WESLEYAN

UNIVERSITY

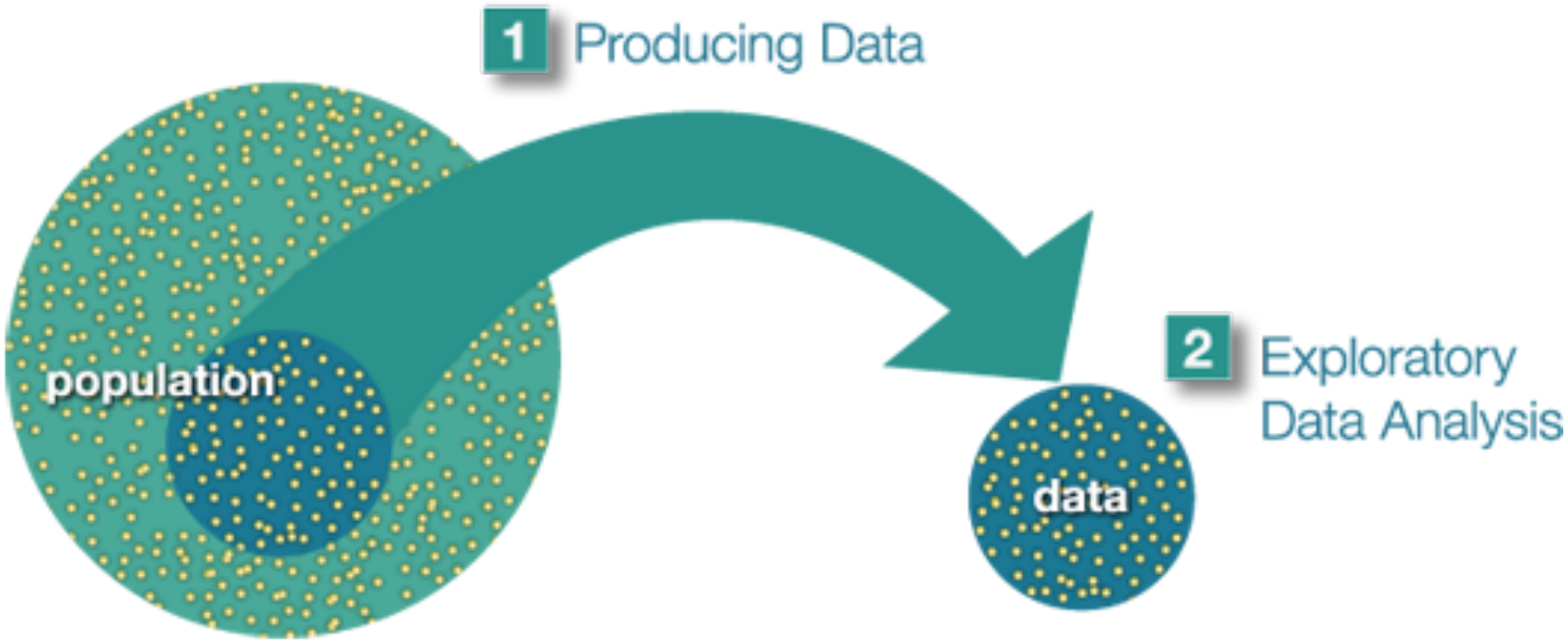
Raw Data

WESLEYAN UNIVERSITY

IDNUM	AGE	SEX	S1Q1C	S1Q1D1	S1Q1D2	S1Q1D3	S1Q1D4	S1Q1D5	S1Q6A	S3AQ1A
1	23	1	1	2	2	2	2	1	8	2
2	28	2	1	2	2	2	2	1	8	2
3	81	2	1	2	2	2	2	1	6	2
4	18	1	1	2	2	2	2	1	8	2
5	36	1	2	2	2	1	2	2	12	2
6	34	2	2	2	2	1	2	2	14	2
7	19	1	2	2	2	1	2	2	10	2
8	84	2	2	2	2	2	2	1	8	2
9	29	2	2	2	2	2	2	1	12	2
10	18	2	1	2	2	2	2	1	8	2
11	68	2	2	2	2	2	2	1	8	1
12	48	2	2	2	2	2	2	1	10	1
13	31	2	2	2	1	2	2	2	12	2
14	55	1	2	2	2	1	2	2	10	1
15	54	2	2	2	2	1	2	2	8	1
16	51	2	2	2	2	2	2	1	14	2
17	38	2	2	2	2	2	2	1	13	1
18	40	1	2	2	2	1	2	2	8	1
19	54	2	1	2	2	2	2	1	7	1
20	21	2	2	2	2	1	2	2	11	2
21	25	2	2	2	2	1	2	2	7	1
22	22	1	2	2	2	1	2	2	11	2
23	77	2	2	2	2	1	2	2	10	1
24	34	2	1	2	2	2	2	1	8	2
25	36	2	1	2	2	2	2	1	8	2
26	41	2	1	2	2	1	2	2	11	2
27	31	2	1	2	2	2	2	1	7	2
28	33	2	2	2	2	1	2	2	8	2
29	37	1	2	1	2	1	2	2	8	2
30	57	1	1	2	2	2	2	1	6	2
31	20	1	2	2	2	2	2	1	10	2
32	29	2	2	2	2	2	2	1	13	1
33	44	2	1	2	2	2	2	1	11	2
34	47	2	1	2	2	2	2	1	5	2
35	32	2	1	2	2	2	2	1	12	1
36	21	1	1	2	2	2	2	1	4	2
37	18	1	1	2	2	2	2	1	11	2
38	54	2	1	2	2	2	2	1	10	2
39	21	2	2	2	2	1	2	2	10	2

