vorticity

Sirish Shrestha

August 13, 2017

R Markdown

Read the data for vorticity analysis

```
setwd("C:/Users/siris/Documents/rstudio-projects/vorticity")

## read all echo data
echo <- read.csv("data.csv", header = TRUE, sep = ",", stringsAsFactors = TRUE)

summary(echo)
```

```
##      ï..ID          Country        Age            Gender
##  Min.   : 1.00   Myao   :49   Min.   :24.00   Min.   :0.0000
##  1st Qu.:24.75   poland :26   1st Qu.:52.00   1st Qu.:0.0000
##  Median :48.50          : 1   Median :61.00   Median :0.0000
##  Mean   :49.67   FS10   : 1   Mean   :60.98   Mean   :0.1087
##  3rd Qu.:75.25   FS11   : 1   3rd Qu.:73.00   3rd Qu.:0.0000
##  Max.   :99.00   FS12   : 1   Max.   :83.00   Max.   :1.0000
##                  (Other):17   NA's   :4       NA's   :4
##       BSA            HRmin             SBP             DBP
##  Min.   :1.50   Min.   : 45.00   Min.   :100.0   Min.   : 50.00
##  1st Qu.:1.75   1st Qu.: 59.95   1st Qu.:119.2   1st Qu.: 64.00
##  Median :1.99   Median : 69.00   Median :130.0   Median : 78.00
##  Mean   :1.95   Mean   : 70.29   Mean   :133.9   Mean   : 76.69
##  3rd Qu.:2.09   3rd Qu.: 78.00   3rd Qu.:144.2   3rd Qu.: 87.25
##  Max.   :2.61   Max.   :114.00   Max.   :185.0   Max.   :106.00
##  NA's   :27     NA's   :16       NA's   :38      NA's   :38
##     Etiology          NYHA           EchoLA         EchoIVS
##  Min.   :1.000   Min.   :1.000   Min.   :28.30   Min.   : 1.350
##  1st Qu.:2.000   1st Qu.:1.000   1st Qu.:35.85   1st Qu.: 8.275
##  Median :2.000   Median :1.000   Median :41.00   Median : 9.750
##  Mean   :2.507   Mean   :1.407   Mean   :41.30   Mean   : 9.590
##  3rd Qu.:3.000   3rd Qu.:1.000   3rd Qu.:46.23   3rd Qu.:10.750
##  Max.   :6.000   Max.   :4.000   Max.   :55.10   Max.   :16.000
##  NA's   :27      NA's   :5       NA's   :50      NA's   :2
##      EchoPW           EchoDd          EchoDs          bEF
##  Min.   : 1.360   Min.   :33.00   Min.   :20.0   Min.   :12.00
##  1st Qu.: 8.000   1st Qu.:44.00   1st Qu.:28.0   1st Qu.:48.50
##  Median : 9.000   Median :48.65   Median :33.0   Median :61.00
##  Mean   : 9.407   Mean   :50.83   Mean   :36.4   Mean   :55.44
##  3rd Qu.:10.975   3rd Qu.:54.85   3rd Qu.:39.0   3rd Qu.:66.50
##  Max.   :16.000   Max.   :99.00   Max.   :91.0   Max.   :75.00
##  NA's   :2        NA's   :2       NA's   :3      NA's   :33
##       RWT             LVMass          EchoLVMI          EchoEDV
##  Min.   :0.1500   Min.   : 16.75   Min.   :  7.261   Min.   : 57.20
##  1st Qu.:0.3200   1st Qu.:130.07   1st Qu.: 68.863   1st Qu.: 88.47
##  Median :0.3500   Median :164.74   Median : 82.227   Median :111.00
```

```
##  Mean   :0.3644   Mean   :181.50   Mean   : 90.237   Mean   :134.34
##  3rd Qu.:0.4200   3rd Qu.:202.18   3rd Qu.:105.706   3rd Qu.:143.90
##  Max.   :0.5900   Max.   :477.98   Max.   :240.188   Max.   :455.50
##  NA's   :87       NA's   :2        NA's   :29        NA's   :2
##     EchoESV          EchoSV          EchoEF       AbnormalEF50
##  Min.   : 18.30   Min.   : 31.50   Min.   :10.00   Min.   :0.0000
##  1st Qu.: 30.00   1st Qu.: 51.00   1st Qu.:46.00   1st Qu.:0.0000
##  Median : 43.35   Median : 62.00   Median :57.00   Median :0.0000
##  Mean   : 69.79   Mean   : 65.27   Mean   :53.72   Mean   :0.2842
##  3rd Qu.: 74.00   3rd Qu.: 77.00   3rd Qu.:65.00   3rd Qu.:1.0000
##  Max.   :366.50   Max.   :129.00   Max.   :78.00   Max.   :1.0000
##  NA's   :2        NA's   :3        NA's   :1       NA's   :1
##   EchoTMfEwave     EchoTMfAwave     TimeTMFDecT     EchoTMfEtoA
##  Min.   : 24.00   Min.   : 21.00   Min.   :115.0   Min.   :0.360
##  1st Qu.: 59.40   1st Qu.: 53.00   1st Qu.:170.2   1st Qu.:0.807
##  Median : 68.50   Median : 67.10   Median :196.0   Median :1.096
##  Mean   : 71.29   Mean   : 67.58   Mean   :207.3   Mean   :1.207
##  3rd Qu.: 81.00   3rd Qu.: 83.00   3rd Qu.:238.0   3rd Qu.:1.408
##  Max.   :132.00   Max.   :120.00   Max.   :362.0   Max.   :6.000
##  NA's   :2        NA's   :4        NA's   :2       NA's   :3
##     EchoLAVI       EchSepTDIe     EchoLatTDIe      EchoAveTDIe
##  Min.   :16.34   Min.   : 2.00   Min.   : 1.000   Min.   : 1.500
##  1st Qu.:26.73   1st Qu.: 5.00   1st Qu.: 6.500   1st Qu.: 5.607
##  Median :33.92   Median : 6.00   Median : 8.190   Median : 7.405
##  Mean   :36.56   Mean   : 6.44   Mean   : 8.419   Mean   : 7.339
##  3rd Qu.:44.68   3rd Qu.: 8.00   3rd Qu.:10.000   3rd Qu.: 9.000
##  Max.   :94.00   Max.   :13.50   Max.   :21.300   Max.   :16.900
##  NA's   :8       NA's   :5       NA's   :11       NA's   :5
##    EchoSeptEe        EchoRVSP        EchoLVM         EchoLVMII
##  Min.   : 4.890   Min.   :21.00   Min.   : 16.75   Min.   :  7.261
##  1st Qu.: 8.188   1st Qu.:27.50   1st Qu.:130.07   1st Qu.: 68.863
##  Median :11.155   Median :32.00   Median :164.74   Median : 82.227
##  Mean   :12.853   Mean   :35.62   Mean   :181.50   Mean   : 90.237
##  3rd Qu.:14.238   3rd Qu.:38.50   3rd Qu.:202.18   3rd Qu.:105.706
##  Max.   :55.000   Max.   :76.00   Max.   :477.98   Max.   :240.188
##  NA's   :6        NA's   :53      NA's   :2        NA's   :29
##   TimePIVRRms     TimePIvEonset   TimePIvEonsetNorRR  TimePIvEpeak
##  Min.   : 586.9   Min.   :242.8   Min.   :227.8      Min.   :356.2
##  1st Qu.: 808.0   1st Qu.:397.8   1st Qu.:434.9      1st Qu.:498.0
##  Median : 898.9   Median :454.7   Median :497.1      Median :564.1
##  Mean   : 928.5   Mean   :446.5   Mean   :491.6      Mean   :552.6
##  3rd Qu.:1054.6   3rd Qu.:497.3   3rd Qu.:559.0      3rd Qu.:608.1
##  Max.   :1704.0   Max.   :614.1   Max.   :680.0      Max.   :738.9
##  NA's   :10       NA's   :12      NA's   :12         NA's   :12
##  TimePIvEpeakNorRR  TimeVSonset     TimePeakVSE       Tracking
##  Min.   :262.3     Min.   :366.0   Min.   :422.0   Min.   :0.0000
##  1st Qu.:542.5     1st Qu.:460.8   1st Qu.:507.2   1st Qu.:0.0000
##  Median :614.2     Median :506.5   Median :583.6   Median :1.0000
##  Mean   :609.8     Mean   :506.2   Mean   :577.2   Mean   :0.5761
##  3rd Qu.:683.4     3rd Qu.:542.5   3rd Qu.:617.4   3rd Qu.:1.0000
##  Max.   :840.6     Max.   :683.0   Max.   :864.0   Max.   :1.0000
##  NA's   :12        NA's   :20      NA's   :15      NA's   :4
##    VortexArea     VortexIntensity    VortexDepth      VortexLength
##  Min.   :0.0930   Min.   :-0.6000   Min.   :0.0990   Min.   :0.2305
```

```
##  1st Qu.:0.2473   1st Qu.:-0.5037   1st Qu.:0.3913   1st Qu.:0.4894
##  Median :0.3257   Median :-0.4479   Median :0.4528   Median :0.5918
##  Mean   :0.3071   Mean   :-0.3243   Mean   :0.4332   Mean   :0.5669
##  3rd Qu.:0.3682   3rd Qu.:-0.3212   3rd Qu.:0.4934   3rd Qu.:0.6539
##  Max.   :0.4549   Max.   : 0.4472   Max.   :0.6118   Max.   :0.8262
##
##  EnergyDissipation VorticityFluctuation KineticEnergyFluctuation
##  Min.   :0.1912    Min.   :0.6954       Min.   :0.8944
##  1st Qu.:0.6624    1st Qu.:0.8361       1st Qu.:1.1506
##  Median :1.0175    Median :0.8740       Median :1.3370
##  Mean   :1.1603    Mean   :0.8733       Mean   :1.3541
##  3rd Qu.:1.5210    3rd Qu.:0.9165       3rd Qu.:1.5201
##  Max.   :3.2823    Max.   :0.9757       Max.   :2.0779
##
##  ShearStressFluctuation KineticEnergyEarlyD KineticEnergyLateD
##  Min.   :-0.3042        Min.   :1.100       Min.   :0.5434
##  1st Qu.:-0.0037        1st Qu.:1.848       1st Qu.:1.5305
##  Median : 0.0919        Median :2.271       Median :1.9211
##  Mean   : 0.1100        Mean   :2.596       Mean   :2.0767
##  3rd Qu.: 0.2127        3rd Qu.:3.119       3rd Qu.:2.5937
##  Max.   : 0.6324        Max.   :7.120       Max.   :4.0359
##                         NA's   :2           NA's   :2
##  KineticEnergyOnsetSystole     LAVI35         SeptTDIe7
##  Min.   :0.2880            Min.   :0.0000   Min.   :0.0000
##  1st Qu.:0.6910            1st Qu.:0.0000   1st Qu.:0.0000
##  Median :0.8125            Median :0.0000   Median :1.0000
##  Mean   :0.8793            Mean   :0.4773   Mean   :0.5714
##  3rd Qu.:1.0286            3rd Qu.:1.0000   3rd Qu.:1.0000
##  Max.   :2.0183            Max.   :1.0000   Max.   :1.0000
##  NA's   :2                 NA's   :8        NA's   :5
##    LatTDIe10        EtoTDIe14          TRVmax       Diastolic4count
##  Min.   :0.0000   Min.   :0.0000   Min.   :0.0000   Min.   :0.000
##  1st Qu.:0.0000   1st Qu.:0.0000   1st Qu.:0.0000   1st Qu.:1.000
##  Median :1.0000   Median :0.0000   Median :0.0000   Median :2.000
##  Mean   :0.6353   Mean   :0.3111   Mean   :0.4419   Mean   :1.969
##  3rd Qu.:1.0000   3rd Qu.:1.0000   3rd Qu.:1.0000   3rd Qu.:3.000
##  Max.   :1.0000   Max.   :1.0000   Max.   :1.0000   Max.   :4.000
##  NA's   :11       NA's   :6        NA's   :53
##  DiastDysfunction3of4   DiastCount     DiastDysfunASE
##  Min.   :0.0000       Min.   :0.0000   Min.   :0.0000
##  1st Qu.:0.0000       1st Qu.:0.0000   1st Qu.:0.0000
##  Median :0.0000       Median :0.0000   Median :0.0000
##  Mean   :0.3958       Mean   :0.6771   Mean   :0.4792
##  3rd Qu.:1.0000       3rd Qu.:1.0000   3rd Qu.:1.0000
##  Max.   :1.0000       Max.   :2.0000   Max.   :1.0000
##
```

