Assignment On PCA : 15.06.2020

**PCA and K-Means Clustering**

df<-read.csv("wine.csv")

df<-scale(df[,-1])

pca<-princomp(df,scores = TRUE)

type(df)

summary(pca)

plot(pca)

pca$scores

Importance of components:

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

Standard deviation 2.1631951 1.5757366 1.1991447 0.9559347 0.92110518

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

SD 0.79878171 0.74022473 0.58867607 0.53596364 0.49949266

P.V 0.04935823 0.04238679 0.02680749 0.02222153 0.01930019

Cum Pr 0.85098116 0.89336795 0.92017544 0.94239698 0.96169717

Comp.11 Comp.12 Comp.13

Standard deviation 0.47383559 0.40966094 0.320619963

Proportion of Variance 0.01736836 0.01298233 0.007952149

Cumulative Proportion 0.97906553 0.99204785 1.000000000

It is inferred that the cumulative proportion of comp1, comp2,comp3,comp4 are 0.3619,.5540,.6652 and .7359 respectively.

Variance of comp1 is 0.36

> plot(pca)

> pca$scores

Comp.1 Comp.2 Comp.3 Comp.4 Comp.5 Comp.6

[1,] 3.30742097 1.43940225 0.165272830 0.215024629 -0.6910933491 0.22325037

[2,] 2.20324981 -0.33245507 2.020757060 0.290538743 0.2569298662 0.92451231

[3,] 2.50966069 1.02825072 -0.980054055 -0.722863199 0.2503269758 -0.54773096

[4,] 3.74649719 2.74861839 0.175696224 -0.566385602 0.3109643979 -0.11410911

[5,] 1.00607049 0.86738404 -2.020987257 0.408613140 -0.2976179585 0.40537608

[6,] 3.04167373 2.11643092 0.627625371 0.514187035 0.6302408998 -0.12308335

[7,] 2.44220051 1.17154534 0.974346376 0.065645327 1.0248708712 0.61837638

[8,] 2.05364379 1.60443714 -0.145870400 1.189253269 -0.0766871685 1.43575612

[9,] 2.50381135 0.91548847 1.765987389 -0.056112076 0.8897471068 0.12881767

[10,] 2.74588238 0.78721703 0.981478855 -0.348398777 0.4672350615 -0.16293204

[11,] 3.46994837 1.29866985 0.421546086 -0.026766256 0.3374229486 0.18238714

[12,] 1.74981688 0.61025577 1.187528444 0.887660354 0.7364950947 0.55149954

[13,] 2.10751729 0.67380561 0.862652985 0.355435369 1.2065252604 0.21447112

[14,] 3.44842921 1.12744948 1.200888789 -0.162001077 2.0174358634 -0.74368299

[15,] 4.30065228 2.09007971 1.260357435 -0.304913067 1.0267961441 -0.79340481

[16,] 2.29870383 1.65787506 -0.217289668 1.436537727 0.4682292849 0.42102574

[17,] 2.16584568 2.32075875 -0.829390256 0.910034178 0.0001146171 0.06634216

[18,] 1.89362947 1.62677993 -0.792677744 1.079335712 0.4374711602 -0.36390428

[19,] 3.53202167 2.51125971 0.484092940 0.907762119 1.1498357895 -0.30302223

[20,] 2.07865856 1.05815307 0.164283255 -0.483633148 -0.8800288984 1.38909922

