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.

**Ans :**

**R Code :**

## PCA

########## Wine Data Set #########

#install.packages("gdata")

#install.packages("xlsx")

library(gdata)

wine <- read.csv('D:\\Data Science\\Excelr\\Assignments\\Assignment\\PCA\\wine.csv')

pca <- princomp(wine[,2:14], cor = TRUE, covmat = NULL)

summary(pca)

pca$scores

#pca$loadings

plot(pca$scores[,1:2], col="BLUE", pch=18,cex=0.5, lwd=3)

text(pca$scores[,1:2], labels = c(1:25), cex = 1)

pc <- pca$scores[,1:3]

######### Hierarchial Clustering

pcs <- scale(pc)

d <- dist(pcs, method = "euclidean") # Computing the distance matrix

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

plot(fit) #display dendrogram

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

#draw dendrogram with red borders around the 5 clusters

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

#attach the cluster number to universities

write.csv(clusters,file = "clusters-hc-type.csv")

fit1 <- hclust(d, method = "centroid") #Computing the algorithm # trying with centroid

plot(fit1) #display dendrogram

groups <- cutree(fit1, k=3) #cut tree into 3 clusters

#draw dendrogram with red borders around the 3 clusters

rect.hclust(fit1, k=3, border = "red")

#attach the cluster number to universities

clusters1=data.frame('Type'=wine[,1],'Cluster'=groups)

write.csv(clusters,file = "clusters1-hc-type.csv")

########### K-means Clustering

#Elbow Chart

wss <- c()

for (i in 2:15) wss[i] <- sum(kmeans(pc, centers = i)$withinss)

plot(1:15,wss,type = "b", xlab = "No of Clusters", ylab = "Avg Distance")

### Cluster algorithm building

km <- kmeans(pc, 3)

km$centers

km$cluster

##Animation

install.packages("animation")

library(animation)

windows()

km <- kmeans.ani(pc, 3)

**Results :**

> pca <- princomp(wine[,2:14], cor = TRUE, covmat = NULL)

> summary(pca)

Importance of components:

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

Standard deviation 2.1692972 1.5801816 1.2025273 0.9586313 0.92370351 0.80103498 0.74231281 0.59033665

Proportion of Variance 0.3619885 0.1920749 0.1112363 0.0706903 0.06563294 0.04935823 0.04238679 0.02680749

Cumulative Proportion 0.3619885 0.5540634 0.6652997 0.7359900 0.80162293 0.85098116 0.89336795 0.92017544

Comp.9 Comp.10 Comp.11 Comp.12 Comp.13

Standard deviation 0.53747553 0.50090167 0.47517222 0.41081655 0.321524394

Proportion of Variance 0.02222153 0.01930019 0.01736836 0.01298233 0.007952149

Cumulative Proportion 0.94239698 0.96169717 0.97906553 0.99204785 1.000000000

> pca$scores

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

[1,] 3.31675081 1.44346263 0.165739045 0.215631188 -0.6930428406 0.22388013 -0.596426546

[2,] 2.20946492 -0.33339289 2.026457374 0.291358318 0.2576546345 0.92712024 -0.053775613

[3,] 2.51674015 1.03115130 -0.982818670 -0.724902309 0.2510331182 -0.54927605 -0.424205451

[4,] 3.75706561 2.75637191 0.176191842 -0.567983308 0.3118415912 -0.11443100 0.383337297

[5,] 1.00890849 0.86983082 -2.026688219 0.409765788 -0.2984575030 0.40651960 -0.444074463

[6,] 3.05025392 2.12240111 0.629395827 0.515637495 0.6320187338 -0.12343056 -0.401653758

[7,] 2.44908967 1.17485013 0.977094891 0.065830505 1.0277619090 0.62012074 -0.052890728

[8,] 2.05943687 1.60896307 -0.146281883 1.192608010 -0.0769034938 1.43980622 -0.032375592

[9,] 2.51087430 0.91807096 1.770969027 -0.056270361 0.8922569767 0.12918105 -0.125285071

[10,] 2.75362819 0.78943767 0.984247490 -0.349381568 0.4685530755 -0.16339165 0.874352245