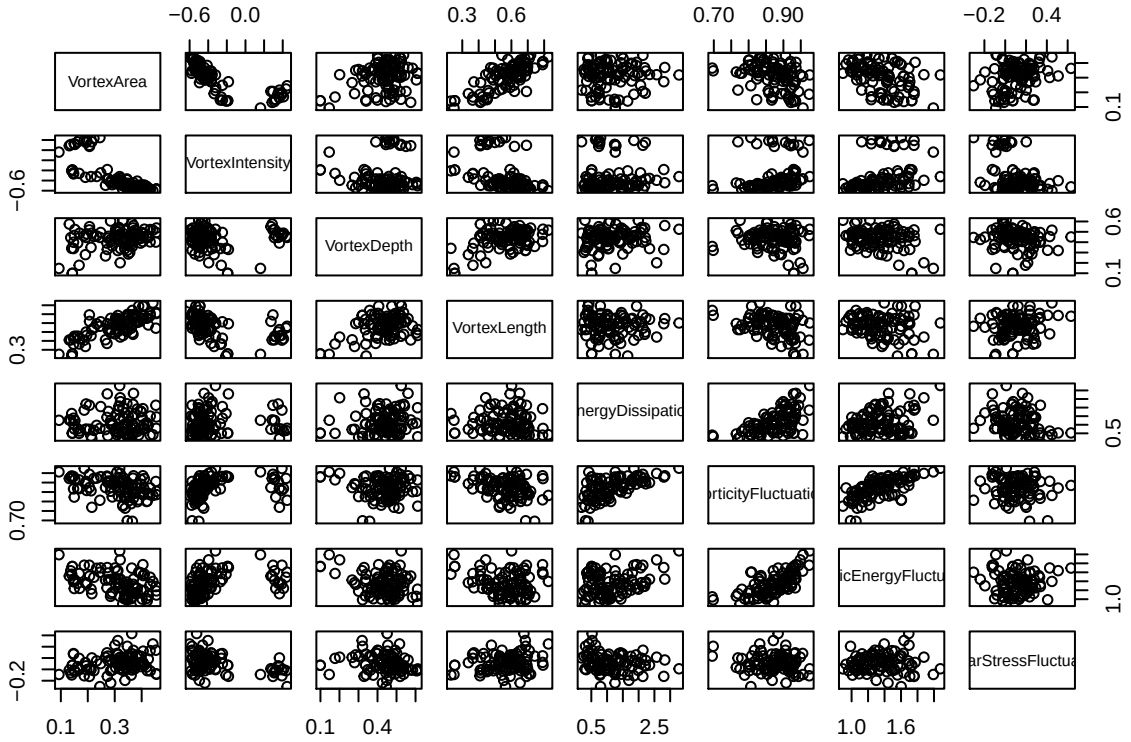
```
str(echo)
```

```
## 'data.frame':    96 obs. of  64 variables:
##  $ ï..ID                   : int  1 4 9 10 15 18 19 28 29 31 ...
##  $ Country                 : Factor w/ 23 levels " ","FS10","FS11",..: 22 22 22 22 22 22 22 22 22 2
##  $ Age                     : int  72 52 55 61 73 61 50 52 77 81 ...
##  $ Gender                  : int  0 0 0 0 0 0 0 0 0 0 ...
##  $ BSA                     : num  1.53 2.09 2.61 1.96 1.65 1.87 1.59 1.54 2 1.74 ...
```

