Web Traffic Time Series Forecasting

library(tidyverse)

## -- Attaching packages --------------------------------------- tidyverse 1.3.1 --

## v ggplot2 3.3.5 v purrr 0.3.4  
## v tibble 3.1.6 v dplyr 1.0.7  
## v tidyr 1.1.4 v stringr 1.4.0  
## v readr 2.1.1 v forcats 0.5.1

## -- Conflicts ------------------------------------------ tidyverse\_conflicts() --  
## x dplyr::filter() masks stats::filter()  
## x dplyr::lag() masks stats::lag()

library(cluster)  
library(haven)  
library(ggdendro)  
library(NbClust)  
library(factoextra)

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

library(klaR)

## Loading required package: MASS

##   
## Attaching package: 'MASS'

## The following object is masked from 'package:dplyr':  
##   
## select

library(data.table)

##   
## Attaching package: 'data.table'

## The following objects are masked from 'package:dplyr':  
##   
## between, first, last

## The following object is masked from 'package:purrr':  
##   
## transpose

library(rlang)

##   
## Attaching package: 'rlang'

## The following object is masked from 'package:data.table':  
##   
## :=

## The following objects are masked from 'package:purrr':  
##   
## %@%, as\_function, flatten, flatten\_chr, flatten\_dbl, flatten\_int,  
## flatten\_lgl, flatten\_raw, invoke, list\_along, modify, prepend,  
## splice

library(dplyr)  
library(NbClust)  
library(ggpubr)  
library(readr)  
theme\_set(theme\_pubr())  
library(corrplot)

## corrplot 0.92 loaded

data <- read.csv("C:/Users/HP/OneDrive/Desktop/africa.csv")  
head(data)

## pop emp emp\_to\_pop\_ratio hc ccon cda  
## 1 28.849621 6.914298 0.2396668 1.547767 99010.172 131991.375  
## 2 45.728313 13.623806 0.2979293 2.086037 316862.656 374842.344  
## 3 33.262043 13.323813 0.4005711 1.498783 29762.877 36584.992  
## 4 8.025703 3.869000 0.4820761 1.356151 6488.693 7006.970  
## 5 1.061468 0.243000 0.2289282 1.676330 7620.188 8419.265  
## 6 4.970367 2.227000 0.4480554 1.703795 5893.620 6921.736  
## cn ck ctfp cwtfp rconna rdana rnna  
## 1 426543.344 0.009373740 0.6489794 0.6261114 95642.992 131823.703 703953.44  
## 2 804602.250 0.019657943 0.7193898 0.6575428 308969.156 379569.906 1669635.62  
## 3 67072.594 0.001670069 0.3086515 0.2926067 36392.965 43032.312 161644.33  
## 4 5790.397 0.000184122 0.2058650 0.2346433 7096.825 7673.088 10142.19  
## 5 16907.236 0.000431469 0.9597649 0.9939848 5540.333 7002.935 41003.62  
## 6 13102.988 0.000326280 0.2080882 0.2166736 5945.256 7349.800 31103.90  
## rkna rtfpna rwtfpna labsh irr delta xr  
## 1 0.5127636 1.0685177 1.0088565 0.4834034 0.1191986 0.04367869 10.625636  
## 2 0.6472401 0.9738633 0.8659422 0.5610615 0.1086111 0.04274033 6.939828  
## 3 0.3999723 0.9857174 0.8166339 0.5060376 0.1883355 0.04584719 800.408508  
## 4 0.4503547 0.7033411 0.7307311 0.7410101 0.0805836 0.04811555 389.696228  
## 5 1.0681652 0.7377213 0.7526957 0.6120788 0.0339911 0.05729877 6.939828  
## 6 0.8475645 0.9595972 0.9199492 0.8062015 0.0100000 0.03019949 711.976257  
## pl\_con pl\_da pl\_gdpo csh\_c csh\_i csh\_g csh\_x  
## 1 0.3074521 0.3091229 0.2959138 0.6185091 0.25088412 0.1346494 0.12141126  
## 2 0.3485862 0.3542197 0.3513103 0.6707266 0.14713579 0.1333800 0.13506839  
## 3 0.3583208 0.3578034 0.3510293 0.6729943 0.18397693 0.1296429 0.03528051  
## 4 0.2811811 0.2933432 0.2936617 0.8692831 0.08773819 0.2291763 0.02413822  
## 5 0.2063527 0.2353755 0.2225078 0.8454817 0.10229653 0.1300420 0.23413709  
## 6 0.2249274 0.2412989 0.2329872 0.7890537 0.16095963 0.1336388 0.05874514  
## csh\_m csh\_r pl\_c pl\_i pl\_g pl\_x pl\_m  
## 1 -0.18202816 0.05657427 0.2897680 0.3141388 0.3886836 0.4656333 0.4819752  
## 2 -0.13060561 0.04429488 0.3236173 0.3850075 0.4741471 0.4940954 0.5201648  
## 3 -0.08773702 0.06584237 0.3687038 0.3555461 0.3044215 0.5012784 0.4876226  
## 4 -0.01792815 -0.19240767 0.2923447 0.4456095 0.2388369 0.4708369 0.5111370  
## 5 -0.29139951 -0.02055784 0.2000526 0.5121438 0.2473135 0.4870323 0.4826454  
## 6 -0.10874363 -0.03365362 0.2261528 0.3351474 0.2176924 0.5116569 0.4663574  
## pl\_n total excl\_energy energy metals\_minerals forestry agriculture  
## 1 0.2788645 327.86 223.05 769.06 268.89 289.58 149.37  
## 2 0.4440915 327.86 223.05 769.06 268.89 289.58 149.37  
## 3 0.3505021 327.86 223.05 769.06 268.89 289.58 149.37  
## 4 0.4028476 327.86 223.05 769.06 268.89 289.58 149.37  
## 5 0.3984682 327.86 223.05 769.06 268.89 289.58 149.37  
## 6 0.3908741 327.86 223.05 769.06 268.89 289.58 149.37  
## fish total\_change excl\_energy\_change energy\_change metals\_minerals\_change  
## 1 876.46 0.2648432 0.01211544 0.7171884 0.09993455  
## 2 876.46 0.2648432 0.01211544 0.7171884 0.09993455  
## 3 876.46 0.2648432 0.01211544 0.7171884 0.09993455  
## 4 876.46 0.2648432 0.01211544 0.7171884 0.09993455  
## 5 876.46 0.2648432 0.01211544 0.7171884 0.09993455  
## 6 876.46 0.2648432 0.01211544 0.7171884 0.09993455  
## forestry\_change agriculture\_change fish\_change growthbucket  
## 1 -0.06117685 0.0739862 0.007575845 0  
## 2 -0.06117685 0.0739862 0.007575845 0  
## 3 -0.06117685 0.0739862 0.007575845 0  
## 4 -0.06117685 0.0739862 0.007575845 0  
## 5 -0.06117685 0.0739862 0.007575845 0  
## 6 -0.06117685 0.0739862 0.007575845 1

