Cluster Analysis

Lilian

30/11/2022

```
health <- read.csv(file = "C:/Users/LILIAN/Desktop/Cluster Analysis/health outcome.csv")
```

```
sum(is.na(health))
```

```
## [1] 321
```

DATA WRANGLING AND EXPLORATION ANALYSIS

```
#DROP INSIGNIFICANT COLUMNS OF THE DATASET
health <- subset(health, select = -c(Infants.exclusively.breastfed....ages.0Ð5.months..2008Ð2013,Deaths
```

```
#df <- df[-c(209,210),]
```

```
head(health,3)
```

```
##       Country Infants_DTP Infants_Measles Infant_Mortality UnderFive_Mortality
## 1      Norway           1               7              2.3                 2.8
## 2 Netherlands           1               4              3.3                 4.0
## 3      Sweden           1               3              2.4                 3.0
##   Female_Adult_mortality Male_Adult_Mortality Tuberculosis_death
## 1                     47                   73                0.1
## 2                     54                   69                0.2
## 3                     43                   69                0.1
##   Life_expectancy Physicians Public_health
## 1            24.0       37.4          9.6
## 2            23.5       31.5         12.9
## 3            24.1       32.7          9.7
```

```
nrow(health)
```

```
## [1] 208
```

```
ncol(health)
```

```
## [1] 11
```

```r
# Add row names
rownames(health) = c(health$Country)
head(health)
```

```
##                         Country Infants_DTP Infants_Measles Infant_Mortality
## Norway                   Norway           1               7              2.3
## Netherlands         Netherlands           1               4              3.3
## Sweden                   Sweden           1               3              2.4
## Korea _Republic Korea _Republic           1               1              3.2
## Luxembourg           Luxembourg           1               5              1.6
## Japan                     Japan           1               5              2.1
##                 UnderFive_Mortality Female_Adult_mortality Male_Adult_Mortality
## Norway                          2.8                     47                   73
## Netherlands                     4.0                     54                   69
## Sweden                          3.0                     43                   69
## Korea _Republic                 3.7                     38                   93
## Luxembourg                      2.0                     50                   79
## Japan                           2.9                     42                   81
##                 Tuberculosis_death Life_expectancy Physicians Public_health
## Norway                         0.1            24.0       37.4          9.6
## Netherlands                    0.2            23.5       31.5         12.9
## Sweden                         0.1            24.1       32.7          9.7
## Korea _Republic                5.4            24.0       21.4          7.2
## Luxembourg                     0.4            23.4       28.2          7.1
## Japan                          1.7            26.1       23.0         10.3
```

```r
newhealth <- health[,-1]
head(newhealth,3)
```

```
##             Infants_DTP Infants_Measles Infant_Mortality UnderFive_Mortality
## Norway                1               7              2.3                 2.8
## Netherlands           1               4              3.3                 4.0
## Sweden                1               3              2.4                 3.0
##             Female_Adult_mortality Male_Adult_Mortality Tuberculosis_death
## Norway                          47                   73                0.1
## Netherlands                     54                   69                0.2
## Sweden                          43                   69                0.1
##             Life_expectancy Physicians Public_health
## Norway                 24.0       37.4          9.6
## Netherlands            23.5       31.5         12.9
## Sweden                 24.1       32.7          9.7
```

```r
health <- newhealth
```

```r
health <- na.omit(health)
head(health,3)
```

```
##             Infants_DTP Infants_Measles Infant_Mortality UnderFive_Mortality
## Norway                1               7              2.3                 2.8
## Netherlands           1               4              3.3                 4.0
## Sweden                1               3              2.4                 3.0
##             Female_Adult_mortality Male_Adult_Mortality Tuberculosis_death
```

```
## Norway                          47                  73               0.1
## Netherlands                     54                  69               0.2
## Sweden                          43                  69               0.1
##             Life_expectancy Physicians Public_health
## Norway                 24.0       37.4          9.6
## Netherlands            23.5       31.5         12.9
## Sweden                 24.1       32.7          9.7
```

```
glimpse(health)
```

```
## Rows: 192
## Columns: 10
## $ Infants_DTP            <int> 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1,~
## $ Infants_Measles        <int> 7, 4, 3, 1, 5, 5, 8, 11, 3, 10, 1, 1, 14, 1, 2,~
## $ Infant_Mortality       <dbl> 2.3, 3.3, 2.4, 3.2, 1.6, 2.1, 3.5, 3.5, 2.1, 3.~
## $ UnderFive_Mortality    <dbl> 2.8, 4.0, 3.0, 3.7, 2.0, 2.9, 4.4, 4.2, 2.6, 3.~
## $ Female_Adult_mortality <int> 47, 54, 43, 38, 50, 42, 57, 52, 51, 38, 57, 41,~
## $ Male_Adult_Mortality   <int> 73, 69, 69, 93, 79, 81, 98, 109, 114, 69, 127, ~
## $ Tuberculosis_death     <dbl> 0.1, 0.2, 0.1, 5.4, 0.4, 1.7, 0.4, 0.5, 0.3, 0.~
## $ Life_expectancy        <dbl> 24.0, 23.5, 24.1, 24.0, 23.4, 26.1, 23.6, 25.1,~
## $ Physicians             <dbl> 37.4, 31.5, 32.7, 21.4, 28.2, 23.0, 29.9, 31.8,~
## $ Public_health          <dbl> 9.6, 12.9, 9.7, 7.2, 7.1, 10.3, 11.2, 11.7, 9.4~
```

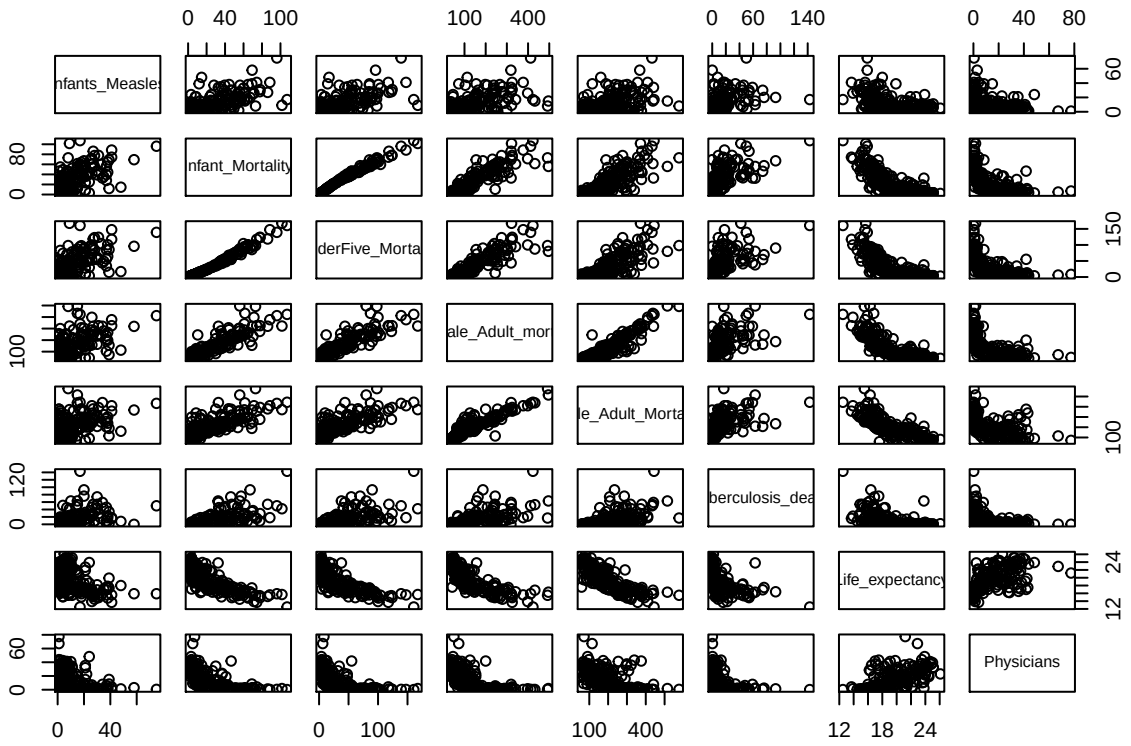
```
summary(health)
```

```
##   Infants_DTP     Infants_Measles Infant_Mortality  UnderFive_Mortality
##  Min.   : 1.000   Min.   : 1.00   Min.   :  1.600   Min.   :  2.00
##  1st Qu.: 1.000   1st Qu.: 3.00   1st Qu.:  7.175   1st Qu.:  8.35
##  Median : 3.000   Median : 7.50   Median : 16.050   Median : 18.85
##  Mean   : 6.359   Mean   :11.68   Mean   : 25.921   Mean   : 35.00
##  3rd Qu.: 8.000   3rd Qu.:17.00   3rd Qu.: 39.650   3rd Qu.: 51.85
##  Max.   :65.000   Max.   :75.00   Max.   :107.200   Max.   :167.40
##  Female_Adult_mortality Male_Adult_Mortality Tuberculosis_death Life_expectancy
##  Min.   : 36.0          Min.   : 59.0        Min.   :  0.00     Min.   :12.50
##  1st Qu.: 69.0          1st Qu.:117.8        1st Qu.:  1.00     1st Qu.:17.00
##  Median :103.0          Median :187.5        Median :  4.45     Median :19.05
##  Mean   :140.7          Mean   :203.2        Mean   : 12.74     Mean   :19.38
##  3rd Qu.:196.5          3rd Qu.:262.2        3rd Qu.: 17.00     3rd Qu.:21.50
##  Max.   :496.0          Max.   :577.0        Max.   :143.00     Max.   :26.10
##    Physicians    Public_health
##  Min.   : 0.10   Min.   : 1.300
##  1st Qu.: 2.30   1st Qu.: 4.675
##  Median :11.70   Median : 6.400
##  Mean   :15.43   Mean   : 6.669
##  3rd Qu.:26.93   3rd Qu.: 8.525
##  Max.   :77.40   Max.   :17.100
```

```
duplicated(health)%>%
head(n = 20)
```

```
##  [1] FALSE FALSE FALSE FALSE FALSE FALSE FALSE FALSE FALSE FALSE FALSE FALSE
## [13] FALSE FALSE FALSE FALSE FALSE FALSE FALSE FALSE
```

```
pairs(health[2:9])
```



CLUSTER ANALYSIS

```
library(cluster)
```

```
## Warning: package 'cluster' was built under R version 4.0.5
```

```
library(factoextra)
```

```
## Warning: package 'factoextra' was built under R version 4.0.5

## Loading required package: ggplot2

## Warning: package 'ggplot2' was built under R version 4.0.5

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

```
health_scale <- scale(health)
head(health_scale)
```

```
##                Infants_DTP Infants_Measles Infant_Mortality
## Norway          -0.6059315      -0.3872758       -1.0108556
```

```
## Netherlands      -0.6059315      -0.6354081       -0.9680605
## Sweden           -0.6059315      -0.7181188       -1.0065761
## Korea _Republic  -0.6059315      -0.8835403       -0.9723400
## Luxembourg       -0.6059315      -0.5526973       -1.0408121
## Japan            -0.6059315      -0.5526973       -1.0194146
##                 UnderFive_Mortality Female_Adult_mortality Male_Adult_Mortality
## Norway                   -0.9123059             -0.9576841            -1.297255
## Netherlands              -0.8783119             -0.8861452            -1.337107
## Sweden                   -0.9066402             -0.9985635            -1.337107
## Korea _Republic          -0.8868104             -1.0496627            -1.097997
## Luxembourg               -0.9349686             -0.9270246            -1.237478
## Japan                    -0.9094731             -1.0087833            -1.217552
##                 Tuberculosis_death Life_expectancy Physicians Public_health
## Norway                  -0.6541151        1.575400  1.5361327     1.1181616
## Netherlands             -0.6489387        1.405048  1.1235741     2.3769875
## Sweden                  -0.6541151        1.609471  1.2074843     1.1563079
## Korea _Republic         -0.3797653        1.575400  0.4173296     0.2026519
## Luxembourg              -0.6385859        1.370978  0.8928209     0.1645057
## Japan                   -0.5712925        2.290878  0.5292099     1.3851853
```

