BNP Project

```
library(data.table)
library(FactoMineR)
library(factoextra)
library(gplots)
```

Exploratory data analysis for BNP Porject

Categorical variables

Looking for correlated categorical variables(Correspondence Analysis)

```
correlated.cat.vars = c()
cats <- categorical.cols
for (col1 in cats) {
  cats <- cats[-1]
  for (col2 in cats) {
    tb <- train[, c(col1, col2), with=FALSE]
    ct <- table(tb) # contingency table
    ca <- CA(ct, graph = FALSE) # correspondence analysis
    trace <- sum(ca$eig$eigenvalue)
    if (sqrt(trace) > 1) {
      correlated.cat.vars <- rbind(correlated.cat.vars, c(col1, col2, sqrt(trace)))
    }
  }
}
correlated.cat.vars <- data.frame(correlated.cat.vars)
colnames(correlated.cat.vars) <- c("var1", "var2", "corr.coeff")
correlated.cat.vars
```

```
##    var1 var2       corr.coeff
## 1    v3  v31 1.00536863164762
## 2    v3  v56 1.17787966506551
## 3   v22  v30 1.22569142296392
## 4   v22  v47 1.31017670015523
## 5   v22  v52 1.43773531229448
## 6   v22  v56 4.78499658073927
## 7   v22  v71 1.01087626758596
## 8   v22  v79 1.83693501591518
## 9   v22  v91  1.0501783338625
## 10  v22 v107 1.05017833386251
## 11  v22 v112 4.53564951702748
## 12  v22 v113  2.6557782212823
## 13  v22 v125 9.12626440332216
## 14  v31  v47  1.0875267653298
## 15  v31  v56 1.30014952990685
## 16  v31  v79  1.2875527694955
## 17  v31 v110 1.08738978098615
## 18  v47  v56 2.09802658938753
## 19  v47  v79 2.64511047809333
```

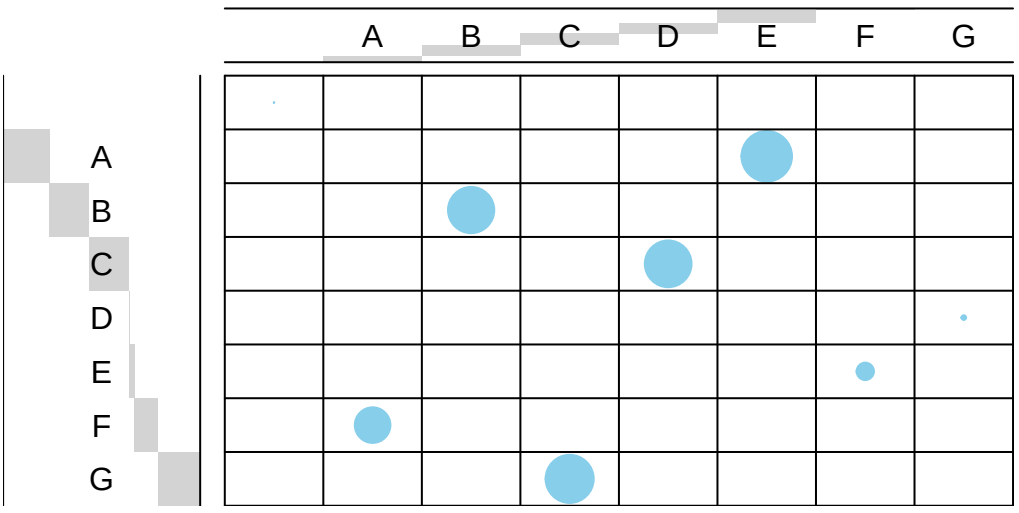
```
## 20  v47 v110  1.4142135623731
## 21  v47 v113 1.10031513083776
## 22  v52  v91 1.00082056913307
## 23  v52 v107 1.00082056913307
## 24  v56  v79 2.58071022655841
## 25  v56 v110 1.05938850903617
## 26  v56 v113 1.20387335191441
## 27  v71  v75 1.30710030588841
## 28  v79 v110 1.41421356237309
## 29  v79 v113 1.13727058901571
## 30  v91 v107 2.64575131106459
## 31 v110 v113 1.06333008154547
## 32 v112 v125 4.69041575982343
```

Categorical variables: v91-v107

```
# Contingency table
tab.v91.v107 <- table(train[,.(v91, v107),])
tab.v91.v107
```

```
##     v107
## v91           A     B     C     D     E     F     G
##         3     0     0     0     0     0     0     0
##   A     0     0     0     0     0 27079     0     0
##   B     0     0 22683     0     0     0     0     0
##   C     0     0     0     0 23157     0     0     0
##   D     0     0     0     0     0     0     0   230
##   E     0     0     0     0     0     0  3206     0
##   F     0 13418     0     0     0     0     0     0
##   G     0     0     0 24545     0     0     0     0
```

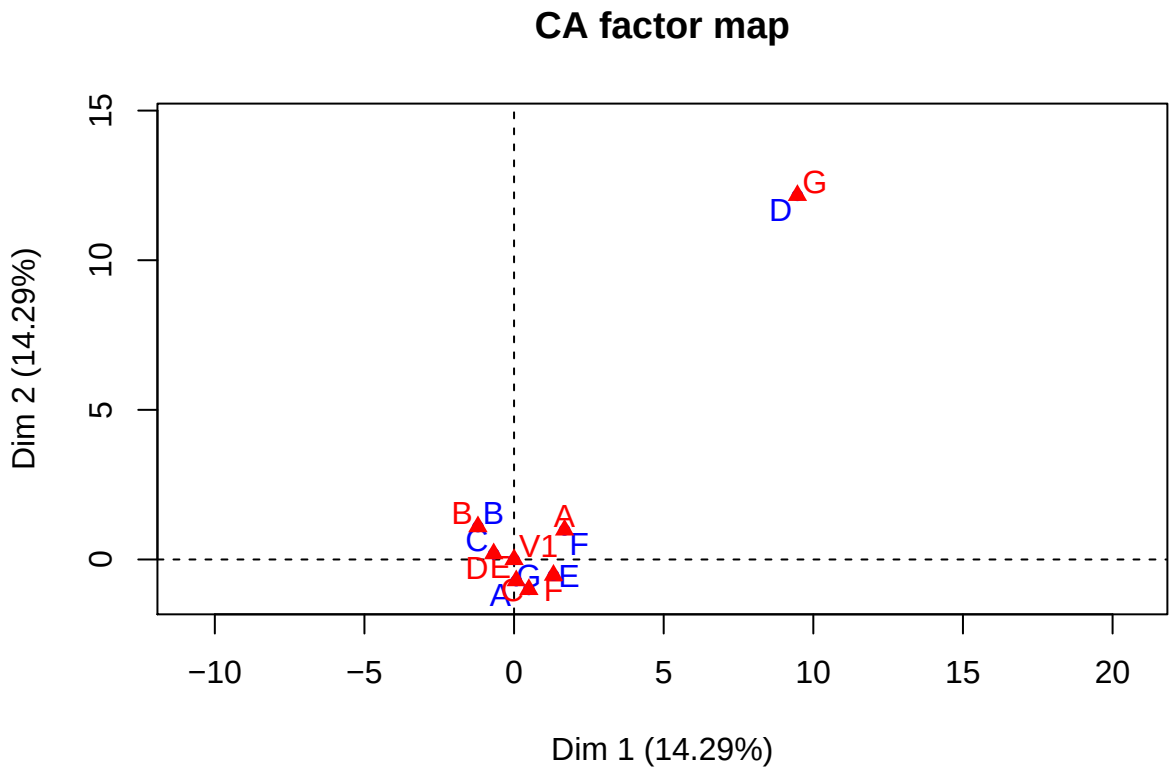
```
# Draw balloon plot
balloonplot(t(tab.v91.v107), main ="Categorical variable: v91-107", xlab ="",
            ylab="", label = FALSE,  show.margins = FALSE)
```



```r
# Chi-squared test
chisq.test(tab.v91.v107)
```

```
##
##  Pearson's Chi-squared test
##
## data:  tab.v91.v107
## X-squared = 800250, df = 49, p-value < 2.2e-16
```

```r
# Correspondence analysis(CA)
ca.v91.v107 <-  CA(tab.v91.v107, graph=TRUE)
```




```r
summary(ca.v91.v107, nb.dec = 2, ncp = 2)
```

```
##
## Call:
## CA(X = tab.v91.v107, graph = TRUE)
##
## The chi square of independence between the two variables is equal to 800247 (p-value =  0 ).
##
## Eigenvalues
##                        Dim.1  Dim.2  Dim.3  Dim.4  Dim.5  Dim.6  Dim.7
## Variance                1.00   1.00   1.00   1.00   1.00   1.00   1.00
## % of var.              14.29  14.29  14.29  14.29  14.29  14.29  14.29
## Cumulative % of var.   14.29  28.57  42.86  57.14  71.43  85.71 100.00
##
## Rows
##      Iner*1000    Dim.1    ctr   cos2    Dim.2    ctr   cos2
##    |    999.97 |   0.00   0.00   0.00 |   0.00   0.00   0.00 |
## A  |    763.13 |   0.07   0.12   0.00 |  -0.71  11.86   0.16 |
```

```
## B  |    801.59 |  -1.21  28.90   0.36 |   1.10  23.89   0.30 |
## C  |    797.44 |  -0.68   9.39   0.12 |   0.19   0.76   0.01 |
## D  |    997.99 |   9.47  18.04   0.18 |  12.17  29.79   0.30 |
## E  |    971.96 |   1.32   4.87   0.05 |  -0.52   0.77   0.01 |
## F  |    882.63 |   1.69  33.43   0.38 |   0.99  11.42   0.13 |
## G  |    785.30 |   0.49   5.25   0.07 |  -1.00  21.52   0.27 |
##
## Columns
##      Iner*1000    Dim.1    ctr   cos2    Dim.2    ctr   cos2
## V1 |    999.97 |   0.00   0.00   0.00 |   0.00   0.00   0.00 |
## A  |    882.63 |   1.69  33.43   0.38 |   0.99  11.42   0.13 |
## B  |    801.59 |  -1.21  28.90   0.36 |   1.10  23.89   0.30 |
## C  |    785.30 |   0.49   5.25   0.07 |  -1.00  21.52   0.27 |
## D  |    797.44 |  -0.68   9.39   0.12 |   0.19   0.76   0.01 |
## E  |    763.13 |   0.07   0.12   0.00 |  -0.71  11.86   0.16 |
## F  |    971.96 |   1.32   4.87   0.05 |  -0.52   0.77   0.01 |
## G  |    997.99 |   9.47  18.04   0.18 |  12.17  29.79   0.30 |
```

These two variables are duplicates. Just keep one of them.

