Week11 R Code - Anly 506-51

Nitin

July 24, 2019

Principal Component Analysis in R

```
mtcars.pca <- prcomp(mtcars[,c(1:7,10,11)], center = TRUE,scale. = TRUE)

summary(mtcars.pca)
```

```
## Importance of components:
##                           PC1    PC2     PC3     PC4     PC5     PC6
## Standard deviation     2.3782 1.4429 0.71008 0.51481 0.42797 0.35184
## Proportion of Variance 0.6284 0.2313 0.05602 0.02945 0.02035 0.01375
## Cumulative Proportion  0.6284 0.8598 0.91581 0.94525 0.96560 0.97936
##                            PC7    PC8     PC9
## Standard deviation     0.32413 0.2419 0.14896
## Proportion of Variance 0.01167 0.0065 0.00247
## Cumulative Proportion  0.99103 0.9975 1.00000
```

```
str(mtcars.pca)
```

```
## List of 5
##  $ sdev    : num [1:9] 2.378 1.443 0.71 0.515 0.428 ...
##  $ rotation: num [1:9, 1:9] -0.393 0.403 0.397 0.367 -0.312 ...
##   ..- attr(*, "dimnames")=List of 2
##   .. ..$ : chr [1:9] "mpg" "cyl" "disp" "hp" ...
##   .. ..$ : chr [1:9] "PC1" "PC2" "PC3" "PC4" ...
##  $ center  : Named num [1:9] 20.09 6.19 230.72 146.69 3.6 ...
##   ..- attr(*, "names")= chr [1:9] "mpg" "cyl" "disp" "hp" ...
##  $ scale   : Named num [1:9] 6.027 1.786 123.939 68.563 0.535 ...
##   ..- attr(*, "names")= chr [1:9] "mpg" "cyl" "disp" "hp" ...
##  $ x       : num [1:32, 1:9] -0.664 -0.637 -2.3 -0.215 1.587 ...
##   ..- attr(*, "dimnames")=List of 2
##   .. ..$ : chr [1:32] "Mazda RX4" "Mazda RX4 Wag" "Datsun 710" "Hornet 4 Drive" ...
##   .. ..$ : chr [1:9] "PC1" "PC2" "PC3" "PC4" ...
##  - attr(*, "class")= chr "prcomp"
```

```
library(devtools)
```

```
## Warning: package 'devtools' was built under R version 3.5.2
```

```
## Warning: package 'usethis' was built under R version 3.5.2
```

```
install_github("vqv/ggbiplot")
```

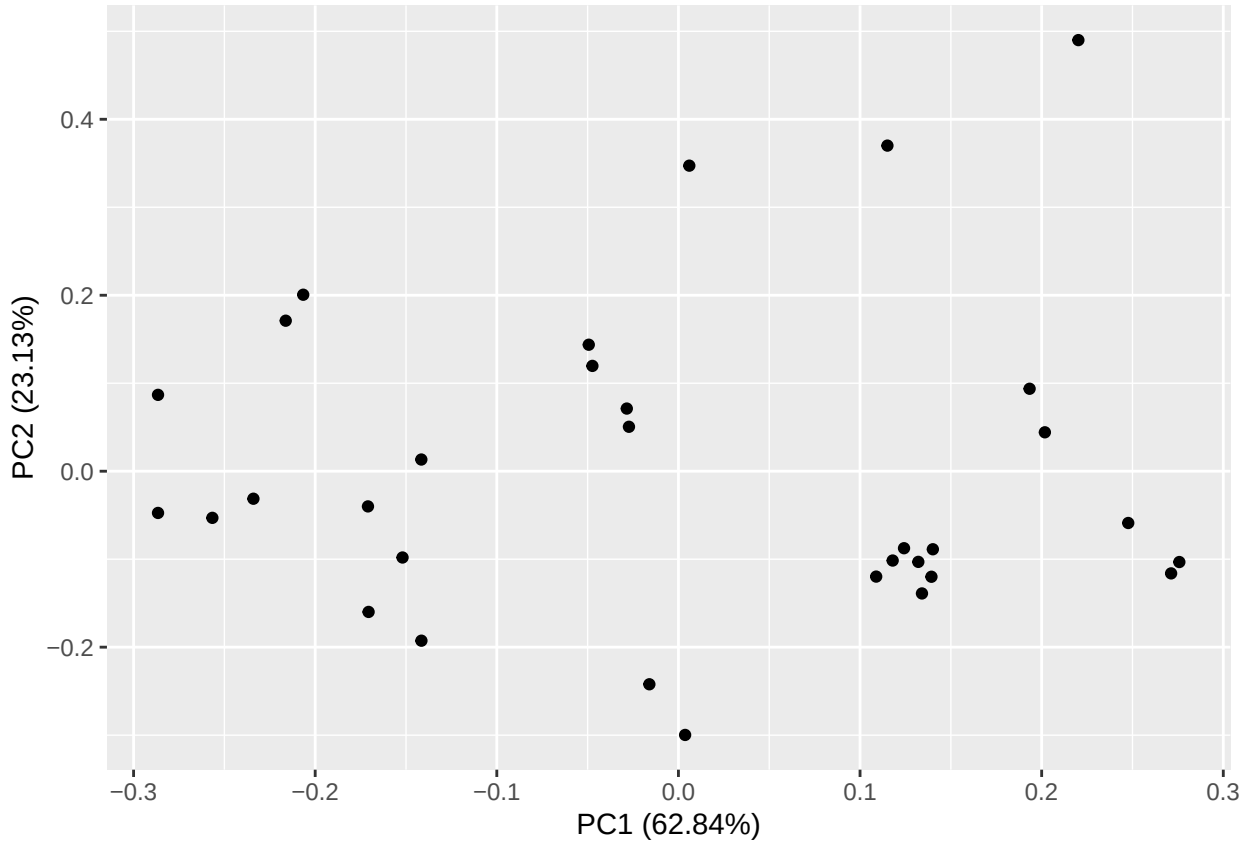
```
## Skipping install of 'ggbiplot' from a github remote, the SHA1 (7325e880) has not changed since last i
##   Use `force = TRUE` to force installation
```

