```
* Homework 2;
options nocenter nonumber nodate;
libname lizard 'p:\bio113';

/* *** make sure the variables are as advertized on page 3 of HW1 ***
proc freq data=lizard.ivh;
   tables hosp--race ivh--pih acs vent dead pda1--dopa4 ptx1-ptx4;
   title 'Homework 1--Original IVH Dataset';
run;

proc means mean min max n nmiss maxdec=0 data=lizard.ivh;
run;
*/

/* here's where I save my new permanent data set (ivh1). Notice it has */
/* all the recoded variables and all the new variables. the original   */
/* data set (ivh) is unchanged.                                        */

data lizard.ivh1; set lizard.ivh;

if id=999999 then id=.;
if t4 > 99.98 then t4=.;
if labor > 9999.8 then labor=.;
if rom > 9999.8 then rom=.;

array X1 hosp--race ivh--pih acs vent dead pda1--dopa4 ptx1-ptx4;  ***    9 ***;
array X2 ga mage apg1 apg5 map1--pco2_4 col1-col4 t4age;          ***   99 ***;
array X3 los fluid1--cry4;                                        ***  999 ***;
array X4 bw wt1-wt4;                                              *** 9999 ***;

do over X1;
   if X1=9 then X1=.;
end;

do over X2;
   if X2=99 then X2=.;
end;

do over X3;
   if X3=999 then X3=.;
end;

do over X4;
   if X4=9999 then X4=.;
end;

if . < bw < 750 then bwcat=1;
else if 750 <= bw < 1000 then bwcat=2;
else if 1000 <= bw < 1250 then bwcat=3;
else if bw >= 1250  then bwcat=4;

if labor=0 then labcat=1;
else if 0 < labor <= 12 then labcat=2;
else if labor > 12 then labcat=3;

if . < rom < 1 then romcat=1;
else if rom >= 1  then romcat=2;
```

```
if . < bw <= 1000 then do;        *** 'nested' if/then ***;
   if . < ga < 26 then gabwcat=1;
   else if 26 <= ga <= 28 then gabwcat=2;
   else if ga > 28 then gabwcat=3;
end;
else if bw > 1000 then do;
   if . < ga < 26 then gabwcat=4;
   else if 26 <= ga <= 28 then gabwcat=5;
   else if ga > 28 then gabwcat=6;
end;

array CO(4) col1-col4;
array CR(4) cry1-cry4;
array WW(4) wt1-wt4;
array AA(4) cc1-cc4;
array PP(4) pctwt1-pctwt4;

do i=1 to 4;   *** all arrays have 4 vars, so process together ***;
   if sum(of CO(i) CR(i)) > 0 then
      AA(i)=1000*sum(of CO(i) CR(i))/WW(i);
      PP(i)=round((100*(bw-WW(i))/bw),0.1);
   end;

map_x1=mean(of map1-map4);
pco2_x1=mean(of pco2_1-pco2_4);
pco2_low=min(of pco2_1-pco2_4);
pco2_hi=max(of pco2_1-pco2_4);

map_x2=(map1+map2+map3+map4)/4;
pco2_x2=(pco2_1+pco2_2+pco2_3+pco2_4)/4;

if apg5 > 0 then apgrat=round((apg1/apg5),.001); *** avoid 0 divide ***;

if apg5 > 0 then a5=(apg5 < 5);

array AE(15) a5 ivh dead dopa1-dopa4 pda1-pda4 ptx1-ptx4;

ae1=0;   *** initialize counter ***;
do i=1 to 15;
   if AE(i)=1 then ae1=ae1+1;   ** increment counter inside loop ***;
end;

ae2=sum(of a5 ivh dead dopa1-dopa4 pda1-pda4 ptx1-ptx4);

drop i;

run;

proc format;
   value bwcat 1='<750g' 2='750-999g' 3='100-1249g' 4='1250+g';
   value labcat 1='none' 2='>0-12 hrs' 3='>12 hrs';
   value romcat 1='< 1 hr' 2='1+ hrs';
   value gabwcat 1='<=1kg/<26wk' 2='<=1kg/26-28wk' 3='<=1kg/>28wk'
             4='>1kg/<26wk' 5='>1kg/26-28wk' 6='>1kg/>28wk';
run;

/*  *** make sure recodes worked ***
proc freq;
   tables hosp--race ivh--pih acs vent dead pda1--dopa4 ptx1-ptx4;
```

```
   title 'Homework 1, Problem 1, Recoded IVH Dataset';
run;

proc means mean min max n nmiss maxdec=0;
