



Daniel Byrd, Ph.D.
University of California Office of the President
CAIR Conference: November 14-16, 2018
Delta Hotels Anaheim Garden Grove
Anaheim, California

Survey Says: How to Create High Quality Surveys to Assist in the Evidence Based Decision Making Process

About IRAP Survey Services

IRAP Survey Services

- **Instrument design** – IRAP's survey experts can walk you through best practices for designing an effective and valid survey, including logic and optimization for the web.
- **Sample determination** – IRAP can assist with selecting a sample population designed to give representative results.
- **Survey administration** – IRAP uses Qualtrics as an online data collection system and to send survey invitations and reminders through a safe, streamlined mailing process. IRAP can also help develop response strategies for marketing your survey.
- **Data analysis** – IRAP's researchers can help examine survey data and provide links to administrative and corporate data to deepen your analysis.
- **Reporting and visualization** – IRAP uses the Qualtrics reporting platform to provide summary data in a timely manner. IRAP can also combine quantitative analysis, skillful writing, and visualization tools such as Tableau to deliver distinctive, impactful reports.

IRAP SURVEYS

- [Response Rates of UC Surveys](#)
- [Survey Calendar](#)

SURVEY RESEARCH

- [Survey Weights](#)
- [Determining Minimum Sample Size](#)
- [Assessing Sample Representativeness](#)

CAMPUS SURVEY SITE

- [Berkeley](#)

Workshop Schedule

20 min: Survey research best practices

40 min: Developing a good survey (hands on activity)

20 min: Sampling and weighting

10 min: Concluding thoughts



The Survey Planning Process

Define Your Goals- What is the purpose of conducting the survey? Who are the decision makers? Who will be surveyed?

Conduct background research- Has anyone else done this type of research? What did they find? Do they have questions that you can use?

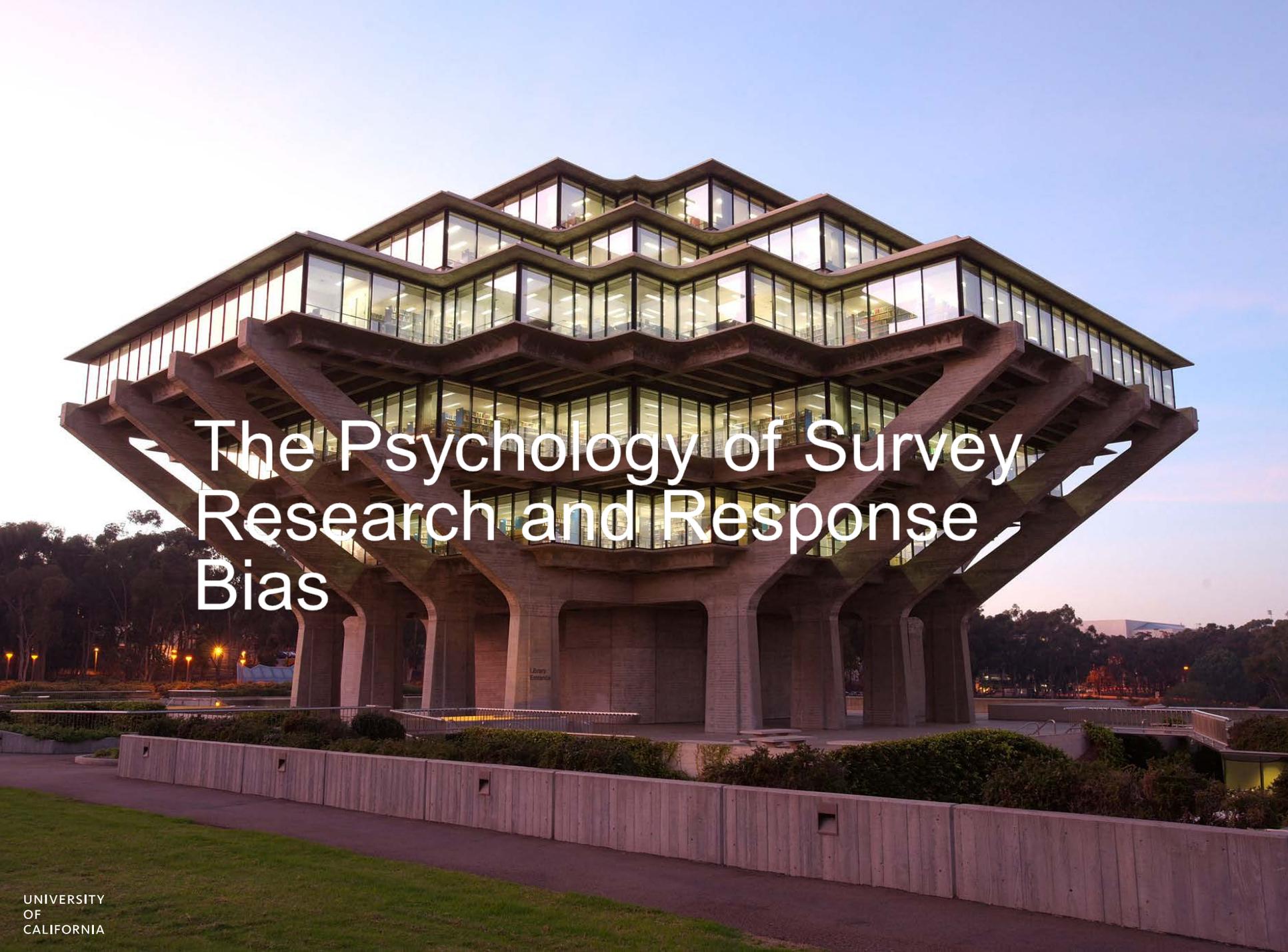
Write survey questions- If needed, write new survey questions, If not, organize the survey using the funnel method and avoid writing biased questions