Plotting PCA (Principal Component Analysis)

```
library(ggfortify)
```

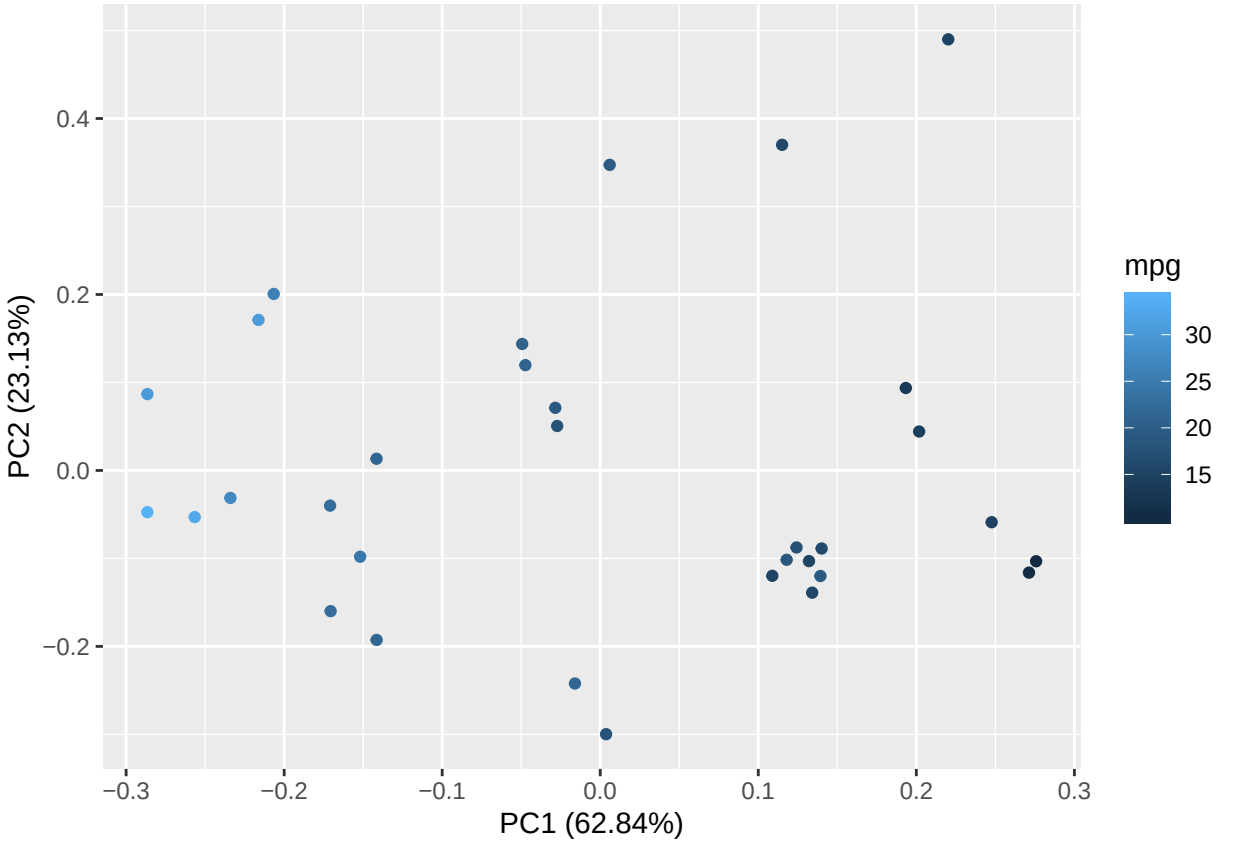
```
## Warning: package 'ggfortify' was built under R version 3.5.3
## Loading required package: ggplot2
## Warning: package 'ggplot2' was built under R version 3.5.3
```

```
df3<-prcomp(mtcars[,c(1:7,10,11)], center = TRUE, scale. = TRUE)
autoplot(df3)
```



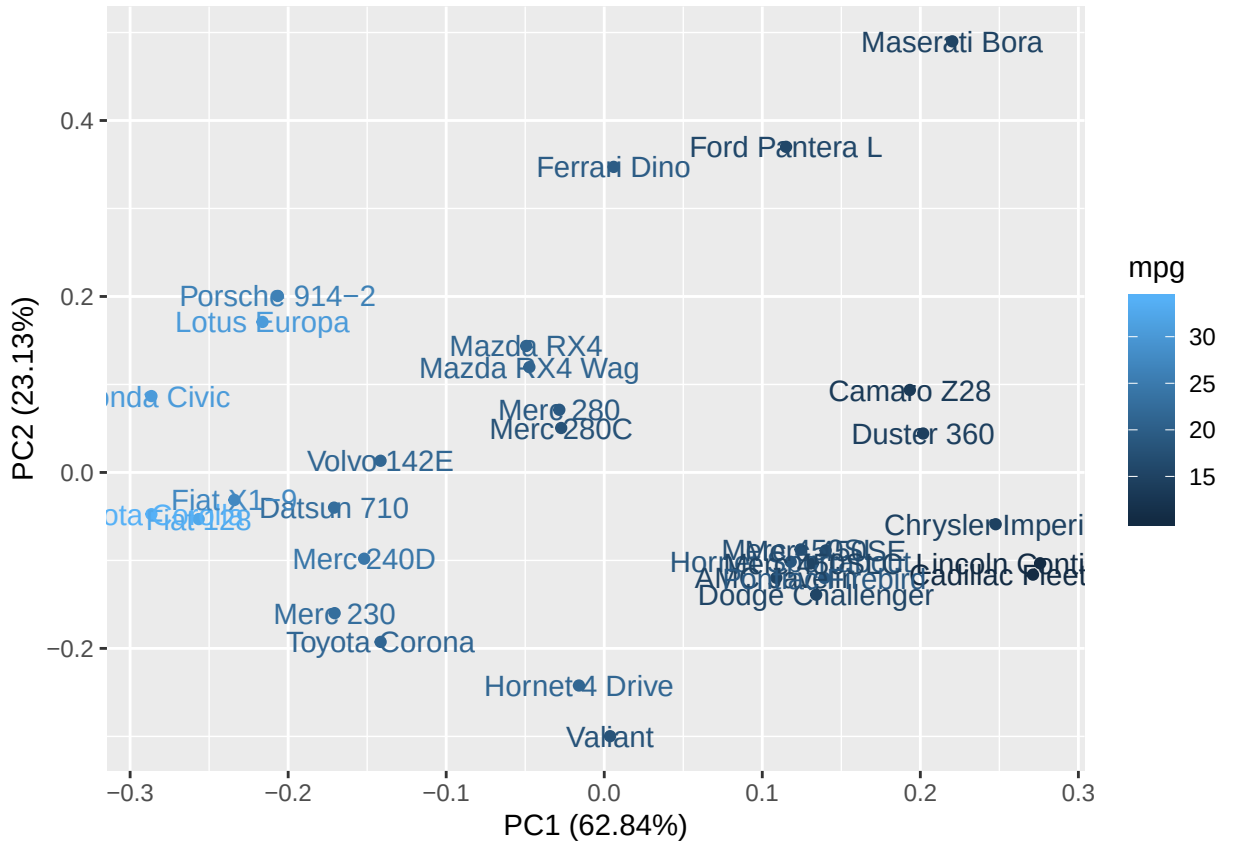
Colorize by non-numeric values

```
autoplot(df3, data = mtcars, colour = 'mpg')
```