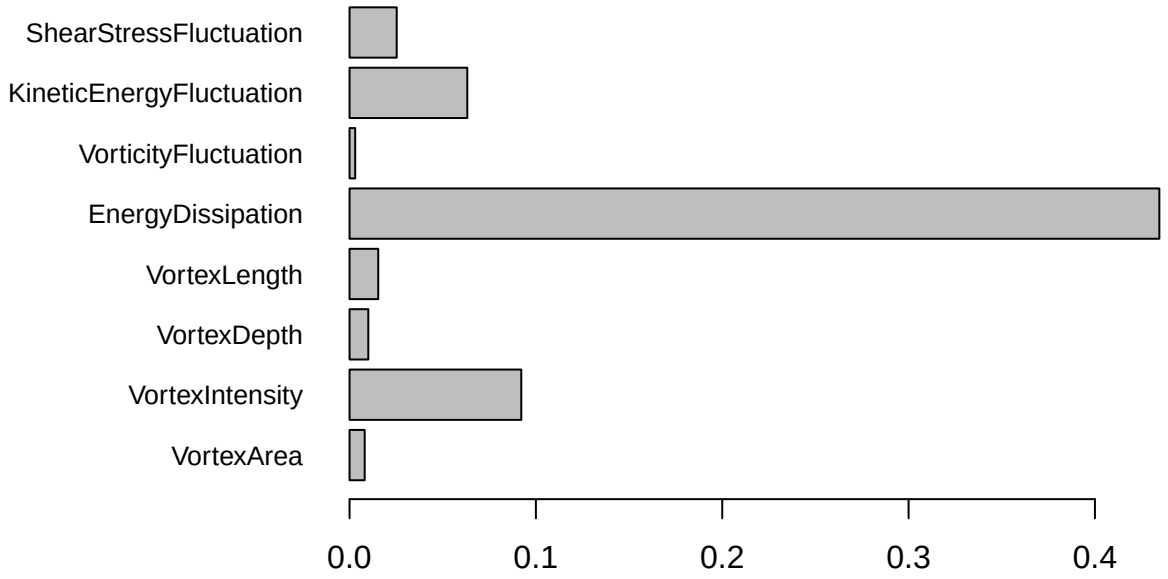
```
##  $ HRmin                     : num  74.3 62.4 74.4 84.9 49 ...
##  $ SBP                       : int  170 140 134 130 170 122 140 110 130 170 ...
##  $ DBP                       : int  90 88 92 80 80 60 88 64 88 64 ...
##  $ Etiology                  : int  NA NA NA NA 4 5 NA 5 1 2 ...
##  $ NYHA                      : int  1 1 1 1 1 1 1 1 1 1 ...
##  $ EchoLA                    : num  NA NA NA NA NA NA NA NA NA NA ...
##  $ EchoIVS                   : num  9 11 10 10 9 11 8 8 10 9 ...
##  $ EchoPW                    : num  9 11 10 11 8 11 8 9 11 9 ...
##  $ EchoDd                    : num  44 46 51 50 58 40 42 39 46 55 ...
##  $ EchoDs                    : num  27 34 34 30 39 25 27 27 34 43 ...
##  $ bEF                       : int  68 64 60 64 61 73 68 57 52 43 ...
##  $ RWT                       : num  NA NA NA NA NA NA NA NA NA NA ...
##  $ LVMass                    : num  128 181 188 194 189 ...
##  $ EchoLVMI                  : num  83.7 86.7 72 99.2 114.7 ...
##  $ EchoEDV                   : num  78 143 116 134 173 109 75 112 153 140 ...
##  $ EchoESV                   : num  22 60 43 53 69 35 26 55 46 74 ...
##  $ EchoSV                    : num  56 83 73 81 104 74 49 57 107 66 ...
##  $ EchoEF                    : num  71 58 63 61 60 68 65 50 70 47 ...
##  $ AbnormalEF50              : int  0 0 0 0 0 0 0 0 0 1 ...
##  $ EchoTMfEwave              : num  61 63 58 76 58 76 73 56 68 60 ...
##  $ EchoTMfAwave              : num  70 51 59 78 43 100 59 50 59 40 ...
##  $ TimeTMFDecT               : int  118 192 229 211 211 189 174 214 185 238 ...
##  $ EchoTMfEtoA               : num  0.871 1.235 0.983 0.974 1.349 ...
##  $ EchoLAVI                  : num  24 27 26 35 41 28 29 29 48 49 ...
##  $ EchSepTDIe                : num  8 9 6 9 6 9 10 9 7 7 ...
##  $ EchoLatTDIe               : num  10 13 10 10 10 10 12 10 7 9 ...
##  $ EchoAveTDIe               : num  9 11 8 9.5 8 9.5 11 9.5 7 8 ...
##  $ EchoSeptEe                : num  7.63 7 9.67 8.44 9.67 8.44 7.3 6.22 9.71 8.57 ...
##  $ EchoRVSP                  : num  27 NA 21 NA 29 32 37 27 42 40 ...
##  $ EchoLVM                   : num  128 181 188 194 189 ...
##  $ EchoLVMII                 : num  83.7 86.7 72 99.2 114.7 ...
##  $ TimePIVRRms               : num  815 966 810 714 1234 ...
##  $ TimePIvEonset             : num  470 465 427 408 474 ...
##  $ TimePIvEonsetNorRR        : num  576 481 527 571 384 ...
##  $ TimePIvEpeak              : num  521 584 546 539 617 ...
##  $ TimePIvEpeakNorRR         : num  639 604 674 755 500 ...
##  $ TimeVSonset               : int  505 501 465 461 504 445 484 599 574 536 ...
##  $ TimePeakVSE               : num  521 584 550 505 617 ...
##  $ Tracking                  : int  1 1 1 1 1 1 1 1 1 1 ...
##  $ VortexArea                : num  0.432 0.337 0.286 0.333 0.455 ...
##  $ VortexIntensity           : num  -0.584 -0.46 -0.402 -0.448 -0.565 ...
##  $ VortexDepth               : num  0.522 0.337 0.406 0.411 0.537 ...
##  $ VortexLength              : num  0.682 0.526 0.523 0.549 0.717 ...
##  $ EnergyDissipation         : num  1.51 1.02 1.01 1.39 1.45 ...
##  $ VorticityFluctuation      : num  0.872 0.825 0.874 0.925 0.854 ...
##  $ KineticEnergyFluctuation : num  1.33 1.07 1.12 1.36 1.19 ...
##  $ ShearStressFluctuation    : num  -0.0267 0.0814 0.1668 0.1603 -0.0858 ...
##  $ KineticEnergyEarlyD       : num  2.16 3.15 2.21 2 3.8 ...
##  $ KineticEnergyLateD        : num  1.77 1.53 3.28 3.43 1.91 ...
##  $ KineticEnergyOnsetSystole: num  0.815 0.525 0.938 0.554 0.868 ...
##  $ LAVI35                    : int  0 0 0 1 1 0 0 0 1 1 ...
##  $ SeptTDIe7                 : int  0 0 1 0 1 0 0 0 0 0 ...
##  $ LatTDIe10                 : int  0 0 0 0 0 0 0 0 1 1 ...
##  $ EtoTDIe14                 : int  0 0 0 0 0 0 0 0 0 0 ...
```

```
##  $ TRVmax                  : int  0 NA 0 NA 0 0 1 0 1 1 ...
##  $ Diastolic4count         : int  0 0 1 1 2 0 1 0 3 3 ...
##  $ DiastDysfunction3of4    : int  0 0 0 0 0 0 0 0 1 1 ...
##  $ DiastCount              : int  0 0 0 0 0 0 0 0 1 2 ...
##  $ DiastDysfunASE          : int  0 0 0 0 0 0 0 0 1 1 ...
```

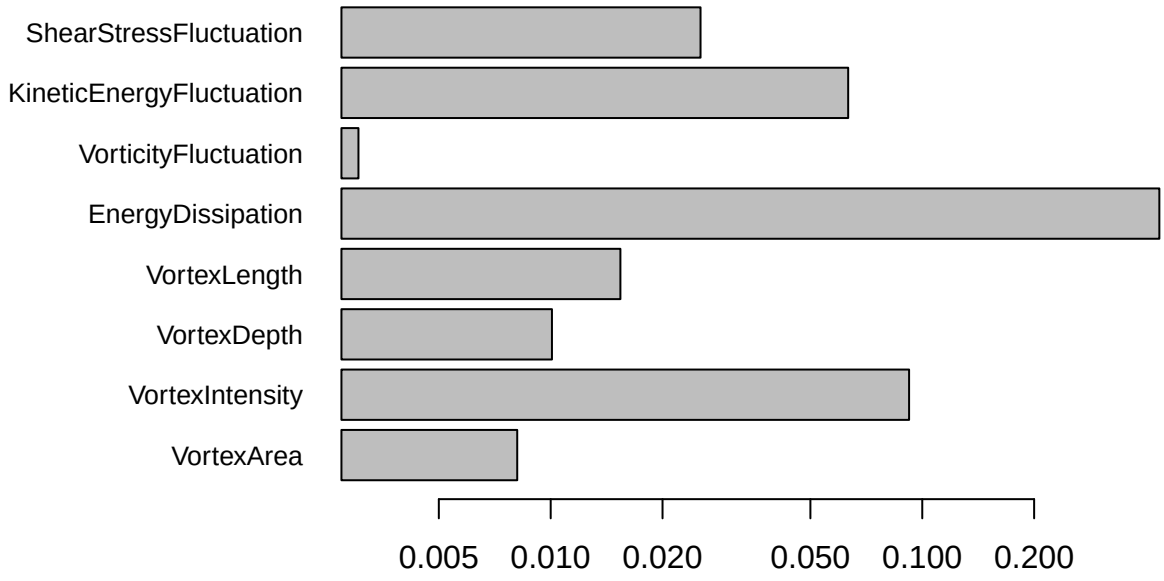
```
## There are several missing data. We only select the ones we care about - the vortex data.
vorticity.data <- echo[,c(45:52)]


#
plot(vorticity.data)
```



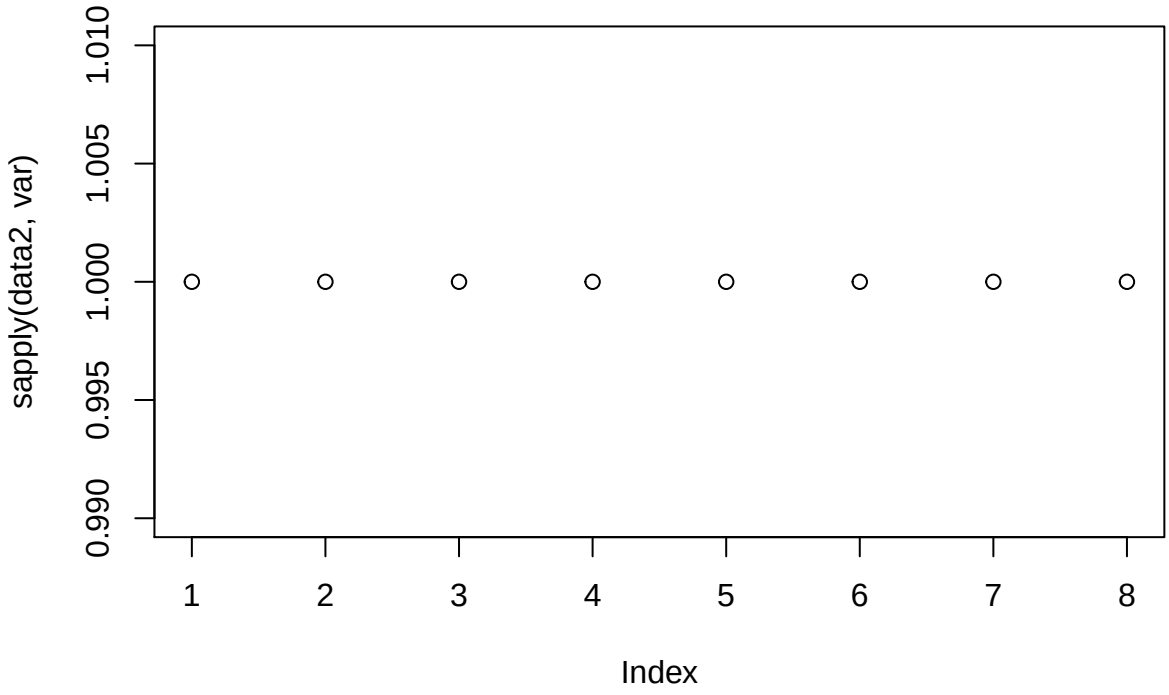

```
mar <- par()$mar
par(mar=mar+c(0,5,0,0))

# seems like the variables are in different scales
barplot(sapply(vorticity.data, var), horiz=T, las=1, cex.names=0.8)
```




```
# in log scale
barplot(sapply(vorticity.data, var), horiz=T, las=1, cex.names=0.8, log='x')
```

