Análisis de Componentes Principales

HectorDelgado

2024-10-29

```
library(readxl)
library("Hmisc")
```

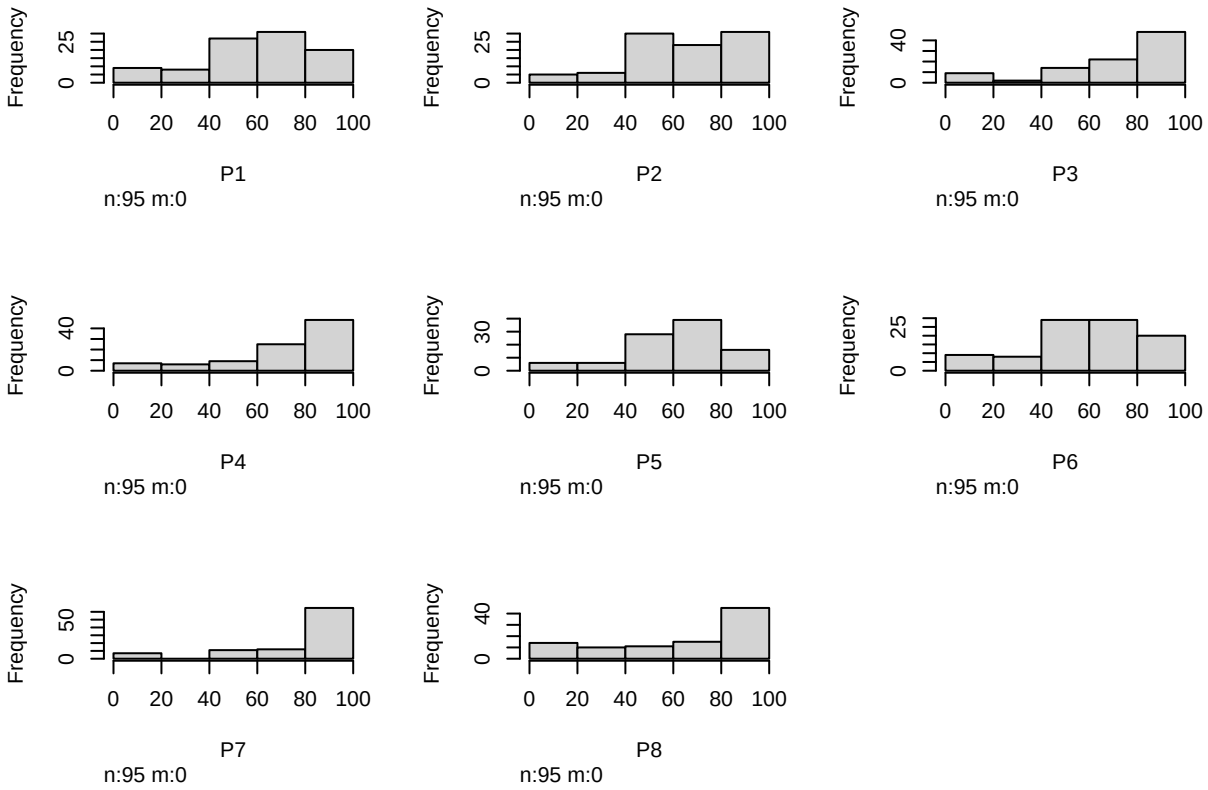
```
##
## Adjuntando el paquete: 'Hmisc'
```

```
## The following objects are masked from 'package:base':
##
##     format.pval, units
```

```
A<-read_excel("C:\\Users\\Man\\Downloads\\SegundaVuelta\\Datos_PD.xlsx")
Datos<-A[ , c("P1", "P2", "P3", "P4", "P5", "P6", "P7","P8")]
Datos
```

```
## # A tibble: 95 x 8
##       P1    P2    P3    P4    P5    P6    P7    P8
##    <dbl> <dbl> <dbl> <dbl> <dbl> <dbl> <dbl> <dbl>
##  1    50    20    95    80    60    60   100    10
##  2    17    27    66    61    58    59    97    29
##  3    70    83    88    74    49    69   100    96
##  4    20    30    90    90    80    50   100    10
##  5    70    80    80    90    75    75    70    95
##  6    77    74    90    88    56    61    92   100
##  7    70    70    60    80    60    60    80   100
##  8    50    50    70    70    70    60    60    70
##  9    78    85    94    72    78    85    82    86
## 10    60    50    90    90    45    65    85    39
## # i 85 more rows
```

```
hist(Datos)
```



