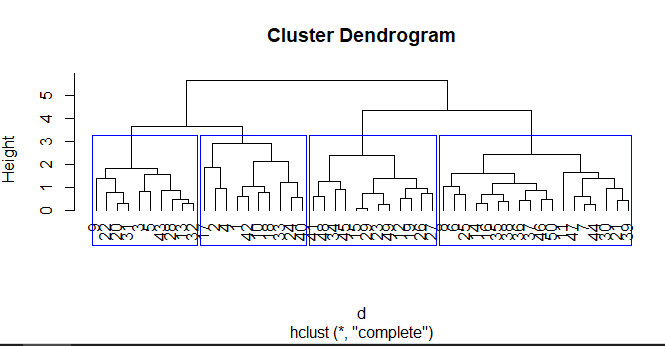
Clustering

1)crime data

|  |
| --- |
| > View(crime\_data)  > mydata<-crime\_data  > mydata  # A tibble: 50 x 5  X1 Murder Assault UrbanPop Rape  *<chr>* *<dbl>* *<dbl>* *<dbl>* *<dbl>*  1 Alabama 13.2 236 58 21.2  2 Alaska 10 263 48 44.5  3 Arizona 8.1 294 80 31  4 Arkansas 8.8 190 50 19.5  5 California 9 276 91 40.6  6 Colorado 7.9 204 78 38.7  7 Connecticut 3.3 110 77 11.1  8 Delaware 5.9 238 72 15.8  9 Florida 15.4 335 80 31.9  10 Georgia 17.4 211 60 25.8  # … with 40 more rows  > normaliza<-scale(mydata[,2:4])  > d<-dist(normaliza,method="euclidean")  > #fit herachical distance  > fit<-hclust(d,method = "complete")  > str(fit)  List of 7  $ merge : int [1:49, 1:2] -15 -23 -7 -13 -20 -14 -35 -21 -36 -28 ...  $ height : num [1:49] 0.0738 0.2291 0.2406 0.269 0.2755 ...  $ order : int [1:50] 9 22 20 31 3 5 43 28 13 32 ...  $ labels : NULL  $ method : chr "complete"  $ call : language hclust(d = d, method = "complete")  $ dist.method: chr "euclidean"  - attr(\*, "class")= chr "hclust"  > fit$order  [1] 9 22 20 31 3 5 43 28 13 32 17 2 4 1 42 10 18 33 24 40 41 48 34 45 15 29 23 49  [29] 12 19 26 27 8 6 25 14 16 35 38 36 37 46 50 11 47 7 44 30 21 39  > fit$labels  NULL  > fit$height  [1] 0.07378321 0.22914130 0.24058516 0.26904429 0.27546206 0.28505482 0.35198035  [8] 0.40223816 0.40783570 0.43724025 0.44597901 0.50323899 0.53507155 0.58009860  [15] 0.59799378 0.60474602 0.68306598 0.69632363 0.72708851 0.73891646 0.74778047  [22] 0.76849905 0.81661679 0.84334033 0.87225487 0.88988966 0.92859848 0.94599383  [29] 0.95822643 1.01928824 1.02177026 1.15286889 1.17786026 1.25362565 1.36657077  [36] 1.36715260 1.38532004 1.53647152 1.60397052 1.62040497 1.79940868 1.86972164  [43] 2.09292374 2.36928661 2.42988795 2.87972882 3.63897074 4.32508691 5.66050667  > plot(fit)  > plot(fit,hang = -1)  E:\data science r studio\Assignment code 1\custing\crime\comp1.PNG  > fit$order  [1] 9 22 20 31 3 5 43 28 13 32 17 2 4 1 42 10 18 33 24 40 41 48 34 45 15 29 23 49  [29] 12 19 26 27 8 6 25 14 16 35 38 36 37 46 50 11 47 7 44 30 21 39 |
|  |
| |  |  |  |  |  | | --- | --- | --- | --- | --- | | >   |  | | --- | | rect.hclust(fit,k=4,border = "blue") | |  | | |  | | --- | | > | | | |



