WESLEYAN

UNIVERSITY

Making decisions with distributions

WESLEYAN UNIVERSITY

Usual Smoking Frequency				
S3AQ3B1	Frequency	Percent	Cumulative Frequency	Cumulative Percent
1	1320	77.37	1320	77.37
2	68	3.99	1388	81.36
3	91	5.33	1479	86.69
4	88	5.16	1567	91.85
5	65	3.81	1632	95.66
6	71	4.16	1703	99.82
9	3	0.18	1706	100.00

STEP 1: Set Appropriate Data to Missing

WESLEYAN UNIVERSITY

Usual Smoking Frequency

S3AQ3B1	Frequency	Percent	Cumulative Frequency	Cumulative Percent
1	1320	77.37	1320	77.37
2	68	3.99	1388	81.36
3	91	5.33	1479	86.69
4	88	5.16	1567	91.85
5	65	3.81	1632	95.66
6	71	4.16	1703	99.82
9	3	0.18	1706	100.00

IF S3AQ3B1=**9** THEN S3AQ3B1 = . ;

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

STEP 1: Set Appropriate Data to Missing

WESLEYAN UNIVERSITY

Usual Smoking Quantity

S3AQ3C1	Frequency	Percent	Cumulative Frequency	Cumulative Percent
1	83	4.87	83	4.87
2	111	6.51	194	11.37
3	114	6.68	308	18.05
4	84	4.92	392	22.98
5	163	9.55	555	32.53
6	60	3.52	615	36.05
7	45	2.64	660	38.69
8	42	2.46	702	41.15
9	6	0.35	708	41.50
10	387	22.68	1095	64.19
. . . .				
40	30	1.76	1693	99.24
60	2	0.12	1695	99.36
80	1	0.06	1696	99.41
98	1	0.06	1697	99.47
99	9	0.53	1706	100.00

IF S3AQ3C1=**99** THEN S3AQ3C1 = . ;

```
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
```

STEP 1: Set Appropriate Data to Missing

WESLEYAN UNIVERSITY

Tobacco Dependence past 12 months

TAB12MDX	Frequency	Percent	Cumulative Frequency	Cumulative Percent
0	810	47.48	810	47.48
1	896	52.52	1706	100.00

```
3650-3650        TAB12MDX                NICOTINE DEPENDENCE IN THE LAST 12 MONTHS
                                         ----------------------------------------
                              38131       0.  No nicotine dependence
                               4962       1.  Nicotine dependence
```

STEP 2: Code in Valid Data

WESLEYAN UNIVERSITY

315-315	S2AQ3		DRANK AT LEAST 1 ALCOHOLIC DRINK IN LAST 12 MONTHS
		26946	1. Yes
		16116	2. No
		31	9. Unknown
397-398	S2AQ8A		HOW OFTEN DRANK ANY ALCOHOL IN LAST 12 MONTHS
		1865	1. Every day
		1210	2. Nearly every day
		2619	3. 3 to 4 times a week
		2914	4. 2 times a week
		3261	5. Once a week
		3557	6. 2 to 3 times a month
		2663	7. Once a month
		1805	8. 7 to 11 times in the last year
		3210	9. 3 to 6 times in the last year
		3637	10. 1 or 2 times in the last year
		16147	BL. NA, former drinker or lifetime abstainer

IF S2AQ8A= **.** THEN S2AQ8A=**11**;

In other words, if **"HOW OFTEN DRANK ANY ALCOHOL IN LAST 12 MONTHS"** is blank (i.e. **shown as a period or dot in a SAS data set**) then we know that the individual drank zero times in the last month and we can give them a valid code.