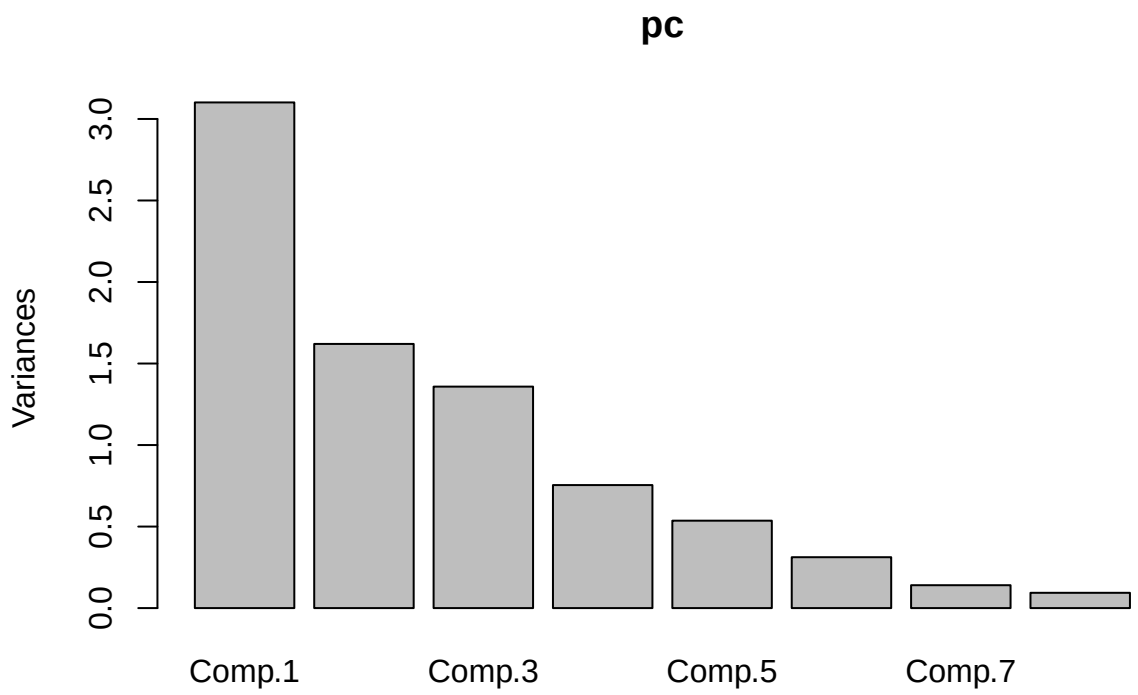



```r
par(mar=mar)



# Scale
data2 <- data.frame(scale(vorticity.data))
# Verify variance is uniform
plot(sapply(data2, var))
```

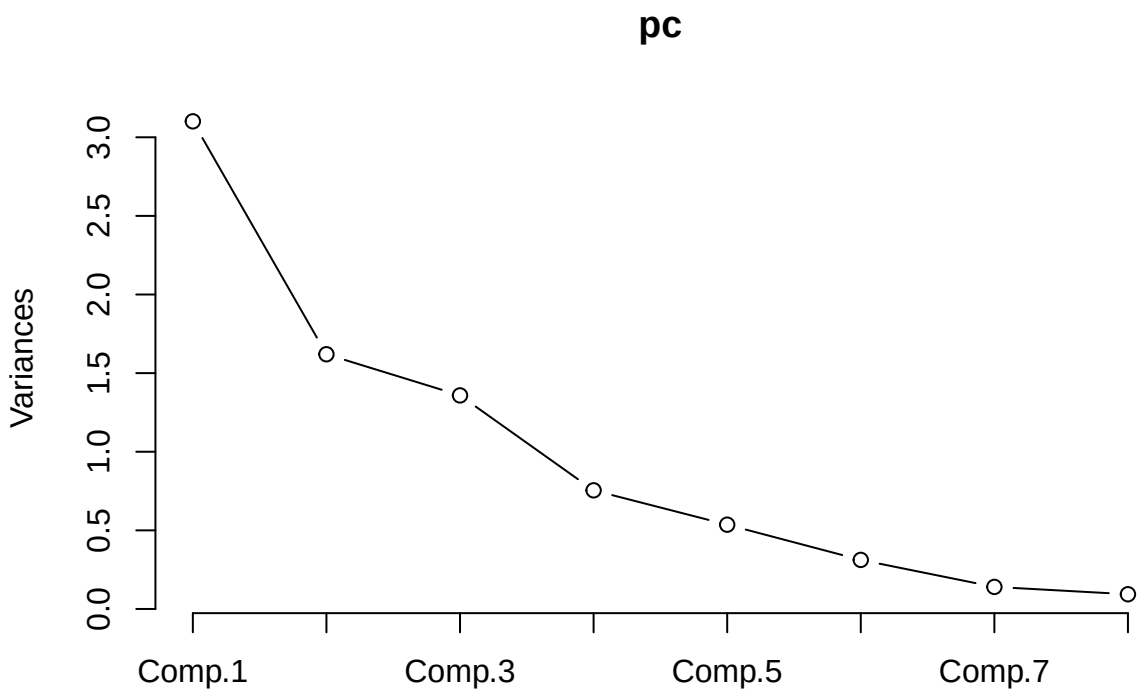


Calculating PCA

```
# Proceed with principal components
pc <- princomp(data2)
plot(pc)
```



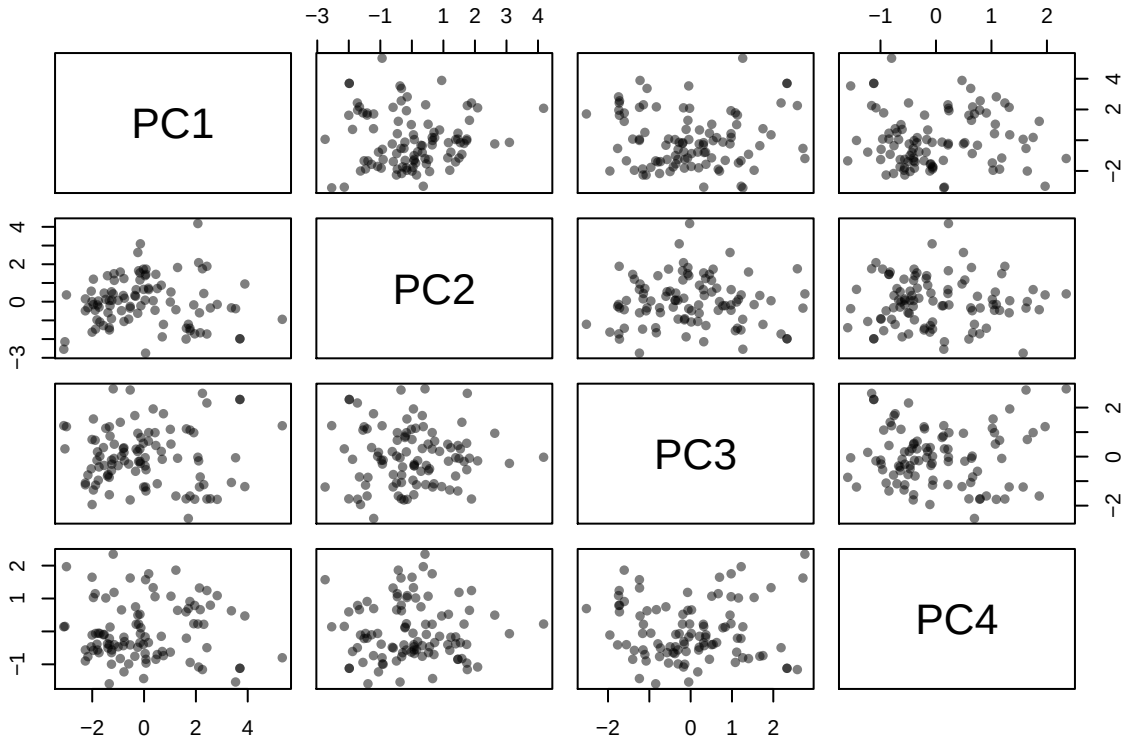

```
plot(pc, type='l')
```




```
summary(pc) # 4 components is both 'elbow' and explains >85% variance
```

```
## Importance of components:
##                           Comp.1    Comp.2    Comp.3     Comp.4     Comp.5
## Standard deviation     1.7610829 1.2728542 1.1654740 0.86863697 0.73218492
## Proportion of Variance 0.3917574 0.2046515 0.1715785 0.09530908 0.06771723
## Cumulative Proportion  0.3917574 0.5964090 0.7679874 0.86329652 0.93101375
##                            Comp.6     Comp.7     Comp.8
## Standard deviation     0.55864939 0.37462455 0.30611837
## Proportion of Variance 0.03942179 0.01772761 0.01183686
## Cumulative Proportion  0.97043554 0.98816314 1.00000000
```

```
# Get principal component vectors using prcomp instead of princomp
pc <- prcomp(data2)

# First for principal components
comp <- data.frame(pc$x[,1:4])
# Plot
plot(comp, pch=16, col=rgb(0,0,0,0.5))
```




```
library(rgl)
# Multi 3D plot
plot3d(comp$PC1, comp$PC2, comp$PC3)
plot3d(comp$PC1, comp$PC3, comp$PC4)
```

