Independent Study

```
library(readxl)
library(openxlsx)
library(tidyr)
library(stringr)
library(tibble)
library(car)
```

```
## Loading required package: carData
```

```
library(showtext)
```

```
## Loading required package: sysfonts
```

```
## Loading required package: showtextdb
```

```
library(dplyr)
```

```
##
## Attaching package: 'dplyr'
```

```
## The following object is masked from 'package:car':
##
##     recode
```

```
## The following objects are masked from 'package:stats':
##
##     filter, lag
```

```
## The following objects are masked from 'package:base':
##
##     intersect, setdiff, setequal, union
```

```
library(CCP)
library(readxl)
library(ggcorrplot)
```

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

```
library(factoextra)
```

```
## Welcome! Want to learn more? See two factoextra-related books at https://goo.gl/ve3WB
a
```

```
library(corrplot)
```

```
## corrplot 0.92 loaded
```

```
library(glmnet)
```

```
## Loading required package: Matrix
```

```
##
## Attaching package: 'Matrix'
```

```
## The following objects are masked from 'package:tidyr':
##
##     expand, pack, unpack
```

```
## Loaded glmnet 4.1-6
```

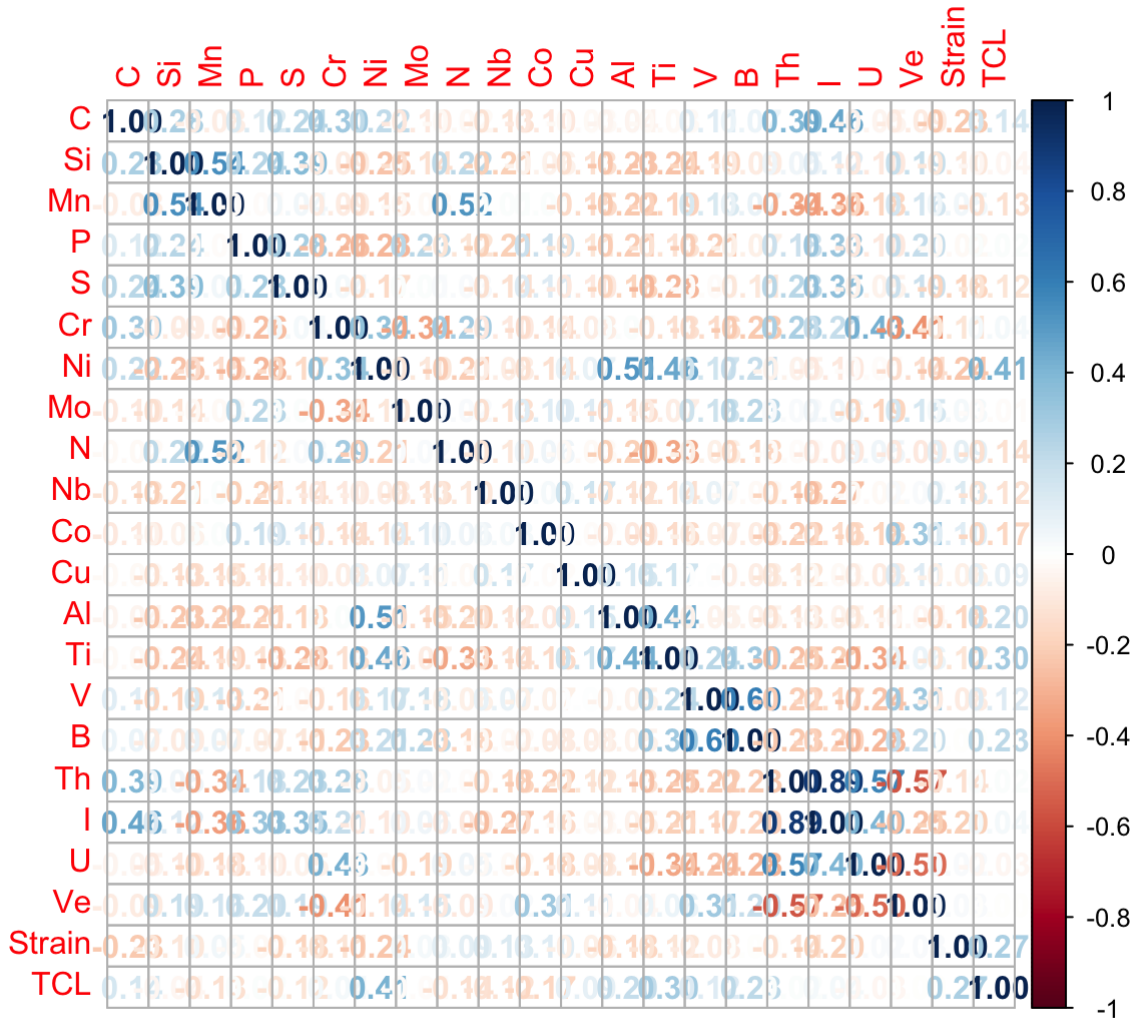
```
library(factoextra)
library(FactoMineR)
require(FactoMineR)
require(factoextra)
setwd("/Users/jettadler/Desktop/Data /Study")
```

```
#Defining Data
rm(list = ls())
stl_raw <- read_xls('steel_data.xls')
al_raw <- read_xls('Al_data.xls')
```

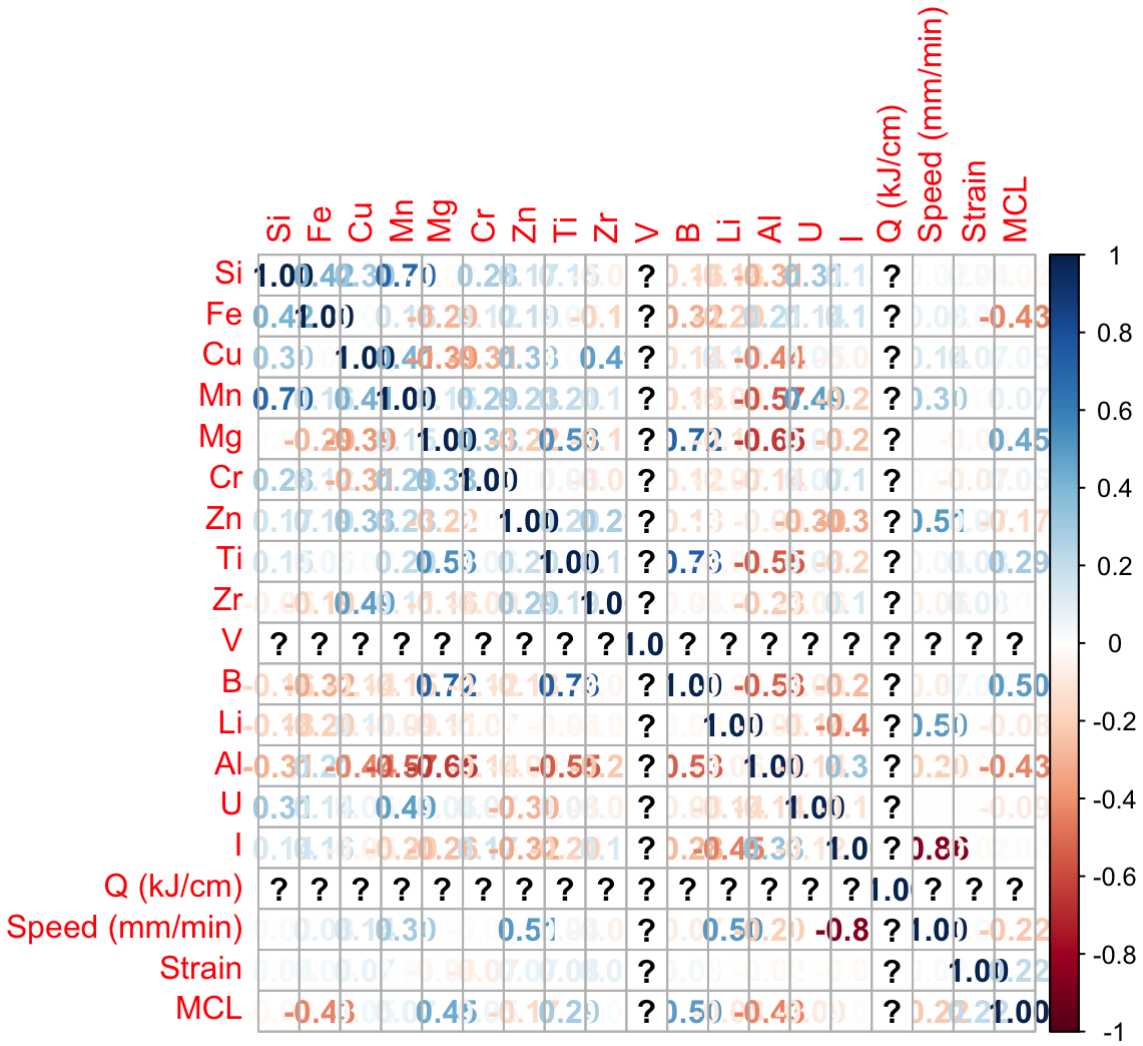
```
#removing blank rows and name/number coulumbs
toDelete <- seq(1, 974, 2)
stl <-  stl_raw[-toDelete, ]
al <-  al_raw[-toDelete, ]
stl <- stl[c(-1)]
al <- al[c(-1)]
```

```
#Seperating Data into Elemental Compositons, Strain, and Crack Length Data
stEl <- stl[c(1:16)]
alEl <- al[c(1:13)]
stSTN <- stl[c(21)]
stTCL <- stl[c(22)]
alSTN <- al[c(18)]
alMCL <- al[c(19)]
```

```
#First Visual check of possible correlations between variables
cors <- cor(stl)
cora <- cor(al)
corrplot(cors, method='number')
```




```
corrplot(cora, method='number')
```


Si
Fe
Cu
Mn
Mg
Cr
Zn
Ti
Zr
V
B
Li
Al
U
I
Q (kJ/cm)
Speed (mm/min)
Strain
MCL
1
0.8
0.6
0.4
0.2
0
-0.2
-0.4
-0.6
-0.8
-1

```
cors
```

```
##                  C           Si           Mn            P           S           Cr
## C       1.00000000  0.27845666 -0.07746384  0.12421010  0.23829076  0.30252074
## Si      0.27845666  1.00000000  0.54288940  0.24403814  0.38630488 -0.08870044
## Mn     -0.07746384  0.54288940  1.00000000 -0.01076547  0.04041910 -0.08914258
## P       0.12421010  0.24403814 -0.01076547  1.00000000  0.28347200 -0.26479798
## S       0.23829076  0.38630488  0.04041910  0.28347200  1.00000000  0.03644738
## Cr      0.30252074 -0.08870044 -0.08914258 -0.26479798  0.03644738  1.00000000
## Ni      0.21548779 -0.24832062 -0.14784543 -0.28337310 -0.16648652  0.34011443
## Mo     -0.09549286 -0.13819168 -0.02845569  0.23045744  0.01192515 -0.34391819
## N      -0.02481124  0.22103907  0.52387534 -0.11579085  0.02805146  0.29194807
## Nb     -0.12579354 -0.20581200  0.01201195 -0.20526057 -0.13866867 -0.09773600
## Co     -0.09669215 -0.05572159  0.01495095  0.18950145  0.11394630 -0.14301930
## Cu     -0.02918592 -0.13300644 -0.15206014 -0.11422006 -0.09795553 -0.07875766
## Al     -0.04188047 -0.23260626 -0.22424775 -0.21209939 -0.18421137  0.01428158
## Ti     -0.02531531 -0.24281444 -0.19008803 -0.12976077 -0.28308274 -0.13446150
## V       0.10629471 -0.18519455  0.13304538 -0.21233624 -0.01149909 -0.15701641
## B       0.06647661 -0.08702766  0.07066607 -0.06543854 -0.10817779 -0.22581597
## Th      0.39039469  0.05270954 -0.33500554  0.18192163  0.23337489  0.27515468
## I       0.46477037  0.11991791 -0.36261756  0.32546704  0.35347395  0.20792318
## U      -0.04636266 -0.09558830 -0.17551355 -0.09585409 -0.05185645  0.42670047
## Ve     -0.08895166  0.19076415  0.15947449  0.20161899  0.18986257 -0.40924212
## Strain -0.22771168 -0.10258401  0.04880527 -0.02243995 -0.18159471 -0.11390472
## TCL     0.14411528 -0.03647010 -0.13099629  0.01561729 -0.11962175  0.03810476
##                 Ni          Mo           N          Nb          Co           Cu
## C       0.21548779 -0.09549286 -0.02481124 -0.12579354 -0.09669215 -0.029185924
## Si     -0.24832062 -0.13819168  0.22103907 -0.20581200 -0.05572159 -0.133006445
## Mn     -0.14784543 -0.02845569  0.52387534  0.01201195  0.01495095 -0.152060141
## P      -0.28337310  0.23045744 -0.11579085 -0.20526057  0.18950145 -0.114220062
## S      -0.16648652  0.01192515  0.02805146 -0.13866867  0.11394630 -0.097955535
## Cr      0.34011443 -0.34391819  0.29194807 -0.09773600 -0.14301930 -0.078757663
## Ni      1.00000000 -0.10408396 -0.21280034 -0.07952737 -0.13510072  0.073374902
## Mo     -0.10408396  1.00000000  0.02217061 -0.13280162  0.10478750  0.110931108
## N      -0.21280034  0.02217061  1.00000000 -0.10490213  0.05665789 -0.025480073
## Nb     -0.07952737 -0.13280162 -0.10490213  1.00000000  0.03159128  0.167369769
## Co     -0.13510072  0.10478750  0.05665789  0.03159128  1.00000000 -0.018090233
## Cu      0.07337490  0.11093111 -0.02548007  0.16736977 -0.01809023  1.000000000
## Al      0.51472603 -0.14673019 -0.20282819 -0.12249223 -0.09141180  0.146781527
## Ti      0.45899169 -0.07394976 -0.33377034 -0.13915043 -0.15716960  0.167676563
## V       0.16933219  0.17510812 -0.06142697  0.07451321 -0.06661316  0.002611036
## B       0.20741000  0.23308727 -0.17500646 -0.02469812 -0.08367700 -0.078963615
## Th     -0.05261927  0.02448561 -0.01815577 -0.18477329 -0.21509654 -0.117651859
## I      -0.09658742  0.05869663 -0.09046160 -0.26644977 -0.15674378 -0.029972280
## U      -0.01035120 -0.18796257  0.05176411 -0.02213752 -0.18049103 -0.081733434
## Ve     -0.14203631  0.14721887 -0.09227065  0.02429409  0.30537700  0.113403144
## Strain -0.23926109  0.03004494  0.09036873  0.12658403  0.10194218 -0.064979957
## TCL     0.41290930 -0.01386736 -0.13861387 -0.12135101 -0.17329686  0.089590436
##                 Al          Ti            V             B          Th
## C      -0.04188047 -0.02531531  0.106294712  0.0664766134  0.39039469
## Si     -0.23260626 -0.24281444 -0.185194549 -0.0870276593  0.05270954
## Mn     -0.22424775 -0.19008803  0.133045383  0.0706660690 -0.33500554
## P      -0.21209939 -0.12976077 -0.212336238 -0.0654385438  0.18192163
## S      -0.18421137 -0.28308274 -0.011499094 -0.1081777913  0.23337489
```

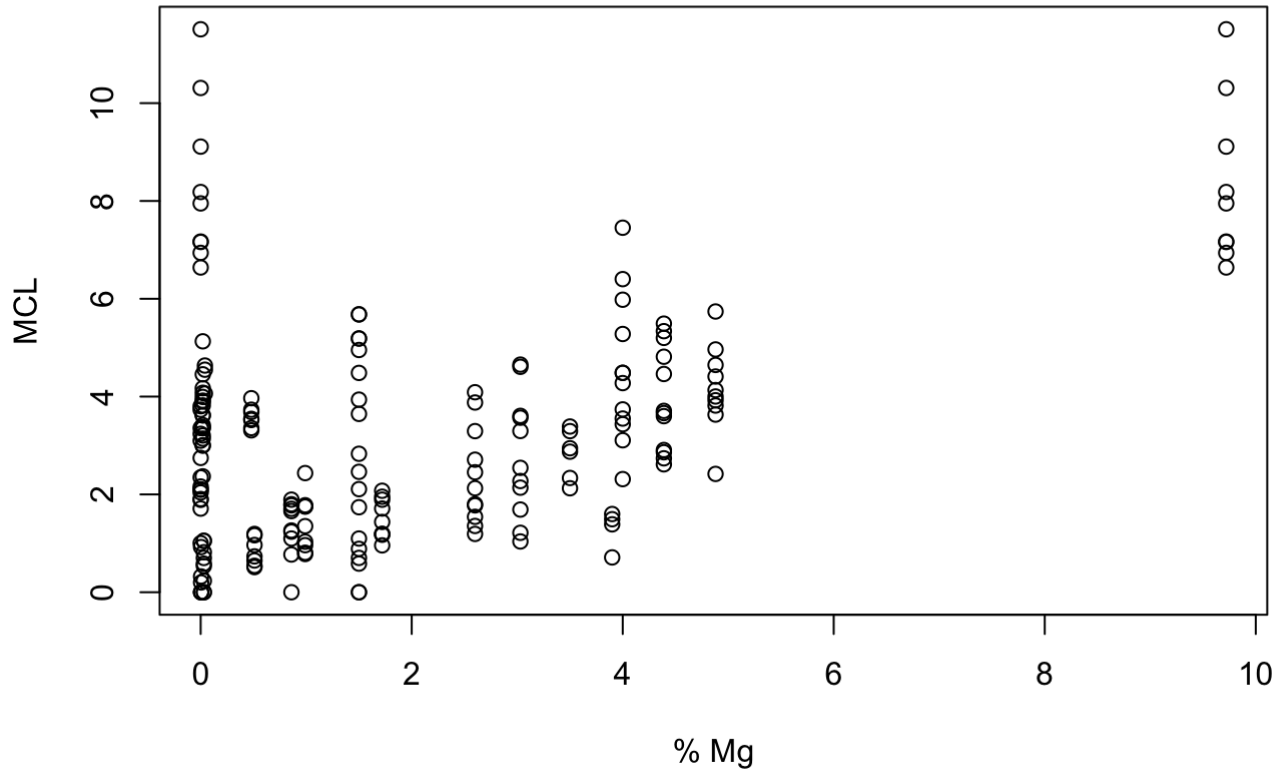
```
## Cr      0.01428158 -0.13446150 -0.157016408 -0.2258159695  0.27515468
## Ni      0.51472603  0.45899169  0.169332190  0.2074099994 -0.05261927
## Mo     -0.14673019 -0.07394976  0.175108125  0.2330872675  0.02448561
## N      -0.20282819 -0.33377034 -0.061426969 -0.1750064628 -0.01815577
## Nb     -0.12249223 -0.13915043  0.074513210 -0.0246981243 -0.18477329
## Co     -0.09141180 -0.15716960 -0.066613158 -0.0836769967 -0.21509654
## Cu      0.14678153  0.16767656  0.002611036 -0.0789636148 -0.11765186
## Al      1.00000000  0.43984444 -0.053377976 -0.0535430587 -0.12812092
## Ti      0.43984444  1.00000000  0.239107468  0.3038814520 -0.25461669
## V      -0.05337798  0.23910747  1.000000000  0.6001763018 -0.21830014
## B      -0.05354306  0.30388145  0.600176302  1.0000000000 -0.22515558
## Th     -0.12812092 -0.25461669 -0.218300141 -0.2251555817  1.00000000
## I      -0.05461621 -0.20797505 -0.172939417 -0.1962436393  0.89051845
## U      -0.11431258 -0.33850179 -0.242033956 -0.2811999863  0.57241764
## Ve     -0.03926660  0.06426317  0.309047426  0.2020865864 -0.56769956
## Strain -0.17956990 -0.11546403 -0.080976829 -0.0007026004 -0.13595857
## TCL     0.20227094  0.29650669  0.115898303  0.2297429217  0.01977808
##                 I           U           Ve        Strain          TCL
## C       0.46477037 -0.04636266 -0.088951656 -0.2277116838  0.144115282
## Si      0.11991791 -0.09558830  0.190764151 -0.1025840072 -0.036470102
## Mn     -0.36261756 -0.17551355  0.159474494  0.0488052736 -0.130996289
## P       0.32546704 -0.09585409  0.201618986 -0.0224399461  0.015617293
## S       0.35347395 -0.05185645  0.189862565 -0.1815947086 -0.119621747
## Cr      0.20792318  0.42670047 -0.409242123 -0.1139047217  0.038104765
## Ni     -0.09658742 -0.01035120 -0.142036314 -0.2392610946  0.412909297
## Mo      0.05869663 -0.18796257  0.147218868  0.0300449419 -0.013867362
## N      -0.09046160  0.05176411 -0.092270649  0.0903687260 -0.138613869
## Nb     -0.26644977 -0.02213752  0.024294094  0.1265840319 -0.121351014
## Co     -0.15674378 -0.18049103  0.305376999  0.1019421843 -0.173296857
## Cu     -0.02997228 -0.08173343  0.113403144 -0.0649799572  0.089590436
## Al     -0.05461621 -0.11431258 -0.039266597 -0.1795698951  0.202270939
## Ti     -0.20797505 -0.33850179  0.064263167 -0.1154640284  0.296506692
## V      -0.17293942 -0.24203396  0.309047426 -0.0809768292  0.115898303
## B      -0.19624364 -0.28119999  0.202086586 -0.0007026004  0.229742922
## Th      0.89051845  0.57241764 -0.567699564 -0.1359585695  0.019778081
## I       1.00000000  0.40114754 -0.250248230 -0.2020144197  0.036293631
## U       0.40114754  1.00000000 -0.502137485  0.0242652899 -0.033752736
## Ve     -0.25024823 -0.50213749  1.000000000  0.0274623077 -0.004297735
## Strain -0.20201442  0.02426529  0.027462308  1.0000000000  0.270733180
## TCL     0.03629363 -0.03375274 -0.004297735  0.2707331797  1.000000000
```

```
cora
```