[21,] 3.11561376 0.78468361 0.363860676 0.025489787 -0.9696782562 0.10662100

[22,] 1.08351361 0.24106354 -0.934325978 -1.027012761 -0.3150828842 1.20760838

[23,] 2.52809263 -0.09158228 0.311055210 0.048255114 0.4283731687 1.01208813

[24,] 1.64036108 -0.51482667 -0.143480354 0.412556253 0.3746630851 0.78229875

[25,] 1.75662066 -0.31625681 -0.887781323 0.114791797 0.5551025444 0.89622074

[26,] 0.98729406 0.93802129 -3.810160004 1.317843719 -0.1585574371 0.26438245

[27,] 1.77028387 0.68424496 0.086456523 0.232251626 1.1397276528 0.56986696

[28,] 1.23194878 -0.08955442 1.382995281 0.494288392 0.3748838960 0.60637766

[29,] 2.18225047 0.68762990 -1.390644041 0.775304769 0.8083039404 0.60037885

[30,] 2.24976267 0.19092336 1.089583673 -0.285347369 0.4817143982 0.33429055

[31,] 2.49318704 1.23734344 -1.382119063 0.365833248 0.6205822759 -0.58081677

[32,] 2.66987964 1.46773335 0.331327094 0.348371162 0.0863507438 -0.16436614

[33,] 1.62399801 0.05255620 0.166658582 0.747206466 0.6342735393 0.04419384

[34,] 1.89733870 1.62846673 -1.168785117 2.333845540 0.1862885543 0.32706515

[35,] 1.40642118 0.69597107 -0.478393534 1.057104252 -0.0988818107 0.70825206

[36,] 1.89847087 0.17621387 -0.449566866 -0.283314357 0.1624910774 0.17415047

[37,] 1.38096669 0.65678714 -0.457149018 1.281044479 0.2044349041 0.61874622

[38,] 1.11905070 0.11378878 0.038997270 0.953710919 0.3412486137 0.40797546

[39,] 1.49796891 -0.76726764 1.422165587 0.755377264 0.1653776245 0.41428362

[40,] 2.52268490 1.79793023 0.342187120 -1.183635402 -1.2942447997 1.53793245

[41,] 2.58081526 0.77742329 0.118144196 -0.474632558 -0.4021874933 -0.72314313

[42,] 0.66660159 0.16948285 0.781158992 -1.309390144 0.3707176678 1.32096536

[43,] 3.06216898 1.15266742 0.311878313 -0.548736715 0.3032701638 0.80564479

[44,] 0.46090897 0.32981177 0.200909754 -1.432453085 -0.2824281812 1.02571004

[45,] 2.09544094 -0.07080918 0.654004547 -0.755613344 -0.4702113166 -0.37076210

[46,] 1.13297020 1.77210849 -0.028624988 -0.689885129 0.0607528293 1.90848431

[47,] 2.71893118 1.18798353 0.538254909 -1.385641944 0.4954968511 1.00815470

[48,] 2.81340300 0.64444071 1.152301905 -0.984884671 0.0099489529 -0.28994752

[49,] 2.00419725 1.24352164 0.057132823 -0.229994321 0.3835690682 -0.52447475

[50,] 2.69987528 1.74703922 0.641304569 0.099949574 0.4933214313 -0.73021856

[51,] 3.20587409 0.16652226 1.968020131 -1.117531517 -0.0122436594 -1.22735670

[52,] 2.85091773 0.74318238 -0.004706226 0.214670996 0.7566740077 0.19158264

[53,] 3.49574328 1.60819732 0.519309620 -0.106236743 0.5999919386 -0.18302807

