Comparison between Principal component analysis and Factor analysis

Cattle morphology data

Cattle morphology data

- ✓ Search for a reduced set of variables for representing the covariance structure among morphologial measurements

- ✓ Breed differentiation

- ✓ Biodiversity

- ✓ N = 300 cattle

- ✓ p = 18 variables

Cattle morphology data

```
data=read.table(»……….. /cow_measurements.csv", sep=";", dec=",", header=TRUE)

cow=data[ , 2:19]

#calcolo della matrice di correlazione
R=cor(cow)

library(psych)
KMO(R)
```

```
 Kaiser-Meyer-Olkin factor adequacy
 Call: KMO(r = R)
 Overall MSA =  0.96
 MSA for each item =
   HG   HD  HLL   HC  HTR LNTN LNLL LNTR LNGR LGTR  LGP LGGA LGGM
 LGGP LGLL   Ct   cn   cs
 0.95 0.95 0.95 0.97 0.98 0.96 0.97 0.97 0.96 0.93 0.88 0.98 0.98 0.90 0.98 0.98
 0.98 0.98
```

# *Cattle morphology data*

eigenvalues=eigen(R)$values

eigenvalues

var_tot=sum(diag(R))

proportion=eigenvalues/var_tot

proportion

```
[1]  12.97478003  2.18218208  0.70041575  0.37587010  0.29972293  0.25827398
[7]   0.20705361  0.16945364  0.15099700  0.12807803  0.11398681  0.09974803
[13]  0.09733661  0.07511081  0.05852570  0.05128265  0.03514135  0.02204090

[1]  0.720821113 0.121232338 0.038911986 0.020881672 0.016651274 0.014348555
[7]  0.011502978 0.009414091 0.008388722 0.007115446 0.006332601 0.005541557
[13] 0.005407589 0.004172823 0.003251428 0.002849036 0.001952297 0.001224494
```