> group<-cutree(fit,k=4)

> group

[1] 1 1 2 1 2 3 3 3 2 1 3 4 2 3 4 3 1 1 4 2 3 2 4 1 3 4 4 2 4 3 2 2 1 4 3 3 3 3

[39] 3 1 4 1 2 3 4 3 3 4 4 3

> membership<-as.matrix(group)

> View(membership)

> final<-data.frame(mydata,membership)

> final

X1 Murder Assault UrbanPop Rape membership

1 Alabama 13.2 236 58 21.2 1

2 Alaska 10.0 263 48 44.5 1

3 Arizona 8.1 294 80 31.0 2

4 Arkansas 8.8 190 50 19.5 1

5 California 9.0 276 91 40.6 2

6 Colorado 7.9 204 78 38.7 3

7 Connecticut 3.3 110 77 11.1 3

8 Delaware 5.9 238 72 15.8 3

9 Florida 15.4 335 80 31.9 2

10 Georgia 17.4 211 60 25.8 1

11 Hawaii 5.3 46 83 20.2 3

12 Idaho 2.6 120 54 14.2 4

13 Illinois 10.4 249 83 24.0 2

14 Indiana 7.2 113 65 21.0 3

15 Iowa 2.2 56 57 11.3 4

16 Kansas 6.0 115 66 18.0 3

17 Kentucky 9.7 109 52 16.3 1

18 Louisiana 15.4 249 66 22.2 1

19 Maine 2.1 83 51 7.8 4

20 Maryland 11.3 300 67 27.8 2

21 Massachusetts 4.4 149 85 16.3 3

22 Michigan 12.1 255 74 35.1 2

23 Minnesota 2.7 72 66 14.9 4

24 Mississippi 16.1 259 44 17.1 1

25 Missouri 9.0 178 70 28.2 3

26 Montana 6.0 109 53 16.4 4

27 Nebraska 4.3 102 62 16.5 4

28 Nevada 12.2 252 81 46.0 2

29 New Hampshire 2.1 57 56 9.5 4

30 New Jersey 7.4 159 89 18.8 3

31 New Mexico 11.4 285 70 32.1 2

32 New York 11.1 254 86 26.1 2

33 North Carolina 13.0 337 45 16.1 1

34 North Dakota 0.8 45 44 7.3 4

35 Ohio 7.3 120 75 21.4 3

36 Oklahoma 6.6 151 68 20.0 3

37 Oregon 4.9 159 67 29.3 3

38 Pennsylvania 6.3 106 72 14.9 3

39 Rhode Island 3.4 174 87 8.3 3

40 South Carolina 14.4 279 48 22.5 1

41 South Dakota 3.8 86 45 12.8 4

42 Tennessee 13.2 188 59 26.9 1

43 Texas 12.7 201 80 25.5 2

44 Utah 3.2 120 80 22.9 3

45 Vermont 2.2 48 32 11.2 4

46 Virginia 8.5 156 63 20.7 3

47 Washington 4.0 145 73 26.2 3

48 West Virginia 5.7 81 39 9.3 4

49 Wisconsin 2.6 53 66 10.8 4

50 Wyoming 6.8 161 60 15.6 3

> View(final)

> final123<-final[c(ncol(final),1:(ncol(final)-1))]