data <- na.omit(data)   
  
head(data)

## pop emp emp\_to\_pop\_ratio hc ccon cda  
## 1 28.849621 6.914298 0.2396668 1.547767 99010.172 131991.375  
## 2 45.728313 13.623806 0.2979293 2.086037 316862.656 374842.344  
## 3 33.262043 13.323813 0.4005711 1.498783 29762.877 36584.992  
## 4 8.025703 3.869000 0.4820761 1.356151 6488.693 7006.970  
## 5 1.061468 0.243000 0.2289282 1.676330 7620.188 8419.265  
## 6 4.970367 2.227000 0.4480554 1.703795 5893.620 6921.736  
## cn ck ctfp cwtfp rconna rdana rnna  
## 1 426543.344 0.009373740 0.6489794 0.6261114 95642.992 131823.703 703953.44  
## 2 804602.250 0.019657943 0.7193898 0.6575428 308969.156 379569.906 1669635.62  
## 3 67072.594 0.001670069 0.3086515 0.2926067 36392.965 43032.312 161644.33  
## 4 5790.397 0.000184122 0.2058650 0.2346433 7096.825 7673.088 10142.19  
## 5 16907.236 0.000431469 0.9597649 0.9939848 5540.333 7002.935 41003.62  
## 6 13102.988 0.000326280 0.2080882 0.2166736 5945.256 7349.800 31103.90  
## rkna rtfpna rwtfpna labsh irr delta xr  
## 1 0.5127636 1.0685177 1.0088565 0.4834034 0.1191986 0.04367869 10.625636  
## 2 0.6472401 0.9738633 0.8659422 0.5610615 0.1086111 0.04274033 6.939828  
## 3 0.3999723 0.9857174 0.8166339 0.5060376 0.1883355 0.04584719 800.408508  
## 4 0.4503547 0.7033411 0.7307311 0.7410101 0.0805836 0.04811555 389.696228  
## 5 1.0681652 0.7377213 0.7526957 0.6120788 0.0339911 0.05729877 6.939828  
## 6 0.8475645 0.9595972 0.9199492 0.8062015 0.0100000 0.03019949 711.976257  
## pl\_con pl\_da pl\_gdpo csh\_c csh\_i csh\_g csh\_x  
## 1 0.3074521 0.3091229 0.2959138 0.6185091 0.25088412 0.1346494 0.12141126  
## 2 0.3485862 0.3542197 0.3513103 0.6707266 0.14713579 0.1333800 0.13506839  
## 3 0.3583208 0.3578034 0.3510293 0.6729943 0.18397693 0.1296429 0.03528051  
## 4 0.2811811 0.2933432 0.2936617 0.8692831 0.08773819 0.2291763 0.02413822  
## 5 0.2063527 0.2353755 0.2225078 0.8454817 0.10229653 0.1300420 0.23413709  
## 6 0.2249274 0.2412989 0.2329872 0.7890537 0.16095963 0.1336388 0.05874514  
## csh\_m csh\_r pl\_c pl\_i pl\_g pl\_x pl\_m  
## 1 -0.18202816 0.05657427 0.2897680 0.3141388 0.3886836 0.4656333 0.4819752  
## 2 -0.13060561 0.04429488 0.3236173 0.3850075 0.4741471 0.4940954 0.5201648  
## 3 -0.08773702 0.06584237 0.3687038 0.3555461 0.3044215 0.5012784 0.4876226  
## 4 -0.01792815 -0.19240767 0.2923447 0.4456095 0.2388369 0.4708369 0.5111370  
## 5 -0.29139951 -0.02055784 0.2000526 0.5121438 0.2473135 0.4870323 0.4826454  
## 6 -0.10874363 -0.03365362 0.2261528 0.3351474 0.2176924 0.5116569 0.4663574  
## pl\_n total excl\_energy energy metals\_minerals forestry agriculture  
## 1 0.2788645 327.86 223.05 769.06 268.89 289.58 149.37  
## 2 0.4440915 327.86 223.05 769.06 268.89 289.58 149.37  
## 3 0.3505021 327.86 223.05 769.06 268.89 289.58 149.37  
## 4 0.4028476 327.86 223.05 769.06 268.89 289.58 149.37  
## 5 0.3984682 327.86 223.05 769.06 268.89 289.58 149.37  
## 6 0.3908741 327.86 223.05 769.06 268.89 289.58 149.37  
## fish total\_change excl\_energy\_change energy\_change metals\_minerals\_change  
## 1 876.46 0.2648432 0.01211544 0.7171884 0.09993455  
## 2 876.46 0.2648432 0.01211544 0.7171884 0.09993455  
## 3 876.46 0.2648432 0.01211544 0.7171884 0.09993455  
## 4 876.46 0.2648432 0.01211544 0.7171884 0.09993455  
## 5 876.46 0.2648432 0.01211544 0.7171884 0.09993455  
## 6 876.46 0.2648432 0.01211544 0.7171884 0.09993455  
## forestry\_change agriculture\_change fish\_change growthbucket  
## 1 -0.06117685 0.0739862 0.007575845 0  
## 2 -0.06117685 0.0739862 0.007575845 0  
## 3 -0.06117685 0.0739862 0.007575845 0  
## 4 -0.06117685 0.0739862 0.007575845 0  
## 5 -0.06117685 0.0739862 0.007575845 0  
## 6 -0.06117685 0.0739862 0.007575845 1