Eigenvalue plot
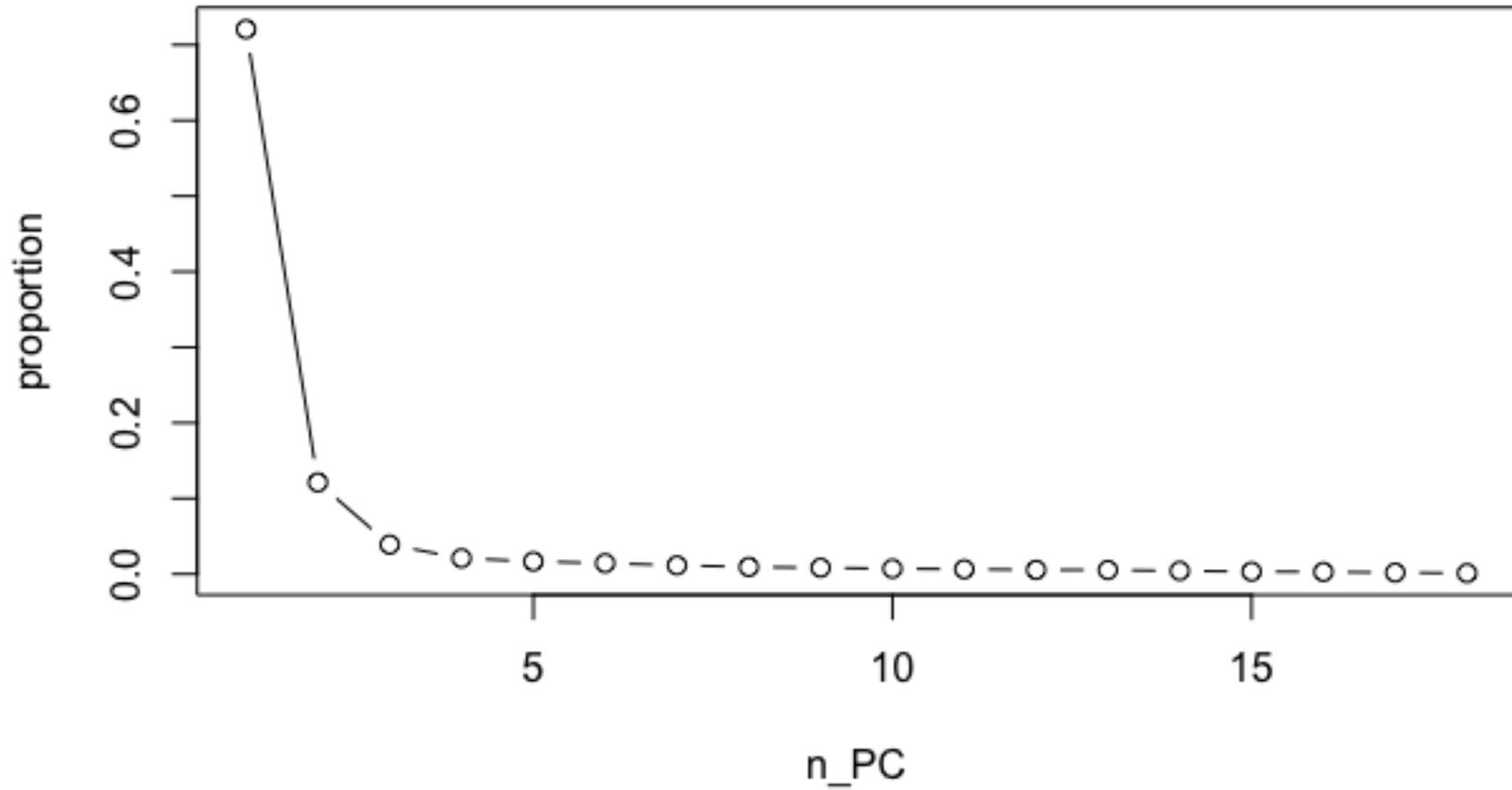
Cattle morphology data
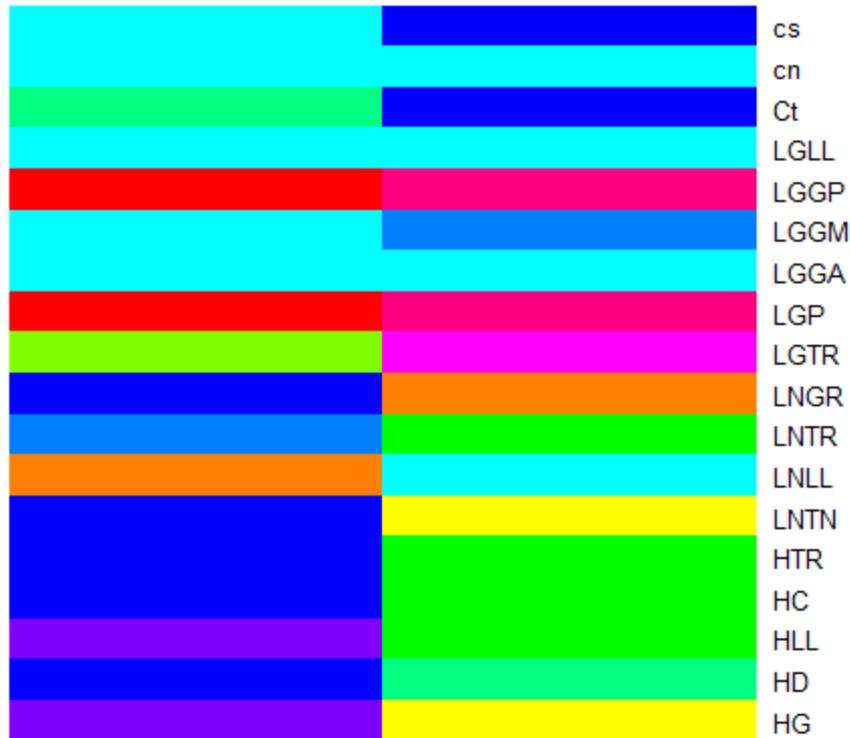
F1   F2

G=data.matrix(factor$loadings)
**heatmap(G,Rowv=NA, Colv=NA, col=rainbow(12),scale="column")**

Score calculation

```
factor=factanal(cow, factors=2, rotation="varimax", scores="regression")
scores=data.frame(factor$scores)


breed=data[ ,1]
graph=cbind(breed,scores)


plot(graph[ ,2],graph[ ,3], xlab="Factor 1", ylab="Factor 2",
col=c("red","green","blue")[unclass(graph[ ,1])],cex=1, pch=19, main="red=Sardo-
Bruna,green=Sardo-modicana,blue=Sarda")
```