Pilot test survey- It isn't a great idea to just launch a survey without pilot testing. Pilot testing helps the researcher to understand, a) if the survey questions make sense, and b) if the skip logic works. Test with both staff members and members of the population you are surveying.

Launch the survey and analyze the data- Lastly, once the survey is debugged and pilot tested, one can launch the survey

Tips for Writing Survey Questions

- **Keep questions short, but not too short-** no filler words (e.g. in order to)
- **Only ask one question at a time-** “How satisfied are you with the cost and quality of your UC education”
- **Minimize demand recall-** people don’t have photographic memories, that data can be messy and hard to interpret
- **Avoid leading questions-** You agree that the UC fosters a welcoming environment
- **Avoid jargon and acronyms and other types of complex language-** “How much did your UC education contribute to your ability to make a coherent argument”
- **Make lists exhaustive-** Include all possible options in your list, that way you won’t have a lot of respondents selecting other
- **Use all positive wording for rating scales-**while conventional wisdom states mix up responses to avoid acquiescent and extreme response bias, these are generally not a big concern. Using mixed wording can distort factor structures (for factor analysis)



The Psychology of Survey Research and Response Bias

Social Psychology and Survey Research

How surveys are constructed can affect how people respond to items- a response bias

- Implicit Social Cognition: the majority of human cognition (mental action) occurs outside of one's awareness
- Priming individuals to think about certain constructs can affect how they respond to a survey.
 - Affects the reliability and generalizability of results
- Classic types of biases
 - Assimilation effects- judgements and contextual information are positively correlated
 - Contrast effects: targets primed negatively are evaluated more positively

Additional Psychological Factors of Importance to Survey Researchers

- **Problem:** Recall questions, in particular, are cognitively taxing, the quality of data can decrease as the number of these questions increase
- **Solution:** Randomize items when possible and limit the number of recall questions
- **Problem:** A lot of research suggests people have a tendency to endorse the status quo (or say I don't know), especially when asked multiple cognitively taxing items
- **Solution:** Limit the number of cognitively taxing items (wordy questions and recall questions) on your survey and randomize

Response Bias: BEWARE

- **Question Order Matters :**

Problem: The order in which items are presented can affect how people respond to items

Solution: randomize order blocks and items within blocks when possible. But still have a logical flow to your questionnaire. Similar questions should be grouped together.