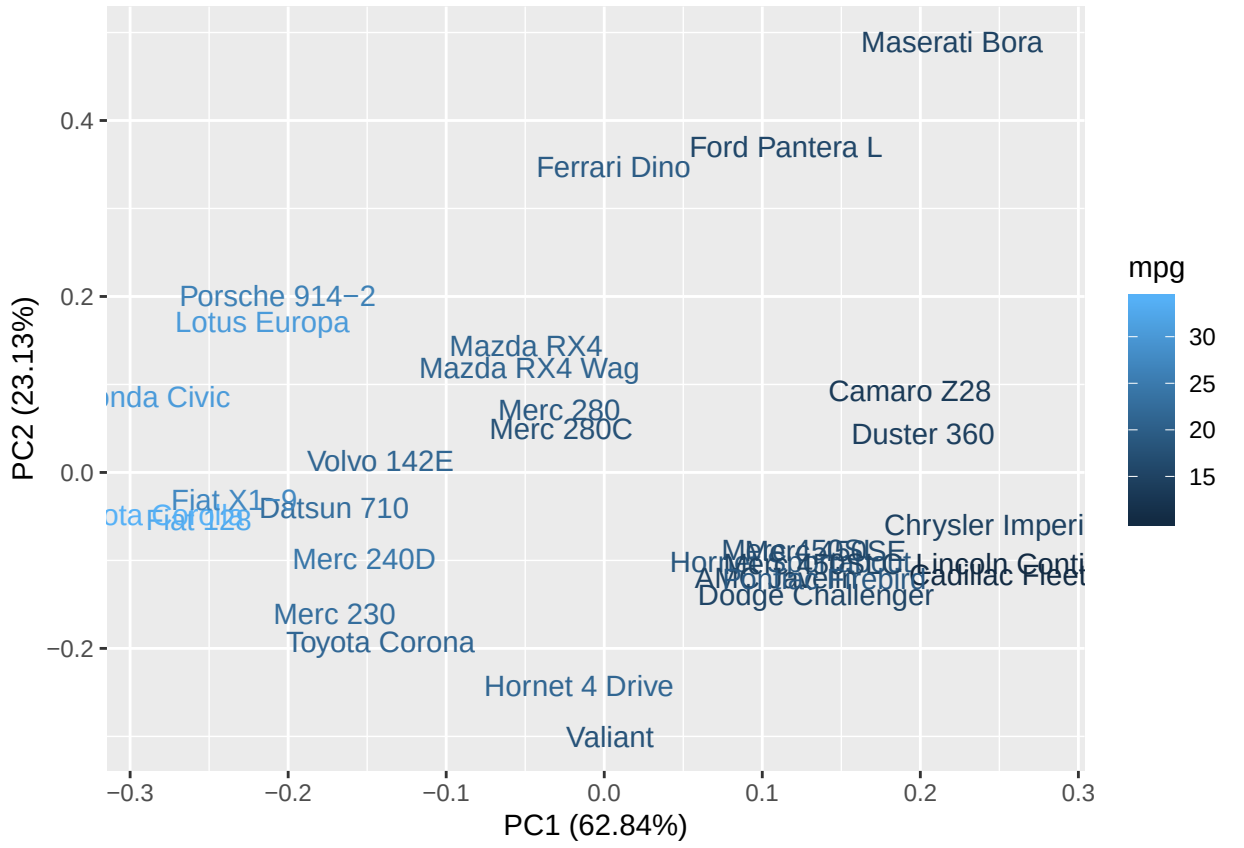
Passing label = TRUE draws each data label using rownames

```
autoplot(df3, data = mtcars, colour = 'mpg', label = TRUE, label.size = 4)
```



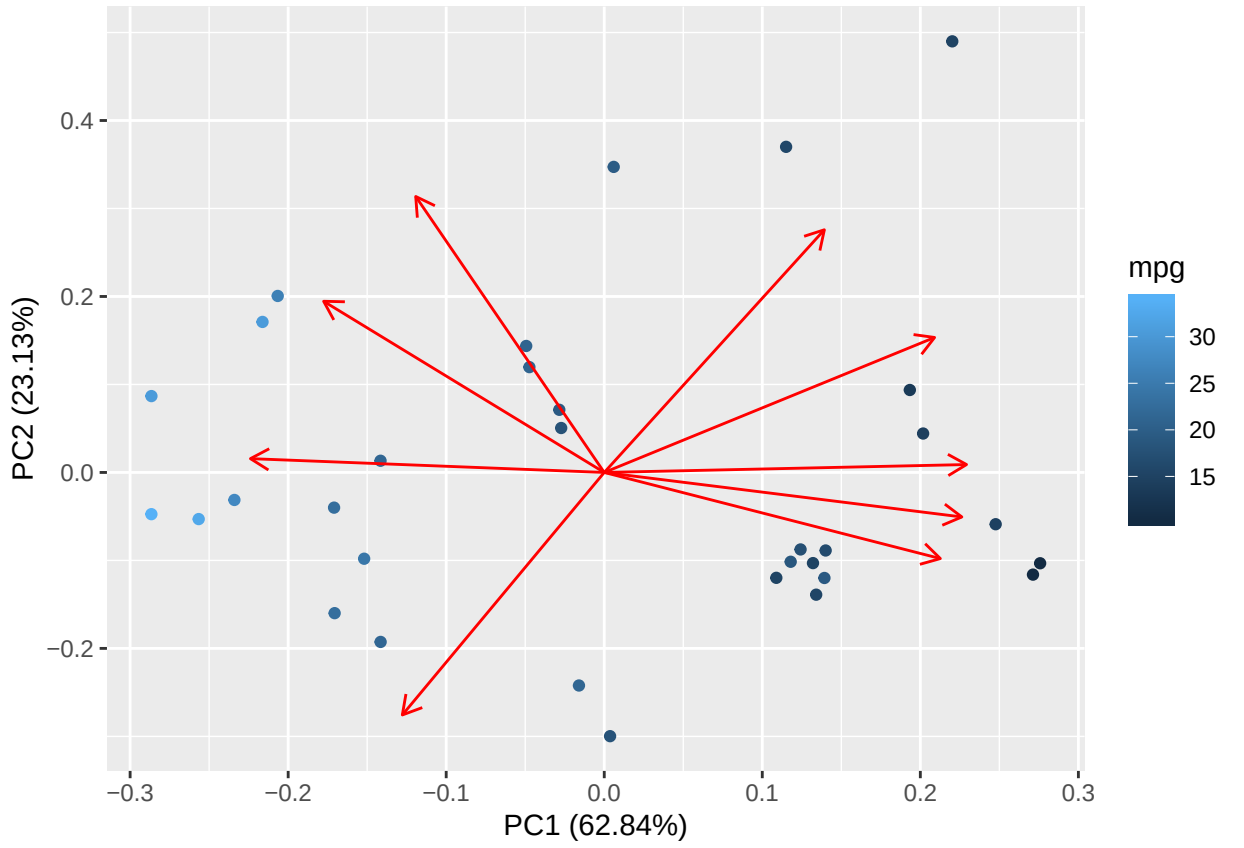
Passing shape = FALSE makes plot without points. In this case, label is turned on unless otherwise specified.

```
autoplot(df3, data = mtcars, colour = 'mpg', shape = FALSE, label.size = 4)
```



