Tony Lim
BIOSTAT 203A LAB 1A
Professor Hilary Aralis
25 Nov 2019

Lab 6

Exercise 1

```
proc sort data = lb.hlth_2009;
 by IND_ID HH_ID;
run;

proc transpose data=lb.hlth_2009 out=hlth_very_long name = WAVE;
 by IND_ID HH_ID;
 var HEADACHE_2004
        HEADACHE_2006
        HEADACHE_2009
        SORETHROAT_2004
        SORETHROAT_2006
        SORETHROAT_2009
        STOMACHACHE_2004
        STOMACHACHE_2006
        STOMACHACHE_2009;
run;

data hlth_very_long (rename = (col1 = INDICATOR));
 set hlth_very_long;
 SYMPTOM = compress(WAVE, '_', 'd');
 WAVE = compress(WAVE, '_', 'a');
run;

proc print data=hlth_very_long (obs=30) noobs;
run;

proc contents data=hlth_very_long;
run;
```

	Alphabetic List of Variables and Attributes			
#	Variable	Type	Len	Label
2	HH_ID	Num	8	HOUSEHOLD ID
4	INDICATOR	Num	8	
1	IND_ID	Num	8	INDIVIDUAL ID
5	SYMPTOM	Char	16	
3	WAVE	Char	16	NAME OF FORMER VARIABLE

Exercise 2

```
proc transpose data=lb.hlth_2009 out=hlth_long_headache name = WAVE;
 by IND_ID HH_ID HH_TYPE;
 var headache_2004
        headache_2006
        headache_2009;
run;

data hlth_long_headache(rename = (col1 = HEADACHE));
 set hlth_long_headache;
 WAVE = compress(WAVE, '_', 'a');
run;

proc transpose data=lb.hlth_2009 out=hlth_long_sorethroat name = WAVE;
 by IND_ID HH_ID HH_TYPE;
 var sorethroat_2004
        sorethroat_2006
        sorethroat_2009;
run;

data hlth_long_sorethroat(rename = (col1 = SORETHROAT));
 set hlth_long_sorethroat;
 WAVE = compress(WAVE, '_', 'a');
run;

proc transpose data=lb.hlth_2009 out=hlth_long_stomachache name = WAVE;
 by IND_ID HH_ID HH_TYPE;
 var stomachache_2004
        stomachache_2006
        stomachache_2009;
run;

data hlth_long_stomachache(rename = (col1 = STOMACHACHE));
 set hlth_long_stomachache;
 WAVE = compress(WAVE, '_', 'a');
run;

proc sort data = hlth_long_headache;
 by IND_ID HH_ID HH_TYPE WAVE;
run;

proc sort data = hlth_long_sorethroat;
 by IND_ID HH_ID HH_TYPE WAVE;
```

```
run;

proc sort data = hlth_long_stomachache;
 by IND_ID HH_ID HH_TYPE WAVE;
run;

data hlth_long;
 merge hlth_long_headache
         hlth_long_sorethroat
         hlth_long_stomachache;
 by IND_ID HH_ID HH_TYPE WAVE;
run;

proc print data = hlth_long (obs = 15) noobs;
run;

proc contents data=hlth_long;
run;
```

#	Variable	Type	Len	Label
	Alphabetic List of Variables and Attributes			
5	HEADACHE	Num	8	
2	HH_ID	Num	8	HOUSEHOLD ID
3	HH_TYPE	Num	8	1=URBAN SITE(U) 2=RURAL SITE(R)
1	IND_ID	Num	8	INDIVIDUAL ID
6	SORETHROAT	Num	8	
7	STOMACHACHE	Num	8	
4	WAVE	Char	13	NAME OF FORMER VARIABLE

Exercise 3

title "Percentage of Individuals Endorsing Each Symptom Set by Household Type";
proc tabulate data = hlth_long;
       class HH_TYPE WAVE;
       var headache sorethroat stomachache;
       table (HH_TYPE = "Household Type (1 = Urban, 2 = Rural)" ALL) * (headache
sorethroat stomachache),
            (wave = "Wave")*(MEAN*F=PERCENT8.2);
       keylabel ALL = "Total";
run;

Percentage of Individuals Endorsing Each Symptom Set by Household Type

		Wave		
		2004	2006	2009
		Mean	Mean	Mean
Household Type (1 = Urban, 2 = Rural)				
1	HEADACHE	6.15%	6.26%	5.42%
	SORETHROAT	11.73%	9.70%	8.83%
	STOMACHACHE	4.67%	3.80%	2.04%
2	HEADACHE	4.63%	4.00%	3.94%
	SORETHROAT	7.91%	6.75%	8.05%
	STOMACHACHE	2.88%	2.39%	1.50%
Total	HEADACHE	5.14%	4.75%	4.43%
	SORETHROAT	9.18%	7.73%	8.31%
	STOMACHACHE	3.47%	2.86%	1.68%