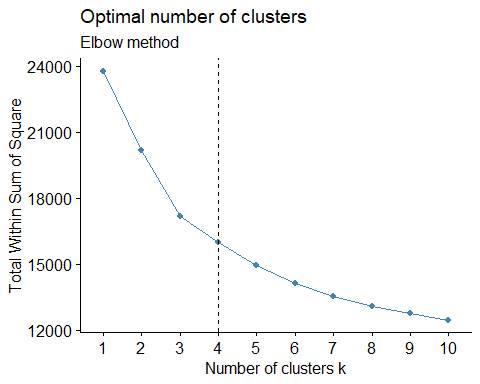
data <- data[-c(50)]  
summary(data)

## pop emp emp\_to\_pop\_ratio hc   
## Min. : 1.061 Min. : 0.243 Min. :0.1982 Min. :1.069   
## 1st Qu.: 3.831 1st Qu.: 1.049 1st Qu.:0.2979 1st Qu.:1.446   
## Median : 10.868 Median : 4.184 Median :0.3688 Median :1.690   
## Mean : 20.186 Mean : 7.121 Mean :0.3579 Mean :1.777   
## 3rd Qu.: 24.221 3rd Qu.: 8.518 3rd Qu.:0.4167 3rd Qu.:2.117   
## Max. :190.886 Max. :65.157 Max. :0.5554 Max. :2.885   
## ccon cda cn ck   
## Min. : 2781 Min. : 2984 Min. : 5790 Min. :0.0001238   
## 1st Qu.: 9117 1st Qu.: 11082 1st Qu.: 24292 1st Qu.:0.0005139   
## Median : 17472 Median : 22228 Median : 64324 Median :0.0013554   
## Mean : 64361 Mean : 80886 Mean : 244224 Mean :0.0045827   
## 3rd Qu.: 58017 3rd Qu.: 69677 3rd Qu.: 188624 3rd Qu.:0.0032271   
## Max. :758455 Max. :896605 Max. :2886312 Max. :0.0418350   
## ctfp cwtfp rconna rdana   
## Min. :0.09862 Min. :0.1078 Min. : 2787 Min. : 2978   
## 1st Qu.:0.30118 1st Qu.:0.2956 1st Qu.: 8798 1st Qu.: 10661   
## Median :0.40065 Median :0.4059 Median : 16610 Median : 21656   
## Mean :0.45442 Mean :0.4534 Mean : 68170 Mean : 86895   
## 3rd Qu.:0.61674 3rd Qu.:0.6035 3rd Qu.: 61146 3rd Qu.: 81661   
## Max. :0.99819 Max. :1.0317 Max. :762127 Max. :908561   
## rnna rkna rtfpna rwtfpna   
## Min. : 8736 Min. :0.1368 Min. :0.4295 Min. :0.3829   
## 1st Qu.: 34658 1st Qu.:0.7100 1st Qu.:0.9470 1st Qu.:0.9006   
## Median : 74131 Median :0.9334 Median :1.0000 Median :0.9788   
## Mean : 296093 Mean :0.9220 Mean :0.9889 Mean :0.9562   
## 3rd Qu.: 248804 3rd Qu.:1.0374 3rd Qu.:1.0385 3rd Qu.:1.0154   
## Max. :2592378 Max. :3.0339 Max. :1.9568 Max. :1.4473   
## labsh irr delta xr   
## Min. :0.1643 Min. :0.01000 Min. :0.02487 Min. : 1.000   
## 1st Qu.:0.4261 1st Qu.:0.06337 1st Qu.:0.04125 1st Qu.: 8.609   
## Median :0.5233 Median :0.10755 Median :0.04632 Median : 152.080   
## Mean :0.5132 Mean :0.11596 Mean :0.04809 Mean : 447.069   
## 3rd Qu.:0.6121 3rd Qu.:0.16136 3rd Qu.:0.05274 3rd Qu.: 550.521   
## Max. :0.8373 Max. :0.43007 Max. :0.09394 Max. :7384.432   
## pl\_con pl\_da pl\_gdpo csh\_c   
## Min. :0.1502 Min. :0.1468 Min. :0.1287 Min. :0.2237   
## 1st Qu.:0.3385 1st Qu.:0.3492 1st Qu.:0.3294 1st Qu.:0.5944   
## Median :0.4093 Median :0.4226 Median :0.4085 Median :0.7083   
## Mean :0.4205 Mean :0.4250 Mean :0.4116 Mean :0.6794   
## 3rd Qu.:0.4902 3rd Qu.:0.4892 3rd Qu.:0.4871 3rd Qu.:0.7927   
## Max. :0.9286 Max. :0.8663 Max. :0.8099 Max. :0.9995   
## csh\_i csh\_g csh\_x csh\_m   
## Min. :0.04995 Min. :0.01655 Min. :0.01221 Min. :-0.78446   
## 1st Qu.:0.14212 1st Qu.:0.13208 