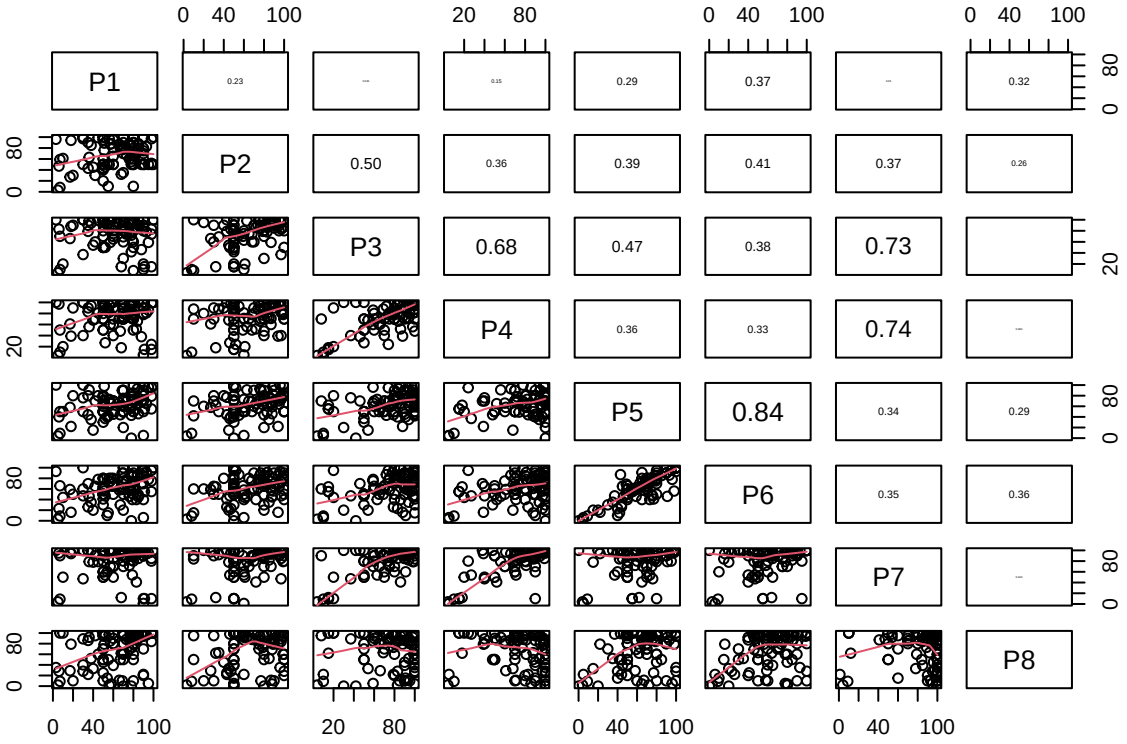
##Gráfico de correlación.

```
panel.cor <- function(x, y, digits=2, prefix="", cex.cor, ...)
{
  usr <- par("usr"); on.exit(par(usr))
  par(usr = c(0, 1, 0, 1))
  r<-abs(cor(x, y))
  txt <- format(c(r, 0.123456789), digits=digits)[1]
  txt <- paste(prefix, txt, sep="")
  if(missing(cex.cor)) cex.cor <- 8/strwidth(txt)
  text(0.5, 0.5, txt, cex = cex.cor**0.15*r)

}
pairs(Datos, lower.panel=panel.smooth, upper.panel=panel.cor, cex.labels =1.3)
```

```
## Warning in par(usr): argument 1 does not name a graphical parameter
## Warning in par(usr): argument 1 does not name a graphical parameter
## Warning in par(usr): argument 1 does not name a graphical parameter
## Warning in par(usr): argument 1 does not name a graphical parameter
## Warning in par(usr): argument 1 does not name a graphical parameter
## Warning in par(usr): argument 1 does not name a graphical parameter
## Warning in par(usr): argument 1 does not name a graphical parameter
## Warning in par(usr): argument 1 does not name a graphical parameter
## Warning in par(usr): argument 1 does not name a graphical parameter
## Warning in par(usr): argument 1 does not name a graphical parameter
## Warning in par(usr): argument 1 does not name a graphical parameter
## Warning in par(usr): argument 1 does not name a graphical parameter
```

```
## Warning in par(usr): argument 1 does not name a graphical parameter
## Warning in par(usr): argument 1 does not name a graphical parameter
## Warning in par(usr): argument 1 does not name a graphical parameter
## Warning in par(usr): argument 1 does not name a graphical parameter
## Warning in par(usr): argument 1 does not name a graphical parameter
## Warning in par(usr): argument 1 does not name a graphical parameter
## Warning in par(usr): argument 1 does not name a graphical parameter
## Warning in par(usr): argument 1 does not name a graphical parameter
## Warning in par(usr): argument 1 does not name a graphical parameter
## Warning in par(usr): argument 1 does not name a graphical parameter
## Warning in par(usr): argument 1 does not name a graphical parameter
## Warning in par(usr): argument 1 does not name a graphical parameter
## Warning in par(usr): argument 1 does not name a graphical parameter
## Warning in par(usr): argument 1 does not name a graphical parameter
## Warning in par(usr): argument 1 does not name a graphical parameter
## Warning in par(usr): argument 1 does not name a graphical parameter
```




```
par(fig=c(0,1,0,1))
```

