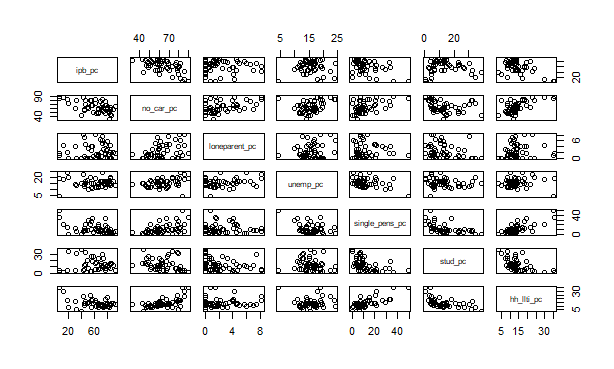
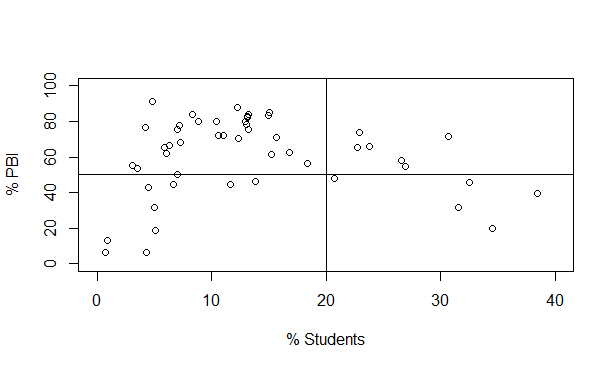
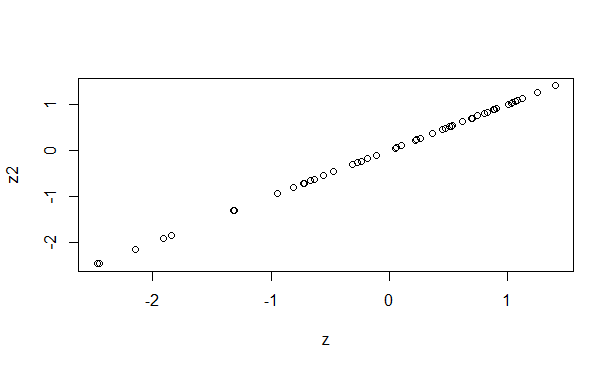
|  |
| --- |
| > my.data <- read.spss("UniWd-ED91data-Vote97.sav", to.data.frame = T)  re-encoding from CP1252  > names(my.data)  [1] "ed91" "Easting" "Northing" "area"  [5] "allppl" "VoteIndex" "wht\_oth\_pc" "ipb\_pc"  [9] "non\_home\_own\_pc" "no\_car\_pc" "overcr\_pc" "loneparent\_pc"  [13] "unemp\_pc" "single\_pens\_pc" "single\_pers\_pc" "ft\_pc"  [17] "stud\_pc" "hh\_llti\_pc" "Multi\_stress" "multistress"  > data <- my.data[, c(8, 10, 12, 13, 14, 17, 18)]  > names(data)  [1] "ipb\_pc" "no\_car\_pc" "loneparent\_pc" "unemp\_pc"  [5] "single\_pens\_pc" "stud\_pc" "hh\_llti\_pc"  > round(cor(data), 3)  ipb\_pc no\_car\_pc loneparent\_pc unemp\_pc single\_pens\_pc stud\_pc  ipb\_pc 1.000 -0.553 0.020 0.008 -0.452 -0.019  no\_car\_pc -0.553 1.000 0.325 0.351 0.473 -0.265  loneparent\_pc 0.020 0.325 1.000 0.200 0.015 -0.270  unemp\_pc 0.008 0.351 0.200 1.000 -0.386 -0.132  single\_pens\_pc -0.452 0.473 0.015 -0.386 1.000 -0.494  stud\_pc -0.019 -0.265 -0.270 -0.132 -0.494 1.000  hh\_llti\_pc -0.504 0.623 0.103 0.005 0.748 -0.550  hh\_llti\_pc  ipb\_pc -0.504  no\_car\_pc 0.623  loneparent\_pc 0.103  unemp\_pc 0.005  single\_pens\_pc 0.748  stud\_pc -0.550  hh\_llti\_pc 1.000  > plot(data) |
|  |
| |  | | --- | | > | |



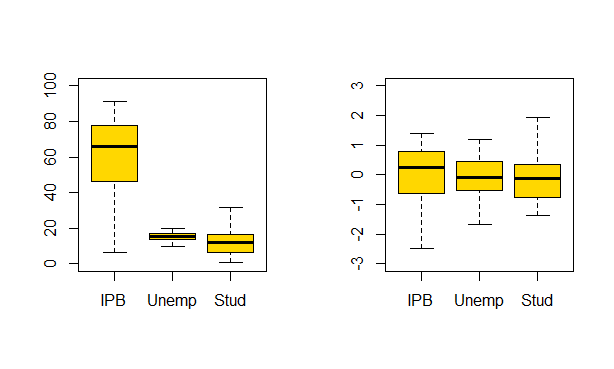
|  |
| --- |
| > plot(data)  > plot(data$stud\_pc, data$ipb\_pc,  + xlab = "% Students", ylab = "% PBI",  + ylim = c(0,100), xlim = c(0, 40))  > abline(h = 50)  > abline(v = 20) |
|  |
| |  | | --- | | > | |



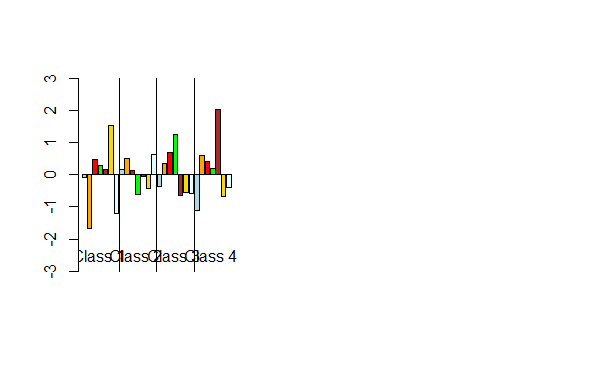
|  |
| --- |
| > z <- scale(data$ipb\_pc)  > head(z)  [,1]  [1,] 0.2201710  [2,] -0.2387659  [3,] 1.4019907  [4,] 0.6996574  [5,] 0.4457579  [6,] 0.8001212 |
|  |
| |  | | --- | | > | |
| > tail(z)  [,1]  [45,] 0.5366320  [46,] 0.4704172  [47,] -0.7164257  [48,] -0.4693760  [49,] 0.1009844  [50,] -0.2702750  > attributes(z)[3]  $`scaled:scale`  [1] 21.89843  > attributes(z)[2]  $`scaled:center`  [1] 60.4186  > z2 <- (data$ipb\_pc - mean(data$ipb\_pc)) / sd(data$ipb\_pc)  > plot(z, z2) |
|  |
| |  | | --- | | > | |



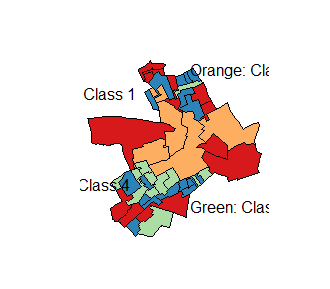
|  |
| --- |