```
health_euc <- dist(health_scale, method = "euclidean")
```

```
round(as.matrix(health_euc)[1:3, 1:3], 1)
```

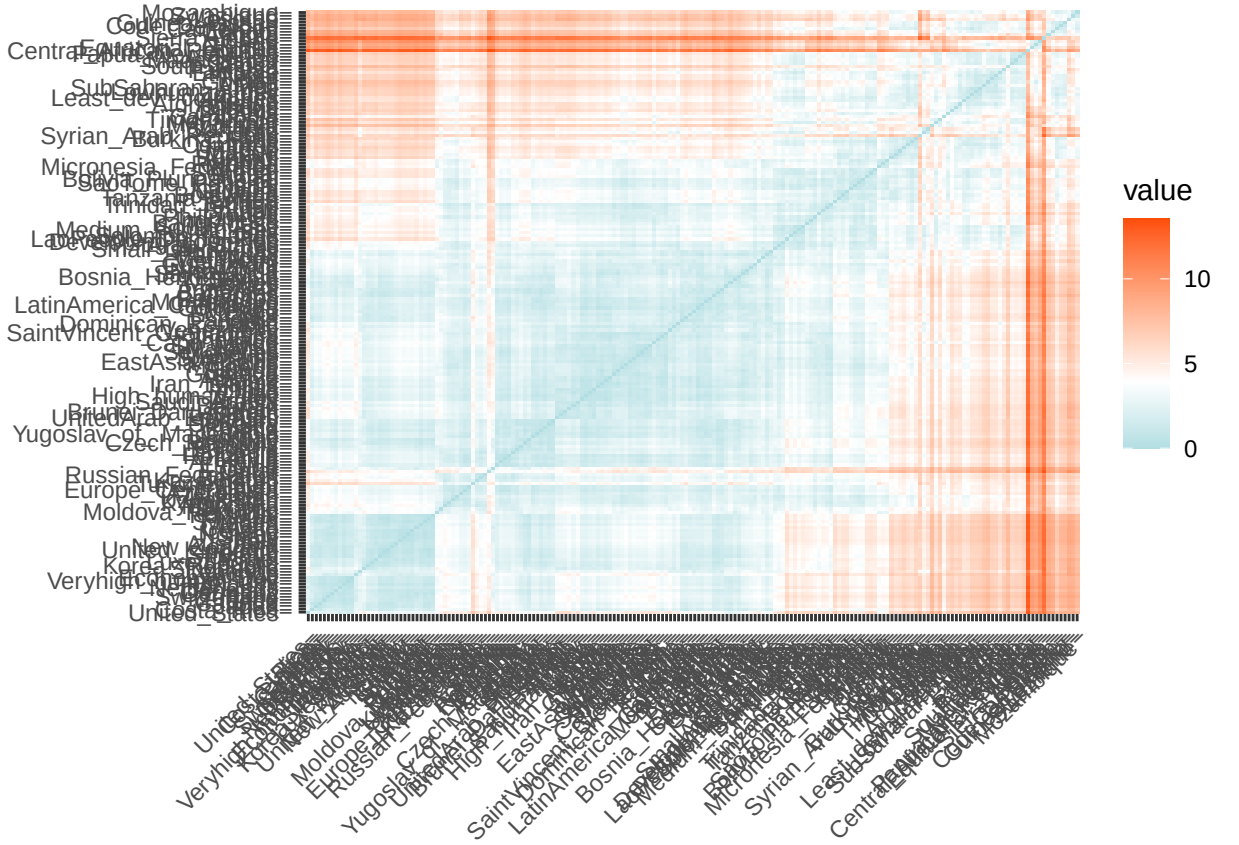
```
##             Norway Netherlands Sweden
## Norway         0.0         1.4    0.5
## Netherlands    1.4         0.0    1.2
## Sweden         0.5         1.2    0.0
```

```
health_cor <- get_dist(health_scale, method = "pearson")
round(as.matrix(health_cor)[1:3, 1:3], 1)
```

```
##             Norway Netherlands Sweden
## Norway         0.0         0.1      0
## Netherlands    0.1         0.0      0
## Sweden         0.0         0.0      0
```

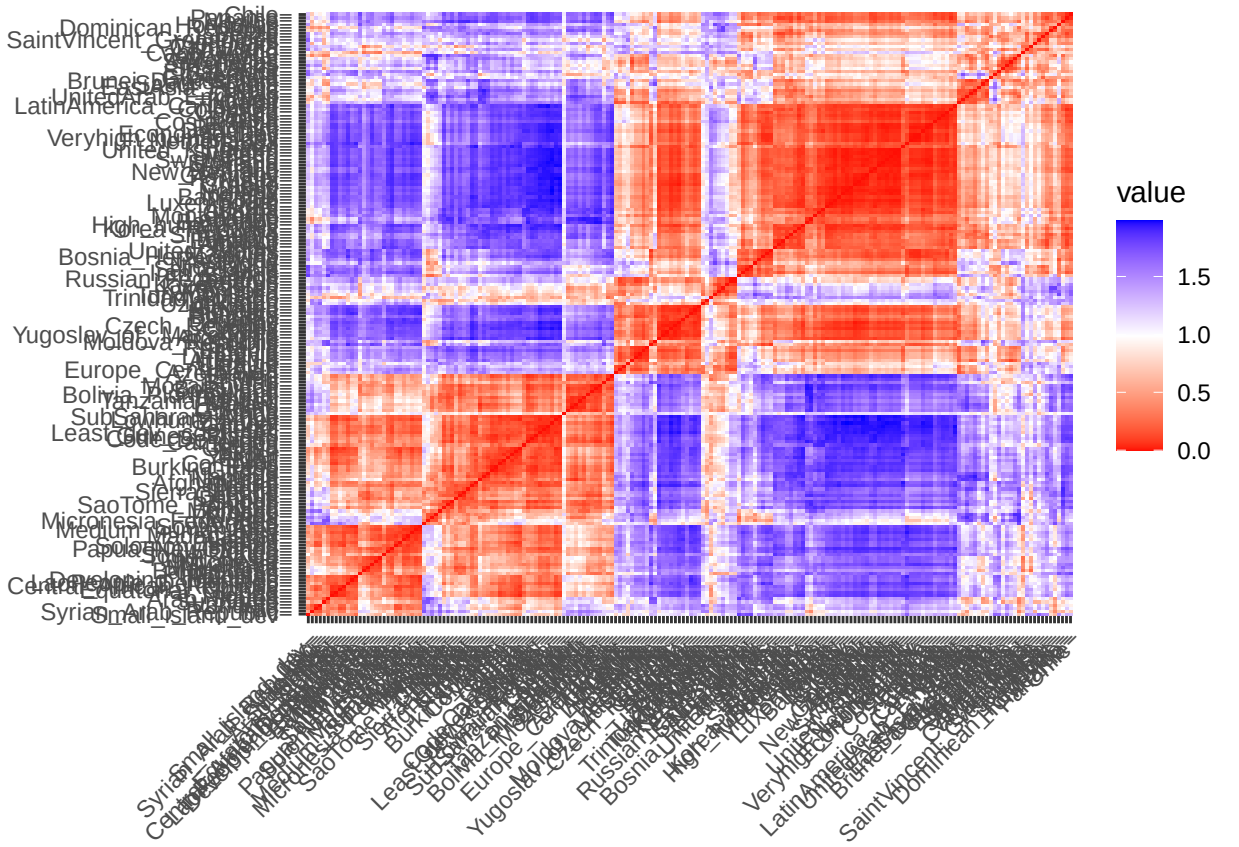
```
#to visualize the distance matrices
#library(factoextra)
```

```
#visualize the euclidean matrix
fviz_dist(health_euc,gradient = list(low = "#00AFBB", mid = "white", high = "#FC4E07"))
```



The red color indicate high similarity while the blue color indicates low similarity. the color level is proportionate to the value of similarity between observations where pure red represents 0 and pure blue represents 1

```
fviz_dist(health_cor)
```



K-MEANS ALGORITHM

```
#Select randomly k objects from the data set as the initial cluster centers or means
health_Kmeans2 <- kmeans(health,centers = 2, nstart = 25)
str(health_Kmeans2)
```

```
## List of 9
##  $ cluster     : Named int [1:192] 1 1 1 1 1 1 1 1 1 1 ...
##   ..- attr(*, "names")= chr [1:192] "Norway" "Netherlands" "Sweden" "Korea _Republic" ...
##  $ centers     : num [1:2, 1:10] 4.29 11.51 8.08 20.65 14.55 ...
##   ..- attr(*, "dimnames")=List of 2
##   .. ..$ : chr [1:2] "1" "2"
##   .. ..$ : chr [1:10] "Infants_DTP" "Infants_Measles" "Infant_Mortality" "UnderFive_Mortality" ...
##  $ totss       : num 4251418
##  $ withinss    : num [1:2] 756122 782988
##  $ tot.withinss: num 1539110
##  $ betweenss   : num 2712308
##  $ size        : int [1:2] 137 55
##  $ iter        : int 1
##  $ ifault      : int 0
##  - attr(*, "class")= chr "kmeans"
```

THE AVAILABLE COMPONENTS LISTED AS THE OUTPUT OF KMEANS ARE AS FOLLOWS:

cluster: A vector of integers (from 1:k) indicating the cluster to which each point is allocated.

centers: A matrix of cluster centers.

totss: The total sum of squares.

withinss: Vector of within-cluster sum of squares, one component per cluster.

tot.withinss: Total within-cluster sum of squares, i.e. sum(withinss).

betweenss: The between-cluster sum of squares, i.e. $totss - tot.withinss$.

size: The number of points in each cluster.

```
health_Kmeans2
```

```
## K-means clustering with 2 clusters of sizes 137, 55
##
## Cluster means:
##   Infants_DTP Infants_Measles Infant_Mortality UnderFive_Mortality
## 1    4.291971        8.080292         14.54672            17.53431
## 2   11.509091       20.654545         54.25273            78.52182
##   Female_Adult_mortality Male_Adult_Mortality Tuberculosis_death
## 1               89.0365             153.3577           6.249635
## 2              269.4182             327.3818          28.894545
##   Life_expectancy Physicians Public_health
## 1        20.65036  19.967153      6.968613
## 2        16.20182   4.134545      5.921818
##
## Clustering vector:
##                   Norway              Netherlands                   Sweden
##                        1                        1                        1
##          Korea _Republic               Luxembourg                    Japan
##                        1                        1                        1
##                  Belgium                   France                  Finland
##                        1                        1                        1
##                    Italy           Czech_Republic                   Greece
##                        1                        1                        1
##                   Cyprus                    Qatar                 Slovakia
##                        1                        1                        1
##                   Poland                    Malta                 Portugal
##                        1                        1                        1
##                  Hungary                  Bahrain                   Kuwait
##                        1                        1                        1
##                  Belarus                     Oman                  Bahamas
##                        1                        1                        1
##               Kazakhstan                 Malaysia                Mauritius
##                        2                        1                        1
##              Seychelles                   Turkey                Sri_Lanka
##                        1                        1                        1
##                   Brazil                  Georgia                  Grenada
##                        1                        1                        1
##                   Jordan   Yugoslav_of_ Macedonia                  Algeria
##                        1                        1                        1
##                  Albania                  Ecuador              Saint_Lucia
##                        1                        1                        1
##                    China                     Fiji                 Thailand
##                        1                        1                        1
##                    Libya                  Tunisia                    Tonga
##                        1                        1                        1
```