1st Qu.:0.05993 1st Qu.:-0.29967   
## Median :0.19326 Median :0.16864 Median :0.13685 Median :-0.18637   
## Mean :0.20677 Mean :0.17406 Mean :0.16854 Mean :-0.22568   
## 3rd Qu.:0.26834 3rd Qu.:0.20880 3rd Qu.:0.22186 3rd Qu.:-0.12670   
## Max. :0.55603 Max. :0.40601 Max. :0.84325 Max. :-0.01793   
## csh\_r pl\_c pl\_i pl\_g   
## Min. :-0.652756 Min. :0.1546 Min. :0.03047 Min. :0.04423   
## 1st Qu.:-0.034658 1st Qu.:0.3467 1st Qu.:0.38328 1st Qu.:0.29905   
## Median : 0.000000 Median :0.4162 Median :0.47176 Median :0.38315   
## Mean :-0.003049 Mean :0.4307 Mean :0.47453 Mean :0.40706   
## 3rd Qu.: 0.041204 3rd Qu.:0.4951 3rd Qu.:0.55890 3rd Qu.:0.49719   
## Max. : 0.634478 Max. :0.9830 Max. :1.18366 Max. :1.20525   
## pl\_x pl\_m pl\_n total   
## Min. :0.3752 Min. :0.4326 Min. :0.06274 Min. :274.7   
## 1st Qu.:0.5363 1st Qu.:0.5412 1st Qu.:0.29473 1st Qu.:355.2   
## Median :0.6406 Median :0.6080 Median :0.35765 Median :485.8   
## Mean :0.6231 Mean :0.5926 Mean :0.39449 Mean :480.4   
## 3rd Qu.:0.7158 3rd Qu.:0.6452 3rd Qu.:0.43316 3rd Qu.:611.6   
## Max. :1.1432 Max. :1.0772 Max. :6.04815 Max. :689.9   
## excl\_energy energy metals\_minerals forestry   
## Min. :205.3 Min. : 586.1 Min. :241.9 Min. :245.6   
## 1st Qu.:271.7 1st Qu.: 802.0 1st Qu.:331.4 1st Qu.:289.6   
## Median :326.9 Median :1285.2 Median :562.3 Median :328.0   
## Mean :315.4 Mean :1205.3 Mean :512.9 Mean :322.4   
## 3rd Qu.:363.5 3rd Qu.:1555.9 3rd Qu.:642.8 3rd Qu.:348.5   
## Max. :415.2 Max. :1985.5 Max. :775.1 Max. :417.6   
## agriculture fish total\_change excl\_energy\_change  
## Min. :149.4 Min. : 843.7 Min. :-0.35945 Min. :-0.17374   
## 1st Qu.:185.8 1st Qu.: 899.2 1st Qu.:-0.05893 1st Qu.:-0.04432   
## Median :213.1 Median : 948.7 Median : 0.06186 Median : 0.01995   
## Mean :221.3 Mean :1009.4 Mean : 0.04454 Mean : 0.03035   
## 3rd Qu.:265.4 3rd Qu.:1095.3 3rd Qu.: 0.20873 3rd Qu.: 0.10579   
## Max. :308.1 Max. :1276.5 Max. : 0.26484 Max. : 0.21835   
## energy\_change metals\_minerals\_change forestry\_change   
## Min. :-0.47702 Min. :-0.15756 Min. :-0.174550   
## 1st Qu.:-0.10294 1st Qu.:-0.05340 1st Qu.:-0.060792   
## Median : 0.05493 Median : 0.04681 Median : 0.003277   
## Mean : 0.08241 Mean : 0.05814 Mean : 0.024774   
## 3rd Qu.: 0.24509 3rd Qu.: 0.15453 3rd Qu.: 0.096810   
## Max. : 0.71719 Max. : 0.47620 Max. : 0.300217   
## agriculture\_change fish\_change   
## Min. :-0.22600 Min. :-0.113858   
## 1st Qu.:-0.02645 1st Qu.:-0.020995   
## Median : 0.03844 Median : 0.006529   
## Mean : 0.03393 Mean : 0.023741   
## 3rd Qu.: 0.11279 3rd Qu.: 0.047421   
## Max. : 0.32088 Max. : 0.165440