EDA

WESLEYAN UNIVERSITY



Syntax (Formal Code)

WESLEYAN UNIVERSITY

```
data boys;
input SES_scout $ delinquent $ count @@;
datalines;
low_yes yes 11 low_yes no 43
low_no yes 42 low_no no 169
med_yes yes 14 med_yes no 104
med_no yes 20 med_no no 132
high_yes yes 8 high_yes no 196
high_no yes 2 high_no no 59
;
proc freq;
 weight count;
 tables SES_scout *delinquent/all expected nocol nopct;
run;
```

Navigating SAS Enterprise Guide

WESLEYAN UNIVERSITY

sas
THE POWER TO KNOW

A random sample of 1,200 U.S. college students were asked the following questions as part of a larger survey: "What is your perception of your own body? Do you feel that you are overweight, underweight, or about right?"

The following table shows part of the data (5 of the 1200 observations);

Body Image

Student	Body Image
student 25	overweight
student 26	about right
student 27	underweight
student 28	about right
student 29	about right

In order to summarize the distribution of a categorical variable, we first create a table of the different values (categories) the variable takes, how many times each value occurs (count) and, more importantly, how often each value occurs (by converting the counts to percentages). Here is the table for our example:

Body Image Distribution

category	Count	Percent
About right	855	$\left(\frac{855}{1200}\right) * 100 = 71.3\%$
Overweight	235	$\left(\frac{235}{1200}\right) * 100 = 19.6\%$
Underweight	110	$\left(\frac{110}{1200}\right) * 100 = 9.2\%$
Total	**n=1200**	**100%**

Use Variable Names in your Program

WESLEYAN UNIVERSITY

3650-3650	TAB12MDX		NICOTINE DEPENDENCE IN THE LAST 12 MONTHS
		38131	0. No nicotine dependence
		4962	1. Nicotine dependence
721-721	CHECK321		CIGARETTE SMOKING STATUS
		9913	1. Smoked cigarettes in the past 12 months
		8078	2. Smoked cigarettes prior to the last 12 months
		22	9. Unknown
		25080	BL. NA, never or unknown if ever smoked 100+ cigarettes
722-722	S3AQ3B1		USUAL FREQUENCY WHEN SMOKED CIGARETTES
		14836	1. Every day
		460	2. 5 to 6 Day(s) a week
		687	3. 3 to 4 Day(s) a week
		747	4. 1 to 2 Day(s) a week
		409	5. 2 to 3 Day(s) a month
		772	6. Once a month or less
		102	9. Unknown
		25080	BL. NA, never or unknown if ever smoked 100+ cigarettes
723-724	S3AQ3C1		USUAL QUANTITY WHEN SMOKED CIGARETTES
		17751	1-98. Cigarette(s)
		262	99. Unknown
		25080	BL. NA, never or unknown if ever smoked 100+ cigarettes

Selecting Observations (optional)

WESLEYAN UNIVERSITY

IDNUM	AGE	SEX	TABLIFEDX	S1Q1C	ETC.
2	21	1	1	1	...
3	18	2	0	2	...
6	25	2	0	1	...
7	23	2	1	2	...
9	23	2	1	2	...
12	19	2	0	2	...

Operations used in SAS Syntax

WESLEYAN UNIVERSITY

Symbol	Mnemonic Equivalent	Definition	Example
=	EQ	equal to	a=3
^=	NE	not equal to	a ne 3
¬=	NE	not equal to	
~=	NE	not equal to	
>	GT	greater than	num>5
<	LT	less than	num<8
>=	GE	greater than or equal to	sales>=300
<=	LE	less than or equal to	sales<=100

Logic statements for sub setting my data

WESLEYAN UNIVERSITY

Smoked in the past

721-721	CHECK321		CIGARETTE SMOKING STATUS
		9913	1. Smoked cigarettes in the past 12 months
		8078	2. Smoked cigarettes prior to the last 12 months
		22	9. Unknown
		25080	BL. NA, never or unknown if ever smoked 100+ cigarettes

IF CHECK321=**1**;

Young adults age 18 to 25

68-69	AGE		AGE
		43079	18-97. Age in years
		14	98. 98 years or older

IF AGE LE **25**;

My research question…

WESLEYAN UNIVERSITY

Are smoking and nicotine dependence associated **among** young adults who have smoked in the past 12 months?

WESLEYAN

UNIVERSITY