* PROBLEM #1 (DATA STEP FROM ASSIGNMENT #2);

libname x 'k:\';

data x.clinical;
infile 'k:\sasclass\data\clinical.dat';
input id $ 1-2
visit mmddyy8.
group $ 11
chol 12-14
sbp 15-17
dbp 18-20
hr 21-22
typevis $ 23;
format visit mmddyy10.;
label id = 'PERSONAL IDENTIFIER'
group = 'STUDY GROUP'
visit = 'DATE OF VISIT'
chol = 'CHOLESTEROL'
sbp = 'SYSTOLIC BLOOD PRESSURE'
dbp = 'DIASTOLIC BLOOD PRESSURE'
hr = 'HEART RATE'
typevis = 'TYPE OF VISIT';
run;

* extra;

proc contents data=x.clinical;
run;

*** USE A CLASS VARIABLE TO GROUP OBSERVATIONS IN PROC MEANS;

title "STATISTICS FOR GROUPED DATA";
proc means data=x.clinical mean median min max nonobs maxdec=1;
class group;
var chol sbp dbp hr;
run;

* PROBLEM #2 (DATA SET FROM ASSIGNMENT #3);

libname d "k:\sasclass\datasets";

*** USE PROC FREQ TO FIND % OF DEATHS WITH AN AUTOPSY;

title "A/ WAS AN AUTOPSY PERFORMED";
proc freq data=d.death99;
table aut;
table res*aut / nocol nopercent;
table res*aut / nocol nopercent missing;
run;

title "B/ COUNTS BY RACE AND ETHNICITY";
proc freq data=d.death99;
table rac*eth / norow nocol nopercent;
* extra;
table rac*eth / list;
run;
title "C/ LOCATION OF CANCER DEATHS";
proc freq data=d.death99;
table pla;
where cau eq : 'C';
run;

title "C/ AIDS DEATHS BY COUNTY";
proc freq data=d.death99 order=freq;
table res;
where cau ge : 'B20' and cau le : 'B24';
run;

*** ADD AGE-AT-DEATH TO THE DATA SET - CREATE A NEW DATA SET;
data new;
set d.death99;
age = (dod - dob) / 365;
run;

title "D/ HEART DISEASE DEATHS";
proc means data=new maxdec=1 mean median;
where (cau ge : 'I00' and cau le : 'I09') or
   cau eq : 'I11' or
   cau eq : 'I13' or
   (cau ge : 'I20' and cau le : 'I51');
class gen;
var age;
run;

*** USE PROC MEANS TO COMPUTE STATISTICS WITHIN GROUPS;

*** USE PROC UNIVARIATE TO CREATE A BOX PLOT;
data cancer99;
infile "k:sasclass\data\cancer99.dat";
input county $ 1-2
   gender $ 3
   age   4-6
   cause $ 7-10
   place $ 11
; run;

title "A/ CANCER DEATHS AMONG MALES";
proc freq data=cancer99 order=freq;
where gender eq '1';
table cause;
format cause $3.;
run;

title "B/ CANCER DEATHS AMONG FEMALES";
proc freq data=cancer99 order=freq;
where gender eq '2';
table cause;
format cause $3.;
run;

title "C/ CANCER DEATHS AMONG MALES - AGE 20-64";
proc freq data=cancer99 order=freq;
where age between 20 and 64 and gender eq '1';
table cause;
format cause $3.;
run;

title "D/ CANCER DEATHS AMONG FEMALES - AGE 20-64";
proc freq data=cancer99 order=freq;
where age between 20 and 64 and gender eq '2';
table cause;
format cause $3.;
run;