Advanced Data Analysis 2, Stat 428/528 Homework 17 Name: _____
Spring 2012 Prof. Erik B. Erhardt

Part I. (90 points) Do all calculations in SAS. Use a word processor of your choice to write a report. Insert computer text output and graphics to support what you are saying, but you need to write something that looks like an academic paper — not a pile of computer output. Turn in a hard copy of your HW in class (i.e., don't email me your HW).

(90$^{\text{pts}}$)

90 pts

1. William W. Howells Craniometric Data Set:
In this HW we will be interested in developing a rule for discriminating between male and female skulls, ignoring possible differences in populations. Initially we will consider $71 = 82 - 11$ potential discriminators. Assume equal prior probabilities for the sexes. Try to make your write-up succinct but complete — use SAS.

Data source: `web.utk.edu/~auerbach/HOWL.htm` in data file `ADA2_HW_17_skull.csv`.
Use `ADA2_HW_17_skull.sas` to read and format the data (this code is included on the course website and at the end of this assignment).
The William W. Howells Craniometric Data Set consists of craniometric measurements taken from 2524 human crania from 28 populations. Up to 82 measurements were obtained from these crania. (In addition, Dr. Howells obtained data from 524 "test" crania (crania not included in the main data set).) Sex was estimated from crania by Dr. Howells and is also provided in both data sets (and additional notes are provided about the provenance of the "test" crania).
Dr. William Howells compiled these data between 1965 and 1980, and published on them in three monographs (see below). Some details about the history of his data and the measurements may be found in his publication from 1996's American Journal of Physical Anthropology 101:441-442, which may be downloaded at web.utk.edu/~auerbach/Howells.pdf.
Any use of these data in presented or published research carries the stipulation that the source of the data be cited. Acceptable citations for the data include the reference of the monographs about these data and of this web site.
Monographs about the Howells Craniometric Data Set:
Howells WW. 1973. Cranial Variation in Man. A Study by Multivariate Analysis of Patterns of Differences Among Recent Human Populations. Papers of the Peabody Museum of Archeology and Ethnology, vol. 67, pp. 259. Cambridge, Mass.: Peabody Museum.
Howells WW. 1989. Skull Shapes and the Map. Craniometric Analyses in the Dispersion of Modern Homo. Papers of the Peabody Museum of Archaeology and Ethnology, vol. 79, pp. 189. Cambridge, Mass.: Peabody Museum.
Howells WW. 1995. Who's Who in Skulls. Ethnic Identification of Crania from Measurements. Papers of the Peabody Museum of Archaeology and Ethnology, vol. 82, pp. 108. Cambridge, Mass.: Peabody Museum.

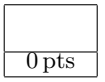
Database Documentation For W.W. Howells' (1973, 1989) Craniometric Series This database consists of the raw variates for W.W. Howells craniometric series (1973, 1989) in ascii format. It consists of 81 craniometric measurements for over 2524 individuals from 28 populations.
Rows: are data on individuals.
Columns: columns are variables (indicator variables and measurement variates). Variables 2-4 are indicator variables, and variables 5-86 are measurement variates (refer to Measurement Notes below regarding measurement scale, etc.). All columns are defined below:

0 pts

Var	Label	Indicator or Measurement
1.	ID	Number in dataset
2.	Sex	Designation of male or female for each individual: Male = 1, Female = 2. See Sex (below).
3.	PopNum	Population affiliation for each individual, coded as 1--28. See Populations (below) for coding designations.
4.	Population	Population name
5.	GOL	Glabello-occipital length
6.	NOL	Nasio-occipital length
7.	BNL	Basion-nasion length
8.	BBH	Basion-bregma height
9.	XCB	Maximum cranial breadth
10.	XFB	Maximum frontal breadth
11.	ZYB	Bizygomatic breadth
12.	AUB	Biauricular breadth
13.	WCB	Minimum cranial breadth
14.	ASB	Biasterionic breadth
15.	BPL	Basion-prosthion length
16.	NPH	Nasion-prosthion height
17.	NLH	Nasal height
18.	JUB	Bijugal breadth
19.	NLB	Nasal breadth
20.	MAB	Palate breadth, external
21.	MDH	Mastoid height
22.	MDB	Mastoid breadth
23.	OBH	Orbit height, left
24.	OBB	Orbit breadth, left
25.	DKB	Interorbital breadth
26.	NDS	Naso-dacryal subtense
27.	WNB	Simotic chord (least nasal breadth)
28.	SIS	Simotic subtense
29.	ZMB	Bimaxillary breadth
30.	SSS	Zygomaxillary subtense
31.	FMB	Bifrontal breadth
32.	NAS	Nasio-frontal subtense
33.	EKB	Biorbital breadth
34.	DKS	Dacryon subtense
35.	IML	Malar length, inferior
36.	XML	Malar length, maximum
37.	MLS	Malar subtense
38.	WMH	Cheek height
39.	SOS	Supraorbital projection
40.	GLS	Glabella projection
41.	STB	Bistephanic breadth
42.	FRC	Nasion-bregma chord (Frontal chord)
43.	FRS	Nasion-bregma subtense (Frontal subtense)
44.	FRF	Nasion-subtense fraction
45.	PAC	Bregma-lambda chord (Parietal chord)
46.	PAS	Bregma-lambda subtense (Parietal subtense)
47.	PAF	Bregma-subtense fraction
48.	OCC	Lambda-opisthion chord (Occipital chord)
49.	OCS	(OCS)
50.	OCF	(OCF)
51.	FOL	Foramen magnum length
52.	NAR	(NAR)
53.	SSR	(SSR)
54.	PRR	(PRR)
55.	DKR	(DKR)
56.	ZOR	(ZOR)
57.	FMR	(FMR)
58.	EKR	(EKR)
59.	ZMR	(ZMR)
60.	AVR	(AVR)
61.	BRR	(BRR)
62.	VRR	(VRR)
63.	LAR	(LAR)
64.	OSR	(OSR)
65.	BAR	(BAR)
66.	NAA	(NAA)
67.	PRA	(PRA)
68.	BAA	(BAA)
69.	NBA	(NBA)
70.	BBA	(BBA)
71.	BRA	(BRA)
72.	SSA	(SSA)
73.	NFA	(NFA)
74.	DKA	(DKA)
75.	NDA	(NDA)
76.	SIA	(SIA)
77.	FRA	(FRA)
78.	PAA	(PAA)
79.	OCA	(OCA)
80.	RFA	(RFA)
81.	RPA	(RPA)
82.	ROA	(ROA)
83.	BSA	(BSA)
84.	SBA	(SBA)
85.	SLA	(SLA)
86.	TBA	(TBA)

Populations

As noted above, population affiliation for each individual is denoted in column 1. Each population is coded by values 1–28 specified below:



```
PopNum Population
    1. Norse: (Medieval), Europe, Oslo
    2. Zalavar: Central Europe, Hungary
    3. Berg: Central Europe, Carinthia, Austria
    4. Teita: East Africa, Kenya
    5. Dogon: West Africa, Mali
    6. Zulu: South Africa
    7. Lake Alexandrina Tribes: South Australia
    8. Tasmanian: Tasmania
    9. Tolai: Melanesia, New Britain
   10. Mokapu: Oahu, Hawaii, Polynesia
   11. Easter Island: Polynesia
   12. Moriori: Chatham Islands, Polynesia
   13. Arikara: (Early) North America
   14. Santa Cruz Island: California, N. America
   15. Yauyos: Peru, South America
   16. Hokkaido: North Japan
   17. North Kyushu: South Japan
   18. Hainan: Haikou City, China
   19. Atayal: Taiwan Aboriginals
   20. Phillipine: Phillipine Islands
   21. Guam: Latte Period
   22. Egypt: Gizeh, 26th-30th Dynasties
   23. San: South Africa
   24. Andaman Islands: Andaman Islands
   25. Ainu: S. and SE. Kokkaido, Japan
   26. Buriat: Siberia
   27. Eskimo: Inugsuk, Greenland
   28. Anyang: Shang Dynasty, China
```

Sex
All populations are represented by both males and females except: 20 Phillipines and 28 Anyang, which have only males in the samples.

Measurement Notes

1. There are very few missing data entries; these are denoted with periods (.).
2. Columns 5–86 are linear measurements. All linear measurements are recorded to the nearest millimeter except: columns WNB and SIS, which are recorded to the nearest 0.1 millimeter.

Sample sizes
In a couple of instances (e.g., PopNum 3 Berg males and 27 Eskimo females) sample sizes in the present data base are off by 1 individual from those published in Howells (1973, 1989). It is unclear whether these are additions or omissions in the original study, or the result of database transformation to the present format. However, given the relatively large N for both of these samples, this is unlikely to significantly affect the outcome of any analysis.