| > par(mfrow = c(1,2))  > boxplot(data[,c(1,4,6)],  + ylim = c(0,100),  + outline = F,  + col = "gold",  + horizontal = F,  + names = c("IPB", "Unemp", "Stud"))  > # calculate z-scores  > data.z <- apply(data, 2, scale)  > # 2nd box plot  > boxplot(data.z[,c(1,4,6)],  + ylim = c(-3,3),  + outline = F,  + col = "gold",  + horizontal = F,  + names = c("IPB", "Unemp", "Stud")) |
|  |
| |  | | --- | | > | |



|  |
| --- |
| > set.seed(1)  > k.res <- kmeans(data.z, 9, iter.max = 100)  > k.res  K-means clustering with 9 clusters of sizes 4, 1, 11, 10, 3, 3, 7, 6, 5  Cluster means:  ipb\_pc no\_car\_pc loneparent\_pc unemp\_pc single\_pens\_pc stud\_pc  1 -1.70941908 1.512167613 -0.2112872 1.2787426 1.0911242 -1.17675975  2 -2.45307958 1.475439829 -0.7085440 -3.1762479 3.6262804 -0.99117980  3 0.83312477 -0.257084213 -0.5330305 0.2987588 -0.3902574 -0.26626884  4 -0.16469671 0.063360136 0.2109119 -0.6326505 1.0092913 -0.62001991  5 -0.07543614 1.403867738 2.0894376 1.7319107 -0.4647102 -0.67775589  6 -1.27262985 0.997288033 -1.0254540 0.2780546 -0.7312943 2.08532330  7 -0.11109862 -0.691311064 -0.5625986 -0.3552519 -0.5923223 1.34520640  8 0.43411324 0.001180312 1.4827217 0.1609281 -0.5267937 0.09737658  9 0.79811188 -1.540231406 -0.5684364 -0.6814532 -0.5791654 0.12079372  hh\_llti\_pc  1 1.9396718  2 3.2246237  3 -0.2489682  4 0.1779199  5 0.4335135  6 -0.6260417  7 -0.7044090  8 -0.3327085  9 -0.5038322  Clustering vector:  [1] 4 4 3 4 3 3 8 8 3 5 1 3 5 5 8 4 1 3 2 1 4 4 1 4 9 3 9 9 3 9 7 7 3 8 7 7 6 7 3 6  [41] 6 7 4 4 8 3 8 4 9 7  Within cluster sum of squares by cluster:  [1] 13.234225 0.000000 14.049345 24.727271 5.660705 4.833859 12.833293 11.594792  [9] 3.469826  (between\_SS / total\_SS = 73.6 %)  Available components:  [1] "cluster" "centers" "totss" "withinss" "tot.withinss"  [6] "betweenss" "size" "iter" "ifault" |
|  |
| |  | | --- | | > | |
| > set.seed(100)  > k.res <- kmeans(data.z, 4, iter.max = 10000)  > round(k.res$centers, 3)  ipb\_pc no\_car\_pc loneparent\_pc unemp\_pc single\_pens\_pc stud\_pc hh\_llti\_pc  1 -0.075 0.169 0.509 -0.416 0.714 -0.586 0.207  2 -1.683 1.531 0.139 0.640 1.247 -1.111 2.048  3 0.496 -1.216 -0.612 -0.373 -0.648 0.605 -0.686  4 0.289 0.169 -0.034 0.358 -0.540 0.423 -0.386  > k.res$cluster  [1] 1 1 4 1 4 4 4 1 3 1 2 4 2 4 4 1 2 1 2 2 1 1 2 1 3 3 3 3 4 3 3 3 4 4 3 4 4 3 4 4  [41] 4 3 1 1 4 4 1 1 3 4 |