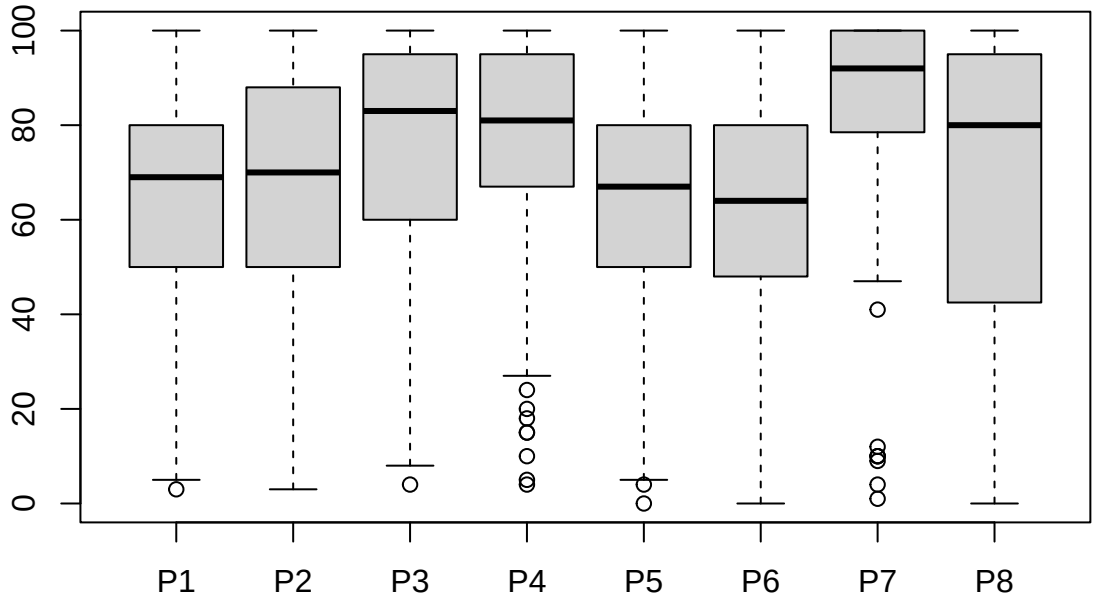
##Se calcula la matriz de varianzas, se extraen los valores de la diagonal y se calculan sus raíces cuadradas para obtener la desviación estándar.

```
V<-var(Datos)
Vdiag<-diag(V, names = FALSE)
desv<-sqrt(Vdiag)
desv
```

```
## [1] 24.61744 24.50005 26.44476 25.70875 22.90850 24.73914 25.93372 33.26191
```

##Para encontrar datos atípicos, se genera un diagrama de caja.

```
boxplot(Datos)
```



##Para cerciorarse de la aplicación del ACP se calcula el índice KMO.

```
library(psych)
```

```
##
## Adjuntando el paquete: 'psych'
```

```
## The following object is masked from 'package:Hmisc':
##
##     describe
```

```
KMO(Datos)
```

```
## Kaiser-Meyer-Olkin factor adequacy
## Call: KMO(r = Datos)
## Overall MSA =  0.74
## MSA for each item =
##   P1   P2   P3   P4   P5   P6   P7   P8
## 0.73 0.85 0.77 0.79 0.68 0.68 0.76 0.75
```

##El siguiente paso es la implementación del ACP para la matriz de correlación.

```
Datos_acp<-prcomp(Datos, center = TRUE, scale = TRUE)
summary(Datos_acp)
```

```
## Importance of components:
##                           PC1    PC2     PC3     PC4     PC5     PC6     PC7
```

```
## Standard deviation     1.8879 1.3051 0.87978 0.86317 0.74216 0.52455 0.50103
## Proportion of Variance 0.4455 0.2129 0.09675 0.09313 0.06885 0.03439 0.03138
## Cumulative Proportion  0.4455 0.6584 0.75517 0.84831 0.91716 0.95155 0.98293
##                            PC8
## Standard deviation     0.36953
## Proportion of Variance 0.01707
## Cumulative Proportion  1.00000
```

```
round(Datos_acp$sdev^2,3)
```