Reading the data. SAS program `ADA2_HW_17_skull.sas` is available on the course website. I insist that you use this (or mimic it in your code) since it correctly chooses the variables for analysis and names them. It is also included at the end of this document.

(a) (15 pts) Do a preliminary analysis comparing male and female skulls on each of the 71 variables. Informally check for normality and constant variability between the sexes. If any transformations are suggested, carry them out and re-check for normality and constant variance. Which of the variables appear to have information for distinguishing male from female skulls? Only present a few plots here! REMARK: do not overwork yourself here.

(b) (15 pts) **Note:** *For parts (b)–(e), assume that the population covariance matrices for sexes are equal.* Using the scale selected for the 71 variables, use STEPDISC to identify the variables that are important for discriminating between male and female skulls. Briefly discuss the output.

(c) (15 pts) Given the variables selected in (b) use CANDISC to identify the linear combination that best distinguishes sexes. Plot the data on the canonical variable and describe the separation that you see between sexes. Do the plots suggest that the canonical variable will allow you to successfully classify new data? Discuss.

(d) (10 pts) Use DISCRIM to see how well the classification rule works on the variables selected in (b) (i.e., the data from which the classification rule was constructed). Note we are doing linear discriminant analysis here.

55 pts

Use this program to perform the classification and discuss the results. Note: from using `proc candisc` above and using the option `out=outdisc`, we can use `can1` to classify sex. See how well it works by looking at the error rate.
Program editor contents:

```
proc discrim data=outdisc testdata=outdisc pool=yes;
  class sex;
  testclass sex;
  var can1;
run;

* Note that using can1 as above gives the same result ;
*   as listing all the variables like this           ;
proc discrim data=skull testdata=skull pool=yes;
  class sex;
  testclass sex;
  var ** list variables retained from stepdisc here **;
run;
```

(e) (15 pts) The following SAS program will break the skull data into two data sets skulltrain and skulltest with essentially the same number of observations. I have described in comments how it works (another way would be to randomly sample two groups of equal size). Use data set skulltrain to develop the linear classification rule based on the variables selected (when you used the entire data set) by `STEPDISC`.
Program editor contents:

```
* split the skull dataset into two output datasets;
*   one to train the classifier and one to test it;
data skulltrain skulltest;
  set skull;
  if _n_ = 2*int(0.5*_n_)
    then output skulltrain;   * even observations are training data set;
    else output skulltest;    * odd  observations are testing  data set;
run;

proc sort data = skulltrain;
  by sex;
run;

proc discrim data=skulltrain testdata=skulltest pool=yes;
  class sex;
  testclass sex;
  var ** list variables retained from stepdisc here **;
run;
```

Then see how well this rule predicts the sex for observations in the test data set skulltest. Contrast the misclassification results with those obtained in (d). This analysis gives a better idea than (d) on how the classification rule will work with new data. Do the two analyses lead to similar conclusions, or does the previous analysis lead to overly optimistic error rates? Discuss.

(f) (10 pts) Test the assumption that the population covariance matrices are equal (with `pool=test`), using the variables selected in STEPDISC. If this hypothesis is rejected, redo parts (d) and (e) using quadratic discrimination (with `pool=no`). Compare the results from the linear and quadratic analyses. Does the quadratic rule improve (decrease) error rates noticeably? Discuss results.

(g) (10 pts) Write a coherent yet concise summary of the analysis.

Appendix: SAS program `ADA2_HW_17_skull.sas` for reading the data.

```
options ls=79 nodate nocenter;

* label the population numbers with their names;
proc format;
  value formatPopNum
     1 = 'Norse: (Medieval), Europe, Oslo'
     2 = 'Zalavar: Central Europe, Hungary'
     3 = 'Berg: Central Europe, Carinthia, Austria'
     4 = 'Teita: East Africa, Kenya'
     5 = 'Dogon: West Africa, Mali'
     6 = 'Zulu: South Africa'
     7 = 'Lake Alexandrina Tribes: South Australia'
     8 = 'Tasmanian: Tasmania'
     9 = 'Tolai: Melanesia, New Britain'
    10 = 'Mokapu: Oahu, Hawaii, Polynesia'
```

