R version 3.5.3 (2019-03-11) -- "Great Truth"

Copyright (C) 2019 The R Foundation for Statistical Computing

Platform: x86\_64-w64-mingw32/x64 (64-bit)

R is free software and comes with ABSOLUTELY NO WARRANTY.

You are welcome to redistribute it under certain conditions.

Type 'license()' or 'licence()' for distribution details.

R is a collaborative project with many contributors.

Type 'contributors()' for more information and

'citation()' on how to cite R or R packages in publications.

Type 'demo()' for some demos, 'help()' for on-line help, or

'help.start()' for an HTML browser interface to help.

Type 'q()' to quit R.

> data <- read.csv(file.choose())

>

> View(data)

> plot(data)

> data1 <- scale(data)

Error in colMeans(x, na.rm = TRUE) : 'x' must be numeric

> data.f = data

> data.f$sex<-NULL

> data.f$smoker<-NULL

> data.f$region<-NULL

> data.f$Bought<-NULL

> View(data.f)

> plot(data.f)

> data.s<-scale(data.f[,-5])

> data.s

age bmi children charges

[1,] -1.186959570 -0.496931620 -0.88013950 0.16460383

[2,] -1.252589246 0.533468313 -0.03617012 -0.92786935

[3,] -0.596292485 0.398305119 1.65176865 -0.73156845

[4,] -0.268144105 -1.408844337 -0.88013950 0.53210706

[5,] -0.333773781 -0.324905737 -0.88013950 -0.77355451

[6,] -0.399403457 -0.876090709 -0.88013950 -0.78149859

[7,] 0.585041684 0.475541229 -0.03617012 -0.45835757

[8,] -0.005625401 -0.525017478 1.65176865 -0.52747479

[9,] -0.005625401 -0.158145952 0.80779927 -0.59053926

[10,] 1.503857149 -0.858537048 -0.88013950 1.03214802

[11,] -0.793181514 -0.791833134 -0.88013950 -0.85610844

[12,] 1.635116501 -0.779545571 -0.88013950 0.95183698

[13,] -0.924440866 0.644056380 -0.88013950 -0.92056975

[14,] 1.241338444 1.595464836 -0.88013950 -0.25296063

[15,] -0.661922161 2.000954417 -0.88013950 1.80243269

[16,] -1.186959570 -1.076202451 -0.03617012 -0.91982070

[17,] 0.978819740 0.008613832 -0.03617012 -0.27410343

[18,] -0.924440866 -1.208732595 -0.88013950 -0.87961266

[19,] 1.241338444 1.679722411 -0.88013950 -0.28815275

[20,] -0.465033133 0.802039334 -0.88013950 1.60250104

[21,] 1.503857149 0.925792648 -0.88013950 -0.09887451

[22,] -0.465033133 0.292983149 -0.03617012 -0.75316847

[23,] -1.252589246 0.591395396 -0.88013950 -0.97028309

[24,] -0.202514429 0.208725574 -0.03617012 1.66479548

[25,] -0.005625401 -0.474989543 0.80779927 -0.60513324

[26,] 1.438227472 -0.528528211 1.65176865 -0.04321899

[27,] 1.700746177 -1.342140423 -0.88013950 -0.01073881

[28,] 1.175708768 0.358809380 0.80779927 -0.16807321

[29,] -0.924440866 -2.342699131 -0.03617012 -0.85222615

[30,] -0.399403457 0.977575949 0.80779927 1.73751881

[31,] -0.990070542 0.854700319 -0.88013950 1.51228244

[32,] -1.252589246 -0.775157155 -0.88013950 -0.89380832

[33,] -1.186959570 -0.374055989 3.33970743 -0.71439264

[34,] 1.700746177 -0.424961608 -0.88013950 -0.05986879

[35,] -0.596292485 0.995129611 -0.03617012 2.63715726

[36,] -1.186959570 -1.809067820 -0.88013950 -0.93508448

[37,] 1.635116501 0.392161337 1.65176865 0.07288340

[38,] -0.727551837 -1.743241589 -0.88013950 -0.88630553

[39,] -0.136884753 1.042524497 -0.03617012 1.81414474

[40,] 1.503857149 1.609507765 -0.88013950 2.41943201

