Internal_lab_Assisgment_Clean

```
library(ggplot2)
library(dplyr)
```

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

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

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

```
library(mice)
```

```
##
## Attaching package: 'mice'
```

```
## The following object is masked from 'package:stats':
##
##     filter
```

```
## The following objects are masked from 'package:base':
##
##     cbind, rbind
```

```
library(tidyr)
library(Rtsne)
```

```
## Warning: package 'Rtsne' was built under R version 4.4.3
```

```
library(ggplot2)
library(factoextra)
```

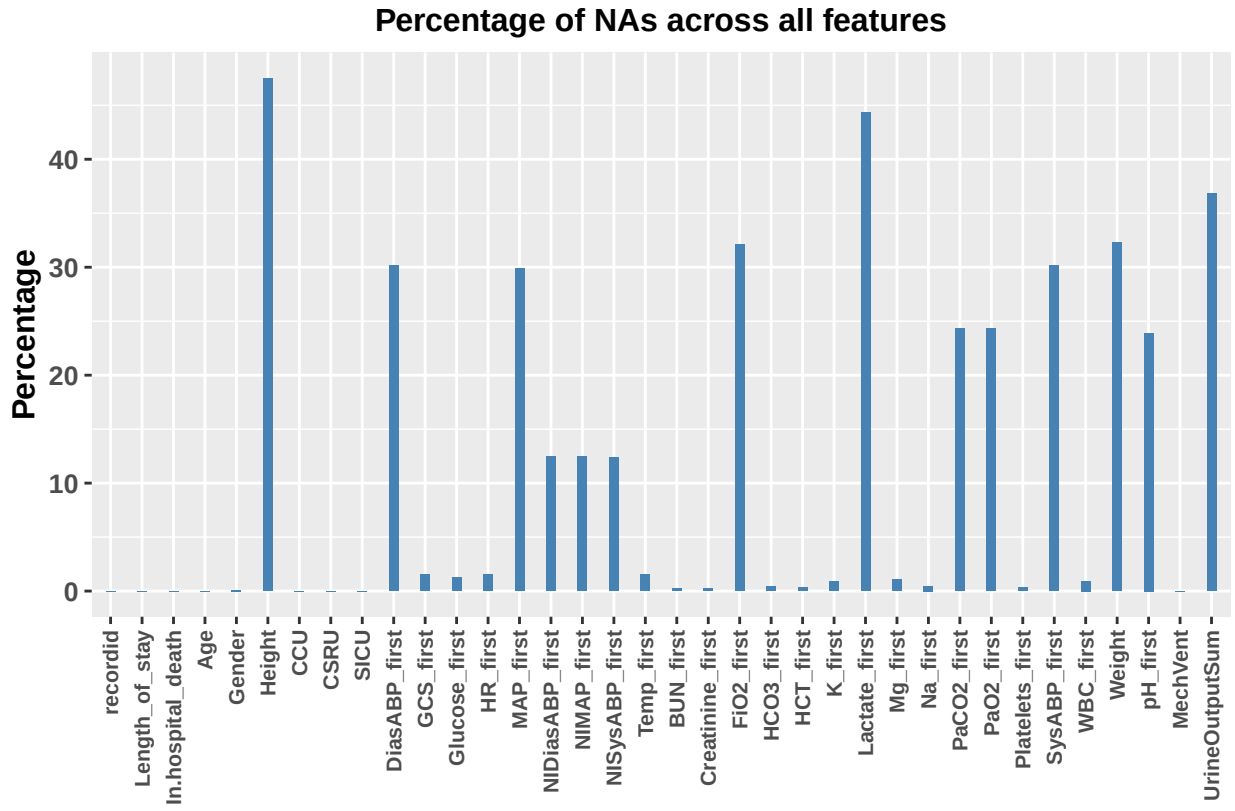
```
## Warning: package 'factoextra' was built under R version 4.4.3
```

