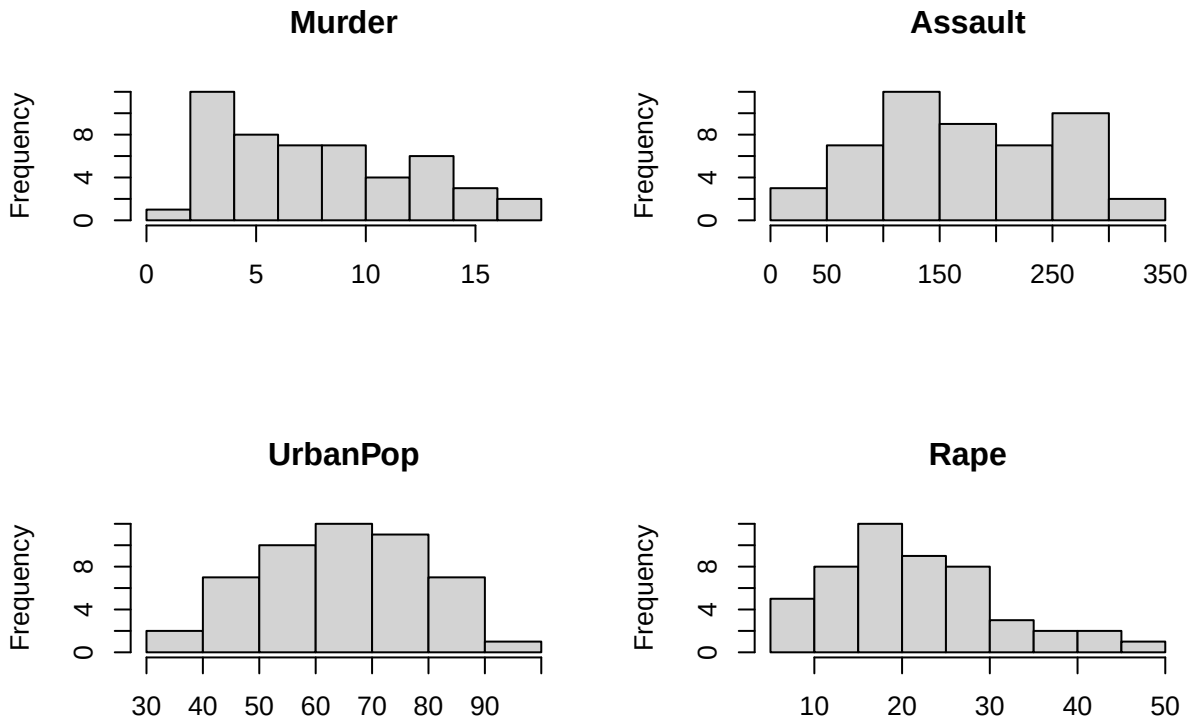
Principal Component Analysis USArrests Dataset and MTCars Dataset

```
data("USArrests")
head(USArrests)
```

```
##            Murder Assault UrbanPop Rape
## Alabama      13.2     236       58 21.2
## Alaska       10.0     263       48 44.5
## Arizona       8.1     294       80 31.0
## Arkansas      8.8     190       50 19.5
## California    9.0     276       91 40.6
## Colorado      7.9     204       78 38.7
```

```
par(mfrow=c(2,2))
for(i in 1:ncol(USArrests)) { hist(USArrests[, i], main = paste(colnames(USArrests[i])), xlab = "") }
```



```
pcaModel <- prcomp(USArrests, scale. = TRUE, center = TRUE)
pcaModel$rotation
```

```
##                 PC1        PC2        PC3         PC4
## Murder   -0.5358995  0.4181809 -0.3412327  0.64922780
## Assault  -0.5831836  0.1879856 -0.2681484 -0.74340748
## UrbanPop -0.2781909 -0.8728062 -0.3780158  0.13387773
## Rape     -0.5434321 -0.1673186  0.8177779  0.08902432
```

```
library(factoextra)
```

```
## Warning: package 'factoextra' was built under R version 4.0.3
```

```
## Loading required package: ggplot2
```