   var ga bw rom labor mage--apg5 los wt1--pco2_4 fluid1--col4 t4 t4age;
run;
*/

proc freq;
   tables bwcat labcat romcat gabwcat;
   tables romcat*(bwcat labcat gabwcat);
   format bwcat bwcat. labcat labcat. romcat romcat. gabwcat gabwcat.;
   title 'Homework 3--Freqs on Categorical Vars';
run;

proc means mean min max n nmiss;
   var cc1-cc4 pctwt1-pctwt4 map_x1 pco2_x1 pco2_low pco2_hi
      map_x2 pco2_x2 apgrat;
   title 'Homework 3--Means on Continuous Vars';
run;

proc freq;
   tables ae1 ae2;
   title 'Homework 3--Freqs on Count Vars';
run;


1    * Homework 2;
2    options nocenter nonumber nodate;
3    libname lizard 'p:\bio113';
NOTE: Libref LIZARD was successfully assigned as follows:
      Engine:        V8
      Physical Name: p:\bio113
4
5    /* *** make sure the variables are as advertized on page 3 of HW1 ***
6    proc freq data=lizard.ivh;
7       tables hosp--race ivh--pih acs vent dead pda1--dopa4 ptx1-ptx4;
8       title 'Homework 1--Original IVH Dataset';
9    run;
10
11   proc means mean min max n nmiss maxdec=0 data=lizard.ivh;
12   run;
13   */
15   /* here's where I save my new permanent data set (ivh1). Notice it has */
16   /* all the recoded variables and all the new variables. the original   */
17   /* data set (ivh) is unchanged.                                        */
18
19   data lizard.ivh1; set lizard.ivh;
20
21   if id=999999 then id=.;
22   if t4 > 99.98 then t4=.;
23   if labor > 9999.8 then labor=.;
24   if rom > 9999.8 then rom=.;
25
26   array X1 hosp--race ivh--pih acs vent dead pda1--dopa4 ptx1-ptx4;  ***    9 ***;
27   array X2 ga mage apg1 apg5 map1--pco2_4 col1-col4 t4age;          ***   99 ***;
28   array X3 los fluid1--cry4;                                        ***  999 ***;
29   array X4 bw wt1-wt4;                                              *** 9999 ***;
```

```
31   do over X1;
32      if X1=9 then X1=.;
33   end;
34
35   do over X2;
36      if X2=99 then X2=.;
37   end;
38
39   do over X3;
40      if X3=999 then X3=.;
41   end;
42
43   do over X4;
44      if X4=9999 then X4=.;
45   end;
46
47   if . < bw < 750 then bwcat=1;
48   else if 750 <= bw < 1000 then bwcat=2;
49   else if 1000 <= bw < 1250 then bwcat=3;
50   else if bw >= 1250  then bwcat=4;
51
52   if labor=0 then labcat=1;
53   else if 0 < labor <= 12 then labcat=2;
54   else if labor > 12 then labcat=3;
55
56   if . < rom < 1 then romcat=1;
57   else if rom >= 1  then romcat=2;
58
59   if . < bw <= 1000 then do;        *** 'nested' if/then ***;
60      if . < ga < 26 then gabwcat=1;
61      else if 26 <= ga <= 28 then gabwcat=2;
62      else if ga > 28 then gabwcat=3;
63   end;
64   else if bw > 1000 then do;
65      if . < ga < 26 then gabwcat=4;
66      else if 26 <= ga <= 28 then gabwcat=5;
67      else if ga > 28 then gabwcat=6;
68   end;
69
70   array CO(4) col1-col4;
71   array CR(4) cry1-cry4;
72   array WW(4) wt1-wt4;
73   array AA(4) cc1-cc4;
74   array PP(4) pctwt1-pctwt4;
75
76   do i=1 to 4;   *** all arrays have 4 vars, so process together ***;
77      if sum(of CO(i) CR(i)) > 0 then
78         AA(i)=1000*sum(of CO(i) CR(i))/WW(i);
79         PP(i)=round((100*(bw-WW(i))/bw),0.1);
80      end;
81
82   map_x1=mean(of map1-map4);
83   pco2_x1=mean(of pco2_1-pco2_4);
84   pco2_low=min(of pco2_1-pco2_4);
