MVA_Ass5.R

mihikagupta

2020-10-15

```
###### Assignment 5 ##############
## Applying Cluster Analysis

#Getting working directory
getwd()
```

```
## [1] "/Users/mihikagupta/Desktop/SEM_2/MVA"
```

```
#Setting directory to load data set
setwd("/Users/mihikagupta/Desktop/SEM_2/MVA")

#Reading the data into a data frame
#df <- read.csv(file = 'US_Acc_June20.csv')
num <- read.csv(file = 'num.csv')
# Performing clustering on the first 500 records for now to achieve easy and quick results and test the
num<-num[1:500,]
attach(num)
# Printing first few columns of data set for inference
#head(df)

## Setting random seed to shuffle data before splitting
set.seed(23)

#Checking number of rows
#rows<-sample(nrow(df))

#Shuffling the data
#mva<-df[rows,]

#Taking the required number of instances from the shuffled data to reduce any biases
#mva<-mva[950000:1000000,]

#Checking the structure of the data set
#str(mva)

# Checking the number of rows and columns in the current uncleaned dataset
#ncol(mva)
#nrow(mva)

# Printing all the column names to find and filter the relevant and irrelevant attributes
#names<-names(mva)
#names
```

```r
## DATA CLEANING ##

#Dropping the surplus attributes which do not contribute to the analysis
#mva <- mva[-c(1:3,7:10,13,14,19,21:23,33,47:49)]

#Checking for any null values in the present data set
# is.na(mva[,])

#Checking which rows have all the values filled and complete
# complete.cases(mva)

#Making a new dataframe with only the rows that have complete information and all values filled
#Mva<-na.omit(mva)
#Mva<-Mva[!(is.na(Mva$Sunrise_Sunset) | Mva$Sunrise_Sunset==""), ]
#Mva<- Mva[complete.cases(Mva),]
#Verifying for missing values in the new dataframe
#complete.cases(Mva)
#unique(Mva$Sunrise_Sunset)
#Checking the number of rows and columns in the new CLEANED dataframe
#ncol(Mva)
#nrow(Mva)

# Creating new dataframe with only the numerical attributes to perform statistical functions
#num<-Mva[,c(1,4,11:15,17,18)]
#write.csv(num,"/Users/mihikagupta/Desktop/SEM_2/MVA/num.csv", row.names = FALSE)
#Scaling the dataset before applying clusturing algorithms

nrow(num)
```

```
## [1] 500
```

```r
ncol(num)
```

```
## [1] 9
```

```r
###### Checking to see for any correlation that might interefere with the data #######
cor(num)
```

```
##                      Severity Distance.mi. Temperature.F. Wind_Chill.F.
## Severity           1.00000000  0.060325232   -0.026274516   -0.03165259
## Distance.mi.       0.06032523  1.000000000   -0.010229758   -0.01679232
## Temperature.F.    -0.02627452 -0.010229758    1.000000000    0.99372259
## Wind_Chill.F.     -0.03165259 -0.016792324    0.993722592    1.00000000
## Humidity...        0.08687654  0.004473866   -0.452662114   -0.43537980
## Pressure.in.       0.06567115 -0.008218439    0.055390478    0.05040739
## Visibility.mi.    -0.01784787 -0.041477419    0.256766004    0.26208308
## Wind_Speed.mph.   -0.06205678  0.009728557    0.004761887   -0.05069851
## Precipitation.in.  0.03617412  0.101174144   -0.104794100   -0.10133502
##                    Humidity... Pressure.in. Visibility.mi. Wind_Speed.mph.
## Severity           0.086876542  0.065671149    -0.01784787    -0.062056779
## Distance.mi.       0.004473866 -0.008218439    -0.04147742     0.009728557
## Temperature.F.    -0.452662114  0.055390478     0.25676600     0.004761887
## Wind_Chill.F.     -0.435379799  0.050407389     0.26208308    -0.050698506
## Humidity...        1.000000000  0.233181450    -0.39892391    -0.137092830
## Pressure.in.       0.233181450  1.000000000    -0.26428984    -0.014390090
## Visibility.mi.    -0.398923914 -0.264289840     1.00000000     0.014696688
```

```
## Wind_Speed.mph.   -0.137092830 -0.014390090     0.01469669     1.000000000
## Precipitation.in.  0.244133880  0.007382003    -0.24473243     0.020686860
##                   Precipitation.in.
## Severity                0.036174121
## Distance.mi.            0.101174144
## Temperature.F.         -0.104794100
## Wind_Chill.F.          -0.101335015
## Humidity...             0.244133880
## Pressure.in.            0.007382003
## Visibility.mi.         -0.244732428
## Wind_Speed.mph.         0.020686860
## Precipitation.in.       1.000000000
```

```r
# Removinf the wind chills attribute #
# since "wind chills had a high correlation with temperature, it is removed
num<-num[-c(4)]

# Checking the new dimensions
nrow(num)
```

```
## [1] 500
```

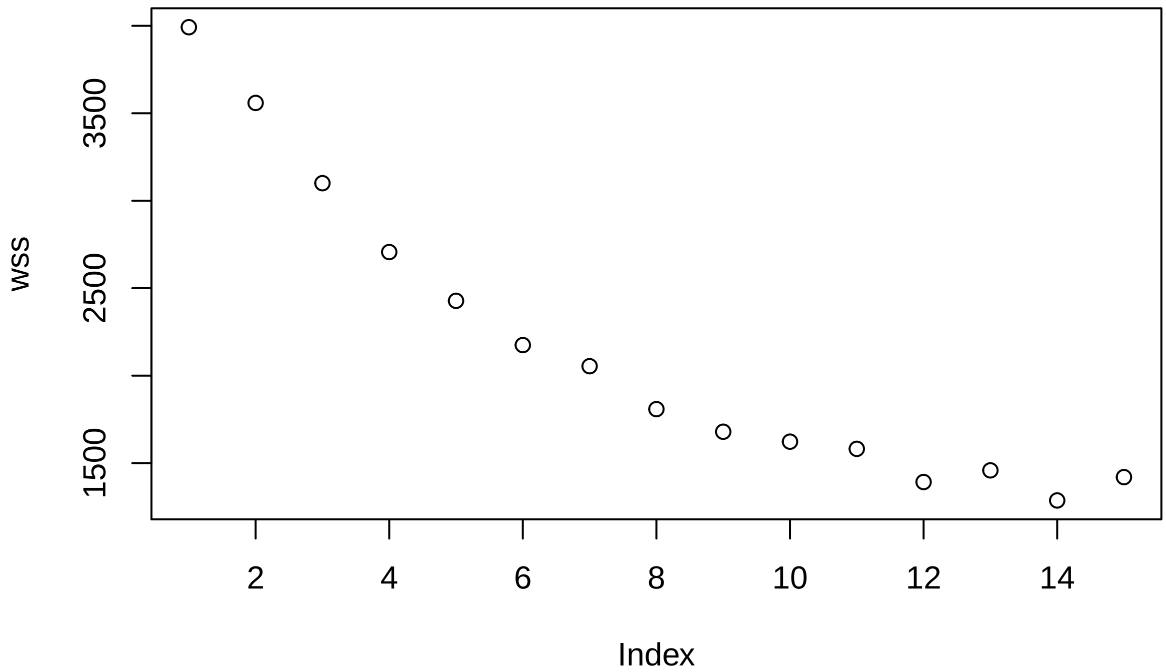
```r
ncol(num)
```

```
## [1] 8
```

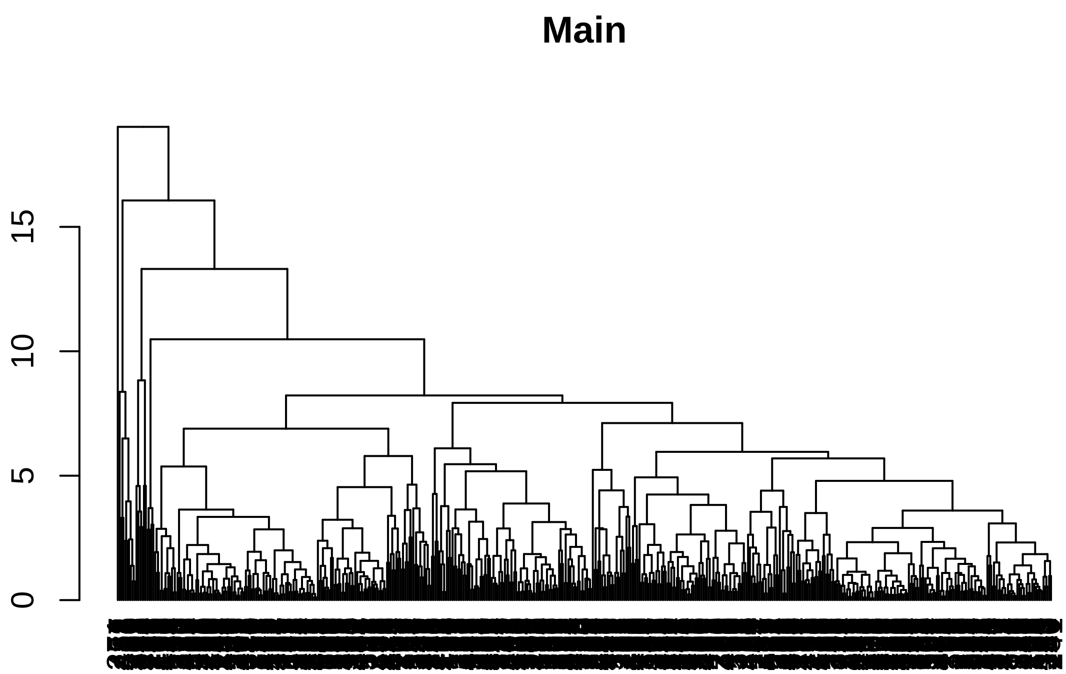
```r
# Scaling the new data set for better accuracies in clustering
num<-scale(num)

# Determine number of clusters
wss <- (nrow(num)-1)*sum(apply(num,2,var))
for (i in 2:15) wss[i] <- sum(kmeans(num,
                                     centers=i)$withinss)
plot(wss)
#here we observe that after 8-10 clusters, the distance remains almost constant, so we fix this number

## Loading the cluster library ###
library(cluster)
library(factoextra)
```

```
## Loading required package: ggplot2
```

```
## Welcome! Want to learn more? See two factoextra-related books at https://goo.gl/ve3WBa
```

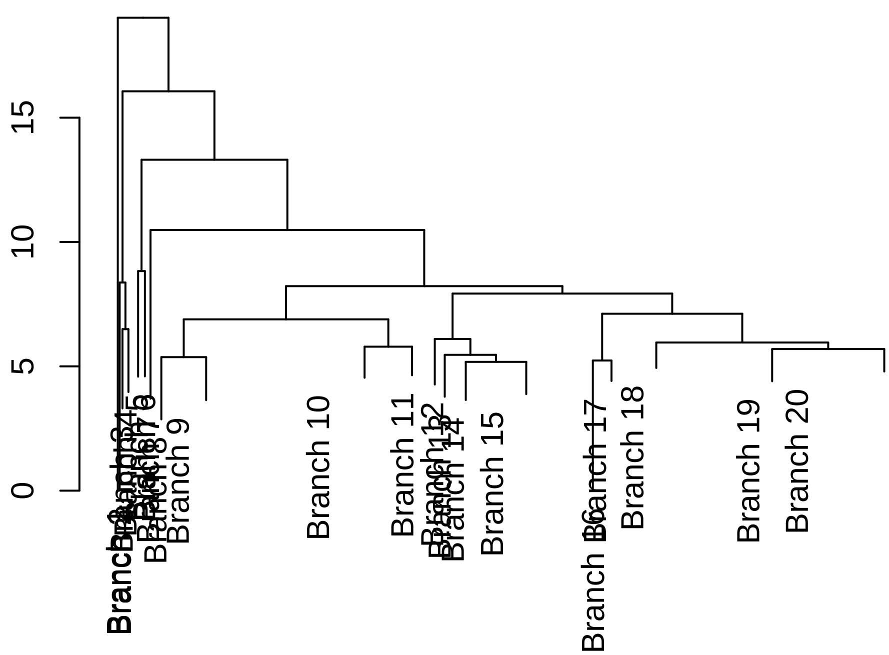



```
# Understanding different methods

# producing a distance matrix of the dataset
numdist<-dist(num,method = "euclidean")

#Sliced dendrograms for better visualizations

hc <- hclust(dist(num))
hcd <- as.dendrogram(hc)
plot(hcd, main="Main")
```



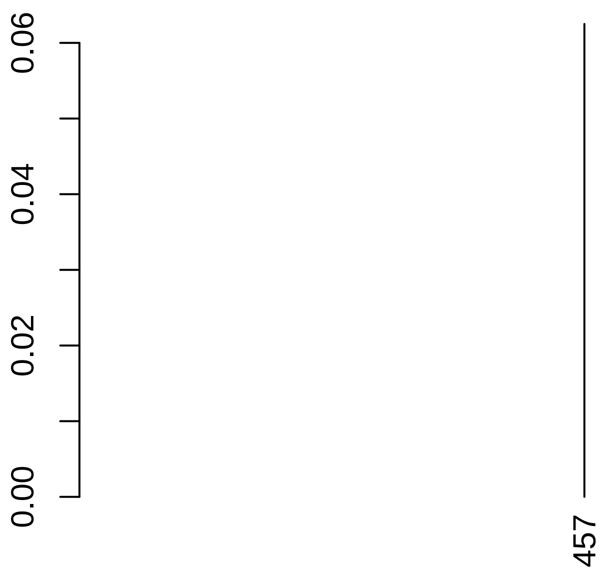
```r
plot(cut(hcd, h=5)$upper,
     main="Upper tree of cut at h=5")
```

Upper tree of cut at h=5



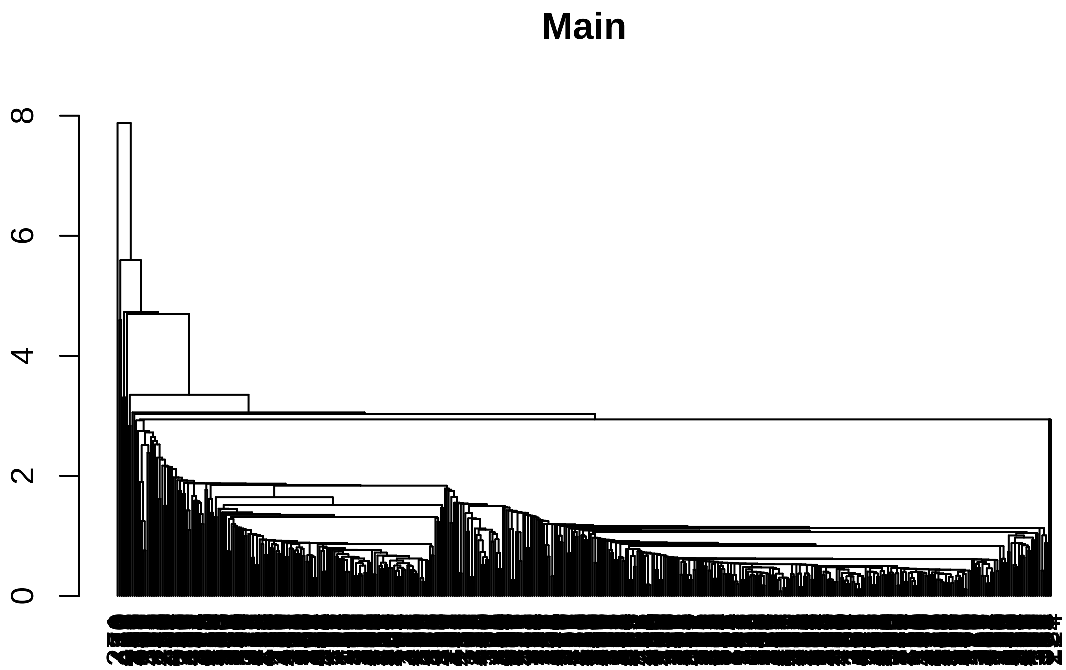

```r
plot(cut(hcd, h=5)$lower[[2]],
     main="Second branch of lower tree with cut at h=5")
```

Second branch of lower tree with cut at h=5

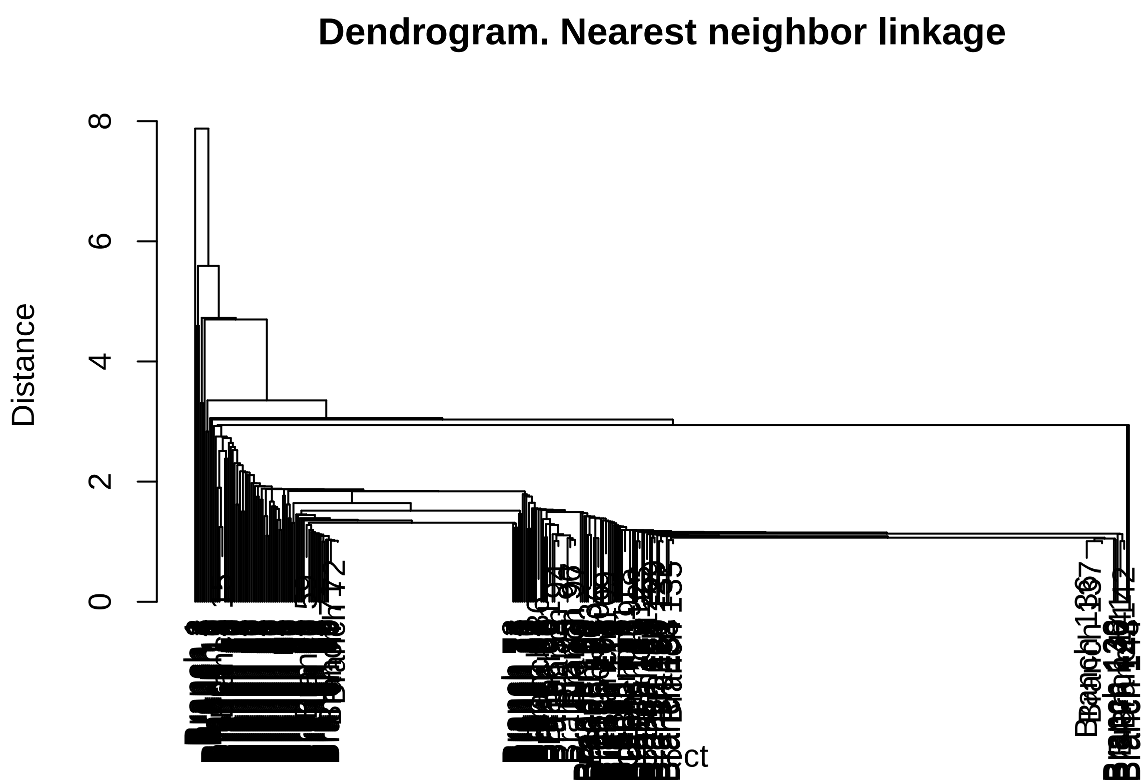



```r
#Single

numsingle<- hclust(numdist, method = "single")
```

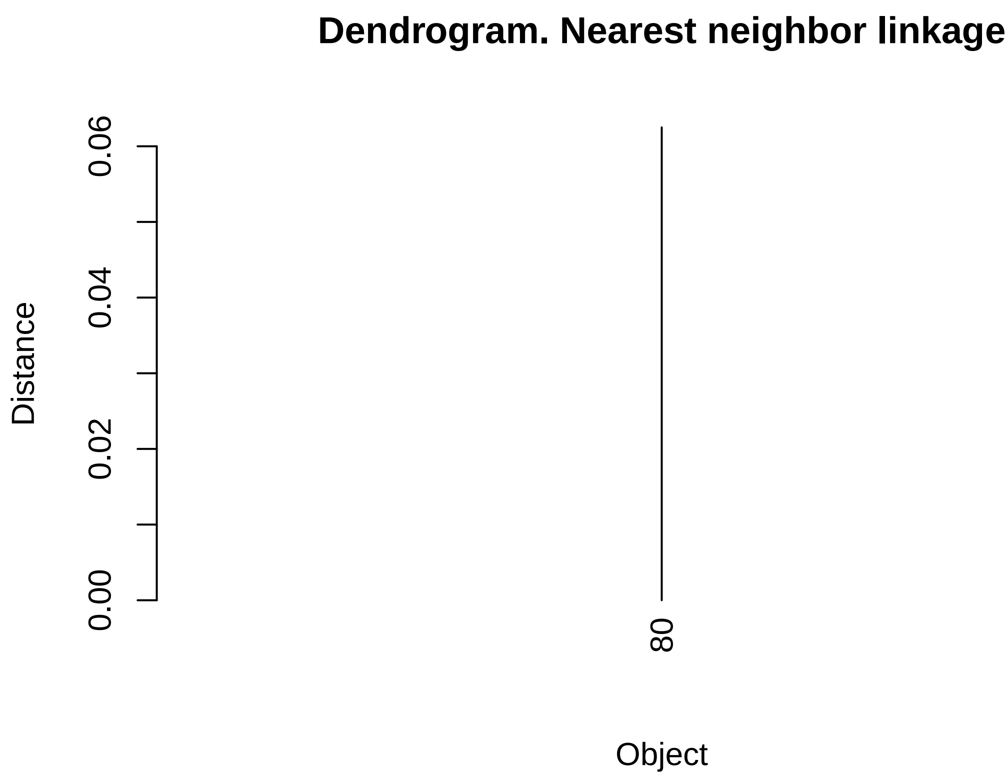
```r
single<-as.dendrogram(numsingle)
plot(single,main="Main")
```