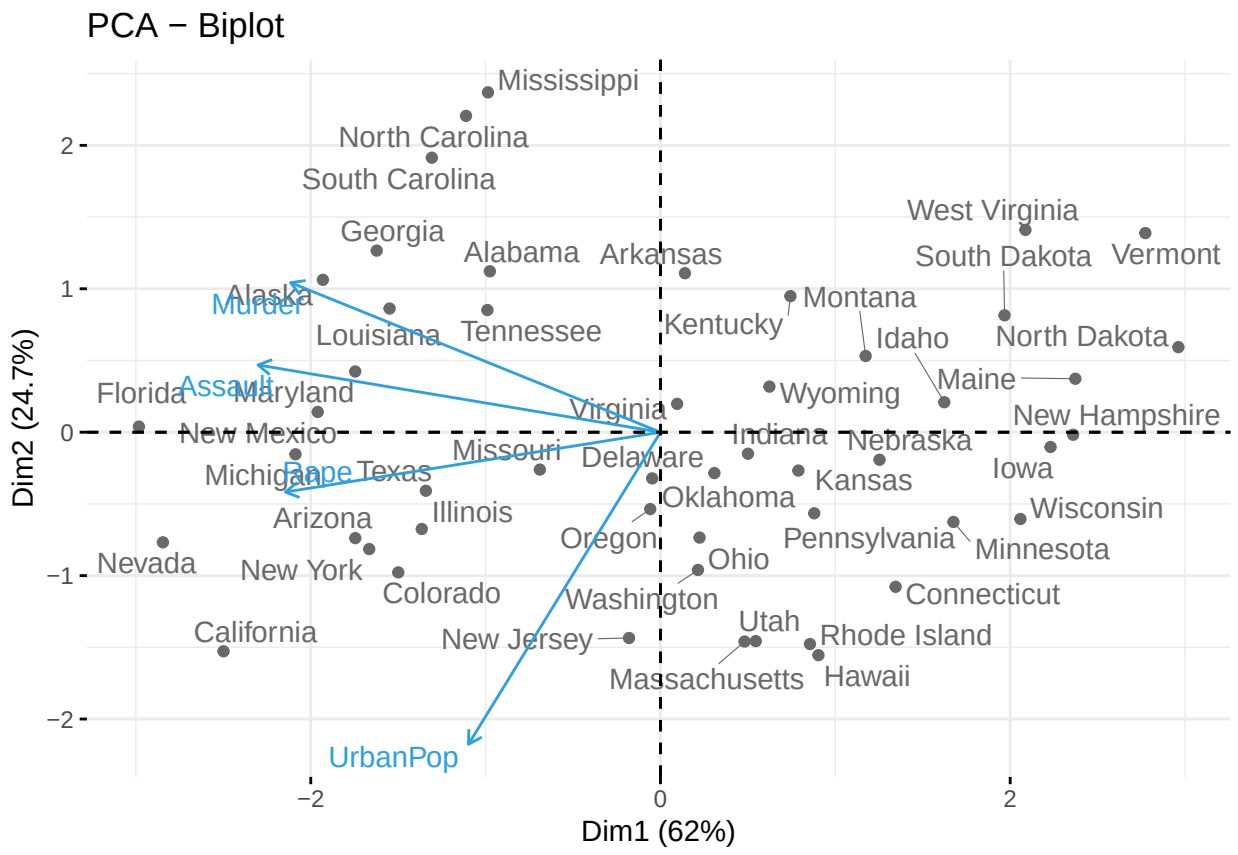
```
## Warning: package 'ggplot2' was built under R version 4.0.3
```

```
## Welcome! Want to learn more? See two factoextra-related books at https://goo.gl/ve3WBa
```