- **Ask demographics last:**

Asking demographic questions can trigger schemas which can affect responses (e.g., stereotype threat). Asking them last reduces this bias

- **Beware of social desirability:**

Asking sensitive items can lead people to respond in socially desirable ways, that often do not reflect their attitudes when measured other ways (e.g., questions about cheating and racial attitudes)
Interpret this data with caution and justify the need for asking these items

More People Favor Civil Unions When Asked After Gay Marriage

Asked first	Legal agreements	%	Gay marriage	%
	Favor	37	Favor	33
	Oppose	55	Oppose	61
	Don't know	<u>8</u>	Don't know	<u>6</u>
		100		100
Asked second	<i>Gay marriage</i>		<i>Legal agreements</i>	
	Favor	30	Favor	45
	Oppose	58	Oppose	47
	Don't know	<u>12</u>	Don't know	<u>8</u>
		100		100
N		780		735

PEW RESEARCH CENTER Oct. 2003.

Sources: [Pew Research](#)

An aerial photograph of the University of California Golden Bears stadium. The stadium is a large, oval-shaped structure with yellow seats. The field is green with white yard lines and the words "GOLDEN BEARS" and "CALIFORNIA" in yellow and blue. The stadium is surrounded by trees and buildings, including a large building with a dome on the right side. The text "Developing a Good Survey Hands on Activity" is overlaid in white on the left side of the image.