Exercise 4

```
data hlth_array_long;
        set lb.hlth_2009;
        array headache_array{3}
                headache_2004
                headache_2006
                headache_2009;
        array sorethroat_array{3}
                sorethroat_2004
                sorethroat_2006
                sorethroat_2009;
        array stomachache_array{3}
                stomachache_2004
                stomachache_2006
                stomachache_2009;
        array wv{3}$ wv1-wv3 ('2004', '2006', '2009');
        do i = 1 to 3;
                HEADACHE = headache_array{i};
                SORETHROAT = sorethroat_array{i};
                STOMACHACHE = stomachache_array{i};
                WAVE = wv{i};
        output;
        end;
        keep HEADACHE SORETHROAT STOMACHACHE IND_ID HH_ID HH_TYPE WAVE;
run;

proc print data = hlth_array_long (obs = 15) noobs;
run;

proc contents data=hlth_array_long;
Run;
```

Alphabetic List of Variables and Attributes			
#	Variable	Type	Len	Label
4	HEADACHE	Num	8	
2	HH_ID	Num	8	HOUSEHOLD ID
3	HH_TYPE	Num	8	1=URBAN SITE(U) 2=RURAL SITE(R)
1	IND_ID	Num	8	INDIVIDUAL ID
5	SORETHROAT	Num	8	
6	STOMACHACHE	Num	8	
7	WAVE	Char	8	

Exercise 5
```
data wv2004 wv2006 (drop = hh_type) wv2009 (drop = hh_type);
 set hlth_array_long;
 select (WAVE);
  when ('2004') output wv2004;
  when ('2006') output wv2006;
  when ('2009') output wv2009;
 end;
run;

proc sort data=wv2004; by IND_ID HH_ID; run;
proc sort data=wv2006; by IND_ID HH_ID; run;
proc sort data=wv2009; by IND_ID HH_ID; run;

data hlth_wide;
        merge wv2004 (rename = (HEADACHE = HEADACHE_2004 STOMACHACHE =
STOMACHACHE_2004 SORETHROAT = SORETHROAT_2004))
                wv2006 (rename = (HEADACHE = HEADACHE_2006 STOMACHACHE =
STOMACHACHE_2006 SORETHROAT = SORETHROAT_2006))
                wv2009 (rename = (HEADACHE = HEADACHE_2009 STOMACHACHE =
STOMACHACHE_2009 SORETHROAT = SORETHROAT_2009));
        drop WAVE;
run;

proc print data = hlth_wide (obs = 8) noobs;
run;
```

IND_ID	HH_ID	HH_TYPE	HEADACHE_2004	SORETHROAT_2004	STOMACHACHE_2004	HEADACHE_2006
211101003002	211101003	1	0	0	0	.
211101003101	211101003	1	.	.	.	0
211101003102	211101003	1	.	.	.	0
211101008001	211101008	1	0	0	0	0
211101008002	211101008	1	0	0	0	0
211101008003	211101008	1	0	0	0	0
211101008005	211101008	1	.	.	.	0
211101008021	211101008	1	0	0	0	0

SORETHROAT_2006	STOMACHACHE_2006	HEADACHE_2009	SORETHROAT_2009	STOMACHACHE_2009
.	.	1	1	1
0	0	.	.	.
0	0	.	.	.
0	0	.	.	.
0	0	0	0	0
0	0	.	.	.
0	0	0	0	0
0	0	0	0	0

Exercise 6

```
data question2;
 set lb.hlth_2009;
 MISS_COUNT_04 = CMISS(HEADACHE_2004);
run;

proc freq data=question2;
 tables MISS_COUNT_04;
run;

data question2;
 set question2;
 if (HEADACHE_2004 ne .) & (HEADACHE_2006 ne .) then WV04_AND_06 = 1;
 else WV04_AND_06 = 0;
run;

proc freq data=question2;
 tables WV04_AND_06;
run;

data question2;
 set question2;
 if (HEADACHE_2004 ne .) & (HEADACHE_2006 ne .) & (HEADACHE_2009 ne .) then
WV04_AND_06_AND_09 = 1;
 else WV04_AND_06_AND_09 = 0;
run;

proc freq data=question2;
 tables WV04_AND_06_AND_09;
run;

data question2;
 set question2;
 if (HEADACHE_2006 ne .) & (HEADACHE_2009 ne .) & (MISS_COUNT_04 eq 1) then
WV06_AND_09 = 1;
 else WV06_AND_09 = 0;
run;

proc freq data=question2;
 tables WV06_AND_09;
run;
```

Individuals with surveys completed at each of the following time points:	N	%
2004 and at least one subsequent time point	9,826	55.5
2004 and 2006	9,120	51.5
2004, 2006, and 2009	6,733	38.0
2006 and 2009 (but not 2004)	1,532	8.65

Exercise 7
title "Individuals with surveys completed at the following time points";
proc freq data=hlth_long;
 where HEADACHE ne .;
 table WAVE/nocum;
run;

Individuals with surveys completed at the following time points

The FREQ Procedure

NAME OF FORMER VARIABLE		
WAVE	Frequency	Percent
2004	12147	33.95
2006	11714	32.74
2009	11922	33.32