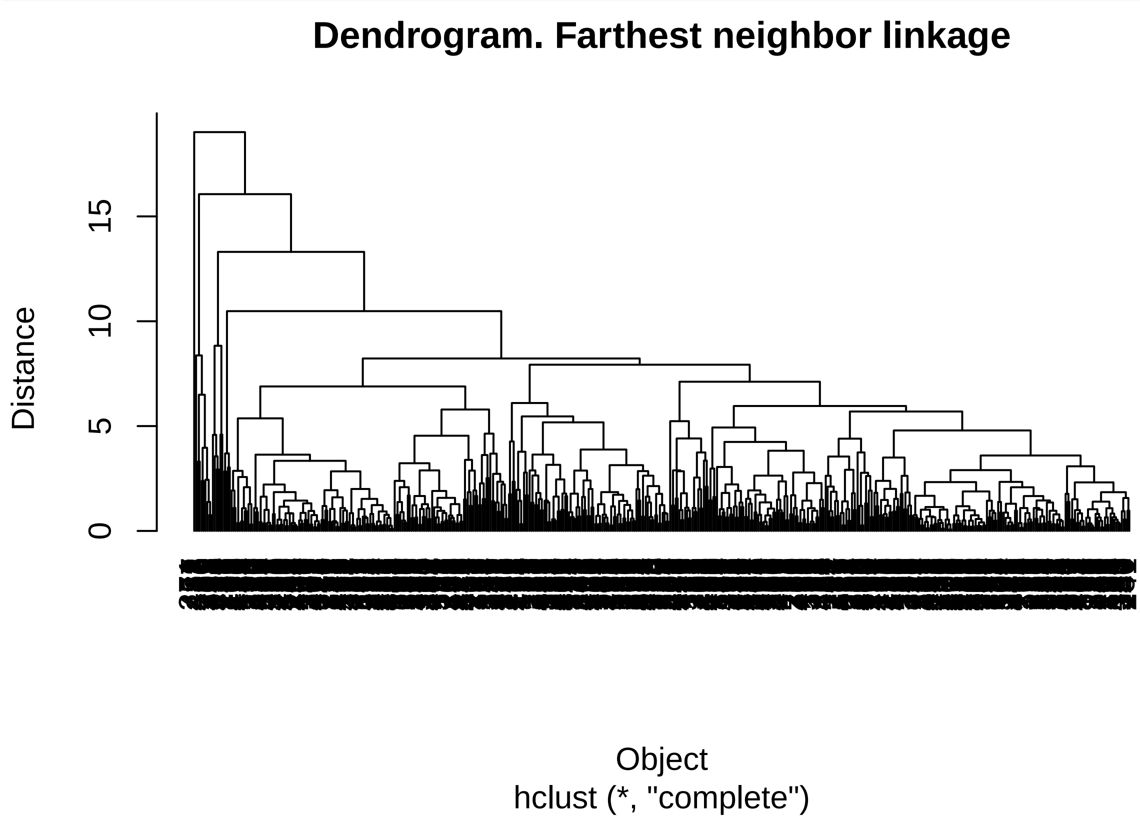

```r
plot(cut(single,h=1)$upper,xlab="Object",ylab="Distance",
        main="Dendrogram. Nearest neighbor linkage")
```



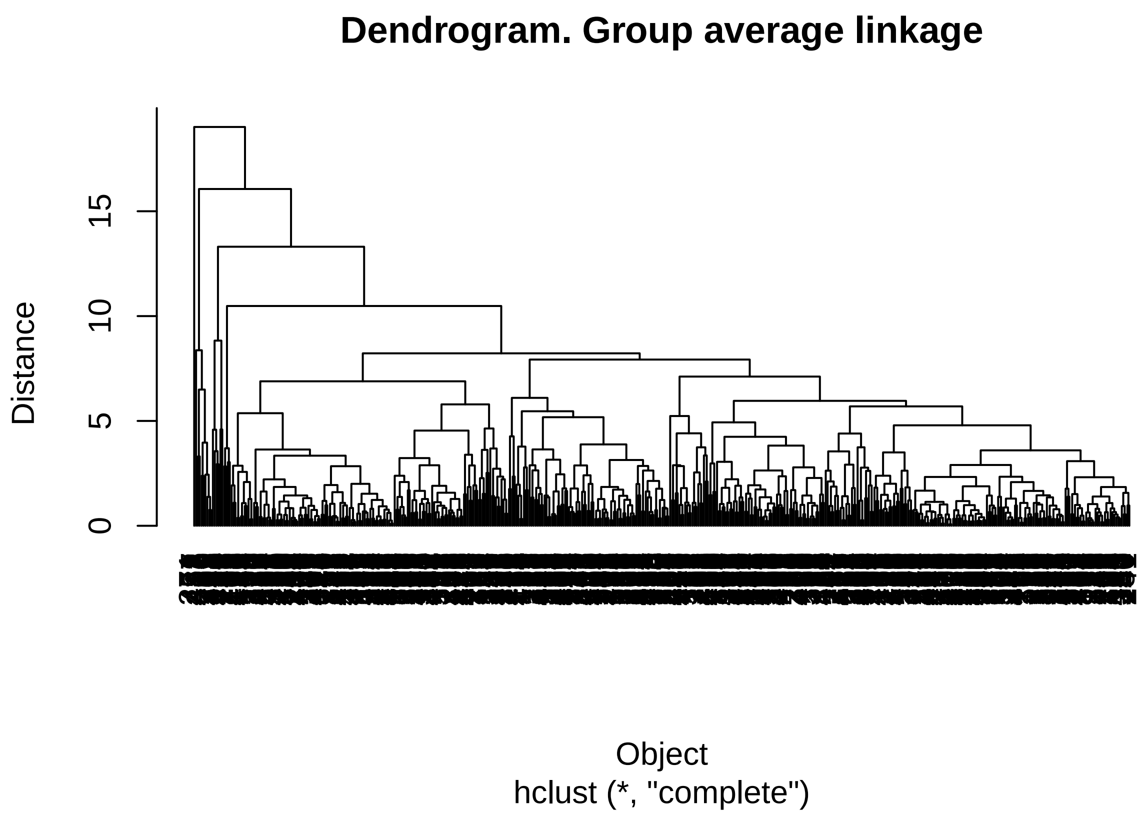

```r
plot(cut(single,h=1)$lower[[2]],xlab="Object",ylab="Distance",
     main="Dendrogram. Nearest neighbor linkage")
```




```r
# Complete link
numcomplete<-hclust(numdist)
plot(numcomplete, hang=-1,xlab="Object",ylab="Distance",
     main="Dendrogram. Farthest neighbor linkage")
```



```r
# Average linkage
numaverage<-hclust(numdist)
plot(numaverage, hang=-1,xlab="Object",ylab="Distance",
     main="Dendrogram. Group average linkage")
```




```r
# Single link clustering
km.res<-kmeans(num,centers = 10,iter.max = 10,nstart = 3)
print(km.res)
```

```
## K-means clustering with 10 clusters of sizes 4, 10, 12, 66, 92, 16, 21, 118, 63, 98
##
## Cluster means:
##      Severity Distance.mi. Temperature.F. Humidity... Pressure.in.
## 1   0.3743416 -0.286109267    -0.23287818 -1.05967239  -3.80953332
## 2   0.1908408  0.826887066    -0.40688757  1.14095972   0.14273235
## 3   0.2214243  4.773803896     0.08429905 -0.03386252   0.02608219
## 4   0.1797195 -0.085547380    -1.30850887  1.02826788   0.23463854
## 5  -0.4434336 -0.167237674     0.90139423 -1.06906195   0.21588584
## 6  -0.3137863 -0.250997367    -1.28100876  0.48714091  -3.06267571
## 7  -0.3683996 -0.182746036     0.62953348 -1.57746213  -2.77625202
## 8  -0.5898150 -0.227487660    -0.11392558  0.54886337   0.37576459
## 9  -0.6014165 -0.126760851     0.10716318 -0.29267170   0.17486842
## 10  1.4603665 -0.007091685     0.21826522  0.02784033   0.30711587
##    Visibility.mi. Wind_Speed.mph. Precipitation.in.
## 1      7.35699377      0.15304465      -0.220774169
## 2     -1.26889254      0.11862939       6.138093483
## 3      0.06005446     -0.06922055      -0.071924271
## 4     -1.34121619      0.12019372       0.082338351
```

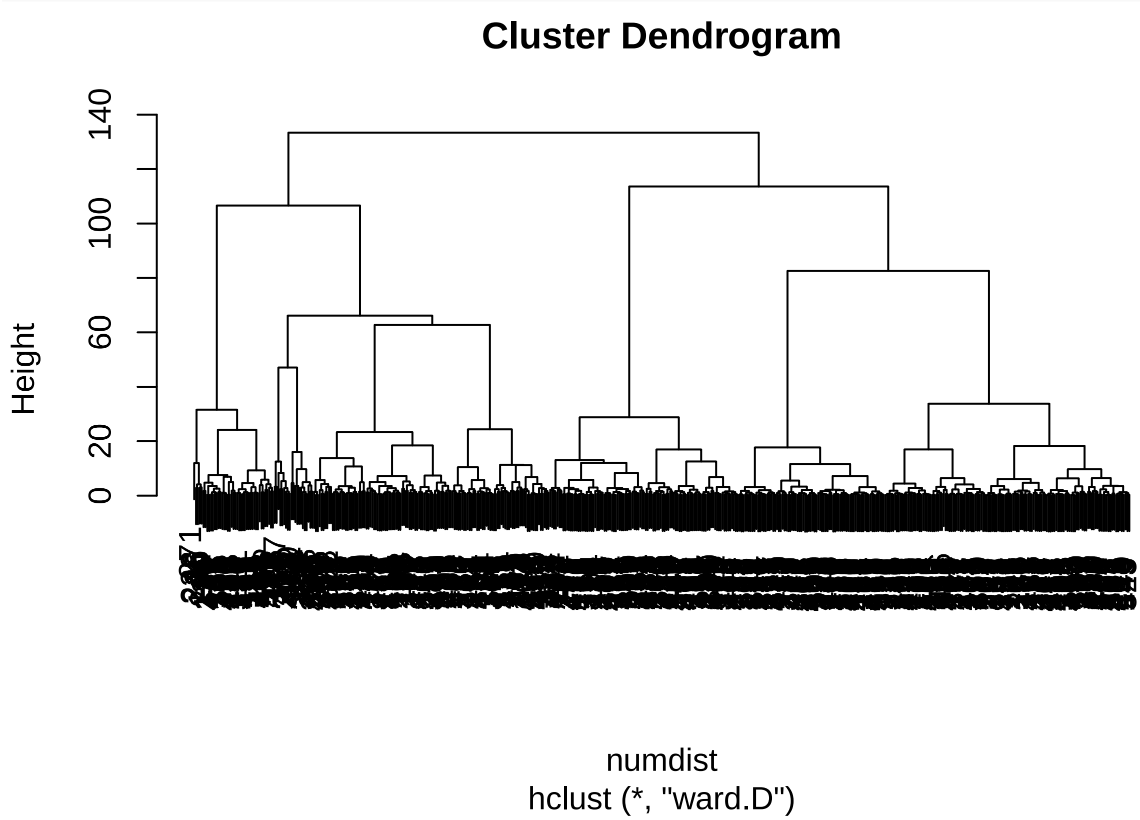
```
## 5      0.22570256     -0.22477502      -0.220774169
## 6     -0.19420836     -0.41695804       0.002500678
## 7      0.37677791     -0.05590513      -0.220774169
## 8      0.13776297     -0.62129863      -0.114813225
## 9      0.20867873      1.67441751      -0.152728501
## 10     0.16416466     -0.12807181      -0.173385222
##
## Clustering vector:
##   1   2   3   4   5   6   7   8   9  10  11  12  13  14  15  16  17  18  19  20
##   9   5  10   8   9   4   4  10   8  10   5  10   5   5   6   2   8   8   9   8
##  21  22  23  24  25  26  27  28  29  30  31  32  33  34  35  36  37  38  39  40
##  10   5   8   5   7  10   8   8   8   5  10   2   4  10   8   7  10  10   3   8
##  41  42  43  44  45  46  47  48  49  50  51  52  53  54  55  56  57  58  59  60
##   5   4   4   7   8   4   8   5   5   8   9   8   8   4   9   1   5  10   6  10
##  61  62  63  64  65  66  67  68  69  70  71  72  73  74  75  76  77  78  79  80
##   6   4   9   9   4  10   4   7  10   5   9   8  10  10   9   4   8   8   7   3
##  81  82  83  84  85  86  87  88  89  90  91  92  93  94  95  96  97  98  99 100
##   5   9   4   8   9  10   4   2   4  10   5  10   8   4  10  10   5   4   6   8
## 101 102 103 104 105 106 107 108 109 110 111 112 113 114 115 116 117 118 119 120
##   5   5   9   9   5   9  10   5  10   7   7   8   8  10   5   8   4   7   8   8
## 121 122 123 124 125 126 127 128 129 130 131 132 133 134 135 136 137 138 139 140
##   9   8  10   8   8   9   7   8   5   8   9   8   8   9   4   5   3   9   8   5
## 141 142 143 144 145 146 147 148 149 150 151 152 153 154 155 156 157 158 159 160
##   7   8   4  10   5   5   4   5   5   4  10   5   8   5   7   7   5  10   4   8
## 161 162 163 164 165 166 167 168 169 170 171 172 173 174 175 176 177 178 179 180
##   4  10   8   9   4   4   7  10   8   4   9   5   5   8   9  10   9   9   4   8
## 181 182 183 184 185 186 187 188 189 190 191 192 193 194 195 196 197 198 199 200
##   6   9   6   9   3  10   8   9   9  10   4   5   8  10  10   9  10  10   9  10
## 201 202 203 204 205 206 207 208 209 210 211 212 213 214 215 216 217 218 219 220
##   4   4   8   4   4  10  10   5   8   5  10   3   7  10  10   5   5   4   9   9
## 221 222 223 224 225 226 227 228 229 230 231 232 233 234 235 236 237 238 239 240
##   7   3   5   3   9   4   8   8   4   4   5   8   8   8   8   8   9   1   8   8
## 241 242 243 244 245 246 247 248 249 250 251 252 253 254 255 256 257 258 259 260
##   8   5   9  10   5   1   8   8   8  10  10   5   5   4   6  10   4   8   8   4
## 261 262 263 264 265 266 267 268 269 270 271 272 273 274 275 276 277 278 279 280
##   8   8   9  10  10   8  10  10   9   8   1  10   2  10   8   4   4   8   4   9
## 281 282 283 284 285 286 287 288 289 290 291 292 293 294 295 296 297 298 299 300
##   9   8   9   5   8  10   5   6   5   7   5   5   9   5   5   5   4   8   8   8
## 301 302 303 304 305 306 307 308 309 310 311 312 313 314 315 316 317 318 319 320
##  10  10   2   8  10   8   8   9   4   9   6  10   8  10  10   5   8   4   9   7
## 321 322 323 324 325 326 327 328 329 330 331 332 333 334 335 336 337 338 339 340
##  10   4   8  10   5  10   4  10   9   6   5   5   8   4   5   7   5   9  10  10
## 341 342 343 344 345 346 347 348 349 350 351 352 353 354 355 356 357 358 359 360
##   3   5  10   8   5   8  10   8   9   8   8   3   5   9  10  10   8   4  10  10
## 361 362 363 364 365 366 367 368 369 370 371 372 373 374 375 376 377 378 379 380
##   9   8   5   8   5   9   9   8   5   9   5   8   2   4  10  10  10   4  10   4
## 381 382 383 384 385 386 387 388 389 390 391 392 393 394 395 396 397 398 399 400
##   8   5   8   5   3   4   5   5   8  10   5   5   8  10   5   8   8  10  10   8
## 401 402 403 404 405 406 407 408 409 410 411 412 413 414 415 416 417 418 419 420
##  10   6   5  10   8   9   6  10   7  10  10   9   2   5  10   5   8   5   8  10
## 421 422 423 424 425 426 427 428 429 430 431 432 433 434 435 436 437 438 439 440
##   9   8   9   5   8   7   9   8   8   4   8   4   5   5   7   6   8   3   9  10
## 441 442 443 444 445 446 447 448 449 450 451 452 453 454 455 456 457 458 459 460
##  10   6   6   5   8   6  10   9   5   9   4   4   5   4  10   4   2   5   5  10
```

```
## 461 462 463 464 465 466 467 468 469 470 471 472 473 474 475 476 477 478 479 480
##   8   5   4   4   8   2  10   8   4   5   8   3  10   9   8   8   4   8   4   2
## 481 482 483 484 485 486 487 488 489 490 491 492 493 494 495 496 497 498 499 500
##   5  10   9   4   5   5   8   8   8   5   5  10  10   9   5   4   8   5  10   8
##
## Within cluster sum of squares by cluster:
##  [1]  57.42334 113.61006 145.09099 265.51365 143.53003  53.86346  71.31594
##  [8] 186.17083 197.33970 251.18168
##  (between_SS / total_SS =  62.8 %)
##
## Available components:
##
## [1] "cluster"      "centers"      "totss"        "withinss"     "tot.withinss"
## [6] "betweenss"    "size"         "iter"         "ifault"
```

```r
# Ward Hierarchical Clustering
fit <- hclust(numdist, method="ward")
```

```
## The "ward" method has been renamed to "ward.D"; note new "ward.D2"
```

```r
# display dendogram
plot(fit)
```




```r
# using the agnes function for all options in a single line code
?agnes
# Agglomerative Nesting (Hierarchical Clustering)
(agn.num <- agnes(num, metric="euclidean", stand=TRUE, method = "single"))
```

```
## Call:     agnes(x = num, metric = "euclidean", stand = TRUE, method = "single")
## Agglomerative coefficient:  0.9225946
```

```
## Order of objects:
##   [1] 1   189 64  308 421 82  85  263 367 177 171 2   148 173 353 335 462 11
##  [19] 459 382 434 458 22  388 296 403 449 49  81  387 102 291 154 30  172 470
##  [37] 57  105 145 29  116 122 348 132 393 169 270 468 350 419 372 35  40  248
##  [55] 53  120 193 77  351 93  100 445 208 418 242 217 491 245 325 210 444 498
##  [73] 252 304 500 112 431 228 300 234 240 364 306 400 488 422 428 478 395 72
##  [91] 84  489 119 239 487 152 363 495 153 160 298 368 104 283 483 220 384 23
## [109] 45  78  405 497 307 437 461 465 417 50  174 278 285 247 133 139 142 266
## [127] 28  357 313 396 227 333 233 97  216 284 24  235 475 101 115 369 392 481
## [145] 236 163 289 371 13  485 14  345 192 391 41  51  223 342 134 269 280 131
## [163] 237 113 203 258 389 47  397 20  146 70  157 5   108 317 370 294 365 433
## [181] 414 295 275 140 232 91  490 52  62  425 471 106 439 486 231 476 243 293
## [199] 354 474 249 125 282 187 17  406 310 366 344 63  429 361 332 196 412 138
## [217] 261 346 323 175 221 320 316 259 209 130 241 225 7   42  204 161 427 164
## [235] 494 281 6   165 18  362 33  327 159 378 65  218 452 191 469 229 334 358
## [253] 430 135 463 147 477 126 149 329 349 129 338 484 87  121 432 201 456 279
## [271] 178 4   230 128 27  381 383 448 3   26  197 404 398 312 37  190 176 411
## [289] 482 86  95  492 379 162 376 151 441 206 326 12  328 360 399 408 69  158
## [307] 109 31  410 48  60  198 253 195 265 359 287 8   274 455 355 493 21  123
## [325] 74  305 214 250 211 268 66  264 420 186 38  200 301 394 347 467 424 453
## [343] 447 34  58  96  267 343 377 401 302 319 90  356 168 277 286 499 297 315
## [361] 43  166 67  260 374 170 117 479 202 257 386 226 454 339 460 254 318 10
## [379] 144 390 73  92  76  276 309 188 251 219 340 114 256 324 194 207 321 15
## [397] 44  59  330 442 181 311 68  110 435 127 183 446 118 36  111 155 336 103
## [415] 167 156 426 46  124 450 54  75  136 337 215 272 89  79  409 213 55  182
## [433] 141 290 180 199 107 415 473 244 179 496 184 99  443 299 150 451 292 331
## [451] 9   39  212 385 222 185 314 19  262 61  205 416 288 407 352 436 402 380
## [469] 464 94  423 98  71  322 341 83  480 375 440 16  32  273 303 373 466 25
## [487] 143 56  238 246 137 224 438 255 80  472 457 88  413 271
## Height (summary):
##     Min.  1st Qu.   Median     Mean  3rd Qu.     Max. 
##  0.09727  0.57043  0.92079  1.30261  1.59595 15.08708 
## 
## Available components:
## [1] "order"     "height"    "ac"        "merge"     "diss"      "call"     
## [7] "method"    "order.lab" "data"
```

```r
#View(agn.num)

#  Description of cluster merging
agn.num$merge
```

```
##        [,1] [,2]
##   [1,]  -78 -405
##   [2,] -403 -449
##   [3,] -169 -270
##   [4,] -461 -465
##   [5,]  -72  -84
##   [6,]  -53 -120
##   [7,]  -35  -40
##   [8,] -294 -365
##   [9,] -148 -173
##  [10,] -296    2
##  [11,] -102 -291
##  [12,] -112 -431
```

```
##  [13,] -217 -491
##  [14,]    8 -433
##  [15,] -306 -400
##  [16,] -382 -434
##  [17,]  -64 -308
##  [18,] -239 -487
##  [19,]  -81 -387
##  [20,]   -2    9
##  [21,]  -57 -105
##  [22,] -174 -278
##  [23,] -228 -300
##  [24,] -227 -333
##  [25,]    1 -497
##  [26,] -335 -462
##  [27,]    6 -193
##  [28,]   15 -488
##  [29,] -307 -437
##  [30,] -106 -439
##  [31,] -425 -471
##  [32,]   11 -154
##  [33,] -392 -481
##  [34,] -201 -456
##  [35,] -132 -393
##  [36,]  -95 -492
##  [37,] -113 -203
##  [38,]   36 -379
##  [39,]   29    4
##  [40,]  -11 -459
##  [41,]   25   39
##  [42,] -240 -364
##  [43,]   27  -77
##  [44,]   42   28
##  [45,]  -26 -197
##  [46,]   12   23
##  [47,] -122 -348
##  [48,] -208 -418
##  [49,]   40   16
##  [50,]  -30 -172
##  [51,]  -49   19
##  [52,]   51   32
##  [53,] -164 -494
##  [54,]  -50   22
##  [55,] -139 -142
##  [56,] -210 -444
##  [57,]  -22 -388
##  [58,]   17 -421
##  [59,] -214 -250
##  [60,]   47   35
##  [61,] -369   33
##  [62,]   57   10
##  [63,] -160 -298
##  [64,] -245 -325
##  [65,]  -36 -111
##  [66,] -363 -495
```

```
##  [67,]   46 -234
##  [68,] -100 -445
##  [69,] -133   55
##  [70,]   62   52
##  [71,] -104 -283
##  [72,] -347 -467
##  [73,]  -37 -190
##  [74,] -216 -284
##  [75,] -428 -478
##  [76,] -156 -426
##  [77,]   56 -498
##  [78,]  -13 -485
##  [79,]   43 -351
##  [80,] -152   66
##  [81,]   70   50
##  [82,]    3 -468
##  [83,]   67   44
##  [84,]  -12 -328
##  [85,]    7 -248
##  [86,]   48 -242
##  [87,]   83 -422
##  [88,]  -51 -223
##  [89,]    5 -489
##  [90,]   84 -360
##  [91,]  -93   68
##  [92,] -350 -419
##  [93,]  -69 -158
##  [94,]   -1 -189
##  [95,]   60   82
##  [96,]   13   64
##  [97,]   86   96
##  [98,]   20 -353
##  [99,]   54 -285
## [100,]   87   75
## [101,]   49 -458
## [102,]   79   91
## [103,] -235 -475
## [104,] -119   18
## [105,] -116   95
## [106,]   71 -483
## [107,]   41 -417
## [108,]   73 -176
## [109,] -192 -391
## [110,]   99 -247
## [111,] -269 -280
## [112,]  105   92
## [113,]   81 -470
## [114,]  -29  112
## [115,]  -82  -85
## [116,]  -45  107
## [117,]   45 -404
## [118,]   69 -266
## [119,]  113   21
## [120,]  108 -411
```

```
## [121,]  101  119
## [122,]  -60 -198
## [123,]  106 -220
## [124,]  110  118
## [125,]  -23  116
## [126,] -399 -408
## [127,] -263 -367
## [128,]   85  102
## [129,]  114 -372
## [130,]  129  128
## [131,]   77 -252
## [132,]  125  124
## [133,]   -8 -274
## [134,]  -28 -357
## [135,]  132  134
## [136,]  121 -145
## [137,] -163 -289
## [138,]   63 -368
## [139,]  122 -253
## [140,]   97  131
## [141,]  130  140
## [142,]   90  126
## [143,]  120 -482
## [144,]  141 -304
## [145,]  136  144
## [146,] -153  138
## [147,]  133 -455
## [148,] -355 -493
## [149,]  145 -500
## [150,]  149  100
## [151,]  150 -395
## [152,]  135 -313
## [153,] -151 -441
## [154,]  151   89
## [155,]  152 -396
## [156,]  154  104
## [157,] -381 -383
## [158,]  156   80
## [159,]  143  -86
## [160,]  158  146
## [161,]  -14 -345
## [162,]  155   24
## [163,]  -70 -157
## [164,]  -59 -330
## [165,]   98   26
## [166,]   58  115
## [167,]  160  123
## [168,]  166  127
## [169,]  117 -398
## [170,]  167 -384
## [171,]   88 -342
## [172,]  171 -134
## [173,]  170  162
## [174,]  153 -206
```

```
## [175,]  169 -312
## [176,] -229 -334
## [177,]   -3  175
## [178,] -101 -115
## [179,]   37 -258
## [180,]   78  161
## [181,]  -41  172
## [182,]  173 -233
## [183,]  182  -97
## [184,]  183   74
## [185,]  177  159
## [186,]  -52  -62
## [187,]   94  168
## [188,]  184  -24
## [189,]  187 -177
## [190,] -200 -301
## [191,]  188  103
## [192,]  191  178
## [193,]  176 -358
## [194,]  192   61
## [195,]  165  194
## [196,]  -31 -410
## [197,]  189 -171
## [198,] -293 -354
## [199,]  137 -371
## [200,] -195 -265
## [201,]  195 -236
## [202,]  201  199
## [203,]  180  109
## [204,]  147  148
## [205,]  202  203
## [206,]  197  205
## [207,]  142   93
## [208,]   38 -162
## [209,]  206  181
## [210,]  208 -376
## [211,]  209  111
## [212,]  185  210
## [213,]  211 -131
## [214,]  213 -237
## [215,]  214  179
## [216,]  215 -389
## [217,]  139  200
## [218,] -310 -366
## [219,]  198 -474
## [220,]  -48  217
## [221,] -110 -435
## [222,]  212  174
## [223,]  207 -109
## [224,] -181 -311
## [225,]  216  -47
## [226,]  -38  190
## [227,] -264 -420
## [228,]  225 -397
```

```
## [229,] -218 -452
## [230,]  228  -20
## [231,]  230 -146
## [232,]  231  163
## [233,]  232   -5
## [234,]  233 -108
## [235,] -261 -346
## [236,]  234 -317
## [237,] -129 -338
## [238,]  223  196
## [239,]  -90 -356
## [240,]  222 -326
## [241,]   14 -414
## [242,]  236 -370
## [243,]  220 -359
## [244,]  226 -394
## [245,] -424 -453
## [246,]  242  241
## [247,]  -21 -123
## [248,]  -73  -92
## [249,]  239 -168
## [250,]  238  243
## [251,] -243  219
## [252,]  240  250
## [253,]  -42 -204
## [254,] -212 -385
## [255,]  186   31
## [256,]  -91 -490
## [257,]  246 -295
## [258,]  164 -442
## [259,]  -67 -260
## [260,]  257 -275
## [261,]  252 -287
## [262,]  260 -140
## [263,]  262 -232
## [264,]  261  204
## [265,] -117 -479
## [266,]  263  256
## [267,]  264  247
## [268,]  267  -74
## [269,] -257 -386
## [270,] -196 -412
## [271,]  244   72
## [272,]  -65  229
## [273,]  268 -305
## [274,]  273   59
## [275,] -125 -282
## [276,] -292 -331
## [277,]  266  255
## [278,]  -66  227
## [279,]  277   30
## [280,]  279 -486
## [281,]  280 -231
## [282,]  274 -211
```

```
## [283,]  281 -476
## [284,]  258  224
## [285,]  282 -268
## [286,] -226 -454
## [287,]  249 -277
## [288,]  285  278
## [289,]  283  251
## [290,]  275 -187
## [291,] -273 -303
## [292,]  288 -186
## [293,]  292  271
## [294,]  289 -249
## [295,] -256 -324
## [296,]  -58  -96
## [297,] -221 -320
## [298,] -127 -183
## [299,]  293  245
## [300,]  294  290
## [301,]   -6 -165
## [302,]  253 -161
## [303,]  299 -447
## [304,]  303  -34
## [305,]  -43 -166
## [306,]  235 -323
## [307,]  305  259
## [308,]  300  -17
## [309,]  308 -406
## [310,]  272 -191
## [311,] -343 -377
## [312,]  311 -401
## [313,] -267  312
## [314,]  304  296
## [315,]  314  313
## [316,]  307 -374
## [317,]  309  218
## [318,]  315 -302
## [319,] -202  269
## [320,]  318 -319
## [321,]  317 -344
## [322,]  316 -170
## [323,]  310 -469
## [324,]  320  287
## [325,]  -68  221
## [326,]  -39  254
## [327,] -339 -460
## [328,]  325  298
## [329,] -159 -378
## [330,]  323  193
## [331,] -103 -167
## [332,]  321  -63
## [333,]  284  328
## [334,]  265  319
## [335,]  332 -429
## [336,]  -44  333
```

