

```

/*****/
/* To run clustering on a mammal teeth example using PROC
   CLUSTER */
/*****/

options nocenter nodate pageno=1 linesize=132;
title h = 1 j = 1 'File: cluster.mammalsteeth.sas';
title2 h = 1 j = 1 'Cluster Analysis of Mammals'' teeth data';
data teeth;
  input mammal $ 1-16
        @21 (v1-v8) (1.);
  label v1='Top incisors'
        v2='Bottom incisors'
        v3='Top canines'
        v4='Bottom canines'
        v5='Top premolars'
        v6='Bottom premolars'
        v7='Top molars'
        v8='Bottom molars';
  cards;
BROWN BAT          23113333
MOLE                32103333
SILVER HAIR BAT   23112333
PIGMY BAT          23112233
HOUSE BAT          23111233
RED BAT            13112233
PIKA               21002233
RABBIT             21003233
BEAVER             11002133
GROUNDHOG          11002133
GRAY SQUIRREL     11001133
HOUSE MOUSE       11000033
PORCUPINE         11001133
WOLF              33114423
BEAR              33114423
RACCOON           33114432
MARTEN            33114412
WEASEL           33113312
WOLVERINE        33114412
BADGER           33113312
RIVER OTTER      33114312
SEA OTTER        32113312
JAGUAR           33113211
COUGAR           33113211
FUR SEAL        32114411
SEA LION        32114411
GREY SEAL       32113322
ELEPHANT SEAL   21114411
REINDEER        04103333
ELK              04103333
DEER             04003333
MOOSE            04003333
;

/* principal components analysis of teeth here we score the
   principal components and output then to data set teeth2 */

```

```

proc princomp data=teeth out=teeth2;
    var v1-v8;
run;

/* average linkage cluster analysis a dendrogram (tree diagram)
is also output */

proc cluster data=teeth2 method=average outtree=ttree
    ccc pseudo rsquare;
    var v1-v8;
    id mammal;
run;

/* --- PROC TREE prints the tree diagram here we also output a
data set, called ttree2 that contains four clusters --- */

proc tree data=ttree out=ttree2 nclusters=4;
    id mammal;
run;

/* --- the next set of statements sort the data sets by variable
mammal and then merge the tree data set (with the cluster scores)
with the teeth data set (with the principal components) --- */

proc sort data=teeth2;
    by mammal;
run;
proc sort data=ttree2;
    by mammal;
run;
data teeth3;
    merge teeth2 ttree2;
    by mammal;
run;

/* --- stuff for plotting --- */

symbol1 c=black f=, v='1';
symbol2 c=black f=, v='2';
symbol3 c=black f=, v='3';
symbol4 c=black f=, v='4';
goptions device=pdfc gsfname=grafout;
    filename grafout 'teeth.pdf';
proc gplot;
    plot prin2*prin1=cluster;
run;

proc sort;
    by cluster;
run;
proc print;
    by cluster;
    var mammal prin1 prin2;
run;

```


File: cluster.mammalsteeth.sas
 2
 Cluster Analysis of Mammals' teeth data

The CLUSTER Procedure
 Average Linkage Cluster Analysis

Eigenvalues of the Covariance Matrix

	Eigenvalue	Difference	Proportion	Cumulative
1	3.76799365	2.33557185	0.5840	0.5840
2	1.43242180	0.91781899	0.2220	0.8061
3	0.51460281	0.08414950	0.0798	0.8858
4	0.43045331	0.30021485	0.0667	0.9525
5	0.13023846	0.03814626	0.0202	0.9727
6	0.09209220	0.04216914	0.0143	0.9870
7	0.04992305	0.01603541	0.0077	0.9947
8	0.03388764		0.0053	1.0000

Root-Mean-Square Total-Sample Standard Deviation = 0.898027
 Root-Mean-Square Distance Between Observations = 3.592106

