PCA

Assignment 1

BP: Perform Principal component analysis and perform clustering using first

3 principal component scores (both heirarchial and k mean clustering(scree plot or elbow curve) and obtain

optimum number of clusters and check whether we have obtained same number of clusters with the original data

(class column we have ignored at the begining who shows it has 3 clusters)df

PROCEDURE:

STEP 1: First we have to Exploratory Data Analysis which can be done by plotting scattered plot, box plots and summary.

summary(crime\_data\_1\_)

Type Alcohol Malic Ash Alcalinity

Min. :1.000 Min. :11.03 Min. :0.740 Min. :1.360 Min. :10.60

1st Qu.:1.000 1st Qu.:12.36 1st Qu.:1.603 1st Qu.:2.210 1st Qu.:17.20

Median :2.000 Median :13.05 Median :1.865 Median :2.360 Median :19.50

Mean :1.938 Mean :13.00 Mean :2.336 Mean :2.367 Mean :19.49

3rd Qu.:3.000 3rd Qu.:13.68 3rd Qu.:3.083 3rd Qu.:2.558 3rd Qu.:21.50

Max. :3.000 Max. :14.83 Max. :5.800 Max. :3.230 Max. :30.00

Magnesium Phenols Flavanoids Nonflavanoids Proanthocyanins

Min. : 70.00 Min. :0.980 Min. :0.340 Min. :0.1300 Min. :0.410

1st Qu.: 88.00 1st Qu.:1.742 1st Qu.:1.205 1st Qu.:0.2700 1st Qu.:1.250

Median : 98.00 Median :2.355 Median :2.135 Median :0.3400 Median :1.555

Mean : 99.74 Mean :2.295 Mean :2.029 Mean :0.3619 Mean :1.591

3rd Qu.:107.00 3rd Qu.:2.800 3rd Qu.:2.875 3rd Qu.:0.4375 3rd Qu.:1.950

Max. :162.00 Max. :3.880 Max. :5.080 Max. :0.6600 Max. :3.580

Color Hue Dilution Proline

Min. : 1.280 Min. :0.4800 Min. :1.270 Min. : 278.0

1st Qu.: 3.220 1st Qu.:0.7825 1st Qu.:1.938 1st Qu.: 500.5

Median : 4.690 Median :0.9650 Median :2.780 Median : 673.5

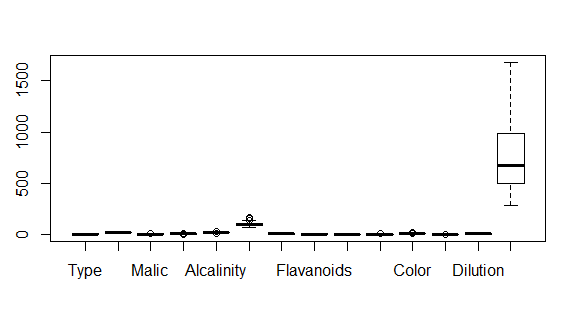
Mean : 5.058 Mean :0.9574 Mean :2.612 Mean : 746.9

3rd Qu.: 6.200 3rd Qu.:1.1200 3rd Qu.:3.170 3rd Qu.: 985.0

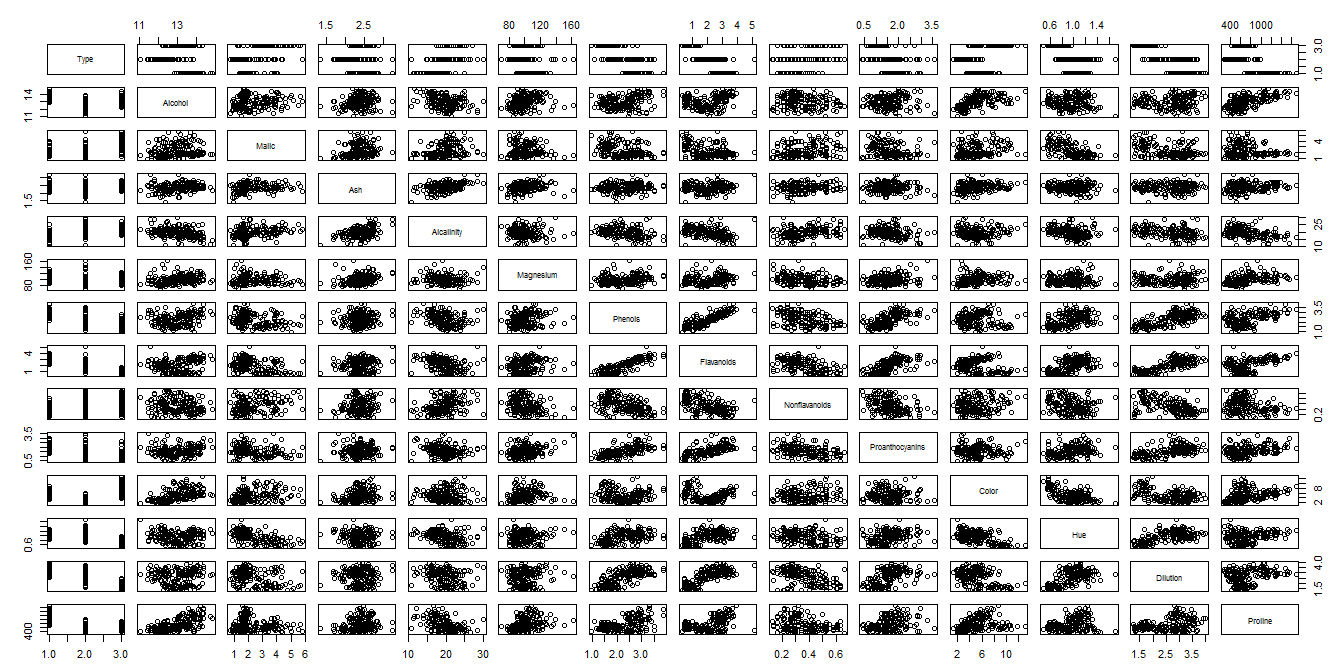
Max. :13.000 Max. :1.7100 Max. :4.000 Max. :1680.0