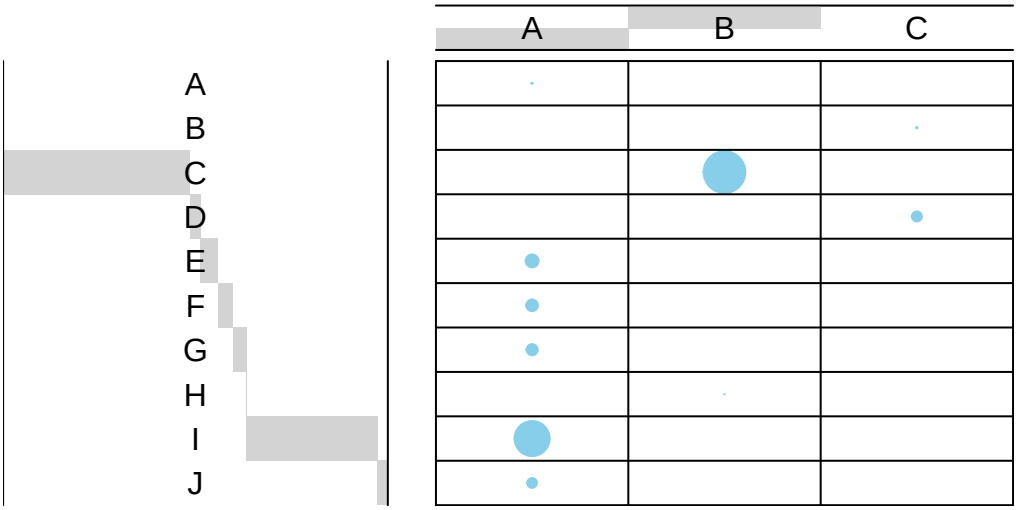
Categorical variables: v47-v110

```
# Contingency table
tab.v47.v110 <- table(train[,.(v47, v110),])
tab.v47.v110
```

```
##    v110
## v47     A     B     C
##   A    38     0     0
##   B     0     0    50
##   C     0 55425     0
##   D     0     0  3157
##   E  5301     0     0
##   F  4322     0     0
##   G  3946     0     0
##   H     0     1     0
##   I 39071     0     0
##   J  3010     0     0
```

```
# Draw balloon plot
balloonplot(t(tab.v47.v110), main ="Categorical variable: v91-107", xlab ="",
            ylab="", label = FALSE,  show.margins = FALSE)
```

Categorical variable: v91−107

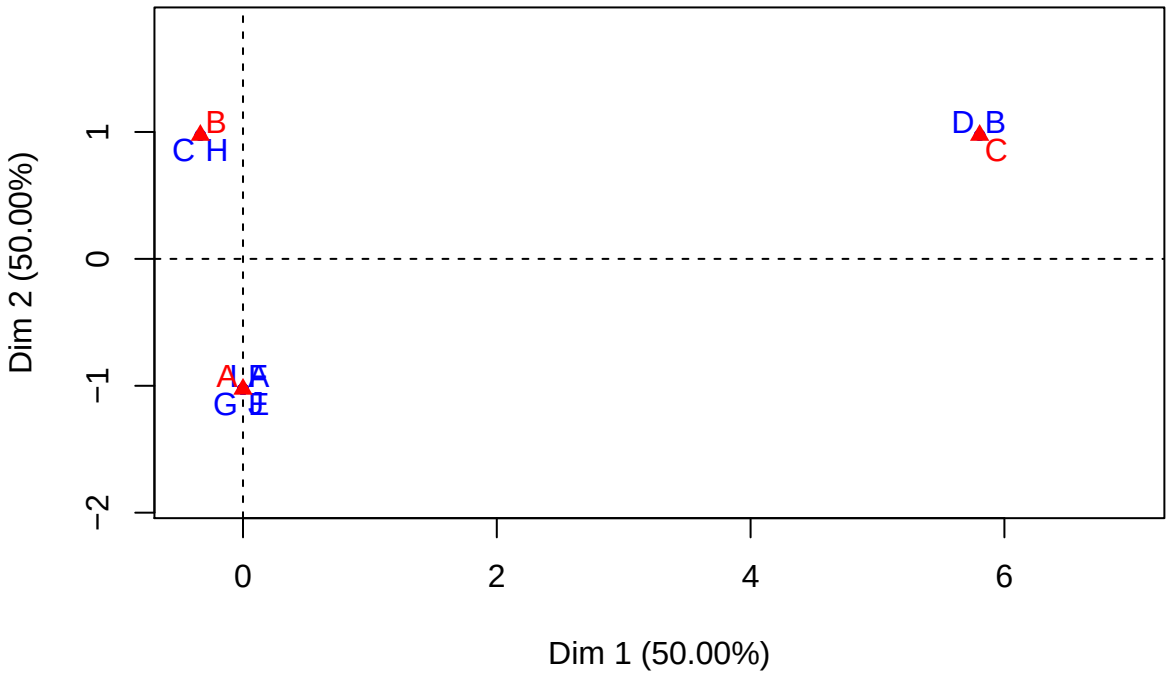



```
# Chi-squared test
chisq.test(tab.v47.v110)
```

```
## Warning in chisq.test(tab.v47.v110): Chi-squared approximation may be
## incorrect
```

```
##
##  Pearson's Chi-squared test
##
## data:  tab.v47.v110
## X-squared = 228640, df = 18, p-value < 2.2e-16
```

```
# Correspondence analysis(CA)
ca.v47.v110 <-  CA(tab.v47.v110, graph=TRUE)
```




```
summary(ca.v47.v110, nb.dec = 2, ncp = 2)
```

```
##
## Call:
## CA(X = tab.v47.v110, graph = TRUE)
##
## The chi square of independence between the two variables is equal to 228642 (p-value =  0 ).
##
## Eigenvalues
##                       Dim.1 Dim.2
## Variance                  1     1
## % of var.                50    50
## Cumulative % of var.     50   100
##
## Rows
##     Iner*1000    Dim.1    ctr   cos2    Dim.2    ctr   cos2
## A |      0.35 |   0.00   0.00   0.00 |  -1.03   0.03   1.00 |
## B |     15.15 |   5.80   1.47   0.97 |   0.97   0.04   0.03 |
## C |    515.16 |  -0.34   5.47   0.11 |   0.97  46.05   0.89 |
## D |    956.79 |   5.80  93.06   0.97 |   0.97   2.62   0.03 |
## E |     48.82 |   0.00   0.00   0.00 |  -1.03   4.88   1.00 |
## F |     39.81 |   0.00   0.00   0.00 |  -1.03   3.98   1.00 |
## G |     36.34 |   0.00   0.00   0.00 |  -1.03   3.63   1.00 |
## H |      0.01 |  -0.34   0.00   0.11 |   0.97   0.00   0.89 |
## I |    359.84 |   0.00   0.00   0.00 |  -1.03  35.98   1.00 |
## J |     27.72 |   0.00   0.00   0.00 |  -1.03   2.77   1.00 |
##
## Columns
##     Iner*1000    Dim.1    ctr   cos2    Dim.2    ctr   cos2
## A |    512.88 |   0.00   0.00   0.00 |  -1.03  51.29   1.00 |
```

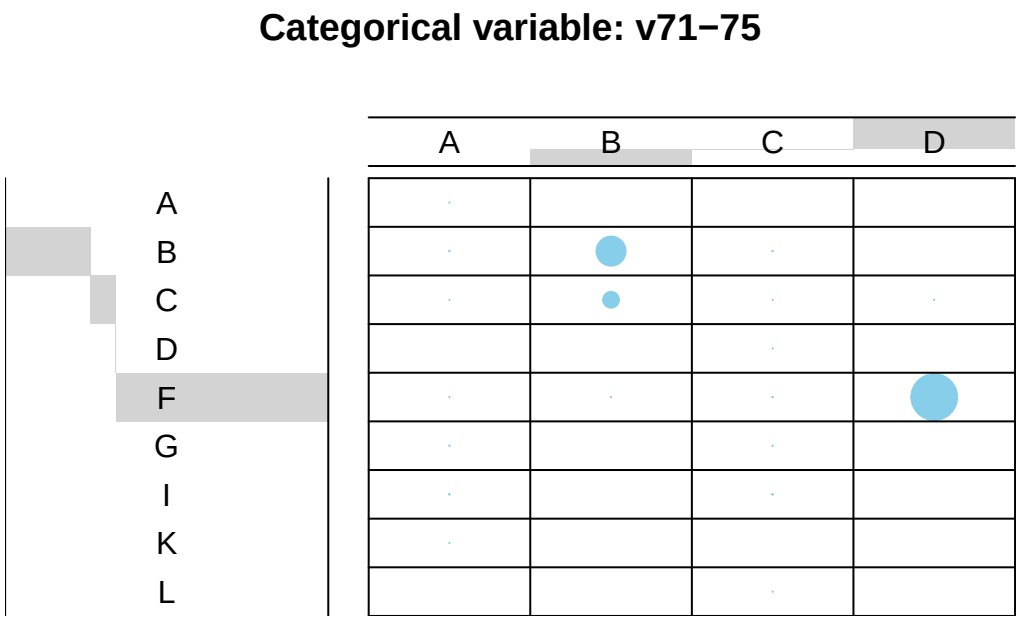
```
## B |    515.17 |  -0.34   5.47   0.11 |   0.97  46.05   0.89 |
## C |    971.95 |   5.80  94.53   0.97 |   0.97   2.66   0.03 |
```

Categorical variables: v71-v75

```
# Contingency table
tab.v71.v75 <- table(train[,.(v71, v75),])
tab.v71.v75
```

```
##    v75
## v71     A     B     C     D
##   A     1     0     0     0
##   B     4 30247     4     0
##   C     1  8944     1     1
##   D     0     0     1     0
##   F     1     1     6 75086
##   G     3     0     2     0
##   I     7     0     9     0
##   K     1     0     0     0
##   L     0     0     1     0
```

```
# Draw balloon plot
balloonplot(t(tab.v71.v75), main ="Categorical variable: v71-75", xlab ="",
            ylab="", label = FALSE,  show.margins = FALSE)
```