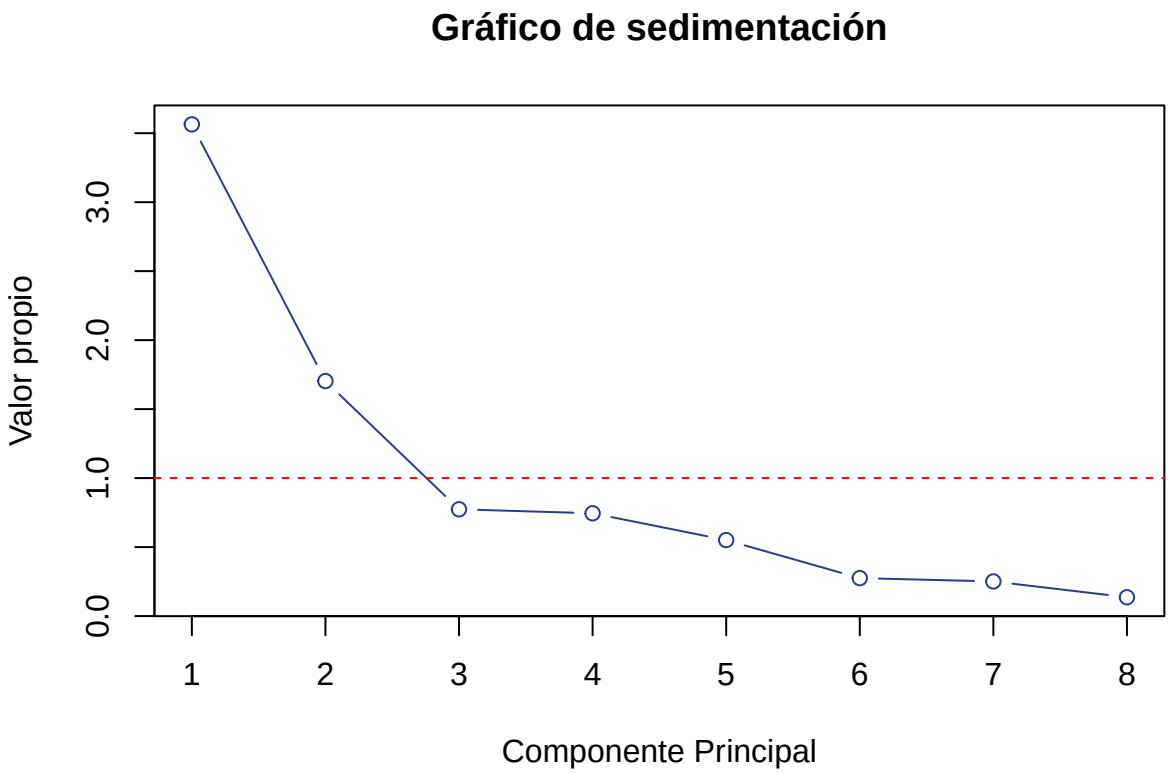
```
## [1] 3.564 1.703 0.774 0.745 0.551 0.275 0.251 0.137
```

```
round(Datos_acp$rotation,4)
```

```
##       PC1     PC2     PC3     PC4     PC5     PC6     PC7     PC8
## P1 0.1925  0.4202  0.5698 -0.6304 -0.0955  0.2196  0.0734  0.0397
## P2 0.3524  0.0665  0.3821  0.5244 -0.6286 -0.2198  0.0316  0.0770
## P3 0.4246 -0.3033  0.0245  0.1669  0.0491  0.7578 -0.1724 -0.3048
## P4 0.3902 -0.3542  0.1679 -0.2379  0.2177 -0.4809 -0.5974 -0.0433
## P5 0.4079  0.2518 -0.5241 -0.0976 -0.1224  0.1347 -0.1911  0.6452
## P6 0.4036  0.3206 -0.4241 -0.1356 -0.0887 -0.2548  0.2277 -0.6414
## P7 0.3944 -0.3768  0.0924 -0.0369  0.3163 -0.1189  0.7135  0.2634
## P8 0.1462  0.5417  0.1935  0.4624  0.6504 -0.0281 -0.0976  0.0235
```

##Se genera el gráfico de sedimentación.

```
plot(Datos_acp$sdev^2, main = "Gráfico de sedimentación",
ylab = "Valor propio", xlab = "Componente Principal",
type = "b", col = "royalblue4")
abline(h = 1, lty = 2, col = "red")
```



##Se calculan los scores y se genera el gráfico bi-plot

```
Z1 <-round(scale(as.matrix(Datos)) %*%
Datos_acp$rotation[,1],4)
Z2 <-round(scale(as.matrix(Datos)) %*%
Datos_acp$rotation[,2],4)
scores <-cbind(Z1, Z2)
colnames(scores) <- c("Z1","Z2")
scores
```

```
##            Z1      Z2
##  [1,] -0.4229 -1.8994
##  [2,] -1.3482 -1.5311
##  [3,]  0.7652  0.1721
##  [4,] -0.2491 -2.3744
##  [5,]  0.9367  0.8185
##  [6,]  0.8251  0.2060
##  [7,] -0.0178  0.7354
##  [8,] -0.7110  0.2748
##  [9,]  1.3822  0.8976
## [10,] -0.1278 -1.1376
## [11,] -0.3963  0.4884
## [12,]  2.5439  1.1262
## [13,]  1.6698 -0.1837
## [14,]  1.3040 -0.2575
## [15,] -2.0950  0.2499
## [16,]  1.6109  0.1163
## [17,]  0.1347  0.0745
```