From the summary we can see that in Proline the difference between mean and max is large so it may be right skewed it can my confirmed by using Box Plot .

It can be confirmed that Assault is right Skewed.



Scatter Plot for the Following data frame is:



STEP 2 : Now forming the clusters using PCA:

On finding the correlation of the data we get :

Alcohol Malic Ash Alcalinity Magnesium

Alcohol 1.00000000 0.09439694 0.211544596 -0.31023514 0.27079823

Malic 0.09439694 1.00000000 0.164045470 0.28850040 -0.05457510

Ash 0.21154460 0.16404547 1.000000000 0.44336719 0.28658669

Alcalinity -0.31023514 0.28850040 0.443367187 1.00000000 -0.08333309

Magnesium 0.27079823 -0.05457510 0.286586691 -0.08333309 1.00000000

Phenols 0.28910112 -0.33516700 0.128979538 -0.32111332 0.21440123

Flavanoids 0.23681493 -0.41100659 0.115077279 -0.35136986 0.19578377

Nonflavanoids -0.15592947 0.29297713 0.186230446 0.36192172 -0.25629405

Proanthocyanins 0.13669791 -0.22074619 0.009651935 -0.19732684 0.23644061

Color 0.54636420 0.24898534 0.258887259 0.01873198 0.19995001

Hue -0.07174720 -0.56129569 -0.074666889 -0.27395522 0.05539820

Dilution 0.07234319 -0.36871043 0.003911231 -0.27676855 0.06600394

Proline 0.64372004 -0.19201056 0.223626264 -0.44059693 0.39335085

Phenols Flavanoids Nonflavanoids Proanthocyanins Color

Alcohol 0.28910112 0.2368149 -0.1559295 0.136697912 0.54636420

Malic -0.33516700 -0.4110066 0.2929771 -0.220746187 0.24898534

Ash 0.12897954 0.1150773 0.1862304 0.009651935 0.25888726

Alcalinity -0.32111332 -0.3513699 0.3619217 -0.197326836 0.01873198

Magnesium 0.21440123 0.1957838 -0.2562940 0.236440610 0.19995001

Phenols 1.00000000 0.8645635 -0.4499353 0.612413084 -0.05513642

Flavanoids 0.86456350 1.0000000 -0.5378996 0.652691769 -0.17237940

Nonflavanoids -0.44993530 -0.5378996 1.0000000 -0.365845099 0.13905701

Proanthocyanins 0.61241308 0.6526918 -0.3658451 1.000000000 -0.02524993

Color -0.05513642 -0.1723794 0.1390570 -0.025249931 1.00000000

Hue 0.43368134 0.5434786 -0.2626396 0.295544253 -0.52181319

Dilution 0.69994936 0.7871939 -0.5032696 0.519067096 -0.42881494

Proline 0.49811488 0.4941931 -0.3113852 0.330416700 0.31610011

Hue Dilution Proline

Alcohol -0.07174720 0.072343187 0.6437200

Malic -0.56129569 -0.368710428 -0.1920106

Ash -0.07466689 0.003911231 0.2236263

Alcalinity -0.27395522 -0.276768549 -0.4405969

Magnesium 0.05539820 0.066003936 0.3933508

Phenols 0.43368134 0.699949365 0.4981149

Flavanoids 0.54347857 0.787193902 0.4941931

Nonflavanoids -0.26263963 -0.503269596 -0.3113852

Proanthocyanins 0.29554425 0.519067096 0.3304167

Color -0.52181319 -0.428814942 0.3161001

Hue 1.00000000 0.565468293 0.2361834

Dilution 0.56546829 1.000000000 0.3127611

Proline 0.23618345 0.312761075 1.0000000

To know the importance of the components :

pcobj <- princomp(wine[-1], cor=TRUE , scores = TRUE , covmat = NULL)

> summary(pcobj)

Importance of components:

Comp.1 Comp.2 Comp.3 Comp.4 Comp.5

Standard deviation 2.1692972 1.5801816 1.2025273 0.9586313 0.92370351

Proportion of Variance 0.3619885 0.1920749 0.1112363 0.0706903 0.06563294

Cumulative Proportion 0.3619885 0.5540634 0.6652997 0.7359900 0.80162293

Comp.6 Comp.7 Comp.8 Comp.9 Comp.10

Standard deviation 0.80103498 0.74231281 0.59033665 0.53747553 0.50090167

Proportion of Variance 0.04935823 0.04238679 0.02680749 0.02222153 0.01930019

Cumulative Proportion 0.85098116 0.89336795 0.92017544 0.94239698 0.96169717

Comp.11 Comp.12 Comp.13

Standard deviation 0.47517222 0.41081655 0.321524394

Proportion of Variance 0.01736836 0.01298233 0.007952149

Cumulative Proportion 0.97906553 0.99204785 1.000000000

> str(pcobj)

List of 7

$ sdev : Named num [1:13] 2.169 1.58 1.203 0.959 0.924 ...

..- attr(\*, "names")= chr [1:13] "Comp.1" "Comp.2" "Comp.3" "Comp.4" ...

$ loadings: loadings [1:13, 1:13] -0.14433 0.24519 0.00205 0.23932 -0.14199 ...

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

.. ..$ : chr [1:13] "Alcohol" "Malic" "Ash" "Alcalinity" ...

.. ..$ : chr [1:13] "Comp.1" "Comp.2" "Comp.3" "Comp.4" ...

$ center : Named num [1:13] 13 2.34 2.37 19.49 99.74 ...

..- attr(\*, "names")= chr [1:13] "Alcohol" "Malic" "Ash" "Alcalinity" ...

$ scale : Named num [1:13] 0.81 1.114 0.274 3.33 14.242 ...

..- attr(\*, "names")= chr [1:13] "Alcohol" "Malic" "Ash" "Alcalinity" ...

$ n.obs : int 178

$ scores : num [1:178, 1:13] -3.32 -2.21 -2.52 -3.76 -1.01 ...

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

.. ..$ : NULL

.. ..$ : chr [1:13] "Comp.1" "Comp.2" "Comp.3" "Comp.4" ...

$ call : language princomp(x = wine[-1], cor = TRUE, scores = TRUE, covmat = NULL)

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

>loadings(pcobj)

Loadings:

Comp.1 Comp.2 Comp.3 Comp.4 Comp.5 Comp.6 Comp.7 Comp.8 Comp.9

Alcohol -0.144 -0.484 -0.207 -0.266 0.214 0.396 -0.509

Malic 0.245 -0.225 0.537 0.537 -0.421

Ash -0.316 0.626 -0.214 -0.143 0.154 0.149 -0.170 0.308

Alcalinity 0.239 0.612 -0.101 0.287 0.428 -0.200

Magnesium -0.142 -0.300 0.131 -0.352 0.727 -0.323 -0.156 -0.271

Phenols -0.395 0.146 0.198 -0.149 -0.406 -0.286

Flavanoids -0.423 0.151 0.152 -0.109 -0.187

Nonflavanoids 0.299 0.170 -0.203 -0.501 -0.259 -0.595 -0.233 -0.196

Proanthocyanins -0.313 0.149 0.399 0.137 -0.534 -0.372 0.368 0.209

Color -0.530 -0.137 -0.419 0.228

Hue -0.297 0.279 -0.428 -0.174 0.106 -0.232 0.437

Dilution -0.376 0.164 0.166 0.184 -0.101 0.266 -0.137

Proline -0.287 -0.365 -0.127 -0.232 -0.158 0.120 0.120 0.576

Comp.10 Comp.11 Comp.12 Comp.13

Alcohol 0.212 -0.226 -0.266

Malic -0.309 0.122

Ash -0.499 -0.141

Alcalinity 0.479

Magnesium

Phenols -0.320 0.304 -0.304 -0.464

Flavanoids -0.163 0.832

Nonflavanoids 0.216 0.117 0.114

Proanthocyanins 0.134 -0.237 -0.117

Color -0.291 0.604

Hue -0.522 0.259

Dilution 0.524 0.601 -0.157

Proline 0.162 0.539

Comp.1 Comp.2 Comp.3 Comp.4 Comp.5 Comp.6 Comp.7 Comp.8 Comp.9

SS loadings 1.000 1.000 1.000 1.000 1.000 1.000 1.000 1.000 1.000

Proportion Var 0.077 0.077 0.077 0.077 0.077 0.077 0.077 0.077 0.077

Cumulative Var 0.077 0.154 0.231 0.308 0.385 0.462 0.538 0.615 0.692

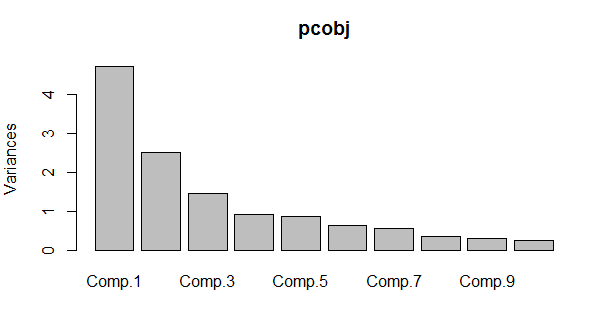
Comp.10 Comp.11 Comp.12 Comp.13

SS loadings 1.000 1.000 1.000 1.000

Proportion Var 0.077 0.077 0.077 0.077

Cumulative Var 0.769 0.846 0.923 1.000

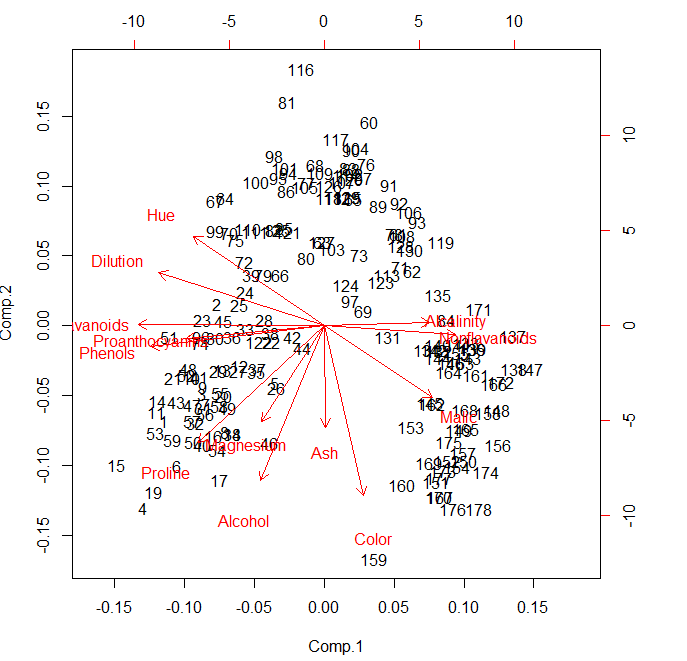
> plot(pcobj)



From the plot we can see that most of the data is occupied in the first 3 columns.

On doing :

biplot(pcobj)



Now we can view the data of 3 major columns by :