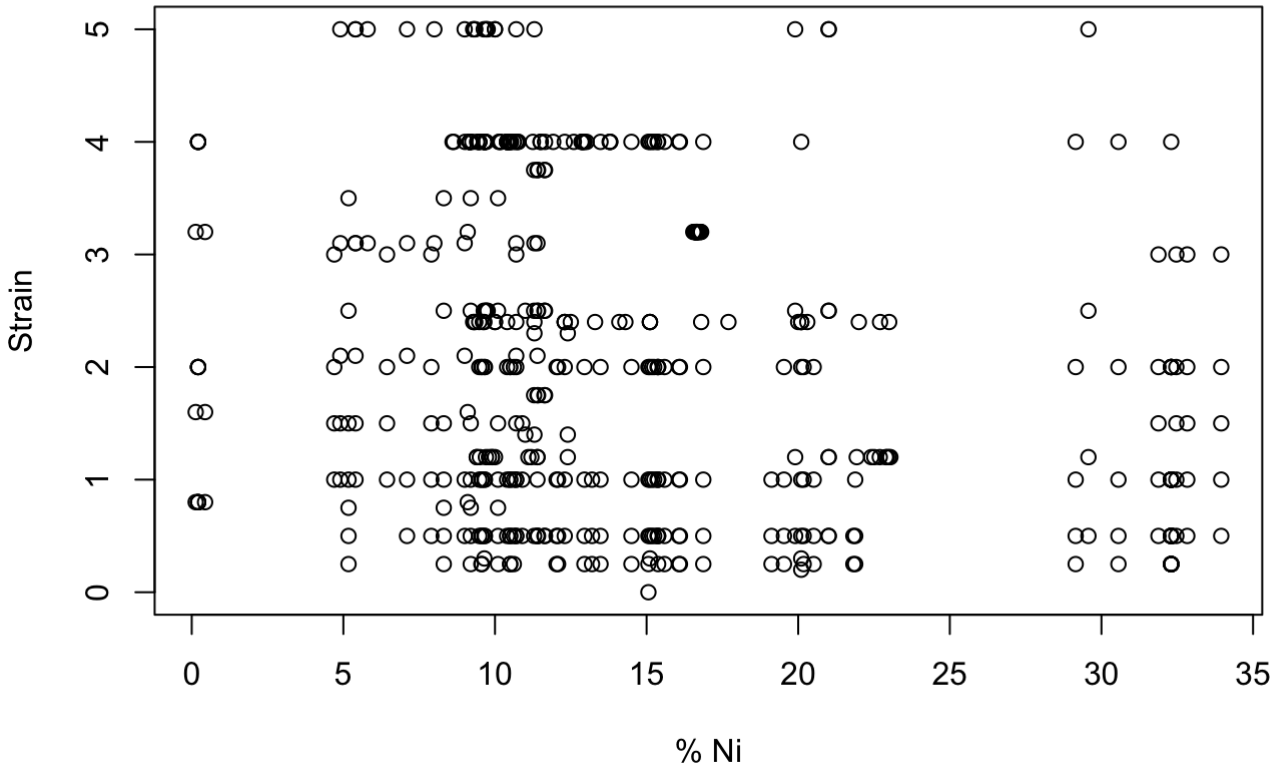
```
##                         Si          Fe          Cu          Mn          Mg
## Si               1.00000000  0.41946565  0.30371090  0.69985686 -0.01679518
## Fe               0.41946565  1.00000000  0.01367361  0.16147267 -0.28617836
## Cu               0.30371090  0.01367361  1.00000000  0.41177930 -0.38803728
## Mn               0.69985686  0.16147267  0.41177930  1.00000000  0.14739982
## Mg              -0.01679518 -0.28617836 -0.38803728  0.14739982  1.00000000
## Cr               0.28442177  0.11760990 -0.30875105  0.28998191  0.33209597
## Zn               0.16886697  0.18712629  0.32661512  0.23479904 -0.21757192
## Ti               0.14720438  0.04773796  0.01882995  0.20480612  0.53313617
## Zr              -0.05186117 -0.18685554  0.48864377  0.11018048 -0.15693241
## V                        NA          NA          NA          NA          NA
## B               -0.16233589 -0.31560459 -0.13845711 -0.15064796  0.71706925
## Li              -0.17604585 -0.20347178  0.10138874 -0.08560685 -0.11009762
## Al              -0.30722007  0.21015028 -0.43847668 -0.56725161 -0.64583767
## U                0.31146101  0.14376163  0.05113820  0.48994437  0.06454867
## I                0.13730297  0.16278287 -0.05140019 -0.20438381 -0.25639121
## Q (kJ/cm)               NA          NA          NA          NA          NA
## Speed (mm/min)   0.02486073  0.07786063  0.13653407  0.29828949  0.00980461
## Strain           0.03654705  0.03700105  0.07475819  0.01044924 -0.03762740
## MCL             -0.02481564 -0.43339428  0.04703515  0.07191182  0.44590688
##                         Cr          Zn          Ti          Zr  V           B
## Si               0.28442177  0.16886697  0.14720438 -0.05186117 NA -0.16233589
## Fe               0.11760990  0.18712629  0.04773796 -0.18685554 NA -0.31560459
## Cu              -0.30875105  0.32661512  0.01882995  0.48864377 NA -0.13845711
## Mn               0.28998191  0.23479904  0.20480612  0.11018048 NA -0.15064796
## Mg               0.33209597 -0.21757192  0.53313617 -0.15693241 NA  0.71706925
## Cr               1.00000000  0.01412643  0.05551047 -0.09409050 NA -0.11616247
## Zn               0.01412643  1.00000000  0.19682255  0.28933882 NA -0.13367152
## Ti               0.05551047  0.19682255  1.00000000  0.18956819 NA  0.73226155
## Zr              -0.09409050  0.28933882  0.18956819  1.00000000 NA -0.03919309
## V                        NA          NA          NA          NA  1           NA
## B               -0.11616247 -0.13367152  0.73226155 -0.03919309 NA  1.00000000
## Li              -0.06601020 -0.00622270 -0.03687814 -0.02227177 NA -0.02749633
## Al              -0.14139048 -0.08588036 -0.54585787 -0.22784879 NA -0.53310298
## U                0.07318558 -0.29584598  0.07620284 -0.06182329 NA -0.07632594
## I                0.16831816 -0.32416979 -0.19510682  0.13477017 NA -0.27719980
## Q (kJ/cm)               NA          NA          NA          NA NA           NA
## Speed (mm/min)  -0.01094396  0.50716711  0.04443688 -0.05772749 NA -0.07126934
## Strain          -0.07356959  0.06596877  0.08489359  0.07694974 NA  0.03121366
## MCL              0.05095885 -0.16555223  0.29152053  0.01311892 NA  0.49887358
##                         Li          Al             U           I Q (kJ/cm)
## Si              -0.176045854 -0.30722007  3.114610e-01  0.13730297        NA
## Fe              -0.203471785  0.21015028  1.437616e-01  0.16278287        NA
## Cu               0.101388738 -0.43847668  5.113820e-02 -0.05140019        NA
## Mn              -0.085606846 -0.56725161  4.899444e-01 -0.20438381        NA
## Mg              -0.110097619 -0.64583767  6.454867e-02 -0.25639121        NA
## Cr              -0.066010205 -0.14139048  7.318558e-02  0.16831816        NA
## Zn              -0.006222700 -0.08588036 -2.958460e-01 -0.32416979        NA
## Ti              -0.036878139 -0.54585787  7.620284e-02 -0.19510682        NA
## Zr              -0.022271770 -0.22784879 -6.182329e-02  0.13477017        NA
## V                        NA          NA            NA          NA        NA
## B               -0.027496334 -0.53310298 -7.632594e-02 -0.27719980        NA
```

```
## Li              1.000000000 -0.05764260 -1.434639e-01 -0.44560128        NA
## Al             -0.057642603  1.00000000 -1.436014e-01  0.33418752        NA
## U              -0.143463859 -0.14360139  1.000000e+00 -0.11904365        NA
## I              -0.445601278  0.33418752 -1.190437e-01  1.00000000        NA
## Q (kJ/cm)                NA          NA            NA          NA         1
## Speed (mm/min)  0.500221599 -0.20118505 -3.739558e-03 -0.86067536        NA
## Strain         -0.007889434 -0.02267967 -9.099931e-05 -0.02044928        NA
## MCL            -0.082222575 -0.43379658 -8.733686e-02  0.01904222        NA
##                Speed (mm/min)        Strain         MCL
## Si                0.024860725  3.654705e-02 -0.02481564
## Fe                0.077860633  3.700105e-02 -0.43339428
## Cu                0.136534070  7.475819e-02  0.04703515
## Mn                0.298289490  1.044924e-02  0.07191182
## Mg                0.009804610 -3.762740e-02  0.44590688
## Cr               -0.010943957 -7.356959e-02  0.05095885
## Zn                0.507167108  6.596877e-02 -0.16555223
## Ti                0.044436878  8.489359e-02  0.29152053
## Zr               -0.057727490  7.694974e-02  0.01311892
## V                          NA            NA          NA
## B                -0.071269339  3.121366e-02  0.49887358
## Li                0.500221599 -7.889434e-03 -0.08222257
## Al               -0.201185049 -2.267967e-02 -0.43379658
## U                -0.003739558 -9.099931e-05 -0.08733686
## I                -0.860675362 -2.044928e-02  0.01904222
## Q (kJ/cm)                  NA            NA          NA
## Speed (mm/min)    1.000000000  2.002035e-02 -0.22277685
## Strain            0.020020351  1.000000e+00  0.21957716
## MCL              -0.222776847  2.195772e-01  1.00000000
```

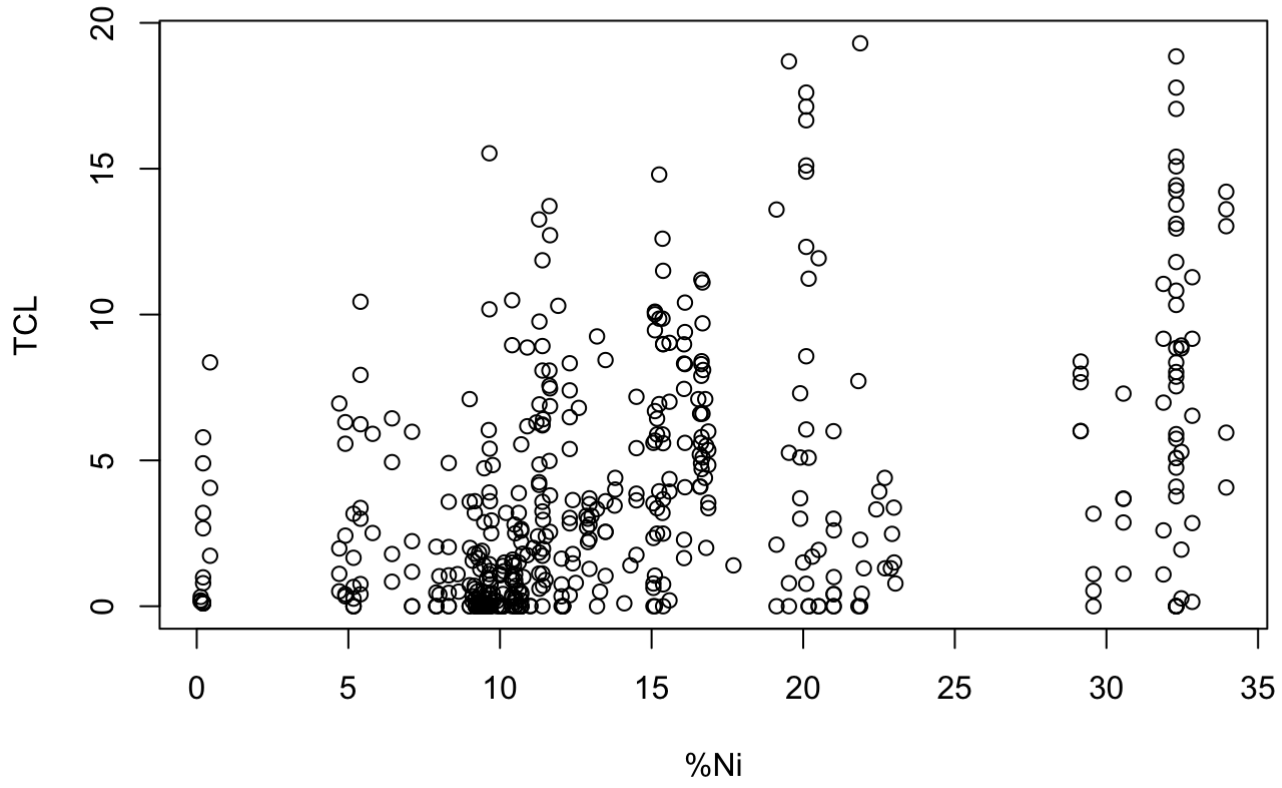
```
plot(al$Mg, al$MCL, xlab = "% Mg", ylab = "MCL")
```




```
plot(stl$Ni, stl$Strain , xlab = "% Ni", ylab = "Strain")
```



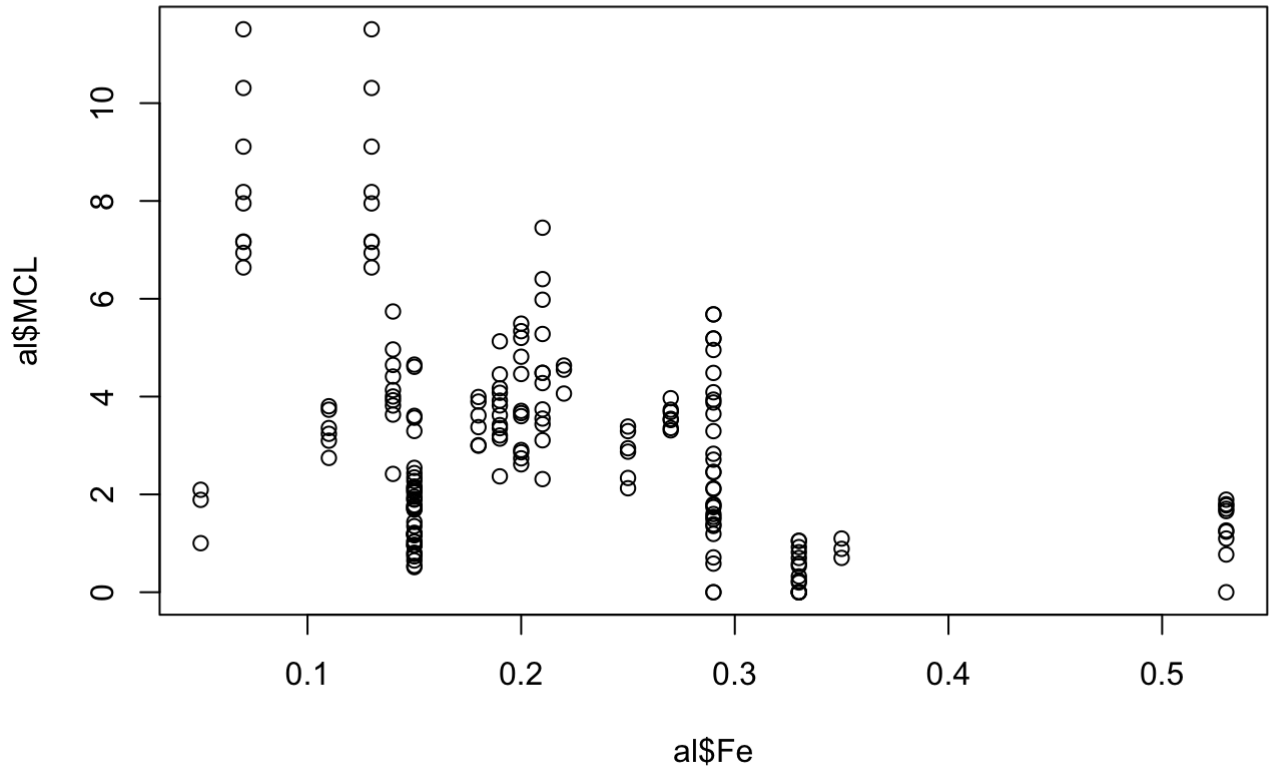

```
plot(stl$Ni, stl$TCL, xlab = "%Ni", ylab = "TCL")
```




```
xlab("%Ni")
```

```
## $x
## [1] "%Ni"
##
## attr(,"class")
## [1] "labels"
```

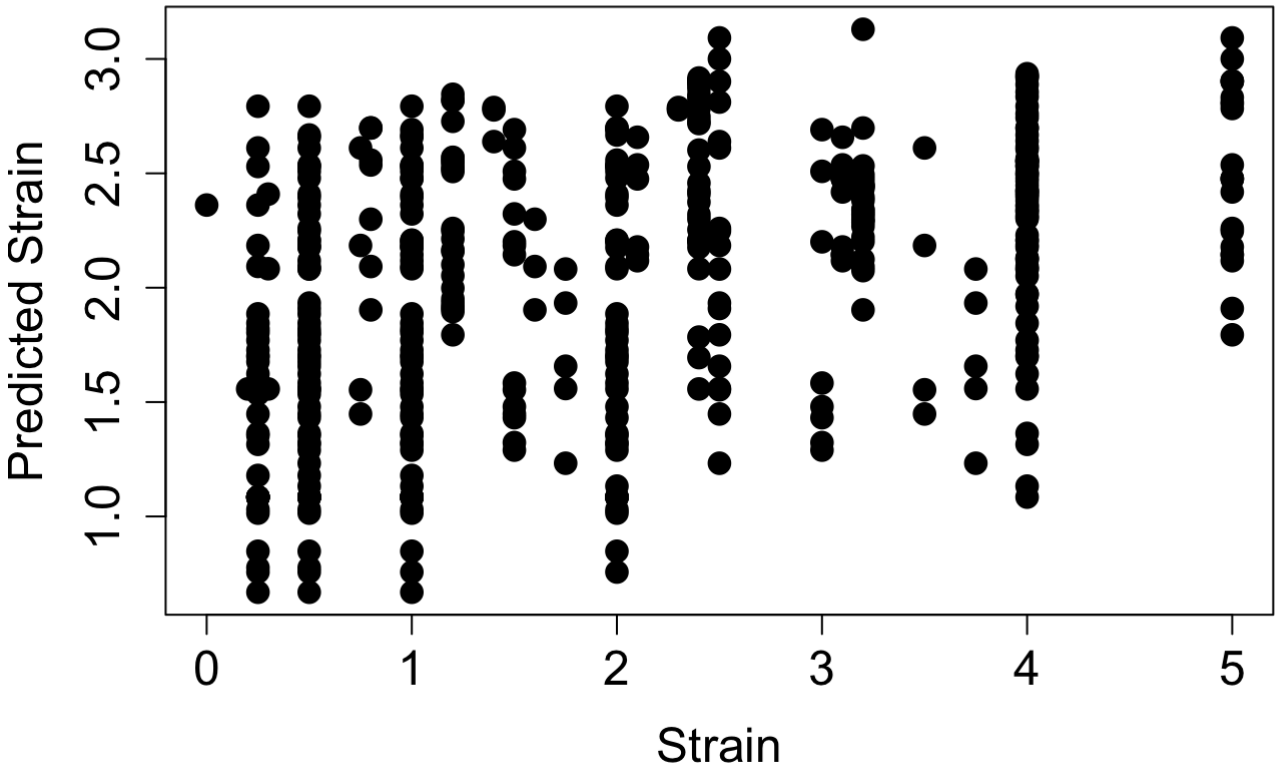
```
plot(al$Fe, al$MCL)
```




```
#Multiple Linear Regression Observing Strain for Steel Dataset
LMstl = lm(Strain ~ C + Si + Mn + P + S + Cr + Ni + Mo + N + Nb + Co + Cu + Al + Ti + V
+ B, data = stl)
summary(LMstl)
```

```
##
## Call:
## lm(formula = Strain ~ C + Si + Mn + P + S + Cr + Ni + Mo + N +
##     Nb + Co + Cu + Al + Ti + V + B, data = stl)
##
## Residuals:
##     Min      1Q  Median      3Q     Max
## -2.5433 -1.0243 -0.2570  0.9557  3.2059
##
## Coefficients:
##              Estimate Std. Error t value Pr(>|t|)
## (Intercept)   3.87275    0.60176   6.436 3.04e-10 ***
## C            -1.19395    0.91073  -1.311   0.1905
## Si           -0.32091    0.15590  -2.058   0.0401 *
## Mn            0.03211    0.08126   0.395   0.6929
## P            -7.40409    7.51020  -0.986   0.3247
## S           -31.55089   10.96724  -2.877   0.0042 **
## Cr           -0.04012    0.02877  -1.394   0.1638
## Ni           -0.02603    0.01140  -2.283   0.0229 *
## Mo           -0.05497    0.06677  -0.823   0.4108
## N             1.63127    1.43070   1.140   0.2548
## Nb            0.29512    0.39216   0.753   0.4521
## Co            1.92190    1.47413   1.304   0.1930
## Cu           -0.17698    0.12385  -1.429   0.1537
## Al           -1.15149    0.49299  -2.336   0.0199 *
## Ti           -0.20741    0.42841  -0.484   0.6285
## V            -1.23342    0.50128  -2.461   0.0142 *
## B            80.31391   49.08260   1.636   0.1024
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Residual standard error: 1.305 on 470 degrees of freedom
## Multiple R-squared:  0.1675, Adjusted R-squared:  0.1391
## F-statistic:  5.91 on 16 and 470 DF,  p-value: 7.467e-12
```

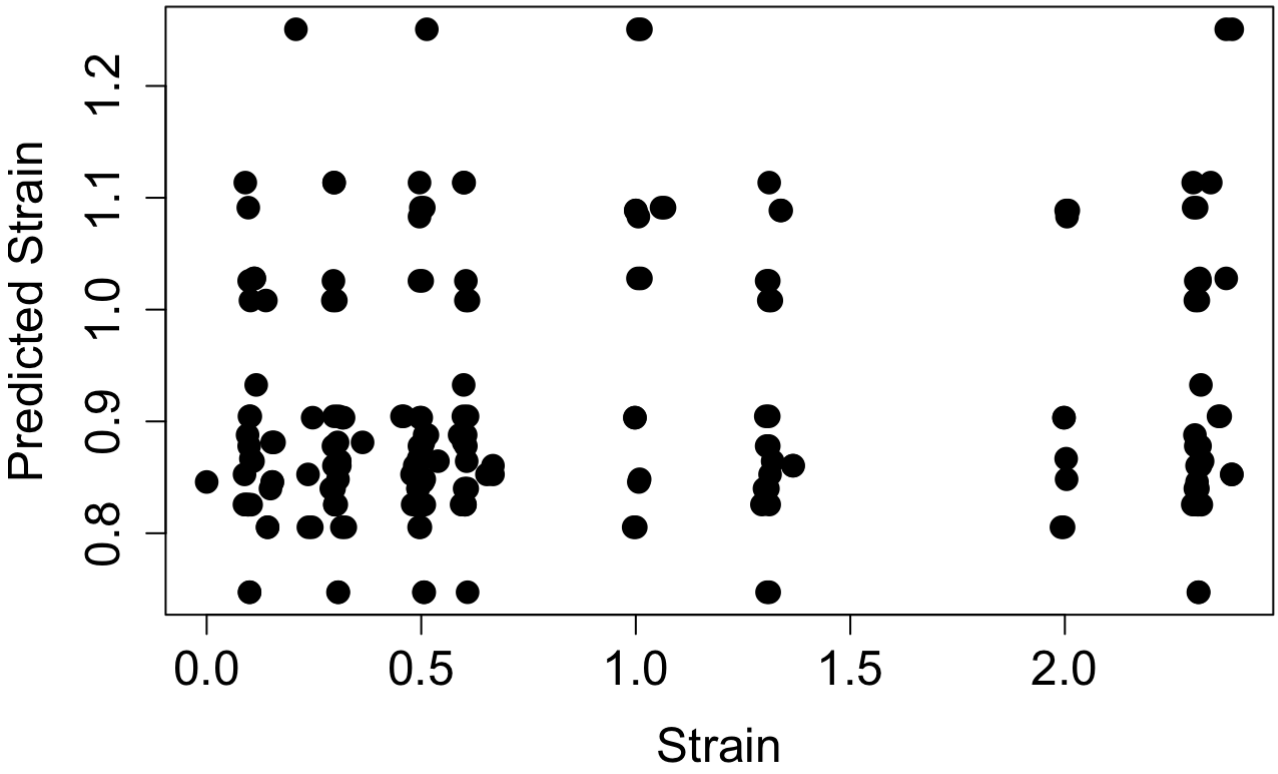
```
preLMmstl=predict(LMstl)
plot(stl$Strain, predict(LMstl), xlab = "Strain",
     ylab = "Predicted Strain", cex.axis = 1.5,
     cex.lab = 1.5, cex = 1.5, pch = 19)
```




```
#MLR Analysis for Strain using the Aluminum Dataset
LMal = lm(Strain ~  Si + Mn + Fe + Mg + Cr + Mn + Zn + Cu + Al + Ti + V + Zr + B + Li, d
ata = al)
summary(LMal)
```

```
##
## Call:
## lm(formula = Strain ~ Si + Mn + Fe + Mg + Cr + Mn + Zn + Cu +
##     Al + Ti + V + Zr + B + Li, data = al)
##
## Residuals:
##     Min      1Q  Median      3Q     Max
## -1.0427 -0.5723 -0.3104  0.4602  1.5648
##
## Coefficients: (2 not defined because of singularities)
##               Estimate Std. Error t value Pr(>|t|)
## (Intercept)  -0.517734  22.830605  -0.023    0.982
## Si            1.093791   2.072920   0.528    0.598
## Mn           -0.153049   0.511311  -0.299    0.765
## Fe            0.066153   0.767541   0.086    0.931
## Mg            0.004737   0.233416   0.020    0.984
## Cr           -0.715096   1.072397  -0.667    0.506
## Zn            0.791344   3.725710   0.212    0.832
## Cu            0.018793   0.246490   0.076    0.939
## Al            0.013137   0.228959   0.057    0.954
## Ti            4.164781   6.154100   0.677    0.499
## V                   NA         NA      NA       NA
## Zr            1.329972   3.366489   0.395    0.693
## B           -32.679515 215.865829  -0.151    0.880
## Li                  NA         NA      NA       NA
##
## Residual standard error: 0.7917 on 183 degrees of freedom
## Multiple R-squared:  0.02249,    Adjusted R-squared:  -0.03627
## F-statistic: 0.3827 on 11 and 183 DF,  p-value: 0.9616
```