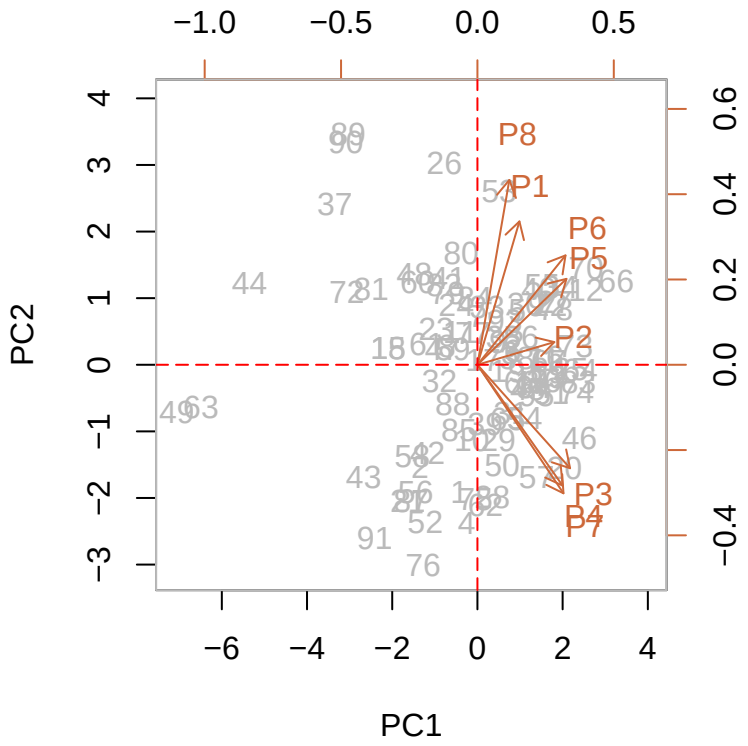
```
## [18,] -2.0950  0.2499
## [19,]  0.7102 -0.0995
## [20,]  2.0472 -1.5537
## [21,] -1.6307 -2.0467
## [22,]  1.6984  0.9027
## [23,] -0.9688  0.5390
## [24,] -0.4937  0.8959
## [25,]  2.1897 -0.0705
## [26,] -0.7769  3.0339
## [27,]  1.8099  0.9391
## [28,]  0.3504 -1.9849
## [29,]  0.4871 -1.1285
## [30,]  1.1163  0.9192
## [31,]  0.7941 -0.7320
## [32,] -0.8915 -0.2486
## [33,]  0.2295  0.8684
## [34,]  1.9477  1.1681
## [35,]  0.6109  0.4142
## [36,]  1.5598 -0.2474
## [37,] -3.3439  2.4115
## [38,]  0.9280  0.0980
## [39,]  0.1872 -0.8885
## [40,]  1.4433  1.1311
## [41,] -0.6832  1.3131
## [42,] -1.1703 -1.3188
## [43,] -2.6693 -1.6854
## [44,] -5.3511  1.2296
## [45,]  1.1850 -0.3615
## [46,]  2.3993 -1.1054
## [47,] -0.8160  0.2449
## [48,] -1.4812  1.3607
## [49,] -7.0560 -0.7124
## [50,]  0.5802 -1.5093
## [51,]  1.7365 -0.4183
## [52,] -1.2228 -2.3573
## [53,]  0.5112  2.6137
## [54,]  1.1099 -0.7942
## [55,]  1.5183  1.1928
## [56,] -1.4481 -1.9058
## [57,]  1.3906 -1.6948
## [58,] -1.5259 -1.3766
## [59,] -0.5741  0.2347
## [60,] -1.3961  1.2486
## [61,] -1.1882  0.3087
## [62,]  0.1899 -2.1136
## [63,] -6.4874 -0.6421
## [64,]  2.4116 -0.0803
## [65,]  0.7118 -0.8184
## [66,]  3.2630  1.2619
## [67,]  1.6682  0.0553
## [68,]  1.0382 -0.2603
## [69,]  1.6500 -0.1162
## [70,]  2.5538  1.4561
## [71,]  1.5336  0.0765
```

```
