

Raking a survey sample using XLSTAT

[demoRaking.xls](#)

Dataset for raking a survey

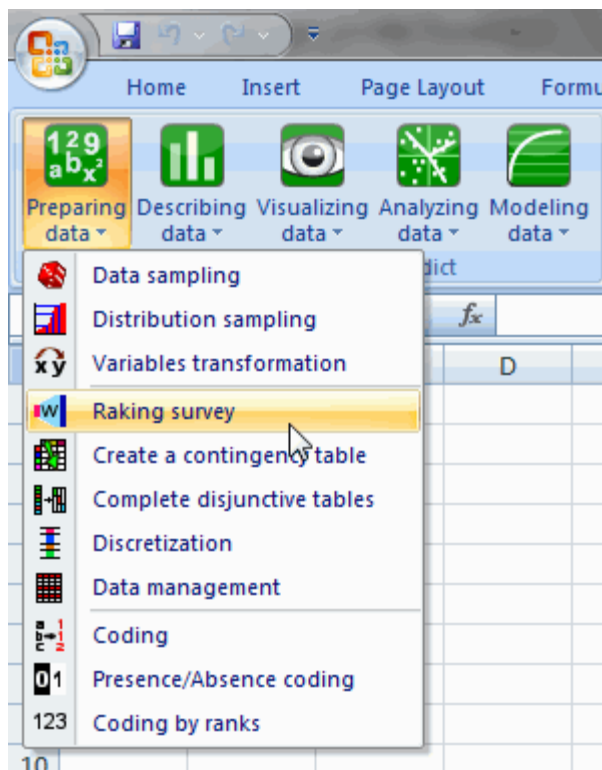
An Excel sheet with both the data and the results can be downloaded by clicking [here](#).

The data correspond to a survey based on 200 employees of a company on their satisfaction at work (simulated data). Two auxiliary variables were included in the survey: gender (1: male, 2: female) and age (1: <30, 2: 30-45 and 3: >45). The proportions of these variables in the entire company are known (marginal control totals). There are 10000 employees in this company.

In the Sat column, you can find a satisfaction score that won't be used in this step of the analysis. We are interested in finding raking weights that can be applied to our survey sample in order to obtain similar proportions for the modalities of the auxiliary variables, in the survey sample and in the population (Deming and Stephan, 1940).

Setting up the raking of a survey

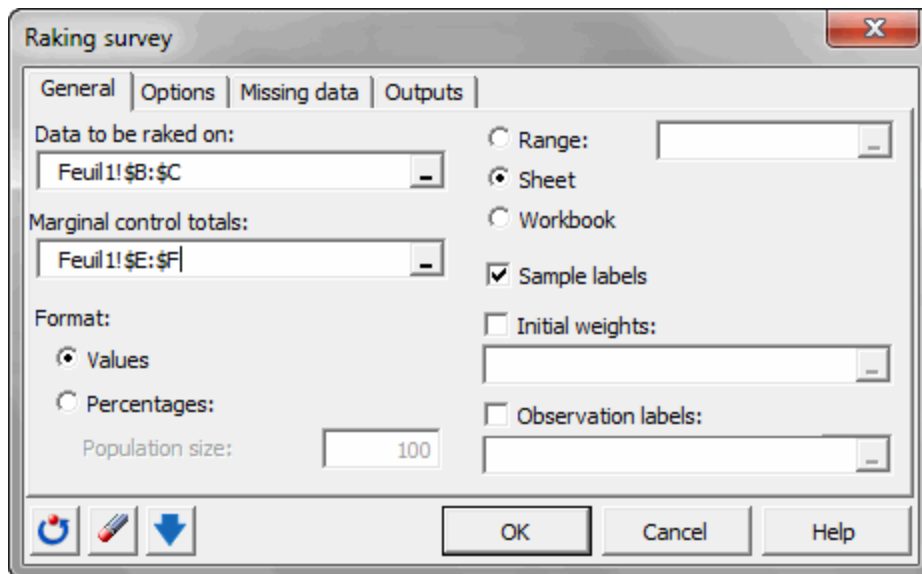
After opening XLSTAT, click the **Preparing data** button in the ribbon and select **Raking survey** (see below).