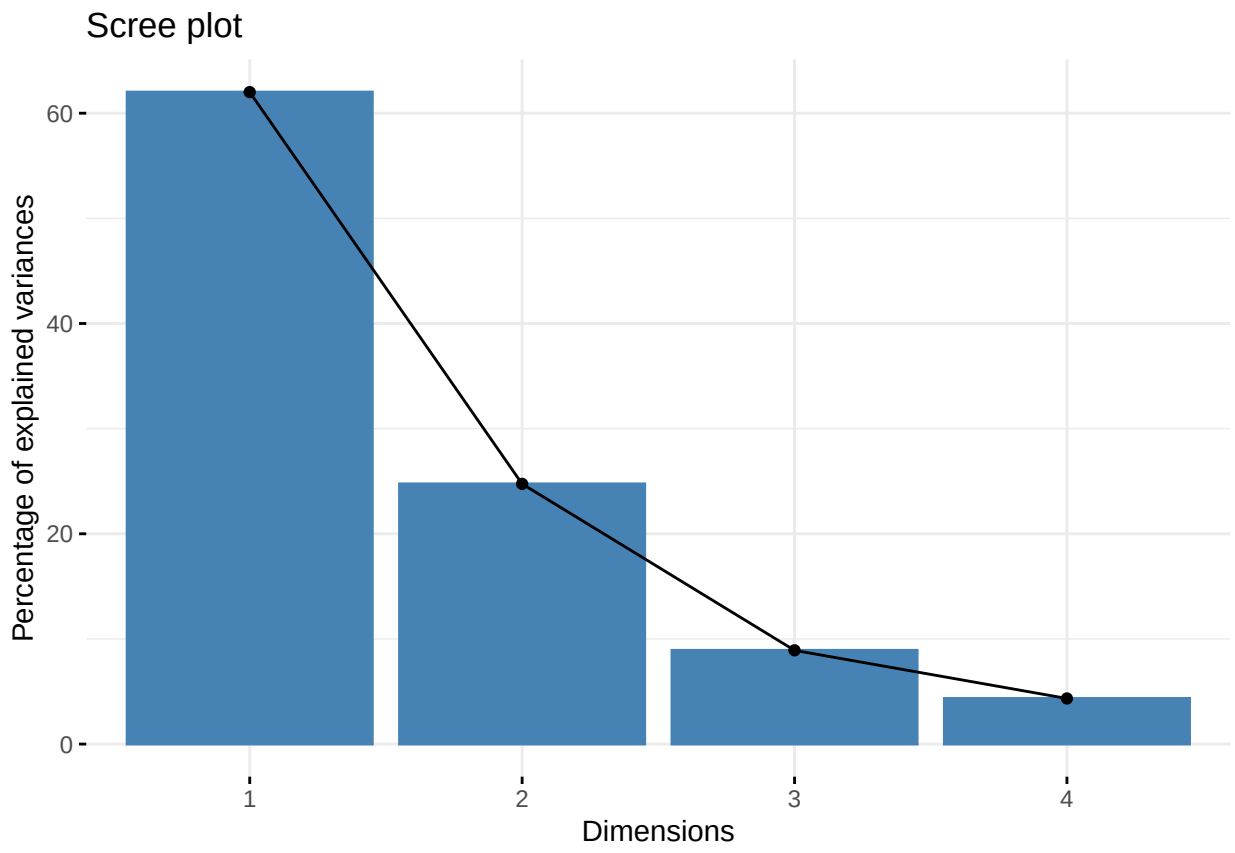
```
fviz_pca_biplot(pcaModel, repel = TRUE,

col.var = "#2E9FDF", # Variables color
col.ind = "#696969" # Individuals color

)
```



```
fviz_eig(pcaModel)
```