## [72,] -3.0647  1.0910
## [73,]  2.2993  0.2838
## [74,]  2.3264 -0.4002
## [75,] -0.0538 -2.0240
## [76,] -1.2699 -3.0072
## [77,] -0.1058  0.4814
## [78,]  1.8358  0.8310
## [79,] -0.7051  1.0827
## [80,] -0.3817  1.6765
## [81,] -2.4946  1.1319
## [82,] -0.7825  1.2010
## [83,]  2.3751 -0.2530
## [84,] -0.0610  1.0080
## [85,] -0.4309 -0.9761
## [86,]  1.0228  0.4212
## [87,] -1.5525 -2.0461
## [88,] -0.5767 -0.5895
## [89,] -3.0308  3.4680
## [90,] -3.0856  3.3485
## [91,] -2.4112 -2.6076
## [92,]  0.6926  0.3419
## [93,]  0.6463  0.6573
## [94,]  1.3219 -0.3345
## [95,]  1.3915 -0.4587
```

```
summary(scores)
```

```
##        Z1                  Z2
##  Min.   :-7.056000   Min.   :-3.007200
##  1st Qu.:-0.930150   1st Qu.:-0.806300
##  Median : 0.229500   Median : 0.098000
##  Mean   :-0.000001   Mean   :-0.000001
##  3rd Qu.: 1.391050   3rd Qu.: 0.900150
##  Max.   : 3.263000   Max.   : 3.468000
```

```
biplot(Datos_acp, choices = 1:2, col = c("gray73","sienna3"),
scale = 0, xlim = c(-7.1, 4), ylim = c(-3.1, 4))
abline(h = 0, v = 0, col = "red", lty = 5)
```




```
Z1 <-round(scale(as.matrix(Datos)) %*%
Datos_acp$rotation[,1],4)
Z3 <-round(scale(as.matrix(Datos)) %*%
Datos_acp$rotation[,3],4)
scores <-cbind(Z1, Z3)
colnames(scores) <- c("Z1","Z3")
scores
```

```
##            Z1      Z3
##  [1,] -0.4229 -1.1349
##  [2,] -1.3482 -1.7778
##  [3,]  0.7652  0.8626
##  [4,] -0.2491 -1.8989
##  [5,]  0.9367  0.1025
##  [6,]  0.8251  0.9493
##  [7,] -0.0178  0.5278
##  [8,] -0.7110 -0.7777
##  [9,]  1.3822  0.0114
## [10,] -0.1278 -0.0021
## [11,] -0.3963  0.8583
## [12,]  2.5439 -0.4516
## [13,]  1.6698  0.7329
## [14,]  1.3040 -0.5584
## [15,] -2.0950 -0.4498
## [16,]  1.6109  0.9009
## [17,]  0.1347 -1.0094
## [18,] -2.0950 -0.4498
```