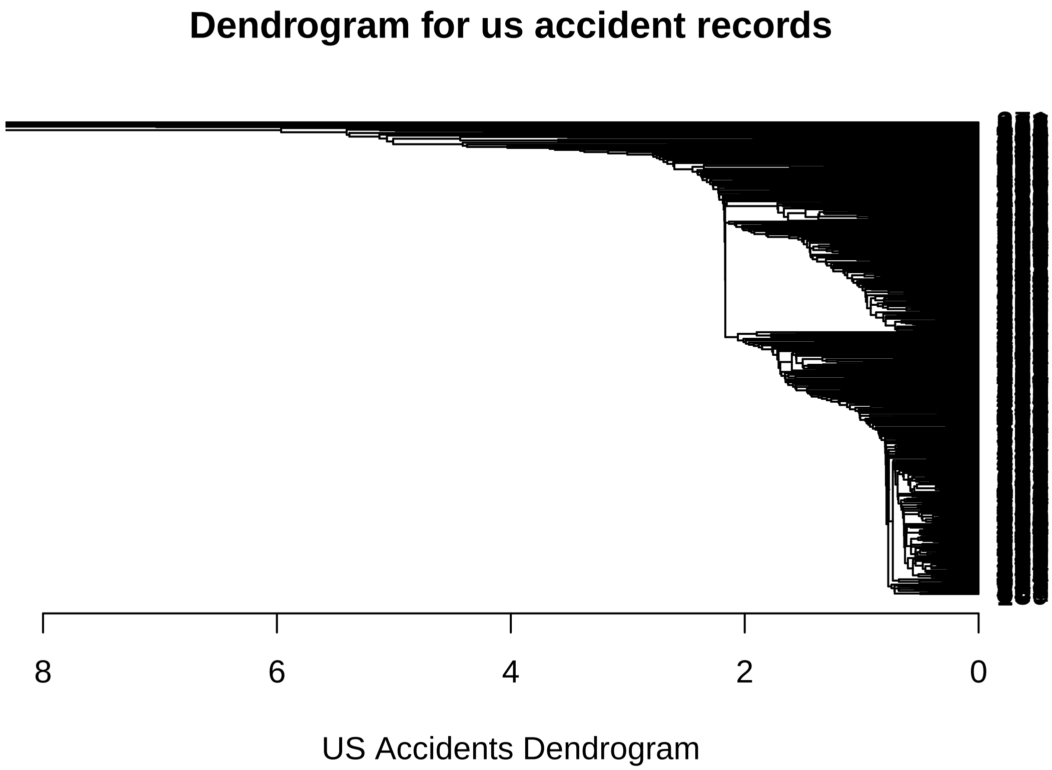
```
## [337,] -286 -499
## [338,]  324  337
## [339,]  -18 -362
## [340,]   34 -279
## [341,]  322  334
## [342,]  338 -297
## [343,]  335 -361
## [344,] -130 -241
## [345,]  343 -332
## [346,]  342 -315
## [347,]  346  341
## [348,]  345  270
## [349,]  -33 -327
## [350,]  347  286
## [351,] -254 -318
## [352,]  297 -316
## [353,]  349  329
## [354,]  348 -138
## [355,]  350  327
## [356,]  354  306
## [357,]  336 -446
## [358,]  355  351
## [359,]  339  353
## [360,]  359  330
## [361,]  358  -10
## [362,] -135 -463
## [363,]   -4 -230
## [364,]  361 -144
## [365,]  -27  157
## [366,]  356 -175
## [367,]  360 -430
## [368,]  367  362
## [369,]  368 -147
## [370,]   -7  302
## [371,]  366  352
## [372,]  371 -259
## [373,]  369 -477
## [374,]   65 -155
## [375,]  301  373
## [376,]  326 -222
## [377,]  364 -390
## [378,]  370 -427
## [379,]  372 -209
## [380,]  -15  357
## [381,]  379  344
## [382,]  381 -225
## [383,]  382  378
## [384,]  383   53
## [385,]  380 -118
## [386,]  374 -336
## [387,]  384 -281
## [388,]  387  375
## [389,]  -79 -409
## [390,]  388 -126
```

```
## [391,] -329 -349
## [392,]  385  386
## [393,]  392  331
## [394,]  390 -149
## [395,]  394  391
## [396,]  395  237
## [397,]  363 -128
## [398,]  397  365
## [399,]  389 -213
## [400,]  -76 -276
## [401,]  377  248
## [402,]  401  400
## [403,] -114  295
## [404,]  396 -484
## [405,]  404  -87
## [406,]  -99 -443
## [407,]  398 -448
## [408,]  402 -309
## [409,]  405 -121
## [410,]  -32  291
## [411,]  408 -188
## [412,]  411 -251
## [413,] -415 -473
## [414,]  409 -432
## [415,]  414  340
## [416,]  415 -178
## [417,]  412 -219
## [418,]  417 -340
## [419,]  418  403
## [420,] -215 -272
## [421,] -136 -337
## [422,] -124 -450
## [423,]  416  407
## [424,]  419 -194
## [425,]  424 -207
## [426,] -107  413
## [427,]  425 -321
## [428,]  393   76
## [429,]  -55 -182
## [430,]  423  427
## [431,]  430  428
## [432,]  431  -46
## [433,]  432  422
## [434,] -380 -464
## [435,]  433  -54
## [436,]  435  -75
## [437,]  436  421
## [438,]  437  420
## [439,]  438  -89
## [440,]  439  399
## [441,]  440  429
## [442,] -141 -290
## [443,]  441  442
## [444,]  443 -180
```

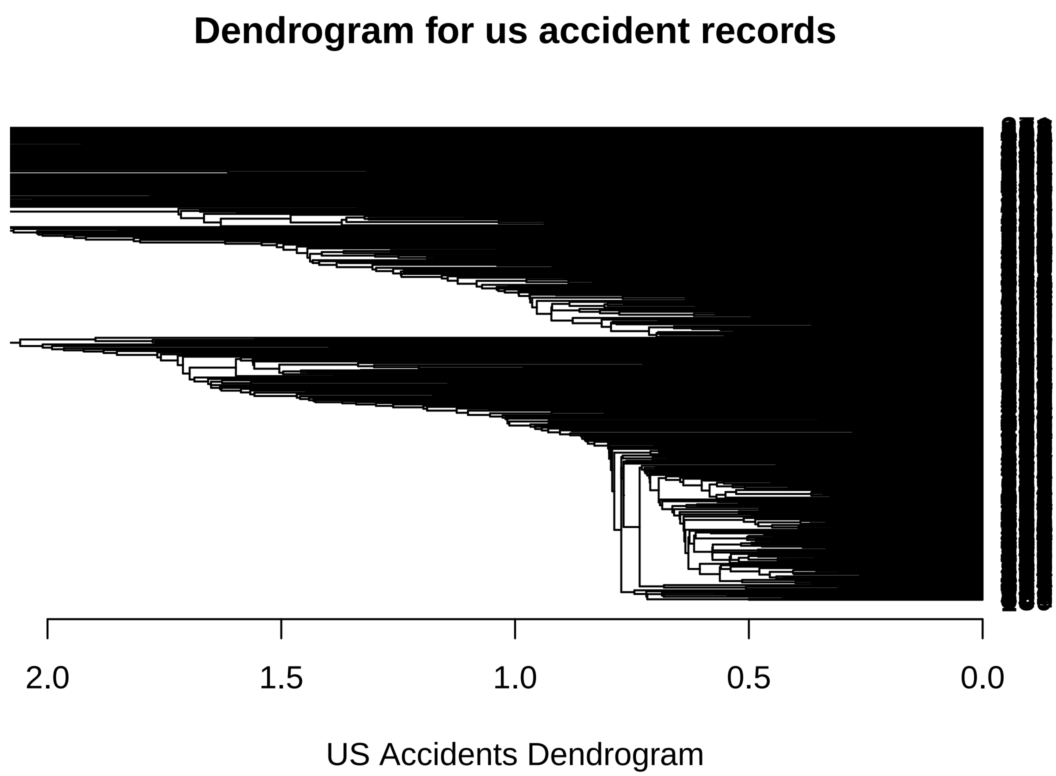
```
## [445,]  444 -199
## [446,]  426 -244
## [447,]  445  446
## [448,]  447 -179
## [449,]  376 -185
## [450,] -288 -407
## [451,]   -9  449
## [452,] -150 -451
## [453,]  448 -496
## [454,]  453 -184
## [455,]  454  406
## [456,]  455 -299
## [457,]  456  452
## [458,]  457  276
## [459,]  458  451
## [460,]  -19 -262
## [461,] -375 -440
## [462,]  459 -314
## [463,]  462  460
## [464,]  -83 -480
## [465,]  463  -61
## [466,]  465 -205
## [467,]  466 -416
## [468,]  467  450
## [469,]  468 -352
## [470,]  469 -436
## [471,]  470 -402
## [472,]  471  434
## [473,]  472  -94
## [474,]  473 -423
## [475,]  474  -98
## [476,] -238 -246
## [477,] -373 -466
## [478,]  -16  410
## [479,]  475  -71
## [480,]  479 -322
## [481,]  480 -341
## [482,]  -56  476
## [483,]  481  464
## [484,]  483  461
## [485,]  478  477
## [486,] -137 -224
## [487,]  -88 -413
## [488,]  484  485
## [489,]  488  -25
## [490,]  486 -438
## [491,]  489 -143
## [492,]  491  482
## [493,]  492  490
## [494,]  493 -255
## [495,]  -80 -472
## [496,]  494  495
## [497,]  496 -457
## [498,]  497  487
```

```
## [499,]  498 -271
```