```
##                Maldives                   Samoa              Turkmenistan
##                       1                       1                         2
##              Uzbekistan                Nicaragua                   Morocco
##                       1                       1                         1
##              Bangladesh         SaoTome_Principe                Swaziland
##                       1                       1                         2
##          Tanzania_United                  Rwanda                     Sudan
##                       2                       2                         2
##                  Gambia             Switzerland                   Germany
##                       2                       1                         1
##                 Ireland           United_States                    Canada
##                       1                       1                         1
##               Singapore          United_Kingdom                  Slovenia
##                       1                       1                         1
##                   Spain            Saudi_Arabia                   Croatia
##                       1                       1                         1
##              Montenegro                 Uruguay                    Serbia
##                       1                       1                         1
##                    Cuba              Costa_Rica              Iran_Islamic
##                       1                       1                         1
##                Mongolia                Botswana                 Indonesia
##                       2                       2                         1
##             Philippines    Bolivia_Plurinational                Kyrgyzstan
##                       1                       1                         1
##     Micronesia_Federated                  Guyana                Tajikistan
##                       1                       2                         1
##            Sierra_Leone                 Burundi            High_human_dev
##                       2                       2                         1
##       Veryhigh_human_dev         EastAsia_Pacific       Europe_CentralAsia
##                       1                       1                         1
##                  Denmark                 Iceland                 Lithuania
##                       1                       1                         1
##       Russian_Federation                    Peru                   Armenia
##                       2                       1                         1
##                 Colombia                 Jamaica                     Egypt
##                       1                       1                         1
##               ElSalvador               Guatemala                    Bhutan
##                       1                       1                         2
##                   Angola                 Lesotho                   Liberia
##                       2                       2                         2
##                  Eritrea             Economic_Dev                   Estonia
##                       2                       1                         1
##        Brunei_Darussalam                  Latvia                   Romania
##                       1                       1                         1
##                 Bulgaria  SaintVincent_Grenadines                    Belize
##                       1                       1                         1
##         Moldova_Republic                 Senegal                    Malawi
##                       1                       2                         2
##                   Israel              Azerbaijan       Bosnia_Herzegovina
##                       1                       1                         1
##                 Suriname                Kiribati                  Cambodia
##                       1                       1                         1
##                 Cameroon              Mauritania    LatinAmerica_Caribbean
##                       2                       2                         1
```

```
##      UnitedArab_Emirates          Trinidad_Tobago                  Namibia
##                        1                        1                        2
##                    Ghana                    Nepal                    Yemen
##                        2                        1                        2
##                     Togo              Burkina_Faso              New_Zealand
##                        2                        2                        1
##                  Austria                Argentina                 Barbados
##                        1                        1                        1
##                   Panama                 Paraguay                Cabo_Verde
##                        1                        1                        1
##           Code_De_Ivoire               Mozambique                 Australia
##                        2                        2                        1
##                    Chile            Guinea_Bissau              Arab_States
##                        1                        2                        1
##                    World       Dominican_Republic         Small_island_dev
##                        1                        1                        1
##         Medium_human_dev     Developing_countries                Venezuela
##                        1                        1                        1
##                   Mexico                  Ukraine                    Congo
##                        1                        1                        2
##                  Myanmar     Least_dev_ countries       LaoPeople_Demo_Rep
##                        2                        2                        1
##                   Uganda               South_Asia                    India
##                        2                        1                        1
##                 Honduras          Solomon_Islands          Papua_NewGuinea
##                        1                        1                        2
##                 Djibouti              Timor_Leste                   Zambia
##                        2                        1                        2
##              Afghanistan                    Congo                   Comoros
##                        2                        2                        2
##                    Benin                    Niger                   Lebanon
##                        2                        2                        1
##                 Ethiopia             Lowhuman_dev                  Viet_Nam
##                        2                        2                        1
##        SubSaharan_Africa                     Iraq                    Kenya
##                        2                        1                        2
##                     Mali                    Gabon                Madagascar
##                        2                        2                        2
##                 Pakistan                  Vanuatu                   Guinea
##                        1                        1                        2
##             South_Africa                  Nigeria     Syrian_Arab_Republic
##                        2                        2                        1
##        Equatorial_Guinea                     Chad Central_African_Republic
##                        2                        2                        2
##
## Within cluster sum of squares by cluster:
## [1] 756122.2 782987.9
##  (between_SS / total_SS =  63.8 %)
##
## Available components:
##
## [1] "cluster"      "centers"      "totss"        "withinss"     "tot.withinss"
## [6] "betweenss"    "size"         "iter"         "ifault"
```

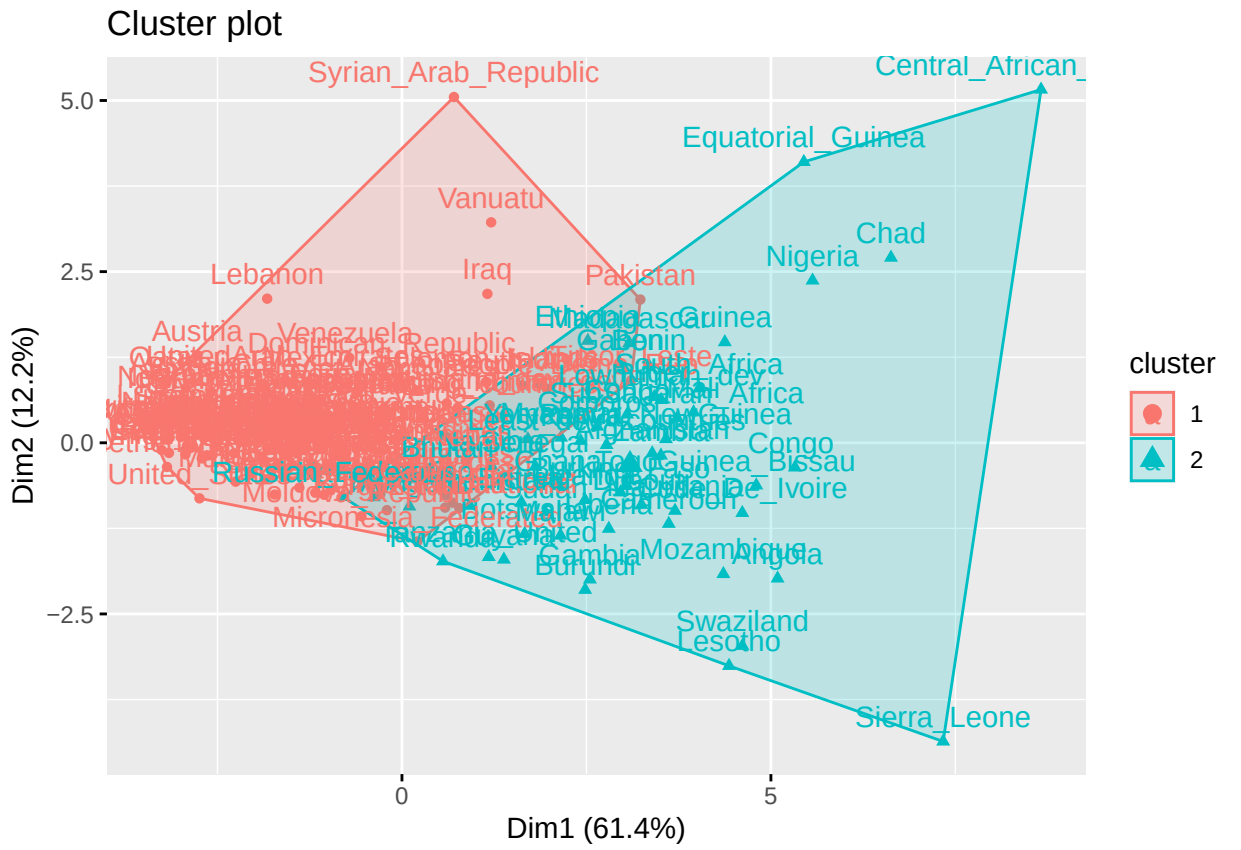
THE RESULT ABOVE SHOWS: That our groupings resulted in 2 cluster sizes of 137 and 55.

Also, the cluster centers (means) for the two groups across the ten variables (Infants_DTP, Infants_Measles, Infant_Mortality, UnderFive_Mortality,Female_Adult_mortality, Male_Adult_Mortality, Tuberculosis_death,Life_expectancy, Physicians, Public_health).

The cluster assignment for each observation (i.e. Nigeria was assigned to cluster 2, Mexico was assigned to cluster 1, etc.). ################

FOR more than two dimensions (variables) fviz_cluster will perform principal component analysis (PCA) and plot the data points according to the first two principal components that explain the majority of the variance.

```
fviz_cluster(health_Kmeans2, data = health)
```



Because the number of clusters (k) must be set before we start the algorithm, it is often advantageous to use several different values of k and examine the differences in the results. We can execute the same process for 3, 4, and 5 clusters, and the results are shown in the figure:

```
#health_Kmeans2 <- kmeans(health, centers = 2, nstart = 25)
#health_Kmeans3 <- kmeans(health, centers = 3, nstart = 25)
#health_Kmeans4 <- kmeans(health, centers = 4, nstart = 25)
#health_Kmeans5 <- kmeans(health, centers = 5, nstart = 25)

# plots to compare
#plot1 <- fviz_cluster(health_Kmeans2, geom = "point", data = health),ggtitle("k = 2")
#plot2 <- fviz_cluster(health_Kmeans3, geom = "point", data = health),ggtitle("k = 3")
#plot3 <- fviz_cluster(health_Kmeans4, geom = "point", data = health),ggtitle("k = 4")
```

```
#plot4 <- fviz_cluster(health_Kmeans5, geom = "point", data = health),ggtitle("k = 5")

library(gridExtra)
```

```
## Warning: package 'gridExtra' was built under R version 4.0.3
```

```
##
## Attaching package: 'gridExtra'
```

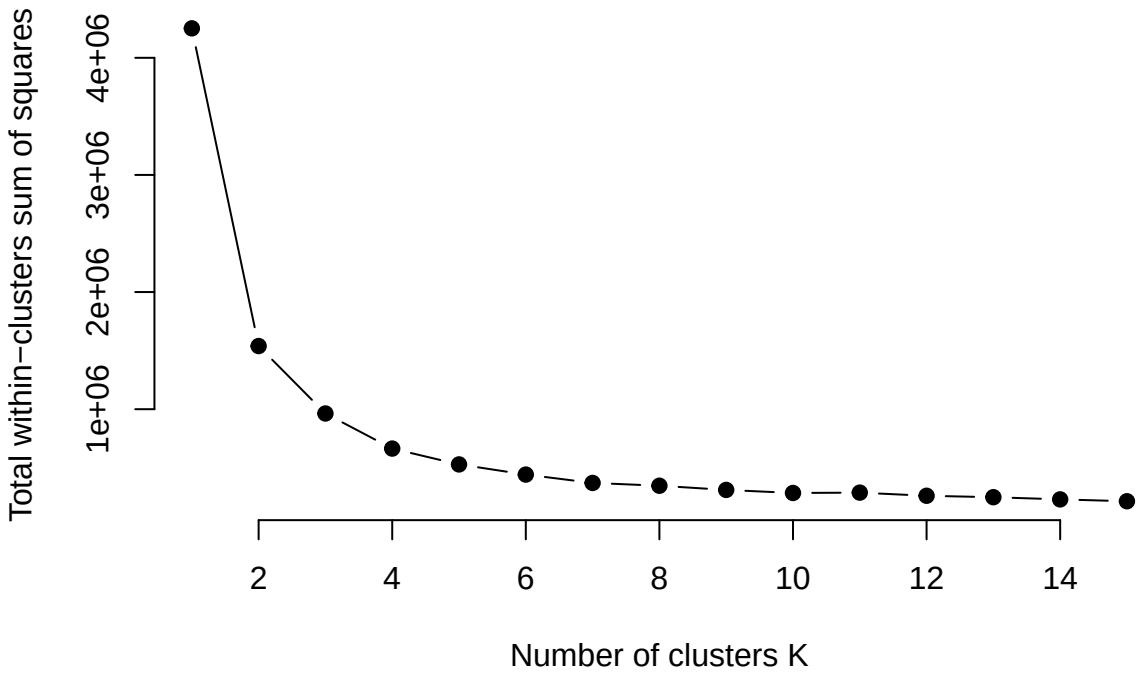
```
## The following object is masked from 'package:dplyr':
##
##     combine
```

```
#grid.arrange(plot1, plot2, plot3, plot4, nrow = 2)
```

```
set.seed(123)

# function to compute total within-cluster sum of square
wss <- function(k) {
  kmeans(health, k, nstart = 10 )$tot.withinss
}