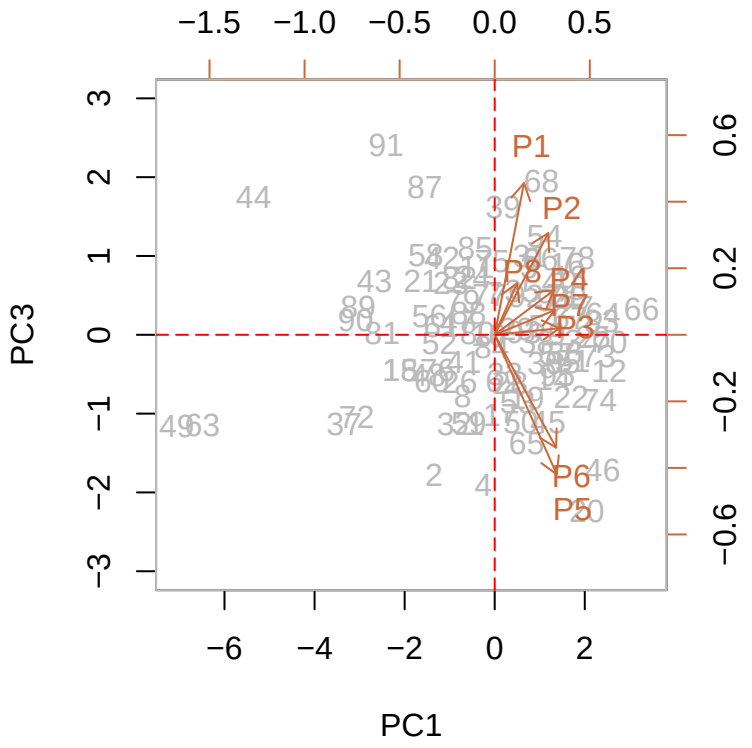
```
## [19,]  0.7102 -0.7986
## [20,]  2.0472 -2.2377
## [21,] -1.6307  0.6883
## [22,]  1.6984 -0.7802
## [23,] -0.9688  0.6566
## [24,] -0.4937  0.7384
## [25,]  2.1897 -0.0288
## [26,] -0.7769 -0.5998
## [27,]  1.8099  0.3602
## [28,]  0.3504 -0.6170
## [29,]  0.4871  0.0747
## [30,]  1.1163 -0.3597
## [31,]  0.7941  0.9936
## [32,] -0.8915 -1.1272
## [33,]  0.2295 -0.4985
## [34,]  1.9477 -0.0812
## [35,]  0.6109  0.5559
## [36,]  1.5598 -0.2367
## [37,] -3.3439 -1.1354
## [38,]  0.9280 -0.0961
## [39,]  0.1872  1.6189
## [40,]  1.4433  0.6669
## [41,] -0.6832 -0.3375
## [42,] -1.1703  0.9716
## [43,] -2.6693  0.6778
## [44,] -5.3511  1.7474
## [45,]  1.1850 -1.1123
## [46,]  2.3993 -1.7110
## [47,] -0.8160  0.1619
## [48,] -1.4812 -0.5121
## [49,] -7.0560 -1.1525
## [50,]  0.5802 -1.1088
## [51,]  1.7365 -0.3299
## [52,] -1.2228 -0.1035
## [53,]  0.5112 -0.8583
## [54,]  1.1099  1.2492
## [55,]  1.5183 -0.3502
## [56,] -1.4481  0.2406
## [57,]  1.3906 -0.2461
## [58,] -1.5259  1.0079
## [59,] -0.5741 -1.1174
## [60,] -1.3961 -0.5844
## [61,] -1.1882  0.1066
## [62,]  0.1899 -0.5914
## [63,] -6.4874 -1.1422
## [64,]  2.4116  0.2615
## [65,]  0.7118 -1.3733
## [66,]  3.2630  0.3282
## [67,]  1.6682  0.2084
## [68,]  1.0382  1.9542
## [69,]  1.6500  0.4742
## [70,]  2.5538 -0.0969
## [71,]  1.5336  0.5064
## [72,] -3.0647 -1.0341
```

```
## [73,]  2.2993 -0.2757
## [74,]  2.3264 -0.8282
## [75,] -0.0538  0.9364
## [76,] -1.2699 -0.4537
## [77,] -0.1058  0.5070
## [78,]  1.8358  0.9753
## [79,] -0.7051  0.4304
## [80,] -0.3817  0.0163
## [81,] -2.4946  0.0262
## [82,] -0.7825  0.7149
## [83,]  2.3751  0.1746
## [84,] -0.0610 -0.1568
## [85,] -0.4309  1.1030
## [86,]  1.0228  0.9639
## [87,] -1.5525  1.8769
## [88,] -0.5767  0.2702
## [89,] -3.0308  0.3485
## [90,] -3.0856  0.1865
## [91,] -2.4112  2.4057
## [92,]  0.6926  0.4923
## [93,]  0.6463  0.0409
## [94,]  1.3219  0.4361
## [95,]  1.3915 -0.5166
```

```
summary(scores)
```

```
##        Z1                  Z3
##  Min.   :-7.056000   Min.   :-2.2377
##  1st Qu.:-0.930150   1st Qu.:-0.5714
##  Median : 0.229500   Median : 0.0163
##  Mean   :-0.000001   Mean   : 0.0000
##  3rd Qu.: 1.391050   3rd Qu.: 0.6062
##  Max.   : 3.263000   Max.   : 2.4057
```

```
biplot(Datos_acp, choices = c(1,3), col = c("gray73","sienna3"),
scale = 0, xlim = c(-7.1, 3.4), ylim = c(-3, 3))
abline(h = 0, v = 0, col = "red", lty = 5)
```




```
Z2 <-round(scale(as.matrix(Datos)) %*%
Datos_acp$rotation[,2],4)
Z3 <-round(scale(as.matrix(Datos)) %*%
Datos_acp$rotation[,3],4)
scores <-cbind(Z2, Z3)
colnames(scores) <- c("Z2","Z3")
scores
```

```
##            Z2      Z3
##  [1,] -1.8994 -1.1349
##  [2,] -1.5311 -1.7778
##  [3,]  0.1721  0.8626
##  [4,] -2.3744 -1.8989
##  [5,]  0.8185  0.1025
##  [6,]  0.2060  0.9493
##  [7,]  0.7354  0.5278
##  [8,]  0.2748 -0.7777
##  [9,]  0.8976  0.0114
## [10,] -1.1376 -0.0021
## [11,]  0.4884  0.8583
## [12,]  1.1262 -0.4516
## [13,] -0.1837  0.7329
## [14,] -0.2575 -0.5584
## [15,]  0.2499 -0.4498
## [16,]  0.1163  0.9009
## [17,]  0.0745 -1.0094
## [18,]  0.2499 -0.4498
```