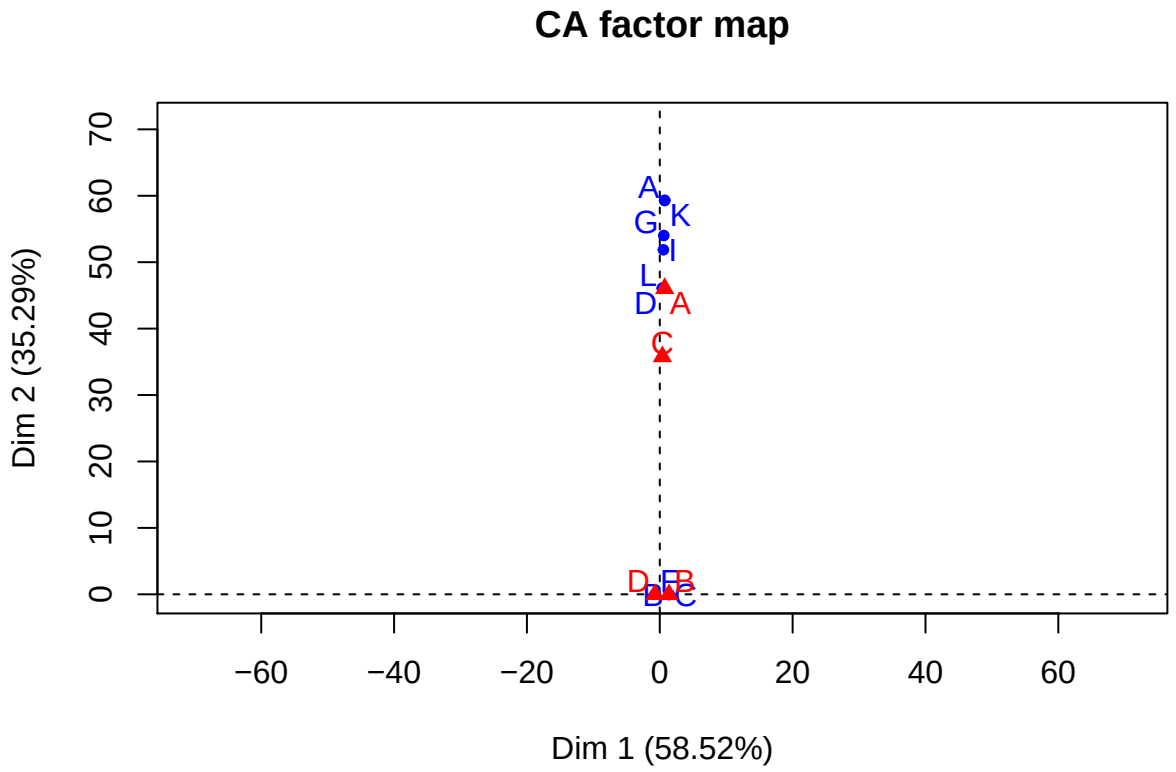

```
# Chi-squared test
chisq.test(tab.v71.v75)
```

```
## Warning in chisq.test(tab.v71.v75): Chi-squared approximation may be
## incorrect
```

```
##
##  Pearson's Chi-squared test
##
## data:  tab.v71.v75
```

```
## X-squared = 195320, df = 24, p-value < 2.2e-16
```

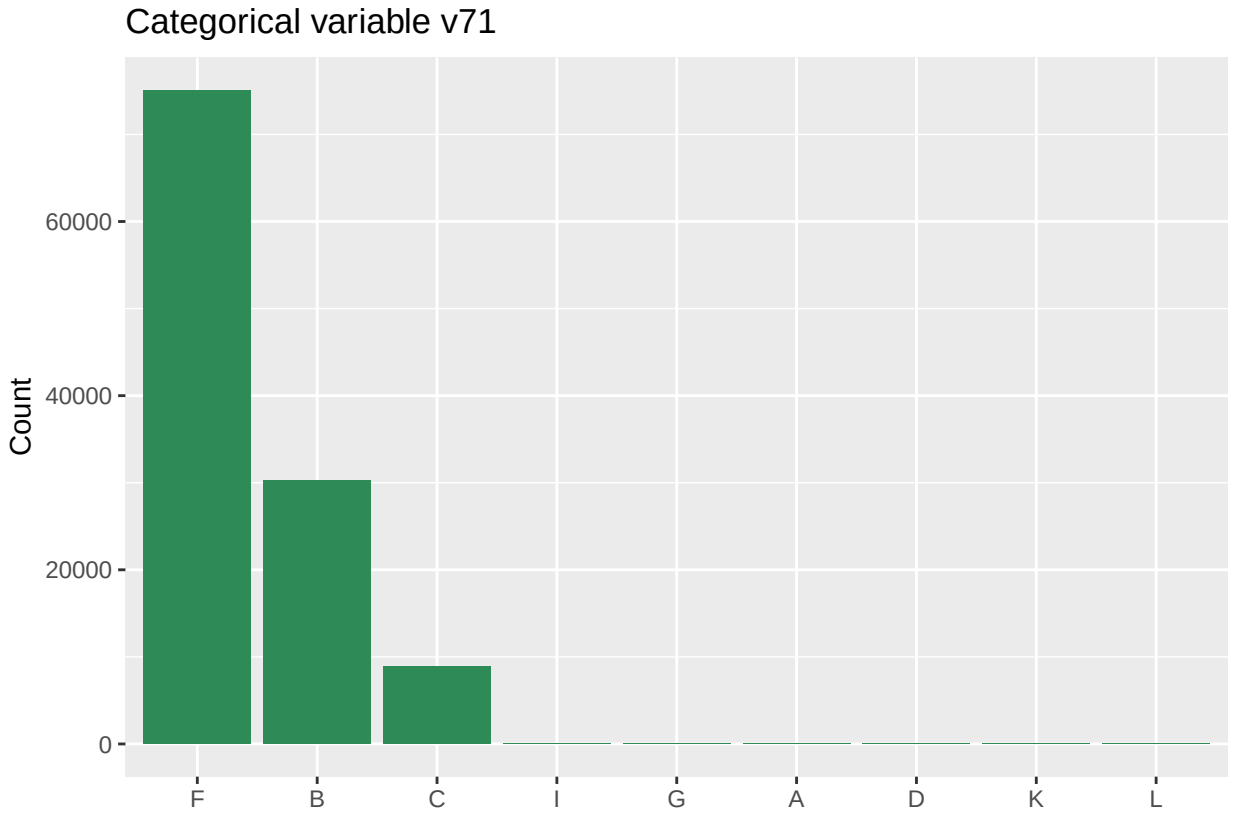
```
# Correspondence analysis(CA)
ca.v71.v75 <-  CA(tab.v71.v75, graph=TRUE)
```

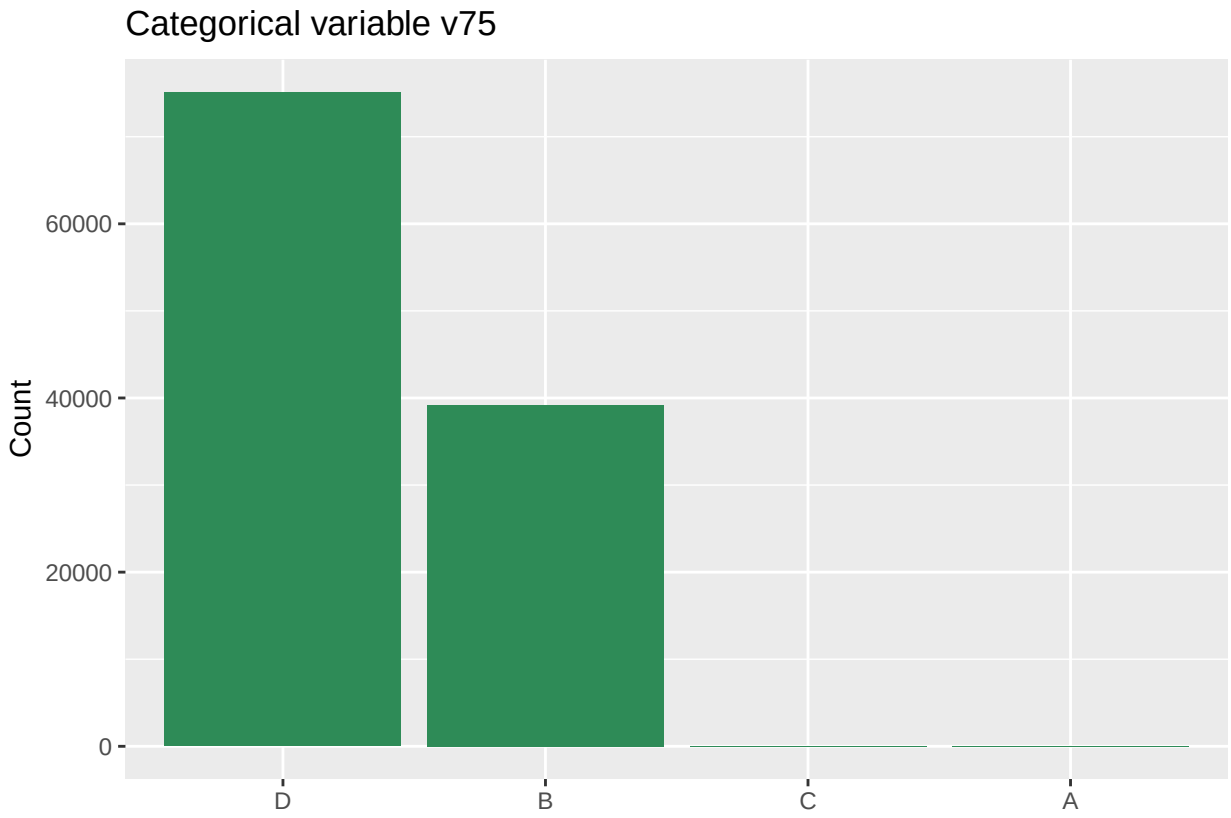



```
summary(ca.v71.v75, nb.dec = 2, ncp = 2)
```

```
##
## Call:
## CA(X = tab.v71.v75, graph = TRUE)
##
## The chi square of independence between the two variables is equal to 195318.7 (p-value =  0 ).
##
## Eigenvalues
##                        Dim.1  Dim.2  Dim.3
## Variance                1.00   0.60   0.11
## % of var.              58.52  35.29   6.19
## Cumulative % of var.   58.52  93.81 100.00
##
## Rows
##     Iner*1000    Dim.1    ctr   cos2    Dim.2    ctr   cos2
## A |     55.55 |   0.74   0.00   0.00 |  59.30   5.10   0.55 |
## B |    506.96 |   1.38  50.70   1.00 |  -0.02   0.02   0.00 |
## C |    149.88 |   1.38  14.99   1.00 |  -0.02   0.01   0.00 |
## D |     41.66 |   0.40   0.00   0.00 |  46.08   3.08   0.45 |
## F |    343.03 |  -0.72  34.31   1.00 |  -0.01   0.01   0.00 |
## G |    133.29 |   0.60   0.00   0.00 |  54.01  21.16   0.96 |
## I |    380.94 |   0.55   0.00   0.00 |  51.87  62.44   0.99 |
## K |     55.55 |   0.74   0.00   0.00 |  59.30   5.10   0.55 |
## L |     41.66 |   0.40   0.00   0.00 |  46.08   3.08   0.45 |
```

```
##
## Columns
##     Iner*1000    Dim.1    ctr   cos2    Dim.2    ctr   cos2
## A |    381.13 |   0.74   0.01   0.00 |  46.05  55.37   0.88 |
## B |    656.87 |   1.38  65.68   1.00 |  -0.03   0.04   0.00 |
## C |    327.44 |   0.40   0.00   0.00 |  35.78  44.58   0.82 |
## D |    343.07 |  -0.72  34.31   1.00 |  -0.01   0.01   0.00 |
```