> View(final123)

|  |
| --- |
| > write.csv(final123, file="final.csv",row.names = F)  > getwd()  [1] "E:/data science r studio/Assignment code 1/Logistic testing"  > aggregate(crime\_data[-1],by=list(final123$membership),mean)  Group.1 Murder Assault UrbanPop Rape  1 1 13.120000 232.1000 53.00000 23.21000  2 2 11.370000 270.1000 79.20000 32.01000  3 3 5.966667 144.6667 73.88889 20.41111  4 4 3.091667 76.0000 52.08333 11.83333  > View(final1)  Error in View : object 'final1' not found  > View(final123) |
|  |
| |  |  |  |  |  |  |  |  |  | | --- | --- | --- | --- | --- | --- | --- | --- | --- | | >view(final123)   |  | | --- | | membership X1 Murder Assault UrbanPop Rape  1 1 Alabama 13.2 236 58 21.2  2 1 Alaska 10.0 263 48 44.5  3 2 Arizona 8.1 294 80 31.0  4 1 Arkansas 8.8 190 50 19.5  5 2 California 9.0 276 91 40.6  6 3 Colorado 7.9 204 78 38.7  7 3 Connecticut 3.3 110 77 11.1  8 3 Delaware 5.9 238 72 15.8  9 2 Florida 15.4 335 80 31.9  10 1 Georgia 17.4 211 60 25.8  11 3 Hawaii 5.3 46 83 20.2  12 4 Idaho 2.6 120 54 14.2  13 2 Illinois 10.4 249 83 24.0  14 3 Indiana 7.2 113 65 21.0  15 4 Iowa 2.2 56 57 11.3  16 3 Kansas 6.0 115 66 18.0  17 1 Kentucky 9.7 109 52 16.3  18 1 Louisiana 15.4 249 66 22.2  19 4 Maine 2.1 83 51 7.8  20 2 Maryland 11.3 300 67 27.8  21 3 Massachusetts 4.4 149 85 16.3  22 2 Michigan 12.1 255 74 35.1  23 4 Minnesota 2.7 72 66 14.9  24 1 Mississippi 16.1 259 44 17.1  25 3 Missouri 9.0 178 70 28.2  26 4 Montana 6.0 109 53 16.4  27 4 Nebraska 4.3 102 62 16.5  28 2 Nevada 12.2 252 81 46.0  29 4 New Hampshire 2.1 57 56 9.5  30 3 New Jersey 7.4 159 89 18.8  31 2 New Mexico 11.4 285 70 32.1  32 2 New York 11.1 254 86 26.1  33 1 North Carolina 13.0 337 45 16.1  34 4 North Dakota 0.8 45 44 7.3  35 3 Ohio 7.3 120 75 21.4  36 3 Oklahoma 6.6 151 68 20.0  37 3 Oregon 4.9 159 67 29.3  38 3 Pennsylvania 6.3 106 72 14.9  39 3 Rhode Island 3.4 174 87 8.3  40 1 South Carolina 14.4 279 48 22.5  41 4 South Dakota 3.8 86 45 12.8  42 1 Tennessee 13.2 188 59 26.9  43 2 Texas 12.7 201 80 25.5  44 3 Utah 3.2 120 80 22.9  45 4 Vermont 2.2 48 32 11.2  46 3 Virginia 8.5 156 63 20.7  47 3 Washington 4.0 145 73 26.2  48 4 West Virginia 5.7 81 39 9.3  49 4 Wisconsin 2.6 