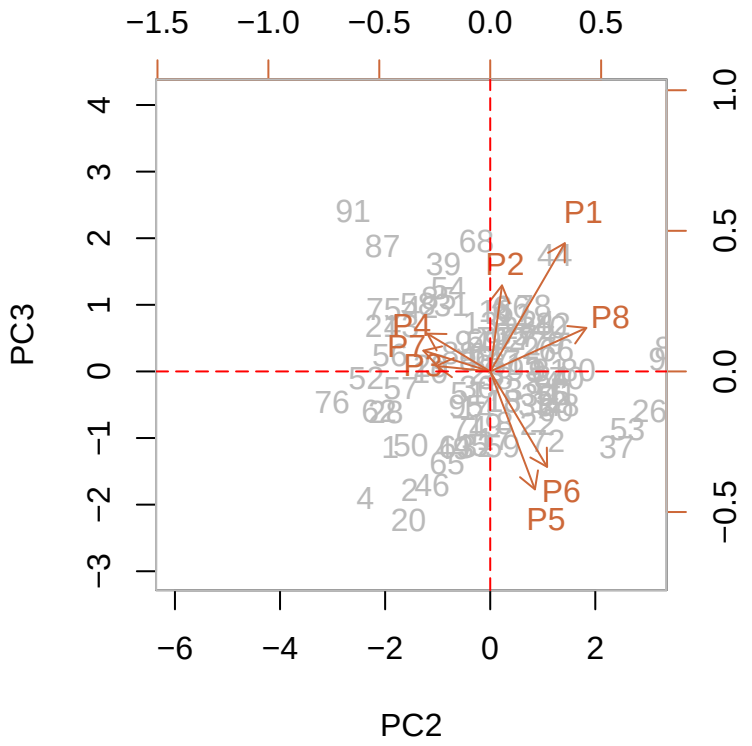
```
## [19,] -0.0995 -0.7986
## [20,] -1.5537 -2.2377
## [21,] -2.0467  0.6883
## [22,]  0.9027 -0.7802
## [23,]  0.5390  0.6566
## [24,]  0.8959  0.7384
## [25,] -0.0705 -0.0288
## [26,]  3.0339 -0.5998
## [27,]  0.9391  0.3602
## [28,] -1.9849 -0.6170
## [29,] -1.1285  0.0747
## [30,]  0.9192 -0.3597
## [31,] -0.7320  0.9936
## [32,] -0.2486 -1.1272
## [33,]  0.8684 -0.4985
## [34,]  1.1681 -0.0812
## [35,]  0.4142  0.5559
## [36,] -0.2474 -0.2367
## [37,]  2.4115 -1.1354
## [38,]  0.0980 -0.0961
## [39,] -0.8885  1.6189
## [40,]  1.1311  0.6669
## [41,]  1.3131 -0.3375
## [42,] -1.3188  0.9716
## [43,] -1.6854  0.6778
## [44,]  1.2296  1.7474
## [45,] -0.3615 -1.1123
## [46,] -1.1054 -1.7110
## [47,]  0.2449  0.1619
## [48,]  1.3607 -0.5121
## [49,] -0.7124 -1.1525
## [50,] -1.5093 -1.1088
## [51,] -0.4183 -0.3299
## [52,] -2.3573 -0.1035
## [53,]  2.6137 -0.8583
## [54,] -0.7942  1.2492
## [55,]  1.1928 -0.3502
## [56,] -1.9058  0.2406
## [57,] -1.6948 -0.2461
## [58,] -1.3766  1.0079
## [59,]  0.2347 -1.1174
## [60,]  1.2486 -0.5844
## [61,]  0.3087  0.1066
## [62,] -2.1136 -0.5914
## [63,] -0.6421 -1.1422
## [64,] -0.0803  0.2615
## [65,] -0.8184 -1.3733
## [66,]  1.2619  0.3282
## [67,]  0.0553  0.2084
## [68,] -0.2603  1.9542
## [69,] -0.1162  0.4742
## [70,]  1.4561 -0.0969
## [71,]  0.0765  0.5064
## [72,]  1.0910 -1.0341
```

```
## [73,]  0.2838 -0.2757
## [74,] -0.4002 -0.8282
## [75,] -2.0240  0.9364
## [76,] -3.0072 -0.4537
## [77,]  0.4814  0.5070
## [78,]  0.8310  0.9753
## [79,]  1.0827  0.4304
## [80,]  1.6765  0.0163
## [81,]  1.1319  0.0262
## [82,]  1.2010  0.7149
## [83,] -0.2530  0.1746
## [84,]  1.0080 -0.1568
## [85,] -0.9761  1.1030
## [86,]  0.4212  0.9639
## [87,] -2.0461  1.8769
## [88,] -0.5895  0.2702
## [89,]  3.4680  0.3485
## [90,]  3.3485  0.1865
## [91,] -2.6076  2.4057
## [92,]  0.3419  0.4923
## [93,]  0.6573  0.0409
## [94,] -0.3345  0.4361
## [95,] -0.4587 -0.5166
```

```
summary(scores)
```

```
##        Z2                  Z3
##  Min.   :-3.007200   Min.   :-2.2377
##  1st Qu.:-0.806300   1st Qu.:-0.5714
##  Median : 0.098000   Median : 0.0163
##  Mean   :-0.000001   Mean   : 0.0000
##  3rd Qu.: 0.900150   3rd Qu.: 0.6062
##  Max.   : 3.468000   Max.   : 2.4057
```

```
biplot(Datos_acp, choices = 2:3, col = c("gray73","sienna3"),
scale = 0, xlim = c(-6, 3), ylim = c(-3, 4.1))
abline(h = 0, v = 0, col = "red", lty = 5)
```


−1.5
−1.0
−0.5
0.0
0.5
1.0
0.5
0.0
−0.5
4
3
2
1
0
−1
−2
−3
PC3
−6
−4
−2
0
2
PC2
P1
P2
P8
P4
P7
P3
P6
P5