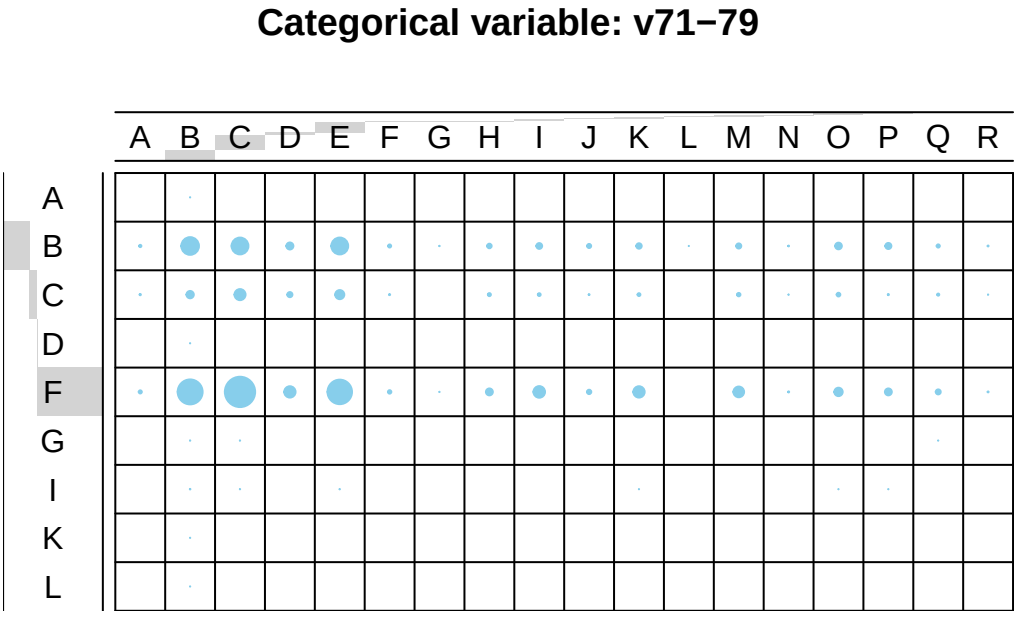
Categorical varibles: v71-v79

```
# Contingency table
tab.v71.v79 <- table(train[,.(v71, v79),])
tab.v71.v79
```

```
##    v79
## v71     A     B     C     D     E     F     G     H     I     J     K
##   A     0     1     0     0     0     0     0     0     0     0     0
##   B   130  8135  7389  1286  7422   243     4   533   867   427   752
##   C    40  1400  3189   677  2152    26     0   233   203    20   210
##   D     0     1     0     0     0     0     0     0     0     0     0
##   F   247 16252 23979  3339 15680   302     2  1238  3491   486  3345
##   G     0     2     2     0     0     0     0     0     0     0     0
##   I     0     8     2     0     3     0     0     0     0     0     1
##   K     0     1     0     0     0     0     0     0     0     0     0
##   L     0     1     0     0     0     0     0     0     0     0     0
##    v79
## v71     L     M     N     O     P     Q     R
##   A     0     0     0     0     0     0     0
##   B     1   666    17  1152   951   257    23
##   C     0   299     3   351    31   111     2
##   D     0     0     0     0     0     0     0
##   F     0  2981    29  1827  1234   637    25
##   G     0     0     0     0     0     1     0
##   I     0     0     0     1     1     0     0
##   K     0     0     0     0     0     0     0
```

```
##   L     0     0     0     0     0     0     0
```

```
# Draw balloon plot
balloonplot(t(tab.v71.v79), main ="Categorical variable: v71-79", xlab ="",
            ylab="", label = FALSE,  show.margins = FALSE)
```



