

Statistical Ecology
Stat 555

Principle Components Analysis

Problem: Size and shape of female painted turtles, *Chrysemys picta marginata*.
Length, width and height were measured in 24 females.

MTB > Covariance 'L' 'W' 'H'.

Covariances: L, W, H

	L	W	H
L	451.3913		
W	272.3913	173.0272	
H	168.6957	103.8967	66.6504

$$\sum S^2 = 451 + 173 + 67 \approx 690$$

MTB > PCA 'L' 'W' 'H';
SUBC> NComponents 3;
SUBC> Covariance;
SUBC> Scores 'PC1' 'PC2' 'PC3';
SUBC> GScree;
SUBC> GScore;
SUBC> GLoading.

$$\sum \lambda = 682 + 6 + 3 \approx 690$$

Principal Component Analysis: L, W, H

Eigenanalysis of the Covariance Matrix

Eigenvalue	681.90	6.40	2.77
Proportion	0.987	0.009	0.004
Cumulative	0.987	0.996	1.000

$$\% \lambda_1 = \frac{681.9}{690} \times 100 = 98.7\%$$

Variable	PC1	PC2	PC3
L	-0.812	0.555	0.181
W	-0.497	-0.820	0.284
H	-0.307	-0.140	-0.941

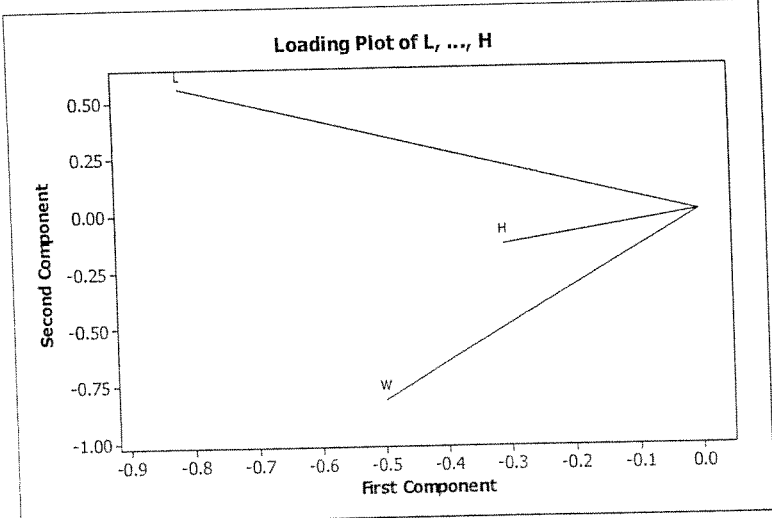
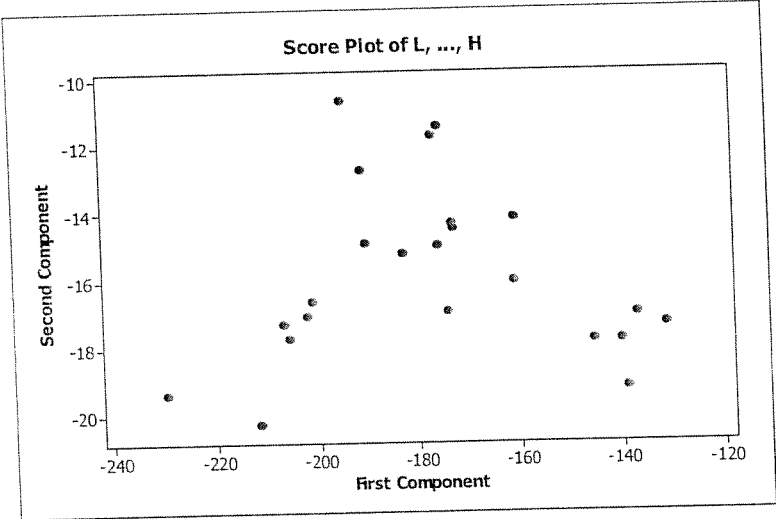
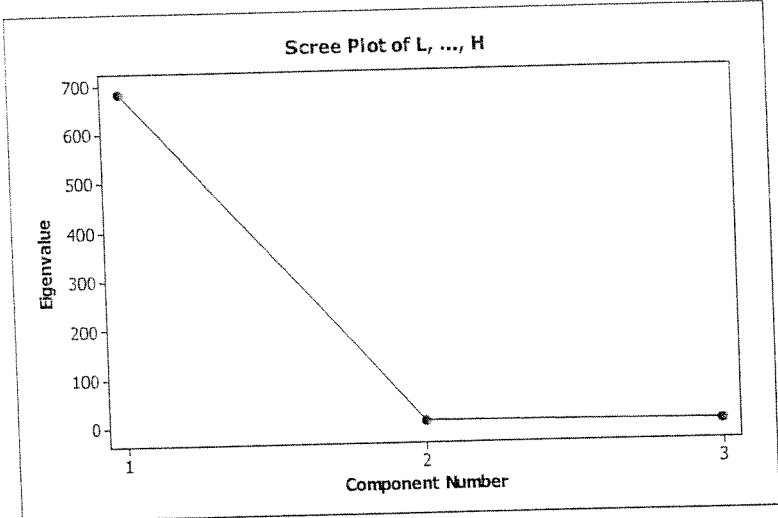
Data:

Component scores:

L	W	H	PC1	PC2	PC3
98	81	38	-131.464	-17.3455	5.02721
103	84	38	-137.014	-17.0295	6.78709
103	86	42	-139.235	-19.2311	3.58965
105	86	40	-140.245	-17.8399	5.83548
109	88	44	-145.712	-17.8211	3.36394
123	92	50	-160.904	-14.1719	1.39257
123	95	46	-161.170	-16.0696	6.01085
133	99	51	-172.808	-14.5003	4.25504
133	99	51	-174.299	-16.9599	5.10755
133	102	51	-174.299	-16.9599	5.10755

$$X_{11} = a_{11} * L + a_{12} * W + a_{13} * H$$

$$= (-0.812)(98) + (-0.497)(81) + (-0.307)(38) = -131.5$$



PROC PRINCOMP: Options

PROC PRINCOMP *< options >* ;

The PROC PRINCOMP statement starts the PRINCOMP procedure and, optionally, identifies input and output data sets, specifies details of the analysis, or suppresses the display of output.

The following list provides details on these options.

<u>COVARIANCE</u>	<u>DATA=SAS-data-set</u>
<u>N=number</u>	<u>NOINT</u>
<u>NOPRINT</u>	<u>OUT=SAS-data-set</u>
<u>OUTSTAT=SAS-data-set</u>	<u>PREFIX=name</u>
<u>SINGULAR=p</u>	<u>STANDARD</u>
<u>VARDEF=DF N WDF WEIGHT WGT</u>	

Turtles.sas

```

data Turtles;
  title 'Male and Female Painted Turtles: Data after Joliceur & Mosimann. 1960.
  Growth 24: 339-354';
  input Sex $1-8 group Length Width Height;
  datalines;
FEMALE01    1          98.0000          81.0000          38.0000
FEMALE02    1          103.0000          84.0000          38.0000
FEMALE03    1          103.0000          86.0000          42.0000
. . . .
MALES 01    2           93.0000          74.0000          37.0000
MALES 02    2           94.0000          78.0000          35.0000
MALES 03    2           96.0000          80.0000          35.0000
. . . .
;
title2 'Plot of Raw Data';
%plotit(data=Turtles, labelvar=Sex,
        plotvars=Length Width, color=black, colors=blue);
run;

proc princomp data=Turtles std out=Prin;
title2;
var Length Width Height;
run;

title2 'Plot of Principal Components';
%plotit(data=Prin, labelvar=Sex,
        plotvars=Prin2 Prin1, color=black, colors=blue);
run;

```

The PRINCOMP Procedure

Observations 48
 Variables 3

Simple Statistics

	Length	Width	Height
Mean	124.6875000	95.45833333	46.33333333
StD	20.4816018	12.71279860	8.39283720

Correlation Matrix

	Length	Width	Height
Length	1.0000	0.9787	0.9647
Width	0.9787	1.0000	0.9617
Height	0.9647	0.9617	1.0000

