

BRFSS Analysis With SAS

Michelle L. Cook, MPH
Center for Health Statistics
October 14, 2009



[Outline]

- Documenting Your Analyses.
- Why...
- Creating Variables in SAS (9.1.3)
 - Race/Ethnicity
 - Underinsured
 - Disability
- A Look at a Few Procedures
- Interpreting Results
- Age-Adjusting in SAS

[Documenting Your Analyses.]

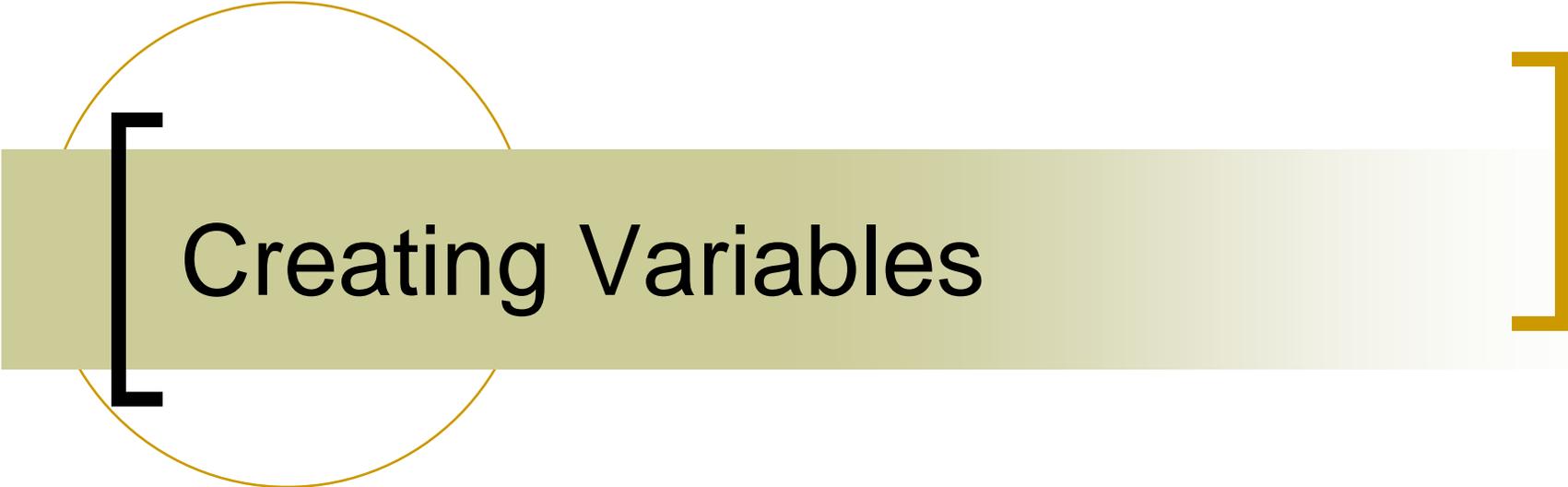
- Syntax files should be stored for your analyses.
- Notations should be made in your syntax files. (At the very least.)
- Should be able to replicate.

[Why...]

- ...do I need to use the appropriate software?
- ...do I need a sample size of at least 50?
- ...do you suppress to 50?
- ...can't I give out preliminary estimates?
- ...must I sort my data?
- ...do I assume WR when it's really not?
- ... is there only one record/PSU per cluster?

[Texas BRFSS SAS Datasets]

- Created from SPSS.
- SPSS v. 17 – something odd
- Value labels provided by SPSS in SAS program.
 - Can run program as is, but I chose to set the data and make the formats permanent.



Creating Variables

Race/Ethnicity

Underinsured

Disability

[Tips]

- It's *ALWAYS* a good idea to have a copy of the questionnaire with you when you are doing analysis.
- Make sure your data is sorted correctly.
- If-Then statements
 - Order is important b/c you may overwrite previous commands.
- Check your frequencies.

[Creating Race/Ethnicity]

- Variable for Race/Ethnicity
- Values of white, black, Hispanic, and other
- Created from multiple questions in BRFSS

c12q02

Are you Hispanic or Latino?

c12q03

Which one or more of the following would you say is your race?

c12q04

Which one of these groups would you say best represents your race?