[54,] 2.21853316 1.86989325 -0.338594715 1.203571449 0.5750221802 0.24526282

[55,] 2.14094846 1.01389147 0.955068628 0.239642188 -0.8776288188 0.36075734

[56,] 2.46238340 1.32526988 -0.511993182 -0.254630339 -0.9095743022 -0.55330865

[57,] 2.73380617 1.43250785 0.610750542 -0.205168685 -0.4330423030 -0.11567147

[58,] 2.16762631 1.20878999 -0.261043221 0.503007163 0.6934656297 0.25626086

[59,] 3.13054925 1.72670828 0.284857863 -0.229918640 -0.0748297503 -0.35057324

[60,] -0.92596992 -3.06484062 4.572166474 1.048885610 -0.4558382584 0.38716053

[61,] -1.53814123 -1.37755758 0.872222677 2.881989998 0.9752572410 -0.03491267

[62,] -1.83108449 -0.82764942 1.601185434 1.448020162 0.2932485783 -0.68219233

[63,] 0.03052074 -1.25923400 1.779388569 1.203182038 0.3924971348 0.60537936

[64,] 2.04449433 -1.91961759 0.007348049 -0.719292864 0.0808984437 -0.76237558

[65,] -0.60796583 -1.90269154 -0.677446941 2.147019606 0.0497621781 0.19082721

[66,] 0.89769555 -0.76176263 -0.571748468 0.677450627 0.1422318296 -1.01466998

[67,] 2.24218226 -1.87929123 2.026124738 -1.403667656 0.7299478607 -1.29973214

[68,] 0.18286818 -2.42031869 1.066736425 -0.127059552 0.5330698330 0.06980397

[69,] -0.81051865 -0.21989369 0.705016628 2.481988187 0.5819183093 0.31654670

[70,] 1.97006319 -1.39933587 1.234793017 1.149065114 -4.1748799067 -0.49161077

[71,] -1.56779366 -0.88249373 0.627228614 1.168550068 -0.9626291085 -0.26865227

[72,] 1.65301884 -0.95402102 -1.947091704 -0.151780756 0.7849469419 0.31533543

[73,] -0.72333196 -1.06065342 -0.080106258 0.075805593 0.1796589739 0.54278556

[74,] 2.55501977 0.25946663 -3.364901977 0.976644436 -1.9752267229 0.31116188

[75,] 1.82741266 -1.28425547 -0.456990911 -0.170868753 -0.9984058503 -0.07763882

[76,] -0.86555129 -2.43722606 1.558935609 0.829121008 -0.7059884717 -0.11204768

Comp.7 Comp.8 Comp.9 Comp.10 Comp.11

[1,] -0.594748831 -0.0649558620 0.639638363 1.018083960 0.450293172

[2,] -0.053624345 -1.0215343244 -0.307977984 0.159252141 0.142256019

[3,] -0.423012184 0.3432478700 -1.174521288 0.113041979 0.285866453

[4,] 0.382258990 -0.6417831052 0.052396617 0.238739150 -0.757447643

[5,] -0.442825306 -0.4155283146 0.325899840 -0.078146041 0.524465629

[6,] -0.400523928 -0.3937826075 -0.151718098 -0.101708908 -0.404444427

[7,] -0.052741950 0.3708876323 -0.455730294 1.013703922 0.441188870

[8,] -0.032284522 -0.2323235967 0.123023282 0.733530845 -0.292729105

[9,] -0.124932651 0.4981726190 0.604882896 0.173616861 0.507501293

[10,] 0.871892740 -0.1501559310 0.229840799 0.178915404 -0.012443071

[11,] -0.247464317 1.2032168356 -0.523098376 -0.213934103 -0.730454090