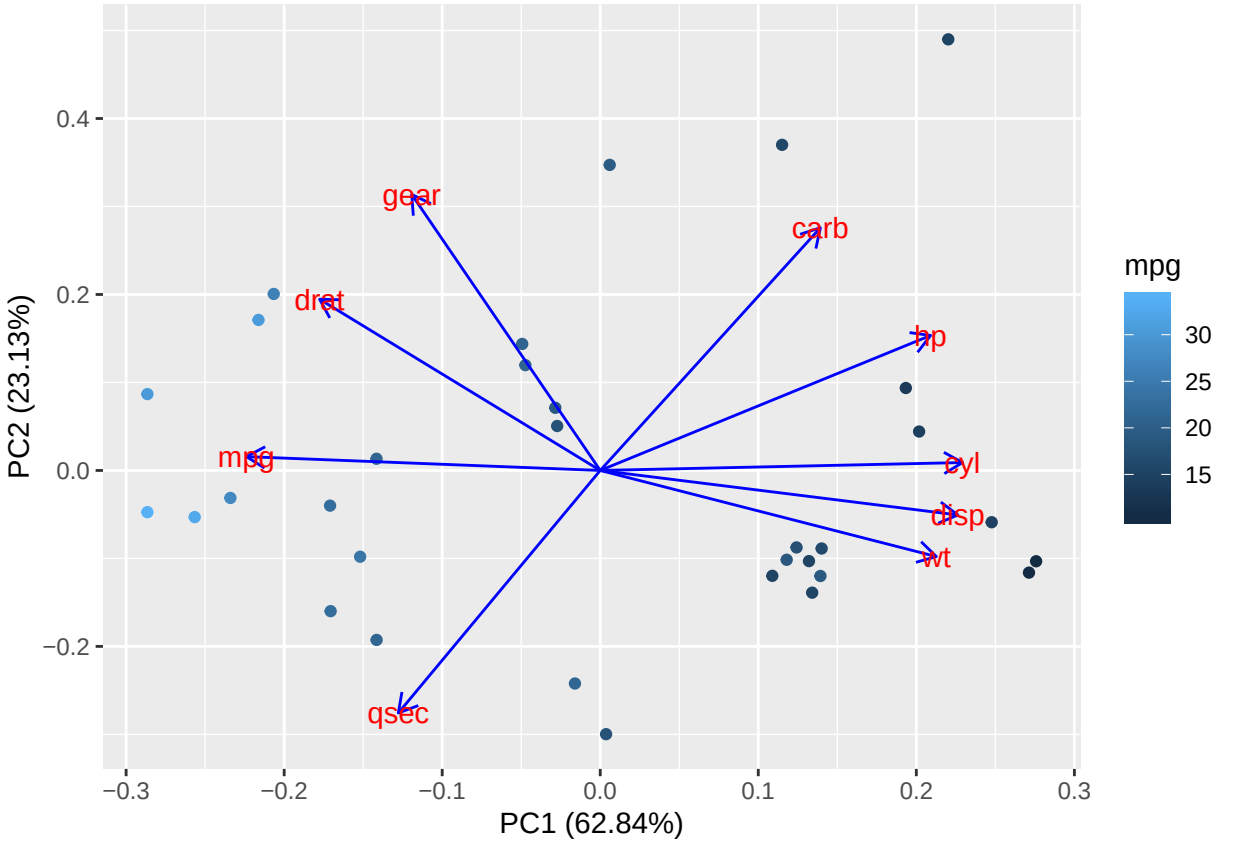
Drawing eigenvectors.

```
autoplot(df3, data = mtcars, colour = 'mpg', loadings = TRUE)
```



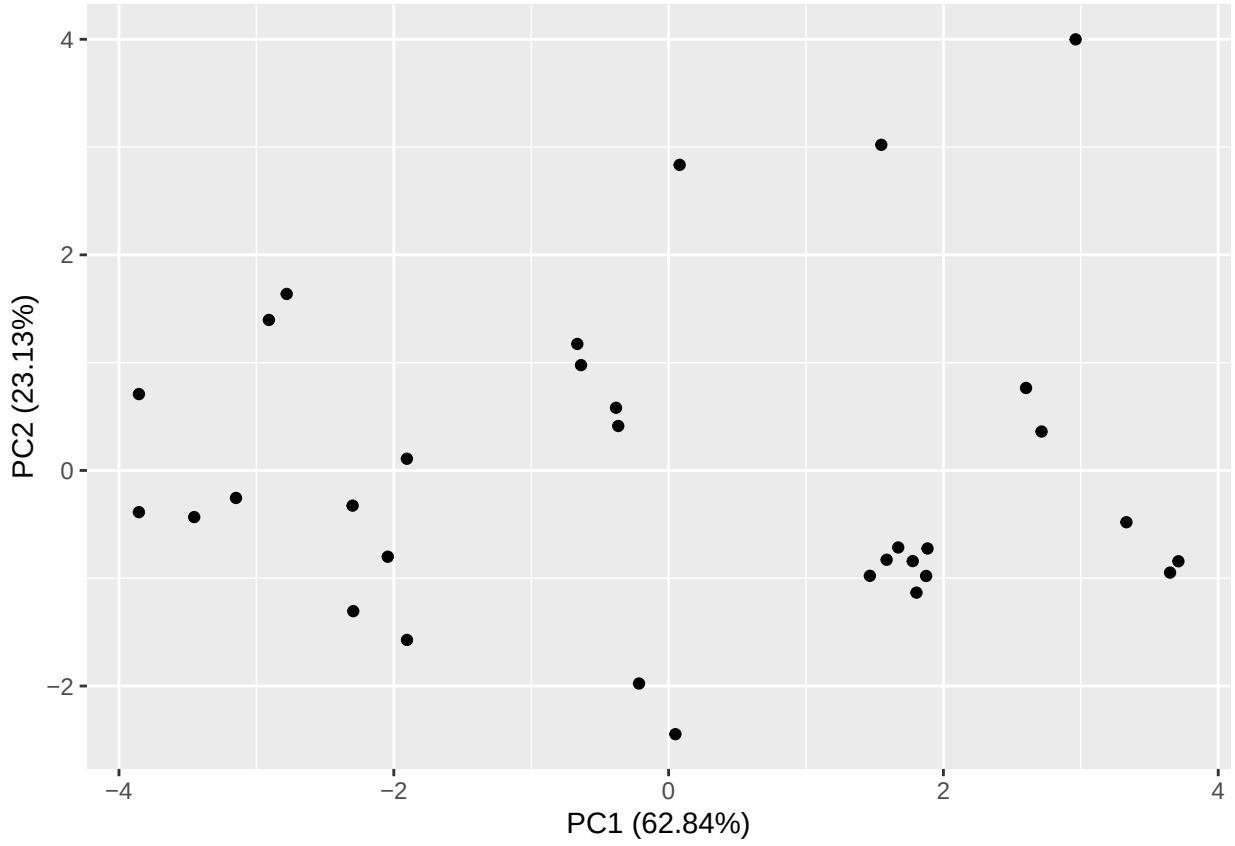
Attaching lables to eigenvector and changing some options.

```
autoplot(df3, data = mtcars, colour = 'mpg',
         loadings = TRUE, loadings.colour = 'blue',
         loadings.label = TRUE, loadings.label.size = 4)
```