Syntax for Race/Ethnicity

```
*CREATE RACE/ETHNICITY*;  
data library.state_08_working2;  
set library.state_08_working;  
if (c12q03b ne 8) then raceth=c12q04;  
else raceth=c12q03a;  
if (c12q04=7 or c12q04=9) then raceth=c12q03a;  
if (c12q03a=7 or c12q03a=9) then raceth=.;  
run;
```

```
data library.state_08_working2;  
set library.state_08_working2;  
if (raceth=1 and c12q02 ne 1) then raceth2=1;  
if (raceth=2 and c12q02 ne 1) then raceth2=2;  
if (c12q02=1) then raceth2=3;  
if (raceth>=3 and raceth<7 and c12q02 ne 1) then raceth2=4;  
run;
```

```
proc format;  
value raceth2f  
1 = 'White/NH'  
2 = 'Black/NH'  
3 = 'Hispanic'  
4 = 'Other/NH';  
run;
```

[Creating Underinsured]

- Variable for three levels of insurance coverage
- Adequately insured, underinsured, & uninsured
- Uses two questions from the survey
- Why we don't always restrict to 18-64 year olds...

c03q01

Do you have any kind of health care coverage, including health insurance, prepaid plans such as HMOs, or government plans such as Medicare?

c03q03

Was there a time in the past 12 months when you needed to see a doctor but could not because of cost?

[Creating Disability]

- Variable for disability
- Creates dichotomist variable
- Combines two questions from the survey

c10q01

Are you limited in any way in any activities because of physical, mental, or emotional problems?

c10q02

Do you now have any health problem that requires you to use special equipment, such as a cane, a wheelchair, a special bed, or a special telephone?

[Definition of “disabil”]

- Disability is when c10q01 is “yes” or c10q02 is “yes.”
- No disability is when c10q01 is “no” AND c10q02 is “no.”

Syntax for Underinsured & Disability

```
data library.state_08_working2;
set library.state_08_working2;
*CREATE UNDERINSURED VARIABLE*;
if (c03q01=1 and c03q03=2) then ins2=1;
if (c03q01=1 and c03q03=1) then ins2=2;
if (c03q01=2) then ins2=3;
*CREATE DISABILITY*;
if (c10q01=1 or c10q02=1) then disabil=1;
if (c10q01=2 and c10q02=2) then disabil=2;
label ins2 = 'Insurance Status, 3-level'
disabil = 'Disability Status';
run;
```

```
proc format;
value ins2f
1 = 'Adequately Insured'
2 = 'Underinsured'
3 = 'No Insurance';
value disabil
1 = 'Has Disability'
2 = 'Does Not Have Disability';
run;
```



SAS Procedures

LIBNAME
PROC FORMAT
PROC CONTENTS
PROC SORT
PROC SURVEYFREQ
PROC SQL

[Using LIBNAME]

- It's a permanent data set and you know the exact location of where it is stored.
- The modified data set will have the same name as the old data set (replacing it), unless otherwise specified.

```
libname library 'c:\documents and  
settings\michelle cook\desktop\brfss  
training\data files\';
```

[PROC FORMAT]

- Setting the data and applying the formats makes the formats permanent otherwise they are temporary.

Temporary format

```
proc format;  
value raceth2f  
1 = 'White/NH'  
2 = 'Black/NH'  
3 = 'Hispanic'  
4 = 'Other/NH';  
run;
```

[PROC CONTENTS]

- Describes the structure of the data set rather than the data values.
- Displays valuable information at the...
 - Data set level
 - Variable level
- Useful options...
 - Position – Output lists the variables by their position in the data set (default is alphabetical).
 - Short – Output is just the variable names in a row by row format.

[PROC SORT]

- Extremely important to have Texas BRFSS data sorted correctly.
 - geostr denstr seqno
 - ststr seqno
- Default is ascending order.
- Data set will be sorted according to first variable, then by the second variable within the values of the first variable.

Examples – PROC CONTENTS & PROC SORT

```
proc contents data=library.state_08_working position;  
run;
```

```
proc contents data=library.state_08_working short;  
run;
```

```
proc sort data=library.state_08_working;  
by ststr seqno;  
run;
```

[PROC SURVEYFREQ]

- Produces one-way to n-way frequency and crosstabulation tables.
- Accounts for complex sample survey data.

```
proc surveyfreq data=library.state_08_working2;  
strata ststr;  
cluster seqno;  
weight cdcfwt;  
tables ins2 disabil / cl;  
format ins2 ins2f. disabil disabil.;  
title1 'Prevalence';  
run;
```

[Interpreting Results]

- Take a moment and think about what the numbers represent.
- The differences between row and column percents.