|  |
| |  | | --- | | > | |
| > col.set = c("lightblue", "orange", "red",  + "green", "brown", "gold", "lightcyan")  > barplot(as.vector(k.res$centers),  + ylim = c(-3, 3), xlim = c(1, 40),  + col = rep(col.set, 4))  > abline(v = 8.5)  > abline(v = 16.8)  > abline(v = 25.3)  > text(4,-2.5, "Class 1")  > text(12.3,-2.5, "Class 2")  > text(20.8,-2.5, "Class 3")  > text(29.1,-2.5, "Class 4") |
|  |
| |  | | --- | | > | |



|  |
| --- |
| > colnames(data.z[,1:7])  [1] "ipb\_pc" "no\_car\_pc" "loneparent\_pc" "unemp\_pc"  [5] "single\_pens\_pc" "stud\_pc" "hh\_llti\_pc" |
|  |
| |  | | --- | | > | |
| > library(GISTools)  > my.shp <- readShapePoly("uni91edspolygons.shp")  Warning message:  use rgdal::readOGR or sf::st\_read  > data.c <- data.frame(my.data, k.res$cluster)  > my.shp@data = data.frame(my.shp@data,  + data.c[match(my.shp@data$ED91, data.c$ed91),])  > summary(my.shp)  Object of class SpatialPolygonsDataFrame  Coordinates:  min max  x 413737 416739  y 431602 434639  Is projected: NA  proj4string : [NA]  Data attributes:  AREA PERIMETER ED91 ed91 Easting  Min. : 6611 Min. : 403.8 08CXGD01: 1 08CXGD01: 1 Min. :414263  1st Qu.: 26590 1st Qu.: 810.6 08CXGD02: 1 08CXGD02: 1 1st Qu.:414766  Median : 42490 Median :1065.2 08CXGD03: 1 08CXGD03: 1 Median :415068  Mean : 91492 Mean :1361.8 08CXGD04: 1 08CXGD04: 1 Mean :415121  3rd Qu.: 85932 3rd Qu.:1508.1 08CXGD05: 1 08CXGD05: 1 3rd Qu.:415504  Max. :629794 Max. :4512.7 08CXGD06: 1 08CXGD06: 1 Max. :416217  (Other) :44 (Other) :44  Northing area allppl VoteIndex wht\_oth\_pc  Min. :431778 Min. : 6611 Min. : 59.0 Min. : 51.42 Min. : 8.67  1st Qu.:432510 1st Qu.: 26590 1st Qu.:233.0 1st Qu.: 90.85 1st Qu.:19.38  Median :432738 Median : 42490 Median :390.0 Median :101.53 Median :32.30  Mean :433103 Mean : 91492 Mean :374.6 Mean :100.61 Mean :36.63  3rd Qu.:433998 3rd Qu.: 85932 3rd Qu.:473.0 3rd Qu.:110.91 3rd Qu.:47.69  Max. :434459 Max. :629794 Max. :859.0 Max. :140.79 Max. :92.00    ipb\_pc non\_home\_own\_pc no\_car\_pc overcr\_pc loneparent\_pc  Min. : 6.29 Min. : 9.52 Min. :32.89 Min. : 0.000 Min. :0.000  1st Qu.:46.90 1st Qu.:26.66 