[12,] 0.433044673 0.9823554302 -0.472697058 0.219663137 -0.041319595

[13,] 0.241914764 0.4602081723 -0.876341336 -0.096233900 -0.053894424

[14,] -1.471622150 0.3793157517 -0.025629849 -0.243964418 1.228335908

[15,] -0.997158084 0.4037524704 -0.837978701 -0.363407777 0.314857945

[16,] 0.180458451 -0.0838800838 -0.403320197 -0.797405742 -0.102395272

[17,] -0.109179799 0.3983112097 0.060883540 0.019458965 -0.078145997

[18,] -0.091388772 -0.1125228438 0.380958636 -0.400192040 -0.308212482

[19,] 0.033370080 0.0354999773 -0.440323908 -0.782370000 -0.915924594

[20,] 0.102183932 -0.5782839955 0.058587532 -0.149627221 0.833375698

[21,] -0.264017729 -0.1850793289 1.314753546 0.361072678 0.459922727

[22,] -0.296096446 0.1061518392 -0.571088240 -0.092172600 0.656976542

[23,] 0.127410238 0.0763985397 0.107958200 0.823195997 0.088831891

[24,] 0.666522005 -0.1947161703 -0.690454667 0.470355630 0.032414377

[25,] 0.621796848 0.3413110454 0.094716751 0.664555677 0.378441498

[26,] -0.480551753 -0.0946245600 0.110082846 0.464647997 0.438274917

[27,] 0.456739864 -0.9061648643 -0.740175072 0.441299105 0.033006111

[28,] 0.361989185 0.2358229894 -0.732726856 0.298589565 -0.863424134

[29,] -0.117601009 0.0745852854 0.458450401 -0.028027557 0.535337077

[30,] 0.157904890 0.4687420346 0.137115228 0.800102233 0.121573571

[31,] 0.414827569 1.0564552256 -0.536880935 -0.319834517 -0.375478395

[32,] 0.532739137 0.5303100989 -0.785302741 -0.218365933 -1.071778784

[33,] -0.977571981 0.4976668676 0.131159086 0.165104152 0.219890759

[34,] -0.981165516 -0.5112019939 0.709719686 0.099328518 -0.623103855

[35,] 0.229092689 0.2268349971 -0.395617175 0.199394968 0.175022934

[36,] 0.655976907 0.3537311273 0.207576115 0.335996604 -0.197951397

[37,] 0.070180035 -0.9580447329 -0.304426016 -0.202780516 1.096826564

[38,] 0.407454131 0.0292585208 -0.851462291 -0.312591091 -0.003796615

[39,] 0.026823287 0.0262811149 -0.277335236 -0.282495670 -0.283059303

[40,] -0.962971858 -0.7667706217 0.719301478 0.022617923 1.190164307

[41,] -0.547587261 -0.7040422599 0.815203971 0.323234830 0.281569393

[42,] -0.027115199 0.3251205326 -0.284283574 -0.072017291 -0.825148162

[43,] 1.008572513 -0.9272066341 -0.217628393 0.229442300 0.441279064

[44,] -0.569786981 -0.5641789109 0.292474332 -0.077972272 0.070522676

[45,] -0.020490323 -0.6225777810 0.298161976 0.428004182 -0.469775701

[46,] -0.172211361 -0.2559916712 0.551859263 0.186825402 -0.592809909

[47,] -0.737349957 0.2425697998 0.445343780 -0.110755261 -0.048497461