[11,] 3.47973668 1.30233324 0.422735217 -0.026841760 0.3383747782 0.18290163 -0.248162384

[12,] 1.75475290 0.61197723 1.190878320 0.890164338 0.7385726591 0.55305525 0.434266241

[13,] 2.11346234 0.67570634 0.865086426 0.356438010 1.2099287234 0.21507612 0.242597176

[14,] 3.45815682 1.13062988 1.204276353 -0.162458063 2.0231268080 -0.74578083 -1.475773419

[15,] 4.31278391 2.09597558 1.263912752 -0.305773190 1.0296926129 -0.79564291 -0.999970947

[16,] 2.30518820 1.66255173 -0.217902616 1.440590027 0.4695501036 0.42221340 0.180967502

[17,] 2.17195527 2.32730534 -0.831729866 0.912601275 0.0001149404 0.06652931 -0.109487782

[18,] 1.89897118 1.63136888 -0.794913792 1.082380388 0.4387052138 -0.36493081 -0.091646569

[19,] 3.54198508 2.51834367 0.485458508 0.910322807 1.1530793383 -0.30387702 0.033464213

[20,] 2.08452220 1.06113799 0.164746678 -0.484997419 -0.8825113544 1.39301770 0.102472181

[21,] 3.12440254 0.78689711 0.364887083 0.025561691 -0.9724136023 0.10692176 -0.264762491

[22,] 1.08657007 0.24174355 -0.936961600 -1.029909841 -0.3159716952 1.21101490 -0.296931698

[23,] 2.53522408 -0.09184062 0.311932659 0.048391236 0.4295815581 1.01494311 0.127769647

[24,] 1.64498834 -0.51627893 -0.143885095 0.413720024 0.3757199648 0.78450552 0.668402184

[25,] 1.76157587 -0.31714893 -0.890285647 0.115115611 0.5566684221 0.89874887 0.623550863

[26,] 0.99007910 0.94066734 -3.820908008 1.321561198 -0.1590047086 0.26512824 -0.481907331

[27,] 1.77527763 0.68617513 0.086700406 0.232906780 1.1429426878 0.57147449 0.458028272

[28,] 1.23542396 -0.08980704 1.386896545 0.495682720 0.3759413985 0.60808818 0.363010313

[29,] 2.18840633 0.68956962 -1.394566881 0.777491811 0.8105840689 0.60207244 -0.117932747

[30,] 2.25610898 0.19146194 1.092657258 -0.286152299 0.4830732567 0.33523354 0.158350321

[31,] 2.50022003 1.24083383 -1.386017855 0.366865220 0.6223328641 -0.58245518 0.415997748

[32,] 2.67741105 1.47187365 0.332261728 0.349353875 0.0865943288 -0.16482980 0.534241930

[33,] 1.62857912 0.05270445 0.167128706 0.749314246 0.6360627488 0.04431850 -0.980329594

[34,] 1.90269086 1.63306043 -1.172082119 2.340429039 0.1868140519 0.32798776 -0.983933266

[35,] 1.41038853 0.69793432 -0.479743025 1.060086216 -0.0991607444 0.71024995 0.229738932

[36,] 1.90382623 0.17671095 -0.450835040 -0.284113552 0.1629494453 0.17464173 0.657827339

[37,] 1.38486223 0.65863985 -0.458438581 1.284658151 0.2050115905 0.62049163 0.070378004

[38,] 1.12220741 0.11410976 0.039107277 0.956401223 0.3422112351 0.40912631 0.408603511

[39,] 1.50219450 -0.76943201 1.426177346 0.757508093 0.1658441349 0.41545226 0.026898952

[40,] 2.52980109 1.80300198 0.343152389 -1.186974296 -1.2978957091 1.54227077 -0.965688286

[41,] 2.58809543 0.77961630 0.118477466 -0.475971440 -0.4033220159 -0.72518303 -0.549131939

[42,] 0.66848199 0.16996094 0.783362548 -1.313083777 0.3717634179 1.32469164 -0.027191688

[43,] 3.07080699 1.15591896 0.312758084 -0.550284635 0.3041256526 0.80791742 1.011417576

