An Explanation of NSSE Weights

Amber D. Lambert
October 2012

Presentation Overview

- The why, what, and how of weighting
- Calculating weights
- Running through the NSSE weighting process
- How to use weights
- How to create your own weights for subgroups

Weighting

- When creating weights, we are trying to identify how many cases each respondent should represent
- We do this to adjust the distribution of cases so that it matches an appropriate target (usually related to the population)
- Adjusting in this way will change our estimates

Why Weight?

- If the following two conditions exist, estimates of institutional and comparison group statistics will contain bias
  - Response rates differ by groups of respondents
  - Groups differ in the way they respond to survey items of interest
- So, we need to weight to arrive at more accurate (i.e. less biased) estimates of institutional and comparison group statistics

Why Does NSSE Weight?

Disproportionate representation

	Population		Respondents
Male	58%	→	33%
Full-time	88%	→	80%

Results vary by sub-groups

Community Service Participation

Gender		Enrollment Status	
Male	20%	Full-Time	80%
Female	60%	Part-Time	40%

What Does Weighting Do?

Mean $\bar{X} = \frac{\sum_{j=1}^{j=n} x_j}{n}$

Example $\bar{X} = \frac{6+5+4+5}{4} = 5$

What Does Weighting Do?

Weighted Mean $\overline{wX} = \frac{\sum_{j=1}^{j=n}(w_j \times x_j)}{\sum_{j=1}^{j=n} w_j}$

Example $\overline{wX} = \frac{(2\times6)+(6\times5)+(3\times4)+(4\times5)}{2+6+3+4}$

$= 4.93$

A Simple Situation

- At institution X, survey response looks like the following for a survey conducted this year:

Survey Response at Institution X

	Pop	Resp	Response Rate	Var1 Mean	Var2 Mean
Males	100	20	20%	2.27	3.25
Females	100	60	60%	2.86	2.94
Total	200	80	40%	2.71	3.02

- Weighting might help us better estimate the institution's mean scores

Types of Weights

Scale & Proportional

- Representative sample created
- **Population** count

Proportional

- Representative sample created
- **Respondent** count

- Both types of weights produce nearly identical item distributions, averages, and st. deviations
- 50% say "very often" using both weights

Creating Weights for Gender

Survey Response at Institution X

	Population Count	%	Respondents Count	%	Response Rate
Males	100	*50%*	20	*25%*	20%
Females	100	*50%*	60	*75%*	60%
Total	200		80		40%

Weighting Up to the Population

$W_m = P_m/R_m =$

$W_f = P_f/R_f =$

Creating Weights for Gender

Survey Response at Institution X

	Population Count	%	Respondents Count	%	Response Rate
Males	100	*50%*	20	*25%*	20%
Females	100	*50%*	60	*75%*	60%
Total	200		80		40%

Weighting Up to the Population

→ $W_m = P_m/R_m = 100/20 =$

$W_f = P_f/R_f =$

Creating Weights for Gender

Survey Response at Institution X

	Population Count	%	Respondents Count	%	Response Rate
Males	100	*50%*	20	*25%*	20%
Females	100	*50%*	60	*75%*	60%
Total	200		80		40%

Weighting Up to the Population

→ $W_m = P_m/R_m = 100/20 = 5$

$W_f = P_f/R_f =$

Creating Weights for Gender

Survey Response at Institution X

	Population		Respondents		Response
	Count	%	Count	%	Rate
Males	100	*50%*	20	*25%*	20%
Females	100	*50%*	60	*75%*	60%
Total	200		80		40%

Weighting Up to the Population

$W_m = P_m/R_m = 100/20 = 5$

$W_f = P_f/R_f = 100/60 =$

Creating Weights for Gender

Survey Response at Institution X

	Population		Respondents		Response
	Count	%	Count	%	Rate
Males	100	*50%*	20	*25%*	20%
Females	100	*50%*	60	*75%*	60%
Total	200		80		40%