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

```
library(cluster)
```

```
file = 'Data/ICU_filtered.csv'
dfICU = read.csv(file, header = TRUE, stringsAsFactors = TRUE)
```

```
pMissDF = setNames(stack(sapply(dfICU, function(x){(sum(is.na(x))/length(x))*100}))[2:1], c('Feature','
p = ggplot(data = pMissDF, aes(x = Feature, y = Value)) +
  geom_bar(stat = 'identity', fill = 'steelblue', width = 0.3) +
  theme(text = element_text(size = 14, face = 'bold'),
  axis.text.x = element_text(angle = 90, hjust = 1, vjust = 0.5)) +
  xlab('') + ylab('Percentage') +
  ggtitle('Percentage of NAs across all features')  +
  theme(plot.title = element_text(size = 12, hjust = 0.5),
        axis.text = element_text(size = 8),
        axis.text.x = element_text(size = 8),
        axis.text.y = element_text(size = 10),
        axis.title = element_text(size = 12, face = "bold"))
p
```




```
dfICU = dfICU %>% select(-c(pMissDF[pMissDF['Value'] > 20, 'Feature']))
```

```
dfICU[dfICU['CCU'] == 1, 'ICU'] = 1
dfICU[dfICU['CSRU'] == 1, 'ICU'] = 2
dfICU[dfICU['SICU'] == 1, 'ICU'] = 3
```

```r
dfICU[(dfICU['CCU'] == 0 ) & (dfICU['CSRU'] == 0) & (dfICU['SICU'] == 0), 'ICU'] = 4
dfICU = dfICU %>% select(-c(CCU, CSRU, SICU, recordid, In.hospital_death, Length_of_stay))
```

```r
str(dfICU)
```

```
## 'data.frame':    7886 obs. of  20 variables:
##  $ Age             : int  54 76 44 68 88 64 68 78 64 74 ...
##  $ Gender          : int  0 1 0 1 0 1 0 0 0 1 ...
##  $ GCS_first       : int  15 3 7 15 15 7 15 15 15 10 ...
##  $ Glucose_first   : int  205 105 141 129 113 264 94 132 113 106 ...
##  $ HR_first        : num  73 88 100 79 93 78 73 111 127 67 ...
##  $ NIDiasABP_first : int  65 38 84 63 41 89 88 51 71 57 ...
##  $ NIMAP_first     : num  92.3 49.3 100.3 86.7 75.3 ...
##  $ NISysABP_first  : int  147 72 133 134 144 129 187 100 138 134 ...
##  $ Temp_first      : num  35.1 35.2 37.8 36.3 37.8 35.8 36.3 38 37.3 34.8 ...
##  $ BUN_first       : int  13 16 8 23 45 15 32 81 21 19 ...
##  $ Creatinine_first: num  0.8 0.8 0.4 0.9 1 1.4 3.4 0.9 0.7 1.1 ...
##  $ HCO3_first      : int  26 21 24 28 18 19 25 18 21 23 ...
##  $ HCT_first       : num  33.7 24.7 28.5 41.3 22.6 41.6 31.9 32.6 28.3 31.5 ...
##  $ K_first         : num  4.4 4.3 3.3 4 6 5.1 3.7 4.2 3.9 4.6 ...
##  $ Mg_first        : num  1.5 3.1 1.9 2.1 1.5 1.7 1.9 2.2 1.6 1.8 ...
##  $ Na_first        : int  137 139 137 140 140 141 140 141 139 141 ...
##  $ Platelets_first : int  221 164 72 391 109 276 325 91 696 141 ...
##  $ WBC_first       : num  11.2 7.4 4.2 11.5 3.8 24 6.2 16.1 15.2 9 ...
##  $ MechVent        : int  0 1 1 0 0 1 0 1 0 1 ...
##  $ ICU             : num  3 2 4 4 4 1 4 4 4 2 ...
```

```r
dfICU = complete(mice(dfICU, m = 40, maxit = 10, pred = quickpred(dfICU, minpuc = 0.25), seed = 500))
```

```
##
##  iter imp variable
##    1   1  Gender  GCS_first  Glucose_first  HR_first  NIDiasABP_first  NIMAP_first  NISysABP_first  Te
##    1   2  Gender  GCS_first  Glucose_first  HR_first  NIDiasABP_first  NIMAP_first  NISysABP_first  Te
##    1   3  Gender  GCS_first  Glucose_first  HR_first  NIDiasABP_first  NIMAP_first  NISysABP_first  Te
##    1   4  Gender  GCS_first  Glucose_first  HR_first  NIDiasABP_first  NIMAP_first  NISysABP_first  Te
##    1   5  Gender  GCS_first  Glucose_first  HR_first  NIDiasABP_first  NIMAP_first  NISysABP_first  Te
##    1   6  Gender  GCS_first  Glucose_first  HR_first  NIDiasABP_first  NIMAP_first  NISysABP_first  Te
##    1   7  Gender  GCS_first  Glucose_first  HR_first  NIDiasABP_first  NIMAP_first  NISysABP_first  Te
##    1   8  Gender  GCS_first  Glucose_first  HR_first  NIDiasABP_first  NIMAP_first  NISysABP_first  Te
##    1   9  Gender  GCS_first  Glucose_first  HR_first  NIDiasABP_first  NIMAP_first  NISysABP_first  Te
##    1   10  Gender  GCS_first  Glucose_first  HR_first  NIDiasABP_first  NIMAP_first  NISysABP_first  T
##    1   11  Gender  GCS_first  Glucose_first  HR_first  NIDiasABP_first  NIMAP_first  NISysABP_first  T
##    1   12  Gender  GCS_first  Glucose_first  HR_first  NIDiasABP_first  NIMAP_first  NISysABP_first  T
##    1   13  Gender  GCS_first  Glucose_first  HR_first  NIDiasABP_first  NIMAP_first  NISysABP_first  T
##    1   14  Gender  GCS_first  Glucose_first  HR_first  NIDiasABP_first  NIMAP_first  NISysABP_first  T
##    1   15  Gender  GCS_first  Glucose_first  HR_first  NIDiasABP_first  NIMAP_first  NISysABP_first  T
##    1   16  Gender  GCS_first  Glucose_first  HR_first  NIDiasABP_first  NIMAP_first  NISysABP_first  T
##    1   17  Gender  GCS_first  Glucose_first  HR_first  NIDiasABP_first  NIMAP_first  NISysABP_first  T
##    1   18  Gender  GCS_first  Glucose_first  HR_first  NIDiasABP_first  NIMAP_first  NISysABP_first  T
##    1   19  Gender  GCS_first  Glucose_first  HR_first  NIDiasABP_first  NIMAP_first  NISysABP_first  T
##    1   20  Gender  GCS_first  Glucose_first  HR_first  NIDiasABP_first  NIMAP_first  NISysABP_first  T
##    1   21  Gender  GCS_first  Glucose_first  HR_first  NIDiasABP_first  NIMAP_first  NISysABP_first  T
```

```
##   1   22  Gender  GCS_first  Glucose_first  HR_first  NIDiasABP_first  NIMAP_first  NISysABP_first  T
##   1   23  Gender  GCS_first  Glucose_first  HR_first  NIDiasABP_first  NIMAP_first  NISysABP_first  T
##   1   24  Gender  GCS_first  Glucose_first  HR_first  NIDiasABP_first  NIMAP_first  NISysABP_first  T
##   1   25  Gender  GCS_first  Glucose_first  HR_first  NIDiasABP_first  NIMAP_first  NISysABP_first  T
##   1   26  Gender  GCS_first  Glucose_first  HR_first  NIDiasABP_first  NIMAP_first  NISysABP_first  T
##   1   27  Gender  GCS_first  Glucose_first  HR_first  NIDiasABP_first  NIMAP_first  NISysABP_first  T
##   1   28  Gender  GCS_first  Glucose_first  HR_first  NIDiasABP_first  NIMAP_first  NISysABP_first  T
##   1   29  Gender  GCS_first  Glucose_first  HR_first  NIDiasABP_first  NIMAP_first  NISysABP_first  T
##   1   30  Gender  GCS_first  Glucose_first  HR_first  NIDiasABP_first  NIMAP_first  NISysABP_first  T
##   1   31  Gender  GCS_first  Glucose_first  HR_first  NIDiasABP_first  NIMAP_first  NISysABP_first  T
##   1   32  Gender  GCS_first  Glucose_first  HR_first  NIDiasABP_first  NIMAP_first  NISysABP_first  T
##   1   33  Gender  GCS_first  Glucose_first  HR_first  NIDiasABP_first  NIMAP_first  NISysABP_first  T
##   1   34  Gender  GCS_first  Glucose_first  HR_first  NIDiasABP_first  NIMAP_first  NISysABP_first  T
##   1   35  Gender  GCS_first  Glucose_first  HR_first  NIDiasABP_first  NIMAP_first  NISysABP_first  T
##   1   36  Gender  GCS_first  Glucose_first  HR_first  NIDiasABP_first  NIMAP_first  NISysABP_first  T
##   1   37  Gender  GCS_first  Glucose_first  HR_first  NIDiasABP_first  NIMAP_first  NISysABP_first  T
##   1   38  Gender  GCS_first  Glucose_first  HR_first  NIDiasABP_first  NIMAP_first  NISysABP_first  T
##   1   39  Gender  GCS_first  Glucose_first  HR_first  NIDiasABP_first  NIMAP_first  NISysABP_first  T
##   1   40  Gender  GCS_first  Glucose_first  HR_first  NIDiasABP_first  NIMAP_first  NISysABP_first  T
##   2   1  Gender  GCS_first  Glucose_first  HR_first  NIDiasABP_first  NIMAP_first  NISysABP_first  To
##   2   2  Gender  GCS_first  Glucose_first  HR_first  NIDiasABP_first  NIMAP_first  NISysABP_first  To
##   2   3  Gender  GCS_first  Glucose_first  HR_first  NIDiasABP_first  NIMAP_first  NISysABP_first  To
##   2   4  Gender  GCS_first  Glucose_first  HR_first  NIDiasABP_first  NIMAP_first  NISysABP_first  To
##   2   5  Gender  GCS_first  Glucose_first  HR_first  NIDiasABP_first  NIMAP_first  NISysABP_first  To
##   2   6  Gender  GCS_first  Glucose_first  HR_first  NIDiasABP_first  NIMAP_first  NISysABP_first  To
##   2   7  Gender  GCS_first  Glucose_first  HR_first  NIDiasABP_first  NIMAP_first  NISysABP_first  To
##   2   8  Gender  GCS_first  Glucose_first  HR_first  NIDiasABP_first  NIMAP_first  NISysABP_first  To
##   2   9  Gender  GCS_first  Glucose_first  HR_first  NIDiasABP_first  NIMAP_first  NISysABP_first  To
##   2   10  Gender  GCS_first  Glucose_first  HR_first  NIDiasABP_first  NIMAP_first  NISysABP_first  T
##   2   11  Gender  GCS_first  Glucose_first  HR_first  NIDiasABP_first  NIMAP_first  NISysABP_first  T
##   2   12  Gender  GCS_first  Glucose_first  HR_first  NIDiasABP_first  NIMAP_first  NISysABP_first  T
##   2   13  Gender  GCS_first  Glucose_first  HR_first  NIDiasABP_first  NIMAP_first  NISysABP_first  T
##   2   14  Gender  GCS_first  Glucose_first  HR_first  NIDiasABP_first  NIMAP_first  NISysABP_first  T
##   2   15  Gender  GCS_first  Glucose_first  HR_first  NIDiasABP_first  NIMAP_first  NISysABP_first  T
##   2   16  Gender  GCS_first  Glucose_first  HR_first  NIDiasABP_first  NIMAP_first  NISysABP_first  T
##   2   17  Gender  GCS_first  Glucose_first  HR_first  NIDiasABP_first  NIMAP_first  NISysABP_first  T
##   2   18  Gender  GCS_first  Glucose_first  HR_first  NIDiasABP_first  NIMAP_first  NISysABP_first  T
##   2   19  Gender  GCS_first  Glucose_first  HR_first  NIDiasABP_first  NIMAP_first  NISysABP_first  T
##   2   20  Gender  GCS_first  Glucose_first  HR_first  NIDiasABP_first  NIMAP_first  NISysABP_first  T
##   2   21  Gender  GCS_first  Glucose_first  HR_first  NIDiasABP_first  NIMAP_first  NISysABP_first  T
##   2   22  Gender  GCS_first  Glucose_first  HR_first  NIDiasABP_first  NIMAP_first  NISysABP_first  T
##   2   23  Gender  GCS_first  Glucose_first  HR_first  NIDiasABP_first  NIMAP_first  NISysABP_first  T
##   2   24  Gender  GCS_first  Glucose_first  HR_first  NIDiasABP_first  NIMAP_first  NISysABP_first  T
##   2   25  Gender  GCS_first  Glucose_first  HR_first  NIDiasABP_first  NIMAP_first  NISysABP_first  T
##   2   26  Gender  GCS_first  Glucose_first  HR_first  NIDiasABP_first  NIMAP_first  NISysABP_first  T
##   2   27  Gender  GCS_first  Glucose_first  HR_first  NIDiasABP_first  NIMAP_first  NISysABP_first  T
##   2   28  Gender  GCS_first  Glucose_first  HR_first  NIDiasABP_first  NIMAP_first  NISysABP_first  T
##   2   29  Gender  GCS_first  Glucose_first  HR_first  NIDiasABP_first  NIMAP_first  NISysABP_first  T
##   2   30  Gender  GCS_first  Glucose_first  HR_first  NIDiasABP_first  NIMAP_first  NISysABP_first  T
##   2   31  Gender  GCS_first  Glucose_first  HR_first  NIDiasABP_first  NIMAP_first  NISysABP_first  T
##   2   32  Gender  GCS_first  Glucose_first  HR_first  NIDiasABP_first  NIMAP_first  NISysABP_first  T
##   2   33  Gender  GCS_first  Glucose_first  HR_first  NIDiasABP_first  NIMAP_first  NISysABP_first  T
##   2   34  Gender  GCS_first  Glucose_first  HR_first  NIDiasABP_first  NIMAP_first  NISysABP_first  T
##   2   35  Gender  GCS_first  Glucose_first  HR_first  NIDiasABP_first  NIMAP_first  NISysABP_first  T
```

```
##   2   36  Gender  GCS_first  Glucose_first  HR_first  NIDiasABP_first  NIMAP_first  NISysABP_first  T
##   2   37  Gender  GCS_first  Glucose_first  HR_first  NIDiasABP_first  NIMAP_first  NISysABP_first  T
##   2   38  Gender  GCS_first  Glucose_first  HR_first  NIDiasABP_first  NIMAP_first  NISysABP_first  T
##   2   39  Gender  GCS_first  Glucose_first  HR_first  NIDiasABP_first  NIMAP_first  NISysABP_first  T
##   2   40  Gender  GCS_first  Glucose_first  HR_first  NIDiasABP_first  NIMAP_first  NISysABP_first  T
##   3   1  Gender  GCS_first  Glucose_first  HR_first  NIDiasABP_first  NIMAP_first  NISysABP_first  Te
##   3   2  Gender  GCS_first  Glucose_first  HR_first  NIDiasABP_first  NIMAP_first  NISysABP_first  Te
##   3   3  Gender  GCS_first  Glucose_first  HR_first  NIDiasABP_first  NIMAP_first  NISysABP_first  Te
##   3   4  Gender  GCS_first  Glucose_first  HR_first  NIDiasABP_first  NIMAP_first  NISysABP_first  Te
##   3   5  Gender  GCS_first  Glucose_first  HR_first  NIDiasABP_first  NIMAP_first  NISysABP_first  Te
##   3   6  Gender  GCS_first  Glucose_first  HR_first  NIDiasABP_first  NIMAP_first  NISysABP_first  Te
##   3   7  Gender  GCS_first  Glucose_first  HR_first  NIDiasABP_first  NIMAP_first  NISysABP_first  Te
##   3   8  Gender  GCS_first  Glucose_first  HR_first  NIDiasABP_first  NIMAP_first  NISysABP_first  Te
##   3   9  Gender  GCS_first  Glucose_first  HR_first  NIDiasABP_first  NIMAP_first  NISysABP_first  Te
##   3   10  Gender  GCS_first  Glucose_first  HR_first  NIDiasABP_first  NIMAP_first  NISysABP_first  T
##   3   11  Gender  GCS_first  Glucose_first  HR_first  NIDiasABP_first  NIMAP_first  NISysABP_first  T
##   3   12  Gender  GCS_first  Glucose_first  HR_first  NIDiasABP_first  NIMAP_first  NISysABP_first  T
##   3   13  Gender  GCS_first  Glucose_first  HR_first  NIDiasABP_first  NIMAP_first  NISysABP_first  T
##   3   14  Gender  GCS_first  Glucose_first  HR_first  NIDiasABP_first  NIMAP_first  NISysABP_first  T
##   3   15  Gender  GCS_first  Glucose_first  HR_first  NIDiasABP_first  NIMAP_first  NISysABP_first  T
##   3   16  Gender  GCS_first  Glucose_first  HR_first  NIDiasABP_first  NIMAP_first  NISysABP_first  T
##   3   17  Gender  GCS_first  Glucose_first  HR_first  NIDiasABP_first  NIMAP_first  NISysABP_first  T
##   3   18  Gender  GCS_first  Glucose_first  HR_first  NIDiasABP_first  NIMAP_first  NISysABP_first  T
##   3   19  Gender  GCS_first  Glucose_first  HR_first  NIDiasABP_first  NIMAP_first  NISysABP_first  T
##   3   20  Gender  GCS_first  Glucose_first  HR_first  NIDiasABP_first  NIMAP_first  NISysABP_first  T
##   3   21  Gender  GCS_first  Glucose_first  HR_first  NIDiasABP_first  NIMAP_first  NISysABP_first  T
##   3   22  Gender  GCS_first  Glucose_first  HR_first  NIDiasABP_first  NIMAP_first  NISysABP_first  T
##   3   23  Gender  GCS_first  Glucose_first  HR_first  NIDiasABP_first  NIMAP_first  NISysABP_first  T
##   3   24  Gender  GCS_first  Glucose_first  HR_first  NIDiasABP_first  NIMAP_first  NISysABP_first  T
##   3   25  Gender  GCS_first  Glucose_first  HR_first  NIDiasABP_first  NIMAP_first  NISysABP_first  T
##   3   26  Gender  GCS_first  Glucose_first  HR_first  NIDiasABP_first  NIMAP_first  NISysABP_first  T
##   3   27  Gender  GCS_first  Glucose_first  HR_first  NIDiasABP_first  NIMAP_first  NISysABP_first  T
##   3   28  Gender  GCS_first  Glucose_first  HR_first  NIDiasABP_first  NIMAP_first  NISysABP_first  T
##   3   29  Gender  GCS_first  Glucose_first  HR_first  NIDiasABP_first  NIMAP_first  NISysABP_first  T
##   3   30  Gender  GCS_first  Glucose_first  HR_first  NIDiasABP_first  NIMAP_first  NISysABP_first  T
##   3   31  Gender  GCS_first  Glucose_first  HR_first  NIDiasABP_first  NIMAP_first  NISysABP_first  T
##   3   32  Gender  GCS_first  Glucose_first  HR_first  NIDiasABP_first  NIMAP_first  NISysABP_first  T
##   3   33  Gender  GCS_first  Glucose_first  HR_first  NIDiasABP_first  NIMAP_first  NISysABP_first  T
##   3   34  Gender  GCS_first  Glucose_first  HR_first  NIDiasABP_first  NIMAP_first  NISysABP_first  T
##   3   35  Gender  GCS_first  Glucose_first  HR_first  NIDiasABP_first  NIMAP_first  NISysABP_first  T
##   3   36  Gender  GCS_first  Glucose_first  HR_first  NIDiasABP_first  NIMAP_first  NISysABP_first  T
##   3   37  Gender  GCS_first  Glucose_first  HR_first  NIDiasABP_first  NIMAP_first  NISysABP_first  T
##   3   38  Gender  GCS_first  Glucose_first  HR_first  NIDiasABP_first  NIMAP_first  NISysABP_first  T
##   3   39  Gender  GCS_first  Glucose_first  HR_first  NIDiasABP_first  NIMAP_first  NISysABP_first  T
##   3   40  Gender  GCS_first  Glucose_first  HR_first  NIDiasABP_first  NIMAP_first  NISysABP_first  T
##   4   1  Gender  GCS_first  Glucose_first  HR_first  NIDiasABP_first  NIMAP_first  NISysABP_first  Te
##   4   2  Gender  GCS_first  Glucose_first  HR_first  NIDiasABP_first  NIMAP_first  NISysABP_first  Te
##   4   3  Gender  GCS_first  Glucose_first  HR_first  NIDiasABP_first  NIMAP_first  NISysABP_first  Te
##   4   4  Gender  GCS_first  Glucose_first  HR_first  NIDiasABP_first  NIMAP_first  NISysABP_first  Te
##   4   5  Gender  GCS_first  Glucose_first  HR_first  NIDiasABP_first  NIMAP_first  NISysABP_first  Te
##   4   6  Gender  GCS_first  Glucose_first  HR_first  NIDiasABP_first  NIMAP_first  NISysABP_first  Te
##   4   7  Gender  GCS_first  Glucose_first  HR_first  NIDiasABP_first  NIMAP_first  NISysABP_first  Te
##   4   8  Gender  GCS_first  Glucose_first  HR_first  NIDiasABP_first  NIMAP_first  NISysABP_first  Te
##   4   9  Gender  GCS_first  Glucose_first  HR_first  NIDiasABP_first  NIMAP_first  NISysABP_first  Te
```

```
##   4   10  Gender  GCS_first  Glucose_first  HR_first  NIDiasABP_first  NIMAP_first  NISysABP_first  T
##   4   11  Gender  GCS_first  Glucose_first  HR_first  NIDiasABP_first  NIMAP_first  NISysABP_first  T
##   4   12  Gender  GCS_first  Glucose_first  HR_first  NIDiasABP_first  NIMAP_first  NISysABP_first  T
##   4   13  Gender  GCS_first  Glucose_first  HR_first  NIDiasABP_first  NIMAP_first  NISysABP_first  T
##   4   14  Gender  GCS_first  Glucose_first  HR_first  NIDiasABP_first  NIMAP_first  NISysABP_first  T
##   4   15  Gender  GCS_first  Glucose_first  HR_first  NIDiasABP_first  NIMAP_first  NISysABP_first  T
##   4   16  Gender  GCS_first  Glucose_first  HR_first  NIDiasABP_first  NIMAP_first  NISysABP_first  T
##   4   17  Gender  GCS_first  Glucose_first  HR_first  NIDiasABP_first  NIMAP_first  NISysABP_first  T
##   4   18  Gender  GCS_first  Glucose_first  HR_first  NIDiasABP_first  NIMAP_first  NISysABP_first  T
##   4   19  Gender  GCS_first  Glucose_first  HR_first  NIDiasABP_first  NIMAP_first  NISysABP_first  T
##   4   20  Gender  GCS_first  Glucose_first  HR_first  NIDiasABP_first  NIMAP_first  NISysABP_first  T
##   4   21  Gender  GCS_first  Glucose_first  HR_first  NIDiasABP_first  NIMAP_first  NISysABP_first  T
##   4   22  Gender  GCS_first  Glucose_first  HR_first  NIDiasABP_first  NIMAP_first  NISysABP_first  T
##   4   23  Gender  GCS_first  Glucose_first  HR_first  NIDiasABP_first  NIMAP_first  NISysABP_first  T
##   4   24  Gender  GCS_first  Glucose_first  HR_first  NIDiasABP_first  NIMAP_first  NISysABP_first  T
##   4   25  Gender  GCS_first  Glucose_first  HR_first  NIDiasABP_first  NIMAP_first  NISysABP_first  T
##   4   26  Gender  GCS_first  Glucose_first  HR_first  NIDiasABP_first  NIMAP_first  NISysABP_first  T
##   4   27  Gender  GCS_first  Glucose_first  HR_first  NIDiasABP_first  NIMAP_first  NISysABP_first  T
##   4   28  Gender  GCS_first  Glucose_first  HR_first  NIDiasABP_first  NIMAP_first  NISysABP_first  T
##   4   29  Gender  GCS_first  Glucose_first  HR_first  NIDiasABP_first  NIMAP_first  NISysABP_first  T
##   4   30  Gender  GCS_first  Glucose_first  HR_first  NIDiasABP_first  NIMAP_first  NISysABP_first  T
##   4   31  Gender  GCS_first  Glucose_first  HR_first  NIDiasABP_first  NIMAP_first  NISysABP_first  T
##   4   32  Gender  GCS_first  Glucose_first  HR_first  NIDiasABP_first  NIMAP_first  NISysABP_first  T
##   4   33  Gender  GCS_first  Glucose_first  HR_first  NIDiasABP_first  NIMAP_first  NISysABP_first  T
##   4   34  Gender  GCS_first  Glucose_first  HR_first  NIDiasABP_first  NIMAP_first  NISysABP_first  T
##   4   35  Gender  GCS_first  Glucose_first  HR_first  NIDiasABP_first  NIMAP_first  NISysABP_first  T
##   4   36  Gender  GCS_first  Glucose_first  HR_first  NIDiasABP_first  NIMAP_first  NISysABP_first  T
##   4   37  Gender  GCS_first  Glucose_first  HR_first  NIDiasABP_first  NIMAP_first  NISysABP_first  T
##   4   38  Gender  GCS_first  Glucose_first  HR_first  NIDiasABP_first  NIMAP_first  NISysABP_first  T
##   4   39  Gender  GCS_first  Glucose_first  HR_first  NIDiasABP_first  NIMAP_first  NISysABP_first  T
##   4   40  Gender  GCS_first  Glucose_first  HR_first  NIDiasABP_first  NIMAP_first  NISysABP_first  T
##   5   1  Gender  GCS_first  Glucose_first  HR_first  NIDiasABP_first  NIMAP_first  NISysABP_first  To
##   5   2  Gender  GCS_first  Glucose_first  HR_first  NIDiasABP_first  NIMAP_first  NISysABP_first  To
##   5   3  Gender  GCS_first  Glucose_first  HR_first  NIDiasABP_first  NIMAP_first  NISysABP_first  To
##   5   4  Gender  GCS_first  Glucose_first  HR_first  NIDiasABP_first  NIMAP_first  NISysABP_first  To
##   5   5  Gender  GCS_first  Glucose_first  HR_first  NIDiasABP_first  NIMAP_first  NISysABP_first  To
##   5   6  Gender  GCS_first  Glucose_first  HR_first  NIDiasABP_first  NIMAP_first  NISysABP_first  To
##   5   7  Gender  GCS_first  Glucose_first  HR_first  NIDiasABP_first  NIMAP_first  NISysABP_first  To
##   5   8  Gender  GCS_first  Glucose_first  HR_first  NIDiasABP_first  NIMAP_first  NISysABP_first  To
##   5   9  Gender  GCS_first  Glucose_first  HR_first  NIDiasABP_first  NIMAP_first  NISysABP_first  To
##   5   10  Gender  GCS_first  Glucose_first  HR_first  NIDiasABP_first  NIMAP_first  NISysABP_first  T
##   5   11  Gender  GCS_first  Glucose_first  HR_first  NIDiasABP_first  NIMAP_first  NISysABP_first  T
##   5   12  Gender  GCS_first  Glucose_first  HR_first  NIDiasABP_first  NIMAP_first  NISysABP_first  T
##   5   13  Gender  GCS_first  Glucose_first  HR_first  NIDiasABP_first  NIMAP_first  NISysABP_first  T
##   5   14  Gender  GCS_first  Glucose_first  HR_first  NIDiasABP_first  NIMAP_first  NISysABP_first  T
##   5   15  Gender  GCS_first  Glucose_first  HR_first  NIDiasABP_first  NIMAP_first  NISysABP_first  T
##   5   16  Gender  GCS_first  Glucose_first  HR_first  NIDiasABP_first  NIMAP_first  NISysABP_first  T
##   5   17  Gender  GCS_first  Glucose_first  HR_first  NIDiasABP_first  NIMAP_first  NISysABP_first  T
##   5   18  Gender  GCS_first  Glucose_first  HR_first  NIDiasABP_first  NIMAP_first  NISysABP_first  T
##   5   19  Gender  GCS_first  Glucose_first  HR_first  NIDiasABP_first  NIMAP_first  NISysABP_first  T
##   5   20  Gender  GCS_first  Glucose_first  HR_first  NIDiasABP_first  NIMAP_first  NISysABP_first  T
##   5   21  Gender  GCS_first  Glucose_first  HR_first  NIDiasABP_first  NIMAP_first  NISysABP_first  T
##   5   22  Gender  GCS_first  Glucose_first  HR_first  NIDiasABP_first  NIMAP_first  NISysABP_first  T
##   5   23  Gender  GCS_first  Glucose_first  HR_first  NIDiasABP_first  NIMAP_first  NISysABP_first  T
```

```
##   5   24  Gender  GCS_first  Glucose_first  HR_first  NIDiasABP_first  NIMAP_first  NISysABP_first  T
##   5   25  Gender  GCS_first  Glucose_first  HR_first  NIDiasABP_first  NIMAP_first  NISysABP_first  T
##   5   26  Gender  GCS_first  Glucose_first  HR_first  NIDiasABP_first  NIMAP_first  NISysABP_first  T
##   5   27  Gender  GCS_first  Glucose_first  HR_first  NIDiasABP_first  NIMAP_first  NISysABP_first  T
##   5   28  Gender  GCS_first  Glucose_first  HR_first  NIDiasABP_first  NIMAP_first  NISysABP_first  T
##   5   29  Gender  GCS_first  Glucose_first  HR_first  NIDiasABP_first  NIMAP_first  NISysABP_first  T
##   5   30  Gender  GCS_first  Glucose_first  HR_first  NIDiasABP_first  NIMAP_first  NISysABP_first  T
##   5   31  Gender  GCS_first  Glucose_first  HR_first  NIDiasABP_first  NIMAP_first  NISysABP_first  T
##   5   32  Gender  GCS_first  Glucose_first  HR_first  NIDiasABP_first  NIMAP_first  NISysABP_first  T
##   5   33  Gender  GCS_first  Glucose_first  HR_first  NIDiasABP_first  NIMAP_first  NISysABP_first  T
##   5   34  Gender  GCS_first  Glucose_first  HR_first  NIDiasABP_first  NIMAP_first  NISysABP_first  T
##   5   35  Gender  GCS_first  Glucose_first  HR_first  NIDiasABP_first  NIMAP_first  NISysABP_first  T
##   5   36  Gender  GCS_first  Glucose_first  HR_first  NIDiasABP_first  NIMAP_first  NISysABP_first  T
##   5   37  Gender  GCS_first  Glucose_first  HR_first  NIDiasABP_first  NIMAP_first  NISysABP_first  T
##   5   38  Gender  GCS_first  Glucose_first  HR_first  NIDiasABP_first  NIMAP_first  NISysABP_first  T
##   5   39  Gender  GCS_first  Glucose_first  HR_first  NIDiasABP_first  NIMAP_first  NISysABP_first  T
##   5   40  Gender  GCS_first  Glucose_first  HR_first  NIDiasABP_first  NIMAP_first  NISysABP_first  T
##   6   1  Gender  GCS_first  Glucose_first  HR_first  NIDiasABP_first  NIMAP_first  NISysABP_first  Te
##   6   2  Gender  GCS_first  Glucose_first  HR_first  NIDiasABP_first  NIMAP_first  NISysABP_first  Te
##   6   3  Gender  GCS_first  Glucose_first  HR_first  NIDiasABP_first  NIMAP_first  NISysABP_first  Te
##   6   4  Gender  GCS_first  Glucose_first  HR_first  NIDiasABP_first  NIMAP_first  NISysABP_first  Te
##   6   5  Gender  GCS_first  Glucose_first  HR_first  NIDiasABP_first  NIMAP_first  NISysABP_first  Te
##   6   6  Gender  GCS_first  Glucose_first  HR_first  NIDiasABP_first  NIMAP_first  NISysABP_first  Te
##   6   7  Gender  GCS_first  Glucose_first  HR_first  NIDiasABP_first  NIMAP_first  NISysABP_first  Te
##   6   8  Gender  GCS_first  Glucose_first  HR_first  NIDiasABP_first  NIMAP_first  NISysABP_first  Te
##   6   9  Gender  GCS_first  Glucose_first  HR_first  NIDiasABP_first  NIMAP_first  NISysABP_first  Te
##   6   10  Gender  GCS_first  Glucose_first  HR_first  NIDiasABP_first  NIMAP_first  NISysABP_first  T
##   6   11  Gender  GCS_first  Glucose_first  HR_first  NIDiasABP_first  NIMAP_first  NISysABP_first  T
##   6   12  Gender  GCS_first  Glucose_first  HR_first  NIDiasABP_first  NIMAP_first  NISysABP_first  T
##   6   13  Gender  GCS_first  Glucose_first  HR_first  NIDiasABP_first  NIMAP_first  NISysABP_first  T
##   6   14  Gender  GCS_first  Glucose_first  HR_first  NIDiasABP_first  NIMAP_first  NISysABP_first  T
##   6   15  Gender  GCS_first  Glucose_first  HR_first  NIDiasABP_first  NIMAP_first  NISysABP_first  T
##   6   16  Gender  GCS_first  Glucose_first  HR_first  NIDiasABP_first  NIMAP_first  NISysABP_first  T
##   6   17  Gender  GCS_first  Glucose_first  HR_first  NIDiasABP_first  NIMAP_first  NISysABP_first  T
##   6   18  Gender  GCS_first  Glucose_first  HR_first  NIDiasABP_first  NIMAP_first  NISysABP_first  T
##   6   19  Gender  GCS_first  Glucose_first  HR_first  NIDiasABP_first  NIMAP_first  NISysABP_first  T
##   6   20  Gender  GCS_first  Glucose_first  HR_first  NIDiasABP_first  NIMAP_first  NISysABP_first  T
##   6   21  Gender  GCS_first  Glucose_first  HR_first  NIDiasABP_first  NIMAP_first  NISysABP_first  T
##   6   22  Gender  GCS_first  Glucose_first  HR_first  NIDiasABP_first  NIMAP_first  NISysABP_first  T
##   6   23  Gender  GCS_first  Glucose_first  HR_first  NIDiasABP_first  NIMAP_first  NISysABP_first  T
##   6   24  Gender  GCS_first  Glucose_first  HR_first  NIDiasABP_first  NIMAP_first  NISysABP_first  T
##   6   25  Gender  GCS_first  Glucose_first  HR_first  NIDiasABP_first  NIMAP_first  NISysABP_first  T
##   6   26  Gender  GCS_first  Glucose_first  HR_first  NIDiasABP_first  NIMAP_first  NISysABP_first  T
##   6   27  Gender  GCS_first  Glucose_first  HR_first  NIDiasABP_first  NIMAP_first  NISysABP_first  T
##   6   28  Gender  GCS_first  Glucose_first  HR_first  NIDiasABP_first  NIMAP_first  NISysABP_first  T
##   6   29  Gender  GCS_first  Glucose_first  HR_first  NIDiasABP_first  NIMAP_first  NISysABP_first  T
##   6   30  Gender  GCS_first  Glucose_first  HR_first  NIDiasABP_first  NIMAP_first  NISysABP_first  T
##   6   31  Gender  GCS_first  Glucose_first  HR_first  NIDiasABP_first  NIMAP_first  NISysABP_first  T
##   6   32  Gender  GCS_first  Glucose_first  HR_first  NIDiasABP_first  NIMAP_first  NISysABP_first  T
##   6   33  Gender  GCS_first  Glucose_first  HR_first  NIDiasABP_first  NIMAP_first  NISysABP_first  T
##   6   34  Gender  GCS_first  Glucose_first  HR_first  NIDiasABP_first  NIMAP_first  NISysABP_first  T
##   6   35  Gender  GCS_first  Glucose_first  HR_first  NIDiasABP_first  NIMAP_first  NISysABP_first  T
##   6   36  Gender  GCS_first  Glucose_first  HR_first  NIDiasABP_first  NIMAP_first  NISysABP_first  T
##   6   37  Gender  GCS_first  Glucose_first  HR_first  NIDiasABP_first  NIMAP_first  NISysABP_first  T
```