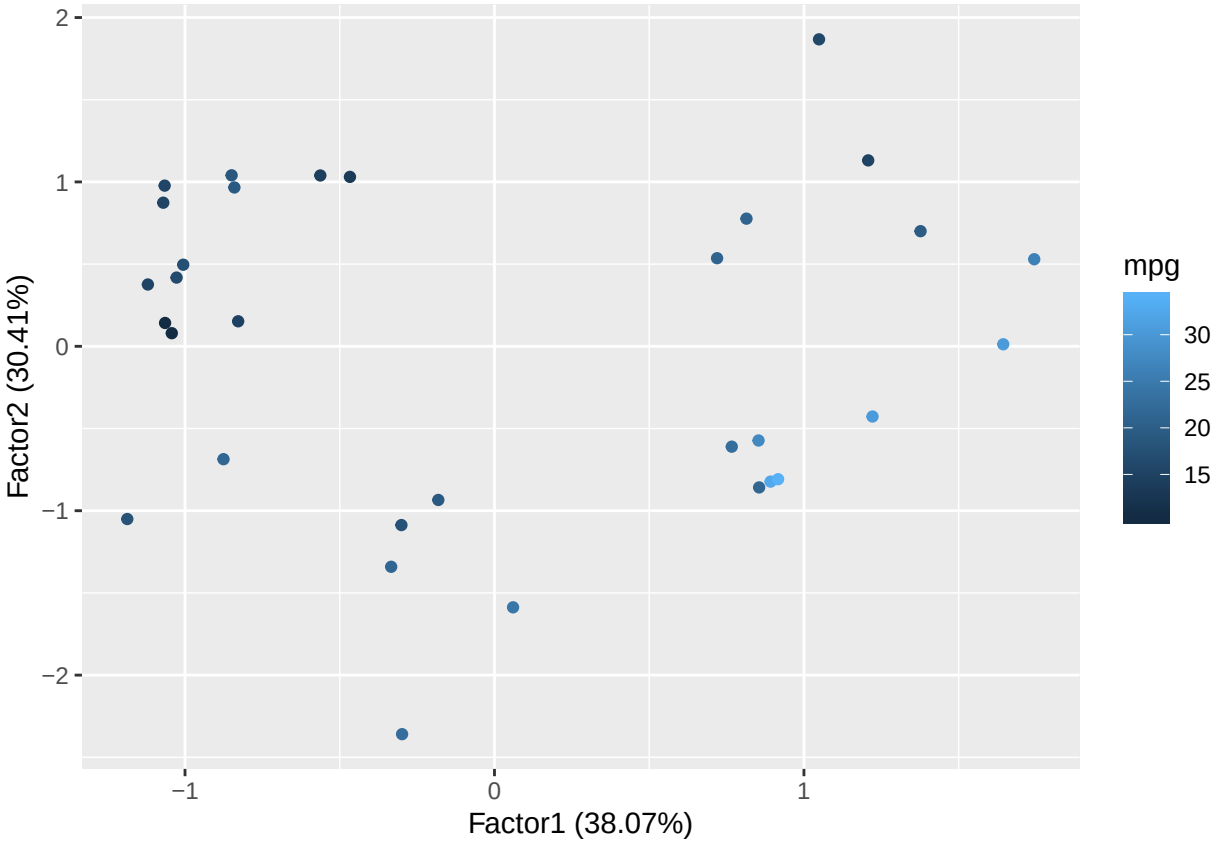
Disable scaling by specifying scale = 0

```
autoplot((df3), scale = 0)
```



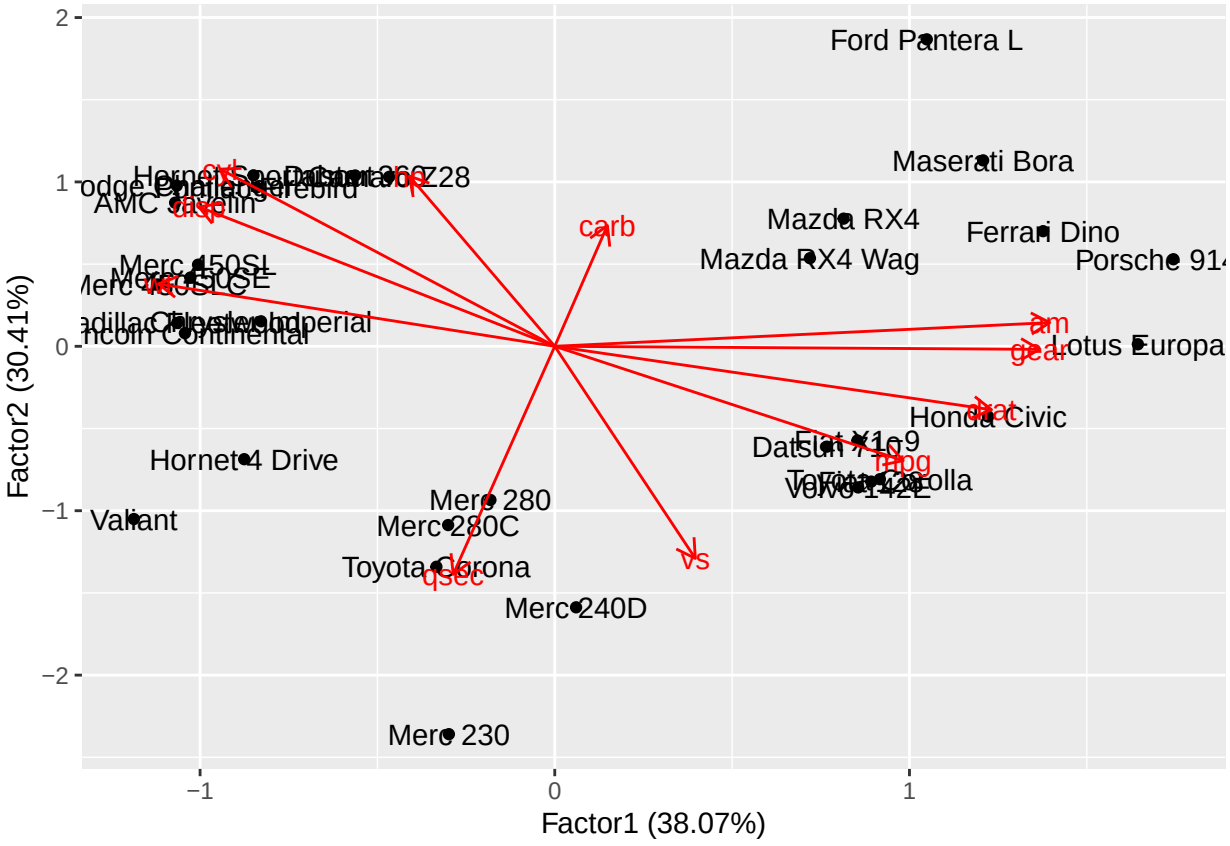
Plotting Factor Analysis

```
d.factanal <- factanal(mtcars, factors = 6, scores = 'regression')
autoplot(d.factanal, data = mtcars, colour = 'mpg')
```