# Compute and plot wss for k = 1 to k = 15
k.values <- 1:15

library(purrr)
```

```
## Warning: package 'purrr' was built under R version 4.0.5
```

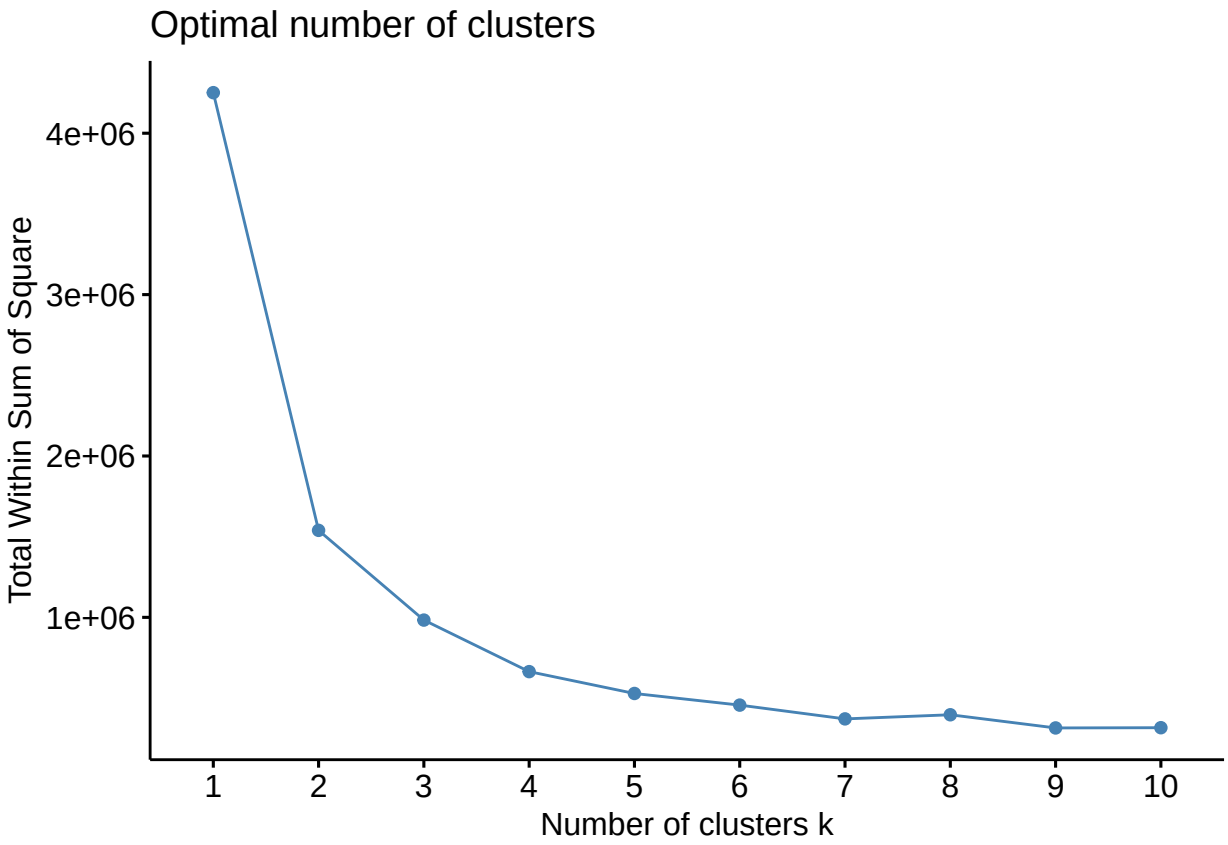
```
# extract wss for 2-15 clusters
wss_values <- map_dbl(k.values, wss)

plot(k.values, wss_values,
       type="b", pch = 19, frame = FALSE,
       xlab="Number of clusters K",
       ylab="Total within-clusters sum of squares")
```



The results suggest that 4 is the optimal number of clusters as it appears to be the bend in the knee (or elbow).

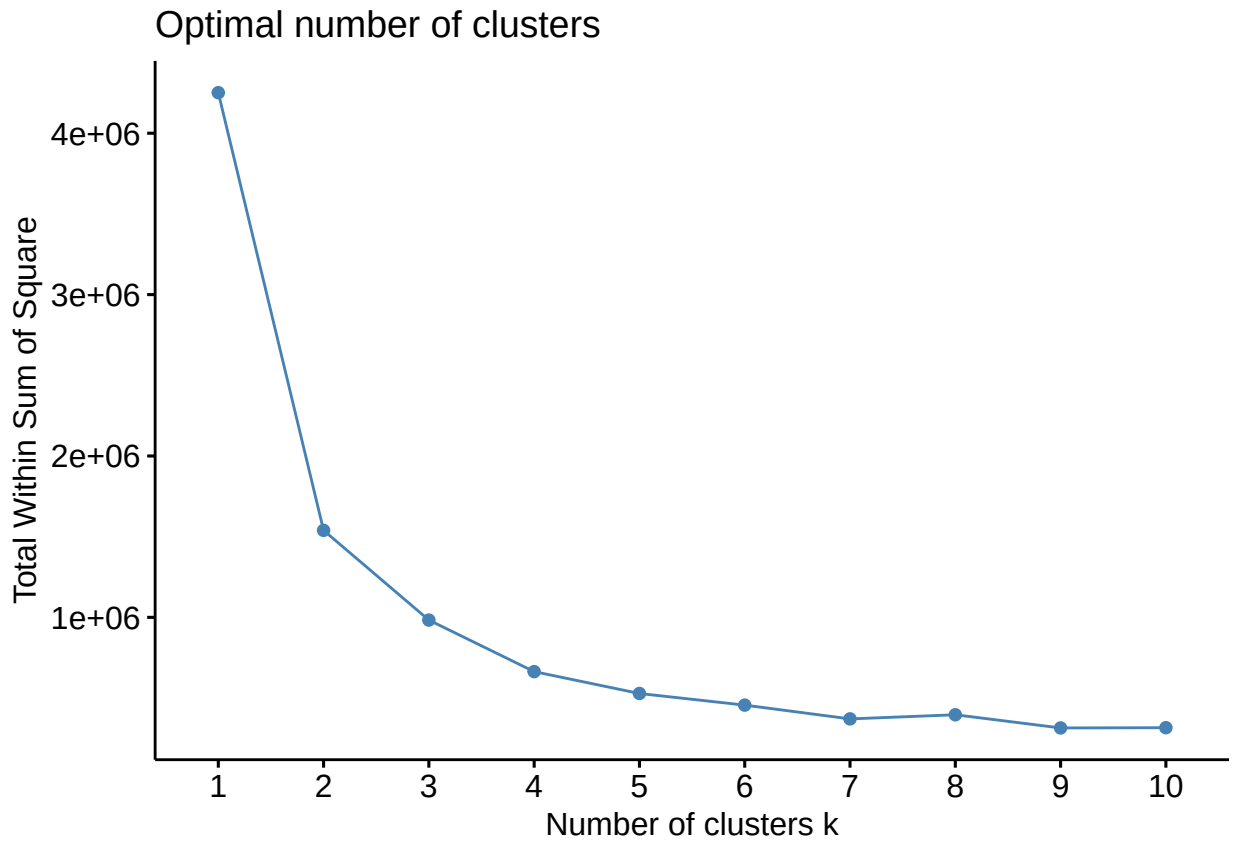
K-MEANS CLUSTERING

```
fviz_nbclust(health, kmeans, method = "wss")
```



The bent is generally is considered as an indicator of the optimal number of clusters. the plot shows variances decreases as the k number increases. a bent on the elbow can be seen as k = 4 , meaning additional clusters beyond 4 has little values.Let's classify our observations into 4 clusters.

```
fviz_nbclust(health, kmeans, method = "wss")
```




```
geom_vline(xintercept = 4, linetype = 2) #the line did not appear bcos of the #sign
```

```
## mapping: xintercept = ~xintercept
## geom_vline: na.rm = FALSE
## stat_identity: na.rm = FALSE
## position_identity
```

```
labs(subtitle = "Elbow method")
```

```
## $subtitle
## [1] "Elbow method"
##
## attr(,"class")
## [1] "labels"
```

To compute k means clustering on a data matrix,We set a seed for random number generator to randomly selected centroids for k means algorthms.

```
# Compute k-means clustering with k = 4
set.seed(123)
final_result <- kmeans(health, 4, nstart = 25)
print(final_result)
```

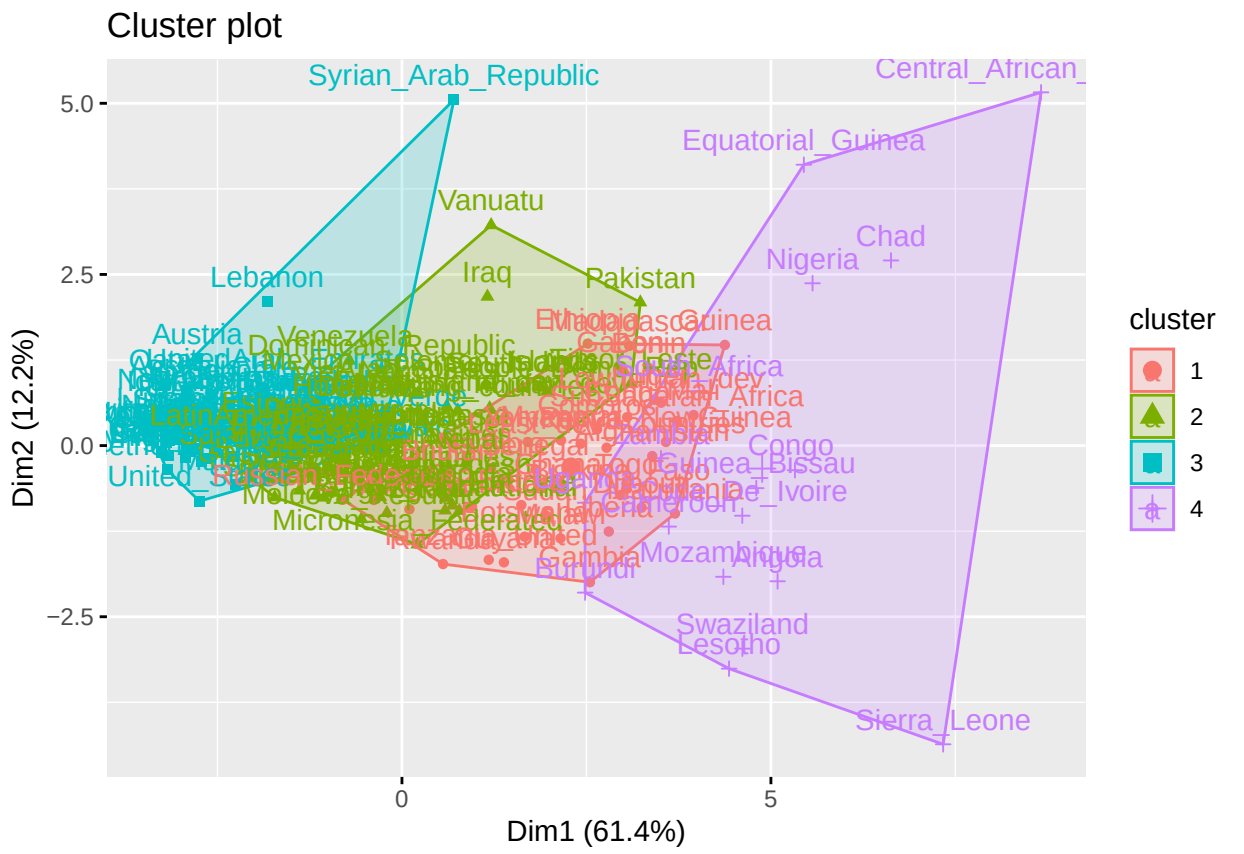
```
## K-means clustering with 4 clusters of sizes 38, 69, 68, 17
```

