options formdlim='-';
libname aaa 'f:\sasclass\data';

proc format;
*** numeric format for numeric variable age;
value age
  0 -< 15  = '<15'
  15 -< 25  = '15-24'
  25 -< 45  = '25-44'
  45 -< 65  = '45-64'
  65 -< 85  = '65-84'
  85 - high  = '85+'
  other  = 'UNKNOWN'
;
*** character format for cardio-vascular disease;
value $cv4fmt
  'I21' - 'I229' = 'ACUTE MYOCARDIAL INFARCTION'
  'I60' - 'I609' = 'CEREBRO-VASCULAR DISEASE'
  'I70' - 'I709' = 'ATHEROSCLEROSIS'
  'I50' - 'I509' = 'HEART FAILURE'
  'I11' - 'I119' = 'HYPERTENSTIVE HEART DISEASE'
  other  = 'OTHER CARDIO-VASCULAR DISEASE'
;
*** alternative for a 3-character cause of death;
value $cv3fmt
  'I21','I22' = 'ACUTE MYOCARDIAL INFARCTION'
  'I60'       = 'CEREBRO-VASCULAR DISEASE'
  'I70'       = 'ATHEROSCLEROSIS'
  'I50'       = 'HEART FAILURE'
  'I11'       = 'HYPERTENSTIVE HEART DISEASE'
  other       = 'OTHER CARDIO-VASCULAR DISEASE'
;
*** character format for character variable race;
value $rac
  '1'     = 'WHITE'
  '2'     = 'BLACK'
  '0','3'-'8','A'-'E' = 'OTHER'
  other  = 'UNKNOWN'
;
*** character format for character variable education;
value $edu
  '00' - '11' = 'DID NOT FINISH HIGH SCHOOL'
  '12'       = 'HIGH SCHOOL'
  '13' - '15' = 'SOME COLLEGE'
  '16' - high = 'COLLEGE+'
  other  = 'UNKNOWN'
;
*** alternative format for numeric variable education;
value edu
  0  - 11  = 'DID NOT FINISH HIGH SCHOOL'
  12        = 'HIGH SCHOOL'
  13 - 15   = 'SOME COLLEGE'
  16 - high  = 'COLLEGE+'
  other     = 'UNKNOWN'
;
*** character format for character variable place;
value $pla
  'D'       = 'HOSPITAL INPATIENT'
  'A'-'C'   = 'OTHER HOSPITAL'
  'E'       = 'OTHER INSTITUTION'
  'F'       = 'AT HOME'
  'G','H','N' = 'OTHER'
  other  = 'UNKNOWN'
;
run;
*** don't alter the original data set - make a NEW one
*** add age at death;

data new;
set aaa.death99;
*calculate age at death;
ad = (dod - dob) / 365;
label aad = 'AGE AT DEATH';
run;

title 'AGE AT DEATH AND EDUCATION*RACE';
proc freq data=new;
table aad edu*rac / missprint norow nocol nopercent;
format aad age. edu $edu. rac $rac.;
run;

title 'DEATHS FROM CARDIO-VASCULAR DISEASE';
proc freq data=new;
where cau eq : 'I';
format cau $cv4fmt.;
run;

title 'MEAN AGE AT DEATH BY PLACE OF DEATH';
proc means data=new maxdec=1 mean median min max;
var aad;
class pla;
format pla $pla.;
run;

*** EXTRA;

*** alternative data step
*** make a new 3-digit cause of death
*** change education to numeric;

data newnew;
length edu 3;
set aaa.death99 (rename=(edu=tempedu));
*** prior to calculation or using a function, check if both numeric
*** variables are non-missing (more 'elegant' - no messages in LOG);
*calculate age at death;
if dod and dob then aad = (dod - dob) / 365;
else aad = .;
edu = input(tempedu,best.);
cau3 = put(cau,$3.);
drop tempedu;
label aad = 'AGE AT DEATH';
run;

proc freq data=newnew;
table edu*rac / missprint norow nocol nopercent;
format edu edu. rac $rac.;
run;

title 'DEATHS FROM CARDIO-VASCULAR DISEASE';
proc freq data=newnew;
table cau3;
where cau eq : 'I';
format cau3 $cv3fmt.;
run;