ST597B Fall 2001: Test Solution

1 a) avgdiet= mean (of dietcar1-dietcar5);
   ranged = max (of dietcar1-dietcar5) - min (of dietcar1-dietcar5);
   or = range(of dietcar1-dietcar5);

b) proc print;
   var age sex dietcar1 dietcar2 dietcar3 dietcar4 dietcar5;

c) libname john 'a:';
   data john.diet;
   set a; run;

d) data b;
   set a;
   if sex=1; run;

e) proc univariate plots cibasic cipctlnormal cipctldf normal alpha=.10;
   var dietcar1;

i) The stem and leaf and boxplot come from the plots option. ii) the normal tests the normal option. A low resolution probability plot comes from the plot option. You could also use probplot dietcar1; to get a high resolution plot. iii) The confidence intervals come from cibasic with the the 90% resulting from the use of alpha=.10. iv) comes from cipctlnormal and v) results from cipctldf.

f) proc chart;
   hbar smkestat/ discrete;
   or
   proc freq;
   tables smkestat;
   proc chart;
   vbar smkestat/discrete;

g) proc corr pearson spearman;
   var dietcar1 dietcar5;

h)
   proc sort;
   by smkestat;
   proc ttest;
   paired dietcar5*dietcar1; var diff;
   by smkestat;
   run;

The solution on the right is if in the data step you define diff = dietcar5 - dietcar1.
i) proc ttest alpha=.01;
class sex;
var avgdiet;

k) proc sort;
by sex;
proc plot;
plot dietcar1*age;
by sex;

l) add to data step
diff = dietcar5 - dietcar1;
if diff gt 0 then new =1;
if diff le 0 then new=0;
if dietcar5 =. then new=.;
if dietcar1 =. then new=.;
proc means clm;
var new;

or use the analyst to work with a t-test for a proportion requesting a confidence interval with new as the response.

m) data a;
infile 'skin.dat';
input id age sex smkestat
dietcar1 dietcar2 dietcar3 dietcar4 dietcar5;
year=1 diet=dietcar1; output;
year=2 diet=dietcar2; output;
year=3 diet=dietcar3; output;
year=4 diet=dietcar4; output;
year=5 diet=dietcar5; output;
keep id sex year diet;
run;

or
data a;
infile 'skin.dat';
input id age sex smkestat @
do i = 1 to 5;
year=i;
input diet@;
output;
end;
keep id sex year diet;
run;

2. input county $ 1-16 nbr 17-19 dist 23-26;
   input county $ 15. nbr dist;
or other combinations.