Cluster History

NCL	-----Clusters Joined-----			FREQ	SPRSQ	RSQ	ERSQ	CCC	PSF	PST2	Norm RMS Dist	T i e
31	BEAVER	GROUNDHOG		2	0.0000	1.00	0	T
30	GRAY SQUIRREL	PORCUPINE		2	0.0000	1.00	0	T
29	WOLF	BEAR		2	0.0000	1.00	0	T
28	MARTEN	WOLVERINE		2	0.0000	1.00	0	T
27	WEASEL	BADGER		2	0.0000	1.00	0	T
26	JAGUAR	COUGAR		2	0.0000	1.00	0	T
25	FUR SEAL	SEA LION		2	0.0000	1.00	0	T
24	REINDEER	ELK		2	0.0000	1.00	0	T
23	DEER	MOOSE		2	0.0000	1.00	0	
22	BROWN BAT	SILVER HAIR BAT		2	0.0025	.998	.	.	190	.	0.2784	T
21	PIGMY BAT	HOUSE BAT		2	0.0025	.995	.	.	109	.	0.2784	T
20	PIKA	RABBIT		2	0.0025	.993	.	.	83.6	.	0.2784	T
19	CL31	CL30		4	0.0050	.988	.	.	57.1	.	0.2784	T
18	CL28	RIVER OTTER		3	0.0033	.984	.	.	51.2	.	0.2784	T
17	CL27	SEA OTTER		3	0.0033	.981	.	.	48.0	.	0.2784	T
16	CL24	CL23		4	0.0050	.976	.	.	43.1	.	0.2784	
15	CL21	RED BAT		3	0.0042	.972	.	.	41.6	1.7	0.341	
14	CL17	GREY SEAL		4	0.0054	.966	.	.	39.6	3.3	0.3594	
13	CL29	RACCOON		3	0.0067	.960	.	.	37.6	.	0.3937	T
12	CL25	ELEPHANT SEAL		3	0.0067	.953	.	.	36.8	.	0.3937	
11	CL18	CL14		7	0.0151	.938	.	.	31.7	6.2	0.4328	
10	CL22	CL15		5	0.0108	.927	.	.	31.1	3.5	0.4402	
9	CL20	CL19		6	0.0200	.907	.	.	28.0	10.7	0.5208	T
8	CL11	CL26		9	0.0229	.884	.	.	26.2	5.9	0.5365	
7	CL8	CL12		12	0.0383	.846	.	.	22.9	6.8	0.6202	
6	MOLE	CL13		4	0.0171	.829	.779	2.77	25.2	5.1	0.6225	
5	CL9	HOUSE MOUSE		7	0.0225	.806	.741	2.86	28.1	4.1	0.6913	
4	CL6	CL7		16	0.0572	.749	.691	1.83	27.9	6.7	0.7143	
3	CL10	CL16		9	0.0761	.673	.614	1.67	29.8	21.3	0.7825	
2	CL3	CL5		16	0.1980	.475	.465	0.20	27.1	18.3	1.0315	
1	CL2	CL4		32	0.4750	.000	.000	0.00	.	27.1	1.1954	

File: cluster.mammalsteeth.sas

3

Cluster Analysis of Mammals' teeth data

CLUSTER=1

Obs	mammal	Prin1	Prin2
1	BEAVER	-3.11568	-1.05493
2	GRAY SQUIRREL	-3.46713	-1.23651
3	GROUNDHOG	-3.11568	-1.05493
4	HOUSE MOUSE	-4.15867	-1.71452
5	PIKA	-2.44624	-1.06101
6	PORCUPINE	-3.46713	-1.23651
7	RABBIT	-2.09478	-0.87943

CLUSTER=2

Obs	mammal	Prin1	Prin2
8	BADGER	1.80728	-0.33141
9	BEAR	1.63582	0.95908
10	COUGAR	1.92490	-1.11384
11	ELEPHANT SEAL	2.29581	-1.39702
12	FUR SEAL	2.79084	-1.01947
13	GREY SEAL	1.23632	-0.68499
14	JAGUAR	1.92490	-1.11384
15	MARTEN	2.49881	0.14660
16	MOLE	-0.42183	0.22803
17	RACCOON	1.68822	0.79958
18	RIVER OTTER	2.15874	-0.14983
19	SEA LION	2.79084	-1.01947
20	SEA OTTER	1.64162	-1.01148
21	WEASEL	1.80728	-0.33141
22	WOLF	1.63582	0.95908
23	WOLVERINE	2.49881	0.14660