85   pco2_hi=max(of pco2_1-pco2_4);
86
87   map_x2=(map1+map2+map3+map4)/4;
88   pco2_x2=(pco2_1+pco2_2+pco2_3+pco2_4)/4;
89
```

```
90   if apg5 > 0 then apgrat=round((apg1/apg5),.001); *** avoid 0 divide ***;
91
92   if apg5 > 0 then a5=(apg5 < 5);
93
94   array AE(15) a5 ivh dead dopa1-dopa4 pda1-pda4 ptx1-ptx4;
95
96   ae1=0;   *** initialize counter ***;
97   do i=1 to 15;
98      if AE(i)=1 then ae1=ae1+1;   ** increment counter inside loop ***;
99   end;
100
101  ae2=sum(of a5 ivh dead dopa1-dopa4 pda1-pda4 ptx1-ptx4);
102
103  drop i;
104
105  run;

NOTE: Missing values were generated as a result of performing an operation on missing
values.   Each place is given by: (Number of times) at (Line):(Column).
    15 at 77:7   193 at 78:37   204 at 79:10   204 at 79:20  204 at 79:24  204 at 79:31
    47 at 82:8    27 at 83:9     27 at 84:10    27 at 85:9   129 at 87:13   43 at 87:18
    49 at 87:23  172 at 88:16    48 at 88:23    37 at 88:30    1 at 90:25    1 at 90:36
NOTE: There were 566 observations read from the data set LIZARD.IVH.
NOTE: The data set LIZARD.IVH1 has 566 observations and 80 variables.
NOTE: DATA statement used:
      real time           0.07 seconds
      cpu time            0.06 seconds

106
107  proc format;
108     value bwcat 1='<750g' 2='750-999g' 3='100-1249g' 4='1250+g';
NOTE: Format BWCAT has been output.
109     value labcat 1='none' 2='>0-12 hrs' 3='>12 hrs';
NOTE: Format LABCAT has been output.
110     value romcat 1='< 1 hr' 2='1+ hrs';
NOTE: Format ROMCAT has been output.
111     value gabwcat 1='<=1kg/<26wk' 2='<=1kg/26-28wk' 3='<=1kg/>28wk'
112                4='>1kg/<26wk' 5='>1kg/26-28wk' 6='>1kg/>28wk';
NOTE: Format GABWCAT has been output.
113  run;

NOTE: PROCEDURE FORMAT used:
      real time           0.13 seconds
      cpu time            0.00 seconds
114
115  /*  *** make sure recodes worked ***
116  proc freq;
117     tables hosp--race ivh--pih acs vent dead pda1--dopa4 ptx1-ptx4;
118     title 'Homework 1, Problem 1, Recoded IVH Dataset';
119  run;
120
121  proc means mean min max n nmiss maxdec=0;
122     var ga bw rom labor mage--apg5 los wt1--pco2_4 fluid1--col4 t4 t4age;
123  run;
124  */
125
126  proc freq;
127     tables bwcat labcat romcat gabwcat;
128     tables romcat*(bwcat labcat gabwcat);
```

```
129     format bwcat bwcat. labcat labcat. romcat romcat. gabwcat gabwcat.;
130     title 'Homework 3--Freqs on Categorical Vars';
131  run;

NOTE: There were 566 observations read from the data set LIZARD.IVH1.
NOTE: PROCEDURE FREQ used:
      real time           0.12 seconds
      cpu time            0.01 seconds

132
133  proc means mean min max n nmiss;
134     var cc1-cc4 pctwt1-pctwt4 map_x1 pco2_x1 pco2_low pco2_hi
135        map_x2 pco2_x2 apgrat;
136     title 'Homework 3--Means on Continuous Vars';
137  run;

NOTE: There were 566 observations read from the data set LIZARD.IVH1.
NOTE: PROCEDURE MEANS used:
      real time           0.01 seconds
      cpu time            0.00 seconds

138
139  proc freq;
140     tables ae1 ae2;
141     title 'Homework 3--Freqs on Count Vars';
142  run;

NOTE: There were 566 observations read from the data set LIZARD.IVH1.
NOTE: PROCEDURE FREQ used:
      real time           0.01 seconds
      cpu time            0.01 seconds
```