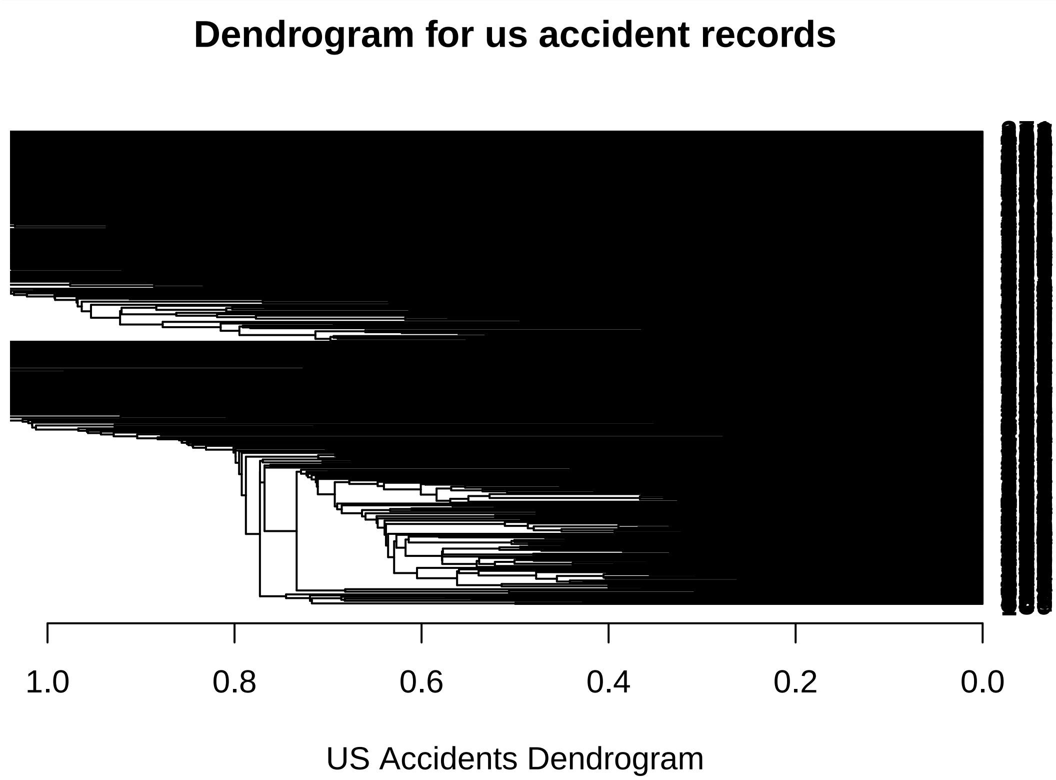
```r
#Dendogram
plot(as.dendrogram(agn.num), xlab= "US Accidents Dendrogram",xlim=c(8,0),
     horiz = TRUE,main="Dendrogram for us accident records")
```




```r
plot(as.dendrogram(agn.num), xlab= "US Accidents Dendrogram",xlim=c(2,0),
     horiz = TRUE,main="Dendrogram for us accident records")
```




```r
plot(as.dendrogram(agn.num), xlab= "US Accidents Dendrogram",xlim=c(1,0),
     horiz = TRUE,main="Dendrogram for us accident records")
```




```r
#Interactive Plots,
#plot(agn.num,ask=TRUE)
```

```r
#plot(agn.num, which.plots=2)


#K-Means Clustering


# K-means, k=2, 3, 4, 5, 6
# Centers (k's) are numbers thus, 10 random sets are chosen


# Computing the percentage of variation accounted for. Two clusters
matstd.num<- scale(num[,1:8])
(kmeans2.num <- kmeans(matstd.num,2,nstart = 10))
```

```
## K-means clustering with 2 clusters of sizes 252, 248
##
## Cluster means:
##     Severity Distance.mi. Temperature.F. Humidity... Pressure.in.
## 1  0.1486065   0.02752796     -0.6182097   0.7626184   0.09499899
## 2 -0.1510034  -0.02797196      0.6281808  -0.7749187  -0.09653123
##   Visibility.mi. Wind_Speed.mph. Precipitation.in.
## 1     -0.3542998      -0.2346050         0.2130170
## 2      0.3600143       0.2383889        -0.2164527
##
## Clustering vector:
##   1   2   3   4   5   6   7   8   9  10  11  12  13  14  15  16  17  18  19  20
##   2   2   2   1   2   1   1   2   1   2   2   1   2   2   1   1   1   1   2   1
##  21  22  23  24  25  26  27  28  29  30  31  32  33  34  35  36  37  38  39  40
##   1   2   1   2   2   2   1   1   1   2   2   1   1   1   1   2   1   2   1   1
##  41  42  43  44  45  46  47  48  49  50  51  52  53  54  55  56  57  58  59  60
##   2   1   1   2   1   1   1   2   2   1   2   1   2   1   2   2   2   1   1   2
##  61  62  63  64  65  66  67  68  69  70  71  72  73  74  75  76  77  78  79  80
##   1   1   1   2   1   1   1   2   1   2   1   1   1   1   1   1   2   1   2   2
##  81  82  83  84  85  86  87  88  89  90  91  92  93  94  95  96  97  98  99 100
##   2   2   1   1   2   2   1   1   1   1   2   1   2   1   1   1   2   1   1   2
## 101 102 103 104 105 106 107 108 109 110 111 112 113 114 115 116 117 118 119 120
##   2   2   2   2   2   2   1   2   1   2   2   2   2   2   2   1   1   2   1   2
## 121 122 123 124 125 126 127 128 129 130 131 132 133 134 135 136 137 138 139 140
##   1   1   2   2   2   2   2   1   2   1   2   1   1   2   1   2   1   1   1   2
## 141 142 143 144 145 146 147 148 149 150 151 152 153 154 155 156 157 158 159 160
##   2   1   1   2   2   2   1   2   2   1   2   2   1   2   2   2   2   1   1   1
## 161 162 163 164 165 166 167 168 169 170 171 172 173 174 175 176 177 178 179 180
##   1   1   2   2   1   1   2   1   1   1   2   2   2   1   2   2   1   1   1   1
## 181 182 183 184 185 186 187 188 189 190 191 192 193 194 195 196 197 198 199 200
##   1   2   2   1   1   2   1   2   2   2   1   2   2   1   2   2   2   2   2   1
## 201 202 203 204 205 206 207 208 209 210 211 212 213 214 215 216 217 218 219 220
##   1   1   2   1   1   1   2   2   1   2   1   2   2   2   1   2   2   1   2   2
## 221 222 223 224 225 226 227 228 229 230 231 232 233 234 235 236 237 238 239 240
##   2   2   2   1   2   1   1   1   1   1   2   1   1   1   2   1   2   2   2   2
## 241 242 243 244 245 246 247 248 249 250 251 252 253 254 255 256 257 258 259 260
##   2   2   2   2   2   2   1   1   1   2   1   2   2   1   1   2   1   2   2   1
## 261 262 263 264 265 266 267 268 269 270 271 272 273 274 275 276 277 278 279 280
##   1   2   2   1   2   1   1   2   2   1   2   1   1   2   1   1   1   1   1   2
## 281 282 283 284 285 286 287 288 289 290 291 292 293 294 295 296 297 298 299 300
```

```
##   1   1   2   2   1   1   2   1   2   2   2   2   2   2   2   2   1   1   1   1
## 301 302 303 304 305 306 307 308 309 310 311 312 313 314 315 316 317 318 319 320
##   2   1   1   1   2   1   1   2   1   2   1   2   1   1   2   2   1   1   2   2
## 321 322 323 324 325 326 327 328 329 330 331 332 333 334 335 336 337 338 339 340
##   1   1   1   2   2   1   1   1   2   1   2   2   1   1   2   2   2   2   1   2
## 341 342 343 344 345 346 347 348 349 350 351 352 353 354 355 356 357 358 359 360
##   1   2   1   1   2   1   2   1   2   1   2   2   2   2   2   1   1   1   1   1
## 361 362 363 364 365 366 367 368 369 370 371 372 373 374 375 376 377 378 379 380
##   1   1   2   1   2   2   2   1   2   2   2   2   1   1   1   1   1   1   1   1
## 381 382 383 384 385 386 387 388 389 390 391 392 393 394 395 396 397 398 399 400
##   1   2   1   2   2   1   2   2   1   1   2   2   1   2   2   1   1   2   2   1
## 401 402 403 404 405 406 407 408 409 410 411 412 413 414 415 416 417 418 419 420
##   1   1   2   2   1   2   1   2   2   2   2   2   1   2   2   2   1   2   1   1
## 421 422 423 424 425 426 427 428 429 430 431 432 433 434 435 436 437 438 439 440
##   2   1   2   2   1   2   1   2   1   1   2   1   2   2   2   1   1   1   2   1
## 441 442 443 444 445 446 447 448 449 450 451 452 453 454 455 456 457 458 459 460
##   1   1   1   2   2   1   2   1   2   2   1   1   2   1   2   1   1   2   2   1
## 461 462 463 464 465 466 467 468 469 470 471 472 473 474 475 476 477 478 479 480
##   1   2   1   1   1   1   2   1   1   2   1   2   1   2   2   1   1   2   1   1
## 481 482 483 484 485 486 487 488 489 490 491 492 493 494 495 496 497 498 499 500
##   2   1   2   1   2   2   2   1   1   2   2   1   2   2   2   1   1   2   1   1
## 
## Within cluster sum of squares by cluster:
## [1] 1824.705 1546.653
##  (between_SS / total_SS =  15.5 %)
## 
## Available components:
## 
## [1] "cluster"      "centers"      "totss"        "withinss"     "tot.withinss"
## [6] "betweenss"    "size"         "iter"         "ifault"
```

```r
perc.var.2 <- round(100*(1 - kmeans2.num$betweenss/kmeans2.num$totss),1)
names(perc.var.2) <- "Perc. 2 clus"
perc.var.2
```

```
## Perc. 2 clus
##         84.5
```

```r
# Computing the percentage of variation accounted for. Three clusters
(kmeans3.num <- kmeans(matstd.num,3,nstart = 10))
```

```
## K-means clustering with 3 clusters of sizes 10, 241, 249
## 
## Cluster means:
##     Severity Distance.mi. Temperature.F. Humidity... Pressure.in.
## 1  0.1908408  0.826887066     -0.4068876   1.1409597    0.1427323
## 2 -0.1776836 -0.030426898      0.6445745  -0.7904393   -0.1152943
## 3  0.1643106 -0.003758989     -0.6075244   0.7192220    0.1058578
##   Visibility.mi. Wind_Speed.mph. Precipitation.in.
## 1     -1.2688925       0.1186294        6.13809348
## 2      0.3609564       0.2647871       -0.21484488
## 3     -0.2983999      -0.2610441       -0.03856755
## 
## Clustering vector:
##   1   2   3   4   5   6   7   8   9  10  11  12  13  14  15  16  17  18  19  20
```

```
##   2   2   2   3   2   3   3   2   3   3   2   3   2   2   3   1   3   3   2   3
##  21  22  23  24  25  26  27  28  29  30  31  32  33  34  35  36  37  38  39  40
##   3   2   3   2   2   2   3   3   3   2   2   1   3   3   3   2   3   2   3   3
##  41  42  43  44  45  46  47  48  49  50  51  52  53  54  55  56  57  58  59  60
##   2   3   3   2   3   3   3   2   2   3   2   3   2   3   2   2   2   3   3   2
##  61  62  63  64  65  66  67  68  69  70  71  72  73  74  75  76  77  78  79  80
##   3   3   3   2   3   3   3   2   3   2   3   3   3   3   3   3   2   3   2   2
##  81  82  83  84  85  86  87  88  89  90  91  92  93  94  95  96  97  98  99 100
##   2   2   3   3   2   2   3   1   3   3   2   3   2   3   3   3   2   3   3   2
## 101 102 103 104 105 106 107 108 109 110 111 112 113 114 115 116 117 118 119 120
##   2   2   2   2   2   2   3   2   3   2   2   2   2   2   2   3   3   2   3   3
## 121 122 123 124 125 126 127 128 129 130 131 132 133 134 135 136 137 138 139 140
##   3   3   2   2   2   2   2   3   2   3   2   3   3   2   3   2   3   3   3   2
## 141 142 143 144 145 146 147 148 149 150 151 152 153 154 155 156 157 158 159 160
##   2   3   3   2   2   2   3   2   2   3   2   2   3   2   2   2   2   3   3   3
## 161 162 163 164 165 166 167 168 169 170 171 172 173 174 175 176 177 178 179 180
##   3   3   2   2   3   3   2   3   3   3   2   2   2   3   2   2   3   2   3   3
## 181 182 183 184 185 186 187 188 189 190 191 192 193 194 195 196 197 198 199 200
##   3   2   2   3   3   2   3   2   2   3   3   2   3   3   2   2   2   2   2   3
## 201 202 203 204 205 206 207 208 209 210 211 212 213 214 215 216 217 218 219 220
##   3   3   2   3   3   3   2   2   3   2   3   2   2   2   3   2   2   3   2   2
## 221 222 223 224 225 226 227 228 229 230 231 232 233 234 235 236 237 238 239 240
##   2   2   2   3   2   3   3   3   3   3   2   3   3   3   2   3   2   2   2   2
## 241 242 243 244 245 246 247 248 249 250 251 252 253 254 255 256 257 258 259 260
##   2   2   2   2   2   2   3   3   3   2   3   2   2   3   3   2   3   2   2   3
## 261 262 263 264 265 266 267 268 269 270 271 272 273 274 275 276 277 278 279 280
##   3   2   2   3   2   3   3   2   2   3   2   3   1   2   3   3   3   3   3   2
## 281 282 283 284 285 286 287 288 289 290 291 292 293 294 295 296 297 298 299 300
##   3   3   2   2   3   3   2   3   2   2   2   2   2   2   2   2   3   3   3   3
## 301 302 303 304 305 306 307 308 309 310 311 312 313 314 315 316 317 318 319 320
##   3   3   1   3   2   3   3   2   3   2   3   2   3   3   2   2   3   3   2   2
## 321 322 323 324 325 326 327 328 329 330 331 332 333 334 335 336 337 338 339 340
##   3   3   3   2   2   3   3   3   2   3   2   2   3   3   2   2   2   2   3   2
## 341 342 343 344 345 346 347 348 349 350 351 352 353 354 355 356 357 358 359 360
##   3   2   3   3   2   3   2   3   2   3   2   2   2   2   2   3   3   3   3   3
## 361 362 363 364 365 366 367 368 369 370 371 372 373 374 375 376 377 378 379 380
##   3   3   2   3   2   2   2   3   2   2   2   2   1   3   3   3   3   3   3   3
## 381 382 383 384 385 386 387 388 389 390 391 392 393 394 395 396 397 398 399 400
##   3   2   3   2   2   3   2   2   3   3   2   2   3   2   2   3   3   2   2   3
## 401 402 403 404 405 406 407 408 409 410 411 412 413 414 415 416 417 418 419 420
##   3   3   2   2   3   2   3   2   2   2   2   2   1   2   3   2   3   2   3   3
## 421 422 423 424 425 426 427 428 429 430 431 432 433 434 435 436 437 438 439 440
##   2   3   2   2   3   2   3   3   3   3   2   3   2   2   2   3   3   3   2   3
## 441 442 443 444 445 446 447 448 449 450 451 452 453 454 455 456 457 458 459 460
##   3   3   3   2   2   3   2   3   2   2   3   3   2   3   2   3   1   2   2   3
## 461 462 463 464 465 466 467 468 469 470 471 472 473 474 475 476 477 478 479 480
##   3   2   3   3   3   1   2   3   3   2   3   2   3   2   2   3   3   2   3   1
## 481 482 483 484 485 486 487 488 489 490 491 492 493 494 495 496 497 498 499 500
##   2   3   2   3   2   2   3   3   3   2   2   3   2   2   2   3   3   2   3   3
## 
## Within cluster sum of squares by cluster:
## [1]  113.6101 1519.3207 1353.0941
##  (between_SS / total_SS =  25.2 %)
## 
```

```
## Available components:
##
## [1] "cluster"      "centers"      "totss"        "withinss"     "tot.withinss"
## [6] "betweenss"    "size"         "iter"         "ifault"
```

```r
perc.var.3 <- round(100*(1 - kmeans3.num$betweenss/kmeans3.num$totss),1)

names(perc.var.3) <- "Perc. 3 clus"
perc.var.3
```

```
## Perc. 3 clus
##         74.8
```

```r
# Computing the percentage of variation accounted for. Four clusters
(kmeans4.num<- kmeans(matstd.num,4,nstart = 10))
```

```
## K-means clustering with 4 clusters of sizes 41, 10, 226, 223
##
## Cluster means:
##     Severity Distance.mi. Temperature.F. Humidity... Pressure.in.
## 1 -0.2746246  0.004734917    -0.09888672  -0.8244867   -2.9857488
## 2  0.1908408  0.826887066    -0.40688757   1.1409597    0.1427323
## 3  0.2119515  0.023318575    -0.61459811   0.7418733    0.2796961
## 4 -0.1728692 -0.061582961     0.65929329  -0.6514306    0.2590899
##   Visibility.mi. Wind_Speed.mph. Precipitation.in.
## 1      0.8790505     -0.07988788       -0.15106885
## 2     -1.2688925      0.11862939        6.13809348
## 3     -0.3308795     -0.26397385       -0.03267005
## 4      0.2306127      0.27689327       -0.21436628
##
## Clustering vector:
##   1   2   3   4   5   6   7   8   9  10  11  12  13  14  15  16  17  18  19  20
##   4   4   4   3   4   3   3   4   3   3   4   3   4   4   1   2   3   3   4   3
##  21  22  23  24  25  26  27  28  29  30  31  32  33  34  35  36  37  38  39  40
##   3   4   3   4   1   4   3   3   3   4   4   2   3   3   4   1   3   4   3   4
##  41  42  43  44  45  46  47  48  49  50  51  52  53  54  55  56  57  58  59  60
##   4   3   3   1   3   3   3   4   4   3   4   3   4   3   4   1   4   3   1   4
##  61  62  63  64  65  66  67  68  69  70  71  72  73  74  75  76  77  78  79  80
##   1   3   3   4   3   3   3   1   3   4   3   3   3   3   3   3   4   3   1   4
##  81  82  83  84  85  86  87  88  89  90  91  92  93  94  95  96  97  98  99 100
##   4   4   3   3   4   4   3   2   3   3   4   3   4   3   3   3   4   3   1   4
## 101 102 103 104 105 106 107 108 109 110 111 112 113 114 115 116 117 118 119 120
##   4   4   1   4   4   4   3   4   3   1   1   4   4   4   4   3   3   1   3   4
## 121 122 123 124 125 126 127 128 129 130 131 132 133 134 135 136 137 138 139 140
##   3   3   4   4   4   4   1   3   4   3   4   3   3   4   3   4   3   3   3   4
## 141 142 143 144 145 146 147 148 149 150 151 152 153 154 155 156 157 158 159 160
##   1   3   3   4   4   4   3   4   4   3   4   4   3   4   1   1   4   3   3   3
## 161 162 163 164 165 166 167 168 169 170 171 172 173 174 175 176 177 178 179 180
##   3   3   4   4   3   3   1   3   3   3   4   4   4   3   4   4   3   4   3   3
## 181 182 183 184 185 186 187 188 189 190 191 192 193 194 195 196 197 198 199 200
##   1   4   1   3   3   4   3   4   4   4   3   4   4   3   4   4   4   4   4   3
## 201 202 203 204 205 206 207 208 209 210 211 212 213 214 215 216 217 218 219 220
##   3   3   4   3   3   3   4   4   3   4   3   4   1   4   3   4   4   3   4   4
## 221 222 223 224 225 226 227 228 229 230 231 232 233 234 235 236 237 238 239 240
##   1   4   4   3   4   3   3   3   3   3   4   3   3   3   4   3   4   1   4   4
```

```
## 241 242 243 244 245 246 247 248 249 250 251 252 253 254 255 256 257 258 259 260
##   4   4   4   4   4   1   3   4   3   4   3   4   4   3   1   4   3   4   4   3
## 261 262 263 264 265 266 267 268 269 270 271 272 273 274 275 276 277 278 279 280
##   3   4   4   3   4   3   3   4   4   3   1   3   2   4   4   3   3   3   3   4
## 281 282 283 284 285 286 287 288 289 290 291 292 293 294 295 296 297 298 299 300
##   3   3   4   4   3   3   4   3   4   1   4   4   4   4   4   4   3   3   3   3
## 301 302 303 304 305 306 307 308 309 310 311 312 313 314 315 316 317 318 319 320
##   3   3   2   3   4   3   3   4   3   4   1   4   3   3   4   4   3   3   4   1
## 321 322 323 324 325 326 327 328 329 330 331 332 333 334 335 336 337 338 339 340
##   3   3   3   4   4   3   3   3   4   1   4   4   3   3   4   1   4   4   3   4
## 341 342 343 344 345 346 347 348 349 350 351 352 353 354 355 356 357 358 359 360
##   3   4   3   3   4   3   4   3   4   3   4   4   4   4   4   3   3   3   3   3
## 361 362 363 364 365 366 367 368 369 370 371 372 373 374 375 376 377 378 379 380
##   3   3   4   3   4   4   4   3   4   4   4   4   2   3   3   3   3   3   3   3
## 381 382 383 384 385 386 387 388 389 390 391 392 393 394 395 396 397 398 399 400
##   3   4   3   4   4   3   4   4   3   3   4   4   3   4   4   3   3   4   4   3
## 401 402 403 404 405 406 407 408 409 410 411 412 413 414 415 416 417 418 419 420
##   3   1   4   4   3   4   1   4   1   4   4   4   2   4   3   4   3   4   4   3
## 421 422 423 424 425 426 427 428 429 430 431 432 433 434 435 436 437 438 439 440
##   4   3   4   4   3   1   3   4   3   3   4   3   4   4   1   1   3   3   4   3
## 441 442 443 444 445 446 447 448 449 450 451 452 453 454 455 456 457 458 459 460
##   3   3   1   4   4   1   4   3   4   4   3   3   4   3   4   3   2   4   4   3
## 461 462 463 464 465 466 467 468 469 470 471 472 473 474 475 476 477 478 479 480
##   3   4   3   3   3   2   4   3   3   4   3   1   3   4   4   3   3   4   3   2
## 481 482 483 484 485 486 487 488 489 490 491 492 493 494 495 496 497 498 499 500
##   4   3   4   3   4   4   4   3   3   4   4   3   4   4   4   3   3   4   3   3
##
## Within cluster sum of squares by cluster:
## [1]  528.9751  113.6101 1118.6016  854.2970
##  (between_SS / total_SS =  34.5 %)
##
## Available components:
##
## [1] "cluster"      "centers"      "totss"        "withinss"     "tot.withinss"
## [6] "betweenss"    "size"         "iter"         "ifault"
```

```r
perc.var.4 <- round(100*(1 - kmeans4.num$betweenss/kmeans4.num$totss),1)
names(perc.var.4) <- "Perc. 4 clus"
perc.var.4
```

```
## Perc. 4 clus
##         65.5
```

```r
# Computing the percentage of variation accounted for. Five clusters
(kmeans5.num <- kmeans(matstd.num,5,nstart = 10))
```

```
## K-means clustering with 5 clusters of sizes 37, 158, 16, 169, 120
##
## Cluster means:
##      Severity Distance.mi. Temperature.F. Humidity... Pressure.in.
## 1 -0.24559346  -0.21842156     -0.2094344  -0.7652081  -3.09931925
## 2 -0.57800420  -0.16871146     -0.5525952   0.7239109   0.25129884
## 3  0.03027763   2.91283925     -0.3041773   0.7462929  -0.04163422
## 4 -0.43458194  -0.08902735      0.7309719  -0.8053124   0.22195759
## 5  1.44476274   0.02648502     -0.1967358   0.3174324   0.31770759
```

```
##   Visibility.mi. Wind_Speed.mph. Precipitation.in.
## 1      0.9466474     -0.03995983       -0.14353314
## 2     -0.2441005     -0.32441259       -0.05798137
## 3     -0.9723333      0.06162912        3.97679296
## 4      0.2277958      0.37654018       -0.21231879
## 5     -0.1616520     -0.09904711       -0.11062524
##
## Clustering vector:
##   1   2   3   4   5   6   7   8   9  10  11  12  13  14  15  16  17  18  19  20
##   4   4   5   2   4   2   2   5   2   5   4   5   4   4   1   3   2   2   4   2
##  21  22  23  24  25  26  27  28  29  30  31  32  33  34  35  36  37  38  39  40
##   5   4   2   4   1   5   2   2   2   4   5   3   2   5   2   1   5   5   2   2
##  41  42  43  44  45  46  47  48  49  50  51  52  53  54  55  56  57  58  59  60
##   4   2   5   1   2   5   2   4   4   2   4   2   2   2   4   1   4   5   1   5
##  61  62  63  64  65  66  67  68  69  70  71  72  73  74  75  76  77  78  79  80
##   1   2   2   4   2   5   5   1   5   4   2   2   5   5   2   5   4   2   1   3
##  81  82  83  84  85  86  87  88  89  90  91  92  93  94  95  96  97  98  99 100
##   4   4   5   2   4   5   2   3   2   5   4   5   4   2   5   5   4   5   1   4
## 101 102 103 104 105 106 107 108 109 110 111 112 113 114 115 116 117 118 119 120
##   4   4   1   4   4   4   5   4   5   1   1   2   4   5   4   2   5   1   2   2
## 121 122 123 124 125 126 127 128 129 130 131 132 133 134 135 136 137 138 139 140
##   2   2   5   4   4   4   1   2   4   2   4   2   2   4   2   4   3   2   2   4
## 141 142 143 144 145 146 147 148 149 150 151 152 153 154 155 156 157 158 159 160
##   1   2   3   5   4   4   2   4   4   5   5   4   2   4   1   1   4   5   2   2
## 161 162 163 164 165 166 167 168 169 170 171 172 173 174 175 176 177 178 179 180
##   2   5   4   4   2   5   1   5   2   5   4   4   4   2   4   5   2   4   2   2
## 181 182 183 184 185 186 187 188 189 190 191 192 193 194 195 196 197 198 199 200
##   1   4   1   2   5   4   2   4   4   5   2   4   2   5   5   4   5   4   4   5
## 201 202 203 204 205 206 207 208 209 210 211 212 213 214 215 216 217 218 219 220
##   2   5   4   2   2   5   4   4   2   4   5   4   1   5   5   4   4   2   5   4
## 221 222 223 224 225 226 227 228 229 230 231 232 233 234 235 236 237 238 239 240
##   4   4   4   3   4   5   2   2   2   2   4   2   2   2   4   2   4   1   2   2
## 241 242 243 244 245 246 247 248 249 250 251 252 253 254 255 256 257 258 259 260
##   4   4   4   5   4   1   2   2   2   5   5   4   4   5   1   4   5   2   4   5
## 261 262 263 264 265 266 267 268 269 270 271 272 273 274 275 276 277 278 279 280
##   2   2   4   5   5   2   5   4   4   2   1   5   3   4   2   5   5   2   2   4
## 281 282 283 284 285 286 287 288 289 290 291 292 293 294 295 296 297 298 299 300
##   2   2   4   4   2   5   4   2   4   1   4   4   4   4   4   4   5   2   2   2
## 301 302 303 304 305 306 307 308 309 310 311 312 313 314 315 316 317 318 319 320
##   5   5   3   2   4   2   2   4   5   4   1   5   2   5   5   4   2   5   4   4
## 321 322 323 324 325 326 327 328 329 330 331 332 333 334 335 336 337 338 339 340
##   5   2   2   5   4   5   2   5   4   2   4   4   2   2   4   1   4   4   5   5
## 341 342 343 344 345 346 347 348 349 350 351 352 353 354 355 356 357 358 359 360
##   5   4   5   2   4   2   5   2   4   2   4   4   4   4   5   5   2   2   5   5
## 361 362 363 364 365 366 367 368 369 370 371 372 373 374 375 376 377 378 379 380
##   2   2   4   2   4   4   4   2   4   4   4   2   3   5   5   5   5   2   5   5
## 381 382 383 384 385 386 387 388 389 390 391 392 393 394 395 396 397 398 399 400
##   2   4   2   4   4   5   4   4   2   5   4   4   2   5   4   2   2   5   5   2
## 401 402 403 404 405 406 407 408 409 410 411 412 413 414 415 416 417 418 419 420
##   5   1   4   5   2   4   1   5   1   5   5   4   3   4   5   4   2   4   2   5
## 421 422 423 424 425 426 427 428 429 430 431 432 433 434 435 436 437 438 439 440
##   4   2   4   4   2   1   2   2   2   2   2   2   4   4   1   1   2   3   4   5
## 441 442 443 444 445 446 447 448 449 450 451 452 453 454 455 456 457 458 459 460
##   5   2   1   4   4   1   5   2   4   4   5   2   4   5   4   2   3   4   4   5
```

```
## 461 462 463 464 465 466 467 468 469 470 471 472 473 474 475 476 477 478 479 480
##   2   4   2   5   2   3   5   2   2   4   2   3   5   4   2   2   2   2   5   3
## 481 482 483 484 485 486 487 488 489 490 491 492 493 494 495 496 497 498 499 500
##   4   5   4   2   4   4   2   2   2   4   4   5   5   4   4   2   2   4   5   2
##
## Within cluster sum of squares by cluster:
## [1] 416.0750 507.6989 452.8908 550.5556 441.7780
##  (between_SS / total_SS =  40.7 %)
##
## Available components:
##
## [1] "cluster"      "centers"      "totss"        "withinss"     "tot.withinss"
## [6] "betweenss"    "size"         "iter"         "ifault"
```

```
perc.var.5 <- round(100*(1 - kmeans5.num$betweenss/kmeans5.num$totss),1)
names(perc.var.5) <- "Perc. 5 clus"
perc.var.5
```

```
## Perc. 5 clus
##         59.3
```

```
# Computing the percentage of variation accounted for. Six clusters
(kmeans6.num <- kmeans(matstd.num,6,nstart = 10))
```

```
## K-means clustering with 6 clusters of sizes 6, 39, 107, 172, 166, 10
##
## Cluster means:
##     Severity Distance.mi. Temperature.F.  Humidity... Pressure.in.
## 1  0.3743416   6.70831691     -0.3958477  0.002130805   -0.3831819
## 2 -0.2608534  -0.22001382     -0.1216707 -0.829591959   -3.0212702
## 3  1.4461919   0.01938662      0.2930756 -0.100668833    0.2974282
## 4 -0.6178428  -0.11510118      0.6152846 -0.618210029    0.2721188
## 5 -0.2557514  -0.13382650     -0.7590297  0.831538521    0.2413979
## 6  0.1908408   0.82688707     -0.4068876  1.140959724    0.1427323
##   Visibility.mi. Wind_Speed.mph. Precipitation.in.
## 1    -0.09061981     -0.06205071        0.07692563
## 2     0.91111568     -0.10506978       -0.14749422
## 3     0.17411628     -0.01951409       -0.17737121
## 4     0.22824898      0.31311561       -0.21246627
## 5    -0.48307310     -0.29207316       -0.00341745
## 6    -1.26889254      0.11862939        6.13809348
##
## Clustering vector:
##   1   2   3   4   5   6   7   8   9  10  11  12  13  14  15  16  17  18  19  20
##   4   4   3   5   4   5   5   3   5   3   4   3   4   4   2   6   5   5   4   5
##  21  22  23  24  25  26  27  28  29  30  31  32  33  34  35  36  37  38  39  40
##   3   4   5   4   2   3   5   5   5   4   3   6   5   3   4   2   3   3   5   4
##  41  42  43  44  45  46  47  48  49  50  51  52  53  54  55  56  57  58  59  60
##   4   5   5   2   5   5   5   3   4   5   4   5   4   5   4   2   4   3   2   3
##  61  62  63  64  65  66  67  68  69  70  71  72  73  74  75  76  77  78  79  80
##   2   5   4   4   5   3   5   2   3   4   5   5   3   3   5   5   4   5   2   1
##  81  82  83  84  85  86  87  88  89  90  91  92  93  94  95  96  97  98  99 100
##   4   4   5   5   4   3   5   6   5   3   4   3   4   5   3   3   4   5   2   4
## 101 102 103 104 105 106 107 108 109 110 111 112 113 114 115 116 117 118 119 120
##   4   4   2   4   4   4   3   4   3   2   2   4   4   3   4   5   5   2   5   4
```

```
## 121 122 123 124 125 126 127 128 129 130 131 132 133 134 135 136 137 138 139 140
##   5   5   3   4   4   4   2   5   4   5   4   5   5   4   5   4   1   5   5   4
## 141 142 143 144 145 146 147 148 149 150 151 152 153 154 155 156 157 158 159 160
##   2   5   5   3   4   4   5   4   4   5   3   4   5   4   2   2   4   3   5   5
## 161 162 163 164 165 166 167 168 169 170 171 172 173 174 175 176 177 178 179 180
##   5   3   4   4   5   5   2   3   5   5   4   4   4   5   4   3   4   4   5   5
## 181 182 183 184 185 186 187 188 189 190 191 192 193 194 195 196 197 198 199 200
##   2   4   2   5   3   3   5   3   4   3   5   4   4   3   3   4   3   3   4   3
## 201 202 203 204 205 206 207 208 209 210 211 212 213 214 215 216 217 218 219 220
##   5   5   4   5   5   3   3   4   5   4   3   4   2   3   3   4   4   5   3   4
## 221 222 223 224 225 226 227 228 229 230 231 232 233 234 235 236 237 238 239 240
##   2   4   4   1   4   5   5   5   5   5   4   5   5   5   4   5   4   2   4   4
## 241 242 243 244 245 246 247 248 249 250 251 252 253 254 255 256 257 258 259 260
##   4   4   4   3   4   2   5   4   5   3   3   4   3   5   2   3   5   4   4   5
## 261 262 263 264 265 266 267 268 269 270 271 272 273 274 275 276 277 278 279 280
##   5   4   4   3   3   5   3   3   4   5   2   3   6   3   4   5   3   5   5   4
## 281 282 283 284 285 286 287 288 289 290 291 292 293 294 295 296 297 298 299 300
##   5   5   4   4   5   3   3   5   4   2   4   4   4   4   4   4   5   5   5   5
## 301 302 303 304 305 306 307 308 309 310 311 312 313 314 315 316 317 318 319 320
##   3   5   6   5   3   5   5   4   5   4   2   3   5   3   3   4   5   5   3   2
## 321 322 323 324 325 326 327 328 329 330 331 332 333 334 335 336 337 338 339 340
##   3   5   5   3   4   3   5   3   4   5   4   4   5   5   4   2   4   4   3   3
## 341 342 343 344 345 346 347 348 349 350 351 352 353 354 355 356 357 358 359 360
##   1   4   3   5   4   5   3   5   4   5   4   3   4   4   3   3   5   5   3   3
## 361 362 363 364 365 366 367 368 369 370 371 372 373 374 375 376 377 378 379 380
##   5   5   4   5   4   4   4   5   4   4   4   4   6   5   3   3   3   5   3   5
## 381 382 383 384 385 386 387 388 389 390 391 392 393 394 395 396 397 398 399 400
##   5   4   5   4   4   5   4   4   5   5   4   4   5   3   4   5   5   3   3   5
## 401 402 403 404 405 406 407 408 409 410 411 412 413 414 415 416 417 418 419 420
##   3   2   4   3   5   4   2   3   2   3   3   4   6   4   3   4   5   4   4   3
## 421 422 423 424 425 426 427 428 429 430 431 432 433 434 435 436 437 438 439 440
##   4   5   4   3   5   2   5   4   5   5   4   5   4   4   2   2   5   1   4   3
## 441 442 443 444 445 446 447 448 449 450 451 452 453 454 455 456 457 458 459 460
##   3   5   2   4   4   2   3   5   4   4   5   5   3   5   3   5   6   4   4   3
## 461 462 463 464 465 466 467 468 469 470 471 472 473 474 475 476 477 478 479 480
##   5   4   5   5   5   6   3   5   5   4   5   1   3   4   4   5   5   4   5   6
## 481 482 483 484 485 486 487 488 489 490 491 492 493 494 495 496 497 498 499 500
##   4   3   4   5   4   4   4   5   5   4   4   3   3   4   4   5   5   4   3   5
##
## Within cluster sum of squares by cluster:
## [1]  83.39929 433.56834 321.26540 484.60073 663.93247 113.61006
##  (between_SS / total_SS =  47.4 %)
##
## Available components:
##
## [1] "cluster"      "centers"      "totss"        "withinss"     "tot.withinss"
## [6] "betweenss"    "size"         "iter"         "ifault"
```

```r
perc.var.6 <- round(100*(1 - kmeans6.num$betweenss/kmeans6.num$totss),1)
names(perc.var.6) <- "Perc. 6 clus"
perc.var.6
```

```
## Perc. 6 clus
##         52.6
```

```r
# Computing the percentage of variation accounted for. Seven clusters
(kmeans7.num <- kmeans(matstd.num,7,nstart = 10))
```

```
## K-means clustering with 7 clusters of sizes 149, 115, 12, 10, 114, 64, 36
##
## Cluster means:
##     Severity Distance.mi. Temperature.F. Humidity... Pressure.in.
## 1 -0.5801087  -0.18488754    -0.51774848  0.71735931   0.26045732
## 2  1.4354547  -0.03825598    -0.18625016  0.35355089   0.32930384
## 3  0.2214243   4.77380390     0.08429905 -0.03386252   0.02608219
## 4  0.1908408   0.82688707    -0.40688757  1.14095972   0.14273235
## 5 -0.3661001  -0.15176808     0.86834199 -0.97347784   0.20572274
## 6 -0.5144903  -0.13288908     0.14990304 -0.30313761   0.17807339
## 7 -0.2373277  -0.21667704    -0.19345008 -0.78252376  -3.14631893
##   Visibility.mi. Wind_Speed.mph. Precipitation.in.
## 1    -0.28630855     -0.44910686       -0.05534099
## 2    -0.17746808     -0.10177788       -0.10583616
## 3     0.06005446     -0.06922055       -0.07192427
## 4    -1.26889254      0.11862939        6.13809348
## 5     0.23112072     -0.25450445       -0.22077417
## 6     0.21745525      1.71383032       -0.15379171
## 7     0.96589378     -0.06683060       -0.14138756
##
## Clustering vector:
##   1   2   3   4   5   6   7   8   9  10  11  12  13  14  15  16  17  18  19  20
##   6   5   2   1   6   1   1   2   1   2   5   2   5   5   7   4   1   1   6   1
##  21  22  23  24  25  26  27  28  29  30  31  32  33  34  35  36  37  38  39  40
##   2   5   1   5   7   2   1   1   1   5   2   4   1   2   1   7   2   2   3   1
##  41  42  43  44  45  46  47  48  49  50  51  52  53  54  55  56  57  58  59  60
##   5   1   2   7   1   2   1   5   5   1   6   1   1   1   6   7   5   2   7   5
##  61  62  63  64  65  66  67  68  69  70  71  72  73  74  75  76  77  78  79  80
##   7   1   6   6   1   2   2   7   2   5   6   1   2   2   6   2   5   1   7   3
##  81  82  83  84  85  86  87  88  89  90  91  92  93  94  95  96  97  98  99 100
##   5   6   2   1   6   2   1   4   1   2   5   2   5   1   2   2   5   2   7   5
## 101 102 103 104 105 106 107 108 109 110 111 112 113 114 115 116 117 118 119 120
##   5   5   6   6   5   6   2   5   2   7   7   1   5   2   5   1   2   7   1   1
## 121 122 123 124 125 126 127 128 129 130 131 132 133 134 135 136 137 138 139 140
##   6   1   2   5   1   6   7   1   5   1   6   1   1   6   1   5   3   6   1   5
## 141 142 143 144 145 146 147 148 149 150 151 152 153 154 155 156 157 158 159 160
##   7   1   1   2   5   5   1   5   5   2   2   5   1   5   7   7   5   2   1   1
## 161 162 163 164 165 166 167 168 169 170 171 172 173 174 175 176 177 178 179 180
##   1   2   5   6   1   2   7   2   1   2   6   5   5   1   6   2   6   6   1   1
## 181 182 183 184 185 186 187 188 189 190 191 192 193 194 195 196 197 198 199 200
##   7   6   7   1   3   6   1   6   6   2   1   5   1   2   2   6   2   5   6   2
## 201 202 203 204 205 206 207 208 209 210 211 212 213 214 215 216 217 218 219 220
##   1   2   5   1   1   2   5   5   1   5   2   3   7   2   2   5   5   1   6   6
## 221 222 223 224 225 226 227 228 229 230 231 232 233 234 235 236 237 238 239 240
##   5   3   5   3   6   2   1   1   1   1   5   1   1   1   5   1   5   7   1   1
## 241 242 243 244 245 246 247 248 249 250 251 252 253 254 255 256 257 258 259 260
##   1   5   6   2   5   7   1   1   1   2   2   5   5   2   7   6   2   1   5   2
## 261 262 263 264 265 266 267 268 269 270 271 272 273 274 275 276 277 278 279 280
##   1   1   6   2   2   1   2   6   6   1   7   2   4   2   1   2   2   1   1   6
## 281 282 283 284 285 286 287 288 289 290 291 292 293 294 295 296 297 298 299 300
##   6   1   6   5   1   2   5   1   5   7   5   5   6   5   5   5   2   1   1   1
```

```
## 301 302 303 304 305 306 307 308 309 310 311 312 313 314 315 316 317 318 319 320
##   2   2   4   1   5   1   1   6   2   6   7   2   1   2   2   5   1   2   6   5
## 321 322 323 324 325 326 327 328 329 330 331 332 333 334 335 336 337 338 339 340
##   2   1   1   2   5   2   1   2   6   1   5   5   1   1   5   7   5   6   2   5
## 341 342 343 344 345 346 347 348 349 350 351 352 353 354 355 356 357 358 359 360
##   3   5   2   1   5   1   2   1   6   1   5   3   5   6   5   2   1   1   2   2
## 361 362 363 364 365 366 367 368 369 370 371 372 373 374 375 376 377 378 379 380
##   1   1   5   1   5   6   6   1   5   6   5   1   4   2   2   2   2   1   2   2
## 381 382 383 384 385 386 387 388 389 390 391 392 393 394 395 396 397 398 399 400
##   1   5   1   5   3   2   5   5   1   2   5   5   1   2   5   1   1   2   2   1
## 401 402 403 404 405 406 407 408 409 410 411 412 413 414 415 416 417 418 419 420
##   2   7   5   2   1   6   7   2   7   2   2   6   4   5   2   5   1   5   1   2
## 421 422 423 424 425 426 427 428 429 430 431 432 433 434 435 436 437 438 439 440
##   6   1   6   5   1   7   6   1   1   1   1   1   5   5   7   7   1   3   6   2
## 441 442 443 444 445 446 447 448 449 450 451 452 453 454 455 456 457 458 459 460
##   2   1   7   5   5   7   2   6   5   6   2   1   5   2   5   1   4   5   5   2
## 461 462 463 464 465 466 467 468 469 470 471 472 473 474 475 476 477 478 479 480
##   1   5   1   2   1   4   2   1   1   5   1   3   2   6   5   1   1   1   2   4
## 481 482 483 484 485 486 487 488 489 490 491 492 493 494 495 496 497 498 499 500
##   5   2   6   6   5   5   1   1   1   5   5   2   2   6   5   1   1   5   2   1
##
## Within cluster sum of squares by cluster:
## [1] 435.1724 393.6987 145.0910 113.6101 218.7422 213.0439 410.8366
##  (between_SS / total_SS =  51.6 %)
##
## Available components:
##
## [1] "cluster"      "centers"      "totss"        "withinss"     "tot.withinss"
## [6] "betweenss"    "size"         "iter"         "ifault"
```

```
perc.var.7 <- round(100*(1 - kmeans7.num$betweenss/kmeans7.num$totss),1)
names(perc.var.7) <- "Perc. 7 clus"
perc.var.7
```

```
## Perc. 7 clus
##         48.4
```

```
# Computing the percentage of variation accounted for. nine clusters
(kmeans9.num <- kmeans(matstd.num,9,nstart = 10))
```

```
## K-means clustering with 9 clusters of sizes 124, 10, 6, 34, 51, 138, 65, 68, 4
##
## Cluster means:
##     Severity Distance.mi. Temperature.F.  Humidity... Pressure.in.
## 1 -0.6023561  -0.08516217      0.8282891 -0.831521130    0.2257714
## 2  0.1908408   0.82688707     -0.4068876  1.140959724    0.1427323
## 3  0.3743416   6.70831691     -0.3958477  0.002130805   -0.3831819
## 4 -0.3272790  -0.21259279     -0.2576174 -0.657605737   -3.0097296
## 5 -0.5791428  -0.07604851     -0.4314103 -0.061386834    0.1601000
## 6 -0.5697566  -0.19017656     -0.4453709  0.739837550    0.2981513
## 7  1.4047690   0.02233057     -0.6647676  0.856224807    0.3343610
## 8  1.4267725  -0.01345174      0.5897004 -0.534381518    0.2853177
## 9  0.3743416  -0.28610927     -0.2328782 -1.059672391   -3.8095333
##   Visibility.mi. Wind_Speed.mph. Precipitation.in.
## 1     0.22878186     -0.07870325       -0.21213127
```

```
## 2   -1.26889254     0.11862939      6.13809348
## 3   -0.09061981    -0.06205071      0.07692563
## 4    0.11893137    -0.13037512     -0.11570365
## 5    0.03447782     1.86638459     -0.11570365
## 6   -0.24831306    -0.54486185     -0.06286384
## 7   -0.47343178    -0.12836626     -0.01742231
## 8    0.21579340     0.01639583     -0.22077417
## 9    7.35699377     0.15304465     -0.22077417
##
## Clustering vector:
##   1   2   3   4   5   6   7   8   9  10  11  12  13  14  15  16  17  18  19  20
##   1   1   8   6   5   6   6   8   6   8   1   8   1   1   4   2   6   6   1   6
##  21  22  23  24  25  26  27  28  29  30  31  32  33  34  35  36  37  38  39  40
##   8   1   6   1   4   8   6   6   6   1   8   2   6   7   6   4   8   8   6   6
##  41  42  43  44  45  46  47  48  49  50  51  52  53  54  55  56  57  58  59  60
##   1   6   7   4   6   7   6   8   1   6   1   6   6   6   5   9   1   7   4   8
##  61  62  63  64  65  66  67  68  69  70  71  72  73  74  75  76  77  78  79  80
##   4   6   5   5   6   7   7   4   7   1   5   6   8   7   5   7   1   6   4   3
##  81  82  83  84  85  86  87  88  89  90  91  92  93  94  95  96  97  98  99 100
##   1   5   7   6   5   8   6   2   6   7   1   8   1   6   7   7   1   7   4   1
## 101 102 103 104 105 106 107 108 109 110 111 112 113 114 115 116 117 118 119 120
##   1   1   5   1   1   1   8   1   8   4   4   6   1   8   1   6   7   4   6   6
## 121 122 123 124 125 126 127 128 129 130 131 132 133 134 135 136 137 138 139 140
##   5   6   8   1   1   5   4   6   1   6   5   6   6   1   6   1   3   5   6   1
## 141 142 143 144 145 146 147 148 149 150 151 152 153 154 155 156 157 158 159 160
##   4   6   6   8   1   1   6   1   1   7   8   1   6   1   4   4   1   7   6   6
## 161 162 163 164 165 166 167 168 169 170 171 172 173 174 175 176 177 178 179 180
##   5   7   1   5   6   7   4   7   6   7   5   1   1   6   5   8   5   5   6   6
## 181 182 183 184 185 186 187 188 189 190 191 192 193 194 195 196 197 198 199 200
##   4   5   4   5   7   8   6   5   1   8   6   1   6   7   8   5   8   8   5   8
## 201 202 203 204 205 206 207 208 209 210 211 212 213 214 215 216 217 218 219 220
##   5   7   1   6   6   8   8   1   6   1   7   1   4   8   7   1   1   6   5   5
## 221 222 223 224 225 226 227 228 229 230 231 232 233 234 235 236 237 238 239 240
##   1   1   1   3   5   7   6   6   6   6   1   6   6   6   1   6   1   9   6   6
## 241 242 243 244 245 246 247 248 249 250 251 252 253 254 255 256 257 258 259 260
##   1   1   5   8   1   9   6   6   6   8   7   1   8   7   4   8   7   6   1   7
## 261 262 263 264 265 266 267 268 269 270 271 272 273 274 275 276 277 278 279 280
##   6   1   1   7   8   6   8   8   1   6   9   8   2   8   6   7   7   6   5   1
## 281 282 283 284 285 286 287 288 289 290 291 292 293 294 295 296 297 298 299 300
##   5   6   1   1   6   7   8   4   1   4   1   1   5   1   1   1   7   6   6   6
## 301 302 303 304 305 306 307 308 309 310 311 312 313 314 315 316 317 318 319 320
##   8   7   2   6   8   6   6   5   7   5   4   8   6   7   8   1   6   7   8   1
## 321 322 323 324 325 326 327 328 329 330 331 332 333 334 335 336 337 338 339 340
##   7   5   6   8   1   7   6   7   5   4   1   1   6   6   1   4   1   1   7   8
## 341 342 343 344 345 346 347 348 349 350 351 352 353 354 355 356 357 358 359 360
##   3   1   8   6   1   6   8   6   5   6   1   8   1   5   8   7   6   6   8   7
## 361 362 363 364 365 366 367 368 369 370 371 372 373 374 375 376 377 378 379 380
##   5   6   1   6   1   5   1   6   1   1   1   6   2   7   7   7   7   6   7   7
## 381 382 383 384 385 386 387 388 389 390 391 392 393 394 395 396 397 398 399 400
##   6   1   6   1   1   7   1   1   6   7   1   1   6   8   1   6   6   8   8   6
## 401 402 403 404 405 406 407 408 409 410 411 412 413 414 415 416 417 418 419 420
##   7   4   1   8   6   5   4   8   4   8   8   5   2   1   8   1   6   1   6   7
## 421 422 423 424 425 426 427 428 429 430 431 432 433 434 435 436 437 438 439 440
##   5   6   5   8   6   4   5   6   6   6   6   6   1   1   4   4   6   3   1   7
```

```
## 441 442 443 444 445 446 447 448 449 450 451 452 453 454 455 456 457 458 459 460
##   8   6   4   1   1   4   8   5   1   1   7   6   8   7   8   5   2   1   1   7
## 461 462 463 464 465 466 467 468 469 470 471 472 473 474 475 476 477 478 479 480
##   6   1   6   7   6   2   8   6   6   1   6   3   8   5   1   6   6   6   7   2
## 481 482 483 484 485 486 487 488 489 490 491 492 493 494 495 496 497 498 499 500
##   1   7   5   5   1   1   6   6   6   1   1   7   8   5   1   6   6   1   7   6
##
## Within cluster sum of squares by cluster:
## [1] 250.36466 113.61006  83.39929 179.46177 199.22258 353.07040 224.68787
## [8] 152.04291  57.42334
##  (between_SS / total_SS =  59.6 %)
##
## Available components:
##
## [1] "cluster"      "centers"      "totss"        "withinss"     "tot.withinss"
## [6] "betweenss"    "size"         "iter"         "ifault"
```

```
perc.var.9 <- round(100*(1 - kmeans9.num$betweenss/kmeans9.num$totss),1)
names(perc.var.9) <- "Perc. 9 clus"
perc.var.9
```

```
## Perc. 9 clus
##         40.4
```

```
# Computing the percentage of variation accounted for. ten clusters
(kmeans10.num <- kmeans(matstd.num,10,nstart = 10))
```

```
## K-means clustering with 10 clusters of sizes 96, 35, 8, 96, 55, 4, 51, 7, 115, 33
##
## Cluster means:
##       Severity Distance.mi. Temperature.F.  Humidity... Pressure.in.
## 1  -0.46670363  -0.08373242     0.87768023 -1.062821807   0.19339608
## 2  -0.33344712  -0.21013380    -0.29371012 -0.608316049  -2.97569649
## 3   0.37434157   6.31879532    -0.47930387  0.219890444  -0.26850890
## 4   1.46387740   0.06231273     0.29775845 -0.002818278   0.31919043
## 5   0.19084080  -0.10589224    -1.07687868  1.159414485   0.24698057
## 6   0.37434157  -0.28610927    -0.23287818 -1.059672391  -3.80953332
## 7  -0.61512337  -0.15293349     0.44074977 -0.478062544   0.23422267
## 8  -0.01887437  -0.24935445    -0.23813527  1.221790457   0.21277893
## 9  -0.59103205  -0.19851669    -0.08773703  0.587538822   0.37029042
## 10  0.01290066  -0.05443549    -1.49362185  0.320180946   0.08257993
##    Visibility.mi. Wind_Speed.mph. Precipitation.in.
## 1       0.2268724    -0.262806373       -0.22077417
## 2       0.1227841    -0.149317966       -0.11870567
## 3      -0.2466753    -0.085711197        1.43145970
## 4       0.1596071    -0.192900382       -0.17239795
## 5      -1.6414298     0.002868981        0.20791354
## 6       7.3569938     0.153044647       -0.22077417
## 7       0.2284551     1.713540361       -0.21376947
## 8      -1.3495186     0.204667532        6.71988394
## 9       0.1302443    -0.649728626       -0.12136832
## 10      0.1285428     1.054098644       -0.06921791
##
## Clustering vector:
##   1   2   3   4   5   6   7   8   9  10  11  12  13  14  15  16  17  18  19  20
```

```
##   7   1   4   9  10   5  10   4   9   4   1   4   1   1   2   8   9   9   7   9
##  21  22  23  24  25  26  27  28  29  30  31  32  33  34  35  36  37  38  39  40
##   4   1   9   1   2   4   9   9   9   1   4   8   5   4   9   2   4   4   9   9
##  41  42  43  44  45  46  47  48  49  50  51  52  53  54  55  56  57  58  59  60
##   1  10   5   2   9   5   9   1   1   9   7   9   9   5   7   6   1  10   2   4
##  61  62  63  64  65  66  67  68  69  70  71  72  73  74  75  76  77  78  79  80
##   2   9  10   7   5   4   5   2   4   1  10   9   4   4  10   5   9   9   2   3
##  81  82  83  84  85  86  87  88  89  90  91  92  93  94  95  96  97  98  99 100
##   1   7   5   9   7   4   5   3   5  10   1   4   9   5   4  10   1   5   2   9
## 101 102 103 104 105 106 107 108 109 110 111 112 113 114 115 116 117 118 119 120
##   1   1  10   7   1   7   4   1   4   2   2   9   9   4   1   9   5   2   9   9
## 121 122 123 124 125 126 127 128 129 130 131 132 133 134 135 136 137 138 139 140
##  10   9   4   9  10   7   2   9   7   9   7   9   9   7   5   1   3  10   9   1
## 141 142 143 144 145 146 147 148 149 150 151 152 153 154 155 156 157 158 159 160
##   2   9   5   4   1   1   5   1   1   5   4   1   9   1   2   2   1   4   5   9
## 161 162 163 164 165 166 167 168 169 170 171 172 173 174 175 176 177 178 179 180
##  10   4   1   7   5   5   2  10   9   5   7   1   1   9   7   4   7   7   5   9
## 181 182 183 184 185 186 187 188 189 190 191 192 193 194 195 196 197 198 199 200
##   2   7   2  10   4   4  10   7   7   4   5   1   9   4   4   7   4   4   7   4
## 201 202 203 204 205 206 207 208 209 210 211 212 213 214 215 216 217 218 219 220
##   5   5   9  10   5   4   4   1   9   1   4   1   2   4   4   1   1   5   7   7
## 221 222 223 224 225 226 227 228 229 230 231 232 233 234 235 236 237 238 239 240
##   1   1   7   3   7   5   9   9   5   5   1   9   9   9   9   9   1   6   9   9
## 241 242 243 244 245 246 247 248 249 250 251 252 253 254 255 256 257 258 259 260
##   9   1   7   4   1   6   9   9   9   4   4   1   4  10   2   4   5   9   9   5
## 261 262 263 264 265 266 267 268 269 270 271 272 273 274 275 276 277 278 279 280
##   9   9   7   4   4   9   4   4   7   9   6   4   8   4   9   5  10   9   5   7
## 281 282 283 284 285 286 287 288 289 290 291 292 293 294 295 296 297 298 299 300
##  10  10   7   1   9   4   1   2   1   2   1   1   7   1   1   1  10   9  10   9
## 301 302 303 304 305 306 307 308 309 310 311 312 313 314 315 316 317 318 319 320
##   4   4   8   9   4   9   9   7   5   7   2   4   9   4   4   1   9  10   7   1
## 321 322 323 324 325 326 327 328 329 330 331 332 333 334 335 336 337 338 339 340
##   4  10   9   4   1   4   5   4   7   2   1   1   9   5   1   2   1   7   4   4
## 341 342 343 344 345 346 347 348 349 350 351 352 353 354 355 356 357 358 359 360
##   3   1   4   9   1   9   4   9   7   9   9   4   1   7   4  10   9   5   4   4
## 361 362 363 364 365 366 367 368 369 370 371 372 373 374 375 376 377 378 379 380
##  10   9   1   9   1   7   7   9   1   7   1   9   8   5   4   4   4   5   4   5
## 381 382 383 384 385 386 387 388 389 390 391 392 393 394 395 396 397 398 399 400
##   9   1   9   1   1   5   1   1   9   4   1   1   9   4   1   9   9   4   4   9
## 401 402 403 404 405 406 407 408 409 410 411 412 413 414 415 416 417 418 419 420
##   4   2   1   4   9  10   2   4   2   4   4   7   3   1   4   1   9   1   9   4
## 421 422 423 424 425 426 427 428 429 430 431 432 433 434 435 436 437 438 439 440
##   7   9   7   1   9   2  10   9   9   5   9   5   1   1   2   2   9   3   7   4
## 441 442 443 444 445 446 447 448 449 450 451 452 453 454 455 456 457 458 459 460
##   4   2   2   1   9   2   4  10   1   7   5   5   1   5   4   5   8   1   1   4
## 461 462 463 464 465 466 467 468 469 470 471 472 473 474 475 476 477 478 479 480
##   9   1   5  10   9   8   4   9   5   1   9   3   4   7   9   9   5   9   5   5
## 481 482 483 484 485 486 487 488 489 490 491 492 493 494 495 496 497 498 499 500
##   1   4   7  10   1   1   9   9   9   1   1   4   4   7   1   5   9   1   4   9
##
## Within cluster sum of squares by cluster:
##  [1] 180.70262 185.78922 138.76397 254.56009 191.38862  57.42334 137.29726
##  [8]  46.64630 180.59779 105.38141
##  (between_SS / total_SS =  63.0 %)
```

```
##
## Available components:
##
## [1] "cluster"      "centers"      "totss"        "withinss"     "tot.withinss"
## [6] "betweenss"    "size"         "iter"         "ifault"
```

```
perc.var.10<- round(100*(1 - kmeans10.num$betweenss/kmeans10.num$totss),1)
names(perc.var.10) <- "Perc. 10 clus"
perc.var.10
```

```
## Perc. 10 clus
##          37
```

```
# Computing the percentage of variation accounted for. twelve clusters
(kmeans12.num <- kmeans(matstd.num,12,nstart = 10))
```

```
## K-means clustering with 12 clusters of sizes 9, 4, 5, 33, 115, 54, 1, 34, 50, 52, 92, 51
##
## Cluster means:
##       Severity Distance.mi. Temperature.F. Humidity... Pressure.in.
## 1   0.06850695   0.95055333    -0.34736573  1.14431910   0.18227478
## 2   0.83309350  -0.28610927     0.12197480 -1.07047039  -3.80360195
## 3  -0.17616074   7.24676722    -0.08673133 -0.09001212  -0.41560673
## 4   0.01290066  -0.05443549    -1.49362185  0.32018095   0.08257993
## 5  -0.59103205  -0.19851669    -0.08773703  0.58753882   0.37029042
## 6   1.59768005   0.21341075    -0.09531789  0.59482095   0.39873295
## 7  -0.54316228  -0.28610927    -0.67973009 -1.04887439  -3.72451710
## 8  -0.38124984  -0.20789923    -0.32410401 -0.59408810  -2.95437016
## 9  -0.65326274  -0.15026998     0.41794847 -0.46319095   0.24095574
## 10  0.12732130  -0.09580825    -1.09908342  1.17136024   0.22531519
## 11 -0.54316228  -0.07610161     0.85510906 -1.02493100   0.19310596
## 12  1.32782597  -0.07237667     0.72318904 -0.74568353   0.22460843
##    Visibility.mi. Wind_Speed.mph. Precipitation.in.
## 1     -1.20847934      0.12436527        6.44770127
## 2      4.77400640      0.19606372       -0.22077417
## 3     -0.15949947     -0.01903164        0.13646559
## 4      0.12854276      1.05409864       -0.06921791
## 5      0.13024433     -0.64972863       -0.12136832
## 6      0.03374625     -0.43551998       -0.13477201
## 7     13.16871534      0.06700651       -0.22077417
## 8      0.04296116     -0.16074152       -0.11570365
## 9      0.22794864      1.70517273       -0.21362937
## 10    -1.68122674      0.03788590        0.23264552
## 11     0.22570256     -0.27901646       -0.22077417
## 12     0.23858447      0.10749504       -0.22077417
##
## Clustering vector:
##   1   2   3   4   5   6   7   8   9  10  11  12  13  14  15  16  17  18  19  20
##   9  11  12   5   4  10   4  12   5  12  11   6  11  11   8   1   5   5   9   5
##  21  22  23  24  25  26  27  28  29  30  31  32  33  34  35  36  37  38  39  40
##   6  11   5  11   2  12   5   5   5  11  12   1  10   6   5   8   6  12   5   5
##  41  42  43  44  45  46  47  48  49  50  51  52  53  54  55  56  57  58  59  60
##  11   4  10   8   5  10   5  12  11   5   9   5   5  10   9   2  11   4   8  12
##  61  62  63  64  65  66  67  68  69  70  71  72  73  74  75  76  77  78  79  80
##   8   5   4   9  10   6  10   8   6  11   4   5   6   6   4  10   5   5   8   3
```

```
##  81  82  83  84  85  86  87  88  89  90  91  92  93  94  95  96  97  98  99 100
##  11   9  10   5   9  12  10   1  10   4  11   6   5  10   6   4  11  10   8   5
## 101 102 103 104 105 106 107 108 109 110 111 112 113 114 115 116 117 118 119 120
##  11  11   4   9  11   9   6  11   6   8   8   5   5  12  11   5   6   8   5   5
## 121 122 123 124 125 126 127 128 129 130 131 132 133 134 135 136 137 138 139 140
##   4   5  12   5   4   9   8   5   9   5   9   5   5   9  10  11   3   4   5  11
## 141 142 143 144 145 146 147 148 149 150 151 152 153 154 155 156 157 158 159 160
##   8   5  10  12  11  11  10  11  11  10  12  11   5  11   8   8  11   6  10   5
## 161 162 163 164 165 166 167 168 169 170 171 172 173 174 175 176 177 178 179 180
##   4   6  11   9  10  10   8   4   5  10   9  11  11   5   9  12   9   9  10   5
## 181 182 183 184 185 186 187 188 189 190 191 192 193 194 195 196 197 198 199 200
##   8   9   8   4   6  12   4   9   9   6  10  11   5   6  12   9  12  12   9   6
## 201 202 203 204 205 206 207 208 209 210 211 212 213 214 215 216 217 218 219 220
##  10  10   5   4  10   6  12  11   5  11   6  11   8  12   6  11  11  10   9   9
## 221 222 223 224 225 226 227 228 229 230 231 232 233 234 235 236 237 238 239 240
##  11  11   9   3   9  10   5   5  10  10  11   5   5   5   5   5  11   2   5   5
## 241 242 243 244 245 246 247 248 249 250 251 252 253 254 255 256 257 258 259 260
##   5  11   9  12  11   2   5   5   5  12   6  11  12   4   8  12  10   5   5  10
## 261 262 263 264 265 266 267 268 269 270 271 272 273 274 275 276 277 278 279 280
##   5   5   9   6  12   5   6  12   9   5   7   6   1  12   5  10   4   5  10   9
## 281 282 283 284 285 286 287 288 289 290 291 292 293 294 295 296 297 298 299 300
##   4   4   9  11   5   6  12   8  11   8  11  11   9  11  11  11   4   5   4   5
## 301 302 303 304 305 306 307 308 309 310 311 312 313 314 315 316 317 318 319 320
##  12   6   1   5  12   5   5   9  10   9   8  12   5   6  12  11   5   4  12  11
## 321 322 323 324 325 326 327 328 329 330 331 332 333 334 335 336 337 338 339 340
##   6   4   5  12  11   6  10   6   9   8  11  11   5  10  11   8  11   9   6  12
## 341 342 343 344 345 346 347 348 349 350 351 352 353 354 355 356 357 358 359 360
##   6  11   6   5  11   5  12   5   9   5   5  12  11   9  12   4   5  10   6   6
## 361 362 363 364 365 366 367 368 369 370 371 372 373 374 375 376 377 378 379 380
##   4   5  11   5  11   9   9   5  11   9  11   5   1  10   6   6   6  10   6  10
## 381 382 383 384 385 386 387 388 389 390 391 392 393 394 395 396 397 398 399 400
##   5  11   5  11  11  10  11  11   5   6  11  11   5  12  11   5   5  12  12   5
## 401 402 403 404 405 406 407 408 409 410 411 412 413 414 415 416 417 418 419 420
##   6   8  11  12   5   4   8  12   8  12  12   9   1  11   6  11   5  11   5   6
## 421 422 423 424 425 426 427 428 429 430 431 432 433 434 435 436 437 438 439 440
##   9   5   9  12   5   8   4   5   5  10   5  10  11  11   8   8   5   3   9   6
## 441 442 443 444 445 446 447 448 449 450 451 452 453 454 455 456 457 458 459 460
##   6   8   8  11   5   8  12   4  11   9  10  10  12   6  12  10   1  11  11   6
## 461 462 463 464 465 466 467 468 469 470 471 472 473 474 475 476 477 478 479 480
##   5  11  10   4   5   1  12   5  10  11   5   3   6   9   5   5  10   5   6  10
## 481 482 483 484 485 486 487 488 489 490 491 492 493 494 495 496 497 498 499 500
##  11   6   9   4  11  11   5   5   5  11  11   6  12   9  11  10   5  11   6   5
## 
## Within cluster sum of squares by cluster:
##  [1] 101.46760  16.79062  62.15281 105.38141 180.59779 142.70698   0.00000
##  [8] 173.43423 131.38535 181.13044 158.43653 103.36376
##  (between_SS / total_SS =  66.0 %)
## 
## Available components:
## 
## [1] "cluster"      "centers"      "totss"        "withinss"     "tot.withinss"
## [6] "betweenss"    "size"         "iter"         "ifault"
```

```r
perc.var.12 <- round(100*(1 - kmeans12.num$betweenss/kmeans12.num$totss),1)
names(perc.var.12) <- "Perc. 12 clus"
perc.var.12
```

```
## Perc. 12 clus
##            34
```

```r
# Computing the percentage of variation accounted for. fifteen clusters
(kmeans15.num <- kmeans(matstd.num,15,nstart = 10))
```

```
## K-means clustering with 15 clusters of sizes 53, 40, 42, 4, 8, 49, 51, 47, 52, 15, 28, 16, 67, 21, 7
##
## Cluster means:
##       Severity Distance.mi. Temperature.F. Humidity... Pressure.in.
## 1   1.56882772  0.141662554     -0.0604856   0.5973111   0.41042712
## 2  -0.72666305 -0.175460984      0.6542543  -0.4517451   0.27954915
## 3  -0.54316228 -0.064107068      1.2816619  -1.2473519   0.14913445
## 4   0.37434157 -0.286109267     -0.2328782  -1.0596724  -3.80953332
## 5   0.37434157  6.318795315     -0.4793039   0.2198904  -0.26850890
## 6  -0.61806056  0.008826541      0.4789735   0.5897770   0.37788875
## 7   1.32782597 -0.072376673      0.7231890  -0.7456835   0.22460843
## 8  -0.58220500 -0.239371355      0.1949587  -0.8264816   0.32311312
## 9   0.12732130 -0.095808248     -1.0990834   1.1713602   0.22531519
## 10 -0.42082843 -0.034253472     -0.3327627  -0.4384275   0.13429663
## 11  0.11219761 -0.063486737     -1.5901439   0.3163011   0.04557095
## 12 -0.31378632 -0.250997367     -1.2810088   0.4871409  -3.06267571
## 13 -0.54316228 -0.236139937     -0.3870597   0.6394813   0.35578931
## 14 -0.36839964 -0.182746036      0.6295335  -1.5774621  -2.77625202
## 15 -0.01887437 -0.249354450     -0.2381353   1.2217905   0.21277893
##    Visibility.mi. Wind_Speed.mph. Precipitation.in.
## 1       0.0295947     -0.43850816      -0.133149323
## 2       0.2344061      1.18980426      -0.220774169
## 3       0.1922788     -0.09687567      -0.220774169
## 4       7.3569938      0.15304465      -0.220774169
## 5      -0.2466753     -0.08571120       1.431459701
## 6       0.1114506     -0.09804626      -0.002055951
## 7       0.2385845      0.10749504      -0.220774169
## 8       0.2372914     -0.54075229      -0.220774169
## 9      -1.6812267      0.03788590       0.232645521
## 10      0.1676789      3.02098272      -0.101694250
## 11      0.1338541      0.96917788      -0.131464230
## 12     -0.1942084     -0.41695804       0.002500678
## 13      0.1573984     -0.94747308      -0.194114486
## 14      0.3767779     -0.05590513      -0.220774169
## 15     -1.3495186      0.20466753       6.719883942
##
## Clustering vector:
##   1   2   3   4   5   6   7   8   9  10  11  12  13  14  15  16  17  18  19  20
##   2   3   7   6   2   9  11   7  13   7   8   1   3   3  12  15  13  13   2  13
##  21  22  23  24  25  26  27  28  29  30  31  32  33  34  35  36  37  38  39  40
##   1   8  13   6  14   7   6  13  13   3   7  15   9   1   6  14   1   7   6   6
##  41  42  43  44  45  46  47  48  49  50  51  52  53  54  55  56  57  58  59  60
##   2  11   9  14  13   9  13   7   3  13   2  13   6   9   2   4   3  11  12   7
##  61  62  63  64  65  66  67  68  69  70  71  72  73  74  75  76  77  78  79  80
```

```
##  12  13  11   2   9   1   9  14   1   3  10   6   1   1  10   9   6  13  14   5
##  81  82  83  84  85  86  87  88  89  90  91  92  93  94  95  96  97  98  99 100
##   3   2   9   6   2   7   9   5   9  11   6   1   6   9   1  11   3   9  12   6
## 101 102 103 104 105 106 107 108 109 110 111 112 113 114 115 116 117 118 119 120
##   3   3  11   2   3   2   1   8   1  14  14   8   6   7   3  13   1  14  13   6
## 121 122 123 124 125 126 127 128 129 130 131 132 133 134 135 136 137 138 139 140
##  11  13   7   8   8  10  14  13   2   6   2  13  13   2   9   8   5  11  13   8
## 141 142 143 144 145 146 147 148 149 150 151 152 153 154 155 156 157 158 159 160
##  14  13   9   7   3   3   9   3   3   9   7   8  13   3  14  14   3   1   9  13
## 161 162 163 164 165 166 167 168 169 170 171 172 173 174 175 176 177 178 179 180
##  11   1   8  10   9   9  14  11  13   9   2   3   3  13   2   7   6  10   9   6
## 181 182 183 184 185 186 187 188 189 190 191 192 193 194 195 196 197 198 199 200
##  12  10  12  11   1   7  11  10   2   1   9   3   6   1   7  10   7   7  10   1
## 201 202 203 204 205 206 207 208 209 210 211 212 213 214 215 216 217 218 219 220
##   9   9   6  11   9   1   7   8   8   6   1   6  14   7   1   8   8   9  10   2
## 221 222 223 224 225 226 227 228 229 230 231 232 233 234 235 236 237 238 239 240
##  14   3   2   5  10   9  13  13   9   9   3  13  13  13   8  13   8   4   8   8
## 241 242 243 244 245 246 247 248 249 250 251 252 253 254 255 256 257 258 259 260
##   6   8   2   7   8   4  13   6  13   7   1   6   7  11  12   7   9   6   6   9
## 261 262 263 264 265 266 267 268 269 270 271 272 273 274 275 276 277 278 279 280
##   6   6   2   1   7  13   1   7   2  13   4   1  15   7   6   9  11  13   9   2
## 281 282 283 284 285 286 287 288 289 290 291 292 293 294 295 296 297 298 299 300
##  11  11   2   8  13   1   7  12   8  14   3   3   2   8   8   8  11  13   6  13
## 301 302 303 304 305 306 307 308 309 310 311 312 313 314 315 316 317 318 319 320
##   7   1  15   6   7  13  13   2   9   2  12   7  13   1   7   3  13  11   7  14
## 321 322 323 324 325 326 327 328 329 330 331 332 333 334 335 336 337 338 339 340
##   1  11   6   7   3   1   9   1  10  12   3   6  13   9   3  14   3   2   1   7
## 341 342 343 344 345 346 347 348 349 350 351 352 353 354 355 356 357 358 359 360
##   5   2   1  13   3   6   7  13  10   6   6   7   3   2   7  11  13   9   1   1
## 361 362 363 364 365 366 367 368 369 370 371 372 373 374 375 376 377 378 379 380
##  11  13   8  13   8   2   2   6   3   2   8   6  15   9   1   1   1   9   1   9
## 381 382 383 384 385 386 387 388 389 390 391 392 393 394 395 396 397 398 399 400
##   6   8   6   8   6   9   2   8   6   1   3   8  13   7   8  13  13   7   7  13
## 401 402 403 404 405 406 407 408 409 410 411 412 413 414 415 416 417 418 419 420
##   1  12   8   7  13  11  12   7  14   7   7  10   5   8   1   3  13   8   6   1
## 421 422 423 424 425 426 427 428 429 430 431 432 433 434 435 436 437 438 439 440
##   2  13   2   7  13  14  11   6  13   9   8   9   8   8  14  12  13   5   2   1
## 441 442 443 444 445 446 447 448 449 450 451 452 453 454 455 456 457 458 459 460
##   1  12  12   8   6  12   7  11   8   2   9   9   7   1   7   9  15   8   8   1
## 461 462 463 464 465 466 467 468 469 470 471 472 473 474 475 476 477 478 479 480
##  13   3   9  11  13  15   7  13   9   3  13   5   1   2   6  13   9   8   1   9
## 481 482 483 484 485 486 487 488 489 490 491 492 493 494 495 496 497 498 499 500
##   8   1   2  11   3   3   8  13   6   3   8   1   7  10   3   9  13   6   1  13
##
## Within cluster sum of squares by cluster:
##  [1] 121.62749  62.46288  60.03897  57.42334 138.76397  91.40352 103.36376
##  [8]  43.43998 181.13044  43.54383  82.30302  53.86346  69.90284  71.31594
## [15]  46.64630
##  (between_SS / total_SS =  69.3 %)
##
## Available components:
##
## [1] "cluster"      "centers"      "totss"        "withinss"     "tot.withinss"
## [6] "betweenss"    "size"         "iter"         "ifault"
```

```r
perc.var.15 <- round(100*(1 - kmeans15.num$betweenss/kmeans15.num$totss),1)
names(perc.var.15) <- "Perc. 15 clus"
perc.var.15
```

```
## Perc. 15 clus
##          30.7
```

```r
# Saving four k-means clusters in a list
clus1 <- matrix(names(kmeans4.num$cluster[kmeans4.num$cluster == 1]),
                ncol=1, nrow=length(kmeans4.num$cluster[kmeans4.num$cluster == 1]))
colnames(clus1) <- "Cluster 1"
clus2 <- matrix(names(kmeans4.num$cluster[kmeans4.num$cluster == 2]),
                ncol=1, nrow=length(kmeans4.num$cluster[kmeans4.num$cluster == 2]))
colnames(clus2) <- "Cluster 2"
clus3 <- matrix(names(kmeans4.num$cluster[kmeans4.num$cluster == 3]),
                ncol=1, nrow=length(kmeans4.num$cluster[kmeans4.num$cluster == 3]))
colnames(clus3) <- "Cluster 3"
clus4 <- matrix(names(kmeans4.num$cluster[kmeans4.num$cluster == 4]),
                ncol=1, nrow=length(kmeans4.num$cluster[kmeans4.num$cluster == 4]))
colnames(clus4) <- "Cluster 4"
list(clus1,clus2,clus3,clus4)
```

```
## [[1]]
##       Cluster 1
##  [1,] "15"
##  [2,] "25"
##  [3,] "36"
##  [4,] "44"
##  [5,] "56"
##  [6,] "59"
##  [7,] "61"
##  [8,] "68"
##  [9,] "79"
## [10,] "99"
## [11,] "103"
## [12,] "110"
## [13,] "111"
## [14,] "118"
## [15,] "127"
## [16,] "141"
## [17,] "155"
## [18,] "156"
## [19,] "167"
## [20,] "181"
## [21,] "183"
## [22,] "213"
## [23,] "221"
## [24,] "238"
## [25,] "246"
## [26,] "255"
## [27,] "271"
## [28,] "290"
## [29,] "311"
## [30,] "320"
```

```
## [31,] "330"
## [32,] "336"
## [33,] "402"
## [34,] "407"
## [35,] "409"
## [36,] "426"
## [37,] "435"
## [38,] "436"
## [39,] "443"
## [40,] "446"
## [41,] "472"
##
## [[2]]
##       Cluster 2
##  [1,] "16"
##  [2,] "32"
##  [3,] "88"
##  [4,] "273"
##  [5,] "303"
##  [6,] "373"
##  [7,] "413"
##  [8,] "457"
##  [9,] "466"
## [10,] "480"
##
## [[3]]
##        Cluster 3
##   [1,] "4"
##   [2,] "6"
##   [3,] "7"
##   [4,] "9"
##   [5,] "10"
##   [6,] "12"
##   [7,] "17"
##   [8,] "18"
##   [9,] "20"
##  [10,] "21"
##  [11,] "23"
##  [12,] "27"
##  [13,] "28"
##  [14,] "29"
##  [15,] "33"
##  [16,] "34"
##  [17,] "37"
##  [18,] "39"
##  [19,] "42"
##  [20,] "43"
##  [21,] "45"
##  [22,] "46"
##  [23,] "47"
##  [24,] "50"
##  [25,] "52"
##  [26,] "54"
##  [27,] "58"
```

```
##  [28,] "62"
##  [29,] "63"
##  [30,] "65"
##  [31,] "66"
##  [32,] "67"
##  [33,] "69"
##  [34,] "71"
##  [35,] "72"
##  [36,] "73"
##  [37,] "74"
##  [38,] "75"
##  [39,] "76"
##  [40,] "78"
##  [41,] "83"
##  [42,] "84"
##  [43,] "87"
##  [44,] "89"
##  [45,] "90"
##  [46,] "92"
##  [47,] "94"
##  [48,] "95"
##  [49,] "96"
##  [50,] "98"
##  [51,] "107"
##  [52,] "109"
##  [53,] "116"
##  [54,] "117"
##  [55,] "119"
##  [56,] "121"
##  [57,] "122"
##  [58,] "128"
##  [59,] "130"
##  [60,] "132"
##  [61,] "133"
##  [62,] "135"
##  [63,] "137"
##  [64,] "138"
##  [65,] "139"
##  [66,] "142"
##  [67,] "143"
##  [68,] "147"
##  [69,] "150"
##  [70,] "153"
##  [71,] "158"
##  [72,] "159"
##  [73,] "160"
##  [74,] "161"
##  [75,] "162"
##  [76,] "165"
##  [77,] "166"
##  [78,] "168"
##  [79,] "169"
##  [80,] "170"
##  [81,] "174"
```

```
##  [82,] "177"
##  [83,] "179"
##  [84,] "180"
##  [85,] "184"
##  [86,] "185"
##  [87,] "187"
##  [88,] "191"
##  [89,] "194"
##  [90,] "200"
##  [91,] "201"
##  [92,] "202"
##  [93,] "204"
##  [94,] "205"
##  [95,] "206"
##  [96,] "209"
##  [97,] "211"
##  [98,] "215"
##  [99,] "218"
## [100,] "224"
## [101,] "226"
## [102,] "227"
## [103,] "228"
## [104,] "229"
## [105,] "230"
## [106,] "232"
## [107,] "233"
## [108,] "234"
## [109,] "236"
## [110,] "247"
## [111,] "249"
## [112,] "251"
## [113,] "254"
## [114,] "257"
## [115,] "260"
## [116,] "261"
## [117,] "264"
## [118,] "266"
## [119,] "267"
## [120,] "270"
## [121,] "272"
## [122,] "276"
## [123,] "277"
## [124,] "278"
## [125,] "279"
## [126,] "281"
## [127,] "282"
## [128,] "285"
## [129,] "286"
## [130,] "288"
## [131,] "297"
## [132,] "298"
## [133,] "299"
## [134,] "300"
## [135,] "301"
```

```
## [136,] "302"
## [137,] "304"
## [138,] "306"
## [139,] "307"
## [140,] "309"
## [141,] "313"
## [142,] "314"
## [143,] "317"
## [144,] "318"
## [145,] "321"
## [146,] "322"
## [147,] "323"
## [148,] "326"
## [149,] "327"
## [150,] "328"
## [151,] "333"
## [152,] "334"
## [153,] "339"
## [154,] "341"
## [155,] "343"
## [156,] "344"
## [157,] "346"
## [158,] "348"
## [159,] "350"
## [160,] "356"
## [161,] "357"
## [162,] "358"
## [163,] "359"
## [164,] "360"
## [165,] "361"
## [166,] "362"
## [167,] "364"
## [168,] "368"
## [169,] "374"
## [170,] "375"
## [171,] "376"
## [172,] "377"
## [173,] "378"
## [174,] "379"
## [175,] "380"
## [176,] "381"
## [177,] "383"
## [178,] "386"
## [179,] "389"
## [180,] "390"
## [181,] "393"
## [182,] "396"
## [183,] "397"
## [184,] "400"
## [185,] "401"
## [186,] "405"
## [187,] "415"
## [188,] "417"
## [189,] "420"
```

```
## [190,] "422"
## [191,] "425"
## [192,] "427"
## [193,] "429"
## [194,] "430"
## [195,] "432"
## [196,] "437"
## [197,] "438"
## [198,] "440"
## [199,] "441"
## [200,] "442"
## [201,] "448"
## [202,] "451"
## [203,] "452"
## [204,] "454"
## [205,] "456"
## [206,] "460"
## [207,] "461"
## [208,] "463"
## [209,] "464"
## [210,] "465"
## [211,] "468"
## [212,] "469"
## [213,] "471"
## [214,] "473"
## [215,] "476"
## [216,] "477"
## [217,] "479"
## [218,] "482"
## [219,] "484"
## [220,] "488"
## [221,] "489"
## [222,] "492"
## [223,] "496"
## [224,] "497"
## [225,] "499"
## [226,] "500"
##
## [[4]]
##        Cluster 4
##   [1,] "1"
##   [2,] "2"
##   [3,] "3"
##   [4,] "5"
##   [5,] "8"
##   [6,] "11"
##   [7,] "13"
##   [8,] "14"
##   [9,] "19"
##  [10,] "22"
##  [11,] "24"
##  [12,] "26"
##  [13,] "30"
##  [14,] "31"
```

```
##  [15,] "35"
##  [16,] "38"
##  [17,] "40"
##  [18,] "41"
##  [19,] "48"
##  [20,] "49"
##  [21,] "51"
##  [22,] "53"
##  [23,] "55"
##  [24,] "57"
##  [25,] "60"
##  [26,] "64"
##  [27,] "70"
##  [28,] "77"
##  [29,] "80"
##  [30,] "81"
##  [31,] "82"
##  [32,] "85"
##  [33,] "86"
##  [34,] "91"
##  [35,] "93"
##  [36,] "97"
##  [37,] "100"
##  [38,] "101"
##  [39,] "102"
##  [40,] "104"
##  [41,] "105"
##  [42,] "106"
##  [43,] "108"
##  [44,] "112"
##  [45,] "113"
##  [46,] "114"
##  [47,] "115"
##  [48,] "120"
##  [49,] "123"
##  [50,] "124"
##  [51,] "125"
##  [52,] "126"
##  [53,] "129"
##  [54,] "131"
##  [55,] "134"
##  [56,] "136"
##  [57,] "140"
##  [58,] "144"
##  [59,] "145"
##  [60,] "146"
##  [61,] "148"
##  [62,] "149"
##  [63,] "151"
##  [64,] "152"
##  [65,] "154"
##  [66,] "157"
##  [67,] "163"
##  [68,] "164"
```