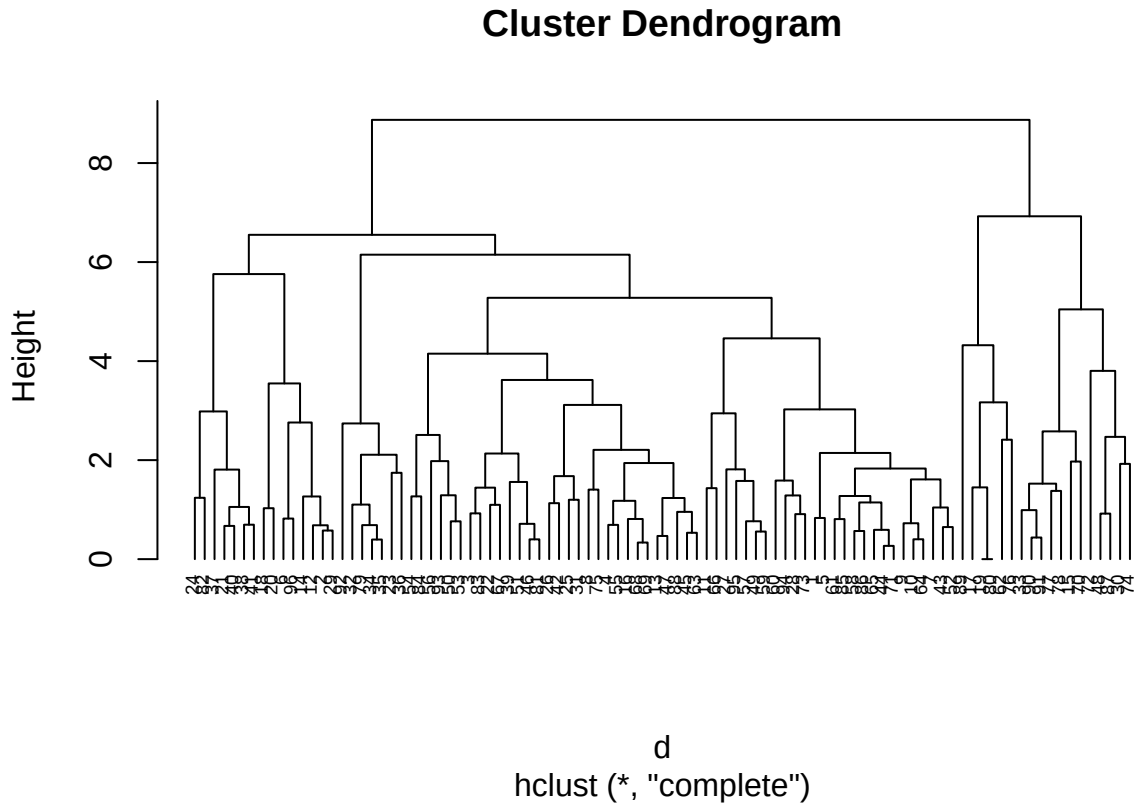
Hierarchical Clustering packages

Load packages

```
suppressPackageStartupMessages(library(tidyverse))
suppressPackageStartupMessages(library(cluster))
suppressPackageStartupMessages(library(factoextra))
suppressPackageStartupMessages(library(dendextend))

library(tidyverse)  # data manipulation
library(cluster)    # clustering algorithms
library(factoextra) # clustering visualization
library(dendextend) # for comparing two dendrograms
```

```
# Dissimilarity matrix
d <- dist(comp, method = "euclidean")

# Hierarchical clustering using Complete Linkage
hc1 <- hclust(d, method = "complete" )

# Plot the obtained dendrogram
```

```r
plot(hc1, cex = 0.6, hang = -1)
```



For Hierarchical Clustering

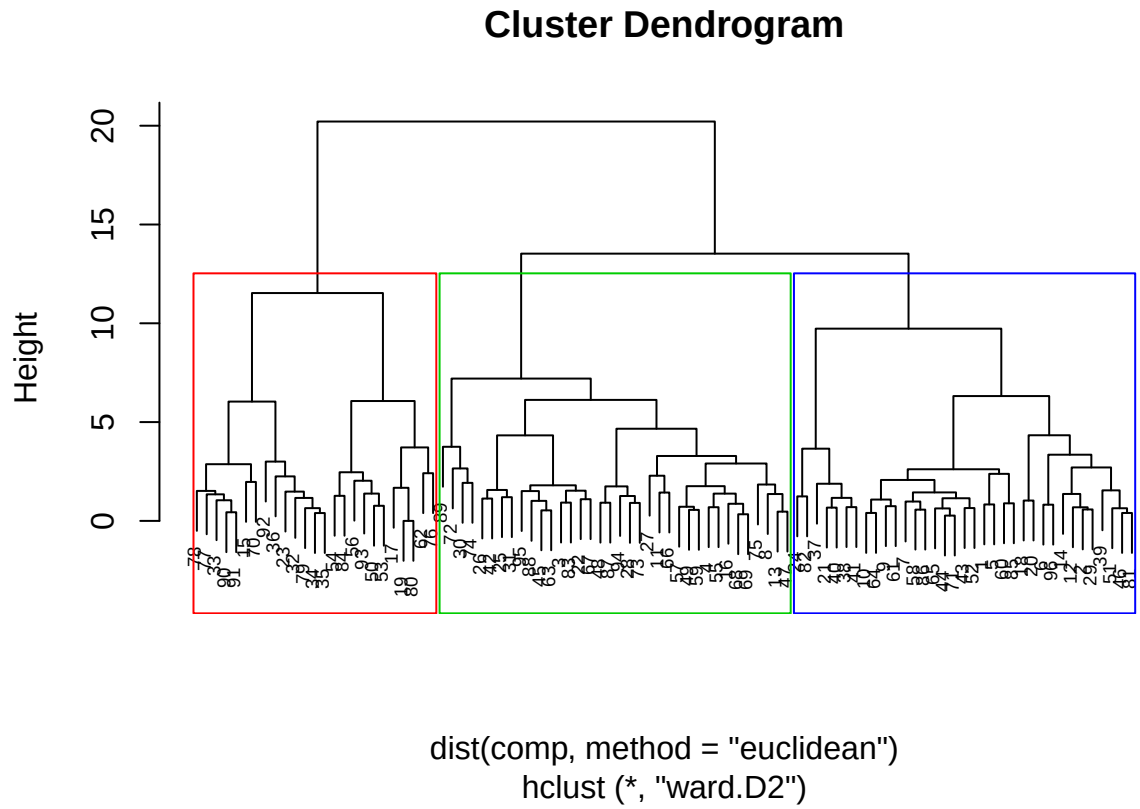
usig Agnes to check the cluster structure

```
##   average    single  complete      ward
## 0.7656333 0.6595636 0.8810995 0.9480419
```

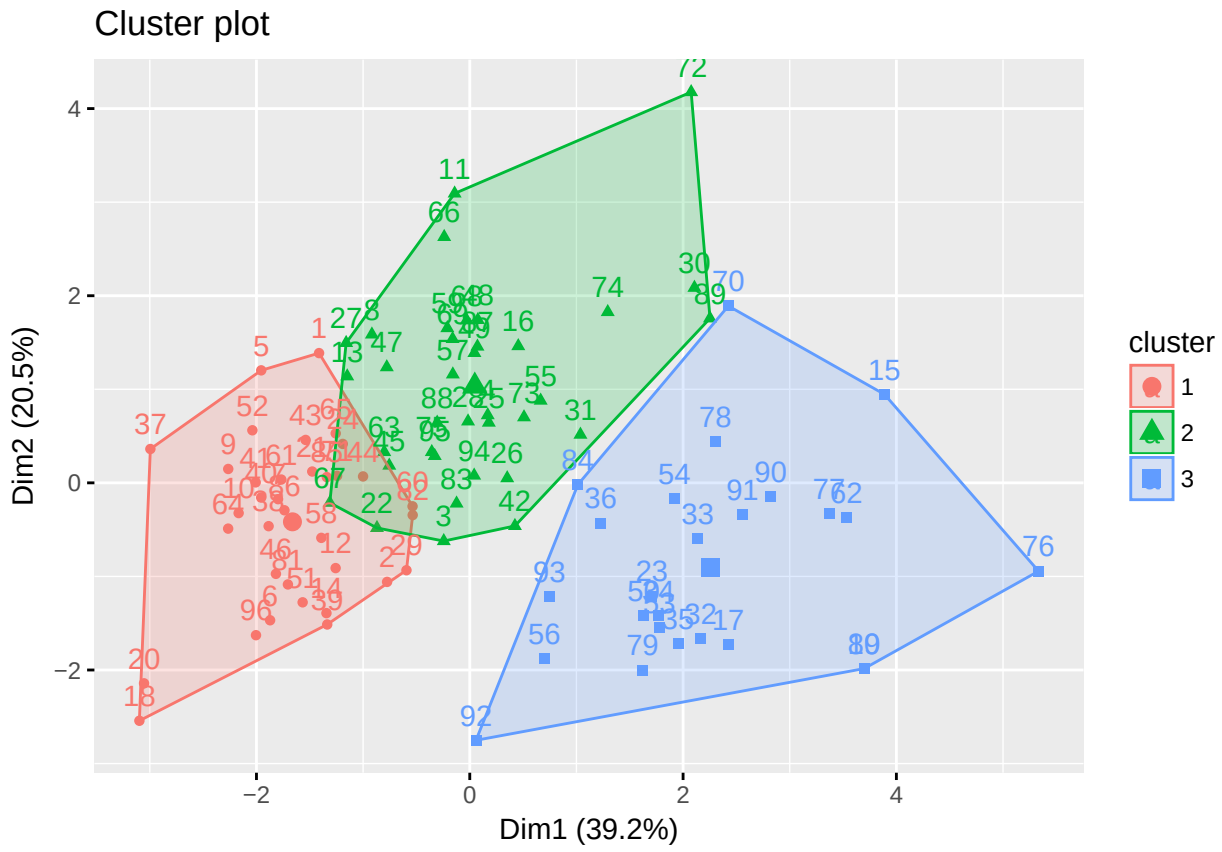
we select Ward's method since it's the strongest structure

```r
library(dplyr)
require(graphics)

# Ward's method
hc5 <- hclust(dist(comp, method = "euclidean"), method = "ward.D2" )

# Cut tree into 4 groups
sub_grp <- cutree(hc5, k = 3)

echo.cluster <- echo %>%
mutate(cluster = sub_grp)
```

```
plot(hc5, cex=0.6)
rect.hclust(hc5, k = 3, border = 2:5)
```