[48,] 0.472580927 -0.1418125697 0.405458879 0.253076120 -0.122643011

[49,] -0.227875796 0.7988747536 0.114958347 -0.027638117 0.103035622

[50,] -0.007598470 0.2505645425 0.177064872 -0.304524717 -0.545559068

[51,] -0.260077530 0.4667300836 -1.228991995 -0.591649070 0.366278474

[52,] 0.352751689 1.0691076152 -0.984183504 0.113833641 0.548138551

[53,] -0.151445420 -1.4171049791 0.401675513 -0.423652543 -0.291255780

[54,] -0.463852190 -0.3981536658 -0.314407352 -0.082250674 -0.308387538

[55,] 0.410789089 -0.2887533222 0.279747442 0.453496853 -0.171942056

[56,] 0.357392565 0.5360108047 -0.135264746 0.155448125 -0.331902617

[57,] 0.039119002 -0.3503989858 0.930462550 0.345876713 0.042558835

[58,] 0.218447363 -0.6101237638 -0.836627063 -0.454038172 -0.083485523

[59,] 0.840703768 -0.6245159754 -0.413469281 -0.202138744 -0.259140456

[60,] 0.092012004 -0.6759025733 0.374915198 -0.414140726 -0.364715885

[61,] -1.055874525 -1.0723508166 0.275221037 -0.580519671 -0.148077570

[62,] -0.253428652 -0.9796897296 0.904720113 -0.570313650 -0.150132358

[63,] 0.569377232 0.6181201431 1.198577375 -0.304343708 -0.251909058

[64,] 1.072811065 -0.2977087673 0.609338190 -1.403500735 -0.008235091

[65,] -0.501705814 -0.0157708190 0.328215591 -1.045635529 0.879811993

[66,] -0.192768964 -0.0928982387 -0.589112288 -0.678113555 0.806354871

[67,] 0.343749093 0.1762906463 0.805356247 -0.374860197 0.154514305

[68,] 1.350954476 0.2256707643 0.401264164 0.106502259 -0.478400643

[69,] -0.366059142 -1.2430922412 0.994047808 0.127668412 -0.231823702

[70,] -1.452448405 1.0395448639 0.373326001 0.414942921 -0.230820881

[71,] -0.150397545 0.7177820779 -1.086273653 0.669455133 -0.227755541

[72,] 1.487434688 1.5266430802 1.155464136 -0.702407629 0.736167009

[73,] 1.594403691 1.0493396300 0.895573153 0.386967535 -0.097219056

[74,] 0.448128905 1.1274521175 1.108799865 -0.133602176 -1.775447165

[75,] 1.371769374 -0.5275233145 -0.450361562 -0.239424241 -1.069751720

[76,] -0.421990367 -0.0647122710 -0.215631225 -0.940274209 0.172542998

Comp.12 Comp.13

[1,] -5.392891e-01 -0.066052305

[2,] -3.871456e-01 0.003626273

[3,] -5.819316e-04 0.021655423

[4,] 2.413388e-01 -0.368444194

[5,] 2.160547e-01 -0.079140320

[6,] 3.783654e-01 0.144747017

[7,] -1.408326e-01 -0.271013687

[8,] -3.785951e-01 -0.109853902

[9,] 6.334624e-01 0.141683863

[10,] -5.487788e-01 -0.042335430

[11,] 8.108354e-02 0.122257108

[12,] 1.621036e-01 0.142368192

[13,] 8.949406e-02 -0.004921359

[14,] 7.701610e-01 0.225186084

[15,] 1.424180e-01 -0.094457503

[16,] -4.963956e-01 -0.065814111

[17,] 4.993885e-01 0.337151717