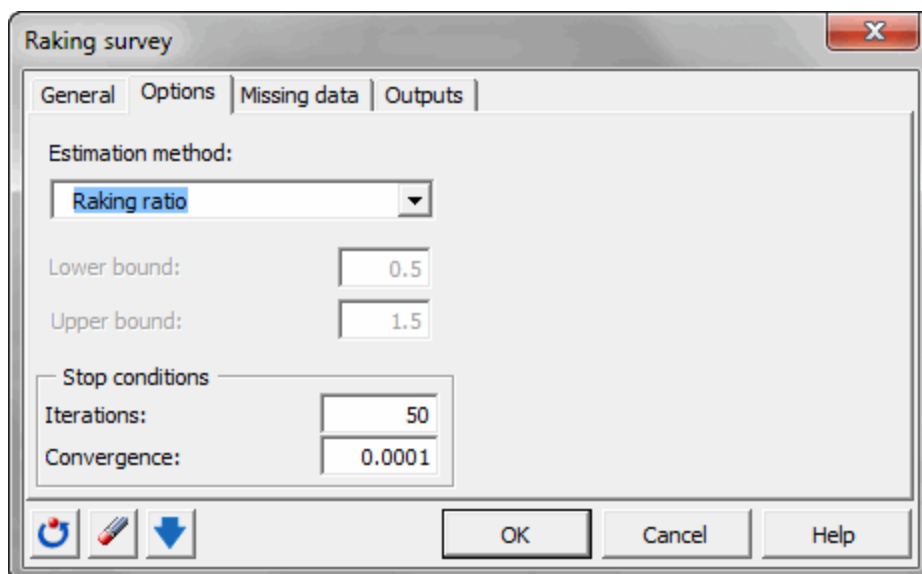
Once you've clicked on the button, the dialog box appears. Select the data on the Excel sheet. You only need to select the **auxiliary variables** (gender and age).

The **marginal control totals** have to be selected all together in the same order as the data to be raked on (one column for each variable; one row for each modality).

Each column has to sum to the same value (here 10000). As we selected the column title for the variables, we left the option **Variable labels** activated.



In the **Options** tab, we select **Raking Ratio** as estimation method.



Once you have clicked on the **OK** button, the computation starts. The results will then be displayed.

Interpreting the results of the raking of a survey

The first results displayed by XLSTAT are the basic statistics associated with the auxiliary variables before raking.

Then, the second table contains the final weights for each observation, the initial auxiliary variables and the weights ratios (final weights / initial weights) (see below for observations 1 to 13).

Final weights:				
	Gender	Age	Final weights	Weights ratio
Obs1	1	2	24,39928531	0,487985706
Obs2	2	3	140,7934679	2,815869358
Obs3	1	2	24,39928531	0,487985706
Obs4	1	1	39,69849571	0,793969914
Obs5	2	2	52,42757195	1,048551439
Obs6	1	2	24,39928531	0,487985706
Obs7	1	2	24,39928531	0,487985706
Obs8	2	2	52,42757195	1,048551439
Obs9	1	3	65,52391927	1,310478385
Obs10	2	2	52,42757195	1,048551439
Obs11	1	2	24,39928531	0,487985706
Obs12	1	1	39,69849571	0,793969914
Obs13	2	2	52,42757195	1,048551439

Then, the basic statistics after raking are displayed. We can see that using the obtained weights, the statistics are equal in the survey sample and in the population.

Summary statistics (after raking):					
Variable	Categories	Frequencies	%	Marginal control totals	Marginal control totals (%)
Gender	1	3800	38,000	3800	38,000
	2	6200	62,000	6200	62,000
Age	1	2000	20,000	2000	20,000
	2	5000	50,000	5000	50,000
	3	3000	30,000	3000	30,000

We have obtained final weights that are adapted and that can be used for further analysis of the employees' satisfaction.