Factor score plots
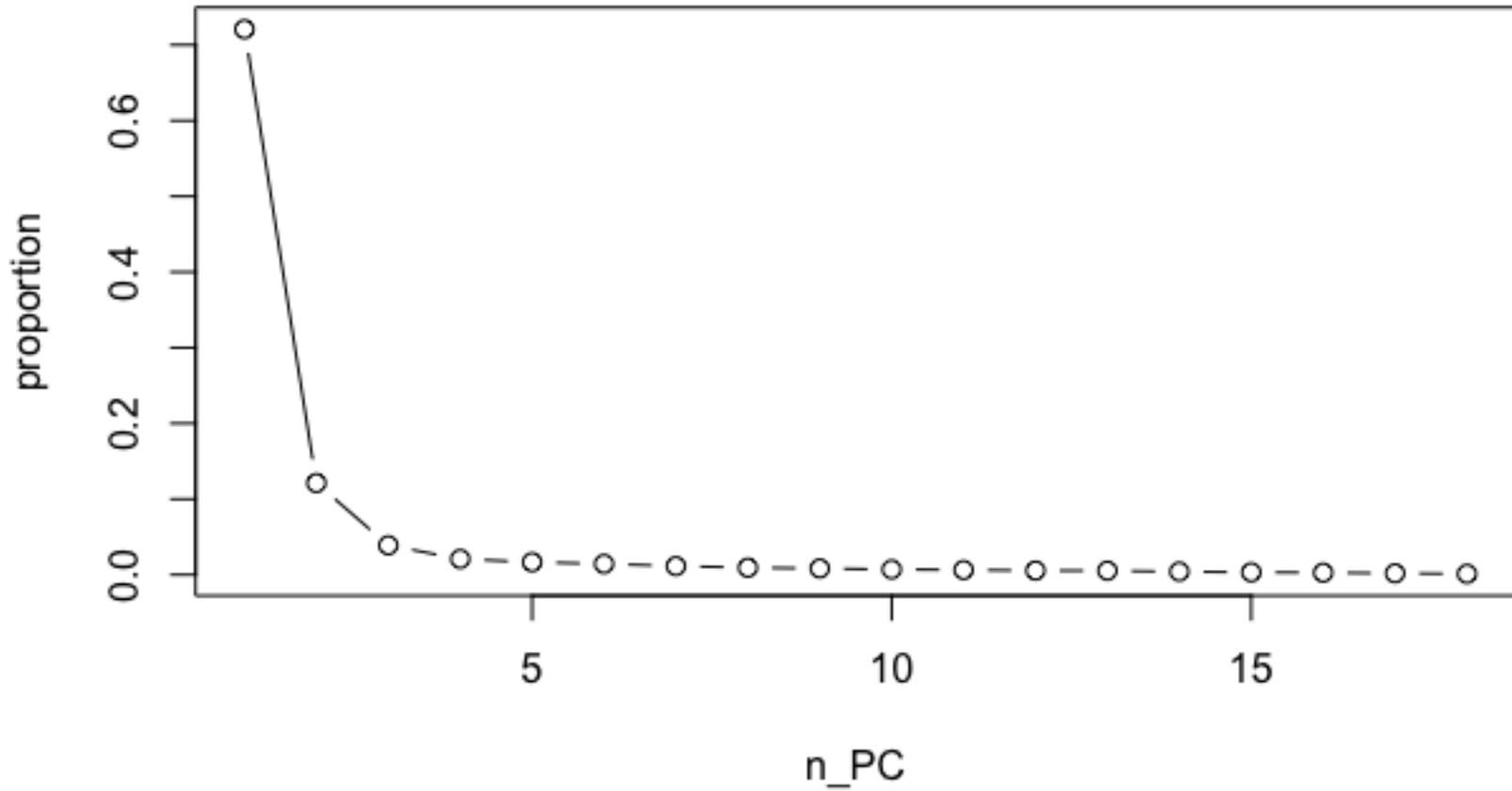
Comparison between PCA and Factor analysis

	Factor1	Factor2
HG	0.956	0.208
HD	0.860	0.456
HLL	0.916	0.330
HC	0.872	0.372
HTR	0.874	0.319
LNTN	0.870	0.192
LNLL	0.237	0.502
LNTR	0.808	0.350
LNGR	0.860	
LGTR	0.383	0.869
LGP	0.154	0.961
LGGA	0.707	0.554
LGGM	0.723	0.576
LGGP	0.121	0.947
LGLL	0.723	0.493
Ct	0.620	0.653
cn	0.658	0.493
cs	0.643	0.660
%var	0.51	0.31

	Comp.1	Comp.2
HG	-0.248	-0.244
HD	-0.264	-0.104
HLL	-0.257	-0.158
HC	-0.255	-0.122
HTR	-0.253	-0.178
LNTN	-0.235	-0.259
LNLL	-0.137	0.273
LNTR	-0.245	-0.135
LNGR	-0.215	-0.328
LGTR	-0.227	0.329
LGP	-0.186	0.467
LGGA	-0.256	0.141
LGGM	-0.259	0.129
LGGP	-0.178	0.491
LGLL	-0.248	0.101
Ct	-0.249	0.118
cn	-0.236	0.179
cs	-0.253	0.122
% var	0.73	0.12