```
##
## Cluster means:
##   Infants_DTP Infants_Measles Infant_Mortality UnderFive_Mortality
## 1    8.763158       17.947368        46.663158           66.010526
## 2    5.246377        9.898551        21.768116           26.430435
## 3    3.323529        6.235294         7.219118            8.507353
## 4   17.647059       26.705882        71.217647          106.488235
##   Female_Adult_mortality Male_Adult_Mortality Tuberculosis_death
## 1              226.47368             289.5789           24.61053
## 2              117.33333             200.1739           10.65942
## 3               60.32353             105.8529            1.77500
## 4              365.41176             411.8824           38.47059
##   Life_expectancy Physicians Public_health
## 1        16.48158   5.286842      5.710526
## 2        19.20580  14.291304      6.097101
## 3        22.11618  25.726471      7.852941
## 4        15.57647   1.558824      6.394118
##
## Clustering vector:
##                   Norway              Netherlands                   Sweden
##                        3                        3                        3
##          Korea _Republic               Luxembourg                    Japan
##                        3                        3                        3
##                  Belgium                   France                  Finland
##                        3                        3                        3
##                    Italy           Czech_Republic                   Greece
##                        3                        3                        3
##                   Cyprus                    Qatar                 Slovakia
##                        3                        3                        3
##                   Poland                    Malta                 Portugal
##                        2                        3                        3
##                  Hungary                  Bahrain                   Kuwait
##                        2                        3                        3
##                  Belarus                     Oman                  Bahamas
##                        2                        3                        3
##               Kazakhstan                 Malaysia                Mauritius
##                        1                        2                        2
##               Seychelles                   Turkey                Sri_Lanka
##                        2                        3                        2
##                   Brazil                  Georgia                  Grenada
##                        2                        2                        2
##                   Jordan   Yugoslav_of_ Macedonia                  Algeria
##                        3                        3                        2
##                  Albania                  Ecuador              Saint_Lucia
##                        3                        2                        2
##                    China                     Fiji                 Thailand
##                        3                        2                        2
##                    Libya                  Tunisia                    Tonga
##                        3                        3                        2
##                 Maldives                    Samoa              Turkmenistan
##                        3                        2                        1
##               Uzbekistan                Nicaragua                  Morocco
##                        2                        2                        2
##               Bangladesh         SaoTome_Principe                Swaziland
```

```
##                      2                      2                      4
##          Tanzania_United                 Rwanda                  Sudan
##                      1                      1                      1
##                 Gambia            Switzerland                Germany
##                      1                      3                      3
##                Ireland          United_States                 Canada
##                      3                      3                      3
##              Singapore         United_Kingdom               Slovenia
##                      3                      3                      3
##                  Spain           Saudi_Arabia                Croatia
##                      3                      3                      3
##             Montenegro                Uruguay                 Serbia
##                      3                      3                      2
##                   Cuba            Costa_Rica           Iran_Islamic
##                      3                      3                      3
##               Mongolia               Botswana              Indonesia
##                      1                      1                      2
##            Philippines   Bolivia_Plurinational             Kyrgyzstan
##                      2                      2                      2
##    Micronesia_Federated                 Guyana             Tajikistan
##                      2                      1                      2
##           Sierra_Leone                Burundi         High_human_dev
##                      4                      4                      3
##      Veryhigh_human_dev       EastAsia_Pacific      Europe_CentralAsia
##                      3                      3                      2
##                Denmark                Iceland              Lithuania
##                      3                      3                      2
##      Russian_Federation                   Peru                Armenia
##                      1                      3                      2
##                Colombia                Jamaica                  Egypt
##                      3                      2                      2
##             ElSalvador              Guatemala                 Bhutan
##                      2                      2                      1
##                 Angola                Lesotho                Liberia
##                      4                      4                      1
##                Eritrea           Economic_Dev                Estonia
##                      1                      3                      2
##       Brunei_Darussalam                 Latvia                Romania
##                      3                      2                      2
##               Bulgaria  SaintVincent_Grenadines                 Belize
##                      2                      2                      3
##        Moldova_Republic                Senegal                 Malawi
##                      2                      1                      1
##                 Israel             Azerbaijan     Bosnia_Herzegovina
##                      3                      2                      3
##               Suriname               Kiribati               Cambodia
##                      2                      2                      2
##               Cameroon             Mauritania  LatinAmerica_Caribbean
##                      4                      1                      2
##     UnitedArab_Emirates        Trinidad_Tobago                Namibia
##                      3                      2                      1
##                  Ghana                  Nepal                  Yemen
##                      1                      2                      1
##                   Togo           Burkina_Faso            New_Zealand
```

```
##                       1                       1                       3
##                 Austria               Argentina                Barbados
##                       3                       3                       3
##                  Panama                Paraguay              Cabo_Verde
##                       3                       2                       3
##          Code_De_Ivoire              Mozambique               Australia
##                       4                       4                       3
##                   Chile           Guinea_Bissau             Arab_States
##                       3                       4                       2
##                   World      Dominican_Republic         Small_island_dev
##                       2                       2                       2
##        Medium_human_dev     Developing_countries              Venezuela
##                       2                       2                       2
##                  Mexico                 Ukraine                   Congo
##                       2                       2                       1
##                 Myanmar     Least_dev_ countries       LaoPeople_Demo_Rep
##                       1                       1                       2
##                  Uganda              South_Asia                   India
##                       4                       2                       2
##                Honduras         Solomon_Islands         Papua_NewGuinea
##                       2                       2                       1
##                Djibouti             Timor_Leste                  Zambia
##                       1                       2                       4
##             Afghanistan                   Congo                  Comoros
##                       1                       4                       1
##                   Benin                   Niger                  Lebanon
##                       1                       1                       3
##                Ethiopia             Lowhuman_dev                Viet_Nam
##                       1                       1                       2
##       SubSaharan_Africa                    Iraq                   Kenya
##                       1                       2                       1
##                    Mali                   Gabon              Madagascar
##                       1                       1                       1
##                Pakistan                 Vanuatu                  Guinea
##                       2                       2                       1
##            South_Africa                 Nigeria     Syrian_Arab_Republic
##                       4                       4                       3
##       Equatorial_Guinea                    Chad Central_African_Republic
##                       4                       4                       4
##
## Within cluster sum of squares by cluster:
## [1] 171201.28 220402.40  93307.34 178776.31
##  (between_SS / total_SS =  84.4 %)
##
## Available components:
##
## [1] "cluster"      "centers"      "totss"        "withinss"     "tot.withinss"
## [6] "betweenss"    "size"         "iter"         "ifault"
```

```
fviz_cluster(final_result, data = health)
```




```
health %>%
  mutate(Cluster = final_result$cluster) %>%
  group_by(Cluster) %>%
  summarise_all("mean")
```

```
## # A tibble: 4 x 11
##   Cluster Infants_DTP Infants_Measles Infant_Mortality UnderFive_Mortality
##     <int>       <dbl>           <dbl>            <dbl>                <dbl>
## 1       1        8.76            17.9             46.7                 66.0
## 2       2        5.25             9.90            21.8                 26.4
## 3       3        3.32             6.24             7.22                 8.51
## 4       4       17.6             26.7             71.2                106.
## # ... with 6 more variables: Female_Adult_mortality <dbl>,
## #   Male_Adult_Mortality <dbl>, Tuberculosis_death <dbl>,
## #   Life_expectancy <dbl>, Physicians <dbl>, Public_health <dbl>
```

OR SIMPLY PUT

```
aggregate(health, by=list(cluster= final_result$cluster), mean)
```

```
##   cluster Infants_DTP Infants_Measles Infant_Mortality UnderFive_Mortality
## 1       1    8.763158       17.947368        46.663158           66.010526
## 2       2    5.246377        9.898551        21.768116           26.430435
## 3       3    3.323529        6.235294         7.219118            8.507353
```

```
## 4       4   17.647059       26.705882        71.217647          106.488235
##   Female_Adult_mortality Male_Adult_Mortality Tuberculosis_death
## 1              226.47368             289.5789           24.61053
## 2              117.33333             200.1739           10.65942
## 3               60.32353             105.8529            1.77500
## 4              365.41176             411.8824           38.47059
##   Life_expectancy Physicians Public_health
## 1        16.48158   5.286842      5.710526
## 2        19.20580  14.291304      6.097101
## 3        22.11618  25.726471      7.852941
## 4        15.57647   1.558824      6.394118
```

This computes summary statistics of data subsets

```
#combine R objects by rows and columns

point_class <- cbind(health, cluster = final_result$cluster)
head(point_class)
```

```
##                 Infants_DTP Infants_Measles Infant_Mortality
## Norway                    1               7              2.3
## Netherlands               1               4              3.3
## Sweden                    1               3              2.4
## Korea _Republic           1               1              3.2
## Luxembourg                1               5              1.6
## Japan                     1               5              2.1
##                 UnderFive_Mortality Female_Adult_mortality Male_Adult_Mortality
## Norway                          2.8                      47                    73
## Netherlands                     4.0                      54                    69
## Sweden                          3.0                      43                    69
## Korea _Republic                 3.7                      38                    93
## Luxembourg                      2.0                      50                    79
## Japan                           2.9                      42                    81
##                 Tuberculosis_death Life_expectancy Physicians Public_health
## Norway                         0.1            24.0       37.4           9.6
## Netherlands                    0.2            23.5       31.5          12.9
## Sweden                         0.1            24.1       32.7           9.7
## Korea _Republic                5.4            24.0       21.4           7.2
## Luxembourg                     0.4            23.4       28.2           7.1
## Japan                          1.7            26.1       23.0          10.3
##                 cluster
## Norway                3
## Netherlands           3
## Sweden                3
## Korea _Republic       3
## Luxembourg            3
## Japan                 3
```

```
 final_result
```

```
## K-means clustering with 4 clusters of sizes 38, 69, 68, 17
##
## Cluster means:
```

```
##   Infants_DTP Infants_Measles Infant_Mortality UnderFive_Mortality
## 1    8.763158       17.947368        46.663158           66.010526
## 2    5.246377        9.898551        21.768116           26.430435
## 3    3.323529        6.235294         7.219118            8.507353
## 4   17.647059       26.705882        71.217647          106.488235
##   Female_Adult_mortality Male_Adult_Mortality Tuberculosis_death
## 1              226.47368             289.5789           24.61053
## 2              117.33333             200.1739           10.65942
## 3               60.32353             105.8529            1.77500
## 4              365.41176             411.8824           38.47059
##   Life_expectancy Physicians Public_health
## 1        16.48158   5.286842      5.710526
## 2        19.20580  14.291304      6.097101
## 3        22.11618  25.726471      7.852941
## 4        15.57647   1.558824      6.394118
##
## Clustering vector:
##                   Norway              Netherlands                   Sweden
##                        3                        3                        3
##          Korea _Republic               Luxembourg                    Japan
##                        3                        3                        3
##                  Belgium                   France                  Finland
##                        3                        3                        3
##                    Italy            Czech_Republic                   Greece
##                        3                        3                        3
##                   Cyprus                    Qatar                 Slovakia
##                        3                        3                        3
##                   Poland                   Malta                 Portugal
##                        2                        3                        3
##                  Hungary                  Bahrain                   Kuwait
##                        2                        3                        3
##                  Belarus                     Oman                  Bahamas
##                        2                        3                        3
##               Kazakhstan                 Malaysia                Mauritius
##                        1                        2                        2
##               Seychelles                   Turkey                Sri_Lanka
##                        2                        3                        2
##                   Brazil                  Georgia                  Grenada
##                        2                        2                        2
##                   Jordan   Yugoslav_of_ Macedonia                  Algeria
##                        3                        3                        2
##                  Albania                  Ecuador              Saint_Lucia
##                        3                        2                        2
##                    China                     Fiji                 Thailand
##                        3                        2                        2
##                    Libya                  Tunisia                    Tonga
##                        3                        3                        2
##                 Maldives                    Samoa             Turkmenistan
##                        3                        2                        1
##               Uzbekistan                Nicaragua                  Morocco
##                        2                        2                        2
##               Bangladesh         SaoTome_Principe                Swaziland
##                        2                        2                        4
##          Tanzania_United                   Rwanda                    Sudan
```