[44,] 0.46220914 0.33074213 0.201476496 -1.436493863 -0.2832248773 1.02860345 -0.571394281

[45,] 2.10135193 -0.07100892 0.655849415 -0.757744839 -0.4715377264 -0.37180798 -0.020548124

[46,] 1.13616618 1.77710739 -0.028705736 -0.691831212 0.0609242058 1.91386791 -0.172697149

[47,] 2.72660096 1.19133469 0.539773261 -1.389550674 0.4968945883 1.01099858 -0.739429932

[48,] 2.82133927 0.64625860 1.155552411 -0.987662912 0.0099770177 -0.29076542 0.473914021

[49,] 2.00985085 1.24702946 0.057293988 -0.230643107 0.3846510705 -0.52595424 -0.228518606

[50,] 2.70749130 1.75196741 0.643113612 0.100231520 0.4947130319 -0.73227842 -0.007619905

[51,] 3.21491747 0.16699199 1.973571680 -1.120683939 -0.0122781973 -1.23081893 -0.260811177

[52,] 2.85895983 0.74527880 -0.004719502 0.215276557 0.7588084943 0.19212307 0.353746760

[53,] 3.50560436 1.61273386 0.520774530 -0.106536424 0.6016844439 -0.18354437 -0.151872629

[54,] 2.22479138 1.87516800 -0.339549850 1.206966579 0.5766442488 0.24595468 -0.465160662

[55,] 2.14698782 1.01675154 0.957762762 0.240318190 -0.8801045044 0.36177500 0.411947876

[56,] 2.46932948 1.32900831 -0.513437453 -0.255348621 -0.9121401022 -0.55486947 0.358400726

[57,] 2.74151791 1.43654878 0.612473396 -0.205747441 -0.4342638634 -0.11599777 0.039229352

[58,] 2.17374092 1.21219984 -0.261779593 0.504426086 0.6954218130 0.25698374 0.219063576

[59,] 3.13938015 1.73157912 0.285661413 -0.230567212 -0.0750408361 -0.35156217 0.843075292

[60,] -0.92858197 -3.07348616 4.585064007 1.051844391 -0.4571241235 0.38825266 0.092271559

[61,] -1.54248014 -1.38144351 0.874683112 2.890119746 0.9780083247 -0.03501116 -1.058853020

[62,] -1.83624976 -0.82998412 1.605702186 1.452104853 0.2940757974 -0.68411671 -0.254143544

[63,] 0.03060683 -1.26278614 1.784408010 1.206576071 0.3936043221 0.60708706 0.570983377

[64,] 2.05026161 -1.92503260 0.007368777 -0.721321903 0.0811266485 -0.76452615 1.075837337

[65,] -0.60968083 -1.90805881 -0.679357938 2.153076091 0.0499025512 0.19136551 -0.503121066

[66,] 0.90022784 -0.76391147 -0.573361302 0.679361634 0.1426330485 -1.01753224 -0.193312742

[67,] 2.24850719 -1.88459248 2.031840193 -1.407627234 0.7320069560 -1.30339853 0.344718768

[68,] 0.18338403 -2.42714611 1.069745560 -0.127417972 0.5345735591 0.07000088 1.354765357

[69,] -0.81280503 -0.22051399 0.707005396 2.488989578 0.5835598311 0.31743964 -0.367091751

[70,] 1.97562050 -1.40328323 1.238276220 1.152306489 -4.1866567422 -0.49299755 -1.456545588

[71,] -1.57221622 -0.88498314 0.628997950 1.171846407 -0.9653445698 -0.26941010 -0.150821799

[72,] 1.65768181 -0.95671220 -1.952584217 -0.152208912 0.7871611831 0.31622496 1.491630563

[73,] -0.72537239 -1.06364540 -0.080332229 0.076019431 0.1801657704 0.54431669 1.598901312

[74,] 2.56222717 0.26019855 -3.374393962 0.979399432 -1.9807986006 0.31203963 0.449393023

[75,] 1.83256757 -1.28787820 -0.458280027 -0.171350754 -1.0012222334 -0.07785783 1.375638971

[76,] -0.86799290 -2.44410119 1.563333179 0.831459859 -0.7079799805 -0.11236375 -0.423180752