STEP 3: Give variables values that are “more logical”

WESLEYAN UNIVERSITY

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

IF S2AQ3B1= **1** THEN USFREQ = **6**;
ELSE IF S2AQ3B1= **6** THEN USFREQ = **1**;
ELSE IF S2AQ3B1= **2** THEN USFREQ = **5**;
ELSE IF S2AQ3B1= **5** THEN USFREQ = **2**;
ELSE IF S2AQ3B1= **3** THEN USFREQ = **4**;
ELSE IF S2AQ3B1= **4** THEN USFREQ = **3**;

The values for my new variable USFREQ
1 = Once a month or less
2 = 2 – 3 Day(s) a month
3 = 1 – 2 Day(s) a week
4 = 3 – 4 Day(s) a week
5 = 5 – 6 Day(s) a week
6 = Every day

STEP 3: Give variables more logical values

WESLEYAN UNIVERSITY

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

IF S3AQ3B1= **1** THEN USFREQMO = **30**;
ELSE IF S3AQ3B1= **2** THEN USFREQMO = **22**;
ELSE IF S3AQ3B1= **3** THEN USFREQMO = **14**;
ELSE IF S3AQ3B1= **4** THEN USFREQMO = **5**;
ELSE IF S3AQ3B1= **5** THEN USFREQMO = **2.5**;
ELSE IF S3AQ3B1= **6** THEN USFREQMO = **1**;

The values USFREQMO (days/mo)
1 = Once a month or less
2.5 = 2 – 3 Day(s) a month
6 = 1 – 2 Day(s) a week (1.5 x 4 wk)
14 = 3 – 4 Day(s) a week (3.5 x 4 wk)
22 = 5 – 6 Day(s) a week (5.5 x 4 wk)
30 = Every day

STEP 4: Create secondary variables

WESLEYAN UNIVERSITY

USFREQMO — Usual smoking days per month

1 = Once a month or less
2.5 = 2 – 3 Day(s) a month
6 = 1 – 2 Day(s) a week (1.5 x 4 wk)
14 = 3 – 4 Day(s) a week (3.5 x 4 wk)
22 = 5 – 6 Day(s) a week (5.5 x 4 wk)
30 = Every day

S3AQ3C1 — Usual quantity when smoked cigarettes

`

1 – 98 usual cigarettes/day

NUMCIGMO_EST = USFREQMO * S3AQ3C1**;**

Creating a secondary variable for the new construct **"ESTIMATED USUAL NUMBER OF CIGARETTTES SMOKED PER MONTH" (NUMCIGMO_EST)** by multiplying usual number of days smoked by usual number of cigarettes smoked.

STEP 4: Creating secondary variables

WESLEYAN UNIVERSITY

4. Are you of Hispanic or Latino origin?			**H1GI4**	num 1
5738	0	no *[skip to Q.6]*		
743	1	yes		

6. What is your race? You may give more than one answer.

White			**H1GI6A**	num 1
2191	0	not marked		
4294	1	marked		

Black or African American **H1GI6B,** American Indian or Native American **H1GI6C**
Asian or Pacific Islander **H1GI6D**

NUMETHNIC=SUM (of H1GI4 H1GI6A H1GI6B H1GI6C H1GI6D);

IF NUMETHNIC GE **2** THEN **ETHNICITY=1**; **/*Multiple race/ethnicity endorsed*/**
ELSE IF H1GI4= **1** THEN **ETHNICITY=2**; **/*Hispanic or Latino*/**
ELSE IF H1GI6B= **1** THEN **ETHNICITY = 3**; **/*Black or African American*/**
ELSE IF H1GI6C= **1** THEN **ETHNICITY = 4**; **/*American Indian or Native American*/**
ELSE IF H1GI6D= **1** THEN **ETHNICITY = 5**; **/*Asian or Pacific Islander*/**
ELSE IF H1GI6A= **1** THEN **ETHNICITY = 6**; **/*White*/**

STEP 5: Group quantities or categories

WESLEYAN UNIVERSITY

AGE	Frequency	Percent	Cumulative Frequency	Cumulative Percent
18	161	9.44	161	9.44
19	200	11.72	361	21.16
20	221	12.95	582	34.11
21	239	14.01	821	48.12
22	228	13.36	1049	61.49
23	231	13.54	1280	75.03
24	241	14.13	1521	89.16
25	185	10.84	1706	100.00

```
IF AGE LE 20 then AGEGROUP = 1;          /*18, 19 and 20 year olds*/
ELSE IF AGE LE  22 THEN AGEGROUP=2;      /* 21 and 22 year olds*/
ELSE AGEGROUP= 3;                        /*23, 24 and 25 year olds*/
```

Creating a grouping variable for age **"AGEGROUP"** by splitting observations near the 33rd and 66th cumulative percent marks (i.e. percentiles).

WESLEYAN
UNIVERSITY