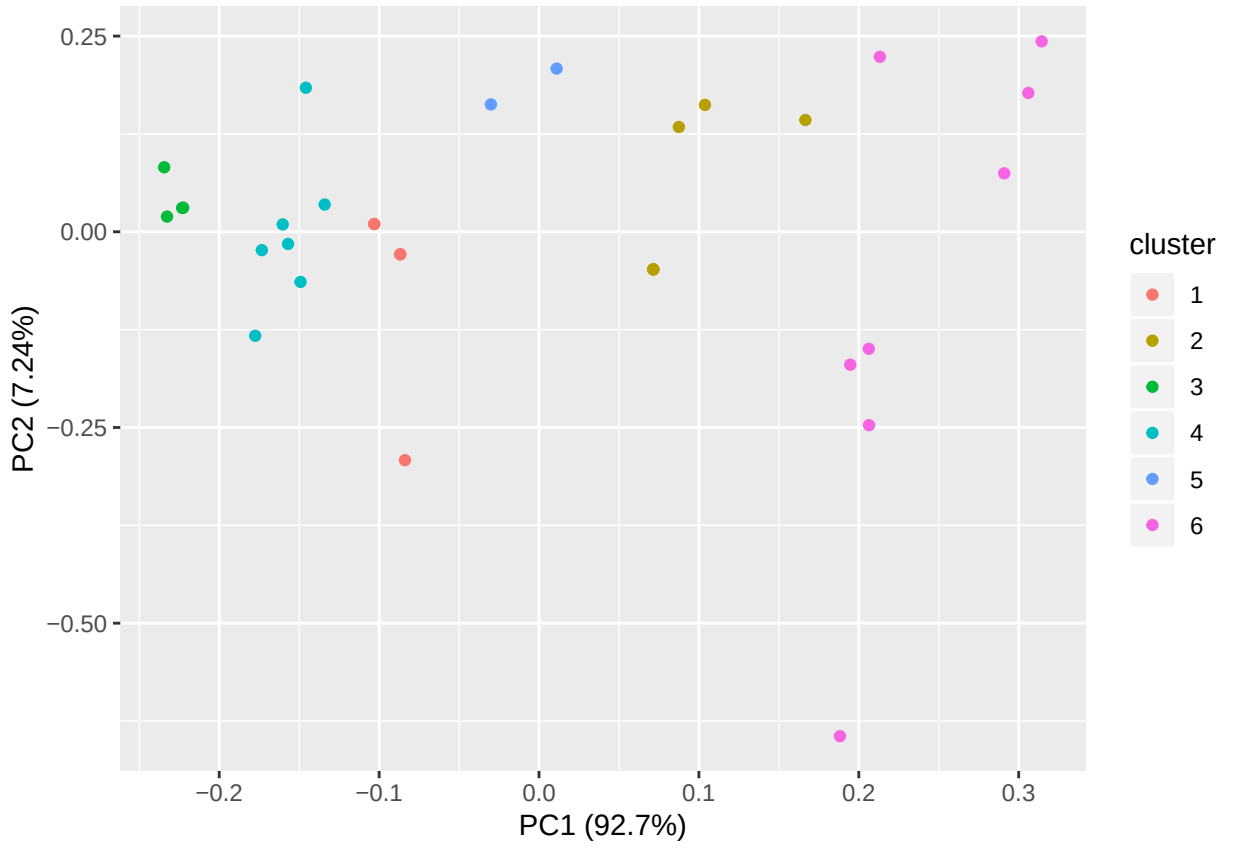
Visualize both the principal component coefficients for each variable and the principal componen.

```
autoplot(d.factanal, label = TRUE, label.size = 4,
         loadings = TRUE, loadings.label = TRUE, loadings.label.size  = 4)
```