Numerical variables

Dealing with missing values

```
library(caret)
library(mice)

# Pattern of missing values
md.pattern(train.num)
```

```
##       v38 v62 v72 v129 v14 v114 v10 v12 v50 v34 v40 v21    v8   v25   v46
## 62561   1   1   1    1   1    1   1   1   1   1   1   1     1     1     1
##   306   1   1   1    1   1    1   1   1   1   1   1   0     1     1     1
##    15   1   1   1    1   1    1   1   1   1   1   1   1     1     1     1
##   615   1   1   1    1   1    1   1   1   1   1   1   1     1     1     1
##    22   1   1   1    1   1    1   1   1   1   1   1   1     1     1     1
##     1   1   1   1    1   1    1   1   1   1   1   1   1     1     1     1
##     2   1   1   1    1   1    1   1   1   1   1   1   1     1     1     1
##     1   1   1   1    1   1    1   1   1   1   1   1   1     1     1     1
##     3   1   1   1    1   1    1   1   1   1   1   1   1     1     1     1
##    31   1   1   1    1   1    1   1   1   1   1   1   1     1     1     1
##    13   1   1   1    1   1    1   0   0   0   0   0   0     1     1     1
##     1   1   1   1    1   1    1   1   1   1   1   1   1     1     1     1
##     5   1   1   1    1   1    1   1   1   1   1   1   1     1     1     1
##     1   1   1   1    1   1    1   1   1   1   1   1   1     1     1     1
##     6   1   1   1    1   1    1   1   1   1   1   1   1     1     1     1
##     1   1   1   1    1   0    1   0   0   0   0   0   0     1     1     1
##     1   1   1   1    1   1    1   1   1   1   1   1   1     1     1     1
```

```
##      3   1   1   1    1   1    1   1   1   1   1   1   1     1     1     1
##      1   1   1   1    1   1    1   1   1   1   1   1   1     1     1     1
##      1   1   1   1    1   1    1   1   1   1   1   1   1     1     1     1
##      1   1   1   1    1   1    1   1   1   1   1   1   1     1     1     1
##      1   1   1   1    1   1    1   1   1   1   1   1   1     1     1     1
##      1   1   1   1    1   1    1   1   1   1   1   1   1     1     1     1
##      2   1   1   1    1   1    1   1   1   1   1   1   1     1     1     1
##      5   1   1   1    1   1    1   1   1   1   1   1   1     1     1     1
##      1   1   1   1    1   1    1   1   1   1   1   1   1     1     1     1
##      1   1   1   1    1   1    1   1   1   1   1   1   1     1     1     1
##      1   1   1   1    1   1    1   1   1   1   1   1   1     1     1     1
##      2   1   1   1    1   1    1   1   1   1   1   1   1     1     1     1
##      5   1   1   1    1   1    1   1   1   1   1   1   1     1     1     1
##      1   1   1   1    1   1    1   1   1   1   1   1   1     1     1     1
##    874   1   1   1    1   1    1   1   1   1   1   1   1     0     0     0
##      1   1   1   1    1   1    1   1   1   1   1   1   1     1     1     1
##      1   1   1   1    1   1    1   1   1   1   1   1   1     1     1     1
##      2   1   1   1    1   1    1   1   1   1   1   1   1     1     1     1
##     36   1   1   1    1   1    1   1   1   1   1   1   1     1     1     1
##   2051   1   1   1    1   1    1   1   1   1   1   1   1     1     1     1
##  47715   1   1   1    1   1    1   1   1   1   1   1   1     0     0     0
##     30   1   1   1    1   1    0   1   1   1   0   0   1     0     0     0
##          0   0   0    0   4   30  84  86  86 111 111 611 48619 48619 48619
##          v54   v63   v89  v124    v5   v36   v81   v82  v108  v109  v117
##  62561     1     1     1     1     1     1     1     1     1     1     1
##    306     1     1     1     1     1     1     1     1     1     1     1
##     15     1     1     1     1     1     1     1     1     1     1     1
##    615     1     1     1     1     1     1     1     1     1     1     1
##     22     1     1     1     1     1     1     1     1     1     1     1
##      1     1     1     1     1     1     1     1     1     1     1     1
##      2     1     1     1     1     1     1     1     1     1     1     1
##      1     1     1     1     1     1     1     1     1     1     1     1
##      3     1     1     1     1     1     1     1     1     1     1     1
##     31     1     1     1     1     1     1     1     1     1     1     1
##     13     1     1     1     1     1     1     1     1     1     1     1
##      1     1     1     1     1     1     1     1     1     1     1     1
##      5     1     1     1     1     1     1     1     1     1     1     1
##      1     1     1     1     1     1     1     1     1     1     1     1
##      6     1     1     1     1     1     1     1     1     1     1     1
##      1     1     1     1     1     1     1     1     1     1     1     1
##      1     1     1     1     1     1     1     1     1     1     1     1
##      3     1     1     1     1     1     1     1     1     1     1     1
##      1     1     1     1     1     1     1     1     1     1     1     1
##      1     1     1     1     1     1     1     1     1     1     1     1
##      1     1     1     1     1     1     1     1     1     1     1     1
##      1     1     1     1     1     1     1     1     1     1     1     1
##      1     1     1     1     1     1     1     1     1     1     1     1
##      2     1     1     1     1     1     1     1     1     1     1     1
##      5     1     1     1     1     0     0     0     0     0     0     0
##      1     1     1     1     1     1     1     1     1     1     1     1
##      1     1     1     1     1     1     1     1     1     1     1     1
##      1     1     1     1     1     1     1     1     1     1     1     1
##      2     1     1     1     1     1     1     1     1     1     1     1
##      5     1     1     1     1     1     1     1     1     1     1     1
```

```
##      1      1      1      1      1      1      1      1      1      1      1      1
##    874      0      0      0      0      0      0      0      0      0      0      0
##      1      1      1      1      1      1      1      1      1      1      1      1
##      1      1      1      1      1      1      1      1      1      1      1      1
##      2      1      1      1      1      1      1      1      1      1      1      1
##     36      1      1      1      1      1      1      1      1      1      1      1
##   2051      1      1      1      1      1      1      1      1      1      1      1
##  47715      0      0      0      0      0      0      0      0      0      0      0
##     30      0      0      0      0      0      0      0      0      0      0      0
##         48619  48619  48619  48619  48624  48624  48624  48624  48624  48624  48624
##          v128    v70    v98   v105    v87     v2     v4    v17    v44    v48    v59
##  62561      1      1      1      1      1      1      1      1      1      1      1
##    306      1      1      1      1      1      1      1      1      1      1      1
##     15      1      1      0      1      1      1      1      1      1      1      1
##    615      1      1      1      1      1      1      1      1      1      1      1
##     22      1      1      1      0      0      1      1      1      1      1      1
##      1      1      0      1      0      0      1      1      1      1      1      1
##      2      1      1      0      0      0      1      1      1      1      1      1
##      1      1      1      0      1      1      1      1      1      1      1      1
##      3      1      0      0      0      0      1      1      1      1      1      1
##     31      1      1      1      1      1      1      1      1      1      1      1
##     13      1      1      1      1      1      1      1      1      1      1      1
##      1      1      1      0      0      0      1      1      1      1      1      1
##      5      1      1      1      1      1      1      1      1      1      1      1
##      1      1      1      0      1      1      1      1      1      1      1      1
##      6      1      1      1      1      1      1      1      1      1      1      1
##      1      1      1      1      1      1      1      1      1      1      1      1
##      1      1      0      0      0      0      1      1      1      1      1      1
##      3      1      0      1      0      0      1      1      1      1      1      1
##      1      1      1      0      0      0      1      1      1      1      1      1
##      1      1      0      0      0      0      1      1      1      1      1      1
##      1      1      1      0      0      0      1      1      1      1      1      1
##      1      1      1      1      1      1      1      1      1      1      1      1
##      1      1      1      1      1      1      1      1      1      1      1      1
##      2      1      1      1      1      1      1      1      1      1      1      1
##      5      0      0      0      1      0      1      1      1      1      1      1
##      1      1      0      0      0      0      1      1      1      1      1      1
##      1      1      1      1      1      1      1      1      1      1      1      1
##      1      1      1      1      1      1      1      1      1      1      1      1
##      2      1      0      0      0      0      1      1      1      1      1      1
##      5      1      1      1      1      1      1      1      1      1      1      1
##      1      1      1      1      1      1      1      1      1      1      1      1
##    874      0      0      0      0      0      1      1      1      1      1      1
##      1      1      1      1      1      1      1      1      1      1      1      1
##      1      1      1      1      1      1      1      1      1      1      1      1
##      2      1      1      1      1      1      1      1      1      1      1      1
##     36      1      1      1      1      1      1      1      1      1      1      1
##   2051      1      1      1      1      1      0      0      0      0      0      0
##  47715      0      0      0      0      0      0      0      0      0      0      0
##     30      0      0      0      0      0      0      0      0      0      0      0
##         48624  48636  48654  48658  48663  49796  49796  49796  49796  49796  49796
##           v61    v64    v76   v101   v106     v1     v6     v7    v13    v18    v26
##  62561      1      1      1      1      1      1      1      1      1      1      1
##    306      1      1      1      1      1      1      1      1      1      1      1
```

```