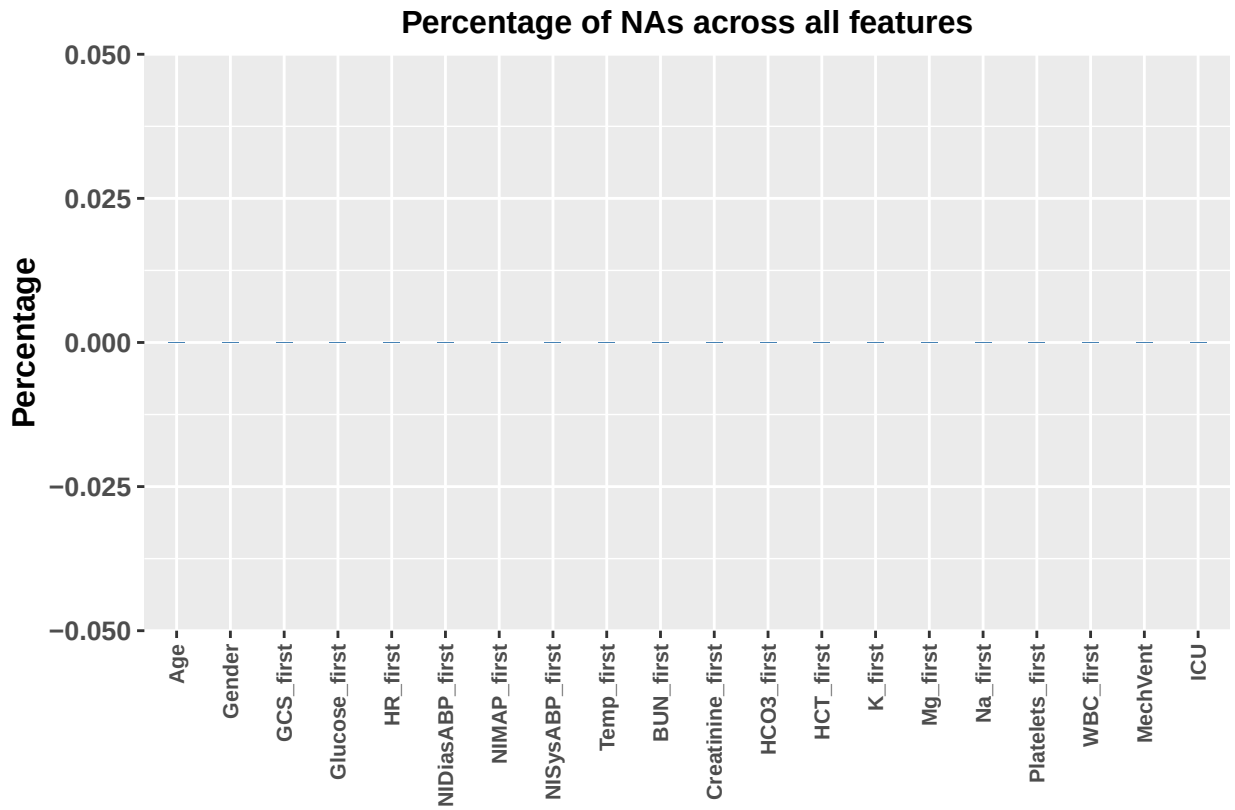
```
##   6   38  Gender  GCS_first  Glucose_first  HR_first  NIDiasABP_first  NIMAP_first  NISysABP_first  T
##   6   39  Gender  GCS_first  Glucose_first  HR_first  NIDiasABP_first  NIMAP_first  NISysABP_first  T
##   6   40  Gender  GCS_first  Glucose_first  HR_first  NIDiasABP_first  NIMAP_first  NISysABP_first  T
##   7   1  Gender  GCS_first  Glucose_first  HR_first  NIDiasABP_first  NIMAP_first  NISysABP_first  Te
##   7   2  Gender  GCS_first  Glucose_first  HR_first  NIDiasABP_first  NIMAP_first  NISysABP_first  Te
##   7   3  Gender  GCS_first  Glucose_first  HR_first  NIDiasABP_first  NIMAP_first  NISysABP_first  Te
##   7   4  Gender  GCS_first  Glucose_first  HR_first  NIDiasABP_first  NIMAP_first  NISysABP_first  Te
##   7   5  Gender  GCS_first  Glucose_first  HR_first  NIDiasABP_first  NIMAP_first  NISysABP_first  Te
##   7   6  Gender  GCS_first  Glucose_first  HR_first  NIDiasABP_first  NIMAP_first  NISysABP_first  Te
##   7   7  Gender  GCS_first  Glucose_first  HR_first  NIDiasABP_first  NIMAP_first  NISysABP_first  Te
##   7   8  Gender  GCS_first  Glucose_first  HR_first  NIDiasABP_first  NIMAP_first  NISysABP_first  Te
##   7   9  Gender  GCS_first  Glucose_first  HR_first  NIDiasABP_first  NIMAP_first  NISysABP_first  Te
##   7   10  Gender  GCS_first  Glucose_first  HR_first  NIDiasABP_first  NIMAP_first  NISysABP_first  T
##   7   11  Gender  GCS_first  Glucose_first  HR_first  NIDiasABP_first  NIMAP_first  NISysABP_first  T
##   7   12  Gender  GCS_first  Glucose_first  HR_first  NIDiasABP_first  NIMAP_first  NISysABP_first  T
##   7   13  Gender  GCS_first  Glucose_first  HR_first  NIDiasABP_first  NIMAP_first  NISysABP_first  T
##   7   14  Gender  GCS_first  Glucose_first  HR_first  NIDiasABP_first  NIMAP_first  NISysABP_first  T
##   7   15  Gender  GCS_first  Glucose_first  HR_first  NIDiasABP_first  NIMAP_first  NISysABP_first  T
##   7   16  Gender  GCS_first  Glucose_first  HR_first  NIDiasABP_first  NIMAP_first  NISysABP_first  T
##   7   17  Gender  GCS_first  Glucose_first  HR_first  NIDiasABP_first  NIMAP_first  NISysABP_first  T
##   7   18  Gender  GCS_first  Glucose_first  HR_first  NIDiasABP_first  NIMAP_first  NISysABP_first  T
##   7   19  Gender  GCS_first  Glucose_first  HR_first  NIDiasABP_first  NIMAP_first  NISysABP_first  T
##   7   20  Gender  GCS_first  Glucose_first  HR_first  NIDiasABP_first  NIMAP_first  NISysABP_first  T
##   7   21  Gender  GCS_first  Glucose_first  HR_first  NIDiasABP_first  NIMAP_first  NISysABP_first  T
##   7   22  Gender  GCS_first  Glucose_first  HR_first  NIDiasABP_first  NIMAP_first  NISysABP_first  T
##   7   23  Gender  GCS_first  Glucose_first  HR_first  NIDiasABP_first  NIMAP_first  NISysABP_first  T
##   7   24  Gender  GCS_first  Glucose_first  HR_first  NIDiasABP_first  NIMAP_first  NISysABP_first  T
##   7   25  Gender  GCS_first  Glucose_first  HR_first  NIDiasABP_first  NIMAP_first  NISysABP_first  T
##   7   26  Gender  GCS_first  Glucose_first  HR_first  NIDiasABP_first  NIMAP_first  NISysABP_first  T
##   7   27  Gender  GCS_first  Glucose_first  HR_first  NIDiasABP_first  NIMAP_first  NISysABP_first  T
##   7   28  Gender  GCS_first  Glucose_first  HR_first  NIDiasABP_first  NIMAP_first  NISysABP_first  T
##   7   29  Gender  GCS_first  Glucose_first  HR_first  NIDiasABP_first  NIMAP_first  NISysABP_first  T
##   7   30  Gender  GCS_first  Glucose_first  HR_first  NIDiasABP_first  NIMAP_first  NISysABP_first  T
##   7   31  Gender  GCS_first  Glucose_first  HR_first  NIDiasABP_first  NIMAP_first  NISysABP_first  T
##   7   32  Gender  GCS_first  Glucose_first  HR_first  NIDiasABP_first  NIMAP_first  NISysABP_first  T
##   7   33  Gender  GCS_first  Glucose_first  HR_first  NIDiasABP_first  NIMAP_first  NISysABP_first  T
##   7   34  Gender  GCS_first  Glucose_first  HR_first  NIDiasABP_first  NIMAP_first  NISysABP_first  T
##   7   35  Gender  GCS_first  Glucose_first  HR_first  NIDiasABP_first  NIMAP_first  NISysABP_first  T
##   7   36  Gender  GCS_first  Glucose_first  HR_first  NIDiasABP_first  NIMAP_first  NISysABP_first  T
##   7   37  Gender  GCS_first  Glucose_first  HR_first  NIDiasABP_first  NIMAP_first  NISysABP_first  T
##   7   38  Gender  GCS_first  Glucose_first  HR_first  NIDiasABP_first  NIMAP_first  NISysABP_first  T
##   7   39  Gender  GCS_first  Glucose_first  HR_first  NIDiasABP_first  NIMAP_first  NISysABP_first  T
##   7   40  Gender  GCS_first  Glucose_first  HR_first  NIDiasABP_first  NIMAP_first  NISysABP_first  T
##   8   1  Gender  GCS_first  Glucose_first  HR_first  NIDiasABP_first  NIMAP_first  NISysABP_first  Te
##   8   2  Gender  GCS_first  Glucose_first  HR_first  NIDiasABP_first  NIMAP_first  NISysABP_first  Te
##   8   3  Gender  GCS_first  Glucose_first  HR_first  NIDiasABP_first  NIMAP_first  NISysABP_first  Te
##   8   4  Gender  GCS_first  Glucose_first  HR_first  NIDiasABP_first  NIMAP_first  NISysABP_first  Te
##   8   5  Gender  GCS_first  Glucose_first  HR_first  NIDiasABP_first  NIMAP_first  NISysABP_first  Te
##   8   6  Gender  GCS_first  Glucose_first  HR_first  NIDiasABP_first  NIMAP_first  NISysABP_first  Te
##   8   7  Gender  GCS_first  Glucose_first  HR_first  NIDiasABP_first  NIMAP_first  NISysABP_first  Te
##   8   8  Gender  GCS_first  Glucose_first  HR_first  NIDiasABP_first  NIMAP_first  NISysABP_first  Te
##   8   9  Gender  GCS_first  Glucose_first  HR_first  NIDiasABP_first  NIMAP_first  NISysABP_first  Te
##   8   10  Gender  GCS_first  Glucose_first  HR_first  NIDiasABP_first  NIMAP_first  NISysABP_first  T
##   8   11  Gender  GCS_first  Glucose_first  HR_first  NIDiasABP_first  NIMAP_first  NISysABP_first  T
```

```
##   8   12  Gender  GCS_first  Glucose_first  HR_first  NIDiasABP_first  NIMAP_first  NISysABP_first  T
##   8   13  Gender  GCS_first  Glucose_first  HR_first  NIDiasABP_first  NIMAP_first  NISysABP_first  T
##   8   14  Gender  GCS_first  Glucose_first  HR_first  NIDiasABP_first  NIMAP_first  NISysABP_first  T
##   8   15  Gender  GCS_first  Glucose_first  HR_first  NIDiasABP_first  NIMAP_first  NISysABP_first  T
##   8   16  Gender  GCS_first  Glucose_first  HR_first  NIDiasABP_first  NIMAP_first  NISysABP_first  T
##   8   17  Gender  GCS_first  Glucose_first  HR_first  NIDiasABP_first  NIMAP_first  NISysABP_first  T
##   8   18  Gender  GCS_first  Glucose_first  HR_first  NIDiasABP_first  NIMAP_first  NISysABP_first  T
##   8   19  Gender  GCS_first  Glucose_first  HR_first  NIDiasABP_first  NIMAP_first  NISysABP_first  T
##   8   20  Gender  GCS_first  Glucose_first  HR_first  NIDiasABP_first  NIMAP_first  NISysABP_first  T
##   8   21  Gender  GCS_first  Glucose_first  HR_first  NIDiasABP_first  NIMAP_first  NISysABP_first  T
##   8   22  Gender  GCS_first  Glucose_first  HR_first  NIDiasABP_first  NIMAP_first  NISysABP_first  T
##   8   23  Gender  GCS_first  Glucose_first  HR_first  NIDiasABP_first  NIMAP_first  NISysABP_first  T
##   8   24  Gender  GCS_first  Glucose_first  HR_first  NIDiasABP_first  NIMAP_first  NISysABP_first  T
##   8   25  Gender  GCS_first  Glucose_first  HR_first  NIDiasABP_first  NIMAP_first  NISysABP_first  T
##   8   26  Gender  GCS_first  Glucose_first  HR_first  NIDiasABP_first  NIMAP_first  NISysABP_first  T
##   8   27  Gender  GCS_first  Glucose_first  HR_first  NIDiasABP_first  NIMAP_first  NISysABP_first  T
##   8   28  Gender  GCS_first  Glucose_first  HR_first  NIDiasABP_first  NIMAP_first  NISysABP_first  T
##   8   29  Gender  GCS_first  Glucose_first  HR_first  NIDiasABP_first  NIMAP_first  NISysABP_first  T
##   8   30  Gender  GCS_first  Glucose_first  HR_first  NIDiasABP_first  NIMAP_first  NISysABP_first  T
##   8   31  Gender  GCS_first  Glucose_first  HR_first  NIDiasABP_first  NIMAP_first  NISysABP_first  T
##   8   32  Gender  GCS_first  Glucose_first  HR_first  NIDiasABP_first  NIMAP_first  NISysABP_first  T
##   8   33  Gender  GCS_first  Glucose_first  HR_first  NIDiasABP_first  NIMAP_first  NISysABP_first  T
##   8   34  Gender  GCS_first  Glucose_first  HR_first  NIDiasABP_first  NIMAP_first  NISysABP_first  T
##   8   35  Gender  GCS_first  Glucose_first  HR_first  NIDiasABP_first  NIMAP_first  NISysABP_first  T
##   8   36  Gender  GCS_first  Glucose_first  HR_first  NIDiasABP_first  NIMAP_first  NISysABP_first  T
##   8   37  Gender  GCS_first  Glucose_first  HR_first  NIDiasABP_first  NIMAP_first  NISysABP_first  T
##   8   38  Gender  GCS_first  Glucose_first  HR_first  NIDiasABP_first  NIMAP_first  NISysABP_first  T
##   8   39  Gender  GCS_first  Glucose_first  HR_first  NIDiasABP_first  NIMAP_first  NISysABP_first  T
##   8   40  Gender  GCS_first  Glucose_first  HR_first  NIDiasABP_first  NIMAP_first  NISysABP_first  T
##   9   1  Gender  GCS_first  Glucose_first  HR_first  NIDiasABP_first  NIMAP_first  NISysABP_first  To
##   9   2  Gender  GCS_first  Glucose_first  HR_first  NIDiasABP_first  NIMAP_first  NISysABP_first  To
##   9   3  Gender  GCS_first  Glucose_first  HR_first  NIDiasABP_first  NIMAP_first  NISysABP_first  To
##   9   4  Gender  GCS_first  Glucose_first  HR_first  NIDiasABP_first  NIMAP_first  NISysABP_first  To
##   9   5  Gender  GCS_first  Glucose_first  HR_first  NIDiasABP_first  NIMAP_first  NISysABP_first  To
##   9   6  Gender  GCS_first  Glucose_first  HR_first  NIDiasABP_first  NIMAP_first  NISysABP_first  To
##   9   7  Gender  GCS_first  Glucose_first  HR_first  NIDiasABP_first  NIMAP_first  NISysABP_first  To
##   9   8  Gender  GCS_first  Glucose_first  HR_first  NIDiasABP_first  NIMAP_first  NISysABP_first  To
##   9   9  Gender  GCS_first  Glucose_first  HR_first  NIDiasABP_first  NIMAP_first  NISysABP_first  To
##   9   10  Gender  GCS_first  Glucose_first  HR_first  NIDiasABP_first  NIMAP_first  NISysABP_first  T
##   9   11  Gender  GCS_first  Glucose_first  HR_first  NIDiasABP_first  NIMAP_first  NISysABP_first  T
##   9   12  Gender  GCS_first  Glucose_first  HR_first  NIDiasABP_first  NIMAP_first  NISysABP_first  T
##   9   13  Gender  GCS_first  Glucose_first  HR_first  NIDiasABP_first  NIMAP_first  NISysABP_first  T
##   9   14  Gender  GCS_first  Glucose_first  HR_first  NIDiasABP_first  NIMAP_first  NISysABP_first  T
##   9   15  Gender  GCS_first  Glucose_first  HR_first  NIDiasABP_first  NIMAP_first  NISysABP_first  T
##   9   16  Gender  GCS_first  Glucose_first  HR_first  NIDiasABP_first  NIMAP_first  NISysABP_first  T
##   9   17  Gender  GCS_first  Glucose_first  HR_first  NIDiasABP_first  NIMAP_first  NISysABP_first  T
##   9   18  Gender  GCS_first  Glucose_first  HR_first  NIDiasABP_first  NIMAP_first  NISysABP_first  T
##   9   19  Gender  GCS_first  Glucose_first  HR_first  NIDiasABP_first  NIMAP_first  NISysABP_first  T
##   9   20  Gender  GCS_first  Glucose_first  HR_first  NIDiasABP_first  NIMAP_first  NISysABP_first  T
##   9   21  Gender  GCS_first  Glucose_first  HR_first  NIDiasABP_first  NIMAP_first  NISysABP_first  T
##   9   22  Gender  GCS_first  Glucose_first  HR_first  NIDiasABP_first  NIMAP_first  NISysABP_first  T
##   9   23  Gender  GCS_first  Glucose_first  HR_first  NIDiasABP_first  NIMAP_first  NISysABP_first  T
##   9   24  Gender  GCS_first  Glucose_first  HR_first  NIDiasABP_first  NIMAP_first  NISysABP_first  T
##   9   25  Gender  GCS_first  Glucose_first  HR_first  NIDiasABP_first  NIMAP_first  NISysABP_first  T
```

```
##   9  26 Gender GCS_first Glucose_first HR_first NIDiasABP_first NIMAP_first NISysABP_first T
##   9  27 Gender GCS_first Glucose_first HR_first NIDiasABP_first NIMAP_first NISysABP_first T
##   9  28 Gender GCS_first Glucose_first HR_first NIDiasABP_first NIMAP_first NISysABP_first T
##   9  29 Gender GCS_first Glucose_first HR_first NIDiasABP_first NIMAP_first NISysABP_first T
##   9  30 Gender GCS_first Glucose_first HR_first NIDiasABP_first NIMAP_first NISysABP_first T
##   9  31 Gender GCS_first Glucose_first HR_first NIDiasABP_first NIMAP_first NISysABP_first T
##   9  32 Gender GCS_first Glucose_first HR_first NIDiasABP_first NIMAP_first NISysABP_first T
##   9  33 Gender GCS_first Glucose_first HR_first NIDiasABP_first NIMAP_first NISysABP_first T
##   9  34 Gender GCS_first Glucose_first HR_first NIDiasABP_first NIMAP_first NISysABP_first T
##   9  35 Gender GCS_first Glucose_first HR_first NIDiasABP_first NIMAP_first NISysABP_first T
##   9  36 Gender GCS_first Glucose_first HR_first NIDiasABP_first NIMAP_first NISysABP_first T
##   9  37 Gender GCS_first Glucose_first HR_first NIDiasABP_first NIMAP_first NISysABP_first T
##   9  38 Gender GCS_first Glucose_first HR_first NIDiasABP_first NIMAP_first NISysABP_first T
##   9  39 Gender GCS_first Glucose_first HR_first NIDiasABP_first NIMAP_first NISysABP_first T
##   9  40 Gender GCS_first Glucose_first HR_first NIDiasABP_first NIMAP_first NISysABP_first T
##  10   1 Gender GCS_first Glucose_first HR_first NIDiasABP_first NIMAP_first NISysABP_first T
##  10   2 Gender GCS_first Glucose_first HR_first NIDiasABP_first NIMAP_first NISysABP_first T
##  10   3 Gender GCS_first Glucose_first HR_first NIDiasABP_first NIMAP_first NISysABP_first T
##  10   4 Gender GCS_first Glucose_first HR_first NIDiasABP_first NIMAP_first NISysABP_first T
##  10   5 Gender GCS_first Glucose_first HR_first NIDiasABP_first NIMAP_first NISysABP_first T
##  10   6 Gender GCS_first Glucose_first HR_first NIDiasABP_first NIMAP_first NISysABP_first T
##  10   7 Gender GCS_first Glucose_first HR_first NIDiasABP_first NIMAP_first NISysABP_first T
##  10   8 Gender GCS_first Glucose_first HR_first NIDiasABP_first NIMAP_first NISysABP_first T
##  10   9 Gender GCS_first Glucose_first HR_first NIDiasABP_first NIMAP_first NISysABP_first T
##  10  10 Gender GCS_first Glucose_first HR_first NIDiasABP_first NIMAP_first NISysABP_first
##  10  11 Gender GCS_first Glucose_first HR_first NIDiasABP_first NIMAP_first NISysABP_first
##  10  12 Gender GCS_first Glucose_first HR_first NIDiasABP_first NIMAP_first NISysABP_first
##  10  13 Gender GCS_first Glucose_first HR_first NIDiasABP_first NIMAP_first NISysABP_first
##  10  14 Gender GCS_first Glucose_first HR_first NIDiasABP_first NIMAP_first NISysABP_first
##  10  15 Gender GCS_first Glucose_first HR_first NIDiasABP_first NIMAP_first NISysABP_first
##  10  16 Gender GCS_first Glucose_first HR_first NIDiasABP_first NIMAP_first NISysABP_first
##  10  17 Gender GCS_first Glucose_first HR_first NIDiasABP_first NIMAP_first NISysABP_first
##  10  18 Gender GCS_first Glucose_first HR_first NIDiasABP_first NIMAP_first NISysABP_first
##  10  19 Gender GCS_first Glucose_first HR_first NIDiasABP_first NIMAP_first NISysABP_first
##  10  20 Gender GCS_first Glucose_first HR_first NIDiasABP_first NIMAP_first NISysABP_first
##  10  21 Gender GCS_first Glucose_first HR_first NIDiasABP_first NIMAP_first NISysABP_first
##  10  22 Gender GCS_first Glucose_first HR_first NIDiasABP_first NIMAP_first NISysABP_first
##  10  23 Gender GCS_first Glucose_first HR_first NIDiasABP_first NIMAP_first NISysABP_first
##  10  24 Gender GCS_first Glucose_first HR_first NIDiasABP_first NIMAP_first NISysABP_first
##  10  25 Gender GCS_first Glucose_first HR_first NIDiasABP_first NIMAP_first NISysABP_first
##  10  26 Gender GCS_first Glucose_first HR_first NIDiasABP_first NIMAP_first NISysABP_first
##  10  27 Gender GCS_first Glucose_first HR_first NIDiasABP_first NIMAP_first NISysABP_first
##  10  28 Gender GCS_first Glucose_first HR_first NIDiasABP_first NIMAP_first NISysABP_first
##  10  29 Gender GCS_first Glucose_first HR_first NIDiasABP_first NIMAP_first NISysABP_first
##  10  30 Gender GCS_first Glucose_first HR_first NIDiasABP_first NIMAP_first NISysABP_first
##  10  31 Gender GCS_first Glucose_first HR_first NIDiasABP_first NIMAP_first NISysABP_first
##  10  32 Gender GCS_first Glucose_first HR_first NIDiasABP_first NIMAP_first NISysABP_first
##  10  33 Gender GCS_first Glucose_first HR_first NIDiasABP_first NIMAP_first NISysABP_first
##  10  34 Gender GCS_first Glucose_first HR_first NIDiasABP_first NIMAP_first NISysABP_first
##  10  35 Gender GCS_first Glucose_first HR_first NIDiasABP_first NIMAP_first NISysABP_first
##  10  36 Gender GCS_first Glucose_first HR_first NIDiasABP_first NIMAP_first NISysABP_first
##  10  37 Gender GCS_first Glucose_first HR_first NIDiasABP_first NIMAP_first NISysABP_first
##  10  38 Gender GCS_first Glucose_first HR_first NIDiasABP_first NIMAP_first NISysABP_first
##  10  39 Gender GCS_first Glucose_first HR_first NIDiasABP_first NIMAP_first NISysABP_first
```

```
##   10   40  Gender  GCS_first  Glucose_first  HR_first  NIDiasABP_first  NIMAP_first  NISysABP_first
```

```
pMissDF = setNames(stack(sapply(dfICU, function(x){(sum(is.na(x))/length(x))*100}))[2:1], c('Feature','V
p = ggplot(data = pMissDF, aes(x = Feature, y = Value)) +
  geom_bar(stat = 'identity', fill = 'steelblue', width = 0.3) +
  theme(text = element_text(size = 14, face = 'bold'),
  axis.text.x = element_text(angle = 90, hjust = 1, vjust = 0.5)) +
  xlab('') + ylab('Percentage') +
  ggtitle('Percentage of NAs across all features')  +
  theme(plot.title = element_text(size = 12, hjust = 0.5),
        axis.text = element_text(size = 8),
        axis.text.x = element_text(size = 8),
        axis.text.y = element_text(size = 10),
        axis.title = element_text(size = 12, face = "bold"))
p
```




```
features = colnames(dfICU)
categorical_features = c('Gender', 'GCS_first', 'MechVent', 'ICU')
continuous_features = features[c(!(colnames(dfICU) %in% categorical_features))]
dfICU_continuous = dfICU %>% select(continuous_features)
```