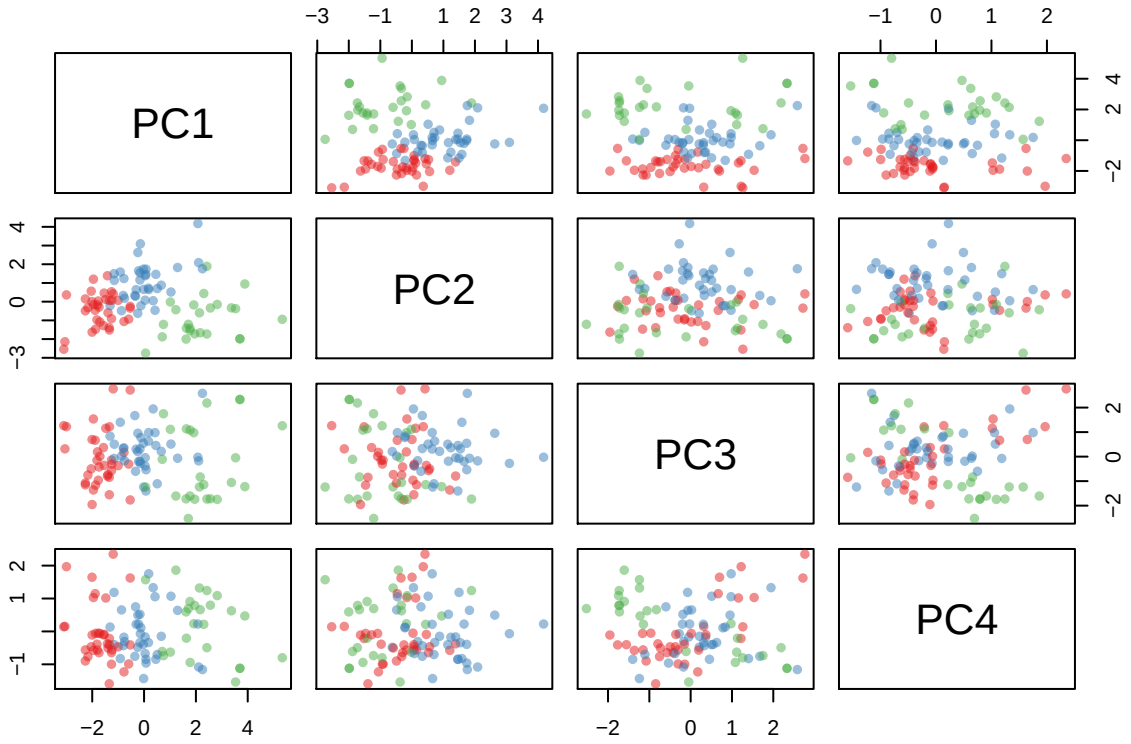



```
fviz_cluster(list(data = vorticity.data, cluster = sub_grp))
```



PCA plot with cluster

```
library(RColorBrewer)

palette(alpha(brewer.pal(9, 'Set1'), 0.5))
plot(comp, col=sub_grp, pch=16)
```




```
# 3D plot
plot3d(comp$PC1, comp$PC2, comp$PC3, col=sub_grp)
plot3d(comp$PC1, comp$PC3, comp$PC4, col=sub_grp)
```

```
# Cluster sizes
sort(table(sub_grp))
```

```
## sub_grp
##  3  1  2
## 25 35 36
```

```
clust <- names(sort(table(sub_grp)))

# First cluster
row.names(echo.cluster[sub_grp==clust[1],])
```

```
##  [1] "15" "17" "19" "23" "32" "33" "34" "35" "36" "50" "53" "54" "56" "62"
## [15] "70" "76" "77" "78" "79" "80" "84" "90" "91" "92" "93"
```

```
# Second Cluster
row.names(echo.cluster[sub_grp==clust[2],])
```

```
##  [1] "1"  "2"  "5"  "6"  "7"  "9"  "10" "12" "14" "18" "20" "21" "24" "29"
## [15] "37" "38" "39" "40" "41" "43" "44" "46" "51" "52" "58" "60" "61" "64"
## [29] "65" "71" "81" "82" "85" "86" "96"
```

```
# Third Cluster
row.names(echo.cluster[sub_grp==clust[3],])
```

```
##  [1] "3"  "4"  "8"  "11" "13" "16" "22" "25" "26" "27" "28" "30" "31" "42"
## [15] "45" "47" "48" "49" "55" "57" "59" "63" "66" "67" "68" "69" "72" "73"
## [29] "74" "75" "83" "87" "88" "89" "94" "95"
```

Analysis of demographics with cluster

Age vs Cluster

```
# Age vs Cluster
aov.age <- aov(echo.cluster$Age ~ echo.cluster$cluster)
summary(aov.age)
```

```
##                       Df Sum Sq Mean Sq F value Pr(>F)
## echo.cluster$cluster   1     12   12.36    0.07  0.792
## Residuals             90  15862  176.24
## 4 observations deleted due to missingness
```

```
#age range
echo.cluster$age.range <- cut(echo.cluster$Age, breaks=c(0,40, 60, 80, 200), right = FALSE, labels = c(


# Age range by cluster
# 1 : 0-39
# 2:  40 - 59
# 3:  60 - 79
# 4:  80 +

age.table <- table(echo.cluster$cluster, echo.cluster$age.range)

chisq.test(age.table)
```

```
## Warning in chisq.test(age.table): Chi-squared approximation may be
## incorrect
```

```
##
##  Pearson's Chi-squared test
##
## data:  age.table
## X-squared = 5.4256, df = 6, p-value = 0.4905
```

NYHA vs Cluster

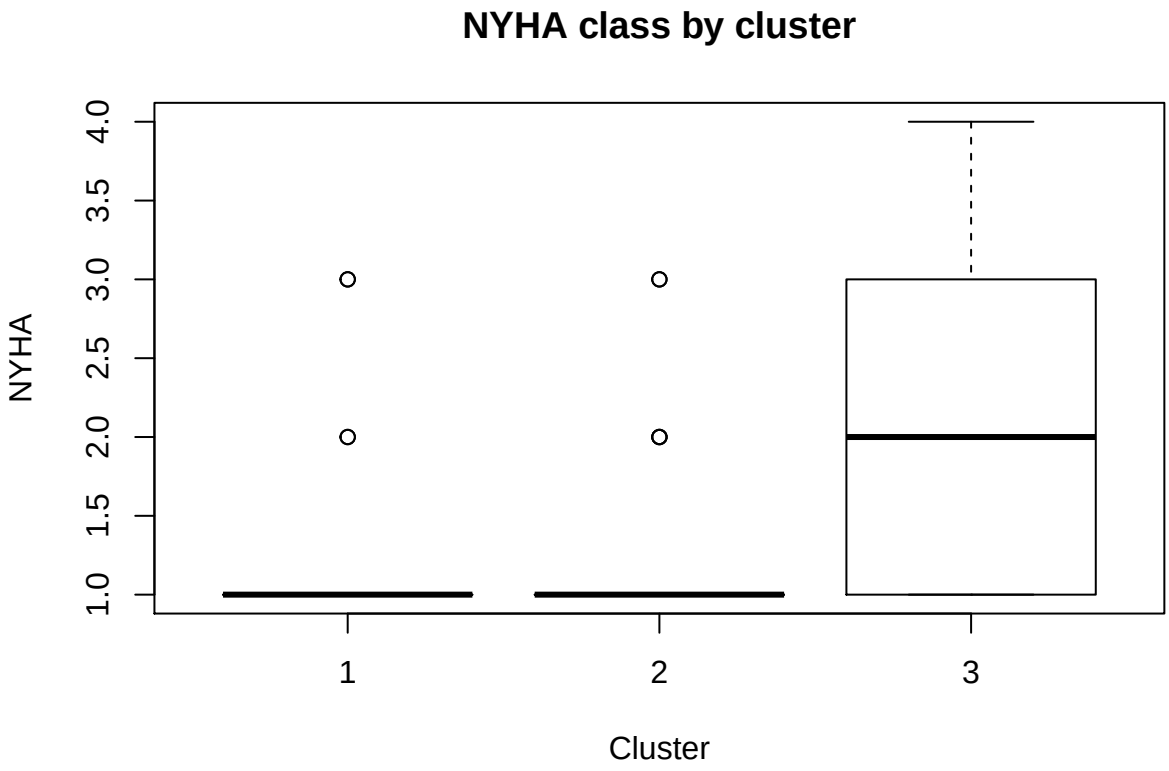
```
nyha.tbl <- table(echo.cluster$cluster, echo.cluster$NYHA)

chisq.test(nyha.tbl)
```

```
## Warning in chisq.test(nyha.tbl): Chi-squared approximation may be incorrect
```

```
##
##  Pearson's Chi-squared test
##
## data:  nyha.tbl
```

```
## X-squared = 22.544, df = 6, p-value = 0.0009645
```

```
boxplot(echo.cluster$NYHA ~ echo.cluster$cluster, xlab = "Cluster", ylab = "NYHA", main = "NYHA class by
```