35 pts

```
    11 = 'Easter Island: Polynesia'
    12 = 'Moriori: Chatham Islands, Polynesia'
    13 = 'Arikara: (Early) North America'
    14 = 'Santa Cruz Island: California, N. America'
    15 = 'Yauyos: Peru, South America'
    16 = 'Hokkaido: North Japan'
    17 = 'North Kyushu: South Japan'
    18 = 'Hainan: Haikou City, China'
    19 = 'Atayal: Taiwan Aboriginals'
    20 = 'Phillipine: Phillipine Islands'
    21 = 'Guam: Latte Period'
    22 = 'Egypt: Gizeh, 26th-30th Dynasties'
    23 = 'San: South Africa'
    24 = 'Andaman Islands: Andaman Islands'
    25 = 'Ainu: S. and SE. Kokkaido, Japan'
    26 = 'Buriat: Siberia'
    27 = 'Eskimo: Inugsuk, Greenland'
    28 = 'Anyang: Shang Dynasty, China'
  ;
run;

* import data from csv file;
proc import out=work.skull
     datafile="F:\Dropbox\UNM\teach\ADA2_stat528\assess\ADA2_HW_17_skull.csv"
     dbms=csv replace;
     getnames=yes;
     datarow=2;
run;

data skull;
  set skull;

  * exclude for lots of missing values;
  *   BRR LAR OSR BAR RFA RPA ROA BSA SBA SLA TBA;

  use = 1; * use the observation in the analysis;
           *   since no missing values;
           * if there is a missing value, set use to 0 below;
           * 0s are missing values in dataset;
    * also, name variables by number v1-v71;
  if GOL = 0 then do; GOL = .; use = 0; end;  v1  = GOL;
  if NOL = 0 then do; NOL = .; use = 0; end;  v2  = NOL;
  if BNL = 0 then do; BNL = .; use = 0; end;  v3  = BNL;
  if BBH = 0 then do; BBH = .; use = 0; end;  v4  = BBH;
  if XCB = 0 then do; XCB = .; use = 0; end;  v5  = XCB;
  if XFB = 0 then do; XFB = .; use = 0; end;  v6  = XFB;
  if ZYB = 0 then do; ZYB = .; use = 0; end;  v7  = ZYB;
  if AUB = 0 then do; AUB = .; use = 0; end;  v8  = AUB;
  if WCB = 0 then do; WCB = .; use = 0; end;  v9  = WCB;
  if ASB = 0 then do; ASB = .; use = 0; end;  v10 = ASB;
  if BPL = 0 then do; BPL = .; use = 0; end;  v11 = BPL;
  if NPH = 0 then do; NPH = .; use = 0; end;  v12 = NPH;
  if NLH = 0 then do; NLH = .; use = 0; end;  v13 = NLH;
  if JUB = 0 then do; JUB = .; use = 0; end;  v14 = JUB;
  if NLB = 0 then do; NLB = .; use = 0; end;  v15 = NLB;
  if MAB = 0 then do; MAB = .; use = 0; end;  v16 = MAB;
  if MDH = 0 then do; MDH = .; use = 0; end;  v17 = MDH;
  if MDB = 0 then do; MDB = .; use = 0; end;  v18 = MDB;
  if OBH = 0 then do; OBH = .; use = 0; end;  v19 = OBH;
  if OBB = 0 then do; OBB = .; use = 0; end;  v20 = OBB;
  if DKB = 0 then do; DKB = .; use = 0; end;  v21 = DKB;
  if NDS = 0 then do; NDS = .; use = 0; end;  v22 = NDS;
  if WNB = 0 then do; WNB = .; use = 0; end;  v23 = WNB;
  if SIS = 0 then do; SIS = .; use = 0; end;  v24 = SIS;
  if ZMB = 0 then do; ZMB = .; use = 0; end;  v25 = ZMB;
  if SSS = 0 then do; SSS = .; use = 0; end;  v26 = SSS;
  if FMB = 0 then do; FMB = .; use = 0; end;  v27 = FMB;
  if NAS = 0 then do; NAS = .; use = 0; end;  v28 = NAS;
  if EKB = 0 then do; EKB = .; use = 0; end;  v29 = EKB;
  if DKS = 0 then do; DKS = .; use = 0; end;  v30 = DKS;
  if IML = 0 then do; IML = .; use = 0; end;  v31 = IML;
  if XML = 0 then do; XML = .; use = 0; end;  v32 = XML;
  if MLS = 0 then do; MLS = .; use = 0; end;  v33 = MLS;
  if WMH = 0 then do; WMH = .; use = 0; end;  v34 = WMH;
  if SOS = 0 then do; SOS = .; use = 0; end;  v35 = SOS;
  if GLS = 0 then do; GLS = .; use = 0; end;  v36 = GLS;
  if STB = 0 then do; STB = .; use = 0; end;  v37 = STB;
  if FRC = 0 then do; FRC = .; use = 0; end;  v38 = FRC;
```