Weighting Up to the Population

$W_m = P_m/R_m = 100/20 = 5$

$W_f = P_f/R_f = 100/60 = 1.67$

Creating Weights for Two Characteristics

- We can use:
  - Cell counts (population and respondents)
    - $W_i = P_i/R_i$

Population Counts

		Characteristic 2	
		Category A	Category B
Characteristic 1	Category A	P_1	P_3
	Category B	P_2	P_4

Respondents Counts

		Characteristic 2	
		Category A	Category B
Characteristic 1	Category A	R_1	R_3
	Category B	R_2	R_4

Example

W(male/full-time) = P_1/R_1
W(male/part-time) = P_3/R_3
W(female/full-time) = P_2/R_2
W(female/part-time) = P_4/R_4

Population Counts			Respondents Counts		
	Full-time	Part-time		Full-time	Part-time
Male	1000	150	Male	200	125
Female	750	100	Female	600	75

Example

W(male/full-time) = 1000/200 =
W(male/part-time) =
W(female/full-time) =
W(female/part-time) =

Population Counts			Respondents Counts		
	Full-time	Part-time		Full-time	Part-time
Male	1000	150	Male	200	125
Female	750	100	Female	600	75

Example

W(male/full-time) = 1000/200 = 5
W(male/part-time) =
W(female/full-time) =
W(female/part-time) =

Population Counts			Respondents Counts		
	Full-time	Part-time		Full-time	Part-time
Male	1000	150	Male	200	125
Female	750	100	Female	600	75

Example

W(male/full-time) = 1000/200 = 5
W(male/part-time) = 150/125 =
W(female/full-time)
W(female/part-time)

Population Counts				Respondents Counts		
	Full-time	Part-time			Full-time	Part-time
Male	1000	150		Male	200	125
Female	750	100		Female	600	75

Example

W(male/full-time) = 1000/200 = 5
W(male/part-time) = 150/125 = 1.2
W(female/full-time)
W(female/part-time)

Population Counts				Respondents Counts		
	Full-time	Part-time			Full-time	Part-time
Male	1000	150		Male	200	125
Female	750	100		Female	600	75

Example

W(male/full-time) = 1000/200 = 5
W(male/part-time) = 150/125 = 1.2
W(female/full-time) = 750/600 =
W(female/part-time)

Population Counts				Respondents Counts		
	Full-time	Part-time			Full-time	Part-time
Male	1000	150		Male	200	125
Female	750	100		Female	600	75

Example

W(male/full-time) = 1000/200 = 5
W(male/part-time) = 150/125 = 1.2
W(female/full-time) = 750/600 = 1.25
W(female/part-time)

Population Counts				Respondents Counts		
	Full-time	Part-time			Full-time	Part-time
Male	1000	150		Male	200	125
Female	750	100		Female	600	75

Example

W(male/full-time) = 1000/200 = 5
W(male/part-time) = 150/125 = 1.2
W(female/full-time) = 750/600 = 1.25
W(female/part-time) = 100/75 =

Population Counts				Respondents Counts		
	Full-time	Part-time			Full-time	Part-time
Male	1000	150		Male	200	125
Female	750	100		Female	600	75

Example

W(male/full-time) = 1000/200 = 5
W(male/part-time) = 150/125 = 1.2
W(female/full-time) = 750/600 = 1.25
W(female/part-time) = 100/75 = 1.33

Population Counts				Respondents Counts		
	Full-time	Part-time			Full-time	Part-time
Male	1000	150		Male	200	125
Female	750	100		Female	600	75

Proportional Weight Calculation

$$W_p = \frac{\text{Percent of Population}}{\text{Percent of Respondents}} = \frac{P_i / P_{total}}{R_i / R_{total}}$$