53 66 10.8  50 3 Wyoming 6.8 161 60 15.6 | |  | | |  |  |  |  |  | | --- | --- | --- | --- | --- | | 2)airline data  ################clustring Airline###########  library(readxl)  eastwest <- read\_excel("airine/eastwest.xlsx")  View(eastwest)  eastwest1<-eastwest[,2:11]  View(eastwest1)  #data normalization beacuse all value is one format  normalization<-scale(eastwest1)  View(normalization)  new<-normalization  View(new)  install(plyr)  library("plyr")   |  | | --- | | * fit<-kmeans(new,3) * str(fit) * List of 9 * $ cluster : int [1:3999] 3 3 3 3 1 3 1 3 1 1 ... * $ centers : num [1:3, 1:10] 0.7325 -0.3548 -0.1211 0.1572 -0.0364 ... * ..- attr(\*, "dimnames")=List of 2 * .. ..$ : chr [1:3] "1" "2" "3" * .. ..$ : chr [1:10] "Balance" "Qual\_miles" "cc1\_miles" "cc2\_miles" ... * $ totss : num 39980 * $ withinss : num [1:3] 18467 5385 6149 * $ tot.withinss: num 30001 * $ betweenss : num 9979 * $ size : int [1:3] 1016 1639 1344 * $ iter : int 3 * $ ifault : int 0   + attr(\*, "class")= chr "kmeans" * fit$cluster * [1] 3 3 3 3 1 3 1 3 1 1 3 1 3 3 3 1 1 3 3 3 1 3 3 3 3 3 3 3 3 3 3 3 1 3 3 1 3 3 3 3 3 1 3 1 3 3 * [47] 3 3 1 3 3 3 1 3 3 3 1 3 3 1 3 3 3 3 3 1 3 1 3 3 3 1 1 3 3 3 3 1 1 1 3 3 3 3 3 1 3 1 3 3 3 3 * [93] 1 3 1 3 3 1 3 3 1 3 3 3 3 1 1 3 1 1 3 3 1 3 3 3 3 1 3 3 3 1 1 3 1 1 1 1 3 1 3 3 3 3 3 1 1 1 * [139] 3 3 3 3 3 1 3 3 3 1 3 3 1 1 1 3 1 1 3 3 1 3 3 1 3 1 3 3 3 1 1 1 1 3 1 3 1 1 3 3 3 3 3 3 3 3 * [185] 3 3 1 3 1 3 1 1 1 3 3 3 1 3 3 1 3 3 3 1 3 3 1 3 3 3 1 3 3 3 3 1 3 3 1 1 1 1 3 3 3 3 3 1 3 1 * [231] 3 1 3 3 3 3 3 3 1 3 1 3 1 1 1 1 3 3 1 3 3 3 3 3 3 1 1 3 3 1 3 3 3 3 3 3 3 3 1 1 3 1 1 3 3 1 * [277] 3 1 3 1 3 1 3 3 1 1 1 3 1 1 1 1 1 3 3 1 1 3 1 3 3 3 1 1 3 1 3 1 3 3 1 1 1 1 3 3 3 3 3 3 3 1 * [323] 1 1 3 3 3 1 1 3 1 3 3 3 1 1 1 3 3 3 1 3 3 3 3 3 3 1 1 3 3 3 3 3 1 3 3 1 3 3 1 3 1 3 3 3 3 3 * [369] 3 1 3 1 1 3 3 1 3 1 3 1 1 3 1 1 1 3 3 1 3 3 3 3 3 3 3 3 3 3 3 3 3 3 3 3 1 1 3 3 3 3 1 1 1 1 * [415] 3 3 3 3 3 3 1 3 3 1 1 1 3 1 3 1 3 3 3 1 3 3 1 1 1 3 3 1 3 3 1 3 3 3 1 3 1 3 1 1 3 3 3 3 3 3 * [461] 3 3 3 3 3 1 1 1 3 3 3 1 1 3 3 1 3 1 1 1 3 1 3 3 3 3 1 3 1 3 3 3 1 3 1 3 3 3 3 1 1 1 3 1 3 1 * [507] 1 3 3 3 1 1 1 3 1 1 3 1 1 1 3 3 1 3 3 3 3 1 1 3 3 1 1 3 1 3 3 1 3 3 3 3 