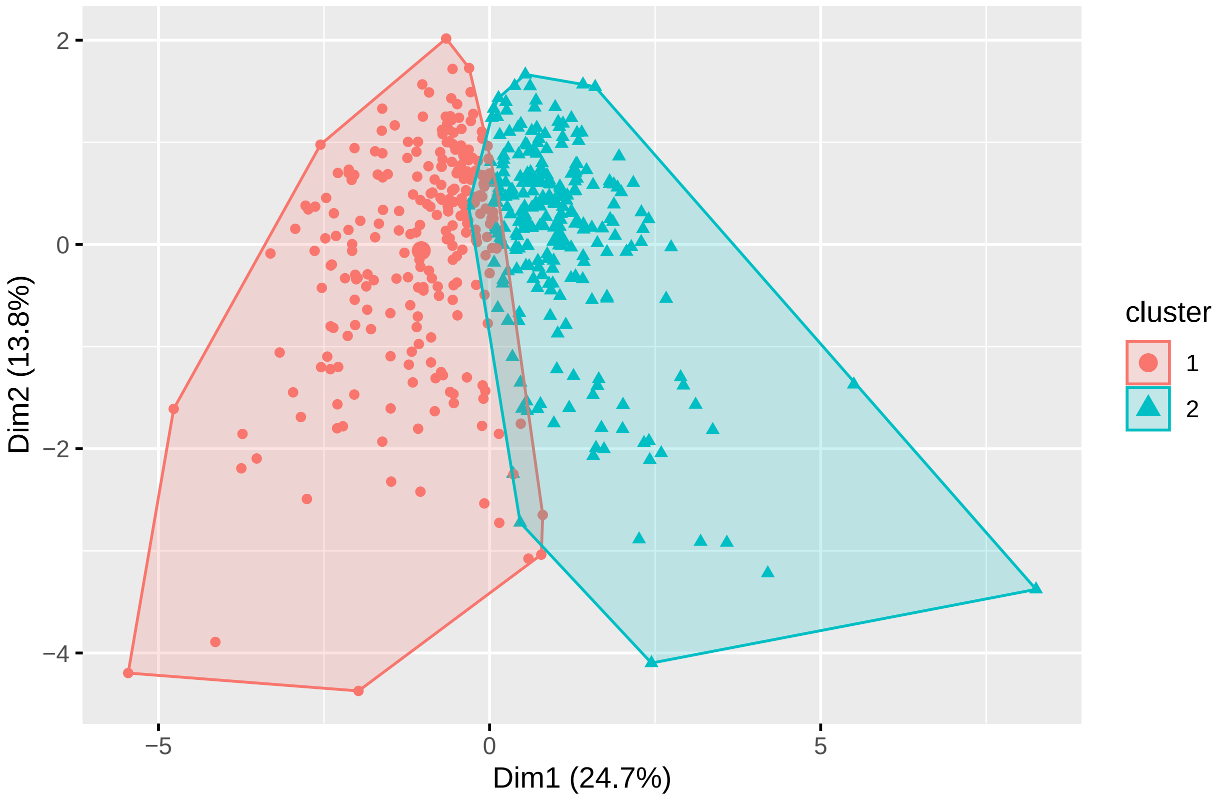
```
##  [69,] "171"
##  [70,] "172"
##  [71,] "173"
##  [72,] "175"
##  [73,] "176"
##  [74,] "178"
##  [75,] "182"
##  [76,] "186"
##  [77,] "188"
##  [78,] "189"
##  [79,] "190"
##  [80,] "192"
##  [81,] "193"
##  [82,] "195"
##  [83,] "196"
##  [84,] "197"
##  [85,] "198"
##  [86,] "199"
##  [87,] "203"
##  [88,] "207"
##  [89,] "208"
##  [90,] "210"
##  [91,] "212"
##  [92,] "214"
##  [93,] "216"
##  [94,] "217"
##  [95,] "219"
##  [96,] "220"
##  [97,] "222"
##  [98,] "223"
##  [99,] "225"
## [100,] "231"
## [101,] "235"
## [102,] "237"
## [103,] "239"
## [104,] "240"
## [105,] "241"
## [106,] "242"
## [107,] "243"
## [108,] "244"
## [109,] "245"
## [110,] "248"
## [111,] "250"
## [112,] "252"
## [113,] "253"
## [114,] "256"
## [115,] "258"
## [116,] "259"
## [117,] "262"
## [118,] "263"
## [119,] "265"
## [120,] "268"
## [121,] "269"
## [122,] "274"
```

```
## [123,] "275"
## [124,] "280"
## [125,] "283"
## [126,] "284"
## [127,] "287"
## [128,] "289"
## [129,] "291"
## [130,] "292"
## [131,] "293"
## [132,] "294"
## [133,] "295"
## [134,] "296"
## [135,] "305"
## [136,] "308"
## [137,] "310"
## [138,] "312"
## [139,] "315"
## [140,] "316"
## [141,] "319"
## [142,] "324"
## [143,] "325"
## [144,] "329"
## [145,] "331"
## [146,] "332"
## [147,] "335"
## [148,] "337"
## [149,] "338"
## [150,] "340"
## [151,] "342"
## [152,] "345"
## [153,] "347"
## [154,] "349"
## [155,] "351"
## [156,] "352"
## [157,] "353"
## [158,] "354"
## [159,] "355"
## [160,] "363"
## [161,] "365"
## [162,] "366"
## [163,] "367"
## [164,] "369"
## [165,] "370"
## [166,] "371"
## [167,] "372"
## [168,] "382"
## [169,] "384"
## [170,] "385"
## [171,] "387"
## [172,] "388"
## [173,] "391"
## [174,] "392"
## [175,] "394"
## [176,] "395"
```