[41,] -0.858811190 -0.725129220 -0.88013950 -0.83270568

[42,] -0.399403457 1.035503033 0.80779927 -0.69551414

[43,] 0.256893303 -1.571215706 -0.03617012 -0.60019129

[44,] -0.005625401 0.012124565 0.80779927 -0.59721628

[45,] 0.060004275 1.109228411 -0.03617012 -0.61408600

[46,] 1.175708768 1.153112565 -0.88013950 0.43451640

[47,] -1.252589246 1.392720045 0.80779927 -0.80767762

[48,] -0.596292485 0.709004928 -0.88013950 -0.79589010

[49,] 1.503857149 -1.088490014 -0.88013950 -0.14203838

[50,] -0.071255077 0.784485673 -0.03617012 1.73738736

[51,] -1.252589246 0.859088734 -0.88013950 -0.89287572

[52,] -1.055700218 0.508893186 0.80779927 -0.79423933

[53,] 0.716301036 -0.479377958 -0.03617012 0.64624507

[54,] -0.071255077 0.649322479 -0.88013950 1.66772848

[55,] 0.191263627 -0.358257694 1.65176865 -0.47139505

[56,] 1.372597796 1.092552433 0.80779927 2.37065304

[57,] 1.372597796 0.192049596 0.80779927 -0.07159601

[58,] -1.252589246 0.166596786 0.80779927 1.41986454

[59,] 1.044449416 -1.378125429 -0.03617012 0.62293307

[60,] -0.202514429 1.159256346 0.80779927 -0.62058258

[61,] 0.388152656 -0.591721392 1.65176865 -0.43200829

[62,] -0.793181514 0.514159285 2.49573804 -0.72759039

[63,] 1.766375853 -1.058648789 -0.03617012 1.12176057

[64,] -0.596292485 -0.841861069 -0.03617012 -0.75432832

[65,] -1.121329894 -1.458872272 -0.88013950 0.00799173

[66,] -1.186959570 -0.321395004 -0.88013950 -0.92659655

[67,] 1.569486825 1.469078473 0.80779927 -0.02636003

[68,] 0.191263627 -0.775157155 -0.03617012 -0.59176674

[69,] 0.191263627 0.958266922 -0.88013950 -0.62558536

[70,] -0.596292485 -1.185035152 1.65176865 0.22068694

attr(,"scaled:center")

age bmi children charges

37.085714 30.730929 1.042857 14600.848988

attr(,"scaled:scale")

age bmi children charges

15.237010 5.696817 1.184877 13876.195663

> km<-kmeans(data.s,3)

> km

K-means clustering with 3 clusters of sizes 25, 23, 22

Cluster means:

age bmi children charges

1 0.68742398 0.3971466 -0.4750342 0.9609260

2 -0.79318151 -0.6485800 -0.6232793 -0.6973723

3 0.04807161 0.2267580 1.1914217 -0.3628904

Clustering vector:

[1] 2 2 3 2 2 2 3 3 3 1 2 1 2 1 1 2 1 2 1 1 1 2 2 1 3 3 1

[28] 3 2 1 1 2 3 1 1 2 3 2 1 1 2 3 2 3 3 1 3 2 1 1 2 3 1 1

[55] 3 1 3 3 1 3 3 3 1 2 2 2 3 2 1 3

Within cluster sum of squares by cluster:

[1] 76.03728 29.33882 46.56739

(between\_SS / total\_SS = 44.9 %)

Available components:

[1] "cluster" "centers" "totss"

[4] "withinss" "tot.withinss" "betweenss"

[7] "size" "iter" "ifault"

> str(km)

List of 9

$ cluster : int [1:70] 2 2 3 2 2 2 3 3 3 1 ...

$ centers : num [1:3, 1:4] 0.6874 -0.7932 0.0481 0.3971 -0.6486 ...

..- attr(\*, "dimnames")=List of 2

.. ..$ : chr [1:3] "1" "2" "3"

.. ..$ : chr [1:4] "age" "bmi" "children" "charges"

$ totss : num 276

$ withinss : num [1:3] 76 29.3 46.6

$ tot.withinss: num 152

$ betweenss : num 124

$ size : int [1:3] 25 23 22

$ iter : int 3

$ ifault : int 0

- attr(\*, "class")= chr "kmeans"

> plot(data.f, col = km$cluster)

> plot(data.f$age~data.f$charges,data = data, col = km$cluster)

> plot(data.f$age~data.f$bmi,data = data, col = km$cluster)

> plot(data.f$age~data.f$children,data = data, col = km$cluster)