CLUSTER=3

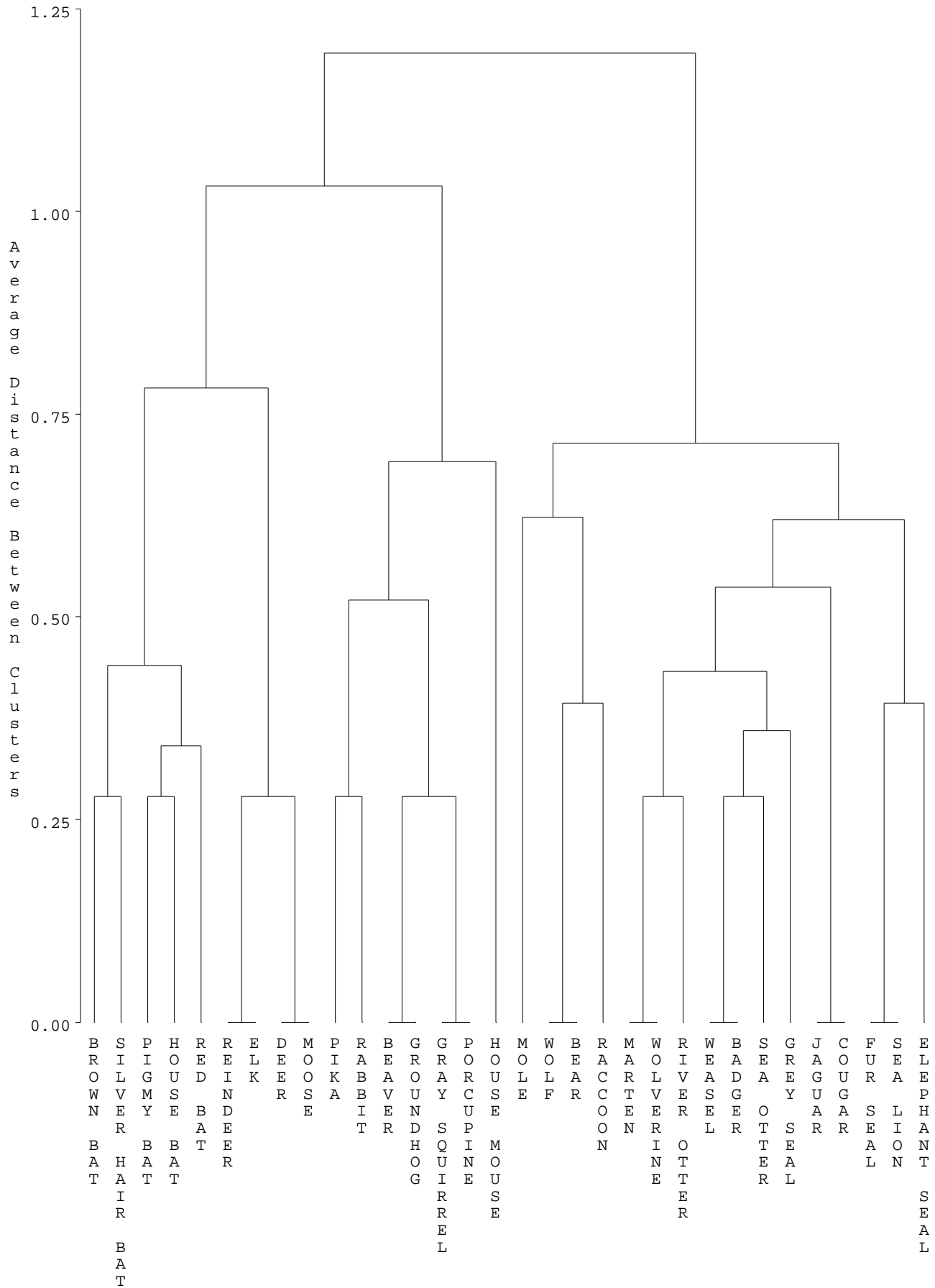
Obs	mammal	Prin1	Prin2
24	DEER	-1.91645	2.06224
25	ELK	-1.07861	2.49571
26	MOOSE	-1.91645	2.06224
27	REINDEER	-1.07861	2.49571

CLUSTER=4

Obs	mammal	Prin1	Prin2
28	BROWN BAT	0.20963	1.11008
29	HOUSE BAT	-0.83336	0.45049
30	PIGMY BAT	-0.48191	0.63206
31	RED BAT	-0.81127	0.93458
32	SILVER HAIR BAT	-0.14183	0.92850

File: cluster.mammalsteeth.sas

Cluster Analysis of Mammals' teeth data



mammal

File: cluster.mammalsteeth.sas

Cluster Analysis of Mammals' teeth data