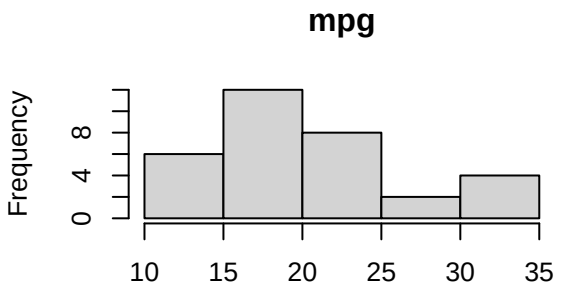


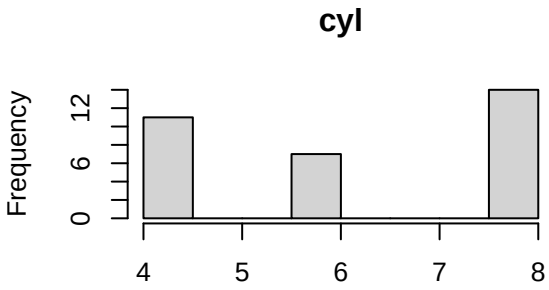
Contoh 2

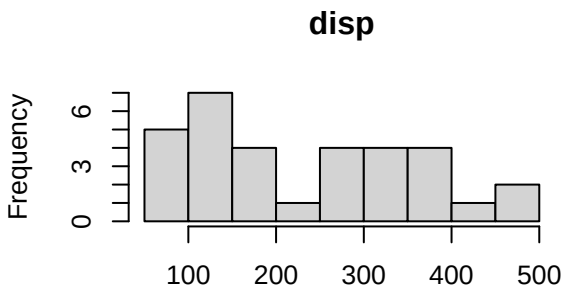
```
dataMPG <- mtcars[, -c(8,9)]
head(dataMPG)
```

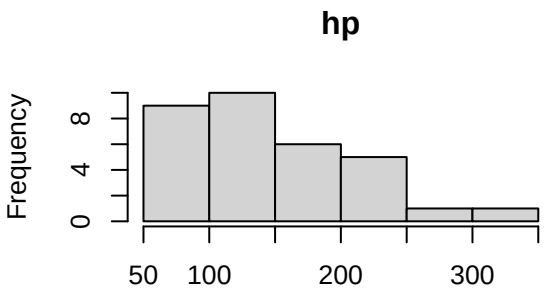
```
##                   mpg cyl disp  hp drat    wt  qsec gear carb
## Mazda RX4         21.0   6  160 110 3.90 2.620 16.46    4    4
## Mazda RX4 Wag     21.0   6  160 110 3.90 2.875 17.02    4    4
## Datsun 710        22.8   4  108  93 3.85 2.320 18.61    4    1
## Hornet 4 Drive    21.4   6  258 110 3.08 3.215 19.44    3    1
## Hornet Sportabout 18.7   8  360 175 3.15 3.440 17.02    3    2
## Valiant           18.1   6  225 105 2.76 3.460 20.22    3    1
```

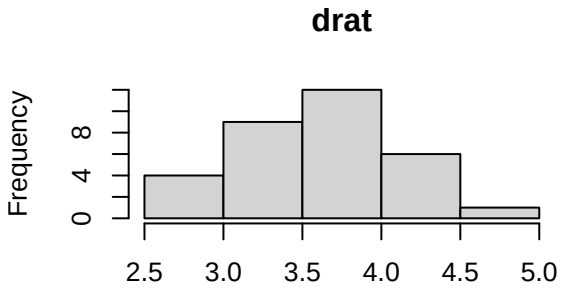
```
par(mfrow=c(2,2))
for(i in 1:ncol(dataMPG)) { hist(dataMPG[, i], main = paste(colnames(dataMPG[i])), xlab = "") }
```