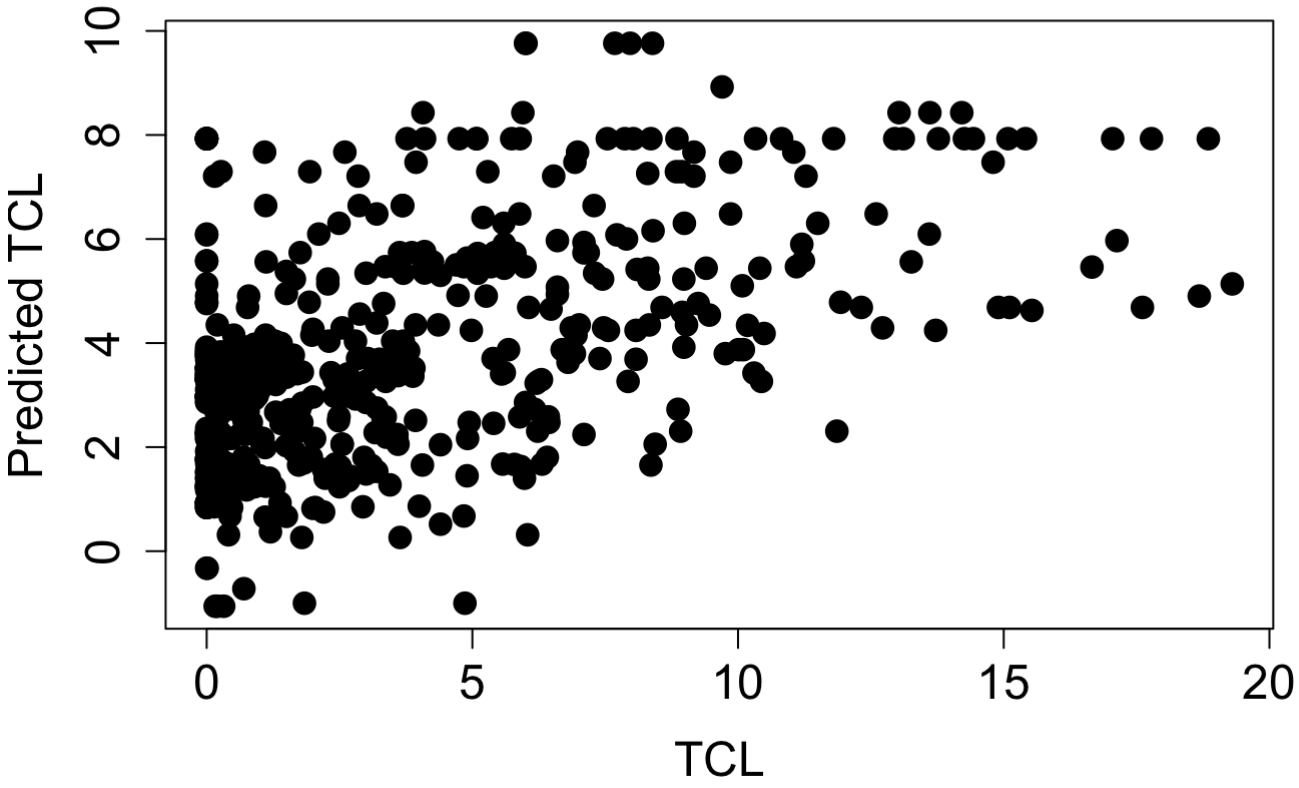
```
preLMal=predict(LMal)
plot(al$Strain, predict(LMal), xlab = "Strain",
     ylab = "Predicted Strain", cex.axis = 1.5,
     cex.lab = 1.5, cex = 1.5, pch = 19)
```




```
#Using MLR to predict Total Crack Length in the Steel Dataset
LMstl_C = lm(TCL ~ C + Si + Mn + P + S + Cr + Ni + Mo + N + Nb + Co + Cu + Al + Ti + V +
B, data = stl)
summary(LMstl_C)
```

```
##
## Call:
## lm(formula = TCL ~ C + Si + Mn + P + S + Cr + Ni + Mo + N + Nb +
##     Co + Cu + Al + Ti + V + B, data = stl)
##
## Residuals:
##     Min      1Q  Median      3Q     Max
## -7.9291 -2.3798 -0.7128  1.6535 14.1621
##
## Coefficients:
##              Estimate Std. Error t value Pr(>|t|)
## (Intercept)   2.33857    1.66958   1.401 0.161965
## C            -1.14612    2.52681  -0.454 0.650338
## Si            1.17301    0.43255   2.712 0.006936 **
## Mn           -0.84531    0.22544  -3.750 0.000199 ***
## P            74.07559   20.83690   3.555 0.000416 ***
## S           -75.77125   30.42840  -2.490 0.013113 *
## Cr           -0.13295    0.07982  -1.666 0.096449 .
## Ni            0.23756    0.03163   7.512 2.97e-13 ***
## Mo           -0.28271    0.18525  -1.526 0.127662
## N             9.23596    3.96945   2.327 0.020402 *
## Nb           -1.00382    1.08804  -0.923 0.356692
## Co          -11.31785    4.08994  -2.767 0.005876 **
## Cu            0.57644    0.34362   1.678 0.094102 .
## Al           -1.09992    1.36779  -0.804 0.421711
## Ti            0.51171    1.18862   0.431 0.667023
## V             1.27378    1.39080   0.916 0.360210
## B           350.08885  136.17870   2.571 0.010453 *
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Residual standard error: 3.621 on 470 degrees of freedom
## Multiple R-squared:  0.2772, Adjusted R-squared:  0.2526
## F-statistic: 11.26 on 16 and 470 DF,  p-value: < 2.2e-16
```

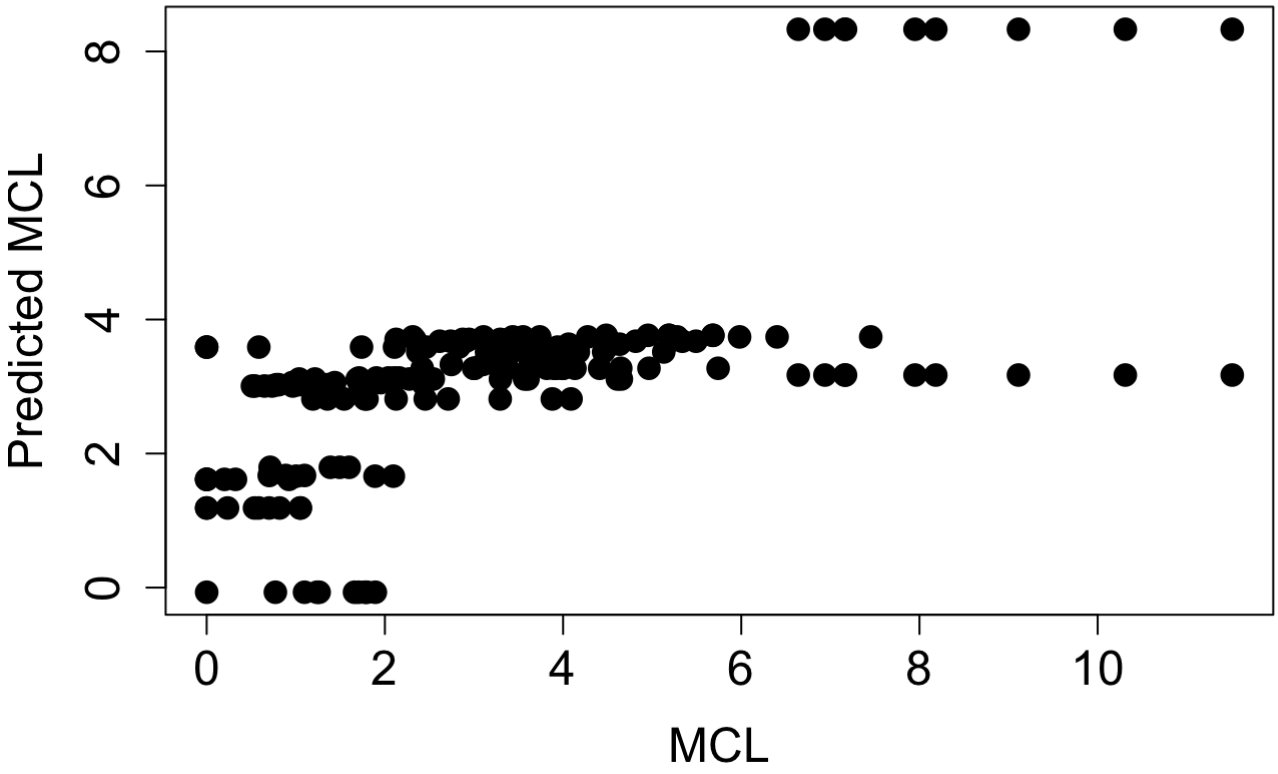
```
preLMstl_C=predict(LMstl_C)
plot(stl$TCL, predict(LMstl_C), xlab = "TCL",
     ylab = "Predicted TCL", cex.axis = 1.5,
     cex.lab = 1.5, cex = 1.5, pch = 19)
```



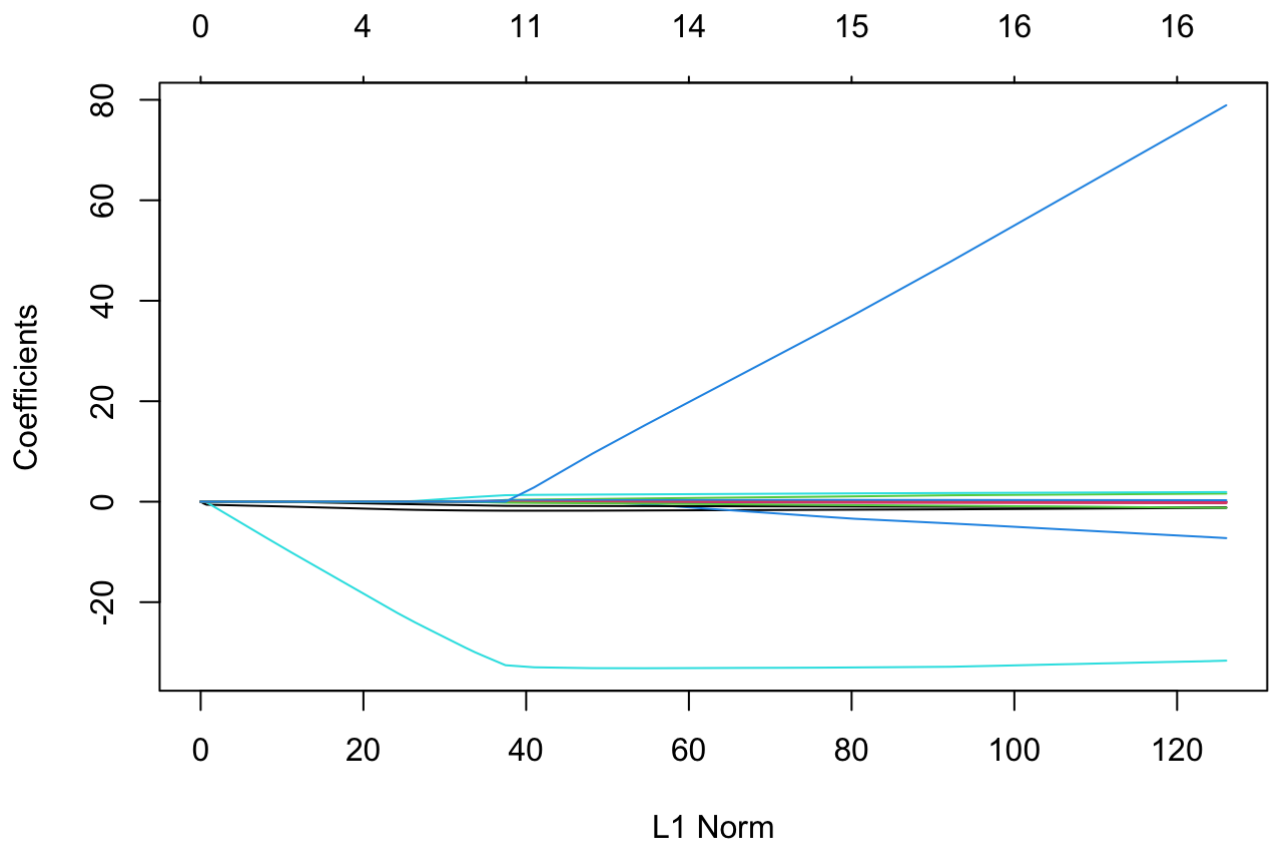

```
#Using MLR to predict Maximum crack Length in Aluminum Dataset
LMal_C = lm(MCL ~  Si + Mn + Fe + Mg + Cr + Mn + Zn + Cu + Al + Ti + V + Zr + B + Li, da
ta = al)
summary(LMal_C)
```

```
##
## Call:
## lm(formula = MCL ~ Si + Mn + Fe + Mg + Cr + Mn + Zn + Cu + Al +
##     Ti + V + Zr + B + Li, data = al)
##
## Residuals:
##     Min      1Q  Median      3Q     Max
## -3.5886 -1.1537 -0.2303  0.7453  8.3386
##
## Coefficients: (2 not defined because of singularities)
##              Estimate Std. Error t value Pr(>|t|)
## (Intercept) -130.3028    51.3230  -2.539  0.01195 *
## Si             2.8753     4.6599   0.617  0.53799
## Mn             1.4725     1.1494   1.281  0.20178
## Fe            -7.2942     1.7254  -4.227 3.73e-05 ***
## Mg             1.3854     0.5247   2.640  0.00900 **
## Cr             6.4416     2.4107   2.672  0.00822 **
## Zn           -12.4408     8.3754  -1.485  0.13916
## Cu             1.6765     0.5541   3.026  0.00284 **
## Al             1.3449     0.5147   2.613  0.00972 **
## Ti             7.1976    13.8344   0.520  0.60351
## V                  NA         NA      NA       NA
## Zr           -11.9012     7.5678  -1.573  0.11754
## B            934.4192   485.2644   1.926  0.05571 .
## Li                 NA         NA      NA       NA
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Residual standard error: 1.78 on 183 degrees of freedom
## Multiple R-squared:  0.4277, Adjusted R-squared:  0.3933
## F-statistic: 12.43 on 11 and 183 DF,  p-value: < 2.2e-16
```

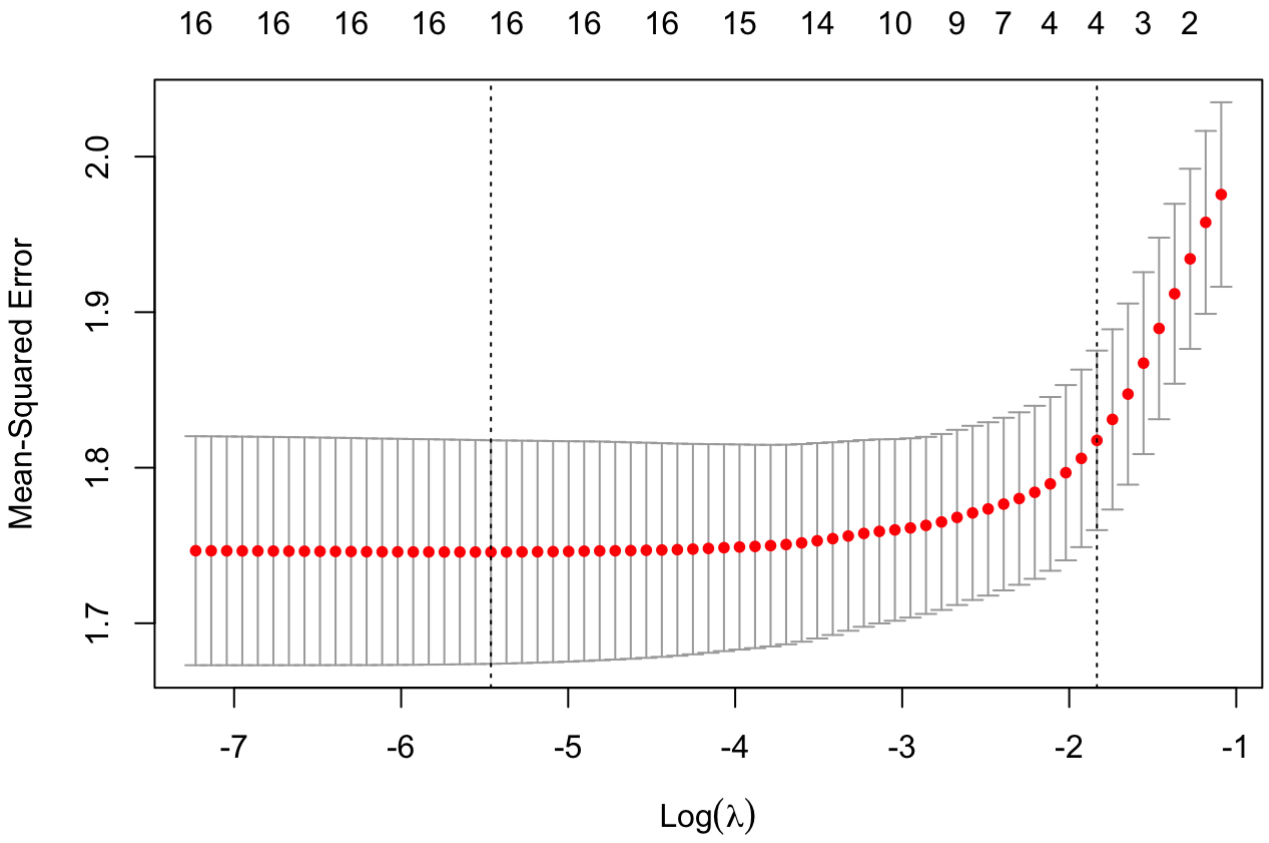
```
preLMal_C=predict(LMal_C)
plot(al$MCL, predict(LMal_C), xlab = "MCL",
     ylab = "Predicted MCL", cex.axis = 1.5,
     cex.lab = 1.5, cex = 1.5, pch = 19)
```




```
#Feature Selection and PCA to Predict Strain Values
#Steel Strain Model
stEl <- as.matrix(stEl)
stSTN <- as.matrix(stSTN)
COMstl <- glmnet(stEl ,stSTN)
plot(COMstl)
```




```
COMfit_stSTN = cv.glmnet(stEl, stSTN, alpha = 1)
plot(COMfit_stSTN)
```




```
COMfit_stSTN$lambda.min
```

```
## [1] 0.004242854
```

```
COMfit_stSTN$lambda.1se
```

```
## [1] 0.1597407
```

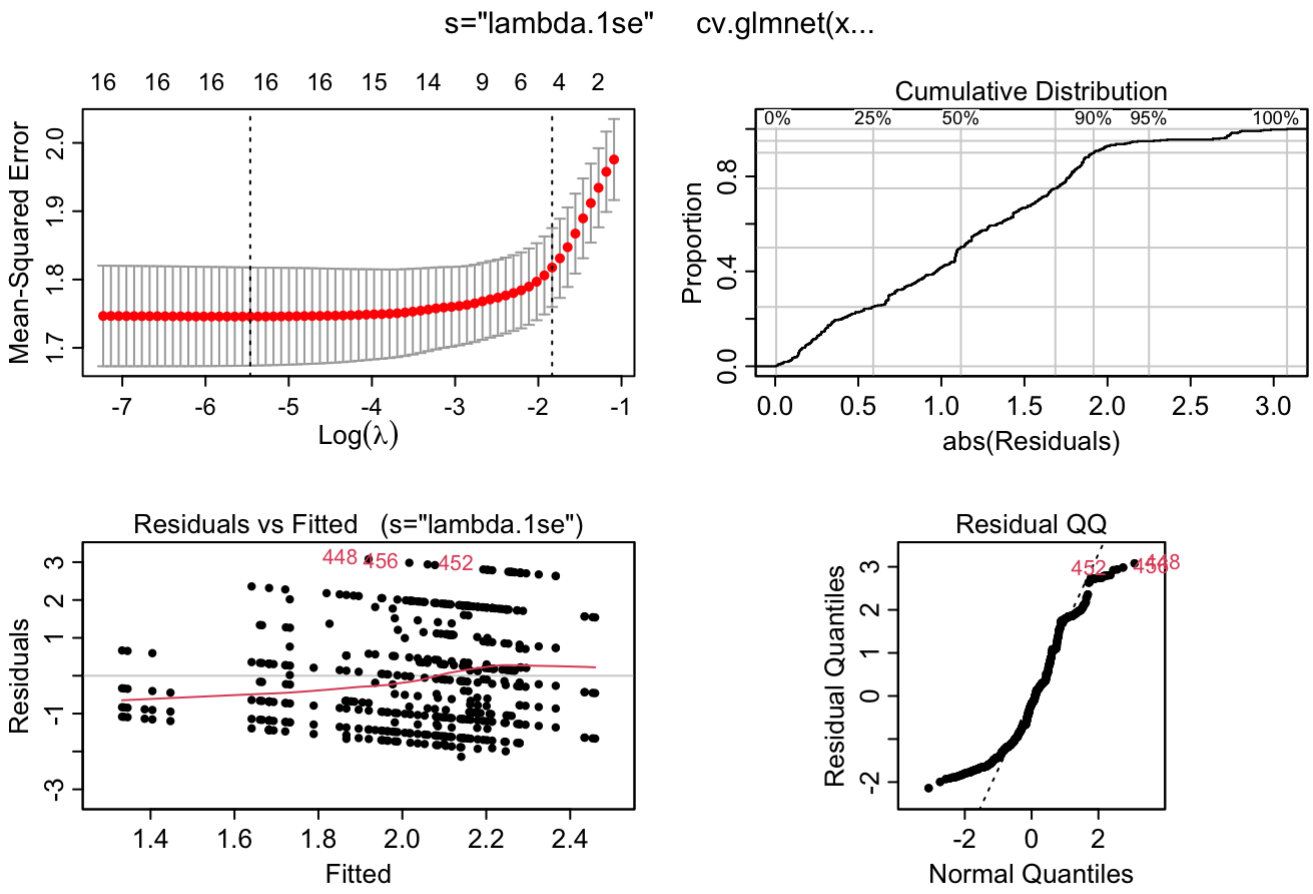
```
library(plotmo)
```

```
## Loading required package: Formula
```

```
## Loading required package: plotrix
```

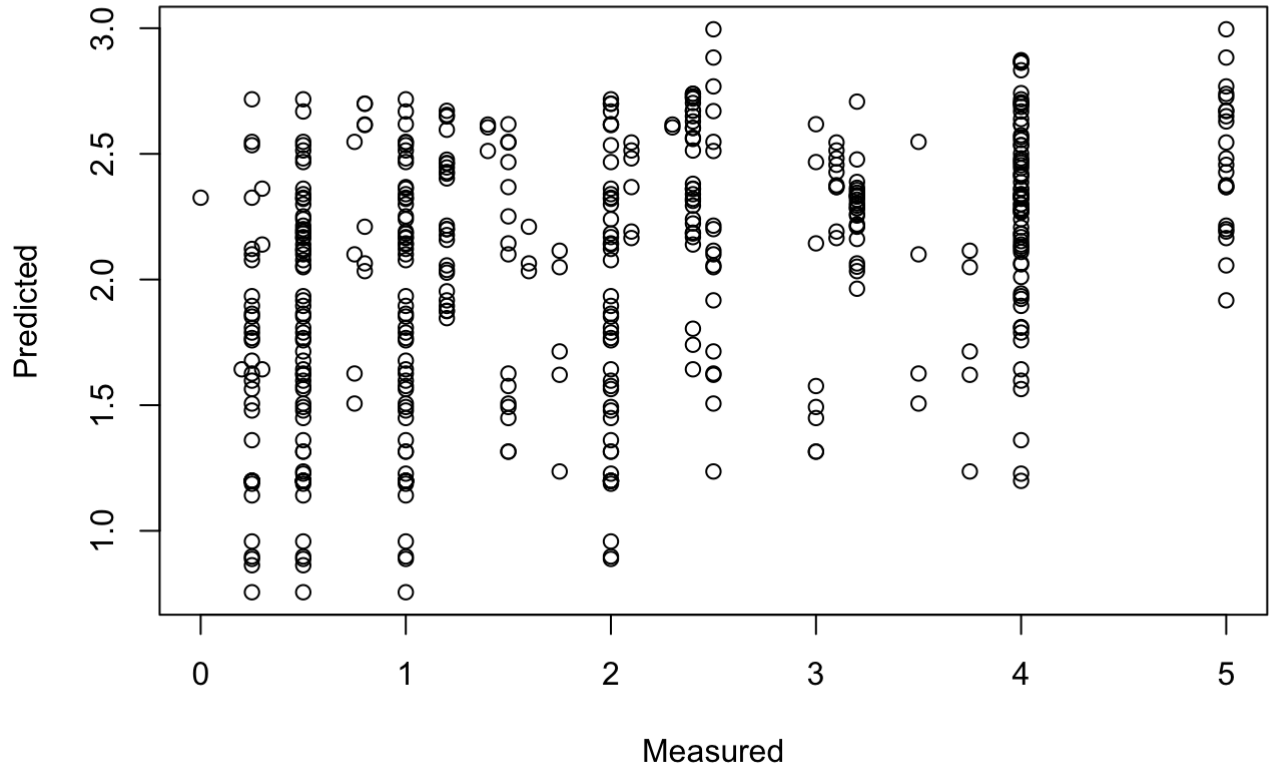
```
## Loading required package: TeachingDemos
```

```
plotres(COMfit_stSTN)
```