[18,] 2.630832e-01 0.557720257

[19,] -1.646237e-01 0.513095770

[20,] -2.870648e-01 0.058544167

[21,] -4.721267e-01 -0.027701209

[22,] -6.807690e-01 -0.303788047

[23,] -4.807459e-01 -0.044807342

[24,] -4.824856e-01 -0.342783088

[25,] -3.131924e-01 -0.226906559

[26,] 1.092096e-01 -0.057938048

[27,] 2.082477e-01 0.002159407

[28,] 3.761563e-01 0.008567548

[29,] -4.943902e-02 -0.198140209

[30,] -1.353494e-01 -0.445353113

[31,] 6.076575e-01 0.085661637

[32,] -2.778477e-03 0.271783807

[33,] 2.035213e-01 0.245083969

[34,] -1.084192e-01 0.054384154

[35,] 2.190564e-01 0.106609539

[36,] -2.322912e-01 0.132380848

[37,] 1.468095e-01 -0.061118487

[38,] 4.010997e-01 0.005741893

[39,] 1.788494e-01 0.318416807

[40,] -1.684139e-01 -0.177785910

[41,] -1.960816e-01 0.076720047

[42,] 7.319633e-02 0.427646271

[43,] 1.858606e-01 -0.022889178

[44,] 4.944707e-02 0.201488532

[45,] -5.603156e-02 0.080838339

[46,] -4.661425e-02 0.061515247

[47,] 2.026562e-01 -0.151639543

[48,] 8.498488e-02 0.188756494

[49,] 3.170517e-01 0.128661414

[50,] -7.390721e-01 0.547126406

[51,] -4.098684e-01 0.080562538

[52,] -3.156914e-01 -0.051318448

[53,] 8.961886e-02 -0.130654479

[54,] -4.517518e-05 -0.159199866

[55,] -1.781425e-01 0.295282582

[56,] 1.697504e-01 -0.241059524

[57,] 5.384920e-02 -0.162028993

[58,] 1.252307e-01 0.111250970

[59,] 5.561413e-01 0.163636436

[60,] 6.006008e-01 -0.490949914

[61,] 3.019641e-01 -0.102504556

[62,] 8.382021e-02 0.315869217

[63,] 1.404344e-01 0.163287677

[64,] 2.302245e-01 -0.385180426

[65,] -2.903585e-01 0.029689070

[66,] 2.034458e-01 0.126114352

[67,] -1.790132e-01 0.119628586

[68,] -1.096439e+00 -0.015744894

[69,] 9.980893e-01 -0.284932394

[70,] -7.207946e-01 -0.495370049

[71,] 4.532478e-01 0.378542475

[72,] 4.622716e-01 -0.338601200

[73,] 1.761363e-01 0.260052575

[74,] 1.073558e-01 -0.248619545

[75,] 4.634538e-01 -1.015292954

pca$scores

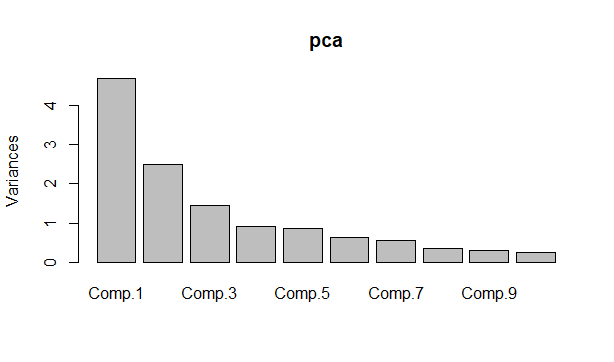
View(pca)

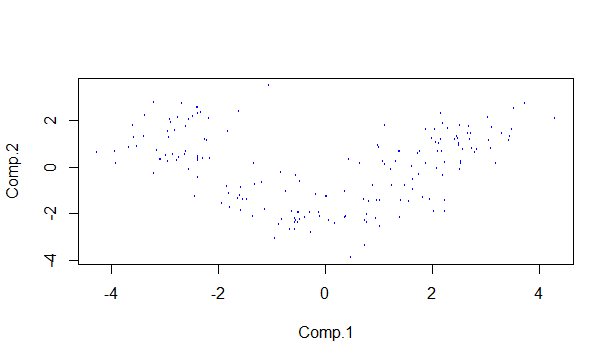
new\_data<-pca$scores[,1:2]

#pca$loadings

loadings(pca)

plot(pca$scores[,4:5],col="Blue",cex = 0.2)





install.packages('factoextra')

install.packages("NbClust")

library(factoextra)

library(cluster)

library(fpc)

library(NbClust)

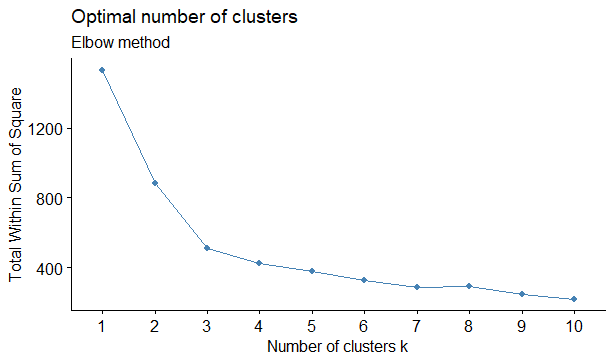
#The first three components are taken for cluster analysis

# Elbow method

pca$scores[,1:3]

###Cluster algorithm building

fviz\_nbclust(pca$scores[,1:3], kmeans, method = "wss") + labs(subtitle = "Elbow method")



From this, it has been inferred that the clusters that are formed are 3.