0 pts

```
if FRS = 0 then do; FRS = .; use = 0; end;  v39 = FRS;
if FRF = 0 then do; FRF = .; use = 0; end;  v40 = FRF;
if PAC = 0 then do; PAC = .; use = 0; end;  v41 = PAC;
if PAS = 0 then do; PAS = .; use = 0; end;  v42 = PAS;
if PAF = 0 then do; PAF = .; use = 0; end;  v43 = PAF;
if OCC = 0 then do; OCC = .; use = 0; end;  v44 = OCC;
if OCS = 0 then do; OCS = .; use = 0; end;  v45 = OCS;
if OCF = 0 then do; OCF = .; use = 0; end;  v46 = OCF;
if FOL = 0 then do; FOL = .; use = 0; end;  v47 = FOL;
if NAR = 0 then do; NAR = .; use = 0; end;  v48 = NAR;
if SSR = 0 then do; SSR = .; use = 0; end;  v49 = SSR;
if PRR = 0 then do; PRR = .; use = 0; end;  v50 = PRR;
if DKR = 0 then do; DKR = .; use = 0; end;  v51 = DKR;
if ZOR = 0 then do; ZOR = .; use = 0; end;  v52 = ZOR;
if FMR = 0 then do; FMR = .; use = 0; end;  v53 = FMR;
if EKR = 0 then do; EKR = .; use = 0; end;  v54 = EKR;
if ZMR = 0 then do; ZMR = .; use = 0; end;  v55 = ZMR;
if AVR = 0 then do; AVR = .; use = 0; end;  v56 = AVR;
if BRR = 0 then do; BRR = .; use = 0; end;
if VRR = 0 then do; VRR = .; use = 0; end;  v57 = VRR;
if LAR = 0 then do; LAR = .; use = 0; end;
if OSR = 0 then do; OSR = .; use = 0; end;
if BAR = 0 then do; BAR = .; use = 0; end;
if NAA = 0 then do; NAA = .; use = 0; end;  v58 = NAA;
if PRA = 0 then do; PRA = .; use = 0; end;  v59 = PRA;
if BAA = 0 then do; BAA = .; use = 0; end;  v60 = BAA;
if NBA = 0 then do; NBA = .; use = 0; end;  v61 = NBA;
if BBA = 0 then do; BBA = .; use = 0; end;  v62 = BBA;
if BRA = 0 then do; BRA = .; use = 0; end;  v63 = BRA;
if SSA = 0 then do; SSA = .; use = 0; end;  v64 = SSA;
if NFA = 0 then do; NFA = .; use = 0; end;  v65 = NFA;
if DKA = 0 then do; DKA = .; use = 0; end;  v66 = DKA;
if NDA = 0 then do; NDA = .; use = 0; end;  v67 = NDA;
if SIA = 0 then do; SIA = .; use = 0; end;  v68 = SIA;
if FRA = 0 then do; FRA = .; use = 0; end;  v69 = FRA;
if PAA = 0 then do; PAA = .; use = 0; end;  v70 = PAA;
if OCA = 0 then do; OCA = .; use = 0; end;  v71 = OCA;
if RFA = 0 then do; RFA = .; use = 0; end;
if RPA = 0 then do; RPA = .; use = 0; end;
if ROA = 0 then do; ROA = .; use = 0; end;
if BSA = 0 then do; BSA = .; use = 0; end;
if SBA = 0 then do; SBA = .; use = 0; end;
if SLA = 0 then do; SLA = .; use = 0; end;
if TBA = 0 then do; TBA = .; use = 0; end;

* label variables;
label ID         = "ID"
      Sex        = "Sex"
      PopNum     = "Population affiliation"
      Population = "Population name"
      v1         = "GOL Glabello-occipital length"
      v2         = "NOL Nasio-occipital length"
      v3         = "BNL Basion-nasion length"
      v4         = "BBH Basion-bregma height"
      v5         = "XCB Maximum cranial breadth"
      v6         = "XFB Maximum frontal breadth"
      v7         = "ZYB Bizygomatic breadth"
      v8         = "AUB Biauricular breadth"
      v9         = "WCB Minimum cranial breadth"
      v10        = "ASB Biasterionic breadth"
      v11        = "BPL Basion-prosthion length"
      v12        = "NPH Nasion-prosthion height"
      v13        = "NLH Nasal height"
      v14        = "JUB Bijugal breadth"
      v15        = "NLB Nasal breadth"
      v16        = "MAB Palate breadth, external"
      v17        = "MDH Mastoid height"
      v18        = "MDB Mastoid breadth"
      v19        = "OBH Orbit height, left"
      v20        = "OBB Orbit breadth, left"
      v21        = "DKB Interorbital breadth"
      v22        = "NDS Naso-dacryal subtense"
      v23        = "WNB Simotic chord (least nasal breadth)"
      v24        = "SIS Simotic subtense"
      v25        = "ZMB Bimaxillary breadth"
      v26        = "SSS Zygomaxillary subtense"
```