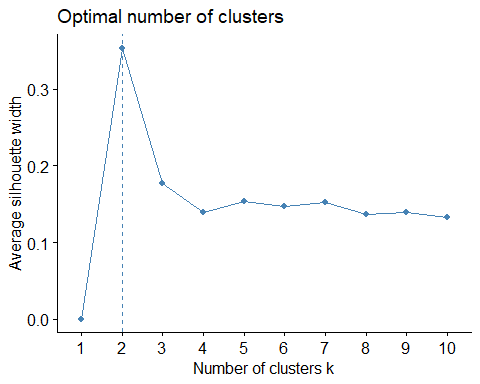
data <- scale(data)  
head(data)

## pop emp emp\_to\_pop\_ratio hc ccon cda  
## 1 0.2884351 -0.02084284 -1.4675585 -0.5144558 0.2672828 0.3260511  
## 2 0.8503560 0.65541858 -0.7441681 0.6915106 1.9477913 1.8754343  
## 3 0.4353322 0.62518188 0.5302381 -0.6242036 -0.2668891 -0.2826393  
## 4 -0.4048291 -0.32778296 1.5422089 -0.9437631 -0.4464256 -0.4713463  
## 5 -0.6366806 -0.69325294 -1.6008896 -0.2264168 -0.4376972 -0.4623359  
## 6 -0.5065466 -0.49328260 1.1198064 -0.1648826 -0.4510159 -0.4718901  
## cn ck ctfp cwtfp rconna rdana rnna  
## 1 0.3858469 0.5835657 0.9418987 0.8504057 0.2014870 0.2724751 0.7956304  
## 2 1.1859440 1.8362240 1.2827661 1.0051971 1.7660093 1.7749480 2.6794284  
## 3 -0.3749115 -0.3547732 -0.7056806 -0.7920190 -0.2330494 -0.2660060 -0.2622754  
## 4 -0.5046048 -0.5357676 -1.2032859 -1.0774741 -0.4479057 -0.4804443 -0.5578171  
## 5 -0.4810779 -0.5056398 2.4464587 2.6620872 -0.4593209 -0.4845085 -0.4976144  
## 6 -0.4891290 -0.5184522 -1.1925232 -1.1659704 -0.4563512 -0.4824049 -0.5169262  
## rkna rtfpna rwtfpna labsh irr delta  
## 1 -1.2027117 0.55500478 0.3739147 -0.19980506 0.04391577 -0.362421102  
## 2 -0.8075151 -0.10468819 -0.6406730 0.32166117 -0.09976110 -0.439567523  
## 3 -1.5341804 -0.02207147 -0.9907259 -0.04781897 0.98212615 -0.184137818  
## 4 -1.3861178 -1.99009041 -1.6005728 1.52999812 -0.48010293 0.002354861  
## 5 0.4294904 -1.75047830 -1.4446405 0.66423762 -1.11237866 0.757351527  
## 6 -0.2188062 -0.20411549 -0.2572627 1.96775188 -1.43794581 -1.470610489  
## xr pl\_con pl\_da pl\_gdpo csh\_c csh\_i  
## 1 -0.54743019 -0.8955495 -0.9898360 -0.9861975 -0.38290552 0.4957157  
## 2 -0.55205329 -0.5697209 -0.6046291 -0.5137980 -0.05429228 -0.6700819  
## 3 0.44319212 -0.4926117 -0.5740180 -0.5161943 -0.04002115 -0.2561060  
## 4 -0.07196307 -1.1036451 -1.1246224 -1.0054020 1.19525683 -1.3375200  
## 5 -0.55205329 -1.6963709 -1.6197701 -1.6121734 1.04547040 -1.1739311  
## 6 0.33227181 -1.5492377 -1.5691736 -1.5228100 0.69035996 -0.5147465  
## csh\_g csh\_x csh\_m csh\_r pl\_c pl\_i pl\_g  
## 1 -0.6247654 -0.3310157 0.3293127 0.6615230 -1.0358276 -1.1243272 -0.1146167  
## 2 -0.6448890 -0.2350992 0.7172742 0.5252829 -0.7870746 -0.6275583 0.4185023  
## 3 -0.7041317 -0.9359276 1.0406996 0.7643530 -0.4557411 -0.8340745 -0.6402416  
## 4 0.8737213 -1.0141820 1.5673781 -2.1009391 -1.0168918 -0.2027567 -1.0493564  
## 5 -0.6978043 0.4606784 -0.4958481 -0.1942600 -1.6951318 0.2636299 -0.9964800  
## 6 -0.6407868 -0.7711313 0.8822136 -0.3395582 -1.5033254 -0.9770628 -1.1812553  
## pl\_x pl\_m pl\_n total excl\_energy energy  
## 1 -1.507530 -1.4482439 -0.298867382 -1.153031 -1.41363 -1.047143  
## 2 -1.234971 -0.9481045 0.128214640 -1.153031 -1.41363 -1.047143  
## 3 -1.166185 -1.3742847 -0.113697013 -1.153031 -1.41363 -1.047143  
## 4 -1.457700 -1.0663344 0.021606559 -1.153031 -1.41363 -1.047143  
## 5 -1.302608 -1.4394673 0.010286698 -1.153031 -1.41363 -1.047143  
## 6 -1.066798 -1.6527789 -0.009342629 -1.153031 -1.41363 -1.047143  
## metals\_minerals forestry agriculture fish total\_change  
## 1 -1.4102 -0.7408666 -1.444797 -0.949043 1.197166  
## 2 -1.4102 -0.7408666 -1.444797 -0.949043 1.197166  
## 3 -1.4102 -0.7408666 -1.444797 -0.949043 1.197166  
## 4 -1.4102 -0.7408666 -1.444797 -0.949043 1.197166  
## 5 -1.4102 -0.7408666 -1.444797 -0.949043 1.197166  
## 6 -1.4102 -0.7408666 -1.444797 -0.949043 1.197166  
## excl\_energy\_change energy\_change metals\_minerals\_change forestry\_change  
## 1 -0.1726087 2.200617 0.2762354 -0.661549  
## 2 -0.1726087 2.200617 0.2762354 -0.661549  
## 3 -0.1726087 2.200617 0.2762354 -0.661549  
## 4 -0.1726087 2.200617 0.2762354 -0.661549  
## 5 -0.1726087 2.200617 0.2762354 -0.661549  
## 6 -0.1726087 2.200617 0.2762354 -0.661549  
## agriculture\_change fish\_change  
## 1 0.3151544 -0.2101098  
## 2 0.3151544 -0.2101098  
## 3 0.3151544 -0.2101098  
## 4 0.3151544 -0.2101098  
## 5 0.3151544 -0.2101098  
## 6 0.3151544 -0.2101098