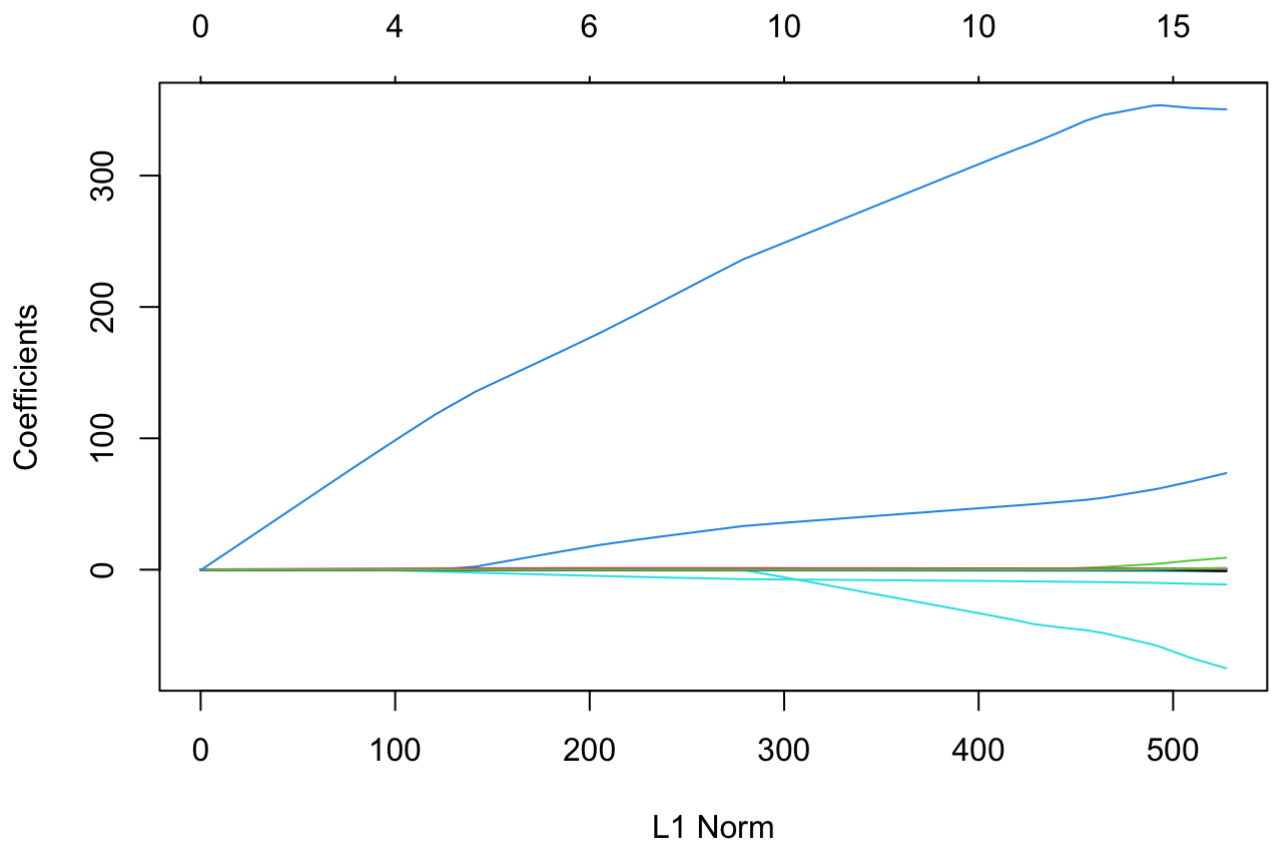



```
pred_stl <- predict(COMstl, stEl, s = .02)
```

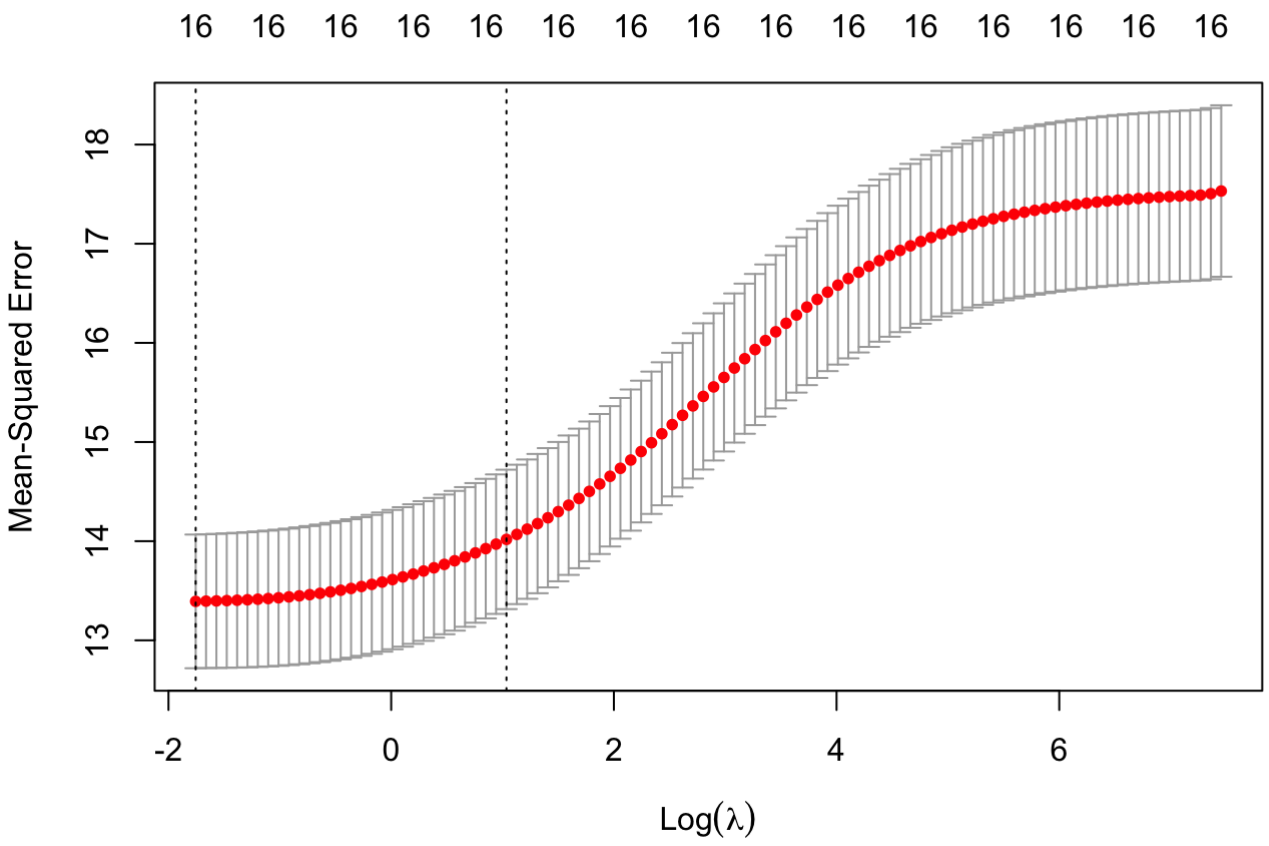
```
stl_t <- pred_stl[,c("s1")]
stl_s <- as.numeric(stSTN)
plot(stl_s,stl_t, xlab="Measured", ylab = "Predicted")
```




```
#Steel TCL Model
stEl <- as.matrix(stEl)
stTCL <- as.matrix(stTCL)
COMstl2 <- glmnet(stEl ,stTCL)
plot(COMstl2)
```




```
COMfit_stTCL = cv.glmnet(stEl, stTCL, alpha = 0)
plot(COMfit_stTCL)
```



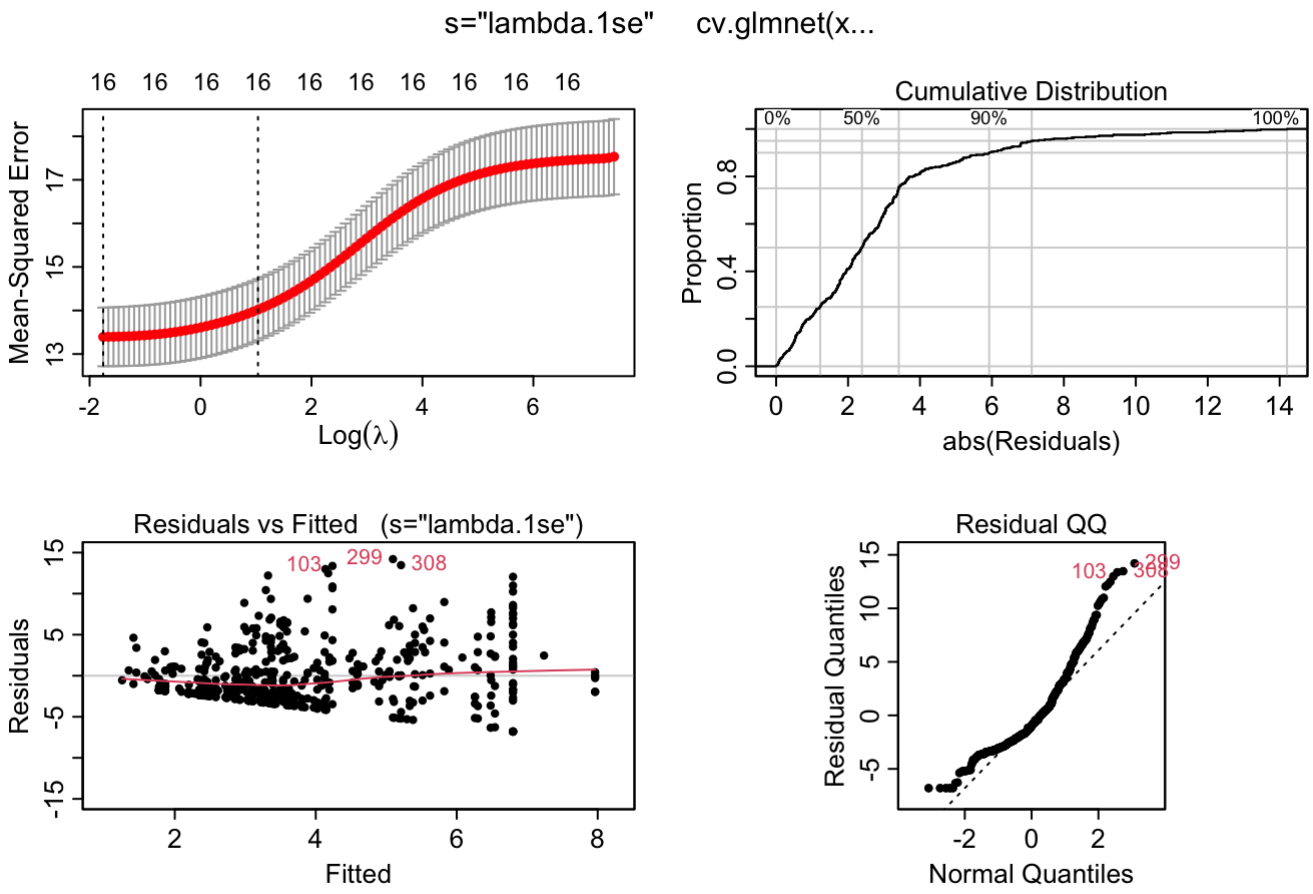

```
COMfit_stTCL$lambda.min
```

```
## [1] 0.1727779
```

```
COMfit_stTCL$lambda.1se
```

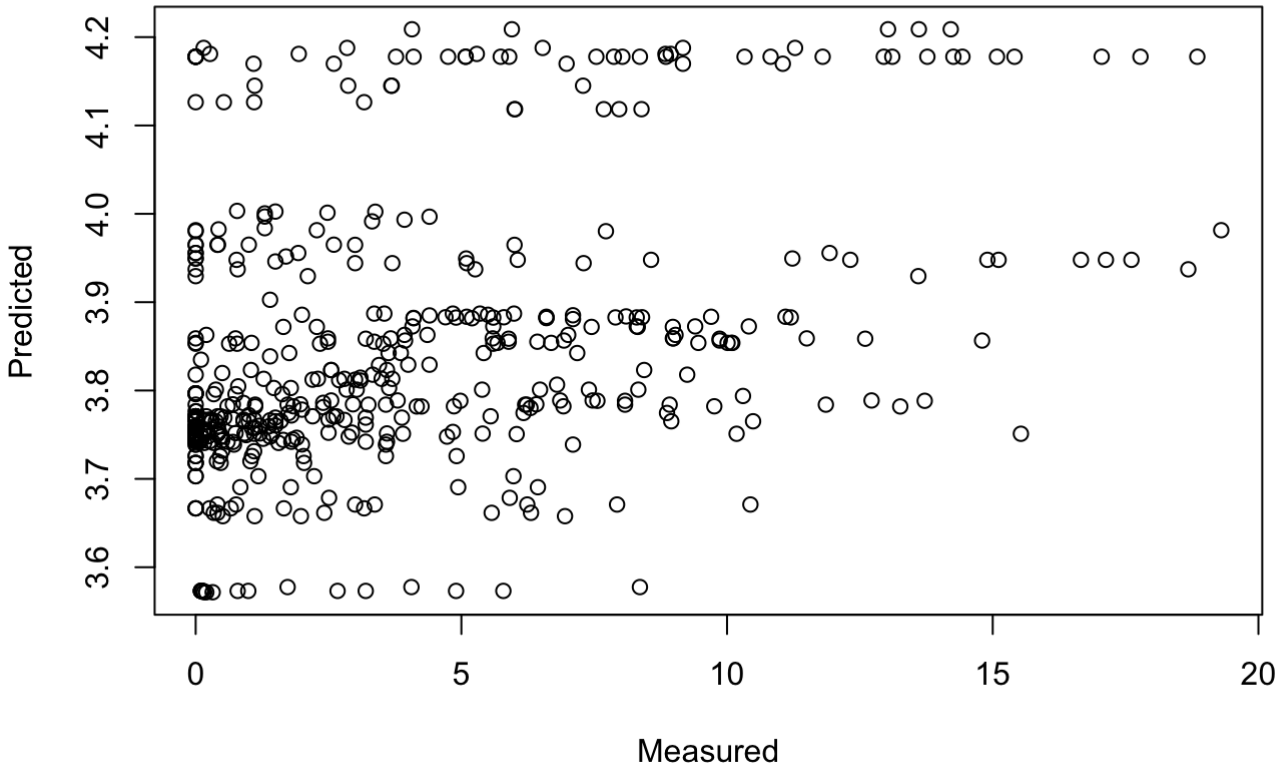
```
## [1] 2.815849
```

```
plotres(COMfit_stTCL)
```

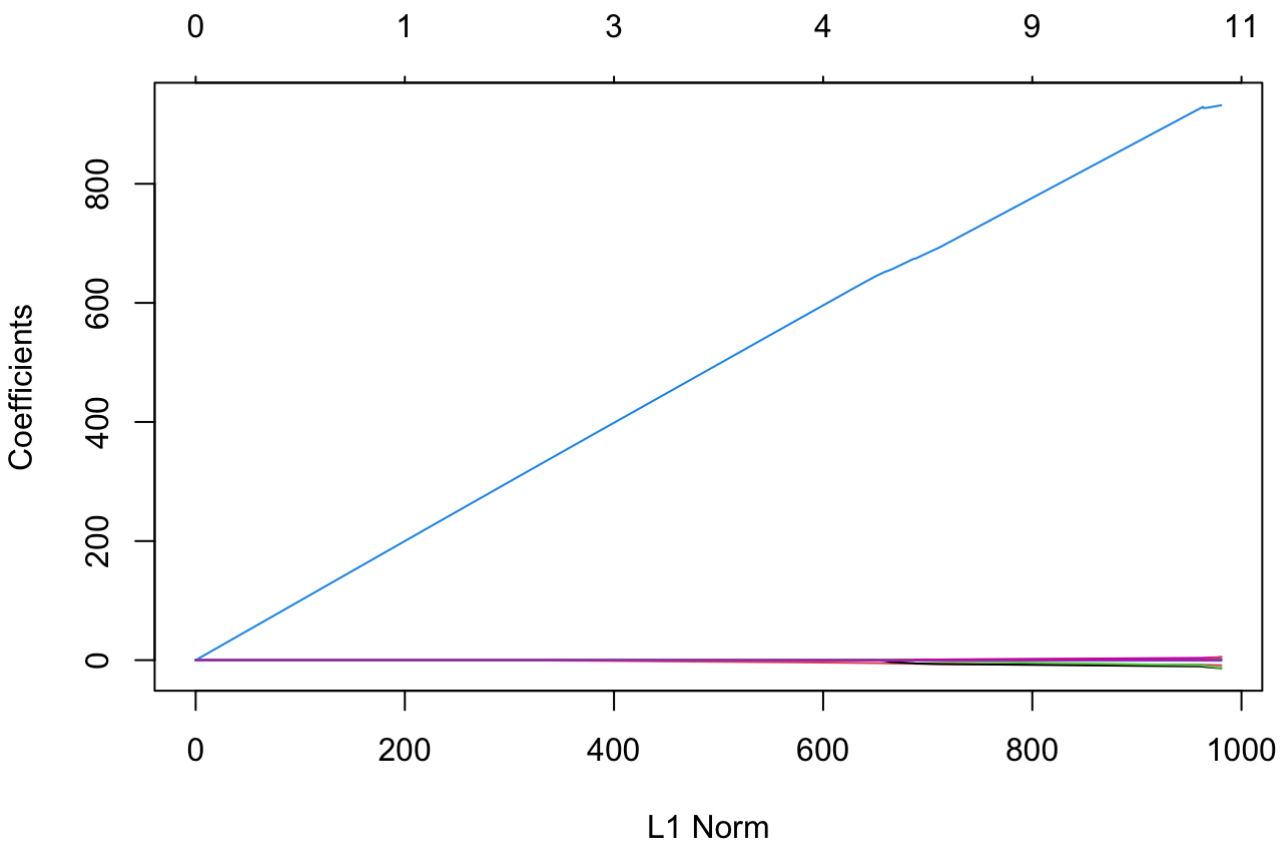



```
pred_stl2 <- predict(COMstl2, stEl)
```

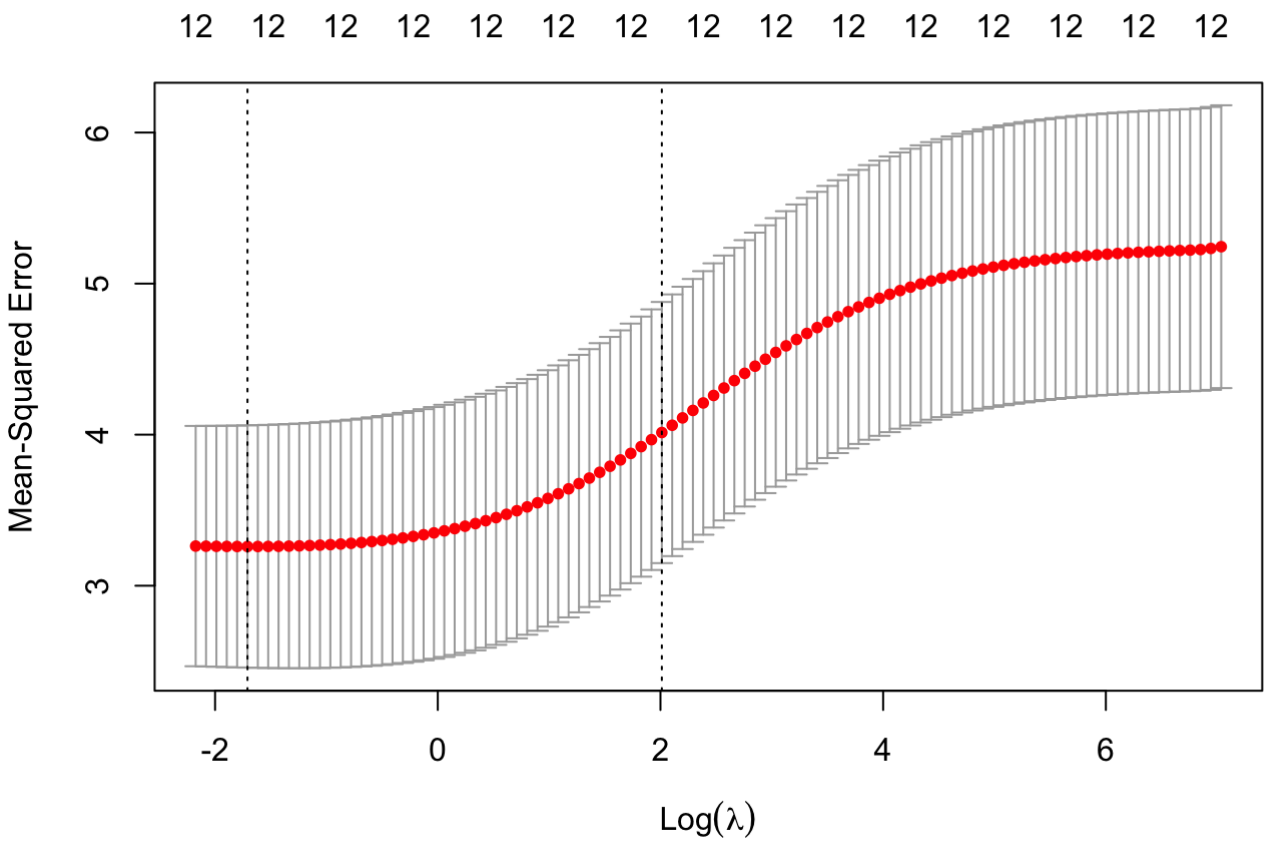
```
stl_t2 <- pred_stl2[,c("s1")]
stl_s2 <- as.numeric(stTCL)
plot(stl_s2,stl_t2, xlab="Measured", ylab = "Predicted")
```




```
alEl <- as.matrix(alEl)
alMCL <- as.matrix(alMCL)
COMal <- glmnet(alEl ,alMCL)
plot(COMal)
```




```
COMfit_al = cv.glmnet(alEl, alMCL, alpha = 0)
plot(COMfit_al)
```




```
COMfit_al$lambda.min
```

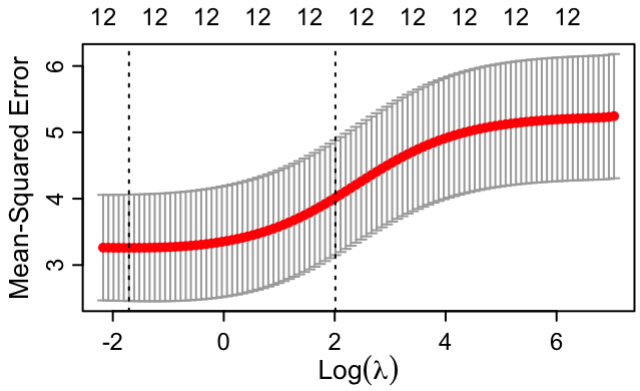
```
## [1] 0.1810444
```

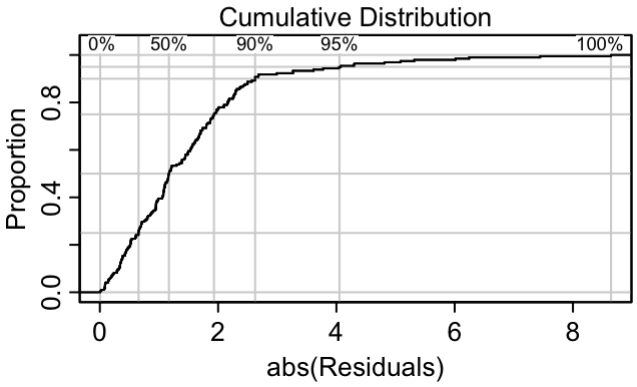
```
COMfit_al$lambda.1se
```

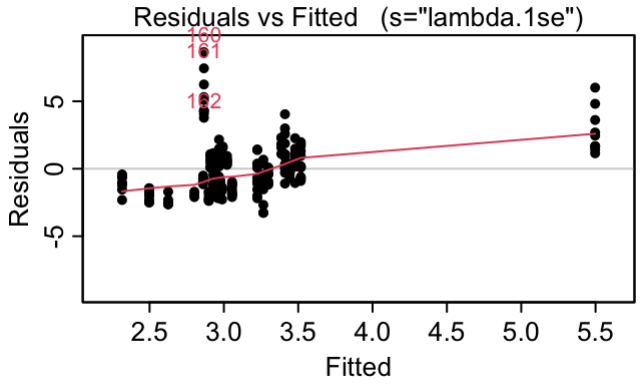
```
## [1] 7.480779
```

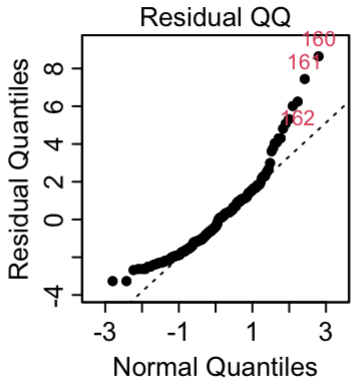
```
plotres(COMfit_al)
```

s="lambda.1se" cv.glmnet(x...