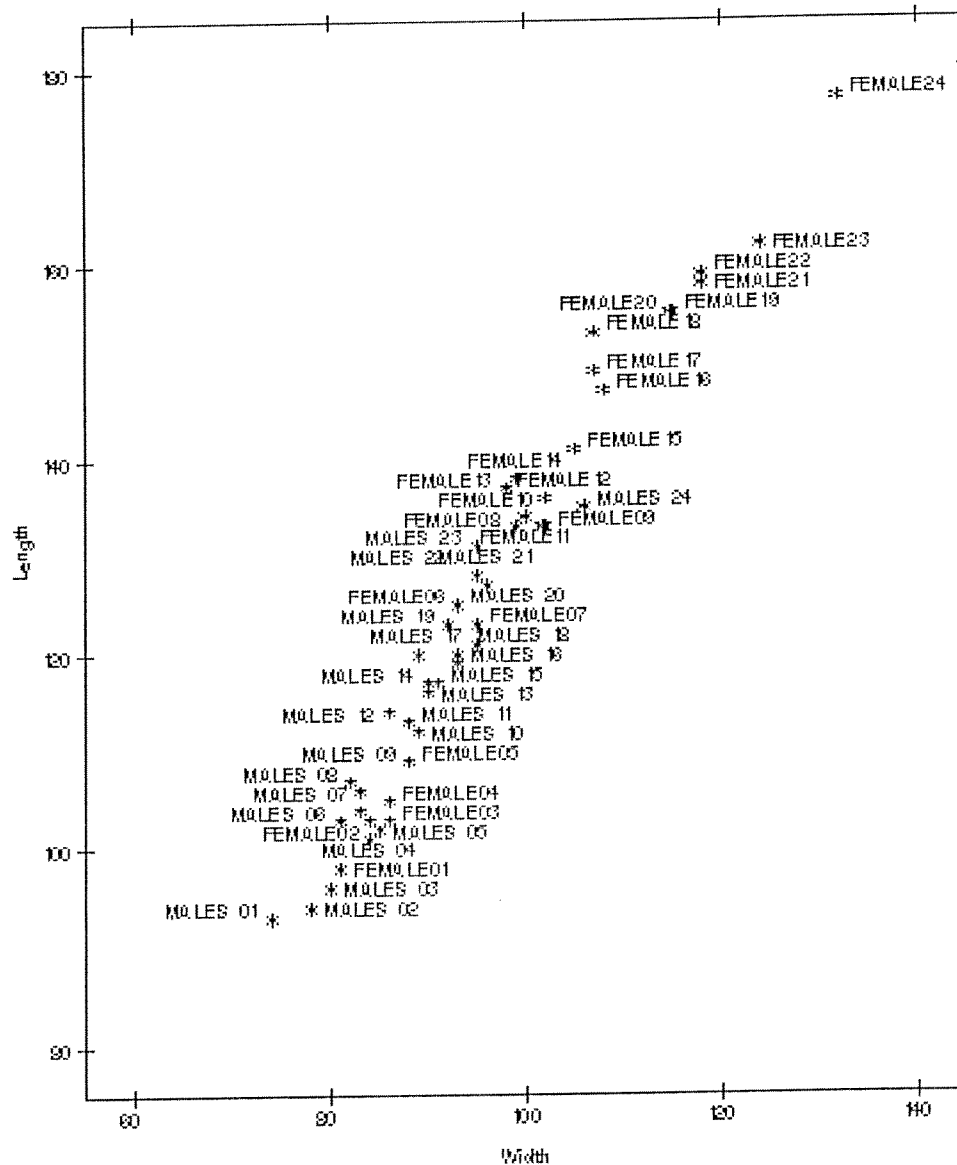
Eigenvalues of the Correlation Matrix

	Eigenvalue	Difference	Proportion	Cumulative
1	2.93673479	2.89464490	0.9789	0.9789
2	0.04208988	0.02091455	0.0140	0.9929
3	0.02117533		0.0071	1.0000

Eigenvectors

	Prin1	Prin2	Prin3
Length	0.578673	-.346635	-.738229
Width	0.578082	-.464173	0.671092
Height	0.575290	0.815100	0.068221

Male and Female Painted Turtles: Data after Jolicoeur & Mosimann, 1960. Growth 24: 333-354
 Plot of Raw Data



Male and Female Painted Turtles: Data after Jolicœur & Mosmann, 1960, Growth 24: 339-354
 Plot of Principal Components