|  |
| --- |
| pcobj$scores[,1:3]  Comp.1 Comp.2 Comp.3  [1,] -3.31675081 -1.44346263 -0.165739045  [2,] -2.20946492 0.33339289 -2.026457374  [3,] -2.51674015 -1.03115130 0.982818670  [4,] -3.75706561 -2.75637191 -0.176191842  [5,] -1.00890849 -0.86983082 2.026688219  [6,] -3.05025392 -2.12240111 -0.629395827  [7,] -2.44908967 -1.17485013 -0.977094891  [8,] -2.05943687 -1.60896307 0.146281883  [9,] -2.51087430 -0.91807096 -1.770969027  [10,] -2.75362819 -0.78943767 -0.984247490  [11,] -3.47973668 -1.30233324 -0.422735217  [12,] -1.75475290 -0.61197723 -1.190878320  [13,] -2.11346234 -0.67570634 -0.865086426  [14,] -3.45815682 -1.13062988 -1.204276353  [15,] -4.31278391 -2.09597558 -1.263912752  [16,] -2.30518820 -1.66255173 0.217902616  [17,] -2.17195527 -2.32730534 0.831729866  [18,] -1.89897118 -1.63136888 0.794913792  [19,] -3.54198508 -2.51834367 -0.485458508  [20,] -2.08452220 -1.06113799 -0.164746678  [21,] -3.12440254 -0.78689711 -0.364887083  [22,] -1.08657007 -0.24174355 0.936961600  [23,] -2.53522408 0.09184062 -0.311932659  [24,] -1.64498834 0.51627893 0.143885095  [25,] -1.76157587 0.31714893 0.890285647  [26,] -0.99007910 -0.94066734 3.820908008  [27,] -1.77527763 -0.68617513 -0.086700406  [28,] -1.23542396 0.08980704 -1.386896545  [29,] -2.18840633 -0.68956962 1.394566881  [30,] -2.25610898 -0.19146194 -1.092657258  [31,] -2.50022003 -1.24083383 1.386017855  [32,] -2.67741105 -1.47187365 -0.332261728  [33,] -1.62857912 -0.05270445 -0.167128706  [34,] -1.90269086 -1.63306043 1.172082119  [35,] -1.41038853 -0.69793432 0.479743025  [36,] -1.90382623 -0.17671095 0.450835040  [37,] -1.38486223 -0.65863985 0.458438581  [38,] -1.12220741 -0.11410976 -0.039107277  [39,] -1.50219450 0.76943201 -1.426177346  [40,] -2.52980109 -1.80300198 -0.343152389  [41,] -2.58809543 -0.77961630 -0.118477466  [42,] -0.66848199 -0.16996094 -0.783362548  [43,] -3.07080699 -1.15591896 -0.312758084  [44,] -0.46220914 -0.33074213 -0.201476496  [45,] -2.10135193 0.07100892 -0.655849415  [46,] -1.13616618 -1.77710739 0.028705736  [47,] -2.72660096 -1.19133469 -0.539773261  [48,] -2.82133927 -0.64625860 -1.155552411  [49,] -2.00985085 -1.24702946 -0.057293988  [50,] -2.70749130 -1.75196741 -0.643113612  [51,] -3.21491747 -0.16699199 -1.973571680  [52,] -2.85895983 -0.74527880 0.004719502  [53,] -3.50560436 -1.61273386 -0.520774530  [54,] -2.22479138 -1.87516800 0.339549850  [55,] -2.14698782 -1.01675154 -0.957762762  [56,] -2.46932948 -1.32900831 0.513437453  [57,] -2.74151791 -1.43654878 -0.612473396  [58,] -2.17374092 -1.21219984 0.261779593  [59,] -3.13938015 -1.73157912 -0.285661413  [60,] 0.92858197 3.07348616 -4.585064007  [61,] 1.54248014 1.38144351 -0.874683112  [62,] 1.83624976 0.82998412 -1.605702186  [63,] -0.03060683 1.26278614 -1.784408010  [64,] -2.05026161 1.92503260 -0.007368777  [65,] 0.60968083 1.90805881 0.679357938  [66,] -0.90022784 0.76391147 0.573361302  [67,] -2.24850719 1.88459248 -2.031840193  [68,] -0.18338403 2.42714611 -1.069745560  [69,] 0.81280503 0.22051399 -0.707005396  [70,] -1.97562050 1.40328323 -1.238276220  [71,] 1.57221622 0.88498314 -0.628997950  [72,] -1.65768181 0.95671220 1.952584217  [73,] 0.72537239 1.06364540 0.080332229  [74,] -2.56222717 -0.26019855 3.374393962  [75,] -1.83256757 1.28787820 0.458280027  [76,] 0.86799290 2.44410119 -1.563333179  [77,] -0.37001440 2.15390698 -2.449386348  [78,] 1.45737704 1.38335177 -0.227306902  [79,] -1.26293085 0.77084953 -1.184224517  [80,] -0.37615037 1.02704340 1.794466295  [81,] -0.76206390 3.37505381 -0.357470056  [82,] -1.03457797 1.45070974 -0.363011773  [83,] 0.49487676 2.38124353 1.335743176  [84,] 2.53897708 0.08744336 0.474251393  [85,] -0.83532015 1.47367055 0.610093576  [86,] -0.78790461 2.02662652 -0.254723404  [87,] 0.80683216 2.23383039 0.772855797  [88,] 0.55804262 2.37298543 2.307611404  [89,] 1.11511104 