3 3 3 3 1 3 3 1 3 3 * [553] 3 3 3 3 3 3 3 3 3 1 3 1 1 1 3 1 3 3 3 3 3 3 1 3 1 3 3 3 1 3 3 3 3 3 3 1 3 3 3 3 3 3 1 3 3 3 * [599] 1 1 1 3 1 1 3 3 3 3 1 3 3 1 3 3 3 1 3 3 1 1 3 3 1 1 1 1 3 1 3 1 1 3 3 3 3 3 3 3 3 1 3 3 3 3 * [645] 3 3 3 3 3 3 3 3 3 1 3 3 3 3 3 3 1 3 1 3 1 3 3 3 3 1 1 3 1 3 1 3 3 3 3 3 3 1 1 1 3 1 3 1 3 1 * [691] 3 1 1 1 3 3 3 1 3 3 3 1 1 3 1 1 3 3 1 3 3 1 3 3 1 3 1 3 1 3 3 3 3 1 3 1 1 3 3 3 1 1 3 3 3 3 * [737] 1 3 3 1 3 1 3 1 3 1 3 3 3 3 1 3 3 3 1 3 1 1 3 3 3 3 3 1 1 1 3 3 3 1 3 3 1 3 1 3 3 3 3 3 1 1 * [783] 3 3 3 3 3 1 3 3 3 1 1 1 3 3 1 3 1 3 1 3 3 3 3 3 1 1 3 3 3 3 1 1 1 3 3 3 1 3 1 3 3 1 3 3 3 3 * [829] 1 3 1 1 1 1 3 1 1 1 3 1 1 1 3 3 1 3 3 1 3 1 1 1 3 3 1 1 3 3 1 3 1 3 3 3 1 1 3 3 3 3 3 3 3 3 * [875] 3 3 3 1 3 3 1 1 3 1 1 3 3 1 3 3 3 1 3 1 3 3 3 3 3 3 1 1 3 3 1 1 3 3 3 1 3 1 3 3 1 3 3 1 3 1 * [921] 3 3 1 3 3 3 3 3 3 3 1 3 1 3 1 1 3 3 3 3 3 3 3 3 3 3 3 1 3 3 3 3 3 3 1 1 3 3 1 3 3 3 1 1 3 3 * [967] 1 3 3 1 3 3 1 3 3 1 1 1 3 1 1 3 3 3 1 3 3 3 3 3 3 3 1 1 3 3 3 1 1 1 * [ reached getOption("max.print") -- omitted 2999 entries ] * final2<-data.frame(eastwest1,fit$cluster)#append * final2 * Balance Qual\_miles cc1\_miles cc2\_miles cc3\_miles Bonus\_miles Bonus\_trans Flight\_miles\_12mo * 1 28143 0 1 1 1 174 1 0 * 2 19244 0 1 1 1 215 2 0 * 3 41354 0 1 1 1 4123 4 0 * 4 14776 0 1 1 1 500 1 0 * 5 97752 0 4 1 1 43300 26 2077 * 6 16420 0 1 1 1 0 0 0 * 7 84914 0 3 1 1 27482 25 0 * 8 20856 0 1 1 1 5250 4 250 * 9 443003 0 3 2 1 1753 43 3850 * 10 104860 0 3 1 1 28426 28 1150 * 11 40091 0 2 1 1 7278 10 0 * 12 96522 0 5 1 1 61105 19 0 * 13 43382 0 2 1 1 11150 20 0 * 14 43097 0 1 1 1 3258 6 0 * 15 17648 0 1 1 1 0 0 0 * 16 28495 0 4 1 1 49442 15 0 * 17 51890 0 4 1 1 48963 16 0 * 18 13958 0 1 1 1 4291 5 0 * 19 91473 0 3 1 1 27408 17 0 * 20 23354 0 3 1 1 10447 5 0 * 21 120576 0 5 1 1 58831 23 250 * 22 185681 2024 1 1 1 13300 16 1800 * 23 20584 0 1 1 1 3450 11 3450 * 24 66275 0 1 1 1 2533 11 150 * 25 205651 500 1 1 1 4025 21 700 * 26 20726 0 1 1 1 1375 4 0 * 27 18521 0 1 1 1 1227 2 1227 * 28 8828 0 1 1 1 0 0 0 * 29 59763 0 3 1 1 33772 20 100 * 30 19221 0 1 1 1 4655 8 500 * 31 177926 0 3 1 1 20797 13 0 * 32 10021 0 1 1 1 0 0 0 * 33 276571 0 4 1 1 42044 23 0 * 34 18047 0 