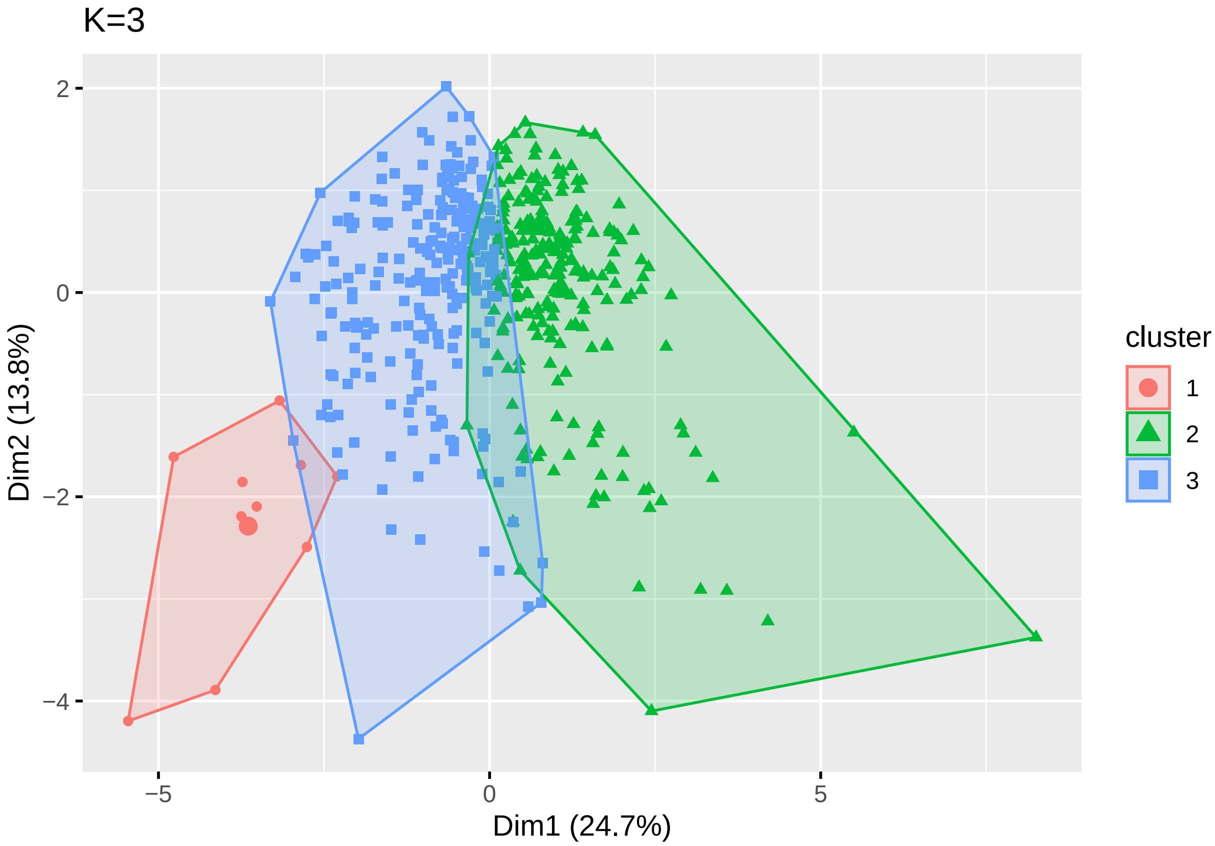
```
## [177,] "398"
## [178,] "399"
## [179,] "403"
## [180,] "404"
## [181,] "406"
## [182,] "408"
## [183,] "410"
## [184,] "411"
## [185,] "412"
## [186,] "414"
## [187,] "416"
## [188,] "418"
## [189,] "419"
## [190,] "421"
## [191,] "423"
## [192,] "424"
## [193,] "428"
## [194,] "431"
## [195,] "433"
## [196,] "434"
## [197,] "439"
## [198,] "444"
## [199,] "445"
## [200,] "447"
## [201,] "449"
## [202,] "450"
## [203,] "453"
## [204,] "455"
## [205,] "458"
## [206,] "459"
## [207,] "462"
## [208,] "467"
## [209,] "470"
## [210,] "474"
## [211,] "475"
## [212,] "478"
## [213,] "481"
## [214,] "483"
## [215,] "485"
## [216,] "486"
## [217,] "487"
## [218,] "490"
## [219,] "491"
## [220,] "493"
## [221,] "494"
## [222,] "495"
## [223,] "498"
```

```r
# Comparing the effect of number of  clusters
fviz_cluster(kmeans2.num,geom = "point",data=matstd.num)+ggtitle("K=2")
```