Comp.8 Comp.9 Comp.10 Comp.11 Comp.12 Comp.13

[1,] -0.0651390947 0.641442706 1.020955853 0.451563395 -5.408104e-01 -0.066238631

[2,] -1.0244159502 -0.308846753 0.159701372 0.142657306 -3.882377e-01 0.003636502

[3,] 0.3442161311 -1.177834471 0.113360857 0.286672847 -5.835732e-04 0.021716510

[4,] -0.6435934984 0.052544421 0.239412605 -0.759584312 2.420196e-01 -0.369483531

[5,] -0.4167004701 0.326819165 -0.078366482 0.525945083 2.166642e-01 -0.079363566

[6,] -0.3948934210 -0.152146076 -0.101995816 -0.405585316 3.794327e-01 0.145155331

[7,] 0.3719338618 -0.457015855 1.016563459 0.442433411 -1.412298e-01 -0.271778184

[8,] -0.2329789537 0.123370316 0.735600047 -0.293554859 -3.796630e-01 -0.110163787

[9,] 0.4995779043 0.606589198 0.174106613 0.508932893 6.352493e-01 0.142083536

[10,] -0.1505795029 0.230489152 0.179420103 -0.012478171 -5.503268e-01 -0.042454853

[11,] 1.2066109660 -0.524573974 -0.214537585 -0.732514613 8.131227e-02 0.122601981

[12,] 0.9851265371 -0.474030480 0.220282780 -0.041436153 1.625609e-01 0.142769795

[13,] 0.4615063644 -0.878813390 -0.096505364 -0.054046453 8.974651e-02 -0.004935242

[14,] 0.3803857560 -0.025702147 -0.244652612 1.231800896 7.723335e-01 0.225821307

[15,] 0.4048914077 -0.840342538 -0.364432906 0.315746122 1.428197e-01 -0.094723956

[16,] -0.0841166995 -0.404457915 -0.799655128 -0.102684116 -4.977959e-01 -0.065999764

[17,] 0.3994347979 0.061055286 0.019513856 -0.078366437 5.007972e-01 0.338102781

[18,] -0.1128402573 0.382033275 -0.401320934 -0.309081912 2.638254e-01 0.559293519

[19,] 0.0356001185 -0.441566009 -0.784576972 -0.918508308 -1.650881e-01 0.514543152

[20,] -0.5799152653 0.058752801 -0.150049302 0.835726551 -2.878746e-01 0.058709313

[21,] -0.1856014155 1.318462308 0.362091221 0.461220114 -4.734585e-01 -0.027779351

[22,] 0.1064512808 -0.572699212 -0.092432608 0.658829794 -6.826894e-01 -0.304644997

[23,] 0.0766140508 0.108262737 0.825518134 0.089082475 -4.821020e-01 -0.044933738

[24,] -0.1952654412 -0.692402357 0.471682447 0.032505814 -4.838466e-01 -0.343750038

[25,] 0.3422738429 0.094983936 0.666430309 0.379509036 -3.140759e-01 -0.227546635

[26,] -0.0948914846 0.110393377 0.465958713 0.439511238 1.095177e-01 -0.058101485

[27,] -0.9087210467 -0.742263017 0.442543957 0.033099217 2.088351e-01 0.002165498

[28,] 0.2364882178 -0.734793791 0.299431850 -0.865859750 3.772174e-01 0.008591716

[29,] 0.0747956816 0.459743634 -0.028106620 0.536847198 -4.957848e-02 -0.198699139

[30,] 0.4700642996 0.137502014 0.802359226 0.121916516 -1.357312e-01 -0.446609401

[31,] 1.0594353591 -0.538395411 -0.320736731 -0.376537574 6.093716e-01 0.085903278

[32,] 0.5318060401 -0.787517986 -0.218981918 -1.074802144 -2.786315e-03 0.272550476

[33,] 0.4990707261 0.131529070 0.165569891 0.220511044 2.040954e-01 0.245775321

[34,] -0.5126440334 0.711721720 0.099608712 -0.624861556 -1.087250e-01 0.054537565