1.80224719 0.959253308  [90,] 0.55572283 2.65754004 0.849126898  [91,] 1.34928528 2.11800147 -0.047652321  [92,] 1.56448261 1.85221452 0.781067031  [93,] 1.93255561 1.55949546 -0.089274676  [94,] -0.74666594 2.31293171 0.114679769  [95,] -0.95745536 2.22352843 0.142444774  [96,] -2.54386518 -0.16927402 0.788696991  [97,] 0.54395259 0.36892655 1.308895932  [98,] -1.03104975 2.56556935 -1.086390174  [99,] -2.25190942 1.43274138 -0.230208244  [100,] -1.41021602 2.16619177 0.748896411  [101,] -0.79771979 2.37694880 -1.568112531  [102,] 0.54953173 2.29312864 -1.498935323  [103,] 0.16117374 1.16448332 1.003713103  [104,] 0.65979494 2.67996119 -0.764920868  [105,] -0.39235441 2.09873171 -0.471850008  [106,] 1.77249908 1.71728847 0.947033174  [107,] 0.36626736 2.16935330 -0.481324235  [108,] 1.62067257 1.35558339 0.287159001  [109,] -0.08253578 2.30623459 -0.463574989  [110,] -1.57827507 1.46203429 1.779645955  [111,] -1.42056925 1.41820664 0.139275829  [112,] 0.27870275 1.93056809 0.078670553  [113,] 1.30314497 0.76317231 1.999596510  [114,] 0.45707187 2.26941561 1.061338968  [115,] 0.49418585 1.93904505 1.323938072  [116,] -0.48207441 3.87178385 1.344271223  [117,] 0.25288888 2.82149237 -0.302639785  [118,] 0.10722764 1.92892204 0.690148243  [119,] 2.43301260 1.25714104 -1.903027404  [120,] 0.55108954 2.22216155 -0.356228830  [121,] -0.73962193 1.40895667 1.125345492  [122,] -1.33632173 -0.25333693 5.345388179  [123,] 1.17708700 0.66396684 3.010221888  [124,] 0.46233501 0.61828818 0.483442366  [125,] -0.97847408 1.44557050 1.481236975  [126,] 0.09680973 2.10999799 0.434826116  [127,] -0.03848715 1.26676211 0.687577913  [128,] 1.59715850 1.20814357 3.361175555  [129,] 0.47956492 1.93884066 1.296507519  [130,] 1.79283347 1.15028810 0.782800173  [131,] 1.32710166 -0.17038923 -1.180013355  [132,] 2.38450083 -0.37458261 -0.723822595  [133,] 2.93694010 -0.26386183 -0.167639816  [134,] 2.14681113 -0.36825495 -0.453301301  [135,] 2.36986949 0.45963481 -1.101399789  [136,] 3.06384157 -0.35341284 -1.099124104  [137,] 3.91575378 -0.15458252 0.221827800  [138,] 3.93646339 -0.65968723 1.712215419  [139,] 3.09427612 -0.34884276 -1.026831413  [140,] 2.37447163 -0.29198035 1.241914333  [141,] 2.77881295 -0.28680487 0.609670124  [142,] 2.28656128 -0.37250784 -0.971643032  [143,] 2.98563349 -0.48921791 0.946952932  [144,] 2.37519470 -0.48233372 -0.252883994  [145,] 2.20986553 -1.16005250 -1.245125226  [146,] 2.62562100 -0.56316076 -0.855961082  [147,] 4.28063878 -0.64967096 -1.458196962  [148,] 3.58264137 -1.27270275 -0.110784038  [149,] 2.80706372 -1.57053379 -0.472527935  [150,] 2.89965933 -2.04105701 -0.495959810  [151,] 2.32073698 -2.35636608 0.437681744  [152,] 2.54983095 -2.04528309 -0.312267999  [153,] 1.81254128 -1.52764595 1.362589782  [154,] 2.76014464 -2.13893235 -0.964628688  [155,] 2.73715050 -0.40988627 -1.190404684  [156,] 3.60486887 -1.80238422 -0.094036861  [157,] 2.88982600 -1.92521861 -0.782322556  [158,] 3.39215608 -1.31187639 1.602025969  [159,] 1.04818190 -3.51508969 1.160038566  [160,] 1.60991228 -2.40663816 0.548559697  [161,] 3.14313097 -0.73816104 -0.090998724  [162,] 2.24015690 -1.17546529 -0.101376932  [163,] 2.84767378 -0.55604397 0.804215218  [164,] 2.59749706 -0.69796554 -0.884939521  [165,] 2.94929937 -1.55530896 -0.983400727  [166,] 3.53003227 -0.88252680 -0.466029128  [167,] 2.40611054 -2.59235618 0.428226211  [168,] 2.92908473 -1.27444695 -1.213358272  [169,] 2.18141278 -2.07753731 0.763782552  [170,] 2.38092779 -2.58866743 1.418044029  [171,] 3.21161722 0.25124910 -0.847129152  [172,] 3.67791872 -0.84774784 -1.339420231  [173,] 2.46555580 -2.19379830 -0.918780960  [174,] 3.37052415 -2.21628914 -0.342569512  [175,] 2.60195585 -1.75722935 0.207581355  [176,] 2.67783946 -2.76089913 -0.940941877  [177,] 2.38701709 -2.29734668 -0.550696197  [178,] 3.20875816 -2.76891957 