fviz\_nbclust(data, kmeans, method = "wss") +  
 geom\_vline(xintercept = 4, linetype = 2)+  
 labs(subtitle = "Elbow method")

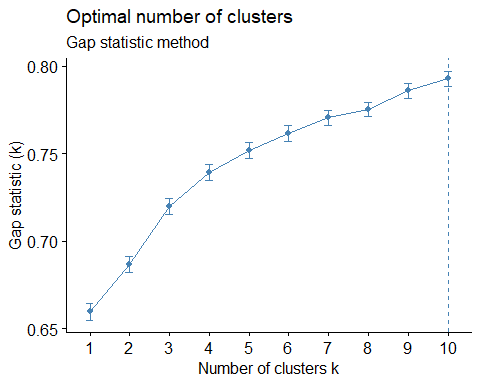


fviz\_nbclust(data, kmeans, method = c("silhouette", "wss", "gap\_stat"))



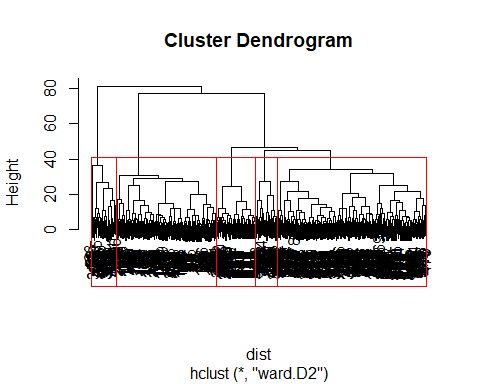
set.seed(123)  
fviz\_nbclust(data, kmeans, nstart = 25, method = "gap\_stat", nboot = 50)+  
 labs(subtitle = "Gap statistic method")

## Warning: did not converge in 10 iterations



dist <- dist(data, method = "euclidean")  
fit <- hclust(dist, method = "ward.D2")

plot(fit)  
  
groups <- cutree(fit, k=5)  
  
rect.hclust(fit, k=5, border="red")



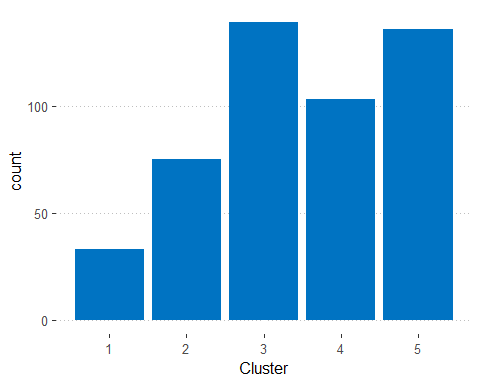
fit <- kmeans(data, 5)  
# get cluster means   
aggregate(data,by=list(fit$cluster),FUN=mean)