1st Qu.:56.25 1st Qu.: 3.340 1st Qu.:0.620  Median :65.79 Median :46.98 Median :62.33 Median : 4.785 Median :1.680  Mean :60.42 Mean :49.11 Mean :64.14 Mean : 4.978 Mean :2.459  3rd Qu.:77.65 3rd Qu.:67.34 3rd Qu.:75.84 3rd Qu.: 6.700 3rd Qu.:4.140  Max. :91.12 Max. :95.35 Max. :89.84 Max. :14.670 Max. :8.250    unemp\_pc single\_pens\_pc single\_pers\_pc ft\_pc stud\_pc  Min. : 4.29 Min. : 0.000 Min. : 3.36 Min. : 8.09 Min. : 0.700  1st Qu.:13.81 1st Qu.: 5.688 1st Qu.:10.06 1st Qu.:19.62 1st Qu.: 6.412  Median :15.40 Median : 8.965 Median :16.87 Median :23.81 Median :12.320  Mean :15.68 Mean :12.140 Mean :25.19 Mean :23.64 Mean :13.503  3rd Qu.:17.21 3rd Qu.:15.307 3rd Qu.:38.35 3rd Qu.:27.23 3rd Qu.:16.500  Max. :24.26 Max. :49.230 Max. :73.33 Max. :33.91 Max. :38.460    hh\_llti\_pc Multi\_stress multistress  Min. : 3.17 No stress : 5 None or little stress:15  1st Qu.:10.20 One stress indicator :10 Multiple stress :35  Median :12.81 Two stress indicators :13  Mean :14.08 Three stress indicators:15  3rd Qu.:15.48 Four stress indicators : 7  Max. :35.84    k.res.cluster  Min. :1.00  1st Qu.:1.00  Median :3.00  Mean :2.68  3rd Qu.:4.00  Max. :4.00 |
|  |
| |  | | --- | | > | |
| > cols <- brewer.pal(4, "Spectral")  > plot(my.shp, col = cols[data.c$k.res.cluster])  > text(413805.7, 434078.5, "Red: Class 1")  > text(416491, 434478.5, "Orange: Class 2")  > text(416403.4,432128.6, "Green: Class 3")  > text(413677.8,432512.2, "Blue: Class 4") |
|  |
| |  | | --- | | > | |



|  |
| --- |
| > aggregate(my.data$VoteIndex,  + by=list(cluster=k.res$cluster), mean)  cluster x  1 1 99.07429  2 2 70.57833  3 3 108.56917  4 4 106.50222 |
|  |
| |  | | --- | | > | |
| > aggregate(my.data[, c(6, 8, 10, 12, 13, 14, 17, 18)],  + by=list(cluster=k.res$cluster), mean)  cluster VoteIndex ipb\_pc no\_car\_pc loneparent\_pc unemp\_pc single\_pens\_pc  1 1 99.07429 58.77286 66.37714 3.6800000 14.18929 19.445000  2 2 70.57833 23.55333 84.45167 2.7933333 17.97333 24.898333  3 3 108.56917 71.27333 47.98917 0.9916667 14.34333 5.515833  4 4 106.50222 66.75056 66.38556 2.3766667 16.96500 6.621111  stud\_pc hh\_llti\_pc  1 8.057857 15.43571  2 3.175000 27.51833  3 19.125833 9.57750  4 17.432778 11.54333 |
|  |
| |  | | --- | | > | |