[35,] 0.2274748714 -0.396733164 0.199957437 0.175516653 2.196743e-01 0.106910271

[36,] 0.3547289604 0.208161662 0.336944410 -0.198509795 -2.329465e-01 0.132754279

[37,] -0.9607472622 -0.305284766 -0.203352536 1.099920580 1.472237e-01 -0.061290895

[38,] 0.0293410556 -0.853864164 -0.313472873 -0.003807324 4.022311e-01 0.005758091

[39,] 0.0263552508 -0.278117565 -0.283292556 -0.283857779 1.793539e-01 0.319315023

[40,] -0.7689335897 0.721330542 0.022681726 1.193521618 -1.688889e-01 -0.178287423

[41,] -0.7060282787 0.817503563 0.324146636 0.282363666 -1.966347e-01 0.076936465

[42,] 0.3260376586 -0.285085504 -0.072220443 -0.827475806 7.340281e-02 0.428852610

[43,] -0.9298221728 -0.218242297 0.230089530 0.442523859 1.863849e-01 -0.022953746

[44,] -0.5657703920 0.293299367 -0.078192223 0.070721612 4.958655e-02 0.202056907

[45,] -0.6243339983 0.299003055 0.429211531 -0.471100882 -5.618962e-02 0.081066374

[46,] -0.2567137930 0.553415991 0.187352414 -0.594482155 -4.674574e-02 0.061688774

[47,] 0.2432540601 0.446600041 -0.111067688 -0.048634267 2.032279e-01 -0.152067300

[48,] -0.1422126059 0.406602629 0.253790017 -0.122988972 8.522461e-02 0.189288954

[49,] 0.8011282835 0.115282630 -0.027716081 0.103326273 3.179460e-01 0.129024352

[50,] 0.2512713550 0.177564351 -0.305383745 -0.547098024 -7.411569e-01 0.548669784

[51,] 0.4680466732 -1.232458834 -0.593318041 0.367311701 -4.110246e-01 0.080789795

[52,] 1.0721234395 -0.986959768 0.114154752 0.549684784 -3.165819e-01 -0.051463211

[53,] -1.4211024623 0.402808592 -0.424847616 -0.292077378 8.987166e-02 -0.131023039

[54,] -0.3992768095 -0.315294258 -0.082482694 -0.309257462 -4.530261e-05 -0.159648950

[55,] -0.2895678606 0.280536576 0.454776113 -0.172427084 -1.786451e-01 0.296115539

[56,] 0.5375228269 -0.135646311 0.155886626 -0.332838875 1.702293e-01 -0.241739524

[57,] -0.3513874193 0.933087273 0.346852390 0.042678888 5.400110e-02 -0.162486058

[58,] -0.6118448498 -0.838987088 -0.455318959 -0.083721026 1.255839e-01 0.111564796

[59,] -0.6262776600 -0.414635628 -0.202708953 -0.259871461 5.577101e-01 0.164098035

[60,] -0.6778092134 0.375972789 -0.415308968 -0.365744704 6.022950e-01 -0.492334825

[61,] -1.0753757897 0.275997402 -0.582157247 -0.148495279 3.028159e-01 -0.102793710

[62,] -0.9824533169 0.907272220 -0.571922436 -0.150555863 8.405666e-02 0.316760246

[63,] 0.6198637859 1.201958418 -0.305202225 -0.252619663 1.408305e-01 0.163748292

[64,] -0.2985485681 0.611057060 -1.407459842 -0.008258321 2.308740e-01 -0.386266973

[65,] -0.0158153066 0.329141448 -1.048585142 0.882293837 -2.911776e-01 0.029772819

[66,] -0.0931602935 -0.590774103 -0.680026432 0.808629501 2.040197e-01 0.126470106

[67,] 0.1767879411 0.807628060 -0.375917632 0.154950171 -1.795182e-01 0.119966044

[68,] 0.2263073545 0.402396082 0.106802689 -0.479750153 -1.099532e+00 -0.015789308

[69,] -1.2465988553 0.996851897 0.128028549 -0.232477648 1.000905e+00 -0.285736154

[70,] 1.0424772953 0.374379109 0.416113426 -0.231472000 -7.228278e-01 -0.496767428