##    15     1     1     1     1     1     1     1     1     1     1     1
##   615     1     1     1     1     1     1     1     1     1     1     1
##    22     1     1     1     1     1     1     1     1     1     1     1
##     1     1     1     1     1     1     1     1     1     1     1     1
##     2     1     1     1     1     1     1     1     1     1     1     1
##     1     1     1     1     1     1     1     1     1     1     1     1
##     3     1     1     1     1     1     1     1     1     1     1     1
##    31     1     1     1     1     1     1     1     1     1     1     1
##    13     1     1     1     1     1     1     1     1     1     1     1
##     1     1     1     1     1     1     1     1     1     1     1     1
##     5     1     1     1     1     1     1     1     1     1     1     1
##     1     1     1     1     1     1     1     1     1     1     1     1
##     6     1     1     1     1     1     1     1     1     1     1     1
##     1     1     1     1     1     1     1     1     1     1     1     1
##     1     1     1     1     1     1     1     1     1     1     1     1
##     3     1     1     1     1     1     1     1     1     1     1     1
##     1     1     1     1     1     1     1     1     1     1     1     1
##     1     1     1     1     1     1     1     1     1     1     1     1
##     1     1     1     1     1     1     1     1     1     1     1     1
##     1     1     1     1     1     1     1     1     1     1     1     1
##     2     1     1     1     1     1     1     1     1     1     1     1
##     5     1     1     1     1     1     1     1     1     1     1     1
##     1     1     1     1     1     1     1     1     1     1     1     1
##     1     1     1     1     1     1     1     1     1     1     1     1
##     1     1     1     1     1     1     1     1     1     1     1     1
##     2     1     1     1     1     1     1     1     1     1     1     1
##     5     1     1     1     1     1     1     1     1     1     1     1
##     1     1     1     1     1     1     1     1     1     1     1     1
##   874     1     1     1     1     1     1     1     1     1     1     1
##     1     1     1     1     1     1     1     1     1     1     1     1
##     1     1     1     1     1     1     1     1     1     1     1     1
##     2     1     1     1     1     1     1     1     1     1     1     1
##    36     1     1     1     1     1     0     0     0     0     0     0
##  2051     0     0     0     0     0     0     0     0     0     0     0
## 47715     0     0     0     0     0     0     0     0     0     0     0
##    30     0     0     0     0     0     0     0     0     0     0     0
##       49796 49796 49796 49796 49796 49832 49832 49832 49832 49832 49832
##         v27   v28   v29   v32   v33   v35   v41   v42   v45   v49   v55
## 62561     1     1     1     1     1     1     1     1     1     1     1
##   306     1     1     1     1     1     1     1     1     1     1     1
##    15     1     1     1     1     1     1     1     1     1     1     1
##   615     1     1     1     1     1     1     1     1     1     1     1
##    22     1     1     1     1     1     1     1     1     1     1     1
##     1     1     1     1     1     1     1     1     1     1     1     1
##     2     1     1     1     1     1     1     1     1     1     1     1
##     1     1     1     1     1     1     1     1     1     1     1     1
##     3     1     1     1     1     1     1     1     1     1     1     1
##    31     1     1     1     1     1     1     1     1     1     1     1
##    13     1     1     1     1     1     1     1     1     1     1     1
##     1     1     1     1     1     1     1     1     1     1     1     1
##     5     1     1     1     1     1     1     1     1     1     1     1
##     1     1     1     1     1     1     1     1     1     1     1     1
##     6     1     1     1     1     1     1     1     1     1     1     1
```

```
##     1     1     1     1     1     1     1     1     1     1     1     1
##     1     1     1     1     1     1     1     1     1     1     1     1
##     3     1     1     1     1     1     1     1     1     1     1     1
##     1     1     1     1     1     1     1     1     1     1     1     1
##     1     1     1     1     1     1     1     1     1     1     1     1
##     1     1     1     1     1     1     1     1     1     1     1     1
##     1     1     1     1     1     1     1     1     1     1     1     1
##     1     1     1     1     1     1     1     1     1     1     1     1
##     2     1     1     1     1     1     1     1     1     1     1     1
##     5     1     1     1     1     1     1     1     1     1     1     1
##     1     1     1     1     1     1     1     1     1     1     1     1
##     1     1     1     1     1     1     1     1     1     1     1     1
##     1     1     1     1     1     1     1     1     1     1     1     1
##     2     1     1     1     1     1     1     1     1     1     1     1
##     5     1     1     1     1     1     1     1     1     1     1     1
##     1     1     1     1     1     1     1     1     1     1     1     1
##   874     1     1     1     1     1     1     1     1     1     1     1
##     1     1     1     1     1     1     1     1     1     1     1     1
##     1     1     1     1     1     1     1     1     1     1     1     1
##     2     1     1     1     1     1     1     1     1     1     1     1
##    36     0     0     0     0     0     0     0     0     0     0     0
##  2051     0     0     0     0     0     0     0     0     0     0     0
## 47715     0     0     0     0     0     0     0     0     0     0     0
##    30     0     0     0     0     0     0     0     0     0     0     0
##       49832 49832 49832 49832 49832 49832 49832 49832 49832 49832 49832
##         v57   v60   v67   v77   v83   v84   v86   v88   v93   v94   v96
## 62561     1     1     1     1     1     1     1     1     1     1     1
##   306     1     1     1     1     1     1     1     1     1     1     1
##    15     1     1     1     1     1     1     1     1     1     1     1
##   615     1     1     1     1     1     1     1     1     1     1     1
##    22     1     1     1     1     1     1     1     1     1     1     1
##     1     1     1     1     1     1     1     1     1     1     1     1
##     2     1     1     1     1     1     1     1     1     1     1     1
##     1     1     1     1     1     1     1     1     1     1     1     1
##     3     1     1     1     1     1     1     1     1     1     1     1
##    31     1     1     1     1     1     1     1     1     1     1     1
##    13     1     1     1     1     1     1     1     1     1     1     1
##     1     1     1     1     1     1     1     1     1     1     1     1
##     5     1     1     1     1     1     1     1     1     1     1     1
##     1     1     1     1     1     1     1     1     1     1     1     1
##     6     1     1     1     1     1     1     1     1     1     1     1
##     1     1     1     1     1     1     1     1     1     1     1     1
##     1     1     1     1     1     1     1     1     1     1     1     1
##     3     1     1     1     1     1     1     1     1     1     1     1
##     1     1     1     1     1     1     1     1     1     1     1     1
##     1     1     1     1     1     1     1     1     1     1     1     1
##     1     1     1     1     1     1     1     1     1     1     1     1
##     1     1     1     1     1     1     1     1     1     1     1     1
##     1     1     1     1     1     1     1     1     1     1     1     1
##     2     1     1     1     1     1     1     1     1     1     1     1
##     5     1     1     1     1     1     1     1     1     1     1     1
##     1     1     1     1     1     1     1     1     1     1     1     1
##     1     1     1     1     1     1     1     1     1     1     1     1
##     1     1     1     1     1     1     1     1     1     1     1     1
```

```
##     2     1     1     1     1     1     1     1     1     1     1     1
##     5     1     1     1     1     1     1     1     1     1     1     1
##     1     1     1     1     1     1     1     1     1     1     1     1
##   874     1     1     1     1     1     1     1     1     1     1     1
##     1     1     1     1     1     1     1     1     1     1     1     1
##     1     1     1     1     1     1     1     1     1     1     1     1
##     2     1     1     1     1     1     1     1     1     1     1     1
##    36     0     0     0     0     0     0     0     0     0     0     0
##  2051     0     0     0     0     0     0     0     0     0     0     0
## 47715     0     0     0     0     0     0     0     0     0     0     0
##    30     0     0     0     0     0     0     0     0     0     0     0
##       49832 49832 49832 49832 49832 49832 49832 49832 49832 49832 49832
##         v99  v103  v104  v111  v126  v127   v11   v15   v39   v43   v53
## 62561     1     1     1     1     1     1     1     1     1     1     1
##   306     1     1     1     1     1     1     1     1     1     1     1
##    15     1     1     1     1     1     1     1     1     1     1     1
##   615     1     1     1     1     1     1     1     1     1     1     1
##    22     1     1     1     1     1     1     1     1     1     1     1
##     1     1     1     1     1     1     1     1     1     1     1     1
##     2     1     1     1     1     1     1     1     1     1     1     1
##     1     1     1     1     1     1     1     1     1     1     1     1
##     3     1     1     1     1     1     1     1     1     1     1     1
##    31     1     1     1     1     1     1     1     1     1     1     1
##    13     1     1     1     1     1     1     1     1     1     1     1
##     1     1     1     1     1     1     1     1     1     1     1     1
##     5     1     1     1     1     1     1     1     1     1     1     1
##     1     1     1     1     1     1     1     1     1     1     1     1
##     6     1     1     1     1     1     1     1     1     1     1     1
##     1     1     1     1     1     1     1     1     1     1     1     1
##     1     1     1     1     1     1     1     1     1     1     1     1
##     3     1     1     1     1     1     1     1     1     1     1     1
##     1     1     1     1     1     1     1     1     1     1     1     1
##     1     1     1     1     1     1     1     1     1     1     1     1
##     1     1     1     1     1     1     1     1     1     1     1     1
##     1     1     1     1     1     1     1     1     1     1     1     1
##     1     1     1     1     1     1     1     1     1     1     1     1
##     2     1     1     1     1     1     1     1     1     1     1     1
##     5     1     1     1     1     1     1     1     1     1     1     1
##     1     1     1     1     1     1     1     1     1     1     1     1
##     1     1     1     1     1     1     1     1     1     1     1     1
##     1     1     1     1     1     1     1     1     1     1     1     1
##     2     1     1     1     1     1     1     1     1     1     1     1
##     5     1     1     1     1     1     1     1     1     1     1     1
##     1     1     1     1     1     1     1     1     1     1     1     1
##   874     1     1     1     1     1     1     1     1     1     1     1
##     1     1     1     1     1     1     1     0     0     0     0     0
##     1     1     1     1     1     1     1     0     0     0     0     0
##     2     1     1     1     1     1     1     0     0     0     0     0
##    36     0     0     0     0     0     0     0     0     0     0     0
##  2051     0     0     0     0     0     0     0     0     0     0     0
## 47715     0     0     0     0     0     0     0     0     0     0     0
##    30     0     0     0     0     0     0     0     0     0     0     0
##       49832 49832 49832 49832 49832 49832 49836 49836 49836 49836 49836
##         v58   v68   v73   v90  v100  v116  v120   v20   v65  v121   v19
```

```