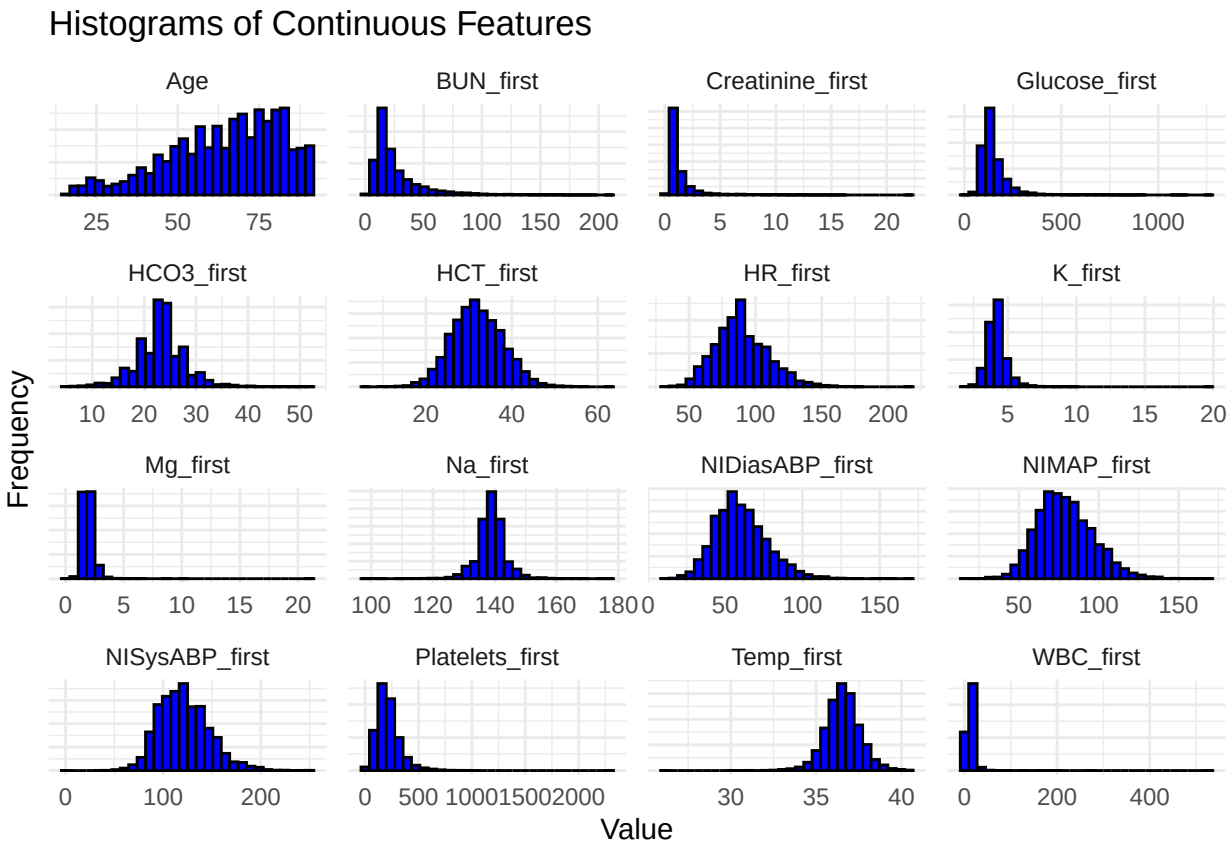
```
## Warning: Using an external vector in selections was deprecated in tidyselect 1.1.0.
## i Please use ‘all_of()‘ or ‘any_of()‘ instead.
##   # Was:
##   data %>% select(continuous_features)
```

```
##
##   # Now:
##   data %>% select(all_of(continuous_features))
##
## See <https://tidyselect.r-lib.org/reference/faq-external-vector.html>.
## This warning is displayed once every 8 hours.
## Call `lifecycle::last_lifecycle_warnings()` to see where this warning was
## generated.
```

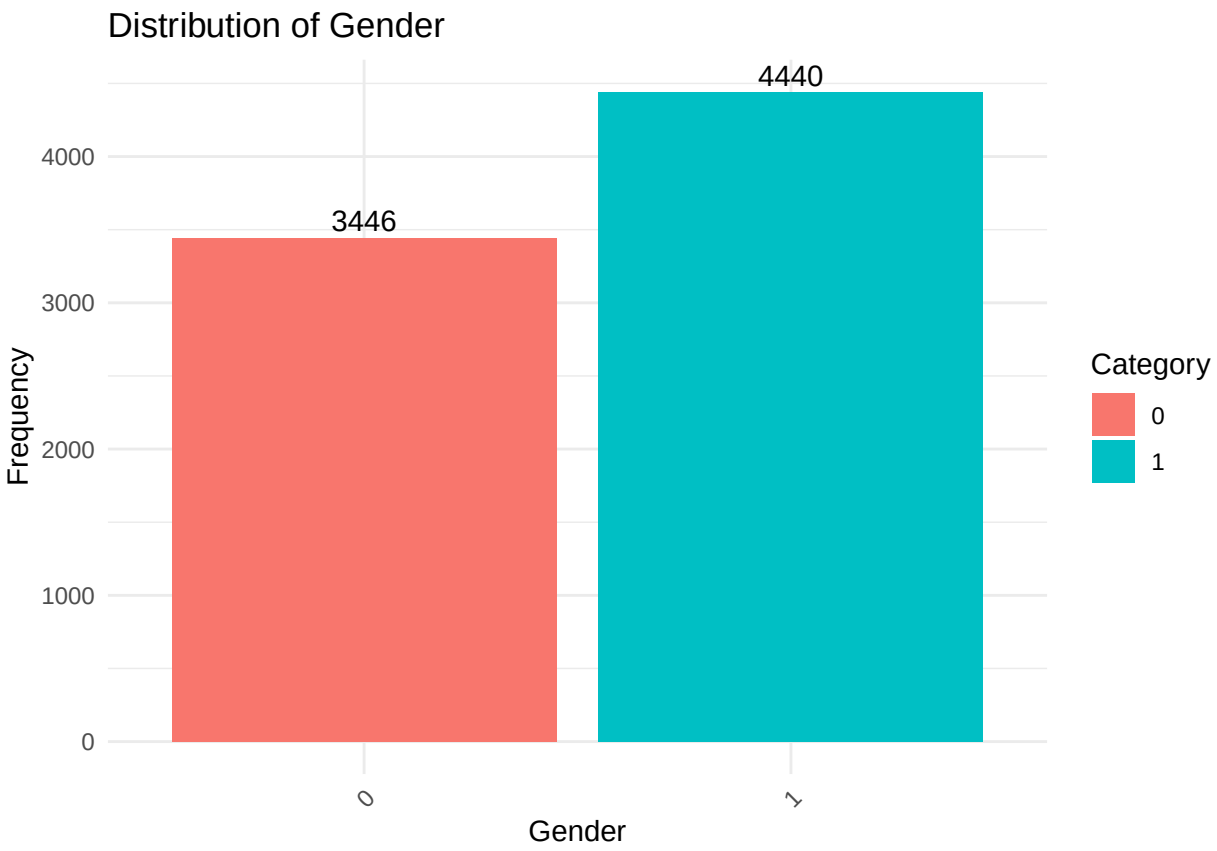
```
dfICU_categorical = dfICU %>% select(categorical_features)
```