1 1 1 100 1 0 * 35 43832 0 1 1 1 0 0 0 * 36 123759 0 4 1 1 50572 23 2500 * 37 23193 0 1 1 1 650 4 550 * 38 68666 0 1 1 1 350 2 350 * 39 59990 0 3 1 1 27878 17 0 * 40 2176 0 1 1 1 0 0 0 * 41 34616 0 1 1 1 1750 4 500 * 42 10470 0 4 1 1 38094 26 0 * 43 60313 0 1 1 1 10000 26 3250 * 44 619393 0 3 1 1 15008 14 0 * 45 121260 0 3 1 1 18493 18 0 * 46 288865 967 1 1 1 23600 14 2000 * 47 92336 0 2 1 1 11214 6 0 * 48 36924 0 1 1 1 5900 6 300 * 49 70312 0 3 1 1 34678 24 500 * 50 17051 0 1 1 1 1150 4 1150 * 51 108137 0 1 1 1 6368 5 6368 * 52 1300 0 1 1 1 370 1 0 * 53 118531 0 4 1 1 44577 38 0 * 54 121395 0 1 1 1 4970 8 650 * 55 38348 0 1 1 1 0 0 0 * 56 14448 0 1 1 1 1625 6 0 * 57 75971 0 4 1 1 34339 14 0 * 58 36298 0 1 1 1 3100 5 600 * 59 38077 0 3 1 1 34024 8 0 * 60 95118 0 3 1 1 23188 23 2200 * 61 134457 0 3 1 1 15588 14 0 * 62 19918 0 1 1 1 17601 11 0 * 63 10120 0 1 1 1 0 0 0 * 64 362642 0 1 1 1 28079 8 0 * 65 80250 0 1 1 1 895 2 0 * 66 53914 0 3 1 1 33767 45 5550 * 67 9375 0 1 1 1 1750 7 0 * 68 83237 0 4 1 1 35287 18 0 * 69 230715 0 3 1 1 24047 12 0 * 70 47457 0 3 1 1 12621 16 0 * 71 84409 5031 2 1 1 15436 16 1150 * 72 252386 0 4 1 1 39787 13 0 * 73 550367 0 3 1 1 12500 13 50 * 74 12646 0 1 1 1 631 4 631 * 75 123867 0 3 1 1 25308 17 0 * 76 129871 0 3 1 1 15776 22 0 * 77 109380 0 1 1 1 7537 16 0 * 78 111157 0 4 1 1 32883 19 0 * 79 49238 0 4 1 1 38037 18 0 * 80 38896 0 5 1 1 76988 16 556 * 81 4340 0 1 1 1 32685 5 0 * 82 88443 0 1 1 1 1200 3 1200 * 83 8454 0 1 1 1 498 9 0 * 84 61990 0 1 1 1 1625 6 0 * 85 24093 0 1 1 1 1750 7 0 * 86 35418 0 5 1 1 58557 18 900 * 87 40284 0 1 1 1 7719 5 0 * 88 609477 0 3 1 1 21422 22 1200 * 89 62553 0 2 1 1 7672 16 700 * 90 48433 0 1 1 1 3500 4 500 * Flight\_trans\_12 Days\_since\_enroll fit.cluster * 1 0 7000 3 * 2 0 6968 3 * 3 0 7034 3 * 4 0 6952 3 * 5 4 6935 1 * 6 0 6942 3 * 7 0 6994 1 * 8 1 6938 3 * 9 12 6948 1 * 10 3 6931 1 * 11 0 6959 3 * 12 0 6924 1 * 13 0 6924 3 * 14 0 6918 3 * 15 0 6912 3 * 16 0 6912 1 * 17 0 6910 1 * 18 0 6905 3 * 19 0 6903 3 * 20 0 6896 3 * 21 2 6896 1 * 22 9 6896 3 * 23 11 6884 3 * 24 1 6884 3 * 25 4 7932 3 * 26 0 7924 3 * 27 2 7917 3 * 28 0 7914 3 * 29 1 7907 3 * 30 1 7896 3 * 31 0 7886 3 * 32 0 7879 3 * 33 0 7872 1 * 34 0 7868 3 * 35 0 7865 3 * 36 5 7865 1 * 37 3 7861 3 * 38 2 7861 3 * 39 0 7854 3 * 40 0 7847 3 * 41 1 7850 3 * 42 0 7840 1 * 43 9 7829 3 * 44 0 7819 1 * 45 0 7808 3 * 46 4 6884 3 * 47 0 6884 3 * 48 2 6879 3 * 49 1 6875 1 * 50 4 6868 3 * 51 5 6844 3 * 52 0 6868 3 * 53 0 6868 1 * 54 2 6889 3 * 55 0 6861 3 * 56 0 6856 3 * 57 0 6869 1 * 58 3 6865 3 * 59 0 6837 3 * 60 7 6865 1 * 61 0 6861 3 * 62 0 6863 3 * 63 0 6864 3 * 64 0 6835 3 * 65 0 6833 3 * 66 29 6826 1 * 67 0 6826 3 * 68 0 6837 1 * 69 0 6826 3 * 70 0 7766 3 * 71 4 7766 3 * 72 0 7787 1 * 73 1 7801 1 * 74 4 7787 3 * 75 0 7759 3 * 76 0 7752 3 * 77 0 7749 3 * 78 0 7771 1 * 79 0 7801 1 * 80 1 7771 1 * 81 0 7733 3 * 82 3 6837 3 * 83 0 6821 3 * 84 0 6833 3 * 85 0 6819 3 * 86 3 6813 1 * 87 0 6827 3 * 88 8 6820 1 * 89 1 6819 3 * 90 1 6807 3 * [ reached 'max' / getOption("max.print") -- omitted 3909 rows ] * final3<-final2[,c(ncol(final2),1:(ncol(final2)-1))] * aggregate(eastwest[,2:11],by= list(fit$cluster),FUN=mean) * Group.1 Balance Qual\_miles cc1\_miles cc2\_miles cc3\_miles Bonus\_miles Bonus\_trans * 1 1 147421.97 265.76673 3.872047 1.007874 1.047244 47963.116 21.536417 * 2 2 37850.51 115.92556 1.339231 1.009762 1.000000 5715.507 7.345943 * 3 3 61394.34 86.52753 1.567708 1.025298 1.000744 7785.713 9.281994 * Flight\_miles\_12mo Flight\_trans\_12 Days\_since\_enroll * 1 1135.8278 3.4045276 4862.013 * 2 229.0415 0.6650397 2280.281 * 3 230.9249 0.7023810 5798.314 | |  | | |  | | --- | | >  install.packages("kselection")  library(kselection)  install.packages("doParallel")  library(doParallel)  > registerDoParallel(cores=3)  > k<-kselection(new,parallel=F,k\_threshold=0.95,max\_centers=14)  > k  f(k) finds 9 clusters  twss<-NULL  for(i in 2:3999){  twss[i]=sum(kmeans(new,centers = i)$tot.withinss)  }  windows()  plot(2:3999,twss,type = "b",xlab = "Number of clusters",ylab = "within groups")  title(sub="k-mean clustering scree-plot")  t<-xy.coords(new)  plot(t)  E:\data science r studio\Assignment code 1\custing\airine\1 plot.PNG  install.packages("Animation")  library(animation)  > t<-xy.coords(new)  > plot(t)  > km<-kmeans.ani(new,4)  E:\data science r studio\Assignment code 1\custing\airine\fina ami.PNG | | | | | |