## 62561     1     1     1     1     1     1     1     1     1     1     1
##   306     1     1     1     1     1     1     1     1     1     1     1
##    15     1     1     1     1     1     1     1     1     1     1     1
##   615     1     1     1     1     1     1     1     1     1     1     1
##    22     1     1     1     1     1     1     1     1     1     1     1
##     1     1     1     1     1     1     1     1     1     1     1     1
##     2     1     1     1     1     1     1     1     1     1     1     1
##     1     1     1     1     1     1     1     1     1     1     1     1
##     3     1     1     1     1     1     1     1     1     1     1     1
##    31     1     1     1     1     1     1     1     1     1     1     1
##    13     1     1     1     1     1     1     1     1     1     1     1
##     1     1     1     1     1     1     1     1     1     1     1     1
##     5     1     1     1     1     1     1     1     1     1     1     1
##     1     1     1     1     1     1     1     1     1     1     1     1
##     6     1     1     1     1     1     1     1     1     1     1     1
##     1     1     1     1     1     1     1     1     1     1     1     1
##     1     1     1     1     1     1     1     1     1     1     1     1
##     3     1     1     1     1     1     1     1     1     1     1     1
##     1     1     1     1     1     1     1     1     1     1     1     1
##     1     1     1     1     1     1     1     1     1     1     1     1
##     1     1     1     1     1     1     1     1     1     1     1     1
##     1     1     1     1     1     1     1     1     0     0     0     1
##     1     1     1     1     1     1     1     1     1     1     1     1
##     2     1     1     1     1     1     1     1     1     1     1     1
##     5     1     1     1     1     1     1     1     1     1     1     1
##     1     1     1     1     1     1     1     1     1     1     1     1
##     1     1     1     1     0     1     1     1     0     0     0     1
##     1     1     1     1     1     1     1     1     1     1     1     0
##     2     1     1     1     1     1     1     1     0     0     0     1
##     5     1     1     1     1     1     1     1     1     1     1     0
##     1     1     1     1     1     1     1     1     1     1     1     0
##   874     1     1     1     1     1     1     1     1     1     1     1
##     1     0     0     0     1     0     0     0     0     0     0     0
##     1     0     0     0     0     0     0     0     0     0     0     0
##     2     0     0     0     0     0     0     0     0     0     0     0
##    36     0     0     0     0     0     0     0     0     0     0     0
##  2051     0     0     0     0     0     0     0     0     0     0     0
## 47715     0     0     0     0     0     0     0     0     0     0     0
##    30     0     0     0     0     0     0     0     0     0     0     0
##       49836 49836 49836 49836 49836 49836 49836 49840 49840 49840 49843
##         v37   v92   v95   v97  v118  v130    v9   v80  v122   v16   v69
## 62561     1     1     1     1     1     1     1     1     1     1     1
##   306     1     1     1     1     1     1     1     1     1     1     1
##    15     1     1     1     1     1     1     1     1     1     1     1
##   615     1     1     1     1     1     1     1     1     1     1     1
##    22     1     1     1     1     1     1     1     1     1     1     1
##     1     1     1     1     1     1     1     1     1     1     1     1
##     2     1     1     1     1     1     1     1     1     1     1     1
##     1     1     1     1     1     1     1     0     0     0     1     1
##     3     1     1     1     1     1     1     1     1     1     1     1
##    31     1     1     1     1     1     1     1     1     1     0     0
##    13     1     1     1     1     1     1     1     1     1     1     1
##     1     1     1     1     1     1     1     0     0     0     1     1
##     5     1     1     1     1     1     1     1     1     1     1     1
```

```
##     1     1     1     1     1     1     1     1     1     1     0     0
##     6     1     1     1     1     1     1     1     1     1     0     0
##     1     1     1     1     1     1     1     1     1     1     1     1
##     1     1     1     1     1     1     1     0     0     0     1     1
##     3     1     1     1     1     1     1     1     1     1     0     0
##     1     1     1     1     1     1     1     1     1     1     0     0
##     1     1     1     1     1     1     1     1     1     1     0     0
##     1     1     1     1     1     1     1     1     1     1     0     0
##     1     1     1     1     1     1     1     1     1     1     0     0
##     1     1     1     1     1     1     1     0     0     0     0     0
##     2     1     1     1     1     1     1     0     0     0     0     0
##     5     1     1     1     1     1     1     1     1     1     1     1
##     1     1     1     1     1     1     1     0     0     0     0     0
##     1     1     1     1     1     1     1     0     0     0     0     0
##     1     0     0     0     0     0     0     1     1     1     0     0
##     2     1     1     1     1     1     1     0     0     0     0     0
##     5     0     0     0     0     0     0     0     0     0     0     0
##     1     0     0     0     0     0     0     1     1     1     0     0
##   874     1     1     1     1     1     1     1     1     1     1     1
##     1     0     0     0     0     0     0     0     0     0     0     0
##     1     0     0     0     0     0     0     0     0     0     0     0
##     2     0     0     0     0     0     0     0     0     0     0     0
##    36     0     0     0     0     0     0     0     0     0     0     0
##  2051     0     0     0     0     0     0     0     0     0     0     0
## 47715     0     0     0     0     0     0     0     0     0     0     0
##    30     0     0     0     0     0     0     0     0     0     0     0
##       49843 49843 49843 49843 49843 49843 49851 49851 49851 49895 49895
##         v78  v115  v131   v23   v51  v123  v119   v85  v102
## 62561     1     1     1     1     1     1     1     1     1        0
##   306     1     1     1     1     1     1     1     1     1        1
##    15     1     1     1     1     1     1     1     1     1        1
##   615     1     1     1     1     1     1     1     1     0        1
##    22     1     1     1     1     1     1     1     1     1        2
##     1     1     1     1     1     1     1     1     1     1        3
##     2     1     1     1     1     1     1     1     1     1        3
##     1     1     1     1     1     1     1     1     1     1        4
##     3     1     1     1     1     1     1     1     1     0        5
##    31     0     0     0     1     1     1     1     1     1        5
##    13     1     1     1     1     1     1     1     1     1        6
##     1     1     1     1     1     1     1     1     1     1        6
##     5     1     1     1     0     0     0     0     0     0        6
##     1     0     0     0     1     1     1     1     1     1        6
##     6     0     0     0     1     1     1     1     1     0        6
##     1     1     1     1     1     1     1     1     1     1        7
##     1     1     1     1     1     1     1     1     1     1        7
##     3     0     0     0     1     1     1     1     1     1        8
##     1     0     0     0     1     1     1     1     1     1        8
##     1     0     0     0     1     1     1     1     1     1        9
##     1     0     0     0     1     1     1     1     1     0        9
##     1     0     0     0     1     1     1     1     1     0        9
##     1     0     0     0     1     1     1     1     1     0        9
##     2     0     0     0     1     1     1     1     0     0       10
##     5     1     1     1     1     1     1     1     1     1       11
##     1     0     0     0     1     1     1     1     1     1       12
```

```
##     1     0     0     0     1     1     1     1     1     0      13
##     1     0     0     0     1     1     1     1     1     0      13
##     2     0     0     0     1     1     1     1     1     1      15
##     5     0     0     0     1     1     1     1     1     0      16
##     1     0     0     0     1     0     0     0     0     0      17
##   874     1     1     1     0     0     0     0     0     0      25
##     1     0     0     0     1     1     1     0     0     0      32
##     1     0     0     0     1     1     1     0     0     0      33
##     2     0     0     0     1     0     0     0     0     0      35
##    36     0     0     0     1     1     1     1     1     1      64
##  2051     0     0     0     0     0     0     0     0     0      81
## 47715     0     0     0     0     0     0     0     0     0     100
##    30     0     0     0     0     0     0     0     0     0     103
##       49895 49895 49895 50675 50678 50678 50680 50682 51316 4967293
```

```
# Imputing missing values with mean
impute.mean <- function(x) {
  x[is.na(x)] <- mean(x, na.rm = TRUE)
  x
}
imputed.train.num <- train.num[, lapply(.SD, impute.mean)]
imputed.train.num
```

```
##                v1       v2       v4        v5       v6       v7         v8
##      1: 1.3357394 8.727474 3.921026  7.915266 2.599278 3.176895 0.01294147
##      2: 1.6306857 7.464411 4.145098  9.191265 2.436402 2.483921 2.30163049
##      3: 0.9438769 5.310079 4.410969  5.326159 3.979592 3.928571 0.01964513
##      4: 0.7974146 8.304757 4.225930 11.627438 2.097700 1.987549 0.17194670
##      5: 1.6306857 7.464411 4.145098  8.742359 2.436402 2.483921 1.49656859
##     ---
## 114317: 1.6306857 7.464411 4.145098  8.742359 2.436402 2.483921 1.49656859
## 114318: 1.6306857 7.464411 4.145098  8.742359 2.436402 2.483921 1.49656859
## 114319: 1.6306857 7.464411 4.145098 10.069277 2.436402 2.483921 0.32332423
## 114320: 1.6306857 7.464411 4.145098 10.106144 2.436402 2.483921 0.30922647
## 114321: 1.6197631 7.932978 4.640085  8.473141 2.351470 2.826766 3.47975411
##                v9       v10      v11      v12      v13       v14      v15
##      1:  9.999999 0.5032815 16.43411 6.085711 2.866830 11.636387 1.355013
##      2:  9.031859 1.3129099 15.44741 6.507647 3.798396 11.636386 2.080911
##      3: 12.666667 0.7658640 14.75610 6.384670 2.505589  9.603542 1.984127
##      4:  8.965516 6.5426695 16.34748 9.646653 3.903302 14.094723 1.945044
##      5:  9.031859 1.0503284 15.44741 6.320087 3.798396 10.991098 2.080911
##     ---
## 114317:  9.031859 1.4442006 15.44741 6.368061 3.798396 11.865255 2.080911
## 114318:  9.031859 6.2363237 15.44741 9.443324 3.798396 14.924483 2.080911
## 114319:  9.031859 2.0787749 15.44741 6.698925 3.798396 12.269012 2.080911
## 114320:  9.031859 1.2910286 15.44741 6.692204 3.798396 12.573678 2.080911
## 114321:  9.629630 0.8533913 14.95979 6.306396 4.195219 11.967826 2.422481
##               v16      v17       v18       v19      v20      v21
##      1: 8.571429 3.670350 0.1067204 0.1488831 18.86928 7.730923
##      2: 4.923222 3.832270 0.8410455 0.2223005 17.77359 6.763110
##      3: 5.882353 3.170847 0.2445410 0.1442584 17.95233 5.245035
##      4: 5.517242 3.610789 1.2241139 0.2316304 18.37641 7.517125
##      5: 4.923222 3.832270 0.8410455 0.2223005 17.77359 6.414567
##     ---
## 114317: 4.923222 3.832270 0.8410455 0.2223005 17.77359 7.088172
```

```