```
##                   1                   1                   1
##              Gambia         Switzerland             Germany
##                   1                   3                   3
##             Ireland       United_States              Canada
##                   3                   3                   3
##           Singapore      United_Kingdom            Slovenia
##                   3                   3                   3
##               Spain        Saudi_Arabia             Croatia
##                   3                   3                   3
##          Montenegro             Uruguay              Serbia
##                   3                   3                   2
##                Cuba          Costa_Rica        Iran_Islamic
##                   3                   3                   3
##            Mongolia            Botswana           Indonesia
##                   1                   1                   2
##         Philippines Bolivia_Plurinational          Kyrgyzstan
##                   2                   2                   2
##  Micronesia_Federated              Guyana          Tajikistan
##                   2                   1                   2
##        Sierra_Leone             Burundi       High_human_dev
##                   4                   4                   3
##   Veryhigh_human_dev     EastAsia_Pacific    Europe_CentralAsia
##                   3                   3                   2
##             Denmark             Iceland           Lithuania
##                   3                   3                   2
##   Russian_Federation                Peru             Armenia
##                   1                   3                   2
##            Colombia             Jamaica               Egypt
##                   3                   2                   2
##           ElSalvador           Guatemala              Bhutan
##                   2                   2                   1
##              Angola             Lesotho             Liberia
##                   4                   4                   1
##             Eritrea        Economic_Dev             Estonia
##                   1                   3                   2
##    Brunei_Darussalam              Latvia             Romania
##                   3                   2                   2
##            Bulgaria SaintVincent_Grenadines              Belize
##                   2                   2                   3
##     Moldova_Republic             Senegal              Malawi
##                   2                   1                   1
##              Israel          Azerbaijan   Bosnia_Herzegovina
##                   3                   2                   3
##           Suriname            Kiribati            Cambodia
##                   2                   2                   2
##           Cameroon          Mauritania LatinAmerica_Caribbean
##                   4                   1                   2
##  UnitedArab_Emirates      Trinidad_Tobago             Namibia
##                   3                   2                   1
##              Ghana               Nepal               Yemen
##                   1                   2                   1
##               Togo        Burkina_Faso         New_Zealand
##                   1                   1                   3
##            Austria           Argentina            Barbados
```

```
##                       3                       3                       3
##                  Panama                Paraguay              Cabo_Verde
##                       3                       2                       3
##          Code_De_Ivoire              Mozambique               Australia
##                       4                       4                       3
##                   Chile           Guinea_Bissau             Arab_States
##                       3                       4                       2
##                   World      Dominican_Republic         Small_island_dev
##                       2                       2                       2
##         Medium_human_dev     Developing_countries               Venezuela
##                       2                       2                       2
##                  Mexico                 Ukraine                   Congo
##                       2                       2                       1
##                 Myanmar     Least_dev_ countries       LaoPeople_Demo_Rep
##                       1                       1                       2
##                  Uganda              South_Asia                   India
##                       4                       2                       2
##                Honduras          Solomon_Islands          Papua_NewGuinea
##                       2                       2                       1
##                Djibouti             Timor_Leste                  Zambia
##                       1                       2                       4
##             Afghanistan                   Congo                   Comoros
##                       1                       4                       1
##                   Benin                   Niger                   Lebanon
##                       1                       1                       3
##                Ethiopia             Lowhuman_dev                Viet_Nam
##                       1                       1                       2
##        SubSaharan_Africa                    Iraq                   Kenya
##                       1                       2                       1
##                    Mali                   Gabon              Madagascar
##                       1                       1                       1
##                Pakistan                 Vanuatu                   Guinea
##                       2                       2                       1
##             South_Africa                 Nigeria     Syrian_Arab_Republic
##                       4                       4                       3
##        Equatorial_Guinea                    Chad Central_African_Republic
##                       4                       4                       4
##
## Within cluster sum of squares by cluster:
## [1] 171201.28 220402.40  93307.34 178776.31
##  (between_SS / total_SS =  84.4 %)
##
## Available components:
##
## [1] "cluster"      "centers"      "totss"        "withinss"     "tot.withinss"
## [6] "betweenss"    "size"         "iter"         "ifault"
```

```
final_result$size
```

```
## [1] 38 69 68 17
```

HIERARCHICAL CLUSTERING

```
#df_scale <- scale(df)
#head(df_scale)
```

```
require(stats)
health_dist <- dist(x = health_scale, method = "euclidean")
```

```
x <- as.matrix(health_dist)[1:6, 1:6]
x
```

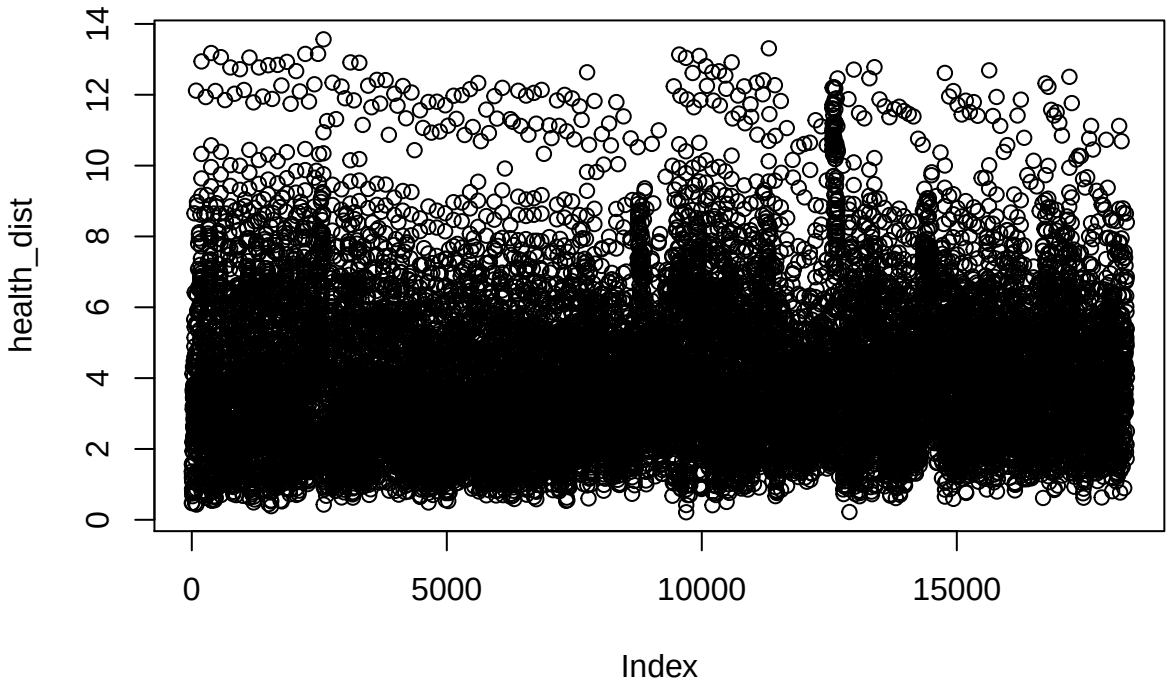
```
##                   Norway Netherlands    Sweden Korea _Republic Luxembourg
## Norway          0.0000000    1.362041 0.4726441       1.5689876  1.1826361
## Netherlands     1.3620413    0.000000 1.2492717       2.3395646  2.2308142
## Sweden          0.4726441    1.249272 0.0000000       1.3034367  1.0882037
## Korea _Republic 1.5689876    2.339565 1.3034367       0.0000000  0.6980569
## Luxembourg      1.1826361    2.230814 1.0882037       0.6980569  0.0000000
## Japan           1.2807704    1.472225 1.0126954       1.4448686  1.5751845
##                    Japan
## Norway          1.280770
## Netherlands     1.472225
## Sweden          1.012695
## Korea _Republic 1.444869
## Luxembourg      1.575184
## Japan           0.000000
```

```
round(x, digits = 3)
```

```
##                 Norway Netherlands Sweden Korea _Republic Luxembourg Japan
## Norway           0.000       1.362  0.473           1.569      1.183 1.281
## Netherlands      1.362       0.000  1.249           2.340      2.231 1.472
## Sweden           0.473       1.249  0.000           1.303      1.088 1.013
## Korea _Republic  1.569       2.340  1.303           0.000      0.698 1.445
## Luxembourg       1.183       2.231  1.088           0.698      0.000 1.575
## Japan            1.281       1.472  1.013           1.445      1.575 0.000
```

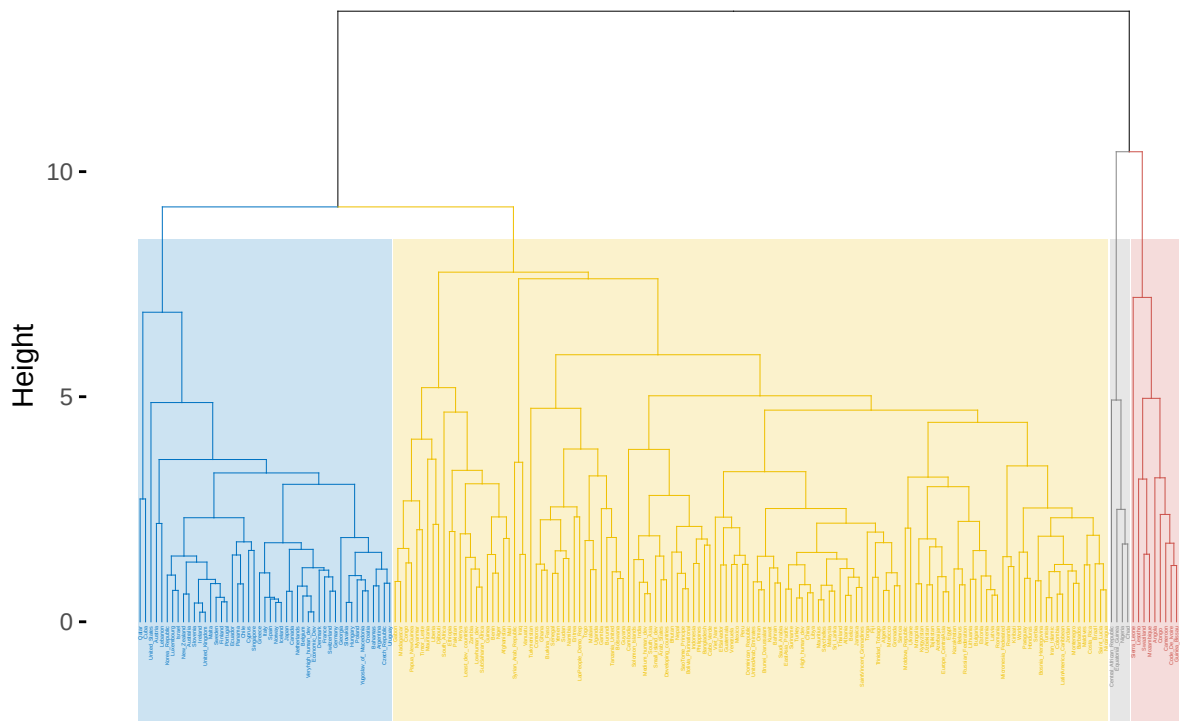
HIERARCHICAL CLUSTER ANALYSIS

```
#LOAD STATS LIBRARY
require(stats)

#FOR HIERARCHICAL CLUSTERING
health_hc <- hclust(d = health_dist,method = "complete")

plot(x = health_dist)
```




```
fviz_dend(x = health_hc, cex = 0.15, lwd = 0.18, k = 4,
 rect = TRUE,
          k_colors = "jco",
         rect_border = "jco",
          rect_fill = TRUE)
```

```
## Warning: `guides(<scale> = FALSE)` is deprecated. Please use `guides(<scale> =
## "none")` instead.
```




```
require(grDevices)
require(scales)
```

```
## Loading required package: scales
```

```
##
## Attaching package: 'scales'
```

```
## The following object is masked from 'package:purrr':
##
##     discard
```

```
## The following object is masked from 'package:readr':
##
##     col_factor
```

```
require("ggsci")
```

```
## Loading required package: ggsci
```

```
## Warning: package 'ggsci' was built under R version 4.0.5
```

```
require(igraph)
```

```
## Loading required package: igraph
```

```
## Warning: package 'igraph' was built under R version 4.0.5
```

```
##
## Attaching package: 'igraph'
```

```
## The following objects are masked from 'package:purrr':
##
##     compose, simplify
```