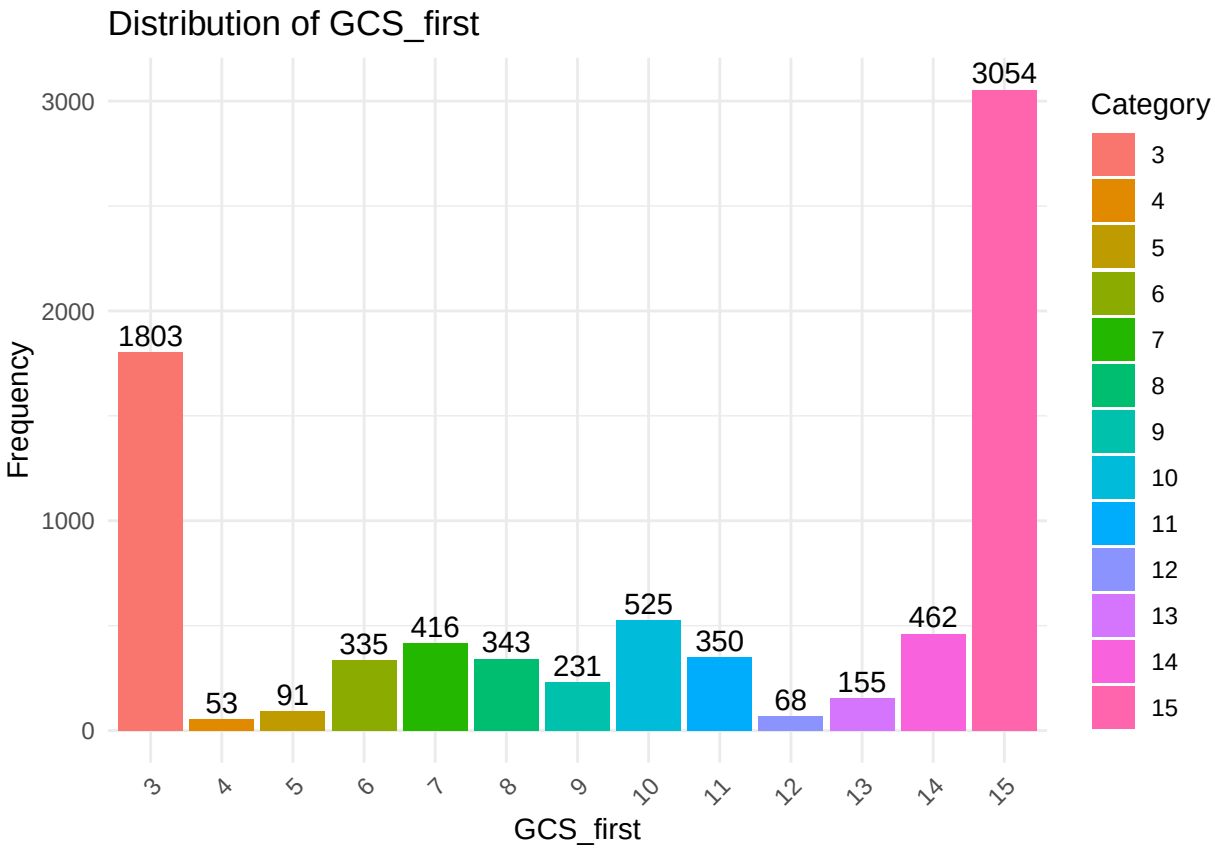
```
## Warning: Using an external vector in selections was deprecated in tidyselect 1.1.0.
## i Please use `all_of()` or `any_of()` instead.
##   # Was:
##   data %>% select(categorical_features)
##
##   # Now:
##   data %>% select(all_of(categorical_features))
##
## See <https://tidyselect.r-lib.org/reference/faq-external-vector.html>.
## This warning is displayed once every 8 hours.
## Call `lifecycle::last_lifecycle_warnings()` to see where this warning was
## generated.
```

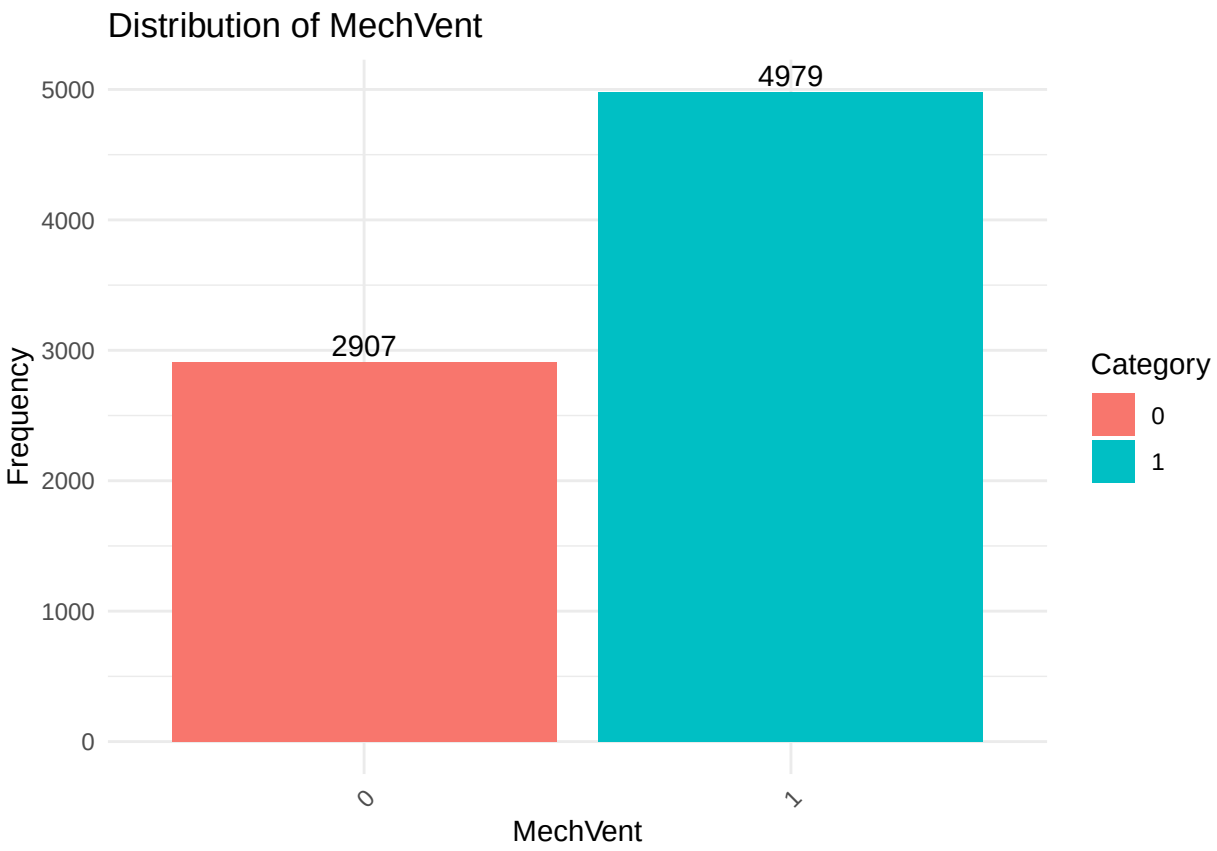
```
dfICU_long = dfICU_continuous %>% pivot_longer(cols = everything(), names_to = "variable", values_to =

ggplot(dfICU_long, aes(x = value)) +
  geom_histogram(bins = 30, fill = "blue", color = "black") +
  facet_wrap(~ variable, scales = "free") +
  labs(title = "Histograms of Continuous Features",
       x = "Value",
       y = "Frequency") +
  theme_minimal() +
  theme(axis.text.y = element_blank(),
        axis.ticks.y = element_blank())
```

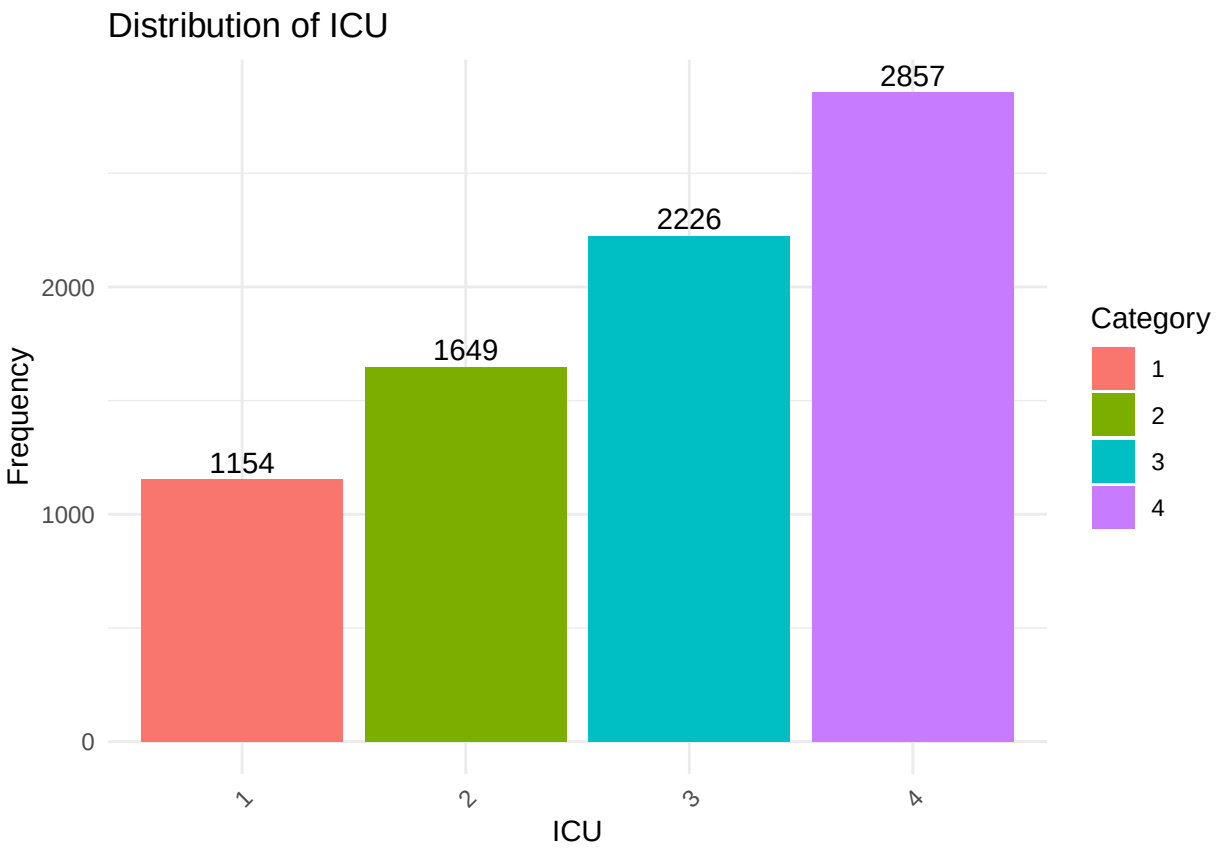



```
categorical_cols <- names(dfICU_categorical)

for (col in categorical_cols) {
  freq_table <- table(dfICU_categorical[[col]])
  freq_df <- as.data.frame(freq_table)
  colnames(freq_df) <- c("Category", "Frequency")

  p <- ggplot(freq_df, aes(x = Category, y = Frequency, fill = Category)) +
    geom_bar(stat = "identity") +
    geom_text(aes(label = Frequency), vjust = -0.3) +
    labs(title = paste("Distribution of", col),
         x = col,
         y = "Frequency") +
    theme_minimal() +
    theme(axis.text.x = element_text(angle = 45, hjust = 1))

  print(p)
}
```


Distribution of Gender
4440
3446
4000
3000
2000
1000
0
Frequency
0
1
Gender
Category
0
1


Distribution of GCS_first
Frequency
GCS_first
Category
3
4
5
6
7
8
9
10
11
12
13
14
15
0
1000
2000
3000
1803
53
91
335
416
343
231
525
350
68
155
462
3054


Distribution of MechVent
5000
4000
3000
2000
1000
0
Frequency
2907
4979
0
1
MechVent
Category
0
1




```
dfICU_categorical <- dfICU[, c("Gender", "GCS_first", "MechVent", "ICU")]
gower_dist_matrix <- daisy(dfICU_categorical, metric = "gower")
```

```
## Warning in daisy(dfICU_categorical, metric = "gower"): binary variable(s) 1, 3
## treated as interval scaled
```

```
gower_dist_matrix <- as.matrix(gower_dist_matrix)
avg_distance <- rowMeans(gower_dist_matrix)

dfICU$categorical_avg_distance <- avg_distance
```

```
features = colnames(dfICU)
categorical_features_1 = c('Gender', 'GCS_first', 'MechVent', 'ICU')
continuous_features_1 = c('Age', 'Glucose_first', 'HR_first', 'NIDiasABP_first', 'NIMAP_first', 'NISysA
```

```
dfICU_continuous_1 = dfICU %>% select(continuous_features_1)
```

```
## Warning: Using an external vector in selections was deprecated in tidyselect 1.1.0.
## i Please use ‘all_of()‘ or ‘any_of()‘ instead.
##   # Was:
##   data %>% select(continuous_features_1)
##
##   # Now:
##   data %>% select(all_of(continuous_features_1))
```

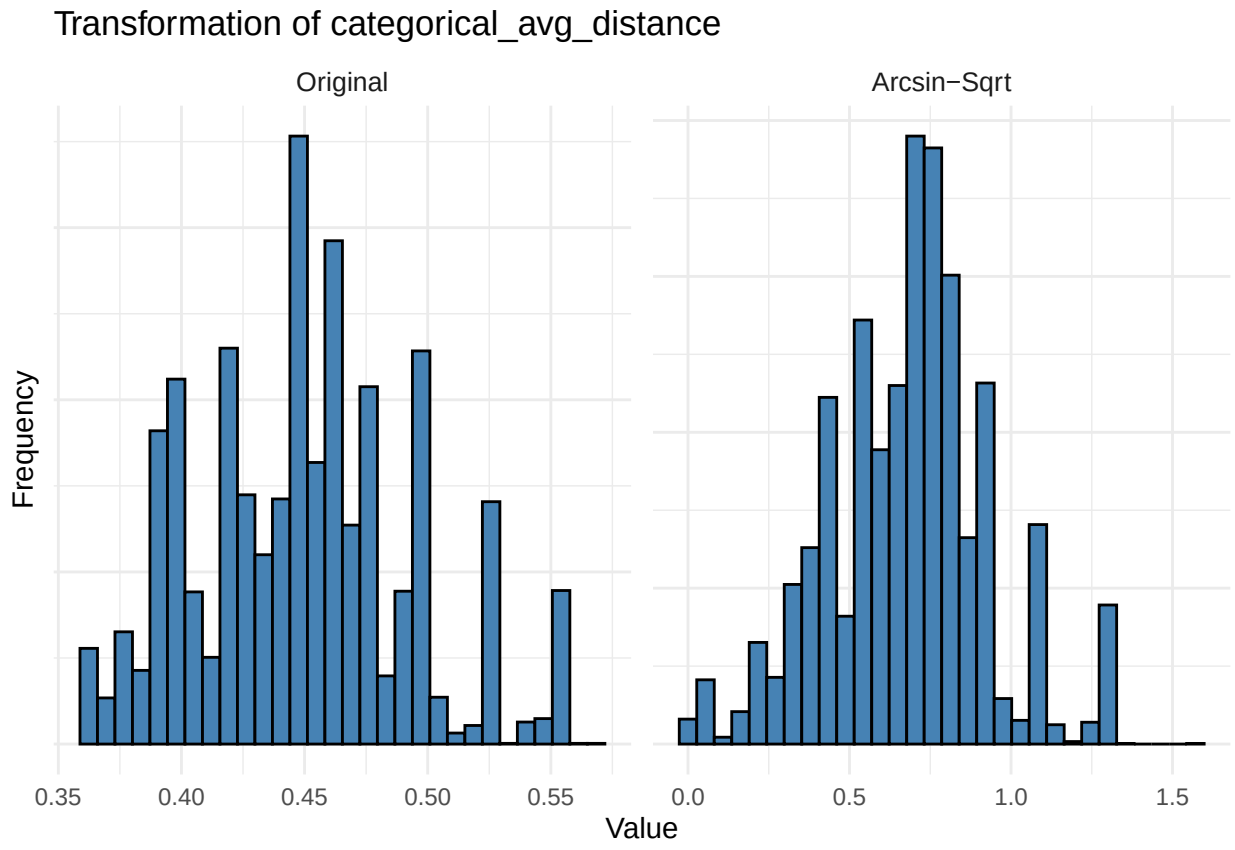
```
##
## See <https://tidyselect.r-lib.org/reference/faq-external-vector.html>.
## This warning is displayed once every 8 hours.
## Call `lifecycle::last_lifecycle_warnings()` to see where this warning was
## generated.
```

```
dfICU_categorical_1 = dfICU %>% select(categorical_features_1)
```

```
## Warning: Using an external vector in selections was deprecated in tidyselect 1.1.0.
## i Please use `all_of()` or `any_of()` instead.
##   # Was:
##   data %>% select(categorical_features_1)
##
##   # Now:
##   data %>% select(all_of(categorical_features_1))
##
## See <https://tidyselect.r-lib.org/reference/faq-external-vector.html>.
## This warning is displayed once every 8 hours.
## Call `lifecycle::last_lifecycle_warnings()` to see where this warning was
## generated.
```