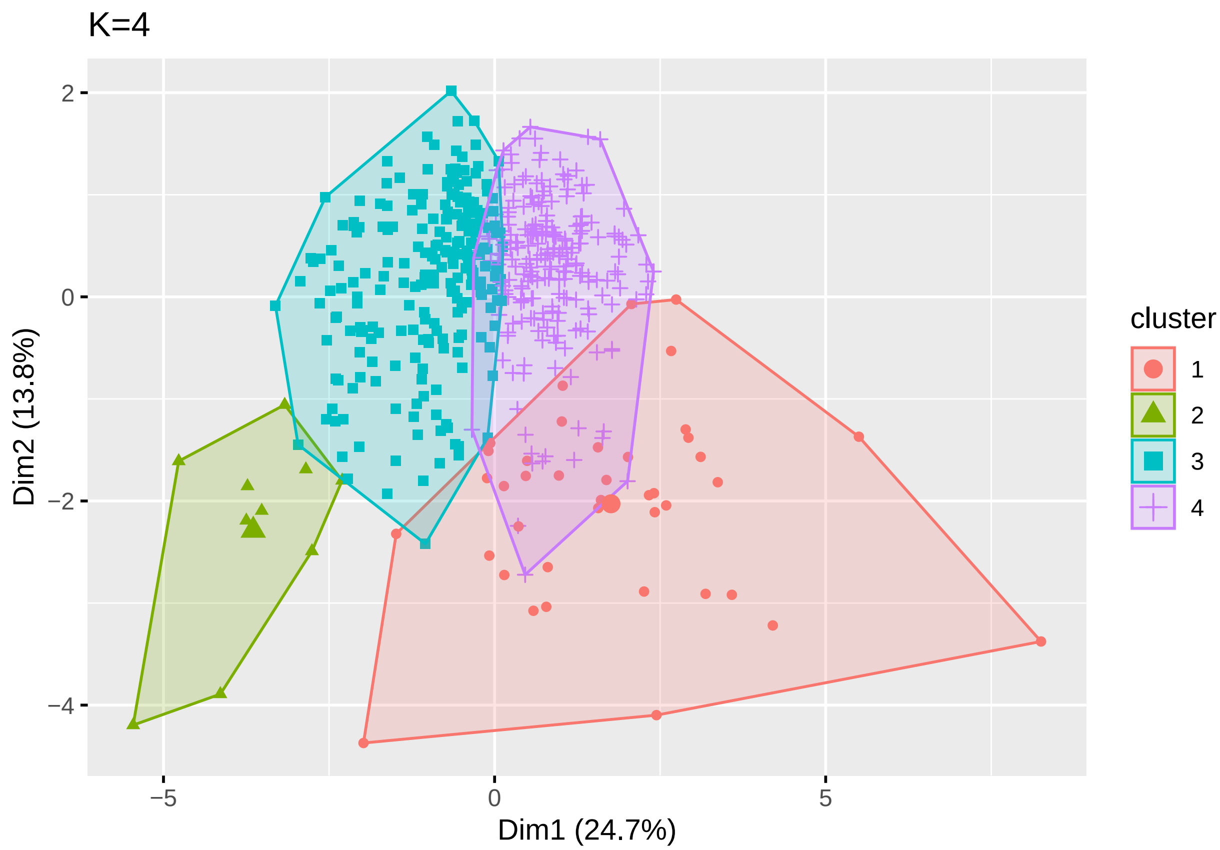
K=2



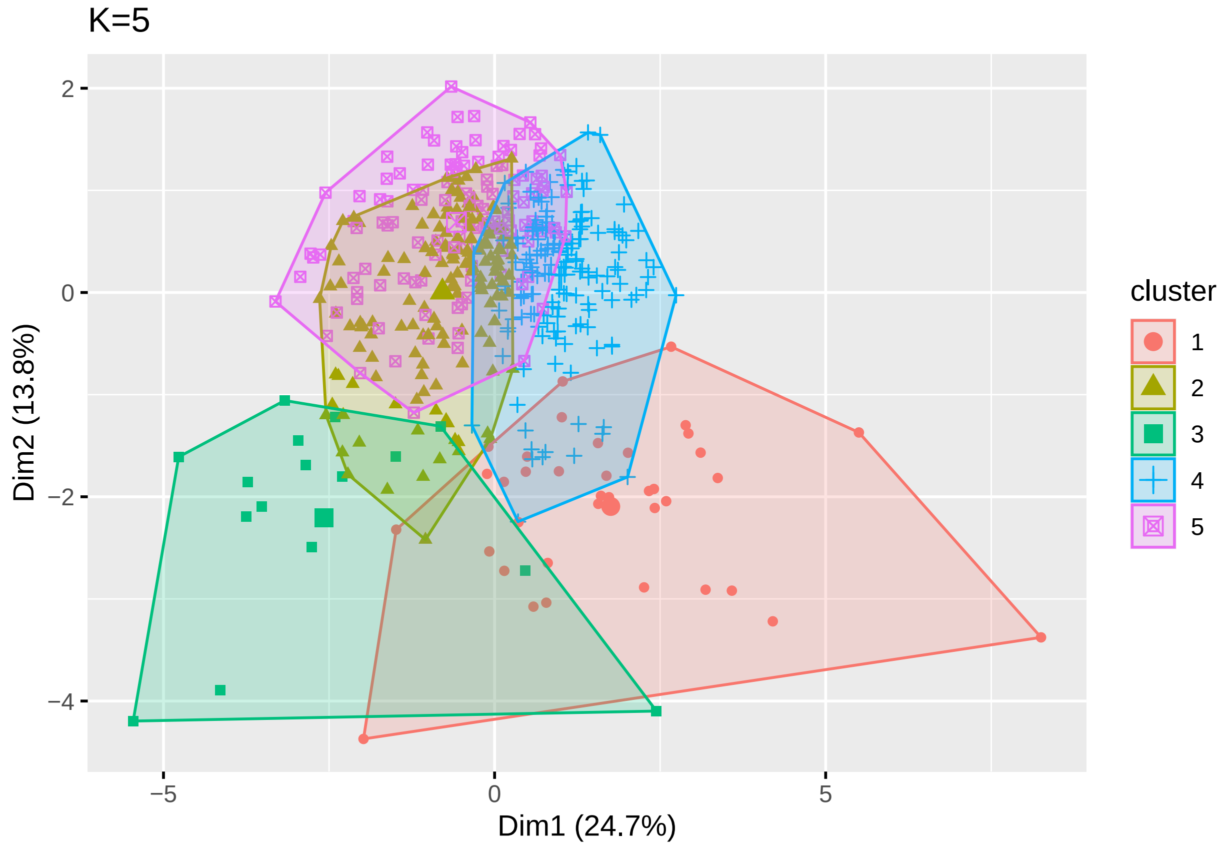

```
fviz_cluster(kmeans3.num,geom = "point",data=matstd.num)+ggtitle("K=3")
```