Population Counts	Full-time	Part-time
Male	1000	150
Female	750	100

Respondents Counts	Full-time	Part-time
Male	200	125
Female	600	75

Proportional Weight Calculation

$$W_{full-time,male} = \frac{\text{Percent of Population}}{\text{Percent of Respondents}} = \frac{P_i / P_{total}}{R_i / R_{total}} = \frac{1000/2000}{200/1000} = \frac{0.5}{0.2} = 2.5$$

Population Counts	Full-time	Part-time
Male	**1000**	150
Female	750	100

Respondents Counts	Full-time	Part-time
Male	**200**	125
Female	600	75

Scale vs. Proportional Weights

Respondents	Unweighted Counts	Scale		Proportional	
		Weight	Weighted Counts	Weight	Weighted Counts
Male/FT	200	5	1000	2.50	500
Male/PT	125	1.2	150	0.60	75
Female/FT	600	1.25	750	0.63	375
Female/PT	75	1.33	100	0.67	50
TOTAL	**1000**		**2000**		**1000**

How does NSSE Weight?

- First year and senior respondents weighted separately using key sub-groups:
  - Gender
  - Full-time status
    - Enrollment size also accounted for

NSSE Weighting: Cells

- Weight by gender (M,F) and enrollment status (FT,PT) within class (FYS,SR) within hundreds of institutions
- Eight cells for each institution

First-Year Students

	Enrollment Status: Full-Time	Part-Time
Male	Cell 1	Cell 3
Female	Cell 2	Cell 4

Seniors

	Enrollment Status: Full-Time	Part-Time
Male	Cell 5	Cell 7
Female	Cell 6	Cell 8

NSSE Weight Calculations

- Scale & Proportional Weight (NSSE Weight2)

$$W_2 = P_i / R_i$$

- Proportional Weight (NSSE Weight1)

$$W_1 = \frac{P_i}{P_{total}} \Big/ \frac{R_i}{R_{total}}$$

Population Counts	Full-time	Part-time
Male	P1	P3
Female	P2	P4

Respondents Counts	Full-time	Part-time
Male	R1	R3
Female	R2	R4

Special Cases

- If a cell has 1, 2, 3, or 4 cases, we assign these cases a WEIGHT2 value equal to the institutional mean of WEIGHT2 (mean based only on cells with 5 or more cases)

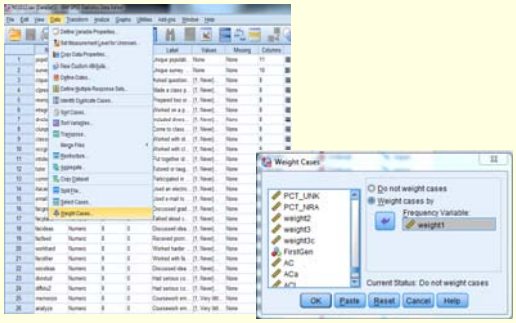
NSSE Weighting: What Do We Want?

- At NSSE we create three weights
  - WEIGHT1: Distribution of cases matches the population distribution for each institution, while the number of respondents is preserved for each institution
  - WEIGHT2: Weights up to the number of cases in the population for each institution (and consequently all institutions)
  - WEIGHT3: Distribution of cases matches the population distribution for each institution, while the number of respondents in the overall population is preserved and the proportion of cases each institution contributes to the total number of cases matches that found in the population numbers

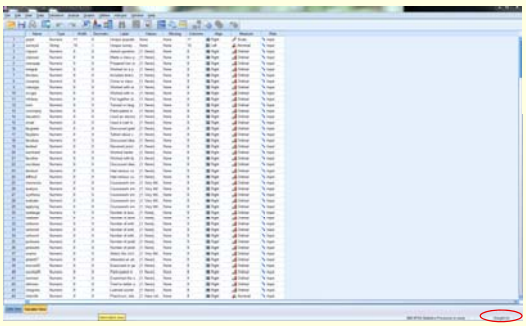
When do you use each?