Comp.1 Comp.2 Comp.3

1 -0.07251558 -1.7725440 -0.192102252

2 -2.71238444 1.1224849 0.238420685

3 2.26940641 0.8920014 0.002144273

> km$cluster

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

[40] 3 3 3 3 3 3 3 3 3 3 3 3 3 3 3 3 3 3 3 3 1 1 2 1 1 1 1 1 1 1 1 1 1 1 3 1 1 1 1

[79] 1 1 1 1 1 2 1 1 1 1 1 1 1 1 1 1 1 3 1 1 3 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1

[118] 1 2 1 1 3 1 1 1 1 1 1 1 1 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2

[157] 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2

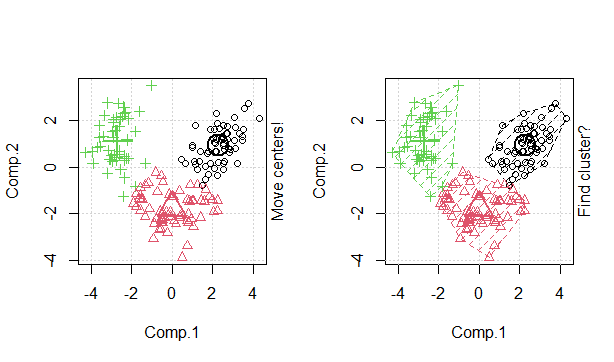
##Animation

install.packages("animation")

library(animation)

ani.options(interval=1)

km <- kmeans.ani(pca$scores[,1:3], 3)



**Hierarchical Clustering after PCA**

d <- dist(pca$scores[,1:3], method = "euclidean") #Computing the distance natrix

as.matrix(d)[1:6, 1:6]

fit <- hclust(d, method="average") # Building the algorithm # try with 'centroid'

plot(fit) # display dendogram

clusters <- cutree(fit, k=3) # cut tree into 3 clusters

table(clusters)

clusters

1 2 3

125 1 52

fit <- hclust(d, method="average") # Building the algorithm # try with 'centroid'

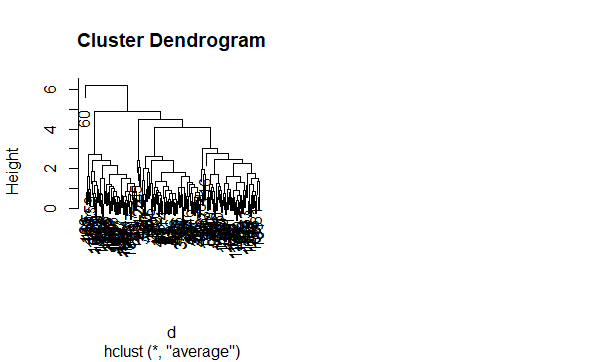
plot(fit) # display dendogram

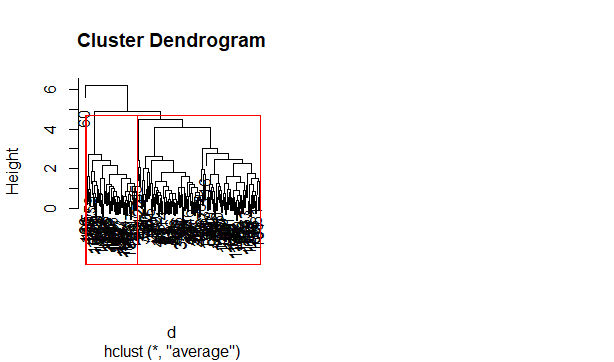
clusters <- cutree(fit, k=3) # cut tree into 3 clusters

table(clusters)

# draw dendogram with red borders around the 3 clusters

rect.hclust(fit, k=3, border="red")





Conclusion: Clusters formed from after the dimensionality reduction using PCA is 3.