```
## The following objects are masked from 'package:dplyr':
##
##     as_data_frame, groups, union
```

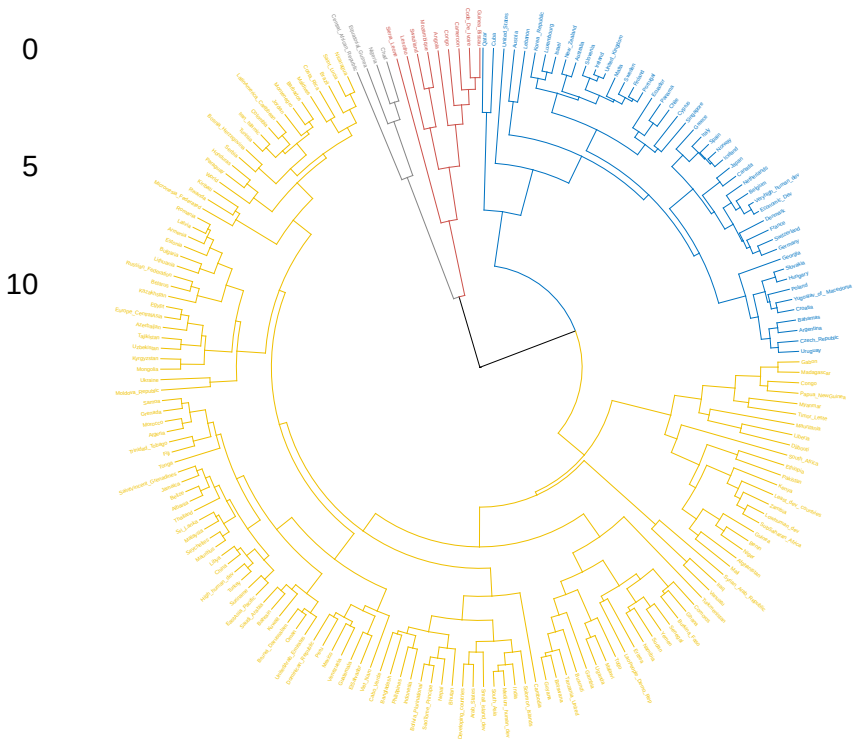
```
## The following objects are masked from 'package:stats':
##
##     decompose, spectrum
```

```
## The following object is masked from 'package:base':
##
##     union
```

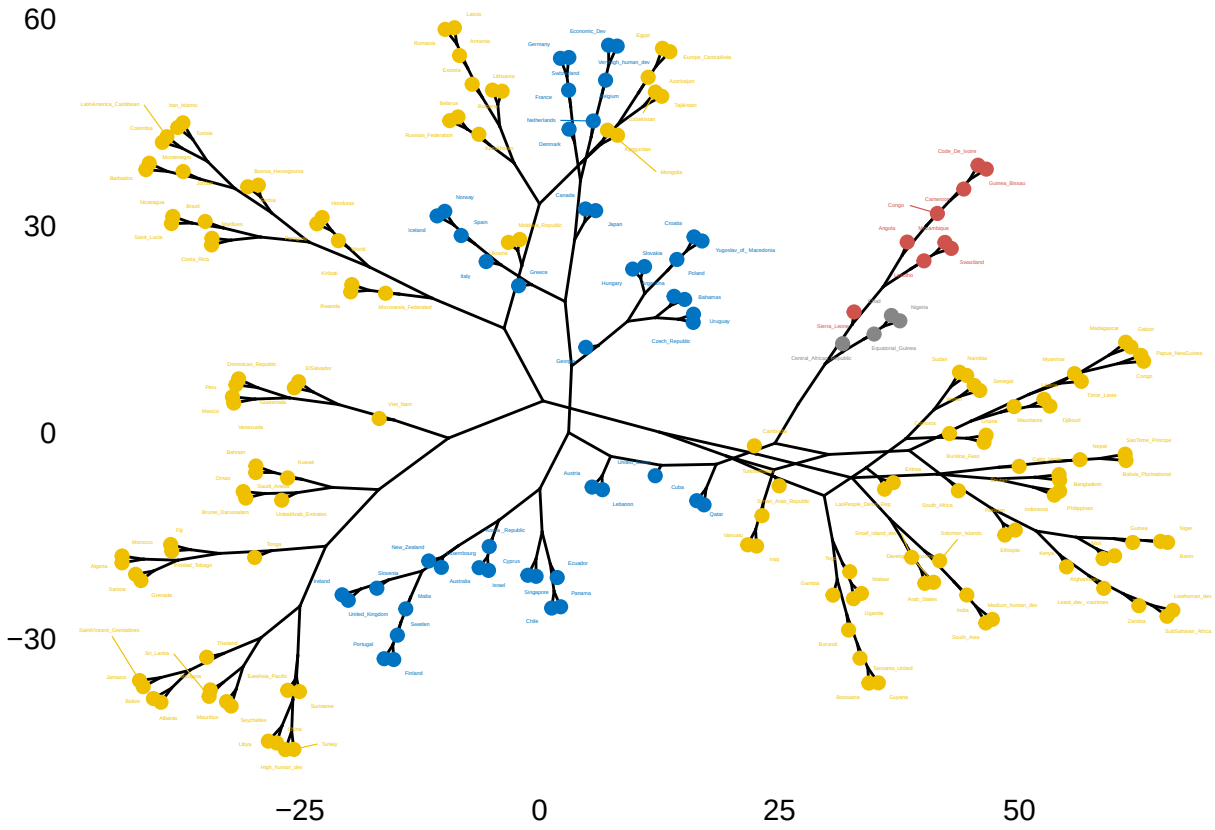
TO ASSIGN COLORS TO THE DENDOGRAM PLOT

```
fviz_dend(x = health_hc, cex = 0.15, lwd = 0.18, k = 4,
rect = TRUE,
          k_colors = "jco",
         rect_border = "jco",
          rect_fill = TRUE,
         type = "circular")
```

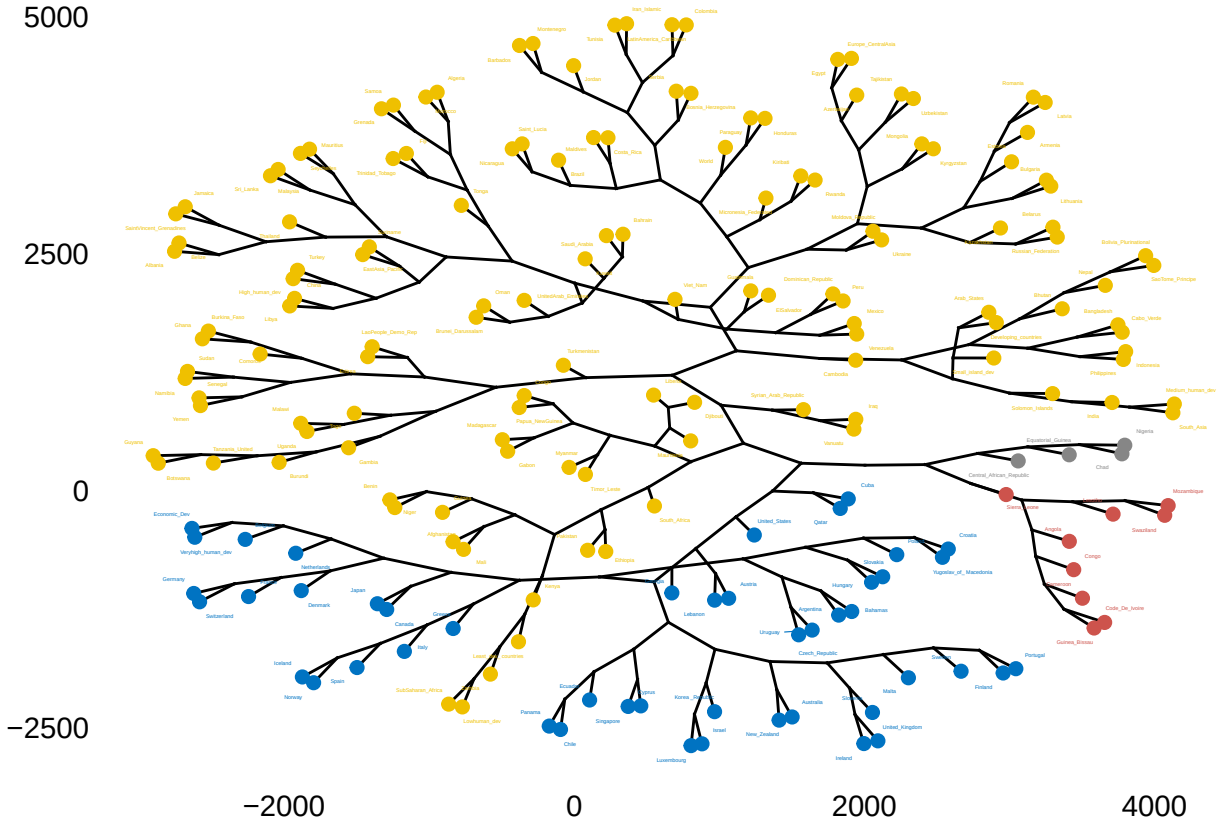
```
## Warning: `guides(<scale> = FALSE)` is deprecated. Please use `guides(<scale> =
## "none")` instead.
```




```
fviz_dend(x = health_hc, cex = 0.15, lwd = 0.18, k = 4,
rect = TRUE,
          k_colors = "jco",
         rect_border = "jco",
          rect_fill = TRUE,
         type = "phylogenic",
         repel = TRUE)
```

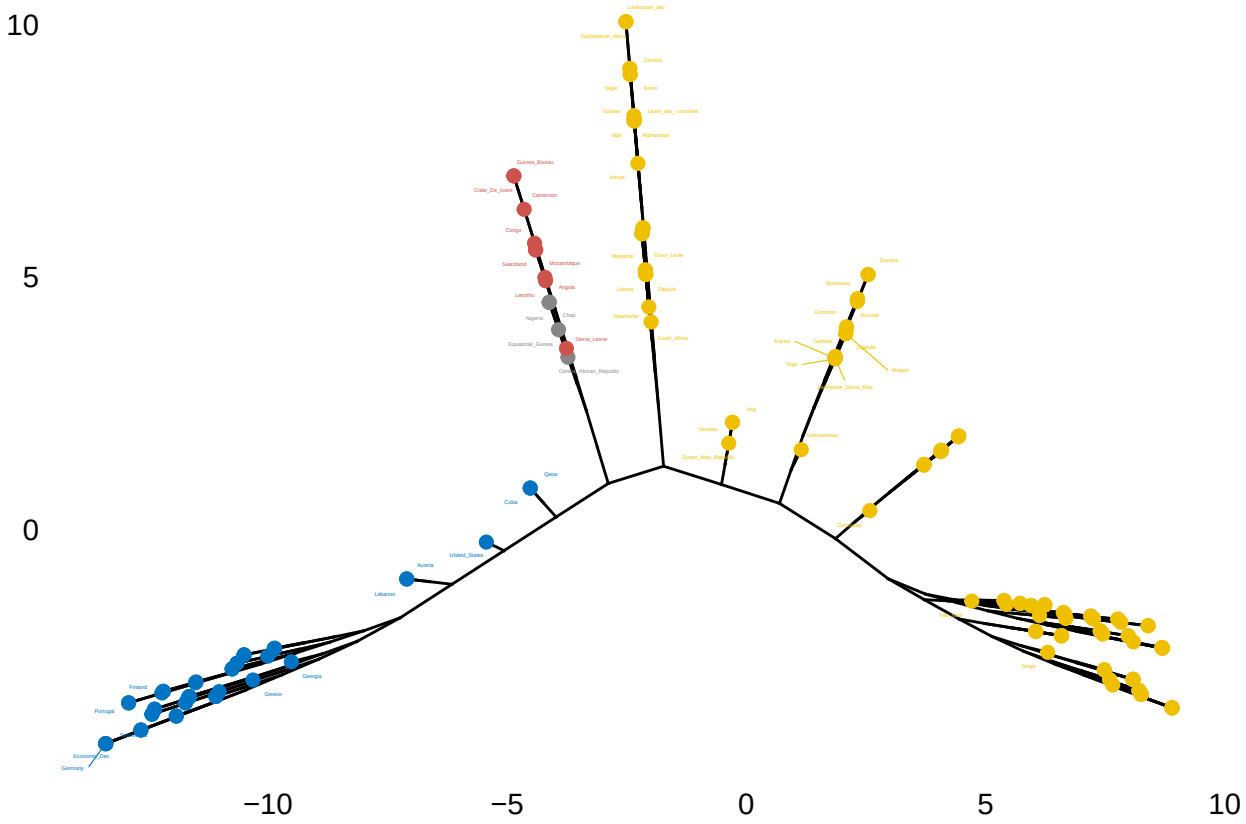