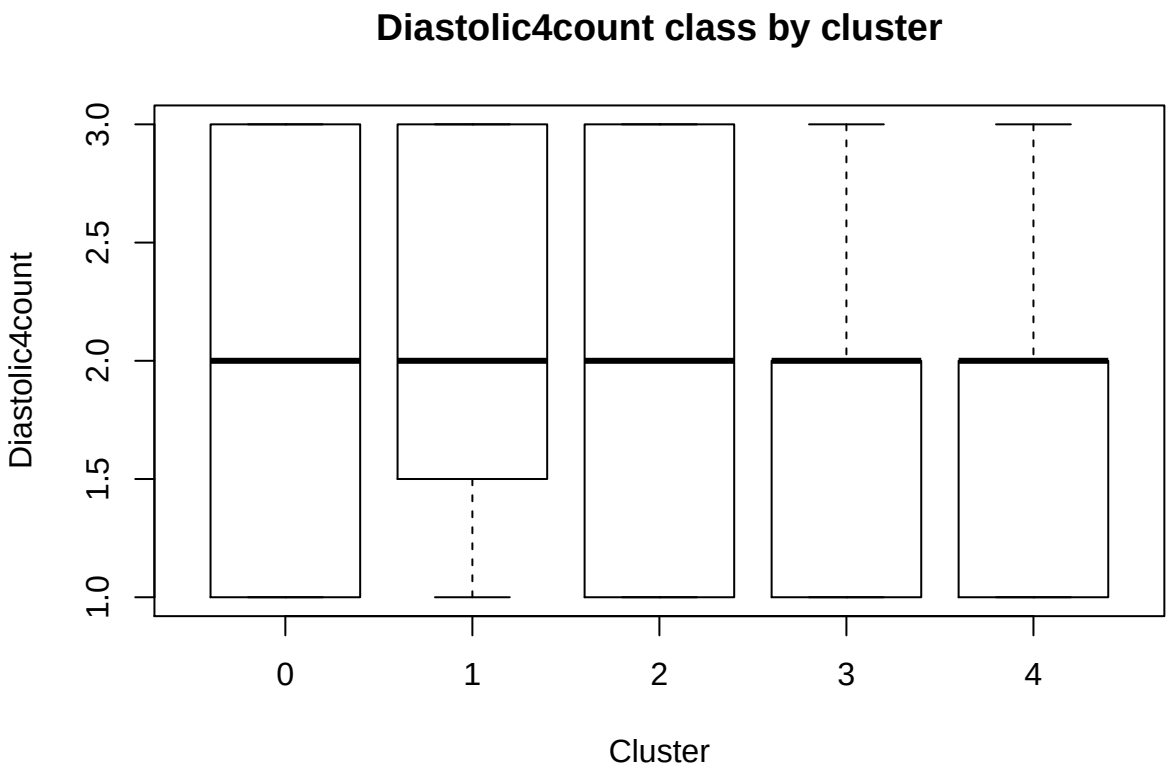
Diastolic 4 Parameter Count

```
dias.tbl <- table(echo.cluster$cluster, echo.cluster$Diastolic4count)

chisq.test(dias.tbl)
```

```
## Warning in chisq.test(dias.tbl): Chi-squared approximation may be incorrect
```

```
##
##  Pearson's Chi-squared test
##
## data:  dias.tbl
## X-squared = 5.8019, df = 8, p-value = 0.6694
```

```
boxplot(echo.cluster$cluster ~ echo.cluster$Diastolic4count, xlab = "Cluster", ylab = "Diastolic4count"
```

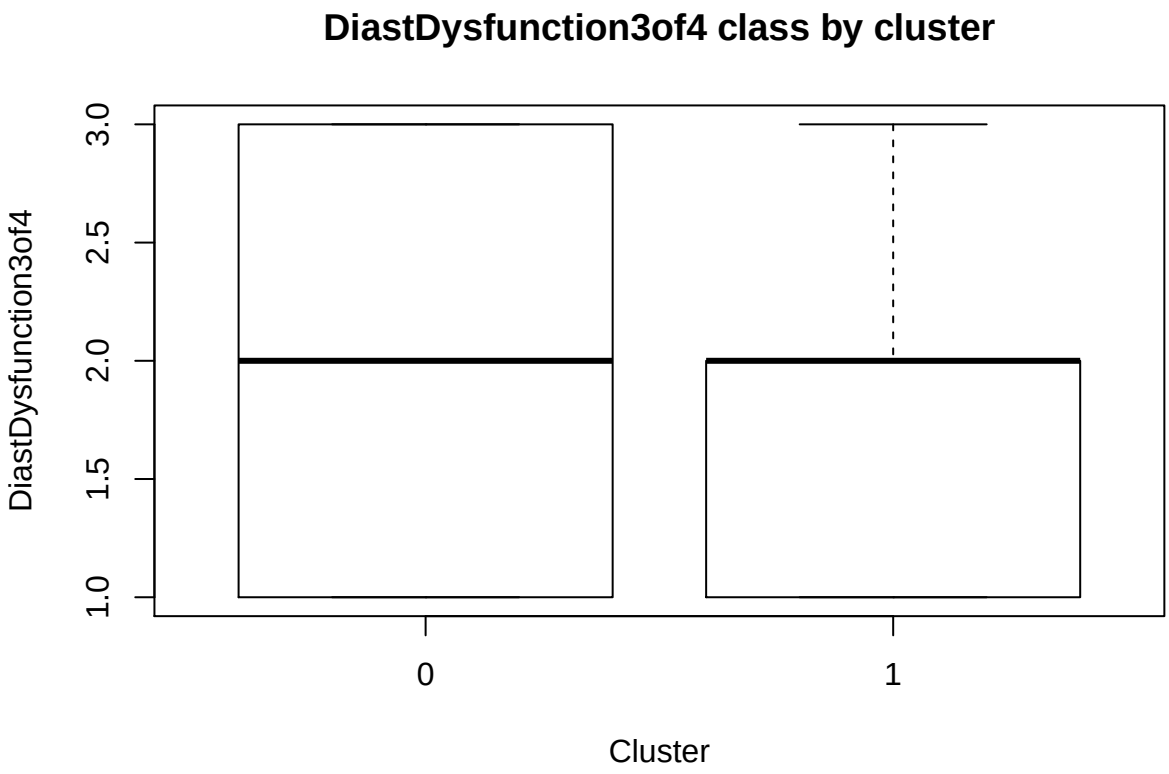


Diastolic Dysfunction 3 or 4 of 4

```
dias.dysfn.tbl <- table(echo.cluster$cluster, echo.cluster$DiastDysfunction3of4)

chisq.test(dias.dysfn.tbl)
```

```
##
##  Pearson's Chi-squared test
##
## data:  dias.dysfn.tbl
## X-squared = 1.9152, df = 2, p-value = 0.3838
```

```
boxplot(echo.cluster$cluster ~ echo.cluster$DiastDysfunction3of4, xlab = "Cluster", ylab = "DiastDysfun
```



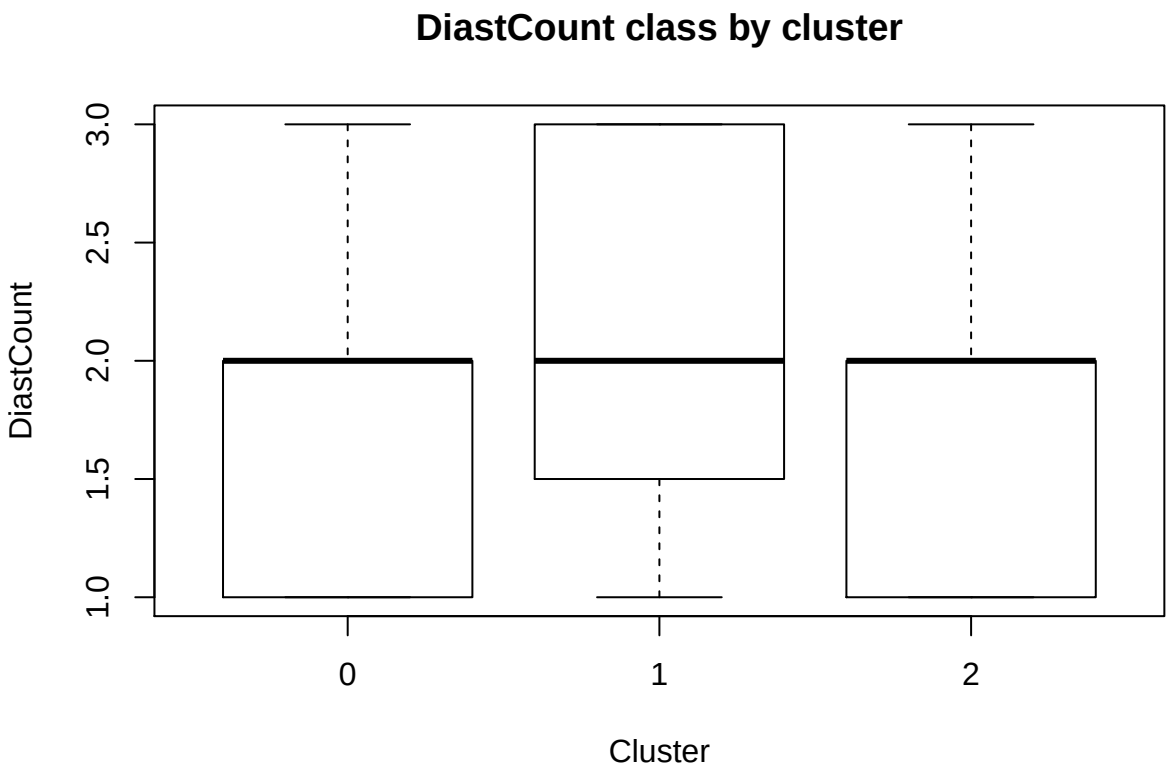
Diastolic Count (low EF or 3/4)

```
low.ef.tbl <- table(echo.cluster$cluster, echo.cluster$DiastCount)

chisq.test(low.ef.tbl)
```

```
## Warning in chisq.test(low.ef.tbl): Chi-squared approximation may be
## incorrect
```

```
##
##  Pearson's Chi-squared test
##
## data:  low.ef.tbl
## X-squared = 4.4928, df = 4, p-value = 0.3434
```

```
boxplot(echo.cluster$cluster ~ echo.cluster$DiastCount, xlab = "Cluster", ylab = "DiastCount", main = "
```