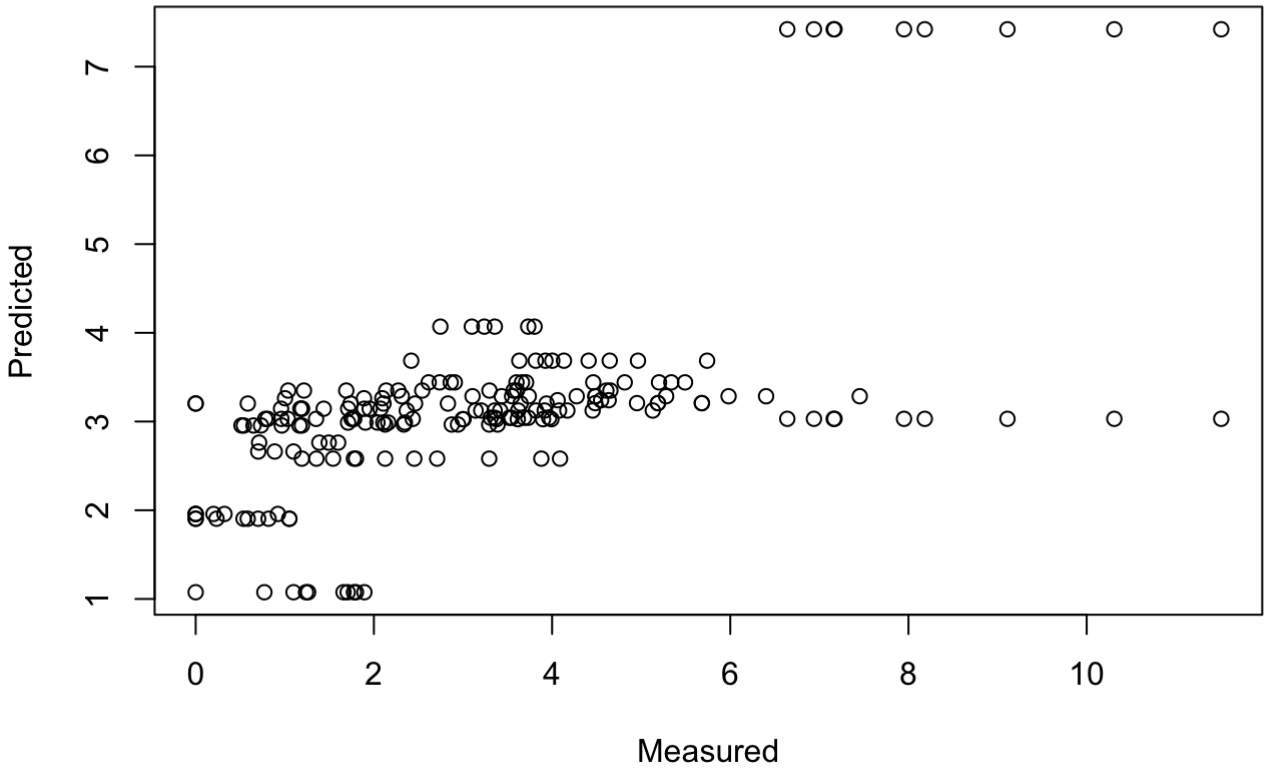







```
pred_al <- predict(COMal, alEl, s = .2)
```

```
al_t2 <- pred_al[,c("s1")]
al_s2 <- as.numeric(alMCL)
plot(al_s2,al_t2, xlab="Measured", ylab = "Predicted")
```




```
#defining dataframes that can be used for PCA
alrf_MCL <- al[c(1:13,19)]
alrf_STN <- al[c(1:13,18)]
stlrf_TCL <- stl[c(1:16,22)]
stlrf_STN <- stl[c(1:16,21)]
```

```
#Using PCA to idenifty sources of varience and hopefully insights into model simplificat
ions
Si <-as.numeric(alrf_MCL$Si)
Fe <-as.numeric(alrf_MCL$Fe)
Cu <-as.numeric(alrf_MCL$Cu)
Mn <-as.numeric(alrf_MCL$Mn)
Mg <-as.numeric(alrf_MCL$Mg)
Cr <-as.numeric(alrf_MCL$Cr)
Zn <-as.numeric(alrf_MCL$Zn)
Ti <-as.numeric(alrf_MCL$Ti)
Zr <-as.numeric(alrf_MCL$Zr)
V <-as.numeric(alrf_MCL$V)
B <-as.numeric(alrf_MCL$B)
Li <-as.numeric(alrf_MCL$Li)
Al <-as.numeric(alrf_MCL$Al)
MCL <-as.numeric(alrf_MCL$MCL)
alSTN <- as.numeric(alrf_STN$Strain)
```

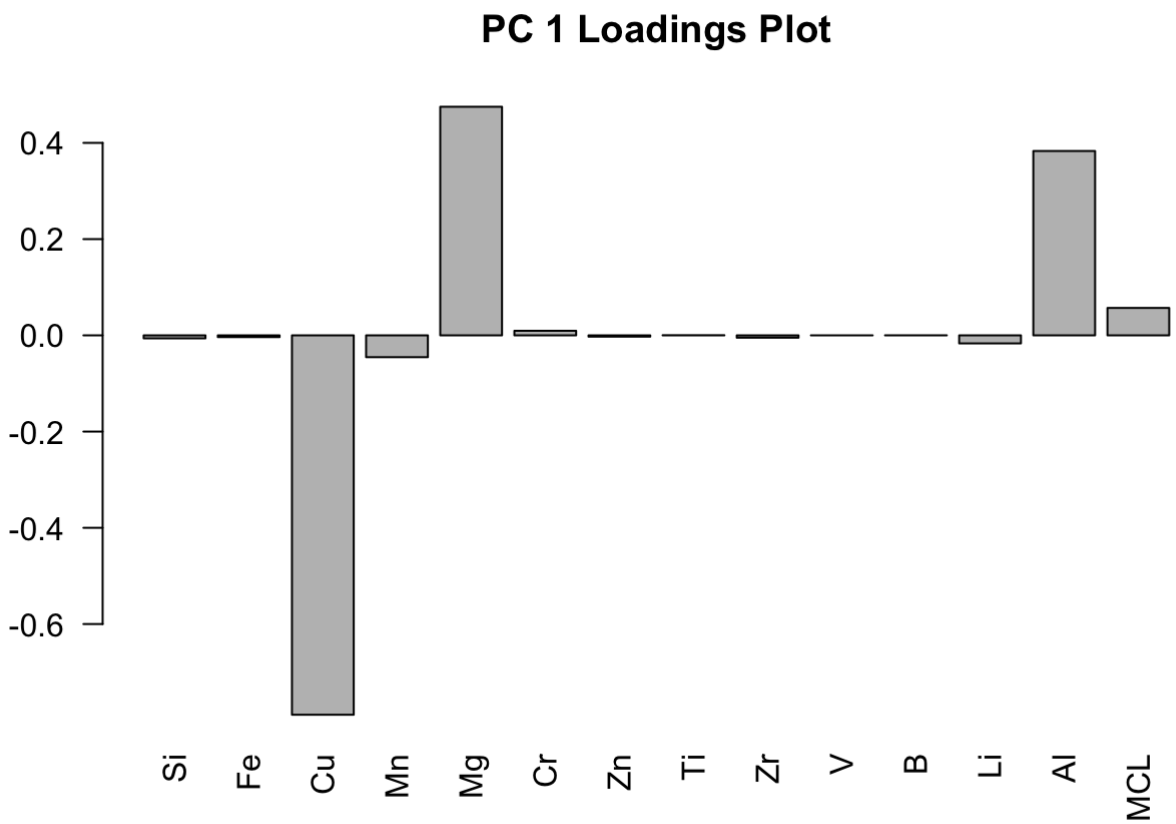
```
#PCA for MCL in Al
pralT <- cbind(Si,Fe,Cu,Mn,Mg,Cr,Zn,Ti,Zr,V,B,Li,Al,MCL)
```

```
pralPCA = prcomp(pralT, scale=F)
```

```
summary(pralPCA)
```

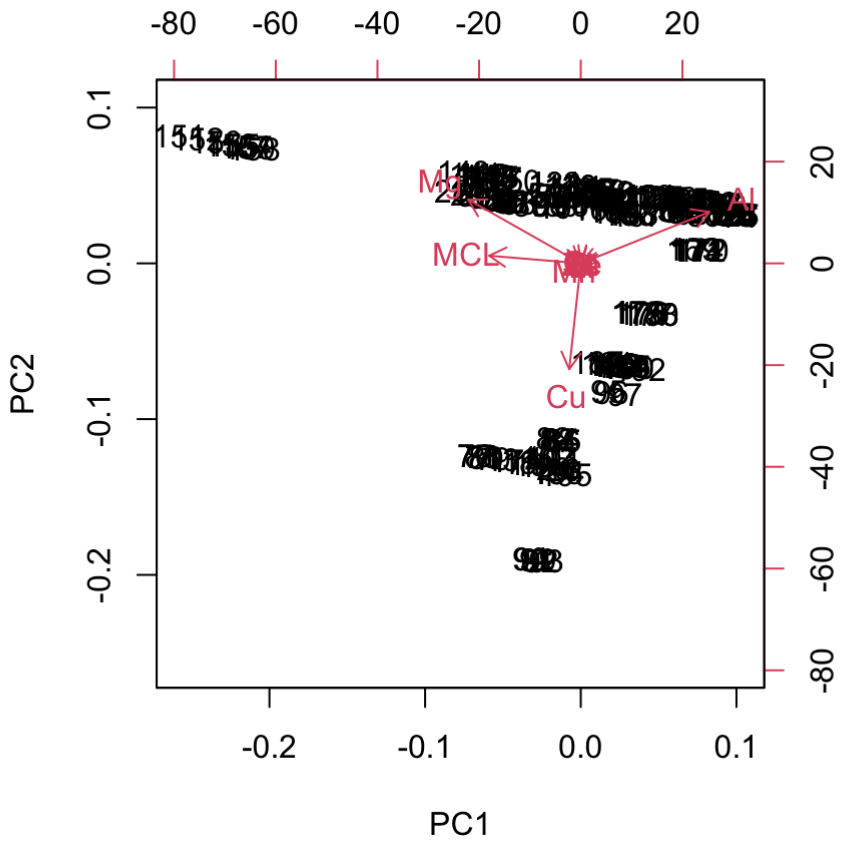
```
## Importance of components:
##                           PC1    PC2    PC3     PC4     PC5     PC6     PC7
## Standard deviation     3.4231 2.3611 1.8349 0.30080 0.25441 0.09686 0.06397
## Proportion of Variance 0.5625 0.2676 0.1616 0.00434 0.00311 0.00045 0.00020
## Cumulative Proportion  0.5625 0.8302 0.9918 0.99615 0.99926 0.99971 0.99991
##                           PC8     PC9    PC10    PC11      PC12     PC13
## Standard deviation     0.02839 0.02624 0.01585 0.01254 0.0002719 1.85e-15
## Proportion of Variance 0.00004 0.00003 0.00001 0.00001 0.0000000 0.00e+00
## Cumulative Proportion  0.99995 0.99998 0.99999 1.00000 1.0000000 1.00e+00
##                           PC14
## Standard deviation     5.653e-19
## Proportion of Variance 0.000e+00
## Cumulative Proportion  1.000e+00
```

```
barplot(pralPCA$rotation[,2], main="PC 1 Loadings Plot", las=2)
```

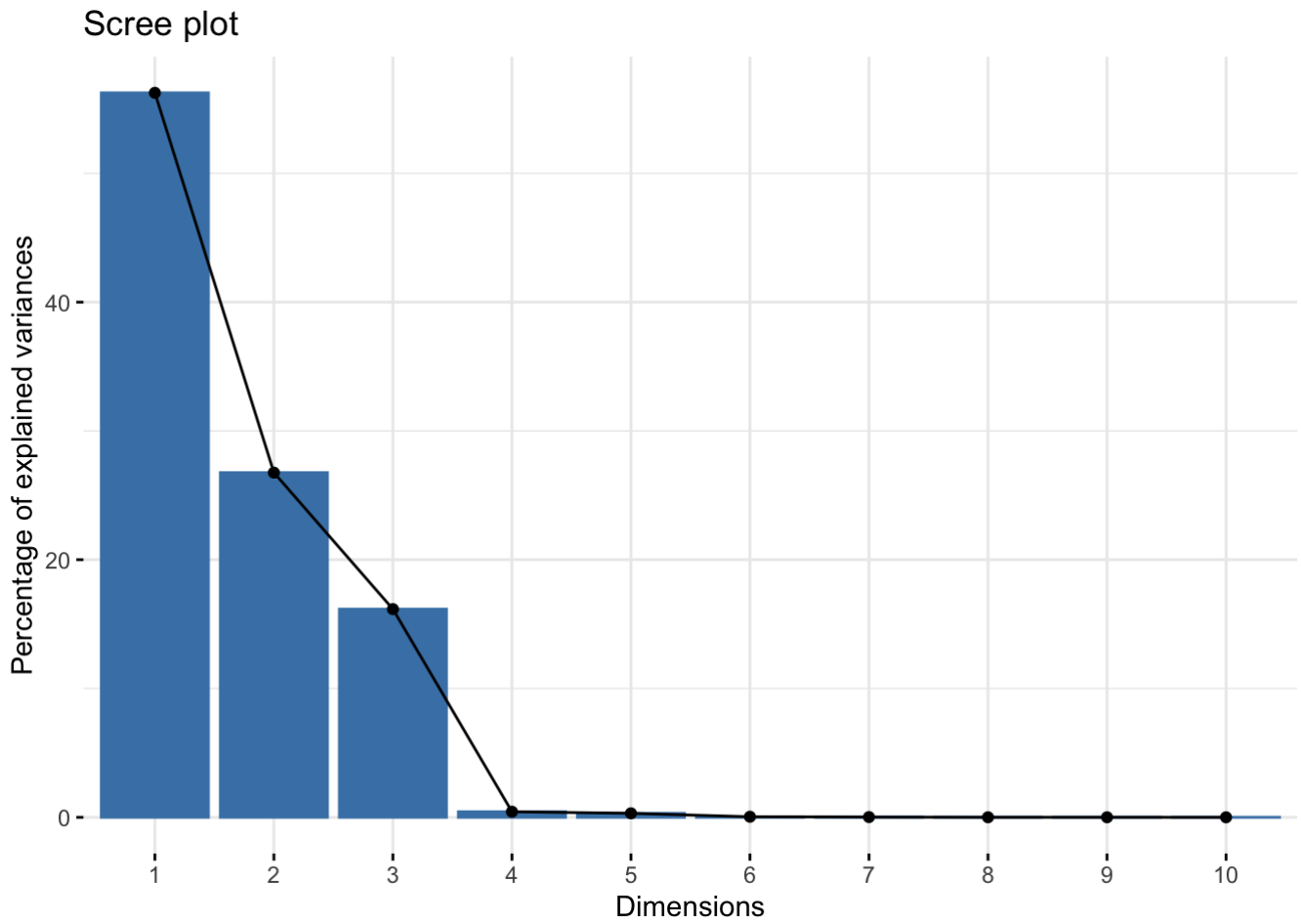


```
biplot(pralPCA)
```

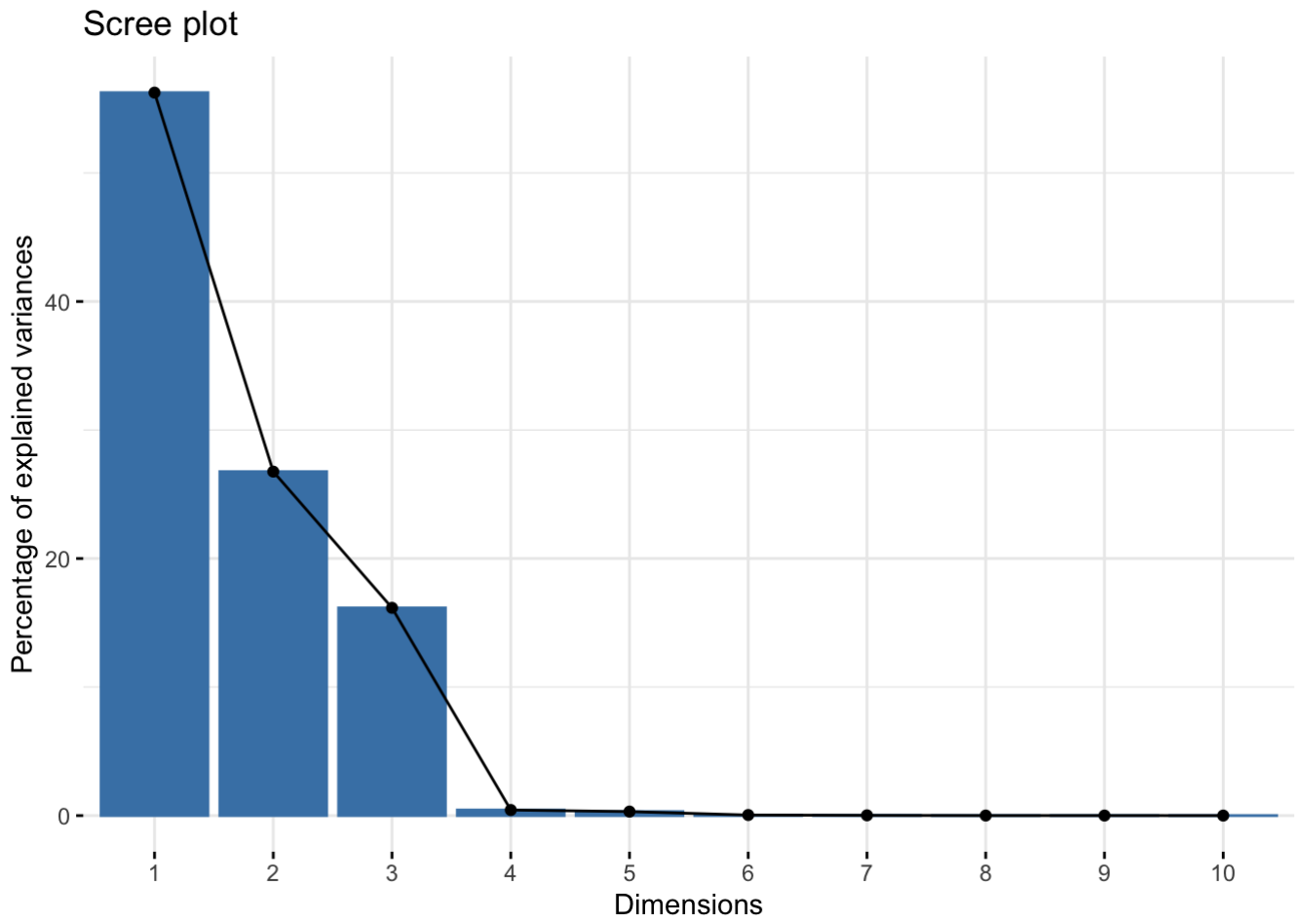
```
## Warning in arrows(0, 0, y[, 1L] * 0.8, y[, 2L] * 0.8, col = col[2L], length =
## arrow.len): zero-length arrow is of indeterminate angle and so skipped

## Warning in arrows(0, 0, y[, 1L] * 0.8, y[, 2L] * 0.8, col = col[2L], length =
## arrow.len): zero-length arrow is of indeterminate angle and so skipped
```



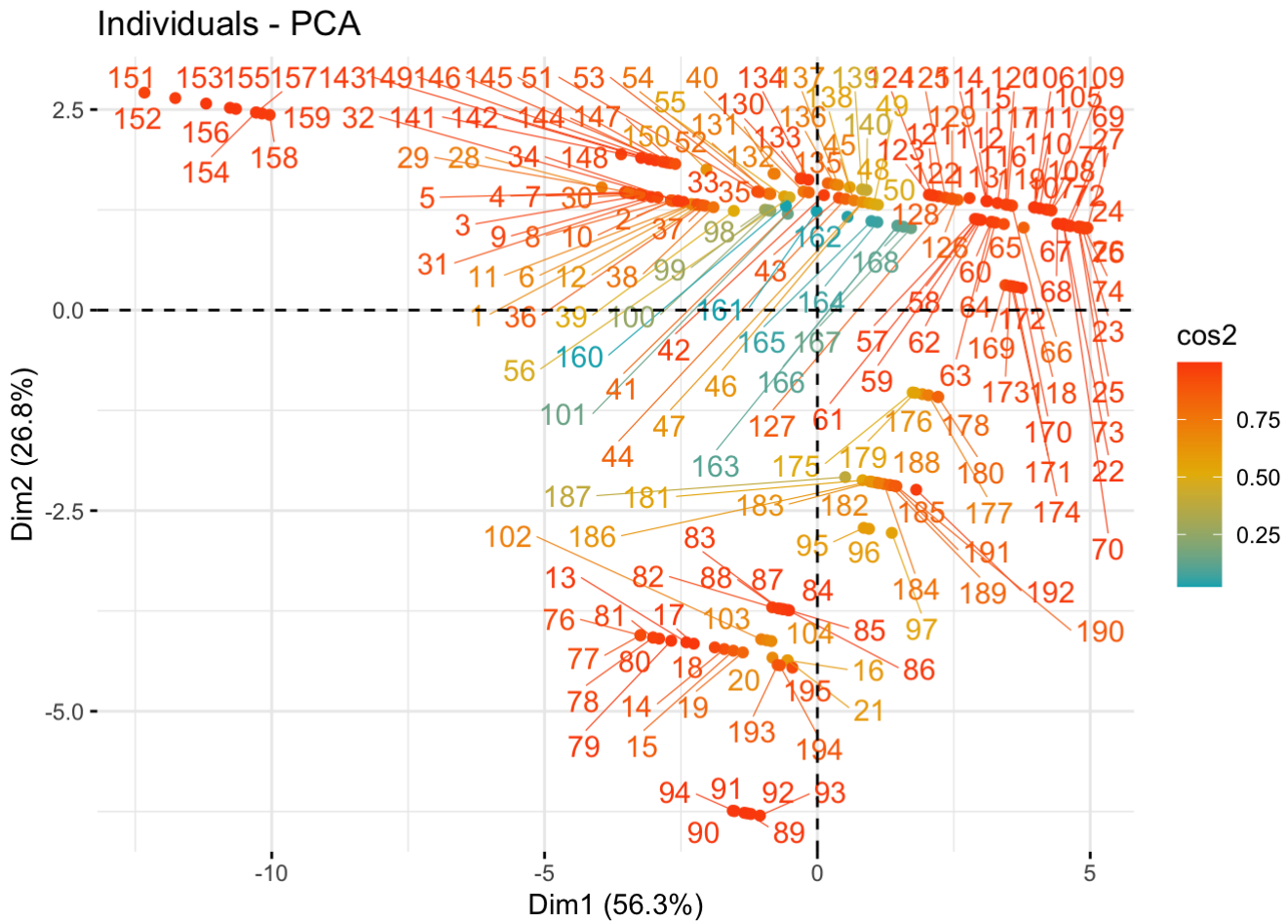

```
require(showtext)
library(FactoMineR)
require(FactoMineR)
require(factoextra)

fviz_eig(pralPCA)
```



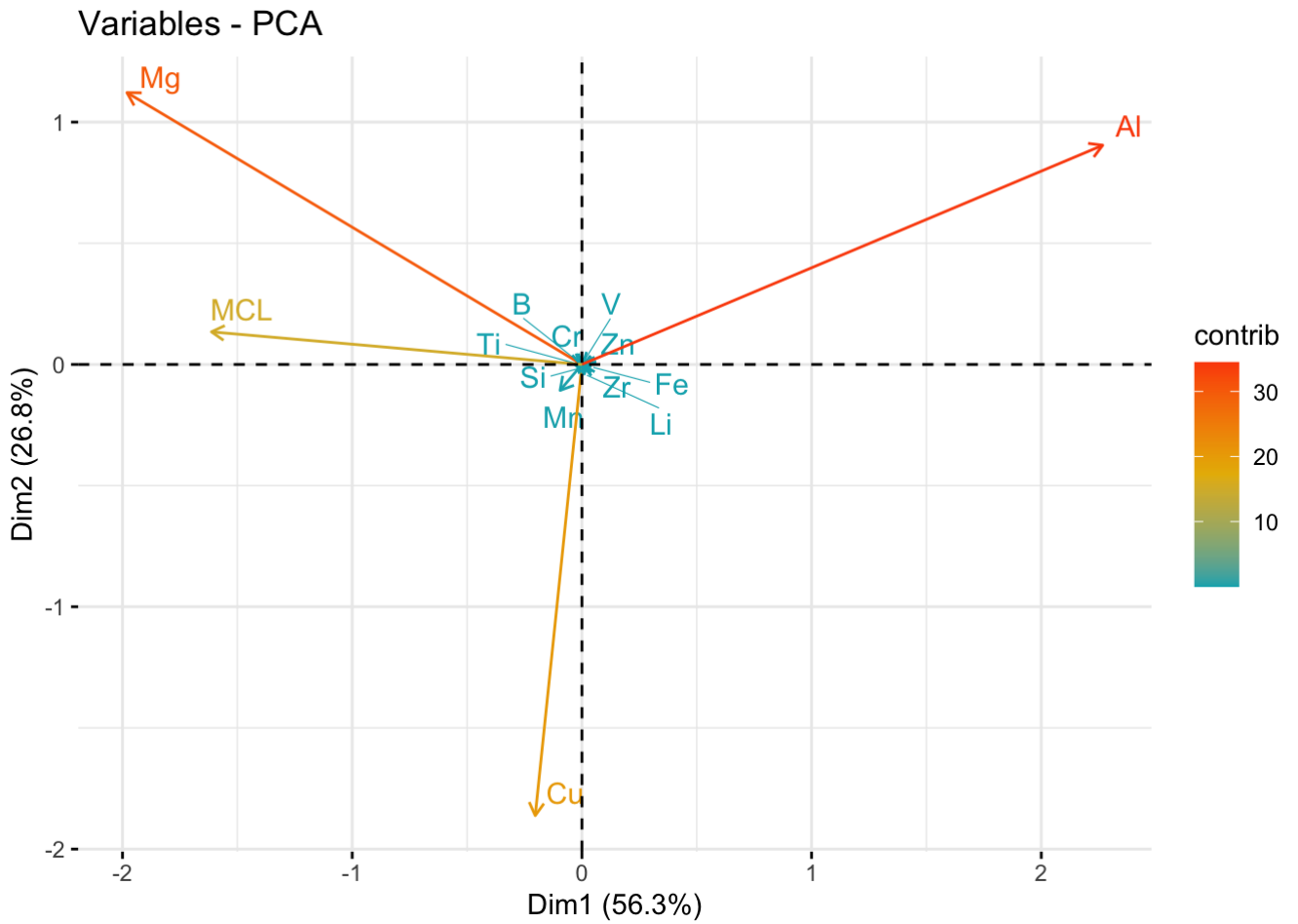

```
library(factoextra)
fviz_eig(pralPCA)
```