- NSSE's three weights
  - WEIGHT1: In most cases, this is the weight to use as it will give you information with the numbers matching the number of NSSE respondents at your institution
  - WEIGHT2: Perhaps you would want to use in specialized cases when you want to look at results in terms of your population numbers (perhaps prediction analyses)
  - WEIGHT3: This weight is not included in your dataset, because it used for analyses done across NSSE participating institutions at the aggregate level.
- Do not use any of the NSSE weights when looking at subgroups

Applying Weights in SPSS



Applying Weights in SPSS



Applying Weights in SPSS

```
*****Apply Weights to Analyses*******.

Weight by WEIGHT1.

* Custom Tables.
CTABLES
  /VLABELS VARIABLES=enrollment ACa ACL SFI EEE SCE DISPLAY=DEFAULT
  /TABLE ACa [MEAN] + ACL [MEAN] + SFI [MEAN] + EEE [MEAN] + SCE [MEAN] BY
  enrollment
  /CATEGORIES VARIABLES=enrollment ORDER=A KEY=VALUE EMPTY=INCLUDE.

Weight off.
```

Creating Weights for Subgroups

- Bad News
  - SPSS cannot create weights for you
  - You create weighting variables on your own
- Good News
  - SPSS can determine cell counts
  - Just apply the formula provided to create your weight
  - You an also use our excel worksheet found here: http://nsse.iub.edu/uploads/Creating_Weights_NSSE_Presentation.xls

Determine Cell Counts with SPSS for Subgroups

```
CROSSTABS
  /TABLES=gender  BY
enrollment
  /FORMAT= AVALUE TABLES
  /CELLS= COUNT
  /COUNT ROUND CELL .
```

gender * enrollment Crosstabulation

Count

		enrollment		
		Part-time	Full-time	Total
gender	Male	34	320	354
	Female	68	578	646
Total		102	898	1000

Population Counts		
	Full-time	Part-time
Male	1000	150
Female	750	100

Respondents Counts		
	Full-time	Part-time
Male	320	34
Female	578	68

Attaching Weights to the Data File

```
******Attach Weights to the data file*****.
IF (gender = 1 and enrollment = 1) WEIGHT = 1.56.
IF (gender = 1 and enrollment = 2) WEIGHT = 2.21.
IF (gender = 2 and enrollment = 1) WEIGHT = .74.
IF (gender = 2 and enrollment = 2) WEIGHT = .65.

FREQ WEIGHT
    /statistics=sum.
```

Check the Weights

Statistics

WEIGHT		
N	Valid	1000
	Missing	0
Sum		1000.36

WEIGHT

		Frequency	Percent	Valid Percent	Cumulative Percent
Valid	.65	578	57.8	57.8	57.8
	.74	68	6.8	6.8	64.6
	1.56	320	32.0	32.0	96.6
	2.21	34	3.4	3.4	100.0
	Total	1000	100.0	100.0	

Weights Can't Do Everything!

- Weights cannot correct all types of survey error. Population estimates also biased by:
  - questionnaire design
  - data collection
  - Differences between responders and nonresponders

To Learn More about Weighting

Basic Overview:

- Maletta, H. (2006). *Weighting*. Retrieved from Raynald Levesque's SPSS website: http://www.spsstools.net/Tutorials/WEIGHTING.pdf\
- On NSSE weighting: http://nsse.iub.edu/2009_Institutional_Report/weighting.cfm

Advanced weighting techniques and implications:

- Thomas, S., Heck, R., & Bauer, K. (2005) Weighting and Adjusting for Design Effects in Secondary Data Analyses. In P. Umbach (Ed.), *Survey Research: Emerging Issues*. New Directions for Institutional Research, Number 127, 51-72.
- Dey, E. L. (1997). Working with low survey response rates: The efficacy of weighting adjustments. *Research in Higher Education*, 38, 215-227.

Questions

- Contact:
  - Amber D. Lambert, NSSE analyst
    - adlamber@indiana.edu