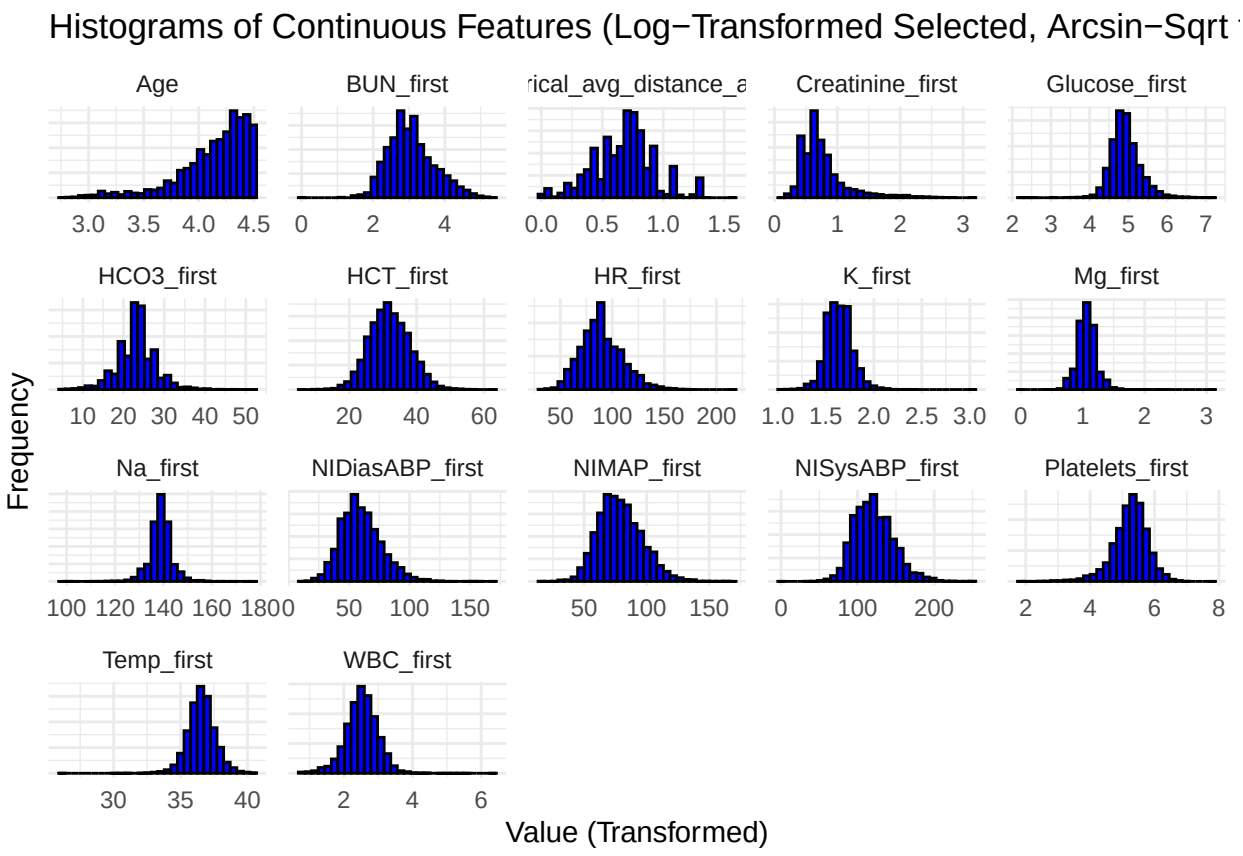
```
feature <- "categorical_avg_distance"

df_transformed <- dfICU_continuous_1 %>%
  select(all_of(feature)) %>%
  mutate(
    asinsqrt_transformed = {
      scaled_value <- (.data[[feature]] - min(.data[[feature]])) /
        (max(.data[[feature]]) - min(.data[[feature]]))
      asin(sqrt(scaled_value))
    }
  )

df_long <- df_transformed %>%
  pivot_longer(
    cols = everything(),
    names_to = "transformation",
    values_to = "value"
  ) %>%
  mutate(
    transformation = factor(
      transformation,
      levels = c(feature, "asinsqrt_transformed"),
      labels = c("Original", "Arcsin-Sqrt")
    )
  )

p <- ggplot(df_long, aes(x = value)) +
  geom_histogram(bins = 30, fill = "steelblue", color = "black") +
  facet_wrap(~ transformation, scales = "free", ncol = 2) +
  labs(
    title = paste("Transformation of", feature),
    x = "Value",
```

```
    y = "Frequency"
  ) +
  theme_minimal() +
  theme(
    axis.text.y = element_blank(),
    axis.ticks.y = element_blank(),
    strip.text = element_text(size = 10)
  )
print(p)
```




```
cols_to_transform <- c("Age", "BUN_first", "Creatinine_first",
                       "Glucose_first", "K_first", "Mg_first",
                       "Platelets_first", "WBC_first")

dfICU_transformed <- dfICU_continuous_1 %>%
  mutate(across(all_of(cols_to_transform), ~ log1p(.))) %>%
  mutate(
    asinsqrt_categorical_avg_distance = {
      scaled_value <- (.data[["categorical_avg_distance"]] - min(.data[["categorical_avg_distance"]])) /
        (max(.data[["categorical_avg_distance"]]) - min(.data[["categorical_avg_distance"]]))
      asin(sqrt(scaled_value))
    }
  ) %>%
  select(-categorical_avg_distance) %>%
  rename(categorical_avg_distance_asinsqrt = asinsqrt_categorical_avg_distance)
```

```r
dfICU_long_transformed <- dfICU_transformed %>%
  pivot_longer(cols = everything(), names_to = "variable", values_to = "value")

ggplot(dfICU_long_transformed, aes(x = value)) +
  geom_histogram(bins = 30, fill = "blue", color = "black") +
  facet_wrap(~ variable, scales = "free") +
  labs(title = "Histograms of Continuous Features (Log-Transformed Selected, Arcsin-Sqrt for categorical
       x = "Value (Transformed)",
       y = "Frequency") +
  theme_minimal() +
  theme(axis.text.y = element_blank(),
        axis.ticks.y = element_blank())
```




```r
# Clean and scale the data
df_clean <- na.omit(dfICU_transformed)
df_scaled <- scale(df_clean)

# ----- 1. PCA Scores -----
pca_result <- prcomp(df_scaled, center = TRUE, scale. = TRUE)

# Get PCA scores (coordinates in the principal component space)
pca_scores <- as.data.frame(pca_result$x)

# View first few rows of PCA scores
head(pca_scores)
```

```
##         PC1         PC2        PC3        PC4         PC5         PC6
## 1 -1.209656  0.05386235  0.5610143 -0.8948439  1.85343945 -0.09274015
## 2  3.186896 -1.18665630  0.9921379 -0.3977424 -0.31310103 -0.84832036
## 3 -2.418289 -1.47571656  1.4257554  2.1333884  0.05732908 -0.50399669
## 4 -1.042649  0.49809046  0.2144297 -1.9975927 -0.44847632  0.12249970
## 5  1.543812  0.70386092  1.3186219  1.8530958 -0.01843421 -0.17405039
## 6 -1.638841  1.82990676 -1.3021321 -1.1317070  1.82877970  0.38048958
##          PC7        PC8         PC9       PC10        PC11       PC12
## 1  0.3344645  0.2908805 -0.24635745  0.2066350  0.88029596  0.5612432
## 2  0.5089148 -0.5112688 -1.19072988 -0.8337952 -0.34819127 -0.6775206
## 3 -0.9572683  0.2934205 -0.80750718  0.2492528 -0.26527291 -1.0034086
## 4  0.1801914 -0.2278417  0.07264626  0.6808625 -0.07419826  0.1374331
## 5 -1.3503789  0.3788029  0.25593323 -0.1364192  0.31640600  2.3549244
## 6  1.1170439 -0.2702633 -0.66749321  0.4096465 -0.19901938  0.6810382
##          PC13       PC14       PC15        PC16        PC17
## 1 -0.01638168 -0.7441039 -0.4193898 -0.07149916  0.05228502
## 2  0.69604041  1.1575390  0.2934766 -0.14498991  0.08917202
## 3 -0.25239995  0.3000169  0.4546271 -0.01305061  0.07388444
## 4  0.57769285  0.1836948 -0.2780335  0.30871350  0.07791634
## 5 -1.61140255  1.3354284 -1.0751449  0.90291424 -0.01643674
## 6 -0.95558130 -0.1059179  0.9750434 -0.49575480  0.06552460
```

```r
# Standard deviations of the principal components
std_devs <- pca_result$sdev

# Variance explained by each PC
var_explained <- std_devs^2
prop_var_explained <- var_explained / sum(var_explained)

# Create a data frame of variance explained
pca_variance_df <- data.frame(
  PC = paste0("PC", 1:length(prop_var_explained)),
  Variance_Explained = prop_var_explained,
  Cumulative_Variance = cumsum(prop_var_explained)
)

# View the variance explained
print(pca_variance_df)
```

```
##      PC Variance_Explained Cumulative_Variance
## 1   PC1        0.163221959           0.1632220
## 2   PC2        0.121921149           0.2851431
## 3   PC3        0.090419533           0.3755626
## 4   PC4        0.081873428           0.4574361
## 5   PC5        0.069304570           0.5267406
## 6   PC6        0.064619507           0.5913601
## 7   PC7        0.060427307           0.6517875
## 8   PC8        0.054411714           0.7061992
## 9   PC9        0.051297807           0.7574970
## 10 PC10        0.048744297           0.8062413
## 11 PC11        0.039851779           0.8460930
## 12 PC12        0.039370374           0.8854634
## 13 PC13        0.035345582           0.9208090
## 14 PC14        0.034059211           0.9548682
```

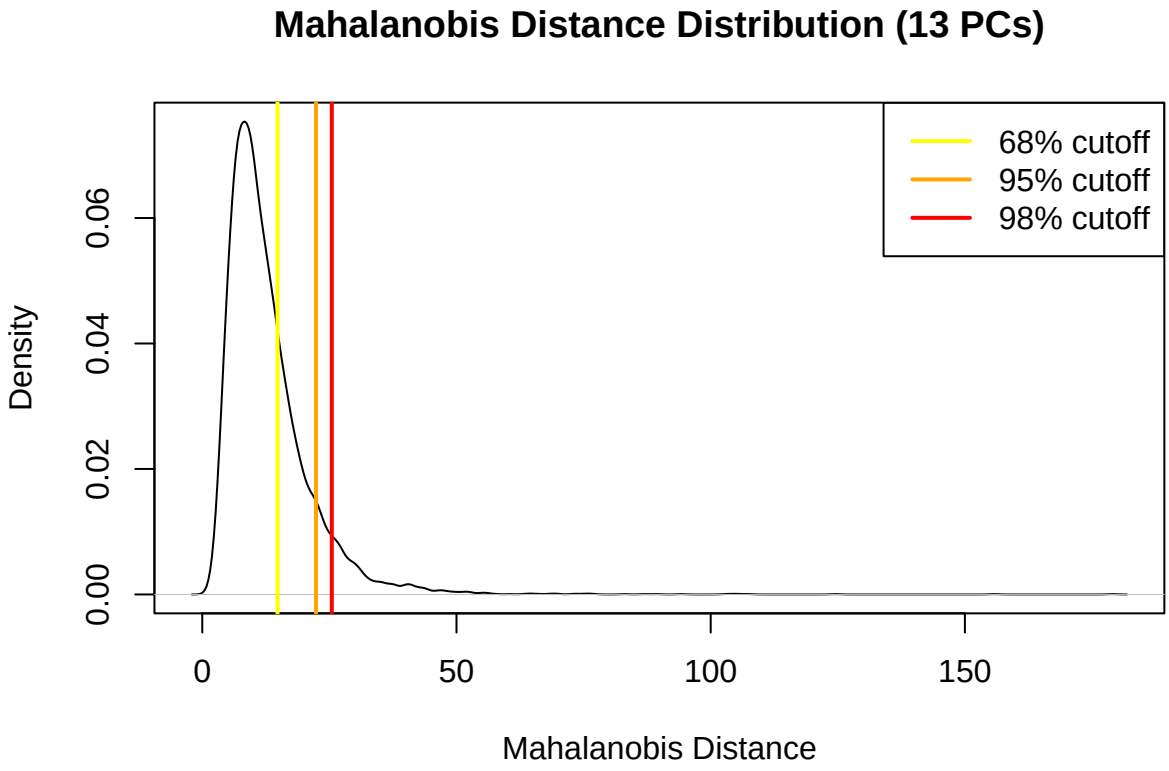
```
## 15 PC15        0.023854198           0.9787224
## 16 PC16        0.013526708           0.9922491
## 17 PC17        0.007750878           1.0000000
```