Data Results Across Software Programs

Prevalence of Insurance Status by Software Program

	SAS		SUDAAN		SPSS - CS	
	%	95% CI	%	95% CI	%	95% CI
Adequately Insured	65.9	(64.2 - 67.5)	65.9	(64.2 - 67.5)	65.9	(64.2 - 67.5)
Underinsured	8.5	(7.6 - 9.5)	8.5	(7.6 - 9.6)	8.5	(7.6 - 9.6)
Uninsured	25.6	(24.0 - 27.2)	25.6	(24.0 - 27.2)	25.6	(24.0 - 27.2)



Age-Adjusting in SAS

SAS Macro

ODS

PROC SQL

[SAS Macro]

- A way of defining parts of or collections of SAS statements which can be carried out repeatedly or which can substitute names of datasets and variables for symbolic names.

[PROC SQL]

- Used to retrieve, update, and report on information from SAS data sets or other database products.
- Combines the functionality of DATA and PROC steps into a single step.
- Can sort, summarize, subset, merge, and concatenate data sets, create new variables.

[Uninsured Example]

Confidence Intervals are close to what you get in SUDAAN, but SUDAAN is still the “Gold Standard” for age-adjusting.

No Health Insurance (Age-Adjusted*) by Software Program

	%	95% CI
SUDAAN	24.8	(23.4 - 26.3)
SAS	24.8	(23.2 - 26.4)

*Age-adjusted to six age groups to show comparison.

[SAS Age-Adjusting Coding]

Zhiwei Zeng

Dept of Public Health

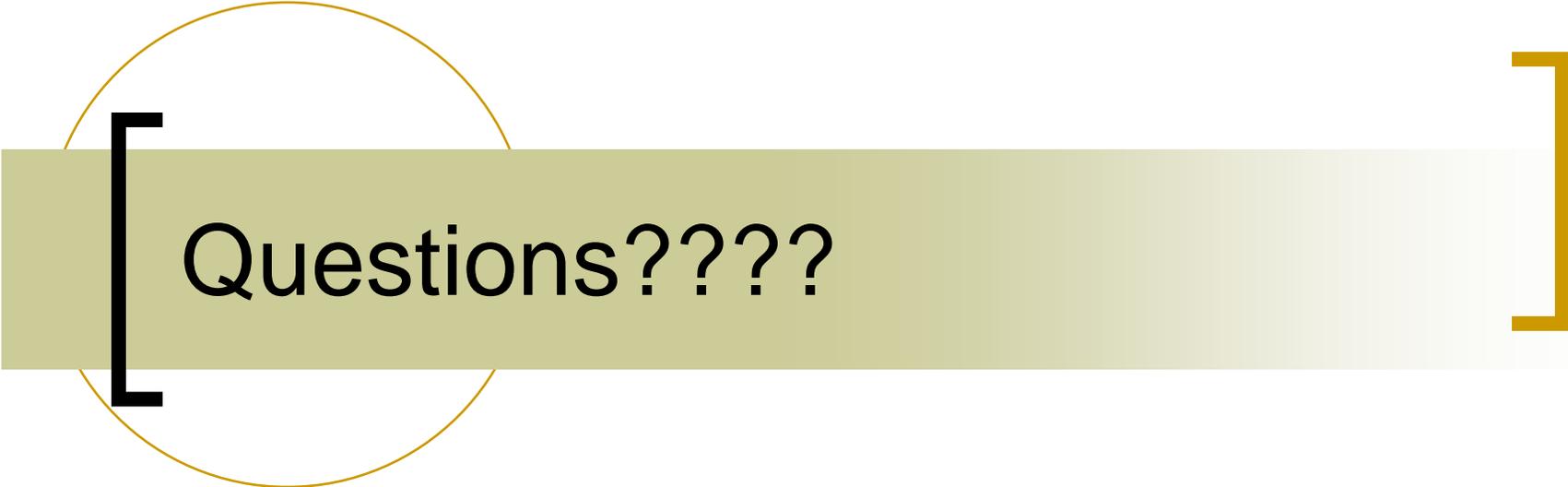
County of Los Angeles

2615 S. Grand Ave. Room 500

Los Angeles, CA 9007

E-mail: zzeng@ladhs.org

Date: Maybe 2008-2009



Questions????