```
fviz_dend(health_hc, cex = 0.15, lwd = 0.18, k = 4,
          rect = TRUE,
          k_colors = "jco",
         rect_border = "jco",
          rect_fill = TRUE,
         type = "phylogenic",
         repel = TRUE,
         phylo_layout = "layout.gem")
```



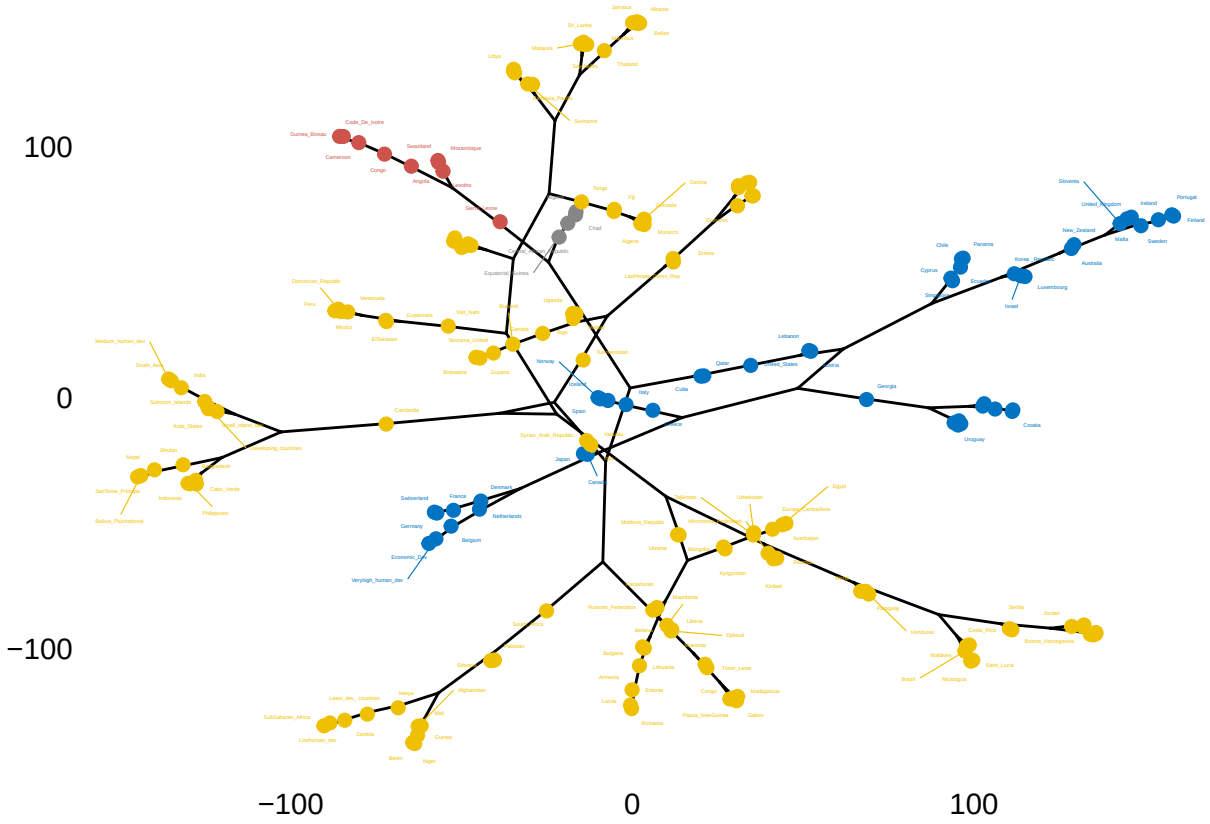

```
fviz_dend(health_hc, cex = 0.15, lwd = 0.18, k = 4,
          rect = TRUE,
          k_colors = "jco",
         rect_border = "jco",
          rect_fill = TRUE,
         type = "phylogenic",
         repel = TRUE,
         phylo_layout = "layout.mds")
```

```
## Warning: ggrepel: 134 unlabeled data points (too many overlaps). Consider
## increasing max.overlaps
```

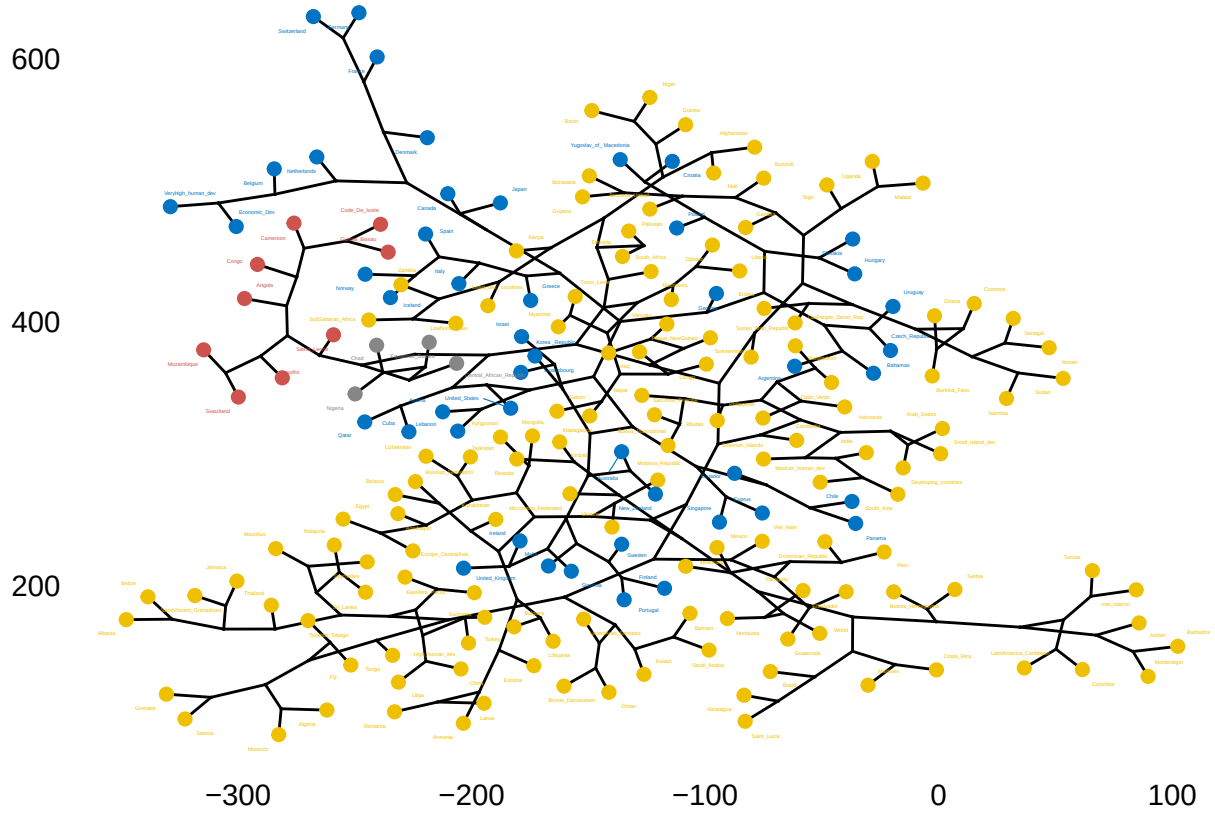



```
fviz_dend(health_hc, cex = 0.15, lwd = 0.18, k = 4,
          rect = TRUE,
          k_colors = "jco",
         rect_border = "jco",
          rect_fill = TRUE,
         type = "phylogenic",
         repel = TRUE,
         phylo_layout = "layout_with_drl")
```

```
## Warning: ggrepel: 30 unlabeled data points (too many overlaps). Consider
## increasing max.overlaps
```

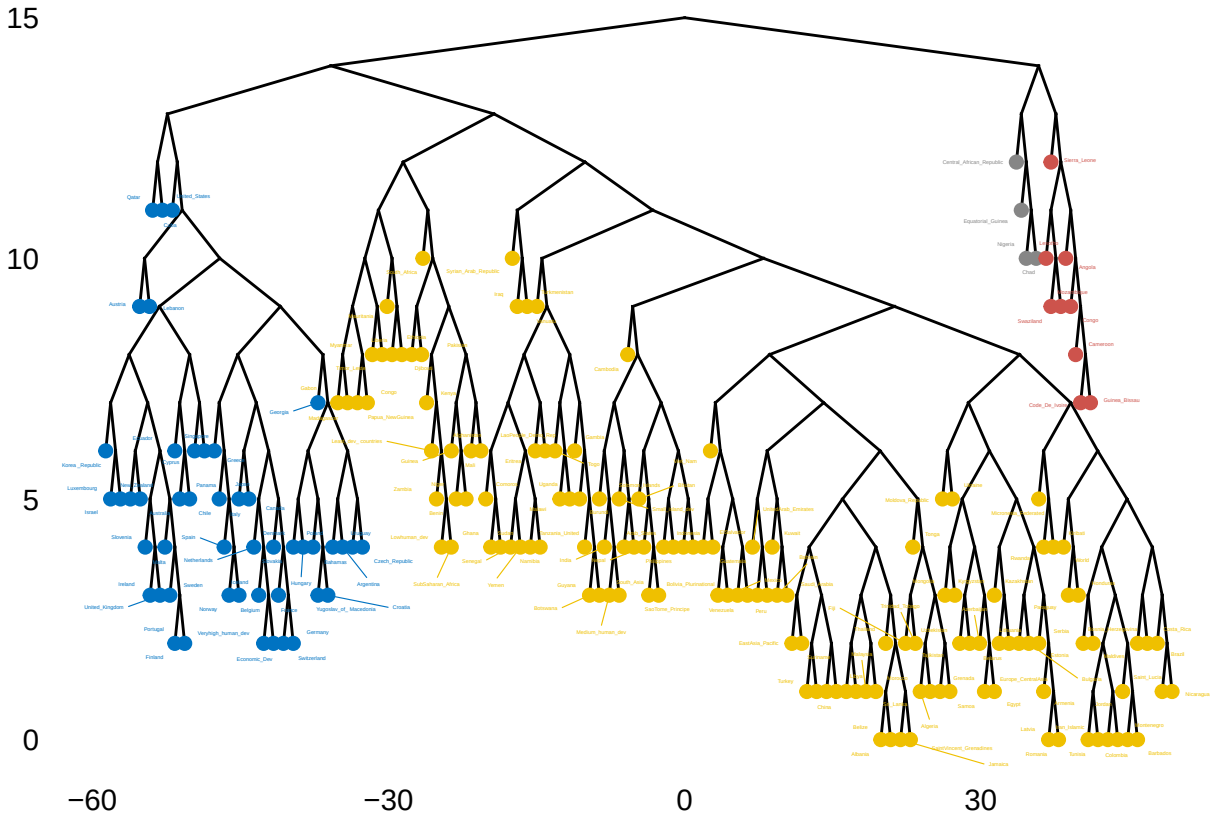



```
fviz_dend(health_hc, cex = 0.15, lwd = 0.18, k = 4,
          rect = TRUE,
          k_colors = "jco",
         rect_border = "jco",
          rect_fill = TRUE,
         type = "phylogenic",
         repel = TRUE,
         phylo_layout = "layout_with_lgl")
```




```
fviz_dend(health_hc, cex = 0.15, lwd = 0.18, k = 4,
          rect = TRUE,
          k_colors = "jco",
         rect_border = "jco",
          rect_fill = TRUE,
         type = "phylogenic",
         repel = TRUE,
         phylo_layout = "layout_as_tree")
```

```
## Warning: ggrepel: 12 unlabeled data points (too many overlaps). Consider
## increasing max.overlaps
```




```
fviz_dend(health_hc, cex = 0.15, lwd = 0.18, k = 4,
          rect = TRUE,
          k_colors = "jco",
          rect_border = "jco",
          rect_fill = FALSE)
```

```
## Warning: `guides(<scale> = FALSE)` is deprecated. Please use `guides(<scale> =
## "none")` instead.
```

Cluster Dendrogram