Developing a Good Survey Hands on Activity

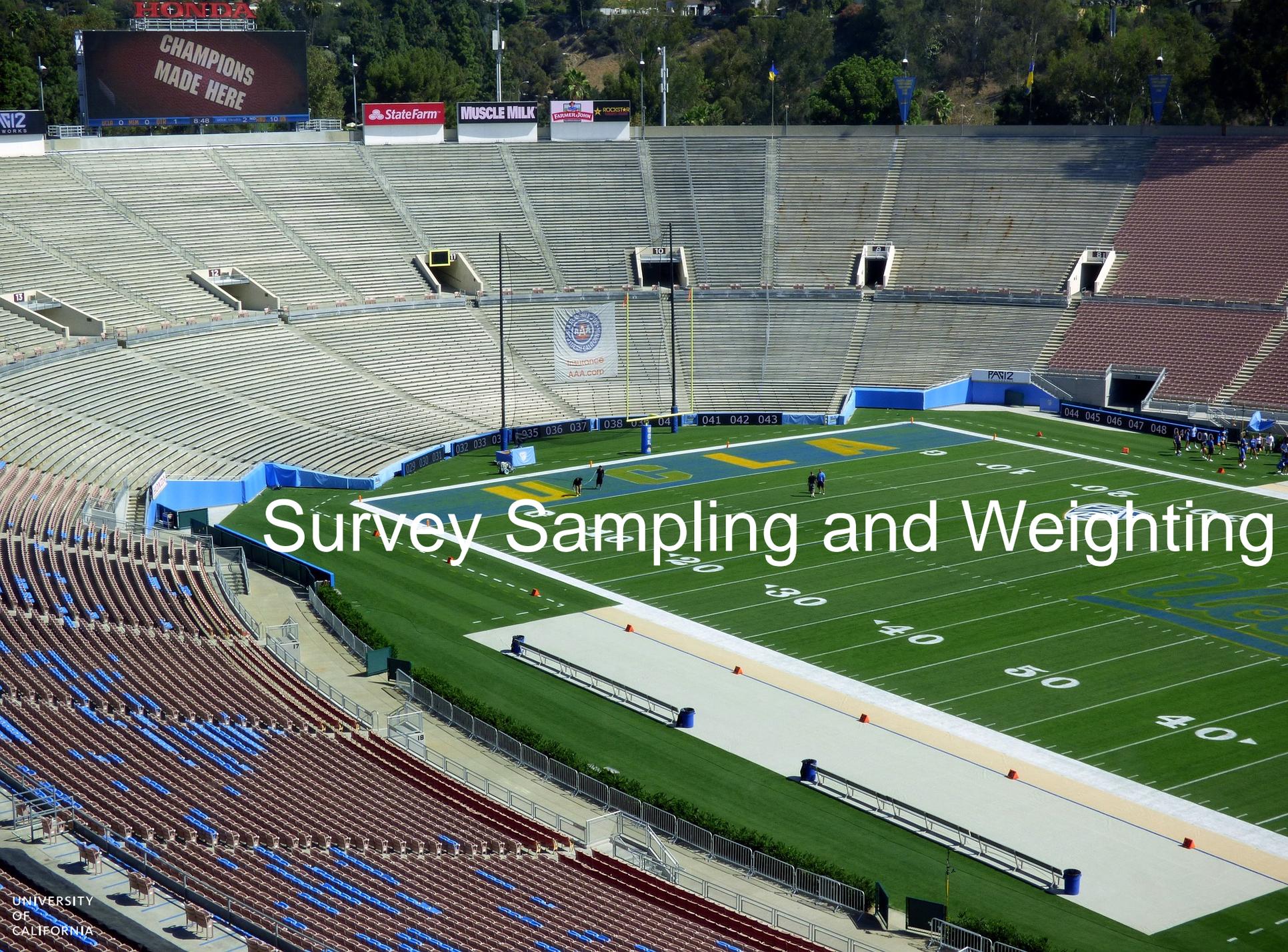
Activity Agenda

20min Small Group Activity

- Break into small groups
- Review the survey instrument
- Decide how the questions should be ordered
- Re-write problematic questions (be prepared to discuss your strategy)

20 min High Level Discussion

- What were the most problematic aspects of the items?
- What was your strategy for revising the items, and the order in which they were presented?



FLORIDA
CHAMPIONS
MADE HERE

NIZ
WORKS

State Farm

MUSCLE MILK

STARBUCKS
ROCKSTAR

AAA
AAA.COM

PANIZ

Survey Sampling and Weighting

Census or Sampling

Census	Sample
Pro: “true” measure of the population	Pro: Able to get results faster
Pro: able to analyze data from sub-groups	Pro: If sampling and weighting are done properly, one can make inferences to the general population
Con: Takes a long time to run these surveys	Con: hard to make inferences about sub populations
Con: these surveys can be more expensive	Con: if sampling is done wrong, it can affect the generalizability of the data



Sampling Types

Sampling Method	Definition	Pros	Cons
Simple Random Sampling	Every member of the population has an equal chance of being sampled	<ul style="list-style-type: none"> You don't need a large sample size. Use Survey Monkey's online calculator 	<ul style="list-style-type: none"> It's hard and almost impossible to analyze data from subpopulations
Cluster sampling	Not done as much in higher education, occurs when researchers use a multi-stage sampling approach. They randomly select geographic areas and then households	<ul style="list-style-type: none"> Method can be helpful when trying to sample smaller populations (e.g., Blacks) 	<ul style="list-style-type: none"> Weighting has to be done to make sure the findings are representative of the population.
Stratified Sampling	Done when researchers seek to have enough respondents from a population of interest. The population is divided into groups and then a SRS is drawn from the groups. Researchers sometimes over sample within certain cells to account for non-response bias for groups	<ul style="list-style-type: none"> Method can be helpful when trying to sample smaller populations (e.g., Blacks) 	<ul style="list-style-type: none"> Weighting has to be done to make sure the findings are representative of the population.

Oversampling with Stratified Sampling

What is oversampling?

- Oversampling is the process by which researchers sample a larger share from a particular group than they would from the population

Why oversample?