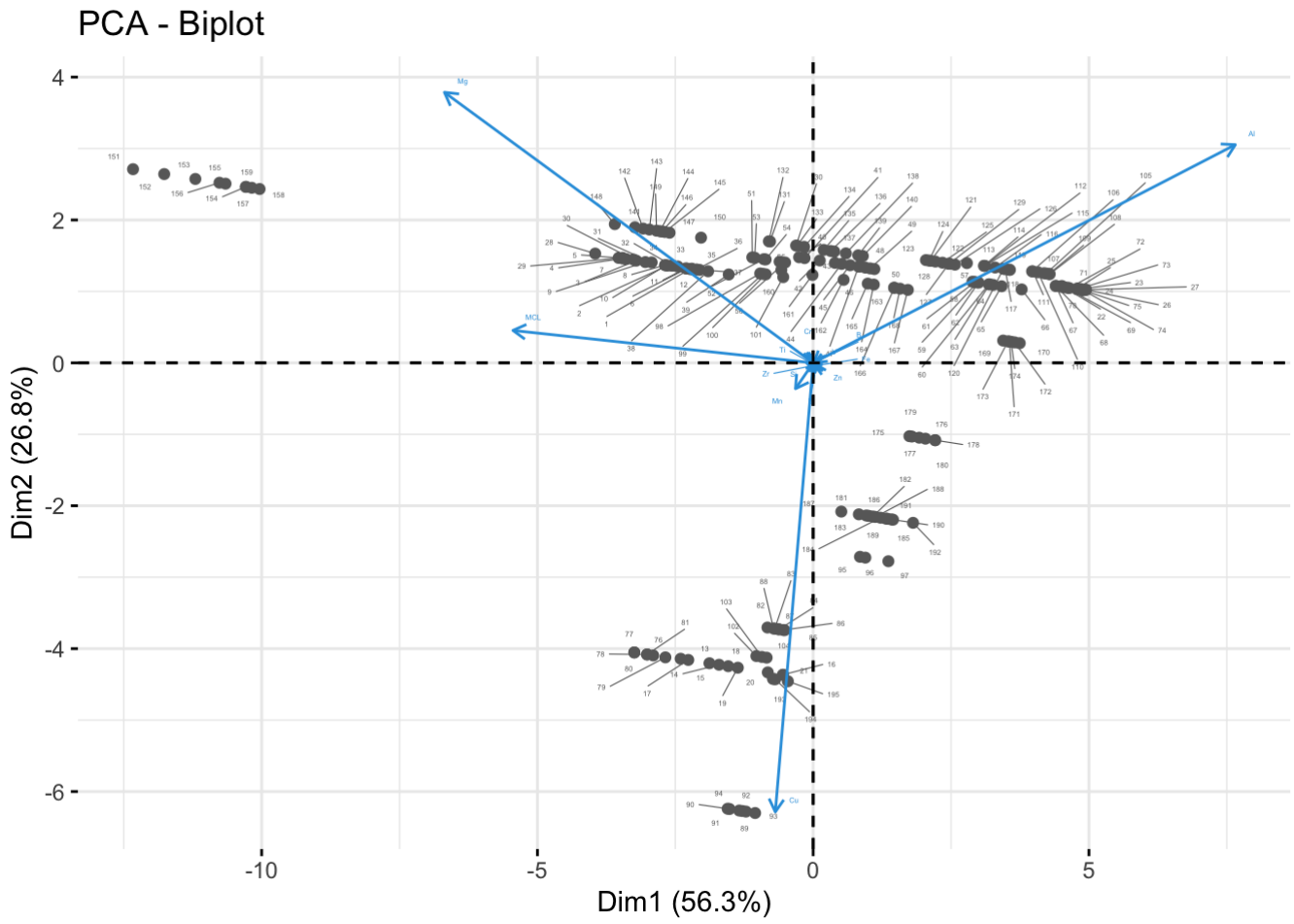

```
fviz_pca_ind(pralPCA,
             col.ind = "cos2", # Color by the quality of representation
             gradient.cols = c("#00AFBB", "#E7B800", "#FC4E07"),
             repel = TRUE     # Avoid text overlapping
             )
```




```
fviz_pca_var(pralPCA,
             col.var = "contrib", # Color by contributions to the PC
             gradient.cols = c("#00AFBB", "#E7B800", "#FC4E07"),
             repel = TRUE     # Avoid text overlapping
             )
```



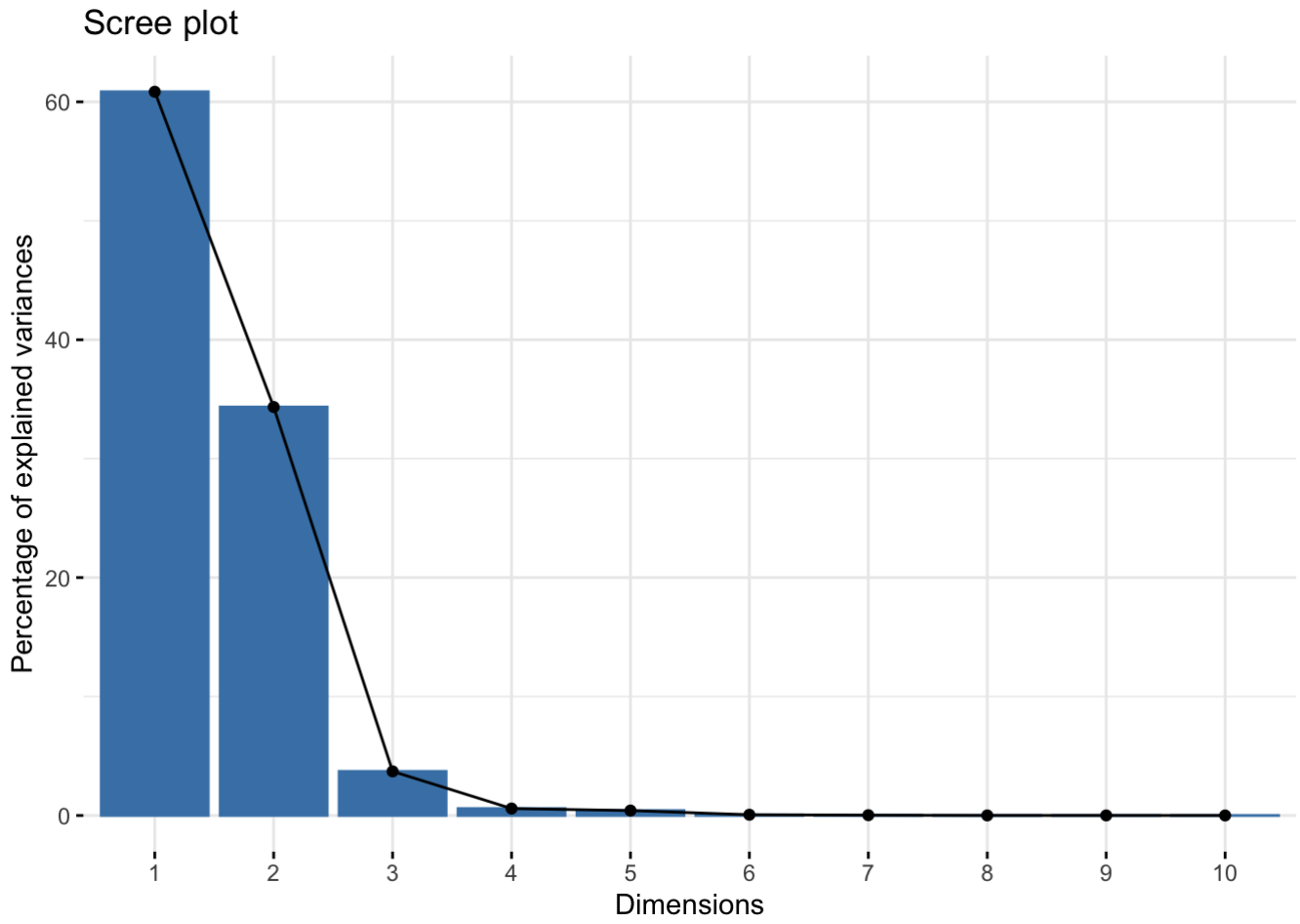

```
fviz_pca_biplot(pralPCA, repel = TRUE, labels = TRUE,
                col.var = "#2E9FDF", # Variables color
                col.ind = "#696969"  # Individuals color
                )
```



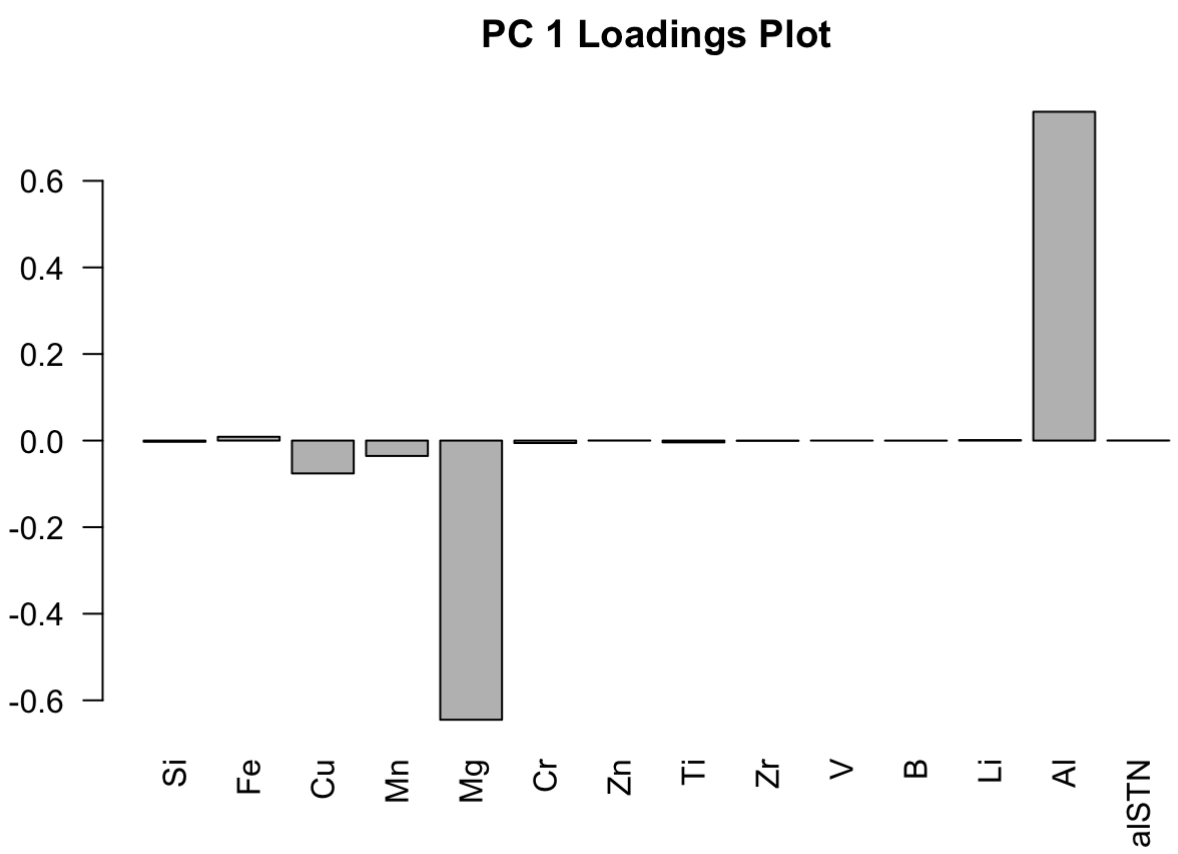

```
#MLR using only Al, Mg, Cu while looking at MCL
LMal2 = lm(MCL ~ Mg + Cu + Al, data = al)
summary(LMal2)
```

```
##
## Call:
## lm(formula = MCL ~ Mg + Cu + Al, data = al)
##
## Residuals:
##     Min      1Q  Median      3Q     Max
## -3.7049 -1.1715 -0.3048  0.7032  9.3522
##
## Coefficients:
##             Estimate Std. Error t value Pr(>|t|)
## (Intercept) -121.6902    35.2737  -3.450 0.000690 ***
## Mg             1.8080     0.3707   4.877 2.26e-06 ***
## Cu             1.6661     0.3944   4.224 3.71e-05 ***
## Al             1.2390     0.3542   3.498 0.000583 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Residual standard error: 1.926 on 191 degrees of freedom
## Multiple R-squared:  0.3006, Adjusted R-squared:  0.2897
## F-statistic: 27.37 on 3 and 191 DF,  p-value: 9.054e-15
```

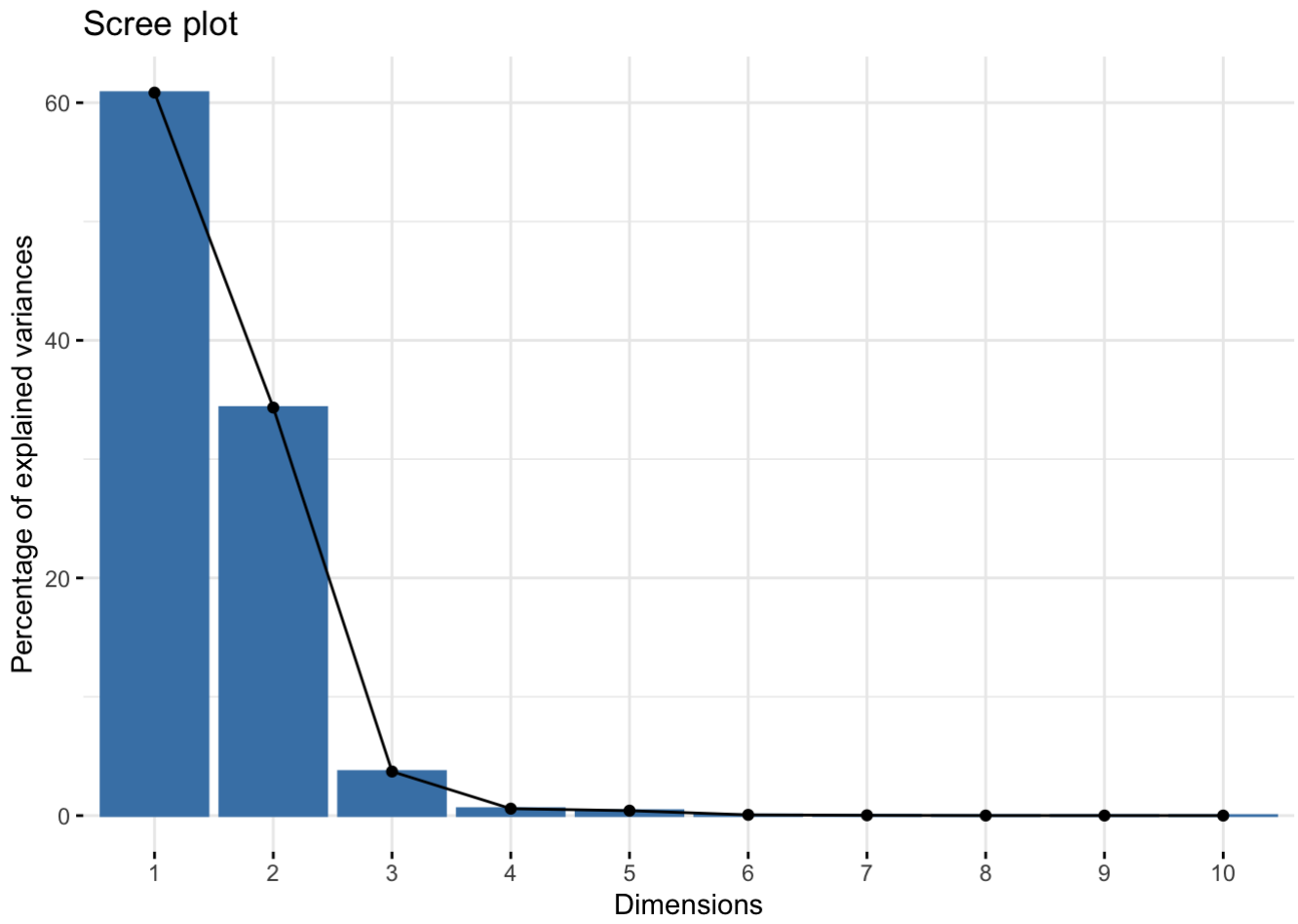
```
pralstn <- cbind(Si,Fe,Cu,Mn,Mg,Cr,Zn,Ti,Zr,V,B,Li,Al,alSTN)
pcaalstn = prcomp(pralstn, scale=F)
fviz_eig(pcaalstn)
```



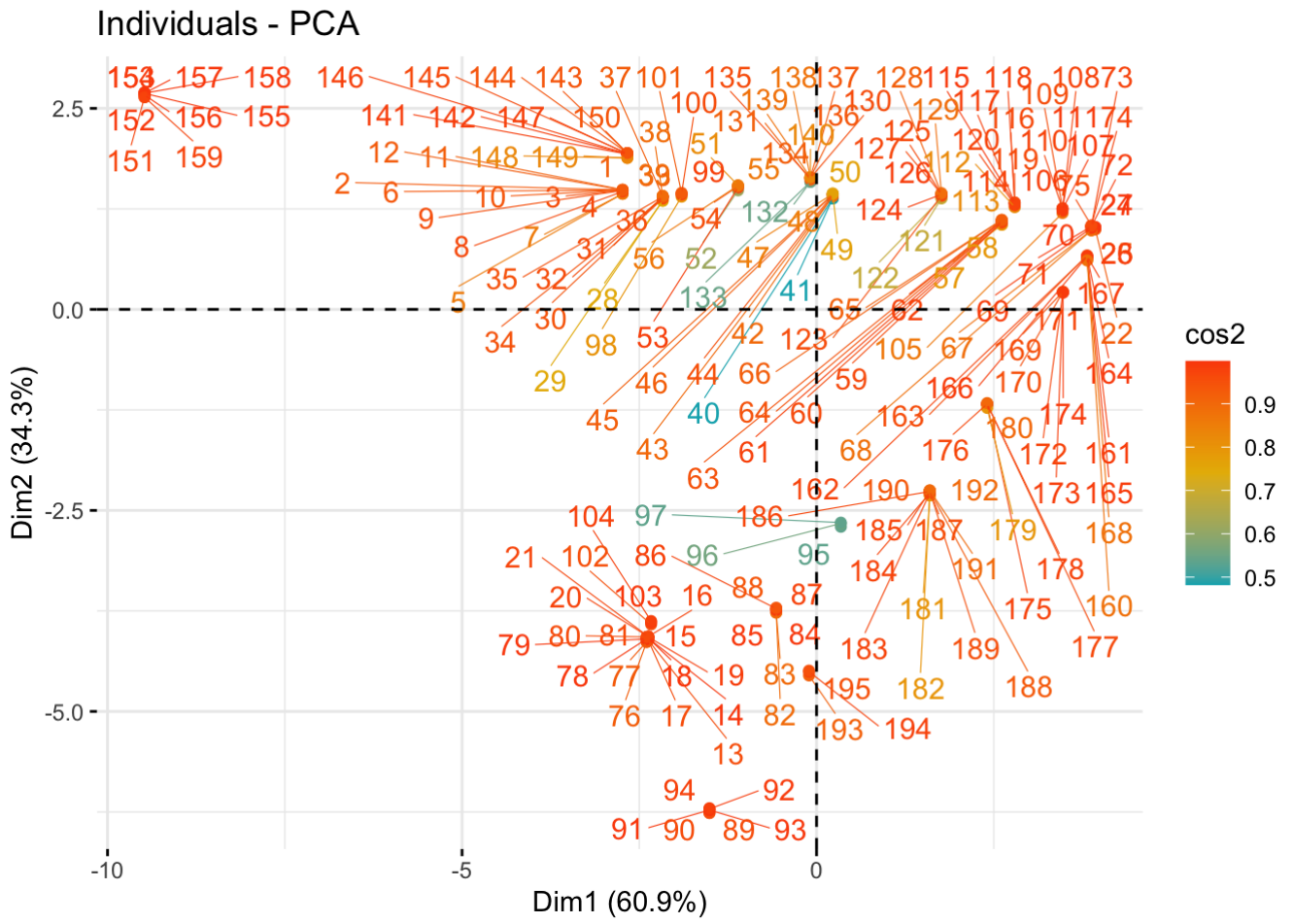

```
barplot(pcaalstn$rotation[,1], main="PC 1 Loadings Plot", las=2)
```