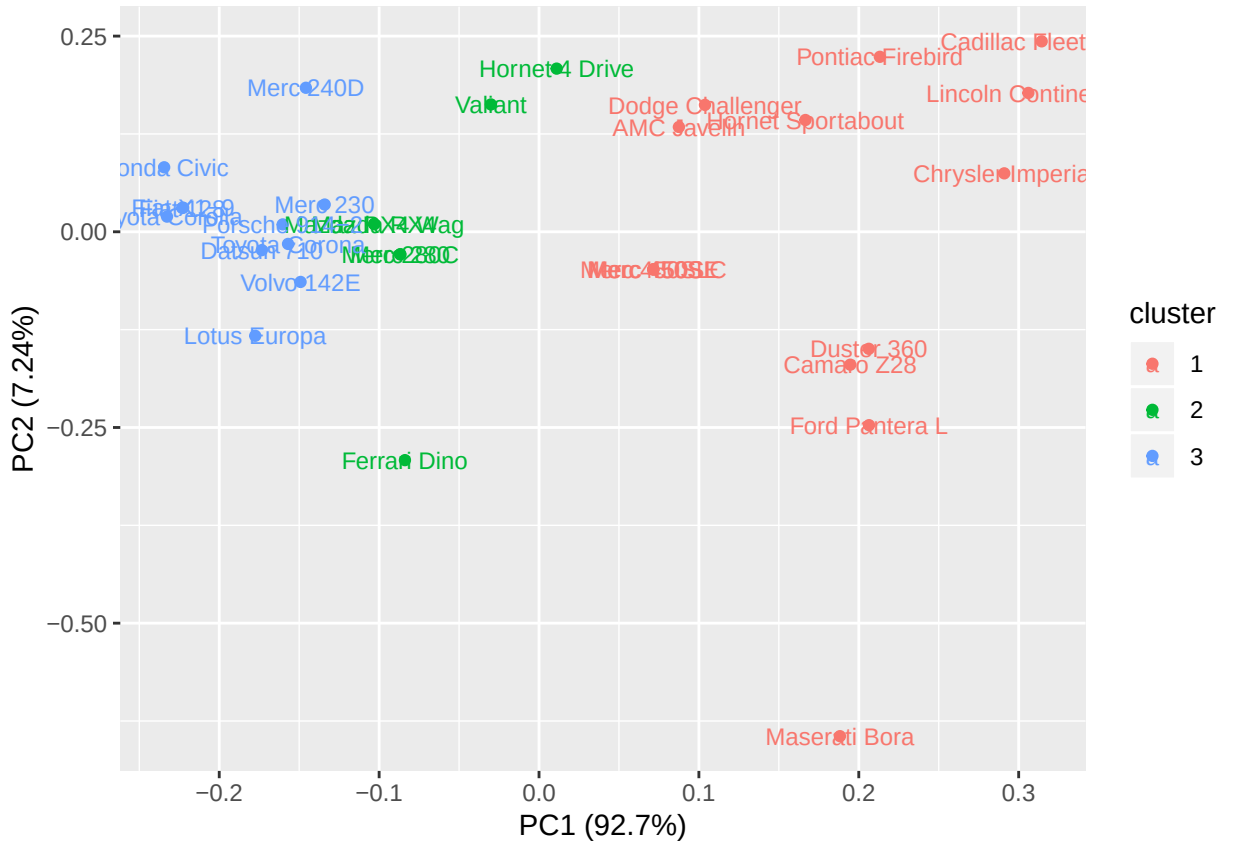


Plotting K-means

```
set.seed(1)
autoplot(kmeans(mtcars, 6), data = mtcars)
```




```
autoplot(kmeans(mtcars, 3), data = mtcars, label = TRUE, label.size = 3)
```