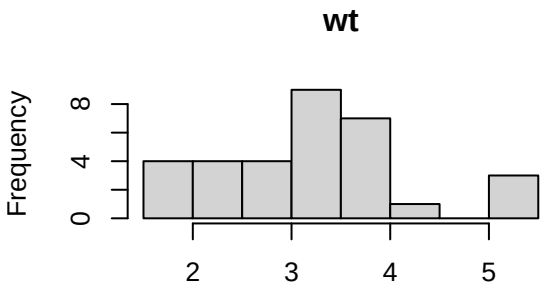

mpg
Frequency
0
4
8
10
15
20
25
30
35

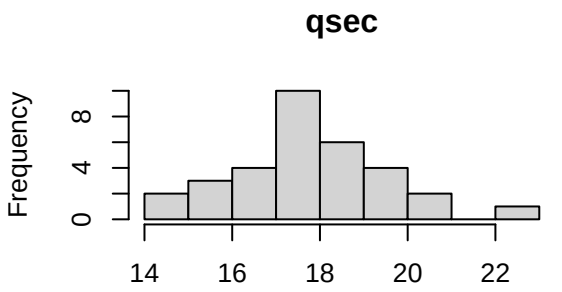

cyl
Frequency
0
6
12
4
5
6
7
8

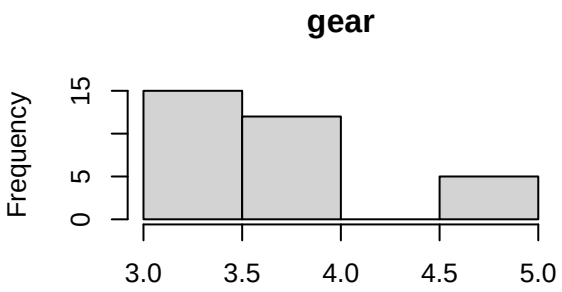

disp
Frequency
0
3
6
100
200
300
400
500

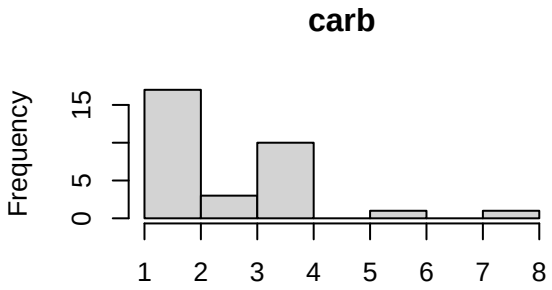

hp
Frequency
0
4
8
50
100
200
300


drat
Frequency
0
4
8
2.5
3.0
3.5
4.0
4.5
5.0


wt
Frequency
0
4
8
2
3
4
5


qsec
Frequency
0
4
8
14
16
18
20
22

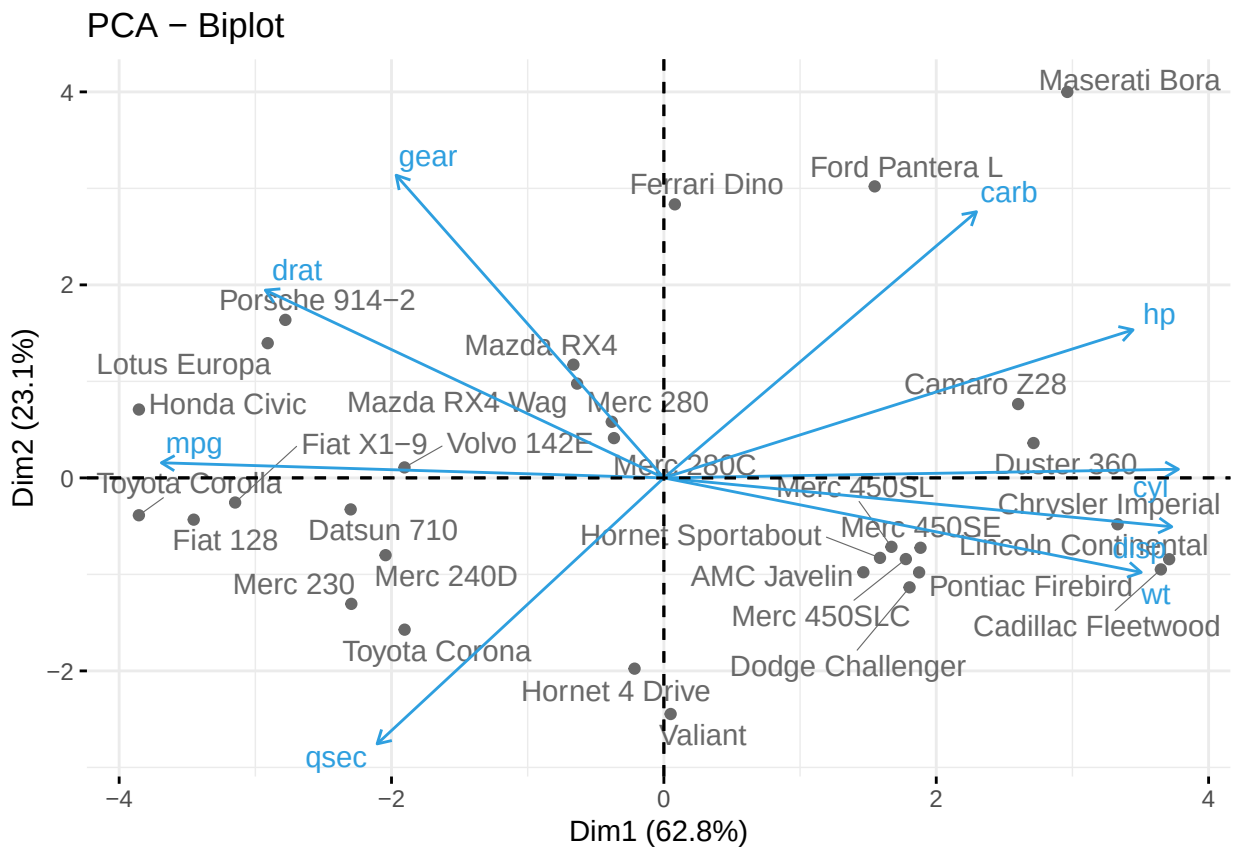

gear
Frequency
0
5
15
3.0
3.5
4.0
4.5
5.0



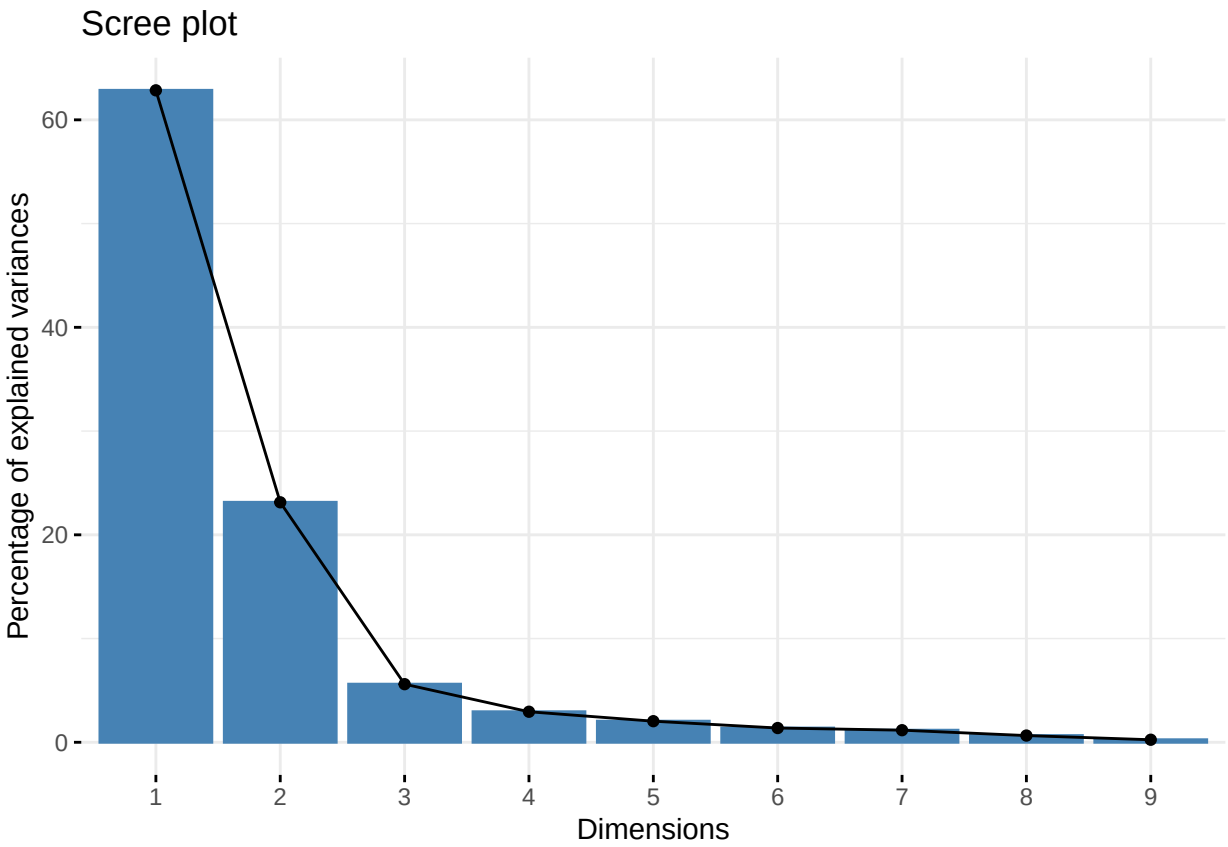

```
mtcarsPca <- prcomp(dataMPG, scale. = TRUE, center=TRUE)
mtcarsPca$rotation
```

```