1.013913664  wine <- cbind(wine ,pcobj$scores[,1:3])  > View(wine)  Type Alcohol Malic Ash Alcalinity Magnesium Phenols Flavanoids Nonflavanoids  1 1 14.23 1.71 2.43 15.6 127 2.80 3.06 0.28  2 1 13.20 1.78 2.14 11.2 100 2.65 2.76 0.26  3 1 13.16 2.36 2.67 18.6 101 2.80 3.24 0.30  4 1 14.37 1.95 2.50 16.8 113 3.85 3.49 0.24  5 1 13.24 2.59 2.87 21.0 118 2.80 2.69 0.39  6 1 14.20 1.76 2.45 15.2 112 3.27 3.39 0.34  7 1 14.39 1.87 2.45 14.6 96 2.50 2.52 0.30  8 1 14.06 2.15 2.61 17.6 121 2.60 2.51 0.31  9 1 14.83 1.64 2.17 14.0 97 2.80 2.98 0.29  10 1 13.86 1.35 2.27 16.0 98 2.98 3.15 0.22  11 1 14.10 2.16 2.30 18.0 105 2.95 3.32 0.22  12 1 14.12 1.48 2.32 16.8 95 2.20 2.43 0.26  13 1 13.75 1.73 2.41 16.0 89 2.60 2.76 0.29  14 1 14.75 1.73 2.39 11.4 91 3.10 3.69 0.43  15 1 14.38 1.87 2.38 12.0 102 3.30 3.64 0.29  16 1 13.63 1.81 2.70 17.2 112 2.85 2.91 0.30  17 1 14.30 1.92 2.72 20.0 120 2.80 3.14 0.33  18 1 13.83 1.57 2.62 20.0 115 2.95 3.40 0.40  19 1 14.19 1.59 2.48 16.5 108 3.30 3.93 0.32  20 1 13.64 3.10 2.56 15.2 116 2.70 3.03 0.17  21 1 14.06 1.63 2.28 16.0 126 3.00 3.17 0.24  22 1 12.93 3.80 2.65 18.6 102 2.41 2.41 0.25  23 1 13.71 1.86 2.36 16.6 101 2.61 2.88 0.27  24 1 12.85 1.60 2.52 17.8 95 2.48 2.37 0.26  25 1 13.50 1.81 2.61 20.0 96 2.53 2.61 0.28  26 1 13.05 2.05 3.22 25.0 124 2.63 2.68 0.47  27 1 13.39 1.77 2.62 16.1 93 2.85 2.94 0.34  28 1 13.30 1.72 2.14 17.0 94 2.40 2.19 0.27  29 1 13.87 1.90 2.80 19.4 107 2.95 2.97 0.37  30 1 14.02 1.68 2.21 16.0 96 2.65 2.33 0.26  31 1 13.73 1.50 2.70 22.5 101 3.00 3.25 0.29  32 1 13.58 1.66 2.36 19.1 106 2.86 3.19 0.22  33 1 13.68 1.83 2.36 17.2 104 2.42 2.69 0.42  34 1 13.76 1.53 2.70 19.5 132 2.95 2.74 0.50  35 1 13.51 1.80 2.65 19.0 110 2.35 2.53 0.29  36 1 13.48 1.81 2.41 20.5 100 2.70 2.98 0.26  37 1 13.28 1.64 2.84 15.5 110 2.60 2.68 0.34  38 1 13.05 1.65 2.55 18.0 98 2.45 2.43 0.29  39 1 13.07 1.50 2.10 15.5 98 2.40 2.64 0.28  40 1 14.22 3.99 2.51 13.2 128 3.00 3.04 0.20  41 1 13.56 1.71 2.31 16.2 117 3.15 3.29 0.34  42 1 13.41 3.84 2.12 18.8 90 2.45 2.68 0.27  43 1 13.88 1.89 2.59 15.0 101 3.25 3.56 0.17  44 1 13.24 3.98 2.29 17.5 103 2.64 2.63 0.32  45 1 13.05 1.77 2.10 17.0 107 3.00 3.00 0.28  46 1 14.21 4.04 2.44 18.9 111 2.85 2.65 0.30  47 1 14.38 3.59 2.28 16.0 102 3.25 3.17 0.27  48 1 13.90 1.68 2.12 16.0 101 3.10 3.39 0.21  49 1 14.10 2.02 2.40 18.8 103 2.75 2.92 0.32  50 1 13.94 1.73 2.27 17.4 108 2.88 3.54 0.32  51 1 13.05 1.73 2.04 12.4 92 2.72 3.27 0.17  52 1 13.83 1.65 2.60 17.2 94 2.45 2.99 0.22  53 1 13.82 1.75 2.42 14.0 111 3.88 3.74 0.32  54 1 13.77 1.90 2.68 17.1 115 3.00 2.79 0.39  55 1 13.74 1.67 2.25 16.4 118 2.60 2.90 0.21  56 1 13.56 1.73 2.46 20.5 116 2.96 2.78 0.20  57 1 14.22 1.70 2.30 16.3 118 3.20 3.00 0.26  58 1 13.29 1.97 2.68 16.8 102 3.00 3.23 0.31  Proanthocyanins Color Hue Dilution Proline Comp.1 Comp.2  1 2.29 5.640000 1.040 3.92 1065 -3.31675081 -1.44346263  2 1.28 4.380000 1.050 3.40 1050 -2.20946492 0.33339289  3 2.81 5.680000 1.030 3.17 1185 -2.51674015 -1.03115130  4 2.18 7.800000 0.860 3.45 1480 -3.75706561 -2.75637191  5 1.82 4.320000 1.040 2.93 735 -1.00890849 -0.86983082  6 1.97 6.750000 1.050 2.85 1450 -3.05025392 -2.12240111  7 1.98 5.250000 1.020 3.58 1290 -2.44908967 -1.17485013  8 1.25 5.050000 1.060 3.58 1295 -2.05943687 -1.60896307  9 1.98 5.200000 1.080 2.85 1045 -2.51087430 -0.91807096  10 1.85 7.220000 1.010 3.55 1045 -2.75362819 -0.78943767  11 2.38 5.750000 1.250 3.17 1510 -3.47973668 -1.30233324  12 1.57 5.000000 1.170 2.82 1280 -1.75475290 -0.61197723  13 1.81 5.600000 1.150 2.90 1320 -2.11346234 -0.67570634  14 2.81 5.400000 1.250 2.73 1150 -3.45815682 -1.13062988  15 2.96 7.500000 1.200 3.00 1547 -4.31278391 -2.09597558  16 1.46 7.300000 1.280 2.88 1310 -2.30518820 -1.66255173  17 1.97 6.200000 1.070 2.65 1280 -2.17195527 -2.32730534  18 1.72 6.600000 1.130 2.57 1130 -1.89897118 -1.63136888  19 1.86 8.700000 1.230 2.82 1680 -3.54198508 -2.51834367  20 1.66 5.100000 0.960 3.36 845 -2.08452220 -1.06113799  21 2.10 5.650000 1.090 3.71 780 -3.12440254 -0.78689711  22 1.98 4.500000 1.030 3.52 770 -1.08657007 -0.24174355  23 1.69 3.800000 1.110 4.00 1035 -2.53522408 0.09184062  24 1.46 3.930000 1.090 3.63 1015 -1.64498834 0.51627893  25 1.66 3.520000 1.120 3.82 845 -1.76157587 0.31714893  26 1.92 3.580000 1.130 3.20 830 -0.99007910 -0.94066734  27 1.45 4.800000 0.920 3.22 1195 -1.77527763 -0.68617513  28 1.35 3.950000 1.020 2.77 1285 -1.23542396 0.08980704  29 1.76 4.500000 1.250 3.40 915 -2.18840633 -0.68956962  30 1.98 4.700000 1.040 3.59 1035 -2.25610898 -0.19146194  31 2.38 5.700000 1.190 2.71 1285 -2.50022003 -1.24083383  32 1.95 6.900000 1.090 2.88 1515 -2.67741105 -1.47187365  33 1.97 3.840000 1.230 2.87 990 -1.62857912 -0.05270445  34 1.35 5.400000 1.250 3.00 1235 -1.90269086 -1.63306043  35 1.54 4.200000 1.100 2.87 1095 -1.41038853 -0.69793432  36 1.86 5.100000 1.040 3.47 920 -1.90382623 -0.17671095  37 1.36 4.600000 1.090 2.78 880 -1.38486223 -0.65863985  38 1.44 4.250000 1.120 2.51 1105 -1.12220741 -0.11410976  39 1.37 3.700000 1.180 2.69 1020 -1.50219450 0.76943201  40 2.08 5.100000 0.890 3.53 760 -2.52980109 -1.80300198  41 2.34 6.130000 0.950 3.38 795 -2.58809543 -0.77961630  42 1.48 4.280000 0.910 3.00 1035 -0.66848199 -0.16996094  43 1.70 5.430000 0.880 3.56 1095 -3.07080699 -1.15591896  44 1.66 4.360000 0.820 3.00 680 -0.46220914 -0.33074213  45 2.03 5.040000 0.880 3.35 885 -2.10135193 0.07100892  46 1.25 5.240000 0.870 3.33 1080 -1.13616618 -1.77710739  47 2.19 4.900000 1.040 3.44 1065 -2.72660096 -1.19133469  48 2.14 6.100000 0.910 3.33 985 -2.82133927 -0.64625860  49 2.38 6.200000 1.070 2.75 1060 -2.00985085 -1.24702946  50 2.08 8.900000 1.120 3.10 1260 -2.70749130 -1.75196741  51 2.91 7.200000 1.120 2.91 1150 -3.21491747 -0.16699199  52 2.29 5.600000 1.240 3.37 1265 -2.85895983 -0.74527880  53 1.87 7.050000 1.010 3.26 1190 -3.50560436 -1.61273386  54 1.68 6.300000 1.130 2.93 1375 -2.22479138 -1.87516800  55 1.62 5.850000 0.920 3.20 1060 -2.14698782 -1.01675154  56 2.45 6.250000 0.980 3.03 1120 -2.46932948 -1.32900831  57 2.03 6.380000 0.940 3.31 970 -2.74151791 -1.43654878  58 1.66 6.000000 1.070 2.84 1270 -2.17374092 -1.21219984  Comp.3  1 -0.165739045  2 -2.026457374  3 0.982818670  4 -0.176191842  5 2.026688219  6 -0.629395827  7 -0.977094891  8 0.146281883  9 -1.770969027  10 -0.984247490  11 -0.422735217  12 -1.190878320  13 -0.865086426  14 -1.204276353  15 -1.263912752  16 0.217902616  17 0.831729866  18 0.794913792  19 -0.485458508  20 -0.164746678  21 -0.364887083  22 0.936961600  23 -0.311932659  24 0.143885095  25 0.890285647  26 3.820908008  27 -0.086700406  28 -1.386896545  29 1.394566881  30 -1.092657258  31 1.386017855  32 -0.332261728  33 -0.167128706  34 1.172082119  35 0.479743025  36 0.450835040  37 0.458438581  38 -0.039107277  39 -1.426177346  40 -0.343152389  41 -0.118477466  42 -0.783362548  43 -0.312758084  44 -0.201476496  45 -0.655849415  46 0.028705736  47 -0.539773261  48 -1.155552411  49 -0.057293988  50 -0.643113612  51 -1.973571680  52 0.004719502  53 -0.520774530  54 0.339549850  55 -0.957762762  56 0.513437453  57 -0.612473396  58 0.261779593  > clust\_data <- wine[,15:17]  > norm\_clust <- scale(clust\_data)  > dist1 <- dist(norm\_clust ,method = "euclidean")  > fit1 <- hclust(dist1 , method = "complete")  > plot(fit1)  C:\Users\admin\Desktop\EDA1.PNG  Forming into groups ,  > groups <- cutree(fit1 , 3)  > mem <- as.matrix(groups)  > final <- cbind(mem , wine)  > View(final)  C:\Users\admin\Desktop\EDA2.PNG  Therefore , 3 cluster have been formed .  We can save the new csv file by :  write.csv(final, file = "pca sol1.csv", row.names = F , col.names = F)  > getwd()  [1] "C:/Users/admin/Documents" |