Homework 3--Freqs on Categorical Vars

The FREQ Procedure

bwcat	Frequency	Percent	Cumulative Frequency	Cumulative Percent
<750g	99	17.55	99	17.55
750-999g	175	31.03	274	48.58
100-1249g	161	28.55	435	77.13
1250+g	129	22.87	564	100.00

Frequency Missing = 2

labcat	Frequency	Percent	Cumulative Frequency	Cumulative Percent
none	126	24.14	126	24.14
>0-12 hrs	237	45.40	363	69.54
>12 hrs	159	30.46	522	100.00

Frequency Missing = 44

romcat	Frequency	Percent	Cumulative Frequency	Cumulative Percent
< 1 hr	274	51.41	274	51.41
1+ hrs	259	48.59	533	100.00

Frequency Missing = 33

gabwcat	Frequency	Percent	Cumulative Frequency	Cumulative Percent
<=1kg/<26wk	110	19.57	110	19.57
<=1kg/26-28wk	144	25.62	254	45.20
<=1kg/>28wk	28	4.98	282	50.18
>1kg/<26wk	2	0.36	284	50.53
>1kg/26-28wk	110	19.57	394	70.11
>1kg/>28wk	168	29.89	562	100.00

Frequency Missing = 4

Table of romcat by bwcat

romcat bwcat

Frequency Percent Row Pct Col Pct	<750g	750-999g	100-1249 g	1250+g	Total
< 1 hr	51	84	77	61	273
	9.60	15.82	14.50	11.49	51.41
	18.68	30.77	28.21	22.34	
	53.68	51.22	50.99	50.41	
1+ hrs	44	80	74	60	258
	8.29	15.07	13.94	11.30	48.59
	17.05	31.01	28.68	23.26	
	46.32	48.78	49.01	49.59	
Total	95	164	151	121	531
	17.89	30.89	28.44	22.79	100.00

Frequency Missing = 35

Table of romcat by labcat

romcat labcat

Frequency Percent Row Pct Col Pct	none	>0-12 hrs	>12 hrs	Total
< 1 hr	92	107	59	258
	18.11	21.06	11.61	50.79
	35.66	41.47	22.87	
	74.80	46.93	37.58	
1+ hrs	31	121	98	250
	6.10	23.82	19.29	49.21
	12.40	48.40	39.20	
	25.20	53.07	62.42	
Total	123	228	157	508
	24.21	44.88	30.91	100.00

Frequency Missing = 58

Table of romcat by gabwcat

romcat gabwcat

Frequency Percent Row Pct Col Pct	<=1kg/<26wk	<=1kg/26-28wk	<=1kg/>28wk	>1kg/<26wk	>1kg/26-28wk	>1kg/>28wk	Total
< 1 hr	41	76	20	0	50	84	271
	7.75	14.37	3.78	0.00	9.45	15.88	51.23
	15.13	28.04	7.38	0.00	18.45	31.00	
	39.81	55.07	76.92	0.00	47.17	54.19	
1+ hrs	62	62	6	1	56	71	258
	11.72	11.72	1.13	0.19	10.59	13.42	48.77
	24.03	24.03	2.33	0.39	21.71	27.52	
	60.19	44.93	23.08	100.00	52.83	45.81	
Total	103	138	26	1	106	155	529
	19.47	26.09	4.91	0.19	20.04	29.30	100.00

Frequency Missing = 37

Homework 3--Means on Continuous Vars

The MEANS Procedure

Variable	Mean	Minimum	Maximum	N	N Miss
cc1	111.5081540	46.0937500	223.5294118	559	7
cc2	147.1401489	81.2949640	348.2142857	485	81
cc3	162.2093978	16.6666667	384.9056604	495	71
cc4	167.9690471	79.1366906	317.6470588	517	49
pctwt1	-0.0255773	-11.5000000	8.2000000	563	3
pctwt2	3.3347737	-20.0000000	19.9000000	486	80
pctwt3	6.7706478	-21.1000000	28.0000000	494	72
pctwt4	7.6317215	-20.2000000	30.3000000	517	49
map_x1	40.8508349	20.0000000	76.0000000	519	47
pco2_x1	36.1266234	17.0000000	68.0000000	539	27
pco2_low	30.7866419	6.0000000	68.0000000	539	27
pco2_hi	40.8812616	19.0000000	85.0000000	539	27
map_x2	39.4181159	25.0000000	63.7500000	345	221
pco2_x2	35.1626214	22.0000000	60.0000000	309	257
apgrat	0.6921145	0	1.3330000	559	7

Homework 3--Freqs on Count Vars

The FREQ Procedure

ae1	Frequency	Percent	Cumulative Frequency	Cumulative Percent
0	276	48.76	276	48.76
1	95	16.78	371	65.55
2	56	9.89	427	75.44
3	38	6.71	465	82.16
4	30	5.30	495	87.46
5	33	5.83	528	93.29
6	24	4.24	552	97.53
7	7	1.24	559	98.76
8	4	0.71	563	99.47
9	1	0.18	564	99.65
10	2	0.35	566	100.00

ae2	Frequency	Percent	Cumulative Frequency	Cumulative Percent
0	276	48.76	276	48.76
1	95	16.78	371	65.55
2	56	9.89	427	75.44
3	38	6.71	465	82.16
4	30	5.30	495	87.46
5	33	5.83	528	93.29
6	24	4.24	552	97.53
7	7	1.24	559	98.76
8	4	0.71	563	99.47
9	1	0.18	564	99.65
10	2	0.35	566	100.00