##             PC1         PC2         PC3          PC4        PC5         PC6
## mpg  -0.3931477  0.02753861 -0.22119309 -0.006126378 -0.3207620  0.72015586
## cyl   0.4025537  0.01570975 -0.25231615  0.040700251  0.1171397  0.22432550
## disp  0.3973528 -0.08888469 -0.07825139  0.339493732 -0.4867849 -0.01967516
## hp    0.3670814  0.26941371 -0.01721159  0.068300993 -0.2947317  0.35394225
## drat -0.3118165  0.34165268  0.14995507  0.845658485  0.1619259 -0.01536794
## wt    0.3734771 -0.17194306  0.45373418  0.191260029 -0.1874822 -0.08377237
## qsec -0.2243508 -0.48404435  0.62812782 -0.030329127 -0.1482495  0.25752940
## gear -0.2094749  0.55078264  0.20658376 -0.282381831 -0.5624860 -0.32298239
## carb  0.2445807  0.48431310  0.46412069 -0.214492216  0.3997820  0.35706914
##              PC7         PC8         PC9
## mpg  -0.38138068 -0.12465987  0.11492862
## cyl  -0.15893251  0.81032177  0.16266295
## disp -0.18233095 -0.06416707 -0.66190812
## hp    0.69620751 -0.16573993  0.25177306
## drat  0.04767957  0.13505066  0.03809096
## wt   -0.42777608 -0.19839375  0.56918844
## qsec  0.27622581  0.35613350 -0.16873731
## gear -0.08555707  0.31636479  0.04719694
## carb -0.20604210 -0.10832772 -0.32045892
```

```
library(factoextra)
fviz_pca_biplot(mtcarsPca, repel = TRUE,
col.var = "#2E9FDF",
col.ind = "#696969"
)
```




```
fviz_eig(mtcarsPca)
```


Scree plot
Percentage of explained variances
60
40
20
0
1
2
3
4
5
6
7
8
9
Dimensions