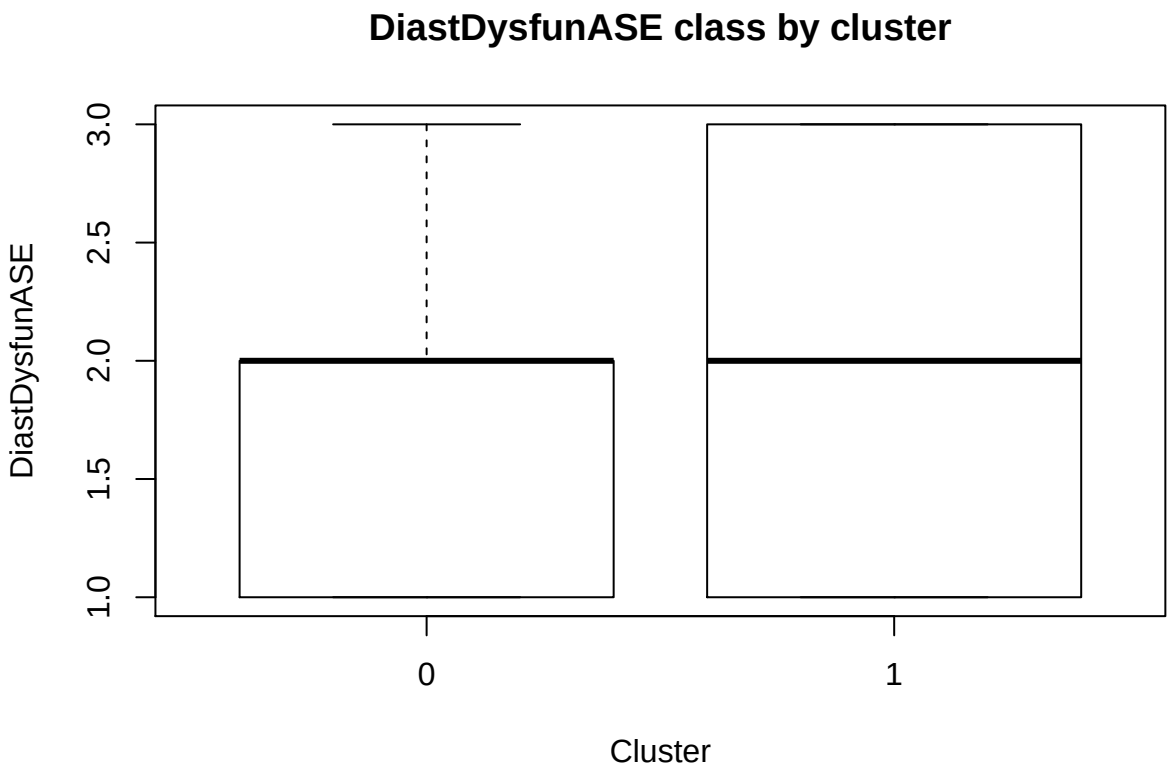


Diastolic Dysfunction ASE (low EF or 3/4)

```
dias.dysfn.ase.tbl <- table(echo.cluster$cluster, echo.cluster$DiastDysfunASE)

chisq.test(dias.dysfn.ase.tbl)
```

```
##
##  Pearson's Chi-squared test
##
## data:  dias.dysfn.ase.tbl
## X-squared = 1.9955, df = 2, p-value = 0.3687
```

```
boxplot(echo.cluster$cluster ~ echo.cluster$DiastDysfunASE, xlab = "Cluster", ylab = "DiastDysfunASE", m
```



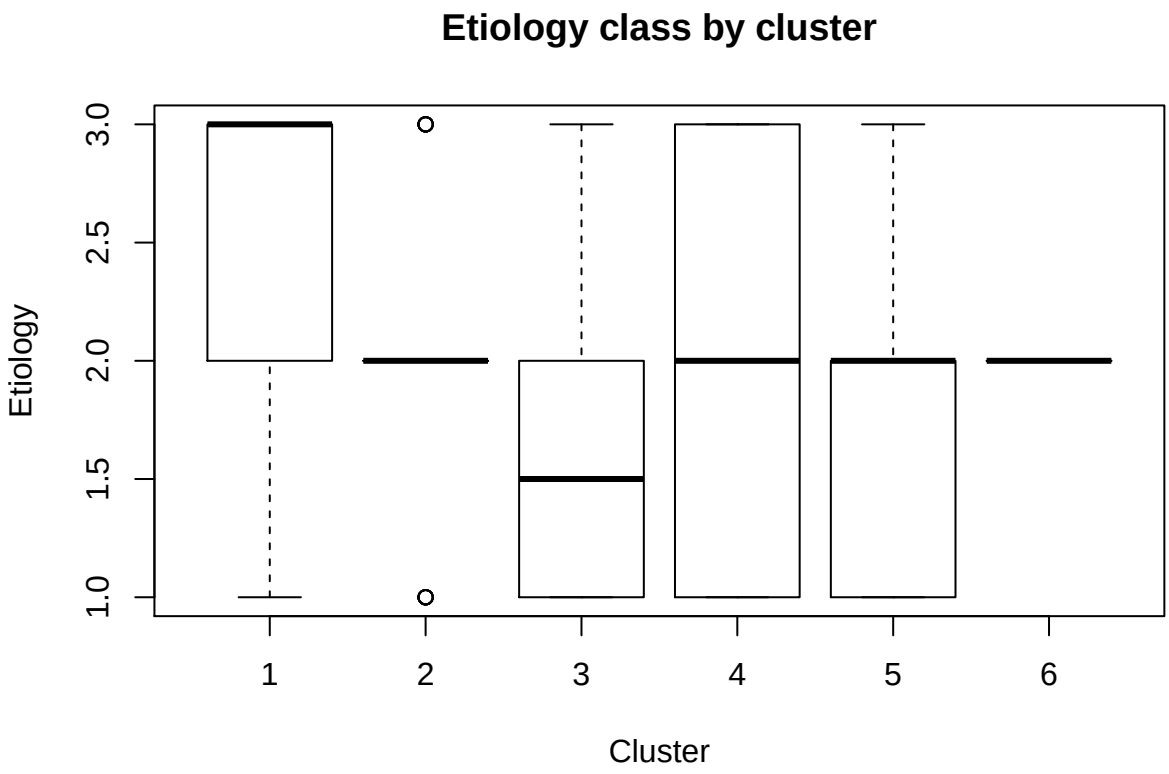
Etiology

```
etiology.tbl <- table(echo.cluster$cluster, echo.cluster$Etiology)

chisq.test(etiology.tbl)
```

```
## Warning in chisq.test(etiology.tbl): Chi-squared approximation may be
## incorrect
```

```
##
##  Pearson's Chi-squared test
##
## data:  etiology.tbl
## X-squared = 13.359, df = 10, p-value = 0.2043
```

```
boxplot(echo.cluster$cluster ~ echo.cluster$Etiology, xlab = "Cluster", ylab = "Etiology", main = "Etiol
```

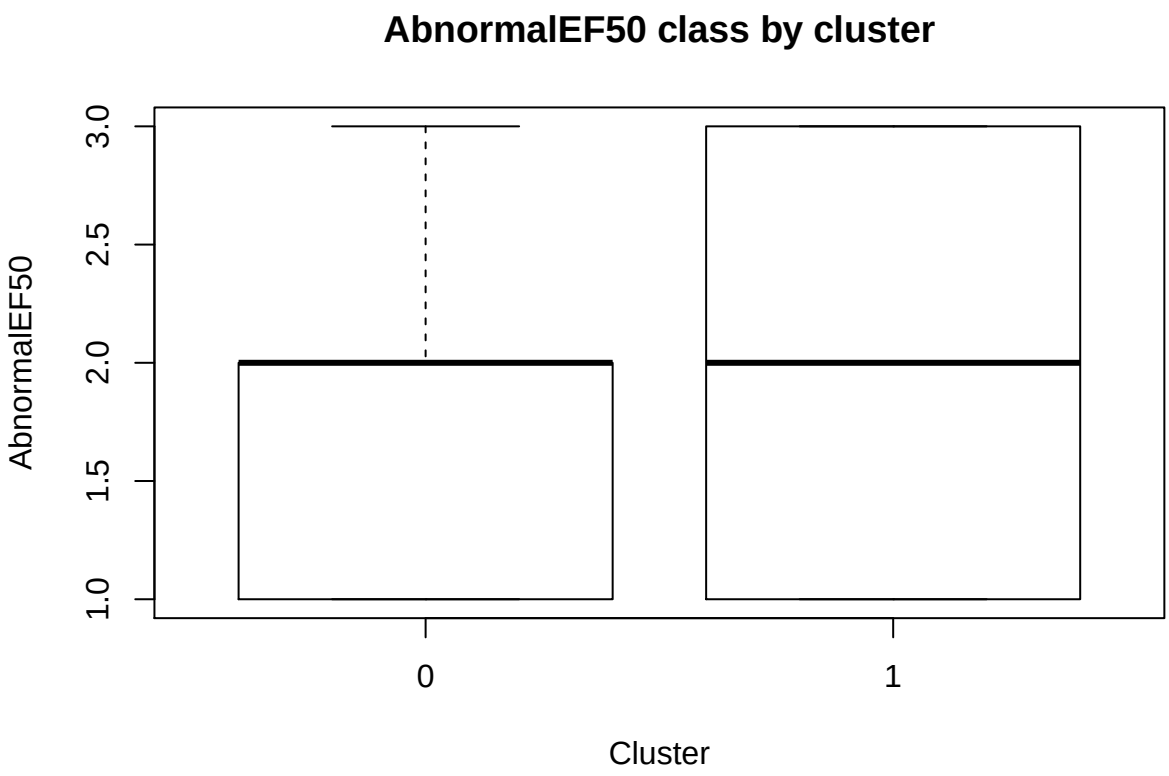


LV Ejection fraction below 40%

```
lv.ejection.tbl <- table(echo.cluster$cluster, echo.cluster$AbnormalEF50)

chisq.test(lv.ejection.tbl)
```

```
##
##  Pearson's Chi-squared test
##
## data:  lv.ejection.tbl
## X-squared = 7.452, df = 2, p-value = 0.02409
```

```
boxplot(echo.cluster$cluster ~ echo.cluster$AbnormalEF50, xlab = "Cluster", ylab = "AbnormalEF50", main
```


AbnormalEF50 class by cluster
AbnormalEF50
1.0
1.5
2.0
2.5
3.0
0
1
Cluster