## Group.1 pop emp emp\_to\_pop\_ratio hc ccon  
## 1 1 2.9095633 2.7350658 -0.56317788 0.6513111 3.2056437  
## 2 2 -0.1618822 -0.1241797 0.14818936 0.1228541 -0.1822919  
## 3 3 -0.2009855 -0.1747037 0.16936403 -0.2345400 -0.2775791  
## 4 4 -0.2624650 -0.2866078 -0.13071035 -0.2694761 -0.2626859  
## 5 5 -0.2125261 -0.1995534 -0.01917475 0.2180131 -0.1946633  
## cda cn ck ctfp cwtfp rconna  
## 1 3.1597746 2.65537123 2.6734135 0.95548208 0.67070506 3.3835969  
## 2 -0.1698565 -0.04710001 -0.1054404 -0.23825698 -0.13224148 -0.2016426  
## 3 -0.2873901 -0.29812732 -0.3303117 -0.23513013 -0.23989990 -0.2867882  
## 4 -0.2787576 -0.33570121 -0.2788097 0.10423145 0.05111786 -0.2745623  
## 5 -0.1681916 -0.05939579 -0.0417933 0.06092362 0.11666023 -0.2087649  
## rdana rnna rkna rtfpna rwtfpna labsh  
## 1 3.3279654 2.9946338 -0.1224659 -0.02898263 0.3104638 -0.34527613  
## 2 -0.1874951 -0.1482333 0.7419732 0.02391373 0.2663280 -0.01459552  
## 3 -0.2967119 -0.3168327 -0.1246406 0.02752377 -0.0907645 0.22230851  
## 4 -0.2878421 -0.2572268 -0.8046015 -0.14877988 -0.5405389 0.19315028  
## 5 -0.1828680 -0.1262597 0.3572969 0.07839280 0.2799402 -0.28166607  
## irr delta xr pl\_con pl\_da pl\_gdpo  
## 1 0.42028914 -0.02289978 -0.45003453 0.7487129 0.8354317 0.8591347  
## 2 -0.07892975 0.10565185 0.14221088 0.2303209 0.2178605 0.2059791  
## 3 -0.05021307 -0.21029103 0.22837326 -0.3291102 -0.2967383 -0.3187549  
## 4 -0.06982779 -0.10349044 -0.05781963 -1.0716965 -1.1312601 -1.1073580  
## 5 0.04575050 0.24060127 -0.15884660 0.8393343 0.8371884 0.8423892  
## csh\_c csh\_i csh\_g csh\_x csh\_m csh\_r  
## 1 -0.06232285 -0.6124905 -0.81614167 0.22497997 0.6807419 -0.07066219  
## 2 -0.03791236 0.2678445 0.28023940 -0.07947007 -0.1681862 -0.02088357  
## 3 0.17947839 -0.2695719 0.09521548 -0.23730321 0.1035948 0.10564398  
## 4 0.14956203 -0.4364738 -0.18347633 -0.19645365 0.4038299 0.01218039  
## 5 -0.26067871 0.6069938 0.08513110 0.38055727 -0.4841520 -0.08853656  
## pl\_c pl\_i pl\_g pl\_x pl\_m pl\_n  
## 1 0.5908600 0.85865602 0.98406129 -0.30092866 0.537559573 1.26516598  
## 2 0.2707672 0.07262018 0.02152577 0.33423717 0.225314418 -0.20473110  
## 3 -0.2748068 -0.07552211 -0.31987783 0.09248932 -0.005684125 -0.05950906  
## 4 -1.0856866 -0.79582289 -0.63900539 -1.28655650 -1.410135867 -0.12020529  
## 5 0.8104260 0.43150859 0.56023617 0.76854531 0.819088532 -0.04222604  
## total excl\_energy energy metals\_minerals forestry agriculture  
## 1 0.1256309 0.14059175 0.1186537 0.13736931 0.1085166 0.1178560  
## 2 -0.5768431 -0.01997784 -0.6731970 0.18360466 -0.2800016 -0.0322598  
## 3 0.2550888 0.11550883 0.2832933 0.08735164 0.4602294 -0.1581331  
## 4 -1.3195191 -1.51429416 -1.2233905 -1.48338310 -1.1674114 -1.2572818  
## 5 1.0262539 1.00570133 0.9794533 0.89958122 0.5418423 1.1030203  
## fish total\_change excl\_energy\_change energy\_change  
## 1 0.09007193 0.006103576 0.04598605 -0.02306954  
## 2 0.57648930 -1.651639655 -1.40558558 -1.43710866  
## 3 0.07461980 0.647538148 0.87372963 0.41511708  
## 4 -0.94128648 0.163675024 -0.25385970 0.43267174  
## 5 0.29684767 0.123567839 0.06323905 0.04616162  
## metals\_minerals\_change forestry\_change agriculture\_change fish\_change  
## 1 0.03368636 0.06838553 -0.01085187 0.0331445  
## 2 -0.96322168 -0.83728246 -1.62690117 -0.9518081  
## 3 0.91956402 0.54201166 0.48692199 0.1832033  
## 4 -0.19173014 -0.39721777 0.07909261 -0.2013925  
## 5 -0.27162661 0.19200935 0.34225739 0.4821324

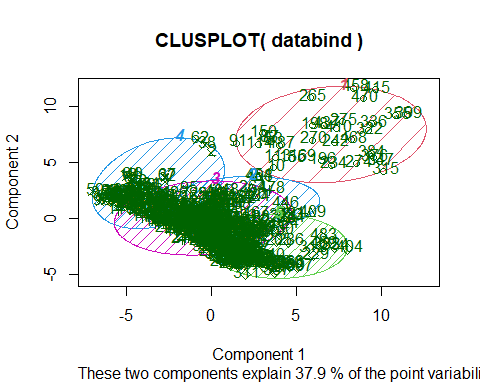
# append cluster assignment  
data <- data.frame(data, fit$cluster)  
  
databind <- cbind(data, Cluster = fit$cluster)  
head(databind)