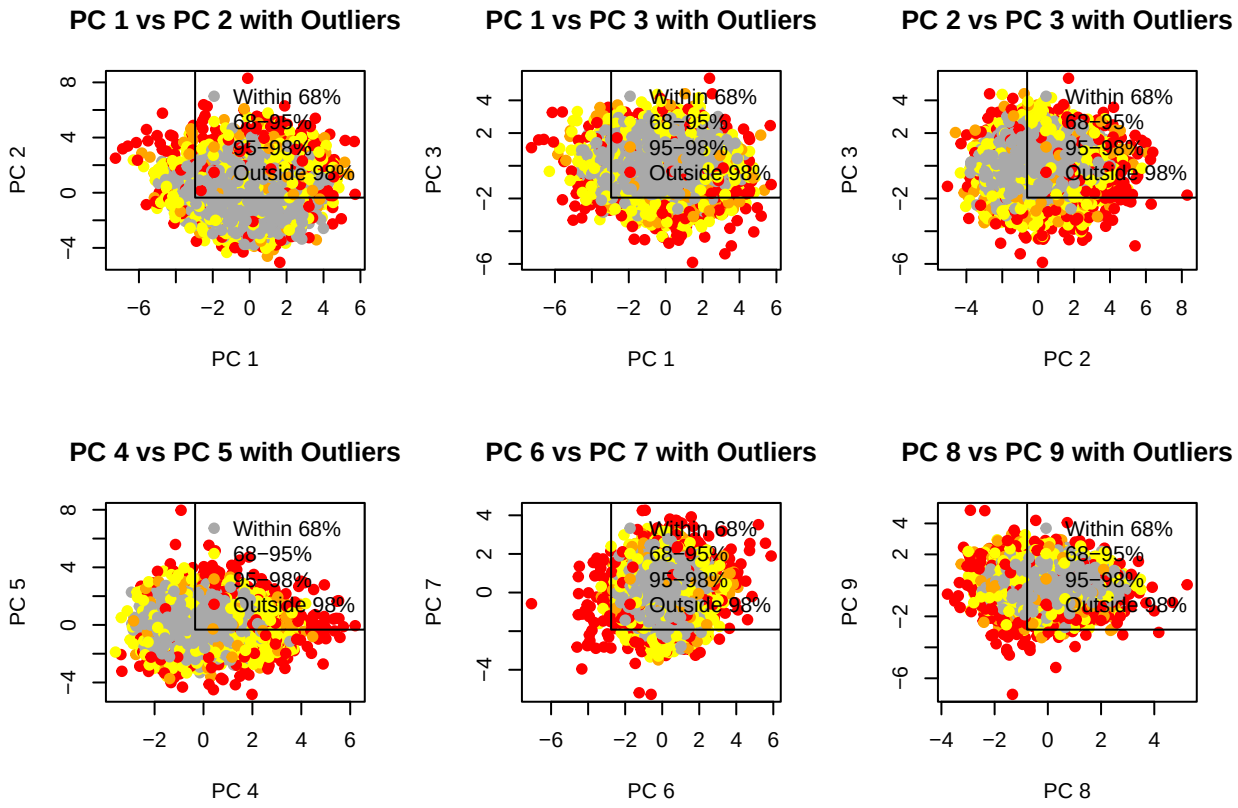
```r
pca_mahalanobis_analysis <- function(df, n_components = 13) {
  # Ensure we're working with numeric data only
  numeric_cols <- sapply(df, is.numeric)
  data_numeric <- df[, numeric_cols]

  # Remove NA values if any
  if(any(is.na(data_numeric))) {
    warning("NA values detected. Using complete cases only.")
    data_numeric <- na.omit(data_numeric)
  }

  # Store row names for later matching
  original_rows <- rownames(data_numeric)

  # Perform PCA
  pca_result <- prcomp(data_numeric, scale. = TRUE)

  # Extract the first n_components PCs
  pc_scores <- pca_result$x[, 1:n_components]

  # Calculate Mahalanobis distances using the PC scores
  mean_vec <- colMeans(pc_scores)
  cov_mat <- cov(pc_scores)

  mahalanobis_dist <- mahalanobis(pc_scores, mean_vec, cov_mat)

  # Create results dataframe
  result_df <- data.frame(
    original_index = as.integer(original_rows),
    mahalanobis_dist = mahalanobis_dist
  )

  # Add outlier flags for different confidence levels
  # For p dimensions, Mahalanobis distance follows chi-square with p degrees of freedom
  result_df$outlier_68 <- mahalanobis_dist > qchisq(0.68, df = n_components)
  result_df$outlier_95 <- mahalanobis_dist > qchisq(0.95, df = n_components)
  result_df$outlier_98 <- mahalanobis_dist > qchisq(0.98, df = n_components)

  # Count outliers
  outlier_counts <- list(
    "Total_observations" = nrow(result_df),
    "68%" = sum(result_df$outlier_68),
    "95%" = sum(result_df$outlier_95),
    "98%" = sum(result_df$outlier_98)
  )

  # Calculate percentages
  outlier_percentages <- list(
    "68%" = round(100 * outlier_counts[["68%"]] / outlier_counts[["Total_observations"]], 2),
    "95%" = round(100 * outlier_counts[["95%"]] / outlier_counts[["Total_observations"]], 2),
```

```
    "98%" = round(100 * outlier_counts[["98%"]] / outlier_counts[["Total_observations"]], 2)
  )

  return(list(
    result_df = result_df,
    pca_result = pca_result,
    pc_scores = pc_scores,
    outlier_counts = outlier_counts,
    outlier_percentages = outlier_percentages
  ))
}

# Plot functions
plot_mahalanobis_density <- function(result, n_components) {
  distances <- result$result_df$mahalanobis_dist

  # Create the density plot
  plot(density(distances),
       main = "Mahalanobis Distance Distribution (13 PCs)",
       xlab = "Mahalanobis Distance",
       ylab = "Density")

  # Add vertical lines for the chi-square cutoffs
  abline(v = qchisq(0.68, df = n_components), col = "yellow", lwd = 2)
  abline(v = qchisq(0.95, df = n_components), col = "orange", lwd = 2)
  abline(v = qchisq(0.98, df = n_components), col = "red", lwd = 2)

  # Add a legend
  legend("topright",
         legend = c("68% cutoff", "95% cutoff", "98% cutoff"),
         col = c("yellow", "orange", "red"),
         lwd = 2)
}

# Plot the first few PCs with outliers highlighted
plot_pc_outliers <- function(result, pc1 = 1, pc2 = 2) {
  pc_data <- result$pc_scores
  outliers <- result$result_df

  # Create PCA plot
  plot(pc_data[, pc1], pc_data[, pc2],
       col = ifelse(outliers$outlier_98, "red",
                    ifelse(outliers$outlier_95, "orange",
                           ifelse(outliers$outlier_68, "yellow", "darkgrey"))),
       pch = 19,
       main = paste("PC", pc1, "vs PC", pc2, "with Outliers"),
       xlab = paste("PC", pc1),
       ylab = paste("PC", pc2))

  # Add legend
  legend("topright",
         legend = c("Within 68%", "68-95%", "95-98%", "Outside 98%"),
         pch = 19,
```

```
         col = c("darkgrey", "yellow", "orange", "red"))
}

# Main function to run the analysis and display results
apply_pca_mahalanobis <- function(df, n_components = 13) {
  # Run the PCA-based Mahalanobis analysis
  cat("Performing PCA and calculating Mahalanobis distances...\n")
  result <- pca_mahalanobis_analysis(df, n_components)

  # Print summary of results
  cat("\n============ PCA-based Outlier Detection Results ============\n")
  cat("Analysis performed using the first", n_components, "principal components\n")
  cat("These components explain approximately 92% of the total variance\n")
  cat("Total observations:", result$outlier_counts[["Total_observations"]], "\n\n")

  cat("Outlier counts and percentages at different confidence levels:\n")
  cat("68% confidence level:", result$outlier_counts[["68%"]], "outliers (",
      result$outlier_percentages[["68%"]], "%)\n")
  cat("95% confidence level:", result$outlier_counts[["95%"]], "outliers (",
      result$outlier_percentages[["95%"]], "%)\n")
  cat("98% confidence level:", result$outlier_counts[["98%"]], "outliers (",
      result$outlier_percentages[["98%"]], "%)\n")

  # Create visualizations
  cat("\nCreating visualizations...\n")

  # Plot Mahalanobis distance distribution
  plot_mahalanobis_density(result, n_components)

  # Plot the first 6 pairs of PCs
  par(mfrow = c(2, 3))
  plot_pc_outliers(result, 1, 2)
  plot_pc_outliers(result, 1, 3)
  plot_pc_outliers(result, 2, 3)
  plot_pc_outliers(result, 4, 5)
  plot_pc_outliers(result, 6, 7)
  plot_pc_outliers(result, 8, 9)
  par(mfrow = c(1, 1)) # Reset to single plot

  # Display examples of extreme outliers
  cat("\nExamples of extreme outliers (top 5 by Mahalanobis distance):\n")
  extreme_indices <- order(result$result_df$mahalanobis_dist, decreasing = TRUE)[1:5]
  extreme_rows <- result$result_df$original_index[extreme_indices]
  print(df[extreme_rows, ])

  cat("\nAnalysis complete.\n")

  # Return the results dataframe with outlier flags
  invisible(result$result_df)
}

# Run the analysis on dfICU_transformed using 13 PCs
outlier_results <- apply_pca_mahalanobis(dfICU_transformed, n_components = 13)
```

```
## Performing PCA and calculating Mahalanobis distances...
##
## ============ PCA-based Outlier Detection Results ============
## Analysis performed using the first 13 principal components
## These components explain approximately 92% of the total variance
## Total observations: 7886
##
## Outlier counts and percentages at different confidence levels:
## 68% confidence level: 2321 outliers ( 29.43 %)
## 95% confidence level: 825 outliers ( 10.46 %)
## 98% confidence level: 541 outliers ( 6.86 %)
##
## Creating visualizations...
```






```
##
## Examples of extreme outliers (top 5 by Mahalanobis distance):
##           Age Glucose_first HR_first NIDiasABP_first NIMAP_first NISysABP_first
## 3941 4.262680      5.273000       76              69       86.33            117
## 6141 4.174387      4.644391      118              85      108.30            155
## 1537 4.382027      6.999422       60              41       62.33            105
## 925  3.713572      5.267858       85              47       54.00             75
## 2637 4.276666      4.595120      109              46       65.00            134
##      Temp_first BUN_first Creatinine_first HCO3_first HCT_first  K_first
## 3941       35.7  2.639057        0.5877867         22      33.7 1.667707
## 6141       38.1  2.397895        0.4054651         20      30.3 3.025291
## 1537       35.0  2.564949        0.5877867         18      32.5 1.335001
## 925        33.7  3.258097        0.8754687         14      27.4 1.791759
## 2637       26.1  3.178054        0.6931472         26      32.5 1.526056
##       Mg_first Na_first Platelets_first WBC_first
## 3941 3.0910425      142        5.225747  2.186051
## 6141 0.7884574      123        5.652489  3.414443
## 1537 0.8329091       98        4.934474  1.704748
## 925  2.3887628      125        4.615121  2.850707
## 2637 1.0296194      139        5.153292  2.054124
##      categorical_avg_distance_asinsqrt
## 3941                         0.6797308
## 6141                         0.5908116
## 1537                         1.2455827
## 925                          0.8279668
## 2637                         0.8925085
```

```
##
## Analysis complete.
```

```r
# Load necessary library
library(ggplot2)

# Assume pca_scores is already available from your prcomp() call.
# Extract the first 13 principal components
pca_13 <- pca_scores[, 1:13]

# ------------------------------
# 1. Compute the Mahalanobis Distance
# ------------------------------

# Calculate the center (mean) and covariance matrix for the PCA-13 data
center <- colMeans(pca_13)
cov_matrix <- cov(pca_13)

# Compute squared Mahalanobis distances for all observations
mahal_sq <- mahalanobis(pca_13, center, cov_matrix)

# For later plotting, also compute the raw Mahalanobis distance (optional)
mahal <- sqrt(mahal_sq)

# ------------------------------
# 2. Define Chi-square Thresholds
# ------------------------------

# Compute chi-square thresholds for 68%, 95%, and 99% using 13 degrees of freedom
thresh_68 <- qchisq(0.68, df = 13)
thresh_95 <- qchisq(0.95, df = 13)
thresh_99 <- qchisq(0.99, df = 13)

# ------------------------------
# 3. Classify Anomalies Based on Thresholds
# ------------------------------

# Create anomaly flags based on the squared Mahalanobis distance
anomaly68 <- ifelse(mahal_sq > thresh_68, "Yes", "No")
anomaly95 <- ifelse(mahal_sq > thresh_95, "Yes", "No")
anomaly99 <- ifelse(mahal_sq > thresh_99, "Yes", "No")

# Add these variables to a new data frame combining the scores and Mahalanobis distances
results <- data.frame(
  pca_13,
  Mahalanobis_Squared = mahal_sq,
  Mahalanobis = mahal,
  Anomaly_68 = anomaly68,
  Anomaly_95 = anomaly95,
  Anomaly_99 = anomaly99
)

# Optionally, you can view a summary of the results
head(results)
```

```
##         PC1         PC2        PC3        PC4         PC5         PC6
## 1 -1.209656  0.05386235  0.5610143 -0.8948439  1.85343945 -0.09274015
## 2  3.186896 -1.18665630  0.9921379 -0.3977424 -0.31310103 -0.84832036
## 3 -2.418289 -1.47571656  1.4257554  2.1333884  0.05732908 -0.50399669
## 4 -1.042649  0.49809046  0.2144297 -1.9975927 -0.44847632  0.12249970
## 5  1.543812  0.70386092  1.3186219  1.8530958 -0.01843421 -0.17405039
## 6 -1.638841  1.82990676 -1.3021321 -1.1317070  1.82877970  0.38048958
##          PC7        PC8         PC9       PC10        PC11       PC12
## 1  0.3344645  0.2908805 -0.24635745  0.2066350  0.88029596  0.5612432
## 2  0.5089148 -0.5112688 -1.19072988 -0.8337952 -0.34819127 -0.6775206
## 3 -0.9572683  0.2934205 -0.80750718  0.2492528 -0.26527291 -1.0034086
## 4  0.1801914 -0.2278417  0.07264626  0.6808625 -0.07419826  0.1374331
## 5 -1.3503789  0.3788029  0.25593323 -0.1364192  0.31640600  2.3549244
## 6  1.1170439 -0.2702633 -0.66749321  0.4096465 -0.19901938  0.6810382
##          PC13 Mahalanobis_Squared Mahalanobis Anomaly_68 Anomaly_95 Anomaly_99
## 1 -0.01638168            6.168766    2.483700         No         No         No
## 2  0.69604041           10.502564    3.240766         No         No         No
## 3 -0.25239995           11.506794    3.392167         No         No         No
## 4  0.57769285            4.837696    2.199476         No         No         No
## 5 -1.61140255           19.507009    4.416674        Yes         No         No
## 6 -0.95558130           11.855388    3.443165         No         No         No
```

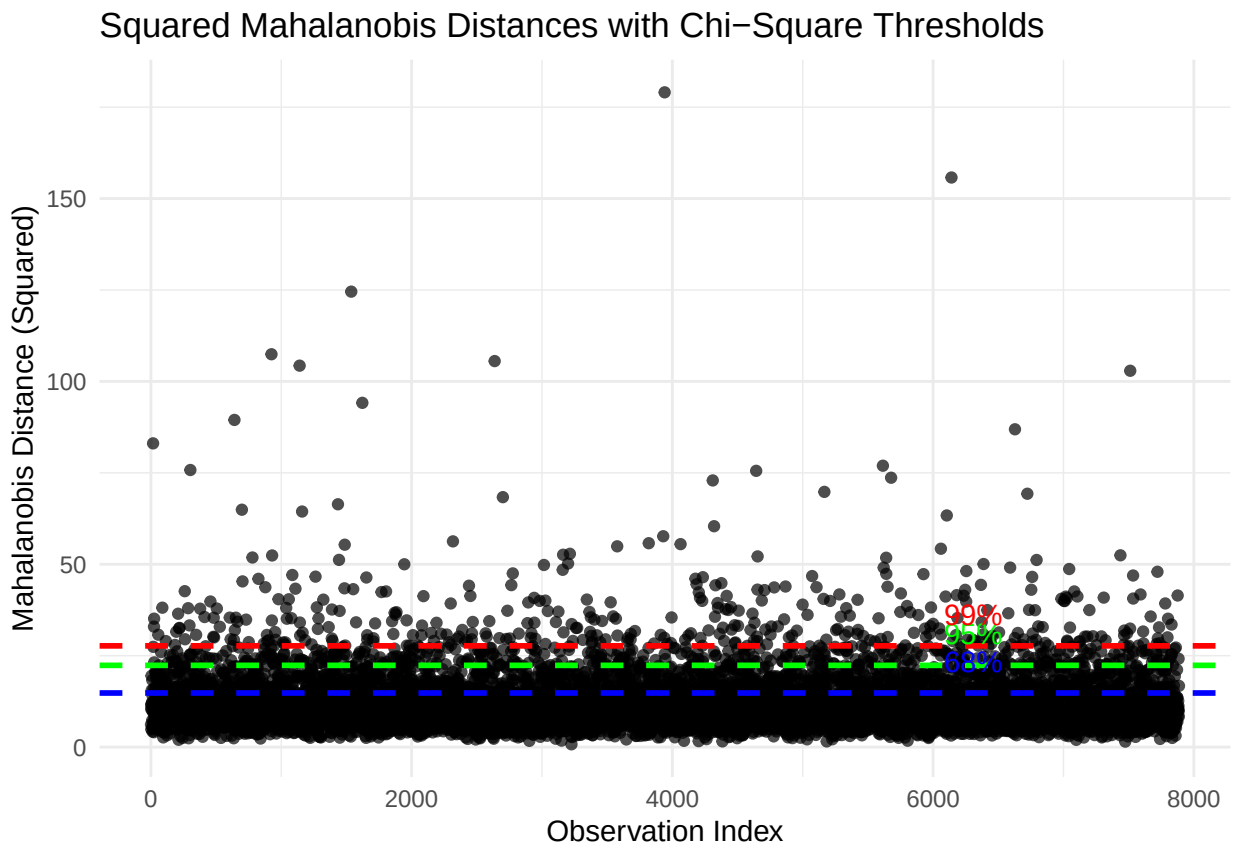
```
summary(results)
```

```
##       PC1                 PC2                PC3                PC4
##  Min.   :-7.27034   Min.   :-5.0507   Min.   :-5.90618   Min.   :-3.5990
##  1st Qu.:-1.12707   1st Qu.:-0.9721   1st Qu.:-0.78410   1st Qu.:-0.7982
##  Median : 0.04763   Median :-0.1446   Median : 0.06516   Median :-0.1273
##  Mean   : 0.00000   Mean   : 0.0000   Mean   : 0.00000   Mean   : 0.0000
##  3rd Qu.: 1.21111   3rd Qu.: 0.8177   3rd Qu.: 0.84837   3rd Qu.: 0.6338
##  Max.   : 5.71665   Max.   : 8.2986   Max.   : 5.35123   Max.   : 6.2031
##       PC5                 PC6                PC7                PC8
##  Min.   :-4.825215   Min.   :-7.05846   Min.   :-5.27500   Min.   :-3.771029
##  1st Qu.:-0.675021   1st Qu.:-0.66630   1st Qu.:-0.64493   1st Qu.:-0.639861
##  Median : 0.004525   Median :-0.02295   Median : 0.01929   Median :-0.004904
##  Mean   : 0.000000   Mean   : 0.00000   Mean   : 0.00000   Mean   : 0.000000
##  3rd Qu.: 0.638149   3rd Qu.: 0.64736   3rd Qu.: 0.65298   3rd Qu.: 0.615424
##  Max.   : 7.965420   Max.   : 5.88539   Max.   : 4.26159   Max.   : 5.236557
##       PC9                PC10               PC11               PC12
##  Min.   :-7.0375   Min.   :-3.78832   Min.   :-4.17157   Min.   :-6.83631
##  1st Qu.:-0.5651   1st Qu.:-0.62227   1st Qu.:-0.52086   1st Qu.:-0.49453
##  Median : 0.0168   Median :-0.00124   Median :-0.01781   Median : 0.01886
##  Mean   : 0.0000   Mean   : 0.00000   Mean   : 0.00000   Mean   : 0.00000
##  3rd Qu.: 0.5949   3rd Qu.: 0.60848   3rd Qu.: 0.50575   3rd Qu.: 0.50870
##  Max.   : 4.8500   Max.   : 3.95913   Max.   : 3.97648   Max.   : 6.63301
##       PC13          Mahalanobis_Squared  Mahalanobis       Anomaly_68
##  Min.   :-3.7426   Min.   :  0.7725    Min.   : 0.8789   Length:7886
##  1st Qu.:-0.5072   1st Qu.:  7.5199    1st Qu.: 2.7422   Class :character
##  Median :-0.0313   Median : 10.9104    Median : 3.3031   Mode  :character
##  Mean   : 0.0000   Mean   : 12.9984    Mean   : 3.4570
##  3rd Qu.: 0.4599   3rd Qu.: 15.9638    3rd Qu.: 3.9955
##  Max.   : 5.4833   Max.   :179.0467    Max.   :13.3808
##   Anomaly_95         Anomaly_99
##  Length:7886        Length:7886
```

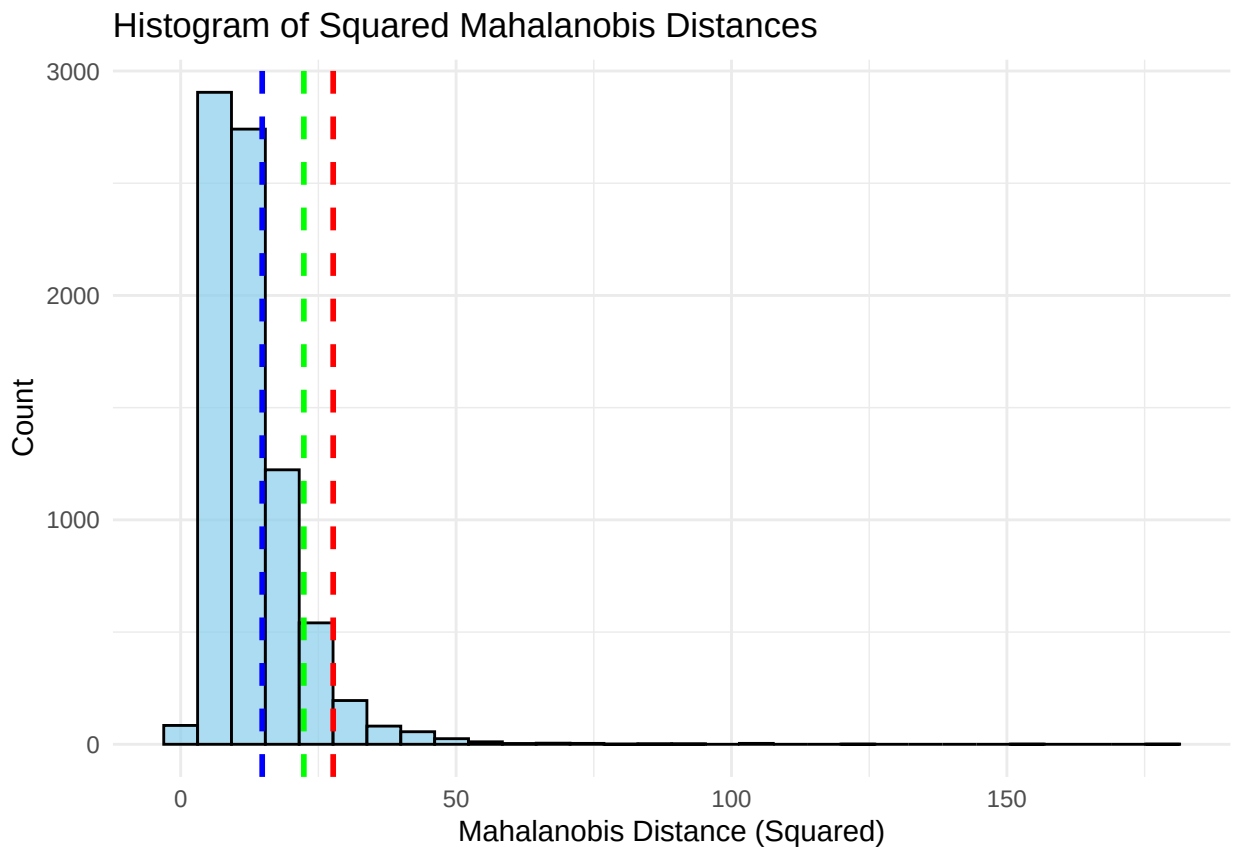
```
##  Class :character   Class :character
##  Mode  :character   Mode  :character
##
##
##
```

```r
# ------------------------------
# 4. Plotting the Mahalanobis Distances with Thresholds
# ------------------------------

# Option A: Scatter plot of squared Mahalanobis distances (by observation index) with threshold lines
df_plot <- data.frame(Index = 1:nrow(results), Mahalanobis_Squared = mahal_sq)

ggplot(df_plot, aes(x = Index, y = Mahalanobis_Squared)) +
  geom_point(alpha = 0.7) +
  geom_hline(yintercept = thresh_68, color = "blue", linetype = "dashed", size = 1) +
  geom_hline(yintercept = thresh_95, color = "green", linetype = "dashed", size = 1) +
  geom_hline(yintercept = thresh_99, color = "red", linetype = "dashed", size = 1) +
  labs(title = "Squared Mahalanobis Distances with Chi-Square Thresholds",
       y = "Mahalanobis Distance (Squared)",
       x = "Observation Index") +
  annotate("text", x = max(df_plot$Index) * 0.8, y = thresh_68,
           label = "68%", color = "blue", vjust = -1) +
  annotate("text", x = max(df_plot$Index) * 0.8, y = thresh_95,
           label = "95%", color = "green", vjust = -1) +
  annotate("text", x = max(df_plot$Index) * 0.8, y = thresh_99,
           label = "99%", color = "red", vjust = -1) +
  theme_minimal()
```

