.626	68.724	-0.098	-0.143

age_rake	Output Weight Sum of Weights	Target Total	Sum of Weights Difference	% of Output Weights	Target % of Weights	Difference in %	Marginal Category Difference in %
<35	5661473.90	5675092	-13618.10	34.762	34.845	-0.084	-0.240
35-44	3589131.23	3587052	2079.23	22.037	22.025	0.013	0.058
45-54	3349591.63	3344514	5077.63	20.567	20.535	0.031	0.152
55-64	2791746.86	2785226	6520.86	17.141	17.101	0.040	0.234
65+	894656.38	894716	-59.62	5.493	5.494	-0.000	-0.007

region	Output Weight Sum of Weights	Target Total	Sum of Weights Difference	% of Output Weights	Target % of Weights	Difference in %	Marginal Category Difference in %
Northeast	3140524.78	3141863	-1338.22	19.283	19.291	-0.008	-0.043
Midwest	3598734.09	3591906	6828.09	22.096	22.054	0.042	0.190
South	5932474.56	5930911	1563.56	36.425	36.416	0.010	0.026
West	3614866.57	3621920	-7053.43	22.195	22.239	-0.043	-0.195

occupation_rake	Output Weight Sum of Weights	Target Total	Sum of Weights Difference	% of Output Weights	Target % of Weights	Difference in %	Marginal Category Difference in %
Physicians & Dentists	679340.00	679340	-0.00	4.171	4.171	-0.000	0.000
NP/PA/Students	211910.00	211910	-0.00	1.301	1.301	-0.000	0.000
Nurses	3027710.00	3027710	-0.00	18.590	18.590	-0.000	0.000
Allied Health Professionals	1297480.00	1297480	-0.00	7.967	7.967	-0.000	0.000
Pharmacists	212610.00	212610	-0.00	1.305	1.305	0.000	0.000
Technicians/Technologists	1496890.00	1496890	0.00	9.191	9.191	0.000	0.000
EMT	160050.00	160050	-0.00	0.983	0.983	-0.000	0.000
Assistants/Aides	3892620.00	3892620	-0.00	23.901	23.901	-0.000	0.000
Admin Support Staff/Manager	3716690.00	3716690	0.00	22.821	22.821	0.000	0.000
Non-clinical support staff	1591300.00	1591300	0.00	9.771	9.771	0.000	0.000

Number of Respondents Who Had Their Weights Decreased by the Trimming: **39.**
Number of Respondents Who Had Their Weights Increased by the Trimming: **0.**

Number of Respondents Who Had Their Weights Decreased to Global High Cap Value (GHCV) : **0.**
Number of Respondents Who Had Their Weights Increased to Global Low Cap Value (GLCV) : **0.**
Number of Respondents Who Had Their Weights Decreased to Individual High Cap Value (IHCV) : **39.**
Number of Respondents Who Had Their Weights Increased to Individual Low Cap Value (ILCV) : **0.**

Raking output weight: **Final_wgt**

Iteration Num	Maximum Absolute Value of Difference in Sum of	Maximum Absolute Value	Coefficient of Variation of Weights at the Completion
1			
2			
3	63793.66	0.3917	1.09765
4	24472.75	0.1503	1.09985
5	21515.32	0.1321	1.10031

Weight	Mean	Min	Max	CV
WEIGHT_ATPT	6680.31	6680.31	6680.31	0.000
Final_wgt	6680.31	268.33	41876.32	1.100