PC score plots

rosso=Sardo-Bruna,verde=Sardo-modicana,blu=Sarda
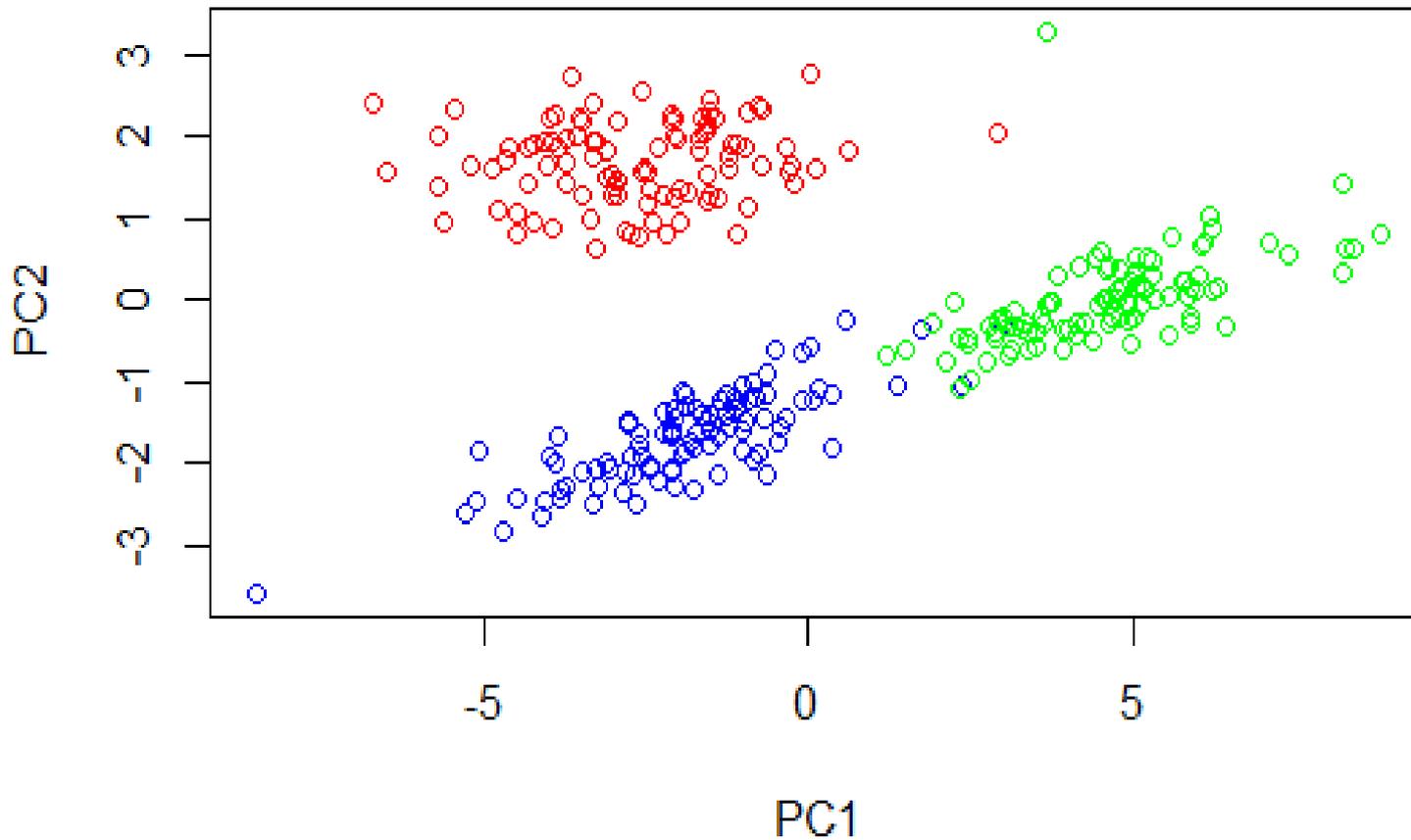
Comparison between PCA and Factor analysis for lactation curve traits

	PC1	PC2		Factor1	Factor2
Latte1	−0.28	0.49	latte1	0.72	0.12
Latte2	−0.35	0.39	latte2	0.88	0.22
Latte3	−0.38	0.24	latte3	0.85	0.37
Latte4	−0.38	0.12	latte4	0.76	0.48
Latte5	−0.39	−0.03	latte5	0.68	0.62
Latte6	−0.38	−0.22	latte6	0.51	0.77
Latte7	−0.34	−0.42	latte7	0.29	0.87
Latte8	−0.30	−0.55	latte8	0.13	0.88
%var	0.69	0.16	%var	0.43	0.37