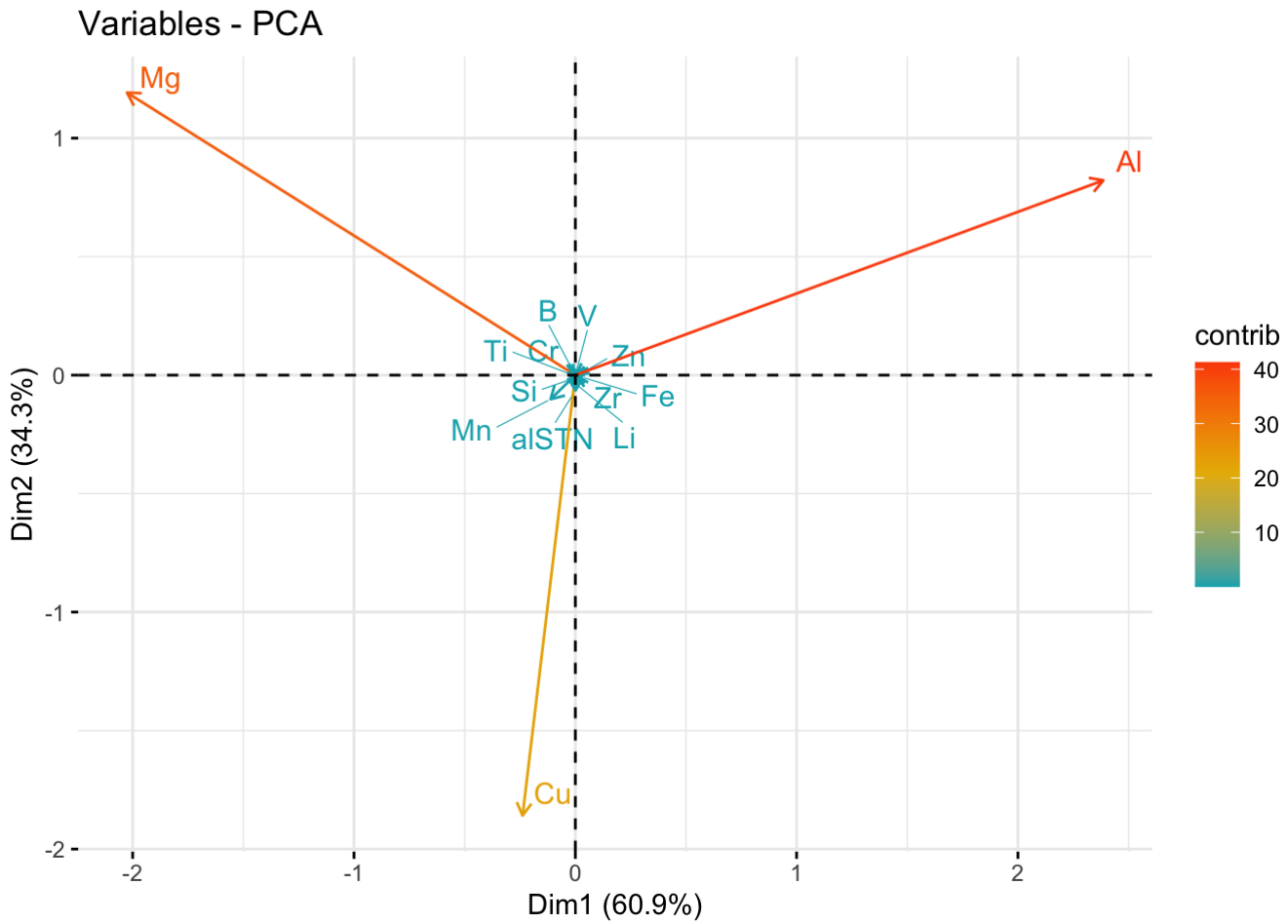

```
fviz_eig(pcaalstn)
```



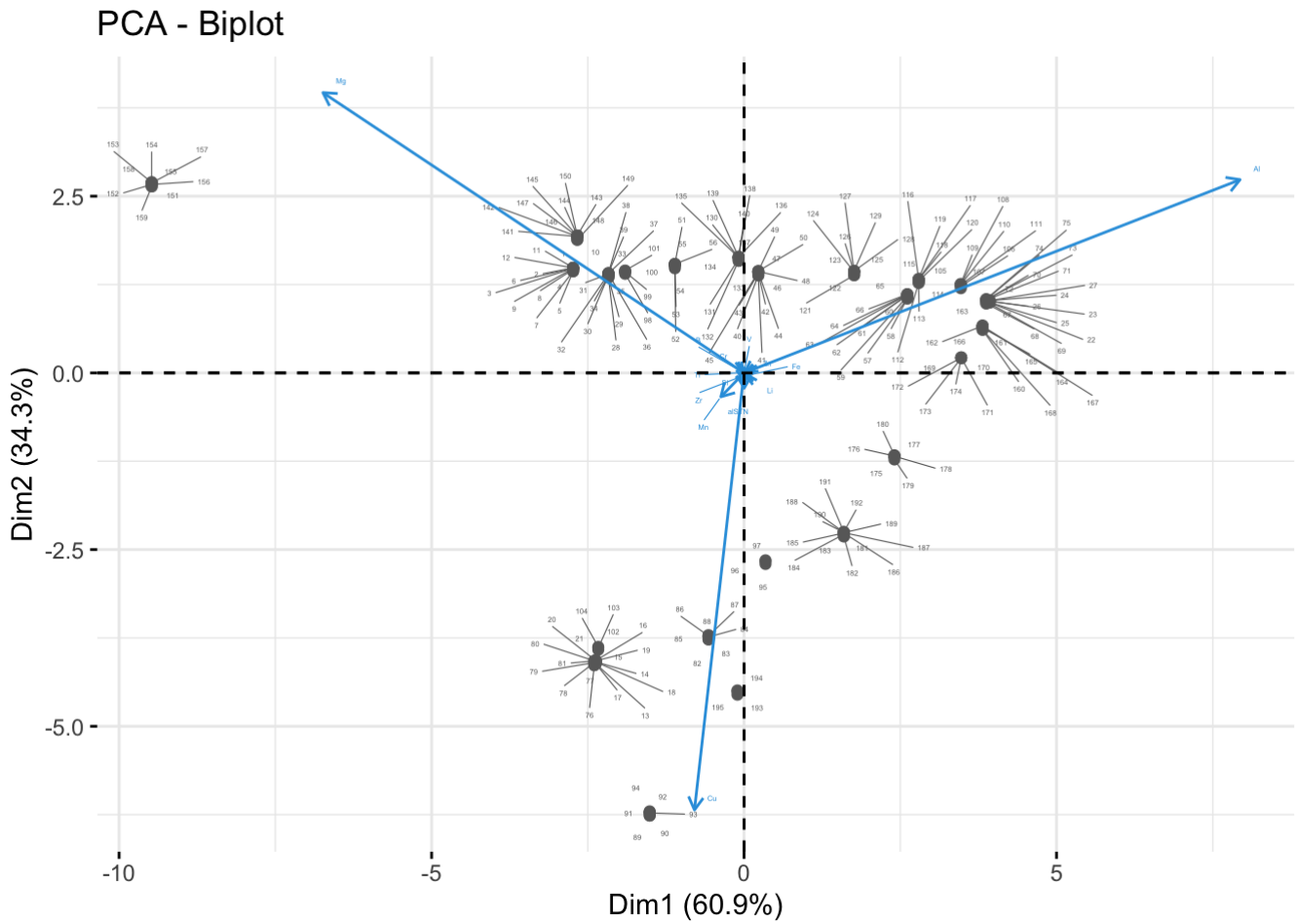

```
fviz_pca_ind(pcaalstn,
             col.ind = "cos2", # Color by the quality of representation
             gradient.cols = c("#00AFBB", "#E7B800", "#FC4E07"),
             repel = TRUE     # Avoid text overlapping
             )
```




```
fviz_pca_var(pcaalstn,
             col.var = "contrib", # Color by contributions to the PC
             gradient.cols = c("#00AFBB", "#E7B800", "#FC4E07"),
             repel = TRUE     # Avoid text overlapping
             )
```




```
fviz_pca_biplot(pcaalstn, repel = TRUE, labels = TRUE,
                col.var = "#2E9FDF", # Variables color
                col.ind = "#696969"  # Individuals color
                )
```



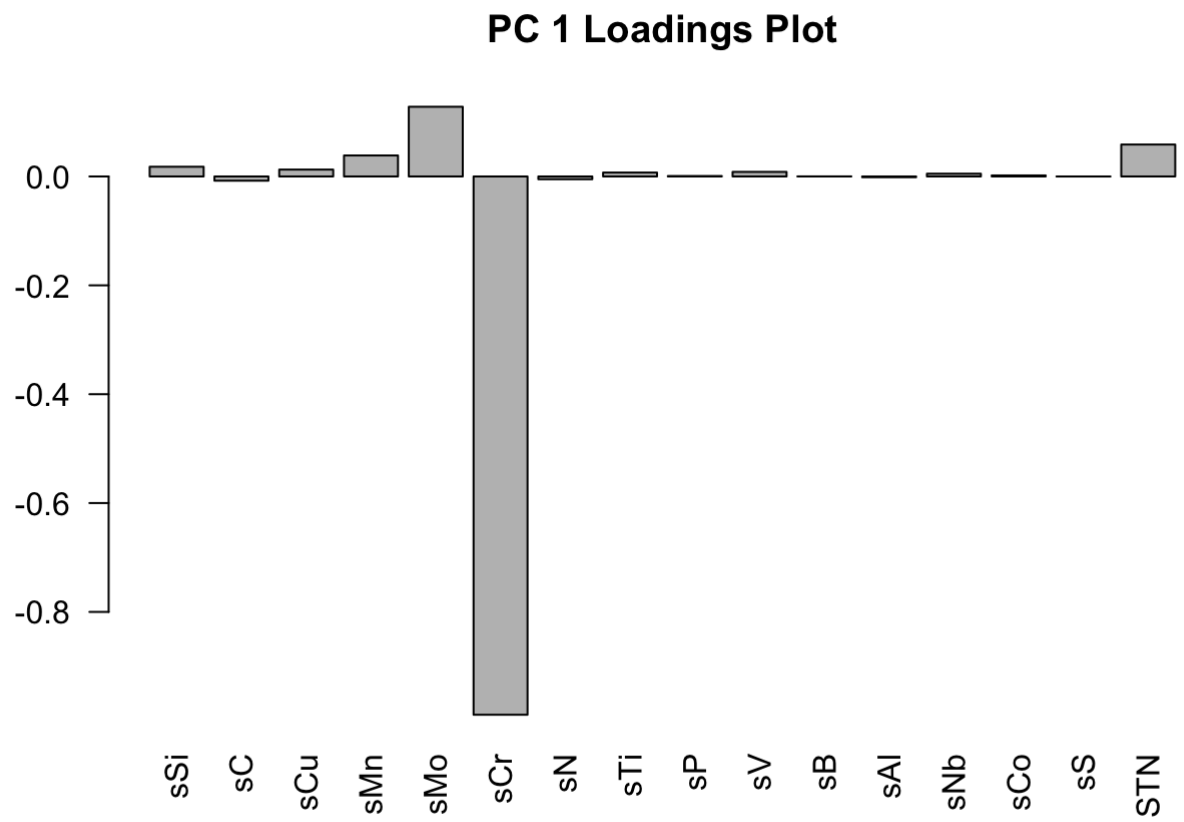

```
sSi <-as.numeric(stlrf_STN$Si)
sC <-as.numeric(stlrf_STN$C)
sCu <-as.numeric(stlrf_STN$Cu)
sMn <-as.numeric(stlrf_STN$Mn)
sMo <-as.numeric(stlrf_STN$Mo)
sCr <-as.numeric(stlrf_STN$Cr)
sN <-as.numeric(stlrf_STN$N)
sTi <-as.numeric(stlrf_STN$Ti)
sP <-as.numeric(stlrf_STN$P)
sV <-as.numeric(stlrf_STN$V)
sB <-as.numeric(stlrf_STN$B)
sNi <-as.numeric(stlrf_STN$Ni)
sAl <-as.numeric(stlrf_STN$Al)
sNb <-as.numeric(stlrf_STN$Nb)
sCo <-as.numeric(stlrf_STN$Co)
sS <-as.numeric(stlrf_STN$S)
STN <-as.numeric(stlrf_STN$Strain)
TCL <- as.numeric(stlrf_TCL$TCL)
```

```
prasstl <- cbind(sSi,sC,sCu,sMn,sMo,sCr,sN,sTi,sP,sV,sB,sAl,sNb,sCo,sS,STN)
```

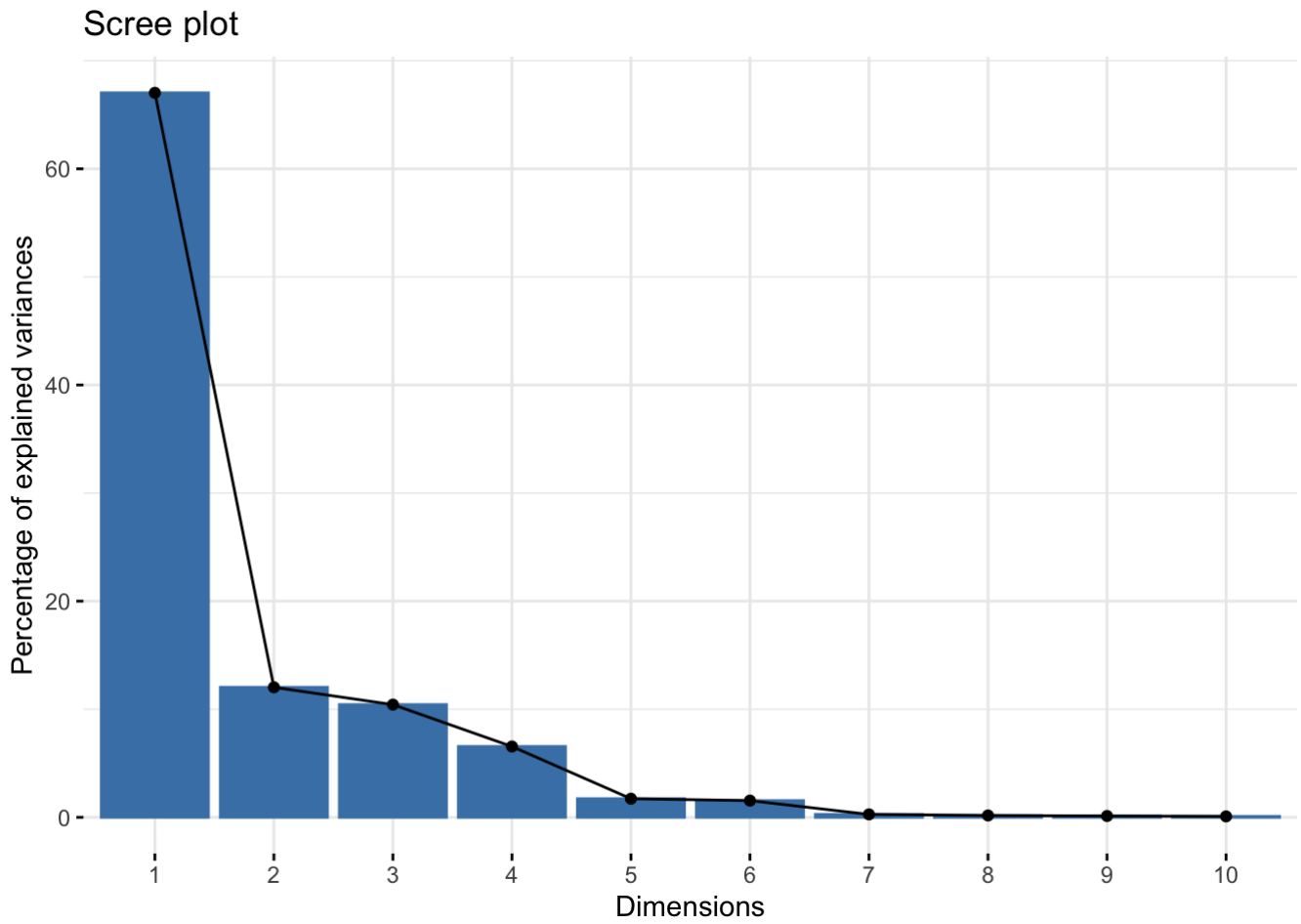
```
pcasstl = prcomp(prasstl, scale=F)

summary(pcasstl)
```

```
## Importance of components:
##                           PC1    PC2    PC3     PC4     PC5     PC6     PC7
## Standard deviation     3.3093 1.4020 1.3052 1.03458 0.53020 0.50246 0.21049
## Proportion of Variance 0.6703 0.1203 0.1043 0.06551 0.01721 0.01545 0.00271
## Cumulative Proportion  0.6703 0.7906 0.8948 0.96035 0.97756 0.99301 0.99572
##                           PC8     PC9    PC10    PC11    PC12    PC13     PC14
## Standard deviation     0.1669 0.14092 0.11861 0.06646 0.04400 0.04068 0.008156
## Proportion of Variance 0.0017 0.00122 0.00086 0.00027 0.00012 0.00010 0.000000
## Cumulative Proportion  0.9974 0.99864 0.99950 0.99977 0.99989 0.99999 1.000000
##                           PC15     PC16
## Standard deviation     0.005316 0.001222
## Proportion of Variance 0.000000 0.000000
## Cumulative Proportion  1.000000 1.000000
```

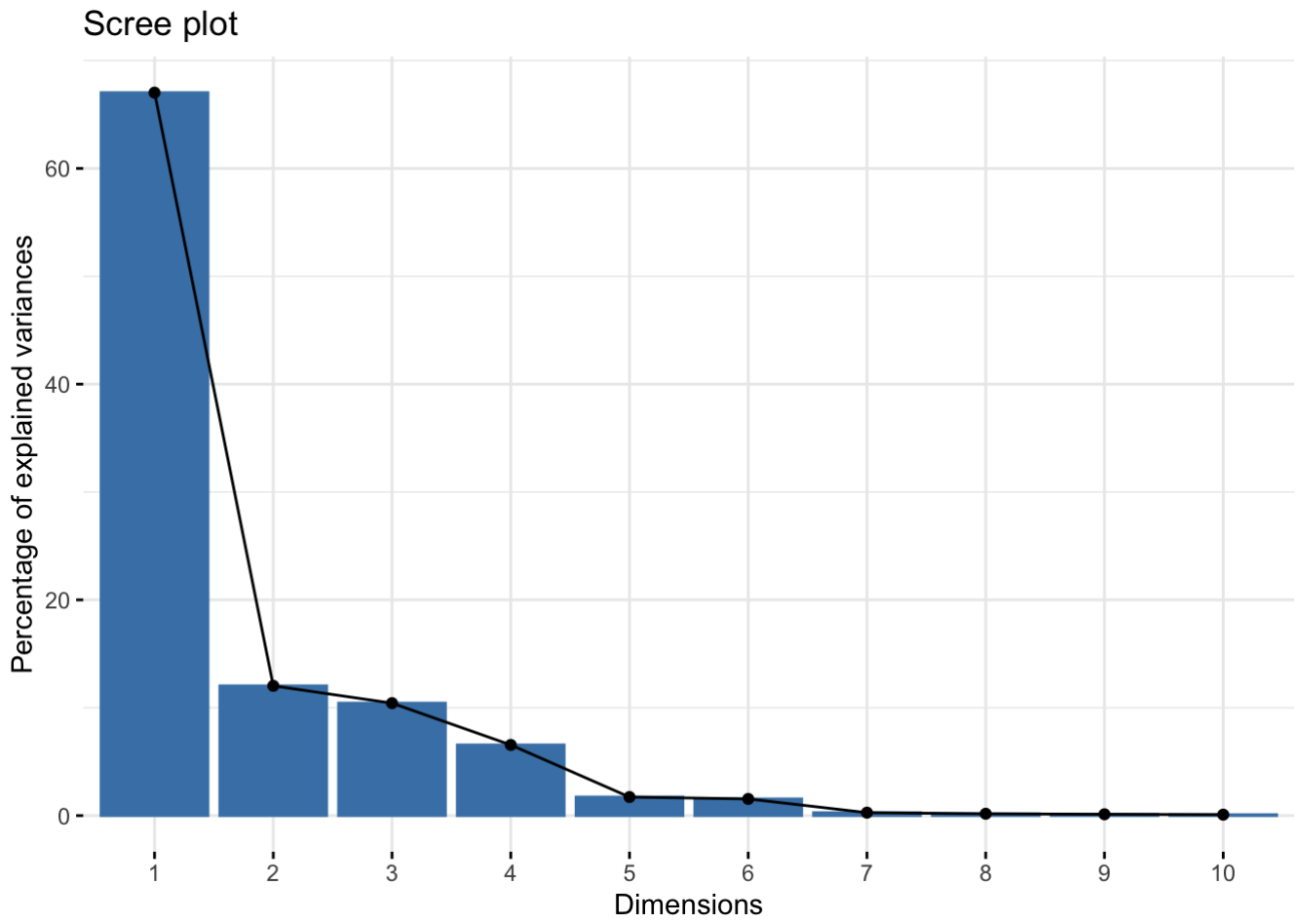
```
barplot(pcasstl$rotation[,1], main="PC 1 Loadings Plot", las=2)
```



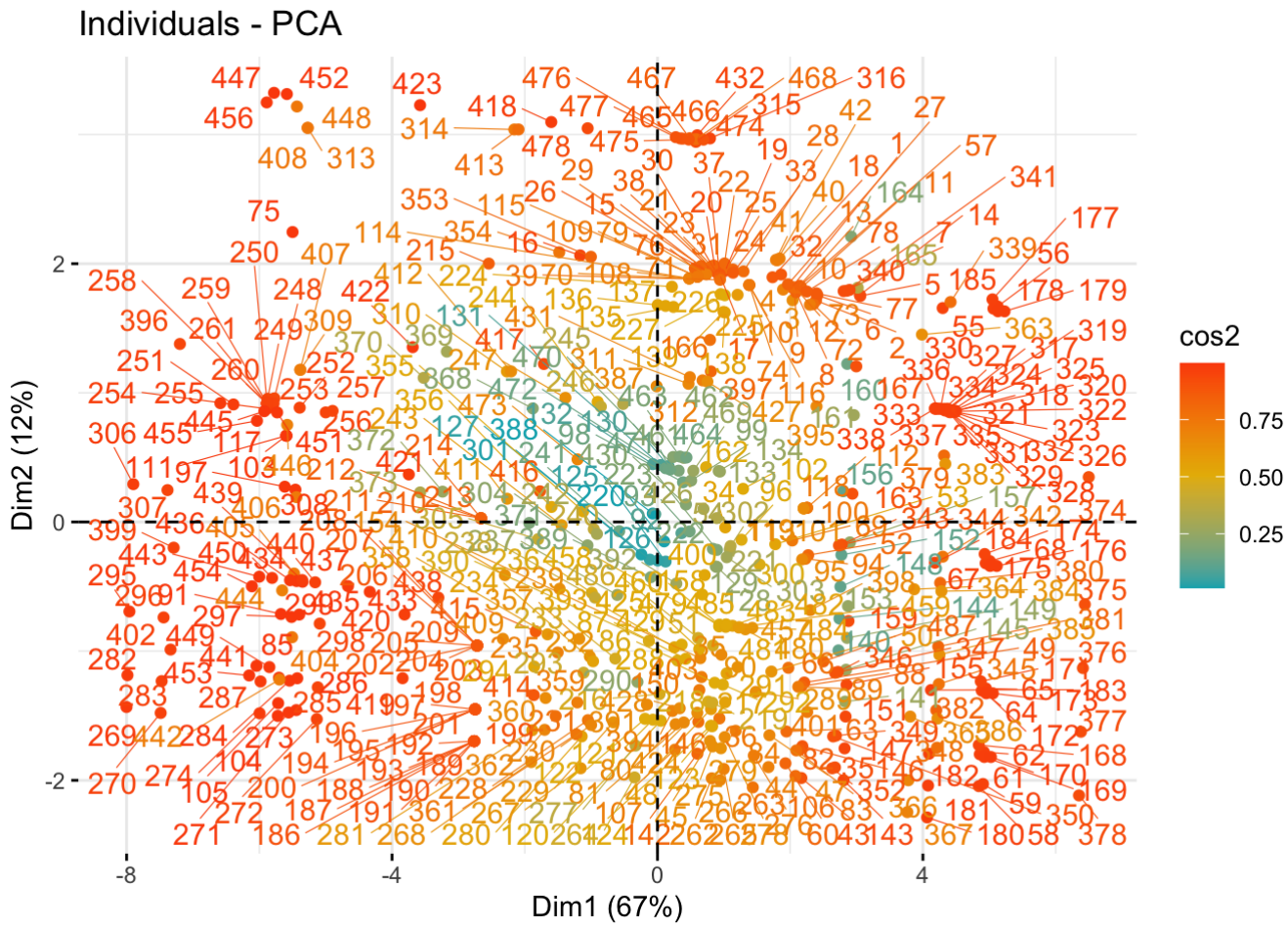

```
fviz_eig(pcasstl)
```



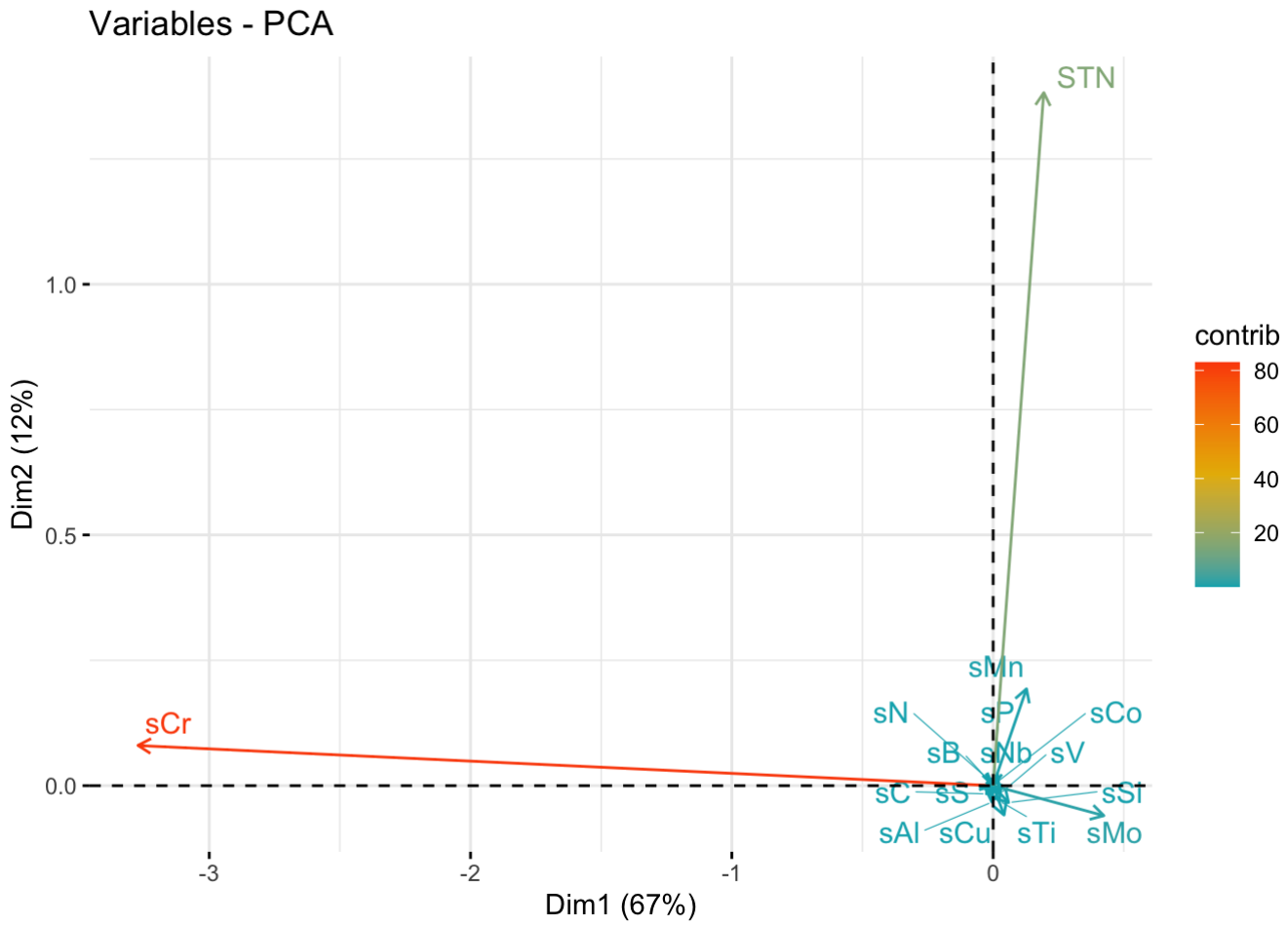

```
library(factoextra)
fviz_eig(pcasstl)
```