[71,] 0.7198068551 -1.089337900 0.671343586 -0.228398013 4.545264e-01 0.379610298

[72,] 1.5309495572 1.158723562 -0.704389036 0.738243648 4.635757e-01 -0.339556353

[73,] 1.0522996913 0.898099457 0.388059124 -0.097493300 1.766331e-01 0.260786152

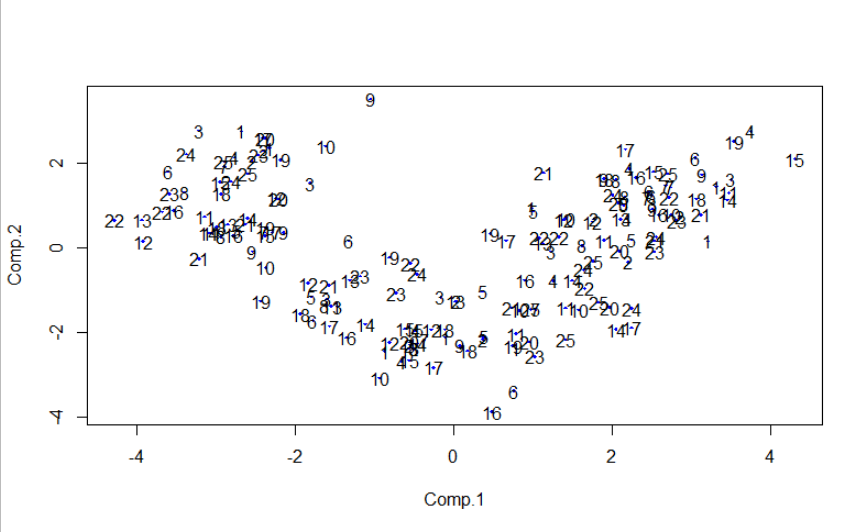
[74,] 1.1306325247 1.111927656 -0.133979052 -1.780455488 1.076587e-01 -0.249320871

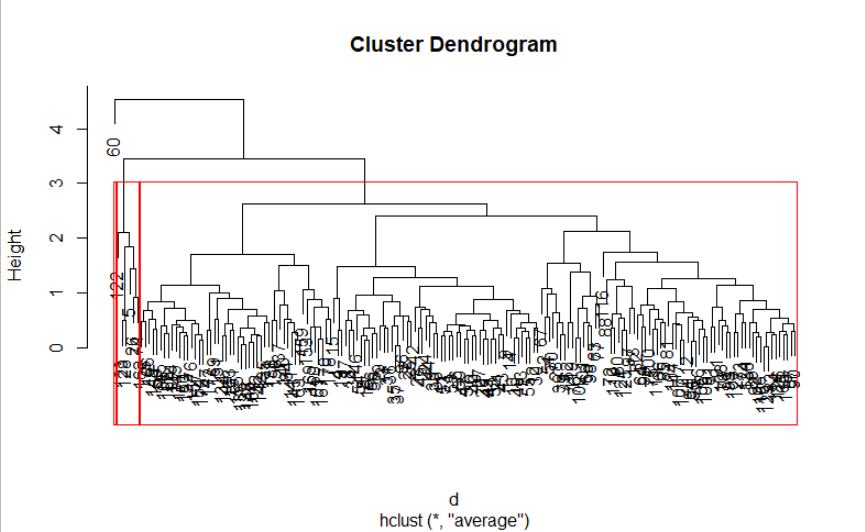
[75,] -0.5290113945 -0.451631978 -0.240099628 -1.072769362 4.647612e-01 -1.018156973