## pop emp emp\_to\_pop\_ratio hc ccon cda  
## 1 0.2884351 -0.02084284 -1.4675585 -0.5144558 0.2672828 0.3260511  
## 2 0.8503560 0.65541858 -0.7441681 0.6915106 1.9477913 1.8754343  
## 3 0.4353322 0.62518188 0.5302381 -0.6242036 -0.2668891 -0.2826393  
## 4 -0.4048291 -0.32778296 1.5422089 -0.9437631 -0.4464256 -0.4713463  
## 5 -0.6366806 -0.69325294 -1.6008896 -0.2264168 -0.4376972 -0.4623359  
## 6 -0.5065466 -0.49328260 1.1198064 -0.1648826 -0.4510159 -0.4718901  
## cn ck ctfp cwtfp rconna rdana rnna  
## 1 0.3858469 0.5835657 0.9418987 0.8504057 0.2014870 0.2724751 0.7956304  
## 2 1.1859440 1.8362240 1.2827661 1.0051971 1.7660093 1.7749480 2.6794284  
## 3 -0.3749115 -0.3547732 -0.7056806 -0.7920190 -0.2330494 -0.2660060 -0.2622754  
## 4 -0.5046048 -0.5357676 -1.2032859 -1.0774741 -0.4479057 -0.4804443 -0.5578171  
## 5 -0.4810779 -0.5056398 2.4464587 2.6620872 -0.4593209 -0.4845085 -0.4976144  
## 6 -0.4891290 -0.5184522 -1.1925232 -1.1659704 -0.4563512 -0.4824049 -0.5169262  
## rkna rtfpna rwtfpna labsh irr delta  
## 1 -1.2027117 0.55500478 0.3739147 -0.19980506 0.04391577 -0.362421102  
## 2 -0.8075151 -0.10468819 -0.6406730 0.32166117 -0.09976110 -0.439567523  
## 3 -1.5341804 -0.02207147 -0.9907259 -0.04781897 0.98212615 -0.184137818  
## 4 -1.3861178 -1.99009041 -1.6005728 1.52999812 -0.48010293 0.002354861  
## 5 0.4294904 -1.75047830 -1.4446405 0.66423762 -1.11237866 0.757351527  
## 6 -0.2188062 -0.20411549 -0.2572627 1.96775188 -1.43794581 -1.470610489  
## xr pl\_con pl\_da pl\_gdpo csh\_c csh\_i  
## 1 -0.54743019 -0.8955495 -0.9898360 -0.9861975 -0.38290552 0.4957157  
## 2 -0.55205329 -0.5697209 -0.6046291 -0.5137980 -0.05429228 -0.6700819  
## 3 0.44319212 -0.4926117 -0.5740180 -0.5161943 -0.04002115 -0.2561060  
## 4 -0.07196307 -1.1036451 -1.1246224 -1.0054020 1.19525683 -1.3375200  
## 5 -0.55205329 -1.6963709 -1.6197701 -1.6121734 1.04547040 -1.1739311  
## 6 0.33227181 -1.5492377 -1.5691736 -1.5228100 0.69035996 -0.5147465  
## csh\_g csh\_x csh\_m csh\_r pl\_c pl\_i pl\_g  
## 1 -0.6247654 -0.3310157 0.3293127 0.6615230 -1.0358276 -1.1243272 -0.1146167  
## 2 -0.6448890 -0.2350992 0.7172742 0.5252829 -0.7870746 -0.6275583 0.4185023  
## 3 -0.7041317 -0.9359276 1.0406996 0.7643530 -0.4557411 -0.8340745 -0.6402416  
## 4 0.8737213 -1.0141820 1.5673781 -2.1009391 -1.0168918 -0.2027567 -1.0493564  
## 5 -0.6978043 0.4606784 -0.4958481 -0.1942600 -1.6951318 0.2636299 -0.9964800  
## 6 -0.6407868 -0.7711313 0.8822136 -0.3395582 -1.5033254 -0.9770628 -1.1812553  
## pl\_x pl\_m pl\_n total excl\_energy energy  
## 1 -1.507530 -1.4482439 -0.298867382 -1.153031 -1.41363 -1.047143  
## 2 -1.234971 -0.9481045 0.128214640 -1.153031 -1.41363 -1.047143  
## 3 -1.166185 -1.3742847 -0.113697013 -1.153031 -1.41363 -1.047143  
## 4 -1.457700 -1.0663344 0.021606559 -1.153031 -1.41363 -1.047143  
## 5 -1.302608 -1.4394673 0.010286698 -1.153031 -1.41363 -1.047143  
## 6 -1.066798 -1.6527789 -0.009342629 -1.153031 -1.41363 -1.047143  
## metals\_minerals forestry agriculture fish total\_change  
## 1 -1.4102 -0.7408666 -1.444797 -0.949043 1.197166  
## 2 -1.4102 -0.7408666 -1.444797 -0.949043 1.197166  
## 3 -1.4102 -0.7408666 -1.444797 -0.949043 1.197166  
## 4 -1.4102 -0.7408666 -1.444797 -0.949043 1.197166  
## 5 -1.4102 -0.7408666 -1.444797 -0.949043 1.197166  
## 6 -1.4102 -0.7408666 -1.444797 -0.949043 1.197166  
## excl\_energy\_change energy\_change metals\_minerals\_change forestry\_change  
## 1 -0.1726087 2.200617 0.2762354 -0.661549  
## 2 -0.1726087 2.200617 0.2762354 -0.661549  
## 3 -0.1726087 2.200617 0.2762354 -0.661549  
## 4 -0.1726087 2.200617 0.2762354 -0.661549  
## 5 -0.1726087 2.200617 0.2762354 -0.661549  
## 6 -0.1726087 2.200617 0.2762354 -0.661549  
## agriculture\_change fish\_change fit.cluster Cluster  
## 1 0.3151544 -0.2101098 4 4  
## 2 0.3151544 -0.2101098 4 4  
## 3 0.3151544 -0.2101098 4 4  
## 4 0.3151544 -0.2101098 4 4  
## 5 0.3151544 -0.2101098 4 4  
## 6 0.3151544 -0.2101098 4 4

library(ggplot2)  
library(ggpubr)  
theme\_set(theme\_pubr())  
ggplot(databind, aes(Cluster)) +  
 geom\_bar(fill = "#0073C2FF") +  
 theme\_pubclean()



library(cluster)  
  
clusplot(databind, fit$cluster, color=TRUE, shade=TRUE,   
 labels=2, lines=0)



library(fpc)  
  
plotcluster(databind, fit$cluster)