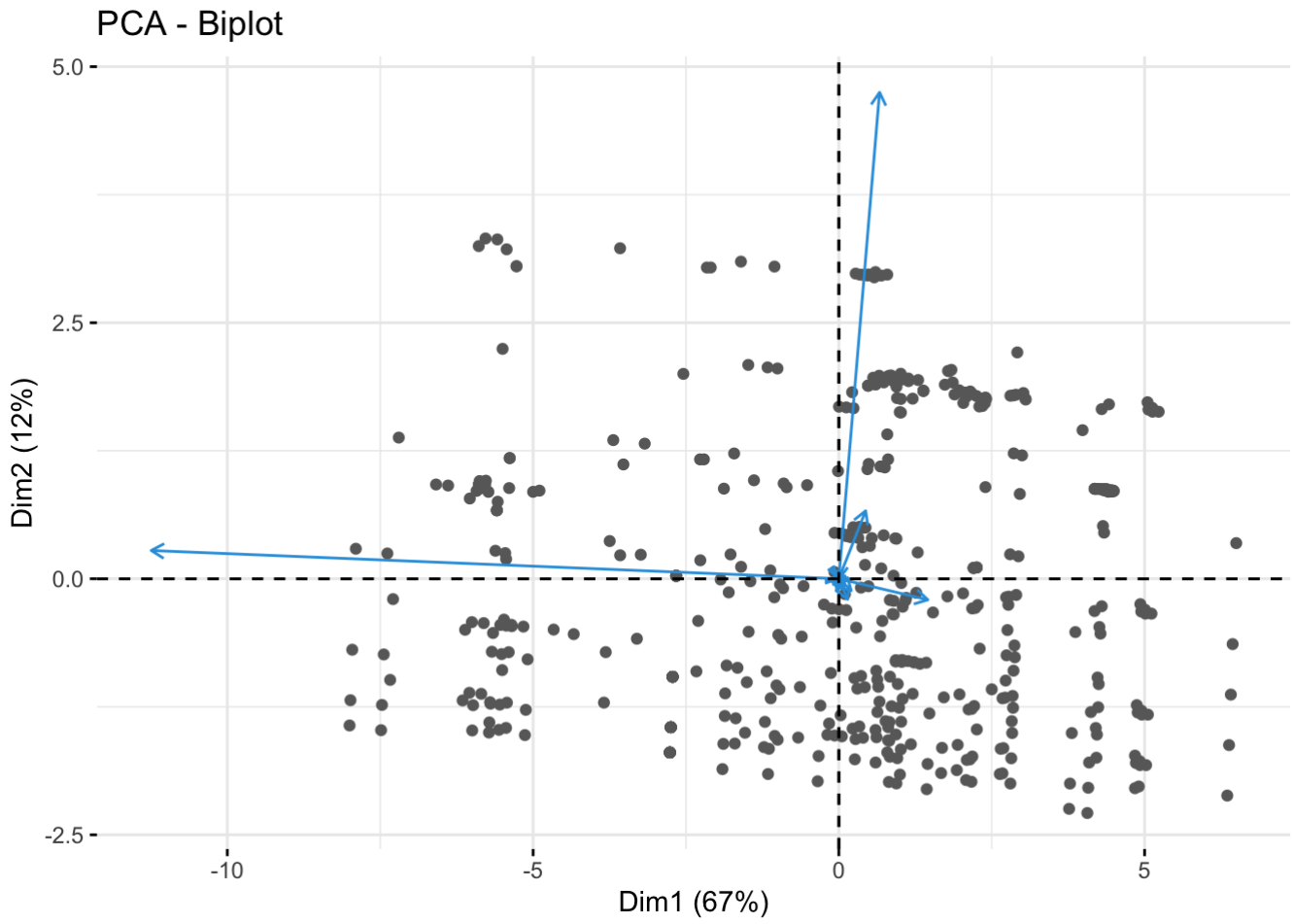

```
fviz_pca_ind(pcasstl,
             col.ind = "cos2", # Color by the quality of representation
             gradient.cols = c("#00AFBB", "#E7B800", "#FC4E07"),
             repel = TRUE     # Avoid text overlapping
             )
```




```
fviz_pca_var(pcasstl,
             col.var = "contrib", # Color by contributions to the PC
             gradient.cols = c("#00AFBB", "#E7B800", "#FC4E07"),
             repel = TRUE     # Avoid text overlapping
             )
```




```
fviz_pca_biplot(pcasstl, repel = TRUE, label = TRUE,
                col.var = "#2E9FDF", # Variables color
                col.ind = "#696969"  # Individuals color
                )
```

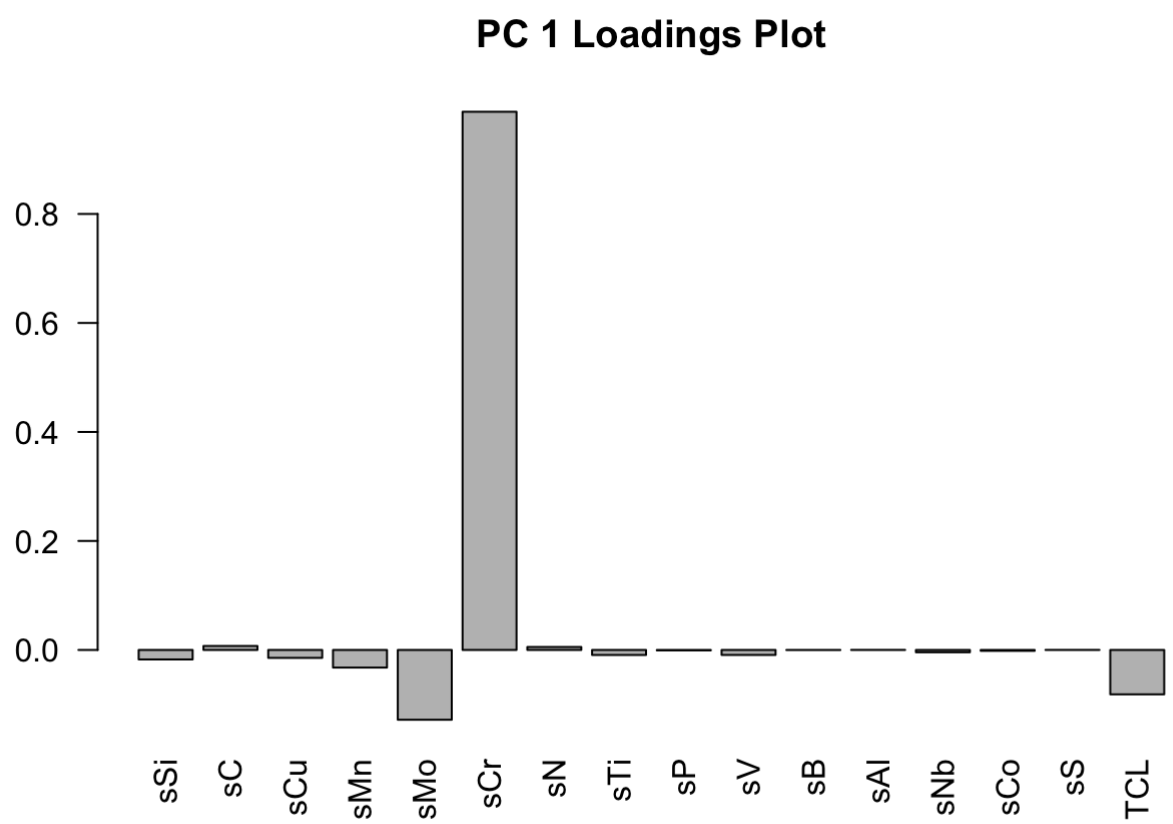



```
prasstl2 <- cbind(sSi,sC,sCu,sMn,sMo,sCr,sN,sTi,sP,sV,sB,sAl,sNb,sCo,sS,TCL)
```

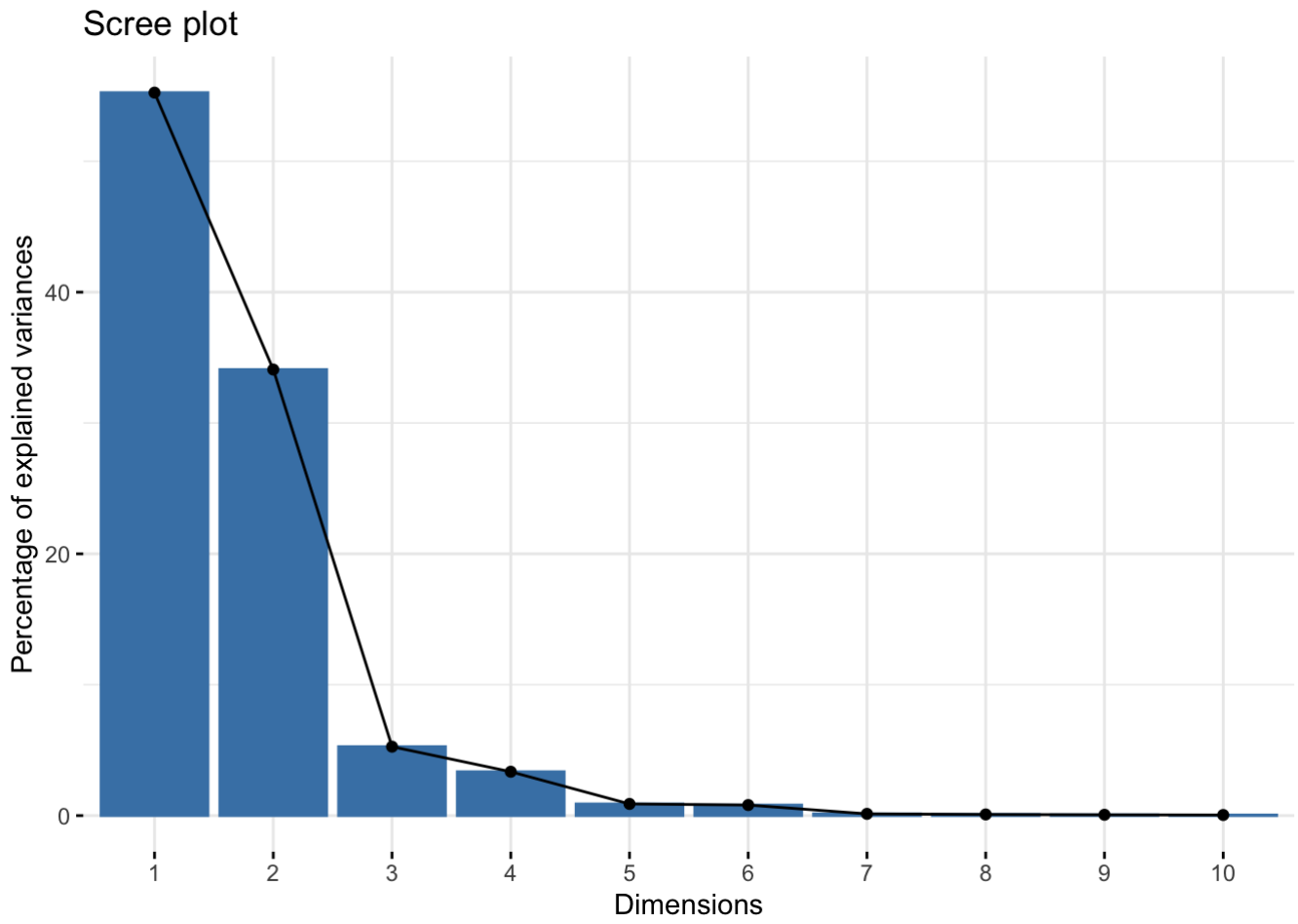
```
pcatstl = prcomp(prasstl2, scale=F)

summary(pcatstl)
```

```
## Importance of components:
##                           PC1    PC2     PC3     PC4     PC5     PC6     PC7
## Standard deviation     4.1983 3.2978 1.29642 1.03397 0.53302 0.50704 0.20541
## Proportion of Variance 0.5525 0.3409 0.05268 0.03351 0.00891 0.00806 0.00132
## Cumulative Proportion  0.5525 0.8933 0.94601 0.97952 0.98842 0.99648 0.99780
##                           PC8     PC9    PC10    PC11    PC12    PC13     PC14
## Standard deviation     0.16650 0.14091 0.12003 0.06655 0.04411 0.04050 0.008163
## Proportion of Variance 0.00087 0.00062 0.00045 0.00014 0.00006 0.00005 0.000000
## Cumulative Proportion  0.99867 0.99929 0.99975 0.99988 0.99995 1.00000 1.000000
##                           PC15     PC16
## Standard deviation     0.005308 0.001207
## Proportion of Variance 0.000000 0.000000
## Cumulative Proportion  1.000000 1.000000
```

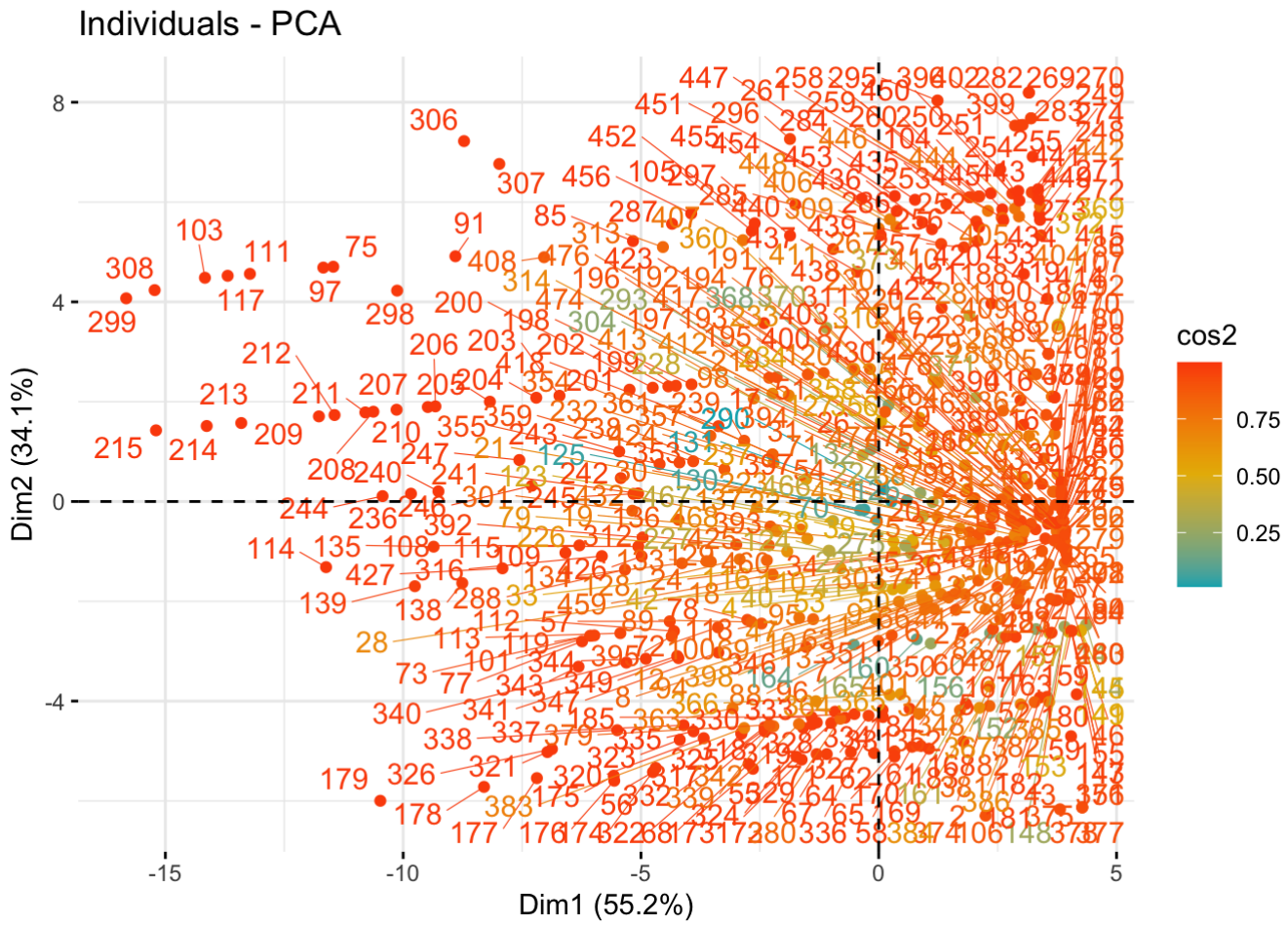
```
barplot(pcatstl$rotation[,2], main="PC 1 Loadings Plot", las=2)
```



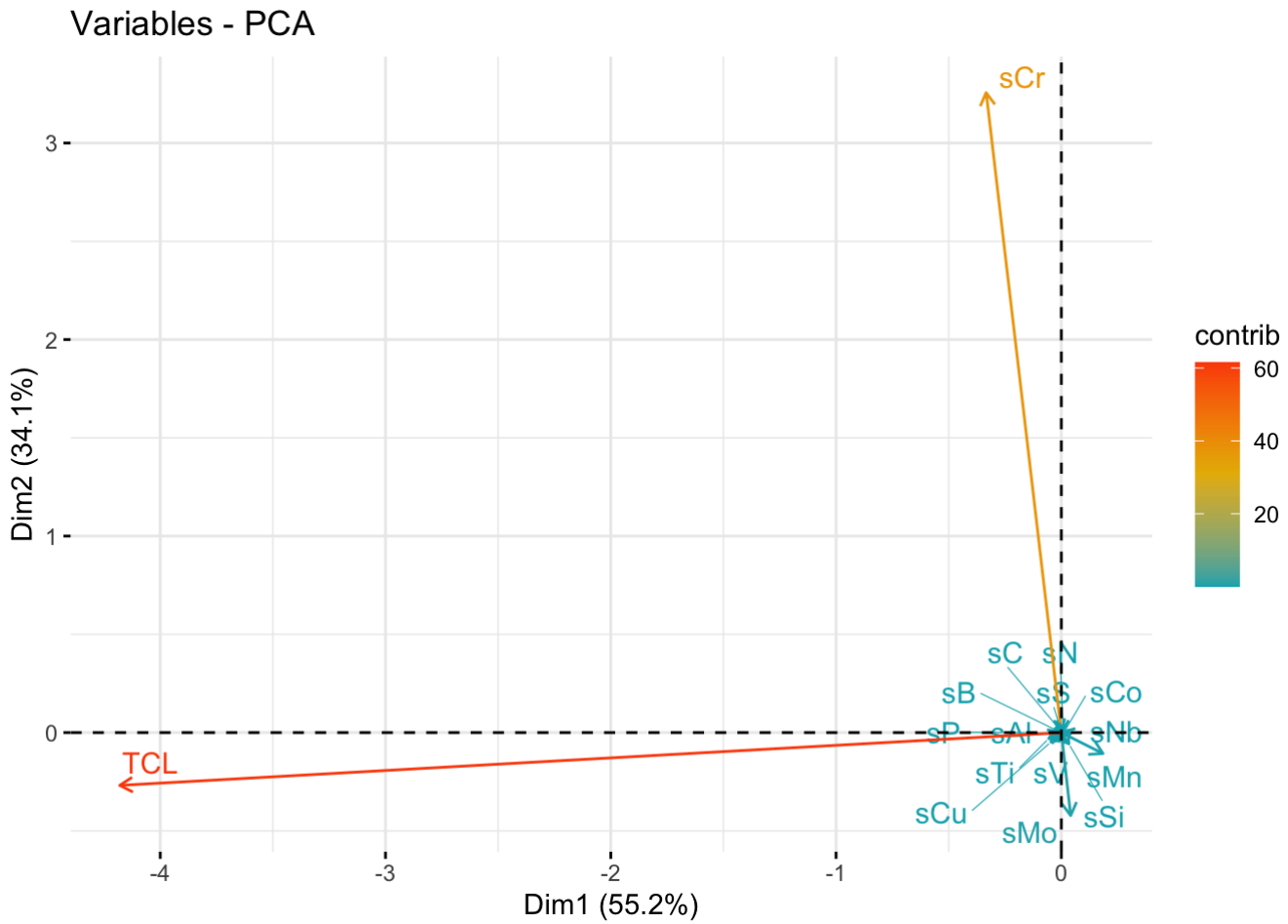

```
fviz_eig(pcatstl)
```



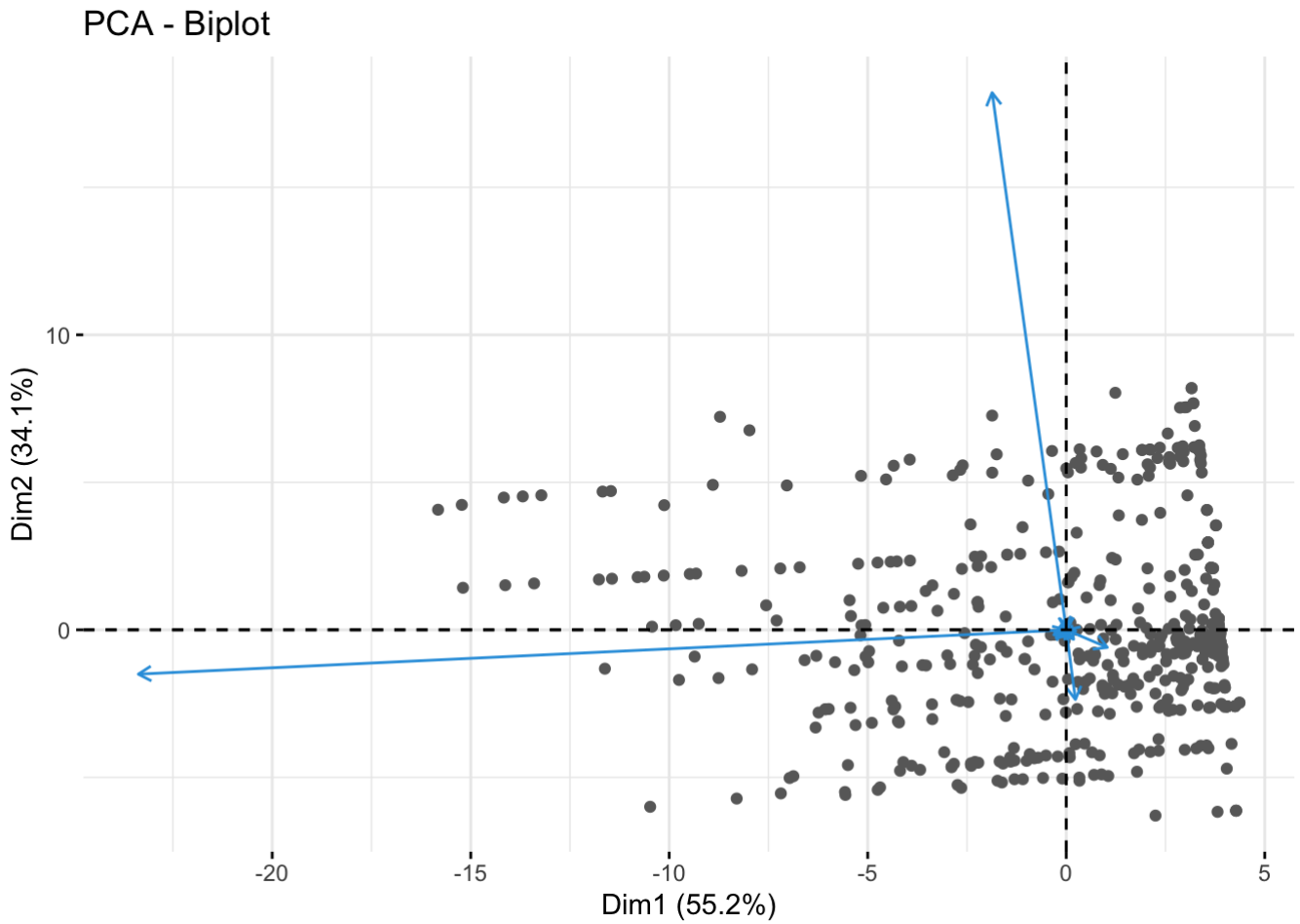

```
fviz_pca_ind(pcatstl,
             col.ind = "cos2", # Color by the quality of representation
             gradient.cols = c("#00AFBB", "#E7B800", "#FC4E07"),
             repel = TRUE     # Avoid text overlapping
             )
```




```
fviz_pca_var(pcatstl,
             col.var = "contrib", # Color by contributions to the PC
             gradient.cols = c("#00AFBB", "#E7B800", "#FC4E07"),
             repel = TRUE     # Avoid text overlapping
             )
```




```
fviz_pca_biplot(pcatstl, repel = TRUE, label = TRUE,
                col.var = "#2E9FDF", # Variables color
                col.ind = "#696969"  # Individuals color
                )
```


PCA - Biplot
10
0
Dim2 (34.1%)
-20
-15
-10
-5
0
5
Dim1 (55.2%)