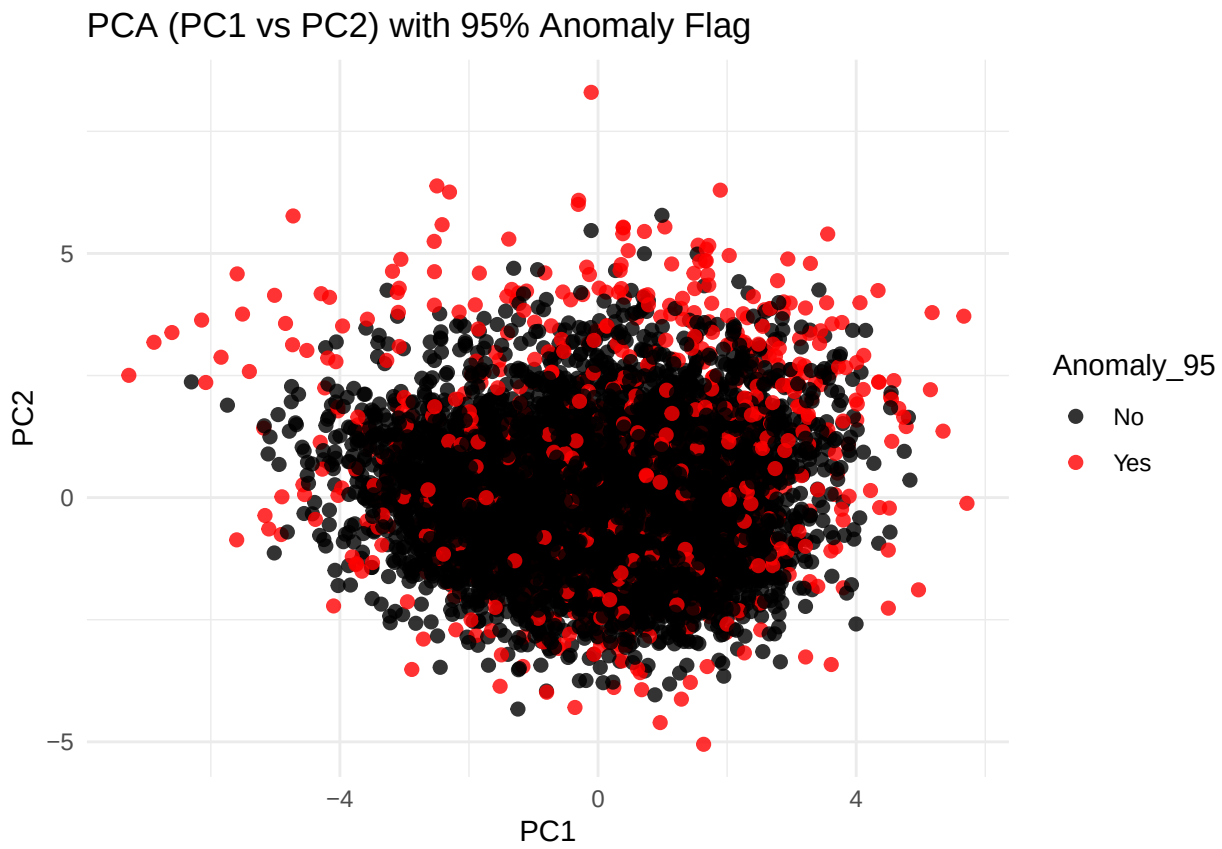
```
## Warning: Using ‘size‘ aesthetic for lines was deprecated in ggplot2 3.4.0.
## i Please use ‘linewidth‘ instead.
## This warning is displayed once every 8 hours.
## Call ‘lifecycle::last_lifecycle_warnings()‘ to see where this warning was
## generated.
```




```r
# Option B: Histogram of squared Mahalanobis distances with overlaid threshold lines
ggplot(df_plot, aes(x = Mahalanobis_Squared)) +
  geom_histogram(bins = 30, fill = "skyblue", color = "black", alpha = 0.7) +
  geom_vline(xintercept = thresh_68, color = "blue", linetype = "dashed", size = 1) +
  geom_vline(xintercept = thresh_95, color = "green", linetype = "dashed", size = 1) +
  geom_vline(xintercept = thresh_99, color = "red", linetype = "dashed", size = 1) +
  labs(title = "Histogram of Squared Mahalanobis Distances",
       x = "Mahalanobis Distance (Squared)",
       y = "Count") +
  theme_minimal()
```




```
# ------------------------------
# 5. (Optional) Plot the PCA-13 Space with Anomalies Highlighted
# ------------------------------
# For a 2D view, you can plot the first two principal components and color points by an anomaly flag.
# Here we use the 95% threshold as an example.
ggplot(results, aes(x = PC1, y = PC2, color = Anomaly_95)) +
  geom_point(alpha = 0.8, size = 2) +
  labs(title = "PCA (PC1 vs PC2) with 95% Anomaly Flag",
       x = "PC1", y = "PC2") +
  scale_color_manual(values = c("No" = "black", "Yes" = "red")) +
  theme_minimal()
```




```
# Install and load necessary package if not already installed
if (!require(GGally)) install.packages("GGally")
```

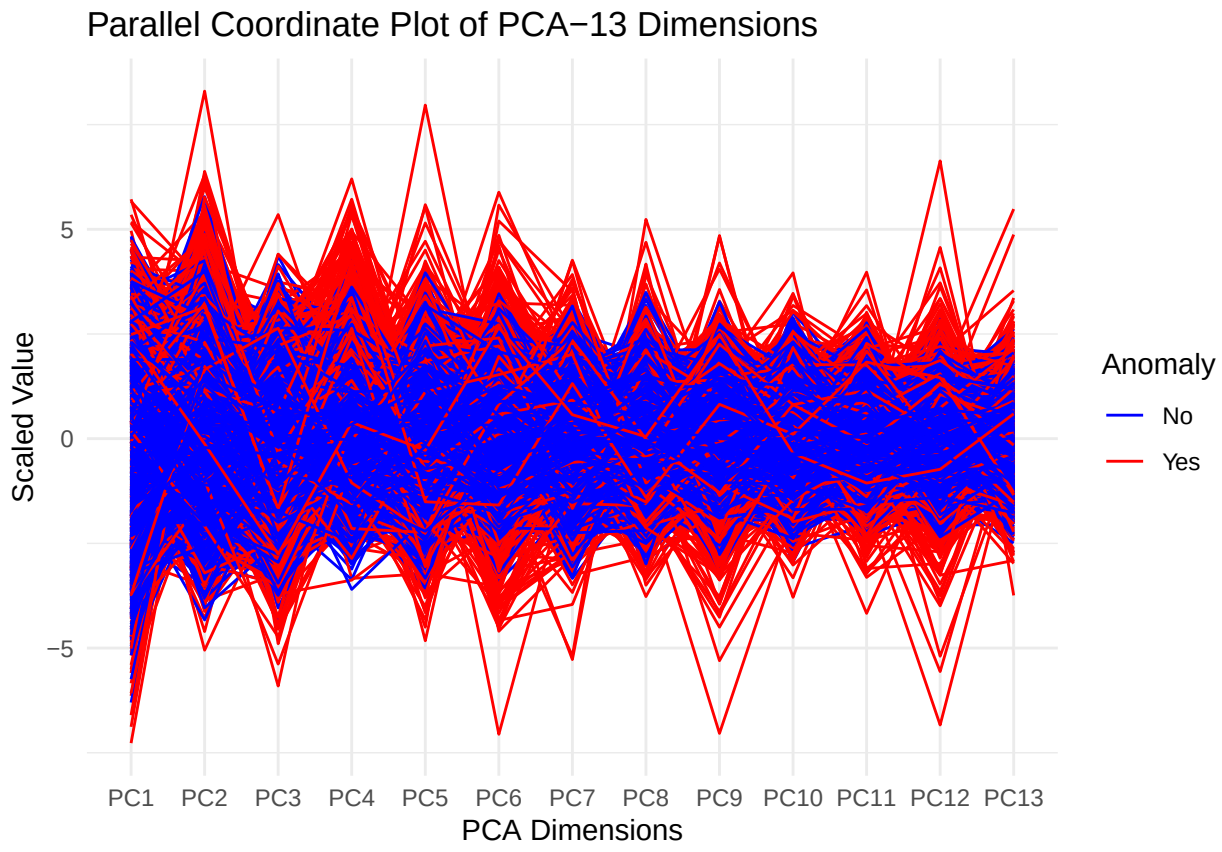
```
## Loading required package: GGally
```

```
## Warning: package 'GGally' was built under R version 4.4.3
```

```
## Registered S3 method overwritten by 'GGally':
##   method from
##   +.gg   ggplot2
```

```
library(GGally)
library(ggplot2)

# Assume 'results' is your data frame that includes the PCA-13 scores and an anomaly indicator.
# For demonstration purposes, create a factor based on the 95% threshold anomaly flag.
results$Anomaly <- as.factor(results$Anomaly_95)

# Create a parallel coordinate plot for the 13 principal components.
# 'columns' identifies the PCA columns, and 'groupColumn' uses the anomaly factor for color grouping.
ggparcoord(data = results,
           columns = 1:13,         # columns corresponding to PCA-13
           groupColumn = "Anomaly", # group by anomaly status
           scale = "globalminmax") +  # scales each variable between its minimum and maximum values
```

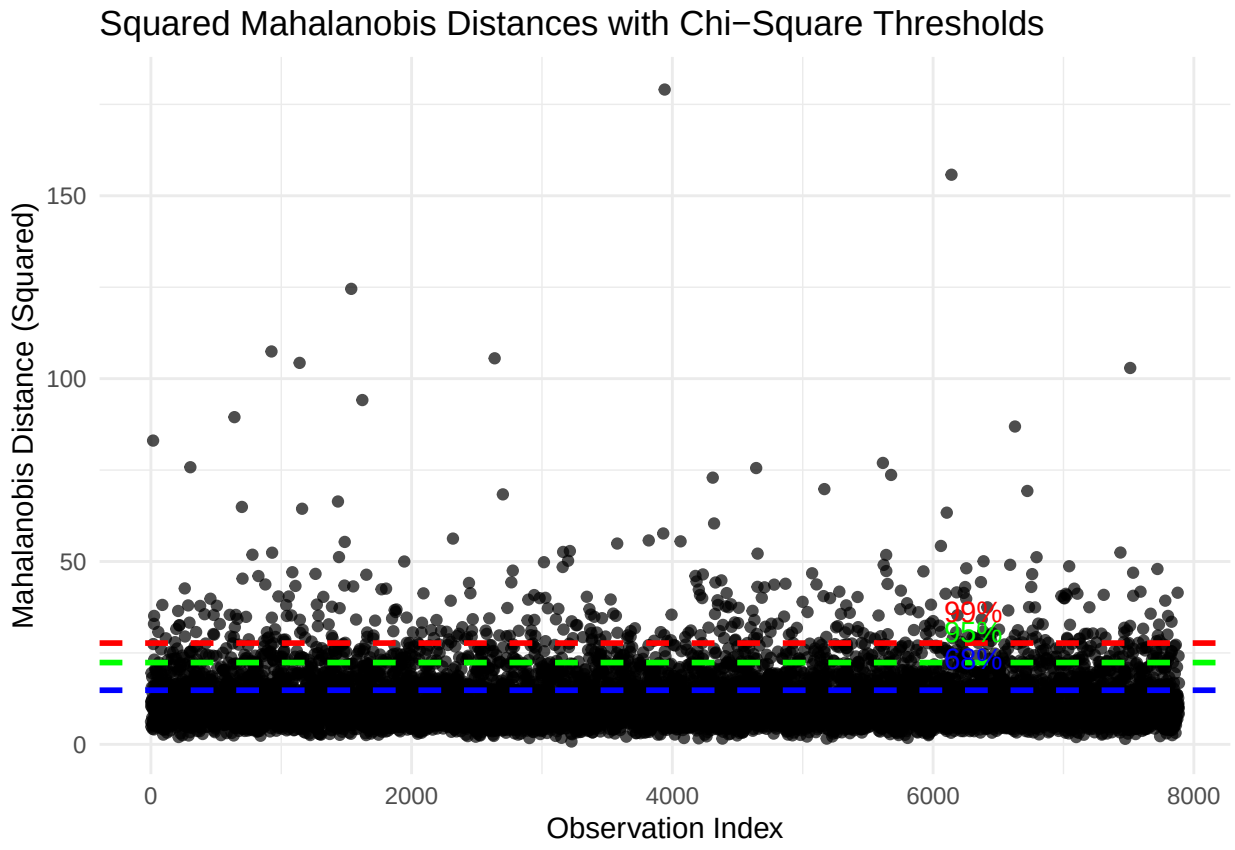
```
  theme_minimal() +
  labs(title = "Parallel Coordinate Plot of PCA-13 Dimensions",
       x = "PCA Dimensions",
       y = "Scaled Value") +
  scale_color_manual(values = c("No" = "blue", "Yes" = "red"))
```




```
# Load necessary libraries
library(ggplot2)

# ------------------------------
# Assume pca_scores is already available from your prcomp() call
# Extract the first 13 principal components
pca_13 <- pca_scores[, 1:13]

# ------------------------------
# 1. Compute the Mahalanobis Distance
# ------------------------------

# Calculate the center (mean) and covariance matrix for the PCA-13 data
center <- colMeans(pca_13)
cov_matrix <- cov(pca_13)

# Compute squared Mahalanobis distances for all observations
mahal_sq <- mahalanobis(pca_13, center, cov_matrix)
# Compute the raw Mahalanobis distances (optional)
```

```r
mahal <- sqrt(mahal_sq)

# ------------------------------
# 2. Define Chi-square Thresholds for 13 dimensions
# ------------------------------

thresh_68 <- qchisq(0.68, df = 13)
thresh_95 <- qchisq(0.95, df = 13)
thresh_99 <- qchisq(0.99, df = 13)

# ------------------------------
# 3. Classify Anomalies Based on Each Threshold
# ------------------------------

# For each threshold, flag observations as "Yes" if the squared Mahalanobis distance
# exceeds the threshold, "No" otherwise.
anomaly68 <- ifelse(mahal_sq > thresh_68, "Yes", "No")
anomaly95 <- ifelse(mahal_sq > thresh_95, "Yes", "No")
anomaly99 <- ifelse(mahal_sq > thresh_99, "Yes", "No")

# Combine the PCA-13 scores and the computed distances/anomaly flags into one data frame.
results <- data.frame(
  pca_13,
  Mahalanobis_Squared = mahal_sq,
  Mahalanobis = mahal,
  Anomaly_68 = anomaly68,
  Anomaly_95 = anomaly95,
  Anomaly_99 = anomaly99
)

# ------------------------------
# 4. Save the Anomaly Results for Each Threshold into Separate CSV Files
# ------------------------------

# Create a directory "Data/" if it doesn't already exist.
if(!dir.exists("Data")){
  dir.create("Data")
}

# Save all the results (if needed)
write.csv(results, file = "Data/anomalies_all.csv", row.names = FALSE)

# Save only those observations flagged as anomalies for each threshold.

# 68% threshold anomalies
results_68 <- results[results$Anomaly_68 == "Yes", ]
write.csv(results_68, file = "Data/anomalies_68.csv", row.names = FALSE)

# 95% threshold anomalies
results_95 <- results[results$Anomaly_95 == "Yes", ]
write.csv(results_95, file = "Data/anomalies_95.csv", row.names = FALSE)

# 99% threshold anomalies
```

```r
results_99 <- results[results$Anomaly_99 == "Yes", ]
write.csv(results_99, file = "Data/anomalies_99.csv", row.names = FALSE)

# ------------------------------
# 5. Plotting the Mahalanobis Distances with Thresholds (Optional)
# ------------------------------

# Create a plot of the squared Mahalanobis distances with threshold lines.
df_plot <- data.frame(Index = 1:nrow(results), Mahalanobis_Squared = mahal_sq)

ggplot(df_plot, aes(x = Index, y = Mahalanobis_Squared)) +
  geom_point(alpha = 0.7) +
  geom_hline(yintercept = thresh_68, color = "blue", linetype = "dashed", size = 1) +
  geom_hline(yintercept = thresh_95, color = "green", linetype = "dashed", size = 1) +
  geom_hline(yintercept = thresh_99, color = "red", linetype = "dashed", size = 1) +
  labs(title = "Squared Mahalanobis Distances with Chi-Square Thresholds",
       y = "Mahalanobis Distance (Squared)",
       x = "Observation Index") +
  annotate("text", x = max(df_plot$Index) * 0.8, y = thresh_68,
           label = "68%", color = "blue", vjust = -1) +
  annotate("text", x = max(df_plot$Index) * 0.8, y = thresh_95,
           label = "95%", color = "green", vjust = -1) +
  annotate("text", x = max(df_plot$Index) * 0.8, y = thresh_99,
           label = "99%", color = "red", vjust = -1) +
  theme_minimal()
```



```r
# Optionally, create a 2D plot of the PCA space (PC1 vs PC2) highlighting anomalies by threshold
ggplot(results, aes(x = PC1, y = PC2, color = Anomaly_95)) +
  geom_point(alpha = 0.8, size = 2) +
  labs(title = "PCA (PC1 vs PC2) with 95% Anomaly Flag",
       x = "PC1", y = "PC2") +
  scale_color_manual(values = c("No" = "black", "Yes" = "red")) +
  theme_minimal()
```