## 114318: 4.923222 3.832270 0.8410455 0.2223005 17.77359 8.455263
## 114319: 4.923222 3.832270 0.8410455 0.2223005 17.77359 6.570625
## 114320: 4.923222 3.832270 0.8410455 0.2223005 17.77359 7.730751
## 114321: 3.168318 4.892720 0.5943879 0.1625732 18.85523 7.496000
##                   v23       v25      v26      v27       v28      v29
##      1: -1.716131e-08 0.1394116 1.720818 3.393503 0.5901219 8.880867
##      2:  1.093088e+00 3.0561440 1.876031 2.743454 5.0933280 8.206416
##      3: -2.785053e-07 0.1139970 2.244897 5.306122 0.8360052 7.499999
##      4: -4.805344e-07 0.1488431 1.308269 2.303640 8.9266621 8.874521
##      5:  1.093088e+00 1.6981288 1.876031 2.743454 5.0933280 8.206416
##     ---
## 114317:  1.093088e+00 1.6981288 1.876031 2.743454 5.0933280 8.206416
## 114318:  1.093088e+00 1.6981288 1.876031 2.743454 5.0933280 8.206416
## 114319:  1.093088e+00 0.2965358 1.876031 2.743454 5.0933280 8.206416
## 114320:  1.093088e+00 0.4713505 1.876031 2.743454 5.0933280 8.206416
## 114321:  5.483099e-08 2.3725858 2.224306 3.277048 3.4167206 8.067542
##              v32      v33      v34      v35      v36       v37 v38
##      1: 1.083033 1.010829 7.270147 8.375452 11.32659 0.4545457   0
##      2: 1.622151 2.161633 3.615077 8.122387 14.57948 0.7414708   0
##      3: 1.454082 1.734693 4.043864 7.959184 12.73052 0.2597401   0
##      4: 1.587644 1.666667 8.703550 8.898468 11.30280 0.4337346   0
##      5: 1.622151 2.161633 6.083151 8.122387 13.37560 0.7414708   0
##     ---
## 114317: 1.622151 2.161633 7.281489 8.122387 13.37560 0.7414708   0
## 114318: 1.622151 2.161633 6.721720 8.122387 13.37560 0.7414708   0
## 114319: 1.622151 2.161633 3.029508 8.122387 13.68885 0.7414708   0
## 114320: 1.622151 2.161633 5.089198 8.122387 12.11553 0.7414708   0
## 114321: 1.707317 2.526579 6.533439 7.179488 15.21767 0.6194700   0
##               v39       v40      v41      v42      v43       v44       v45
##      1: 4.01208777  7.711453 7.653429 12.70758 2.015505 10.498338 9.848672
##      2: 1.23718376 14.305766 7.182551 12.92497 2.216597 10.795169 9.142231
##      3: 7.37896421 13.077201 6.173469 12.34694 2.926830  8.897561 5.343819
##      4: 0.28732173 11.523045 7.931035 12.93582 1.470878 12.708574 9.670823
##      5: 1.23718376 10.138920 7.182551 12.92497 2.216597 10.795169 9.142231
##     ---
## 114317: 1.23718376  8.640475 7.182551 12.92497 2.216597 10.795169 9.142231
## 114318: 1.23718376 14.348455 7.182551 12.92497 2.216597 10.795169 9.142231
## 114319: 1.23718376 16.002461 7.182551 12.92497 2.216597 10.795169 9.142231
## 114320: 1.23718376 11.952825 7.182551 12.92497 2.216597 10.795169 9.142231
## 114321: 0.07718111  9.228710 6.979363 12.55785 2.600536 12.199975 6.733072
##               v46      v48      v49       v50      v51      v53        v54
##      1: 0.1135606 12.17173 8.086643 0.8994200 7.277792 16.74797 0.03709633
##      2: 2.4499589 12.53802 8.016547 1.3792101 7.198159 15.71130 1.12946855
##      3: 0.1260346 12.71133 6.836734 0.6045041 9.637627 15.10204 0.08557286
##      4: 0.1083869 12.19485 8.591954 3.3291765 4.780357 16.62169 0.13972117
##      5: 1.6305254 12.53802 8.016547 1.3645359 7.198159 15.71130 1.25385626
##     ---
## 114317: 1.6305254 12.53802 8.016547 2.0350670 7.198159 15.71130 1.25385626
## 114318: 1.6305254 12.53802 8.016547 3.2690201 7.198159 15.71130 1.25385626
## 114319: 0.2514433 12.53802 8.016547 2.4106815 7.198159 15.71130 0.06681885
## 114320: 0.4590391 12.53802 8.016547 0.8216566 7.198159 15.71130 0.37975392
## 114321: 2.7336173 11.10409 7.379612 1.0006610 6.315512 14.85271 2.07959069
##               v55      v57        v58       v59      v60      v61 v62
##      1: 1.2996383 3.971118  0.5298022 10.890984 1.588448 15.85815   1
```

```
##      2: 1.5595562 4.077828  7.7016531 10.587945 1.714294 14.58303   2
##      3: 0.7653052 4.030613  4.2774557  9.105481 2.151361 16.07560   1
##      4: 1.1781613 3.965517  1.7321022 11.777912 1.229246 15.92739   1
##      5: 1.5595562 4.077828  7.7016531 10.587945 1.714294 14.58303   1
##     ---
## 114317: 1.5595562 4.077828  7.7016531 10.587945 1.714294 14.58303   1
## 114318: 1.5595562 4.077828  7.7016531 10.587945 1.714294 14.58303   1
## 114319: 1.5595562 4.077828  7.7016531 10.587945 1.714294 14.58303   1
## 114320: 1.5595562 4.077828  7.7016531 10.587945 1.714294 14.58303   1
## 114321: 1.8386485 4.015009 16.4090371  9.197136 2.063790 15.90806   1
##               v63      v64      v65      v67      v68       v69       v70
##      1: 0.1534611 6.363189 18.30393 9.314079 15.23179 17.142857 11.784549
##      2: 2.5447365 6.343713 15.84756 9.287275 17.56412  9.449335 12.053353
##      3: 0.1236432 5.517949 16.37721 8.367347 11.04046  5.882353  8.460654
##      4: 0.1402597 6.292979 17.01165 9.703065 18.56813  9.425288 13.594728
##      5: 1.6873273 6.343713 15.84756 9.287275 17.56412  9.449335 12.269960
##     ---
## 114317: 1.6873273 6.343713 15.84756 9.287275 17.56412  9.449335 12.269960
## 114318: 1.6873273 6.343713 15.84756 9.287275 17.56412  9.449335 12.269960
## 114319: 0.2818862 6.343713 15.84756 9.287275 17.56412  9.449335  9.402630
## 114320: 0.4628394 6.343713 15.84756 9.287275 17.56412  9.449335 15.058248
## 114321: 2.1350054 7.924031 16.02907 9.330831 19.16766  9.702969 13.346576
##         v72      v73      v76      v77       v78      v80       v81
##      1:   1 1.614988 2.230940 7.292418  8.571429 3.000000  7.528326
##      2:   2 2.433303 2.405056 7.307366 13.334482 2.209700  7.277655
##      3:   3 2.413618 1.963971 5.918368 11.764705 3.333334 10.194433
##      4:   2 2.272541 2.188198 8.213602 13.448277 1.947261  4.797873
##      5:   1 2.433303 2.405056 7.307366 13.334482 2.209700  7.287174
##     ---
## 114317:   1 2.433303 2.405056 7.307366 13.334482 2.209700  7.287174
## 114318:   2 2.433303 2.405056 7.307366 13.334482 2.209700  7.287174
## 114319:   3 2.433303 2.405056 7.307366 13.334482 2.209700  6.070124
## 114320:   1 2.433303 2.405056 7.307366 13.334482 2.209700  4.947657
## 114321:   1 2.541331 3.103171 7.404627 15.643565 1.693121  8.249026
##              v82       v83      v84      v85       v86       v87      v88
##      1:  8.861647 0.6498199 1.299638 1.707317 0.8664262  9.551836 3.321300
##      2:  3.430691 2.1738077 1.607956 2.822253 1.2201841  9.848004 1.924184
##      3:  8.266200 1.5306113 1.530613 2.429906 1.0714292  8.447465 3.367346
##      4: 13.315819 1.6810343 1.379310 1.587045 1.2428166 10.747144 1.408046
##      5:  6.208356 2.1738077 1.607956 2.822253 1.2201841 10.180216 1.924184
##     ---
## 114317:  6.208356 2.1738077 1.607956 2.822253 1.2201841 10.180216 1.924184
## 114318:  6.208356 2.1738077 1.607956 2.822253 1.2201841 10.180216 1.924184
## 114319:  3.380753 2.1738077 1.607956 2.822253 1.2201841 11.603652 1.924184
## 114320:  2.522596 2.1738077 1.607956 2.822253 1.2201841 15.415573 1.924184
## 114321:  3.411930 2.9643522 1.200751 2.869380 1.4071285 10.326888 1.425892
##               v89       v90       v92      v93      v94       v95
##      1: 0.09567836 0.9053423 0.4422517 5.814018 3.517720 0.4620187
##      2: 2.67858429 0.9669126 0.5823668 5.475185 3.852883 0.6657576
##      3: 0.11138775 0.8114466 0.2714800 5.156559 4.214944 0.3096565
##      4: 0.03905132 1.0424254 0.7639246 5.498902 3.423944 0.8325182
##      5: 1.51842520 0.9669126 0.5823668 5.475185 3.852883 0.6657576
##     ---
## 114317: 1.51842520 0.9669126 0.5823668 5.475185 3.852883 0.6657576
```

```
## 114318: 1.51842520 0.9669126 0.5823668 5.475185 3.852883 0.6657576
## 114319: 0.25759035 0.9669126 0.5823668 5.475185 3.852883 0.6657576
## 114320: 0.42636864 0.9669126 0.5823668 5.475185 3.852883 0.6657576
## 114321: 2.41545188 0.7567042 0.4535285 6.735038 3.417862 0.5321365
##              v96      v97       v98      v99     v100     v101     v102
##      1: 7.436824 5.454545  8.877414 1.191337 19.47020 8.389237 2.757375
##      2: 6.457952 7.622554  8.303967 1.250721 12.09162 6.866414 2.890289
##      3: 5.663265 5.974026 11.588858 0.841837 15.49133 5.879353 3.292788
##      4: 7.375480 6.746988  6.942002 1.334611 18.25635 8.507281 2.503055
##      5: 6.457952 7.622554  7.667624 1.250721 12.09162 6.866414 2.890289
##     ---
## 114317: 6.457952 7.622554  7.667624 1.250721 12.09162 6.866414 2.890289
## 114318: 6.457952 7.622554  7.667624 1.250721 12.09162 6.866414 2.890289
## 114319: 6.457952 7.622554 11.939015 1.250721 12.09162 6.866414 2.890289
## 114320: 6.457952 7.622554  6.702722 1.250721 12.09162 6.866414 2.890289
## 114321: 5.991245 7.964602  5.930010 1.421721  3.51962 5.715783 4.115456
##             v103     v104        v105     v106     v108      v109
##      1: 4.374296 1.574039 0.007293816 12.57918 2.382692  3.930922
##      2: 5.296716 2.642828 1.505334661 11.79136 1.825361  4.247858
##      3: 5.924457 1.668401 0.008274619 11.67057 1.375753  1.184211
##      4: 4.872157 2.573664 0.113967370 12.55427 2.230754  1.990131
##      5: 5.296716 2.642828 1.081045222 11.79136 2.152620  4.181284
##     ---
## 114317: 5.296716 2.642828 1.081045222 11.79136 2.152620  4.181284
## 114318: 5.296716 2.642828 1.081045222 11.79136 2.152620  4.181284
## 114319: 5.296716 2.642828 0.251911633 11.79136 2.458328  5.834772
## 114320: 5.296716 2.642828 0.337262290 11.79136 3.311224 11.175868
## 114321: 3.630999 2.576789 2.103316936 13.91935 1.397909  3.051251
##              v111      v114      v115     v116      v117     v118
##      1: 0.4332129 15.634907  2.857144 1.951220  6.592012 5.909091
##      2: 3.3653137 10.308044 10.548051 2.291218 10.595357 8.364651
##      3: 3.3673476 11.205561 12.941177 3.129253  3.478911 6.233767
##      4: 2.6436782 13.777666 10.574713 1.511063  4.949609 7.180722
##      5: 3.3653137 14.097099 10.548051 2.291218  8.303857 8.364651
##     ---
## 114317: 3.3653137 15.392866 10.548051 2.291218  8.303857 8.364651
## 114318: 3.3653137 11.248736 10.548051 2.291218  8.303857 8.364651
## 114319: 3.3653137  8.893134 10.548051 2.291218 12.743060 8.364651
## 114320: 3.3653137 12.381113 10.548051 2.291218 15.357450 8.364651
## 114321: 3.9024390 14.635298 10.297030 2.790698  8.969215 8.584070
##                  v119      v120      v121     v122     v123        v124
##      1: -6.297423e-07 1.0596026 0.8035719 8.000000 1.989780 0.035753685
##      2:  3.168970e+00 1.2912179 2.7375960 6.822439 3.549938 0.598895646
##      3: -2.792745e-07 2.1387283 2.2388065 9.333333 2.477596 0.013451914
##      4:  5.655086e-01 1.1662808 1.9565207 7.018256 1.812795 0.002267384
##      5:  3.168970e+00 1.2912179 2.7375960 6.822439 3.549938 0.919811985
##     ---
## 114317:  3.168970e+00 1.2912179 2.7375960 6.822439 3.549938 0.919811985
## 114318:  3.168970e+00 1.2912179 2.7375960 6.822439 3.549938 0.919811985
## 114319:  3.168970e+00 1.2912179 2.7375960 6.822439 3.549938 0.156764229
## 114320:  3.168970e+00 1.2912179 2.7375960 6.822439 3.549938 0.490657553
## 114321:  7.093470e-01 0.7372185 4.0169483 7.936508 2.944285 3.135204728
##             v126     v127     v128 v129      v130     v131
##      1: 1.804126 3.113719 2.024285    0 0.6363645 2.857144
```

```
##      2: 1.672658 3.239542 1.957825    0 1.9257635 1.739389
##      3: 1.773709 3.922193 1.120468    2 0.8831175 1.176472
##      4: 1.415230 2.954381 1.990847    1 1.6771076 1.034483
##      5: 1.672658 3.239542 2.030373    0 1.9257635 1.739389
##     ---
## 114317: 1.672658 3.239542 2.030373    0 1.9257635 1.739389
## 114318: 1.672658 3.239542 2.030373    1 1.9257635 1.739389
## 114319: 1.672658 3.239542 2.417606    2 1.9257635 1.739389
## 114320: 1.672658 3.239542 3.526650    0 1.9257635 1.739389
## 114321: 1.943149 4.385553 1.604493    0 1.7876103 1.386138
```

```
findLinearCombos(imputed.train.num)$linearCombos # Remove 110
```

```
## [[1]]
## [1] 110  33  54  62
```

```
findLinearCombos(imputed.train.num[,c(-110),])
```

```
## $linearCombos
## list()
##
## $remove
## NULL
```

References:

1. http://www.sthda.com/english/wiki/correspondence-analysis-in-r-the-ultimate-guide-for-the-analysis-the-visualization-a correspondence-analysis-ca