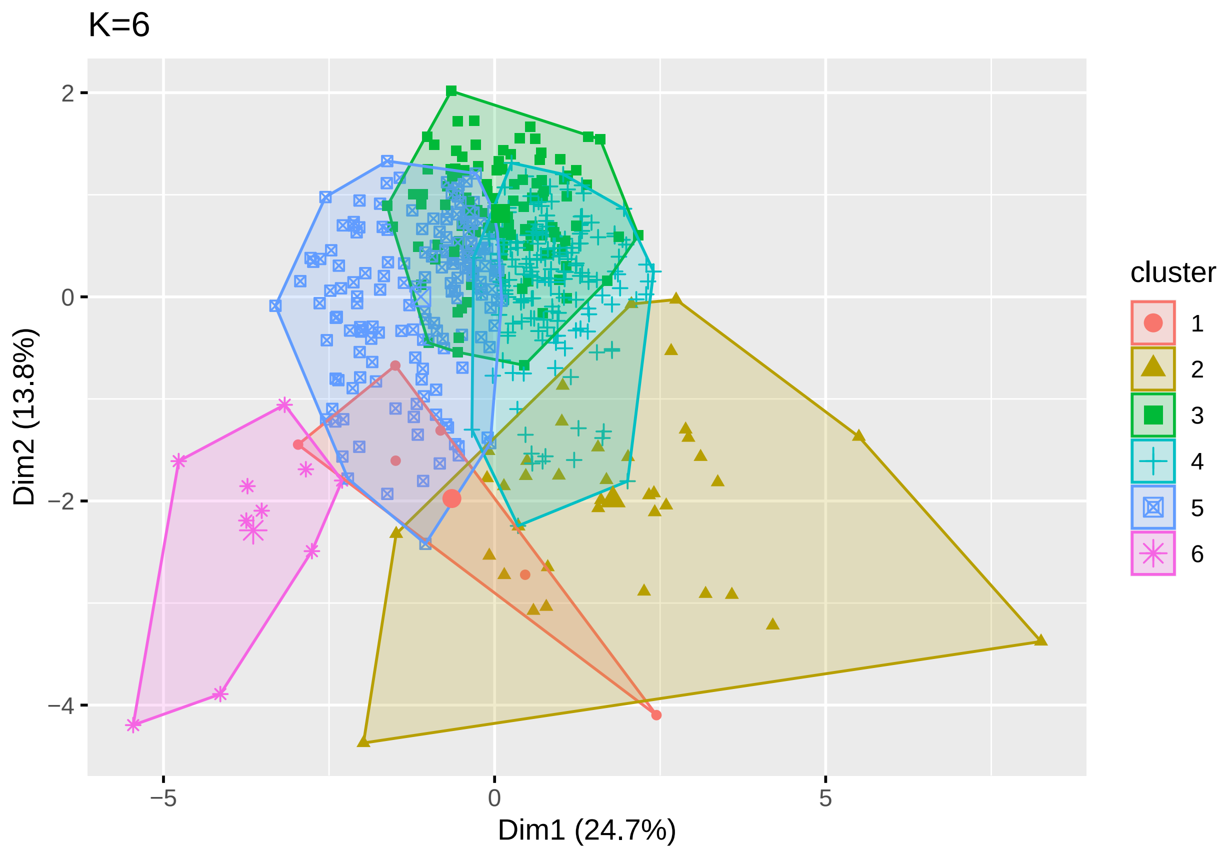

```
fviz_cluster(kmeans4.num,geom = "point",data=matstd.num)+ggtitle("K=4")
```



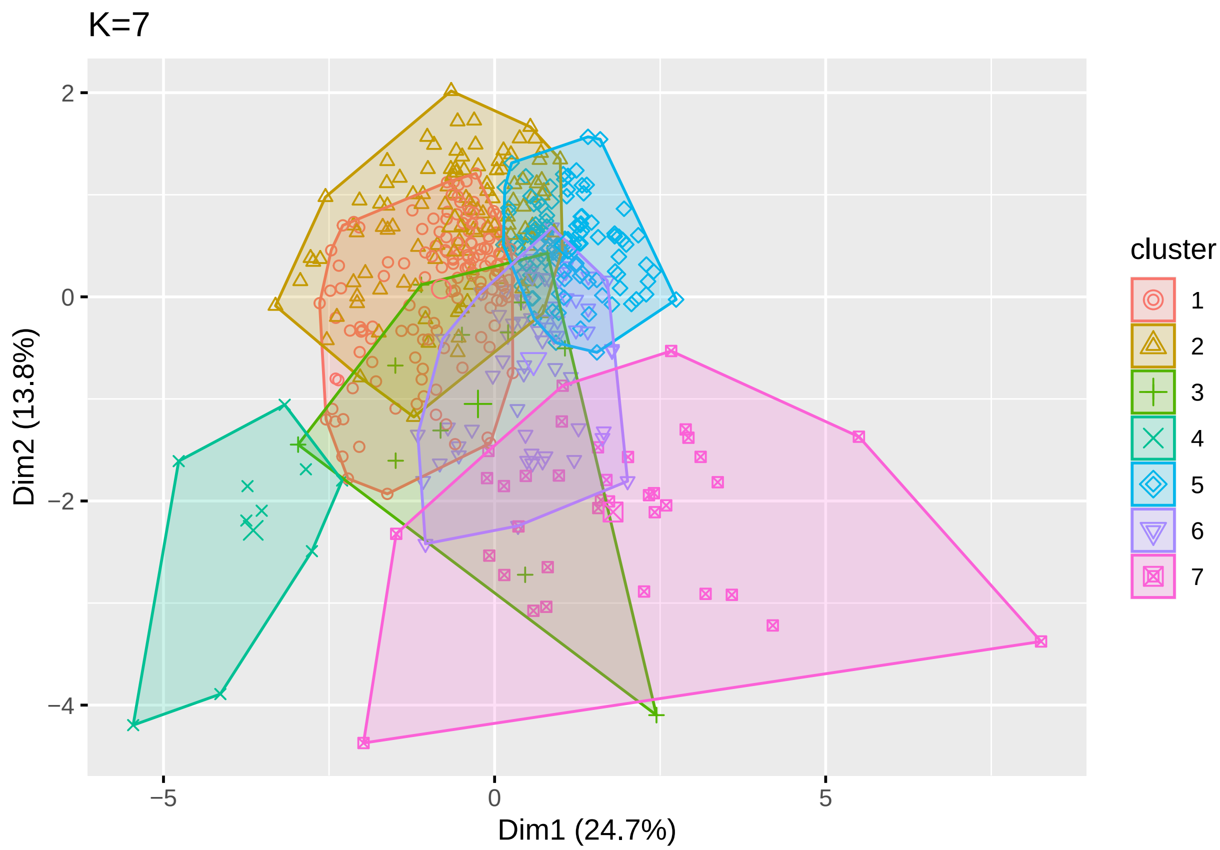

```r
fviz_cluster(kmeans5.num,geom = "point",data=matstd.num)+ggtitle("K=5")
```



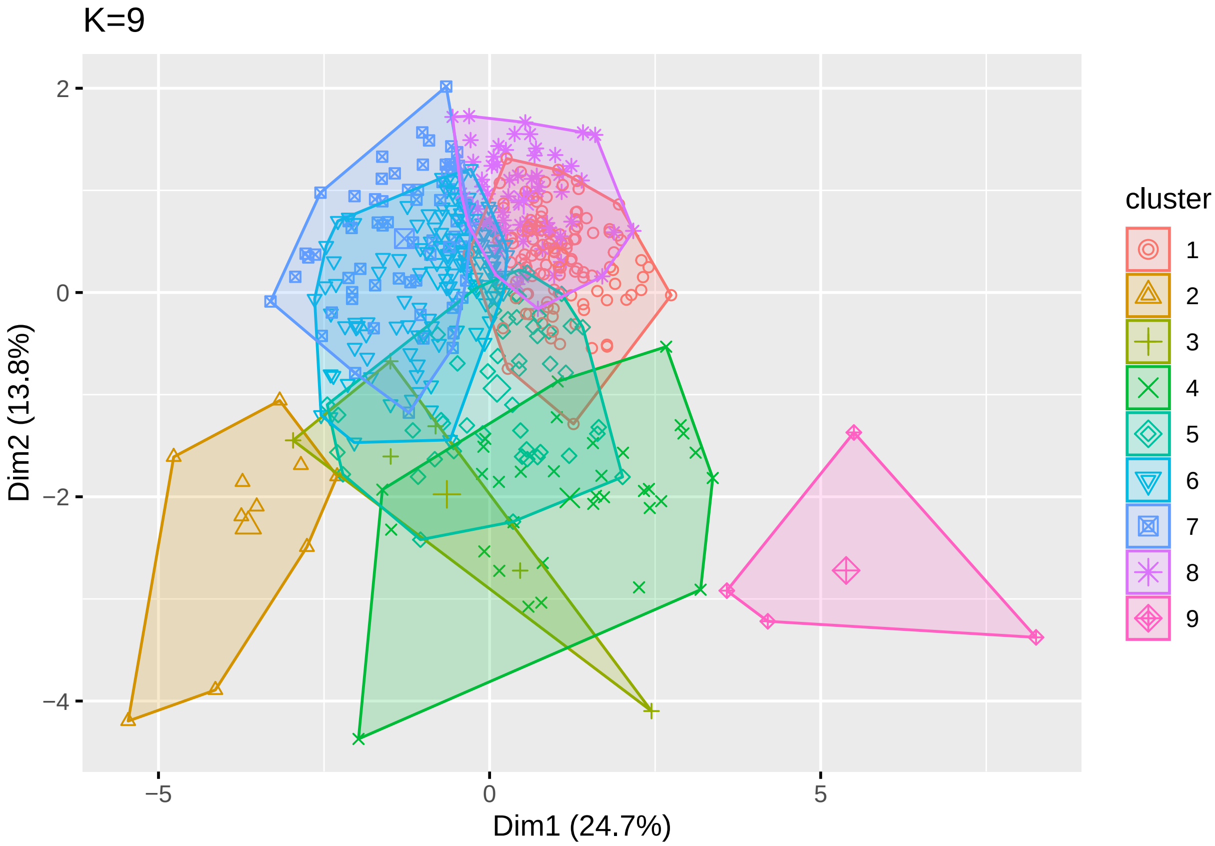

```
fviz_cluster(kmeans6.num,geom = "point",data=matstd.num)+ggtitle("K=6")
```



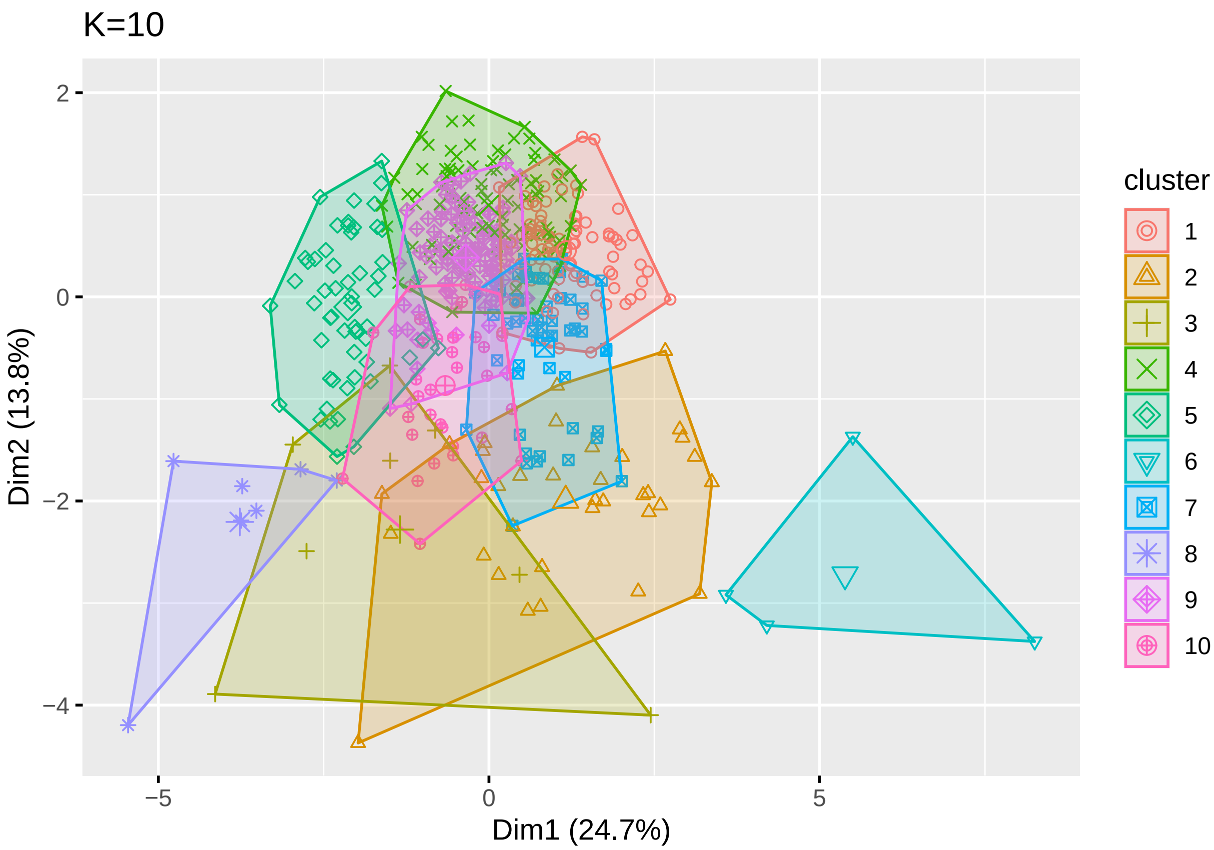

```
fviz_cluster(kmeans7.num,geom = "point",data=matstd.num)+ggtitle("K=7")
```



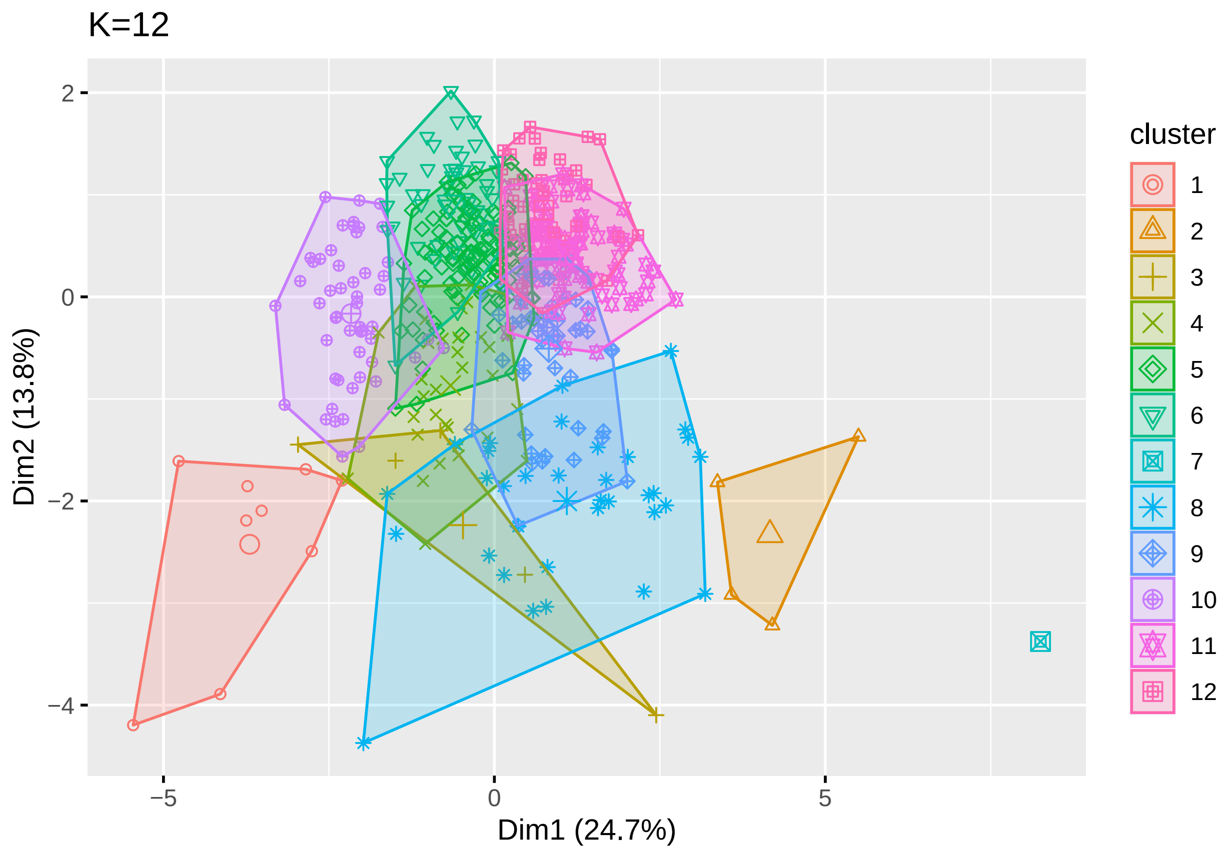

```
fviz_cluster(kmeans9.num,geom = "point",data=matstd.num)+ggtitle("K=9")
```



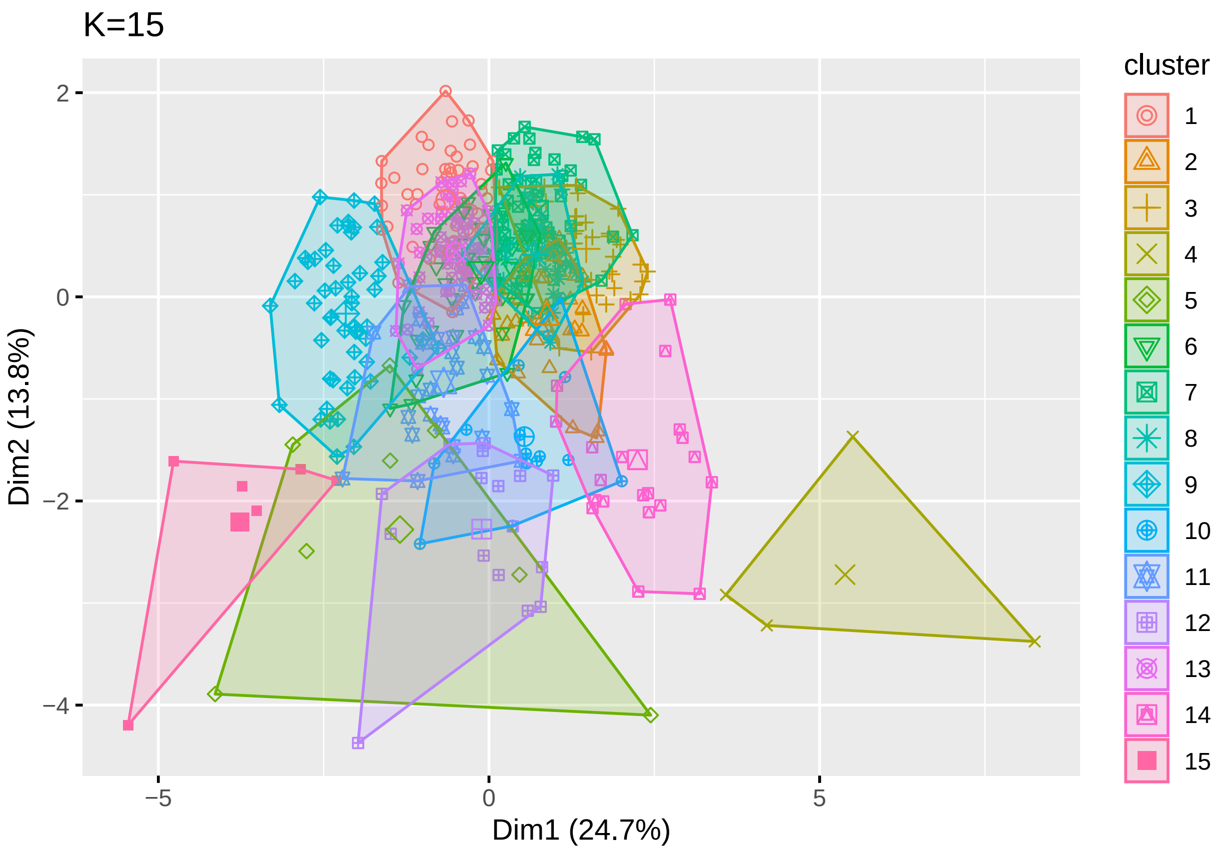

```
fviz_cluster(kmeans10.num,geom = "point",data=matstd.num)+ggtitle("K=10")
```




```
fviz_cluster(kmeans12.num,geom = "point",data=matstd.num)+ggtitle("K=12")
```




```
fviz_cluster(kmeans15.num,geom = "point",data=matstd.num)+ggtitle("K=15")
```


K=15
cluster
1
2
3
4
5
6
7
8
9
10
11
12
13
14
15
Dim2 (13.8%)
Dim1 (24.7%)
2
0
−2
−4
−5
0
5