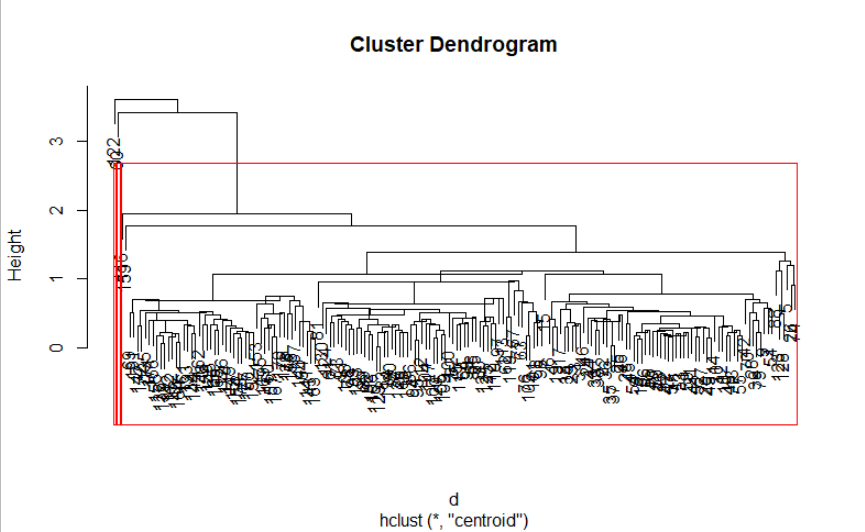
[76,] -0.0648948165 -0.216239495 -0.942926609 0.173029721 -5.997641e-01 0.272452796

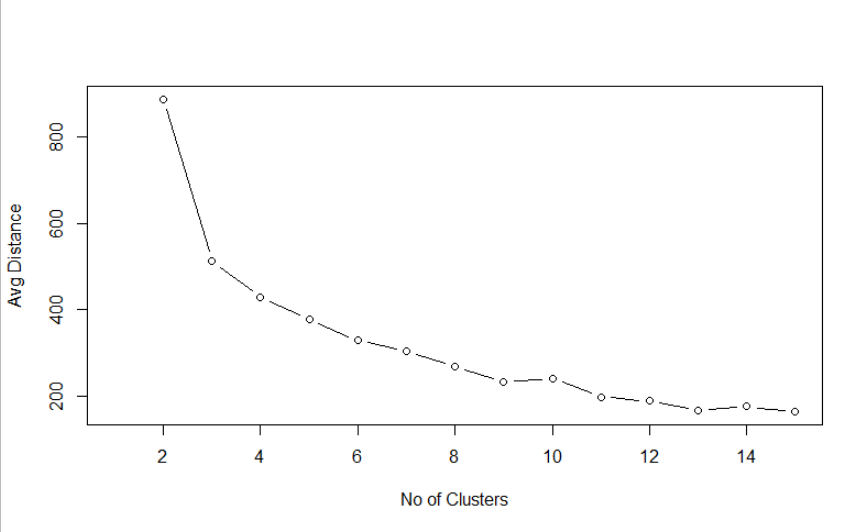
[ reached getOption("max.print") -- omitted 102 rows ]

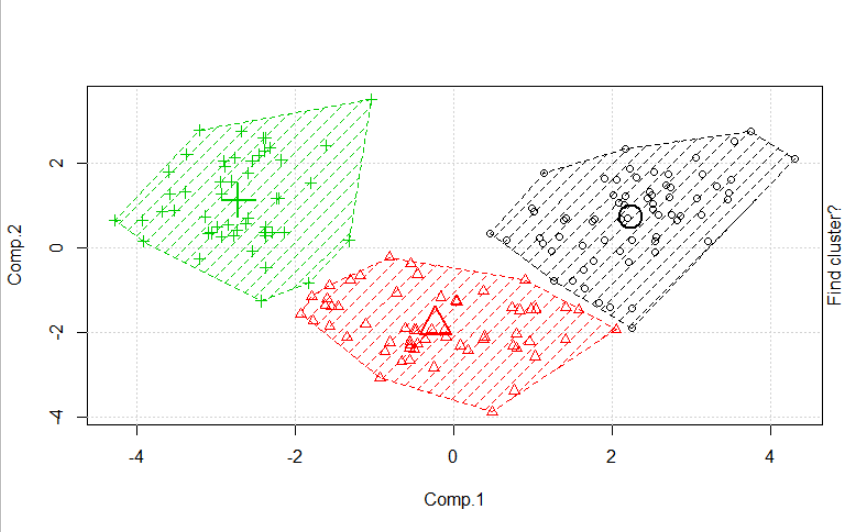
**Plots :**











**Inference :**

Performed PCA and with 3 major components done the clustering analysis and the results are given in separate csv file.