- It allows researchers to analyze data from sub-populations which might not be possible if drawn proportionally

How to work with over-samples

- Design weights take care of the over sampling so that the weights make the data representative of the population

Sample Size Calculator for Stratified Random Sampling

The table below gives sampling rates we will use for each of the cells above.

	ses=1	ses=2	ses=3
female=0	.70	.50	.70
female=1	.70	.50	.70

```
proc sort data = "D:hsb2";  
by female ses ;  
run;
```

```
proc surveysselect data = "D:hsb2" out = samp3 method = srs  
                    samprate = (.7 .5 .7 .7 .5 .7) seed = 9876;  
strata female ses;  
run;
```

```
proc sort data = samp3;  
by female ses;  
run;
```

Survey Weighting

Objective: Adjust for both sampling and non-response bias

Outcome: Descriptive statistics should “match” the population

Issue: Weights are great for descriptive statistics, but the jury is still out on their use with inferential statistics



Types of Weights

Design Weights: Design weights are used to correct for both over and under sampling within the population

Post-Stratification Weights: Post-stratification weights are used to correct for the error that is introduced when respondents based on a particular characteristic(s) may be more or less likely to respond to a survey when compared to others



The Logistic Approach to Weighting

Design Weights-for sample surveys

- Obtain your population file. This is the file from which you drew your sample from
- Keep in this file your demographic variables used for sampling and your student ID variable. Left join this data to your sample file.
- Create a sampled variable, where anyone who was sampled gets a 1 and anyone who wasn't gets a 0
- Run a logistic regression model with the sample variable as the dependent variable and the variables used in your stratified sampling (e.g., Race/ethnicity discipline) as your predictor variables.
- Save the predicted probabilities



The Logistic Approach to Weighting

Post Stratification Weights-census and sampled surveys

- Obtain your sample (or population file for census surveys) file
- Left join that file with your respondent file on student ID
- Create a respondent variable, where anyone who responded to the survey gets a 1 and anyone who didn't gets a 0
- Check your data to see if the distribution of respondents differs from the sampled (or whole population for census) population.
- Run a logistic regression model with the respondent variable as the dependent variable, and the variables used in your descriptive analysis (e.g., race/ethnicity discipline) as your predictor variables.
- Save the predicted probabilities



The Logistic Approach to Weighting

Create the final weights

- **For Sampled Surveys:** multiply the sampled probability with the respondent probability to get the combined probability, then use the formula below
- **For Census Surveys:** simply use the formula below



Final_Weight= Combined probability(Population total/ Sum predicted probability)*

Weighting Results

Population

ACADEMIC_YR				
ACADEMIC_YR	Frequency	Percent	Cumulative Frequency	Cumulative Percent
2001	2730	30.27	2730	30.27
2008	3101	34.38	5831	64.65
2013	3188	35.35	9019	100.00

Respondents

ACADEMIC_YR				
ACADEMIC_YR	Frequency	Percent	Cumulative Frequency	Cumulative Percent
2001	418	18.42	418	18.42
2008	746	32.88	1164	51.30
2013	1105	48.70	2269	100.00

Weighted Counts

ACADEMIC_YR				
ACADEMIC_YR	Frequency	Percent	Cumulative Frequency	Cumulative Percent
2001	2953.094	32.74	2953.094	32.74
2008	2910.274	32.27	5863.368	65.01
2013	3155.632	34.99	9019	100.00

Additional Survey Resources

- [Jon Krosnick](#) Professor of Communications, Political Science and Psychology at Stanford University
- An introduction to [factor analysis](#)
- [ICPSR](#) at the University of Michigan
- [Pew Research](#) Survey Methodology
- My contact information: Daniel.Byrd@ucop.edu

Final Thoughts

How can the material discussed today be useful in your work?

For more information about survey research at the University of California visit <https://www.ucop.edu/institutional-research-academic-planning/services/survey-services/index.html>