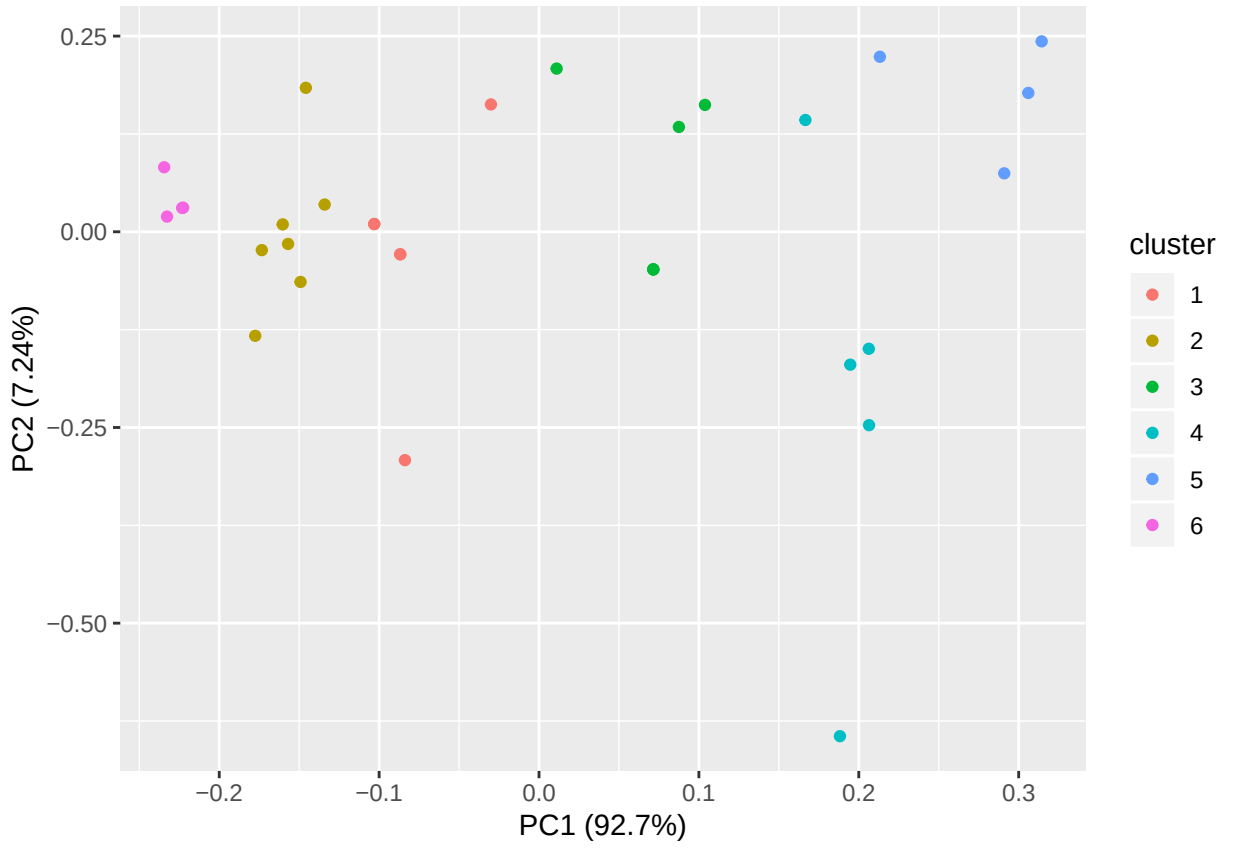


Plotting cluster package

```
library(cluster)
```

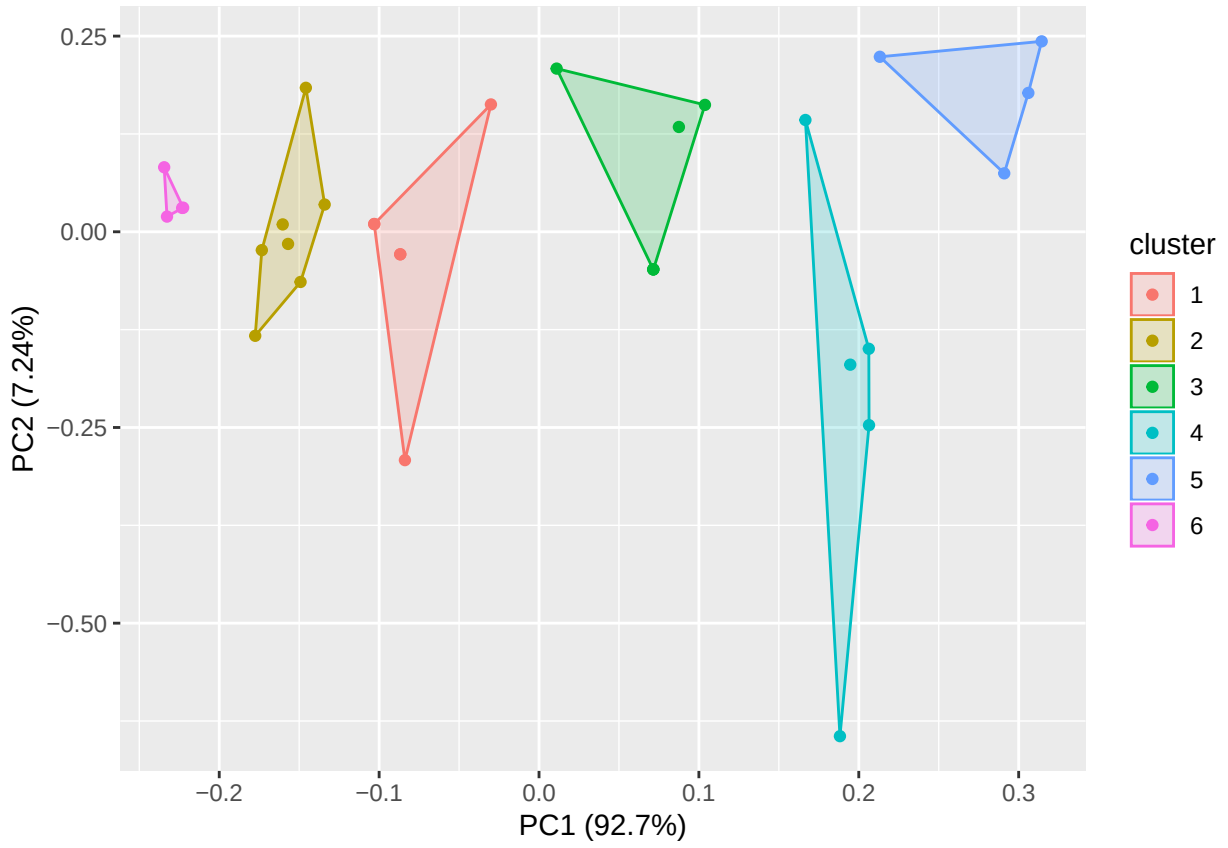
```
## Warning: package 'cluster' was built under R version 3.5.3
```

```
autoplot(clara(mtcars[-6], 6))
```



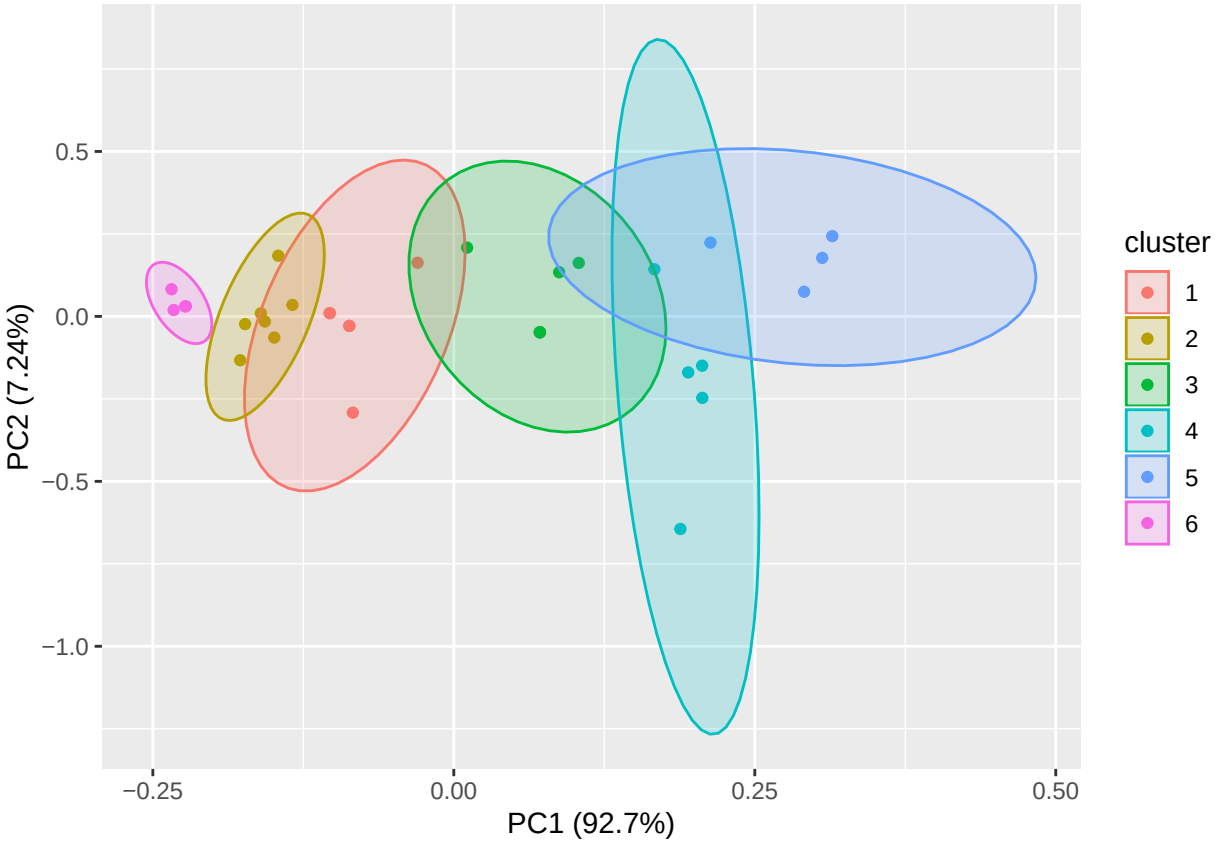
Drawing convex for each cluster

```
autoplot(fanny(mtcars[-6], 6), frame = TRUE)
```



Drawing ellipse for each cluster

```
autoplot(pam(mtcars[-6], 6), frame = TRUE, frame.type = 'norm')
```


0.5
0.0
−0.5
−1.0
PC2 (7.24%)
−0.25
0.00
0.25
0.50
PC1 (92.7%)
cluster
1
2
3
4
5
6