0 pts

```
        v27         = "FMB Bifrontal breadth"
        v28         = "NAS Nasio-frontal subtense"
        v29         = "EKB Biorbital breadth"
        v30         = "DKS Dacryon subtense"
        v31         = "IML Malar length, inferior"
        v32         = "XML Malar length, maximum"
        v33         = "MLS Malar subtense"
        v34         = "WMH Cheek height"
        v35         = "SOS Supraorbital projection"
        v36         = "GLS Glabella projection"
        v37         = "STB Bistephanic breadth"
        v38         = "FRC Nasion-bregma chord (Frontal chord)"
        v39         = "FRS Nasion-bregma subtense (Frontal subtense)"
        v40         = "FRF Nasion-subtense fraction"
        v41         = "PAC Bregma-lambda chord (Parietal chord)"
        v42         = "PAS Bregma-lambda subtense (Parietal subtense)"
        v43         = "PAF Bregma-subtense fraction"
        v44         = "OCC Lambda-opisthion chord (Occipital chord)"
        v45         = "OCS (OCS)"
        v46         = "OCF (OCF)"
        v47         = "FOL Foramen magnum length"
        v48         = "NAR (NAR)"
        v49         = "SSR (SSR)"
        v50         = "PRR (PRR)"
        v51         = "DKR (DKR)"
        v52         = "ZOR (ZOR)"
        v53         = "FMR (FMR)"
        v54         = "EKR (EKR)"
        v55         = "ZMR (ZMR)"
        v56         = "AVR (AVR)"
           BRR      = "BRR (BRR)"
        v57         = "VRR (VRR)"
           LAR      = "LAR (LAR)"
           OSR      = "OSR (OSR)"
           BAR      = "BAR (BAR)"
        v58         = "NAA (NAA)"
        v59         = "PRA (PRA)"
        v60         = "BAA (BAA)"
        v61         = "NBA (NBA)"
        v62         = "BBA (BBA)"
        v63         = "BRA (BRA)"
        v64         = "SSA (SSA)"
        v65         = "NFA (NFA)"
        v66         = "DKA (DKA)"
        v67         = "NDA (NDA)"
        v68         = "SIA (SIA)"
        v69         = "FRA (FRA)"
        v70         = "PAA (PAA)"
        v71         = "OCA (OCA)"
           RFA      = "RFA (RFA)"
           RPA      = "RPA (RPA)"
           ROA      = "ROA (ROA)"
           BSA      = "BSA (BSA)"
           SBA      = "SBA (SBA)"
           SLA      = "SLA (SLA)"
           TBA      = "TBA (TBA)"
        ;

  * assign value format to PopNum;
  format PopNum formatPopNum.;

  * delete observations with missing values;
  if use = 0 then delete;

  * drop original variables from dataset, keeping only v1-v71;
  drop GOL NOL BNL BBH XCB XFB ZYB AUB WCB ASB BPL NPH NLH JUB NLB
       MAB MDH MDB OBH OBB DKB NDS WNB SIS ZMB SSS FMB NAS EKB DKS IML
       XML MLS WMH SOS GLS STB FRC FRS FRF PAC PAS PAF OCC OCS OCF FOL
       NAR SSR PRR DKR ZOR FMR EKR ZMR AVR BRR VRR LAR OSR BAR NAA PRA
       BAA NBA BBA BRA SSA NFA DKA NDA SIA FRA PAA OCA RFA RPA ROA BSA
       SBA SLA TBA;
run;

* If successful, these notes appear in the log window                    ;
*   NOTE: There were 2524 observations read from the data set WORK.SKULL. ;
*   NOTE: The data set WORK.SKULL has 1855 observations and 76 variables. ;

* At last, the dataset is ready for analysis! ;
```

0 pts